Ultrasound Image Auto-Segmentation: Trainable Weka Segmentation

CHE4180 S1 2020 - Chemical Engineering Project

Supervised by Dr. Simon Corrie

IMAGEJ

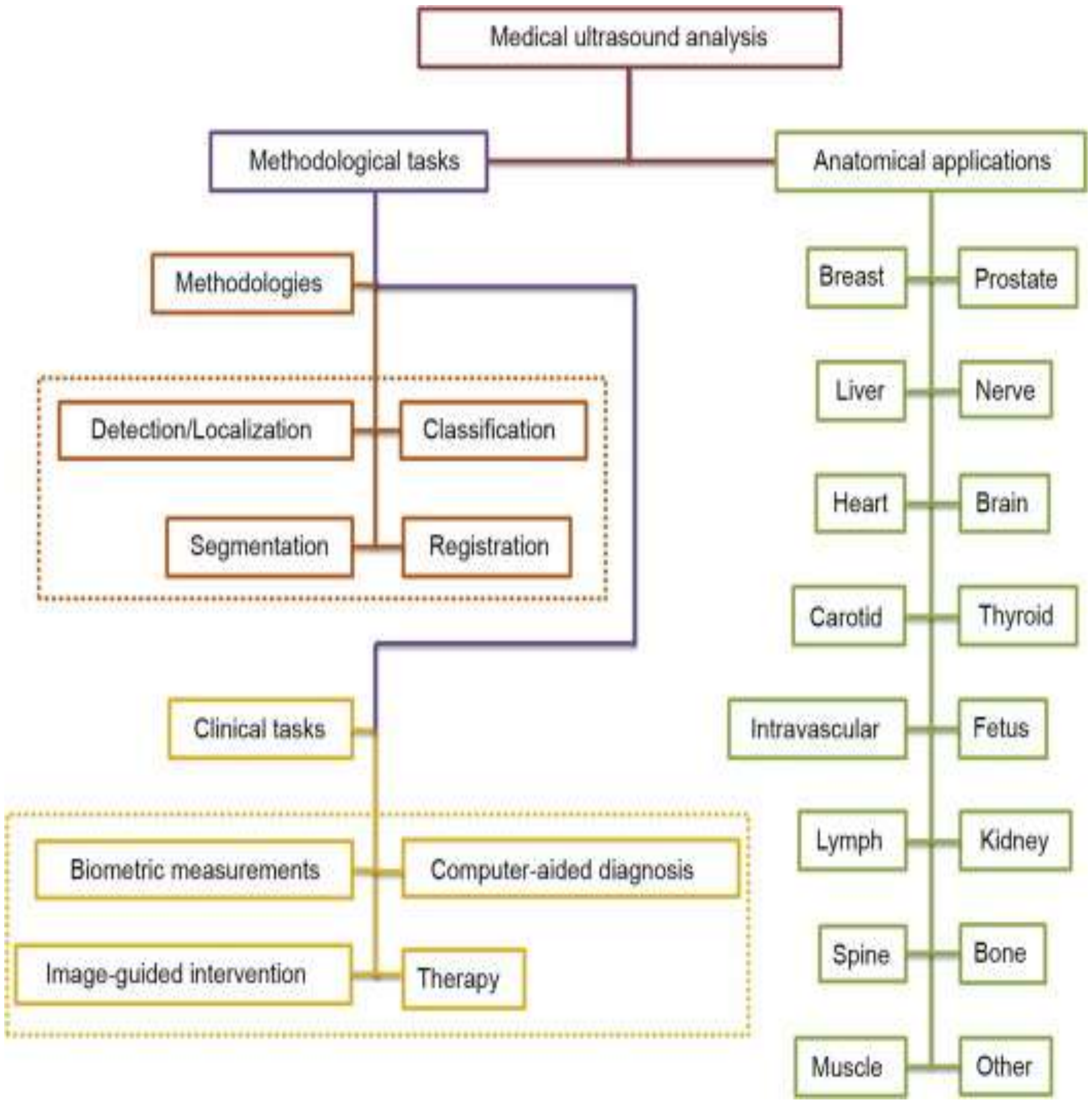
An open platform for scientific image analysis

Presented by:

George Baihan Wang

Mark Leqi Zhao

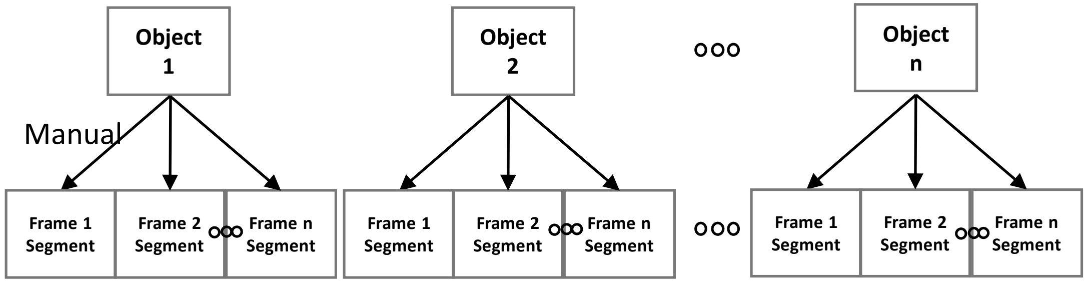
BACKGROUND:
US imaging and TWS



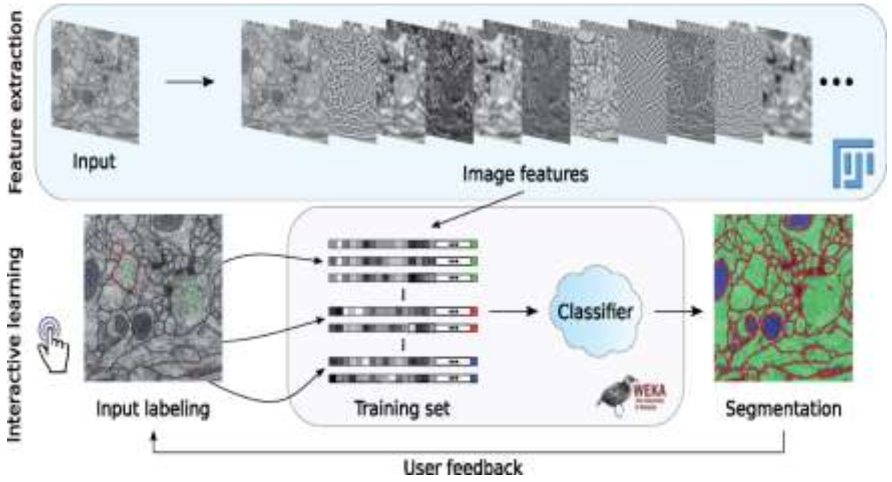
(Liu et al., 2019)

