THALES

Enhancing long range Automatic Target recognition using spatial context

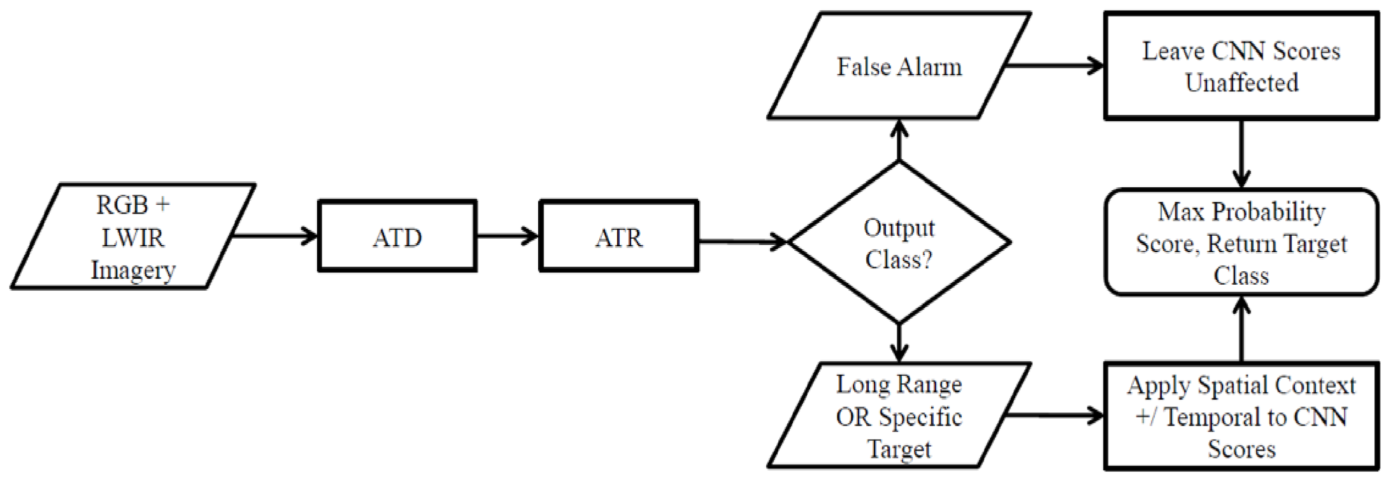
Rachael Abbott, Iain Rodger, Barry Connor and Neil Robertson

QUEEN'S UNIVERSITY BELFAST

Introduction

This poster presents a high-performing automatic target recognition system which can be used for long- range surveillance scenarios. The main novelty of our system is that it uses contextual information from RGB images to help classify targets in long range real world LWIR images. This contextual framework provides additional information of an object's surrounding environment, leading to a significant increase in long-range target recognition accuracy. This work will be of interest to the defence community as a high-performing automatic recognition system is a highly sought-after capability.

Overview of ATDR system



Candidate detections can be generated via an ATD process and then are fed to the trained CNN. If the maximum class score is a false alarm, do nothing. If the target is a long-range class, remove FA and long range scores from CNN vector. Re-weight using spatial context. If real object class returned, re-weight CNN scores using spatial context.

Contextual framework

The following components are needed to generate spatial context information:

1. An algorithm for semantic segmentation and spatial sampling feature
2. Prior scene knowledge
3. The probability for each object to exist given its surrounding context

1. Semantic segmentation and spatial sampling feature

- We use ***efficient graph based segmentation*** to perform region segmentation. ***Colour intensity, textural information and spatial context*** are extracted from the different regions after segmentation, then fed through the SVM to be assigned the correct labels.
- The Stanford dataset (Gould et al., 2009) is to compute the feature vectors which are used to train our SVM. The SVM can label five regions: sky, bush/tree/grass, road, water and building.
- **When an object is detection is made it is sampled above, below, left and right of the center of the bounding box at pixel locations [1, 20, 40, 100, 200].**

2. Prior scene knowledge

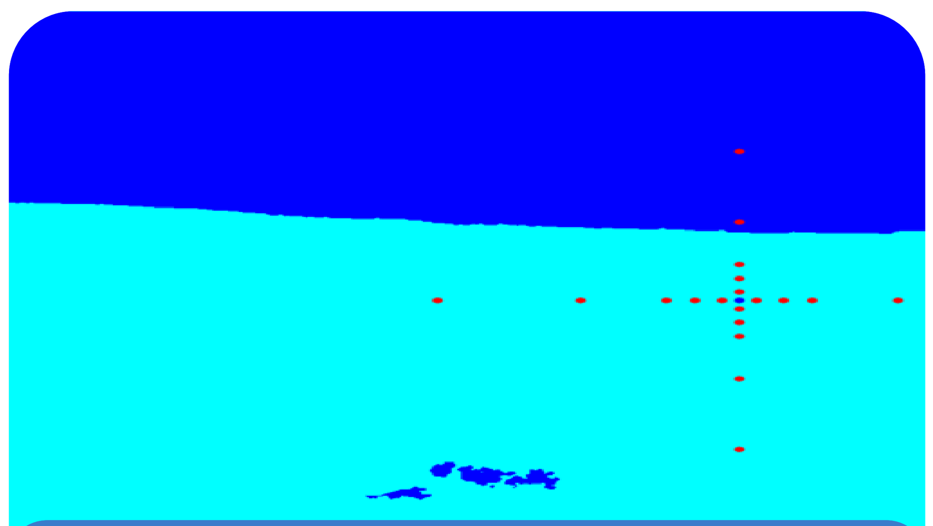
The Prior probabilities, $P_o(R_c|I_k)$, are found at the 20 locations for each object class using images from ImageNet. Twenty images of each object class are used.

