CSL 601 Final Project: Cervical Cancer Screening

Authors

Jatin Goyal
Indian Institute of Technology Ropar
2015csb1014@iitrpr.ac.in

Unit Patel
Indian Institute of Technology Ropar
2015csb1038@iitrpr.ac.in

Abstract

The type of a patient's cervix determines the type of pre-cancer treatments the patient can undergo, and the medical community would benefit from an efficient method to classify a patient by their cervix type. Kaggle and Mobile ODT have published a collection of several thousand specular photographs of cervixes, each labeled as one of three types. We present our work in developing a convolutional neural network (CNN) to classify the cervix images in this dataset. We constructed and trained our Keras CNN model. We discuss the performance of our model. Our most successful model achieved an average classification accuracy of 53.73%.

Introduction

The earlier the signs of cervical cancer are detected, the easier the treatment path will be for the patient. This treatment path varies for women based on the physiological differences in their cervix. Rural or understaffed clinics would benefit from a way of quickly and accurately classifying patients based on their cervix types. Cervical cancer tends to begin in cells within the transformation zone, which could be completely ectocervical and visible (Type 1), partially endocervical but visible (Type 2), or partially endocervical and not fully visible (Type 3) (see Figure 0). Cervix types 2 and 3 may require different screening or treatment due to the placement and hidden view of precancerous lesions.

The input to our model is a photograph of the cervix taken through a vaginal speculum. The output is the probability distribution over the three classes, from which we extract the most likely class. Quantitatively, the goal is to
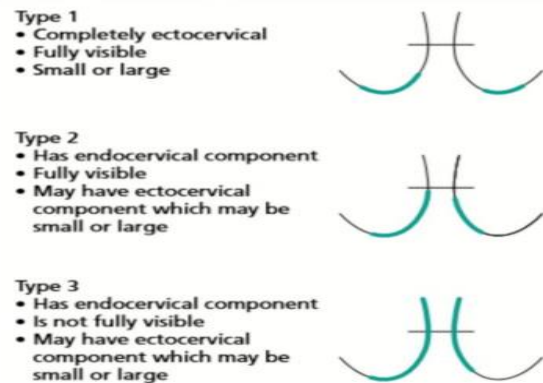
minimize the cross-entropy loss J for the classification:

$$J = -1/N * \sum\sum y(ij) * \log(pij)$$

where i varies from 1 to N, where N is the number of testing data points, i varies from 1 to C, C is the number of classes (3, in this case), y is the one-hot vector for the correct class, and p(ij) is the predicted probability that data point i has class j. Our project is to use deep learning to automate and improve this important classification process.

This project was inspired by a public Kaggle competition, and the dataset is provided on Kaggle's website.

Related Work

A huge amount of research has been conducted in the field of Deep Learning and Image processing for the classification or segmentation of the medical images. This area is still an active area in the field of research. One recent attempt at cervical cancer classification combined image features from the last fully connected layer of pre-trained AlexNet with biological features extracted from a Pap smear to make the prediction [1]. Deep learning has also

been used for other types of cancer detection. A convolutional neural network (CNN) following OxfordNet's structure was used to detect mammographic lesions[2]. Apart from neural networks some research are also carried into the field of image segmentation as this process helped in decreasing the classification error and thereby increasing the accuracy. Automated cervix and cervical cell segmentation is another important area of study. One method takes care to remove glare from the photo and uses K-nearest neighbors (KNN) with images pre-segmented by a distance metric based off of the histogram of oriented gradients to locate the most similar bounding boxes and averages them [3]. A model by researchers at Medical College of Georgia also used glare removal, K-means clustering, and texture features to segment the different cell types around the cervix [4]. A similar method fed color and cell area features into K-means to segment the cervix [5]. Another group performed cervix segmentation by first transforming the image from RGB to luminosity, red-green chromaticism, and blue-yellow chromaticism, and then ran K-means and selected the largest region [6]. One group found that using a CNN to segment cervical cell cytoplasm and nuclei outperformed traditional filters and classification methods, especially when multiple cells were in the picture[7].