BACKGROUND:
US imaging and TWS

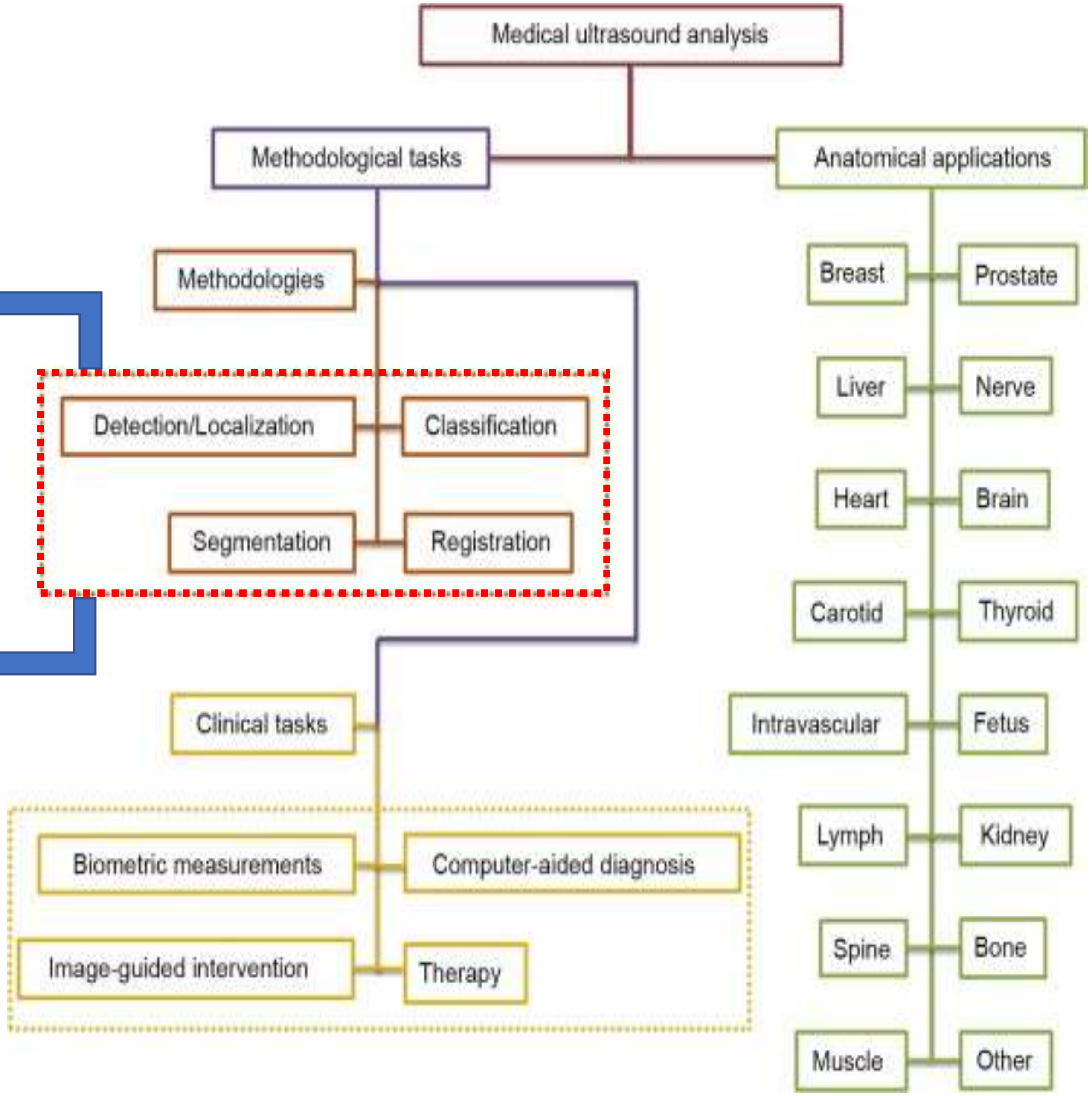
Manual Operation



Automatic Operation (Trainable Weka Segmentation)

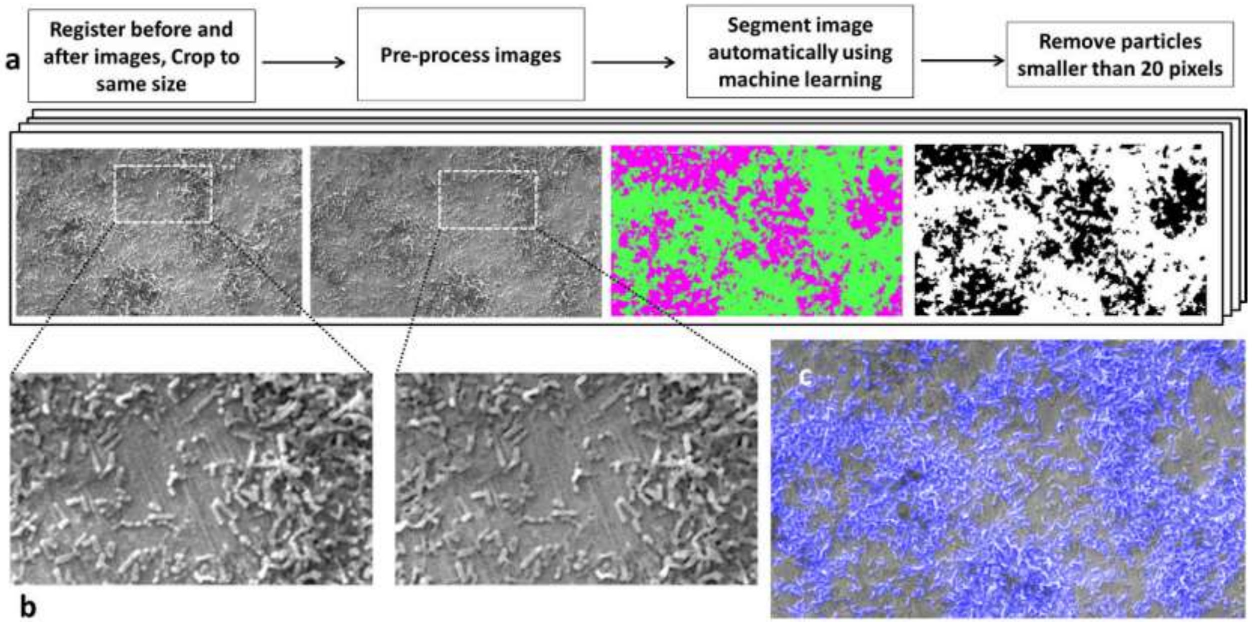


(Carreras et al., 2016)

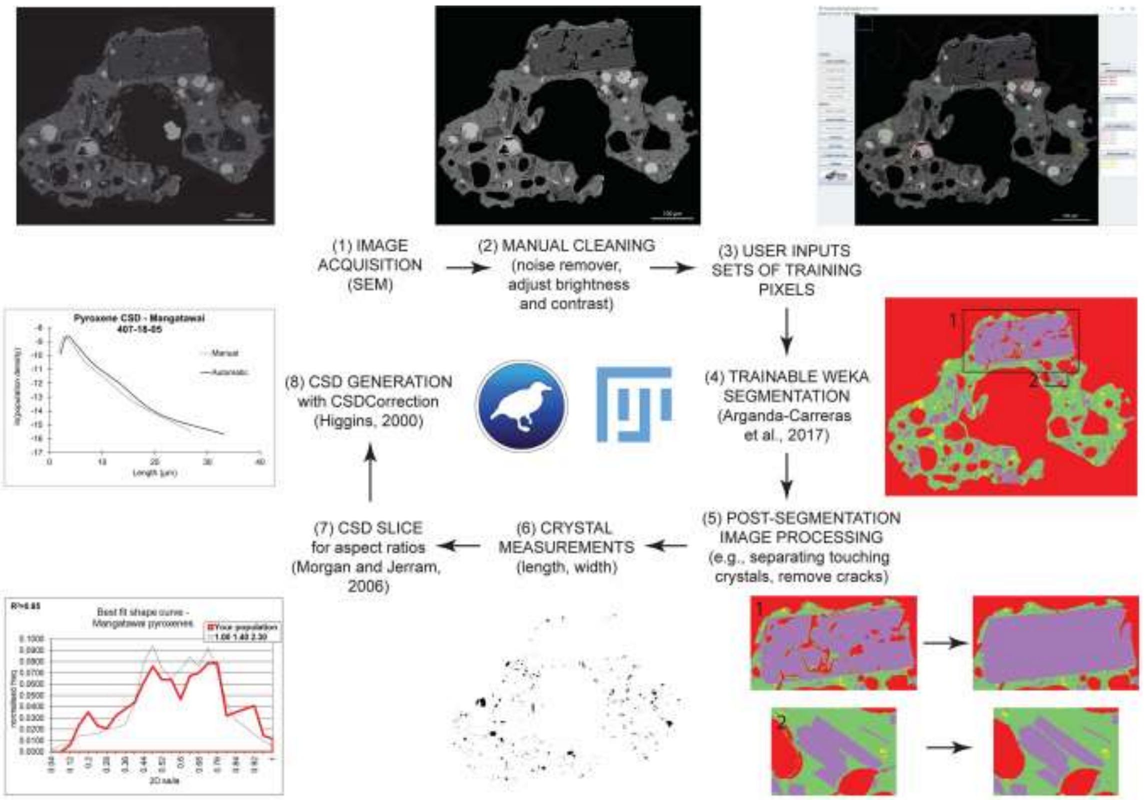


(Liu et al., 2019)

LITERATURE REVIEW: Application of TWS



TWS on SEM image of biomaterial surfaces
(Vyas et.al., 2016)



TWS on SEM image of volcanic rocks
(Lormand et.al., 2018)

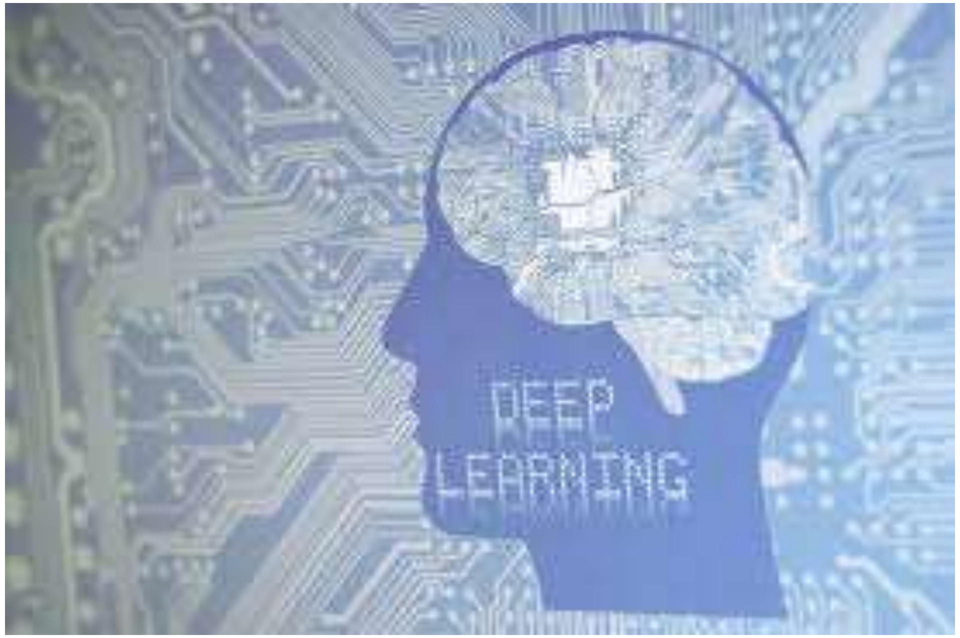
Problem Statement:
What is the performance of TWS on Ultrasound Image Auto-Segmentation

OBJECTIVES

1 Will **auto-segmentation** by TWS on ultrasound images give comparable results to **manual** segmentation?

2 Will different Machine Learning models **(Random Forest vs Naïve Bayesian)** affect segmentation results?

3 Will the application of **generalised model** give comparable results to the ordinary approach of training models on each dataset **individually**?

