Digestive System Pathological Image Classification Using Dual Stage CNN

Gonella S Uma Maheswara Rao, Mishra Akash Brijeshchandra, Mohd Shoyab, Gautam Kumar

Dr. Sridhar Swaminathan

***Abstract*—This paper consists of the Deep Learning solution of the real-world medical problem of colorectal cancer using dual stage CNN Colonoscopy pathology examination can find type of colon tumor from small tissue slices. Pathologists need to daily examine hundreds of tissue slices, which is a time consuming and exhausting work. Here we propose a challenge task aiming at automatic classification of the whole tissue (benign vs. malignant). As the medical images are heavy to handle and need more effort to process, we follow the patch wise approach where we process the image of particular class separately as patches to extract the features and later we load the whole images for classification based on the patch wise model. We took the challenge from Digest Path 2019 Challenge.**

***Index Terms*—Pathological Images, Colorectal Cancer, Dual Stage CNN, Patch wise Learning, Colonoscopy Tissue Classification.**

I. Introduction

Colorectal cancer is the #3 diagnosed cancer in the US and the #2 cause of cancer related death in men and women. Many digestive cancers are found in later stages, but colorectal cancer can be found and treated in earlier stages or even prevented due to screening exams. The American Cancer Society reports that even though colorectal cancer is found more often in people ages 50 and older, it’s on the rise in people under age 50. A study from 1974-2013 found a 1% increase per year in the colon cancer incidence of people 20-49; and an increase per year in the new cases of rectal cancer in adults 20-39 (3%) and 40-54 (2%).So, by this we can get an idea that early diagnosis of this leads to prevention of the disease. Hence the usage of deep learning techniques in the field of medical diagnosis can save the lives of so many hopeless people. In rural areas there will be no facility of radiologist availability not only in rural areas but also in some well-established towns so in such places deploying our model can increase the medical diagnosis. Although a radiologist is available, analyzing the scans of plenty of images is a hectic task so this model can make the work of an radiologist easy.

II. Related Work

Computer Aided Diagnosis (CAD) is burning topic so using deep learning techniques are already followed on different diseases. So many models are built for liver tumor segmentation, kidney tumor segmentation etc. and particularly for the topic of digestive system colonoscopy tissue segmentation this is the first ever work since the Digest Path 2019 challenge is the first ever public challenge on the Digestive system images. This model was not built from scratch rather it is made by adopting the architecture of the Imaging lab ICIAR 2018 breast classification model.

Generally, all the types of classification involve classification into malignant/benign or in some types of cancers there may be some more classes, so all the classification will be of similar kind but differ in the type of cancer and data as some type of cancer data need to processed in different style but the end task is of same type.Our work is like [1] where they are classifying the cancer into four types, but we need to classify to two classes and some different type of data handling.

III. Methodology

Given an image with high resolution of around 20000 x 30000 pixels we need to classify the image into Malignant, Benign as the given image is the MRI scan of Digestive System of a person.

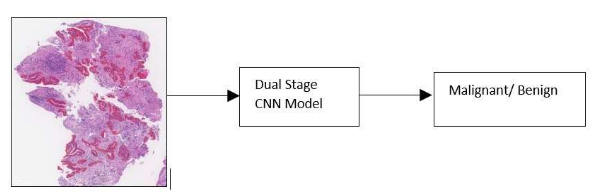

Figure 1 . Problem statement

Patch wise CNN:

We are given with the high-resolution image of 20000 x 30000 for such kind of images traditional methods will be computationally intensive and does not yield more accuracy so we adapt a different approach to deal with the high-resolution images which also help in the process of extracting the features. What exactly the patch wise model is doing is dividing the images into patches of fixed size. Now we will train the model patch wise to extract the features and we will do it by loading the images of classes separately so that the features will be extracted according to the category. So, by training the model patch wise we are getting two advantages one is we are not losing the detailed features of the image and also, we can easily feed our data into the model as no complex model for image is required in our case it's just only about the type of data processing we adopted. Patch wise CNN is first stage of our training process in the later stages it is followed by an image wise model which classifies the images into two classes. Given the pathological image we divide into patches of fixed size by sliding a window of size k x k and with stride s on an image. We choose the value of patch and stride according to the memory available in GPU. Generally we try to take the overlapping patches to not to lose the detailed information in the image. The stride and patch size are kind of hyperparameters as they decide the size of minimum sample to be analyzed and the amount of information we need to include from the sample. As an outline patch wise model is kind of taking a sample from a microscopy image and analyzing it in the direction of our problem. One more important point to note is that in the patch wise model we don't have any discussion based on class, the classes are only for the whole images. We train the patch wise network based on the label of the image which it is made into patches and use the categorical cross entropy loss based on the label of the image. In the Architecture of patch wise CNN, we are removing the classification layer and just doing 1 x 1 convolution to feed this output to input of the image wise CNN. The size of the feature maps extracted are based on the stride and the patch size.

Image wise CNN:

Image wise CNN is used to classify the images based on the local information from the patch wise CNN and the global information shared between different patches. In image wise CNN we also extract patches but not like patch wise CNN we take the stride so that the patches will not overlap. If we take the overlapping patches there may be chance of redundant features in the image wise CNN. The patches we got from the image wise CNN will be concatenated with the feature maps we got from the patch wise CNN and this will be fed as input to the image wise CNN. The image wise network will be trained against image labels based on the categorical cross entropy loss. Once the both models get trained, we then use both to predict the class label of an image.

Network Architecture:

Network architecture is taken from [2] the same network architecture was used without change as it works good in this kind of problem too. In patch wise network we have 3 x 3 convolutional layers in series followed by pooling layer where the number of channels gets doubled after each down sampling. All convolutional layers are followed by batch normalization and ReLU non-linearity. In the place of conventional max pooling we use 2 x 2 kernel with stride 2 as it works better than the max pooling.

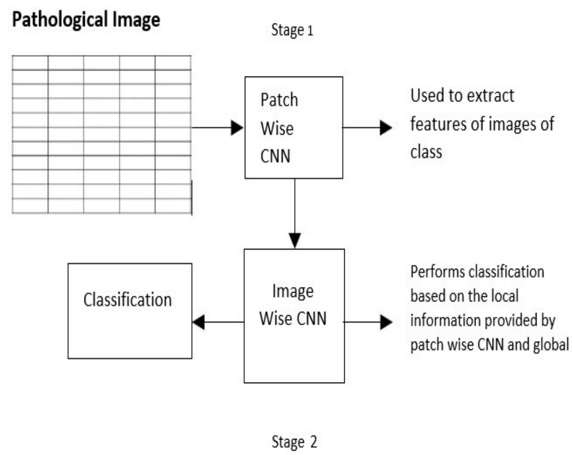

Figure 2. Dual Stage CNN

As discussed earlier at the end of the patch wise model we do 1 x 1 convolution to extract the feature maps that need to be concatenated with the non-overlapping patches generated from the image wise network and fed into the architecture of image wise network. The size of feature maps extracted will be dependent on the patch size and stride we have taken. Batch Normalization and global average pooling are structural regularizes there will be no need of any additional regularization like drop out etc. the patch wise model architecture is as follows.

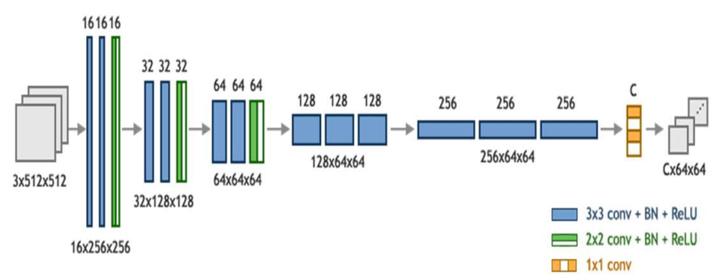

Figure 3. Patch Wise Network Architecture

In the image wise network, we follow somewhat similar kind of architecture we followed in the patch wise network. There will be series of 3 x 3 convolutional layers followed by 2 x 2 convolutional layers with stride of 2 for down sampling. Each layer is followed with batch normalization and ReLU activation units. There will be a 1 x 1 convolution at the end to extract the feature maps and also a SoftMax layer at the end to perform classification. But in the image wise network we can have a problem of overfitting, so we use drop out regularization at a rate of 0.5 and we also use early stopping to prevent the overfitting. As discussed, the patch size and stride are the hyper parameters so from experimental results it is suggestable to use patch size = 512 and patch stride = 256.

