Influence Of Multiple Instance Learning On The Generation Of Cost-effective Computational Pathology Algorithms

Keywords: Computational Pathology, Cost-effectiveness, Healthcare AI

UNIVERSITAT DE BARCELONA

AUTHORS

Manuel Cossio1,2 & Ramiro Gilardino1

HE-XPERTS CONSULTING

AFFILIATIONS

1. *HE-Xperts Consulting LLC, Miami, FL, USA.*
2. *Universitat de Barcelona, Barcelona, Spain*

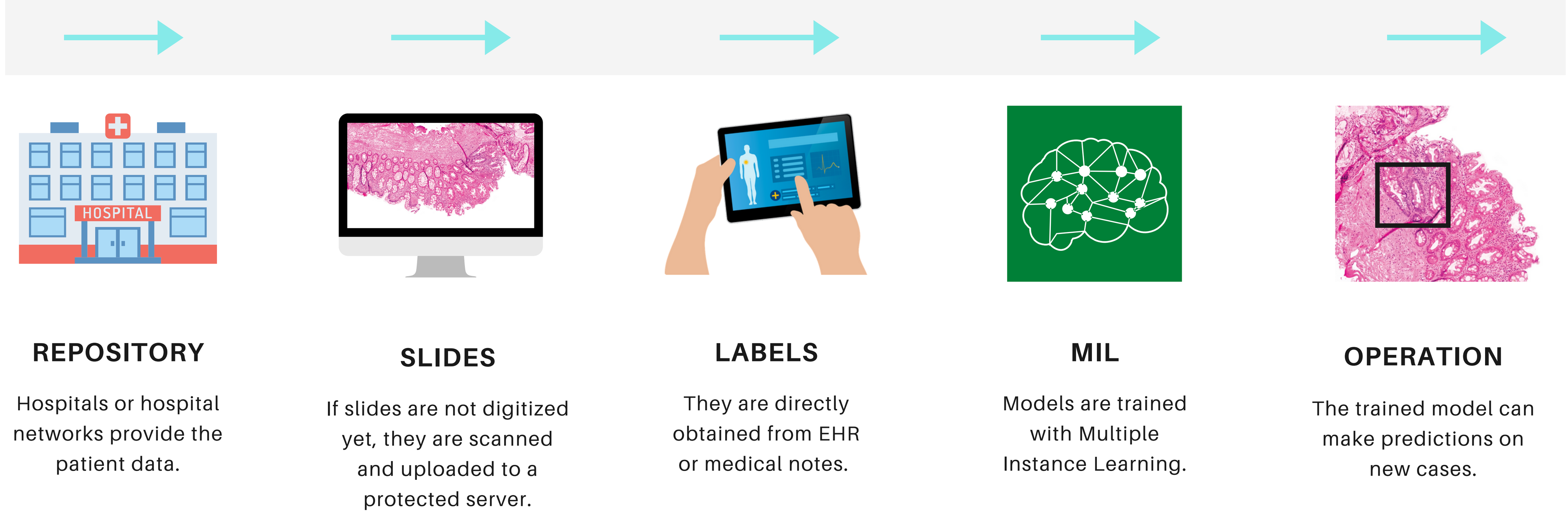

Figure 1. Global process of training and testing vision models with Multiple Instance Learning.

01. Introduction

In the development of computational pathology models, a significant challenge is the availability, cost, and time required for skilled labor, such as trained physicians, to generate pathologic tissue masks for digital slides, also known as Whole Slide Images (WSI). The multiple instance learning (MIL) technique provides a potential solution to this challenge, as it enables training models using only the patient's diagnosis (pathological or non-pathological) associated with the slide, without the need for tissue delimitation via a mask (1,2). This allows for the direct use of electronic health record files as input for training, thereby reducing the need for manual annotation by a pathologist (Figure 1).

02. Objective

The objective of this work is to evaluate the potential of the multiple instance learning (MIL) technique in reducing the cost of building computer vision algorithms (reduce use of healthcare resources HRU) applied to pathology images.

