**ABSTRACT**

Cancer is a disease characterized by abnormal cell growth in the human body. Cancer is evaluated by histopathological examination, which is important for further treatment planning. The tubule formation, mitotic cell count and nuclear pleomorphic are three parameter used for cancer grading. Mitotic cell (MC) count is one of important factor in cancer diagnosis from histopathological images.MC detection is very challenging task in cancer diagnosis because mitotic cell are small objects with a large variety of shapes. The aim is to evaluate performances of SVM (Support Vector Machine) classifier and Bayesian classifier in cancer diagnosis. This proposed work consists of three modules: 1) Pre-processing, 2) MC detection and segmentation, and 3) MC classification.MC detection and segmentation are performed by Bayesian modelling and local region threshold method. The segmented mitotic cell is classified by both SVM classifier and Bayesian classifier and their performance is evaluated.

**INDEX**

1. Technologies Used
2. Work Done
3. Code
4. Results and Snapshots
5. Conclusion
6. References

**Technologies Used:**

**MATLAB** (**mat**rix **lab**oratory) is a [multi-paradigm](https://en.wikipedia.org/wiki/Multi-paradigm_programming_language) [numerical computing](https://en.wikipedia.org/wiki/Numerical_analysis) environment and [fourth-generation programming language](https://en.wikipedia.org/wiki/Fourth-generation_programming_language). MATLAB allows [matrix](https://en.wikipedia.org/wiki/Matrix_(mathematics)) manipulations, plotting of [functions](https://en.wikipedia.org/wiki/Function_(mathematics)) and data, implementation of [algorithms](https://en.wikipedia.org/wiki/Algorithm), creation of [user interfaces](https://en.wikipedia.org/wiki/User_interface), and interfacing with programs written in other languages, including [C](https://en.wikipedia.org/wiki/C_(programming_language)), [C++](https://en.wikipedia.org/wiki/C%2B%2B), [C#](https://en.wikipedia.org/wiki/C_Sharp_(programming_language)), [Java](https://en.wikipedia.org/wiki/Java_(programming_language)), [Fortran](https://en.wikipedia.org/wiki/Fortran) and [Python](https://en.wikipedia.org/wiki/Python_(programming_language)).

MATLAB was used in this project in order to carry out various operations and manipulations on the image.

MATLAB was used to generate the dataset and divide the image into two sets cancerous and non-cancerous.

**WORK DONE**

The following steps have been taken in order to process the images.

1. The first step in Pre-processing is to choose grayscale channel. The grayscale image generally has the highest contrast between structures even in the presence of different backgrounds.
2. Contrast enhancements improve the perceptibility of objects in the scene by enhancing the brightness difference between objects and their backgrounds. Histogram equalization is a technique for adjusting image intensities to enhance contrast. Histogram Equalization implies mapping one distribution to another distribution (a wider uniform distribution of intensity values) so the intensity values are spreaded over the whole range
3. In best fitted ellipse calculation, an ellipse is fitted based on the boundary points of a candidate region using the direct least-squares fitting algorithm.
4. In Local region determination, a local circular area with an enlarged radius (2 × Lj) is determined to recover the intact nuclei.
5. A local threshold (denoted as T local) for this local circular area is calculated by using the Otsu thresholding method.
6. In Local threshold calculation, the local threshold T local is used to segment its corresponding local area into nuclei and background region.
7. Two commonly used texture-based features, Tamura textures and Haralick-based textures are extracted to measure the textural information on the surface of the regions.
8. The final module of proposed method is classification of MC. The classification framework that can deal with the imbalanced class distribution is employed to address the grade of cancer after MC detection. Here we use convolutional neural networks.

