Preparation of Whole Slide Images for usage in Neural Networks

Master Thesis

Submitted in partial fulfillment of the requirements for the degree of

Master of Science (M.Sc.)
in Applied Computer Science

at the

Berlin University of Applied Sciences (HTW)

Hochschule für Technik und Wirtschaft Berlin
University of Applied Sciences

First Supervisor: Prof. Dr. Peter Hufnagl

Second Supervisor: Diplom Informatiker Benjaming Voigt

Submitted by:

Sascha Nawrot (B.Sc.)

Berlin, April 12, 2016

Abstract

This is the abstract.

Preface

Hello, this is the preface

Contents

Chapter 1

Introduction

1.1 Motivation

The medical discipline of pathology is in a digital transformation. Instead of looking at tissue samples through the means of traditional light microscopes it now is possible to digitalize those samples. This digitalization is done with the help of a so called slide scanner. The result of such an operation is a*whole slide image* [2]. The digital nature of whole slide images opens the door to the realm of image processesing and analysis which yields certain benefits, such as the use of image segmentation and registration methods to support the pathologist in his/her work.

A very promising area of image analysis are the so called *neural networks*[1], which follow a machine learning approach. Their use, especially in the area of image classification and object recognition, made big breakthroughs possible in the recent past, e.g. Karpathy and Fei-Fei [1], who created a neural network which is capable of describing an image or a scene with written text (see fig. 1.1 for some examples).

There is enormous potential in the use of neural networks in the digital pathology as well, but to transfer these algorithms and technologies, certain hindrances must be overcome. One of those hindrances is the need for proper training samples and an established ground truth[2]. While generally there are huge amounts of whole slide images, most of them won't be usable without further preparation as a training sample. One way of preparing them are annotations. These can be added to the whole slide images, stored and later used for training.

Therefore the goal of this thesis is to give tools into the hands of pathologists to annotate whole slide images and save those annotations in such a way that they will be usable later in the combination with neural networks.

[1] See chapter 2.1.x

[2] *Ground truth* refers to a training sample whose outcome is known and validated as true.

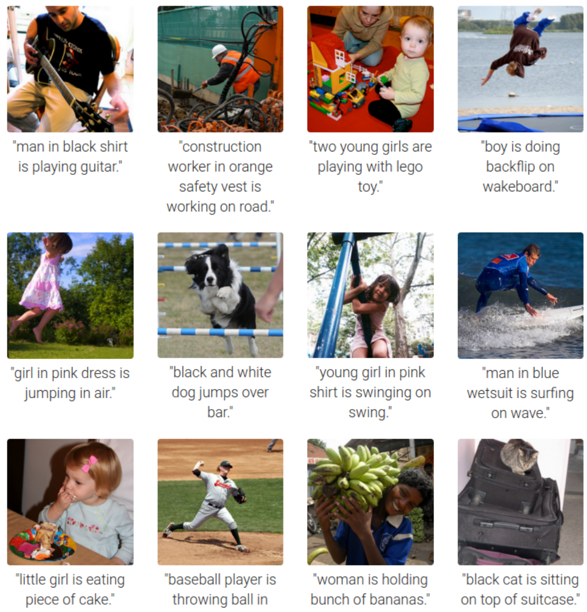

Figure 1.1: Example results of the in [1] introduced model (source: http://cs.stanford.edu/people/karpathy/deepimagesent/

1.2 Research Objective

The objective of this thesis is the conceptualization and implementation of tools for whole slide images[3], which allow for their annotation and a further usage in neural networks. As a requirement, the tools have to be implemented in the form of microservices[4], each one with their own short documentation, including instructions for installation, usage and some examplary use cases. To achieve this goal, the implementation of 3 microservices is necessary.

The first microservice needs to be capable of converting a given set of image formats into the so called *Deep Zoom Image Format*[5]. The supported image formats are *.bif, .mrxs, .ndpi, .scn, .svs, .svslide, .tif, .tiff, .vms* and *.vmu*, in accordance with the capabilities of the Openslide framework [3].

The second microservices task is to give a tool at hand with which a pathologist will be able to annotate all those whole slide images which were converted using the first microservice. Furthermore, the made annotations need to be persisted together with the highest resolution of the corresponding image.

The third microservice will be responsible for preparing the annotated whole slide images for the further usage in neural networks. For that purpose the ser-

[3]See chapter 2.1.x
[4]See chapter 2.1.x
[5]See chapter 2.1.x

vice needs to be capable of dividing a single annotated whole slide image into multiple tiles, with the choice of either using the whole image or just the annotated areas. Furthermore, each tile needs enough information to reconstruct the whole image again afterwards.

1.3 About this thesis

Apart from the *Introduction*, there are 5 more chapters in this thesis.

Chapter 2 - Background defines some terminoligy and the general, required process chain which are all necessary to understand further chapters of this thesis. Furthermore, 3 microservices will be introduced in short.

Chapter 3 - Methodology gives an overview over the current state of research for each microservice, as well as best practices.

Chapter 4 - Implementation goes into further details about how each microservice is implemented and which software and frameworks were used for that.

Chapter 5 - Discussion will introduce a measurement for each microservice to measure its success. It will discuss the test setup as well as list the results.

