Adapting Models to Signal Degradation using Distillation

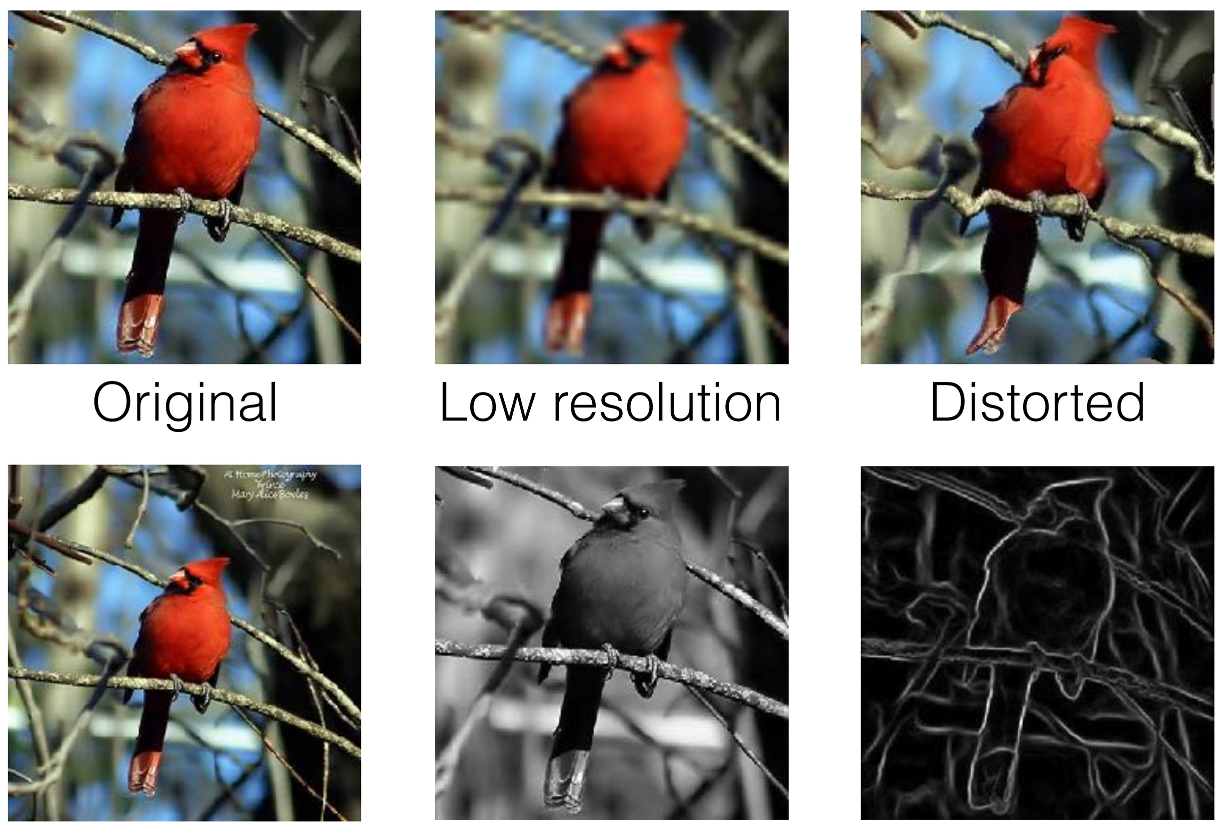
Jong-Chyi Su Subhransu Maji
University of Massachusetts, Amherst

BMVC 2017 LONDON

Motivation

Convolutional neural networks are effective at several visual recognition tasks. However, they are prone to the degradation of image quality.

Various forms of image degradation:



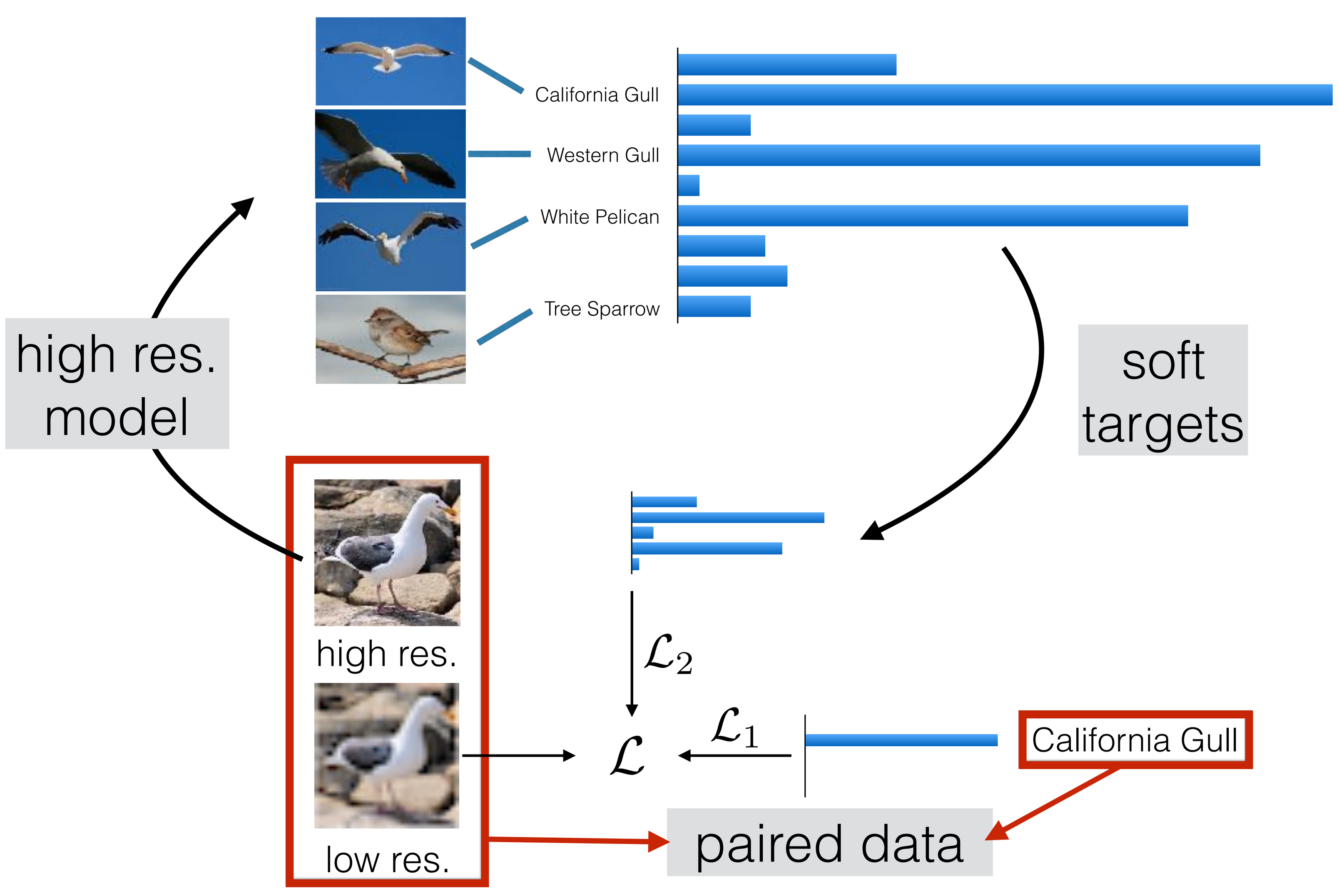
Original · Low resolution · Distorted · Unlocalized · Grayscale · Line drawing

Prior Work

- Problem-specific solutions
  - Super-resolution, Image colorization, etc.
- Domain adaptation
  - DSLR ⇨ webcam images, high ⇨ low resolution images
- Transfer learning
  - ImageNet ⇨ CUB low resolution data

Our Approach

- **Cross Quality Distillation (CQD)**
- Assumption: "paired" training data is available
  1. Train a model on high quality data
  2. Train a model on the low quality data to match:
     (1) The target labels
     (2) The "soft targets" produced by the first model



Soft targets provide more information per example leading to better knowledge transfer across domains [1, 2]

References

[1] Distilling knowledge in a neural network. G. Hinton, O. Vinyals, and J. Dean, *Deep Learning and Representation Learning Workshop, NIPS*, 2014.
[2] Model compression. C. Buciluă, R. Caruana, and A. Niculescu-Mizil, *SIGKDD*, 2006.

