Deep High-Resolution U-Net Latent Diffusion Model for Histopathological Image Synthesis

Sun, Ka Chon (School: Kao Yip Middle School)
Chan, Hang Kin (School: Kao Yip Middle School)
Xiao, Ian I (School: Kao Yip Middle School)

Oral Squamous Cell Carcinoma (OSCC) is a highly invasive malignant tumor with significant health risks worldwide. Following the success of deep learning models over the past decade. Some studies demonstrate that deep learning-based models achieve 80%–90% accuracy in classifying OSCC. However, training deep learning models for classification requires large amounts of annotated data, which is expensive to collect and label. Our project introduces a novel Deep High-Resolution U-Net Latent Diffusion Model to generate high-quality histopathological images to enhance OSCC classification. By leveraging a ChatGPT-4o-based generator-discriminator framework, which generates detailed text prompts from unlabeled images. We integrate advanced conditioning embeddings by concatenating the CLIP text encoder and the PLIP text encoder outputs with HIPT-generated attention masks. The proposed HR-UNet architecture, adapted from the UNet design of Stable Diffusion 2.1, captures cellular details while maintaining multi-scale contextual information. We also fine-tuned the VAE decoder using a novel LPPIPS (Learned PLIP-image-encoder-based Perceptual Image Patch Similarity) loss function for better adapted medical images. Our approach improves the Fr ´echet Inception Distance (FID) from 111.7 to 24.7, and the OSCC classification accuracy from 91.17% to 96.25% by adding synthetic images.

Awards Won: