



Segmentation and Classification of Prostate Glands into Benign, Grade 3, 4 or 5 Tumors

Ashima Jain, Rahul Ghosh, Amit Sethi

Department of Electronics and Electrical Engineering

Objective

Given the biopsy slides of the prostate glands, segment the glands and classify them into benign, grade 3, 4 or 5 tumors using Convolutional Neural Network.

Introduction

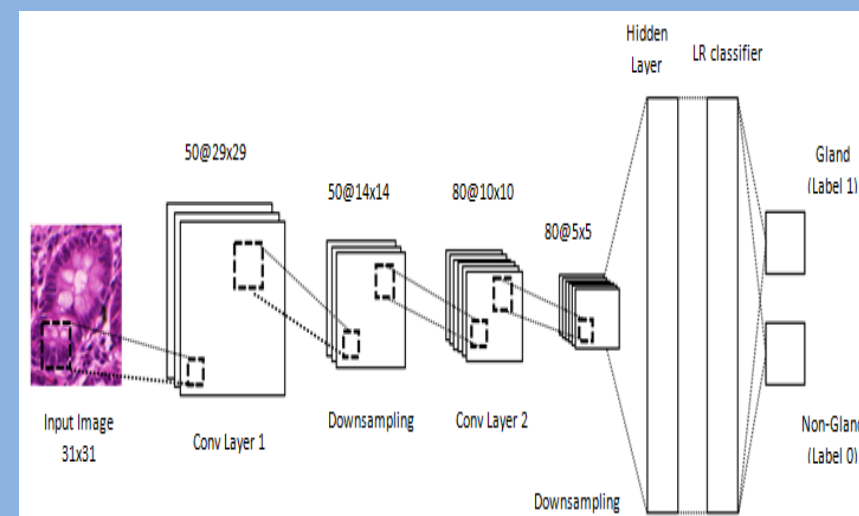
Prostate cancer caused by the growth of cancer cells in the prostate gland of the men reproductive system is now the second most prevalent cancer among men. A grade 1 pattern is very well differentiated whereas grade 5 is poorly differentiated. In grade 1 and 2 (benign patterns), the glands have large lumen with prominent nuclei. Grade 3 glands are more circular with smaller lumen and thin nuclei boundaries. The glands start to lose their architecture as we move to grade 4 whereas grade 5 glands have poorly defined units. This grading is done manually by pathologists, which is tedious as well as error prone. Thus, an automated system of segmentation and classification of the gland tumors is the need of the hour. The techniques that have been proposed in the literature so far are highly dependent on the presence of lumen or nuclei around the gland. The purpose of this project is to overcome this difficulty by using deep learning techniques.

Convolutional Neural Network

CNNs consists of multiple convolutional layers which are collection of small filters that take only a small portion of the original image as input and process over it. The outputs of these filters of a convolutional layer are pooled to have a better representation of the given image and passed as input to the next convolutional layer. The same set of weights of a filter is used over the entire image which reduces the training time and memory required. The complexity of the extracted features increases as the number of layers increase. These learned features can be passed through a hidden layer with a linear or nonlinear activation function. The output is used as input to a logistic regression model, multi-layer perceptron or other classifiers to solve binary or multi-class classification problems.

Proposed methodology

The dataset consisted of H&E stained biopsy slides of colon tissue of size 775x522. The binary image was generated where the glands were marked in white while the background in black. A fixed number of patches of size 31x31 were selected from every image such that the ratio of gland vs non-gland centre pixels is 1:1. The CNN consisted of 2 convolutional layers, 1 hidden layer and a logistic regression layer. The size and the number of filters were experimented upon to achieve the minimum test and validation error. Pooling was done which reduced the output of the convolution layer by a factor of 2. Dropout was also added with probabilities 0.15 and 0.25 in the convolutional layers. Batch processing was used for training and testing with a batch size of 420 and the number of epochs was set to 200. The learning used was dynamic.



Architecture of the Convolutional Neural Network

Results

The following table lists the validation and test error obtained for the different parameters chosen for the CNN architecture.

S.No.	Dataset size (Number of patches)			Convolutional Layer 1 Filters		Convolutional Layer 2 Filters		Learning Rate	Decrement in Learning Rate	Error (in %)	
	Training	Validation	Test	Number	Size	Number	Size			Validation	Test
1.	11760	2940	2940	20	5x5	50	5x5	0.001	-	15.07	25.54
2.	11760	2940	2940	30	5x5	50	5x5	0.001	-	15.34	25.00
3.	117600	2940	2940	30	3x3	50	5x5	0.001	0.25	12.99	16.83
4.	117600	29400	29400	30	3x3	50	5x5	0.001	0.25	21.50	30.80
5.	117600	11760	11760	30	5x5	80	5x5	0.001	0.25	13.66	14.44
6.	117600	11760	11760	30	3x3	80	5x5	0.001	0.25	12.67	14.78
7.	117600	11760	11760	20	3x3	50	5x5	0.001	0.25	14.47	14.51
8.	117600	11760	11760	20	5x5	50	5x5	0.001	0.25	14.37	15.51
9.	117600	11760	11760	100	3x3	80	5x5	0.001	0.25	11.26	14.50
10.	117600	11760	11760	120	3x3	80	5x5	0.001	0.25	11.82	14.97
11.	117600	11760	11760	50	3x3	80	5x5	0.001	0.00025	11.07	13.62

The below figures depict the original, annotated and the obtained segmented output for each biopsy slide.

