Deep Learning and Computational Histopathology

LITERATURE REVIEW

Introduction

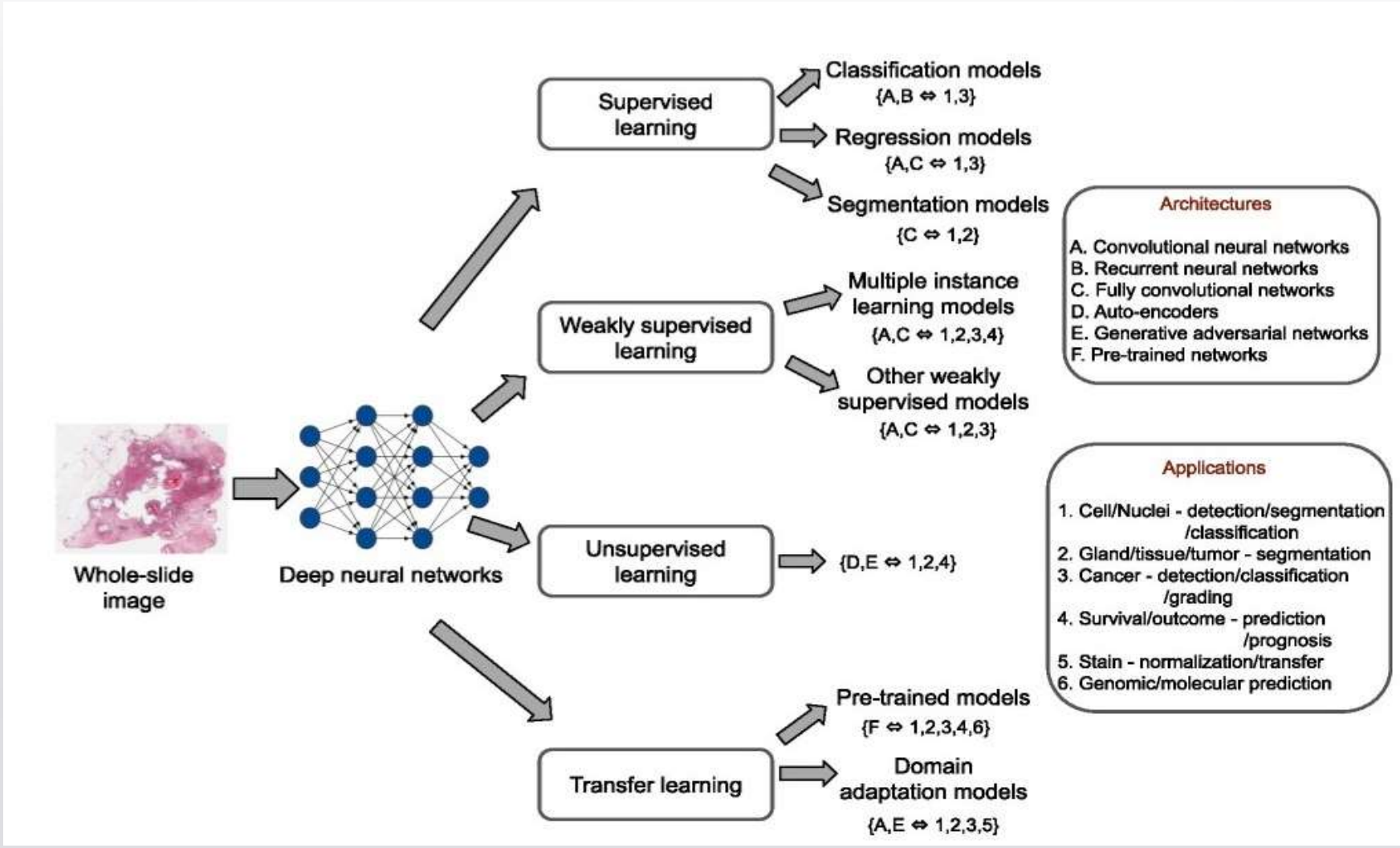
- Through this review we shall discuss the various deep learning methodologies that have been applied to Histopathological Images, along with the tasks they solve
- This is essentially a survey of several relevant papers to gauge the progress made in the field
- The review will also help us understand the various challenges we face when it comes to dealing with Whole slide Images and how various researchers deal with them for various tasks.

