A2 Final Report

Saneyar Khazin, Simon Bateman

Introduction

Histopathology (or histopathology) refers to the study and diagnosis diseases related to the tissues and involves observing tissues under a microscope (Royal college of Pathologists).

Histopathologists, doctors involved in the study and observation of tissues, provide a diagnostic service for cancer; they handle, and study tissues and cells removed from suspicious lumps, identify the nature of the abnormality, and categorize it as either malignant or benign. If the diagnosed as malignant, they provide information to about the type of cancer, its grade and, for some cancers, its responsiveness to certain treatments (Royal college of Pathologists).

From this, we can come to understand how critical this process can be. However, detection and classification of cell nuclei in histopathology images of cancerous tissue is a challenging task due to cellular heterogeneity. Deep learning approaches have shown some promising results on histopathology images in various studies.

In this project, we attempt to create two separate machine learning model based on Deep Neural Networks to aid in predicting i) whether based on the given images, a cell is cancerous or not and ii) the cell type of the provided image.

VGG Model

The original VGG model was an attempt at improving the original architecture of highly popular ImageNet. Some of the major changes include fixing other parameters of the architecture and steadily increasing the depth of the network by adding more convolutional layers, which is made possible by using small (3 x 3) convolution filters in all layers. A VGG model is structured to consist of many blocks, with a block consisting of: convolution + activation + pooling. These block will also downscale the images due to the use of the max pooling layers.

The models we have used in this assignment are a modified version of the original VGG Model. While in the original paper, the images sizes used were 224 x 244, our images were a lot smaller at 27 x 27. This meant we would have to make a couple of adjustments to get our model running well.

As per Simoyan *et al.*, only a small amount of preprocessing is done, which includes substracting the mean RGB value, computed on the training set. The image is then passed through a stack of convolutional layers with small receptive field: 3 x 3. Convolution striding is fixed to 1 pixel for 3 x 3 convolutional layers, which translates to `padding = same`.

Since the we were dealing with two different types of classifcation models here (cancerous or not versus cell type), compile parameters were different in both the cases. Spatial pooling was carried out by 5 max-pooling layers in the original paper. But since the images we were supplied were much smaller, we used only 2-3 layers.

Simonyan *et al.* carried out the training by optimizing multinomial logistic regression objective using mini-batch gradient descent. Regularization was done through weight decay, i.e using L_2 regularization as well as

dropout regularizationfor the first 2 fully connected layers, with a dropout ratio of 0.5. The learning rate was set to 10^{-2} initially and was set to decrease by a factor of 10 when the validation set accuracy stopped improving. Overall, the learning rate decreased 3 times and was stopped after 74 epochs.

Since we were dealing with two seperate problems here, our development process was slightly different than the original paper. For the binary classification, we used `sigmoid` activation function for the final layer since we were dealing with binary values and `relu` was used for classifying the cell type. The loss function that was used were also different in both cases, with `Adagrad` being used for cancer classification and `sgd` for cell type. The trajectory for model development were similar in all other aspects, save a few steps.

The cell type classification problem had frustrating problem, in that the accuracy and loss values were volatile and jumping all over the place. Multiple attempts were made to solve, ranging from regularization to a time based decay. this but to no avail. This issue was resolved to some extent by setting the learning rate at 0.001 instead of the 0.01, but it didn't help solve the issue completely. Another problem we came across for cell type classification was that the VGG models had an issue of overfitting. This was resolved by using data augmentation but at the cost of a tiny drop in overall performance.

The cancerous or not classification problem seemed to do really well with just the baseline VGG model. In fact, adding regularization seemed only to drop the overall performance. The problem we faced here was that baseline model with 3 convolutional layers took a lot of time to train. Therefore, we tried reduce the overall time by reducing the convolutions along with some regularization. This did indeed help cut down the time, but also reduced recall to much lower than the baseline model, i.e the base model with only one layer.

Overall, in their paper, Simoyan *et al.* demonstrated that their model generalized well to a wide range tasks and datasets, matching and at times outperforming more some of the more complex recognition models. This was evident n the case of classifying images on to find out if the patient had cancer or not. Even among all the other models that were implemented, the VGG architecture was the best performing for cell type classification.

