Classification of Lung and Colon Cancer Histopathological Images

Sina Ghofrani Majelan Parastoo Sotoudeh Sharifi Farzaneh Ghasemi Javid

Abstract

Cancer is a disease when abnormal cells begin developing wildly. One of the most dangerous and deadly diseases worldwide is Colon and Lung cancer. This disease is hazardous because it can lead to the patient's death. The possibility of occurrence of colon and lung cancer at the same time is very low, but if the disease is not diagnosed and treated, there is the risk of metastasis between the colon and lung. So, according to the importance of classifying histopathological images of lung and colon, we have done comprehensive research in this field. In this project, we will use transfer learning to classify the histopathological images of the LC25000 dataset. We tried three different pretrained models as the feature descriptor, including NASNet Mobile, VGG16, and ResNet. Additionally, we adopt the SE Block as an attention mechanism in the classifier of the network. This block enables the network to select more discriminative features by weighting each channel adaptively. As for the evaluation of the research, we will use some standard metrics, including F1-score, Accuracy, Precision, and Recall. We also plotted the attention maps and feature maps of different network layers in the experimental results section of this report.
keywords: Histopathological Images, Transfer learning, Attention Mechanism, LC25000
https://github.com/ParastooSotoudeh/Lung-and-Colon-Histopathology-Images

1. Introduction

Cancer is a type of disease when abnormal cells begin developing wildly in the human body. This process can start in any organ of the body. Due to the high death rate resulting from this disease and the problems in diagnosing in traditional methods, machine learning models tried to help humans to speed up the decision-making process and increase accuracy. By using representative images to train a machine learning model, it will be able to recognize patterns from labeled photographs. In this study, an artificial intelligence-based approach that detects cancer types using a histopathological image dataset (LC25000) is proposed. This dataset consists of histopathological images of the lung and colon.

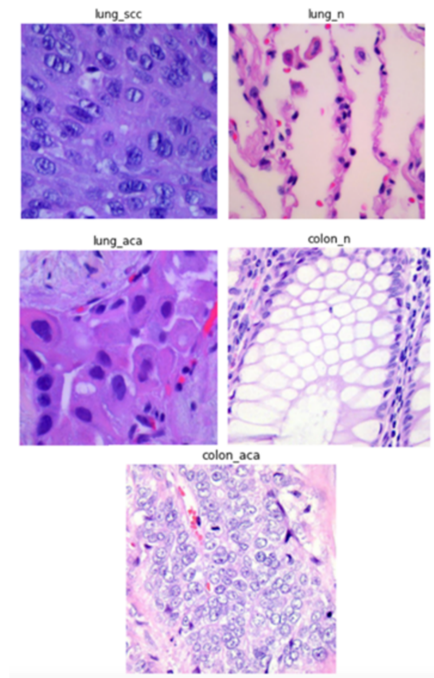

Figure 1. Sample images of the LC25000 Dataset.

The LC25000 dataset contains 25,000 colored jpeg images. 20% of the dataset was used as test data, and 80% was allocated as training data. Images are divided into five classes where each class has 5,000 images of the following histologic entities: colon adenocarcinoma, benign colonic tissue, lung adenocarcinoma, lung squamous cell carcinoma, and benign lung tissue (figure 1). The size of the images is set to 768 x 768 pixels. For the preprocessing, we scaled the images and normalized them (Figure 1).
Additionally, three different networks including VGG16, NASNetMobile, and Resnet were applied to the model separately for training and evaluation, and their performance was compared.

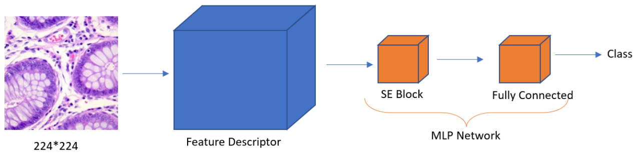

Figure 2. Methodology.

Architecture	Parameters
Resnet50	23,587,812
VGG16	14,714,688
NasNet Mobile	4,269,716

Table 1. Comparison of the number of parameters.

2. Methodology

We split the methodology into three major sections. In the first part, we will talk about data preparation. The second part will focus on the pre-trained networks used as the feature extractor. In the third part, we will talk about the MLP classifier at the end of the network.

2.1. Data Preparation

The only pre-processing steps applied to the Images before feeding them to the network are rescaling and resizing. The shape of the images of the LC25000 dataset is 768*768, and we resize the images to 224*224. We also rescale the pixel values to the range of zero to one.

2.2. Feature Descriptor (CNN)

In deep learning, a number of different CNN architectures are proposed, which differ in structure and parameters. Some of these architectures are trained on the ImageNet data set. In this research, we used three different architectures as the feature extractor, namely, NASNet Mobile, VGG, and ResNet. a comparison of the number of parameters for these architectures is presented in Table 1.

2.3. Multilayer Perceptron

In this section, a squeeze and excitation (SE) block, a GlobalAveragePooling, and multiple fully connected layers are employed. The first block in MLP design is the SE block which is used as the attention mechanism. Since the last convolutional layer in a CNN has highly class-specific information, we employed the SE block as a content-aware mechanism to weight each channel adaptively. SE block introduces a block for CNN that improves channel affinities at a less computational cost. The general idea in this block is to add parameters to each convolution block so that the network is capable of adjusting the weights for each feature map automatically. In short, it can be said that this block improves the interdependency of the channels by weighting the channels.

3. Experimental Results

3.1. Data Set Setting and Evaluation Metrics

We split the LC25000 data set into two parts. The first part, consisting of 80% of the total images, is used for training, while the second part, consisting of the remaining 20%, is used for validation and testing. In this research, we used different metrics to evaluate the methods, including F1-score, Accuracy, Precision, and Recall.
Moreover, in order to perform a comprehensive study and investigate the method, we plotted the feature maps of some layers of the network. Another interesting study in this reseach is plotting the attention map of the activated class by computing the gradient of the activated class with respect to the last convolutional layer of the feature descriptor. We plot the mentioned attention maps by adopting the Grad Cam method of the Keras Library.

3.2. Parameter setting

In this research, The Adam optimizer was used for training all of the networks. Since the used networks are deep, there is a high probability that the networks will overfit. So, we applied data shuffling in the training process in order to prevent overfitting. It is worth mentioning that the batch size is set to 128, and the learning rate is set at 0.0001.

3.3. Experimental Results and Comparisons

In this section, we will evaluate the proposed method. As mentioned earlier, we utilized three different pre-trained models as the feature descriptor. We have frozen these networks in the training process, and only the weights of the classifier were updated during the training. The results are presented in Table 2. According to the results proposed in table 2, among three feature descriptors, using the Nasnet-Mobile leads to achieving the best results. It is an interesting result because NasNet mobile has much fewer parameters compared to other networks.
Since the best results were obtained by the Nasnet-Mobile network, in the following, we will only report the results of this network. Figure 2 shows the graph of the accuracy and loss of the network during training. It is obvious that the network has been fully converged after 15 epochs. Figure 3 shows the confusion matrix of the network. According to the confusion matrix, Class2 (luna-aca) is the most difficult class for the network.
Figure 4 shows the attention map of the activated class (class 0) followed by the feature maps of different layers of the network for a sample input image. According to the attention map, it seems that the network pays attention to the purple points of the picture when the class 0 ishas activated.

Network	Class	Precision	Recall	F1-Score	Accuracy
NasNet	0	1.0	1.0	1.0	0.99
NasNet	1	1.0	1.0	1.0	0.99
NasNet	2	0.99	0.98	0.98	0.99
NasNet	3	1.0	1.0	1.0	0.99
NasNet	4	0.98	0.99	0.99	0.99
Resnet50	0	0.84	0.84	0.84	0.84
Resnet50	1	0.88	0.82	0.85	0.84
Resnet50	2	0.79	0.74	0.76	0.84
Resnet50	3	0.93	0.94	0.93	0.84
Resnet50	4	0.79	0.87	0.83	0.84
VGG16	0	1.0	0.99	0.99	0.99
VGG16	1	1.0	1.00	1.0	0.99
VGG16	2	0.97	0.98	0.98	0.99
VGG16	3	1.0	1.0	1.0	0.99
VGG16	4	0.98	0.98	0.98	0.99

Table 2. The results of training different models.

