Perhaps the primary objective could be to demonstrate the smallest number of pixels on target, coupled with the associated 360 horizon scan time, for the most suitable machine learning approaches to be effective (CNN or other candidates) in our application space. Effective could mean scanning the 360 horizon as fast or faster than a human with similar or better ROC performance. Also include some effort to do technology forecasting/roadmapping to try and predict how this is expected to change in say the next 10 years. I very much like the idea of introducing the target pixel size as a variable when evaluating the performance of CNN or other approaches. As for the scanning idea, we might be able to use the panning platform that Sham purchased before but did not have the chance to use.

Forgot about that. I was thinking more conceptually, rather than actually doing 360 panning. If we consider a non rotating mast with fixed staring cameras, then the 360 scan time is that for one camera as the others can be done simultaneously. This can be compared to a human who can look at a spatially compressed representation (eg our two strip display modality) which will be fast but with poor ranges, or manually search through the scene sector by sector which is slower but better ranges. The metric we choose should include time and range (pixels on target) for the task of scanning the 360 horizon.

Because of limitation of training dataset, we might consider using the Unity platform developed by Richard to help generate the training sample for the CNN.

Worth considering. We also have some validated Cameosim scenes that we could use which are more realistic. Although since Richard determined that simulated imagery struggled to reach adequate functional accuracy for the tasks of recognition and identification, I think we need to use a real data set that is representative of our application. Since we don’t have a lot of suitable periscope imagery we need to look elsewhere or collect our own. We want imagery taken from near the surface with distant surface vessels, preferably with location data. Installing some cameras on the Armaments Jetty is one option.

Do you think there is a radiometric validity of target producing a certain signature while being scanned by a EO sensor in the manner as you described? If so, I am sure we can explore the use of CNN to learn this target traces.

Do you mean are there some specific signature traits ? Possible, particularly if we are talking hyperspectral EO.  However we do know that contrast and clutter varies significantly with atmosphere, observer-solar geometry and sea state.

One case for comparison maybe when raising a panoramic mast that also has a rotating high magnification sensor. A CNN machine may allow automatic scanning of the low magnification panoramic scene and training of the high magnification sensor onto areas of interest within a few seconds in the one exposure of the mast. Where a human may need to raise the mast once to capture the panoramic imagery, then lower the mast and look through the imagery, then raise the mast a second time to look at points of interest with the high magnification sensor. One exposure versus two is a big thing. Yes, we can certainly explore the concept of target handover from pano to rotating high magnification systems

 We may be able to leverage what is employed in the domain of Infrared Search and Track (IRST) systems employed on warships. These systems can detect missiles and aircraft with very low pixels on target. Are the cooled IR sensor technology the same between the IRST and submarine mast? A literature search re IRST detection and tracking techniques is a good idea.

The sensor technology will be similar although we don’t have the same space available for optics and stabilisation systems so there are additional constraints.

 The secondary objective could be to demonstrate the benefit of CNN’s in reducing operator workloads and improving response times for recognition/identification tasks in high contact density situations. Once again also looking at trends  10 years ahead. The extension to target recognition and identification is always worth considering.

Thank you for your thoughts.

Perhaps this may be just semantics as you pointed out. I approach this from the pixel resolution perspective with the old school Johnson criteria where target detection can be said to be achieved by a human observer with a minimum resolvable resolution of 1 +/- 0.25 line pair. This translates to 2 pixels on a digital camera. I felt that it will be a significant challenge for any computer vision techniques to achieve a reasonable detection ROC with only a minimum of 2x2 pixels of information. Of course if this not the primary requirement, and that we focus on early detection of targets (based on your definition below), it will be interesting to investigate CNN for target detection.

Hi Tele

Thanks for putting the proposal together. I’m still working my way through the references to improve my limited understanding yet before going further I did note the focus on recognition rather than detection.

Recognition *is* important for submarine operations, yet since this is a Sea1000 funded project we need to link it’s outputs to impacting the build of a new submarine. The success of automatic detection approaches for objects in panoramic imagery is linked to decisions regarding which masts we buy for Sea1000. Whereas automatic recognition, whilst still of great interest, does not strongly impact on decisions surrounding submarine building, so I may have some difficulty justifying the funding against SEA1000.

If CNN’s are not strong performers with detection of low observable objects, such as in our application, then we may need to think again.

Although it may just be semantics: A definition of detection; alerting the operator that there is an object of interest present, and localising it. Objects of interest include surface vessels, markers, manmade structures etc, but do not include birds, waves, clouds or land masses for example. Hence in this context, recognition is complicit in separating objects of interest from all objects.

Cheers,

steve