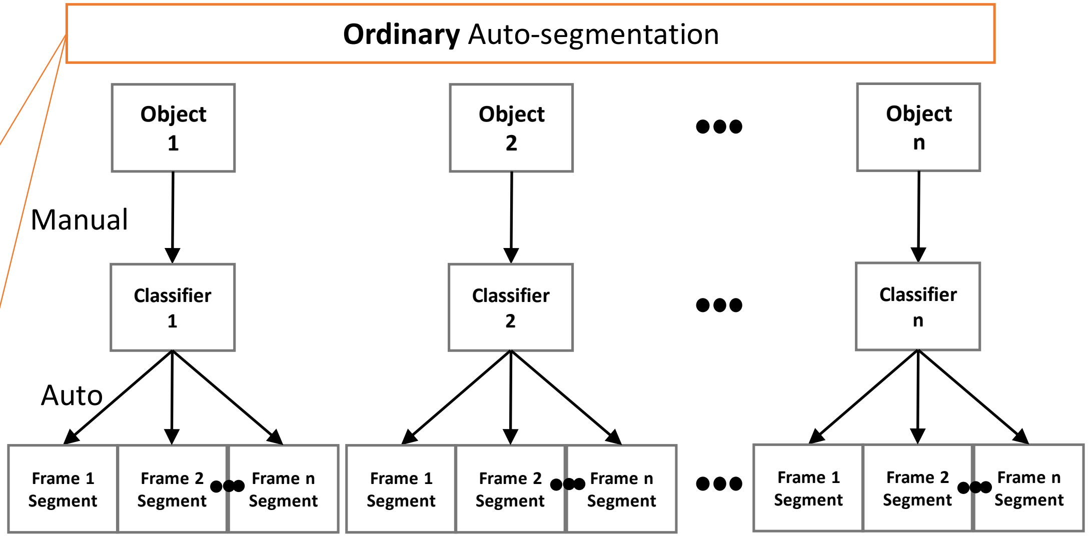


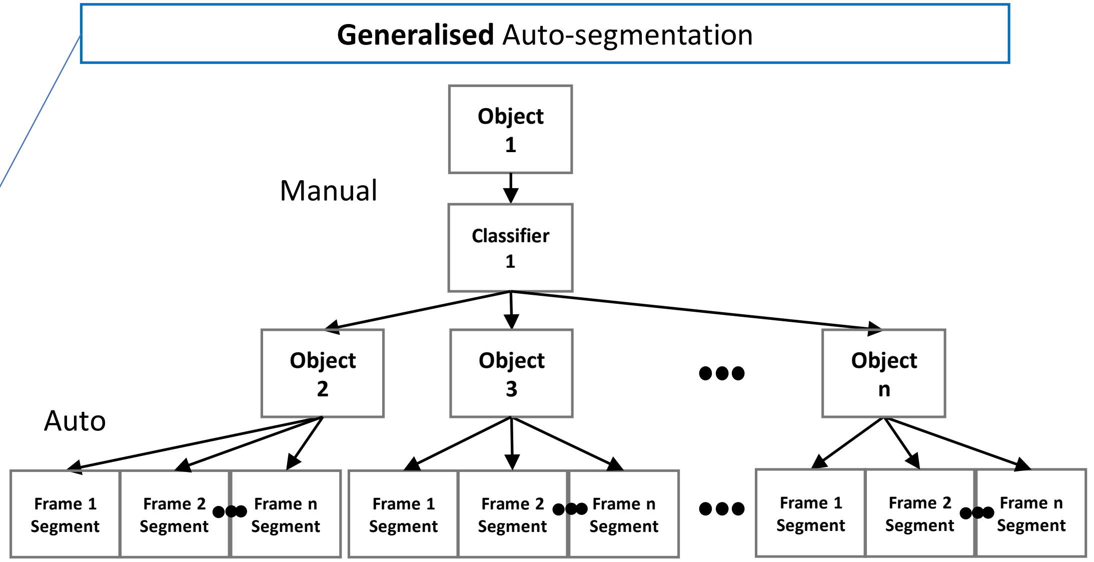
OBJECTIVES

1 Will **auto-segmentation** by TWS on ultrasound images give comparable results to **manual** segmentation?

2 Will different Machine Learning models **(Random Forest vs Naïve Bayesian)** affect segmentation results?

3 Will the application of **generalised model** give comparable results to the ordinary approach of training models on each dataset **individually**?



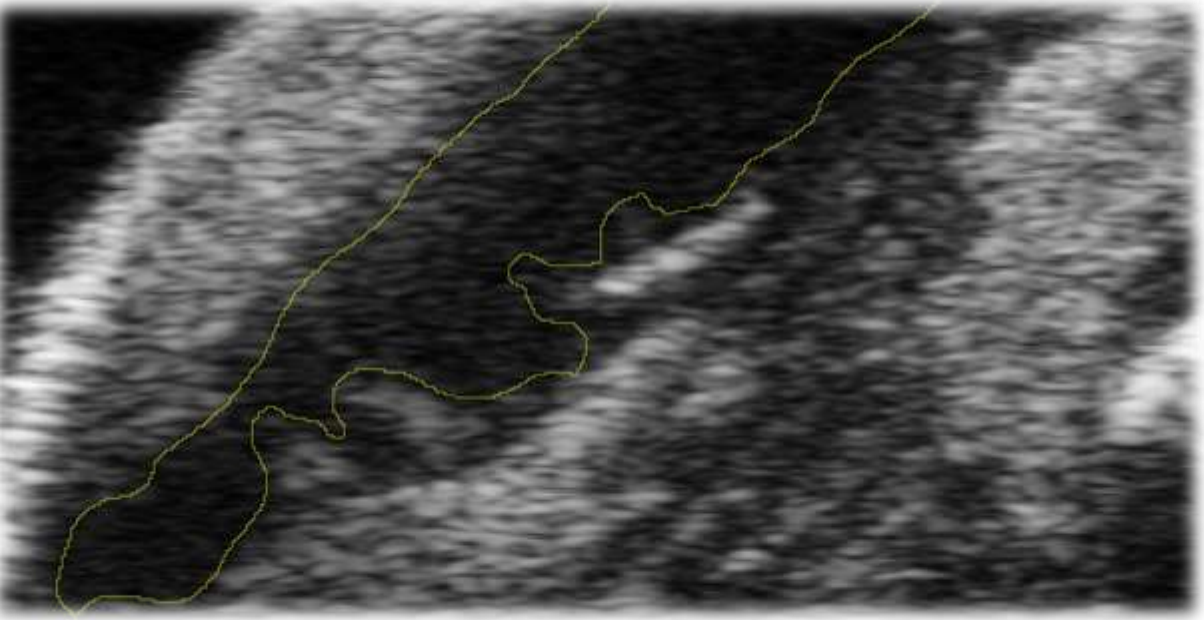


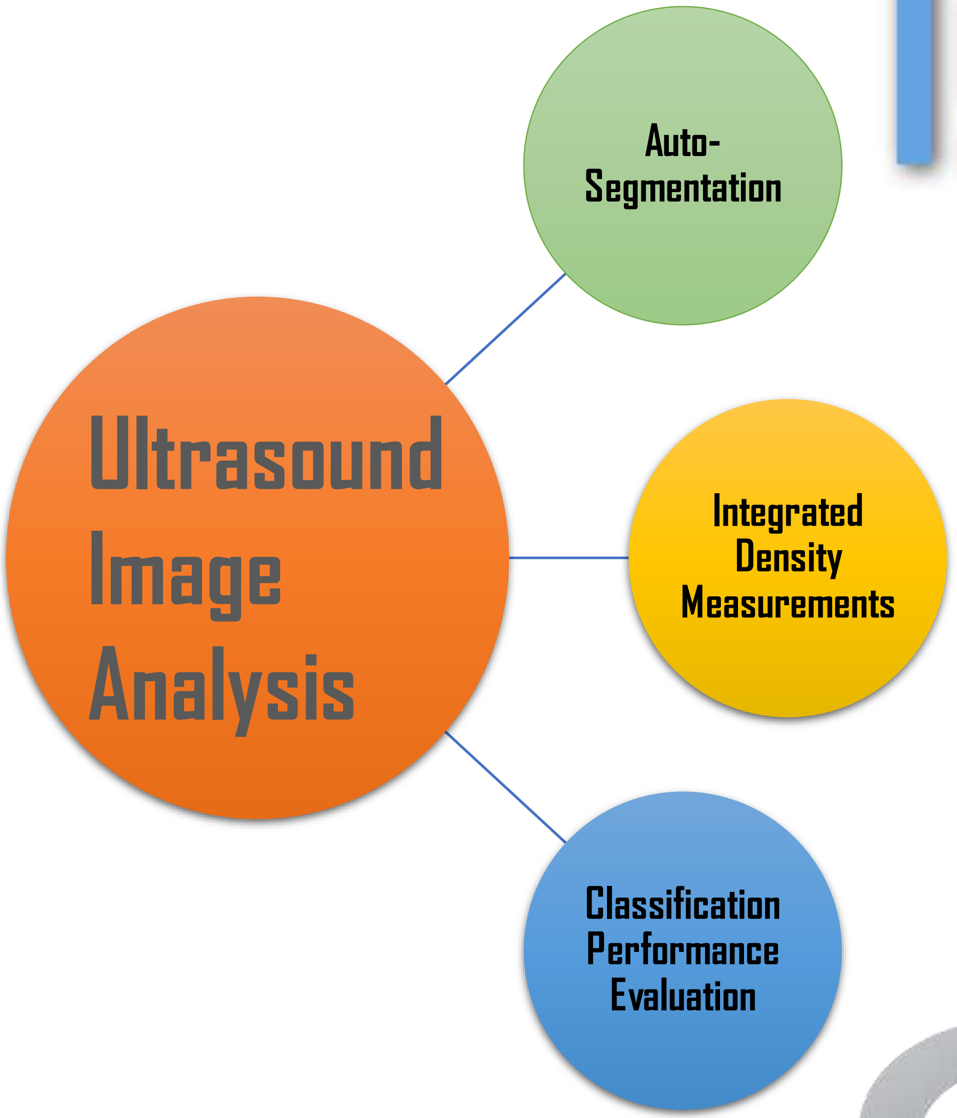
RESEARCH SCOPE

Raw Ultrasound Image Datasets: **Walker et.al. (2020)**

- Nanosensor Engineering Lab (NEL)
- Department of Chem.Eng., Monash Uni

- pH-responsive nanoparticle-injected lab mouse
- 4 image stacks taken at 0, 5, 10 and 15 mins
- 100 frames per image stack





