

A NEW PARADIGM TO DO AND UNDERSTAND THE RACE ANALYSES IN SWIMMING: THE APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS

Marc Elipot^{1,2,3}, Jess Coronas^{1,3}, Ash Hall⁴, Brandon Victor⁴, Matthias Langer⁴, Stuart Morgan^{2,4}, Zhen He⁴ and Mark Osborne^{1,3}

Swimming Australia, Australia¹

Australian Institute of Sport, Australia²

Australian Analysis and Research Group for Optimisation in Swimming³

La Trobe University – Department of Computer Science and Information Technology, Australia⁴

This article aims to present the benefits of using a deep learning approach to perform race analyses during domestic and international swimming championships. The procedure currently used to perform these race analyses is mostly manual and requires important human resources to annotate the videos and produce the performance reports. Recent technological and scientific developments now allow using ultra high quality cameras (4K) and machine learning algorithms to automatise the detection of the key events and greatly improve the video processing. Such a process helps the collection of data with a higher accuracy, the deployment of a more flexible and reliable setup, the access to the more variables such as the swimmers' instantaneous position and velocity, and the redistribution of the human resources to more effective actions.

KEYWORDS: Convolutional neural networks, Machine learning, Race analysis, Swimming.

INTRODUCTION:

For the last 30 years, the data produced by the various competition race analyses systems have been crucial to provide significant information to the coaches and sport scientists, helping them to review and subsequently improve the athletes performance. However, the systems currently used have important limitations and the introduction of technologies based on convolutional neural networks seems to present promising alternatives.

RACE ANALYSIS PROTOCOLS AND CURRENT LIMITATIONS:

Since some early works done in the 1970's, most the international and national championships race analyses are performed using complex video cameras system. The initial protocols were based on the use of a set of fixed analogue cameras located in the stands and shooting multiple selected areas of the pool (Craig and Pendergast, 1979). These setups require a large number of cameras and a significant work force to process the data and obtain the results (Mason and Cossor, 2001).

In the early 2000's, technical improvements allowed developing setups based on the use of digital panning cameras and portable computers. An operator pans a camera so that he can follow a single swimmer (Roig and Lopez, 2018). These protocols allow the performance analyst to have a better insight to the swimmers racing strategies and skills (Roig and Lopez, 2018). However, three major limiting factors exist:

- 1- The use of a panning camera to record the performance of a single swimmer in the pool forces requires many operators to cover all the participants swimming the same race.
- 2- The cameras are most of the time calibrated following a basic procedure using the laneropes of the pool. This operation assumes that the laneropes are correctly assembled (consistent distance between every buoys; same number of buoys, ...) and remain perfectly steady. Furthermore, as the cameras are panning, the calibration has to be re-done every time the analyst wants to measure a key distance (Break out distance, ...).
- 3- Most of the video annotations required (stroke detection, breaths, ...) are manually identified by the analyst. This process can take time and the accuracy and

consistency are always a major concern.

Consequently, despite hardware technology evolving, these processes are still massively in use during both national and international meets.

NEW PROTOCOL:

More recently, some evolutions to perform the competition race analyses have been introduced (Elipot et al., 2010; Veiga and Roig, 2016), using automatic tracking technologies and more advanced calibrations. However, despite some great features, these systems were still limited by their use of the laneropes and/or by the obligation of putting multiple cameras above the pool, making it difficult to deploy, especially during international meets.

The last scientific and technological updates allow considering the new option that we develop in this current article. This option uses a single fixed UHD camera (3840×2160) with a wide-angle lens (8-16mm), a high specification latest generation laptop computer (Proc: i9-8950HK, RAM: 64Go, Graph card: GeForce RTX2080), a photogrammetric algorithm for the camera calibration and a deep learning algorithm for the automatic detection of the main key events and the swimmer's head tracking and allows:

- 1- Shooting the whole pool (up to 10 lanes long course) using a single camera and a single operator
- 2- Processing the videos with minimal human interactions
- 3- A high degree of accuracy, without relying on laneropes to calibrate the camera
- 4- Accessing new additional data such as the instantaneous position and velocity of the swimmer during the entire race.

CAMERA CALIBRATION ALGORITHM:

The camera calibration algorithm proposed in this new approach is a photogrammetric algorithm (pinhole model), combined with a non-linear iterative estimation. To perform the calibration, the analyst doesn't have to use the lanerope buoys. The system only requires the biomechanist to mark the 4 corners of the pool or of the specific zone of interest (figure 1). Using these coordinates, the algorithm computes the camera coefficients (internal and external) allowing extracting the locations of the tracked points in the 2-dimensional pool environment. However, whilst the calibration algorithm only returns the swimmers' position along the x and y axis, it also allows discriminating lateral from vertical movements (so that vertical movements in breaststroke and butterfly are not seen as lateral movements).

