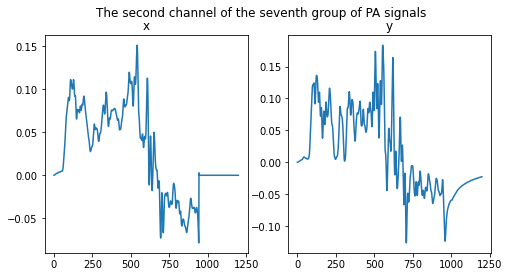
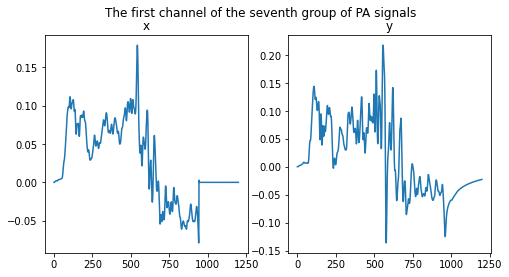
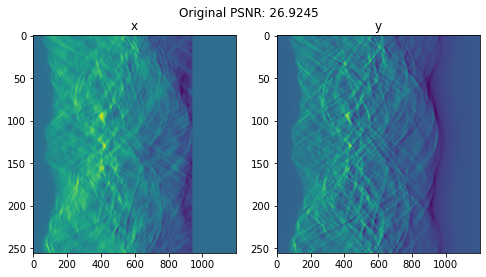
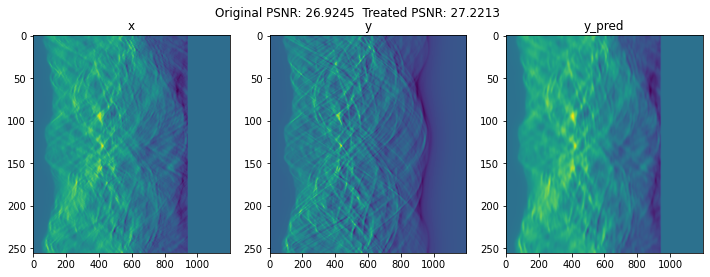
以第七组数据为例（为了与之后机器学习结果保持一致），首先我们将数据可视化。



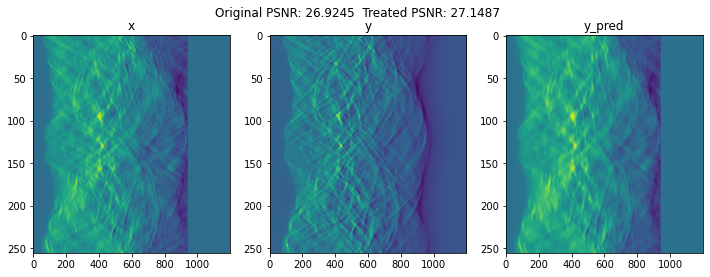
以第一个和第二个通道为例，我们可以看出对于单个通道而言，噪声并没有存在太大的规律性。但是当我们将256个通道作为y轴将整组数据进行可视化后，我们在y轴方向能发现不同通道信号的连续性与关联性。



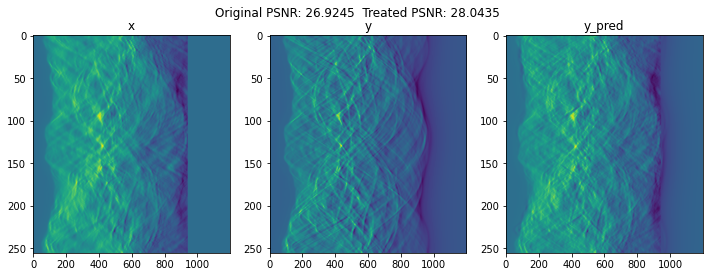
为了让项目更快入手，我们尝试了一些传统的去噪方法。首先是最简单的均值滤波，以5\*5的核为例，在每个数据点周围取平均值来替代原有数据。可以看出，均值滤波对于去噪的效果并不理想，甚至会使得较大值附近的数据点产生更大的偏差。



