1 Problem Statement

How can we use zero-shot and transfer learning to better denoise images?

2 Deep Learning Book

Relevant chapters from DLB [1]

- **7.7 Multitask Learning**
  - The model can generally be divided into two kinds of parts and associated parameters:
    1. Task-specific parameters (which only benefit the examples of their task to achieve good generalization). These are the upper layers of the neural network in figure 7.2
    2. Generic parameters, shared across all the tasks (which benefit from the pooled data of all the tasks). These are the lower layers of the neural network in figure 7.2
  - From the point of view of deep learning, the underlying prior belief is the following: *among the factors that explain the variations observed in the data associated with the different tasks, some are shared across two or more tasks.*
- **7.13 Adversarial Training**
  - Adversarial examples also provide a means of accomplishing semi-supervised learning
  - Approach encourages the classifier to learn a function that is robust to small changes anywhere along the manifold where the unlabeled data lie
  - The assumption motivating this approach is that different classes usually lie on the disconnected manifolds, and a small perturbation should not be able to jump from one class manifold to another class manifold
- **15 Representation Learning**
  - Training with supervised learning techniques on the labeled subset often results in severe overfitting
  - Semi-supervised learning offers the chance to resolve this overfitting problem by also learning from the unlabeled data
  - Specifically, we can learn good representations for the unlabeled data, and then use these representations to solve the supervised learning task
- **15.2 Transfer Learning and Domain Adaptation**

- The learner must perform two or more different tasks, but we assume that many of the factors that explain the variations in P_1 are relevant to the variations that need to be captured for learning P_2
- Typically understood in a supervised learning context, where the input is the same but the target may be of a different nature
- Two extreme forms of transfer learning are *one-shot learning* and *zerof-shot learning*, sometimes also called *zero-data learning*.
  * Only one labeled example of the transfer task is given for one-shot learning, while no labeled examples are given at all for the zero-shot learning task
  * Zero-data learning [2] and zero-shot learning [3, 4]

3 Papers

3.1 Zero-Shot Learning

- CleanNet: Transfer Learning for Scalable Image Classifier Training With Label Noise [5]
  - In this paper, we study the problem of learning image classification models with label noise. Existing approaches depending on human supervision are generally not scalable as manually identifying correct or incorrect labels is time-consuming, whereas approaches not relying on human supervision are scalable but less effective. To reduce the amount of human supervision for label noise cleaning, we introduce CleanNet, a joint neural embedding network, which only requires a fraction of the classes being manually verified to provide the knowledge of label noise that can be transferred to other classes. We further integrate CleanNet and conventional convolutional neural network classifier into one framework for image classification learning. We demonstrate the effectiveness of the proposed algorithm on both of the label noise detection task and the image classification on noisy data task on several large-scale datasets. Experimental results show that CleanNet can reduce label noise detection error rate on held-out classes where no human supervision available by 41.5% compared to current weakly supervised methods. It also achieves 47% of the performance gain of verifying all images with only 3.2% images verified on an image classification task. Source code and dataset will be available at kuanghuei.github.io/CleanNetProject.

3.2 Image Denoising

- **Deep Learning for Image Denoising: A Survey** [6]
  - Since the proposal of big data analysis and Graphic Processing Unit (GPU), the deep learning technology has received a great deal of attention and has been widely applied in the field of imaging processing. In this paper, we have an aim

to completely review and summarize the deep learning technologies for image denoising proposed in recent years. Morever, we systematically analyze the conventional machine learning methods for image denoising. Finally, we point out some research directions for the deep learning technologies in image denoising.

