

YOLO v4它來了：接棒者出現，速度效果雙提升

© 2020-04-24

兩個月前，YOLO 之父 Joseph Redmon 表示，由於無法忍受自己工作所帶來的負面影響，決定退出計算機視覺領域 (http://mp.weixin.qq.com/s?__biz=MzA3MzI4MjgzMw==&mid=2650781020&idx=1&sn=0cb4ae88c603ec778ef5acc1228fb3c1&chksm=871a7122b06df834f5d273b5ae8e8cc)

此事引發了極大的熱議，其中一個懸念就是：我們還能等到 YOLO v4 面世嗎？

當我們都以為再也沒有希望的時候，YOLO v4 卻悄無聲息地來了。這一目標檢測神器出現了新的接棒者！

數小時前，YOLO 的官方 Github 賬號更新了 YOLO v4 的 arXiv 鏈接與開源代碼鏈接，迅速引起了 CV 社區的關注。

📖 README.md

Yolo-v4 and Yolo-v3/v2 for Windows and Linux

(neural network for object detection) - Tensor Cores can be used on [Linux](#) and [Windows](#)

Paper Yolo v4: <https://arxiv.org/abs/2004.10934>

More details: <http://pjreddie.com/darknet/yolo/>

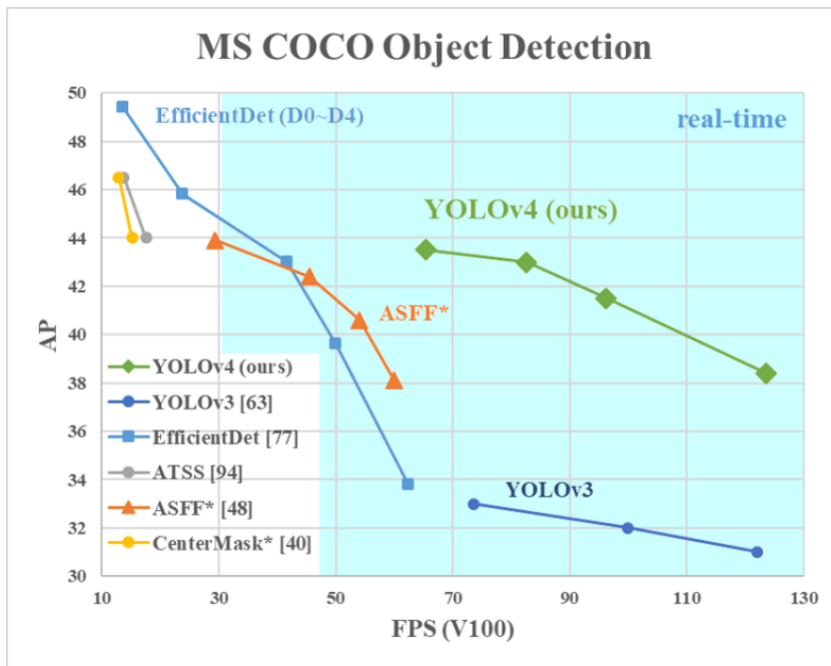
 PASSED build error build passing contributors 54 license Unlicense DOI 10.5281/zenodo.3693999

- YOLO v4 論文：<https://arxiv.org/abs/2004.10934>
- YOLO v4 開源代碼：<https://github.com/AlexeyAB/darknet>

YOLO v4 的作者共有三位：Alexey Bochkovskiy、Chien-Yao Wang 和 Hong-Yuan Mark Liao。其中一作 Alexey Bochkovskiy 是位俄羅斯開發者，此前曾做出 YOLO 的 windows 版本。

那麼，YOLOv4 性能如何呢？

在相關論文中，研究者對比了 YOLOv4 和當前最優目標檢測器，發現 YOLOv4 在取得與 EfficientDet 同等性能的情況下，速度是 EfficientDet 的二倍！此外，與 YOLOv3 相比，新版本的 AP 和 FPS 分別提高了 10% 和 12%。



接下來，我們看下 YOLO V4 的技術細節。

許多特徵可以提高 CNN 的準確率，然而真正實行起來，還需要在大型數據集上對這些特徵組合進行實際測試，並且對測試結果進行理論驗證。某些特徵僅在某些模型上運行，並且僅限於特定的問題，或是隻能在小型數據集上運行；而另外有些特徵（如 批歸一化和殘差連接）則適用於大多數模型、任務和數據集。

那麼，如何利用這些特徵組合呢？

YOLOv4 使用了以下特徵組合，實現了新的 SOTA 結果：

- 加權殘差連接 (WRC)
- Cross-Stage-Partial-connection · CSP
- Cross mini-Batch Normalization · CmBN
- 自對抗訓練 (Self-adversarial-training · SAT)
- Mish 激活 (Mish-activation)
- Mosaic 數據增強
- DropBlock 正則化
- CloU 損失

據介紹，YOLOv4 在 MS COCO 數據集上獲得了 43.5% 的 AP 值 (65.7% AP50)，在 Tesla V100 上實現了 ~65 FPS 的實時速度。

該研究的主要貢獻如下：

1. 建立了一個高效強大的目標檢測模型。它使得每個人都可以使用 1080Ti 或 2080Ti 的 GPU 來訓練一個快速準確的目標檢測器。
2. 驗證了當前最優 Bag-of-Freebies 和 Bag-of-Specials 目標檢測方法在檢測器訓練過程中的影響。
3. 修改了 SOTA 方法，使之更加高效，更適合單 GPU 訓練。這些方法包括 CBN、PAN、SAM 等。

YoloV4 如何實現這麼好的效果？

YoloV4 的基本目標是提高生產系統中神經網絡的運行速度，同時為並行計算做出優化，而不是針對低計算量理論指標（BFLOP）進行優化。YoloV4 的作者提出了兩種實時神經網絡：

- 對於 GPU，研究者在卷積層中使用少量組（1-8 組）：CSPResNeXt50 / CSPDarknet53；

- 對於 VPU，研究者使用了分組卷積（grouped-convolution），但避免使用 Squeeze-and-excitement（SE）塊。具體而言，它包括以下模型：EfficientNet-lite / MixNet / GhostNet / MobileNetV3。

YOLOv4 包含以下三部分：

- 骨幹網絡：CSPDarknet53
- Neck：SPP、PAN
- Head：YOLOv3

具體而言，YOLO v4 使用了：

- 用於骨幹網絡的 Bag of Freebies（BoF）：CutMix 和 Mosaic 數據增強、DropBlock 正則化和類標籤平滑；
- 用於骨幹網絡的 Bag of Specials（BoS）：Mish 激活、CSP 和多輸入加權殘差連接（MiWRC）；
- 用於檢測器的 Bag of Freebies（BoF）：CloU-loss、CmBN、DropBlock 正則化、Mosaic 數據增強、自對抗訓練、消除網格敏感性（Eliminate grid sensitivity）、針對一個真值使用多個錨、餘弦退火調度器、優化超參數和隨機訓練形狀；
- 用於檢測器的 Bag of Specials（BoS）：Mish 激活、SPP 塊、SAM 塊、PAN 路徑聚合塊和 DIoU-NMS。

架構選擇

該研究的目標是找出輸入網絡分辨率、卷積層數量、參數量（濾波器大小*濾波器*通道/組）和層輸入數量（濾波器）四者之間的最優平衡。

次要目標則是挑選能夠增加 感受野的額外塊（additional block），以及針對不同級別的檢測器從不同骨幹層中挑選最佳的參數聚合方法，如 FPN、PAN、ASFF 和 BiFPN 網絡。

研究者在 CSPDarknet53 上添加了 SPP 塊，因為它能夠極大地增加 感受野，分離出最顯著的上下文特徵，並且幾乎沒有降低網絡運行速度。他們針對不同級別的檢測器從不同骨幹層中挑選 PANet 作為參數聚合方法，而放棄了 YOLOv3 中使用的 FPN 網絡。

最後，研究者選擇了 CSPDarknet53 骨幹網絡、SPP 額外模塊、PANet 路徑聚合 neck 和 YOLOv3（基於錨的）head 作為 YOLOv4 的整體架構。

BoF 和 BoS 的選擇

為了提升目標監測的訓練效果，CNN 使用了以下方法：

- 激活函數：ReLU、leaky-ReLU、parametric-ReLU、ReLU6、SELU、Swish、Mish；
- 邊界框迴歸損失（Bounding box regression loss）：MSE、IoU、GIoU、CloU、DIoU；
- 數據增強：CutOut、MixUp、CutMix；
- 正則化方法：DropOut、DropPath、Spatial DropOut、DropBlock；
- 通過均值和方差的歸一化網絡激活函數：批歸一化（BN）、跨 GPU 批歸一化（CGBN 或 SyncBN）、濾波器響應歸一化（FRN）、交叉迭代批歸一化（CBN）；
- 跳躍連接方式：殘差連接、加權殘差連接、多輸入加權殘差連接、Cross stage 局部連接（CSP）。

