Predict Microsatellite Instability Status from Histological Images of Colorectal Cancer using Deep Learning

[AUTHORS] Shengjia Chen, Ruoxun Zi

1. Motivation and Related Work

Microsatellite instability (MSI) is a molecular maker of deficient DNA mismatch repair. It was found to have diagnostic, prognostic and therapeutic implications in colorectal cancer (CRC) (Vilar and Gruber, 2010; Overman et al., 2017), which is the third most common and second deadly cancer worldwide (Bray et al., 2018). Testing all CRC patients for MSI is recommended by multiple professional societies, which requires either an immunohistochemical analysis or a PCR-based assay (Cerretelli et al., 2020). However, universal testing for MSI has not been implemented due to cost and resource limitations (Shaikh et al., 2018). Recently, several studies have investigated the potential of deep learning to predict MSI directly from hematoxylin and eosin (H&E) stained slides. ResNet-18 was selected in many studies to assign an MSI likelihood to each tile within the tumor area. Different methods were used to classify MSI in the whole-slide image (WSI) level. Kather et al. (2020) determined the MSI status of each slide by the majority of its constituent tiles. Cao et al. (2020) used multiple instance learning to train the CNN for WSI classification with improved overall predictive accuracy. In this work, we proposed a deep learning model to classify MSI versus microsatellite stability (MSS) for a WSI.

2. Data

The dataset contains 152,078 image tiles from histological images of CRC cancer patients in The Cancer Genome Atlas (TCGA), which was created from WSIs by Kather (Kather, 2019). The original WSIs are available from the public repositories at NIH. The histological images came from 360 patients. The dataset is divided into training and testing sets by 70% and 30% on a patient level, which results in 260 and 100 patients in the training and testing set, separately, and 93,408 training tiles and 58,670 testing tiles. The training set is balanced with the same number of MSI and MSS image tiles. Each tile has the same size of 224x224 pixels. The label is assigned on a patient level.

3. Proposed Method

The proposed deep learning framework in this study will be Multiple instance learning (MIL) RNN models. (Campanella et al., 2019).

- MIL Feature Representation
    1. Inference. Pass all tiles of each slide through the training data. A classifier is trained at the tile level using CNN (ResNet architecture) to rank the tiles in each slide according to their probability of being positive (MSI as 1, MSS as 0). (He et al., 2015)
    2. Learning. Learn a tile-level representation that can linearly separate the discriminative tiles in positive slides from all other tiles. This representation will be used as input to slide-level aggregation.
- Slide-level Aggregation: select the group of top-ranking tiles per slide and aggregate them (max-pooling, naive multiscale aggregation, random forests, RNN). The most suspicious tiles in each slide are passed to the model to predict the final slide-level classification as MSI or MSS.

4. Evaluation Metrics

The AUROC will be used as the performance metric to evaluate our binary classification model.

5. Timeline

- Step1 (03/08 - 03/14): Data exploratory analysis
- Step2 (03/14 - 03/28): Train tile-level CNN model to get the probability of MSI for each tile.
- Step3 (03/29 - 04/12): Train RNN model to get the prediction on whole-slide level
- Step4 (04/13 - 05/01): Model improvement and results evaluation.

References

Freddie Bray, Jacques Ferlay, Isabelle Soerjomataram, Rebecca L. Siegel, Lindsey A. Torre, and Ahmedin Jemal. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6):394–424, 2018. ISSN 1542-4863. doi: 10.3322/caac.21492. URL https://onlinelibrary.wiley.com/doi/abs/10.3322/caac.21492. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3322/caac.21492.

Gabriele Campanella, Matthew G. Hanna, Luke Geneslaw, Allen Miraflor, Vitor Werneck Krauss Silva, Klaus J. Busam, Edi Brogi, Victor E. Reuter, David S. Klimstra, and Thomas J. Fuchs. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8):1301–1309, August 2019. ISSN 1078-8956, 1546-170X. doi: 10.1038/s41591-019-0508-1. URL http://www.nature.com/articles/s41591-019-0508-1.

