Deep learning for histopathological image analysis

Cédric Wemmert[*], Jonathan Weber[†], Friedrich Feuerhake[‡], Germain Forestier[†]

[*] ICube, University of Strasbourg, France
[†] IRIMAS, University of Haute-Alsace, France
[‡] Hannover Medical School, Germany and University Clinic of Freiburg, Germany

Abstract. Anatomical Pathology dates back to the 19th century when Rudolf Virchow introduced his concept of cellular pathology and when the technical improvements of light microscopy enabled wide-spread use of structural criteria to define diseases. Since then, the quality of optical instruments has been constantly evolving. However the central element of the diagnostic process remains the knowledge and experience of pathologists visually classifying observations according to internationally agreed guidelines (e.g., *World Health Organisation* (WHO) classification), and much of the pre-analytical steps of specimen preparation (e.g., fixation, embedding, sectioning, staining) is only partially automated and still requires many manual steps. Thanks to the recent advent and cost-effectiveness of digital scanners, tissue histopathology slides can now be fully digitized and stored as *Whole Slide Images* (WSI). With the availability and analysis of a much larger set of variables combined with sophisticated imaging and analytic techniques, the traditional paradigm of pathology based on visually descriptive microscopy can be complemented and substantially improved by digital pathology, utilizing screen-based visualization of digital tissue sections and novel analysis tools potentially combining the conventional evaluation by pathologists with a computer-based diagnostic aid system. A central element of such evolving medical utilities and decision support systems will be image analysis, a field in which *Deep Learning* (DL) has recently made immense progress, notably the work of Lecun *et al.* [33] on *Convolutional Neural Networks* (CNNs) and especially the development of very large *Artificial Neural Networks* (ANNs) that are revolutionizing the field. Indeed, they have surpassed all existing image processing methods in most fields (segmentation, object detection, classification, etc.). All current methods applied to histopathological image analysis will be presented as well as the future technological issues and challenges of this discipline.

Keywords: machine learning, digital pathology, whole slide images, deep learning

1 Introduction

Pathology is the gold standard for the diagnostic evaluation and the understanding of many of the underlying biological and pathophysiological mecha-

nisms. This typically involves visual evaluation by pathologists of a sample of cells under a microscope, to identify structural tissue properties. Currently, visual evaluation of microscopic specimens is largely an unassisted process, and pathologist's accuracy is established through long training, standardization and benchmarking, quality control by peer reviews, and personal experience. But this field has experienced several technological revolutions in recent years with the advent of virtual microscopy (conversion of glass slides into high-resolution images *i.e.* digital slides), often referred to as *Digital Pathology* (DP). In this domain, important efforts have been made to design image analysis tools, to identify for example basic biological structures (*e.g.* stroma, immune cells), in order to facilitate the task of biologists for (semi-) automated digital slide interpretation. Digital Pathology is currently regarded as one of the most promising avenues of diagnostic medicine in order to achieve better and faster diagnosis, prognosis and prediction of important diseases. With the recent advent of *Whole-Slide Imaging* (WSI) *i.e.* the scanning of entire slides, the field of digital pathology produces daily a massive amount of images with related metadata (*e.g.* patient information, diagnosis, treatment).

At the same time, automatic image analysis algorithms have recently made extraordinary progress, particularly with the advent of the *Deep Learning* (DL) methods introduced by Lecun *et al.* [32]. Indeed, the performance of these methods has exploded in recent years, in particular allowing the detection, classification and segmentation of objects of interest in images with very high precision [21]. Although the technical progress holds great promise for digitization and improvements regarding standardization and increasing efficiency, the majority of institutes for pathology still prefer the conventional microscopy approach and only few hospitals are moving towards a totally digital pathology service [59]. Apart from the huge investment cost for the hospital and the difficulty of changing the daily practices of an entire department, there are also many challenges to overcome in order to integrate WSI in routine diagnostic workflows. Indeed, these images contain billion of pixels and are highly heterogeneous (Fig. 1), in terms of signal (acquisition devices, inter-centric variability [49]), and of semantics (e.g. multiple types of cells), causing practical difficulties for the processing/visualization by conventional algorithms. As pointed out in [28], the main difficulties to analyze histopathological images are their huge size, the lack of sufficient labeled data, the different levels of magnification resulting in different levels of details, the nature itself of WSI (which are more like texture images), the color variations and presence of artifacts.

The rest of the chapter is organized as follows: In section 2, we first present current existing DL models for WSI analysis and explain their strength and limits. In section 3, all the challenges and opportunities, induced by this emerging domain, are discussed. Finally in the last section, we conclude that even if DL has proven its efficiency in the field of histological images analysis, its broad use and acceptance in clinical routine remains challenging.
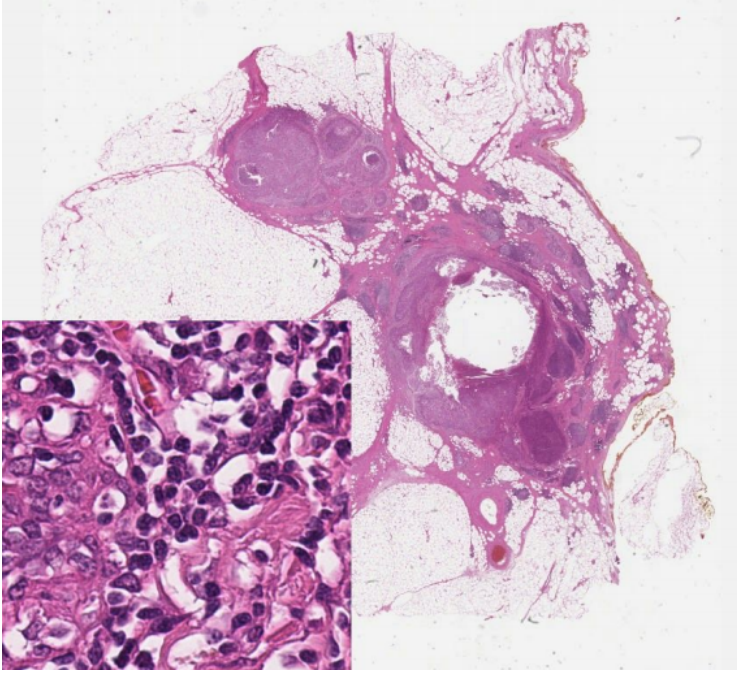
Fig. 1: Example of digitized WSI of a breast cancer surgical resection after previous diagnostic biopsy and subsequent neoadjuvant chemotherapy stained with Haematoxylin&Eosin (18000 × 15000 pixels).

2 Current deep learning models for digital pathology

Numerous DL approaches have been proposed in the last years [23,36]. In this section, we present the most popular ones and analyze their strengths and weaknesses after having introduced what is DL and how it can be applied to solve two major tasks in histopathological image analysis: classification and segmentation.

2.1 What is deep learning?

DL methods are a family of *Machine Learning* (ML) methods that are based on a representation of a model to be estimated in the form of a so-called *deep Artificial Neural Networks* (ANN). The concept of learning from an ANN is not new and dates back to the 1960s with Rosenblatt's definition [48] of the perceptron (Fig. 2). The learned model is a function of the form: $y = f(\sum_{i=1}^{n} \Phi_i x_i)$.

Subsequently, the 1980s saw the emergence of multilayer ANNs [38]. Each *neuron* is an elementary processing unit and the entire ANN allows a representation of a non-linear model. The learning is based on the *gradient backpropagation* method [50], which remains very computationally intensive (Fig. 3).
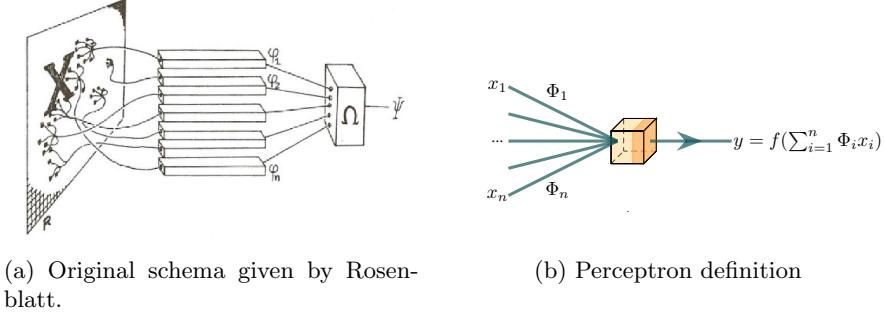
(a) Original schema given by Rosen- (b) Perceptron definition
blatt.

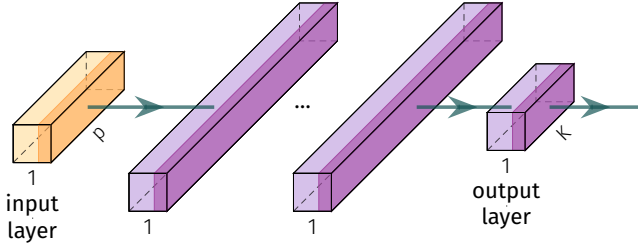Fig. 2: The perceptron proposed by Rosenblatt in 1958.

Fig. 3: A multilayer perceptron with multiple layers of neurons, p neurons in the input layer and K neurons in the output layer.

It is in the 2010's, following the democratization of *Graphics Processing Units* (GPU) cards allowing to have very important computing powers, that this technology got popular, with ANNs containing many layers and offering exceptional performances on many problems. In particular, in the field of image analysis (object detection, classification, segmentation, etc.), the so-called *Convolutional Neural Networks* (CNN) make it possible to achieve unprecedented levels of efficiency. An example of this type of architecture is given in Fig. 4.