**Code:**

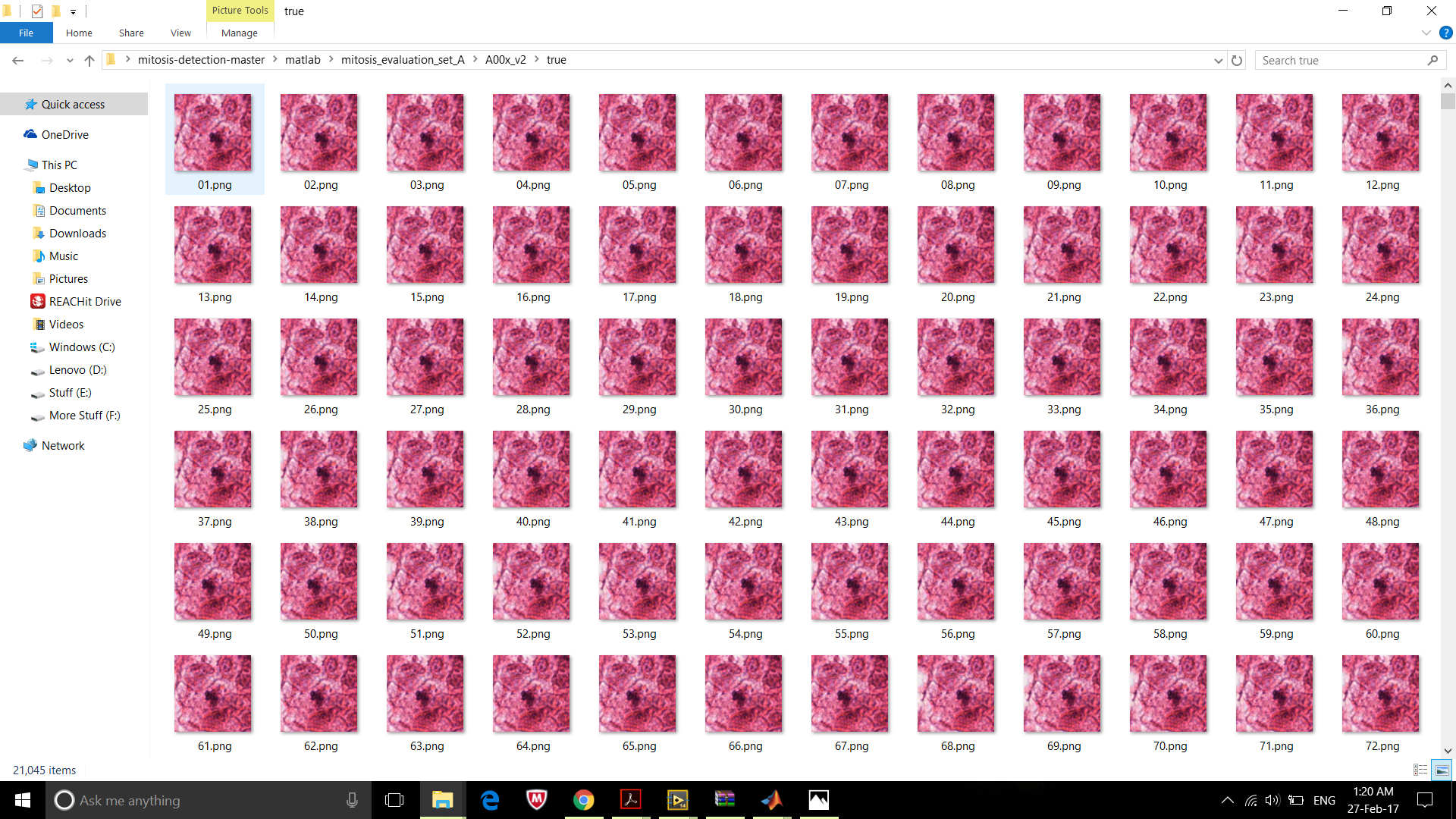
* **Generating dataset:**

|  |
| --- |
| clear |
|  | close all |
|  |  |
|  | % path to folder containing training data |
|  | old\_trainset\_folder = '/home/andrew/mitosis/MITOS/testing/'; |
|  |  |
|  | % make training set directories |
|  | new\_trainset\_folder = '/home/andrew/mitosis/mitosis-test/'; |
|  | if exist(new\_trainset\_folder, 'dir') |
|  | rmdir(new\_trainset\_folder, 's') |
|  | end |
|  | mkdir([new\_trainset\_folder 'true']) |
|  | mkdir([new\_trainset\_folder 'false']) |
|  |  |
|  | % count the total number of images for the waitbar |
|  | N = 0; |
|  | for j=1:5 |
|  | % find the number of images in the folder |
|  | folder = [old\_trainset\_folder 'A' num2str(j-1, '%02u') '\_v2/']; |
|  | n = length(dir([folder '\*.csv'])); |
|  | N = N + n; |
|  | end |
|  |  |
|  | % iterate over the twelve patients |
|  | M = 0; |
|  | P = 0; |
|  | k = 0; |
|  | h = waitbar(0,'Creating training set ... 0 %'); |
|  | for j=1:5 |
|  | % find the number of images in the folder |
|  | folder = [old\_trainset\_folder 'A' num2str(j-1, '%02u') '\_v2/']; |
|  | image\_files = dir([folder '\*.bmp']); |
|  | csv\_files = dir([folder '\*.csv']); |
|  | n = length(image\_files); |
|  |  |
|  | % create new data set |
|  | for i=1:n |
|  | image\_file = [folder image\_files(i).name]; |
|  | csv\_file = [folder csv\_files(i).name]; |
|  | [m, p] = add\_dataset(image\_file, csv\_file, new\_trainset\_folder); |
|  | M = M + m; |
|  | P = P + p; |
|  | k = k + 1; |
|  | waitbar(k/N,h,['Creating training set ... ' num2str(100\*k/N) ' %']); |
|  | end |
|  | end |
|  |  |
|  | close(h) |
|  |  |
|  | disp(['Created ' num2str(P) ' window training images from ' num2str(N) ' large training images containing ' num2str(M) ' mitotic figures.']) |