Introduction

The various Learning Schemes that are used in context of DL applied to computational pathology which we shall discuss are

Supervised Learning	Weakly supervised Learning	Unsupervised Learning	Transfer Learning

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


Whole-slide image
Deep neural networks
Supervised learning
Classification models
{A,B ⇔ 1,3}
Regression models
{A,C ⇔ 1,3}
Segmentation models
{C ⇔ 1,2}
Weakly supervised learning
Multiple instance learning models
{A,C ⇔ 1,2,3,4}
Other weakly supervised models
{A,C ⇔ 1,2,3}
Unsupervised learning
{D,E ⇔ 1,2,4}
Transfer learning
Pre-trained models
{F ⇔ 1,2,3,4,6}
Domain adaptation models
{A,E ⇔ 1,2,3,5}
Architectures
A. Convolutional neural networks
B. Recurrent neural networks
C. Fully convolutional networks
D. Auto-encoders
E. Generative adversarial networks
F. Pre-trained networks
Applications
1. Cell/Nuclei - detection/segmentation /classification
2. Gland/tissue/tumor - segmentation
3. Cancer - detection/classification /grading
4. Survival/outcome - prediction /prognosis
5. Stain - normalization/transfer
6. Genomic/molecular prediction

Challenges

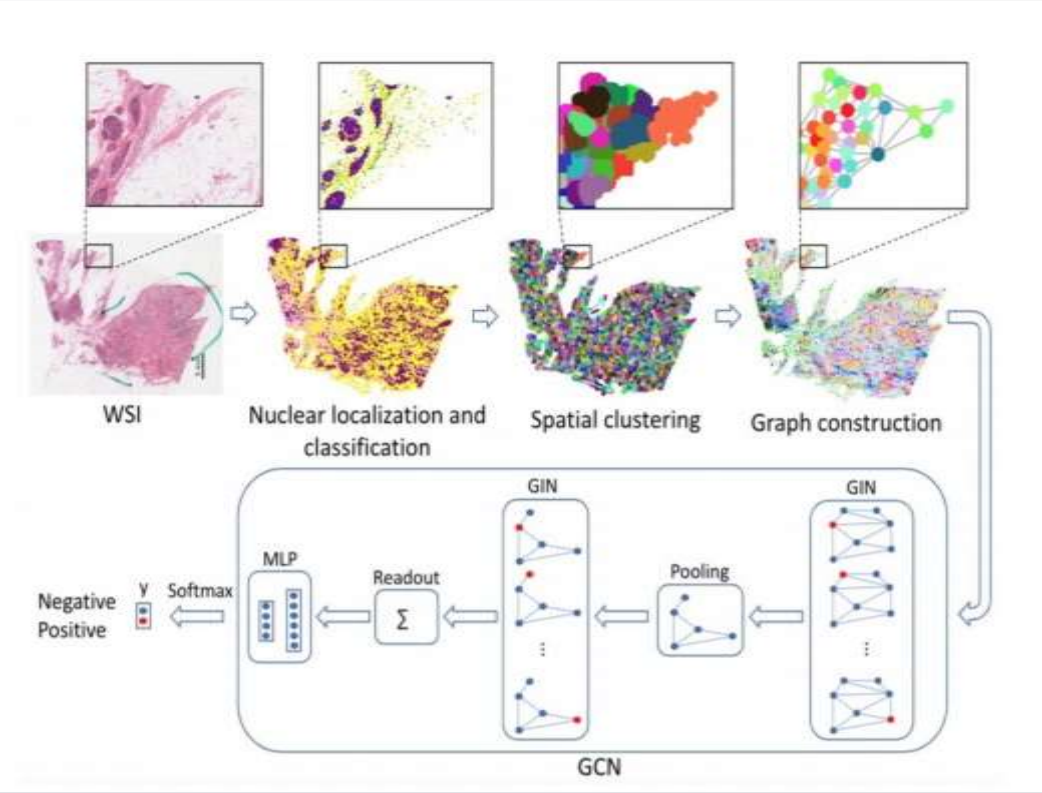
- The analysis of Whole Slide Digital pathology images using Deep learning poses a unique set of challenge
- The primary challenge being the large size of Digital Slides. They are computationally intractable and resizing them to low resolution leads to loss of information.
- Most of the research deals with this by breaking down the images into smaller tiles/patches for training and processing
- Patch based approaches have their own issues. They do not capture visual context. Segregation of individual patch level predictions into slide level labels is also a major issue

