Quantification of Biological Systems Using MATLAB for Image Processing

Matloob Khushi, PhD

Image Processing

Project # 1:

- Proteins and Telomeres Colocalisation
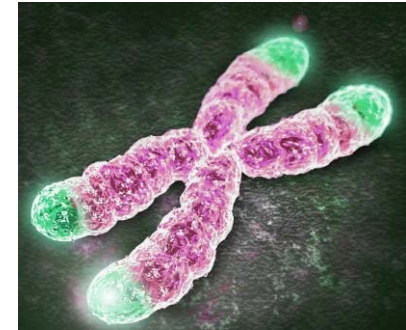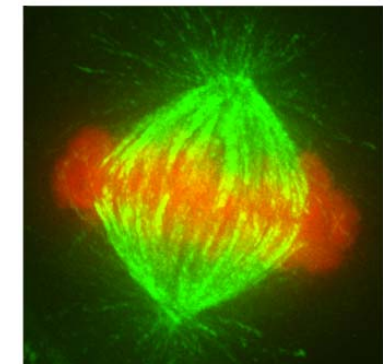
Project # 2:

- DNA and Mitotic Spindle Features Quantification

CHILDREN'S
MEDICAL
RESEARCH
INSTITUTE

Healthier kids, brighter futures

Colocalisation Background

- Existing tools to study protein colocalisation use Pearson correlation or Mander's overlap coefficients to calculate the degree of colocalisation. These methods have a number of limitations.

- These methods have been shown to be greatly affected by the background noise.

- Most tools cannot automatically select a region of interest (ROI) and thus hinders analysis of a large number of images.

- Coefficient-based methods do not clearly report whether two signals are colocalised within a ROI, nor do they report the precise number of colocalisations in a specific region.

