Content Based Image Retrieval in Digital Pathology

Kieran, D., Wang, Y., Crookes, D., & Hamilton, P. (2012). Content Based Image Retrieval in Digital Pathology. Poster session presented at 11th European Congress on Telepathology and 5th International Congress on Virtual Microscopy, Venice, Italy.

Document Version:
Early version, also known as pre-print

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date:16. Feb. 2017

Content Based Image Retrieval in Digital Pathology

Declan Kieran[1], Yinhai Wang[1], Danny Crookes[2], Peter Hamilton[1].

[1]Bio-imaging & Informatics, Centre for Cancer Research & Cell Biology, Queens University Belfast.

[2]Institute of Electronics, Communications and Information Technology, Queens University Belfast.

Overview

The proposed CBIR system works in the following way:

i) An end-user is able to select a region of interest/concern from a candidate digital slide

ii) A robust set of textural and spectral features are calculated on the selected region

iii) This feature vector derived from the user-given image region is then trained to form a Support Vector using one-class Support Vector Machine (SVM) classification

iv) A large set of virtual slides from a database is then queried

v) Corresponding feature vectors for every region of the digital slides stored in the database are calculated

vi) Pattern recognition is performed using the previous trained Support Vector and SVM for all feature vectors

vii) The result from SVM, the so called decision value is then used as indication regarding how similar a region of an image in the database is to the candidate user selected region

viii) Using the similarity metric, the top most similar images are retrieved from the archive.

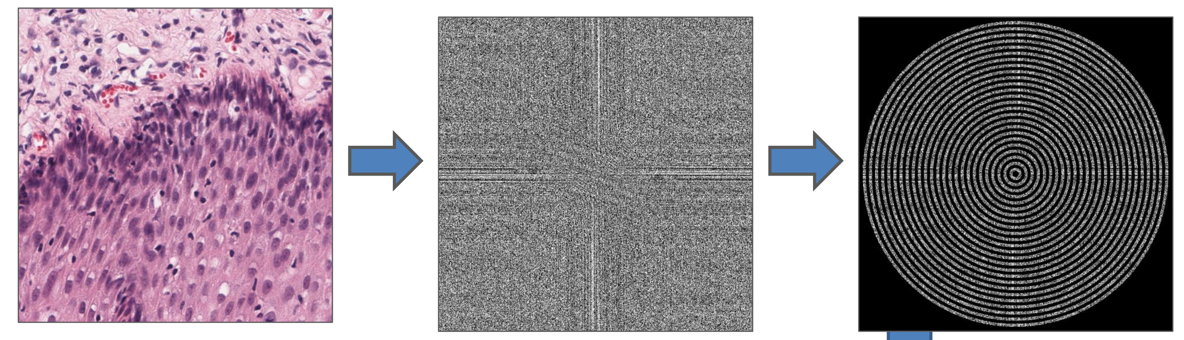
Features

Texture Measurements
$m = \frac{1}{n}\sum_{i=1}^{n} x_i$
$\sigma = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - m)\,2\right)\frac{1}{2}$
$smoothness = 1 - 1/(1+\sigma^2)$
$third\ moment = \sum_{i=1}^{n}(x_i - m)^3 p(x_i)$
$uniformity = \sum_{i=1}^{n} p^2(x_i)$
$entropy = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i)$
$skewness = \frac{E(x-m)^3}{\sigma^3}$
$kurtosis = \frac{E(x-m)^4}{\sigma^4}$