  - *4.1 The challenges of deep learning technologies in image denoising*
    1. Current deep learning denoising methods only deal with AWGN, which are not effective for real noisy images, such as low light images.
    2. They can't use a model to deal with all the low level vision taks, such as image denoising, image super-resolution, image blurring, and image deblocking.
    3. They can't use a model to address the blind Gaussian noise

- **Correction by Projection: Denoising Images with Generative Adversarial Networks** [7]

  - Generative adversarial networks (GANS) transform low-dimensional latent vectors into visually plausible images. If the real dataset contains only clean images, then ostensibly, the manifold learned by the GAN should contain only clean images. In this paper, we propose to denoise corrupted images by finding the nearest point on the GAN manifold, recovering latent vectors by minimizing distances in image space. We first demonstrate that given a corrupted version of an image that truly lies on the GAN manifold, we can approximately recover the latent vector and denoise the image, obtaining significantly higher quality, comparing with BM3D. Next, we demonstrate that latent vectors recovered from noisy images exhibit a consistent bias. By subtracting this bias before projecting back to image space, we improve denoising results even further. Finally, even for unseen images, our method performs better at denoising than BM3D. Notably, the basic version of our method (without bias correction) requires no prior knowledge on the noise variance. To achieve the highest possible denoising quality, the best performing signal processing based methods, such as BM3D, require an estimate of the blur kernel.

- **Very Deep Convolutional Networks for Large-Scale Image Recognition** [8]

  - In this work we investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3x3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16-19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localisation and classification tracks respectively. We also show that our representations generalise well to other datasets, where they achieve state-of-the-art results. We have made our two best-performing ConvNet models publicly available to facilitate further research on the use of deep visual representations in computer vision.

- **Universal Denoising Networks : A Novel CNN Architecture for Image Denoising** [9]

– We design a novel network architecture for learning discriminative image models that are employed to efficiently tackle the problem of grayscale and color image denoising. Based on the proposed architecture, we introduce two different variants. The first network involves the convolutional layers as a core component, while the second one relies instead on non-local filtering layers and thus it is able to exploit the inherent non-local self-similarity property of natural images. As opposed to most of the existing deep network approaches, which require the training of a specific model for each considered noise level, the proposed models are able to handle a wide range of noise levels using a single set of learned parameters, while they are very robust when the noise degrading the latent image does not match the statistics of the noise used during training. The latter argument is supported by results that we report on publicly available images corrupted by uknown noise and which we compare against solutions obtained by competing methods. At the same time the introduced netowrks achieve excellent results under additive white Gaussian noise (AWGN), which are comparable to those of the current state-of-the-art network, while they depend on a more shallow architecture with the number of trained parameters being one order of magnitude smaller. These properties make the proposed netowrks ideal candidates to serve as sub-solvers on restoration methods that deal with general inverse imaging problems such as deblurring, demosaicking, superresolution, etc.

References

[1] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[2] Hugo Larochelle and Y Bengio. Classification using discriminative restricted boltzmann machines. pages 536–543, 01 2008.

[3] Mark Palatucci, Dean Pomerleau, Geoffrey Hinton, and Tom M. Mitchell. Zero-shot learning with semantic output codes. In *Proceedings of the 22Nd International Conference on Neural Information Processing Systems*, NIPS'09, pages 1410–1418, USA, 2009. Curran Associates Inc.

[4] Richard Socher, Milind Ganjoo, Hamsa Sridhar, Osbert Bastani, Christopher D. Manning, and Andrew Y. Ng. Zero-shot learning through cross-modal transfer, 2013.

[5] Kuang-Huei Lee, Xiaodong He, Lei Zhang, and Linjun Yang. Cleannet: Transfer learning for scalable image classifier training with label noise. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[6] Chunwei Tian, Yong Xu, Lunke Fei, and Ke Yan. Deep learning for image denoising: A survey, 2018.

[7] Subarna Tripathi, Zachary C. Lipton, and Truong Q. Nguyen. Correction by projection: Denoising images with generative adversarial networks. *CoRR*, abs/1803.04477, 2018.

[8] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

[9] Stamatios Lefkimmiatis. Universal denoising networks : A novel cnn-based network architecture for image denoising. *CoRR*, abs/1711.07807, 2017.