Segmentation, Classification and Counting of Colon Nuclei using Hybrid Deep Learning Model

Ankush D. Jamthikar, Pranay Dumbhare, Vedant M. Phuse, Yash Dubey, Himanshu Padole, Deep Gupta

***Abstract*—Colon nuclei detection is important to mitigate abnormalities or diseases like cancer in its early stages. The objective of this study is to perform the colon nuclei segmentation, classification, and counting of the nuclei (cellular composition). This study uses the baseline 'HoVer-Net' model proposed by the organizers of the CoNIC2022 and modified it to incorporate handcrafted features. The revised model is termed a hybrid deep learning model that combines both deep features and handcrafted features. The handcrafted features are generated from the Local Binary Patterns and the Histogram of Oriented Gradients feature descriptors. The hybrid deep learning model is trained for 50 epochs and validated on the CoNIC2022 training data. The proposed hybrid deep learning model shows an improvement over the baseline 'HoVer-Net' model in both segmentation and classification as well as nuclei counting tasks**

***Index Terms*—Colon nuclei segmentation, HoVer-Net, hybrid model**

I. Introduction

The manual assessment of Haematoxylin and Eosin (H&E) stained histology slides suffers from intra- and inter-observer variability. To overcome the challenges in the visual assessment of tissues, there is a growing interest in digital pathology, which uses a class of histology images that are used to generate the whole slide images (WSI). Each WSI contains several nuclei of cells of different types, which can be analyzed to predict the clinical outcomes such as survival or grade and the type of diseases, such as cancer. Efficient segmentation and detection of different nuclei cell types can provide important information about the different tissues that contribute to the disease growth. This is the idea behind the CoNIC2022 challenge, which specifically required segmenting and classifying the colon nuclei as well as finding the cellular composition by counting the different types of nuclei present in the cells. Nuclei segmentation is a challenging task because of the heterogeneous nuclear shapes, sizes, structures, as well as overlapping nuclei clusters.

Similar to nuclei segmentation, nuclei classification is also an important task that can identify and classify different kinds of cells. Currently, the colon cancer diagnosis is based on the human examination of relevant histopathological images by trained pathologists. This process is time-consuming and the interpretation is subject to variability between different observers and even between the different images by the same observer.

Artificial intelligence (AI) is gaining popularity in almost every domain and has greatly contributed to the field of healthcare including computational pathology. In recent years, several automated AI-based systems have been proposed for nuclei detection and identification tasks. Hamad et al. [1] proposed a convolution neural network-based system for colon nuclei classification using Haematoxylin and Eosin (H&E) stained histology images. Another study used the spatially constrained convolutional neural network to perform nuclei detection that predicts the probability of any pixel being the centre of the nucleus [2]. Hofener et al. [3] used a deep learning-based approach to predict the proximity map of image pixels with the center of nuclei. This approach was primarily used to identify nuclei from the histology images. Most recently, Graham e al. [4] proposed a model called HoVer-Net that uses a deep learning strategy for simultaneously segmenting and classifying nuclear occurrences in histology images. The HoVer-Net model is based on the prediction of horizontal and vertical distances between nuclear pixels and their centres of mass, which are then used to separate the grouped nuclei. The nuclear type is then determined for each segmented instance using a specialized up-sampling branch.

In this proposed study, we used the HoVer-Net as a baseline model and extended its structure by combing the deep features and the handcrafted features to segment and classify the nuclei images. Although the deep learning-based model is self-sufficient to extract features from images (so-called deep features), recent studies have also shown the improvement in the segmentation and classification tasks using the handcrafted features [5] Therefore, we hypothesized that combining handcrafted features with deep features may result in an improvement in nuclei segmentation and classification.

II. Methods

A. Data Preprocessing and Augmentations

This study used the original dataset provided by CoNIC 2022 challenge. The dataset is comprised of 4981×256×256×3 RGB patches extracted from the original Lizard dataset. Besides this, the dataset also included the 4981 segmentation and classification maps. Both types of patches are available in the form of .npy array format, which can be easily recognised and read by a python-based library. Once the images are read, augmentation is performed using Gaussian blur and median blur to the images, as well as perturbing the hue, saturation, contrast, and brightness of the input images.

B. Proposed Model

In this study, we used the baseline HoVer-Net model and tried to extend its architecture by including the handcrafted

All authors with the Visvesvaraya National Institute of Technology Nagpur 440010, India.