在訓練激活函數時，因為 PReLU 和 SELU 更難訓練，ReLU6 是專為量化網絡設計的，所以從候選列表裏刪除了這幾個函數。

額外改進

為了使檢測器更適合在單個 GPU 上進行訓練，研究者還做出了以下額外的設計與改進：

- 提出新型數據增強方法 Mosaic 和自對抗訓練（SAT）；
- 在應用遺傳算法時選擇最優超參數；
- 修改現有方法，使新方法實現高效訓練和檢測——modified SAM、modified PAN 和 Cross mini-Batch Normalization（CmBN）。

新型數據增強方法 **Mosaic** 混合了 4 張訓練圖像，而 **CutMix** 只混合了兩張輸入圖像，具體如下圖 3 所示：



圖 3：Mosaic 表示的一種新型數據增強方法。

自對抗訓練（SAT）也是一種新的數據增強方法，它包括兩個階段。第一個階段中，神經網絡更改原始圖像；第二階段中，訓練神經網絡以正常方式在修改後的圖像上執行目標檢測任務。

CmBN 是 CBN 的改進版，它僅收集單個批次內 mini-batch 之間的統計數據。

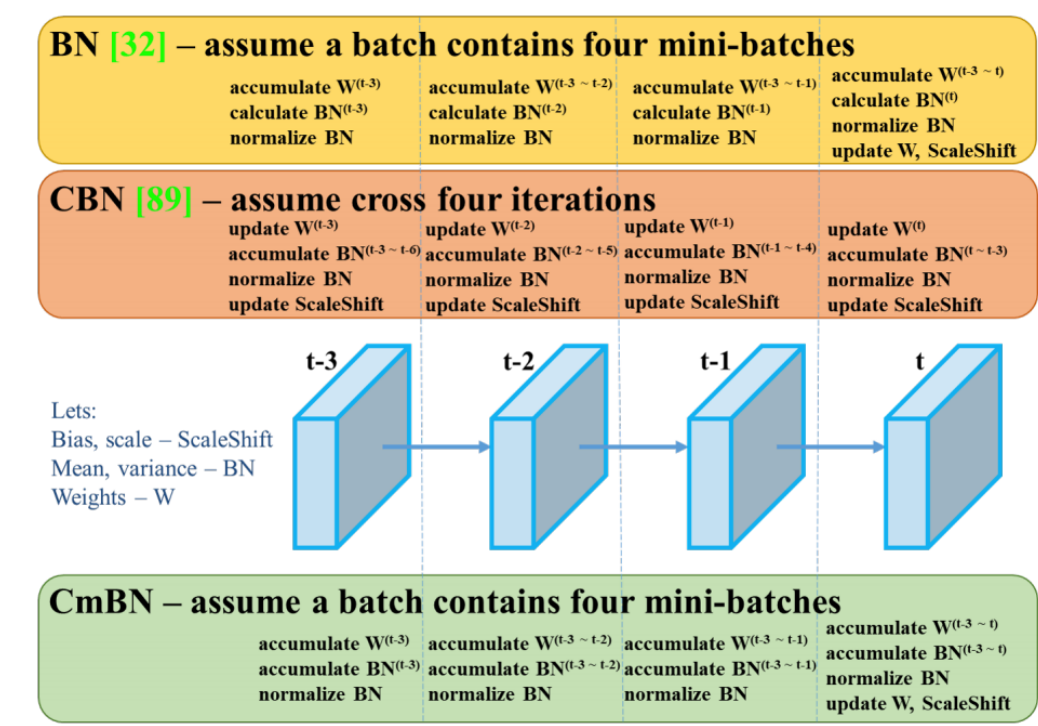


圖 4：CmBN 圖示。

研究者還將 SAM 從空間注意力機制（spatial-wise attention）修改為點注意力機制（point-wise attention），並將 PAN 中的捷徑連接替換為級聯，如下圖 5、6 所示：

