Project AI for Medical Image Analysis (8P361)

Assignment 1

Group 04

Full Name	Student ID
Geert Boer	xxxxxx
Levi Geraedts	1890735
Jelle van Namen	1894102
Dirk Sollewijn Gelpke	xxxxxxx

Eindhoven, February 19, 2025

Contents

1 Exercises **2**
- 1.1 Exercise 1 . . . 2
- 1.2 Exercise 2 . . . 2
- 1.3 Exercise 3 . . . 2
- 1.4 Exercise 4 . . . 3
- 1.5 Exercise 5 . . . 4

2 Appendices **4**

1 Exercises

1.1 Exercise 1

What is the clinical utility of evaluating the presence of metastases in sentinel lymph nodes in breast cancer patients? In other words, how is this information used in the clinical decision making process for breast cancer patients?

The lymphatic system is a network of vessels and organs that transport immune system cells and waste products trough a clear fluid called lymph [1]. This fluid is filtered trough lymph nodes, which are present throughout the body. These nodes can trap bacteria, viruses, and most importantly in this context: damaged and abnormal cells [1]. By definition, sentinel lymph nodes are the first lymph nodes that receive lymph-borne metastatic cells . In other words, they are the first lymph nodes trough which cancer spreads from a primary tumor [1, 2, 3]. Detecting the presence of metastases in sentinel lymph nodes is therefore a crucial step in determining the stage of breast cancer: If no metastasis is found in the sentinel lymph nodes, it suggest that the cancer has not yet spread to other nearby lymph nodes or organs. Dissection of axillary lymph nodes other than the sentinel node can therefore be omitted, which significantly reduces the incidences of dissection related symptoms such as postoperative lymph-edema [3]. Conversely, the presence of metastases in sentinel lymph nodes suggests a possibility of the cancer having spread to other lymph nodes or organs [1, 3]. This information helps doctors determine the extent to which the cancer has spread trough the body (staging), which can help in choosing an appropriate treatment plan [1]. For example, a positive sentinel lymph node in older women is often associated with increased use of chemotherapy, radiotherapy, and axillary lymph node dissection [4].

1.2 Exercise 2

How does the introduction of whole-slide imaging change the typical workflow of a pathology lab?

Whole-slide imaging (WSI) is a technology that has the ability to digitally scan entire slides of microscopic data in high resolution [5]. Meaning that microscopic experiments can be translated into digital images useful for further research. WSI is therefore a very promising tool for computer aided diagnostics like AI or machine learning algorithms [6]. These algorithms need a lot of image data to perform the task of predicting pathological patterns and WSI can provide this large amount of images.

It is quite clear that the introduction of WSI changed the normal workflow of a pathology lab because it is way more time efficient. Before the introduction of WSI, pathologists had to examine and analyze microscopic data through physical microscopic based sessions. These sessions were often very time consuming and non-efficient and required a certain skill of using a microscope. With WSI the focus is more on the ability to analyze and intepret the image data than microscope skill and it gives the opportunity for more people to look at the same microscopic image at the same time. [6]

1.3 Exercise 3

The PatchCamelyon dataset is derived from the CAMELYON16 dataset of whole-slide images. Describe how a neural network classification model trained on small image patches can be applied to larger, whole-slide images with the goal of detecting metastases.

Because of the large size of the whole-slide images, and the low number of available whole-slide images, training a neural network on an entire slide is difficult [7]. To solve this, the whole-slide

images are divided in smaller patches on which a classification model can more easily be trained . This classification model will give a value between 0 and 1 representing the probability that the patch contains metastasis. When applied on a whole-slide image, this model can create a heat map of metastasis probability [7, 8]. In this heatmap, each pixel represents a patch, and has a value between 0 and 1 computed by the patch-based classification model. This heatmap can then be further processed into a final prediction of tumor presence in the whole-slide image. For example, geometrical and morphological features can be extracted and combined with a random forest classifier to predict metastasis presence on a whole-slide basis [8]. Additionally, thresholding can be used to determine the locations of the tumors on the whole-slides. The entire framework therefore consists of a patch-based classification stage and a heatmap-based post-processing stage [8].

