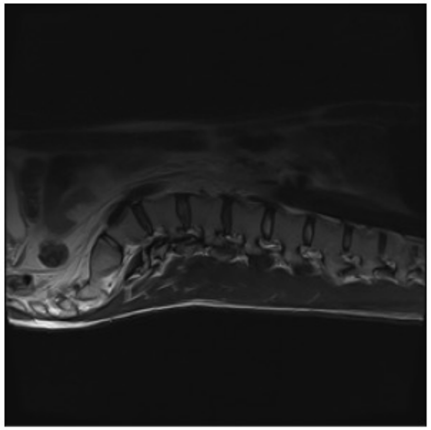
Application of Hybrid Network of UNet and Feature Pyramid Network in Spine Segmentation

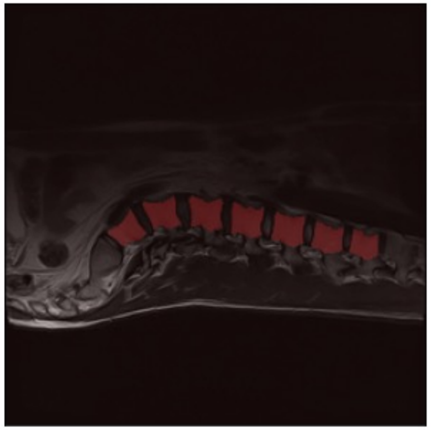
Presenter: Xingxing Liu

Authors: Xingxing Liu, Wenxiang Deng, Yang Liu

Background ---- Spine Segmentation

- Spine diseases and injuries has afflicted people, especially the elders, around the world and caused huge financial burdens to the patients.
- Spine segmentation can provide useful information for the diagnosis of spine diseases and injuries and help doctors to determine the plan for medical treatment.
- However, spine segmentation is usually labor-intensive in the past.



