

I'll describe two lines of work from my group on fine-grained recognition. The first is a class of deep neural network architectures called bilinear CNNs that are remarkably effective at fine-grained recognition tasks such as identifying the species of a bird, or the make and model of a car. These generalize classical texture representations (e.g., Bag-of-Visual-Words or Fisher vectors) but are end-to-end trainable. The second is a technique for adapting models to various forms of signal degradation. Current techniques for fine-grained recognition are severely affected by loss in image resolution, or when objects have increased variation in pose and viewpoint. The proposed technique systematically generates low-quality data from existing labelled datasets and transfers the representations learned on high-quality data to a model trained on low-quality data using a technique called "distillation". This way of transferring representations outperforms existing techniques for domain adaptation.