Chapter 7
Deep Learning Models for Semantic Multi-Modal Medical Image Segmentation

V. R. S. Mani
https://orcid.org/0000-0002-3234-0486
National Engineering College, Kovilpatti, India

ABSTRACT

In this chapter, the author paints a comprehensive picture of different deep learning models used in different multi-modal image segmentation tasks. This chapter is an introduction for those new to the field, an overview for those working in the field, and a reference for those searching for literature on a specific application. Methods are classified according to the different types of multi-modal images and the corresponding types of convolution neural networks used in the segmentation task. The chapter starts with an introduction to CNN topology and describes various models like Hyper Dense Net, Organ Attention Net, UNet, VNet, Dilated Fully Convolutional Network, Transfer Learning, etc.

INTRODUCTION

A deep neural network has extremely large number of layers of neurons, forming a hierarchical feature representation. The number of layers now rises to over 2,000. With this extremely large modeling capacity, a deep network can effectively remember all possible transformations after good training with an extremely large training data set and make smart predictions, they identify from new data sets which were untrained earlier. Thus, deep learning is used in many computer vision and medical imaging tasks. Convolutional Neural Network (CNN) models like AlexNet, VGGNet, RESnet or GoogLeNet have shown accurate results in various competitive bench marks. Different ideas were explored regarding CNN design, like use of different activation functions, loss functions, and optimization and regularization functions. Approximately there are seven different categories of CNN architectures which were used widely in different applications. All these architectures are based on depth, width, spatial exploitation,

DOI: 10.4018/978-1-6684-7544-7.ch007

multi path, attention, channel boosting and feature map exploitation. In CNN, the topology is divided into different learning stages like convolutional layer, non-linear processing units, and sub sampling layers. Multiple transformations are done in each layer using a group of convolutional kernels. Locally correlated features are extracted in the convolutional operation.

Semantic segmentation explains how each pixel of an image is labelled with a class label, (such as water, sand, grass, sky, ocean, or car). Semantic segmentation applications include: Autonomous driving, Robotic Vision, Remote Sensing, Medical Imaging, Video Processing, etc., Due to the tremendous advancements in multimodal image acquisition devices, the amount of imaging data to be analyzed is increasing day by day. Medical imaging data is heterogeneous and multidimensional and also sensitive to human error and varies across different subjects. The segmentation and interpretation of medical imaging data is therefore a time consuming process. Semantic segmentation relates each pixel of an image to a class label. Though closely related to semantic segmentation, medical image segmentation includes specific challenges that need to be addressed, such as the scarcity of labeled data, the high class imbalance found in the ground truth and the high memory demand of three-dimensional images. In this chapter, CNN-based method with three-dimensional filters applied to different structural and functional images are to be discussed in depth.

MULTI MODAL BRAIN MRI IMAGE SEGMENTATION:

Brain MR image segmentation is a key function in radio therapy and image guided surgery. Quantitative analysis of brain abnormality requires an accurate segmentation of MR images of brain which is a complex and hard challenge. The heterogeneous nature of the lesions which include its large scale variability in phrases of length, shape and area make segmentation assignment extremely difficult. Manual segmentation the use of a human expert is the high-quality approach that is time eating, tedious, high priced and impractical in larger research and introduces inter observer variability. A couple of image sequences with varying contrast want to be taken into consideration for identifying whether a selected location is a lesion or not. And additionally the extent of expert knowledge and enjoy impacts the segmentation accuracy. In mind imaging studies special modality pix are mixed to enhance the short comings of character imaging techniques.T1 weighted photos produce correct evaluation among gray remember (GM) and white matter (WM) tissues. T2 weighted and Proton Density (PD) photographs enables in visualizing the lesions and different abnormalities Fluid Attenuated Inversion recovery (aptitude) photographs improve the photograph assessment of white remember lesions on account of a couple of sclerosis [2]. To enhance the accuracy in brain photo segmentation, fusing distinct modality photograph is important. Fusing multi modality photograph is important within the case of toddler brains which has poor assessment.

Recently unique Deep studying techniques based totally on 2d CNNs (Pereira et al., 2015) are utilized in mind image segmentation duties. In second CNN primarily based strategies the 3-d brain segmentation is finished by using processing individual 2nd slices independently which is a non optimal method. CNNs can research both the capabilities and classifiers simultaneously from statistics. Currently densely related networks are utilized in scientific image segmentation (). Most of the present Multi modal CNN segmentation techniques comply with an early fusion strategy in which MRI T1, T2 and fractional anisotropy (FA) pix are certainly merged at the input of the network. These methods count on that the relationship among specific modalities is sincerely linear. Some different methods study complimentary facts from the other modalities. But the courting among different modalities are tons extra complicated.