References

1. Sudharshan, P. J., Petitjean, C., Spanhol, F., Oliveira, L. E., Heutte, L., & Honeine, P. (2019). Multiple instance learning for histopathological breast cancer image classification. Expert Systems with Applications, 117, 103-111.
2. Ilse, M., Tomczak, J., & Welling, M. (2018, July). Attention-based deep multiple instance learning. In International conference on machine learning (pp. 2127-2136). PMLR.

03. Methods

We conducted a targeted literature review until January 2023, searching for: "multiple instance learning" AND "attention" AND "whole slide imaging pathology" OR "WSI pathology". A data extraction grid was created to analyze the following variables: proportion of applications per disease or medical specialty, number of WSI with global label, type of dataset, HRU component addressed, and registered metrics for performance. Categorical data is presented as percentage and continuous data as means.

04. Results

62 articles underwent full text screening and data extraction. Therapeutic area/ medical specialties included: Oncology: 57 (91%), Gastroenterology: 3 (5%), Hematology: 1(2%), Infectious Diseases: 1(2%). Subsets for oncology included: Breast (28), Lung (14) and Gastrointestinal (12) cancer (Figure 3). The WSI samples employed ranged from 24 to 20,229 (mean: 226). 57 (91%) articles applied labels extracted from diagnostic registries, 27 (43%) articles used data from the Cancer Genome Atlas (TCGA, Figure 2), and 39 (58%) mentioned any element involving HRU optimization (reduction in number of trained physicians: 26, reduction in physician hours to diagnosis: 8; both: 5; Figure 4). Very few articles (6) mentioned the type and any estimated cost of image storage for training and testing. The most commonly used metric was accuracy (37, 60%), with a maximum (0.99), minimum(0.68), and average (0.89).

