Editorial Commentary

Role of artificial intelligence in integrated analysis of multi-omics and imaging data in cancer research

Nam Nhut Phan[1,2], Amrita Chattopadhyay[2], Eric Y. Chuang[2,3]

[1]Bioinformatics Program, Taiwan International Graduate Program, Institute of Information Science, Academia Sinica, Taipei; [2]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei; [3]Biomedical Technology and Device Research Laboratories, Industrial Technology Research Institute, Hsinchu

Correspondence to: Eric Y. Chuang. Department of Electrical Engineering, Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 10617. Email: chuangey@ntu.edu.tw.

Comment on: Saltz J, Gupta R, Hou L, *et al.* Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. Cell Rep 2018;23:181-93.e7.

Submitted Nov 03, 2019. Accepted for publication Dec 05, 2019.
doi: 10.21037/tcr.2019.12.17
View this article at: http://dx.doi.org/10.21037/tcr.2019.12.17

In recent years, machine learning and deep learning-based approaches, two sub-fields of artificial intelligence, have emerged as key components in biomedical data analyses (1-5). They can be applied to image segmentation, identifying insertion/deletion mutations, protein alignments, and so on. Several studies have integrated pathological image data with genomics data. Yuan *et al.* have quantitatively analyzed image data to better describe and validate the independent prognostic factors in estrogen receptor-negative breast cancer (6). Another study by Copper *et al.* also used histopathology images and genomics data to identify prognostic factors in breast cancer (7). Other types of cancers such as prostate cancer (8), renal cell carcinoma (9), low grade glioma (10), and non-small cell lung cancer (11), just to name a few, have also been studied by approaches integrating (multi-) omics data with pathology images.

The literature on deep learning methods used to assist cancer diagnosis and predict patient outcomes enables us to observe the exploding trend in this field (12-15). Massive amounts of published research and large numbers of clinical trials have illustrated the reliability and practicality of machine learning approaches, particularly deep learning. Various studies have employed deep learning methods to auto-detect and classify benign nuclei from cancer cells (1,16,17), to identify and quantify the rate and amount of mitosis (18). Deep learning has also been used for tissue origin classification, nuclear grading, precision medicine matching trials (1,19,20), classification of ancient and modern DNA (21), and drug repurposing (22).

For tissue quantification, there are two primary methods, namely handicraft features and the unsupervised approach (23). The former method consists of domain-agnostic and domain-inspired features (24,25) whereas the latter uses an automatic approach to identify distinguishing features (26). Domain-agnostic features focus on nuclear appearance, gland shape, object size, tissue texture, and architecture, while the domain-inspired features focus on certain particular domains, such as disease and organ origin (27). There have been studies that have applied these methods in prostate cancer and triple negative breast cancer (TNBC) samples (27). Gland architecture has been correlated by the domain-inspired approach (25) with aggressiveness of intermediate-risk prostate cancer. Another study calculated the number of intra-tumor lymphocytes, adjacent lymphocytes, and distant site tumor lymphocytes from TNBC (26). These studies found that these cell types and numbers can be used as independent prognostic predictors of disease-specific survival in TNBC (26). Tissue microarrays have also been used to predict colorectal cancer patient outcomes by deep learning approaches (28). The advantage of deep learning is that it is quick and seamless, although feature interpretability is missing (23).

In recent years, various tools related to pattern recognition have also been developed, and huge numbers of datasets are now readily available for public use. Many archives and databases for radiological and pathological images have been established as well, such as The Cancer Imaging Archive (TCIA) and the Cancer Digital Slide

Archive (CDSA), both of which facilitate image data analyses. Taking advantage of these databases and archives, many studies have been published with MRI and/or CT imaging incorporated with biological pathways and cellular morphology to further characterize a disease (29-34). These radiological data could potentially aid in determining the molecular subtypes of cancer.

Furthermore, radiological data could be linked to gene expression and/or mutation profiles to identify distinct cellular subtypes within the same cancer. Radiological data comprising hundreds of thousands of cells within a patient, once coupled with gene expression, could decipher the multiple dimensional features of the tumor, which is not achievable with genomic data alone (35). Consequently, the integration of radiological features with genomic data undoubtedly has a crucial role to play in improving diagnostic, prognostic, and predictive power in comparison to conventional approaches such as immunohistochemical assays. There are a couple of research papers that have applied radiological and genomic data to discriminate prostate cancer tissues from benign tissues, thereby enhancing information related to prostate cancer aggressiveness (36). Another study conducted with lung adenocarcinoma integrated CT images to predict the metastatic potency driving cells to distant organs (37).

Radiomics has been shown to be powerful in parallel with genetic markers with extracted semantic and agnostic features. Integrating multiple platforms to bridge radiomics with genomics could lead to better characterization of disease. This is of particular value for better treatment decisions and correct explanation of biological and treatment heterogeneity. Especially for cancer therapy, image-aided decision-making is crucial. For instance, the integrated radiomics and pathological features corresponding to a specific breast cancer molecular signature can provide prognosis markers and surrogates to predict patient outcomes, drug responsiveness, and eventually enhance treatment efficacy. Prospective studies using artificial intelligence as the predominant tool for classification and/or prediction tasks, along with omics and imaging data, could certainly facilitate and accelerate research output and accuracy. However, machine learning requires informed human supervision, as results without proper interpretation are not of much value.

Acknowledgments

We thank Melissa Stauffer, PhD, for editing the manuscript.

Funding: None.