CHILDREN'S
MEDICAL
RESEARCH
INSTITUTE

Healthier kids, brighter futures

ROI-based background intensity selection

	Red	Green	Blue	RGB
GM847				
JFCF-6/T.5K-sc1				
U-2 OS				

CHILDREN'S MEDICAL RESEARCH INSTITUTE
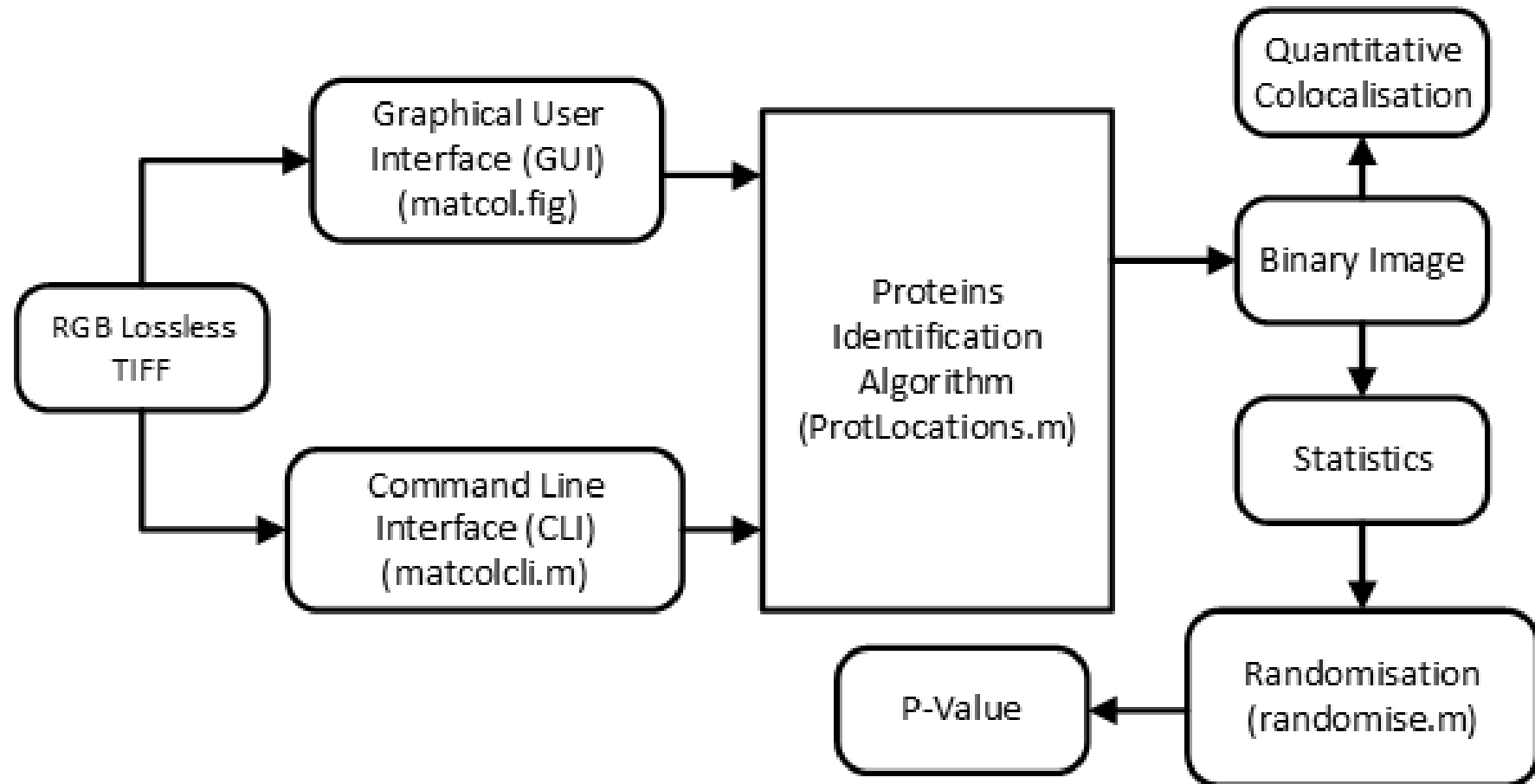
Healthier kids, brighter futures
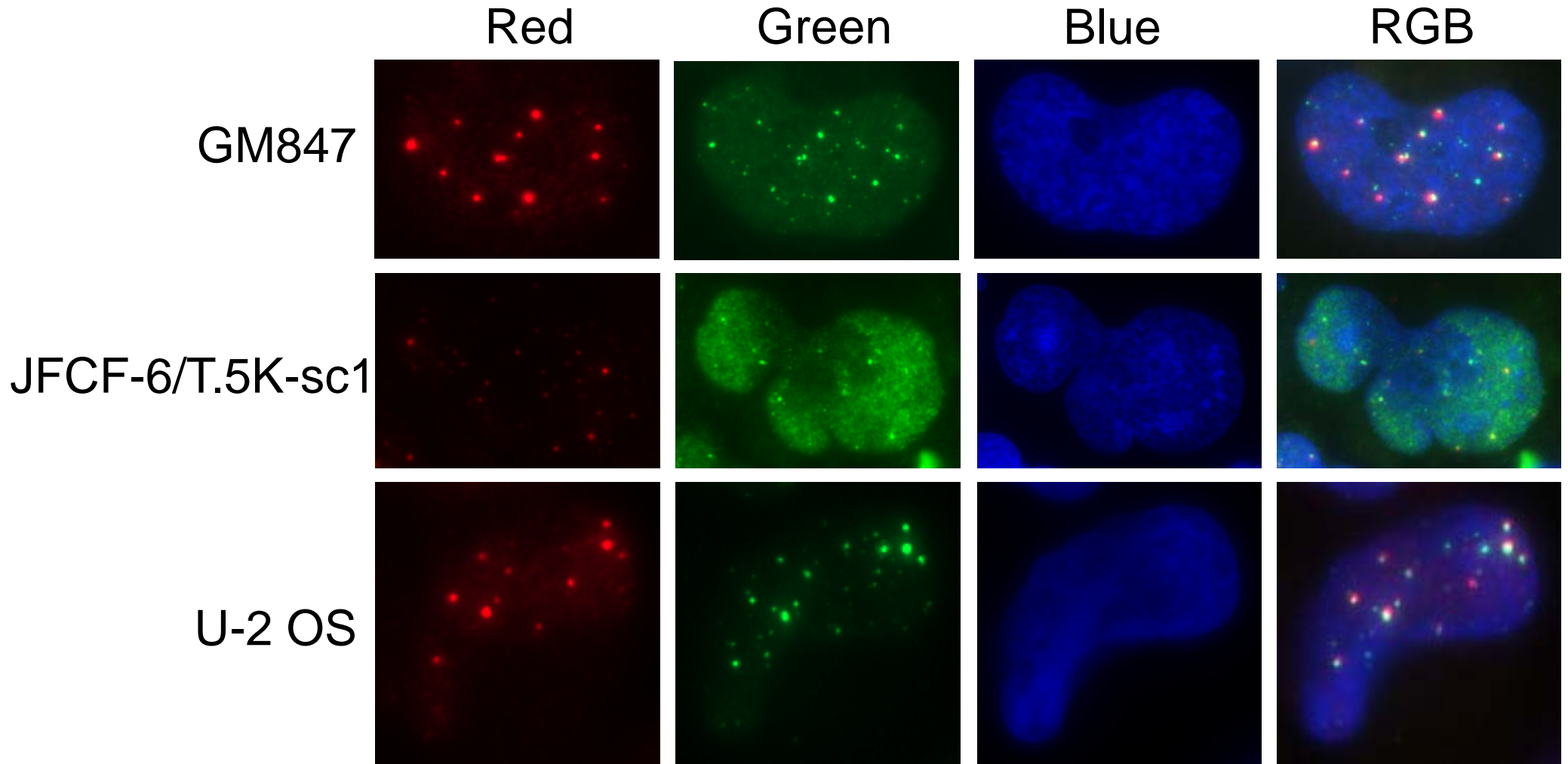
MatCol GUI

Windows of the red, green and blue channels, and their respective binary versions are provided.

