

# Multimix：通過改良 ResizeMix 方法進行目標檢測的新型數據增強方法

## Multimix：A Novel Data Augmentation for Object Detection

### by Modifying ResizeMix Method

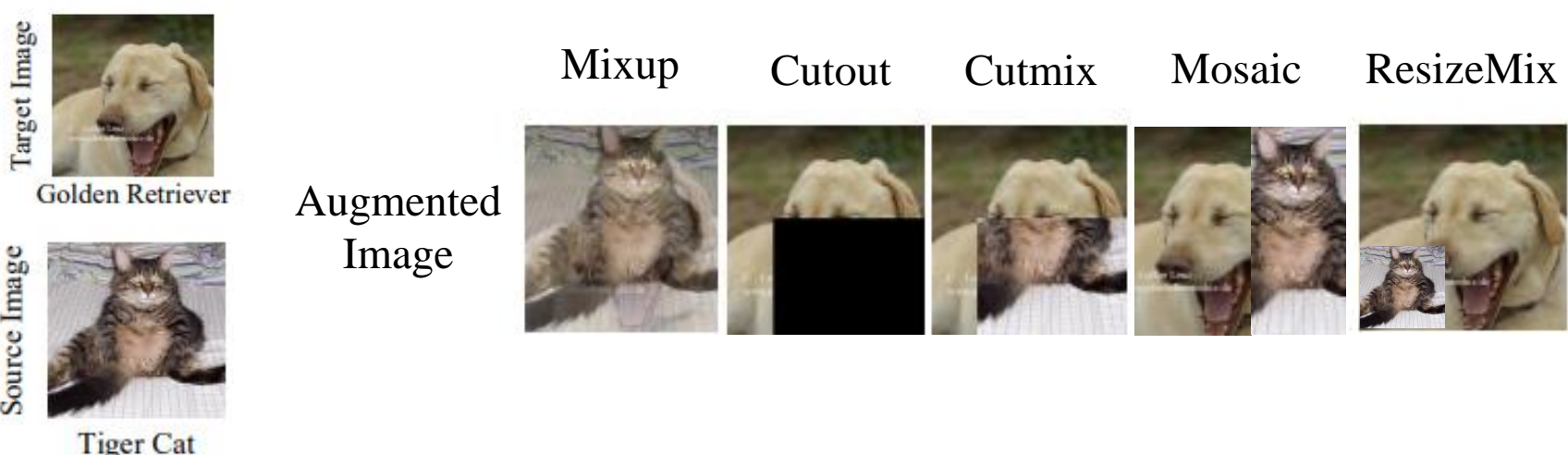
PEI-CHEN CHEN

College of Electrical and Computer Engineering, National Yang Ming Chiao Tung University

## Introduction

數據增強是一種增加數據多樣性的強大技術，可以有效提高神經網絡在圖像識別任務中的性能和泛化能力。目前已有多種數據增強技術，如圖一所示，而最近基於數據混合的增強策略取得了良好的模型表現，但多數的數據增強技術對於分類器上能夠有很好的表現，在物件偵測上卻不一定適合，因為影像 dropout 對於物件偵測上會相對不穩定。而其中，ResizeMix 技術能夠應用在物件偵測上，並且比 Cutout 的效果來的良好，它使用一種簡單但有效的方法來進行影像增強，方法是將一張圖像隨機縮放並將其黏貼到另一張圖像上。

為了進一步提升 ResizeMix 的性能，我們提出了 Multimix，透過多種組合的黏貼方式，增加影像的多樣性，並且使模型對物件遮掩具有穩定辨識率，在沒有額外計算成本的情況下，使一張影像擁有多種特徵供模型學習。