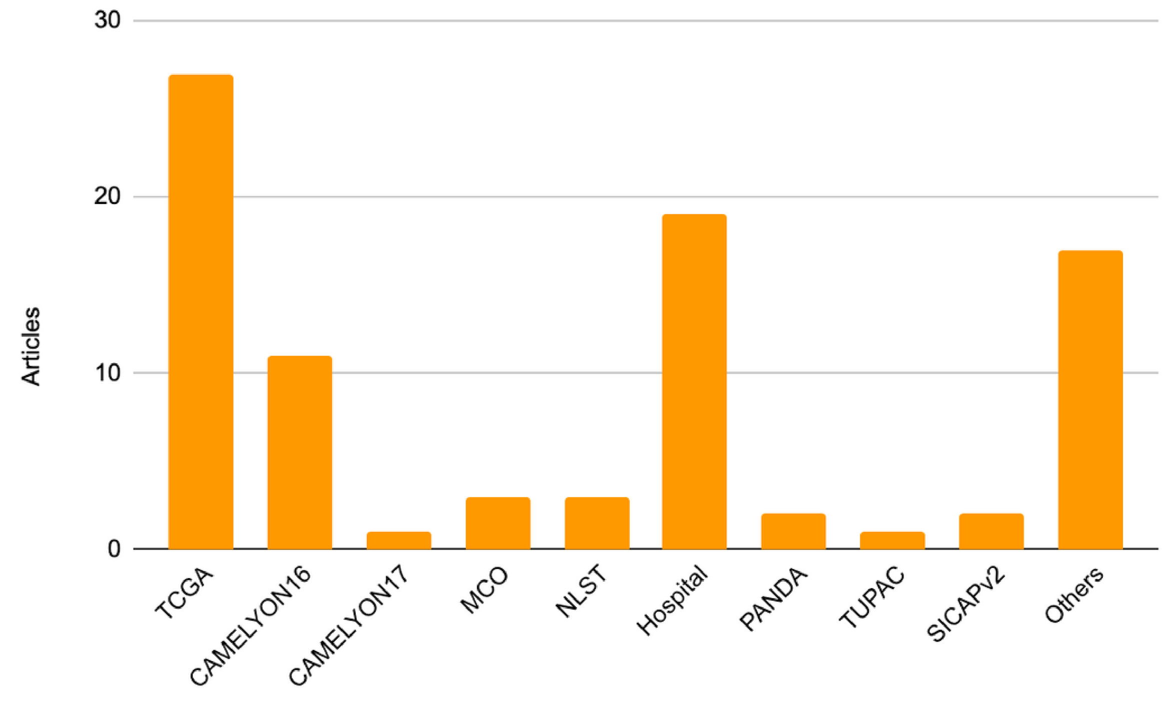

Figure 2. Use of different medical imaging datasets per article . TCGA, The Cancer Genome Atlas, MCO, Molecular and Cellular Study, NLST, The National Lung Screening Trial, PANDA, Prostate Cancer Grade Assessment, TUPAC, Tumor Proliferation Assessment Challenge, SICAPv2, Prostate Cancer with Gleason Score.

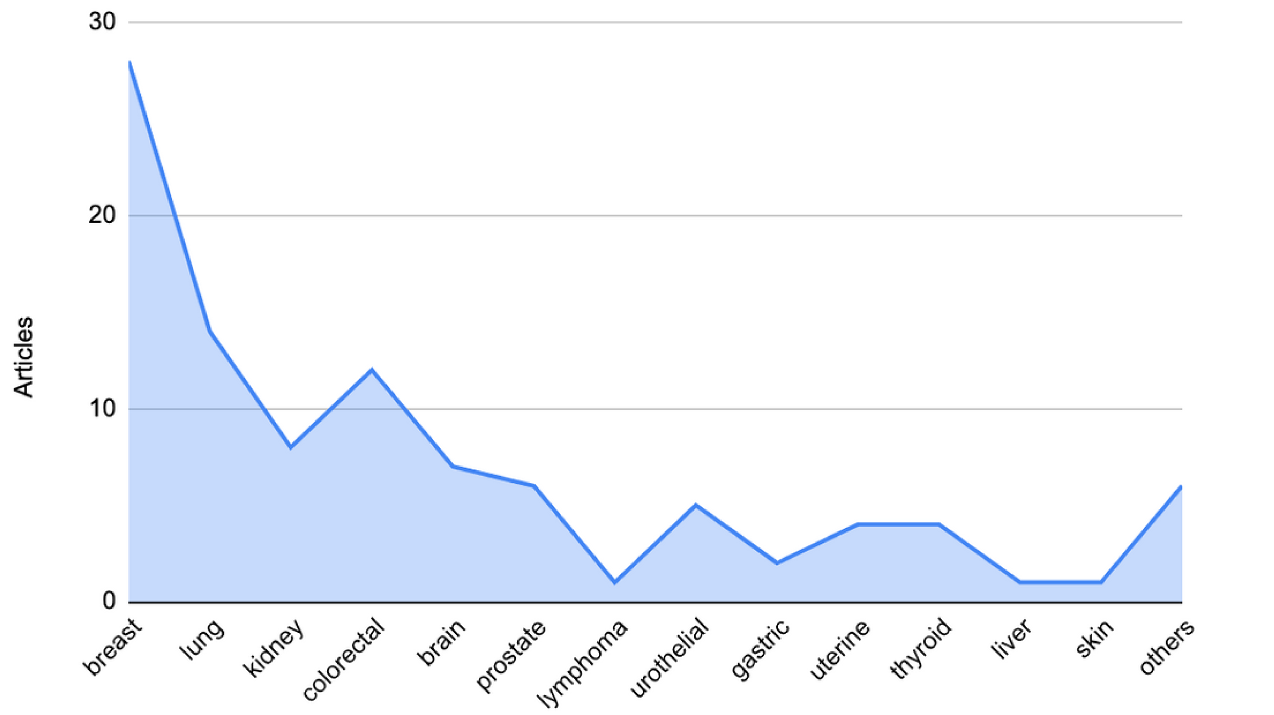

Figure 3. Source of digital slides for oncology studies. The horizontal axis of the figure represents the type of cancer from which the images were obtained to construct the computational models.

