8P361 Project Imaging - BIA Group 4 | Assignment 1
Myrthe van den Berg, Rick van Bergen, Nils Bodenstaff, Cris Claessens
and Jeanine van Ingen

Exercise 1
What is the clinical utility of evaluating the presence of metastases in sentinel lymph nodes (SLNs) in breast cancer patients? In other words, how is this information used in the clinical decision-making process for breast cancer patients?

Due to their connection to the axillary lymphatic system, SLNs are believed to be first involved by the potential spreading of disease. Their status of metastatic involvement is an important determinant of stage and need for further surgical dissection and/or nodal irradiation [1]. The evaluation of metastases in SNLs is used in the clinic to estimate the chance of developing a tumor in different body parts, as breast cancer mostly spreads through the lymphatic system. Metastasis to increasing numbers of nodes results in larger metastases, extracapsular growth of metastases, and a higher axillary level of involvement. In other words, the likelihood of tumors being present elsewhere in the body increases with the size of the metastases [2].
In cases of early-stage and clinically node-negative breast cancer, SLNs can demonstrate a range of involvement by metastatic spread from macrometastases (>2 mm), micrometastases (>0.2 to≤2.0mm), and isolated tumor cells or ITCs (≤0.2 mm or 200 cells). An SLN is considered involved by metastatic disease if there is clear infiltration of tumor cells with morphologic features consistent with the patient's primary breast cancer [1].

Exercise 2
How does the introduction of whole-slide imaging change the typical workflow of a pathology lab?

Compared to the general workflow of how tissue sections are prepared and viewed under a microscope, the digital workflow requires additional equipment, trained personnel, and specific quality control steps. The pathologist plays a key role both in the process of slide digitization and in the subsequent data generation via the use of appropriate image analysis algorithms. The advantages of transitioning to a digital workflow include ease of slides and cases sharing, standardization of teaching, organization of archived digitized slides, and extraction of complex data in a highly reproducible fashion. Since the data is digitized and can be shared more easily, it can also be used for deep learning [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.

In each local patch, feature extraction and classification between cancer and non-cancer areas can be performed. The goal of feature extraction is to extract useful information for machine learning tasks after which a convolutional neural network can start an analysis. Features and classifiers are simultaneously optimized in deep learning and features learned in deep learning often outperform other traditional features in histopathological image analysis. Aggregation can be used to get the value of the whole image. For that cause, small

patches can be used to train the neural network, after which the DNN can be applied to whole-slide images [4].

Exercise 4
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.

When running the script (provided in the .zip file), the random images that are displayed are shown in figure 1.
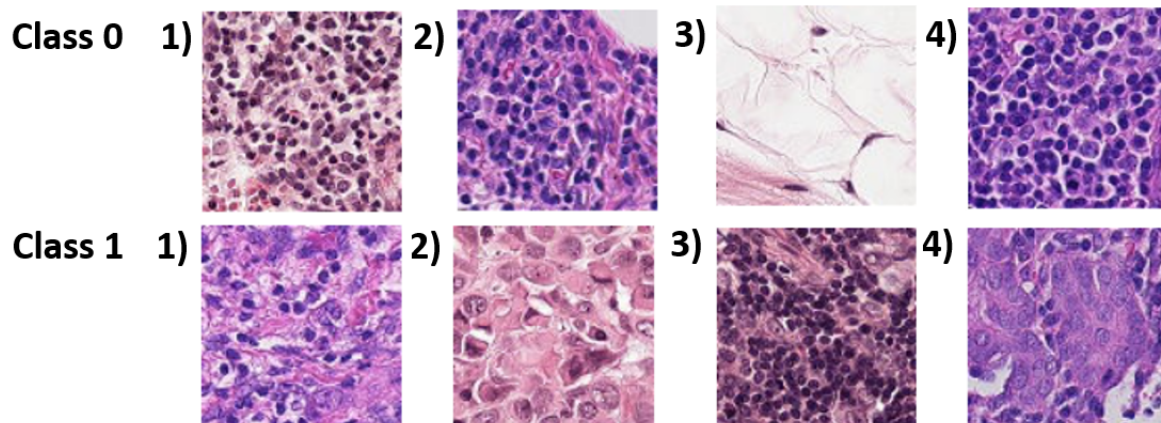

Figure 1, the four random images, from class 0 and class 1 of training data, that are shown after running the script.

The nuclei in the images in class 1 are bigger and less circular compared to the nuclei in class 0, this can best be seen when comparing, in figure 1, image 1 from class 0 and image 3 from class 1. While considering the literature [5], we assume that class 0 contains patches without metastases and class 1 with metastases.

Exercise 5
Make an account on Kaggle and subscribe to the Patch-CAMELYON challenge. You will use this account to submit results for evaluation from your main project work. You can also make accounts for every team member and submit results as a team. Note that you do not have to download the dataset again from Kaggle (the version that we have prepared is a bit easier to work with and can still be used to submit results).

We choose to submit results by using the account of Rick van Bergen, username: Rick van Bergen.

References

[1] Sunil W. Dutta, Ashley Volaric, John T. Morgan, Zachary Chinn, Kristen A. Atkins, Einsley-Marie Janowski, Pathologic Evaluation and Prognostic Implications of Nodal Micrometastases in Breast Cancer, Seminars in Radiation Oncology, Volume 29, Issue 2, Pages 102-110, ISSN 1053-4296, https://doi.org/10.1016/j.semradonc.2018.11.001.

[2] Donegan, W.L., 1997. Tumor-related prognostic factors for breast cancer. CA: a cancer journal for clinicians, 47.

[3] Aeffner, F., Zarella, M. D., Buchbinder, N., Bui, M. M., Goodman, M. R., Hartman, D. J., Bowman, D. (2019). Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. Journal of pathology informatics, 10, 9. DOI:10.4103/jpi.jpi_82_18

[4] Komura, D., & Ishikawa, S. (2018). Machine Learning Methods for Histopathological Image Analysis. Computational and structural biotechnology journal, 16, 34–42. DOI:10.1016/j.csbj.2018.01.001

[5] Seyfried, T. N., & Huysentruyt, L. C. (2013). On the origin of cancer metastasis. Critical reviews in oncogenesis, 18(1-2), 43–73. doi:10.1615/critrevoncog.v18.i1-2.40