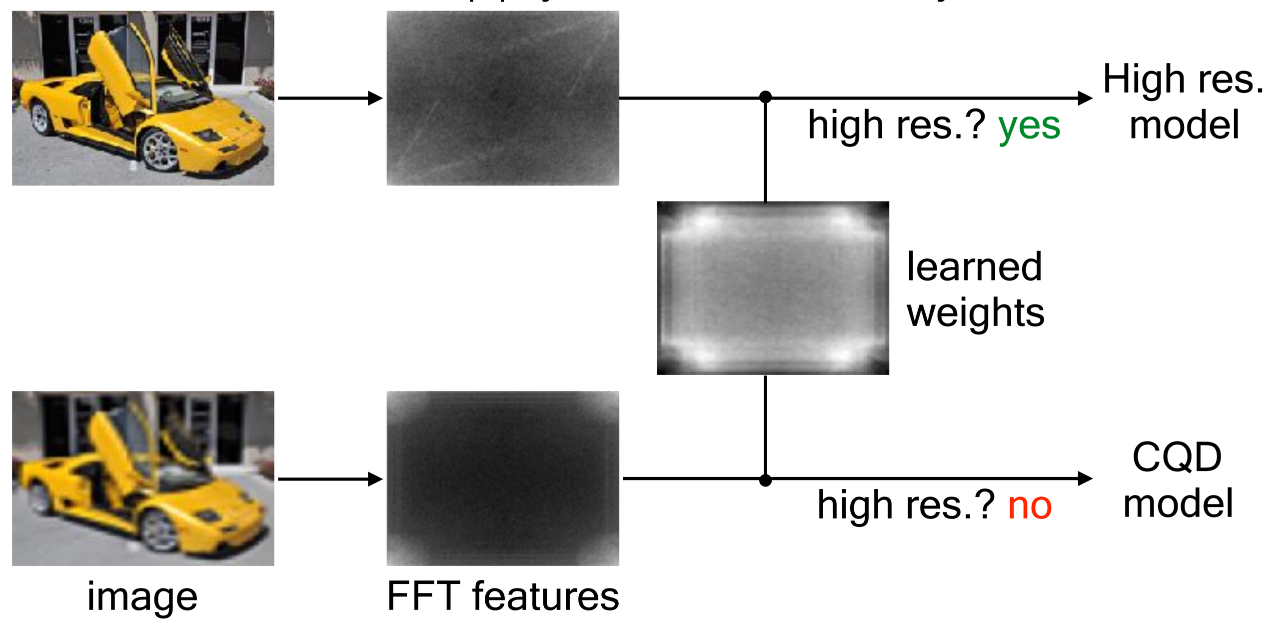
Results

Accuracy on CUB low quality dataset

Model	Localization	Low resolution	Line drawing	Local. + low res.
Oracle	*67.0*	*67.0*	*67.0*	*67.0*
No adaptation	57.4	39.4	1.9	24.9
Fine-tuning	60.8	61.0	29.2	46.2
Data augmentation	63.6	62.2	32.5	51.7
Staged training	62.4	62.3	30.4	50.4
CQD	**64.4**	**64.4**	**34.1**	**52.7**

Multi-Resolution Model

1. Train a model to predict image quality.
2. For each image:
   (1) If it's high resolution, apply model trained on high res. data
   (2) If it's low resolution, apply model trained by CQD



Model	Test on high res.	Test on low res.	Test on high res. + low res.
Train on high res. data	59.3	7.6	33.5
CQD	31.8	48.8	40.3
Train on high res. + low res.	57.3	47.3	52.3
Multi-resolution	-	-	**54.0**
Oracle	*59.3*	*48.8*	*54.1*

Multi- resolution model is better than using one model trained on high res. + low res. data.

Distillation and Model Compression

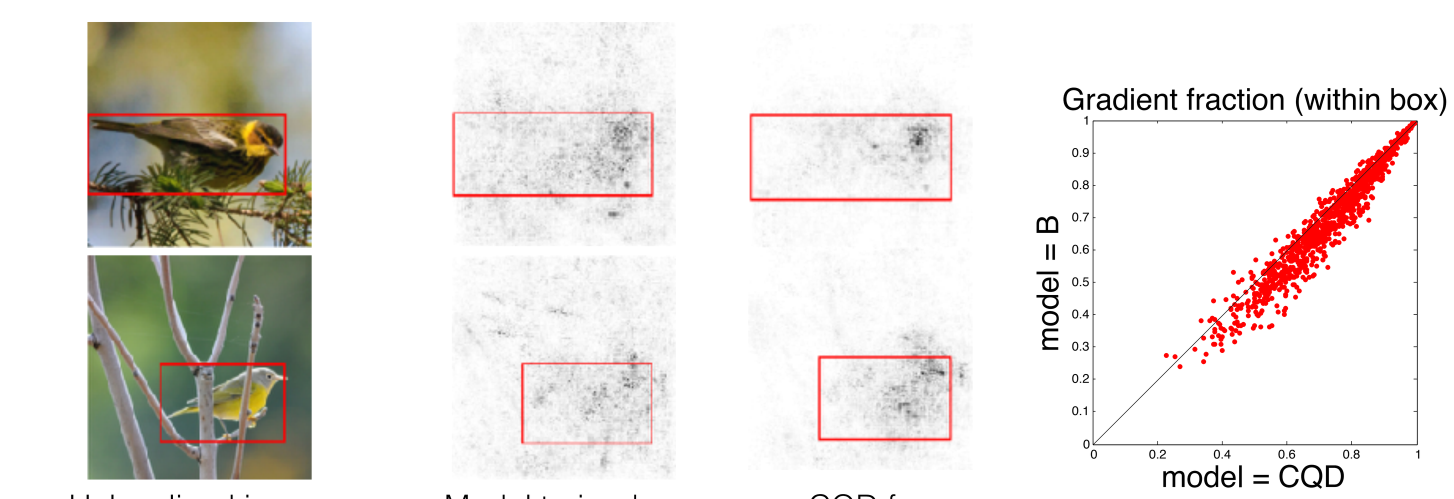
A deeper CNN trained on high-quality data can be distilled to a shallow CNN trained on low-quality data.

Model distilled from → to	Training → Testing A → A	B → B	CQD → B
VGG-m → VGG-m	67.0	60.8	63.7
VGG-16 → VGG-m	-	-	**64.6**
VGG-16 → VGG-16	74.9	69.5	72.4

A: CUB localized
B: CUB unlocalized

Visualization

We visualize the gradient of an image w.r.t. true class label. By distilling from localized images to unlocalized images, the gradient is more focused inside the bounding-boxes.



Unlocalized image
Red box: Localized image
Model trained on unlocalized images
CQD from localized images

Distillation from a localized image can help us localize the object on an unlocalized image.