The problem: Given Anteroposterior(AP) and Lateral(LAT) views of the spinal cord (X-ray images), we have to label the given spine as being ‘Damaged’ or ‘Normal’.

The training data consists of Anteroposterior(AP) and Lateral(LAT) views along with the corresponding Segmentation masks (Ap\_Pedicle, Ap\_Spinous\_Process, Ap\_Vertebra for AP view, and Lat\_Anterior\_Vertebral\_Line, Lat\_Disk\_Height, Lat\_Posterior\_Vertebral\_Line, Lat\_Spinous\_Process, Lat\_Vertebra for the LAT view). The AP and LAT view images are 3-channel RGB images (.jpg) and the corresponding masks are one-channnel binary images (.png). The dataset has 328 training examples for the class ‘Damaged’, and 350 training examples for ‘Normal’.

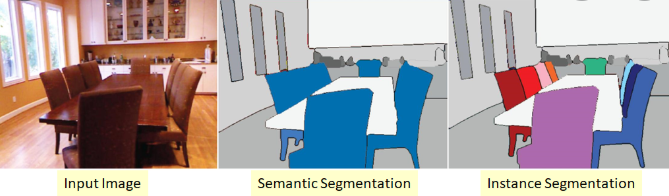
The proposed methodology is to use Image Segmentation techniques to generate segmentation masks for the given test images, and then use those masks for the classification task.

Image segmentation is a pixel-level classification task. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. It makes analysis and further processing of the image easy.

There are two types of segmentation tasks: Semantic Segmentation and Instance Segmentation.

In semantic segmentation, the different objects belonging to the same class are considered the same entity, and hence, assigned the same label.

Instance segmentation is more robust, and is able to identify different objects belonging to the same class as differently.



*(*[*https://datascience.stackexchange.com/questions/52015/what-is-the-difference-between-semantic-segmentation-object-detection-and-insta*](https://datascience.stackexchange.com/questions/52015/what-is-the-difference-between-semantic-segmentation-object-detection-and-insta)*)*

The given problem is an instance segmentation problem, where corresponding to each Anteroposterior(AP) and Lateral(LAT) view, we have to segment out the relevant masks.

We have used UNET for the purpose.

The [UNET](https://arxiv.org/abs/1505.04597)was developed by Olaf Ronneberger et al. for Bio Medical Image Segmentation. The architecture contains two paths. First path is the contraction path (also called as the encoder) which is used to capture the context in the image. The encoder is just a traditional stack of convolutional and max pooling layers. The second path is the symmetric expanding path (also called as the decoder) which is used to enable precise localization using transposed convolutions. Thus it is an end-to-end fully convolutional network (FCN), i.e. it only contains Convolutional layers and does not contain any Dense layer because of which it can accept image of any size.

*(https://towardsdatascience.com/understanding-semantic-segmentation-with-unet-6be4f42d4b47)*

The architecture of the UNET is as follows:

