# 使用YOLO提高实时目标检测能力 使用YOLO提高实时目标检测能力 原文链接： 一个关于实时目标检测的新观点 近年来，在深度学习的帮助下，目标检测领域取得了巨大的进展。目标检测是在一个图片中识别物体并在物体周围画一个框标出来，即确定它们的位置。这在计算机视觉邻域是一个十分重要的问题，因为它经常应用于自动驾驶汽车的安全和追踪的应用上。 早先的目标检测方法通常是将图片按一定顺序分离成不同的部分，这就导致完成的每一部分和最后的结果没有连接，就是将图像中的对象画一个框包围。一个端对端框架在降低一个联合的图中的识别错误方面是一个很好的加爵办法，它不仅是训练模型使其有更高的准确率，还可以提高它的识别速度。

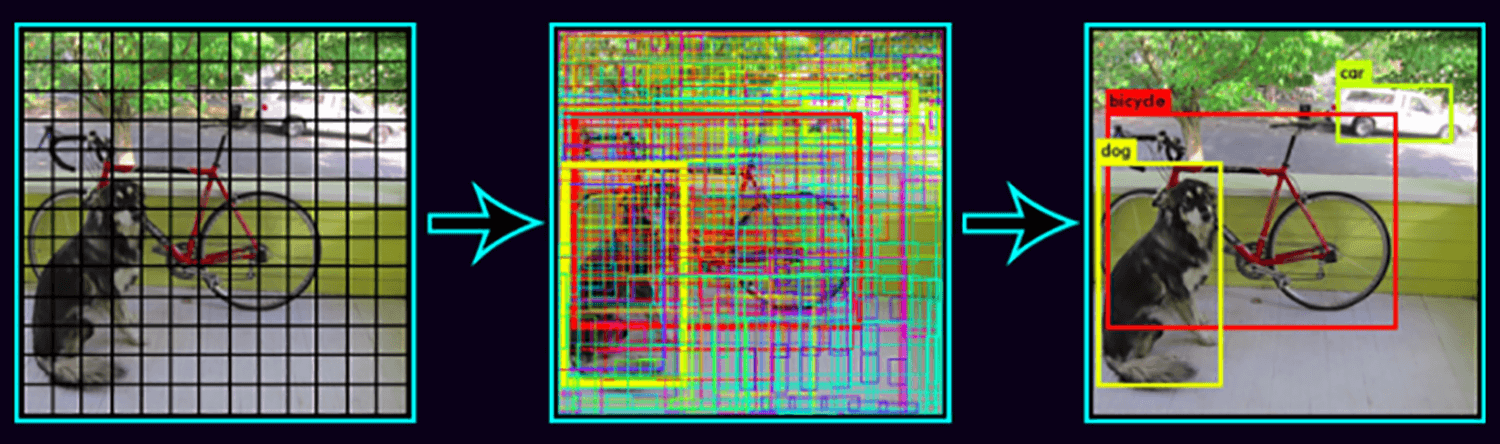
原文链接：[Improving Real-Time Object Detection with YOLO](https://blog.statsbot.co/real-time-object-detection-yolo-cd348527b9b7?from=hackcv&hmsr=hackcv.com&utm_medium=hackcv.com&utm_source=hackcv.com)

## 一个关于实时目标检测的新观点

近年来，在深度学习的帮助下，目标检测领域取得了巨大的进展。目标检测是在一个图片中识别物体并在物体周围画一个框标出来，即确定它们的位置。这在计算机视觉邻域是一个十分重要的问题，因为它经常应用于自动驾驶汽车的安全和追踪的应用上。

早先的目标检测方法通常是将图片按一定顺序分离成不同的部分，这就导致完成的每一部分和最后的结果没有连接，就是将图像中的对象画一个框包围。一个端对端框架在降低一个联合的图中的识别错误方面是一个很好的加爵办法，它不仅是训练模型使其有更高的准确率，还可以提高它的识别速度。

这就是You Only Look Once（YOLO）算法可以应用的场景。Varun Agrawa 告诉的团队为什么YOLO比别的算法更适合在目标检测问题上应用。  
这就是You Only Look Once（YOLO）算法可以应用的场景。Varun Agrawa 告诉的团队为什么YOLO比别的算法更适合在目标检测问题上应用。  
img  
深度学习已经被证明了使图像分类问题最有效的工具，并已经在这个问题上具备了人工级别的能力。早期的检测方法利用这种能力将目标检测转换为分类问题之一，即识别图像目标属于哪种类型的目标。  
这样的方式是通过2个步骤完成的：  
1.第一步包括产生数以万计的想法，他们都只是图像上特定的矩形区域，也被称为边界框。边界框可以围绕图像中的对象进行，也可以不。而对这些想法进行过滤是第二步的目标。



img

[Illustration source](https://arxiv.org/abs/1506.02640)

深度学习已经被证明了使图像分类问题最有效的工具，并已经在这个问题上具备了人工级别的能力。早期的检测方法利用这种能力将目标检测转换为分类问题之一，即识别图像目标属于哪种类型的目标。