METHODLOGY: Experimental Design

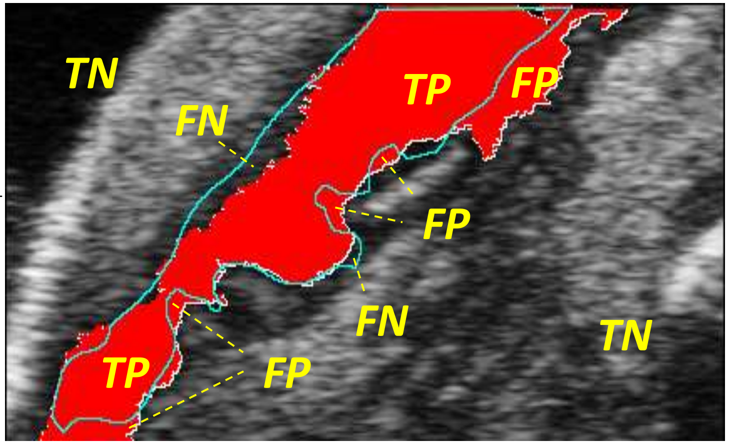
***Control Set**: Manual Segmentation Results*

Experimental Design

Objectives	Methodology
Manual *vs* **Auto-segmentation**	Comparison of Integrated Density (APE)
Random Forest *vs* **Naïve Bayesian**	➢ Binary Classification Features ➢ Receiver Operating Characteristics (ROC) Analysis
Generalised *vs* **Ordinary Approach**	

Absolute Percentage Error (APE%)

$$= \left| \frac{AI\ IntDen\ -\ Manual\ IntDen}{Manual\ IntDen} \right| \times 100\%$$



Binary Features:
- *True Positive (TP)*
- *True Negative (TN)*
- *False Positive (FP)*
- *False Negative (FN)*

TWS Segmentation

Manual Segmentation

METHODLOGY: Statistical Analysis

Receiver Operating Characteristic (ROC) Analysis

Binary Performance Metrics

Sensitivity $\frac{TP}{TP+FN}$

level of under-prediction of desired pixels

Specificity $\frac{TN}{TN+FP}$

Level of under-prediction of undesired pixels

Precision $\frac{TP}{TP+FP}$

level of over-prediction of desired pixels

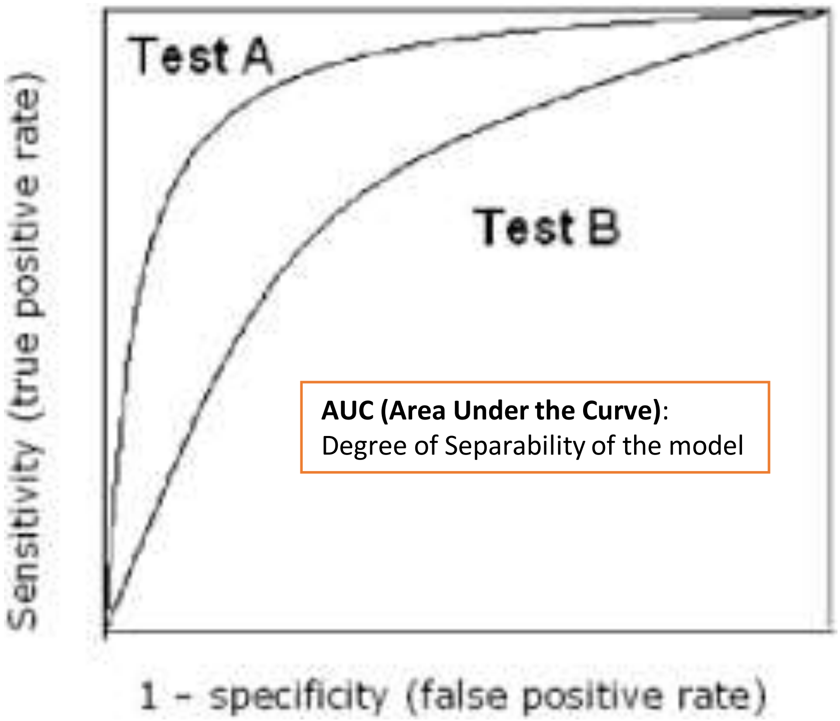
F-score $\frac{1}{\frac{0.5}{Prec}+\frac{0.5}{Sens}}$

Harmonic mean of sens. and prec.
Overall Performance

$'ROCR'$ *Package in R* (Sing, 2005)

ROCR v1.0-7

ROC Curve



AUC (Area Under the Curve):
Degree of Separability of the model

$'ggplot2'$ *Package in R* (Wickham, 2016)

ggplot2

Note: The R script code can be found at https://github.com/georgewang229/CHE4180-Project/blob/master/CHE4180_Project_DataAnalysis.R