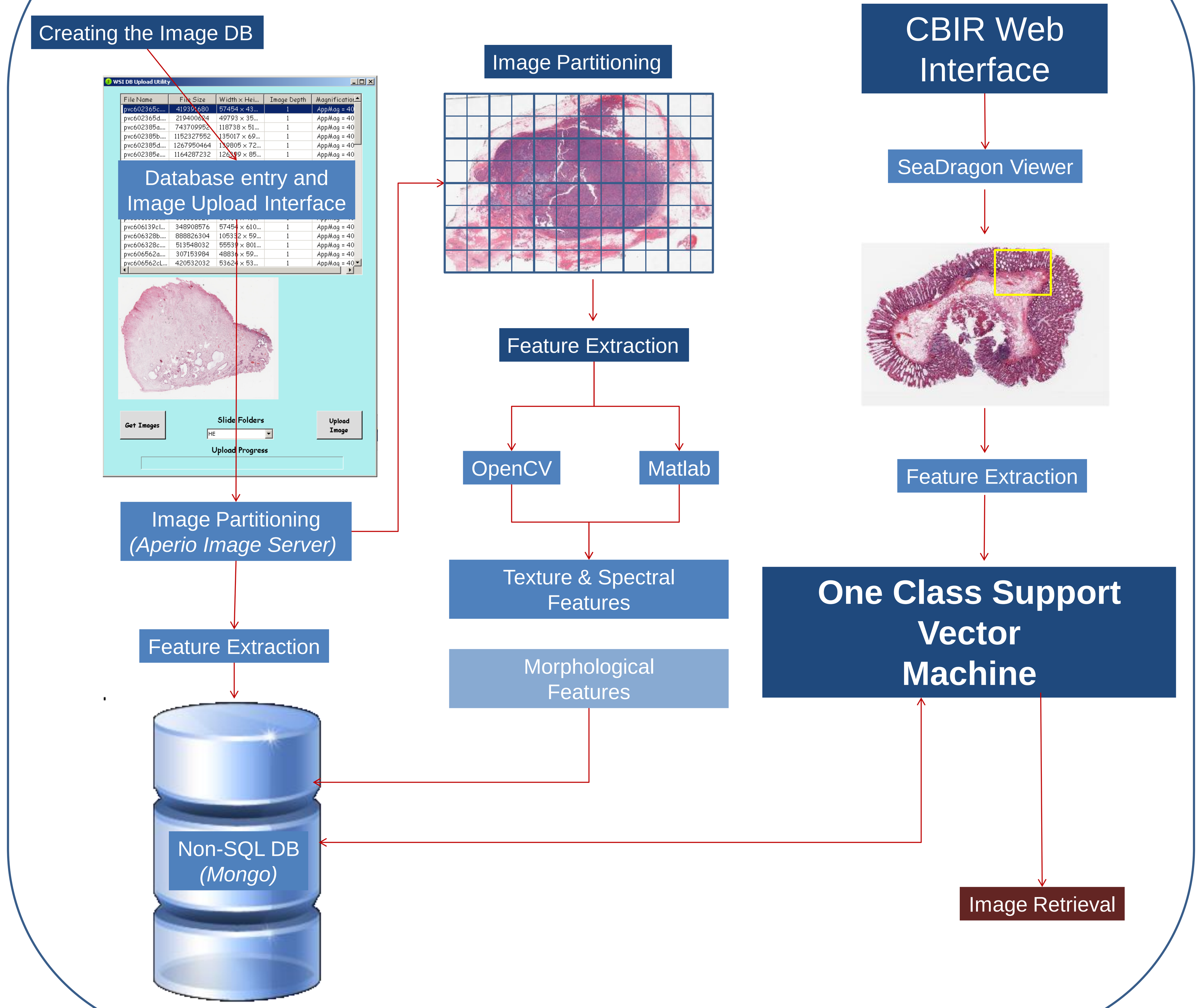
2D Invariant Moments
Normalised central moment → $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}}$, where $\gamma = \frac{p+q}{2} + 1$
$\phi_1 = \eta_{20} + \eta_{02}$
$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$
$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$
$\phi_4 = (\eta_{30} - \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} - \eta_{12})[(\eta_{30} - \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{21} + \eta_{03})^2 - (\eta_{21} - \eta_{03})^2]$
$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$