Dataset

Kaggle provided the dataset of the images consisting three types of cervical images. The distribution of the images is shown in the Figure1.
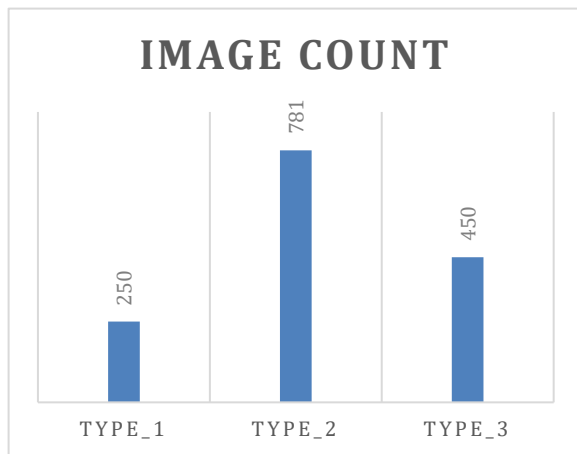

Figure1. Distribution of training dataset over classes

Our training set contained total of 1481 images. The training data was split into the training and validation data in the ratio 0.7:0.3. (**Note**: *Validation and test datasets are in fact same for our project as the true labels for actual test dataset were not provided publically by kaggle*) All the

images are coloured and have different size. We also tried to check the heterogeneity of data by projecting the data onto the first two principal components, but the data was very homogenous so we could not apply clustering as shown in Figure2.
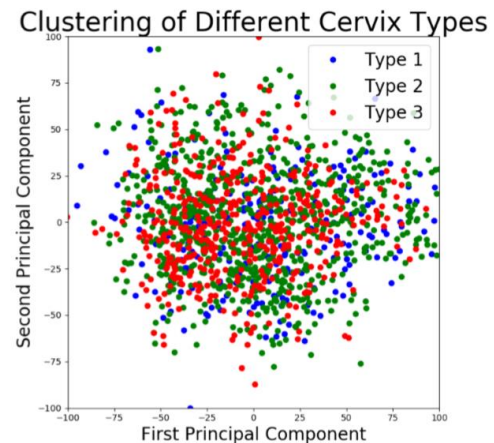

Figure2. Distribution of training data over first two Principal components.

Methodology

This section contains the architectural details of our best trained CNN model which is obtained after tuning several parameters like number of convolutional layers, number of epochs, dropout rate, batch size, etc. which output probability of image corresponding to each class.

Preprocessing

All the image features were firstly normalized. As all the images were of non-uniform size, they were scaled down to a uniform size 32x32. The new data was stored in the numpy files for reusability and time saving.

Model Architecture

After rescaling of the input image it was convolved with three convolutional blocks with following properties:
c1 – 32 convolutional layer, size 8x8, stride value 4, Activation function RELU, and Dropout Layer with p = 0.2.
c2 – 32 convolutional layer size 4x4, stride value 4, Activation function RELU, and Dropout Layer with p = 0.2.
c3 – 32 convolutional layer size 3x3, stride value 4, Activation function RELU, and Dropout Layer with p = 0.2.

And the last two layers were fully connected layers with L2- regularization to avoid overfitting which was attained by the previous convolutional layers(Figure3). Also as the model over fitted in our previous trials, we introduced the

