CNN-based Approach for Cervical Cancer Classification in Whole-Slide Histopathology Images

Ferdaous Idlahcen [1], M. Majid Himmi [1], Abdelhak Mahmoudi [1,2]

[1] Computer Sciences, Applied Mathematics,
AI & Pattern Recognition Lab (LIMIARF)
Faculty of Sciences
Mohammed V University, Rabat, Morocco
[2] École Normale Supérieure (ENS)
Mohammed V University, Rabat, Morocco

Background

1. According to IARC/WHO, Cervical Cancer (CC) will cause 460 000 deaths per year by 2040 in the absence of decisive actions (∼50% over 2018 levels with ∼90% of deaths in LMICs, particularly sub-saharan african women).

Figure 1: The Uterine Cervix

2. Histopathological assessment is the cornerstone for non-invasive/invasive cancers diagnosis/treatments and still crucial for cervical cancer screening programs evaluation. (Teixeira and Vasconcelos, 2019; Bulten et al., 2011).

Background

3. "Numerous publications show the clinical and economic benefits of obtaining a second opinion for Pathology specimens ... In terms of cancer, changes can be from cancer to benign (or vice versa) or from one type of cancer to another, which could have a significant impact on treatment and prognosis ..."

— Jonathan Epstein, M.D.

Director of Surgical Pathology

Johns Hopkins Hospital

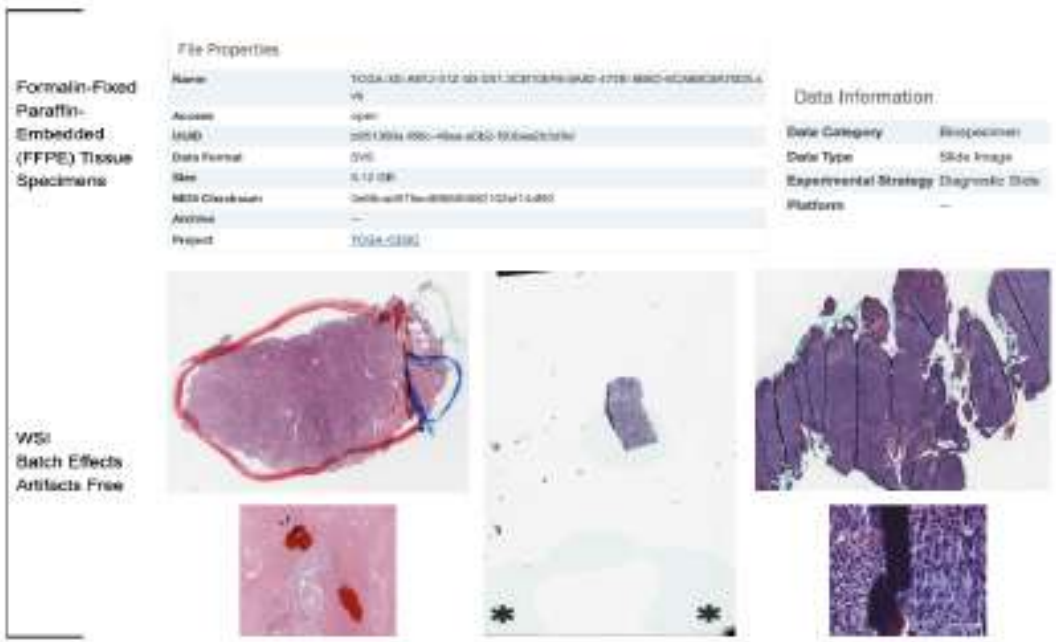
Background

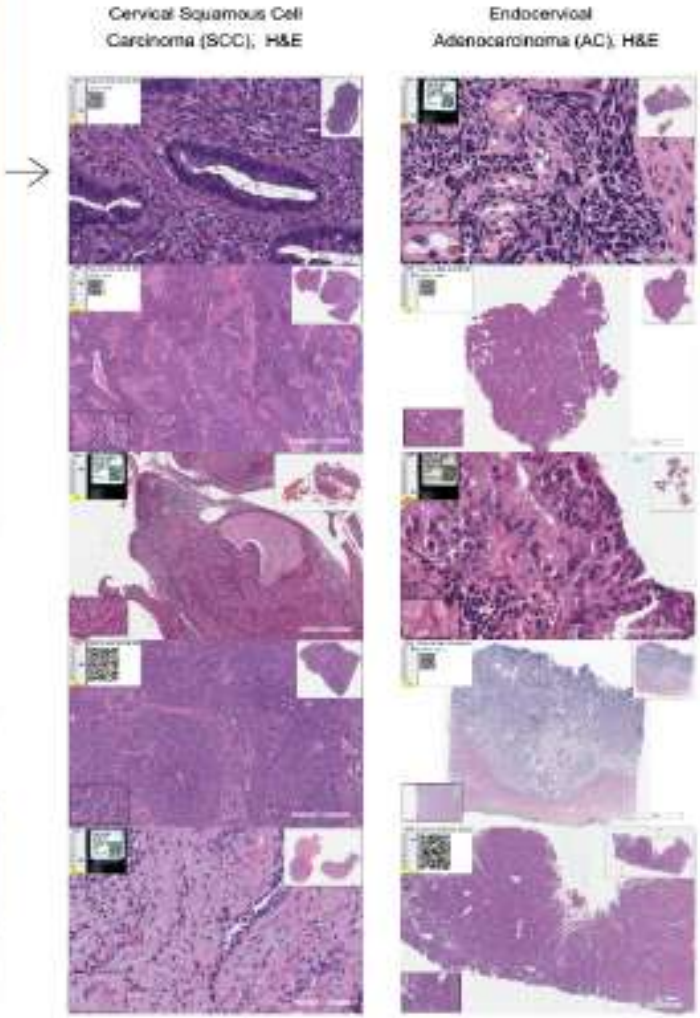
4. The emergence of Whole-Slide Imaging (WSI), i.e. virtual microscopy, has led to an unprecedented synergy between AI and digital pathology to address several issues like misdiagnoses.

