Spatially Variant Methods in Image Reconstruction

Ian Hudson

December 10, 2021

Abstract

Currently convolutional methods are used to solve image reconstruction problems. However, convolutions have the inherent property of being spatially invariant. We look at a model that is said to have better spatially variant denoising properties and test to see if it really does. The method for testing the spatially variant denoising ability is extracting the part of the model that selectively

1 Introduction

Denoising is an important problem in computer imaging as it allows for reconstruction of information that was lost due to hardware or some other reason. In this report, we focus on reconstructing images with spatially variant noise. The noise of an image is considered spatially variant if the "type" or level of noise differs at different locations in the image. One real world example of this is in MRI imaging. MRI imaging equipment slowly samples the Fourier space of the image it is taking and rarely fully samples the space. The undersampling of the Fourier space causes artifacts in the pixel space, such as ghosting. Ghosting artifacts appear as repeated versions of some main object placed side by side and next to the original object, making them a type of spatially variant noise. (https://mriquestions.com/ghosting.html)

2 Denoising Methods

2.1 Current Methods

Currently, popular methods for denoising images use convolutional neural networks (CNN), which applies a fixed kernel to the entire image. CNNs lack the ability to differentiate kernel transformations based on local information. This makes them inherently bad at reconstructing images where local information is important for determining the transformation needed to reconstruct the image.

2.2 Novel Methods for Denoising with Spatial Variance

Methods designed to handle spatially variant noise typically use kernels to transform images, just like CNN. However, instead of a fixed kernel, these methods find ways to dynamically select a kernel based on the information of the image where the kernel is going to be applied. For example, in Involution, a special function is learned that uses a single pixel's values to generate the kernel for that pixel. [LHW+21]

2.2.1 UDVD

Unified Dynamic Convolutional Network for Variational Degradations (UDVD) is another model that claims to have better spatially variant denoising properties. UDVD has two main parts, a feature map and a refinement network. The feature map extracts high level data about the input image, which is then fed to the refinement network. The refinement network consists of multiple dynamic blocks, which are where the per-pixel kernels are generated. The first step in the dynamic $k \times k$-kernel generation is concatenating the feature information with the input image, and feeding it to a convolutional layer that outputs a "dynamic block" with dimensions $H \times W \times k^2$. The kernel for each pixel is a flattening of the k^2 values for that pixel position in the dynamic block. [XTT+20]. We are interested in seeing if the dynamic kernel generation in UDVD is really able to spatially select good kernels for denoising. This is achieved by using the same dynamic kernel generation in different models that are spatially invariant, such as DnCNN.

3 DnCNN with Dynamic Kernel

The model that we will test is a variant of DnCNN that includes a dynamic kernel inspired by UDVD. DnCNN is a denoising methods that uses CNN, so adding a dynamic kernel should improve performance in spatially variant tasks compared to regular DnCNN.

3.1 Description of Models

Some preliminary tests were performed to determine the best placement for the dynamic kernel within the DnCNN model. Those tests showed that the dynamic kernel should be placed at the tail of the network for the best performance. We create two models that differ in the complexity of dynamic kernel generation. In the first model we use a single convolutional layer to generate the dynamic kernel (DnCNN_DK), see Figure 1. In the second model, multiple convolutional layers with multiple channels and batch normalization are used in the dynamic kernel generator (DnCNN_DDK). In order to maintain similar complexity across all models, we remove one layer from the CNN of the second model, see Figure 2.

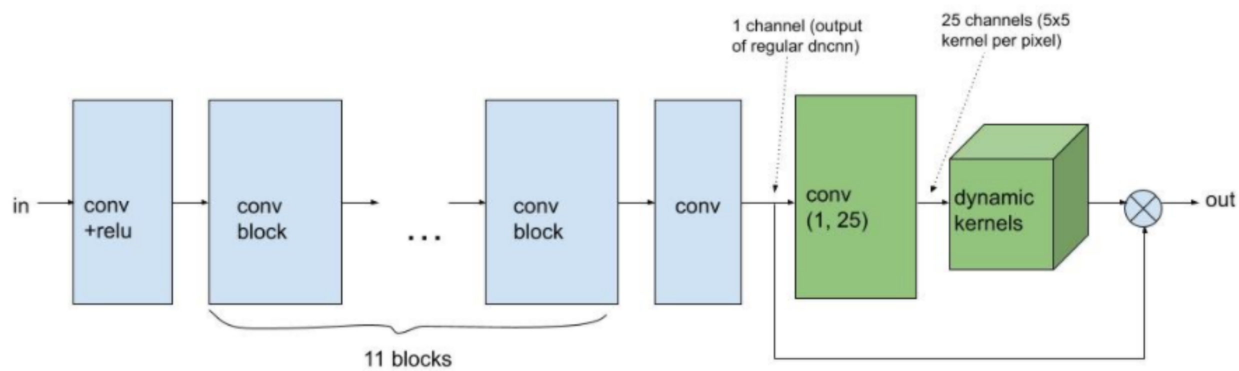

Figure 1: DnCNN with single layer dynamic kernel generation.