MACHINE LEARNING TRACKING ALGORITHM:

One of the key aspects of this new protocol is the machine learning algorithm allowing the automatic detection of the swimmer's head and of the main key events (strokes, breakouts, breaths, ...). The machine learning algorithm proposed in this work is a two steps convolutional neural network (CNN) (Girshick, 2015; Bewley et al., 2016). The first phase of the algorithm aims to produce small crops approximately around the swimmer's head using the initial UHD video frames and through a set of multiple convolutions. The second phase of the algorithm consists in applying to the initial crops a second CNN to refine the head detection and identify its centre. At the same time, the crops are also used to perform the key event detection, such as the stroke detection, through another CNN module. The frame-wise predictions of the head location are then formed into continuous tracks associated with a single swimmer trajectory in the pool (greedy algorithm).

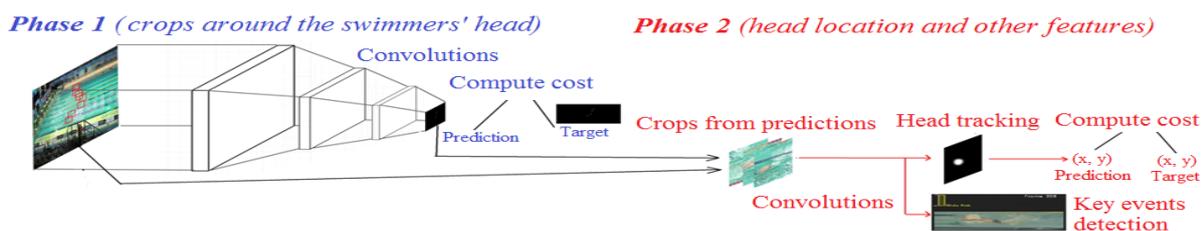


Figure 2: Machine learning algorithm main architecture.

TRAINING AND VALIDATION:

For this algorithm to perform with a high level of accuracy and reliability, the different CNN models were trained using video clips collected in many different pools, for all the strokes and with different types of populations. These clips were all manually annotated and use as an input to train the CNNs. A second set of videos was also collected (3 pools, approximately 50 000 frames, different conditions, ...) to verify the accuracy and reliability of this new protocol. Initial results clearly show that the data processed using the machine learning algorithm and a photogrammetric calibration display: 1) the median error for the head detection is less than 3cm and over 90% of the head detection were within 9cm (on the validation set); 2) On the training database, the stroke detection is performed within 1 frame and within 2 frames on the validation database. To our knowledge, no study fully tested the accuracy of the other systems measuring the instantaneous position of the swimmer's head. Veiga and Roig (2016), using the InThePool 2.0 system, only indicated a root mean square error for the head position calculation of 5cm (using a manual digitising of the lane rope buoys as a calibration reference). Elipot et al. (2010) also reported a similar error of 3 to 8cm for the head location.

A separate validation exercise was also conducted to verify the validity of the data measured with this new method. For this exercise, 32 races (8 per stroke, 4 men and 4 women, and recorded during 2 different international meets) were processed using both traditional panning camera protocol and the new deep learning method. Table 1 displays the average absolute differences for every analysed parameter. Results clearly show that the data collected with both methods are very similar and don't show any difference between the 4 strokes. No differences were either observed between the men and women, between the two different pools and between the different sections of the pool. To our knowledge, no other studies propose a complete accuracy and reliability analysis of any panning camera system.

The main two drawbacks of this current system are related to: 1- the portability: the system requires using a rather large professional 4K camera attached to a high and heavy tripod to ensure the best stability. The system also works better when being plugged to a power source. 2- The qualitative feedback: The system's accuracy is not impacted by the camera location in the stand. However its position determines the quality of the videos given to the coaches. However the benefits of this new system largely overcome its inconveniences and will be deployed to perform the race analyses during the next World Championships for the Australian National Team.

Table 1: Comparison between a traditional panning camera system and the deep-learning protocol

Differences in	Butterfly (8 races)	Backstroke (8 races)	Breaststroke (8 races)	Freestyle (8 races)	Average
Start Time (start signal to 15m, in s)	0.16	0.15	0.17	0.10	0.14
TT (5in to 10 out, in s)	0.06	0.09	0.07	0.05	0.07
Average stroke rate 15-25m (cycle/min)	0.81	0.25	0.64	0.31	0.50
Average velocity 15-25m (m/s)	0.02	0.01	0.04	0.02	0.02
Average stroke rate 35-45m (cycle/min)	0.80	0.24	0.33	0.60	0.49
Average velocity 35-45m (m/s)	0.05	0.01	0.02	0.02	0.02
Average stroke rate 65-75m (cycle/min)	0.79	0.32	0.55	0.49	0.54
Average velocity 65-75m (m/s)	0.04	0.02	0.02	0.03	0.03
Average stroke rate 75-85m (cycle/min)	0.43	0.34	0.47	0.50	0.43
Average velocity 75-85m (m/s)	0.03	0.01	0.04	0.02	0.02

All distances measured using the head position.