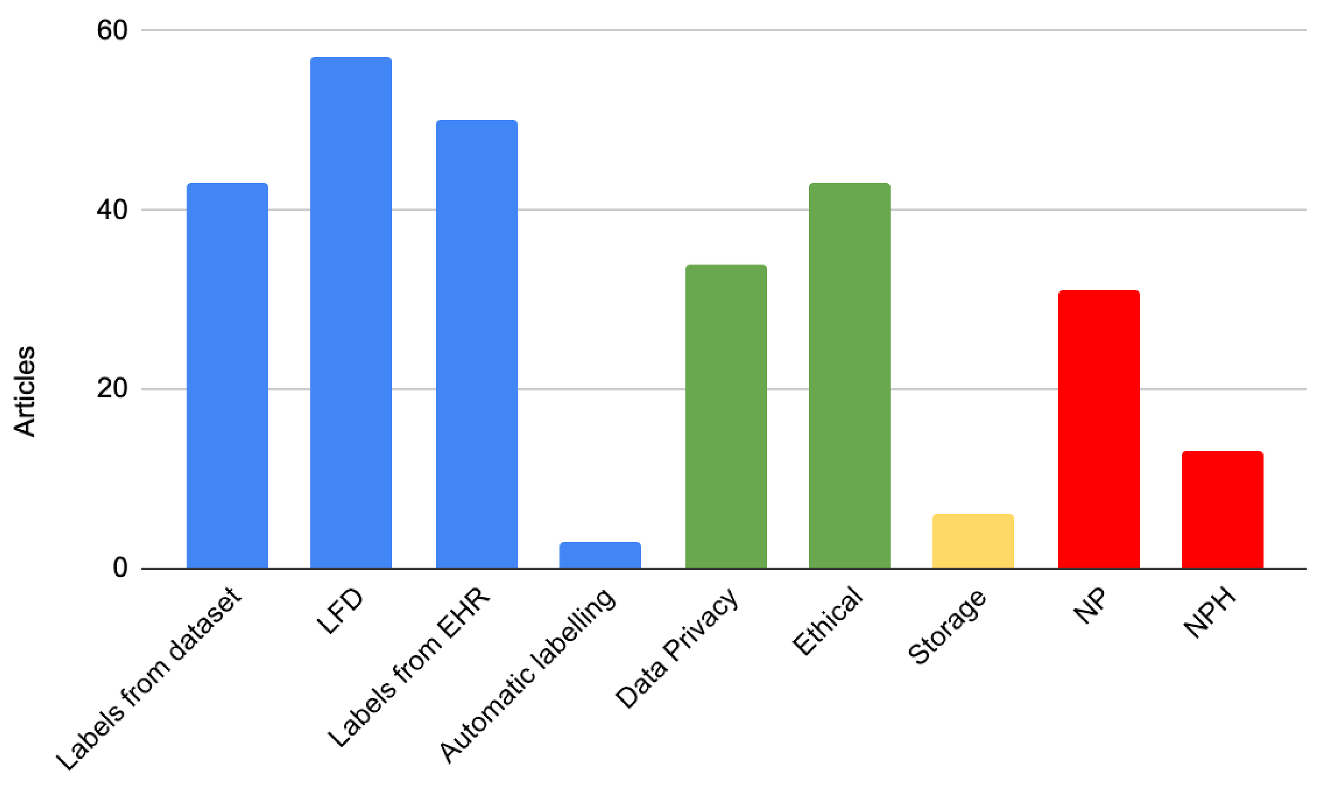

Figure 4. Different parameters considered and analyzed in the articles. The blue columns represent the different label sources (LFD, labels from diagnostic). The green columns are related to data protection. The yellow column represents the cost of storage. The red columns represent the number of trained physicians (NP) and the number of trained physician hours (NPH) used (healthcare resources, HRU).

05. Conclusion

The results of our data analysis indicate that the field of oncology is the primary focus of MIL research, and that the most commonly used database is the TCGA. However, despite the widespread use of MIL, only a few of the reviewed articles included data on parameters related to the use of trained human resources or the cost of image storage. Thus, further studies are required to accurately evaluate the cost-effectiveness of this technique and identify relevant variables that reflect its true economic impact.