Frugal Deep Learning for Multi-Resolution Image Analysis

Marie Reinbigler *

SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris,
19 place Marguerite Perey
91120 Palaiseau - France
marie.reinbigler@telecom-sudparis.eu

Résumé

In the biological field, analysis of histological images, *i.e.* slices of organs observed by microscopy, is key to improving our understanding of diseases. Such multi-resolution images feature very high resolutions and can reach sizes of 100.000x100.000 pixels [12], *i.e.* up to 40GB of uncompressed data. Consequently, manual analysis by biologists is time-consuming tedious work.

In this context, our aim is to take advantage of their pyramidal multi-resolution structure (Figure 1) to automatically perform analysis using frugal local equipment, offering an efficient, affordable, and privacy-preserving solution. The core of our approach is to build a scaling processing architecture composed of a pair of differentiable blocks.

In a first step, the entire histological image is divided into tiles that are analyzed separately, as usually observed in works like in [5]. Instead of extracting one set of features at the highest resolution level, each tile is converted into a hierarchical tree storing extracted features according to their resolution level (see Figure 1). The set of extracted trees composes a representative graph where each locally extracted feature is placed in the global context of the histological image, similarly to works [2, 1, 8, 6]. Our approach distinguishes itself from previous work by exploiting the pyramidal structure of the histological image to retrieve information at the lowest possible resolution. Figure 2 illustrates our multi-resolution analysis process.

In a second step, the information extracted representing the histological image will be interpreted as a whole. Attempts like [4] use graph convolutional neural networks.

We strive to maximize limited resources utilization by developing a dynamic scheduling strategy constrained by the execution context (*e.g.* a hospital or laboratory) and an unpredictable workload induced by the analysis of one tile, whose lowest resolution enabling information retrieval is not known in advance. Consequently, we plan to develop a dynamic scheduling strategy based on gossip-protocol [9, 11] and stealing strategies [3, 7, 10] inspired by the distributed systems field, leveraging application level knowledge to better anticipate the workload of each tile analysis.

The approach, studied for a specific biological case, can be generalized to any multi-resolution image, e.g satellite or geographical images.

Mots-clés : Analyse de données Haute Performance, Calcul haute performance, Apprentissage profond, Numérique frugal, Système décentralisé

---

*. Le texte a été relu par Elisabeth Brunet.

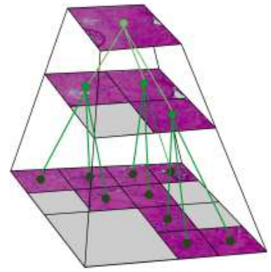

FIGURE 1 – Pyramidal image structure and extracted tree.

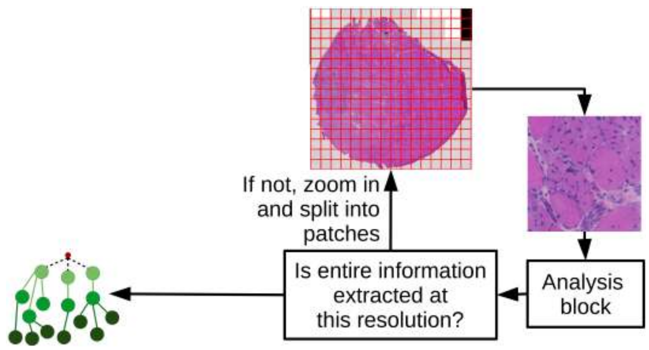

FIGURE 2 – Dynamic multiresolution representative graph building pipeline

Bibliographie

1. Adnan (M.), Kalra (S.) et Tizhoosh (H. R.). – Representation Learning of Histopathology Images using Graph Neural Networks. – In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 4254–4261, Seattle, WA, USA, juin 2020. IEEE.
2. Ahmedt-Aristizabal (D.), Armin (M. A.), Denman (S.), Fookes (C.) et Petersson (L.). – A survey on graph-based deep learning for computational histopathology. *Computerized Medical Imaging and Graphics*, vol. 95, 2022, p. 102027.
3. Delgado (P.), Dinu (F.), Kermarrec (A.-M.) et Zwaenepoel (W.). – Hawk : Hybrid datacenter scheduling. – In *Proceedings of the 2015 USENIX Conference on Usenix Annual Technical Conference, USENIX ATC '15*, USENIX ATC '15, p. 499–510, USA, 2015. USENIX Association.
4. Hou (W.), Yu (L.), Lin (C.), Huang (H.), Yu (R.), Qin (J.) et Wang (L.). – $H^2 - MIL$: Exploring Hierarchical Representation with Heterogeneous Multiple Instance Learning for Whole Slide Image Analysis. *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, n 1, Jun. 2022, pp. 933–941.
5. Kassani (S. H.), Kassani (P. H.), Wesolowski (M. J.), Schneider (K. A.) et Deters (R.). – Classification of histopathological biopsy images using ensemble of deep learning networks. – In *Proceedings of the 29th Annual International Conference on Computer Science and Software Engineering, CASCON '19*, CASCON '19, p. 92–99, USA, 2019. IBM Corp.
6. Konda (R.), Wu (H.) et Wang (M. D.). – Graph Convolutional Neural Networks to Classify Whole Slide Images. – In *2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC)*, pp. 754–758, Madrid, Spain, juillet 2020. IEEE.

7. Kumar (V.), Murthy (K.), Sarkar (V.) et Zheng (Y.). – Optimized distributed work-stealing. – In *2016 6th Workshop on Irregular Applications : Architecture and Algorithms (IA3)*, pp. 74–77, 2016.
8. Lu (W.), Graham (S.), Bilal (M.), Rajpoot (N.) et Minhas (F.). – Capturing Cellular Topology in Multi-Gigapixel Pathology Images. – In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1049–1058, juin 2020. – ISSN : 2160-7516.
9. Menon (H.) et Kalé (L.). – A distributed dynamic load balancer for iterative applications. – In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, *SC '13*, SC '13, New York, NY, USA, 2013. Association for Computing Machinery.
10. Paudel (J.), Tardieu (O.) et Amaral (J.). – On the merits of distributed work-stealing on selective locality-aware tasks. – pp. 100–109, 10 2013.
11. Wuhib (F.), Stadler (R.) et Spreitzer (M.). – A gossip protocol for dynamic resource management in large cloud environments. *IEEE Transactions on Network and Service Management*, vol. 9, n2, 2012, pp. 213–225.
12. Xu (Y.), Jia (Z.), Wang (L.-B.), Ai (Y.), Zhang (F.), Lai (M.) et Chang (E. I.-C.). – Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*, vol. 18, n1, 5 2017, pp. 1–17.