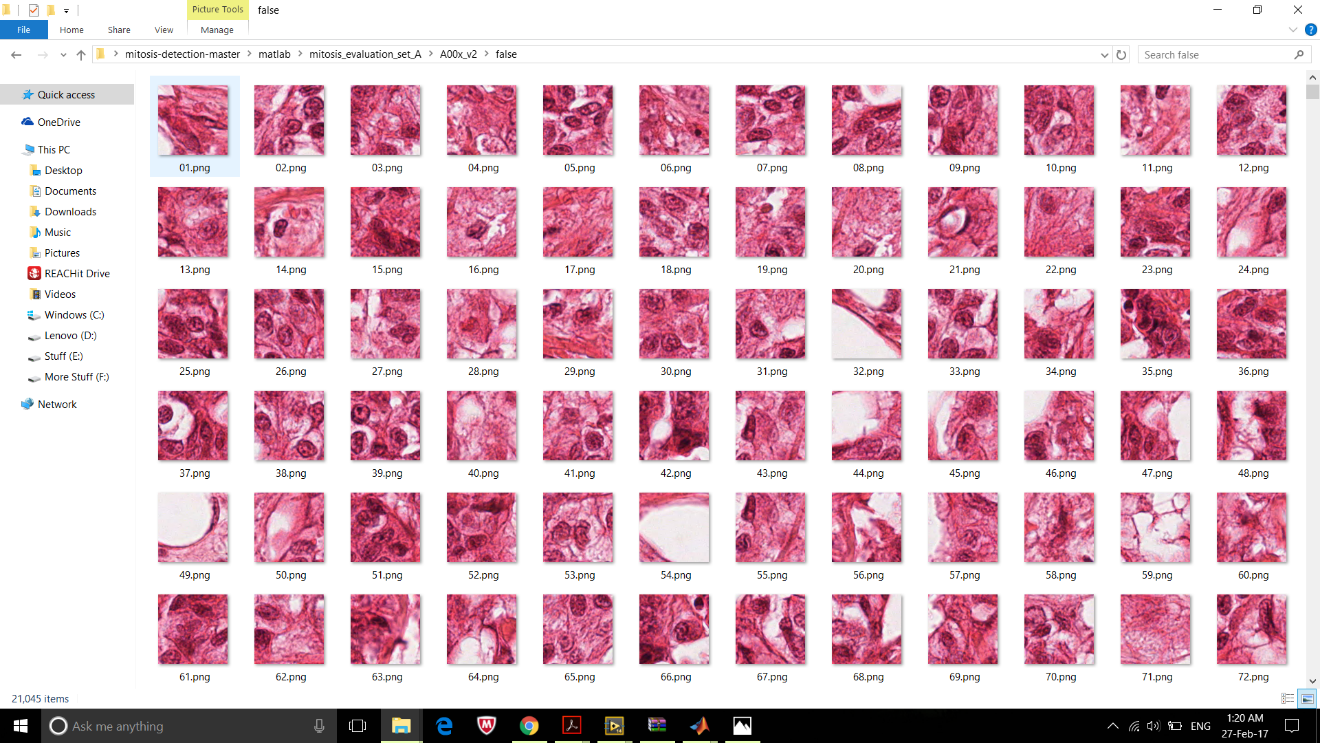
* **Plotting the detections**

|  |
| --- |
| function plot\_detections(imageFile, csvFile, threshold) |
|  |  |
|  | A = imread(imageFile); |
|  | results = csvread(csvFile); |
|  |  |
|  | % cut results off at the threshold |
|  | for i=1:size(results,1) |
|  | if results(i,3) < threshold |
|  | results = results(1:i-1,:); |
|  | break |
|  | end |
|  | end |
|  |  |
|  | % mark detections |
|  | for i=1:size(results,1) |
|  | A = insertShape(A, 'FilledCircle', [results(i,2) results(i,1) 10], 'Opacity', 1); |
|  | end |
|  |  |
|  | imwrite(A,'detections.bmp') |

