Horizontal View Spine MRI Denoise For Medical Judgement

Yicheng An
Stanford University
yicheng@stanford.edu

Chao Wang
Stanford University
cwang17@stanford.edu

Abstract

In this project, we'll explore how to make the diagnosis for prolapse of lumbar inter-vertebral disc disease and draw the MSU classification line given horizontal view spine MRI images. More importantly, we'll explore trying to enhance the machine learning performance by using preprocessing method. These methods include MRI image correction, image enhancement and image denoising. Although machine learning stuff prevails today, computational imaging technique also plays an import role when processing image.

1. Introduction

Computational Imaging is applied to many different areas these days. Magnetic Resonance Imaging (MRI) is a medical analysis of noninvasive nature that uses magnetic fields and radio waves to visualize the internal structure of tissues and organs in the body. One of the most significant advantage of having MRI is that it does not cause any harmful radiation to the patient. The use of MRI as a diagnosing technique continues to grow and have become much faster that it rivals the speed of Computed Tomography. Disease diagnosis by MRI or other medical imaging requires a lot of time, labour and professional knowledge. The visual quality of the MRI plays a very important role during clinical diagnosis.

Nowadays, Lumbar disc herniation is influencing more and more people. The spinal bone bulges out and presses the cord and nerve, which causes the disease and seriously impact people's lives. There is an urgent demand of the diagnosis and treatment for lumbar disc herniation, which is the reason why we do this project. We hope our approaches will ease the workload of radiologists and doctors and increase the treatment efficiency significantly. We aim to design a MRI processing pipeline and apply different computational imaging technologies to improve the image quality, which could help the doctors to diagnose, improve the efficiency of the process in hospitals and increase the accuracy of the auto-diagnosis system.

2. Related work

Non-Local Means (NLM) is analyzed and adapted to reduce this noise in MR magnitude images. NLM is highly dependent on the setting of its parameters. People try to find the optimal parameter selection for MR magnitude image denoising. [3] Also, the filter has been adapted to fit with specific characteristics of the noise in MR image magnitude images (i.e. Rician noise). From the results over synthetic and real images we can conclude that this filter can be successfully used for automatic MR denoising.

Varghees, Manikandan and Gini [6] present an automated, adaptive image denoising method for removal of Rician noise from MRI images. The proposed method is based on the discretized total variation (TV) minimization model and the local noise estimation technique. The regularization parameter of the TV-based denoising method is adapted based on the standard deviation of noise in MRI image.

Denoising is an important step that needs to be taken before the image is analysed for diagnosis. [5] The images acquired through MRI scans can adversely affect quantitative image analysis due to the noise hindrances. Denoising is challenging because denosing might cause artifacts and blur images. In order to fix this problem, different filtering methods are used for the denoising of MRI images. Median, Gaussian and Gabor have been discussed based on PSNR, Energy, Entropy, Variance, Correlation and Contrast. And the appropriate filtering method is identified for the image correction to bring in the evaluation strategies and application.

An auto-diagnosis system could be helpful in clinical applications for disc herniation. [2] A computer-aided diagnosis framework for lumbar spine with a two-level classification scheme for disc herniation diagnosis was developed using heterogeneous classifiers: a perceptron classifier, a least mean square classifier, a SVM classifier, and a k-Means classifier. The computer-aided framework works well to diagnose herniated discs in MRI scans.

3. Problem statement

Currently we are authorized by some Chinese hospitals to use their private vertical-view spine lumbar dataset to help them build up a machine learning based diagnosis system. Since doctors and radiologists spend a lot of time every day to watch medical images like MRI or CT, and every one of those images could take minutes to hours to give a final conclusion on whether there is a disease area or not. So far you could guess the potential value of such a system. It could significantly decrease the workload of doctors and radiologists, and could even increase the diagnose precision because computer never gets tired.

Out main task is to build a model to 1.decide whether a given lumbar MRI image has disease in it. 2.draw the main vertical line of medical MSU standard in the spinal cord section. And also make the decision/drawing precision as high as possible so that it has real value for doctors instead of becoming a high-tech toy.

