Project Report (May 15, 2021)

Ankita Ghosh[1] **and Sahil Khose**[2]

[1]Research Assistant, ghoshankita0907@gmail.com , CSE, MIT Manipal
[2]Research Assistant, sahilkhose18@gmail.com, ICT, MIT Manipal

ABSTRACT

In the last report, we shed a light on the present state of Deep Learning in histopathology, we discussed Histopathology, Digital Histopathology, the possibilities of Machine Learning in the area, the various applications, followed by a detailed discussion of the challenges involved in working with Digital Microscopic Slide Images and in the application of Deep Learning Algorithms to them.
In this report, we shall be discussing in greater detail the applicability of Deep Learning to Histopathology from a methodological perspective along with the tasks it helps accomplish using relevant work for illustration.

Introduction

The applicability of deep learning can be studied in terms of the tasks it performs or in terms of the learning paradigm, which is the classification we shall be using in this writeup. The different learning algorithms, viz a viz Deep Learning for histopathology, along with the tasks are visualized in the following overview.

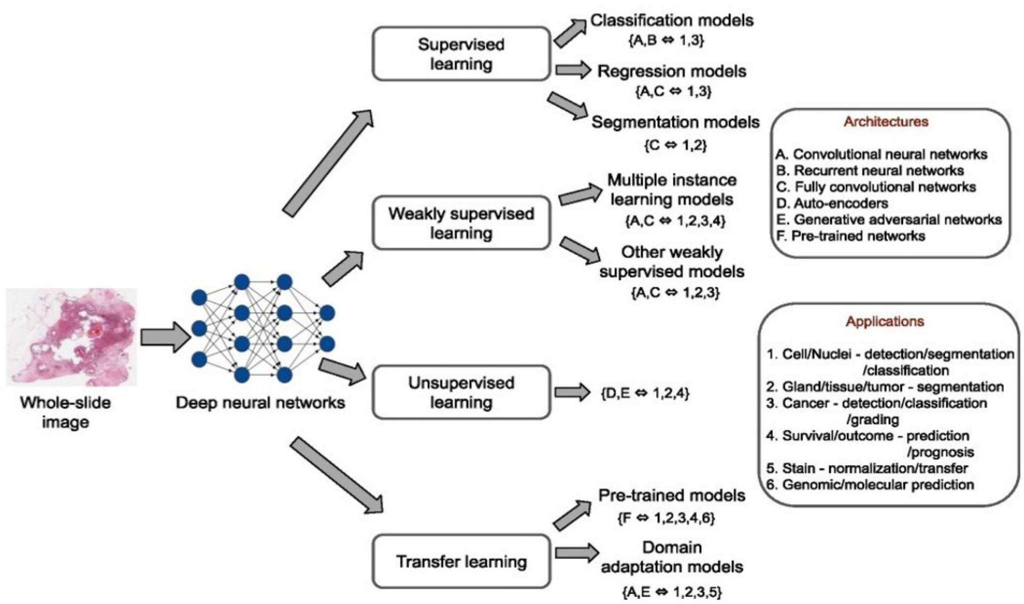

Figure 1. Source: Srinidhi et al(2019)

Based on these, a number of DL models have been proposed in the literature that are traditionally based on convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GANs), auto-encoders(AEs) and other variants.

Supervised Learning

Among the supervised learning techniques, we identify three major canonical deep learning models based on the nature of tasks that are solved in digital histopathology: Classification, Regression, and Segmentation.

Supervised Classification

- It can be further subdivided into local and global level classification. Local level classification entails identifying cells or nuclei in patches of the whole slide image. Deep Learning has proven extensively successful in pixel-wise prediction by sliding window approach over image patches that are annotated by pathologists as regions containing objects of interest(cells/nuclei) or background.

- Qaiser et al in their paper used Persistent Homology Profiles as distinguishing features in order to segment colon tumor regions by classifying patches as tumor regions or normal ones. Persistent Homology profiles are compact mathematical feature representations of a region that are distinctive as well as robust to scale, perturbations in input data, dimension, and coordinates.

- They used PHP of training dataset in combination with features extracted using CNN and then employed Random Forest regressions on them separately followed by a multi-stage ensemble strategy for the final classification. This hybrid approach proved to be both accurate and highly efficient wrt inference speed.

- In global level classification, most of the published work focusses on a patch-based classification approach for whole-slide level disease prediction task. It can involve both patch level localization as well as whole slide level classification or grading of disease. The main disadvantage of these methods is the relatively long computational time required to carry out a dense patch-wise prediction over an entire WSI. Different works have approached this problem in different ways, some using heuristic sampling strategies to more recent ones using task-driven visual attention based coarse processing.

- Xu et al in their work adaptively select a sequence of coarse regions from the raw image by a hard visual attention algorithm, and then for each such region, it is able to investigate the abnormal parts based on a soft-attention mechanism. A recurrent network is then built on top to classify the image region and also to predict the location of the image region to be investigated at the next time step. This way, only a fraction of pixels need to be investigated for the classification

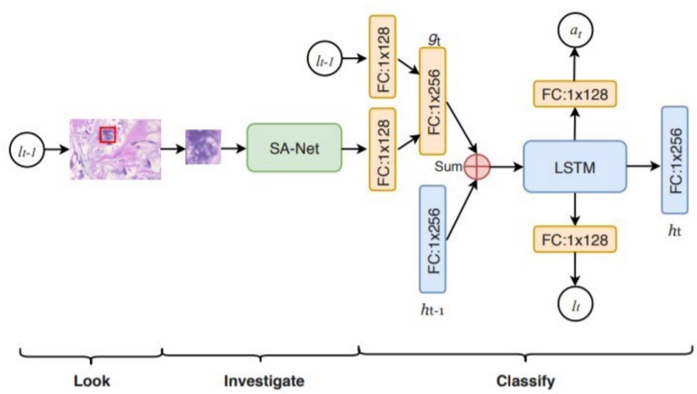

Figure 2. Source: Xu et al

- Advantages of using visual attention-based models for Whole Slide Image global classification task are:

  - The model tries to learn only the most relevant diagnostically useful areas for disease prediction as it enforces a region selection mechanism.