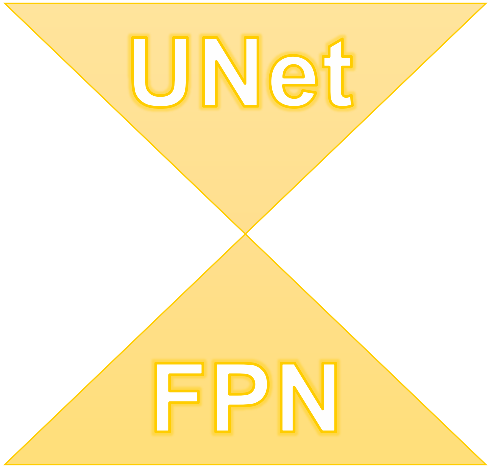


Background ---- UNet & FPN

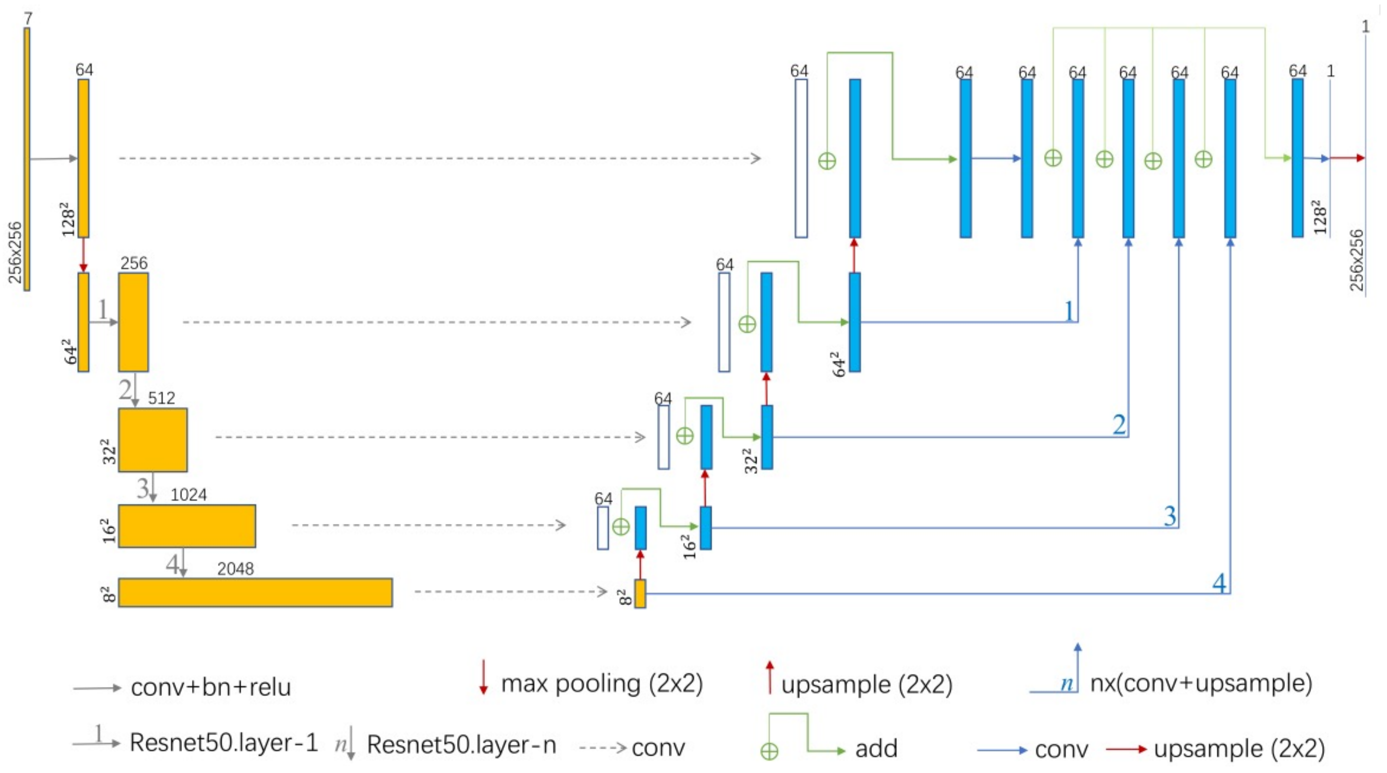
- Deep learning-based image segmentation develop rapidly in recent years. Many deep learning-based image segmentation algorithms outperform traditional algorithms.
- UNet is a well-known deep network, and it has been widely used for medical image segmentation, due to its symmetrical encoder-decoder structure, it has achieved great segmentation performance. Many UNet-based networks have been developed for different applications.
- Feature pyramid network (FPN) has the ability of extracting features from images with multiple scales, so it's also widely used for image segmentation.

Motivation

- Based on these background knowledge, we want to develop an effective deep learning-based spine segmentation algorithm, the output of which can help doctors diagnose spine diseases and injuries.
- A network which incorporates UNet and FPN can not only preserve high-resolution image features, but also combine feature maps of multiple scales to achieve better segmentation performance .



Methodology ---- Network Structure



Res50-UNet

Methodology ---- Network Structure

- The left part of Res50_UNet has an overall “U” structure. However, blocks of resnet replace the original convolution blocks on the contracting path and feature map addition replaces the original concatenation on the expansive path.
- FPN synthesizes all the output feature maps on the expansive path to generate the final segmentation mask.
- Segmentation mask of the same size with the input image is generated as the final output.

Methodology ---- Dataset

- A dataset containing spine MRI slices of 23 patients was used to train the network and test the trained modules. The spine MRIs of each patient contain at least 7 vertebral bodies of the lower spine (T11 – L5).
- The spine MRIs of each patient has 39 slices, so there are 897 MRI slices in the dataset.
- The training set contains the spine MRIs of 19 patients, and the spine MRIs of the left 4 patients were used for testing the trained modules. So, the size ratio of the training set to the testing set is 19:4.