To train the model $f_\Theta(x) \to y$, the algorithm needs a training set composed of labeled samples: $(x_1, y_1), \ldots, (x_n, y_n)$. Then, the training of the ANN consists in an optimization problem of a loss function $\Delta(\hat{y}, y) \in \mathbb{R}^+$ as follows:

$$\Theta^* = \arg\min_\Theta \frac{1}{n} \sum \Delta(f_\Theta(x_i), y_i)$$

2.2 Deep learning for classification

The classification task is the classical task for ANNs, in digital pathology application it aims to label patches extracted from WSIs. For this purpose, CNNs
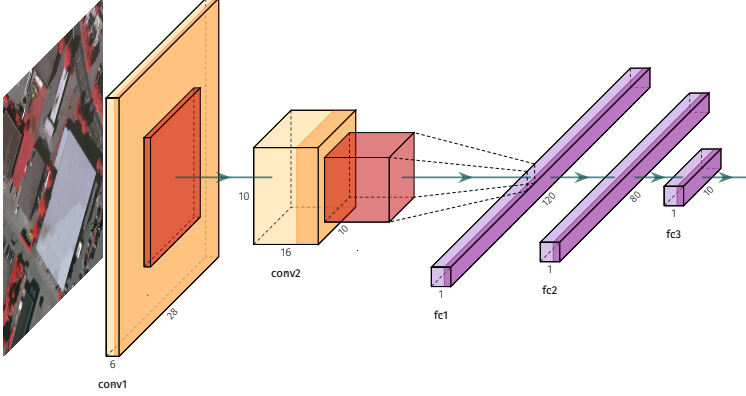
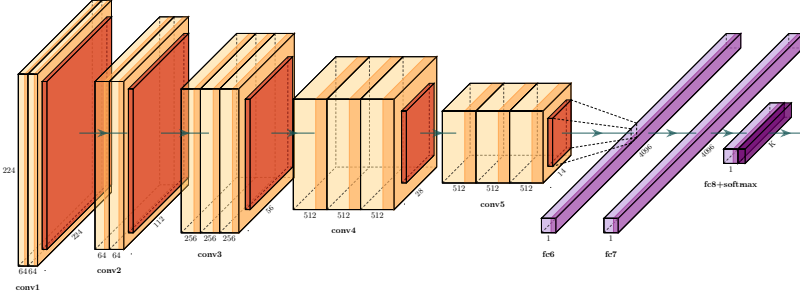Fig. 4: Example of CNN, the *LeNet* architecture[33].

Fig. 5: Example of deep CNN for classification, the VGG16 architecture [55].

are the most effective and most widely used ANNs today. They have a similar methodology to traditional supervised image classification methods: they receive input images, extract features of each of them, and then train a classifier (fully connected multi-layer perceptron like in Fig. 3) on them. However, features are automatically learned by the ANN. Indeed, during the training phase, the classification error is minimized in order to optimize the classifier parameters and the features to be extracted from the image (in the form of convolutions applied at different levels of image resolution). The first architecture developed was the *LeNet* ANN (Fig. 4) by LeCun *et al.* [33]. Soon, more convolution layers were added to achieve deeper architectures, allowing better results like VGG16 proposed in [55] (Fig. 5).

This type of architecture has been widely applied to histopathological image analysis for different tasks mainly in cancer: detection of particular cells or regions of interest, tissue classification, scoring.

Detection Mitosis detection is an important topic in cancer diagnostic, in [52] the authors proposed an interesting approach which uses both DL and hand-crafted features. The idea is to train a CNN, composed of five convolutional layers and two fully connected layers to classify mitosis image patches and non mitosis image patches and to combine it with 55 hand-crafted features plugged to the first fully connected layer. Adding hand-crafted features greatly improves the DL approach results.

In [12], the authors compared the efficiency of classical pre-defined features and features learned from a DL architecture, more precisely an autoencoder CNN [6], for *basal-cell carcinoma cancer* detection. The learned representations performs better than the the pre-defined and allows better cancer prediction.

Scoring *Tumor Proportional Scoring* (TPS) plays an important role in the identification of *non-small-cell-lung-cancer*, as it represents the level of the *Programmed Death-Ligand 1* (PD-L1) expression [67]. In [26], the authors propose to use an *Auxiliary Classifier - Generative Adversarial Network* (AC-GAN) [42] which works as a classical *Generative Adversarial Network* (GAN) [19] but instead of just using noise as input for the generator also use hot-encoded desired class information. To be able to produce TPS score, discriminator not only indicates if the input patch is fake or real but also predict it class (positive tumor cell region or negative tumor cell region) which is further used to compute the TPS score. The performance of this network are good and allows to quickly obtain the TPS score in seconds.

In [53] used DL classification to estimate the spatial organization of immune cells in the tumor microenvironment. For this, they generate maps of *Tumor-Infiltrating Lymphocytes* (TILs) using CNNs enabling to evaluate a prognostic factors, like the Immunoscore [17], that quantify such spatial TILs densities in different tumor regions.

Tissue classification Gecer *et al.* [18] proposes a cascade architecture, first four fully CNNs aims to detect salient regions of interest which are then classified by a CNN into five different type of diagnostics.

Alternatively, in [11], the authors aims to classify tissue between *adenocarcinoma*, squamous cell *carcinoma* and normal tissues using *inception-v3* [62] ANN. Moreover, their ANN is able to predict six of the ten most commonly mutated genes in *adenocarcinoma* which could have a great impact on treatments.

In [70], the authors deal with the problem of not having enough annotated histological data for a specific tissue type. To tackle this problem, they propose to train a *GoogLeNet* architecture [63] on a different annotated dataset (with other tissue type) and then fine-tuning it on their dataset with less annotations.

Deep Multiple Instance Learning [45] is a rising topic in biomedical ML. The basic idea is to use multiple labels to annotate an area, with multiple examples, the network will learn which structure belongs to which labels. It is useful for WSI classification. It has been used for classification of breast histopathology [13] and precursor lesions of *esophageal adenocarcinoma* [66].

2.3 Deep learning for segmentation

While CNNs have proven their efficiency for classification tasks, by splitting images into small patches to be classified, specific architectures have quickly emerged to obtain more accurate detection of objects of interest with their outlines [47,4].

These architectures consist of two parts (Fig. 6), one for encoding the information contained in the image given to the input layer (composed of a succession of encoder layers) and the other for decoding (composed of a succession of decoder layers). Each encoder layer applies convolution, batch normalization and non-linearity, then applies a max pooling on the result. Decoders are similar to encoders, but they oversample their input, using indices stored from the encoding step. After the final decoder, the output is sent to a classifier which gives the final prediction corresponding to the segmentation of the image given as input (each channel of the prediction corresponds to a class of objects to segment).
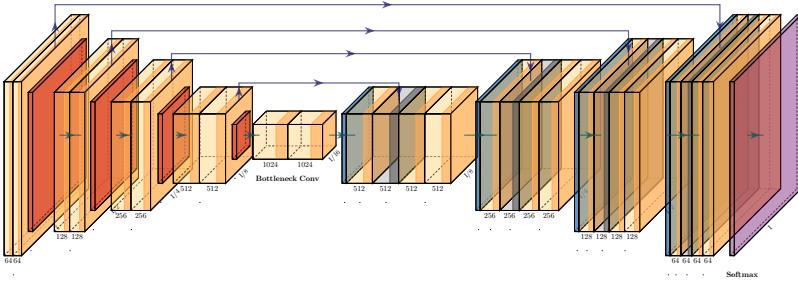

Fig. 6: Example of CNN for segmentation, the UNet architecture[47].

More recently, He *et al.* [21] proposed a novel approach (Mask-R-CNN) to detect objects in an image while simultaneously generating a high-quality segmentation mask for each instance. The architecture of this ANN is presented on Fig. 7. It is based on the Faster-R-CNN network proposed in [46]. The Mask-R-CNN network is composed of the classical R-CNN performing the classification and bounding box regression and of a branch for predicting segmentation masks on each instance extracted.

All these methods have been applied to histopathological data mainly for two types of applications: the segmentation of cell nuclei to go beyond an estimation of cell density by colorimetric deconvolution of images, and the segmentation of larger or composite objects of interest (tumor, glomeruli, lobules, etc.).

Cell segmentation The analysis of histopathological images and the resulting diagnosis are mainly done by quantifying the immune or cancerous cells present in the biopsy or by observing the morphology of the cells. To go beyond simply estimating the density of cells or to analyze the morphology of the cell nucleus,
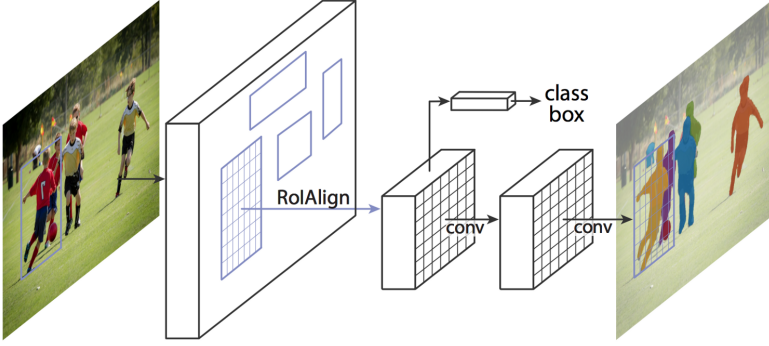
Fig. 7: The Mask-R-CNN architecture for instance segmentation [21].

it is important to segment each cell precisely and individually. This is why a lot of work is being conducted on cell segmentation in WSI of any type (brightfield or immunofluorescence).

Naylor *et al.* [39] studied multiple DL segmentation methods on a same dataset (publicly available [1]). Three different approaches were compared: FCN [35], *PangNet* [43] and *DeconvNet* [41]. The authors proposed a post-processing of the probability map which resulted in a F-score of around 0.8 for the individual cell segmentation.

Many recent work have been proposed based on new architectures dedicated to this specific problem [51,40,57], using adversarial ANN [3] or Mask-R-CNN [71].

Large regions of interest or composite objects segmentation The ANNs presented in the previous section have also been applied to larger objects segmentation. Indeed, they are very efficient at capturing the texture and shape of composite or complex objects, since they are based on convolutional layers at several levels of resolution (encoder part of the ANN).

For example, they were successfully applied many times in breast cancer for region of interest segmentation (stroma, tumor area, etc.) [60,29,72] or for objects segmentation like lobules [2] or glands [10]. Other organs or pathologies were also investigated like, for example, colon cancer [9,24,68,64,56], brain cancer [73,61] or glomeruli segmentation in renal pathology [65,14,25].

3 Challenges and opportunities

As seen before, artificial learning methods, notably ANNs, have demonstrated their interest in the field of computer vision and more specifically in the analysis

---

[1] https://peterjacknaylor.github.io/

of histopathological images. However, there remains many challenges to be overcome before seeing the emergence of tools in clinical routine that are sufficiently reliable, generic and whose diagnostic proposals are explicable. The main challenge, far more important than the definition of new learning or image analysis algorithms, concerns the data themselves. In fact, the majority of new methods that are emerging in computer vision and that could be applied to medical images come up against a problem of access to data, the creation of a sufficiently large learning set (annotations) and/or data quality (heterogeneity of data, etc.) in the majority of projects.

3.1 Annotations

Acquiring annotated data is difficult and time consuming in all ML project. While large databases of annotated simple objects (e.g. car, dog, car, bridge, etc.) are now available through public datasets like *ImageNet* [15] or COCO [34], large repositories of annotated biomedical images are still rare. This is probably due to the high degree of qualification required to discriminate between pathology from images compared to annotating a dog or a cat in an image.

This problem is particularly present in DL where ANN models generally require thousands of images to be trained efficiently. Furthermore, medical images are sensitive data, and they also requires a carefully data privacy policy to be able to share and distribute the images. In addition, expensive hardware is often required to capture the images compared to everyday objects that can be obtained using a simple smart-phone. However, despite these limitations, some current projects are now available and contain medical images, like *The Cancer Genome Atlas Program* (TCGA) [69].

However, the data are rarely annotated and only the raw images are available. Another point concerns the high degree of class unbalancing that can appear in medical image applications. For example, for many tasks, (e.g. mitosis detection), a negative label ("no mitosis") is much more frequent than the positive label ("mitosis"). This can disturb the training of an ANN that will naturally over-fit the most present class.

Multiple solutions are currently developed by the ML community to try to cope with these limitations. For example, data augmentation [44] which consists in creating synthetic data from real data is most of the time used in current DL model. Performing data augmentation allows to perform an explicit regularization of the ANN. An alternative to obtain more annotated data is to use crowdsourcing [20], which consists in using a population of annotators to perform annotations manually. It has been shown that the level of expertise of the annotators can be limited [20] and reliable annotation can be obtained by merging of combining multiple "weak" annotations of the same image (i.e. the power of the crowd [1]). Finally, domain adaptation or transfer learning [54] consists in using a pretrained ANN instead of using a brand new, randomly initialized model. This is particularly promising as it allows to avoid starting "from scratch" every time a new task has to be performed. It has been shown in [27] that results ob-

tained using pretrained ANNs are quite competitive compared to newly trained ones.

Finally, recent results [58,8] have demonstrated that models can be transferred and adapted to solve different, but related, tasks.

3.2 Multiple stainings

An important part of digital pathology is the analysis of multiple digitized WSI from differently stained tissue sections. It is common practice to mount consecutive sections containing corresponding microscopic structures on glass slides, and to stain them differently to highlight specific tissue components. These multiple staining modalities result in very different images but include a significant amount of consistent image information. DL approaches have recently been proposed to analyze these images in order to automatically identify objects of interest for pathologists.

The analysis and integration of information from different stainings is usually performed with reference to a specific organ, structure, or pattern observed in the tissue. For example, to diagnose pathologies such as breast cancer or kidney allograft rejection it is necessary to study the inflammatory microenvironment of the organ. In these cases, the relevant information is the distribution of immune cells (e.g. macrophages or lymphocytes) in relation to important structures of the organ, such as glomeruli for the kidney, or lobules for the breast.

To automatically perform such an analysis, the structure of interest (glomeruli, lobules, etc.) should be detected in each section irrespective of the individual staining modality. In order to avoid having to annotate each staining, strategies have been proposed in [31] to build an ANN that is robust to color variations between the WSI.

Another source of heterogeneity comes from slides originating from different hospital or centers. Indeed, depending of the operator preparing the slide, section thickness, chemical formulations, lab protocols etc. two images, even using the same staining, can look very different [37]. In order to address this problem of inter-centric heterogeneity, several teams proposed approaches based on domain adaptation [16,30].

3.3 Generative Adversarial Network

A *Generative Adversarial Network* (GAN) is a class of ML systems invented by Ian Goodfellow [19]. These models are composed of a generative ANN that generates candidates and a discriminate network that evaluates them. The generative ANN learns to map from a latent space to a data distribution, while the discriminative ANN distinguishes candidates produced by the generator from the true data. GAN is use extensively in computer vision and are starting to be used in DP too. For example, Neslihan *et al.* [5] proposed a method that uses dimension reduction and conditional adversarial generative ANNs to transform unstained hyperspectral tissue image to their *Haematoxylin&Eosin* (H&E) equivalent. The

goal is to create a virtual digital H&E staining that could automate some of the tasks in the diagnostic pathology workflow.

Zanjani *et al.* [74] also explored how GAN could be used to normalize stainings. By replacing the latent representation of a source image with those extracted from a template image in the trained model, the proposed model can generate a new color copy of the source image while preserving the important tissue structures.