  - The model complexity is independent of the size of WSI.

- Another recent work for global classification by Halicek, Martin, et al.[3] perform patch-based localization and whole slide classification for Squamous Cell Carcinoma(SCC) and Thyroid Cell Carcinoma using CNN using an entirely different approach.

- A ground-truth binary mask of the cancer area was produced from each outlined histology slide. The WSIs and corresponding ground-truths were down-sampled by a factor of four using nearest-neighbor interpolation. The downsampled slides were then broken into patches of 101 x 101 size. To ensure generalization the number of image patches was augmented by 8x by applying 90-degree rotations and reflections to develop a more robust diagnostic method. Additionally, to establish a level of color-feature invariance and tolerance to differences in HE staining between slides, the hue, saturation, brightness, and contrast of each patch were randomly manipulated to make a more rigorous training paradigm before being fed to the Inception-v4 model for detecting head and neck cancer.

Supervised Regression

- In this method, we focus on directly regressing the likelihood of pixel being the center of an object for detection or localization of objects. Regression, unlike classification, gives us a continuous value, usually probability score instead of simply a class label as output. Regression helps in better detection by enforcing topological constraints such as assigning higher probability values to pixels near the object center.

- Regression also helps with challenges faced in cell/nuclei detection arising due to highly irregular appearance and them occurring as overlapping clumps resulting in problems separating them. Deep regression models proposed in the literature are mainly based on either CNN or Fully Convolutional Network(FCN) architectures.

- The paper by Graham et al on HoVer-Net is one of the most seminal works in the entire area of research. It proposes a unified FCN model for simultaneous nuclear instance segmentation and classification. It leverages the instance-rich information encoded within the vertical and horizontal distances of nuclear pixels to their centers of mass. These distances are then utilized to separate clustered nuclei, resulting in an accurate segmentation, particularly in areas with overlapping instances.

- Then, for each segmented instance, the network predicts the type of nucleus via a devoted up-sampling branch. The network is composed of three parallel branches that are used for three different tasks. We have corresponding ground truth annotations of the data for each of the three branches.

  - Then, for each segmented instance, the network predicts the type of nucleus via a devoted up-sampling branch. The network is composed of three parallel branches that are used for three different tasks. We have corresponding ground truth annotations of the data for each of the three branches.

  - whereas the Horizontal-Vertical(HoVer) branch predicts the horizontal and vertical distances of nuclear pixels to their centers of mass. Here the colors represent the gradation of the distance of each nuclear pixel from the center of mass.

  - Blue represents positive distance up to +1 and means the pixel lies on the left side of the COM in case of the horizontal map and above the COM when it comes to the vertical mapping. Similarly, red represents negative distance up to -1 and means the pixel lies on the Right/bottom side of COM accordingly.

  - Then, the Nuclear Classification(NC) branch(optional) predicts the type of nucleus for each pixel

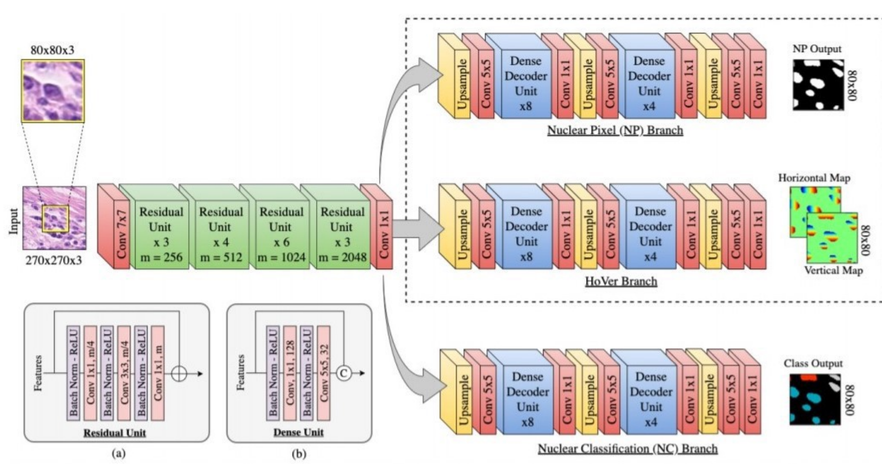

Figure 3. Source: Graham et al

- In particular, the NP and HoVer branches jointly achieve nuclear instance segmentation by first separating nuclear pixels from the background (NP branch) and then separating touching nuclei (HoVer branch). This is the same model that was used for the localization and clustering of tissues step for modeling Whole Slide Images as graphs for subsequent learning using Graph Neural Networks, as discussed in the previous report.

Discussion

In this report, we discussed in greater detail the applicability of Deep Learning to Histopathology from a methodological perspective along with the tasks it helps accomplish. Now we have a broader grasp of different approaches for DL in Histopathology which will help us proceed with our research.