

# Machine Learning Project

Hsuan-Ying Hsieh  
Department of Applied Mathematics  
National Chung Hsing University  
Taichung, Taiwan  
g108053115@mail.nchu.edu.tw

## I. INTRODUCTION

這是第一次機器學習報告。

## II. RELATED WORK

由於第一次拿到的是跟 Siamese 相關的論文，因此想由此著手，以下是為這次報告選出的七篇將要研讀的論文，在此做簡單的介紹。

### A. Siamese Neural Networks for One-shot Image Recognition - 2015 出處未知

學習好的特徵在機器學習上有幾項缺點，在計算上有些浩大工程，且在小樣本時可能很難去證明它。此篇運用了一個獨特的架構，兩個 input 分別是兩張圖片，放入相同權重的神經網絡中並且強調只學習一次，即可拿來與沒看過的圖片做 testing，output 是兩張圖片的內容物相似度，再做相似度排名，判斷是否是相同內容物，此篇概念圖，顯示在 Fig 1。此篇將網絡的預測能力不僅適用於新數據，而且適用於來自未知分佈的全新類別。

#### Training



#### Testing

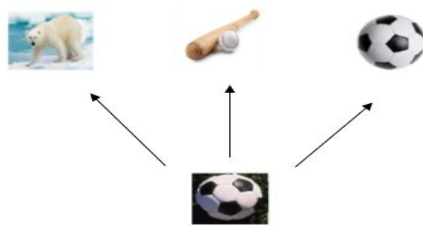


Fig. 1.

### B. Siamese Instance Search for Tracking - CVPR 2016

此篇並非提出一個新的架構，而是以當時的基礎，提出在不需要更新模型，沒有遮擋檢測，沒有跟踪器組合，沒有幾何匹配，並且依然保有不錯的跟踪性能的跟踪器。一旦經由學習後，matching function 就能直接使用，無需進行任何調整即可跟踪以前看不見的目標，在此所提出的跟踪器，甚至是沒有完整的出現在鏡頭的目標，也能重新確定目標位置。此篇粗略流程圖，顯示在 Fig 2。

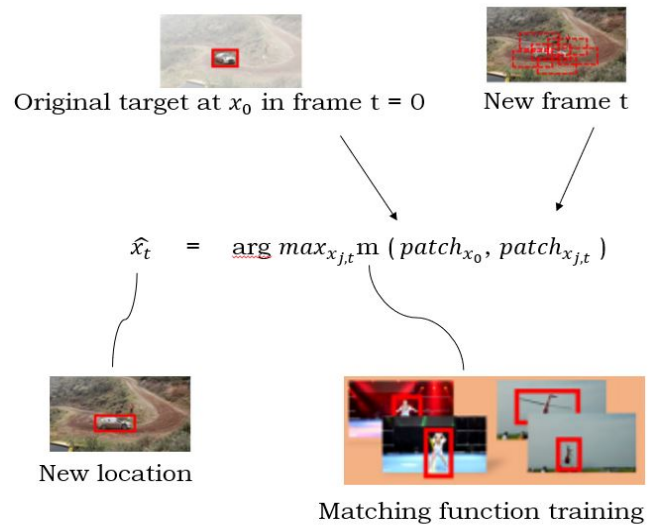
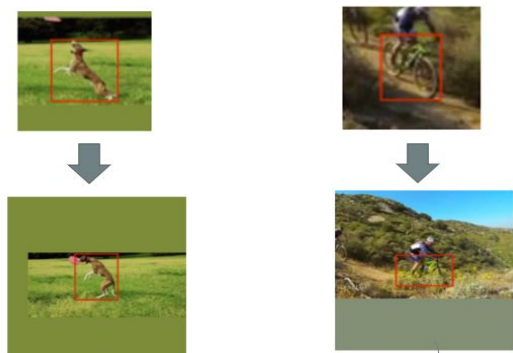


Fig. 2.

### C. Fully-Convolutional Siamese Networks for Object Tracking - CVPR 2016

此篇針對的問題是，如果事先不知道要跟踪的對象，則必須在線執行 SGD 以調整網絡的權重，從而嚴重影響系統的速度，因此提出了一個方法，在初始離線階段把 convolutional network 視為相似性學習問題，然後在跟踪時對此進行線上簡單估計，要跟踪的目標圖像和新的搜索範圍經過 Siamese network 後得到各自的特徵，對兩者進行一些運算後則會同樣得到一個相對應的圖，其每一個 pixel 的值對應了搜索範圍中與目標圖像一樣大的對應區域，並顯示跟踪目標的機率。此篇粗略概念圖，顯示在 Fig 3。

Training pairs extracted from the same video



exemplar image and corresponding search image from same video.

Fig. 3.

#### D. Triplet Loss in Siamese Network for Object Tracking - ECCV 2018

此文提出了一種新穎的 Triplet Loss，作者認為 SiamFC 的 Loss function 只考慮 instance 是否足夠正確，忽略了 positive instance 和 negative instance 之間的關係。因此將其添加到 Siamese network 框架中代替成對損失進行訓練，從而提取表達性深層特徵進行對象跟踪。在不添加任何輸入的情況下，我們的方法能夠利用更多元素進行訓練，以通過結合原始樣本來實現更強大的功能。此篇架構圖，顯示在 Fig 4。

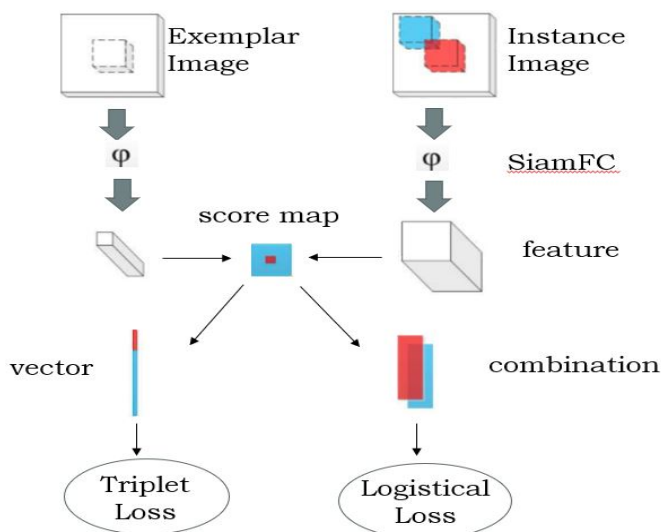


Fig. 4.

#### E. Structured Siamese Network for Real-Time Visual Tracking - ECCV 2018

此篇想解決的問題是，由於深度神經網絡的現有方法大多從全局角度描述目標外觀，對於非剛性外觀變化和部分

遮擋非常敏感，因此提出一種局部結構學習方法來避免此問題，該方法同時考慮目標的局部模式及其結構關係。將局部模式檢測模塊設計為自動識別目標對象的局部區域。此篇架構圖，顯示在 Fig 5。

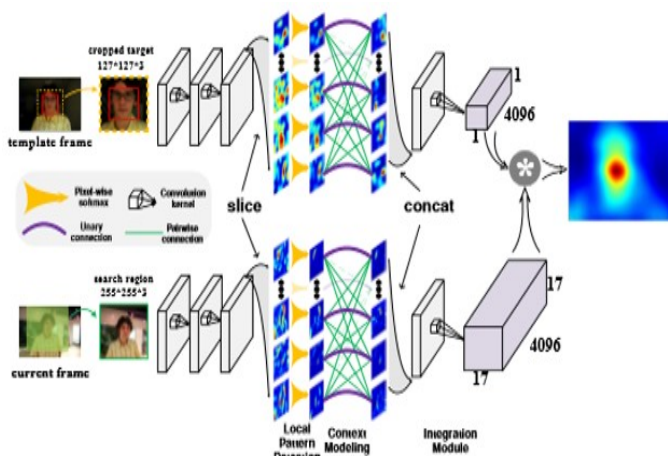


Fig. 5.

#### F. Distractor-aware Siamese Networks for Visual Object Tracking - ECCV 2018

大多數 Siamese Networks 方法中使用的功能只能將 foreground 與 non-semantic backgrounds 分開。semantic backgrounds 始終被視為乾擾因素，在訓練數據中的 non-semantic backgrounds 和 semantic backgrounds 的干擾物背景的數據不平衡做進一步的學習。藉由引入現有數據充實正樣本提升跟踪器的 generalization ability，充實 Semantic negative pairs 來提升跟踪器的 discriminative ability。此篇粗略概念圖，顯示在 Fig 6。

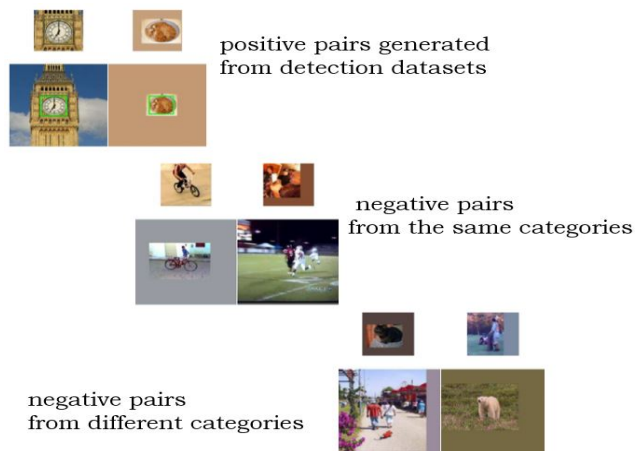


Fig. 6.

*G. High Performance Visual Tracking with Siamese Region Proposal Network - CVPR 2018*

此篇針對大多數跟踪器無法以 real-time speed 獲得佳性能，作者提出了 Siamese region proposal network，它能够利用大尺度的圖對端到端訓練。也就是說，這個結構包含用於特提取的 Siamese subnetwork 和 region proposal subnetwork，其中 region proposal subnetwork 包含分類和回歸。在跟踪階段，作者提出的方法被稱為 one-shot detection task。有了這些改良，傳統的多尺度測試和線上微調可以被捨棄，這樣也提高了速度。

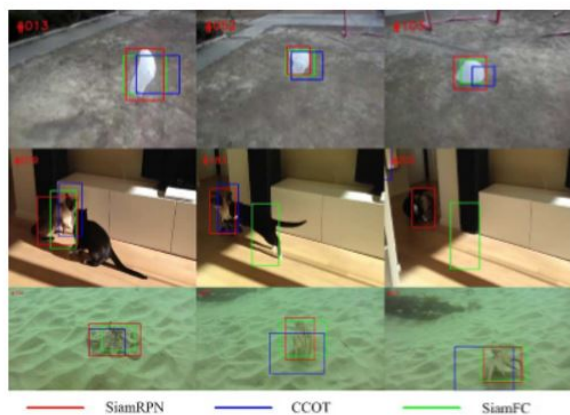


Fig. 7. SiamRPN is able to predict the shape more precisely than SiamFC , CCOT when target' s shape is severely changing.