

基於域隨機化之多模態訓練資料生成模擬器

Domain Randomized Multi-Model Training Data Simulator

林子祺

指導教授：朱宏國 教授

一、摘要

前在發展深度學習時，如果需要物件偵測相關的應用，常需要大量已標注的圖片去訓練神經網路，然而獲得已標注的資料需要消耗大量的人力與時間。如果能夠在虛擬的環境中合成出已標注的資料，就能夠更穩定，更快速地獲得更多樣的資料。

我們基於 Unreal Engine 4(UE4)上的 NVIDIA Deep Learning Dataset Synthesizer(NDDS)，建立了 Multi-Model Training Data Simulator (MMTD Simulator)，透過 MMTD Simulator 使用者只需要簡單的設置環境就能夠快速產生各式已標注的資料。

為了提高虛擬資料的擬真度，我們在 NDDS 的架構上增加了域隨機化相關的功能，讓合成出來的虛擬資料能夠更多樣。

最後，我使用 MMTD Simulator 生成了一組標注有 bounding box 的虛擬資料，並使用這筆虛擬資料在 YOLOv3 神經網路上訓練出一個模型。經過測試，這個由虛擬資料訓練出的模型有能力在真實的圖片中分辨出需要的物件，這驗證了 MMTD Simulator 所產生的資料能夠用於神經網路的訓練。

二、內容

2.1 動機

目前在發展深度學習時，如果需要物件偵測相關的應用，常需要大量已標注的圖片去訓練神經網路，然而獲得已標注的資料需要消耗大量的人力與時間。如果能夠在虛擬的環境中合成出已標注的資料，就能夠更穩定，更快速地獲得更多樣的資料。

但是虛擬資料有其缺點，使用電腦渲染出的虛擬資料在光影、材質和模型上一定不如真實資料精細，在訓練時可能會導致 overfitting。如何彌補虛擬資料和真實資料之間的差異就是我們要解決的問題。

2.2 研究目的

我們打算使用域隨機化的手法來彌補虛擬資料和真實資料之間的差異。域隨機化就是在生成

虛擬資料時，隨機地變化攝影機位置、物件的材質和貼圖、光線的手法。域隨機化建立在一個假設下：如果模型在訓練時看過的虛擬物件差異足夠大，最終訓練出來的模型，對於物件的辨識能力比較容易推廣到在真實物件上。

2.3 研究方法

我們基於 Unreal Engine 4(UE4)上的插件：NVIDIA Deep Learning Dataset Synthesizer(NDDS)，開發了 Multi-Model Training Data Simulator (MMTD Simulator)，透過 MMTD Simulator 使用者只需要簡單的設置環境就能夠快速產生域隨機化的標注資料，除了 NDDS 原有的域隨機化功能以外，我們為了提高虛擬資料的擬真度，增加了更多域隨機化相關的功能，讓合成出來的虛擬資料能夠更多樣。

圖1 域隨機化虛擬訓練資料示意圖



2.4 研究結果

了確認 MMTD Simulator 的功效，我使用 MMTD Simulator 生成了3000張標注有 bounding box 的虛擬資料，為了減少虛擬資料和真實資料的差距 我使用 Occlusion Filter 的功能來去除重疊的酒瓶。最後我使用這筆虛擬資料在 YOLOv3神經網路上訓練出一個偵測酒瓶位置的模型。

表1 模型訓練參數

# of training data	# of validation data	Optimizer	Lr	Epochs	Batch-size
3000	300	SGD	0.0004	50	4

(資料來源：YOLOv3訓練結果)

