



Deep learning powers cancer diagnosis in digital pathology

Yunjie He, Hong Zhao ^{*}, Stephen T.C. Wong ^{*}

Systems Medicine and Bioengineering Department, Houston Methodist Cancer Center, Houston, TX, 77030, USA

ARTICLE INFO

Keywords:

Digital pathology
Microscopy image
AI
Deep learning
Graph neural networks
Cancer diagnosis

ABSTRACT

Technological innovation has accelerated the pathological diagnostic process for cancer, especially in digitizing histopathology slides and incorporating deep learning-based approaches to mine the subvisual morphometric phenotypes for improving pathology diagnosis. In this perspective paper, we provide an overview on major deep learning approaches for digital pathology and discuss challenges and opportunities of such approaches to aid cancer diagnosis in digital pathology. In particular, the emerging graph neural network may further improve the performance and interpretability of deep learning in digital pathology.

Diagnostic and surgical pathology relies on human interpretations of patterns and morphology. The process is slow, error-prone, and largely unaided by digital technology of the 21st century. Coupling the rapid improvement of deep learning models, vastly enhanced computing power, and reduced storage costs with the rising burden of cancer incidence, the shrinking number of pathologists, the wide-availability of whole slide digitization techniques, and the increasing pathologist case load while reimbursements are dropping, the potential of deploying deep learning in automating the interpretation of digitized pathology slides is tremendous. Tissue or biopsy slides provide an important basis for pathological diagnosis. Many clinical tasks requiring high reproducibility, low fault tolerance, and a high degree of dependence on experienced readers are suitable for applying deep learning for classification, segmentation, or object detection (Wong, 2019; Kononenko, 2001; Jiang et al., 2020; Bera et al., 2019).

Before digging in the digital pathology, let us briefly review the history of deep learning from the proposal of the multilayer perceptron (MP) neuron mathematical model in 1943 (McCulloch and Pitts, 1943) to now, where it has shined in several fields. In 1958, Rosenblatt proposed the first generation of neural network single-layer perceptron (Rosenblatt, 1958), which can distinguish basic shapes such as triangles and squares. But in 1969, Minsky proved that a single-layer perceptron could not solve the XOR problem and did not have a good learning mechanism (Minsky and Papert, 1971). In 1986, the second-generation neural network proposed by Rumelhart et al. (1986), with the back-propagation training method and sigmoid nonlinear mapping, provided an effective solution and triggered the second wave of neural networks. In 1989, LeCun and colleagues (LeCun et al., 1989) invented a

convolutional neural network (CNN) and applied it to digital handwriting recognition. However, the backpropagation method has a serious gradient vanishing problem, and the development of neural networks once fell into a trough. Then, in 2006, Hinton proposed a pre-training method by the auto-encoder to suppress the vanishing gradient problem and resurrected interest in CNN (Hinton and Salakhutdinov, 2006). In 2011, the Rectified Linear Unit (ReLU) activation function proposed by Glorot et al. effectively suppressed the vanishing gradient problem (Glorot et al., 2011). In 2012, Hinton et al. participated in the ImageNet image recognition competition and won the championship through the CNN network AlexNet (Krizhevsky et al., 2012) that far exceeded the classification performance of the second-place support vector machine (SVM) method. Methods such as ReLU and Dropout used in the AlexNet model to suppress the vanishing gradient problem and reduce overfitting can still be found in the recently developed deep learning models. Since then, deep learning has ushered in explosive growth. There are improvements made in the convergence process of deep learning, such as batch normalization (Ioffe and Szegedy, 2015), transfer learning (Pan and Yang, 2009), and skip-connection in ResNet (He et al., 2016a; Szegedy et al., 2016; He et al., 2016b). Meanwhile, there are diversification of tasks, such as RNN-based time series deep learning model (Bojanowski et al., 2015; Hochreiter and Schmidhuber, 1997; Schuster and Paliwal, 1997; Salehinejad et al., 2017), used for image generation generative adversarial networks (GANs) (Goodfellow et al., 2014; Mirza and Osindero, 2014; Radford et al., 2015; Salimans et al., 2016; Isola et al., 2017; Zhu et al., 2017; Karras et al., 2017; Zhang et al., 2017; Brock et al., 2018; Karras et al., 2019; Gui et al., 2020), object detection models (Zou et al., 2019; Jiao

^{*} Corresponding author.