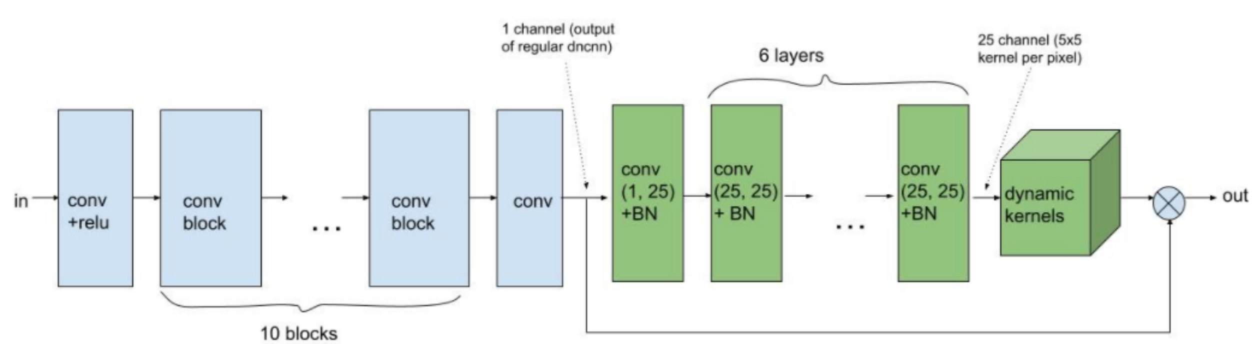

Figure 2: DnCNN with deep dynamic kernel generation.

4 Experiments

4.1 Dataset and Noise Map

To test the spatial variant properties of our models, we train them on knee MRI images obtained from NYU fastMRI [19] with different levels of Gaussian noise applied to different parts of the image. We divide our training images into four equal quadrants apply a Gaussian noise with standard deviation determined by the quadrant we are in. For our tests we used standard deviation values of 15, 25, 50 and 100, where our pixel values range from 0 to 255.

4.1.1 Preprocessing

The ground truth images are first preprocessed by mapping the smallest value in each image to 0, the largest value in each image to 1, and all other values linearly between 0 and 1. This is to account for the significantly different pixel value ranges the ground truth images have.

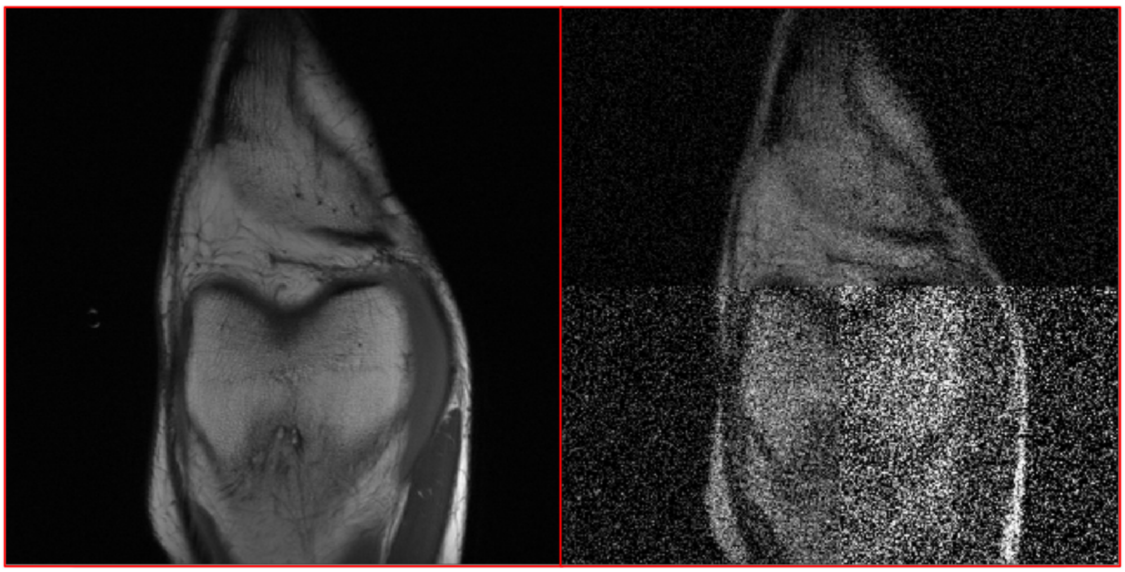

Figure 3: Example of knee MRI ground truth (left) and noisy image (right).

4.2 Results and Analysis

Model	Trainable parameters	PSNR	SSIM
DnCNN	$408 * 10^3$	29.42	0.852
DnCNN_DK	$408 * 10^3$	29.17	0.822
DnCNN_DDK	$405 * 10^3$	29.57	0.874

There is no significant improvement in PSNR or SSIM when using dynamic kernels, only a slight improvement with deep dynamic kernel. However, when we look at some images from the test set where DnCNN w/ deep dynamic does better than DnCNN in terms of PSNR and SSIM, we notice that it performs significantly better in the background portions image, see Figure 4. This is likely the reason for improvements in our testing metrics, because we see no noticeable improvement in the reconstruction of the main object. One reason for the lack of improvement may be that the dynamic kernel is generated using convolution layers which are inherently spatially invariant, so the dynamic kernel that it generates may be unable to "escape" the spatial invariance of convolution.