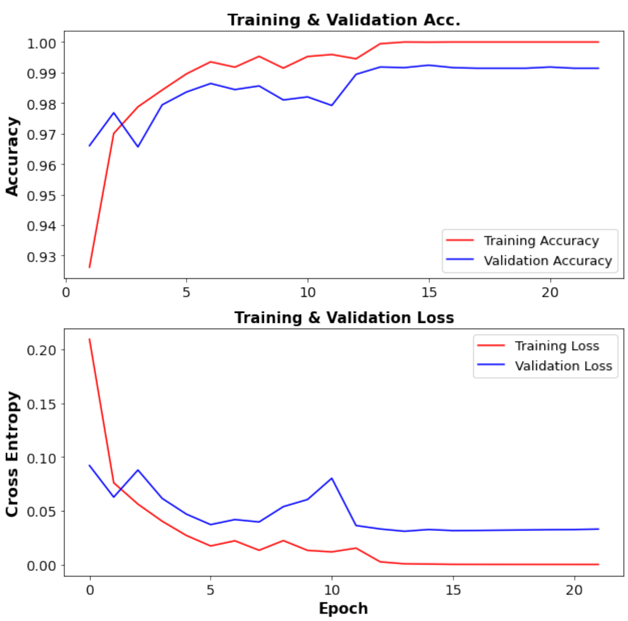

Figure 3. graph of the loss and the accuracy.

4. Conclusions

In this research, we used three different pre-trained networks as the feature descriptor for classifying the histopathological images of the LC25000 dataset. We employed the SE block as an attention mechanism in the designed MLP network, which is utilized at the end of the pre-trained models. Among all the pre-trained models, the Nasnet-Mobile achieved the best results (training accuracy:1.0 and validation acuracy:0.99). It is worth mentioning that the NasNet-Mobile has much fewer parameters compared to other networks we employed in this study. We also performed different studies on the Nasnet-Mobile network to show the attention maps and feature maps of different layers of the network.

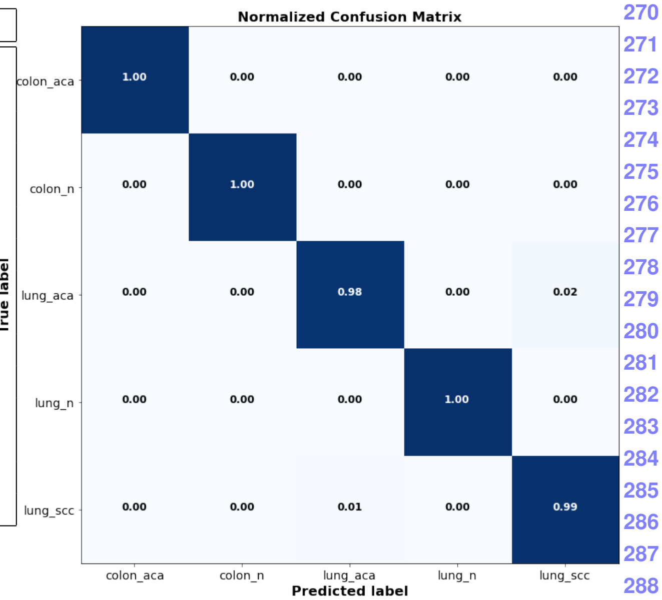

Figure 4. Confusion Matrix.

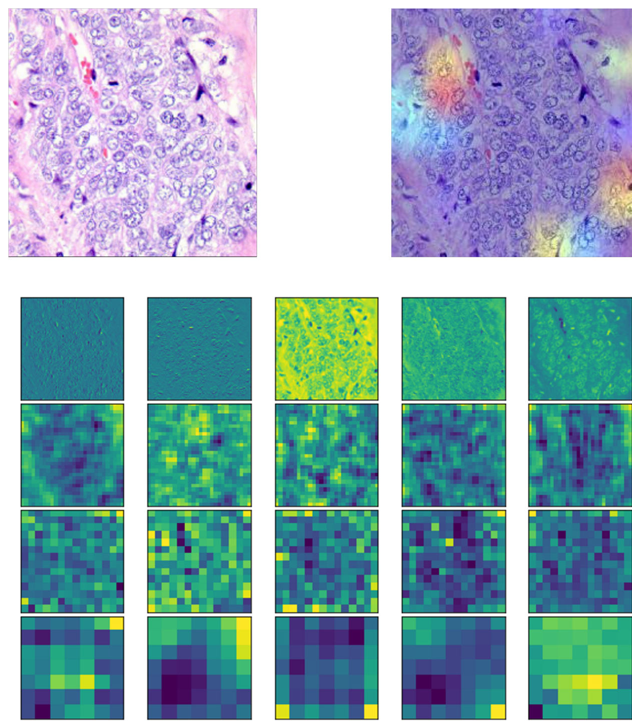

Figure 5. Attention map of the activated class and feature maps of different layers of the network.