Methodology ---- Training

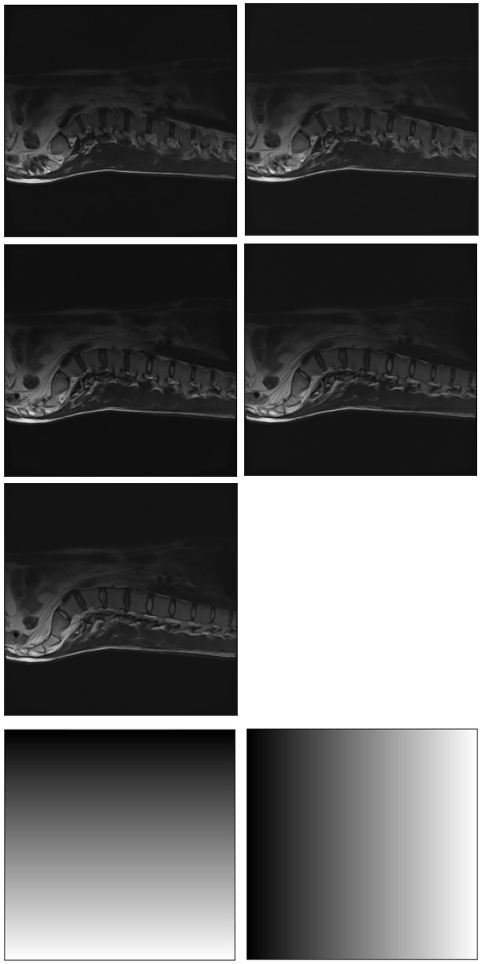
❖ Training hyper parameters:

- Learning rate: 0.005
- Optimizer: Adam
- Batch size: 16

❖ Loss:

$$-\ln\left(\frac{2|(GT \cap SR) + Smooth|}{|GT| + |SR| + Smooth}\right)$$

❖ The input contains 7 slices, including 5 consecutive spine MRIs and 2 coordinate maps.