Alternatively, Burlingame *et al.* [7] proposed the SHIFT method that uses GAN to translate histopathological images to immunofluorescent images. This method has the potential to improve our understanding of the mapping of histological and morphological profiles into protein expression profiles. decision-making.

GAN can also be used to cope with the lack of annotated data (see section 3.1). For example, Hou *et al.* [22] used GAN to synthesize histopathological images, in order to train supervised CNN with the generated data. The synthetic images are generated with the mask corresponding to the expected segmentation and adapted to the reference style. This enables to boost the performance of the trained CNN by using onthe-fly generated adversarial examples.

4   Conclusion

Thanks to the recent advent and increasing cost-effectiveness of digital scanners, tissue histopathology slides can now be completely digitized and stored as digital images. With the availability and analysis of a much larger set of variables combined with sophisticated imaging and analytic techniques, the traditional paradigm of pathology and microscopy could quickly be complemented and potentially also partially replaced by digital pathology, based on a screen-based visualization of digital tissue sections and an analysis combining a pathologist and a computer-based diagnostic aid system. In this chapter, we presented challenges and opportunities of using DL techniques to process these data. We discussed current DL models for digital pathology and discussed the problems related to the acquisition of annotations and the heterogeneity of the data (e.g. slides coming from multiples centers, different types of stainings, etc.). With the broad access to state-of-art of DL models and the availability of efficient computation systems, DL for digital pathology will continue to make progress in the coming years.

Acknowledgement

References

1. Shadi Albarqouni, Christoph Baur, Felix Achilles, Vasileios Belagiannis, Stefanie Demirci, and Nassir Navab. Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE transactions on medical imaging*, 35(5):1313–1321, 2016.

2. Grégory Apou, Nadine S Schaadt, Benoît Naegel, Germain Forestier, Ralf Schönmeyer, Friedrich Feuerhake, Cédric Wemmert, and Anne Grote. Detection of lobular structures in normal breast tissue. *Computers in biology and medicine*, 74:91–102, 2016.

3. Assaf Arbelle and Tammy Riklin Raviv. Microscopy cell segmentation via adversarial neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 645–648. IEEE, 2018.

4. Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 39(12):2481–2495, 2017.

5. Neslihan Bayramoglu, Mika Kaakinen, Lauri Eklund, and Janne Heikkila. Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 64–71, 2017.

6. Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828, Aug 2013.

7. Erik A Burlingame, Adam A Margolin, Joe W Gray, and Young Hwan Chang. Shift: speedy histopathological-to-immunofluorescent translation of whole slide images using conditional generative adversarial networks. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 1058105. International Society for Optics and Photonics, 2018.

8. Jongwon Chang, Jisang Yu, Taehwa Han, Hyuk-jae Chang, and Eunjeong Park. A method for classifying medical images using transfer learning: a pilot study on histopathology of breast cancer. In *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–4. IEEE, 2017.

9. Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical image analysis*, 36:135–146, 2017.

10. Hao Chen, Xiaojuan Qi, Lequan Yu, and Pheng-Ann Heng. Dcan: deep contour-aware networks for accurate gland segmentation. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 2487–2496, 2016.

11. Nicolas Coudray, Paolo Santiago Ocampo, Theodore Sakellaropoulos, Navneet Narula, Matija Snuderl, David Fenyö, Andre L Moreira, Narges Razavian, and Aristotelis Tsirigos. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature medicine*, 24(10):1559, 2018.

12. Angel Alfonso Cruz-Roa, John Edison Arevalo Ovalle, Anant Madabhushi, and Fabio Augusto González Osorio. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In Kensaku Mori, Ichiro Sakuma, Yoshinobu Sato, Christian Barillot, and Nassir Navab, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, pages 403–410, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

13. Kausik Das, Sailesh Conjeti, Abhijit Guha Roy, Jyotirmoy Chatterjee, and Debdoot Sheet. Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 578–581. IEEE, 2018.

14. Thomas de Bel, Meyke Hermsen, Bart Smeets, Luuk Hilbrands, Jeroen van der Laak, and Geert Litjens. Automatic segmentation of histopathological slides of renal tissue using deep learning. In *Medical Imaging 2018: Digital Pathology*, volume 10581, page 1058112. International Society for Optics and Photonics, 2018.

15. Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.

16. Michael Gadermayr, Martin Strauch, Barbara Mara Klinkhammer, Sonja Djudjaj, Peter Boor, and Dorit Merhof. Domain adaptive classification for compensating variability in histopathological whole slide images. In *International Conference on Image Analysis and Recognition*, pages 616–622. Springer, 2016.

17. Jérôme Galon, Anne Costes, Fatima Sanchez-Cabo, Amos Kirilovsky, Bernhard Mlecnik, Christine Lagorce-Pagès, Marie Tosolini, Matthieu Camus, Anne Berger, Philippe Wind, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*, 313(5795):1960–1964, 2006.

18. Baris Gecer, Selim Aksoy, Ezgi Mercan, Linda G. Shapiro, Donald L. Weaver, and Joann G. Elmore. Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks. *Pattern Recognition*, 84:345 – 356, 2018.

19. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

20. Anne Grote, Nadine S Schaadt, Germain Forestier, Cédric Wemmert, and Friedrich Feuerhake. Crowdsourcing of histological image labeling and object delineation by medical students. *IEEE transactions on medical imaging*, 2018.

21. Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.

22. Le Hou, Ayush Agarwal, Dimitris Samaras, Tahsin M Kurc, Rajarsi R Gupta, and Joel H Saltz. Unsupervised histopathology image synthesis. *arXiv preprint arXiv:1712.05021*, 2017.

23. Andrew Janowczyk and Anant Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of pathology informatics*, 7, 2016.

24. Philipp Kainz, Michael Pfeiffer, and Martin Urschler. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. *PeerJ*, 5:e3874, 2017.

25. Shruti Kannan, Laura A Morgan, Benjamin Liang, McKenzie G Cheung, Christopher Q Lin, Dan Mun, Ralph G Nader, Mostafa E Belghasem, Joel M Henderson,

Jean M Francis, et al. Segmentation of glomeruli within trichrome images using deep learning. *Kidney International Reports*, 2019.

26. Ansh Kapil, Armin Meier, Aleksandra Zuraw, Keith E Steele, Marlon C Rebelatto, Günter Schmidt, and Nicolas Brieu. Deep semi supervised generative learning for automated tumor proportion scoring on nsclc tissue needle biopsies. *Scientific reports*, 8(1):17343, 2018.

27. Brady Kieffer, Morteza Babaie, Shivam Kalra, and Hamid R Tizhoosh. Convolutional neural networks for histopathology image classification: Training vs. using pre-trained networks. In *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 1–6. IEEE, 2017.

28. Daisuke Komura and Shumpei Ishikawa. Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16:34–42, 2018.

29. V Kovalev, A Kalinovsky, and V Liauchuk. Deep learning in big image data: Histology image classification for breast cancer diagnosis. In *Big Data and Advanced Analytics, Proc. 2nd International Conference, BSUIR, Minsk*, pages 44–53, 2016.

30. Maxime W Lafarge, Josien PW Pluim, Koen AJ Eppenhof, Pim Moeskops, and Mitko Veta. Domain-adversarial neural networks to address the appearance variability of histopathology images. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 83–91. Springer, 2017.

31. Thomas Lampert, Odyssée Merveille, Jessica Schmitz, Germain Forestier, Friedrich Feuerhake, and Cédric Wemmert. Strategies for training stain invariant cnns. *arXiv preprint arXiv:1810.10338*, 2018.

32. Yann LeCun, Yoshua Bengio, and Geoffrey E. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

33. Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

34. Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.

35. Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015.

36. Anant Madabhushi and George Lee. Image analysis and machine learning in digital pathology: Challenges and opportunities, 2016.

37. Raphaël Marée, Loïc Rollus, Benjamin Stévens, Renaud Hoyoux, Gilles Louppe, Rémy Vandaele, Jean-Michel Begon, Philipp Kainz, Pierre Geurts, and Louis Wehenkel. Collaborative analysis of multi-gigapixel imaging data using cytomine. *Bioinformatics*, 32(9):1395–1401, 2016.

38. James L McClelland, David E Rumelhart, PDP Research Group, et al. Parallel distributed processing. *Explorations in the Microstructure of Cognition*, 2:216–271, 1986.

39. Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Nuclei segmentation in histopathology images using deep neural networks. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 933–936. IEEE, 2017.

40. Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE transactions on medical imaging*, 38(2):448–459, 2018.

41. Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1520–1528, 2015.

42. Augustus Odena, Christopher Olah, and Jonathon Shlens. Conditional image synthesis with auxiliary classifier gans. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 2642–2651. JMLR. org, 2017.

43. Baochuan Pang, Yi Zhang, Qianqing Chen, Zhifan Gao, Qinmu Peng, and Xinge You. Cell nucleus segmentation in color histopathological imagery using convolutional networks. In *2010 Chinese Conference on Pattern Recognition (CCPR)*, pages 1–5. IEEE, 2010.

44. Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621*, 2017.

45. G. Quellec, G. Cazuguel, B. Cochener, and M. Lamard. Multiple-instance learning for medical image and video analysis. *IEEE Reviews in Biomedical Engineering*, 10:213–234, 2017.

46. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015.

47. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 - 18th International Conference Munich, Germany, October 5 - 9, 2015, Proceedings, Part III*, pages 234–241, 2015.

48. Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

49. Marie-Christine Rousselet, Sophie Michalak, Florence Dupré, Anne Croué, Pierre Bedossa, Jean-Paul Saint-André, and Paul Calès. Sources of variability in histological scoring of chronic viral hepatitis. *Hepatology*, 41(2):257–264, 2005.

50. David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning internal representations by error propagation. Technical report, California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

51. Monjoy Saha and Chandan Chakraborty. Her2net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation. *IEEE Transactions on Image Processing*, 27(5):2189–2200, 2018.

52. Monjoy Saha, Chandan Chakraborty, and Daniel Racoceanu. Efficient deep learning model for mitosis detection using breast histopathology images. *Computerized Medical Imaging and Graphics*, 64:29–40, 2018.

