

Cancer Metastasis Detection via Effective Contrastive Learning

Subjects: **Computer Science, Artificial Intelligence**

Contributor: Haixia Zheng , Yu Zhou , Xin Huang

The metastasis detection in lymph nodes via microscopic examination of H&E stained histopathological images is one of the most crucial diagnostic procedures for breast cancer staging. The manual analysis is extremely labor-intensive and time-consuming because of complexities and diversities of histopathological images. Deep learning has been utilized in automatic cancer metastasis detection in recent years. The success of supervised deep learning is credited to a large labeled dataset, which is hard to obtain in medical image analysis. Contrastive learning, a branch of self-supervised learning, can help in this aspect through introducing an advanced strategy to learn discriminative feature representations from unlabeled images.

convolutional neural network

contrastive learning

self-supervision

deep learning

1. Introduction

Cancer is currently one of the major causes of death for people all over the world. It is estimated that 14.5 million people have died of cancer, and by 2030 this figure is expected to exceed 28 million. The most common cancer for women is breast cancer. Every year, 2.1 million people around the world are diagnosed with breast cancer, according to World Health Organization (WHO) [1]. Due to the high rate of mortality, considerable efforts were made in the past decade to detect breast cancer from histological images so as to improve survival through early breast tissue diagnosis.

Since lymph node is the first position of breast cancer metastasis, metastasis identification of lymph node is one of the most essential criteria for early detection. In order to analyze the characteristics of tissues, pathologists examine tissue slices under the microscope [2]. The tissue slices are traditionally directly observed with a histopathologist's naked eyes and visual data are assessed manually based on prior medical knowledge. The manual analysis is highly time consuming and labor expensive due to the intricacies and diversities of histopathological images. At the same time, highly depending on histopathologist's expertise, workload, and current mood, the manual diagnostic procedure is subjective and limited repeatability. In addition, in the face of escalating demands for diagnostics with increased cancer incidence, there is a serious shortage of pathologists [3]. Hundreds of biopsies must be diagnosed daily by pathologists, thus it is almost impossible to thoroughly examine the entire slides. However, if only regions of interest are investigated, the chance of incorrect diagnosis may increase. To this end, in order to increase the efficiency and reliability of pathological examination, it is required to develop automatic detection techniques.

However, automated metastasis identification in sentinel lymph node from whole-slide image (WSI) is extremely challenging for the following reasons: first, the hard imitations in normal tissues usually look similar in morphology to metastatic areas, which leads to many false positives; second, the great varieties in biological structures and textures of metastatic and background areas; third, the varied circumstances of histological image processing, such as staining, cutting, sampling and digitization, enhance the variations of the appearance of image. This usually happens while tissue samples are taken at different time points or from different patients. Last but not least, WSI is incredibly huge, around $100,000 \text{ pixels} \times 200,000 \text{ pixels}$, and may not be directly input into any emerging method for cancer identification. Therefore, one of the major issues for automatic detection algorithms is how to analyze such a large pixel image effectively.

Artificial Intelligence (AI) technologies have developed rapidly in recent years. Especially in computer vision, image processing, and analysis, they have achieved outstanding breakthroughs. In histopathological diagnosis, AI has also exhibited potential advantages. With the help of AI-assisted diagnostic approaches, valuable information about diagnostics may be speedily extracted from big data, alleviating the workload of pathologists. At the same time, AI-aided diagnostics have more objective analysis capabilities and can avoid subjective discrepancies of manual analysis. To a certain extent, the use of artificial intelligence can not only improve work efficiency, but also reduce the rate of misdiagnosis by pathologists.

In the past few decades, a lot of works for breast histology image recognition have been developed. Early research used hand-made features to capture tissue properties in a specific area for automatic detection [4][5][6]. However, hand-made features are not sufficiently discriminative to describe a wide variety of shapes and textures. Recently, a deep Convolutional Neural Network (CNN) has been utilized to detect cancer metastases that can learn more effective feature representation and obtain higher detection accuracy in a data-driven approach [7][8][9]. The primary factor that may degrade the performance of CNN-based detection methods is the insufficiency of training samples, which may cause overfitting during the training process. In most medical circumstances, it is unrealistic to require understaffed radiologists to spend time creating such huge annotation sets for every new application. Therefore, in order to address the problem of lack of sufficient annotated data samples, it is critical to build less data-hungry algorithms capable of producing excellent performance with minimal annotations.

