Hi,

The topic of our group is image denoising. The dataset we are using is the cropped Yale Face Database contains images of different people under different lighting conditions. There are three samples from our dataset: these three are the same person under different lighting conditions.

The problem we are trying to solve is to restore original images from noisy ones. One real world application area could be medical images. For example, image denoising could be very helpful in some medical diagnose such as restoring obscured parts in MRI scan.

In our project, we tried both supervised and unsupervised methods, which consist of Markov Random Field, Total Variation Minimization, Low-Rank representation and, U-net and Deep Image Prior. I'll briefly talk about them later. For ethical implication, I don’t think our methods have any ethical problems, since image denoising is a well-established area.

The goal of image denoising is to recover the original image from a corrupted one. In order to restore the original image, we need to make assumption(a.k.a prior knowledge) about the original image. First, we use the low-variation assumption

MRF

We used an Ising model for our MRF method, which consists of the X, the observed pixel values and Y, the prediction pixel value. We assume that the probability is an energy model, for the energy function, consider the truth-observation potential, ~~which encode the difference between truth and observation~~ and neighbor potential, ~~which encodes the variation between neighbor pixels.~~ Then the denoising task becomes the inference task where we estimate P(Y|X). We use Gibbs sampling for inference, where for each Y_ij we sample Y_ij given all other variables. ~~And recall that in a Markov network, a variable is independent of other variables~~ given its Markov blanket. ~~Thus~~ we can easily compute the conditional of a

Y_ij from its neighbors. And we iteratively sample each Y_ij. The result of MRF are shown here. From top down, we add the noise and compute the PSNR to compare the performance.

TV Minimization

In MRF, we kind of take into consideration the variation in prediction image and the difference between the prediction and observation. So why don't we throw away the probability part of MRF? so we came up with a simple optimization problem as shown here. The X hat is our prediction and X is the noisy image. The total variation computes the difference between each variable and its neighbors. So what this optimization problem says is, we want the prediction to have low variation, while keeping it close to the observation. And we want to balance these to objectives with a hyperparameter lambda. Here is the result of this optimization problem. The result is pretty good, and solving this optimization problem is really fast.

Compressd Sensing

Low-rank representation

In previous assumptions, we only used one image. So, can we get use information from other images to denoise one image? In our dataset, we have information containing 64 different persons, and for each people, we have images under different lighting conditions. Our motivation is, for example, we have 25 images on the screen, there are only 5 people, roughly speaking, the degree of freedom of the images are 5, and other freedom come from the noise or corruption. Thus we assume the data come from a low-rank subspace. So we used the low-rank representation for image denoising. Here, we flatten all images and stack them along the column axis. Since there are many images and only a few people, we assume the matrix X can be decomposed as a low rank matrix plus some error term. ~~To make X low rank, we can multiply X by a low rank matrix Z.~~ And here is the optimization problem, ~~the X is the images, XZ is the low-rank representation of~~

~~images, and E is the error term~~. But minimizing matrix rank is np hard, the rank minimization problem can be approximated by a nuclear norm minimization, and we can solve this by augmented Lagrangian. Here is the result of low-rank representation. We can see that for small noise and medium noise, though the PSNR score is not very high, the image looks pretty good. But for large noise, LR representation is not as good as other methods. However, low rank representation has some interesting side effect. ~~Let's see these images,~~ the left part is the original image, and the right part is the low-rank representation. We can see that this method ~~magically~~ recover some corruption in the image. In the first image, the shade near the nose is eliminated, and in the second image, it kind of recovered the left eye. Maybe we can use Low-rank representation for image restoration, but we did not have time to further investigate this.

Unet

In low rank representation, we use a lot of images. And when we have a lot of data, a natural idea is to use deep learning. So we also tried deep learning method. We used a U-net, where we add noise to image as input and the original images as label. We use 2k images as training set, and 200 images for validation set. For the input, we randomly add noise to the original image. Here is the result of the U-net model. The PSNR values are higher than all previous models in for all kinds of noises, and the visual effect is also very good.

Deep Image Prior

However, one drawback of deep learning Is that it need a lot of data to have good performance. ~~Can we get rid of this data-dependency? Our answer is yes.~~ If we look and our previous methods, all of them can be formulated into the form on the slide. We want to make the prediction image close to the noisy image under some regularization. In MRF and Total Variation Minimization, the regularization is that the prediction image has low variation, and in low rank representation, we assume the original images come from a low-rank subspace. In deep image prior, we want the prediction to be very close to the noisy image, subject to the regularization that the prediction is produced by a deep CNN. Here we use the architecture of CNN as prior. The motivation is, in Deep CNN, we stack some convolution layers together and wish it to extract some semantic meaning of the image, and they are pretty good at this than other architectures. So, the architecture of the Deep CNN may encodes prior knowledge about how an image should look like. In Deep Image prior, we input a random noise and output an image, then optimize the weights of the CNN to make the prediction close to the noisy image. And here is the result of the Deep Image Prior.

In conclusion, the U-net model achieves the highest PSNR value. For the deep image prior model, though it's not good as the U-net model, it gives a good result without using a lot of data~~ (actually it only uses one image)~~. And the total variation minimization, it can not beat the UNet based on PSNR score, but it can also be a good choice because of the cheap computation. As for low rank representation, even it has very low PSNR score, the visual effect is not bad and it is promising in other perspectives.