TRANSFER LEARNING FROM NUCLEUS DETECTION TO CLASSIFICATION IN HISTOPATHOLOGY IMAGES

Safoora Yousefi*†  Yao Nie*

* Roche Tissue Diagnostics, Santa Clara CA  †Emory University, Department of Computer Science, Atlanta GA

ABSTRACT
Despite significant recent success, modern computer vision techniques such as Convolutional Neural Networks (CNNs) are expensive to apply to cell-level prediction problems in histopathology images due to difficulties in providing cell-level supervision. This work explores the transferability of features learned by an object detection CNN (Faster R-CNN) to nucleus classification in histopathology images. We detect nuclei in these images using class-agnostic models trained on small annotated patches, and use the CNN representations of detected nuclei to cluster and classify them. We show that with a small training dataset, the proposed pipeline can achieve superior nucleus detection and classification performance, and generalizes well to unseen stain types.

Index Terms— Cell detection, cell classification, transfer learning, convolutional neural networks

1. INTRODUCTION
Nucleus detection is a core task in computational pathology and refers to the identification and localization of individual cells in histopathological images. Such images come in varied sizes, resolutions, stain types, and are often crowded with visually heterogeneous, overlapping cells. Among the challenges that are commonly faced in nucleus detection from these images are the difficulty of obtaining ground truth annotations, and transferring a model tuned for a specific resolution or stain to a dataset with different characteristics.

We approach nucleus detection as the general object detection problem where the goal is to find and classify a variable number of objects on an image. Object detection is a well studied problem in machine learning and computer vision where recently deep convolutional neural networks have shattered performance benchmarks. However, most of the advances made in this field assume availability of large training datasets, and are designed to handle relatively small images with a few objects in them, so applying these algorithms to whole slide images is not straightforward.

Building upon the Faster-RCNN deep object detection model [1] we propose an efficient nucleus classification pipeline that can be used for AI-assisted ground truth annotation. We train a Faster-RCNN for class-agnostic nucleus detection, and use the learned model for detection of and feature extraction from new classes of nuclei. Extracted features are later used to cluster nuclei into homogeneous subsets that can be labeled in bulk by human annotator, reducing the burden of annotation from an order of number of nuclei to the number of clusters. We demonstrate the ability of the learned model to detect nuclei in datasets with mixed sizes and resolutions, and to generalize to stain types that it has not seen during training.

2. RELATED WORK
Adaptive nucleus localization and classification has been approached previously as object detection, semantic segmentation, or regression. Regression random forests were used in [2] to predict nucleus centers. Object counting regression methods have also been applied to nucleus detection [3]. CNNs were used for supervised nucleus segmentation in microscopy images [4, 5] and supervised nucleus detection and classification in histopathology images [6].

CNNs are known to learn features that are generalizable to new tasks, and image representations derived from pre-trained CNNs have become the state of the art in computer vision tasks such as instance retrieval [7] and cell classification [8]. To our knowledge, this is the first attempt to measure the transferability of nucleus detection CNN features to nucleus classification task.

3. DATASETS
Pre-training dataset: We used a Faster-RCNN model released along with the Tensorflow Object Detection API [9]. The model was pre-trained on the COCO dataset [10] which is a large scale general object detection dataset with 300K images from 80 categories of animals, furniture, vehicles, etc.

Fine-tuning datasets: In order to familiarize the pre-trained Faster-RCNN with the unique characteristics of histopathology images, we fine-tuned it on the following immunohistochemistry (IHC) bright field image datasets annotated by our trained research associates with class-agnostic bounding box annotations. These images are originally several thousands of pixels in each dimension, therefore, to facilitate annotation we extracted random patches of 64x64 or 128x128 pixels
### 5. EXPERIMENTS

Fine-tuning the pre-trained Faster-RCNN model on different datasets resulted in the following models:
- **Model Zero**: Faster RCNN model with Resnet-50 feature extractor pre-trained on COCO dataset.
- **HER2**: Model Zero fine-tuned on nucleus detection on Her2 data.
- **DUALISH**: Model Zero fine-tuned on nucleus detection on Dual ISH data.
- **Model A**: Model Zero fine-tuned on nucleus detection on Combined data.
- **Model B**: Model Zero fine-tuned on detection and classification on BB data.
- **Model C**: Model A fine-tuned on detection and classification on BB data.

#### 5.1. Nucleus detection

A summary of detection (and classification, where class labels are present) performances of our models is given in Table 2. In Fig. 1, a few examples of detected vs. ground truth nuclei are presented. It can be seen that model A sometimes detects nuclei missed by the annotator and objects of ambiguous nature, resulting in model penalization.

In Fig. 2, examples of nucleus detection results on large inference images are given. Such detected nuclei are used in later stages of the pipeline as described in the next subsections.