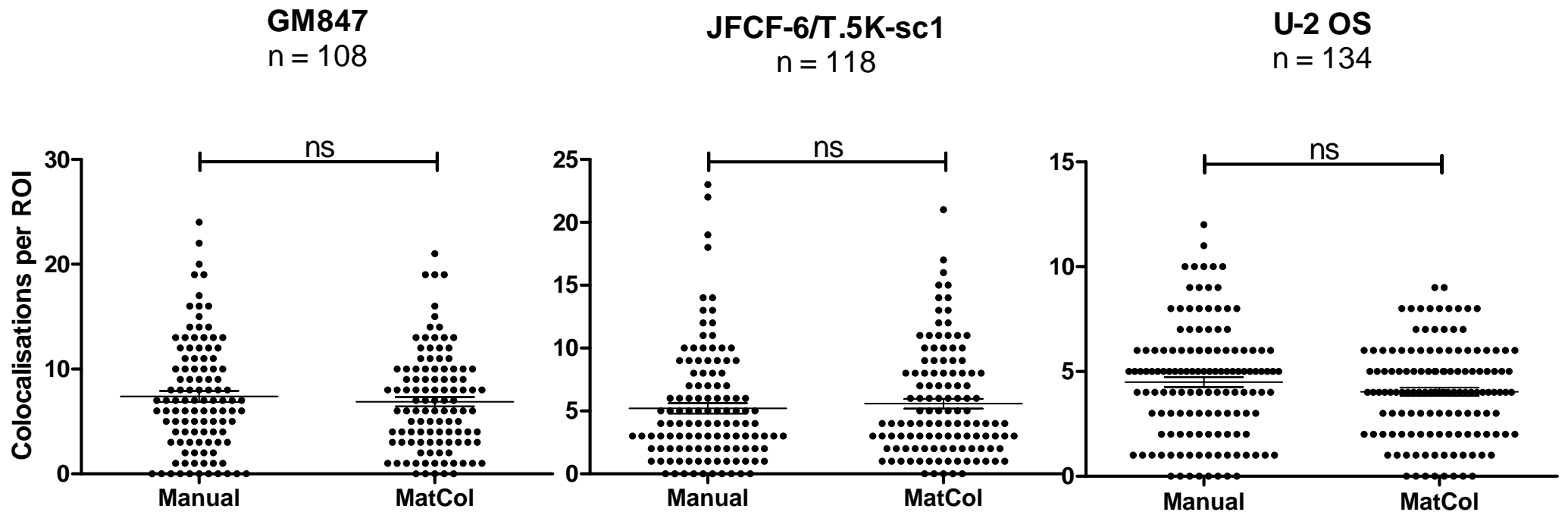Colocalisation are shown in yellow
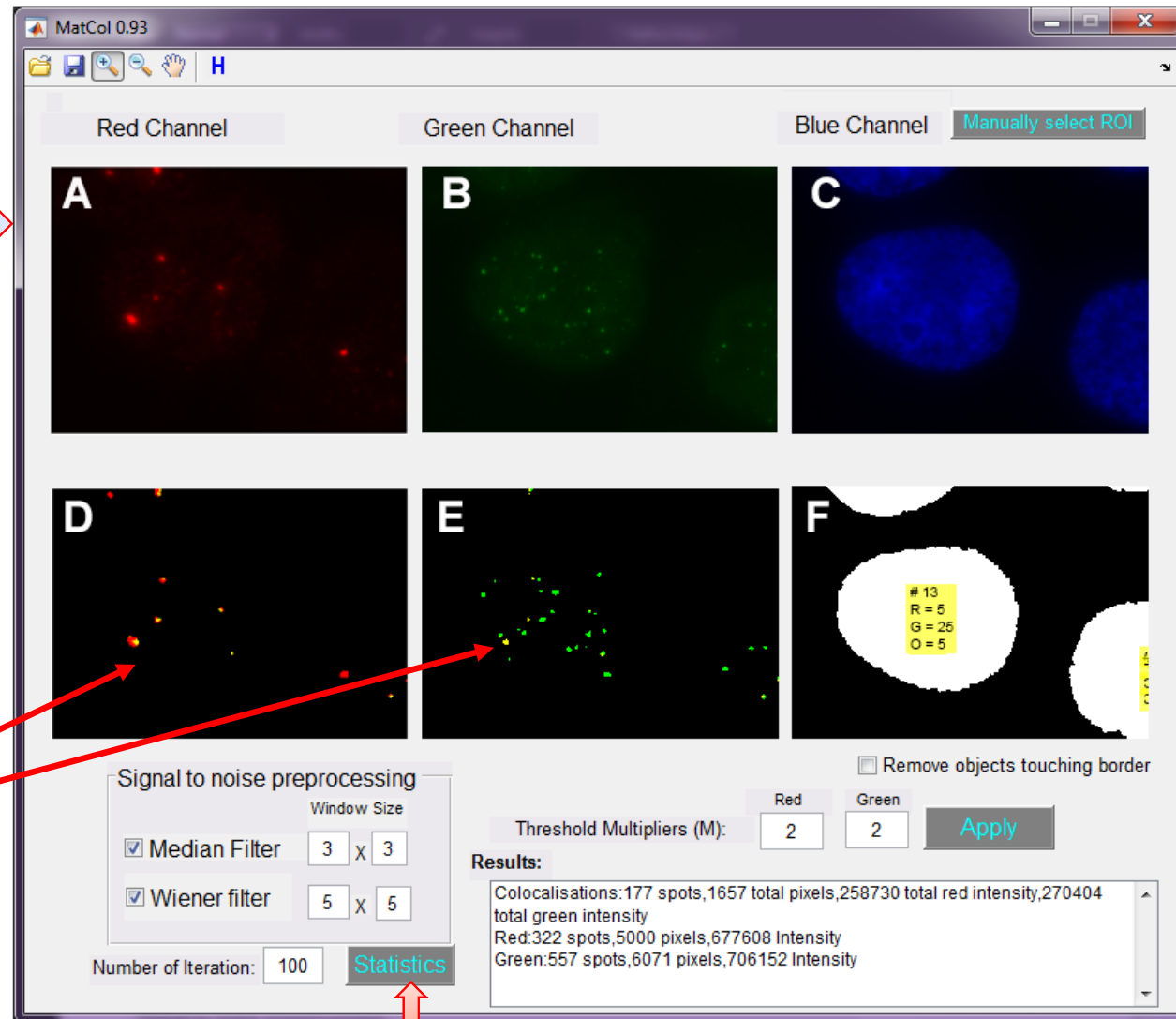
P-Value is calculated by the Student's t-test.
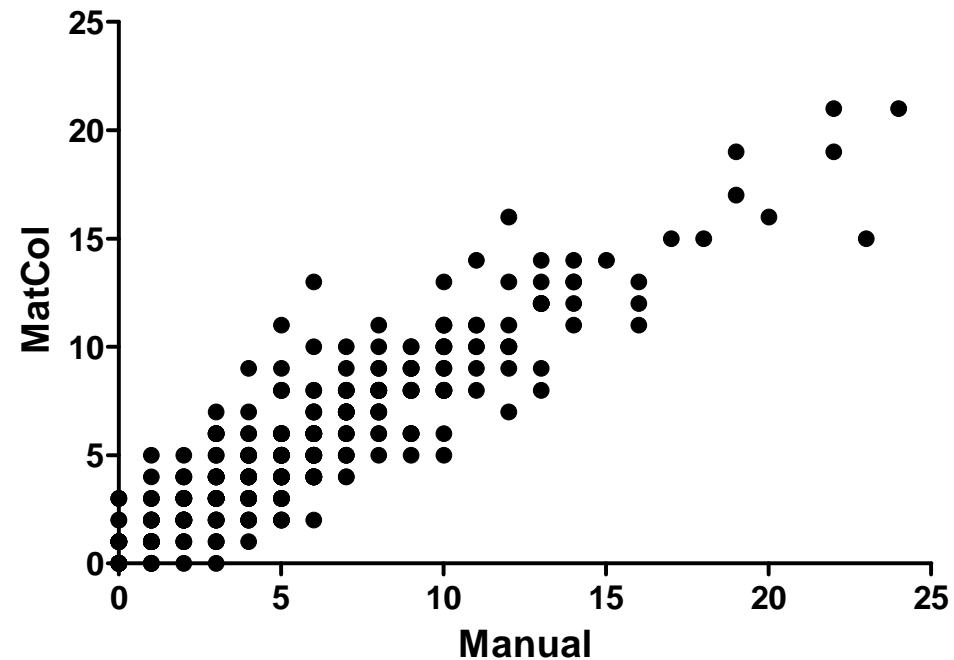
Modular software design

Non-significant difference between manual and MatCol quantification

Column graphs of manual versus MatCol colocalisations per ROI for each cell line. ns = not significant using unpaired Student's t-test.

Significant correlation between the manual and automated quantification

XY plot of the combined colocalisation data of three cell lines. A significant correlation between the manual and automated MatCol colocalisation was found (Pearson=0.91, P<0.0001).

MatCol Colocalisation Summary

- MatCol is a novel and user-friendly tool that addresses the need to study the colocalisation of two biological features.

- MatCol has enabled us to efficiently, automatically, and without bias quantify colocalisations.

- MatCol reports colocalisation as a quantity independent of intensity.

- MatCol enables the measurement of statistical significance of the observed colocalisation of two fluorescence signals against overlap by random chance.

Quantification of DNA and Mitotic Spindles

- Untreated sample (control)

- Luciferase (+ve control)

- Clathrin heavy chain Knockdown (CHC KD)

Sample Images

	Mitotic Spindle	Chromosomes	Merged

Control

Luciferase

CHC KD

Healthier kids, brighter futures
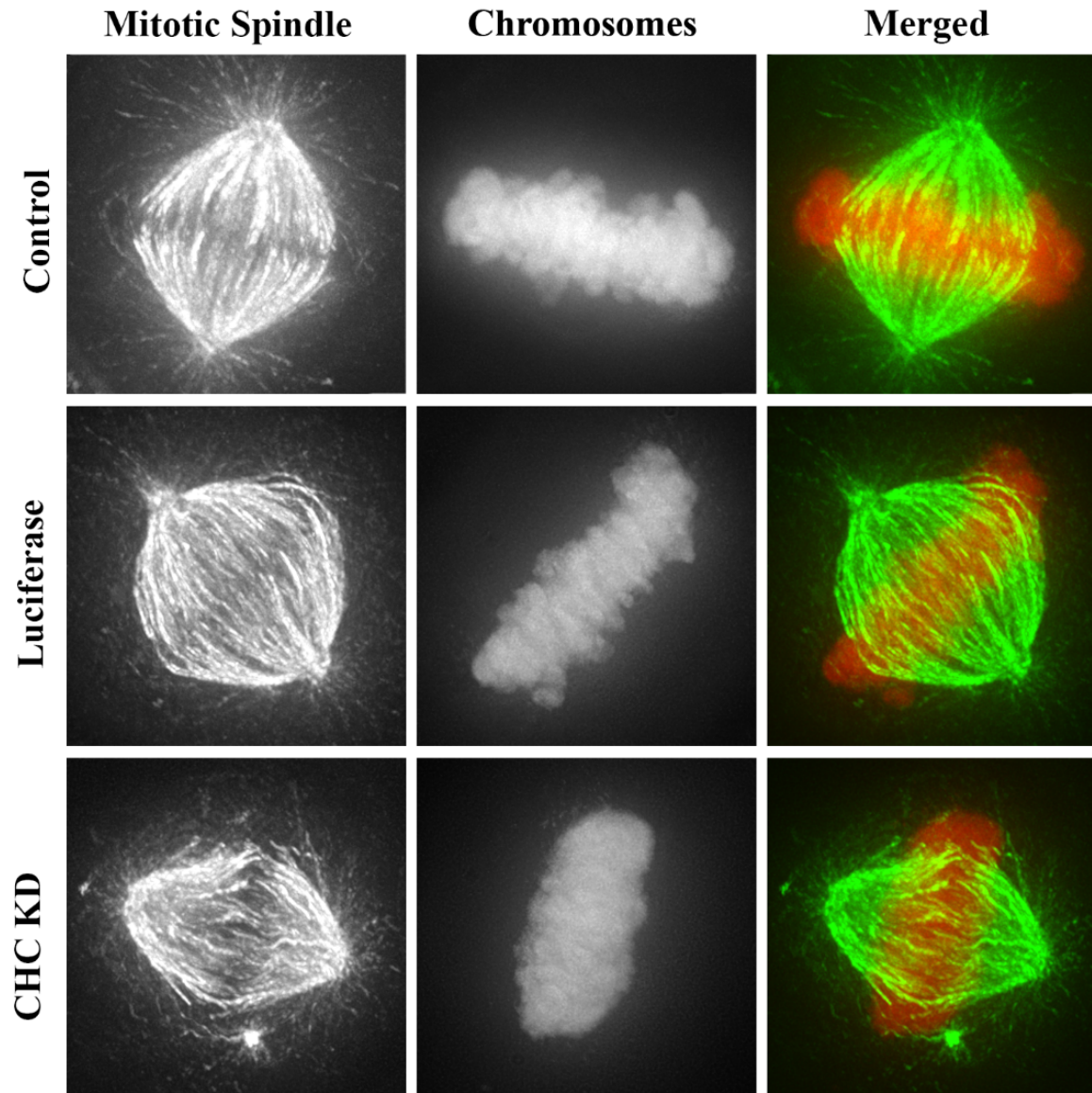
Properties Measured