3.1. Disease detail

As we all know lumbar inter-vertebral disc herniation is a very common disease, the vertical view of your spine is shown in figure 1.

Figure 1. Vertical view of spine, MRI scans from left to right

As we could see the disc between spinal bones bulges out and compresses the nerve, that's the reason why people having such disease could feel pains. Then let's see an example of horizontal view in figure 2, in which MRI scans from up to down. We will mainly deal with such horizontal view images in the project.

In the horizontal view image the center circular object is a single bone or disc in the vertical image, the the 'V' shape white object below the bone is spinal cord. In figure 2 we could see the disc(circular black object) compresses the spinal cord, causing the disease. The red bounding box has circled out the disease area.

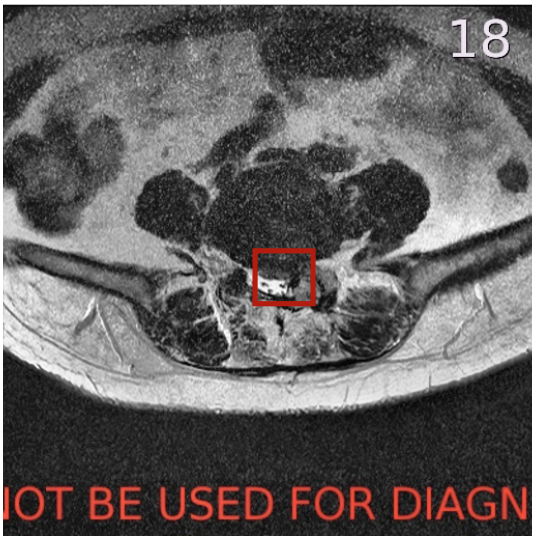

Figure 2. Horizontal view of spine, MRI scans from up to down

3.2. MSU standard

The MSU Classification is a simple and reliable method to objectively measure herniated lumbar disc.-[4] It simply classifies herniation size as 1-2-3 and location as A-B-C, with inter-examiner reliability of 98%. The most frequent types of herniation selected for surgery in each series were types 2-B and 2-AB, suggesting the combined importance of both size and location. When used in correlation with appropriate clinical findings, the MSU Classification can provide objective criteria for surgery that may lead to a higher percentage of good to excellent outcomes. The results show that MSU classification has high reliability for disease diagnosis based on MRI.

The MSU Classification takes into account both the size of disc herniation and its location within the various constraints posed by the local anatomy. It employs a single intra-facet line as a reference point to measure the disc herniation at the level of maximum extrusion, where the most impact on neurologic structures likely occurs.

To further qualify location of the disc herniation, the lesion is described as A, B, or C to more exactly locate the position that is routinely, but less accurately, reported as central, lateral or far lateral. Three points are placed along the intra-facet line, dividing it into four equal quarters (Fig. 3). We draw perpendicular lines through each of these points to create the right and left lateral and right and left central quadrants. Zone-A is represented by the right and left central quadrants. Zone-B is represented by the right and left lateral quadrants. Zone-C is represented at the level of the

foramen by the area that extends beyond the medial margin of either facet joint, past the borderline of the lateral quadrants. When a size-2 disc herniation occurs within either of the two more constrained zones-B and zone-C, it typically exerts a greater impact on the spinal nerves.