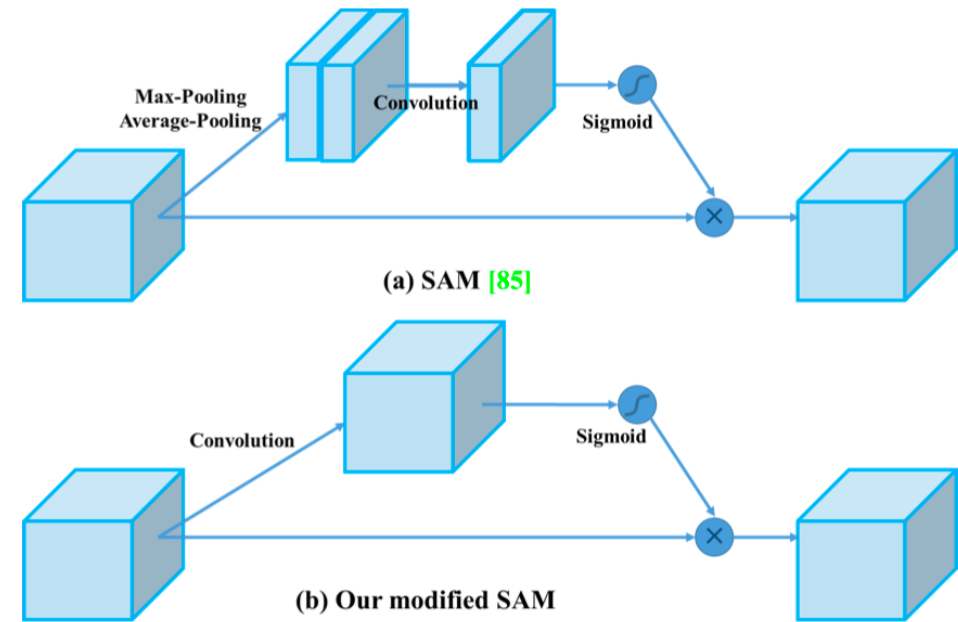


圖 5：Modified SAM。

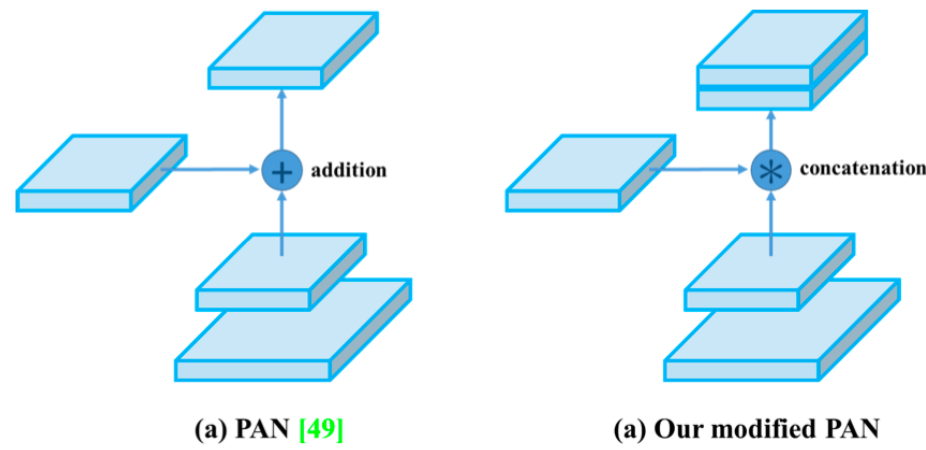


圖 6：Modified PAN。

實驗結果

YOLO v4 與其他 SOTA 目標檢測器的對比結果如下圖 8 所示。從圖上可以看出，YOLOv4 位於帕累托最優曲線上，並在速度和準確性上都優於最快和最精準的檢測器。