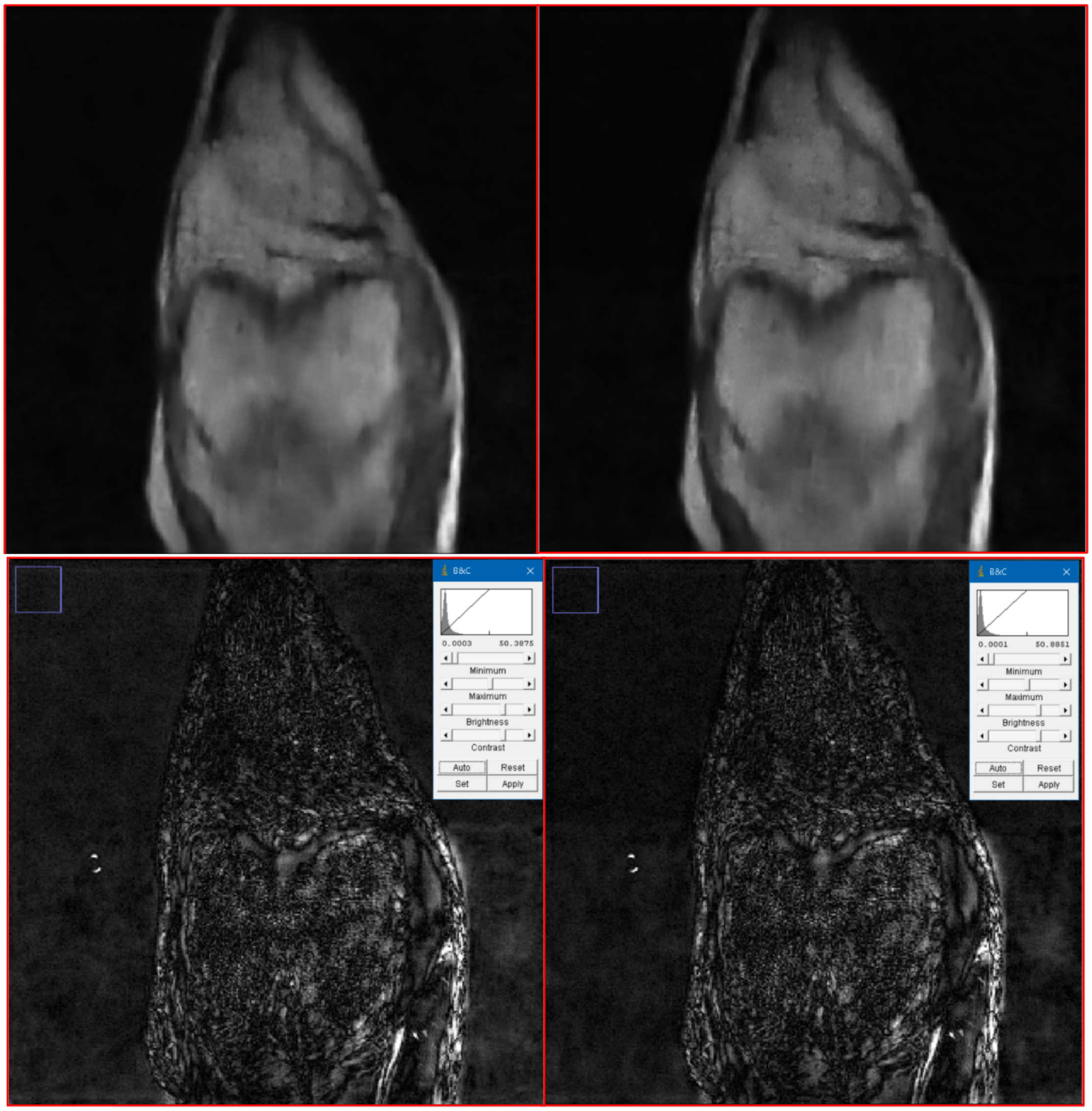

Figure 4: Reconstruction of image using DnCNN (top left) and DnCNN_DDK (top right) with PSNR values of 29.82 and 30.21, respectively. Absolute value of error for DnCNN (bottom left) and DnCNN_DDK (bottom right).

References

[*19*] Jure Zbontar *et al.* fastmri: An open dataset and benchmarks for accelerated mri, 2019.

[LHW+21] Duo Li, Jie Hu, Changhu Wang, Xiangtai Li, Qi She, Lei Zhu, Tong Zhang, and Qifeng Chen. Involution: Inverting the inherence of convolution for visual recognition. 2021.

[XTT+20] Yu-Syuan Xu, Shou-Yao Roy Tseng, Yu Tseng, Hsien-Kai Kuo, and Yi-Min Tsai. Unified dynamic convolutional network for super-resolution with variational degradations. 2020.