

Data Science Final Report

1. Motivation

目前，全球機器人市場持續增長，全球機器人產業市場規模逾兩百七十億美元，但開發目標卻是讓機器人執行人類原本就能完成的事，如說話、跳舞、拉小提琴，甚至主持婚禮。

美國「國防部高級研究計劃局」(Defense Advanced Research Projects Agency，簡稱 DARPA) 的研究人員深感機器人在救災方面的局限，故舉辦為期三年的 SubT 地下搜救競賽。

此次研究目標將針對隧道等嚴峻環境之地區進行搜救、探勘，運用 AI 影像辨識技術及定位系統，結合 RRT 路徑規劃應用於自走車、自走飛船等載具，進入隧道並建置地圖及標示出相關物體。

而在比賽當中，Artifact Search 為比賽的主要項目，必須偵測到未知環境中的目標物並將其座標回報出來，故本次專題將用 GAN、FCN 及 ERFNet 做出一個好的預測模型並結合 ROS 在 Real time 的場域環境下運行。

2. Problem formulation

此次比賽的 Artifact Search 主要是達成，「如何在無光隧道進行影像辨識找尋目標物及生還者」，因此擬定策略並找到合適的神經網路修改是非常重要的。一個關鍵，而照主辦單位公布的 Artifact，Survivor 及 Phone 是具有溫度的物品，依據溫度這條線索我們如何在嚴苛的環境下增加準度就是本次研究的主要問題。



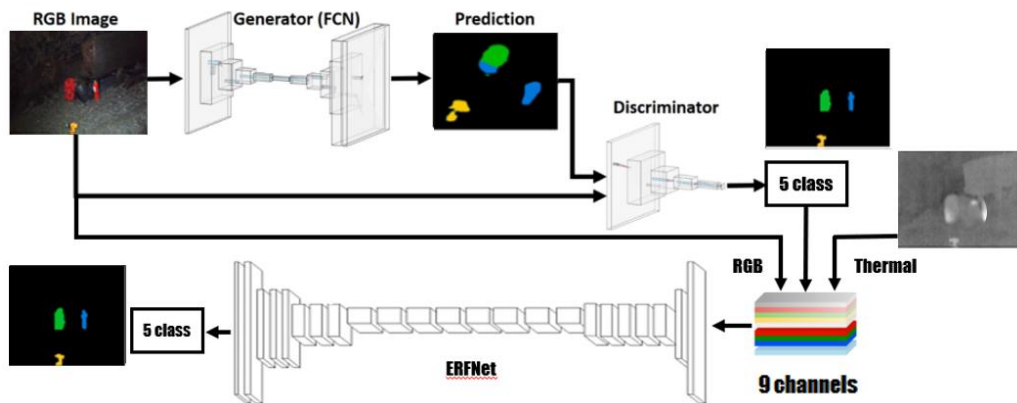
(a)Artifact



(b)Sensor on Robot

3. Proposed Model

本次專題提出結合 Camera D435 及 Thermal Sensor 的神經網絡來有效使用溫度這項資訊，讓 perception 部分更準確。

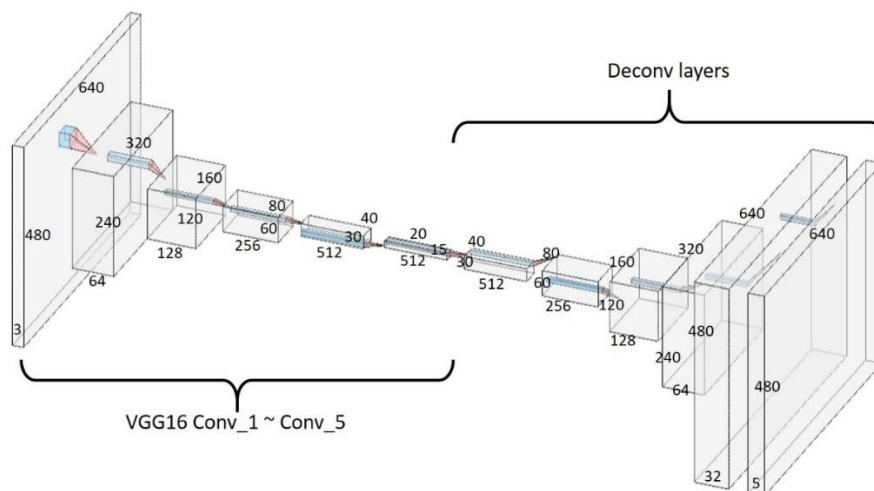


(c) Dual-Stream Network

上層網路使用 FCN-Pix2Pix 處理 RGB 生成 Mask，下層使用 ERFNet 加入 Thermal 資訊讓準確率提高

- FCN-Pix2Pix

■ Generator:

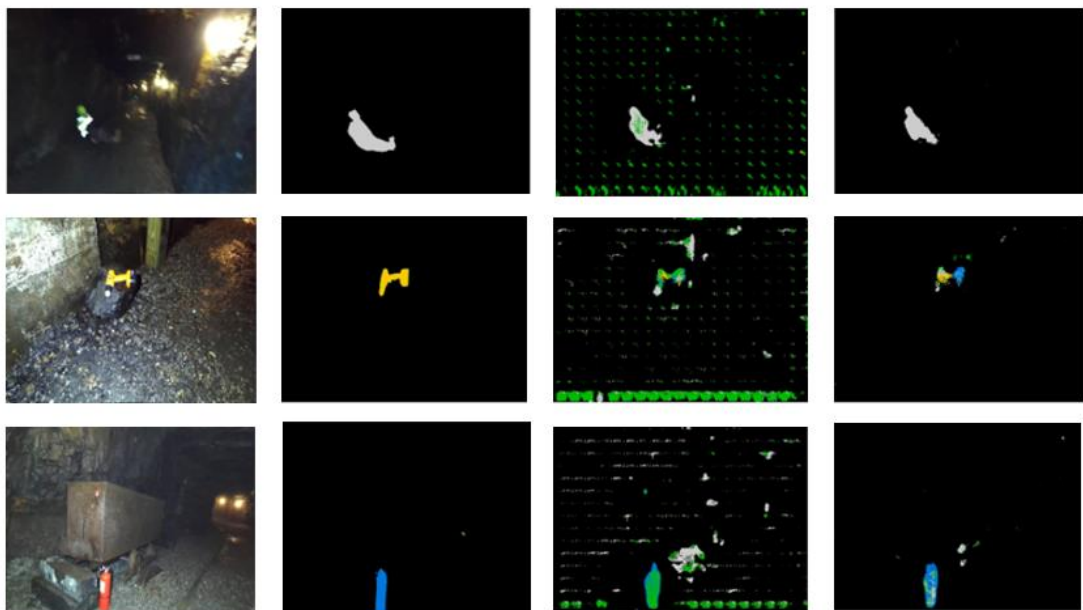


此處使用 Fully Convolutional Network (FCN)來生成 images 的 mask

Input 為 $480 \times 640 \times 3$ (RGB)

Output 為 $480 \times 640 \times 5$ (extinguisher, backpack, drill, survivor,
and background)

但由於在光線不足或雜亂的環境下每一個 pixel 都去預測就會出現許多的
雜訊，如圖(d)，所以除了 FCN 之外還使用了 Patch GAN 的 Discriminator
去消除雜訊。



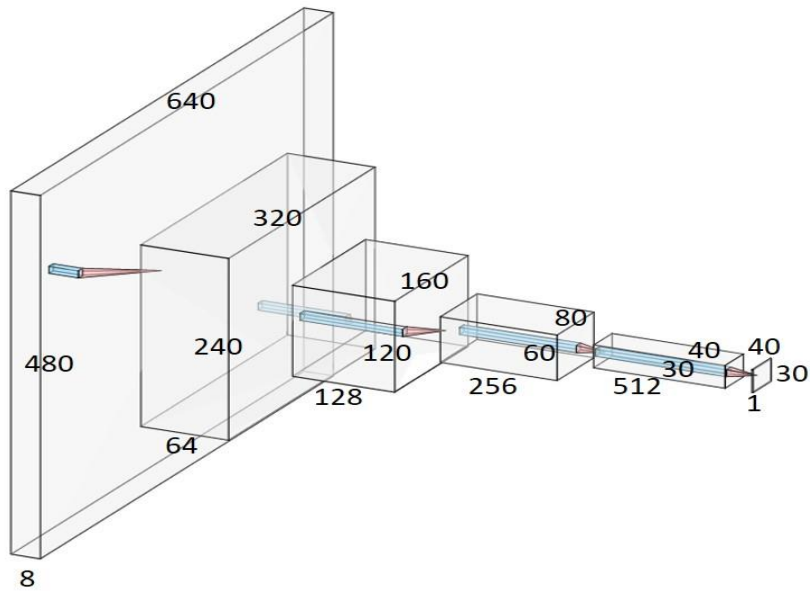
(d) raw images

label

FCN

FCN-Pix2Pix

■ Discriminator:

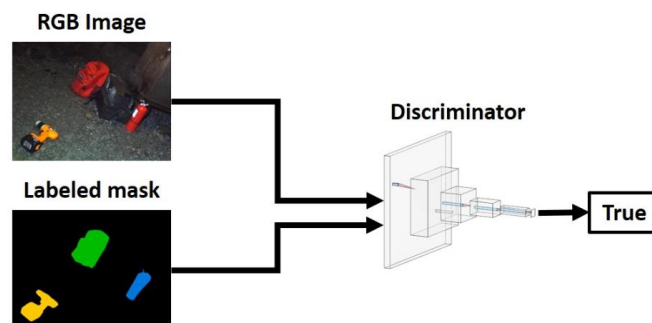
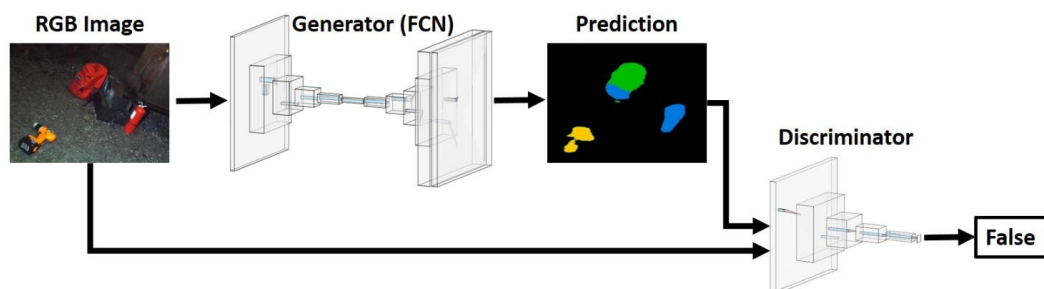


此處使用 Patch Discriminator 來判斷生成的 mask 跟原始 label 的差異

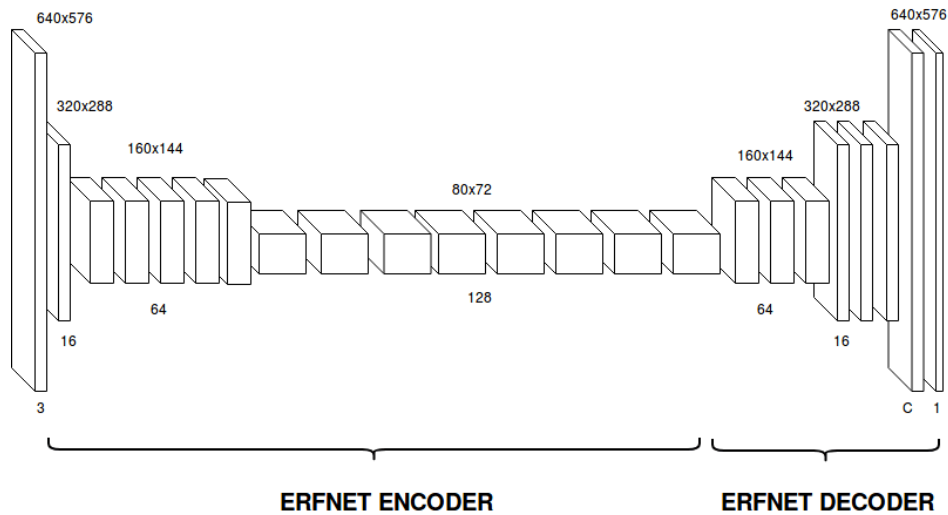
盡量讓 Discriminator 判斷不出真假。

Input 為 $480 \times 640 \times 8$ (RGB + 5 Channel)

Output 為 $30 \times 40 \times 1$



- ERFNet :



RESULTS ON PST900 DATASET: ASIDE FROM MFNET AND RTFNET, THE RGB-T MODE INDICATES A NAIVE FUSION OF THERMAL BY ADDING IT AS A FOURTH CHANNEL TO THE NETWORK INPUT. WE ALSO LIST INFERENCE LATENCY (IN MS) ON AN NVIDIA AGX XAVIER.

Dataset: PST900 Dataset								
Network	Mode	Background	Fire-Extinguisher	Backpack	Hand-Drill	Survivor	mIoU	ms (Xavier)
ERFNet	RGB	0.9869	0.6118	0.6528	0.4240	0.4169	0.6185	30
ERFNet	RGB-T	0.9873	0.5879	0.6808	0.5276	0.3438	0.6255	31
MAVNet	RGB	0.9822	0.2831	0.5850	0.3367	0.0901	0.4551	16
MAVNet	RGB-T	0.9789	0.2258	0.5152	0.3194	0.3473	0.4774	17
UNet	RGB	0.9843	0.4928	0.6364	0.4026	0.2337	0.5499	12
UNet	RGB-T	0.9795	0.4296	0.5289	0.3827	0.3164	0.5274	12
Fast-SCNN	RGB	0.9857	0.3454	0.6679	0.2063	0.2053	0.4822	18
Fast-SCNN	RGB-T	0.9851	0.3548	0.6460	0.1550	0.2168	0.4715	18
MFNet	RGB-T	0.9863	0.6035	0.6427	0.4113	0.2070	0.5702	23
RTFNet-50	RGB-T	0.9884	0.4349	0.7058	0.0100	0.2800	0.4840	42
RTFNet-152	RGB-T	0.9892	0.5203	0.7530	0.2537	0.3643	0.5761	127
Ours: RGB Stream	RGB	0.9883	0.6814	0.6990	0.5151	0.4989	0.6765	18
Ours: Naive	RGB-T	0.9852	0.6253	0.5961	0.4708	0.4339	0.6223	20
Ours: Full	RGB-T	0.9885	0.7012	0.6920	0.5360	0.5003	0.6836	42

(e) PST900 paper predict network result

此處參考 paper (PST900: RGB-Thermal Calibration, Dataset and

Segmentation Network) 中的圖表資訊選擇使用 mIoU 最高的 ERFNet 當作第二

層 Network。

Input 為 $480 \times 640 \times 9$ (5 class + RGB + Thermal)

Output 為 $480 \times 640 \times 5$

4. Experimental Results & Conclusions

Dataset: PST900						
Network	Mode	Extinguisher	Backpack	Drill	Survivor	mIOU
PST900	RGBT	0.7012	0.6920	0.5360	0.5003	0.6836
FCN-Pix2Pix	RGB	0.6036	0.7872	0.5638	0.6572	0.6529
FCN-Pix2Pix-ERFNet	RGB (5 channels)	0.6275	0.7803	0.6115	0.6235	0.6631
FCN-Pix2Pix-ERFNet	RGBT (6 channels)	0.6076	0.7956	0.6400	0.6403	0.6709
FCN-Pix2Pix-ERFNet	RGBT (9 channels)	0.7512	0.9210	0.8633	0.9027	0.8635