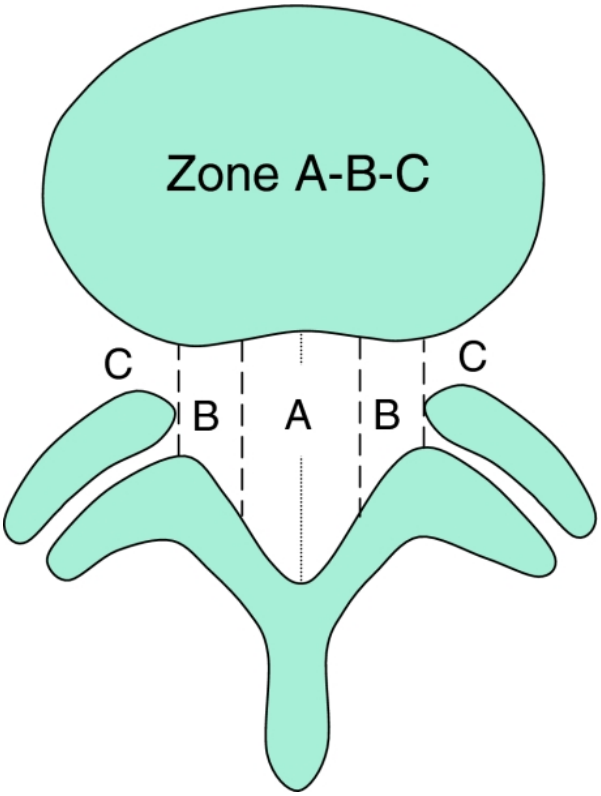

Figure 3. Zoning the disc for location. Lesions have more impact in tighter zone-B and -C

After applying MSU to our data, the result is shown as follow.

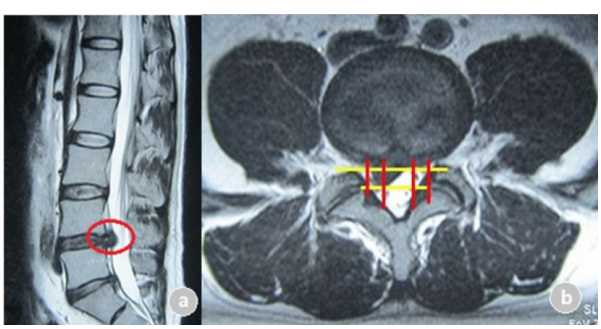

Figure 4. Zoning the disc for location. Lesions have more impact in tighter zone-B and -C

4. Diagnosis model

Based on the problem statement, we plan to divide the whole task into several sub-tasks. The processing pipeline is shown below in figure 4.

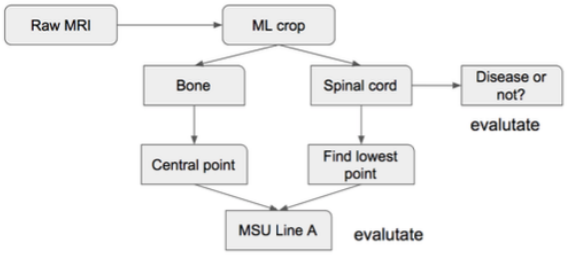

Figure 5. Machine learning decision pipeline

For disease diagnosis purpose, since the disease only happens in the spinal cord area, we want to first crop out spinal cord to ignore those useless parts in MRI images.

And for drawing main MSU vertical line, we chose the end point to be 1.center of circular bone 2. lowest optimal point in spinal cord. And the visual effect is shown in figure 6. As we could see we need both bone section and spinal cord section to extract end points and draw the line.

Combining both disease diagnosis purpose and MSU purpose, we need to crop out bone area and spinal cord area at the same time, using machine learning method.

After getting these 2 crops, we find end points in them and draw the line. Then train a binary classifier on spinal cord crop section to conclude whether or not it has disease.