这样的方式是通过2个步骤完成的：

1.第一步包括产生数以万计的想法，他们都只是图像上特定的矩形区域，也被称为边界框。边界框可以围绕图像中的对象进行，也可以不。而对这些想法进行过滤是第二步的目标。

2.第二步是一个图像分类器会对边界框内包括的子图像进行分类，判断它是否是一个特别的目标类型的一部分又或者只是简单的无目标的或者只是背景。  
2.第二步是一个图像分类器会对边界框内包括的子图像进行分类，判断它是否是一个特别的目标类型的一部分又或者只是简单的无目标的或者只是背景。  
这两步过程虽然非常精确，但是仍然存在一些缺陷，例如效率问题，这是由于生成了大量的想法，并且缺乏对想法生成和分类的联合优化。这导致每个阶段都不能真正理解全局，而只是被困在自己的小问题中，从而限制了他们整体的性能。  
YOLO到底是什么  
这就是YOLO引入的地方。YOLO代表着你只看一次，是一种基于目标识别算法的深度学习。在2016年由华盛顿大学的 开发出来的。

这两步过程虽然非常精确，但是仍然存在一些缺陷，例如效率问题，这是由于生成了大量的想法，并且缺乏对想法生成和分类的联合优化。这导致每个阶段都不能真正理解全局，而只是被困在自己的小问题中，从而限制了他们整体的性能。

### **YOLO到底是什么**

这就是YOLO引入的地方。YOLO代表着你只看一次，是一种基于目标识别算法的深度学习。在2016年由华盛顿大学的 [Joseph Redmon and Ali Farhadi](https://arxiv.org/abs/1506.02640) 开发出来的。

YOLO系统背后的理论基础不再是过去的那种传递多个可能的子图片，而是仅向深度学习系统传递一次完整的图片。接着，你将会得到所有的边界框以及目标类别分类。这是YOLO的基本设计策略，这是在目标识别领域的一个崭新的视角。  
YOLO系统背后的理论基础不再是过去的那种传递多个可能的子图片，而是仅向深度学习系统传递一次完整的图片。接着，你将会得到所有的边界框以及目标类别分类。这是YOLO的基本设计策略，这是在目标识别领域的一个崭新的视角。  
YOLO工作的方式是将图像细分成NXN网格，或者更具体地说，在原始文件7x7网格中的每个网格单元，也称为锚，表示一个分类器，它负责在潜在对象的周围生成K个边界框，潜在对象的基本真值中心落在该网格单元内（本文中K是2），并将其分类为正确的对象。

YOLO工作的方式是将图像细分成NXN网格，或者更具体地说，在原始文件7x7网格中的每个网格单元，也称为锚，表示一个分类器，它负责在潜在对象的周围生成K个边界框，潜在对象的基本真值中心落在该网格单元内（本文中K是2），并将其分类为正确的对象。

注意，边界框不限于网格单元内，它可以在图像的边界内扩展，以容纳它认为负责检测的对象。这意味着在当前版本的YOLO中，系统生成98个不同大小的边界框，以适应场景中的各种对象。  
注意，边界框不限于网格单元内，它可以在图像的边界内扩展，以容纳它认为负责检测的对象。这意味着在当前版本的YOLO中，系统生成98个不同大小的边界框，以适应场景中的各种对象。  
性能和结果  
对于更密集的对象检测，用户可以根据他们的需要将K或N设 置为更高的数字。然而，在当前的配置中，我们有一个系统，该系统能够输出大量围绕对象的边界框，并且基于图像的空间布局将它们分类到各种类别。  
在运行时间里只有一次通过图像。因此，联合检测和分类可以更好地优化学习目标（损失函数）和实时性能。

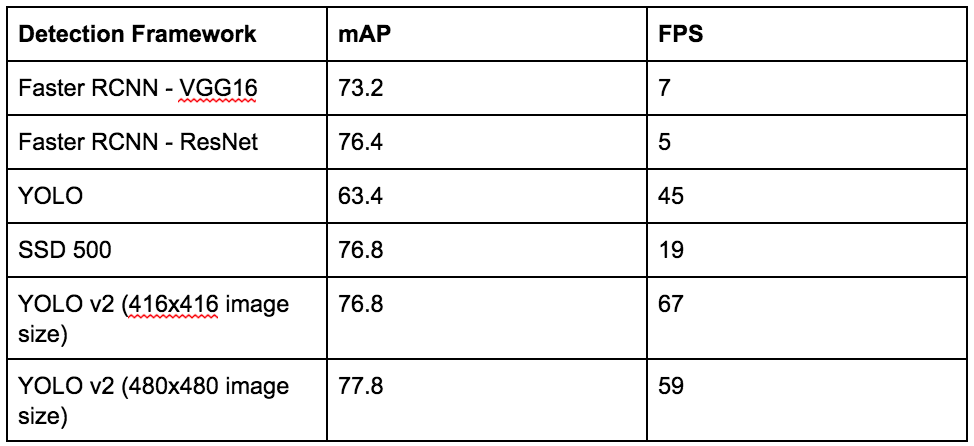
### 性能和结果

对于更密集的对象检测，用户可以根据他们的需要将K或N设 置为更高的数字。然而，在当前的配置中，我们有一个系统，该系统能够输出大量围绕对象的边界框，并且基于图像的空间布局将它们分类到各种类别。