1.4 Exercise 4

Download and unzip the training and validation (4 GB), and testing (1 GB) subsets of the Patch-CAMELYON dataset. Note that the unzipping process might take a while due to the large number of files in the archives. Write a small Python script that reads and displays a few images from the two classes. Visually describe and compare the appearance of the tissue in the patches with and without metastases.

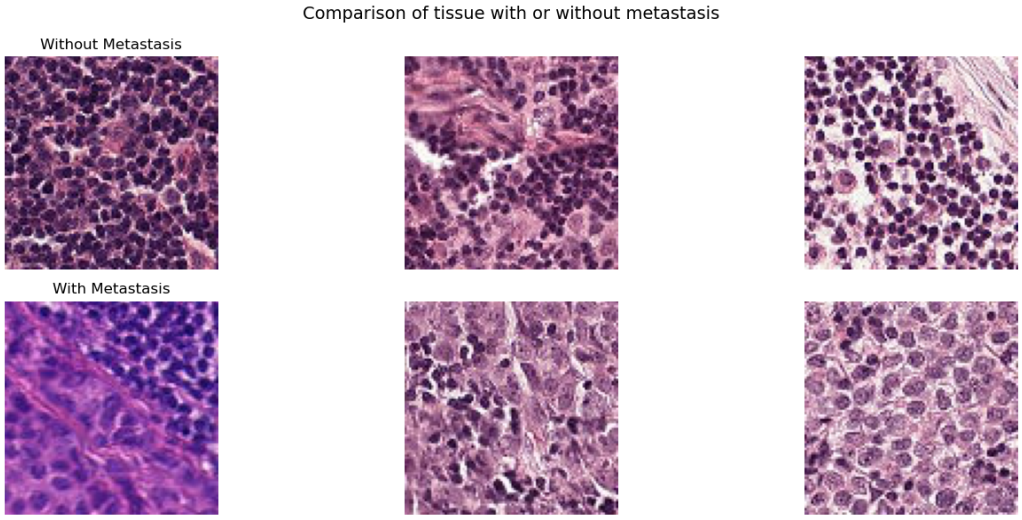

Figure 1: figure showing 3 images for both with and without metastases

The images shown in figure 1 are quite different mainly because they are randomly chosen from a set of training data. On first sight is it quite difficult to spot some differences between the tissues with and without metastases. However, when examined more carefully it can be stated that the tissues with metastases have larger nuclei on average.

