Project Imaging Group 10 – Assignment 1

Helgers, V.M.J. 1329332
Jong, de R.L.P.D. 1328328
Korsten, T. 1340522
Moharir, S. 1296256

Question 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 assessment of the breast cancer stage with testing the presence of metastasis in histological analysis of sentinel axillary lymph nodes (SNLs) in breast cancer patients is an essential part to inform clinical management [1]. The lymph node status is one of the many prognostic factors for breast cancer. The prognostic factor plays an important role in the clinical decision making process for breast cancer patients, in other words the treatments of the patients is dependent on the prognostic factors [2].

Breast cancer patients may benefit from adjuvant therapy for the early-stage operable cancer. These therapies are very costly and have serious side effects. This way patients that are very sensitive for recurrence can have systematic adjuvant therapy, whereas the other patients with less risk can be spared from the treatment and their side-effects that will offer no benefit. In addition, the prognostic factors can help improving the clinical testing can by having comparable treatment groups and providing markers for measurement of the success or failure of therapies [3].

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

With whole slide imaging (WSI), images of tissue sections that have been scanned at various magnifications are digitally manipulated. These manipulations give the user an image in which they can zoom in on areas of interest, and so simulate the examination of samples on glass slides under a traditional microscope [4].

The introduction of WSI can change the typical workflow of a pathology lab in terms of performance, work-flow efficiency, and access to pathology services in under resourced locations.

Firstly, by using WSI systems, the quality of the images will be better, and any disruptions otherwise caused by unintentional movement of the slides are eliminated, thus streamlining the workflow. The pathologist can easily navigate through the slides and images and observe the cells in detail without disturbances.

Next to this, exchanging and accessing the information on these slides is also made easier for the pathologist. They can obtain and share images through a database, thereby reducing the time that the pathologist normally would have to spend on matching slides to paper requisitions and looking for glass slides for signing out. By using WSI the time spent productively is therefore increased.

Lastly, these images can also be accessed on remote locations where transportation of samples might be difficult [5].

Question 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

The detection of objects typically consists of localization and identification of small objects in the full image. Objects are detected by focusing on small parts of the image and classifying them in two or more classes. In the small image patches, the Region of Interest has already been determined. In case of a larger, whole-slide image, this still needs to be done before the neural network model can be applied.

Localization through 2D image classification with convolutional neural networks (CNNs) seems to be the most popular strategy overall to identify organs, regions and landmarks, with good results [6]. The most object detection systems are made of multiple CNNs and uses pixel classification after which postprocessing is applied. These CNNs are used to identify the patches in the whole slide image that could be of interest for the diagnosis, and can for example distinguish between tissue cells and air bubbles and thereby eliminate areas that are definitely not of interest. Once the target areas are found with these CNNs, the neural network classification model can be applied to the whole-slide images [6].

Question 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.

A few images are being displayed in the Jupyter-notebook file that is included in the zip-file. Below, the 18 images that are displayed in the notebook are shown. The 9 images in the left square show histological preparations of lymph node tissue without metastases. The 9 images in the right square show preparations that do contain metastases. However there are also some outliers in both groups.

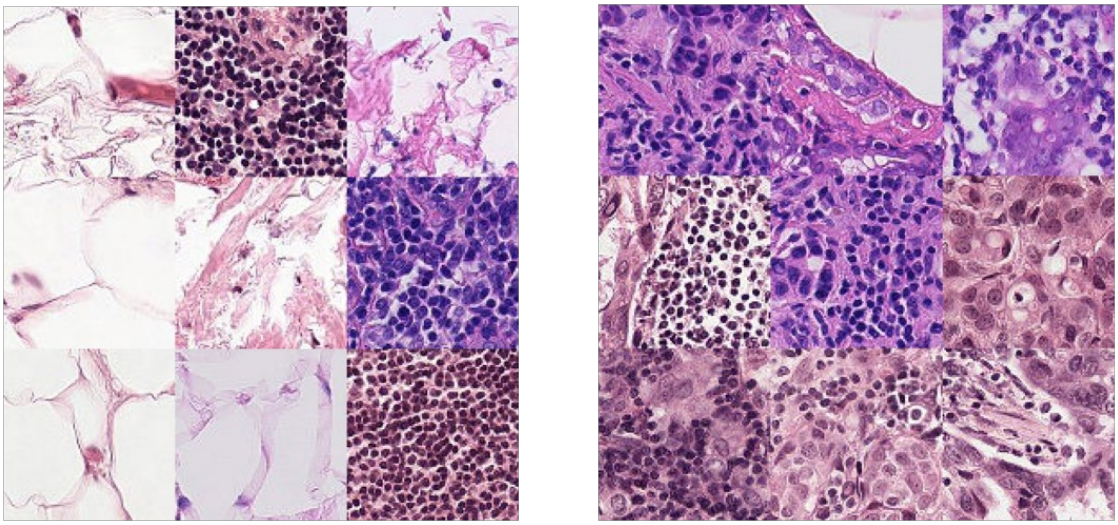

Figure 1: Image patches from histological preparations of lymph node tissue. The images in the left square do not contain metastases, whereas the images in the right square do contain metastases.

It can be seen that tissue containing metastases have pleomorphic, protruding nucleoli and numerous mitosis visible the images [7]. In the tissue that contains metastases, it can furthermore be seen that the cancerous cells have spread throughout the entire tissue and have a high cell density.

Question 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).

The choice was made to make one Kaggle account for the team. The name of this account is as follows: biagroup10.

Bibliography

[1] Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., van der Laak, J. A. W. M. ,the CAMELYON16 Consortium, 2017. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA, 318.

[2] Veta, M., et al., 2015. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Medical image analysis, 20.

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

[4] Whole slide imaging | DermNet NZ. (2018). DermNet NZ. https://dermnetnz.org/topics/whole-slideimaging/#:%7E:text=Whole%20slide%20imaging%20is%20the,slides%20under%20a%20traditional%20microscope.

[5] Ghaznavi, F., Evans, A., Madabhushi, A., & Feldman, M. (2013). Digital Imaging in Pathology: Whole-Slide Imaging and Beyond. Annual Review of Pathology: Mechanisms of Disease, 8(1), 331–359. https://doi.org/10.1146/annurev-pathol-011811-120902

[6] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. and Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis, 42.

[7] Nascimento, R., Otoni, K. 2020 Histological and molecular classification of breast cancer: what do we know?.