FULLY CONVOLUTIONAL NEURAL NETWORK FOR VOLUMETRIC MEDICAL IMAGE SEGMENTATION

VIVA QUESTIONS

Group members:
MAHAK BAFNA
RIYA RALHAN
MOHANA SIDDHARTHA CHIVUKULA

1- What is this paper about ?

A1- This paper is purely based on CNN and we propose an approach to 3D image segmentation based on a volumetric, fully convolutional neural network. We presented an approach that performs segmentation on MRI prostate volume in a fast and accurate manner.

2- What CNN?

A2- CNN is a type of artificial neural network used in image recognition and processing. It is specially designed to process pixel data. CNN uses the concept of deep learning.

3- Why CNN?

A3. We use CNN and not any other type of neural networks due to the following advantages:

- CNN learns the filters automatically without mentioning it explicitly. These filters help in extracting the right and relevant features from the input data
- CNN captures the spatial features from an image. Spatial features refer to the arrangement of pixels and the relationship between them in an image. They help us in identifying the object accurately, the location of an object, as well as its relation with other objects in an image
- It is also includes the concept of parameter sharing

3. What is convolution?

4. What was the result ?

A4-

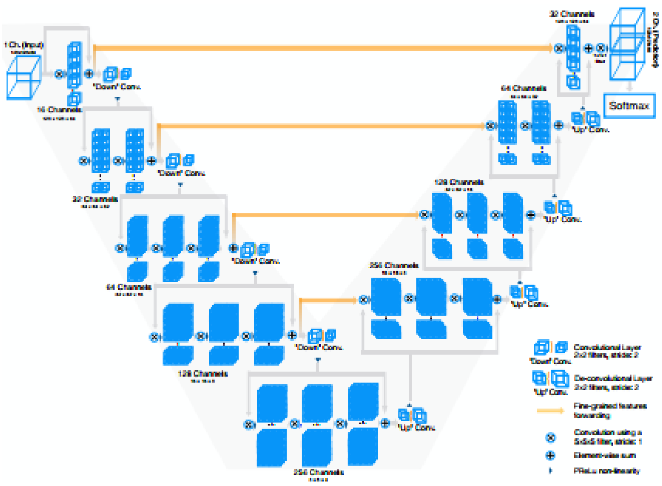
5. What is the dice coefficient?

6. Method explanation.

A6. The left side of the network is divided in different stages that operate at different resolutions. Each stage comprises one to three convolutional layers.we formulate each stage such that it learns a residual function: the input of each stage is (a) used in the convolutional layers and processed through the non-linearities and (b) added to the output of the last convolutional layer of that stage in order to enable learning a residual function. As confirmed by our empirical observations, this architecture ensures convergence in a fraction of the time required by a similar network that does not learn residual functions. The convolutions performed in each stage use volumetric kernels having size 5×5×5 voxels. As the data proceeds through different stages along the compression path, its resolution is reduced. This is performed through convolution with 2 × 2 × 2 voxels wide kernels applied with stride 2. Since the second operation extracts features by considering only non overlapping 2×2×2 volume patches, the size of the resulting feature maps is halved. This strategy serves a similar purpose as pooling layers that, other works discouraging the use of max-pooling operations in CNNs, have been replaced in our approach by convolutional ones. Moreover, since the number of feature channels doubles at each stage of the compression path of the V-Net, and due to the formulation of the model as a residual network, we resort to these convolution operations to double the number of feature maps as we reduce their resolution

7. Diagram of method

A7-



8. What is kernel, stride, padding and pooling ?

A8

classmate
Date
Page

Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation

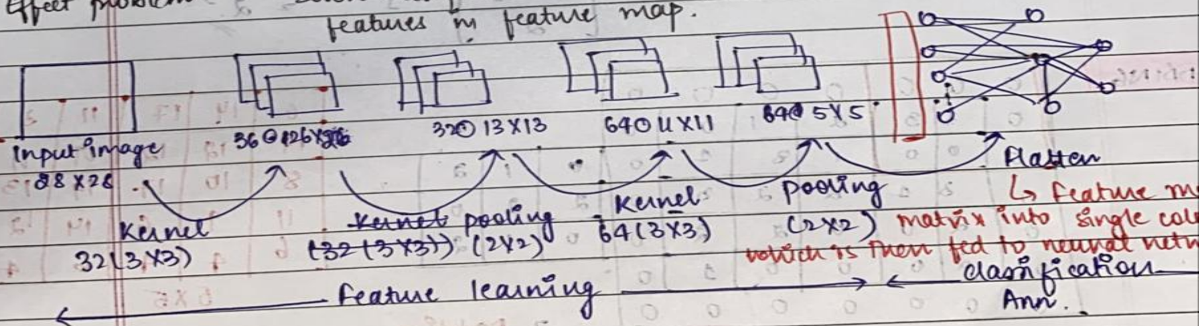
CNN is a type of artifical neural network used in image recognition and processing, specially designed to process pixel data. CNN uses the concept of deep learning.