Figure 2: Digitization of Pathology Specimens

5. WSIs come with several obstacles that limit AI potentials exploit in computational pathology, e.g. single WSI dimension, clinical features variability, lack of annotated data, etc. (Tizhoosh and Pantanowitz, 2018).

Dataset Acquisition

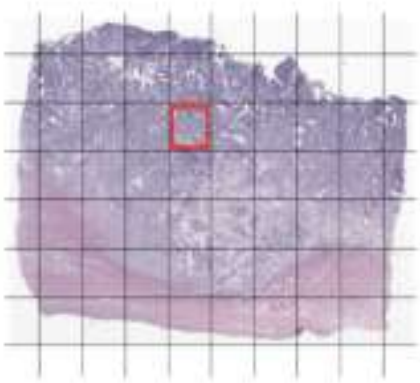


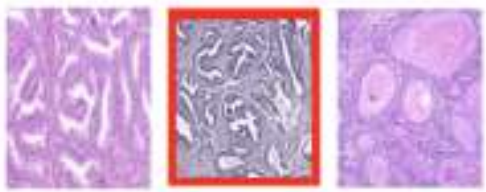


Workflow

1. Import Data

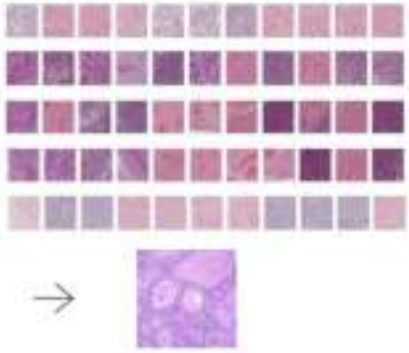
2. WSI Tile - Based Approach





> 12 412 ROIs, 600 selected
1024 x 1024 ROIs, 20x / 224 x 224 ROIs, 20x
> 90% of Tissue, one histologic criterias at least

3. Stain Normalization & Data Split



$$\exp\left(\ln RGB_{x,y}\begin{bmatrix}0.655 & 0.715 & 0.290\\ 0.083 & 0.985 & 0.115\end{bmatrix}^{T}\right) = HE_{x,y}$$

Table 1: Statistics of data split

Train	Validation	Test
216 SCC	24 SCC	60 SCC
216 AC	24 AC	60 AC

4. Augmentation

5. Training

Table 2: VGG-16 network

Layer	Size
conv x2	224 x 224 x 64
pool	112 x 112 x 64
conv x2	112 x 112 x 128
pool	56 x 56 x 128
conv x3	56 x 56 x 256
pool	28 x 28 x 256
conv x3	28 x 28 x 512
pool	28 x 28 x 512
conv x3	14 x 14 x 512
pool	7 x 7 x 512
[Fully-connected]	

BNF
Acc
97 %

Classifier

Table 3: Performance (%) on the test set

Accuracy	Precision	Recall	F1
98.26	96.8	99	97.9

Contribution

The major contribution focuses on the potential of a pre-trained CNN approach for an Hematoxylin & Eosin histopathology image analysis of cervical cancer to acquire efficient classification despite the data-poorness and unlabelled problems. Hence, the challenge relied on the non trivial task of WSIs processing and the VGG16-CNN pre-trained model fine tuning from a totally different data on which it was trained on.

Future Work : Integration of proteogenomics, clinical data and WSI to detect cancer patterns.

Ferdaous Idlahcen
idlahcenferdaous.s@gmail.com

Bibliography

Bulten, J., Horvat, R., Jordan, J., Herbert, A., Wiener, H., and Arbyn, M. (2011). European guidelines for quality assurance in cervical histopathology. *Acta Oncologica*, 50(5):611–620.

Teixeira, A. K. S. and Vasconcelos, J. L. A. (2019). Histopathological profile of patients diagnosed with malignant tumors assisted in a hospital of reference of agreste pernambucano. *Jornal Brasileiro de Patologia e Medicina Laboratorial*, 55(1).

Tizhoosh, H. and Pantanowitz, L. (2018). Artificial intelligence and digital pathology: Challenges and opportunities. *Journal of Pathology Informatics*, 9(1):38.