人工智能与商业创新中期报告

Project 1、病理图像识别

姓 名: 孟 念 Muennighoff Niklas
学 号: 1800092850
班 号: 02839210

人工智能与商业创新
(秋季, 2021)

北京大学
光华管理学院
彭一杰老师
高华西助教

2021 年 11 月 5 日

北京大学
PEKING UNIVERSITY

目 录

1 Introduction **3**

2 Progress: Image Classification **4**

2.1 Dummy Classifier 4

2.2 CNN 4

3 Progress: Image Segmentation **5**

3.1 Binary Background Segmentation 5

3.2 Mutli-class Segmentation 6

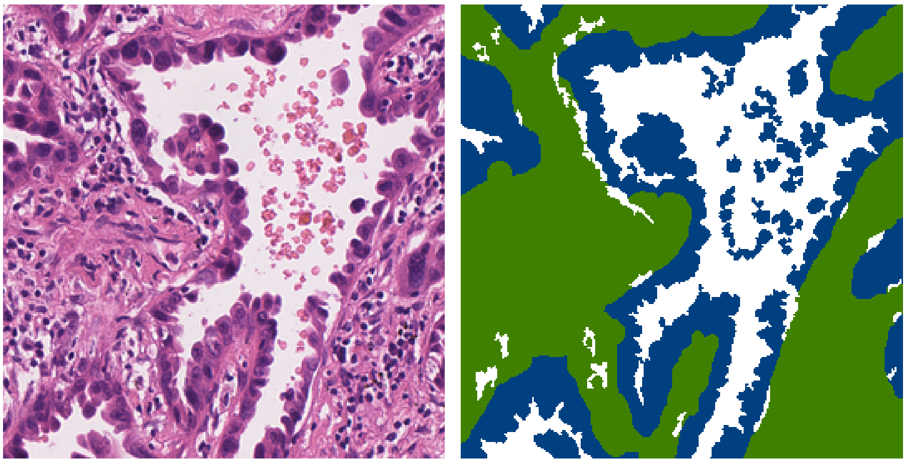
4 Conclusion **7**

1 Introduction

This work constitutes a progress report for the WSSS4LUAD Challenge dataset. The dataset consists of 10,000 images. The task as outlined in the course materials can be split into two parts:

- **Image Classification**: In this module the goal is to classify each image as to whether it contains a tumor epithelial tissue, tumor stromal tissue or normal tissue or any combination of these.
- **Image Segmentation**: In this module, the different tissue types and the background need to be annotated on a pixel-level basis for each image.

In the following, my progress on both tasks will be presented. In the final report, to be released at the end of the class, my entire work will be presented including the code.



(a) Ground truth image (b) Ground truth mask

图 1: An example image and mask from the validation set. For the mask, white (255, 255, 255) corresponds to non-labeled background, turquoise (64, 128, 0) to stroma, dark blue (0, 64, 128) to tumor. The label is hence **T+S** (Tumor + Stroma).

2 Progress: Image Classification

I have written two approaches to image classification. A dummy classifier to act as a baseline and a Convolutional Neural Network (CNN).

2.1 Dummy Classifier

Type	**N**	**S**	**T**	**S+T**	**N+T**	**T+S+N**
Train	1832	1680	1181	5393	4	1
Val	2	2	0	30	1	5

表 1: Image label distribution, N = Normal, T = Tumor, S = Stroma. Labels for the validation set were not provided in the dataset, but I extracted them with a custom script. The script will be provided in the final report.

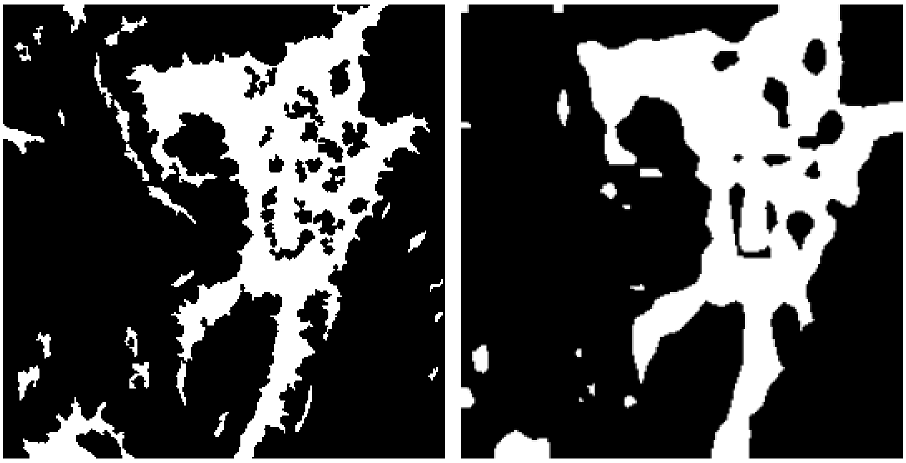
From Table 1 we can see that the dataset is imbalanced. The Tumor + Stroma Tissue class makes up around 50% of the training dataset and 75% of the validation set. Hence, a simple dummy classifier that always predicts class 3 can achieve 75% classification accuracy on the validation set.

2.2 CNN

After the dummy classifier, I used a CNN to see if we can outperform this strong dummy baseline on the validation dataset. I use a pretrained resnet18 [1], replace the final linear layer and finetune all parameters. I did not perform extensive hyperparameter tuning and just train for five epochs on the entire training set with an Adam optimizer and a learning rate of 1e-4. The model achieves its maximum validation accuracy of 75% after the fourth epoch. Figure 1 shows an example image and ground truth labels from the validation set. The trained classifier correctly predicts **S+T** for the displayed image.

The CNN classifier achieves the same accuracy as the dummy classifier so is it any better? I also computed the validation weighted F1-Scores for both classifiers and the dummy classifier has an F1-Score of 64.28%, where as the CNN model has 78.55% F1-Score. This is because the CNN model treats false positives and false negatives better than the dummy classifier.

[1] Deep residual learning for image recognition, He et al.



(a) Ground truth (b) Prediction

图 2: Binary mask prediction

The calculation of accuracy does not directly consider these. Therefore, we can conclude that the CNN is indeed better than our dummy.

3 Progress: Image Segmentation

I split the Image Segmentation tasks into two separate tasks.

- In 3.1, I try to differentiate the unlabeled background from the labeled foreground in a binary segmentation task.
- In 3.2, I try to differentiate the background and all classes in the foreground in a multi-class segmentation setup.

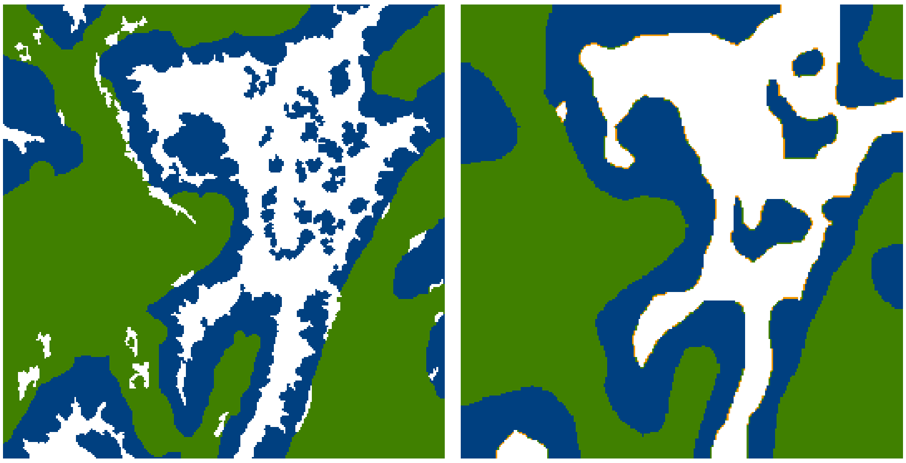
Since segmentation labels are only provided for the validation dataset, I split it into 38 training images and 2 validation images.

For both set-ups, I use a pre-trained DeepLab V3 model [2]. I adjust the output neurons in the final linear layer for each of the two cases and finetune the full model for 20 epochs. All models are trained with cross entropy loss and the Adam optimizer with a 1e-4 learning rate.

3.1 Binary Background Segmentation

For binary prediction, I set the output neurons in the final layer to 1. I add a sigmoid function to compute an output between 0 and 1 according

[2] Rethinking Atrous Convolution for Semantic Image Segmentation, Chen et al.



(a) Ground truth　　(b) Prediction

图 3: Mutli-class mask prediction

to:

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

If $\sigma > 0.5$, I define the output as foreground, else as background. Figure 2 shows a ground truth example from the validation set and the corresponding model prediction. The background segmentation model achieves a validation accuracy of 81.54% after 20 epochs.

3.2 Mutli-class Segmentation

For multi-class prediction, I set the output neurons to 4 and simply choose the neuron with the highest activation as the predicted class (background, normal, stroma, tumor).

Figure 3 shows a ground truth example from the validation set and the corresponding model prediction. The multi-class segmentation model achieves a validation accuracy of 74.64% after 20 epochs. It is expected that the accuracy is lower than that of the binary background segmentation. This is because, the model still needs to perform the same background segmentation like in section 3.1. In addition, it needs to further segment the foreground into 3 distinct classes (normal, stroma and tumor tissue). Hence the performance is capped at the accuracy of the background segmentation model.

4 Conclusion

In this report, I have presented my progress on the class project. I have finished all tasks outlined in the project introduction including the optional segmentation part. For the final report, I will add more example predictions from the data and model graphs to provide a more in-depth model explanation. Further the code will be provided in the final report.