Challenges

- **Sparse coding of pathology slides** for learning features and inferring representations of cancer pathology slides
- LSTM, Conditional Random Fields are used to model correlation between neighboring patches, Cascaded CNNs, Visual Attention model are also used to capture contextual information.
- A very novel(CVPR 2020) work titled "**Capturing Cellular Topology in Multi-Gigapixel Pathology Images"** addresses this issue by using a graph convolutional neural network based model (called Slide Graph)

Using Graph Neural Network

- It works by modelling a WSI image as a single graph and using GNN
- They use Regression models(Hovernet, discussed later) to localize and classify individual nuclei
- Performed spatial clustering on these nuclei and model each cluster as a node, with edges capturing possible tissue signaling mechanisms
- The graphs are then used for classification



Other Challenges

- Lack of data is another big challenge. It might arise due to general unavailability of sufficient slides, or due to problems associated with Annotation.
- Tiling helps in multiplying the dataset when it comes to classifying patches. Also Image augmentation techniques such as affine transformations, colour and noise variations are widely used to increase data and introduce variability for generalization.
- Unsupervised and weakly supervised algorithms also help when we don't have sufficient labelled data

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 based models

Supervised Classification

- **Local Level classification** entails identifying cells or nuclei in patches of whole slide image.
- Deep learning has proven extensively successful in pixel wise prediction by sliding window approach to train on small image patches
- These image patches are often annotated by the pathologist as a region containing an object of interest (e.g., cells/nuclei) or a background

Supervised Classification

- Qaiser et al in their paper titled "**Fast and Accurate Tumor Segmentation of Histology Images using Persistent Homology and Deep Convolutional Features"** used Persistent Homology Profiles as distinguishing features in order to segment colon tumor regions by classifying patches as tumor regions or normal ones.
- They use PHP of training dataset in combination with features extracted using CNN and then employ Random Forest regression separately followed by multi-stage ensemble strategy for final classification.
- This Hybrid approach is both accurate and efficient wrt inference speed

Supervised Classification

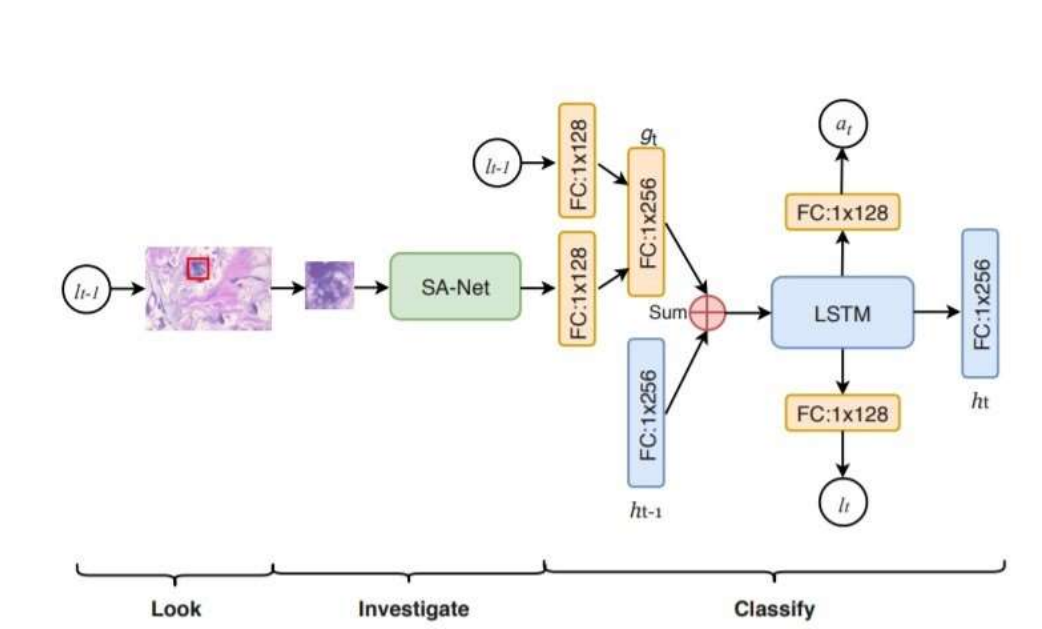
- In **global level classification**, most of the published work focusses on patch-based classification approach for whole-slide level disease prediction task.
- Localization and/or WSI level disease identification/grading
- The main disadvantage of these methods is the relatively long computational time required to carry out a dense patch-wise prediction over an entire WSI.\
- Some works combined sampling strategies for patches along with CNN training to address this.

