A generalised framework for detailed classification of swimming paths in the Morris Water Maze

Avgoustinos Vouros

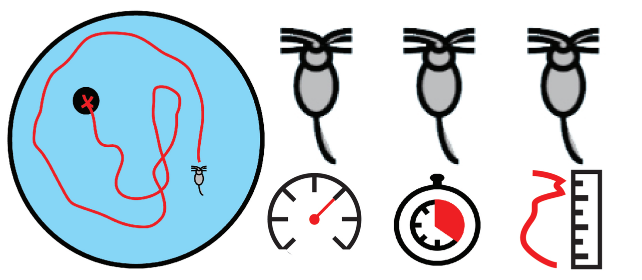
The University of Sheffield, Department of Computer Science, Faculty of Engineering

avouros1@sheffield.ac.uk

1. Introduction

The Morris Water Maze (MWM) is one of the most commonly used tasks in behavioural neuroscience for studying spatial learning and memory. In this experiment a rodent is placed inside an arena full with opaque water and is tasked to find a hidden platform. After a number of trials the rodent should find the platform in less time since it has 'learnt' its location. Many studies have examined how rodents behave inside the arena in order to discover general principles about learning that can be applied to other species, including humans.

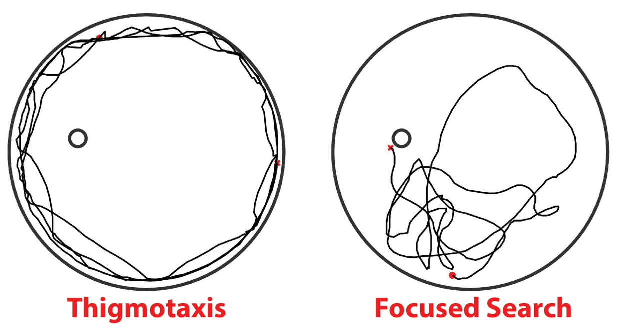
Performance Measurements (speed, latency, path length)



Pros:
• Easy to use.
• Easy to implement.

Cons:
• Quantification issues.

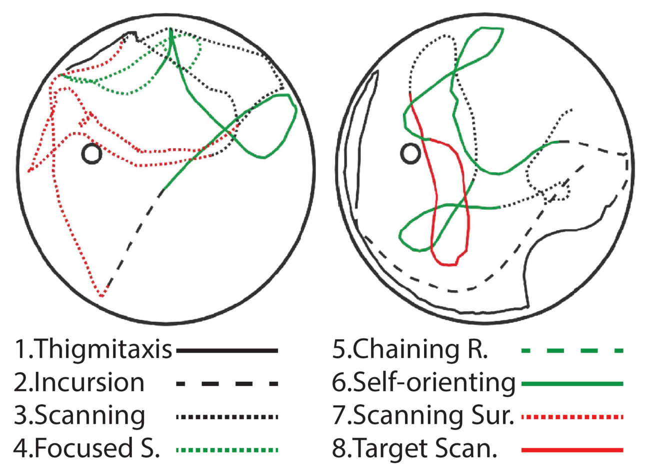
Full Trajectories Classification



Pros:
• Ability to quantify results.
• Informative about the actual animal behaviour.

Cons:
• Complex.
• Subjective.
• Requires machine learning knowledge.

Detailed Trajectories Classification (Gehring et. al.)



Pros:
• Ability to quantify results.
• More informative about the actual animal behaviour.
• Ability to detect more differences between animal groups.

Cons:
• More complex.
• Prone to error.
• Somewhat subjective.
• Requires machine learning knowledge.

2. Objectives

Following the previous work of Gehring et. al.[1] this particular work is focused on creating a more generic classification framework capable of automatic detailed classification of animal swimming paths inside the Morris Water Maze.

Major issues:

- Trajectory Segmentation: manual, length of each segment as well as an overlap percentage needs to be specified.
- Classification: manual, requires tuning and a number of data to be labelled.
- Analysis Conclusion: difficult to define a ground truth about the specific dataset as the classification of each segmentation can lead to entirely different results.