在运行时间里只有一次通过图像。因此，联合检测和分类可以更好地优化学习目标（损失函数）和实时性能。

事实上，YOLO的结果是非常理想的。在具有挑战性的上，YOLO以每秒45帧的速度运行时，实现了63.4（每100帧中）的平均精度或mAP。相比之下，根据现有模型，Faster-RCNN VGG 16实现了73.2的mAP ，但是仅以最大每秒7帧的速度运行，效率降低了6倍。  
事实上，YOLO的结果是非常理想的。在具有挑战性的上，YOLO以每秒45帧的速度运行时，实现了63.4（每100帧中）的平均精度或mAP。相比之下，根据现有模型，Faster-RCNN VGG 16实现了73.2的mAP ，但是仅以最大每秒7帧的速度运行，效率降低了6倍。  
你可以在下面的图表中看到YOLO与其他模型的不同。  
img  
YOLO在牺牲一些精度的条件下，它可以以每秒155帧运行，虽然只有在52.7的mAP。  
因此，YOLO的主要优势是其在实时速度下的目标检测中稳定的良好性能。这使它能在机器人、自动驾驶汽车和无人机等系统中使用，在这些系统中，时间至关重要。

你可以在下面的图表中看到YOLO与其他模型的不同。



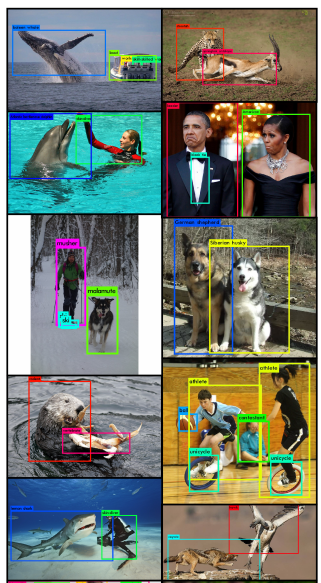
img

*YOLO在牺牲一些精度的条件下，它可以以每秒155帧运行，虽然只有在52.7的mAP。*

因此，YOLO的主要优势是其在实时速度下的目标检测中稳定的良好性能。这使它能在机器人、自动驾驶汽车和无人机等系统中使用，在这些系统中，时间至关重要。

### YOLOv2 的框架 YOLOv2 的框架 最近，同样的研究人员发布了新的YOLOv2框架，该框架利用最近在深层学习网络设计中得到的结果来构建更有效的网络，并且使用Faster-RCNN的锚箱思想来减轻网络的学习问题。 img YOLOv2有着更好的检测结果，在Pascal VOC检测数据集上以78.6mAP实现最先进的性能，而其他系统，例如改进的Faster-RCNN(Faster-RCNN ResNet)和，在相同的测试数据上仅实现76.4mAP和76.8mAP。 关键的区别是性能速度。性能最好的YOLVO2模型运行在40 FPS相比5 FPS的 FAST-RCNN RESNET 虽然 SSD500以45FPS运行，但是具有mAP 76.8（与SSD500相同）的低分辨率版本的YOLOv2在67FPS下运行，由于YOLOv2的设计选择，这向我们展示了YOLOv2的高性能能力。

最近，同样的研究人员发布了新的YOLOv2框架，该框架利用最近在深层学习网络设计中得到的结果来构建更有效的网络，并且使用Faster-RCNN的锚箱思想来减轻网络的学习问题。



img

[插图来源](http://www.pjreddie.com/)

YOLOv2有着更好的检测结果，在Pascal VOC检测数据集上以78.6mAP实现最先进的性能，而其他系统，例如改进的Faster-RCNN(Faster-RCNN ResNet)和[SSD500](https://arxiv.org/pdf/1512.02325.pdf)，在相同的测试数据上仅实现76.4mAP和76.8mAP。

*关键的区别是性能速度。性能最好的YOLVO2模型运行在40 FPS相比5 FPS的 FAST-RCNN RESNET*

虽然 SSD500以45FPS运行，但是具有mAP 76.8（与SSD500相同）的低分辨率版本的YOLOv2在67FPS下运行，由于YOLOv2的设计选择，这向我们展示了YOLOv2的高性能能力。

### 最后的一些感想

综上所述，YOLO在实时性能上运行时表现出了显著的性能增益，这是在资源匮乏的深度学习算法时代一个重要的中间环节。随着我们向着更加自动化的未来迈进，像YOLO和SSD500这样的系统已经准备好迎来大步的进步，逐步实现伟大的AI梦想。

### 一些重要的阅读文章

* [You Only Look Once: Unified, Real-Time Object Detection](https://arxiv.org/abs/1506.02640)
* [The PASCAL Visual Objects Challenge: A Retrospective](http://host.robots.ox.ac.uk/pascal/VOC/pubs/everingham15.pdf)
* [SSD: Single Shot Multibox Detector](https://arxiv.org/pdf/1512.02325.pdf)