Supervised Classification

- Some recent studies (Qaiser and Rajpoot, 2019; BenTaieb and Hamarneh,2018; Xu et al., 2019) explored task-driven visual attention models
- Xu et al in their work titled "**LOOK, INVESTIGATE, AND CLASSIFY: A DEEP HYBRID ATTENTION METHOD FOR BREAST CANCER CLASSIFICATION"** 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 to make decisions to classify the image region and also to predict the location of the image region to be investigated at the next time step. Only a fraction of pixels then need to be investigated

Advantages of Attention based models

- 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



Head and Neck Cancer Detection

- A recent work titled "**Head and Neck Cancer Detection in Digitized Whole-Slide Histology Using Convolutional Neural Networks**" deals with this.
- They used a patch based localization and whole slide classification for Squamous cell carcinoma and Thyroid Cell carcinoma using CNN.
- 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

Head and Neck Cancer Detection

- To ensure generalization the number of image patches were 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 H&E staining between slides, the hue, saturation, brightness, and contrast of each patch were randomly manipulated to make a more rigorous training paradigm.
- The proposed method is able to detect and localize primary head and neck SCC on WSI with an AUC of 0.916 for patients in the SCC testing group and 0.954 for patients in the thyroid carcinoma testing group. Moreover, the proposed method is able to diagnose WSI with cancer versus normal slides with an AUC of 0.944 and 0.995 for the SCC and thyroid carcinoma testing groups, respectively

Supervised Regression

- In this method, we focus on directly regressing the likelihood of pixel being the center of an object to detection or localization of objects.
- Regression helps in better detection by enforcing topological constraints such as assigning higher probability values to pixels near 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 FCN architectures

Supervised Regression

- Graham et al in their paper titled "**HoVer-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images"** proposed 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 centres of mass
- These distances are then utilised 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 NP branch predicts whether or not a pixel belongs to the nuclei or background,
- whereas the HoVer branch predicts the horizontal and vertical distances of nuclear pixels to their centres of mass.
- Then, the NC branch predicts the type of nucleus for each pixel.
- 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).