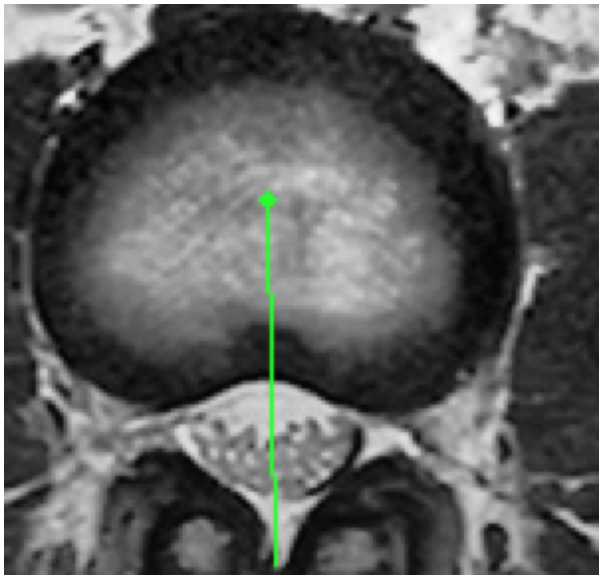

Figure 6. Visual effect of A-line and corresponding end points

4.1. Machine learning crop

In the machine learning crop process, we modified YOLO algorithm to simultaneously localize and identify the circular bone and white spinal cord. The modified YOLO network is shown in figure 7. You can see the changes of

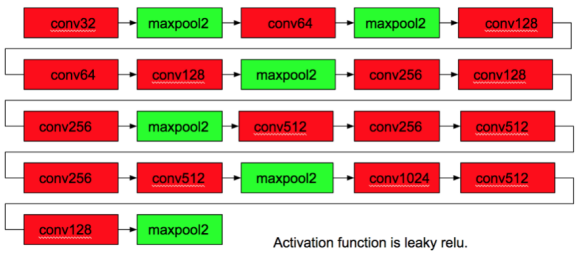

Figure 7. Modified YOLO network

number of filters are pretty gentle, which is pretty reasonable.

And next modification is in the loss function, the original YOLO penalized the square root of the width and height of the bounding boxes,the loss function is shown in figure 8. But we deleted the square root option so that the bounding box could be bigger in order to include more details in spinal cord section. And also since we are going to use the center of bounding box to represent the mass center of circular bone, bigger bounding box could increase the precision for this.

$$
\begin{aligned}
&\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(x_i - \hat{x}_i\right)^2 + \left(y_i - \hat{y}_i\right)^2 \\
&+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(\sqrt{w_i} - \sqrt{\hat{w}_i}\right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i}\right)^2 \\
&+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(C_i - \hat{C}_i\right)^2 \\
&+ \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i\right)^2 \\
&+ \sum_{i=0}^{S^2} \mathbb{1}_{i}^{\text{obj}} \sum_{c \in \text{classes}} \left(p_i(c) - \hat{p}_i(c)\right)^2
\end{aligned}
$$

Figure 8. YOLO loss function

4.2. Binary classifier

After we get the crop of spinal cord section, we could train a binary classifier according to the image label. We used ImageNet's VGG16 as pre-trained model and freeze its first 10 layers, only train last several layers, and add another fully connected layer at last to do the classification work.

4.3. Find end points and draw the line

After using machine learning to get the crops, we use the center of bone's bounding box as bone's mass center, which is the upper end point of A-line. Then for the spinal cord, we thresholds the cropped spinal cord first and apply contour detection to find contours of main spinal cord. Then find the minimum tip point of 'V' shape cord, which is another end point of the A-line in MSU classification.

5. New model with computational image preprocessing

After finishing the above diagnosis model, we then decided to add preprocessing block in order to further improve the total performance. Because on the one hand conventionally MRI images need preprocessing which is mentioned in related work section. On the other hand our MRI data images don't have that high resolution, but both disease detection and line(end points) drawing require lots of detail information, and we need to use preprocessing method to enhance or separate these detail information to let machine learning learn better.

So adding preprocessing method is pretty necessary and could possibly be helpful for improving the accuracy of diagnosis model.

5.1. Overview of the new model

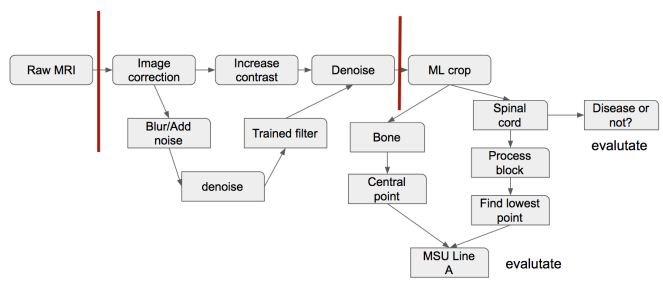

Figure 9. New model with preprocessing block

As is shown in figure 9, we add several processing steps in the block, the following subsection will discuss it in detail.

