**Vision Transformation on the same dataset.**

Vision transformation is used on the same dataset to test its performance and compare with the method is original paper. The model is trained with acceleration of GPU. Cervical cancer dataset is split into a training set and a test set with ratio of 80:20. Custom dataset SiPakMedDataset is created with inherits to apply specific transformations to the images. Data augmentation is applied to the training dataset using torchvision transforms such as RandomResizedCrop, RandomHorizontalFlip, RandomRotation, and CenterCrop. The vision transformer (ViT-B/16) pretrained model is loaded from torchvision with ImageNet weights. Then the output layer of the model is modified to match the number of classes (5) in the SiPakMed dataset and Softmax activation is applied. Additionally, We applied early stopping criterion for our model.

图表, 折线图

描述已自动生成

图表, 折线图

描述已自动生成

Based on the training and validation accuracy and loss data, we can analyze the performance of this vision transformation model. The model learns effectively from the data, as the training accuracy increases consistently over the epochs, reaching around 97.75% by epoch 7. The training loss, on the other hand, is decreasing, suggesting that the model is optimizing well. The validation accuracy also shows a general upward trend, with a few fluctuations in between. The highest validation accuracy is achieved at epoch 8, with around 95%. The validation loss is generally decreasing, but there are fluctuations, indicating that the model might be sensitive to the variations in the validation dataset. From the graphs above, it can be concluded that the model performs well on the given cervical cancer dataset. However, the model might be overfitting since the training accuracy is significantly higher than the validation accuracy in some epochs.