E-mail addresses: hzhao@houstonmethodist.org (H. Zhao), st Wong@houstonmethodist.org (S.T.C. Wong).

<https://doi.org/10.1016/j.compmedimag.2020.101820>

Received 8 June 2020; Received in revised form 5 October 2020; Accepted 3 November 2020

Available online 11 December 2020

0895-6111/© 2020 Elsevier Ltd. All rights reserved.

et al., 2019; Liu et al., 2020), natural language processing (NLP) models (Torfi et al., 2020; Otter et al., 2020; Qiu et al., 2020; Minaee et al., 2020; Hu, 2019), and reinforcement learning models (Li et al., 2019; Moerland et al., 2020; Luong et al., 2019). Deep learning systems even surpasses humans in many applications, e.g., AlphaGo zero Silver et al. (2017).

Digital pathology initially referred to digitized histopathology slides using whole slide imaging technology. A large-scale comparison study on tumor specimens from over 1,992 patients concluded that the equivalent diagnostic performance between the pathologists using digitized whole-slide images (WSI) and the pathologists viewing conventional tissue slides (Mukhopadhyay et al., 2018). However, studies have shown promise of integrating deep learning-based approaches with human pathologist's diagnoses in improving diagnostic accuracy while reducing human error rate in cancer diagnosis (Wang et al., 2016; Steiner et al., 2018; Dimitriou et al., 2019). The fast evolving field of digital pathology now explores the use of deep learning-based approaches for the classification and diagnosis of digitized images (Bera et al., 2019; Niazi et al., 2019; Janowczyk and Madabhushi, 2016; Sig-naevsky et al., 2019; Becker et al., 2020).

For example, Nagpal et al. (2019) designed a two-stage deep learning model to help pathologists differentiate between benign and malignant prostate tumors and grade the cancer quantitatively. The first stage was a deep convolutional neural network-based regional Gleason pattern classification. It was trained using 912 slides with 112 million human pathologist-annotated images. At the second stage, the classifications of human pathologists on 1,159 slides were used for training the model. The reference standard included a dataset of 331 slides from 331 patients that were reviewed by independent pathologists and genitourinary pathologists. Performance of the deep learning model vs. 29 human pathologists were compared. The mean accuracy of the deep learning model was 0.7, significantly superior to the accuracy of 0.61 ($p = 0.002$) by the 29 pathologists. In terms of the Gleason Grade decision thresholds, the receiver operating characteristic (ROC) curves by the deep learning model achieved 0.95–0.96 at each threshold. Especially at Gleason Grade ≥ 4 , the deep learning model showed greater sensitivity and specificity than 9 out of 10 pathologists (Nagpal et al., 2019). On the other hand, Raverot et al. (2020) developed a multistage mitotic-cell-detection method based on Faster region convolutional neural network (Faster R-CNN) (Ren et al., 2015), post-processing and deep convolutional neural networks (deep CNNs) for breast cancer detection. Two breast cancer mitotic cell-detection image datasets, i.e., ICPR 2012 MITOSIS (Roux et al., 2013) and ICPR 2014 (Roux et al., 2014), were used to train the Faster R-CNN model. In stage 1, an image was input into the trained Faster R-CNN detector. In stage 2, post-processing was performed to reduce the false positives based on statistical, texture, shape, and color features. In stage 3, independently trained Resnet-50 (He et al., 2016a) and Densenet-201 (Huang et al., 2017) scores were fused by deep CNNs, and the final classification of the mitotic and non-mitotic cells was performed in stage 4 (Raverot et al., 2020). This method obtained high detection accuracy of mitotic cells in multiple datasets including ICPR 2012, ICPR 2014, and TUPAC16 datasets, and the precision was higher than other ten different methods published previously.