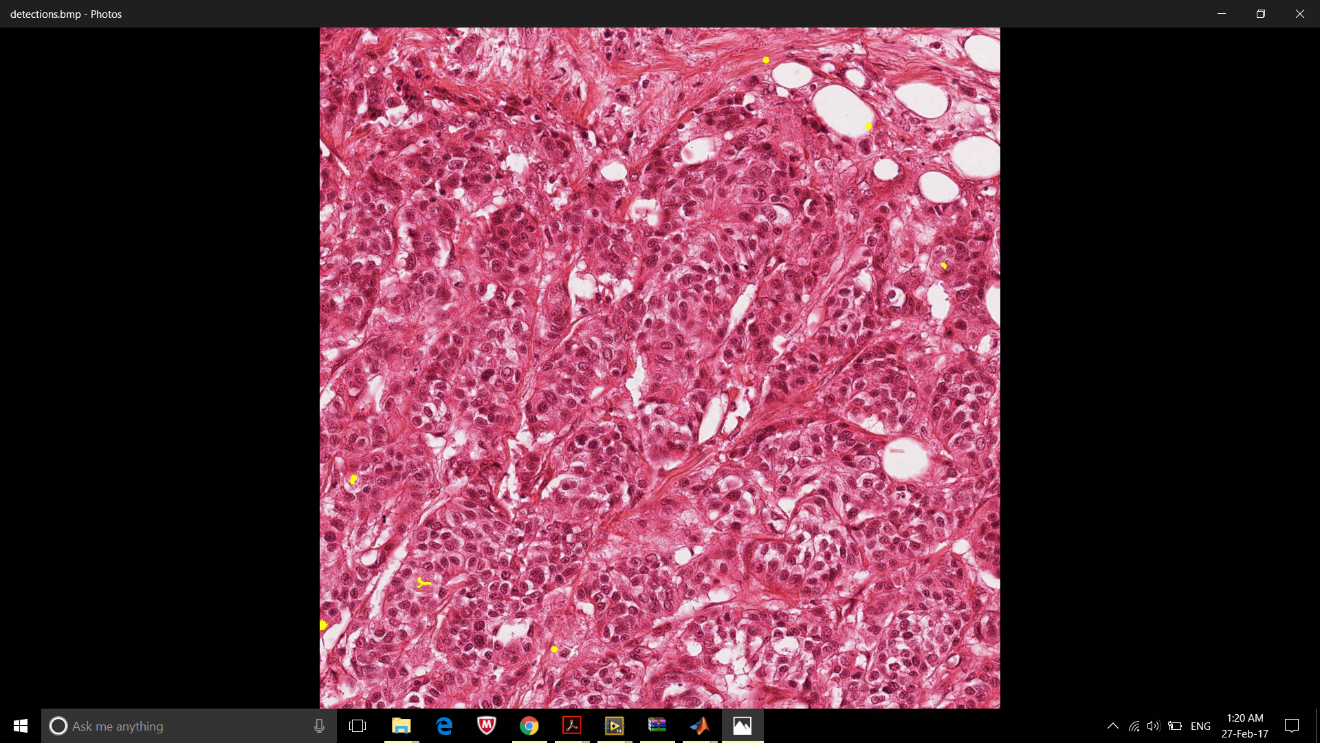
**Results and Snapshots:**



**All the divided images which are cancerous**

****

**All the divided images which are non-cancerous**



**Detecting the cancerous parts of an image (yellow parts)**

**CONCLUSION:**

A main objective of this project is to propose an effective technique for mitosis segmentation and detection in histopathological images. The plan of proposed method is to perform the study of Convolutional Neural Networks classifier on cancer diagnosis after segmentation and feature extraction steps in histopathological image.

In segmentation steps, an efficient local threshold method proposes to obtain intact nuclei regions in mitotic cell. Dimension features extracted from the candidate regions to perform classification step on cancer diagnosis. The final classification step can be used to address the grade of cancer after MC detection.

**REFERENCES:**

1. **Dataset** - <http://ludo17.free.fr/mitos_2012/download.html>
2. Learning-based mitotic cell detection in histopathological images - <http://anker.iwr.uni-heidelberg.de/publications/mip/techrep/sommer_12_learning-based.pdf>
3. http://in.mathworks.com
4. Anna Henrik Israyelyam MD, The development of molecular diagnostics for breast cancer, Yerevan state medical university, pp.1-26, August 2003.
5. <https://en.wikipedia.org/wiki/Histopathology>