dropout rate after every layer. The three convolutional layers were used to extract more and more features from the images. Keras CNN framework with the tensor flow in backend was used to implement our model.
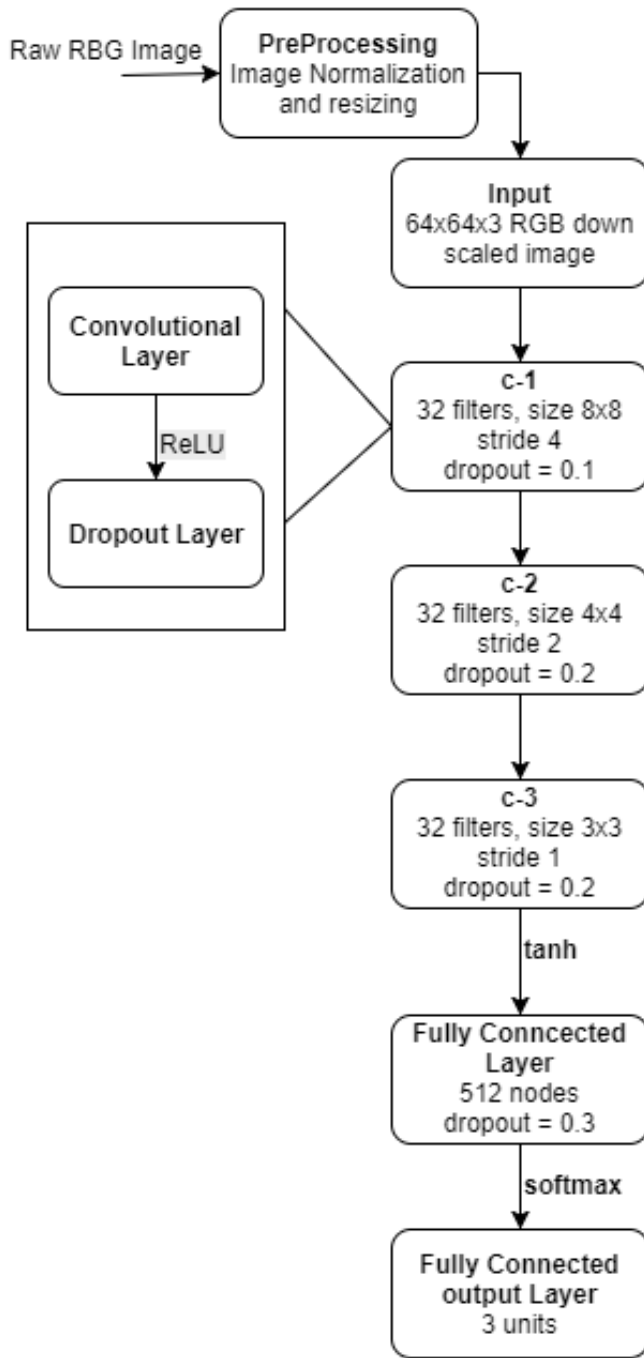

Figure 3: Details of model architecture

The requirement of fully connected component arose because we need to use collected data needed for classification. In other words, the convolutional layers have extract-

ed the meaningful features from the image. Now we need a system that uses these meaningful features to classify dataset. And the weights were initialized randomly.

Results and Discussion

The experiments were conducted to measure the overfitting of the model. In this experiment we plotted the graph of training and the test accuracy and the. Firstly observed that that after certain number of epochs training accuracy rose significantly while the test accuracy dipped and also the
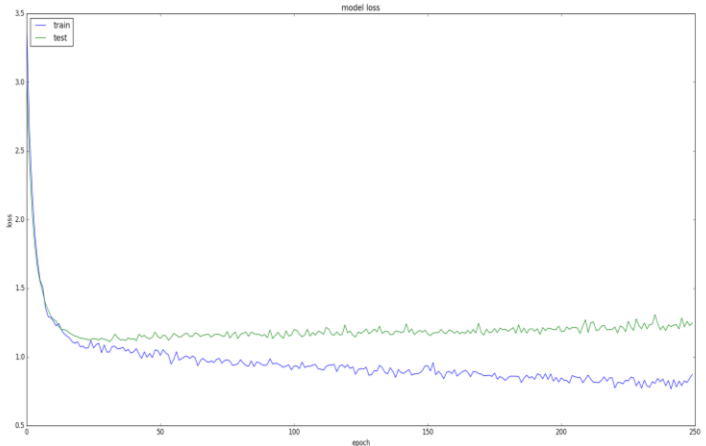
difference between the training and the validation loss increased. Hence there was clear overfitting in the model as shown in figure4 and figure5.