ELEMENTS OF ANALYSIS EXAMPLE:

Despite the usual data provided by any other race analysis systems, this new deep-learning based solution enables the analysts to access a complete new set of parameters to analyse and understand swimmers' performances.

Swimmer's trajectory: The analysis of the swimmer's trajectory is a great source of information for the analysts. Many swimmers, especially during long distance events, turn in the lane instead of swimming in the middle of it. The instantaneous head tracking allows calculating for example the real distance covered by the swimmer during the race.

Also, whilst the calibration used in this project is a 2D (x, y) transformation of the real world, it also allows estimating the swimmer's head elevation (z) at every stroke. The head position on the z-axis is returned expressed in an arbitrary but consistent unit (from 0 to 1 rather than in real world meters) allowing the analysts to estimate the head vertical movement (especially during the breaths in simultaneous strokes) and make suggestions on the overall swimmer's position in the water.

Stroke length: The system also provides the analyst the stroke length for every stroke (rather than an average value over a section of the race). Such data allows making conclusions on elements of the swimmer's technique such as the right-left arm symmetry or the stroke characteristics immediately after the breakout and approaching the wall.

Instantaneous velocity: The instantaneous velocity is also one of the new key parameters provided by the deep-learning based approach. It provides data to perform a complete analysis of the inter-cycle velocity variation. It also gives the analyst a much deeper and more specific insight on technical elements such as the effect of the breaths on the swimmer's velocity, the role of some arm-leg coordination or the swimmer's overall mechanical efficiency. The instantaneous velocity is also a critical piece of information to build accurate and efficient race models, especially in short distance events for which the average velocity over a rather large section of the race is not sensitive enough.

CONCLUSION:

This new system is based on the use of a single high quality camera capturing the entire pool within its field of view and thus all swimmers at the same time, an optimised photogrammetric camera calibration algorithm and the machine learning approach for the automatic detection of the main features of a swimming race. It enables the swimming analysts to access accurate racing data (without having to rely on the laneropes), new additional data not previously available to coaches (instantaneous velocity, stroke rate and length, trajectory analysis, ...), and significantly reduces the processing time providing more opportunities for the analysts to work with the coaches and athletes in interpreting rather than collecting the data and be more effective in their work.

REFERENCES

- Bewley, A., Zongyuan, G., Lionel, O., Ramos, F., Upcroft, B. (2016). Simple online and realtime tracking. In Proceedings of the IEEE International Conference on Image Processing (ICIP).
- Craig, A.B. and Pendergast, D.R. (1979). Relationships of stroke rate, distance per stroke and velocity in competitive swimming. *Medicine and Science in Sports and Exercise*, 11, 278-283.
- Elipot, M., Dietrich, G., Hellard, P., Houel, N. (2010). Cinalysis: A New Software for swimming race analysis. *Procedia Engineering*, 2(2), 3467.
- Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 1440-1148).
- Mason, B. and Cossor, J. (2000). What can we learn from competition analysis at the 1999 Pan Pacific Swimming Championships?, In R. Sanders and Y. Hong (eds.), Proceedings of XVIII Symposium on Biomechanics in Sports: Applied Program: Application of Biomechanical Study in Swimming (pp. 75–82). Hong Kong: International Society of Biomechanics in Sports
- Roig, A., Lopez, R. (2018). Segment coordination to ensure the efficacy of body levers in alternative strokes. In H. Takagi, Y. Ohgi, Y. Sengoku, and T. Gonjo (eds.), Proceedings of the XIIIth International Symposium on Biomechanics and Medicine in Swimming (pp. 108-114). University of Tsukuba, Masanobu Iseri.
- Veiga, S., Roig, A. (2016). Underwater and surface strategies of 200m world level swimmers. *Journal of Sports Sciences*, 34(8), 766-771.

ACKNOWLEDGEMENTS: The author would like to thank the Queensland Academy of Sport, the Victoria Institute of Sport and the New South Wales Institute of Sport, and more especially Mrs. Elaine Tor, Mr. Ryan Hodierne and Mr. Nick Smith for their help in collecting the data and information required for this work.