Progress since last week:

I've been studying deep learning, machine learning and reading papers on other object trackers trying to understand other ways to approaching the problem.

One particular paper:

**Learning Spatial-Aware Regressions for Visual Tracking**, they explain and study the pros and cons of KCF (Kernelized Correlation Filter) and CNN and combine the two methods to produce better results.

Specifically:

**KCF**, which is essentially **KRR** (Kernelized Ridge Regression) with cyclical rotated samples is:

1. Efficient at utilizing large numbers of negative samples during training.

2. Computationally expensive because it has to produce a kernel for each region of interest.

3. Take a region of interest as the input, assuming it as equally importance without concern for its relationship with other areas, which makes it very difficult to exploit the structural information of the target.

4. Produces unwanted boundary effects by the cyclically rotated samples.

**CNN** based trackers:

1. Have too many learning parameters.

2. Tends of overfit data.

To exploit the complimentary cons of these two object tracking methods, they show that the KRR model can be implemented as a CNN, making it solvable efficiently. Secondly, they propose a CNN where they focus on localized regions based on a spatially regularized filter kernel. They then combine the results from the two methods to locate the object.

**Link to paper**: <https://arxiv.org/pdf/1706.07457.pdf>

I haven't tested their code in the ECO paper yet, I'm hoping to actually modify code in the next week.