3. Probability given context

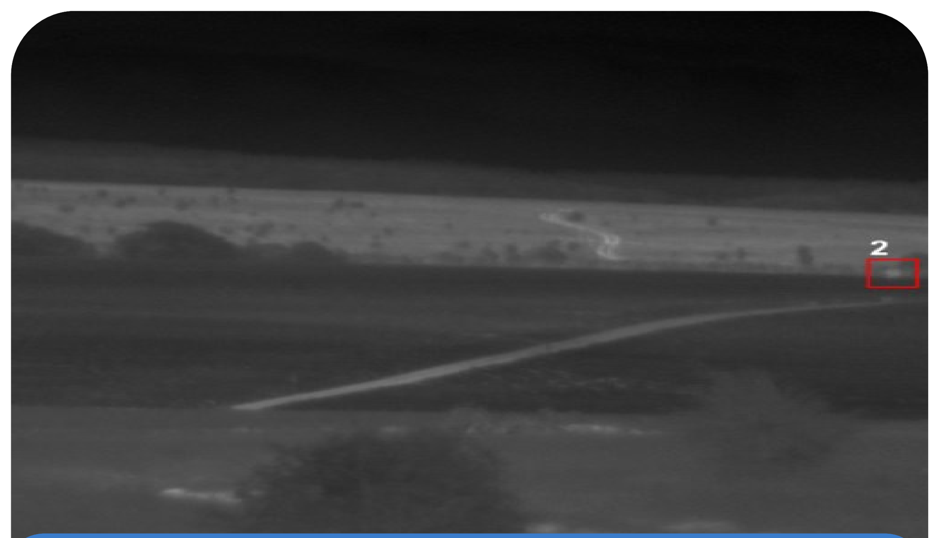
The prior probabilities, $P_o(R_c|I_k)$, of each of the 20 locations are summed up and divided by n (the total number of locations) to calculate the probability given context:

$$P(O|C) = \frac{1}{n}\sum P_o(R_c|I_k)$$

Input Scene, AD + CNN
1) AD Detection → Run through CNN → Returns Far Field Class as highest probability.

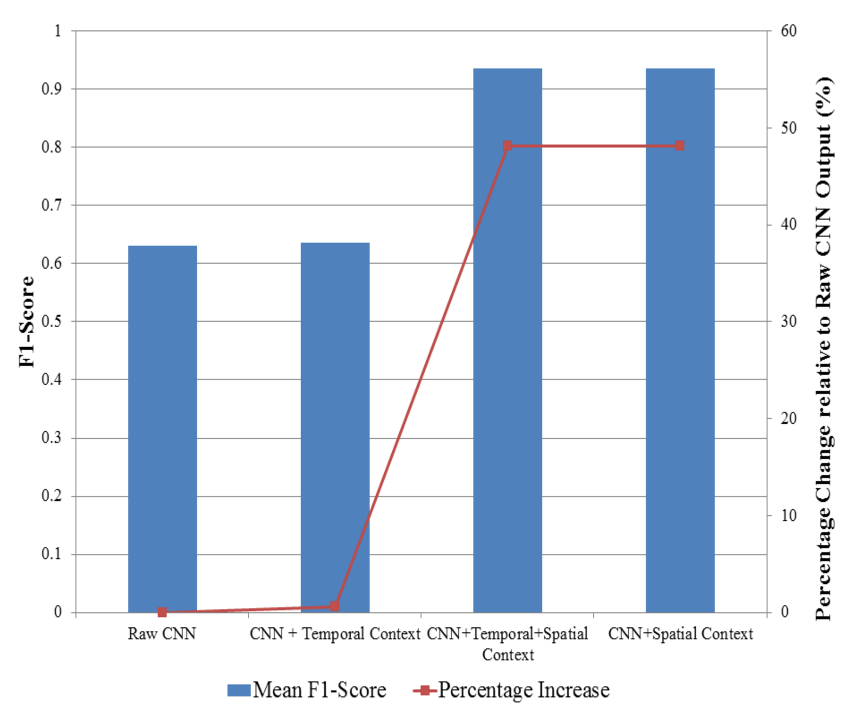


Segmentation with spatial context feature.
2) Use Spatial Context Feature to sample from semantic Colour Segmentation. Generate Probability for each object class to exist.



Same Scene, Output AD + CNN + CX
3) Reweight CNN scores based on the spatial context probability and find highest class. This is now correctly a Land Vehicle

Results



The multi-axis plot shows mean F1 Scores for the different variants of classification algorithm in our final experiment. As we can see, there is a marked improvement gained from spatial context incorporation. This is highlighted by the red line showing the percentage increase in F1 Score relative to the raw CNN output.

Conclusion

- We have successfully created an ATDR system for enhancing target recognition in long-range surveillance scenarios using multi-modal data.
- We have achieved this by using state of the art machine learning techniques in the form of a highly accurate CNN LWIR classifier and incorporating a subtle amount of information from RGB imagery.
- This work is a step towards allowing analysts to make safer and more informed decisions in dangerous defence domains.

THALES