Morphological features computed using deep learning for an annotated digital DLBCL image set

Damir Vrabac[1], Akshay Smit[1], Rebecca Rojansky[2] , Yasodha Natkunam[2] , Ranjana H. Advani[2] , Andrew Y. Ng[1] , Sebastian Fernandez-Pol[2], and Pranav Rajpurkar[1]
[1]Department of Computer Science, Stanford University [2]Department of Pathology, Stanford University

Introduction

- Diffuse Large B-Cell Lymphoma is the most common type of non-Hodgkin lymphoma[1].
- The best clinical prognosticator (IPI score) is insufficient to guide therapeutic decision-making for individual patients[2].
- The Hans algorithm distinguishes the more favorable GCB subtype from the non-GCB subtype [3].
- Double-hit and double-expressor lymphomas with concurrent aberrations in MYC and BCL2 or BCL6 correlate with an aggressive clinical course [4].
- Classification based on morphology has been challenging due to histomorphologic heterogeneity.

Objectives

1. Identify prognostically relevant features on H&E sections.
2. Provide a cost-effective alternative to current classification methods.
3. Provide a dataset of annotated slides with immunohistochemical and outcome data.

Methods

- 170 de novo, CD20+ DLBCL patients treated with with R-CHOP with clinical data from Stanford Cancer Institute[5].
- 7 tissue microarrays composed of duplicate 0.6-mm diameter cores, 0.4 um formalin-fixed, paraffin-embedded (FFPE) sections stained with H&E, CD10, BCL6, MUM1, BCL2, and MYC.
- Scanned at 40x magnification (0.25 μm per pixel), Aperio AT2 scanner (Leica Biosystems, Nussloch, Germany).
- Pathologists annotate representative area
- Non-overlapping 224x224 pixel patches extracted within ROIs.
- HoVer-Net[6] deep learning model to segment tumor cell nuclei and compute geometric descriptors.
- Statistical Analysis: Cox Proportional Hazards model with Follow-up Status as indicator of censoring, and overall survival as time to event or censoring.

Results

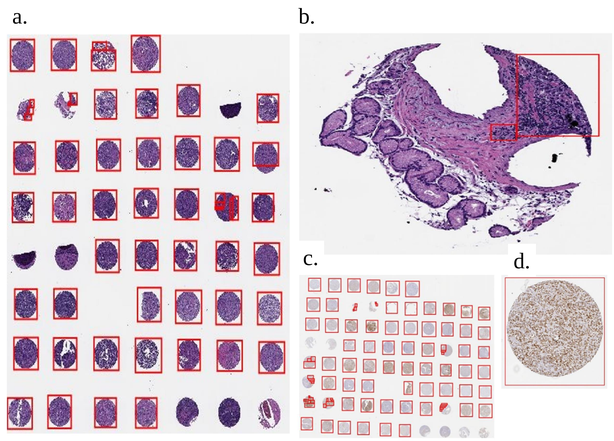

Figure 1. **Tissue microarrays (TMAs) with region-of-interest (ROI) annotations**. A) H&E stained TMA. The red rectangles denote ROIs annotated by a pathologist. B) A single core from the TMA in a) showing ROIs. C) BCL6 stained TMA, containing cores from the same patients as a). D) A single annotated core from the TMA in c).

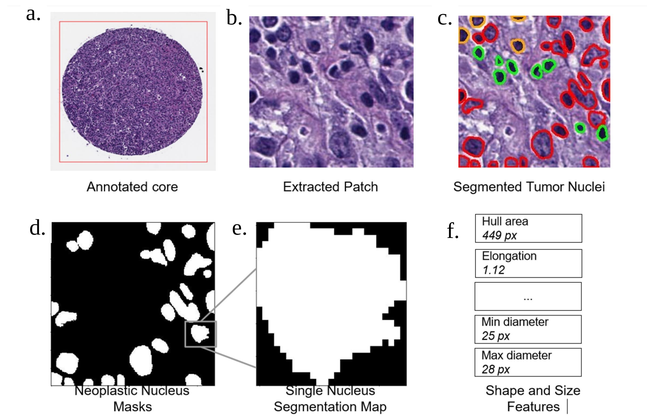

Figure 2. **Data pipeline for a single core from an H&E stained tissue microarray (TMA)**. In a) the red rectangle is the pathologist-annotated ROI. In c) red corresponds to cell nuclei classified as “neoplastic” by HoVer-Net. Green corresponds to “inflammatory” and orange corresponds to “non-neoplastic epithelial”.

