

Prediction of HER2 Score in Breast Cancer using Synthetic IHC Derived from H&E Images

Xingwei Wang^{1,*}, Auranuch Lorsakul^{1,*}, Margaret Zhao^{1,*}, Yao Nie^{1,*}, Hartmut Koeppen^{2,*}, Hauke Kolster^{3,*} and Kandavel Shanmugam^{4,*}

1 Imaging and Algorithms, Digital Pathology, R&D, RTD, Santa Clara, California

2 Pathology Office, Genentech, South San Francisco, California

3 Digital Pathology, Oncology Biomarker Development, Genentech, South San Francisco, California

4 Medical Innovation, RTD, Tucson, Arizona

*: All authors equally contributed and this work achieved the 3rd place winner in the 2020 OneRoche Hackathon event

Background: The current practice to diagnose human epidermal growth factor receptor 2 (HER2) positive breast cancer commonly relies on hematoxylin and eosin stain (H&E). To confirm a breast cancer diagnosis, extra multiple tissue sections for immunohistochemistry (IHC) slides are required. In this preliminary study, we explored the feasibility of predicting HER2 scores from H&E stained tissue sections by generating synthetic HER2-IHC images based on the H&E images.

Methods: We used a deep-learning architecture, a conditional Generative Adversarial Network (cGAN), to generate the synthetic HER2-IHC images from H&E images. Using the IRIS platform, a pathologist first created the ground-truth HER2 scores (0+, 1+, 2+, and 3+) within annotated tumor regions. Then, image patches from corresponding tumor areas in the IHC and H&E images were extracted and aligned. Multiple experiments for training/testing were performed for four different HER2 scores, using 7,472 and 1,900 patches of size 128x128/256x256, respectively. We used 80/20 split for training/testing, applying Adam optimizer, learning rate 0.0002, epochs 100/200 for small/large patches.

Results: To test the feasibility of using synthetic IHC images to predict HER2 scores, we selected three synthetic and three real stained image patches for each of the four HER2 scores. The images were randomly presented to one pathologist, who evaluated the HER2 scores and determined which IHC image was a tissue-based stain (“Real”) or synthetic (“Fake”) image. In the preliminary result, the pathologist reached an accuracy of 11/24 (45.8%) in identifying the correct image type, which shows that synthetic images were indistinguishable to the pathologist. The accuracy in predicting HER2 scores for image types “Synthetic” and “Real” was 10/12 (83.33%) and 11/12 (91.67%).

Conclusions: The synthetic HER2-stained IHC images created by a trained cGAN network were indistinguishable from real HER2-IHC images in our experiment. The HER2 scores evaluated based on the synthetic IHC images have a high degree of consistency with the scores based on the real IHC images. This suggests that the synthetic HER2-IHC images might be used as an early read before the availability of traditional tests. A further application can be the generation of large-scale image data for algorithm verification and network training.