Length & area	Intensity	Texture based Analysis	Other properties
Area	Mean intensity	Entropy	Orientation
Convex area	Median intensity	Standard deviation	Percent Density
Compactness	Total intensity		Fractal Dimension
Eccentricity			Euler No.
Perimeter			
Solidity			
Extent			
Major axis Length			
Minor axis length			

CHILDREN'S MEDICAL RESEARCH INSTITUTE

Healthier kids, brighter futures
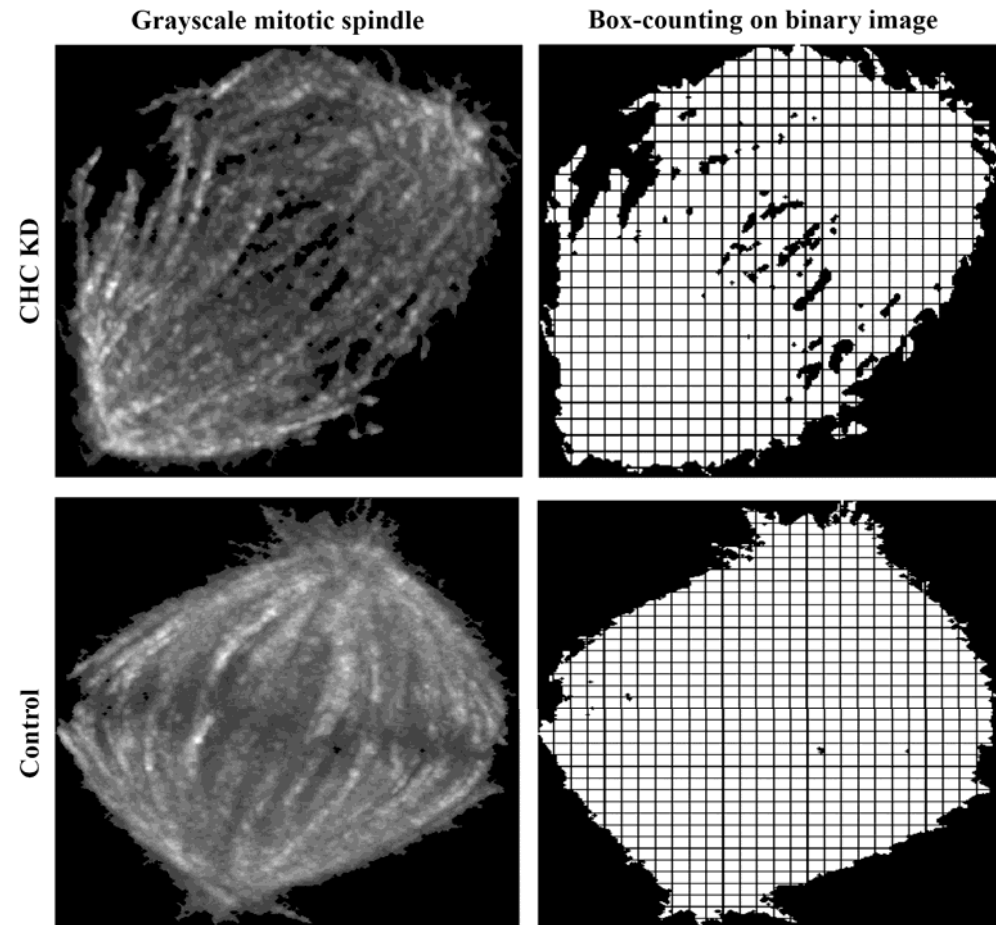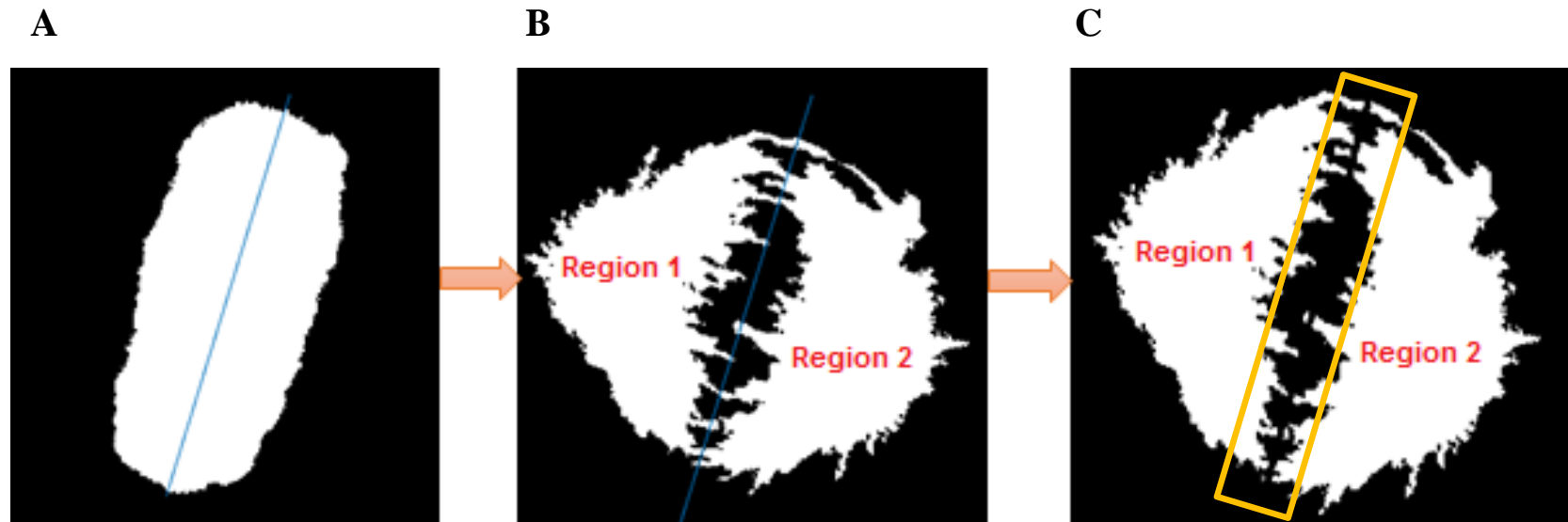
Computing the fractal dimension

