Exam Report

The BCEloss that my best model got on the test set is 4.17. Two network architecture, inception v3 and resnet50, were tried to get my best model. Resnet50 performed better on the dataset, brought the BCEloss from 9.11 to 8.48 when implemented together with input normalization. Further, image augmentation and more precise image preprocessing brought the BCEloss to 6.37. In the end, after analyzing incorrectly labeled images and customizing the image augmentation procedure, my best model brought BCEloss down to 4.17. Here I'll dive more into detail:

1. Network architecture
   Two network architecture, inception v3 and resnet50, and input normalization were tried. The mean and standard deviation of all the images in each channel were computed to normalize inputs, and thus to get a larger learning rate work better.

2. Image preprocessing
   After scanning the image dataset, I observed a wide range of hue and brightness among the images: the color tone of some images is bluer compared to others which are more purple, and some images are brighter compared to others. Therefore, the original BGR color channels were converted to HSV and further, the hue, saturation, and value channels were manipulated to get image augmentation. Moreover, instead of resizing the images into square matrices, which could cause distortion, the ratio of the image was kept and the central part was cropped.

3. Image augmentation
   There are three levels of image augmentation implemented:
   1) Patching more minority classes of cells: replacing red blood cells in the image with the minority classes of cells or simply adding more minority classes of cells to random places of images.
   2) Flipping, rotation, adding Gaussian noise or Gaussian blur, and color-shifting images of minority labels.
   3) Sample pairing: synthesizing a new sample from one image by overlaying another image randomly chosen from the training data (i.e., taking an average of two images for each pixel). This technique works as regularization, reducing overfitting, and is valuable for tasks with a limited amount of training data, such as medical imaging tasks.

4. Incorrectly labeled images analysis
   After analyzing the incorrectly label images, true labels and predicted labels, I found the model was not performing well on the class 'difficult' cells, therefore, more image augmentations were customized for 'difficult' cells, such as patching more 'difficult' cells on the image, or simply flipping or color shifting images labeled with 'difficult' cells.