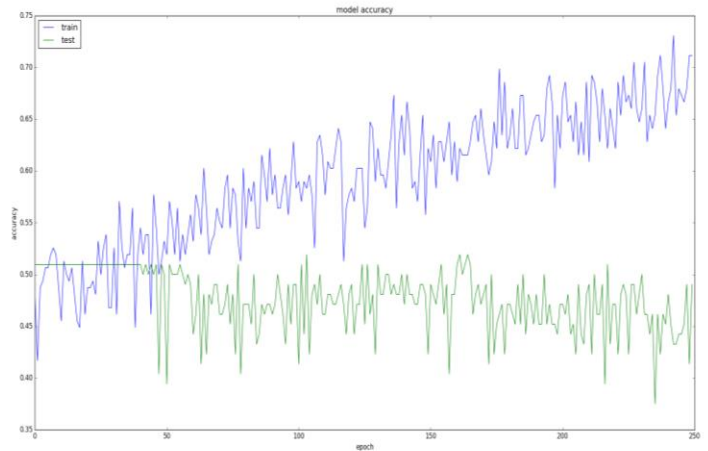Figure 4: Plot between the number of epochs and the training and validation accuracy.

Figure 5: Plot between number of epochs and the training and validation error

To counter this situations we decreased the number of epochs and introduced the dropout layer in the c-1convolutional block to possibly minimize the unnecessary

feature extraction that might be present only in the training dataset. Further to remove the possible overfitting while classifying the dataset using the extracted features we introduced a dropout rate in the fully connected layer. We also implemented L2 regularization in the final two layers. Hence after making these assumptions and implementing them we found ourselves successful and the result were as expected as shown in figure 6 and 7.
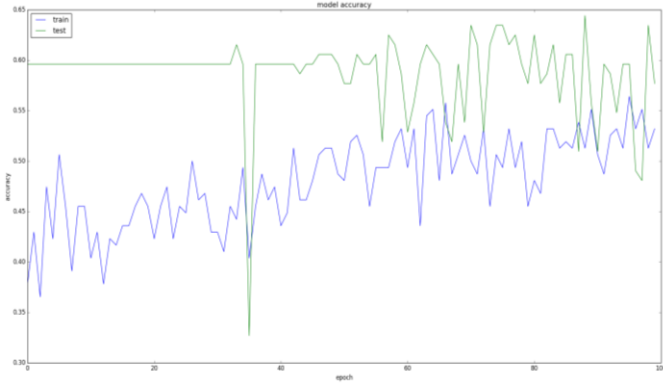

Figure 6: Plot between the number of epochs and the training and validation accuracy.
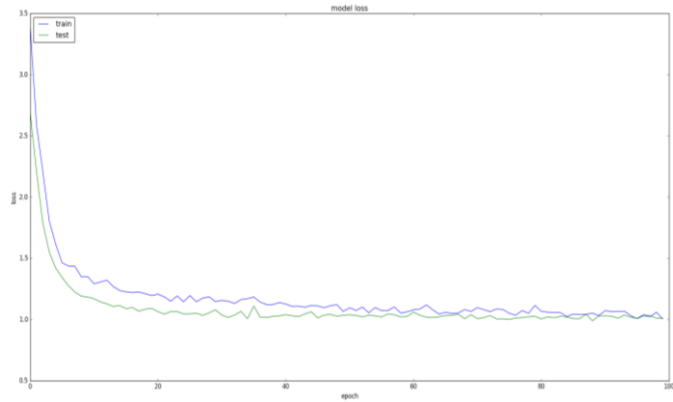