5.2. Image correction

5.2.1 Bias correction

So we develop a bias correction function to handles MRI bias and artifact corrections. Bias field signal is a low-frequency and very smooth signal that corrupts MRI images. [1] Because of the inhomogeneities in the magnetic fields of the MRI machine, there is often a bias field. Images

are blurred by a bias field and high frequency components such as edges and contours are reduced. We need a preprocessing step to correct for the effect of bias field before further processing. We try different approaches to correct the bias field such as N4 bias correction, N3 bias correction, estimating bias fields by parametric surface fitting and an iterative removal of bias field based on modifying the objective function. For example, the formula of N4 bias correction is as follow.

$$u^n = u^{n-1} - S^*\{u^{n-1} - E[u|u^{n-1}]\} \tag{1}$$

where S{} is a different B-spline approximator.

5.2.2 Intensity standardization

MRI intensities do not have a fixed meaning, even for the same protocol, body region, patient, scanner. It causes some problems for diagnosis and analysis. Since simple linear scaling does not help we need to develop a non-linear scaling method. Intensity mapping function is shown as follow.

$$\tau(x) = \begin{cases} \lceil \mu_s + (x-\mu_i)\frac{s_1-\mu_s}{p_{1i}-\mu_i} \rceil, if m_{1i} \le x \le \mu_i \\ lceil \mu_s + (x-\mu_i)\frac{s_2-\mu_s}{p_{2i}-\mu_i} \rceil, if \mu_i \le x \le m_{2i} \end{cases} \tag{2}$$

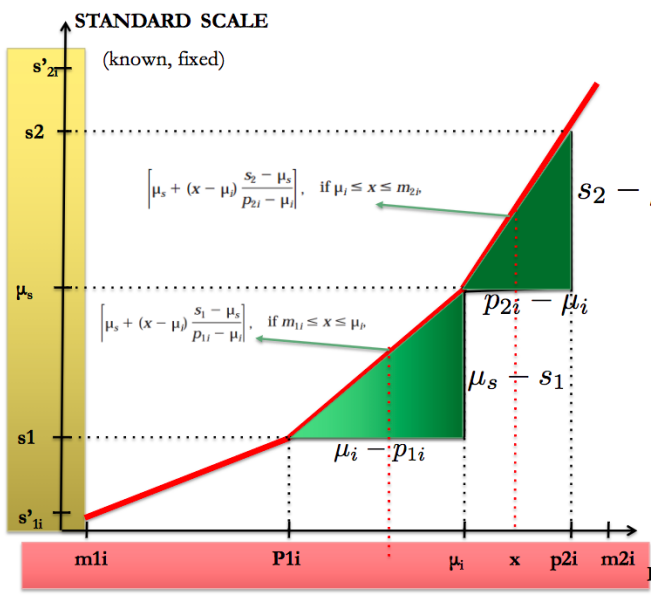

Figure 10. Image Scale

5.2.3 Bias correction

5.3. Increase contrast

Since the disease area is located at the boarder of black bone and white spinal cord, and the corrected MRI image was not very clear in detail because of a low resolution, we need to increase the contrast to enhance the boarder's visual effect, otherwise if human couldn't see it clearly we could expect machine learning to do better.

So we used 2 methods to increase the contrast, the first one we used is Histograms Equalization in YUV channel. And apply histogram equalized operation on Y(Luma) channel, which is a global histogram equalization method.

The second method we used is Contrast Limited Adaptive Histogram Equalization(CLAHE). We apply CLAHE in LAB color space, image is divided into small blocks called "tiles". Then each of these blocks are histogram equalized as usual. So in a small area, histogram would confine to a small region (unless there is noise). If noise is there, it will be amplified. To avoid this, contrast limiting is applied. If any histogram bin is above the specified contrast limit, those pixels are clipped and distributed uniformly to other bins before applying histogram equalization.