Solutions:

- Trajectory Segmentation: manual, but based on the arena dimensions.
- Classification: boosted with the use of majority voting which nullifies the need for manual tuning.
- Analysis Conclusion: consistent results among different segmentations.

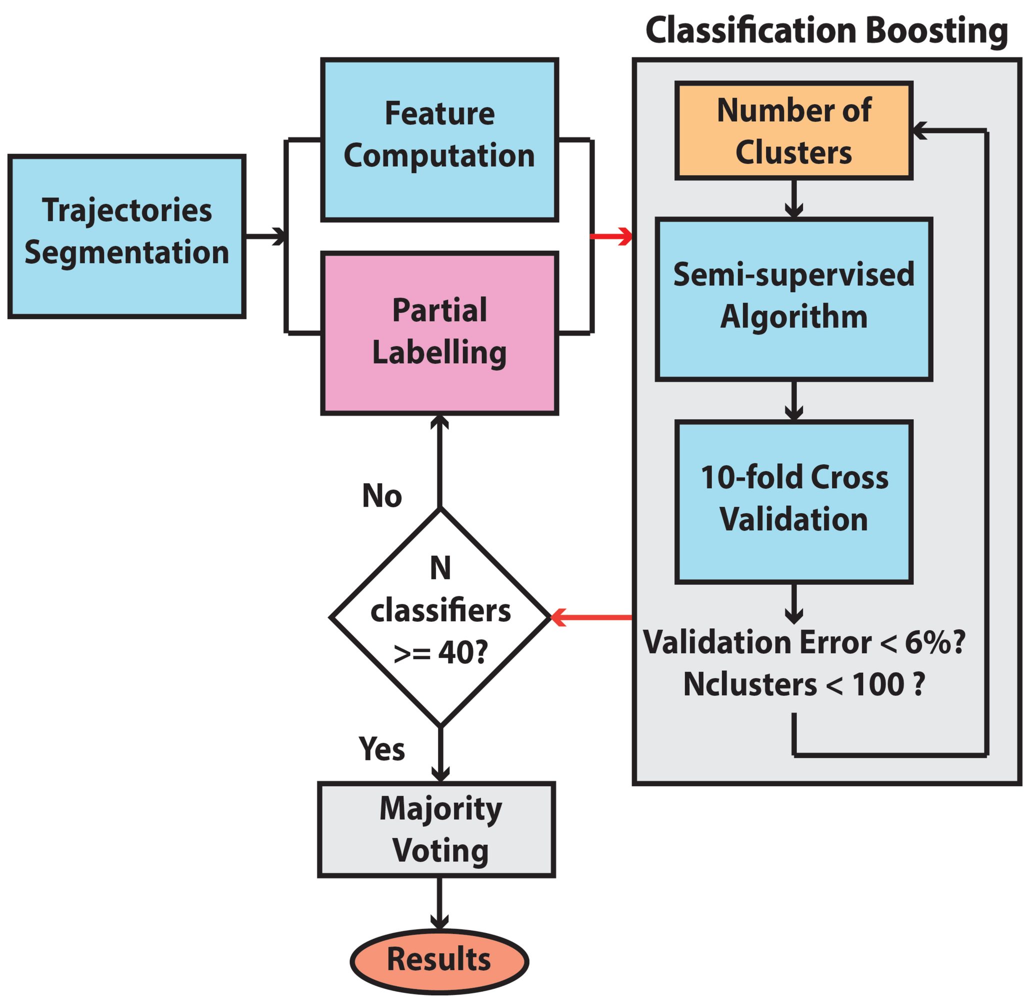
3. Methodology

Classification boosting:

- Classification problem too complex.
- Single algorithmic classification solutions are unable to achieve high performance.
- An ensemble of weak classifiers can be used to form a strong classification by combining each individual's opinion [2].