Figure 5: Plot between number of epochs and the training and validation error

Summary

The best validation accuracy achieved by our model was 57.68% and the average validation accuracy obtained was 53.73%. The label to the test data images was given the label of the most likelihood class predicted by our model. With limited resources, the accuracy obtained by our model is pretty decent. Some of the trial runs of our model are shown in the table

Exp. No.	Drop-out (Layer1)	Drop-out (Layer2)	Convolutional Layers	L2 Regularizers	Loss
1	0.0	0.0	2	0.0	1.347434
2	0.1	0.1	2	0.0	1.297843
3	0.1	0.1	2	0.005	1.263782
4	0.0	0.1	3	0.005	1.301283
5	0.1	0.2	3	0.005	1.210935
6	0.1	0.1	3	0.01	1.183974
7	0.1	0.2	3	0.01	1.119738

Table: Trial runs with varying parameters

Conclusions and Future Work

The main challenging part of the work was to increase the accuracy, which was approximately 50% in our first attempt. Even a small increase of 2-3% was obtained after applying various architectures with varying convolutional layer and the number of fully connected nodes. The reason to this behavior of the model can be explained from the fact that the data we obtained was not perfectly classified because few images had their duplicate with different labels. Also there were some bad images which were totally green, which were useless for our model as the skin segment of images was not clearly identifiable. Moreover, we also noticed that even when applying the same models again and again on the dataset there was a high variance in performance. The argument to this may be derived from the dependencies of weight with their initializations. Additionally, we didn't have enough background in image processing to really take advantage of the data we had. If we had more time we would focus on the segmenting the cervical part of images and then use them to train our model, instead we focused more on the architectural tuning. We would like to do segmenting and labeling regions of anatomical interest within the cervigrams. In particular, we would focus on the extraction of the cervix region and fine detection of the cervix boundary[9,10]. This would be a great kick start to our preprocessing step. We learned a lot from the project about deep neural network and data analysis. This motivates us to solve future problems in the real world with the help of what we have learned in the class.

References

[1] Xu, Tao, et al. "Multimodal Deep Learning for Cervical Dysplasia Diagnosis." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer International Publishing, 2016.

[2] Kooi, Thijs, et al. "Large scale deep learning for computer aided detection of mammographic lesions." Medical image analysis 35 (2017): 303-312.

[3] Song, Dezhao, et al. "Multimodal Entity Coreference for Cervical Dysplasia Diagnosis." IEEE transactions on medical imaging 34.1 (2015): 229-245.

[4] Li, Wenjing, et al. "Automated image analysis of uterine cervical images." Medical Imaging. International Society for Optics and Photonics, 2007.

[5] Srinivasan, Yeshwanth, et al. "A probabilistic approach to segmentation and classification of neoplasia in uterine cervix images using color and geometric features." Medical Imaging. International Society for Optics and Photonics, 2005.

[6] Das, Abhishek, Avijit Kar, and Debasis Bhattacharyya. "Elimination of specular reflection and identification of ROI: The first step in automated detection of Cervical Cancer using Digital Colposcopy." Imaging Systems and Techniques (IST), 2011 IEEE International Conference on. IEEE, 2011.

[7] Song, Youyi, et al. "A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei." Engineering in Medicine and Biology Society (EMBC), 2014 36th annual international conference of the IEEE. IEEE, 2014.

[8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. "Rethinking the inception architecture for computer vision." arXiv preprint arXiv:1512.00567. (2015)

[9] ImagesHayit Greenspan*, Shiri Gordon, Gali Zimmerman, Shelly Lotenberg, Jose Jeronimo, Sameer Antani, andRodney "Long Automatic Detection of Anatomical Landmarks in Uterine Cervix Images". IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 28, NO. 3, MARCH 2009

[10] User:Chattob https://www.kaggle.com/chattob/cervix-segmentation-gmm

[11]CompetitionSite https://www.kaggle.com/chattob/cervix-segmentation-gmm