- Fractals are infinitely complex patterns that are self-similar across different scales
- Describes irregularity of an object

Grayscale mitotic spindle   Box-counting on binary image

CHC KD

Control

$$fractal\ dimenison = \frac{\log(no.\ of\ self similar\ pieces)}{\log(magnification\ factor)}$$

Comparing the area between two regions

Samples	df	P-value
Control & Luciferase	99	0.691
Control & CHC KD	95	0.004
Luciferase & CHC KD	92	0.003

Healthier kids, brighter futures

Two sample t-test

No.	Image Property	Control and Luciferase		Control & CHC KD		Luciferase and CHC KD	
		T/F*	P-value	T/F*	P-value	T/F*	P-value
1	Area	0	0.5572	1	7.00E-16	1	2.15E-18
2	ConvexArea	0	0.5118	1	1.70E-09	1	9.48E-13
3	Compactness	0	0.3248	1	0.0319	1	0.0019
4	Eccentricity	0	0.0754	1	0.013	1	5.14E-05
5	Entropy	1	0.048	0	0.1115	0	0.5209
6	EulerNumber	1	0.0039	0	0.9733	1	6.79E-04
7	Fractal_Dimension	0	0.4241	1	3.34E-05	1	7.30E-07
8	Intensity(mean)	0	0.0649	0	0.2	0	0.4216
9	Intensity(median)	0	0.5671	0	0.7591	0	0.7391
10	Intensity(total)	1	0.0291	1	1.21E-04	1	1.78E-08
11	Major Axis Length	0	0.8679	1	0.0017	1	4.84E-04
12	Minor Axis Length	1	0.0062	1	1.68E-12	1	3.59E-17
13	Orientation	0	0.3414	0	0.5828	0	0.1131
14	Percent Density	0	0.9067	0	0.2974	0	0.2471
15	Perimeter	0	0.2137	0	0.1464	0	0.7524
16	Solidity	0	0.6201	1	1.88E-06	1	1.72E-06
17	Standard Deviation	1	0.0108	1	0.022	0	0.5278
18	Extent	0	0.5817	1	0.0028	1	4.00E-04
19	Satellites	0	0.4358	0	0.3811	0	0.1258