53. Joel Saltz, Rajarsi Gupta, Le Hou, Tahsin Kurc, Pankaj Singh, Vu Nguyen, Dimitris Samaras, Kenneth R Shroyer, Tianhao Zhao, Rebecca Batiste, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell reports*, 23(1):181–193, 2018.

54. Hoo-Chang Shin, Holger R Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M Summers. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE transactions on medical imaging*, 35(5):1285–1298, 2016.

55. Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

56. Korsuk Sirinukunwattana, Josien PW Pluim, Hao Chen, Xiaojuan Qi, Pheng-Ann Heng, Yun Bo Guo, Li Yang Wang, Bogdan J Matuszewski, Elia Bruni, Urko Sanchez, et al. Gland segmentation in colon histology images: The glas challenge contest. *Medical image analysis*, 35:489–502, 2017.

57. Jie Song, Liang Xiao, Mohsen Molaei, and Zhichao Lian. Multi-layer boosting sparse convolutional model for generalized nuclear segmentation from histopathology images. *Knowledge-Based Systems*, 176:40–53, 2019.

58. Yang Song, Ju Jia Zou, Hang Chang, and Weidong Cai. Adapting fisher vectors for histopathology image classification. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 600–603. IEEE, 2017.

59. Nikolas Stathonikos, Mitko Veta, André Huisman, and Paul J van Diest. Going fully digital: Perspective of a dutch academic pathology lab. *Journal of pathology informatics*, 4, 2013.

60. Hai Su, Fujun Liu, Yuanpu Xie, Fuyong Xing, Sreenivasan Meyyappan, and Lin Yang. Region segmentation in histopathological breast cancer images using deep convolutional neural network. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pages 55–58. IEEE, 2015.

61. Z Swiderska-Chadaj, T Markiewicz, J Gallego, G Bueno, B Grala, and M Lorent. Deep learning for damaged tissue detection and segmentation in ki-67 brain tumor specimens based on the u-net model. *Bulletin of the Polish Academy of Sciences. Technical Sciences*, 66(6), 2018.

62. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2818–2826, June 2016.

63. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

64. Jing Tang, Jun Li, and Xiangping Xu. Segnet-based gland segmentation from colon cancer histology images. In *2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pages 1078–1082. IEEE, 2018.

65. Maja Temerinac-Ott, Germain Forestier, Jessica Schmitz, Meyke Hermsen, JH Bräsen, Friedrich Feuerhake, and Cédric Wemmert. Detection of glomeruli in renal pathology by mutual comparison of multiple staining modalities. In *Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis*, pages 19–24. IEEE, 2017.

66. Jakub M Tomczak, Maximilian Ilse, Max Welling, Marnix Jansen, Helen G Coleman, Marit Lucas, Kikki de Laat, Martijn de Bruin, Henk Marquering, Myrtle J van der Wel, et al. Histopathological classification of precursor lesions of esophageal adenocarcinoma: A deep multiple instance learning approach. In *MIDL 2018*, 2018.

67. Margarita Udall, Maria Rizzo, Juliet Kenny, Jim Doherty, SueAnn Dahm, Paul Robbins, and Eric Faulkner. Pd-l1 diagnostic tests: a systematic literature review of scoring algorithms and test-validation metrics. *Diagnostic pathology*, 13(1):12, 2018.

68. Yves-Rémi Van Eycke, Cédric Balsat, Laurine Verset, Olivier Debeir, Isabelle Salmon, and Christine Decaestecker. Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise ihc biomarker quantification: A deep learning approach. *Medical image analysis*, 49:35–45, 2018.

69. John N Weinstein, Eric A Collisson, Gordon B Mills, Kenna R Mills Shaw, Brad A Ozenberger, Kyle Ellrott, Ilya Shmulevich, Chris Sander, Joshua M Stuart, Cancer Genome Atlas Research Network, et al. The cancer genome atlas pan-cancer analysis project. *Nature genetics*, 45(10):1113, 2013.

70. Tian Xia, Ashnil Kumar, Dagan Feng, and Jinman Kim. Patch-level tumor classification in digital histopathology images with domain adapted deep learning. In

2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 644–647. IEEE, 2018.

71. Xinpeng Xie, Yuexiang Li, Menglu Zhang, and Linlin Shen. Robust segmentation of nucleus in histopathology images via mask r-cnn. In *International MICCAI Brainlesion Workshop*, pages 428–436. Springer, 2018.

72. Jun Xu, Xiaofei Luo, Guanhao Wang, Hannah Gilmore, and Anant Madabhushi. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing*, 191:214–223, 2016.

73. Yan Xu, Zhipeng Jia, Yuqing Ai, Fang Zhang, Maode Lai, I Eric, and Chao Chang. Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 947–951. IEEE, 2015.

74. Farhad Ghazvinian Zanjani, Svitlana Zinger, Babak Ehteshami Bejnordi, Jeroen AWM van der Laak, and Peter HN de With. Stain normalization of histopathology images using generative adversarial networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 573–577. IEEE, 2018.