

# 二維影像深度預測

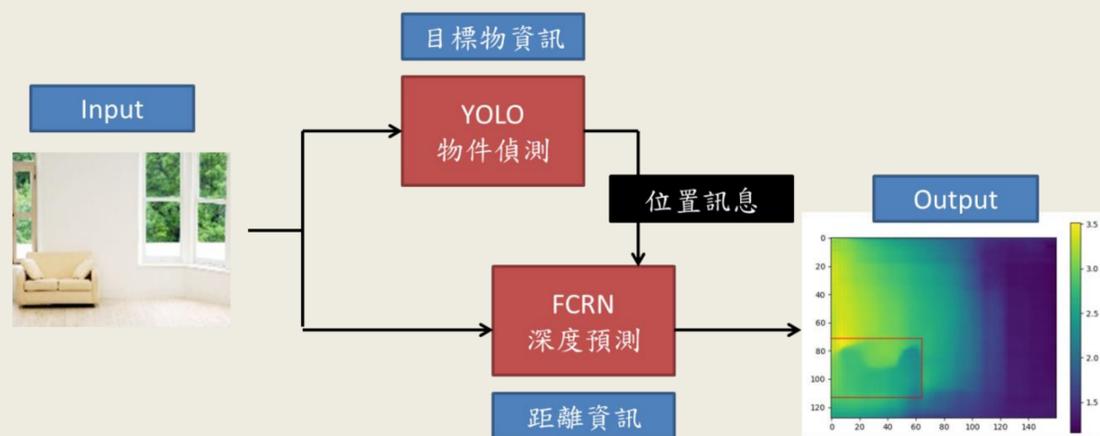
# Depth Prediction On Two Dimensional Image

## Introduction

這個專題，我們把問題專注在自動移動的機器上。自動移動的機器，像是自動駕駛、無人機，或是居家環境中，能夠自動移動的機器人。這些機器最重要的就是辨識出目標物體的位置，再來就是能夠知道距離，並且移動到目標物體的位置。我們根據這個情境，提出的方法。利用物件偵測，加上深度預測來進行。最後成果是，當一張照片輸入後，可以偵測物體位置，並且獲取距離資訊的照片。

指導教授: 孫民 教授  
組別: A22  
組員: 鄭敬儒  
王立友  
陳經賢

## Design



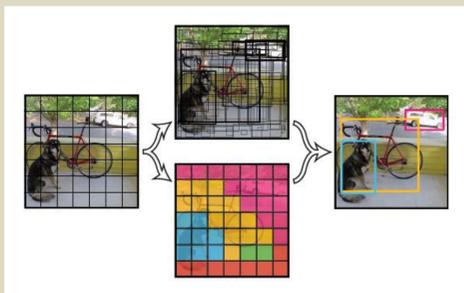
這個系統主要分為兩部分:

1. YOLO的物件偵測 (取得目標物體位置資訊)
2. FCRN的深度預測 (取得距離目標物及其他物體的距離資訊)

## YOLO

首先一張輸入的照片會被切成 $s \times s$ 塊的方格，YOLO此處切為 $7 \times 7$ 個方格，這些方格我們稱為grid cell，每個grid cell會根據那個方格裡的东西預測各種物體的機率( $c$ 個種類的物體)

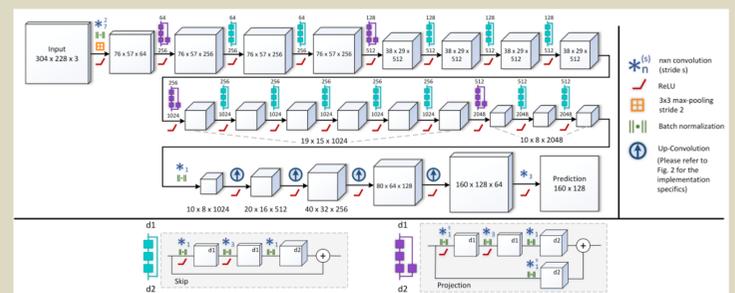
每個grid cell也會根據裡面的東西來預測出 $B$ 個框，這 $B$ 個框就是預測來框住物體的，也就是預測物體的位置會在哪裡。這 $B$ 個框包含5個資訊，框的中心 $(x, y)$ 、框的長寬 $(w, h)$ 、框的置信度，所謂的置信度就是那個框有物體且框的準不準的指標



## FCRN

首先，用完全卷積架構來深度預測，加上up-sampling blocks可讓輸出畫質更精細，也同時只需要較少的參數、相較現在技術少一個數量級的資料，而且還比所有標準基準數據集表現更好。

再來，對up-convolution提出一個更有效的方法，然後結合殘差網路的概念去創造up-projection blocks，可達到對特徵地圖有效的upsampling



## Result