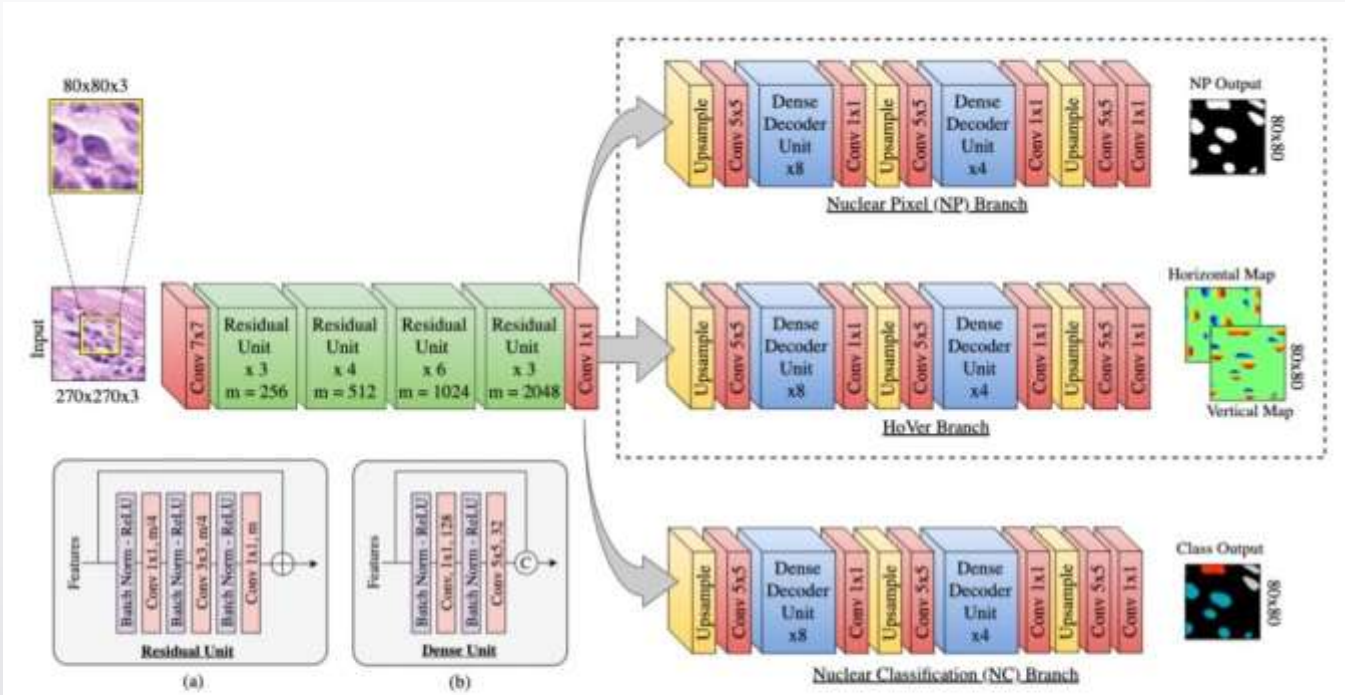


Fig. 2: Overview of the proposed architecture. (a) (Pre-activated) residual unit, (b) dense unit. m indicates the number of feature maps within each residual unit. The yellow square within the input denotes the considered region at the output. When the classification labels aren't available, only the up-sampling branches in the dashed box are considered.

Weakly Supervised Learning

- The idea of weakly supervised learning (WSL) is to exploit coarse-grained (image-level) annotations to automatically infer fine-grained (pixel/patch-level) information
- This is well suited for histopathology domain where pixel level annotations are not readily available as compared to coarse level annotations
- One popular form of WSL is Multiple Instance Learning where we have bags labeled as positive or negative, and each bags consist of instances whose label is to be predicted