Good

Neutral

Bad

Clustering Image Properties
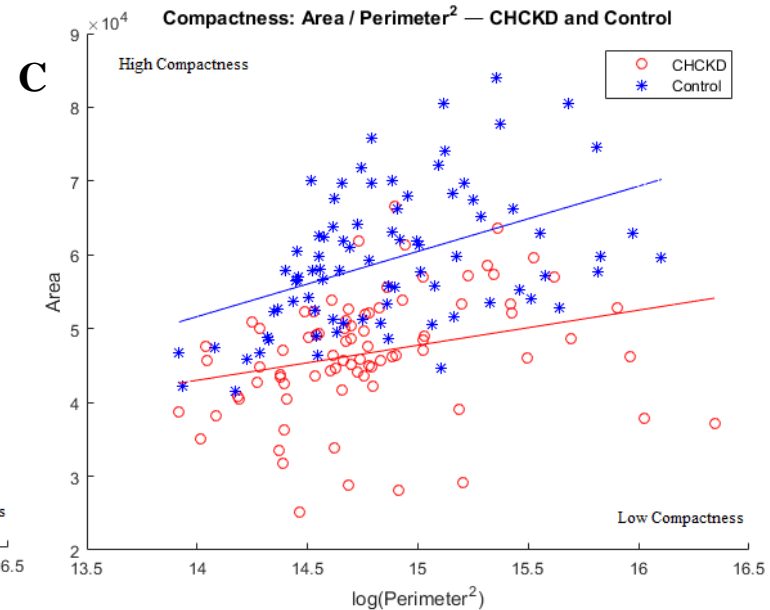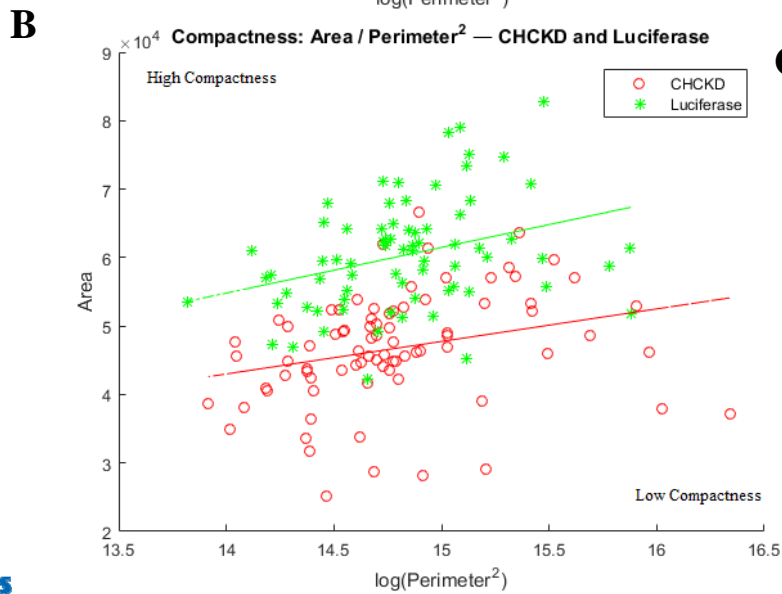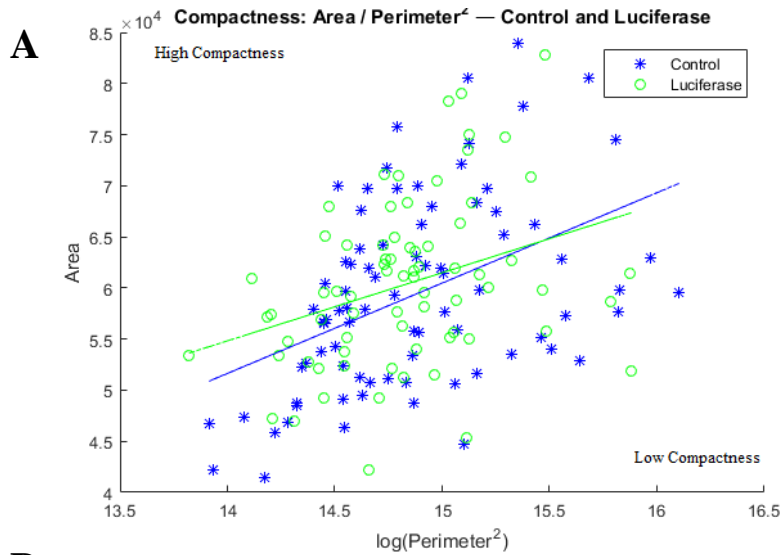
Pearson correlation coefficient heatmap