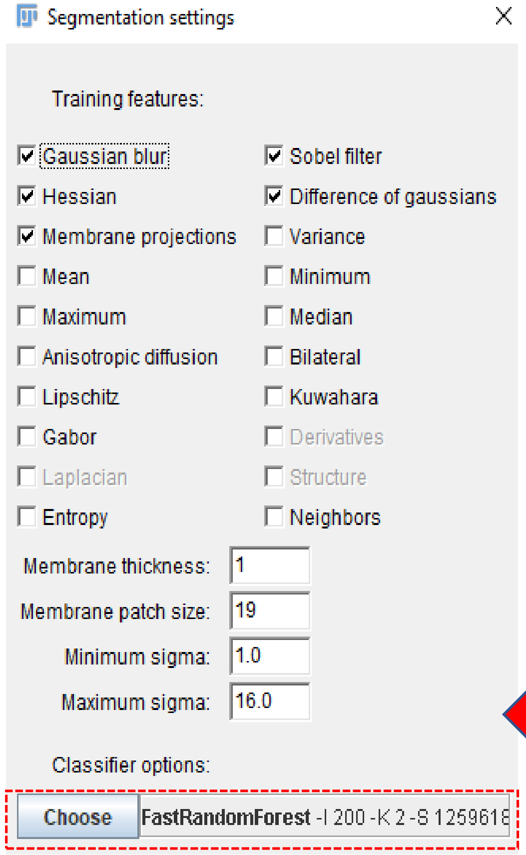
METHODLOGY: Image Segmentation

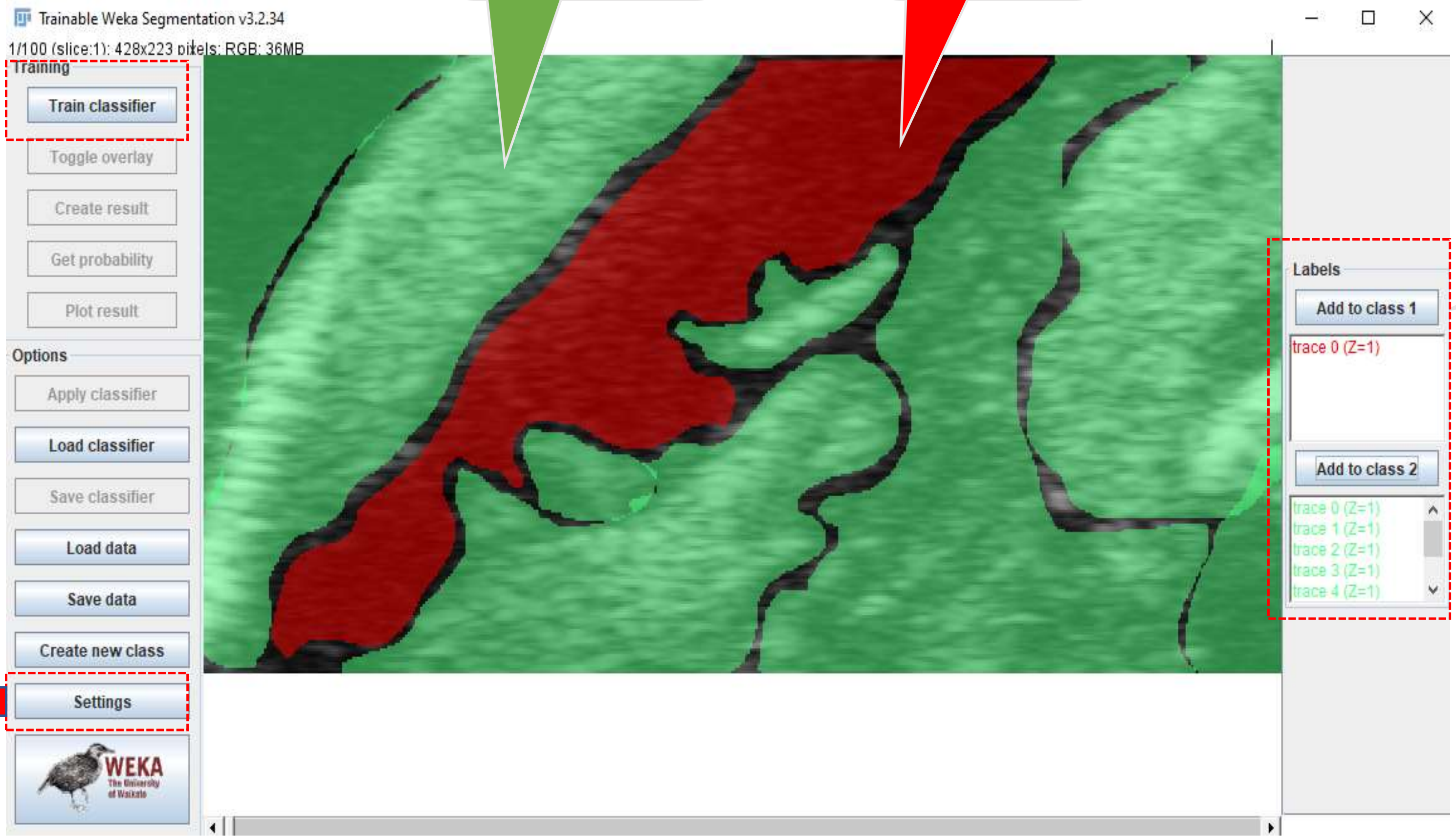
Auto-Segmentation

Undesired Region

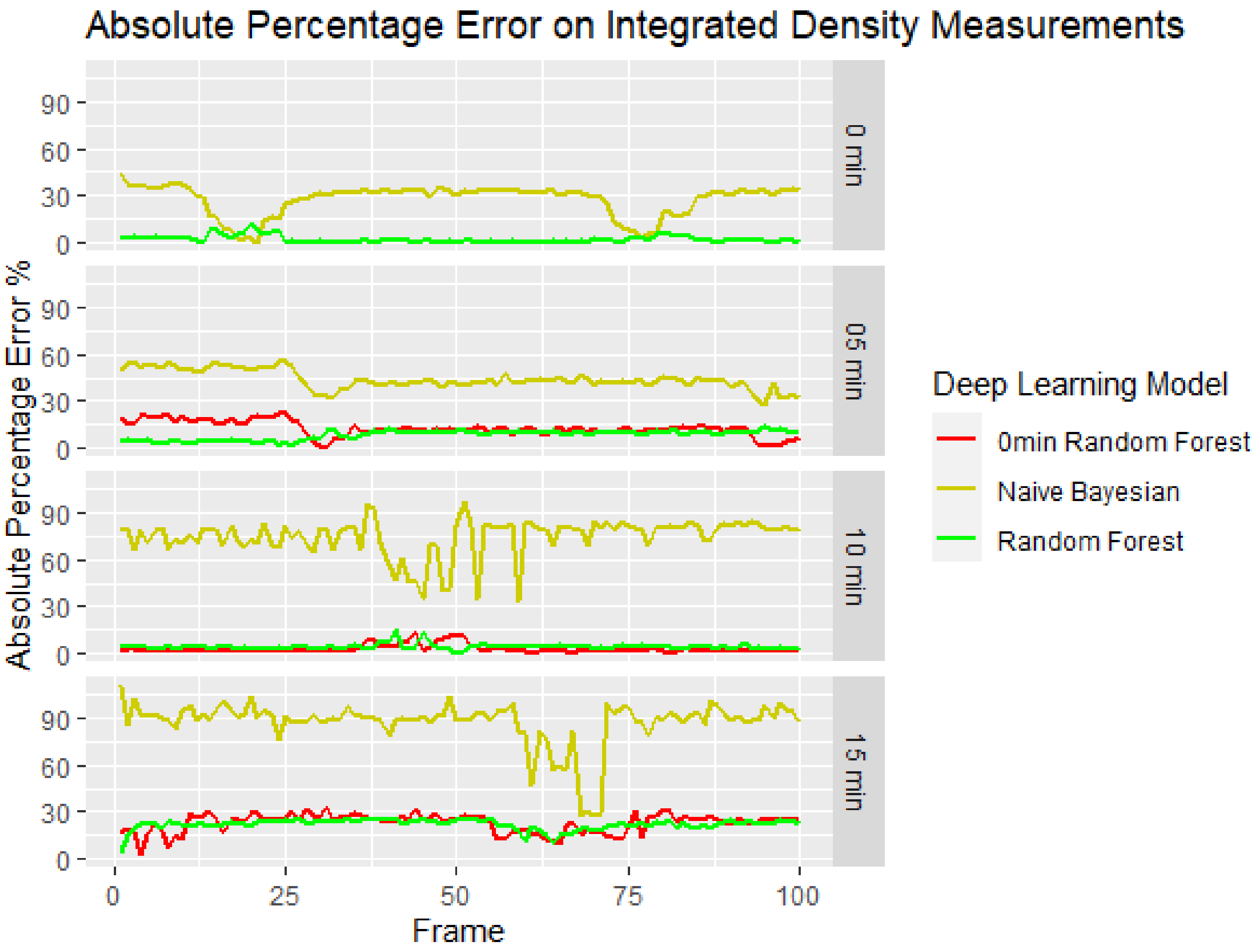
Desired Region

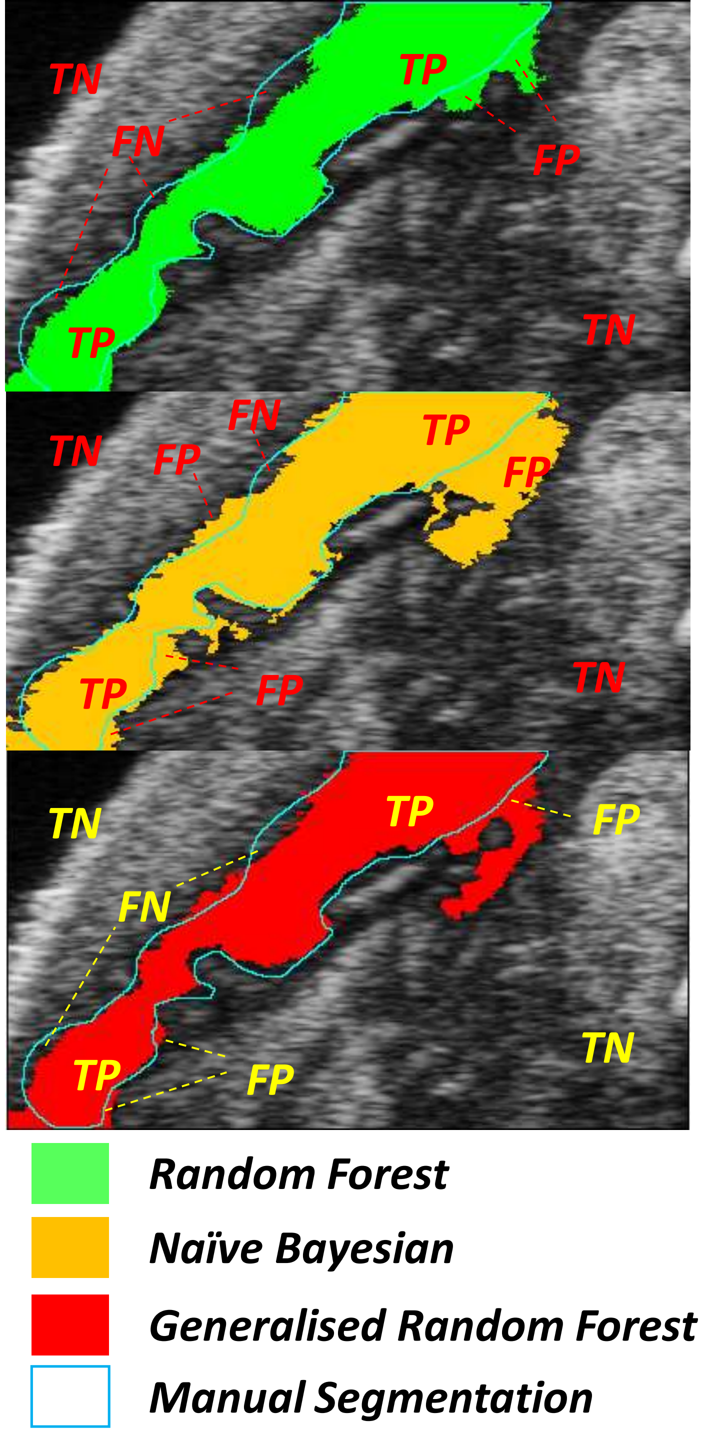
Binary Classification





Results & Discussion: Integrated Density





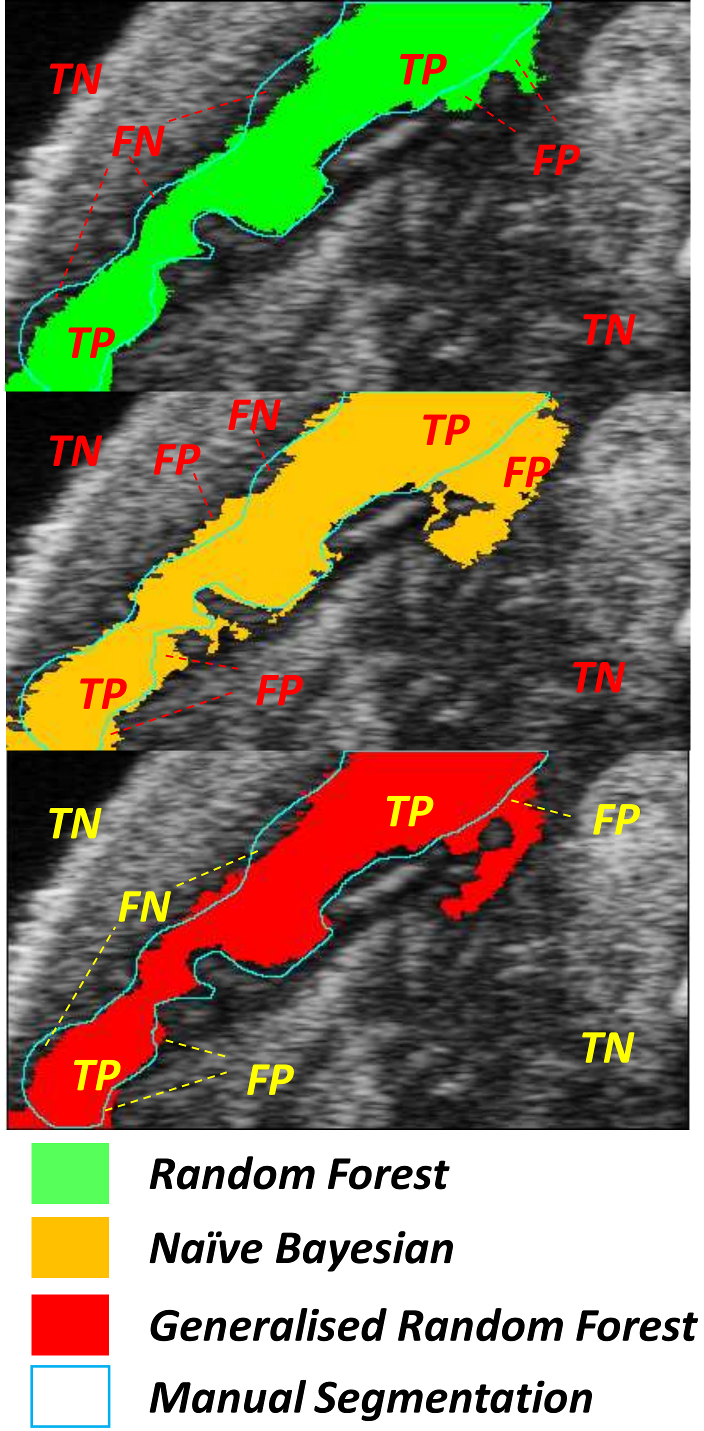
- Random Forest: Better flexibility and Good generalisation ability

Results & Discussion: Performance Metrics

Model	Average Sensitivity*	Average Specificity*	Average Precision*	Average F-score*
Random Forest	0.854 ± 0.045	0.972 ± 0.010	0.878 ± 0.036	0.865 ± 0.022
Naïve Bayesian	0.948 ± 0.016	0.988 ± 0.042	0.661 ± 0.079	0.776 ± 0.052
Generalised RF	0.886 ± 0.011	0.944 ± 0.016	0.791 ± 0.054	0.834 ± 0.031