#### 5.2. Nucleus clustering

During the process of learning to detect and localize cells, Model A learns features that can facilitate clustering of nuclei. Although model A was not provided with any class labels, these features yield better clustering results compared to the original RGB values of cells (See Fig. 3).
Table 2. Nucleus detection and classification performance on held-out validation data.

| Validation Data | Model  | mAP@50  |
|-----------------|--------|---------|
| Her 2           | HER2   | 78%     |
| Dual ISH        | DUALISH| 86%     |
| Combined        | Model A| 85%     |
| BB              | Model B| 70%     |
| BB              | Model C| 72%     |

Fig. 1. Top: Examples of nucleus detection results using model A on Her2 and Dual ISH images. Bottom: Ground truth annotations for the corresponding patches.

5.3. Clustering of unseen nuclei classes

In this section, we investigate the transferability of class-agnostic nucleus detectors to clustering nucleus classes/stain types that the model has not seen during training. Fig. 4 summarizes clustering performance on BB data using representations extracted from different layers of Model A as well as the original RGB representation of nuclei. With only two clusters, Faster-RCNN features have a 8% advantage over RGB representation; This means if we cluster nuclei into 2 clusters using Faster-RCNN features and assign the majority class to all nuclei in each cluster, we get 98% classification accuracy compared to 90% if we use RGB representations instead.

Fig. 2. Model A inference on a 796x806 Dual ISH image.

Fig. 3. Qualitative nucleus clustering performance. a. Features extracted from block 1 of Model A enable meaningful clustering of nuclei. b. Clustering of cells based on their original RGB pixel values does not look meaningful.

5.4. Transferring further

BBYP images are stained by Discovery Yellow and Discovery Purple chromogens to identify 5 classes of cells of interest, i.e., ki67+/- tumor cells, Ki67+/- T-Cells and CD8+ T-Cells. Some of these classes are easily distinguishable by color, while others are defined by both color and morphology or context (See Fig. 5). We hypothesize that since CNN features provide a more global view of objects due to their multi-level abstraction compared to original image pixels, they should provide better distinction between nucleus classes that look similar in individual appearance but can be distinguished by contextual information. As shown in Figure 6, clustering BBYT nuclei using Model B’s features yields a 15% advantage over using original RGB representation of these nuclei, and continue to outperform them as we increase the number of clusters, supporting our initial hypothesis. Using features from other models led to similar results.
Fig. 4. Evaluation of clustering based on Model A representation of BB validation data.

Fig. 5. Examples of T-Cells (right) and tumor cells (left) in BBYP dataset. Ki67+ and Ki67- tumor cells have the same color as Ki67+ and Ki67- T-Cells, respectively. These cells can be differentiated based on cell size, shape, context, etc.

6. CONCLUSION

Our proposed pipeline facilitates ground truth labeling by detecting nuclei in histopathology images and clustering detected nuclei into a few homogeneous subsets that can be annotated in bulk, reducing the burden of ground truth labeling from an order of number of cells in the dataset to the number of clusters.

We show that features learned from the COCO dataset are transferable for nucleus clustering and achieve over 98% cluster homogeneity on BB validation data. Using only 1800 bounding box annotations, we demonstrate that with the right training, Faster R-CNN can handle histopathology images of different magnifications and sizes at the same time, and perform inference on images 100x the size of training patches (the limitation here is not algorithmic but memory-related).

7. REFERENCES

[1] S. Ren, K. He, et al., “Faster r-cnn: Towards real-time object detection with region proposal networks,” in NIPS, 2015.

[2] P. Kainz, M. Urschler, et al., “You should use regression to detect cells,” in MICAI, 2015.

[3] W. Xie, J.A. Noble, et al., “Microscopy cell counting and detection with fully convolutional regression networks,” Computer methods in biomechanics and biomedical engineering: Imaging & Visualization, 2018.

[4] S. Akram, J. Kannala, et al., “Cell segmentation proposal network for microscopy image analysis,” in Deep Learning and Data Labeling for Medical Applications. Springer, 2016.

[5] J. Chen, S. Banerjee, et al., “Neuron segmentation using deep complete bipartite networks,” in MICCAI, 2017.

[6] R. Bidart, M. Gangeh, et al., “Localization and classification of cell nuclei in post-neoadjuvant breast cancer surgical specimen using fully convolutional networks,” in Medical Imaging 2018: Digital Pathology, 2018.

[7] A. Salvador, X. Giro-i Nieto, et al., “Faster r-cnn features for instance search,” in CVPR Workshops, 2016.

[8] N. Bayramoglu and J. Heikkila, “Transfer learning for cell nuclei classification in histopathology images,” in ECCV Workshops, 2016.

[9] J. Huang, V. Rathod, et al., “Speed/accuracy trade-offs for modern convolutional object detectors,” in IEEE CVPR, 2017.

[10] Tsung-Yi Lin, Michael Maire, Serge Belongie, et al., “Microsoft coco: Common objects in context,” in ECCV, 2014.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, et al., “Scikit-learn: Machine learning in Python,” Journal of Machine Learning Research, 2011.

[12] Christoph F Eick, Nidal Zeidat, and Zhenghong Zhao, “Supervised clustering-algorithms and benefits,” in Tools with Artificial Intelligence, 2004. IEEE International Conference on, 2004.