▲圖一 不同數據增強技術視覺化效果

## Related Work

表一比較目前效果最顯著的數據增強技術之間的主要區別，可以發現要將數據混合方法應用在物件偵測上式較有難度的，一般數據混合方法可將影像分為 Target Image 及 Source Image，透過隨機裁減 Source Image 部分影像或是特徵較明顯的部分，黏貼至 Target Image 上，而其中，ResizeMix 則與大家不同，它是透過將Source image 縮放黏貼至 Target image 上，它完整保留了 Source Image 的所有影像資訊，因此，我們透過改良 ResizeMix，將 Source Image 的檢測框也完整保留，使得一張影像同時擁有兩張影像的特徵檢測框，並且按照隨機比例不按照長寬比縮放，並隨機黏貼至 Target Image 上。

然而，因為數據增強技術多半專注在分類任務上，文獻都未提及在物件偵測上，是否有避免掉 Target Image 檢測框被過度覆蓋的問題，因此我們設定一個面積比例閾值，也就是 Target Image 的檢測框不可以被黏貼的 Source Image 覆蓋率超過 50% 以上，以避免不切實際的預測效果。

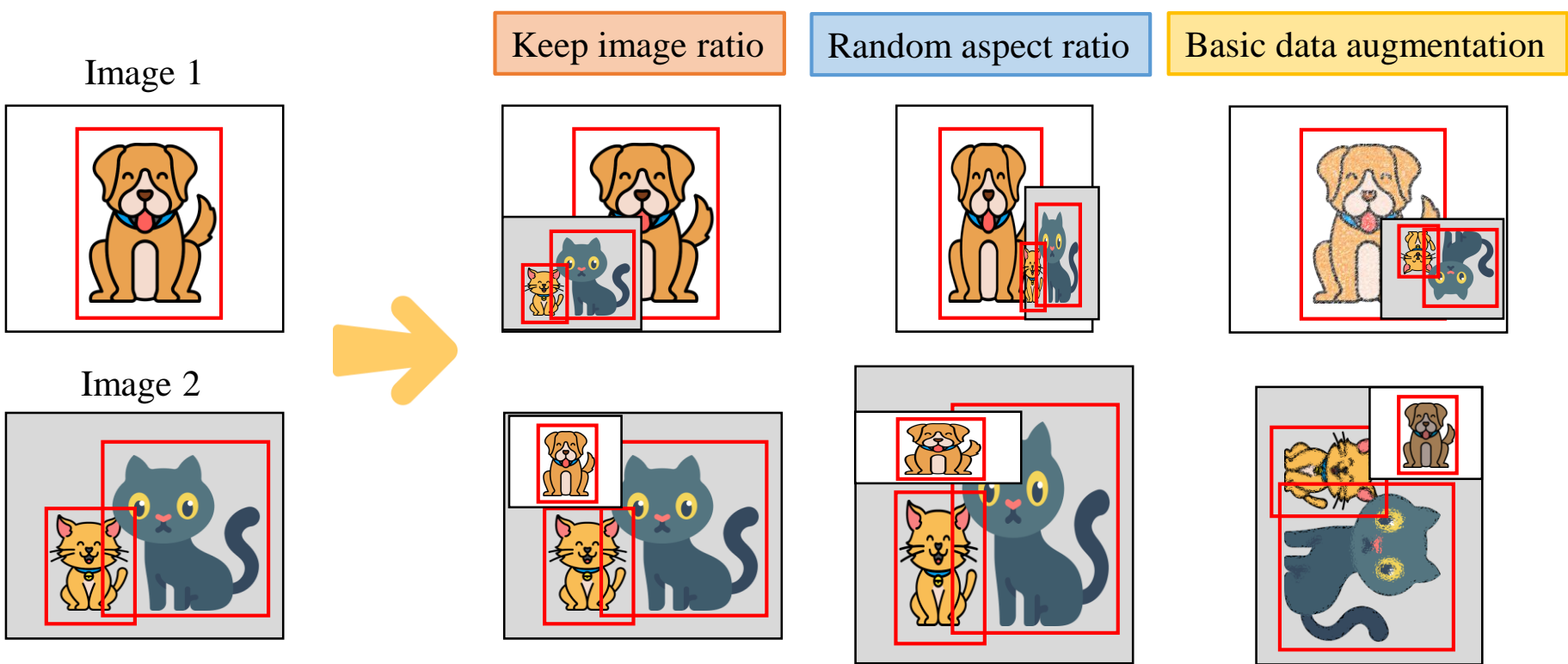
Approach	Mixed?	Double-Image?	Applicable to object detection?
Mixup	✓	✓	✗
Cutout	✗	✗	✓
CutMix	✓	✓	✗
Mosaic	✓	✓	✗
SelfMix	✓	✗	✓
ResizeMix	✓	✓	✓
Multimix	✓	✓	✓