Spatially Constrained CNN

Sirinukunwattana *et al.* uses a locality sensitive deep approach of a Neighbouring Ensemble Predictor (NEP) coupled with CNN to more accurately predict the class labels of detected cell nuclei. This process used 2 premises: the distance from the centre, and a weighted ensemble of local predictions. This process allocates higher probability to pixels closed to the centre of the nucleus that those further away. This enables the model to detect and assign class labels to multiple nuclei. Prior to the research done by Sirinukunwattana *et al.* , most of the cell detection methods rely on shape of nuclei as well as stability of the features for cell detection and texture of the nuclei for classification. The methodology used by Sirinukunwattana *et al.* and Xie *et al.* are closely related, in which they use an SC-CNN (Spatially Constrained Convolution Neural Network). SC-CNN contains a layer that is designed to predict the centroid location of the nuclei as well as the confidence of the locations.

This used an overall image of the data allowing the neural network training to allocate priority to specific pixel in the photo. This approach did not consider the impact of varying locations of cell nuclei.

Other papers used a combination of Difference of Gaussian followed by other processing to detect nuclei, Bi and multilateral symmetric to detect nuclei. But these failed to detect thin fibroblasts and small nuclei as they replied on nuclei shaped and texture for detection. Using a basic VGG architecture did not provide greater accuracy, as was demonstrated by reducing the number of convolutions before passing the results onto a Neural Network layer.

Their approach potentially did not require the preparation of slides so the nucleus was position near the centre of the slide thus offering a more generic model. We chose simple NLP network as the size of the images (27 x27 x 3) was small enough the total image could be processed as an input later with as this was simpler and

produce good results on the data set. However, it wasn't possible to replicate this model in its entirety as the authors of the model used algorithms that weren't made public for **PARAMETER ESTIMATION (S1)** AND **SPATIALLY CONSTRAINED (S2)** layers. Without these layers, the model was not usable. Therefore, we were left with the Neural Network architecture which was outlined by Sirinukunwattana *et al.*. But this did not produce any favorable results in both cases.

Score and Evaluation

Identify whether patient has cancer or not

Model	Accuracy	F1 score	Precision	Recall
Baseline Model	0.87	0.85	0.89	0.71
Baseline + l2 regularization	0.85	0.83	0.90	0.73
Baseline + dropout	0.87	0.85	0.88	0.73
Baseline VGG	0.90	0.89	0.90	0.81
VGG + regularization	0.87	0.85	0.91	0.70
SC-CNN	0.65	0.39	0.00	0.00
SC-CNN + dropout	0.65	0.39	0.00	0.00

Identify the cell type

Model	Accuracy	F1 score	Precision	Recall
Baseline Model	0.69	0.59	0.56	0.88
Baseline + l2 regularization	0.20	0.08	0.00	0.00
Baseline VGG	0.70	0.63	0.75	0.60
VGG + regularization	0.71	0.66	0.67	0.71
VGG + data augmentation	0.70	0.60	0.65	0.72
SC-CNN	0.40	0.14	0.00	0.00

Final thoughts

One of the more surprising things we found out was that a base neural network with minimal layers and zero regularization seemed to perform quite well with the binary classification, while giving descent performance on the multiclass problem.

While we were able to replicate the CNN model outlined by Sirinukunwattana *et al.*, we weren't able to replicate the exact algorithm used by them to classify the images. Therefore, we based our model on the VGG model. The initial baseline VGG model was giving good accuracy as well as recall for the cancer classification problem, and decent results for the cell type classification problem. But this model proved to be problematic, in that with the cancer classification problem, it was very time intensive, and for the cell type classification, there seemed to be a problem of overfitting.

In the end, we settled with VGG as the best performing architecture here. The rationale behind this is that model was said to be an improvement the popular ImageNet model. Simoyan *et al.* states at the end their paper that the model ended up matching, and at times exceeding performance compared to other similar models. This was evident with the first part of the assignment for classifying cancer cells. It was also the best performing model for cell type classification as well, though the accuracy and F1 score were lower than the first part.

References

Sirinukunwattana, K., Raza, S., Tsang, Y., Snead, D., Cree, I. and Rajpoot, N., 2016. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Transactions on Medical Imaging, 35(5), pp.1196-1206.

Simonyan, K. and Zisserman, A., 2014. Very Deep Convolutional Networks for Large Scale Image Recognition. arXiv e-prints, pp.arXiv:1409, 1556.

Xie, Y., Xing, F., Kong, X., Su, H. and Yang, L., 2015. Beyond classification: Structured regression for robust cell detection using convolutional neural network. Medical Image Computing and Computer-Assisted Intervention—MICCAI, pp.358–365.