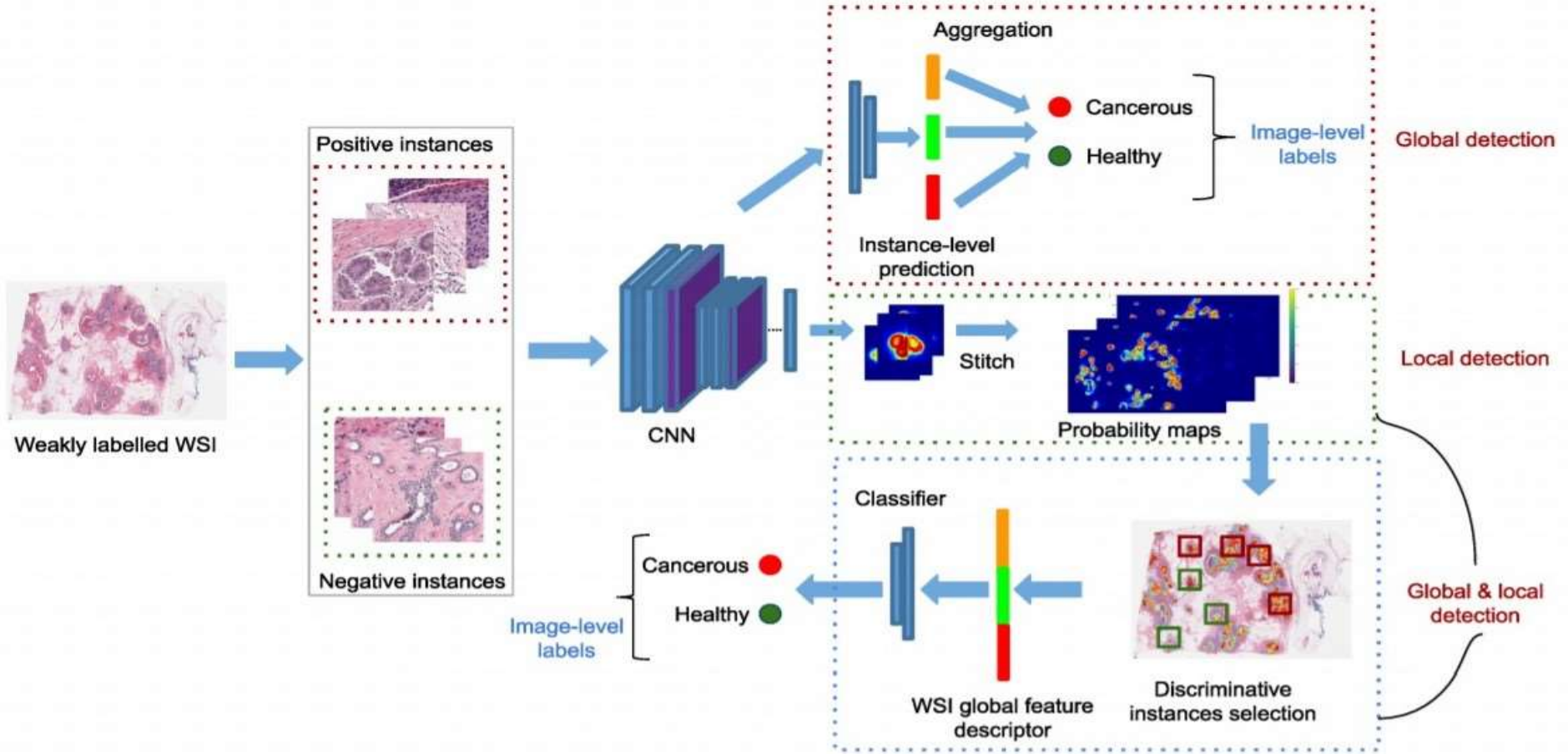
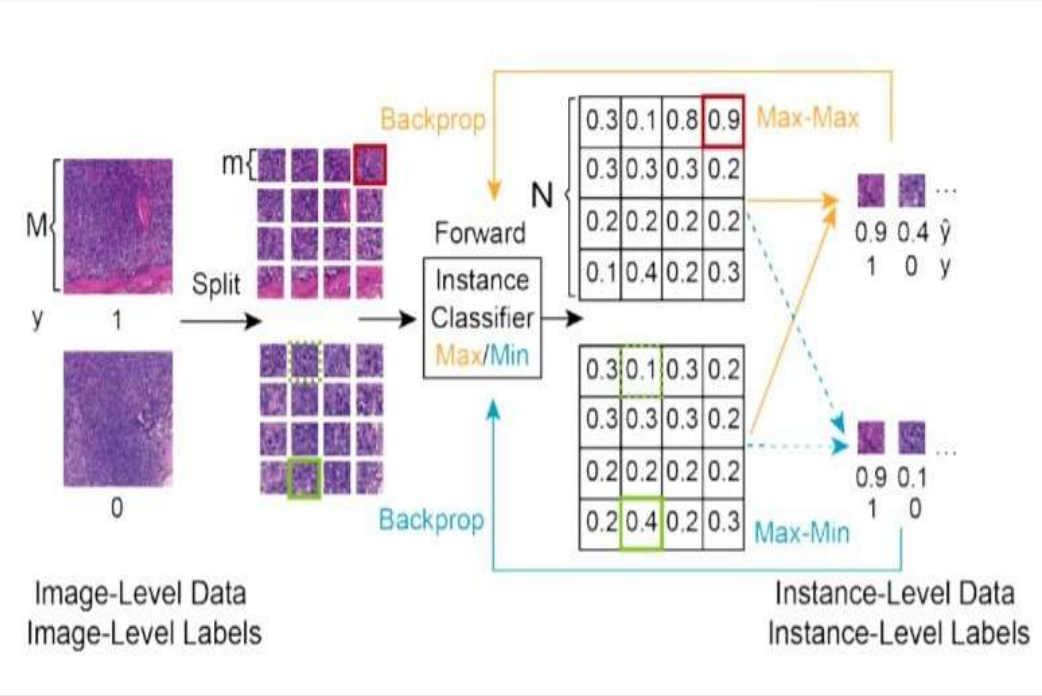


Figure 4: An overview of weakly supervised learning models.

Weakly Supervised Learning

- The paper "**CAMEL: A Weakly Supervised Learning Framework for Histopathology Image Segmentation"** proposes a weakly supervised algorithm for pixel level segmentation with only weak image level labels.
- Using combined multiple instance learning (cMIL)-based label enrichment, CAMEL splits the image into latticed instances and automatically generates instance-level labels.
- After label enrichment, the instance-level labels are further assigned to the corresponding pixels, producing the approximate pixel level labels and making fully supervised training of segmentation models possible

- In cMIL, two MIL-based classifiers with different instance selection criteria (Max-Max and Max-Min) are used to select instances to construct the instance-level dataset
- We choose ResNet-50 [11] as the classifier. The two MIL-based classifiers are trained separately under the same configuration in the forward pass, we use the MaxMax (or Max-Min for the other classifier) criterion to select one instance from each bag based on their predictions, and the prediction of the selected instance is regarded as the prediction of the image