Methodology ---- Evaluation

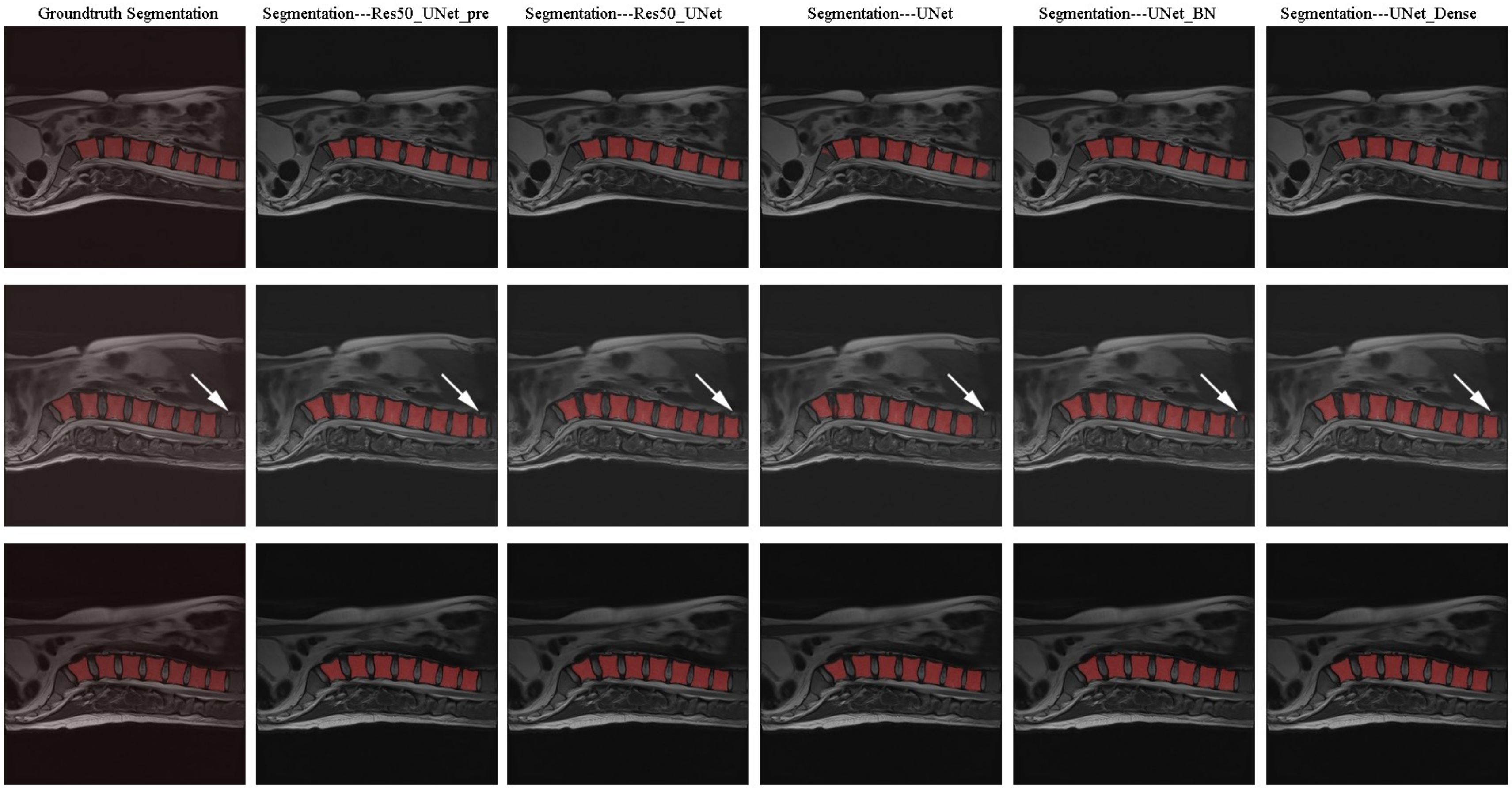
- For evaluation and comparison of the segmentation performance of Res50_UNet, we have also implemented original UNet, UNet_BN and UNet_Dense. They are trained on the same dataset, and their segmentation performance are evaluated qualitatively and quantitatively.
- We found that using pretrained blocks of resnet can improve the final segmentation results, so we also evaluate the performance of Res50_UNet with pretrained resnet blocks.
- For qualitative performance comparison, we randomly selected 3 different MRI slices from 3 patients and display the segmentation mask generated by different trained networks.