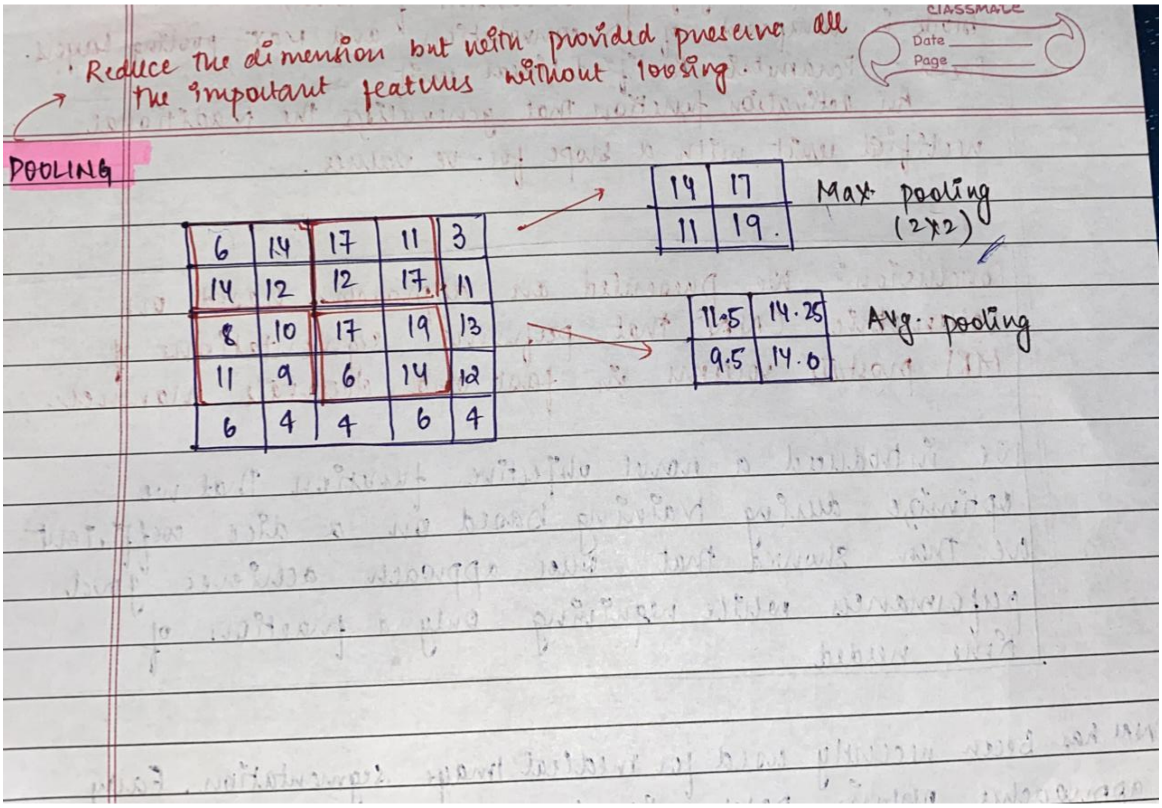
This paper is purely based on CNN and we propose an approach to 3D image segmentation based on volumetric, fully convolutional neural network.

Dice coefficient: 2 * the area of overlap divided by the total number of pixels in both the images.

Convolutional Neural Network

1. Kernel = filter = Feature detector : Extract features from images. It is a matrix that moves over the input data, performs dot product and gets output. It moves on input data by stride value.
2. Stride → The amount of movement b/w applications of filter to the input image is refered to as stride.
3. padding → To fix the border Effect problem.
4. Pooling → Down sample the detection of features in feature map.
5. Flatten





9. What is dice loss layer?

A9-Dice pixel classification layer provides a categorical label for each image pixel or voxel using generalised Dice loss. The layer uses generalised Dice loss to alleviate the problem of class imbalance in semantic segmentation problems.

10. Formula explanation

A10

In this work we propose a novel objective function based on dice coefficient, which is a quantity ranging between 0 and 1 which we aim to maximise. The dice coefficient D between two binary volumes can be written as

$$D = \frac{2\sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}$$

where the sums run over the N voxels, of the predicted binary segmentation volume pi ∈ P and the ground truth binary volume gi ∈ G. This formulation of Dice can be differentiated yielding the gradient.

$$\frac{\partial D}{\partial p_j} = 2\left[\frac{g_j\left(\sum_i^N p_i^2 + \sum_i^N g_i^2\right) - 2p_j\left(\sum_i^N p_i g_i\right)}{\left(\sum_i^N p_i^2 + \sum_i^N g_i^2\right)^2}\right]$$

.

11. Training involved

A11. During every training iteration, we fed as input to the network randomly deformed versions of the training images by using a dense deformation field obtained through a 2 × 2 × 2 grid of control-points and B-spline interpolation.This augmentation has been performed "on-the-fly", prior to each optimisation iteration, in order to alleviate the otherwise excessive storage requirements.Additionally we vary the intensity distribution of the data by adapting, using histogram matching, the intensity distributions of the training volumes used in each iteration, to the ones of other randomly chosen scans belonging to the dataset.