Self-supervised learning is a new unsupervised learning paradigm that does not require data annotations. In this entry, researchers propose a multi-task setting where researchers train the backbone model through joint supervision from the supervised detection target-task and an additional self-supervised contrastive learning task, as shown in **Figure 1**. Unlike most multi-task cases, where the goal is to achieve desired performance on all tasks at the same time, researchers' aim is to enhance the performance of the backbone model through exploiting the supervision from the additional contrastive learning task. More specifically, researchers extend the initial training loss with an extra self-supervised contrastive loss. As a result, the artificially augmented training task contributes to learning a more diverse set of features. Furthermore, researchers can incorporate unlabeled data to the training process, since self-supervision does not need labeled data. Through increasing the number and diversity of training data in this semi-supervised manner, one may expect to acquire stronger image features and achieve further performance improvement.

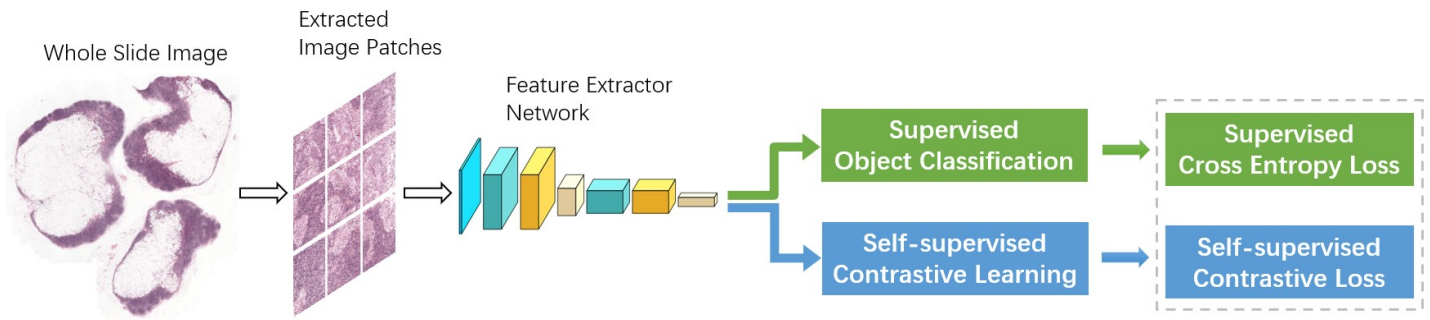


Figure 1. Overview of the proposed architecture.

2. Breast Cancer Detection

In earlier years, most designed approaches employed hand-crafted features. Spanhol et al. demonstrate classification performance based on several hand-made textural features for distinguishing malignant from benign [4]. Some works merged two or more hand-made features to enhance the accuracy of detection. In [5], graph, haralick, Local Binary Patterns (LBP), and intensity features were used for cancer identification in H&E stained histopathological images. The histopathological images were represented via fusing color histograms, LBP, SIFT, and some efficient kernel features, and the significance of these pattern features was also studied in [6]. However, it needs considerable efforts to design and validate these hand-made features. In addition, the properties of tissues with great variations in morphologies and textures cannot properly be represented, and consequently their detection performance is poor.

With the emergence of powerful computers, deep learning technology has made remarkable progress in a variety of domains, including natural language understanding, speech recognition, computer vision and image processing [10]. These methods have also been successfully employed in various modalities of medical images for detection, classification, and segmentation tasks [11]. Bejnordi et al. built a deep learning system to determine the stromal features of breast tissues associated with tumor for classifying Whole Slide Images (WSIs) [12]. Spanhol et al. utilized AlexNet to categorize breast cancer in histopathological images to be malignant and benign [7]. Bayramoglu et al. developed two distinct CNN architectures to classify breast cancer of pathology images [8]. Single-task CNN was applied to identify malignant tumors. Multi-task CNN has been used for analyzing the properties of benign and malignant tumors. The hybrid CNN unit designed by Guo et al. could fully exploit the global and local features of image, and thus obtain superior prediction performance [9]. Lin et al. proposed a dense and fast screening architecture (ScanNet) to identify metastatic breast cancer in WSIs [13][14]. In order to fully capture the spatial structure information between adjacent patches, Zanjani et al. [15][16] applied the conditional random field (CRF), whereas Kong et al. [17] employed 2D Long Short-Term Memory (LSTM) on patch features, respectively, which are first obtained from a CNN classifier. As the limited number of training samples in medical applications may be insufficient to learn a powerful model, some methods [18][19][20] transferred deep and rich feature hierarchies learned from a large number of cross-domain images, for which training data could be easily acquired.