Unsupervised Learning

- The goal of unsupervised learning is to learn something useful about the underlying data structure without the use of labels.
- Most unsupervised approaches aim to maximize the probability distribution of the data, subject to some constraints, in order to limit the solution space and to achieve a desired grouping/clustering
- A common technique is to transform the data into a lower dimensional subspace, followed by aggregation of feature representations into mutually exclusive or hierarchical clusters

Unsupervised Learning

- Autoencoders are the most popular choice for dimensionality reduction
- As these methods are constructed to disentangle relationships between samples in the dataset for grouping (or clustering), a successful unsupervised learning method can also improve the interpretability of a model, by examining how the model groups items into separate categories

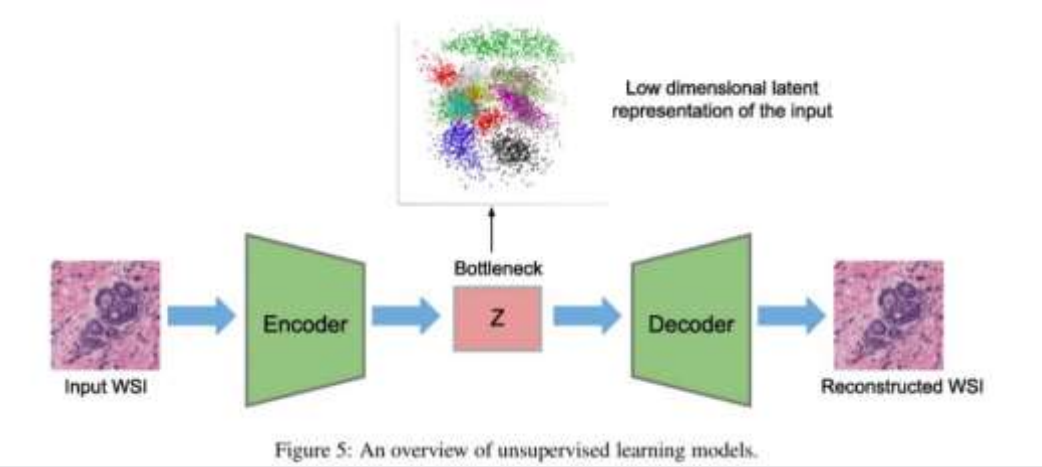
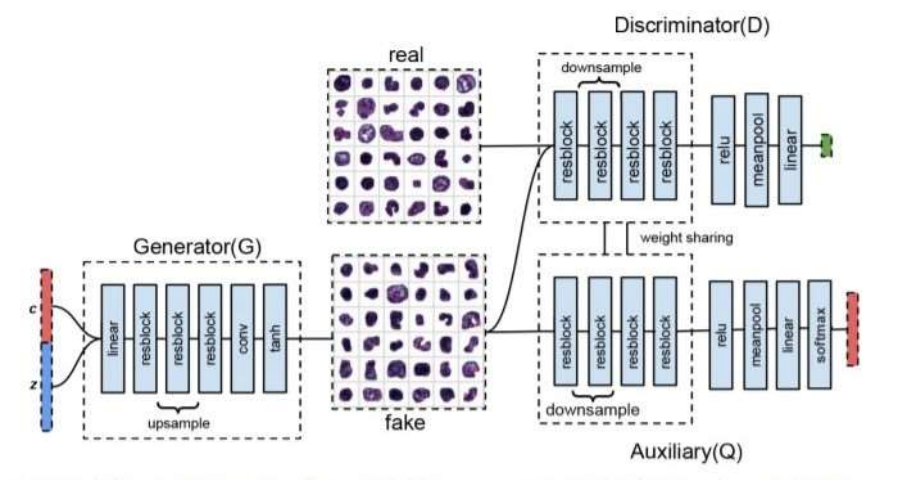
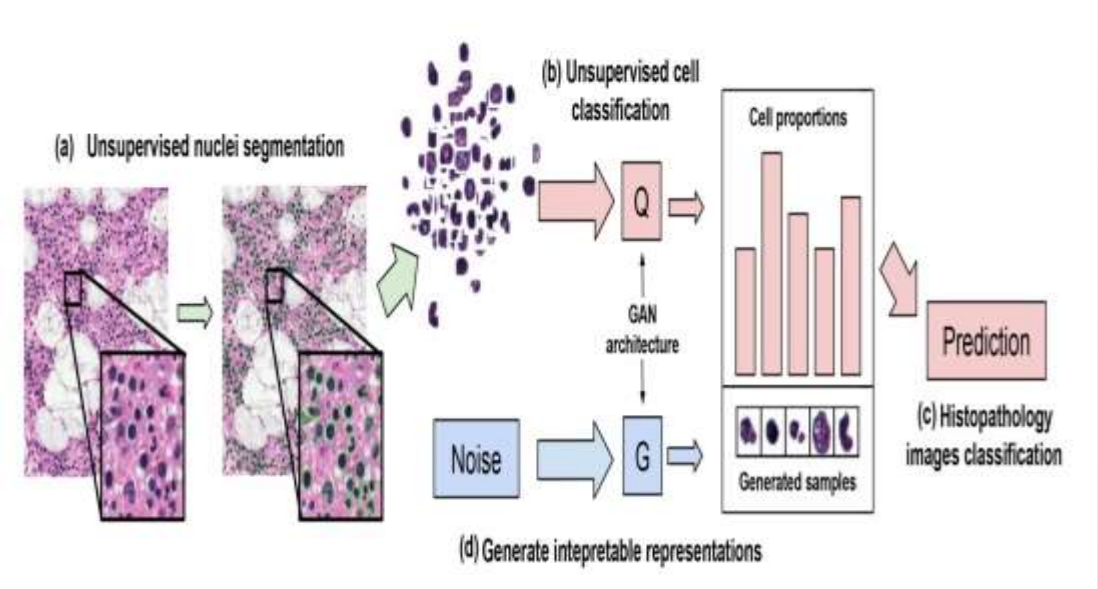


