Digital Pathology

732A64 Master's Thesis

David Hrabovszki
Linköping University

22 February 2021

Contents

1 Introduction

It is estimated that around 300,000 new brain and nervous system cancer cases occurred in 2020 worldwide. Around 250,000 deaths occurred from this type of cancer in the same year (Sung (n.d.)). The World Health Organization classifies tumors into grades based on their malignancy, where grade I is the least malignant and grade IV is the most malignant (Louis (n.d.a)). Grade II and III cancers are called Lower Grade Gliomas (LGG), and grade IV cancers are called Glioblastoma or Glioblastome Multiforme (GBM) (Ostrom (n.d.)).

It is important to diagnose cancer types correctly, because treatment options and survival

expectancy depend largely on how malignant a tumor is and what characteristics it has. There are histological differences between different types, which helps the expert pathologist in the decision making. Grade I lesions have the possibility of cure after surgery alone, grade II tumors are more infiltrative, can progress to higher grades, and often recur, and grade III is reserved for cancer that has some evidence of malignancy. The treatment of grade III lesions usually include radiation and chemotherapy. Grade IV tumors are malignant, active, necrosis-prone (death of the tissue), progress quickly and often cause fatality (Louis (n.d.b)).

!!Glioblastome is different from LGG, because it can take mulitple shapes, include necrosis, etc.!!

!! Histology studies the microscopic anatomy of biological tissues. What is WSI? Digital pathology allows us to analyze WSIs with a computer, therefore facilitate the work of an expert pathologist.!!

1.1 Related Work

2 Data

In this paper, Whole Slide Images from The Cancer Genome Atlas (TCGA) are used. The dataset is publicly available, and contains tissues from GBM and LGG brain cancer types from many different clinics. There are 860 examples of GBM and 844 examples of LGG available as Diagnostic Slides. The images are labeled as a whole, therefore no pixel-wise annotation is available. The files can be more than 3 GB in size, and their resolution is ofter higher, than 100,000 x 100,000. This is why they are saved in a special format (svs), that allows for storage of such large files.

The images were scanned at multiple different resulutions, which can be separately obtained thanks to the special file format. Not all images have the same highest magnification level, however. All LGG classified scans were recorded at 40x magnification (0.25 μm/pixel), while 77% of GBM scans have only 20x magnification (0.5 μm/pixel) available. In order to analyze them together, all images need to be obtained at the same level.

Since the images are so large, it is impossible to process them as a whole, therefore patches or tiles are extracted from them, that are easier to handle for a neural network.

8 References

Louis. n.d.a.

———. n.d.b. https://doi.org/%20https://doi.org/10.3322/caac.21551.

Ostrom. n.d. https://doi.org/doi:10.1093/neuonc/noy131.

Sung. n.d. https://doi.org/doi:%2010.3322/caac.21660.

9 Appendix