圖2 模型訓練曲線

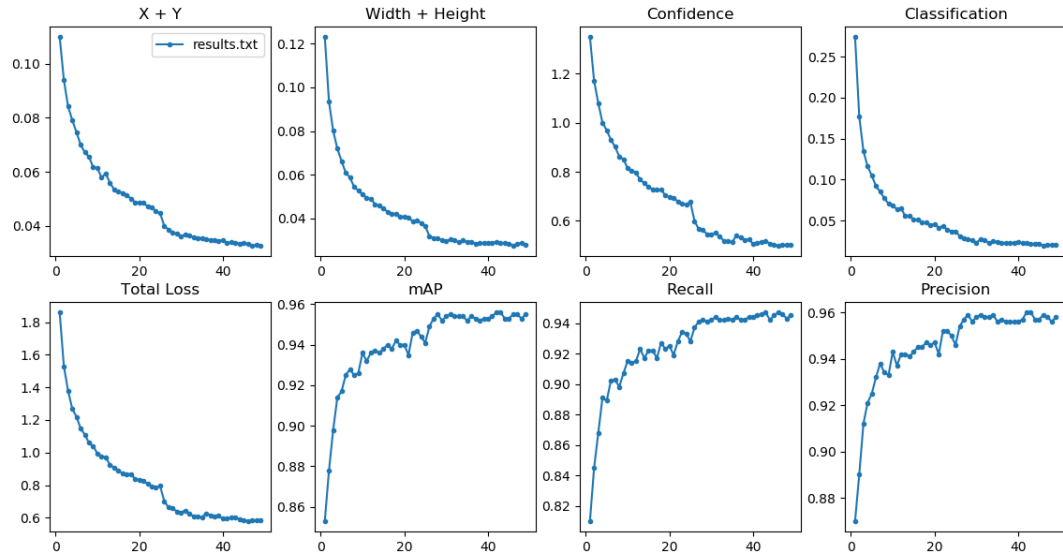


表2 酒瓶偵測模型 Average Precision

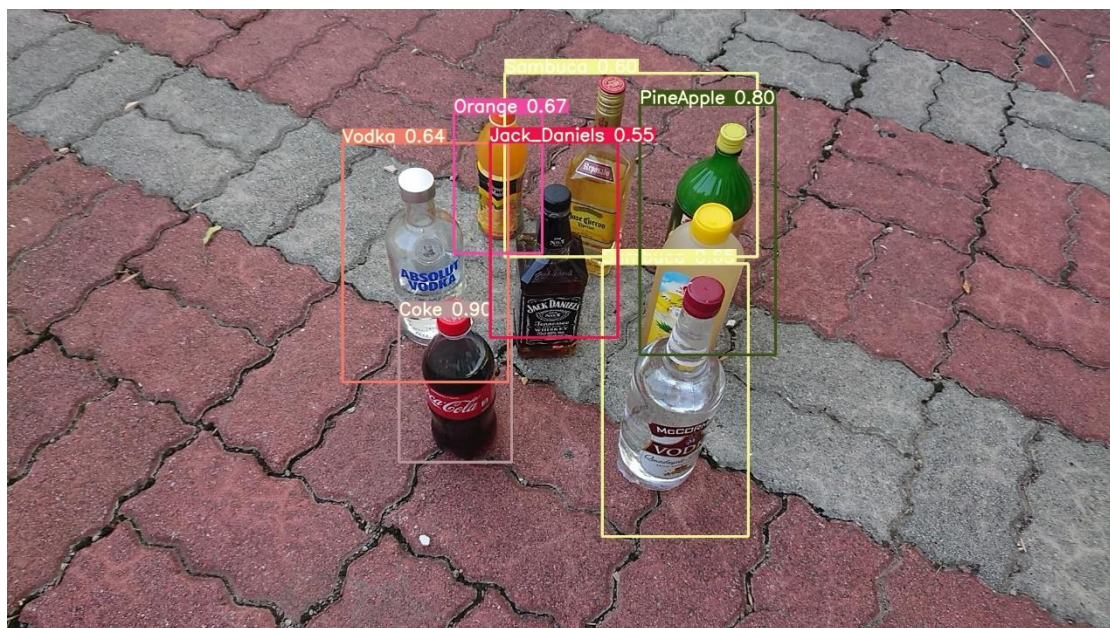
coke	JB	Sambuca	Vodka
0.9719	0.9669	0.9549	0.9106

表3 酒瓶偵測模型 Average Precision

Orange	PineApple	Jack_Daniels	Jose_Cuervo
0.9549	0.9719	0.9318	0.9665

為了驗證模型的功效，我請實驗室的同學幫我蒐集了一系列有關酒瓶的真實相片，並用將這些圖片丟進訓練好的模型中測試。考慮到這是完全由虛擬資料訓練出的模型，我認為準確度相當的高。如果加入真實資料去座訓練，對 bounding box 的偵測應該能更準確。

圖3 真實圖片在酒瓶偵測模型上的測試結果



經過測試，這個由虛擬資料訓練出的模型有能力在真實的圖片中分辨出需要的物件，這驗證了 MMTD Simulator 所產生的資料能夠用於神經網路的訓練。

2.5 總結

MMTD Simulator 用域隨機化手法產生虛擬訓練資料的，操作直觀，不具程式基礎、對 Unreal Engine 4 不熟悉的人也可以輕鬆上手。而且資料生成資料快速，可以在一小時內產生上萬張的訓練資料。

我們的主要貢獻就是在 NDDS 域隨機化的功能上加入了更多功能，以便產生更加貼近現實的資料。

最後我們也在 YOLOv3 上訓練了一個酒瓶模型，證明由 MMTD 產生出的資料是能夠實際應用在物件偵測神經網路的訓練上。

三、心得

經過一年的跌跌撞撞，專題總算到了一個段落，為了這次專題接觸了許多以前接觸不到的東西。為了能夠生成域隨機化的資料學習了如何使用 Unreal Engine 4，為了驗證 MMTD Simulator 功效學習了對深度學習相關的知識，在每週 Meeting 的壓力之下學習如何規劃時間，在每週 meeting 的過程中學會了要如何編排簡報和說話方式才能完整傳達心中的想法。雖然很累，但獲得了太多太多平常修課修不到的東西，完全沒有後悔進了 CGV LAB 做了這個專題。