Spectral Measurements of Texture
2D Fourier Spectrum → $F(u,v) = \int\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(ux+vy)}\,dx\,dy$
2D Spectral Measures of Texture → $S(r,\theta)$, where r is a radial direction and $\theta$is a curve centred around the DC channel

- **Below gives an illustration of how spectral measurements of texture are taken**
- **Using these spectral bands provides a mean of performing very fast pseudo-segmentation**
- **This allow for higher level measurements of structure and pattern to be taken by taking texture measurements directly from these spectral bands within the Fast Fourier Transform of a given image**

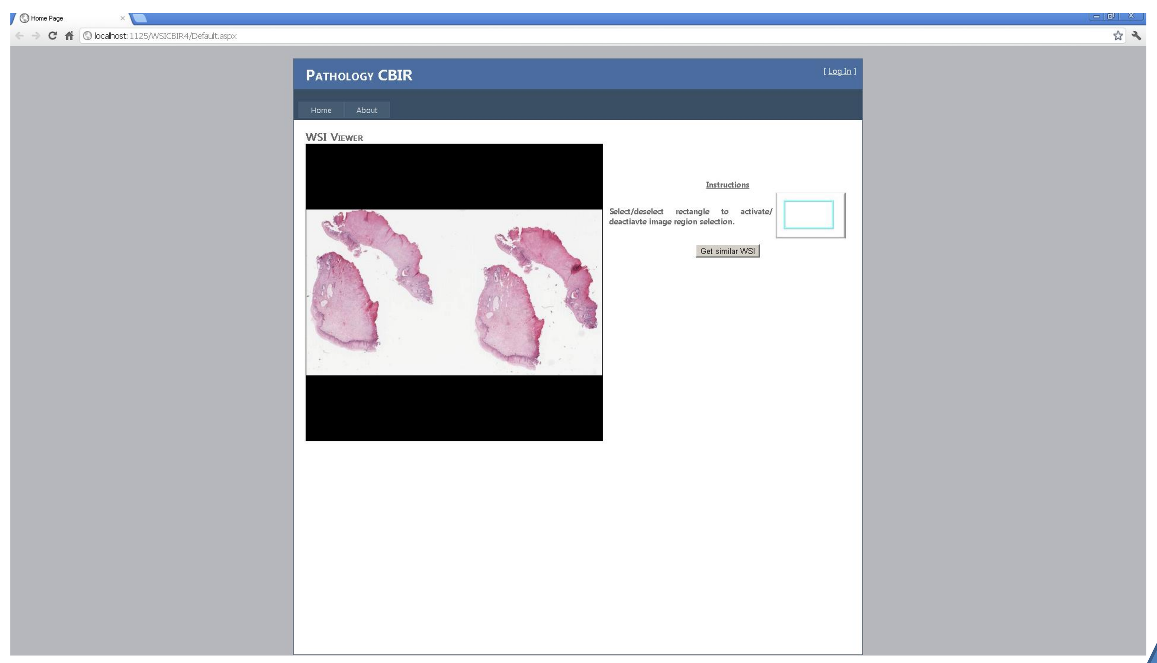


System Architecture

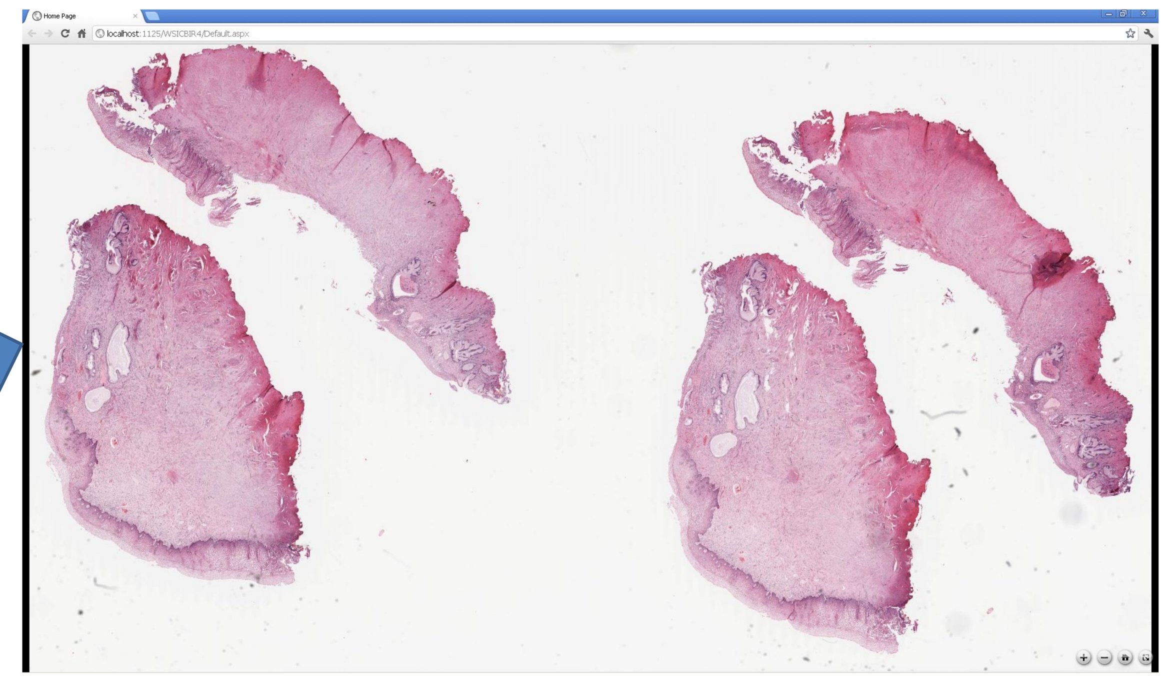


CBIR Web Interface

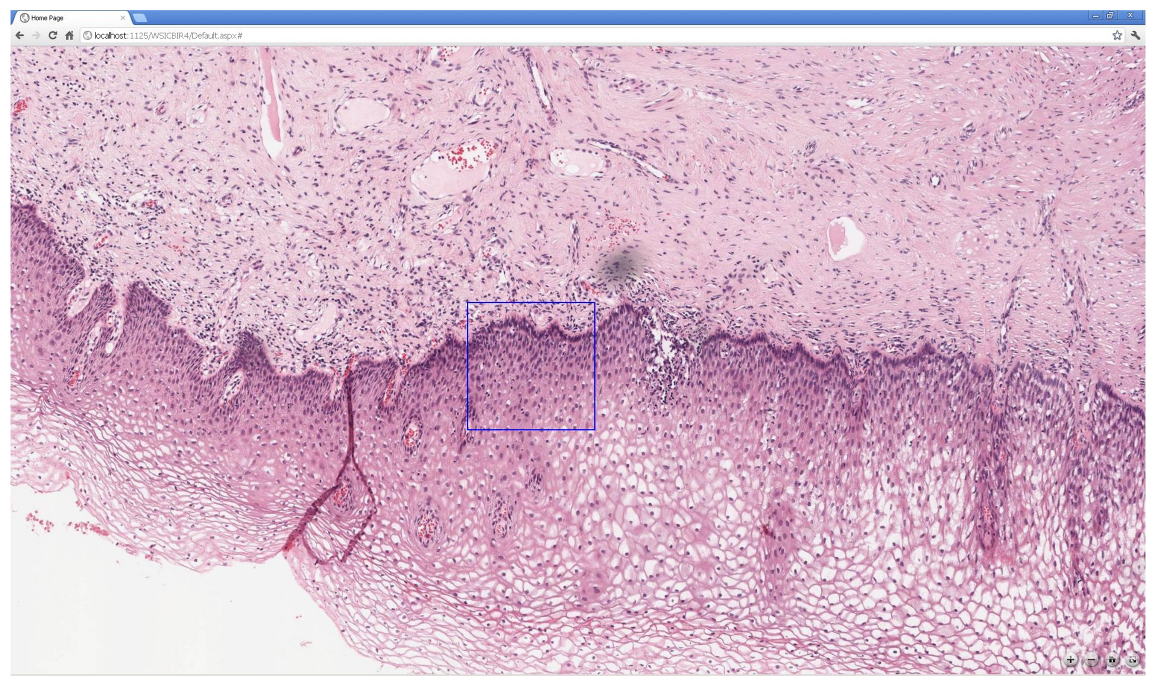
Main page



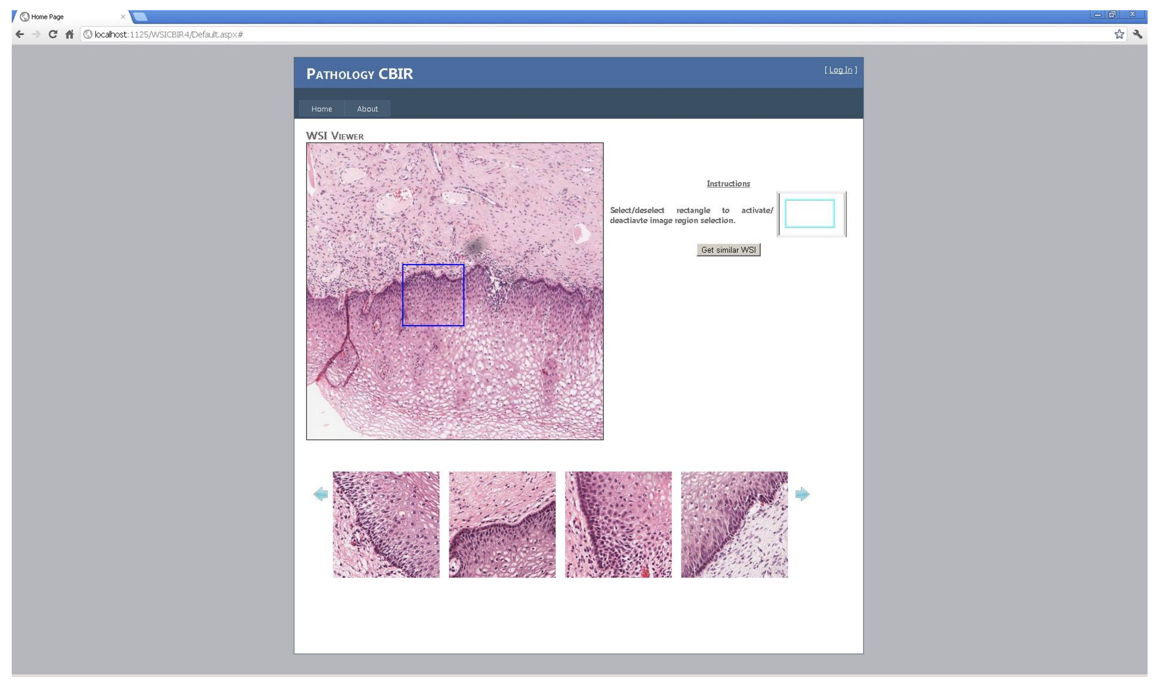
Full Screen Slide Navigation



Region of Interest Selected



Results Returned from Search of Image DB



Conclusions

CBIR has been shown to be feasible for WSI using texture and spectral feature measurements with a One Class SVM used as a classifier.

Further work needs to be developed to support high throughput analysis and evaluation on large image libraries. The computational complexity of working with such large imagery as well as the associated feature calculation is substantial.

It is clear the massively parallel nature of the problem can be exploited to provide a fast, real-time manageable CBIR system.