然后我们尝试了高斯滤波，以11\*11的核为例，对数据点周围的点赋予不同权重来获得新的值，从结果上看等同于对图像进行了模糊处理，这样的滤波方式对于有规律、服从一定分布的高斯噪声有效果，但是对于PA信号的相对无规律噪声效果不明显。



PA信号噪声的特殊性迫使我们放弃了用传统信号处理手段进行降噪，使得我们将视线转向了机器学习。作为前置的试验过程，我们尝试以前六组数据作为已知，从中获取噪声的平均值来作为噪声规律的体现，将这个平均值从第七组数据中减去，发现效果较传统信号处理方法有显著提升。这就证明了虽然单组PA信号的噪声规律性较差，但在多组PA信号中，我们可以找到统一的去噪方法，也就是证明了机器学习的可行性。



Dataset

We are given eight paired human brain photoacoustic sinogram.

每一对sinogram由the sinogram with noise和the corrected sinogram as ground truth组成。

其中，每一幅sinogram中包含256个通道的信号，每一个通道的信号提供了1200个数据点。数据点的取值范围约为-0.2到0.2.

以第七组数据为例（为了与之后机器学习结果保持一致），对256个通道的噪声做快速傅里叶变化与可视化，我们可以看出噪声大多集中在低频和高频，原因在于信号本身也是集中在低频和高频。而且当我们取其中的1~4通道的噪声进行特例分析时，可以得知特定通道的噪声是杂乱无规律的。

宏观上，对比带噪信号与ground truth可知，较大的数据点受噪声影响而变得更大，较小的数据点受噪声影响与相邻数据点差距减小。

补充上，带噪信号的每个通道实际上只有945个数据点，相比ground truth缺少了最后的255个数据点。

Baseline method

均值滤波

Average (or mean) filtering is a method of ‘smoothing’ images by reducing the amount of intensity variation between neighbouring pixels.

The average filter works by moving through the image pixel by pixel, replacing each value with the average value of neighbouring pixels, including itself.

高斯滤波

高斯滤波是一种线性平滑滤波，也就是对整幅图像进行加权平均的过程，每一个像素点的值，都由其本身和邻域内的其他像素值经过加权平均后得到。这个权重由高斯函数得出，理由在于其将噪声的分布视作正态分布，在这种前提条件下高斯滤波才会有较好的表现。

MMN

模拟机器学习的条件，我们把前6对PA信号作为已知条件，从中获取噪声的平均值来作为噪声规律的体现。从而将减去噪声的平均值作为信号处理的算法。

Experiments

首先我们可以看出传统的信号处理去噪的方法效果不佳，无论是直观上看出还是数值体现。这其中的原因在于信号处理的去噪方法对于噪声类型有限制。均值滤波适用于平滑变化的图像中有突变的噪点，高斯滤波适用于满足正态分布的高斯噪声。而PA信号中，数据既是剧烈变化，而且噪声也没有规律可循。所以这两种方法在图像上只是做了模糊处理，甚至在较大值附近对噪声有放大的情况。通过这样的模糊处理，使得PSNR和SSIM有小量提升。

MMN则取得了一定的效果。推断其原因在于本次给定的数据集也是由人工生成，这种方式生成的噪声在不同组之间有一定规律。但是在实际情况中，我们也不难想象，人颅骨的厚薄程度、对超声波的影响在不同人的相同位置也是有规律的。正是这种潜在规律的存在，证明了神经网络的可行性，因为通过大量的学习、运算，神经网络定然能发现这种规律。

DuRN的最佳效果体现在将信号快速变化的边缘清晰化、将偏大的信号值还原，并试图对带噪信号缺失的后255个数据点进行重建，这样子的效果是上述方法完全无法实现的。