At this point, we can state that CNNs have greatly improved the diagnosis based on histopathological slides in digital pathology (Acs et al., 2020). However, in most situations the image is Euclidean while the content is usually not. Traditional CNN needs to learn the content through multiple layers. Due to cropping, however, traditional CNN cannot obtain information outside the field of view, resulting in performance degradation. Moreover, it is precisely because of this process of extracting non-Euclidean information that the learning process of CNN is difficult to comprehend. There are many research papers regarding the interpretability of deep learning (Gilpin et al., 2018; Chakraborty et al., 2017; Ahmad et al., 2018; Serag et al., 2019), and notably the emerging development of graph neural network (GNN) has

brought us a promising path (Zhou et al., 2018; Wu et al., 2020).

The first appearance of graphs in the field of mathematics can be traced back to the Königsberg bridges problem (Barnett, 2005). A graph is a kind of data structure that models a set of objects (nodes) and their relationships (edges). It can be stated that GNN is to graphs what CNN is to image processing. In the development of GNN, traditional deep learning methods have also been analogously applied to GNN, such as Attention mechanism (Veličković et al., 2017), Long short-term memory (LSTM) control gate (Si et al., 2019), etc. GNN is widely used in social networks, recommendation systems, physical systems, chemical molecular properties prediction, knowledge maps, and many other fields (Zhou et al., 2018). As a data structure with powerful performance capabilities, graphs combining with deep learning would become the next hot spot in the digital pathology.

The feature extraction ability of GNN is also stronger than CNN. For instance, Zhou et al. (2019) developed automatic methods for grading colorectal cancer (CRC). The authors stated that CRC grading usually uses small image patches, and it is impossible to obtain information on the micro-architecture of the entire tissue during the grading process. The authors proposed a cell-graph convolutional neural network (CGC-Net) model that combines the adaptive GraphSAGE graph convolution structure to obtain multi-level features (Hamilton et al., 2017). In comparison with traditional CNN, the proposed model achieves much better accuracy of 97 %. Under the theme of interpretability, GNN clearly shows its understanding of the relationship between objects in the graph without the support of additional models. For example, Sureka et al. (2020) pointed out that the medical community needs interpretable models to conduct due diligence and improve understanding of disease mechanisms and treatment, which is difficult for CNN to do. Thus, the authors adopted an approach to model a histology tissue as a graph of nuclei and developed a GNN framework based on the attention mechanism and node occlusion for disease diagnosis. The method highlights the relative contribution of each cell nucleus in the image and fits the mental model of the pathologists.

In summary, recent development and applications of deep learning have progressively transformed digital pathology, particularly with promising GNN models. Although GNN has certain shortcomings, such as over-smoothing and scalability, the relatively easy-to-analyze structure and stronger generalization ability of GNN reveal its potential in digital pathology.

Funding

This work was funded by NIHR01CA238727 to H.Z and S.T.C.W, and NIH U01CA253553, NIH R01CA251710, John S. Dunn Research Foundation, and T.T. and W.F. Chao Foundation to S.T.C.W.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

The authors would like to thank Drs. Dorothy Lewis and Rebecca Danforth for proofreading the manuscript.

References

- Acs, B., Rantalainen, M., Hartman, J., 2020. Artificial intelligence as the next step towards precision pathology. *J. Intern. Med.* 288 (Jul. (1)), 62–81. <https://doi.org/10.1111/joim.13030>. Epub 2020 Mar 3. PMID: 32128929.
- Ahmad, M.A., Eckert, C., Teredesai, A., 2018. Interpretable machine learning in healthcare. *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*.
- Barnett, J.H., 2005. Early writings on graph theory: Euler circuits and the Königsberg bridge problem. Part II: Historical Projects in Discrete Mathematics and Computer Science of Resources for Teaching Discrete Mathematics: Classroom Project, History Modules, and Articles. *MAA Notes* 74, 197–208.

- Becker, J.U., et al., 2020. Artificial intelligence and machine learning in nephropathology. *Kidney Int.* 98 (1), 65–75. <https://doi.org/10.1016/j.kint.2020.02.027>. ISSN 0085-2538.
- Bera, K., et al., 2019. Artificial intelligence in digital pathology—new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* 16 (11), 703–715.
- Bojanowski, P., Joulin, A., Mikolov, T., 2015. Alternative Structures for Character-Level RNNs. *arXiv preprint arXiv:1511.06303*.
- Brock, A., Donahue, J., Simonyan, K., 2018. Large Scale Gan Training for High Fidelity Natural Image Synthesis. *arXiv preprint arXiv:1809.11096*.
- Chakraborty, S., et al., 2017. Interpretability of deep learning models: a survey of results. In: 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). IEEE.
- Dimitriou, N., Arandjelović, O., Caie, P.D., 2019. Deep learning for whole slide image analysis: an overview. *Front. Med.* 6.
- Gilpin, L.H., et al., 2018. Explaining explanations: an overview of interpretability of machine learning. In: 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE.
- Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*.
- Goodfellow, I., et al., 2014. Generative adversarial nets. *Advances in Neural Information Processing Systems*.
- Gui, J., et al., 2020. A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications. *arXiv preprint arXiv:2001.06937*.
- Hamilton, W., Ying, Z., Leskovec, J., 2017. Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems*.
- He, K., et al., 2016a. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- He, K., et al., 2016b. Identity mappings in deep residual networks. In: *European Conference on Computer Vision*. Springer.
- Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. *Science* 313 (5786), 504–507.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780.
- Hu, D., 2019. An introductory survey on attention mechanisms in NLP problems. In: *Proceedings of SAI Intelligent Systems Conference*. Springer.
- Huang, G., et al., 2017. Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *arXiv preprint arXiv:1502.03167*.
- Isola, P., et al., 2017. Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Janowczyk, A., Madabhushi, A., 2016. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. *J. Pathol. Inform.* 7.
- Jiang, Y., et al., 2020. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun.* 40 (4), 154–166.
- Jiao, L., et al., 2019. A survey of deep learning-based object detection. *IEEE Access* 7, 128837–128868.
- Karras, T., et al., 2017. Progressive Growing of Gans for Improved Quality, Stability, and Variation. *arXiv preprint arXiv:1710.10196*.
- Karras, T., Laine, S., Aila, T., 2019. A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Kononenko, I., 2001. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif. Intell. Med.* 23 (1), 89–109.
- Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*.
- LeCun, Y., et al., 1989. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* 1 (4), 541–551.
- Li, G., et al., 2019. Human-centered reinforcement learning: a survey. *IEEE Trans. Hum. Mach. Syst.* 49 (4), 337–349.
- Liu, L., et al., 2020. Deep learning for generic object detection: a survey. *Int. J. Comput. Vis.* 128 (2), 261–318.
- Luong, N.C., et al., 2019. Applications of deep reinforcement learning in communications and networking: a survey. *IEEE Commun. Surv. Tutor.* 21 (4), 3133–3174.
- McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5 (4), 115–133.
- Minaee, S., et al., 2020. Deep Learning Based Text Classification: A Comprehensive Review. *arXiv preprint arXiv:2004.03705*.
- Minsky, M., Papert, S.A., 2017. *Perceptrons: An Introduction to Computational Geometry*. MIT Press.
- Mirza, M., Osindero, S., 2014. Conditional Generative Adversarial Nets. *arXiv preprint arXiv:1411.1784*.
- Moerland, T.M., Broekens, J., Jonker, C.M., 2020. Model-Based Reinforcement Learning: A Survey. *arXiv preprint arXiv:2006.16712*.
- Mukhopadhyay, S., et al., 2018. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (pivotal study). *Am. J. Surg. Pathol.* 42 (1), 39.