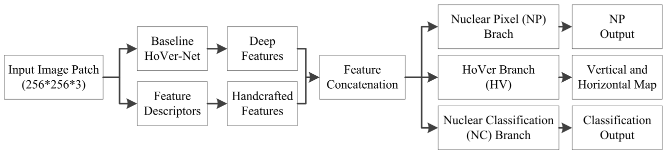

Fig. 1. Global view of the proposed study

features. The revised architecture is termed as hybrid deep learning or hybrid HoVer-Net model. There are two parts to the hybrid HoVer-Net model: (i) encoder and (ii) decoder. The encoder hosts the ResNet50-based deep learning model and the decoder contains the three different parallel layered architectures decoding the features into segmentation and classification maps. The ResNet50-based encoder provides the image representations of size 2048×32×32. These image representations are also called deep features, that are extracted by the encoder. These deep features are nothing but 2-dimensional feature maps. We have also added a branch parallel to the ResNet50 network that extracted the image representations from the popular predefined feature descriptors. These feature descriptors are local binary patterns (LBP) and the histogram of gradients (HoG). Figure 1 shows the global view of the proposed model. Out of three decoder architectures, the nuclear pixel (NP) branch predicts whether or not a pixel belongs to the nuclei or background. Similarly, the HoVer branch predicts the horizontal and vertical distances of nuclear pixels to their centres of mass. Then finally, the Nuclear Classification (NC) branch predicts the type of nucleus for each pixel and helps in finding the cellular composition by counting the nuclei.

C. Handcrafted Feature Descriptors

The handcrafted features used in our existing model are Local Binary Patterns (LBP) [6] and the Histograms of Oriented Gradients (HoG) [7]. The Local Binary Patterns (LBP) is a simple yet very efficient texture descriptor, which labels the pixels of an image by thresholding the neighbourhood of each pixel and considers the result as a binary number. LBP detects microstructures such as edges, lines, spots, flat areas, which can be estimated by the histogram. LBP descriptor initially converts the input RGB image into grayscale by averaging the pixel intensities. Then, a neighbourhood of each pixel is thresholded with the intensity value of the centre pixel. If the neighbouring pixel intensity is greater than or equal to the center pixel intensity, then the corresponding pixel is replaced by 1 otherwise 0. The binarized neighbourhood of the centre pixel is then converted to the decimal value, which is finally used to replace the center pixel intensity level. This process is repeated for entire greyscale image. The revised LBP feature map is of the size 1 x 1024, which is then converted into a shape that can be matched with the deep features for feature concatenation in further steps. In this study, the entire LBP feature maps are converted into 1 x 32 x 32 feature maps, which are concatenated with deep feature maps.

Besides, LBP, the Histograms of Oriented Gradients (HoG) is also an important feature descriptor, which focuses on the structure or the shape of objects in the input image. HOG can provide the edge directions. This is done by extracting the gradient and orientation (magnitude and direction) of the edges. Steps involved in HOG feature extraction are:

1) Preprocessing and reshaping the input image to size 64 × 128
2) Calculating gradients (direction x and y) for a particular pixel
3) Calculating the magnitude and orientation of gradients
4) Creating histograms using gradients and orientation
5) Normalizing the gradients
6) Calculating the features for the complete image

The output feature vector from the HOG descriptor has dimensions 1 x 34596, which is further converted to the feature maps of size 33 x 32 x 32. Combining deep features with the LBP and HoG feature maps has resulted in the final handcrafted feature maps with the dimension of 2082 x 32 x 32 . These feature maps are then fed to the three decoder architectures. The hybrid HoVer-Net model is trained for 50 epochs and validated with the fold index 0 of the training dataset as provided in the baseline HoVer-Net script of the CoNIC2022 challenge.

III. RESULTS AND DISCUSSION

The proposed hybrid deep learning-based model is compared with the baseline hovernet model on the basis of multi-class panoptic quality (mPQ+) and panoptic quality (PQ) for nuclei segmentation and classification. On training for 10 epochs, the proposed hybrid model demonstrated an improvement of 1.71% (0.606 vs. 0.595) and 1.59% (0.4435 vs. 0.4365) in PQ and MPQ, respectively, over the baseline hovernet model.

The nuclei cellular composition predicted using the proposed hybrid model and the baseline model is compred using the multi-class coefficient of determination (R2). The proposed model demonstrated an improvement of 8.5% in R2 value over the baseline hybrid model (0.775355 vs. 0.714375).

IV. CONCLUSION

This short study provided the hybrid model of the hovernet by comining the deep features with the two types handcrafted features. The hybrid model has demonstrated the marginal improvement in multi-class panoptic quality and panoptic quality metrics, which are the key metrics used for segmentation and classification tasks.

REFERENCES

[1] A. Hamad, F. Bunyak, and I. Ersoy, “Nucleus classification in colon cancer h&e images using deep learning,” *Microscopy and Microanalysis*, vol. 23, no. S1, pp. 1376–1377, 2017.

[2] K. Sirinukunwattana, S. E. A. Raza, Y.-W. Tsang, D. R. Snead, I. A. Cree, and N. M. Rajpoot, “Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images,” *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1196–1206, 2016.

[3] H. Höfener, A. Homeyer, N. Weiss, J. Molin, C. F. Lundström, and H. K. Hahn, “Deep learning nuclei detection: A simple approach can deliver state-of-the-art results,” *Computerized Medical Imaging and Graphics*, vol. 70, pp. 43–52, 2018.

[4] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Medical Image Analysis*, vol. 58, p. 101563, 2019.

[5] H. Khan, P. M. Shah, M. A. Shah, S. ul Islam, and J. J. Rodrigues, "Cascading handcrafted features and convolutional neural network for iot-enabled brain tumor segmentation," *Computer Communications*, vol. 153, pp. 196–207, 2020.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[7] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1. Ieee, 2005, pp. 886–893.