Segmentation in Skeletal Scintigraphy Images with Deep Learning

Konrad Gjertsson

Ever since the discovery of the now famous X-rays at the end of the 19th century, an important line of work conducted by physicians has been the study and analysis of medical images. Today there exists a large number of methods for producing images of human entrails, where physicians have become very adept at deciphering different body parts and organs in the all but clear images. There are still however, room for improvements in both the quality of the images as well as the performance of the physicians in their ability to extract relevant information from the images.

Studies performed with different physicians show that the interobserver agreement is low when working with medical images for both segmentation tasks – i.e. being able to tell which pixels belong to the region of interest and which belong to the background – and classification tasks – i.e. in what stage is the cancer of the patient in. While increasing the quality of the images is a costly undertaking, an easier approach would be to increase the performance of the physicians by generating automatic measurements from said images. These would of course be reproducible, consistent and accurate and serve to aid the physicians in their work. One such measurement is the Bone Scan Index (BSI), derived at Exini Diagnostics AB. The BSI is calculated by dividing the mass of a patient's metastases and their entire skeletal mass. This index gives a good indication of the progression of a patient's metastatic disease over time, e.g. prostate cancer. A case study which compared the performance of different physicians concluded that the physicians performed much better, and with much less variability, when given the patient's BSI as a basis for their judgments.

A problem with the BSI is that it is very expensive to produce manually since it requires a physician to extract all necessary information by hand from the images. This is a lengthy and tiresome process since it is necessary to classify each pixel in said images as either part of the skeleton, a metastasis or the background. This means there is room for in silico methods which are able to automatically produce the index. As of the writing of this article, such automatic methods exist and are commercially available in the Exini BSI software. Today, the automatic segmentations are performed through traditional image analysis: non-rigid image registration using the Morphon method. While such a method shows good and consistent performance, it is believed that better and even more accurate segmentations can be achieved through models which falls into the domain of machine learning, and more specifically deep learning.

Machine learning is the art of constructing computer programs that can perform certain tasks without explicitly commanding them to do so in their source code. Instead, computers are *taught* to perform certain tasks, which means that they *learn* from experience and improve over time. Such models are applied in a wide range of different domains, where image analysis has proven one of the most successful in recent years. This is largely due to a rising sub-field within machine learning known as deep learning. Models created with deep learning aim to mimic the architecture and constituents found in the biological brain where artificial neurons are intertwined and connected in such a way that they together form an artificial neural network. The reasoning behind such structures/models are simple: it works for animals, which are able to reason, recognize faces, patterns, speech, etc. So if a neural network works for biological entities, why would it not work for such tasks for artificial agents as well? The answer is that it does work for artificial constructions, it works very well and is a key component in artificial intelligence. When one constructs very deep such networks with a large number of layers of artificial neurons stacked upon each other with up to hundreds of millions tweakable parameters one calls such networks for deep neural networks.

In recent years such models which have been employed for a large range of image related tasks such as being able to tell which objects exist in an image, where the objects are, what pixels belong to the identified objects and which do not etc. The problem that has attracted the most attention in the last few years is the classification of objects in images, which means that most deep neural networks are fine-tuned to perform well in such a task. If one wants to perform segmentation rather than classification, as in the calculation of the BSI, a common approach is to copy such optimized models as found for classification tasks and modify them to make them work for segmentation instead. Artificial neural networks are often very specialized for the tasks they are designed for so such alterations is far from trivial. In the last few years however, groundbreaking research have shown that such changes are possible and powerful artificial networks have been constructed for inter alia segmentation. Recent research also indicates that deep artificial neural networks are very powerful models also when it comes to analysis in medical images. Recent breakthroughs in deep learning have allowed researchers to accurately perform pixels-wise segmentations of objects in amongst other things skeletal scintigraphy images. The goal is to replace the segmentations previously performed with traditional image analysis to generate more accurate, automatic segmentations which give physicians reliable measurements when they make decisions with regards to patients' health based on the information found in medical images.