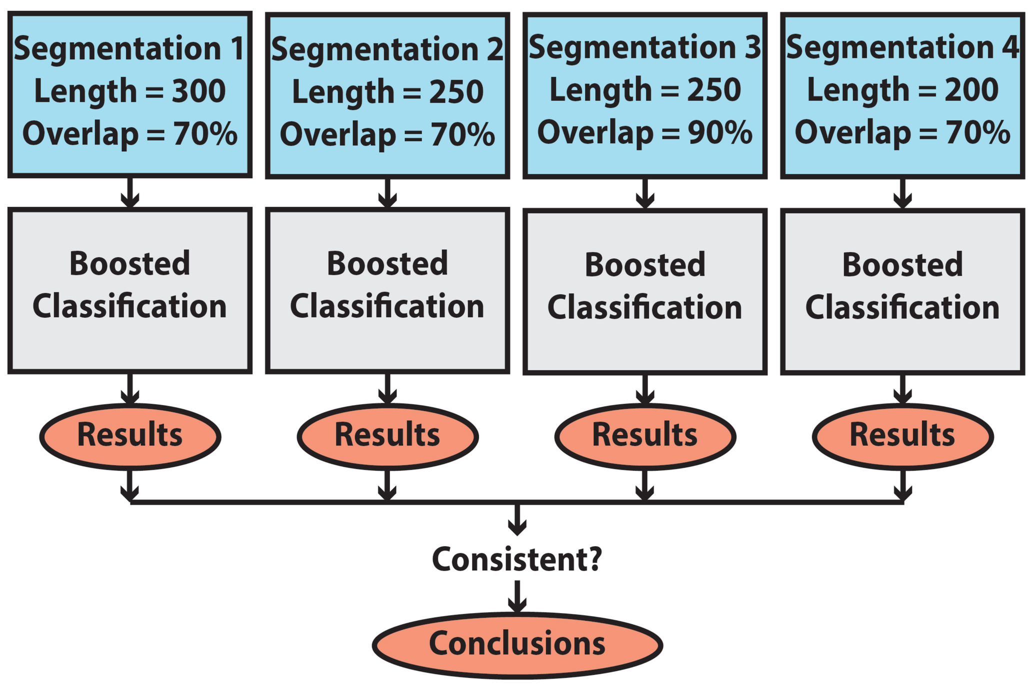
Segmentation and Labelling:

- Segment length: 2R to 2.5R. Lower will create segments too small to be assigned to a class; larger will cause the segments to fall under multiple classes.
- Number of labels should be around 7-10% of the data.