Footnote

Provenance and Peer Review: This article was commissioned by the Editorial Office, *Translational Cancer Research*. It has not been sent for external peer review.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2019.12.17). EYC serves as the unpaid Editor-in-Chief of Translational Cancer Research. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. Saltz J, Gupta R, Hou L, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. Cell Rep 2018;23:181-93.e7.
2. Kourou K, Exarchos TP, Exarchos KP, et al. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2014;13:8-17.
3. Vial A, Stirling D, Field M, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. Transl Cancer Res 2018;7:803-16.
4. Cruz-Roa A, Gilmore H, Basavanhally A, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. Sci Rep 2017;7:46450.
5. Xu J, Luo X, Wang G, et al. A deep convolutional neural network for segmenting and classifying epithelial

and stromal regions in histopathological images. Neurocomputing 2016;191:214-23.
6. Yuan Y, Failmezger H, Rueda OM, et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. Sci Transl Med 2012;4:157ra143.
7. Cooper LA, Kong J, Gutman DA, et al. Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images. Lab Invest 2015;95:366-76.
8. Robinson D, Van Allen EM, Wu YM, et al. Integrative clinical genomics of advanced prostate cancer. Cell 2015;161:1215-28.
9. Cheng J, Zhang J, Han Y, et al. Integrative Analysis of Histopathological Images and Genomic Data Predicts Clear Cell Renal Cell Carcinoma Prognosis. Cancer Res 2017;77:e91-e100.
10. Cancer Genome Atlas Research Network, Brat DJ, Verhaak RG, et al. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. N Engl J Med 2015;372:2481-98.
11. Yu KH, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 2016;7:12474.
12. Zhang Z, Beck MW, Winkler DA, et al. Opening the black box of neural networks: methods for interpreting neural network models in clinical applications. Ann Transl Med 2018;6:216.
13. Xiong Y, Ba X, Hou A, et al. Automatic detection of mycobacterium tuberculosis using artificial intelligence. J Thorac Dis 2018;10:1936-40.
14. Klang E. Deep learning and medical imaging. J Thorac Dis 2018;10:1325-8.
15. Yang Y, Feng X, Chi W, et al. Deep learning aided decision support for pulmonary nodules diagnosing: a review. J Thorac Dis 2018;10:S867-75.
16. Sirinukunwattana K, Ahmed Raza SE, Yee-Wah Tsang, et al. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 2016;35:1196-206.
17. Xu J, Xiang L, Liu Q, et al. Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. IEEE Trans Med Imaging 2016;35:119-30.
18. Romo-Bucheli D, Janowczyk A, Gilmore H, et al. A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers. Cytometry A 2017;91:566-73.
19. Coudray N, Ocampo PS, Sakellaropoulos T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med 2018;24:1559.
20. Korbar B, Olofson AM, Miraflor AP, et al. Deep learning for classification of colorectal polyps on whole-slide images. J Pathol Inform 2017;8:30.
21. Zou J, Huss M, Abid A, et al. A primer on deep learning in genomics. Nat Genet 2019;51:12-8.
22. Aliper A, Plis S, Artemov A, et al. Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. Mol Pharm 2016;13:2524-30.
23. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: Challenges and opportunities. Med Image Anal 2016;33:170-5.
24. Veta M, van Diest PJ, Willems SM, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med Image Anal 2015;20:237-48.
25. Lee G, Sparks R, Ali S, et al. Co-occurring gland angularity in localized subgraphs: predicting biochemical recurrence in intermediate-risk prostate cancer patients. PloS One 2014;9:e97954.
26. Yuan Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. J R Soc Interface 2015. doi: 10.1098/rsif.2014.1153.
27. Jafari-Khouzani K, Soltanian-Zadeh H. Multiwavelet grading of pathological images of prostate. IEEE Trans Biomed Eng 2003;50:697-704.
28. Bychkov D, Turkki R, Haglund C, et al. Deep learning for tissue microarray image-based outcome prediction in patients with colorectal cancer. Proceedings 2016;9791, Medical Imaging 2016: Digital Pathology; 979115.
29. Lu CF, Hsu FT, Hsieh KLC, et al. Machine learning–based radiomics for molecular subtyping of gliomas. Clin Cancer Res 2018;24:4429-36.
30. Schellinger PD, Meinck HM, Thron A. Diagnostic accuracy of MRI compared to CCT in patients with brain metastases. J Neurooncol 1999;44:275-81.
31. Fiebach J, Schellinger P, Jansen O, et al. CT and diffusion-weighted MR imaging in randomized order: diffusion-weighted imaging results in higher accuracy and lower interrater variability in the diagnosis of hyperacute

ischemic stroke. Stroke 2002;33:2206-10.
32. Antoch G, Vogt FM, Freudenberg LS, et al. Whole-body dual-modality PET/CT and whole-body MRI for tumor staging in oncology. JAMA 2003;290:3199-206.
33. Grossmann P, Gutman DA, Dunn WD, et al. Imaging-genomics reveals driving pathways of MRI derived volumetric tumor phenotype features in Glioblastoma. BMC Cancer 2016;16:611.
34. Colen RR, Vangel M, Wang J, et al. Imaging genomic mapping of an invasive MRI phenotype predicts patient outcome and metabolic dysfunction: a TCGA glioma phenotype research group project. BMC Med Genomics 2014;7:30.
35. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563-77.
36. Wibmer A, Hricak H, Gondo T, et al. Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. Eur Radiol 2015;25:2840-50.
37. Coroller TP, Grossmann P, Hou Y, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. Radiother Oncol 2015;114:345-50.

Cite this article as: Phan NN, Chattopadhyay A, Chuang EY. Role of artificial intelligence in integrated analysis of multi-omics and imaging data in cancer research. Transl Cancer Res 2019;8(8):E7-E10. doi: 10.21037/tcr.2019.12.17