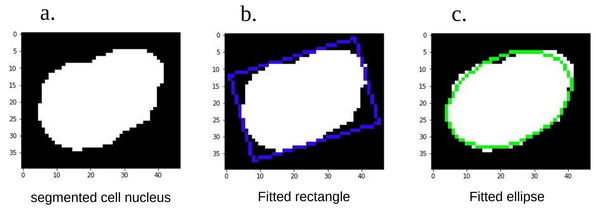

Figure 3. **Rectangle and ellipse fitted to a single segmented tumor nucleus**. a) a binary segmentation image for a tumor cell nucleus. B) rotated rectangle fit to the nucleus. c) rotated ellipse fit to the nucleus.

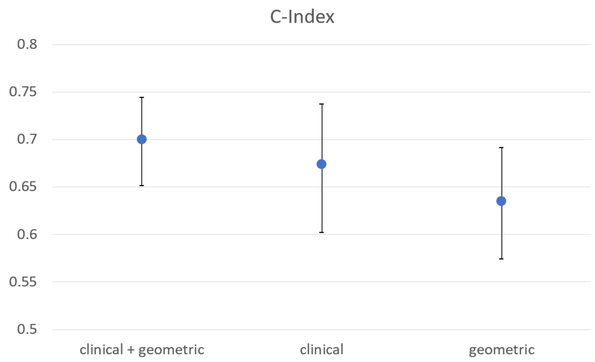

Figure 4. 95% two-sided confidence interval (CI) for the optimism-corrected C-index using the non-parametric percentile bootstrap method[30] with 1000 bootstrap replicates.

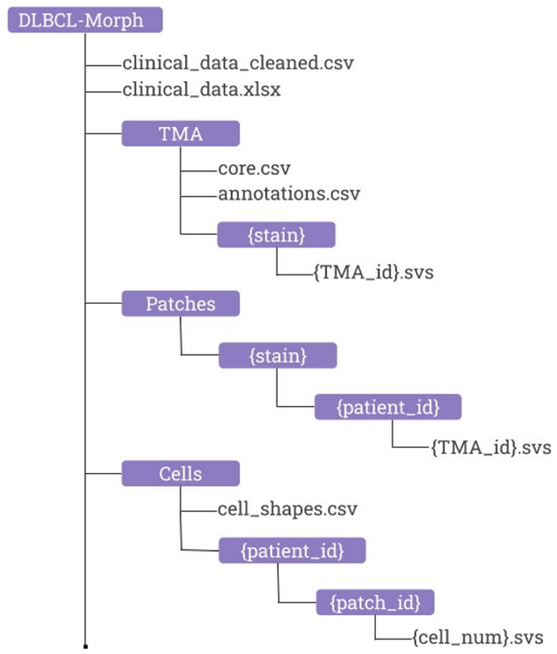

Figure 5. The directory structure of DLBCL Morph

Discussion

- Geometric features alone allow significantly better than random prognostic outcome prediction that is as good as clinical features.
- Trend that clinical and geometric features combined achieved higher performance than the clinical features alone.

Conclusions

- Geometric features computed from H&E-stained sections can provide a significant signal to determine prognostic outcome independent of clinical features.

Future Directions

- Further evaluation on external datasets and prospectively in future studies
- This dataset including annotations is available here: https://arxiv.org/abs/2009.08123

References

1. Horvat, M. *et al.* Diffuse large b-cell lymphoma: 10 years’ real-world clinical experience with rituximab plus cyclophos- phamide, doxorubicin, vincristine and prednisolone. *Oncol. Lett.* 10.3892/ol.2018.7774 (2018).
2. Zhou, Z. *et al.* An enhanced international prognostic index (NCCN-IPI) for patients with diffuse large b-cell lymphoma treated in the rituximab era. *Blood* **123**, 837–842, 10.1182/blood-2013-09-524108 (2014).
3. Alizadeh, A. A. *et al.* Distinct types of diffuse large b-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511, 10.1038/35000501 (2000).
4. Riedell, P. A. & Smith, S. M. Double hit and double expressors in lymphoma: Definition and treatment. *Cancer* **124**, 4622–4632, 10.1002/cncr.31646 (2018).
5. Rosenwald, A. *et al.* Prognostic significance of myc rearrangement and translocation partner in diffuse large b-cell lymphoma: A study by the lunenburg lymphoma biomarker consortium. *J. Clin. Oncol.* **37**, 3359–3368, 10.1200/JCO.19. 00743 (2019). PMID: 31498031, https://doi.org/10.1200/JCO.19.00743.
6. Graham, S. *et al.* Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images (2018). 1812.06499.

Contact Information: Rebecca Rojansky, MD, PhD | rebecca.rojansky@stanford.edu