3. Self-Supervised Contrastive Learning

Self-supervised learning is a new unsupervised learning paradigm. Recent research has shown that, by minimizing a suitable unsupervised loss during training, self-supervised learning can obtain valuable representations from unlabeled data [21][22][23][24]. The resulting network is a valid initialization for subsequent tasks.

The current revival of self-supervised learning started with intentionally devised annotation-free pretext tasks, such as colorization [25], jigsaw puzzle solving [23], relative patch prediction [21], and rotation prediction [24][26]. Although more complex networks and longer training time can yield good results [27], these pretext tasks more or less depend on ad-hoc heuristics, limiting the generality of learnt representations.

Contrastive learning is a discriminative technique that uses contrastive loss [28] to group similar instances closer together and dissimilar instances far apart from each other [29][30][31][32][33], as indicated in **Figure 2**. Similarity is defined in an unsupervised manner. It is usually considered that various transformations of an image are similar [34]. Ref. [35] employed domain-specific knowledge of videos to model the global contrastive loss. Authors in [36][37][38] maximized mutual information (MI) between global and local features from different layers of an encoder network, which is comparable to contrastive loss in implementation [39]. Some works utilized memory bank [29] or momentum contrast [30][31] to obtain more negative samples in each batch. MoCo [30], SimCLR [33], and SwAV [40] with modified algorithms generated similar performance with the state-of-the-art supervised method on the ImageNet dataset [41].