Rui Cao, Fan Yang, Si-Cong Ma, Li Liu, Yu Zhao, Yan Li, De-Hua Wu, Tongxin Wang, Wei-Jia Lu, Wei-Jing Cai, Hong-Bo Zhu, Xue-Jun Guo, Yu-Wen Lu, Jun-Jie Kuang, Wen-Jing Huan, Wei-Min Tang, Kun Huang, Junzhou Huang, Jianhua Yao, and Zhong-Yi Dong. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in Colorectal Cancer. *Theranostics*, 10(24):11080–11091, September 2020. ISSN 1838-7640. doi: 10.7150/thno.49864. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7532670/.

Guia Cerretelli, Ann Ager, Mark J Arends, and Ian M Frayling. Molecular pathology of Lynch syndrome. *The Journal of Pathology*, 250(5):518–531, 2020. ISSN 1096-9896. doi: 10.1002/path.5422. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/path.5422. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/path.5422.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep Residual Learning for Image Recognition. December 2015. URL https://arxiv.org/abs/1512.03385v1.

Jakob Nikolas Kather. Histological images for MSI vs. MSS classification in gastrointestinal cancer, FFPE samples, February 2019. URL https://zenodo.org/record/2530835. Type: dataset.

Jakob Nikolas Kather, Lara R. Heij, Heike I. Grabsch, Chiara Loeffler, Amelie Echle, Hannah Sophie Muti, Jeremias Krause, Jan M. Niehues, Kai A. J. Sommer, Peter Bankhead, Loes F. S. Kooreman, Jefree J. Schulte, Nicole A. Cipriani, Roman D. Buelow, Peter Boor, Nadina Ortiz-Brüchle, Andrew M. Hanby, Valerie Speirs, Sara Kochanny, Akash Patnaik, Andrew Srisuwananukorn, Hermann Brenner, Michael Hoffmeister, Piet A. van den Brandt, Dirk Jäger, Christian Trautwein, Alexander T. Pearson, and Tom Luedde. Pan-cancer image-based detection of clinically actionable genetic alterations. *Nature Cancer*, 1(8):789–799, August 2020. ISSN 2662-1347. doi: 10.1038/s43018-020-0087-6. URL http://www.nature.com/articles/s43018-020-0087-6.

Michael J Overman, Ray McDermott, Joseph L Leach, Sara Lonardi, Heinz-Josef Lenz, Michael A Morse, Jayesh Desai, Andrew Hill, Michael Axelson, Rebecca A Moss, Monica V Goldberg, Z Alexander Cao, Jean-Marie Ledeine, Gregory A Maglinte, Scott Kopetz, and Thierry André. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *The Lancet Oncology*, 18(9):1182–1191, September 2017. ISSN 14702045. doi: 10.1016/S1470-2045(17)30422-9. URL https://linkinghub.elsevier.com/retrieve/pii/S1470204517304229.

Talha Shaikh, Elizabeth A. Handorf, Joshua E. Meyer, Michael J. Hall, and Nestor F. Esnaola. Mismatch Repair Deficiency Testing in Patients With Colorectal Cancer and Nonadherence to Testing Guidelines in Young Adults. *JAMA Oncology*, 4(2):e173580, February 2018. ISSN 2374-2437. doi: 10.1001/jamaoncol.2017.3580. URL https://doi.org/10.1001/jamaoncol.2017.3580.

Eduardo Vilar and Stephen B. Gruber. Microsatellite instability in colorectal cancer—the stable evidence. *Nature Reviews Clinical Oncology*, 7(3):153–162, March 2010. ISSN 1759-4774, 1759-4782. doi: 10.1038/nrclinonc.2009.237. URL http://www.nature.com/articles/nrclinonc.2009.237.