Sensitivity: Level of under-prediction of desired region
Specificity: Level of under-prediction of un-desired region
Precision: Level of over-prediction of desired region
F-score: Overall Performance

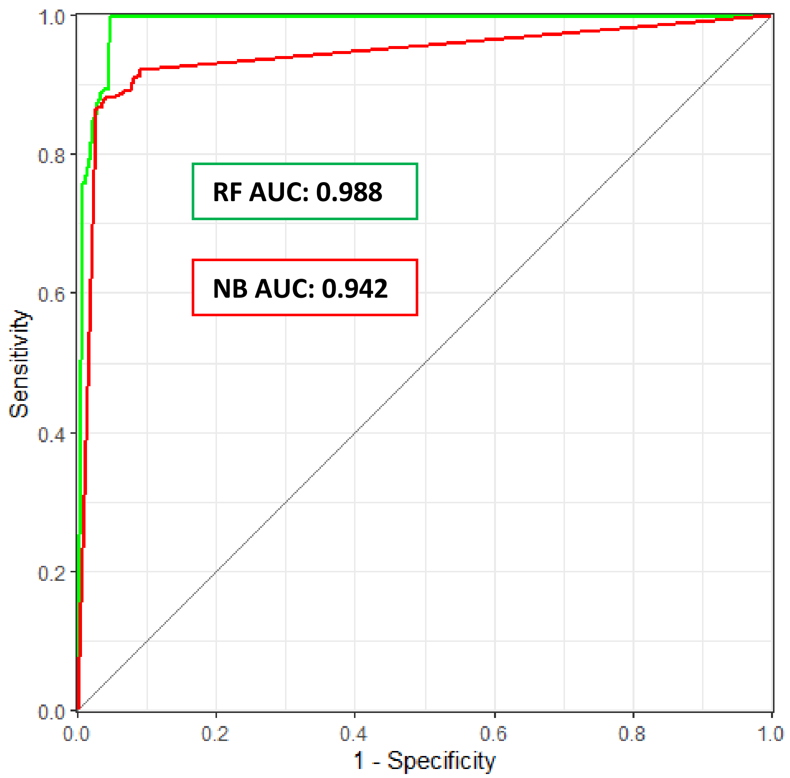
Random Forest: In general, Best **F-score**: 0.865 (Overall Performance)
→ Best Model in terms of Performance Metrics
Generalised RF: Good **F-score**: 0.834 (Overall Performance)
→ Satisfactory Performance

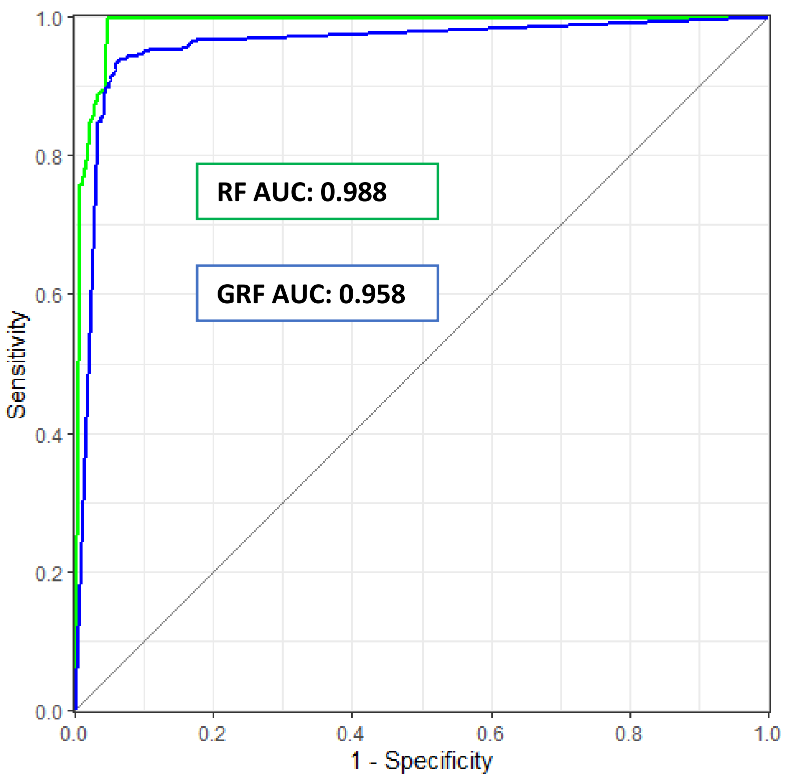


Results & Discussion: ROC curves & AUC

Model / Time	Random Forest	Naïve Bayesian	Generalised RF
AUC*	0.988	0.942	0.958

AUC (Area Under the Curve): Degree of Separability of the model





Random Forest
Naïve Bayesian
Generalised RF

- Random Forest: higher **Area Under Curve** (0.988) than NB (0.942) → Better Model in terms of AUC
- Generalised RF: Slightly lower **Area Under Curve** (0.958) than RF (0.988) → Satisfactory Performance

LIMITATION & FUTURE PRESPECTIVE

Technological Limitation

Reproducibility

Poor Interpretability

3D Ultrasound Image

Limited Sample Images

Better algorithms to be developed

CONCLUSION

- TWS gives satisfactory auto-segmentation results
- Random Forest >>> Naïve Bayesian
- Generalised Model is viable, but less reliable

ACKOWNLEDGEMENT

We would like to acknowledge **Dr. Simon Robert Corrie** and **Dr. Julia Ann-Therese Walker** from the Nanosensor Engineering Lab (NEL) in the Department of Chemical Engineering at Monash University for providing experimental data and assistance throughout the project.

THANK YOU!

George and Mark

REFERENCES

I. Arganda-Carreras et al., "Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification", *Bioinformatics*, vol. 33, no. 15, pp. 2424-2426, 2017. Available: 10.1093/bioinformatics/btx180.

J. Walker, X. Wang, K. Peter, K. Kempe and S. Corrie, "Dynamic Solid-State Ultrasound Contrast Agent for Monitoring pH Fluctuations In Vivo", *ACS Sensors*, vol. 5, no. 4, pp. 1190-1197, 2020. Available: 10.1021/acssensors.0c00245.

L. Breiman, "Random Forests", *Machine Learning*, vol. 45, no. 3, pp. 261-277, 2001. Available: 10.1023/a:1017934522171.

R. Yager, "An extension of the naive Bayesian classifier", *Information Sciences*, vol. 176, no. 5, pp. 577-588, 2006. Available: 10.1016/j.ins.2004.12.006.

N. Vyas, R. Sammons, O. Addison, H. Dehghani and A. Walmsley, "A quantitative method to measure biofilm removal efficiency from complex biomaterial surfaces using SEM and image analysis", *Scientific Reports*, vol. 6, no. 1, 2016. Available: 10.1038/srep32694.

T. Sing, O. Sander, N. Beerenwinkel and T. Lengauer, "ROCR: visualizing classifier performance in R", *Bioinformatics*, vol. 21, no. 20, pp. 3940-3941, 2005. Available: 10.1093/bioinformatics/bti623.

Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

S. Narkhede, "Understanding AUC - ROC Curve", *Medium*, 2018. [Online]. Available: https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5.