1.5 Exercise 5

2 Appendices

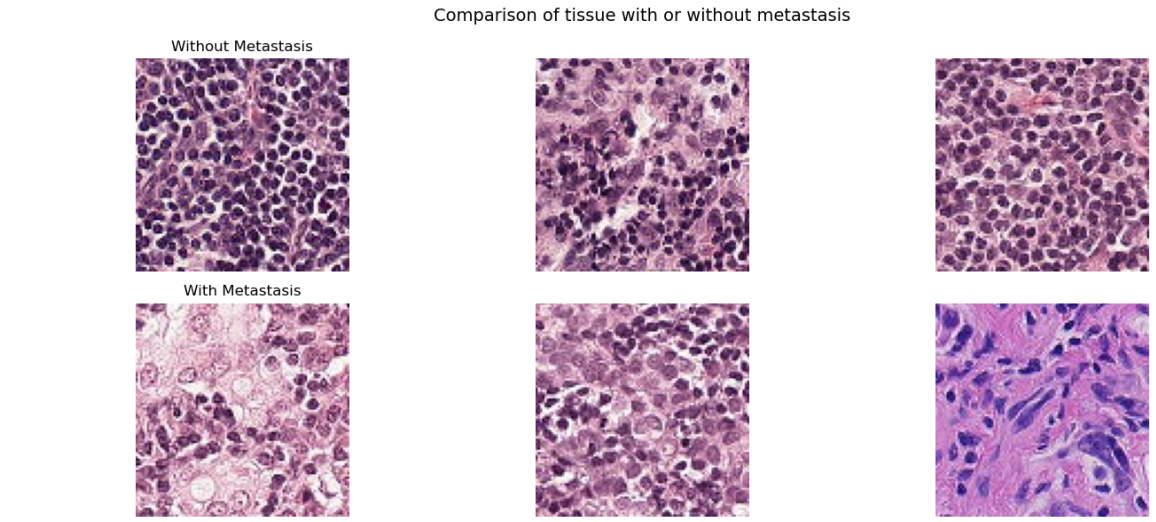

Figure 2: number of images per row=3

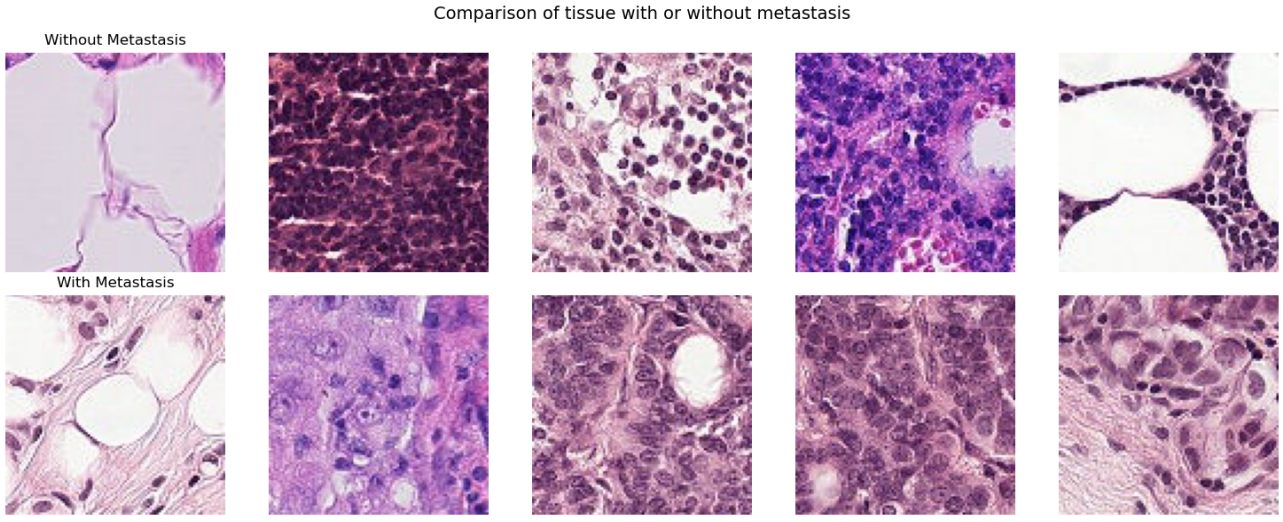

Figure 3: number of images per row =5

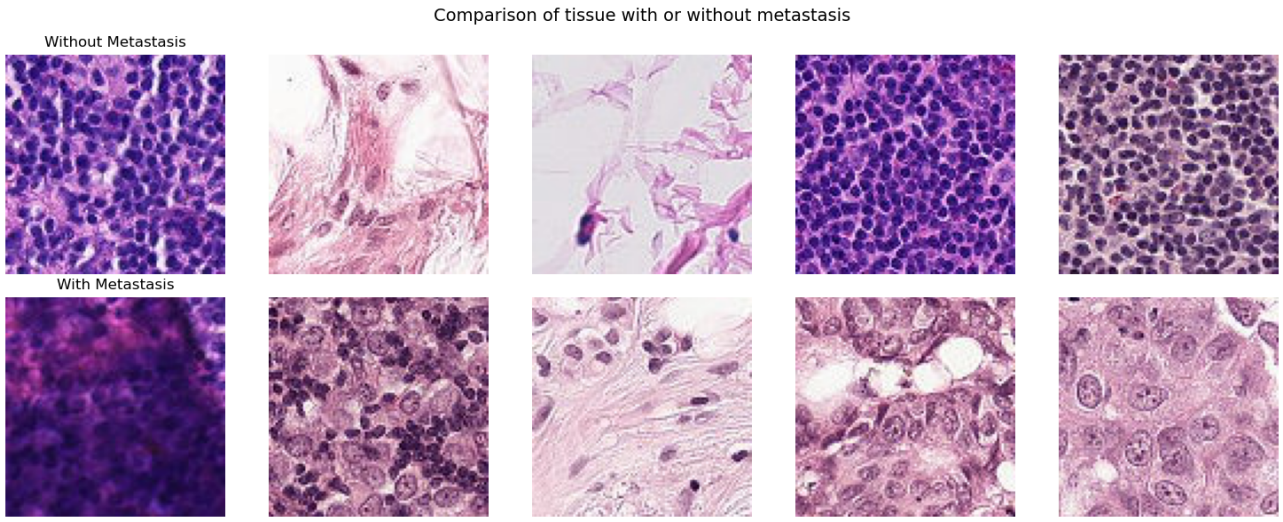

Figure 4: number of images per row =5

References

[1] National Cancer Institute. Sentinel lymph node biopsy fact sheet, June 2019.

[2] Francesco Giammarile, Sergi Vidal-Sicart, Diana Paez, Olivier Pellet, Estrada-Lobato Enrique, Miriam Mikhail-Lette, Olga Morozova, Maria Camila Navarro Marulanda, Diana Ivonne Rodríguez Sanchez, Roberto C. Delgado Bolton, Renato A. Valdés Olmos, and Giuliano Mariani. Sentinel lymph node methods in breast cancer. *Seminars in Nuclear Medicine*, 52(5):551–560, March 2022.

[3] Hiroya Takeuchi and Yuko Kitagawa. Sentinel node and mechanism of lymphatic metastasis. *Annals of Vascular Diseases*, 5(3):249–257, January 2012.

[4] Gordon R. Daly, Gavin P. Dowling, Eman Hamza, Ma'en AlRawashdeh, Aisling Hegarty, Tarnum Naz, Jason McGrath, Sindhuja Naidoo, Eithne Downey, Abeeda Butt, Colm Power, and Arnold D. K. Hill. Does sentinel lymph node biopsy influence subsequent management decisions in women with breast cancer 70 years old? *Clinical Breast Cancer*, 24(6):510–518.e4, May 2024.

[5] Molecular Machines & Industries. Whole slide imaging — digitally scan & archive in hi-res, March 30 2022. Accessed: February 19, 2025.

[6] N. Kumar, R. Gupta, and S. Gupta. Whole slide imaging (wsi) in pathology: Current perspectives and future directions. *Journal of Digital Imaging*, 33(4):1034–1040, 2020.

[7] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E. Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q. Nelson, Greg S. Corrado, Jason D. Hipp, Lily Peng, and Martin C. Stumpe. Detecting cancer metastases on gigapixel pathology images, March 2017.

[8] Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H. Beck. Deep learning for identifying metastatic breast cancer, June 2016.

[9] OpenAI ChatGPT. Chatgpt: Ai language model by openai, February 2025. Accessed: YYYY-MM-DD.

AI statement

ChatGPT [9] was used during the writing of this assignment in the following ways:

- Source conversion to BibTex.
- Exploratory queries to get a quick overview and starting point for some of the questions.
- During coding to create code snippets, debug and make the code more readable.
- Final suggestions on how to improve the texts grammar / flow.