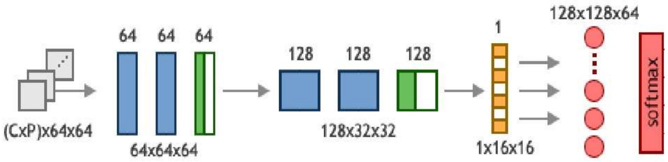

Figure 4. Image wise model Architecture

IV. Experimental Results

This challenge is taken from Digest Path Challenge 2019 and the dataset is taken from there Digestpath. The dataset consists of the scan images of the digestive system of different patients in jpg format separated by classes. Data Augmentation is based on the size of dataset if the dataset contains ample number of images no need of data augmentation anyway data augmentation helps in better training of the model and helps in some learning some invariant features.

Model	Patchwise Model			Imagewise Model	
Methods	Sum	Max	Maj	CNN	Ensemble
Model's Accuracy	70.4	71.2.	69.72	72.61	75.3

Figure 5. Experimental Results

We have experimented the problem with different kind of models, but this model gives us the best possible results by training the model for 30 epochs. We have used Adam optimizer and batch size can be 64 initially and the learning rate was initially 0.001 with a decay of 0.1 for every 20 epochs. The validation can be of your choice but should not be more than 30% since the model does not train on more data can leads to decrease in accuracy.

We need to train the model simultaneously patch wise and image wise for the classification in the image wise network for classification we have three methods to apply on the data we get from the patch wise model to decide the class label they are majority vote, maximum probability, sum of probabilities. It is better to experiment with all three methods that which is giving the best accuracy.

As a part of experimenting we can use CNN with SVM or some ensemble models for the classification which gives more accuracy than just using the Image wise CNN for classification. Ensemble models help in classifying from the features extracted from the patch wise CNN and from the global information shared between different patches. It can be further used to improve the performance of the model by including it in a better way possible.

V. Conclusion & Future work

In this Paper we have presented a model which classifies the images of digestive system cancer and as the image size is more to decrease the computational complexity, we have adopted this model and the image gets trained patch wise and image wise as discussed above we get the label of classification as malignant or benign.

In future work we can try to improve the model to detect the infected part automatically and classifies it as a malignant or benign. In the model we can experiment with different kind of possibilities to further increase the accuracy. The actual task was somewhat different from what we have done that in the challenge it requires the part of image where the colonoscopy tissue was and to segment it and later classify it as a malignant or benign tumor. So we can make the model in such way taking this idea as base.

As a part of future work, we can also increase the performance of our model by adapting any better classification model or an ensemble model along with the image wise CNN we are using which makes the task of classification much more efficient. As of now in our model we are just using image wise CNN for classification without any classifier so this can be improved by introducing a suitable ensemble model or a good classifier

References

[1] Nazari, K., Aminpour, A., & Ebrahimi, M. (2018, June). Two-stage convolutional neural network for breast cancer histology image classification. In *International Conference Image Analysis and Recognition* (pp. 717-726). Springer, Cham.

[2Two-Stage Convolutional Neural Network for Breast Cancer Histology Image Classification. ICIAR 2018 Grand Challenge on Breast Cancer Histology images (BACH) :

[3] Brook, A., El-Yaniv, R., Isler, E., Kimmel, R., Meir, R., Peleg, D.: Breast cancer diagnosis from biopsy images using generic features and sums. IEEE Transactions on Information Technology in Biomedicine (2006)

[4] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

[5][Spring Enberg, J.T., Rosovsky, A., Brox, T., Riedmiller, M.: Striving for simplicity: The all convolutional net. arXiv preprint arXiv:1412.6806 (2014)

[6] Srivastava, N., Hinton, G.E., Izhevsk, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. Journal of machine learning research 15(1) (2014) 1929–1958