▲表一 數據增強技術之間的主要區別

## MultiMix

我們提出的 MultiMix 數據增強技術可以解決數據不足的問題，並且擁有多種組合的可能性，位址也都隨機決定，但是是以不覆蓋底層影像的檢測框面積百分比 50% 以上的前提，如圖二所示，即便只有兩張影像，通過隨機生成與多種組合的方式，將兩張影像進行數據混合。

對於 Target Image 來說，重要的資訊就是檢測框所涵蓋的範圍，其餘的背景處被用來黏貼另一張影像，使得整張影像包含原始兩張完整的影像的標記資訊。



▲圖二 Multimix 概念結構圖

## GUI Development

為了使 MultiMix 多種組合方式方便使用，因此開發了一個使用者介面，透過簡單的介面與按鈕設計，完成客製化數據增強的資料集，影像與標記檔都會即時產生，介面如圖三所示。

針對 Target Image 可決定是否要按照比例縮放或是是否要執行基本的數據增強技術(如：翻轉、雜訊、亮度)，而針對 Source Image 則可以決定是否要維持原始大小，以及是否按照比例縮放，也可使用基本數據增強技術，甚至可以選擇是否要黏貼兩張 Source Image 至 Target Image 上，而最終還可以設定需要生成幾倍的影像。



▲圖三 實現 Multimix 的介面

## Experiment Results

為了檢測 MultiMix 在物件偵測上的性能，我們使用 coco10 作為訓練資料集，並使用 YOLOv5 作為模型訓練，在訓練期間，訓練集固定為 1000 張影像，驗證集則為 260 張影像，並且將影像大小固定至 640，Epoch 設定為 25，不使用預訓練權重。

實作結果如表二所示，我們的 MultiMix 在物件檢測下顯示出強大的泛化能力，不僅比未進行數據增強的模型表現來的好，也比 ResizeMix 與 Cutout 效果來的好，特別是在維持比例縮放的偵測上，但相較於原始數據，MultiMix 數據增強無論是哪種組合的方式，都能達到一定程度上的模型泛化能力，其中最好的模型表現在 mAP@0.5 高了 2.8%，而在 mAP@[0.50:95] 則高了 0.3%。

Dataset		mAP@.5 (%)	mAP@[.5 ,.95] (%)
Basic data		42.6	20.6
ResizeMix		38.6	17.8
Cutout		43.1	19.4
MultiMix	Keep image ratio	<b>45.4</b>	<b>20.8</b>
	Random aspect ratio	<b>44.6</b>	<b>20.9</b>
	Double image	<b>44.9</b>	20.1

▲表二 MultiMix、ResizeMix 與 Cutout 之間的泛化能力比較

## Conclusions

許多數據增強方式都是針對分類器所設計，並沒有一個數據增強的方式是針對物件偵測，因此我們設計了 MultiMix，專門應用在物件偵測的任務上，而透過實際測試結果，也證實它在模型泛化表現上能夠大幅提升。透過多重數據混合的方式，將一張影像包含多張影像的標記物件，在不增加運算量的情況下，可以減少 batch size，也可以減少模型學習時間，同時又可以避免過度擬合，能夠擁有更好的準確度及泛化能力。