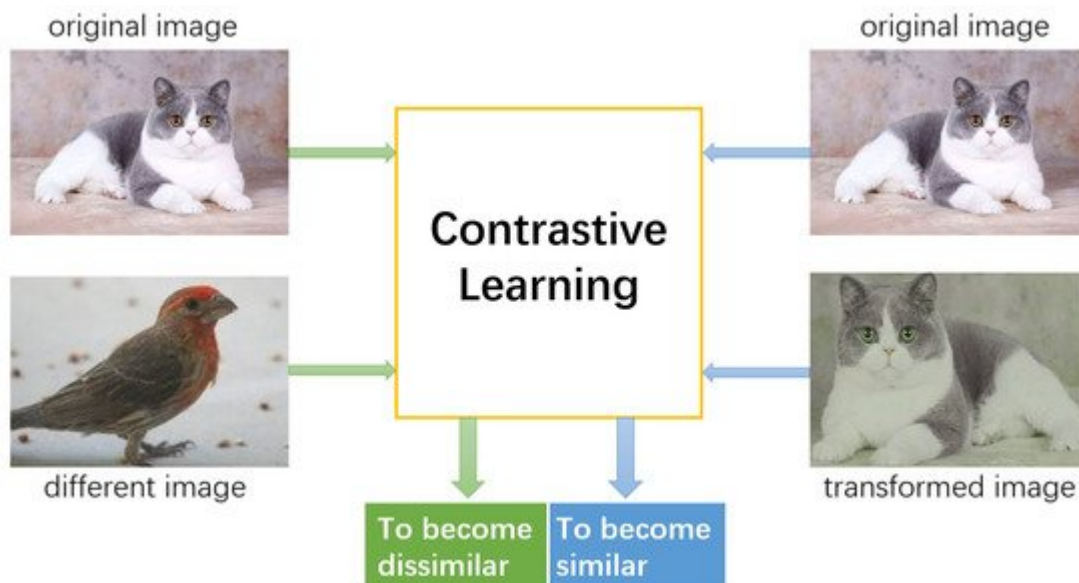


Figure 2. The core idea of contrastive learning: pushing the representations of original and transformed images closer together while separating the representations of original and different images far apart from each other.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 2018, 68, 394–424.
2. Ramos-Vara, J.A. Principles and methods of immunohistochemistry. *Methods Mol. Biol.* 2011, 691, 83–96.
3. Humphreys, G.; Ghent, A. World laments loss of pathology service. *Bull. World Health Organ.* 2010, 88, 564–565.
4. Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L. A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Trans. Biomed. Eng.* 2015, 63, 1455–1462.
5. Cruz-Roa, A.A.; Ovalle, J.; Madabhushi, A.; Osorio, F. A Deep Learning Architecture for Image Representation, Visual Interpretability and Automated Basal-Cell Carcinoma Cancer Detection. In *Proceedings of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention*, Nagoya, Japan, 22–26 September 2013; pp. 403–410.
6. Kandemir, M.; Hamprecht, F.A. Computer-aided diagnosis from weak supervision: A benchmarking study. *Comput. Med. Imaging Graph. Off. J. Comput. Med. Imaging Soc.* 2015, 42, 44–50.
7. Spanhol, F.; Oliveira, L.S.; Cavalin, P.R.; Petitjean, C.; Heutte, L. Deep features for breast cancer histopathological image classification. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Banff, AB, Canada, 5–8 October 2017; pp. 1868–1873.
8. Bayramoglu, N.; Kannala, J.; Heikkilä, J. Deep learning for magnification independent breast cancer histopathology image classification. In *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico, 4–8 December 2016; pp. 2440–2445.
9. Guo, Y.; Dong, H.; Song, F.; Zhu, C.; Liu, J. Breast Cancer Histology Image Classification Based on Deep Neural Networks. In *International Conference Image Analysis and Recognition*; Springer: Cham, Switzerland, 2018; Volume 10882, pp. 827–836.
10. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Essen, B.C.V.; Awwal, A.A.S.; Asari, V.K. The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. *arXiv* 2018, arXiv:1803.01164.
11. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; der Laak, J.A.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* 2017, 42, 60–88.
12. Ehteshami Bejnordi, B.; Linz, J.; Glass, B.; Mullooly, M.; Gierach, G.; Sherman, M.; Karssemeijer, N.; van der Laak, J.; Beck, A. Deep learning-based assessment of tumor-associated stroma for diagnosing breast cancer in histopathology images. In *Proceedings of the IEEE 14th International Symposium on Biomedical Imaging*, Melbourne, VIC, Australia, 18–21 April 2017; pp. 929–932.

13. Lin, H.; Chen, H.; Dou, Q.; Wang, L.; Qin, J.; Heng, P.A. ScanNet: A Fast and Dense Scanning Framework for Metastatic Breast Cancer Detection from Whole-Slide Images. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 539–546.
14. Lin, H.; Chen, H.; Graham, S.; Dou, Q.; Rajpoot, N.; Heng, P.A. Fast ScanNet: Fast and Dense Analysis of Multi-Gigapixel Whole-Slide Images for Cancer Metastasis Detection. *IEEE Trans. Med. Imaging* 2019, 38, 1948–1958.
15. Zanjani, F.G.; Zinger, S.; With, P. Cancer detection in histopathology whole-slide images using conditional random fields on deep embedded spaces. In Proceedings of the Digital Pathology, Houston, TX, USA, 6 March 2018.
16. Yi, L.; Wei, P. Cancer Metastasis Detection with Neural Conditional Random Field. *arXiv* 2018, arXiv:1806.07064.
17. Kong, B.; Xin, W.; Li, Z.; Qi, S.; Zhang, S. Cancer Metastasis Detection via Spatially Structured Deep Network. In International Conference Image Analysis and Recognition; Springer: Cham, Switzerland, 2017; pp. 236–248.
18. Xie, J.; Liu, R.; Luttrell, J.; Zhang, C. Deep Learning Based Analysis of Histopathological Images of Breast Cancer. *Front. Genet.* 2019, 10, 80.
19. de Matos, J.; de Souza Britto, A.; Oliveira, L.; Koerich, A.L. Double Transfer Learning for Breast Cancer Histopathologic Image Classification. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.
20. Kassani, S.H.; Kassani, P.H.; Wesolowski, M.J.; Schneider, K.A.; Deters, R. Breast Cancer Diagnosis with Transfer Learning and Global Pooling. In Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea, 16–18 October 2019; pp. 519–524.
21. Doersch, C.; Gupta, A.; Efros, A.A. Unsupervised Visual Representation Learning by Context Prediction. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1422–1430.
22. Pathak, D.; Krähenbühl, P.; Donahue, J.; Darrell, T.; Efros, A.A. Context Encoders: Feature Learning by Inpainting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2536–2544.
23. Noroozi, M.; Favaro, P. Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles. In Proceedings of the ECCV, Amsterdam, The Netherlands, 11–14 October 2016.
24. Gidaris, S.; Singh, P.; Komodakis, N. Unsupervised Representation Learning by Predicting Image Rotations. *arXiv* 2018, arXiv:1803.07728.