Chapter 6 - Conclusion will interpret the Results from Chapter 5 and analyze them closer. Furthermore, it will give an idea of what steps are to be taken next in the future.

Chapter 2

Background

2.1 Definition of terms

2.1.1 Deep Zoom Image Format

The Deep Zoom Image Format (.dzi) is an xml-based file format maintained by Microsoft to improve performance and quality in the handling of large image files. Therefore an image will be represented in a tiled pyramid scheme (see fig. 2.1).

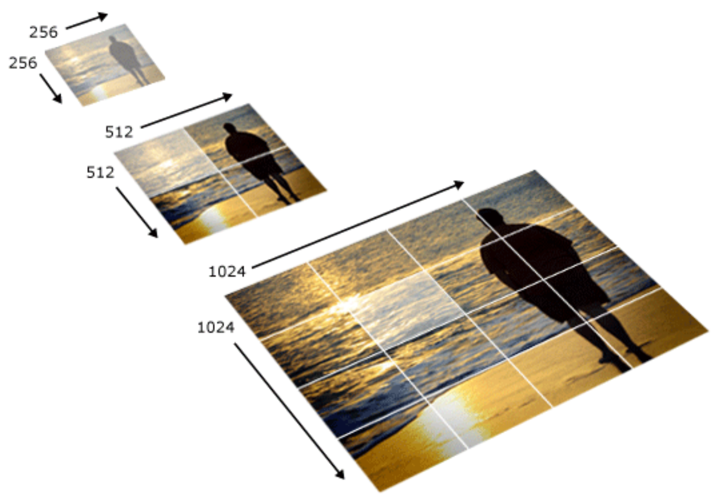

Figure 2.1: of the dzi pyramid image representation (source: https://i-msdn.sec.s-msft.com/dynimg/IC141135.png)

As seen in fig. 2.1 there are multiple versions of a single image in different resolutions. The idea behind this is, that if a user wants to see a whole picture zoomed out or as a small thumbnail, it is not necessary to load the image file in its highest resolution. To save bandwidth a version with a smaller resolution is loaded. If the user wishes to zoom in on a specific area of the image, a version with a higher resolution is loaded. Once again, however, it is not necessary to load the whole image, since only a fraction of it will be visible. For this reason there are tiles of the image which are loaded instead (see highlited tiles in fig. 2.1) [4].

Each resolution in the pyramid is called a *level*. At each level the image is scaled down by the factor 4 (2 in each dimension). In other words, a level can be defined as an image with the resolution 2*level for height and width, resulting in a resolution of (2*level)*(2*level). Levels are counted from level 0 (1*1 Pixel) [4]. E.g. the levels shown in fig. 2.1 are:

- level 8 ($2^8 = 256$) for the 256^2 pixel image
- level 9 ($2^9 = 512$) for the 512^2 pixel image
- level 10 ($2^10 = 1024$) for the 1024^2 pixel image

2.1.2 Microservice

2.1.3 Machine Learning

2.1.4 Neural Networks

2.2 Process chain

2.2.1 Description

2.2.2 Definition of Conversion Service

2.2.3 Definition of Annotation Service

2.2.4 Definition of Tesselation Service

Chapter 3

Methodology

3.1 Conversion Service

3.1.1 Literature review

3.1.2 Chosen methods

3.2 Annotation Service

3.2.1 Literature review

3.2.2 Chosen methods

3.3 Tessellation Service

3.3.1 Literature review

3.3.2 Chosen methods

Chapter 4

Implementation

4.1 Implementation of Conversion Service

4.1.1 Used technologies and frameworks

4.1.2 Documentation

4.2 Implementation of Annotation Service

4.2.1 Used technologies and frameworks

4.2.2 Documentation

4.3 Implementation of Tessellation Service

4.3.1 Used technologies and frameworks

4.3.2 Documentation

Chapter 5

Discussion

5.1 Conversion Service Test

5.1.1 Setup

5.1.2 Results

5.2 Annotation Service Test

5.2.1 Setup

5.2.2 Results

5.3 Tessellation Service Test

5.3.1 Setup

5.3.2 Results

Chapter 6

Conclusion

6.1 Results

6.2 Conclusion

6.3 Future tasks

Bibliography

[1] Li Fei-Fei A. Karpathy. *Deep Visual-Semantic Alignments for Generating Image Descriptions*. Department of Computer Science, Standford University, 2015. http://cs.stanford.edu/people/karpathy/cvpr2015.pdf.

[2] Toby Cornish. An introduction to digital whole slide imaging and whole slide image analysis. http://www.hopkinsmedicine.org/mcp/PHENOCORE/CoursePDFs/2013/13%2019%20Cornish%20Digital%20Path.pdf, July 2013. Accessed: 12.04.201600.

[3] A. Goode et al. Openslide: A vendor-neutral software foundation for digital pathology. *J pathol inform*, 4(1), September 2013. http://download.openslide.org/docs/JPatholInform_2013_4_1_27_119005.pdf.

[4] Microsoft. Deep zoom file format overview. https://msdn.microsoft.com/en-us/library/cc645077(v=vs.95).aspx. Accessed: 11.04.2016.

List of Figures

List of Tables