Figure 5: An overview of unsupervised learning models.

Unsupervised learning

- **"Unsupervised Learning for Cell-level Visual Representation in Histopathology Images with Generative Adversarial Networks"** proposes a unified generative adversarial networks architecture to perform robust cell-level visual representation learning in an unsupervised setting
- An auxiliary network maximizes mutual information between generated images and random variables in order to be able to perform unsupervised classification of cell features.
- Later Based on proportions of varieties of cellular elements, image level classification is performed.





Unsupervised Learning

Transfer learning

- The most popular and widely adopted technique in digital pathology is the use of transfer learning approach
- In transfer learning, the goal is to extract knowledge from one domain and apply it to another domain by relaxing the assumption that the train and test set must be independent and identically distributed.
- Two major ways in which TL is applied in histopathology
  - Using Pre trained CNN as a feature extractor
  - Fine Tuning

Transfer Learning

- Papers like "**Cytokeratin-Supervised Deep Learning for Automatic Recognition of Epithelial Cells in Breast Cancers Stained for ER, PR, and Ki-67"** have applied transfer learning for automated epithelial cell detection.
- A partially pre-trained deep convolutional neural network was fine-tuned using image batches from 152 patient samples of invasive breast tumors
- A good discrimination of epithelial cells was achieved (AUC of mean ROC = 0.93), which was well in concordance with pathologists' visual assessment

References

1. Qaiser, T., Tsang, Y.W., Taniyama, D., Sakamoto, N., Nakane, K., Epstein, D., Rajpoot, N., 2019b. Fast and accurate tumor segmentation of histology images using persistent homology and deep convolutional features. Medical Image Analysis 55, 1–14.
2. Xu, B., Liu, J., Hou, X., Liu, B., Garibaldi, J., Ellis, I.O., Green, A., Shen, L., Qiu, G., 2019. Look, investigate, and classify: A deep hybrid attention method for breast cancer classification, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 914–918.
3. Graham, S., Vu, Q.D., Raza, S.E.A., Azam, A., Tsang, Y.W., Kwak, J.T., Rajpoot, N., 2019b. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Medical Image Analysis 58, 101563.
4. Xu, G., Song, Z., Sun, Z., Ku, C., Yang, Z., Liu, C., Wang, S., Ma, J., Xu, W., 2019. CAMEL: A weakly supervised learning framework for histopathology image segmentation, in: Proceedings of the IEEE International Conference on Computer Vision, pp. 10682–10691.
5. Valkonen, M., Isola, J., Isola, J., Ylinen, O., Muhonen, V., Saxlin, A., Tolonen, T., Nykter, M., Ruusuvuori, P., 2019. Cytokeratin-supervised deep learning for automatic recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67. IEEE Transactions on Medical Imaging , 1–1

THE END