Theoretically the second method is better because 1.every histogram is confined to a certain region, which means related areas are equalized instead of everything is equalized together. 2. It suppresses noise during equalization process. And in final model we chose the second method. The comparison result result image is shown in figure 11 and figure 12.

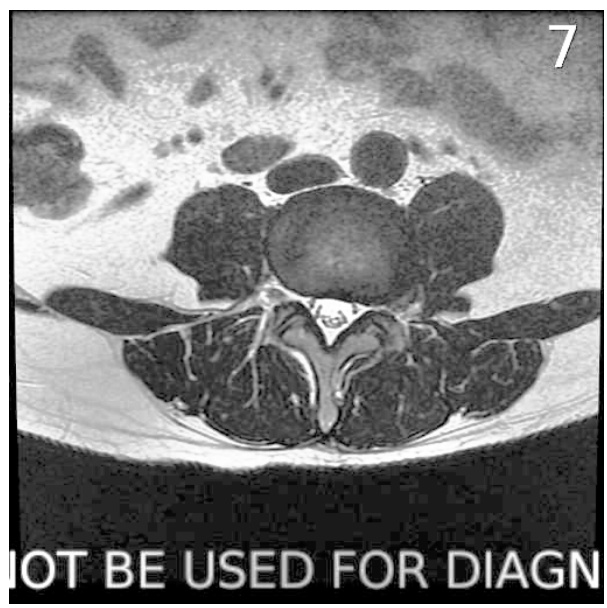

Figure 11. Contrast increased image with method 1

5.4. Simulate noise condition and train filter

After increasing the contrast, we get lots of Gaussian white noise in the image, as we can see in figure 8. So we need to further use filters to remove them. Since noise mainly comes from increasing contrast, we could regard them as Gaussian white noise.

Figure 12. Contrast increased image with method 2

In order to design the appropriate filter to eliminate these Gaussian noise, we added Gaussian noise to the corrected images to "simulate" the production of noise during contrast increase process. We cant directly use high contrast images to get filters since PSNR between corrected images and high contrast images is meaningless. And if our filter works well on the artificial noisy image, we could assume they also work well on contrast increased images. Thus we could use PSNR here as an evaluation metric for our designed filter.

We have tried several filters including average filter, bilateral filter, Gaussian filter and a combination filter(median + Wiener + bilateral). And we need to choose the one with best performance(highest PSNR) in the final model. The combination filter wins at last.

5.5. Process block for spinal cord lowest tip detection

Based on our initial results, the contour detection on spinal cord didn't perform very well, since the spinal cord surface has different depths and texture, the whole spinal cord could be segmented into several pieces of contours, figure 13 shows an example. Therefore we need some method to "blur" the image on some degree to make spinal cord as one whole piece of contour. At the same time, we need to control the blur degree to prevent some other pieces of small contours merge into the spinal cord's contour. Based on our here we used non-local-means + bilateral filter to do the blur job. The corresponding result image is shown in figure 14.