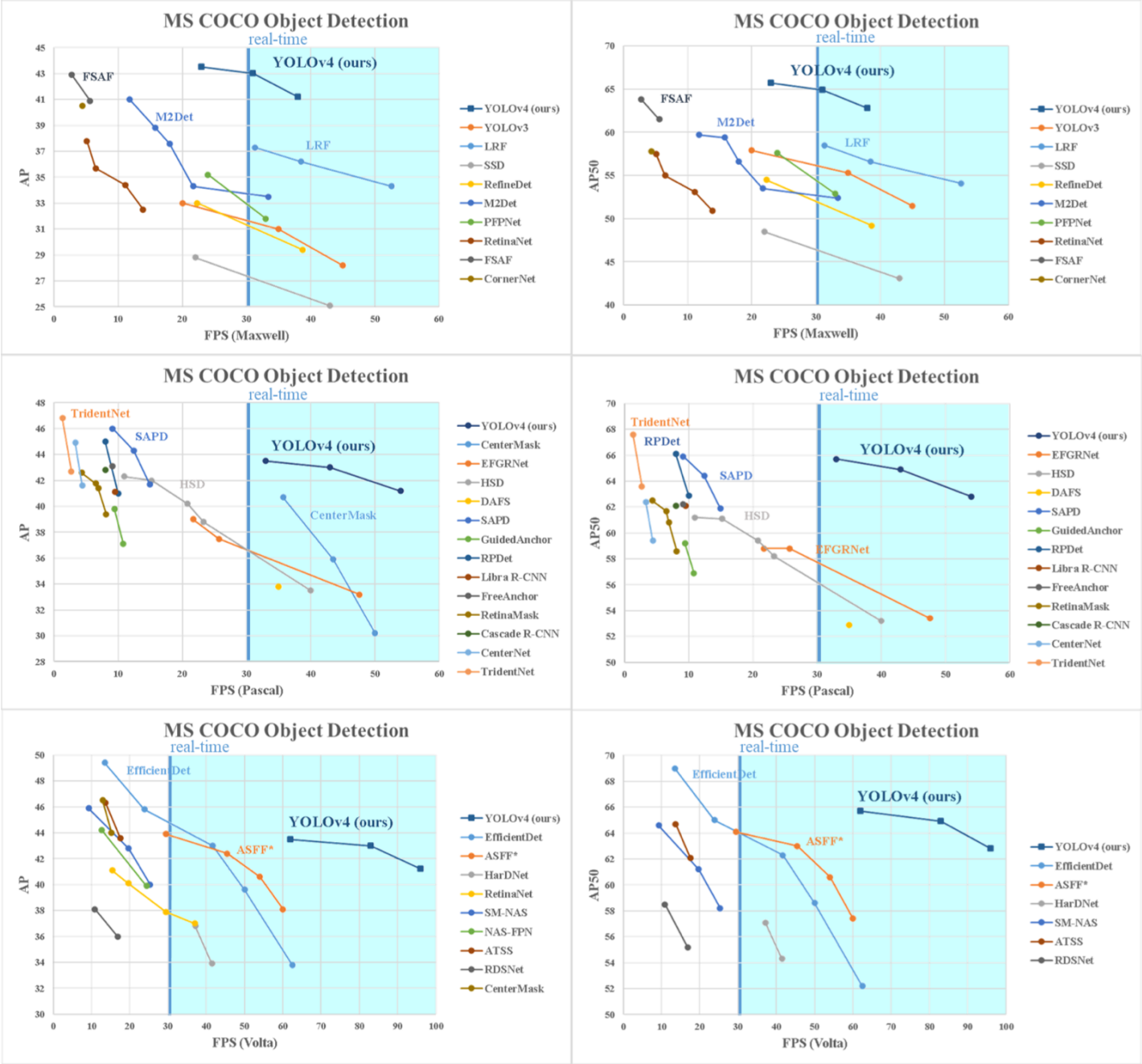


圖 8：不同目標檢測器的速度和準確性對比結果。

文章來源：機器之心 (<https://www.jiqizhixin.com/articles/2020-04-24-4>)

喜歡這篇文章嗎？快分享吧！

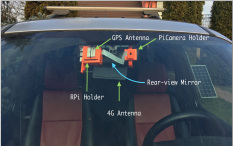
猜你會想看



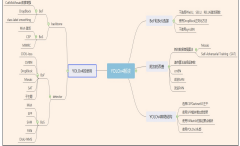
7 Papers & Radios | YOLO v4它來了；北航 ManqaGAN生成久保...



「我為什麼反對AI論文強制開源代碼」



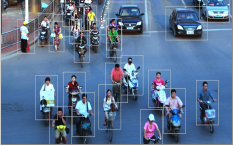
你有特斯拉，我有樹莓派，純手工打造車載車牌識別檢測系統，家...



YOLOv4算法解讀（思維導圖）和論文翻譯



CVPR 2020 | 可擴展且高效，谷歌提出目標檢測「新標杆」



PyTorch版EfficientDet比官方TF實現快25倍？這個GitHub項目數天...

by