25. Zhang, R.; Isola, P.; Efros, A.A. Colorful Image Colorization. In Proceedings of the ECCV, Amsterdam, The Netherlands, 11–14 October 2016.
26. Chen, T.; Zhai, X.; Ritter, M.; Lucic, M.; Houlsby, N. Self-Supervised GANs via Auxiliary Rotation Loss. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 12146–12155.
27. Kolesnikov, A.; Zhai, X.; Beyer, L. Revisiting Self-Supervised Visual Representation Learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 16–20 June 2019; pp. 1920–1929.
28. Hadsell, R.; Chopra, S.; LeCun, Y. Dimensionality Reduction by Learning an Invariant Mapping. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; Volume 2, pp. 1735–1742.
29. Wu, Z.; Xiong, Y.; Yu, S.X.; Lin, D. Unsupervised Feature Learning via Non-parametric Instance Discrimination. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3733–3742.
30. He, K.; Fan, H.; Wu, Y.; Xie, S.; Girshick, R.B. Momentum Contrast for Unsupervised Visual Representation Learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 9726–9735.
31. Misra, I.; van der Maaten, L. Self-Supervised Learning of Pretext-Invariant Representations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 6706–6716.
32. Tian, Y.; Krishnan, D.; Isola, P. Contrastive Multiview Coding. In Proceedings of the ECCV, Glasgow, UK, 23–28 August 2020.
33. Chen, T.; Kornblith, S.; Norouzi, M.; Hinton, G.E. A Simple Framework for Contrastive Learning of Visual Representations. arXiv 2020, arXiv:2002.05709.
34. Dosovitskiy, A.; Springenberg, J.T.; Riedmiller, M.; Brox, T. Discriminative Unsupervised Feature Learning with Convolutional Neural Networks. In Proceedings of the Advances in Neural Information Processing Systems 27 (NIPS), Montreal, QC, Canada, 8–13 December 2014.
35. Tschannen, M.; Djolonga, J.; Ritter, M.; Mahendran, A.; Houlsby, N.; Gelly, S.; Lucic, M. Self-Supervised Learning of Video-Induced Visual Invariances. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 13803–13812.
36. Bachman, P.; Hjelm, R.D.; Buchwalter, W. Learning Representations by Maximizing Mutual Information Across Views. In Proceedings of the NeurIPS, Vancouver, BC, Canada, 8–14 December 2019.

37. Hénaff, O.J.; Srinivas, A.; Fauw, J.D.; Razavi, A.; Doersch, C.; Eslami, S.M.A.; van den Oord, A. Data-Efficient Image Recognition with Contrastive Predictive Coding. arXiv 2020, arXiv:1905.09272.
38. Hjelm, R.D.; Fedorov, A.; Lavoie-Marchildon, S.; Grewal, K.; Trischler, A.; Bengio, Y. Learning deep representations by mutual information estimation and maximization. arXiv 2019, arXiv:1808.06670.
39. Tschannen, M.; Djolonga, J.; Rubenstein, P.K.; Gelly, S.; Lucic, M. On Mutual Information Maximization for Representation Learning. arXiv 2019, arXiv:1907.13625.
40. Caron, M.; Misra, I.; Mairal, J.; Goyal, P.; Bojanowski, P.; Joulin, A. Unsupervised Learning of Visual Features by Contrasting Cluster Assignments. arXiv 2020, arXiv:2006.09882.
41. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, F.F. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Miami, FL, USA, 20–25 June 2009.

Retrieved from <https://encyclopedia.pub/entry/history/show/65842>