4. Validation

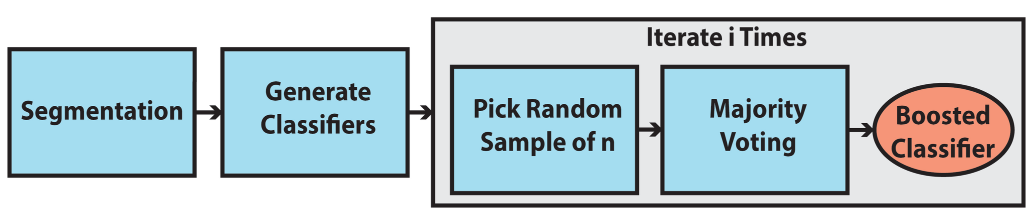
Robustness on segmentation tuning

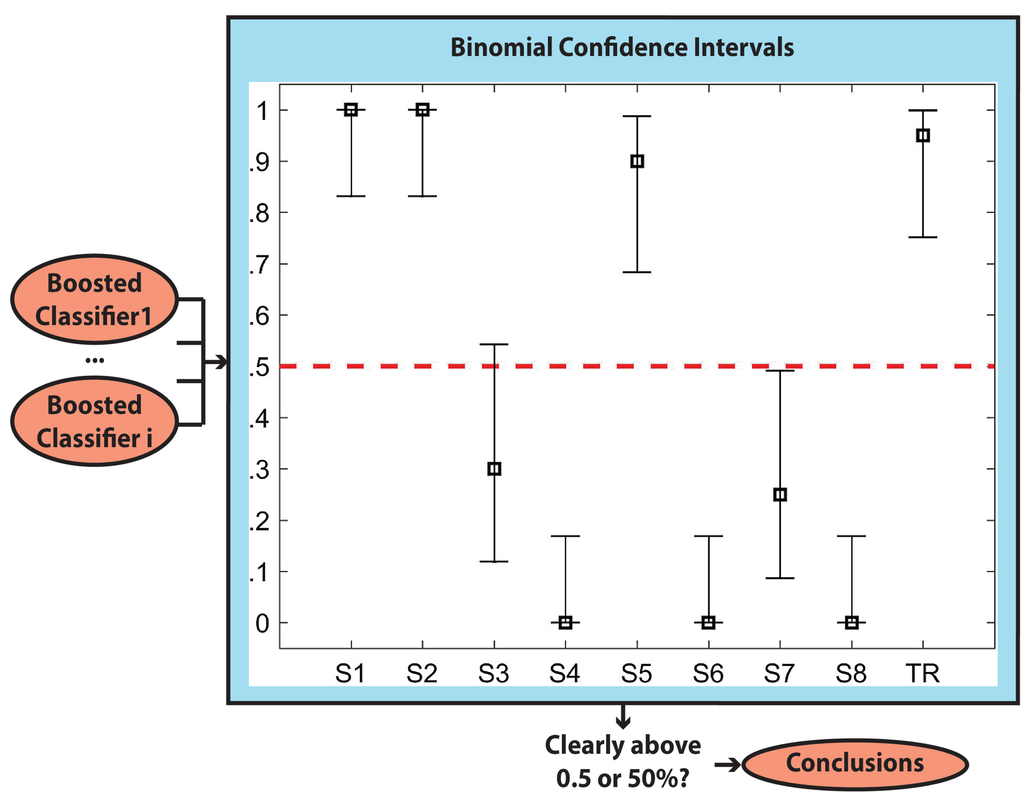


What do the "Results" represent?

Given 2 animal groups (stress and control) and 8 possible strategies we investigate if there are any significant differences between them. This is accomplished with the use of the non-parametric Friedman test as the depending variable being measured is ordinal (2 equal groups perform the same number of trials). Thus the Friedman test for each strategy will give a yes or no answer to the null-hypothesis: "There aren't any significant differences when comparing the stress group with the control group on strategy A".

Confidence on forming the conclusion





Why the confidence intervals on boosted classifiers and not the simple ones?

After experimentation it is proved that even if only 5 classifiers are merged the classification results are greatly enhanced. Nevertheless, there are still prone to noise. Confidence intervals clearly above 0.5 or 50% ensures that the result is not due to chance.

5. Results and Discussion

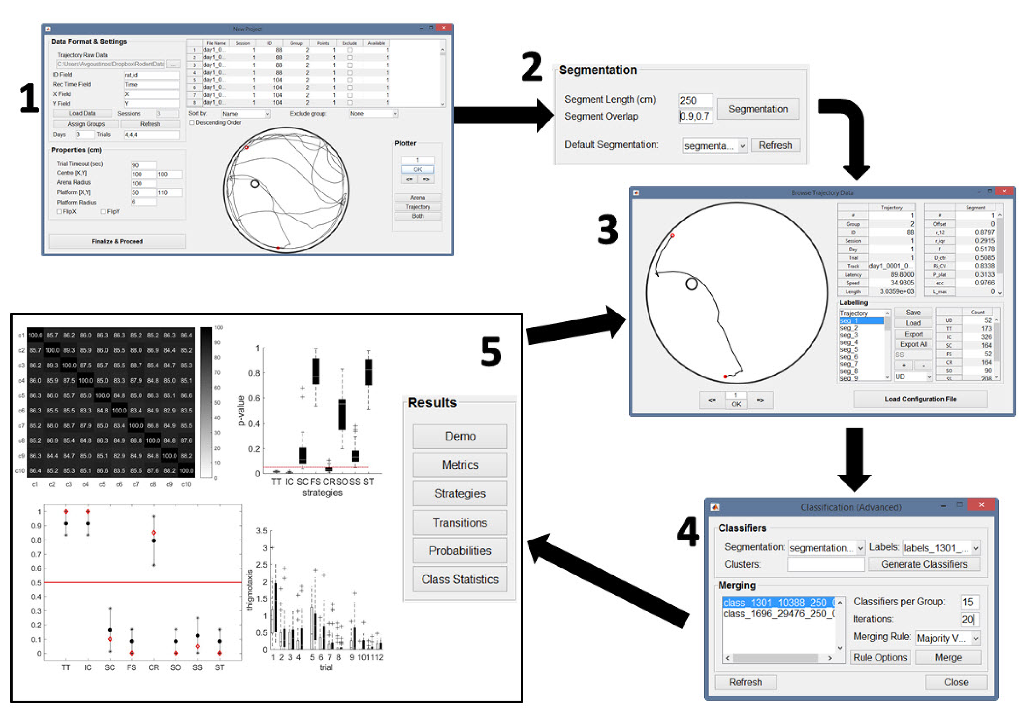
The classification of all the segmentations leads to the following conclusions:

- There are significant differences between the two animal groups on the strategies of thigmotaxis, incursion, chaining response as well as the number of transition between strategies.
- The classification of segmentation with segment length equal to 3R failed to capture the difference on the chaining response strategy. This is because shorter strategies are overshadowed by more common ones.

Compared with the results of the previous publication [1] the new framework:

- Achieves consistent classification results over different segmentation tunings.
- It does not require any machine learning knowledge in order to be used.
- It is far more robust to error. In the previous publication difference in the scanning strategy had also been reported due to a random occurrence.
- It is easy to be generalised and be applied to any MWM trajectory dataset.

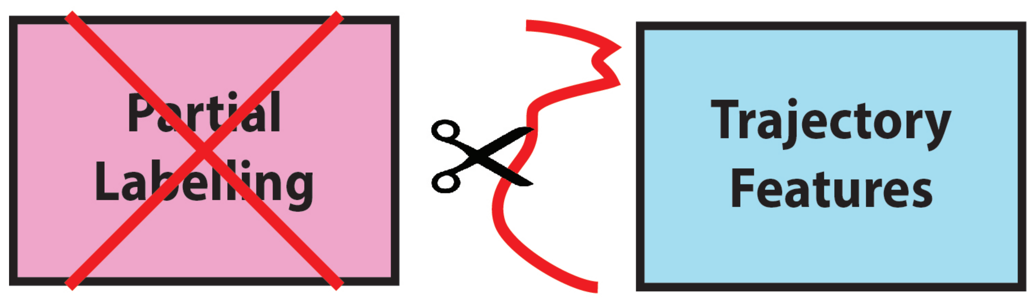
The RODA Software



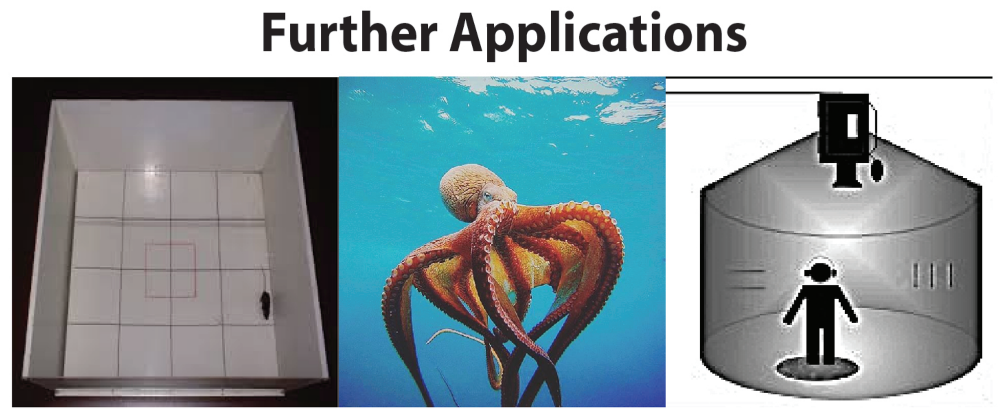
Currently in use by researchers of the Nencki Institute of Experimental Biology in Warsaw, Poland.

6. Future Work

Further Improvements



Further Applications



References

[1] T. Gehring, G. Luksys, C. Sandi and E. Vasilaki Detailed classification of swimming paths in the Morris Water Maze: multiple strategies within one trial. *Scientific reports*, 5, 2015.

[2] A. Jurek, Y. Bi, S. Wu, C. Nugent Classification by cluster analysis: A new meta-learning based approach. *International Workshop on Multiple Classifier Systems*, 259–268, 2011.