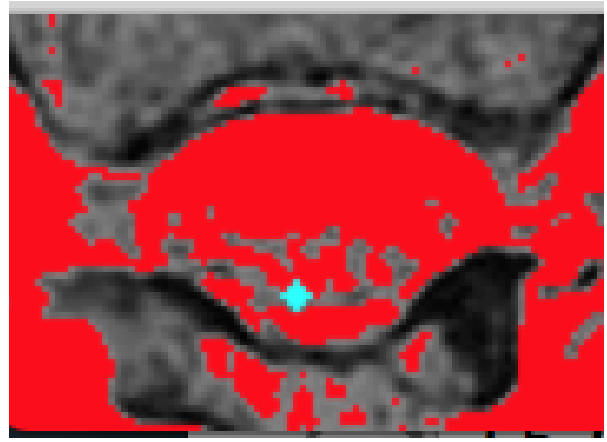

Figure 13. Wrong lowest point detection

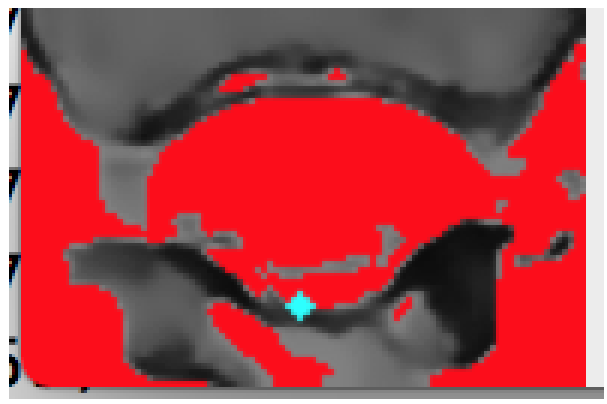

Figure 14. Wrong lowest point detection

6. Results

The results and decision in image correction, contrast increasing method and spinal cord lowest tip detection has been discussed above We got them mainly by theoretical induction and experiment prove. But the filter choice cannot be dealt with totally by theory. The following table shows the performance of potential filters, evaluated by PSNR.

Filters	PSNR(average on 100 images)
gaussian-2D	23.31
average	24.10
bilateral	14.18
Median+wiene+bilateral	24.36

Figure 15. Table of average PSNR for different filters over 100 images

We could see the combination filter achieves maximum PSNR, but we still need to choose the best K value for wiener filter, we sampled at 0.001, 0.01, 0.1, 1, 2 and get the following plot versus PSNR. Therefore the best performance happens at K = 0.1.

And once K is determined, we get the whole new model complete. We need to compare the old draw-line accuracy and diagnosis accuracy with the new draw-line accuracy and diagnosis accuracy. According to our test, the final results are shown in the following table. As we could see after we add the preprocessing pipeline, both line localization accuracy and diagnosis accuracy increases, especially line localization accuracy increases significantly, which implies line drawing depends more on preprocessing pipeline than disease diagnosis.

	Original images	Processed images
Line localization accuracy	90%	96%

	Original images	Processed images
Diagnosis accuracy	82%	84%

Figure 16. Old results vs. New results

7. Analysis and evaluation

The result mentioned above makes sense theoretically. For filter selection, median filter deals with Gaussian noise pretty well as we all know. And also according to our homework Wiener filter works better on sparsity background, which fits the condition of our MRI images. Finally bilateral filter preserves edges, which works well since disease happens at the boarder area of bone and spinal cord.

For diagnosis, since the preprocessing pipeline enhances the general quality of MRI images, like making boarder area more obvious and suppressing the noise. The classification accuracy will go up but won't be too much.

For line localization, line is determined by 2 end points, the upper one of them is determined using machine learning method, the other downside point is determined by computational imaging method, so our image processing would boost the downside point detection pretty reasonably. As a result the whole line's accuracy will be increased. (we measure the line accuracy by sending them to doctors and let them judge whether it's correct or not)

8. Discussion

Compared to end-to-end machine learning method, the performance of the algorithm has a noticeable improvement after the MRI preprocessing, such as denoising, image enhancement and bias correction. We should apply the knowledge from the course and use appropriate algorithm based on the property of our dataset. We can see that computational imaging approaches can significantly improve the image quality and help the machine methods to achieve a better result. In the future, we will try more advanced methods to process the data and improve the performance.

References

[1] J. Juntu. Bias field correction for mri images, 2005.

[2] V. C. Koh, Jaehan and G. Dhillon. Disc herniation diagnosis in mri using a cad framework and a two-level classifier, 2012.

[3] L. J. J. Manjon J V, Carbonell-Caballero J. Mri denoising using non-local means. *Medical image analysis*, 12(4):514–523, 2008.

[4] L. W. Mysliwiec. Msu classification for herniated lumbar discs on mri: toward developing objective criteria for surgical selection, 2010.

[5] J. Seetha and S. S. Raja. Denoising of mri images using filtering methods, 2016.

[6] G. R. Varghees V N, Manikandan M S. Adaptive mri image denoising using total-variation and local noise estimation. *Advances in Engineering*, 14(1):506–511, 2012.