CHC data
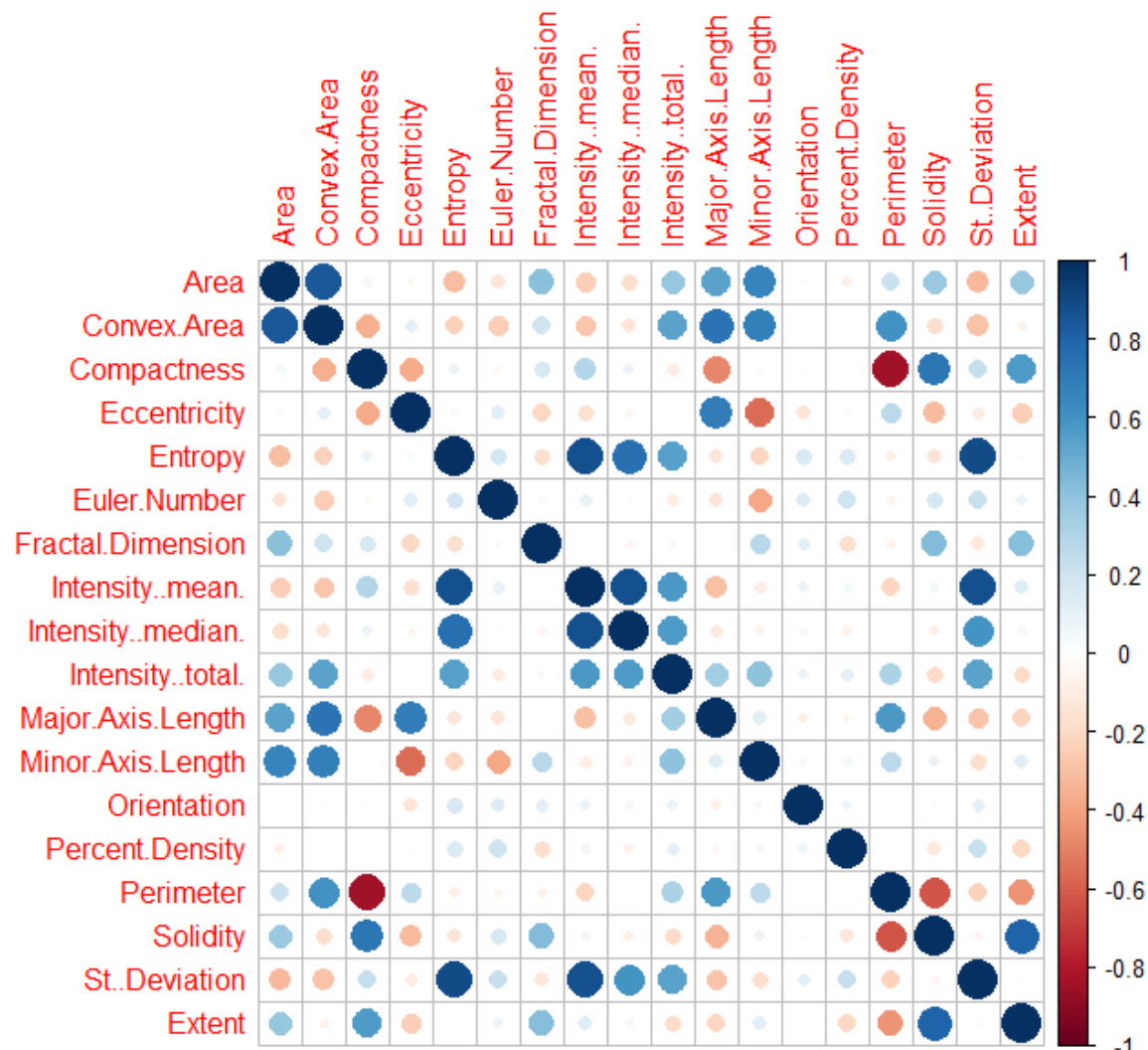
Highly correlated fields could be excluded from the prediction model

Random Forest
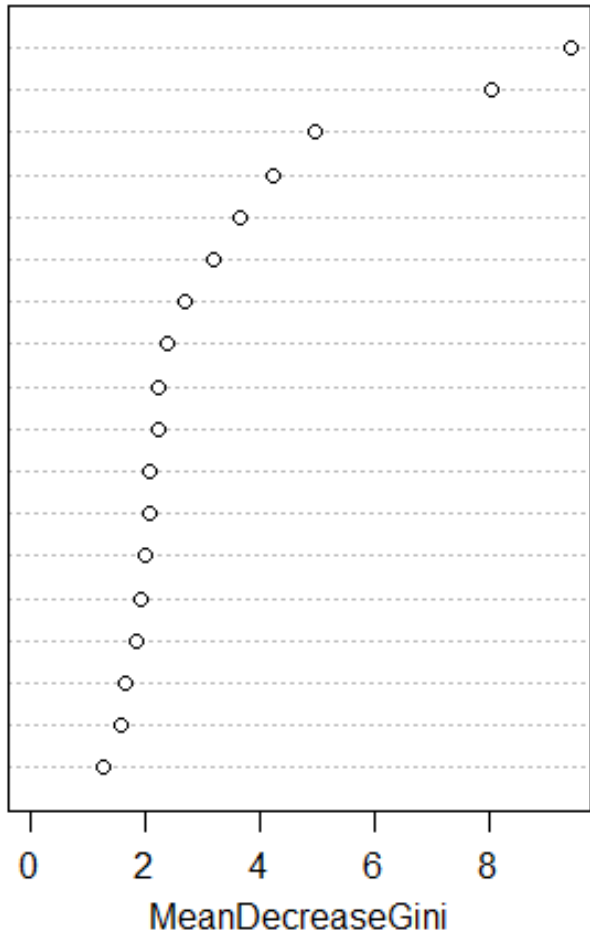
Accuracy	Sensitivity	Specificity
0.80	0.714	0.862

Summary

- The programs provide a means of quantifying image characteristics rapidly

- High throughput image analysis reduces labour, sampling error and subjectivity

- Automatic image processing can detect changes not discernible to the human eye

Acknowledgement

- A/Prof. Jonathan Arthur
- Dr. Erdahl Teber
- Dr. Christine E. Napier
- Dr. Christine Smyth
- Dr. Neftali Flores-Rodriguez
- A/Prof. Mean Chircop
- Mr. Imraan Dean
- Prof. Roger R. Reddel

Healthier kids, brighter futures

Thank you!

شكرا

Díky

תודה

Vielen Dank

धन्यवाद

Ευχαριστώ

Merci

ขอบคุณ

Bedankt

Köszönettel

PhD in Bioinformatics interested?

✉ mkhushi@mkhushi.com.au

Thank you
Danke
Khawp khun
Xie xie
Yum ... Mahalo
Salamat
Juspa... zaña
Spacibo Obrigada
Arigato

CHILDREN'S
MEDICAL
RESEARCH
INSTITUTE

Healthier kids, brighter futures