(f) 同樣使用 PST900 Dataset 並用不同 Network Predict Result

5 channel = 5 class output

6 channel = 5 class output + thermal

9 channel = 5 class output + thermal + RGB

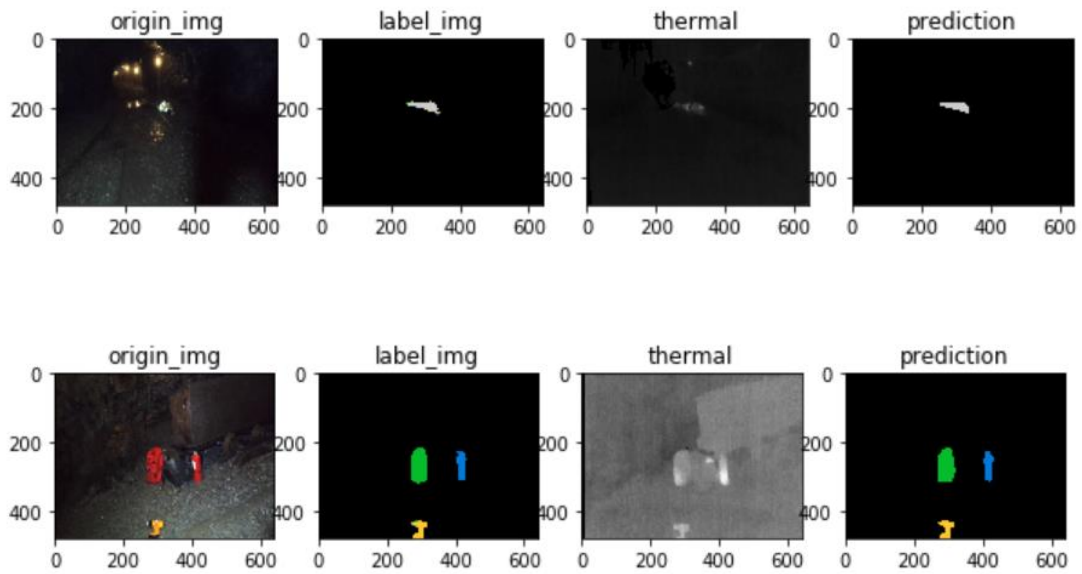
- FCN-Pix2Pix

如果單純使用上層 Network (FCN-Pix2Pix) mIOU 是低於 PST900 Network 的 (0.6529 < 0.636)，而其中比較奇怪的是 PST900 的 survivor 只有 0.5 IOU，因為假人是有溫度的照理說應該效果要更好但卻是所有 Artifact 最差的。

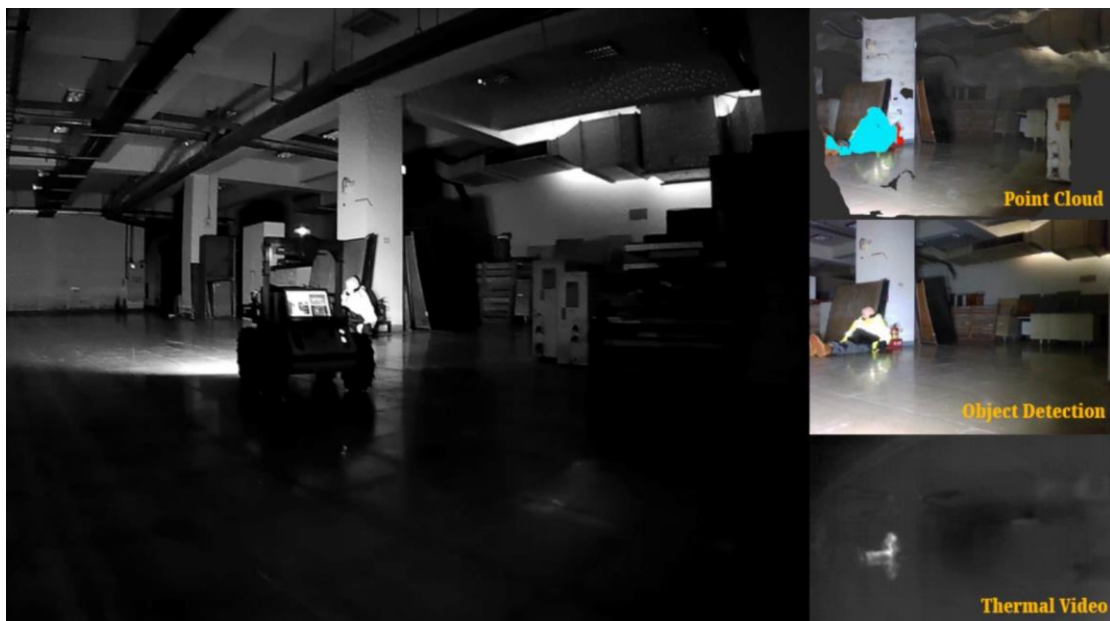
- FCN-Pix2Pix-ERFNet

加了第二層 Network 之後 survivor 的 IOU PST900 相比從最差的變成第二好的，表示增加 thermal 資訊及 Network 之後效果有變好。以及其他沒有溫度的物品 IOU 也相對提升許多。我想跟再將 RGB 丟進去 train 有極大的因素，可以看到圖表中只將 Output + Thermal 丟入第二層 train 效果並沒有提升很多直

到多加 RGB 資訊後表現才大幅提升。



(f) Predict result



(g) All sensor install on husky and run predict model

0510727 張博凱

0510737 黃瑞得

6. Teamwork Assignment

0510727 張博凱 (50%)

0510737 黃瑞得 (50%)