- Nagpal, K., et al., 2019. Development and validation of a deep learning algorithm for improving Gleason scoring of prostate cancer. *Npj Digit. Med.* 2 (1), 1–10.
- Niazi, M.K.K., Parwani, A.V., Gurcan, M.N., 2019. Digital pathology and artificial intelligence. *Lancet Oncol.* 20 (5), e253–e261.
- Otter, D.W., Medina, J.R., Kalita, J.K., 2020. A survey of the usages of deep learning for natural language processing. *IEEE Trans. Neural Netw. Learn. Syst.*
- Pan, S.J., Yang, Q., 2009. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22 (10), 1345–1359.
- Qiu, X., et al., 2020. Pre-Trained Models for Natural Language Processing: A Survey. *arXiv preprint arXiv:2003.08271*.
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. *arXiv preprint arXiv:1511.06434*.
- Raverot, V., et al., 2020. Age-and sex-specific TSH upper-limit reference intervals in the general French population: there is a need to adjust our actual practices. *J. Clin. Med.* 9 (3), 792.
- Ren, S., et al., 2015. Faster r-cnn: towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*.
- Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* 65 (6), 386.
- Roux, L., et al., 2013. Mitosis detection in breast cancer histological images an ICPR 2012 contest. *J. Pathol. Inform.* 4.
- Roux, L., et al., 2014. Mitos and atypia. *Image Pervasive Access Lab (IPAL)*. Agency Sci. Technol. & Res. Inst. Infocom Res., Singapore. Tech. Rep.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323 (6088), 533–536.
- Salehinejad, H., et al., 2017. Recent Advances in Recurrent Neural Networks. *arXiv preprint arXiv:1801.01078*.
- Salimans, T., et al., 2016. Improved techniques for training gans. *Advances in Neural Information Processing Systems*.
- Schuster, M., Paliwal, K.K., 1997. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* 45 (11), 2673–2681.
- Serag, A., et al., 2019. Translational AI and deep learning in diagnostic pathology. *Front. Med.* 6.
- Si, C., et al., 2019. An attention enhanced graph convolutional lstm network for skeleton-based action recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- Signaevsky, M., et al., 2019. Artificial intelligence in neuropathology: deep learning-based assessment of tauopathy. *Lab. Invest.* 99 (7), 1019.
- Silver, D., et al., 2017. Mastering the game of go without human knowledge. *Nature* 550 (7676), 354–359.
- Steiner, D.F., et al., 2018. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *Am. J. Surg. Pathol.* 42 (12), 1636.
- Sureka, M., et al., 2020. Visualization for Histopathology Images Using Graph Convolutional Neural Networks. *arXiv preprint arXiv:2006.09464*.
- Szegedy, C., et al., 2016. Inception-v4, Inception-Resnet and the Impact of Residual Connections on Learning. *arXiv preprint arXiv:1602.07261*.
- Torfi, A., et al., 2020. Natural Language Processing Advancements by Deep Learning: A Survey. *arXiv preprint arXiv:2003.01200*.
- Velicković, P., et al., 2017. Graph Attention Networks. *arXiv preprint arXiv:1710.10903*.
- Wang, D., et al., 2016. Deep Learning for Identifying Metastatic Breast Cancer. *arXiv preprint arXiv:1606.05718*.
- Wong, S.T., 2019. The adoption of artificial intelligence in cancer pathology and imaging. *Cancer and Society*. Springer, pp. 113–120.
- Wu, Z., et al., 2020. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.*
- Zhang, H., et al., 2017. Stackgan: text to photo-realistic image synthesis with stacked generative adversarial networks. *Proceedings of the IEEE International Conference on Computer Vision*.
- Zhou, J., et al., 2018. Graph Neural Networks: A Review of Methods and Applications. *arXiv preprint arXiv:1812.08434*.
- Zhou, Y., et al., 2019. Cgc-net: cell graph convolutional network for grading of colorectal cancer histology images. *Proceedings of the IEEE International Conference on Computer Vision Workshops*.
- Zhu, J.-Y., et al., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the IEEE International Conference on Computer Vision*.
- Zou, Z., et al., 2019. Object Detection in 20 Years: A Survey. *arXiv preprint arXiv:1905.05055*.