Methodology ---- Evaluation

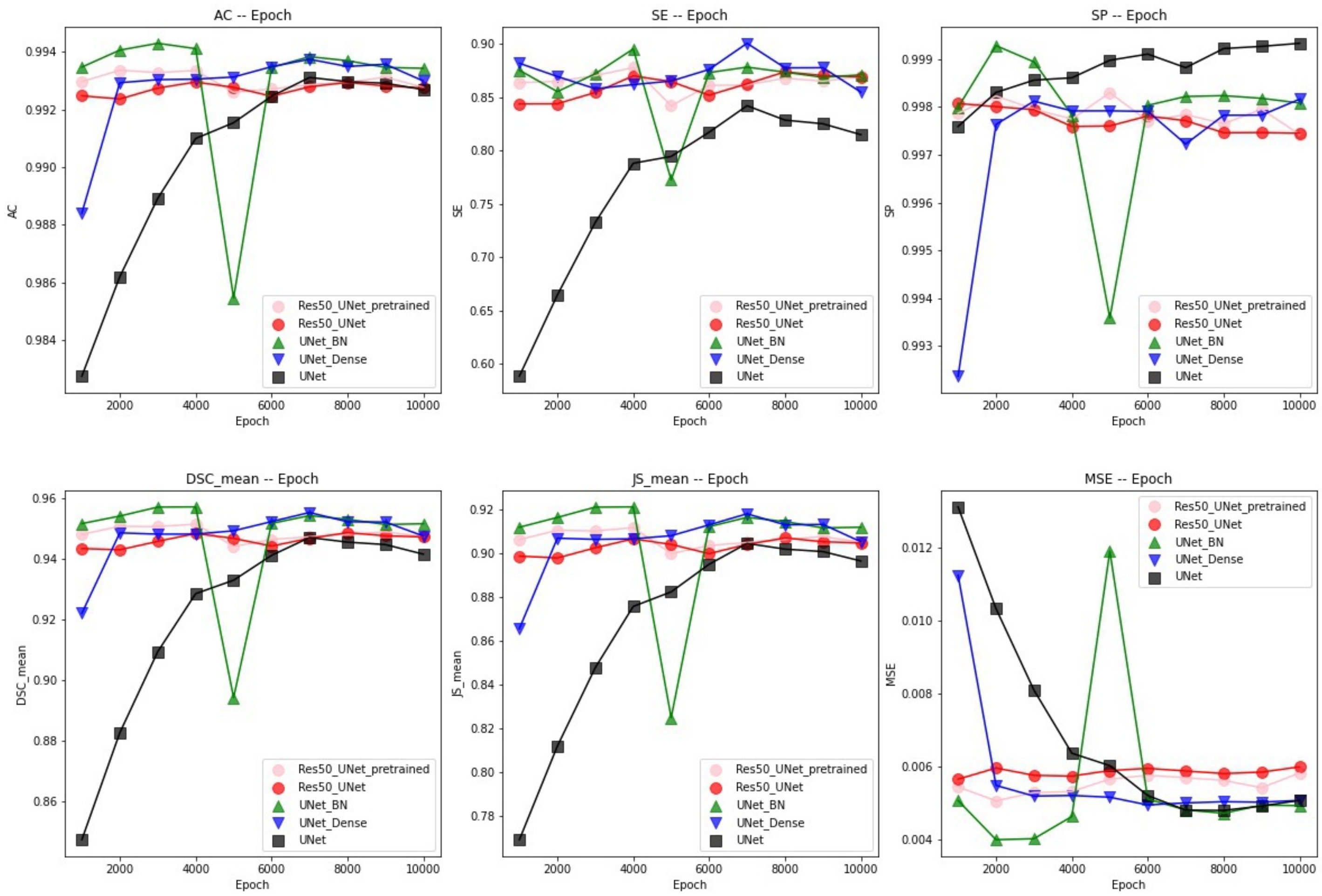
❖ We use 6 different metrics to quantitatively evaluate the segmentation performance of different networks. Here are the definitions of 6 metrics:

- $accuracy\ (AC) = \frac{TP+TN}{TP+TN+FP+FN}$
- $sensitivity\ (SE) = \frac{TP}{TP+FN}$
- $specificity\ (SP) = \frac{TN}{TN+FP}$
- $Dice\ similarity\ coefficient\ (DSC) = \frac{2|GT \cap SR|}{|GT| + |SR|}$
- $Jaccard\ similarity\ (JS) = \frac{|GT \cap SR|}{|GT \cup SR|}$
- $mean\ square\ error (MSE) = \frac{1}{N}\sum_{i=1}^{N}(GT_i - SR_i)^2$

Result ---- Qualitative Evaluation



Result ---- Quantitative Evaluation



Discussion

- ❖ Res50_UNet has achieved accurate spine segmentation performance.
- ❖ The accuracy (AC) of Res50_UNet is higher than 99.5% with only 1000 epochs, which showed the potential for Res50_UNet in spine MRI segmentation when a low number of epochs is desirable.
- ❖ Pretrained Res50_Unet outperforms non-pretrained Res50_Unet.

Conclusion & Future Prospect

- ❖ This study has demonstrated the feasibility of applying Res50_UNet in spine segmentation. The network integrates the characteristics of FPN and UNet. High accuracy was achieved with a low number of epochs. These results have shown the potential for Res50_UNet in spine MRI segmentation, especially when a low number of epochs is desirable.
- ❖ In the future, we plan to optimize the network structure further and investigate the application of Res50_UNet in other domains of spine segmentation.

Thanks.