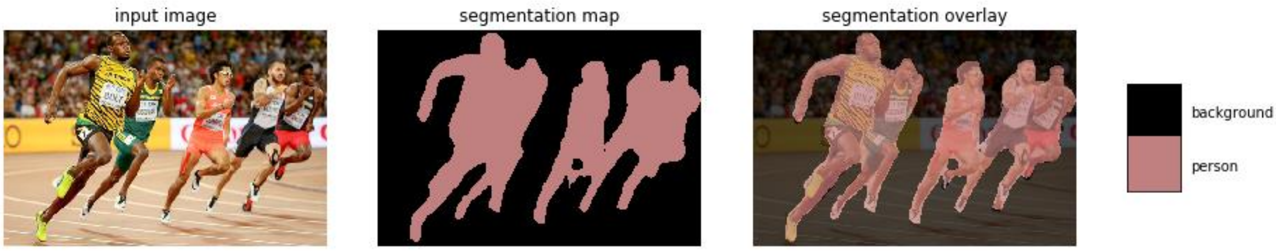
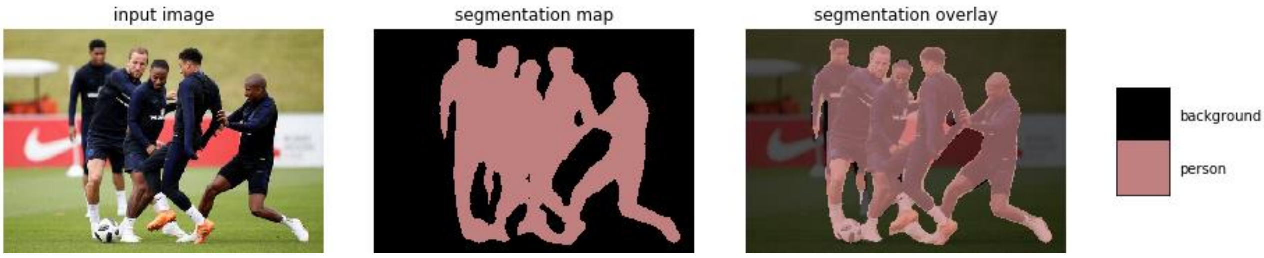
Recently I've been working on a project where we were looking for a markerless method that we could use for kinematic analysis. Eventually I found a good solution, and our first paper using this method is out now. But along the way I got to play with some really cool methods.

There are some very different ways of doing this, depending on the level of detail you are after. For example, here are some images I processed using the Deeplab segmentation method from tensorflow:



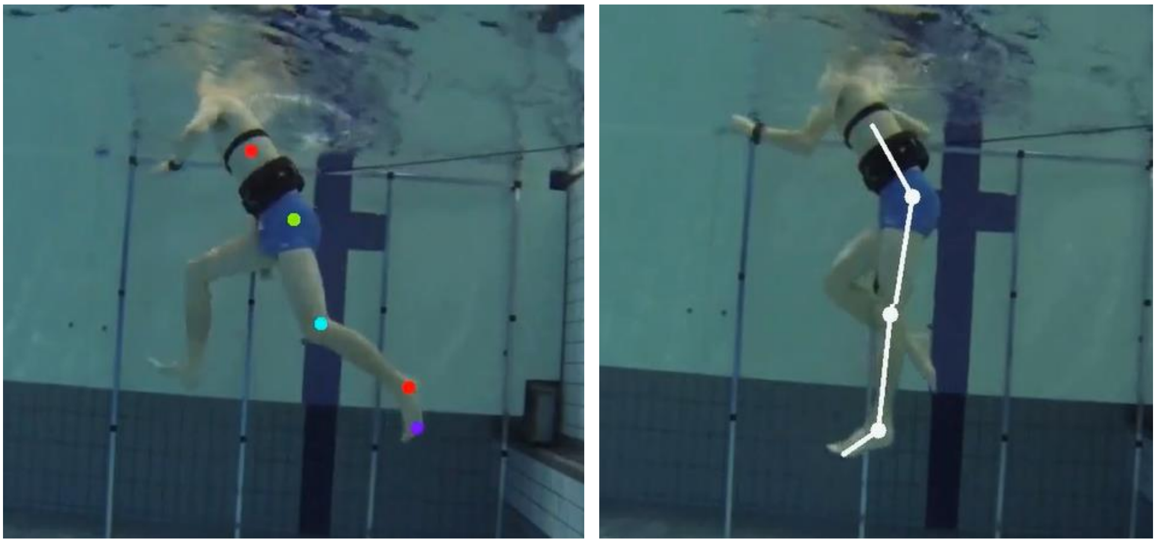


With this kind of 'segmentation' method, the aim is to first 'cut out' the object(s) of interest. From there we would need to a bit of extra processing, but in theory it is possible to calculate the angle between two parts of a segment. Clearly, this method is very good at what it does (I did many more tests than just these 2 images). The problem is that in the case of kinematic analysis, we need extremely high precision, and to me it seemed that this kind of technique probably wouldn't give me that, nor would the angle computation be straightforward.

A more promising example comes from the use of so-called 'pose estimation' methods, where the goal is more closely aligned to that of someone looking to measure human joint angles. Below are images processed with the method of Cao et al., which creates multiple stick figures so that you can follow several people at the same time, even in real-time (if you've got the hardware).

Again very promising, and it may be that this method (or another- there are loads of promising algorithms out there) could be modified for our purposes. However, at about the same time, I discovered the DeepLabCut method, which shows some astounding results in different animal species, so I decided to play with this method further.

Here's a very brief snapshot of what happened when I trained DeepLabCut on some data we had from deep water running:



On the left you see the markers which were placed by the trained deep neural network. On the right I have added a stick figure to show the alignment a bit better. Compared to a human labeler, we achieved a mean error of around 1cm, which is very promising. I am also confident we can improve upon this, because so far we have only tested it with default settings, and our data were collected with a GoPro rather than a state of the art camera.

Our first paper with this method is out now, and there will be plenty more from us in this space...