

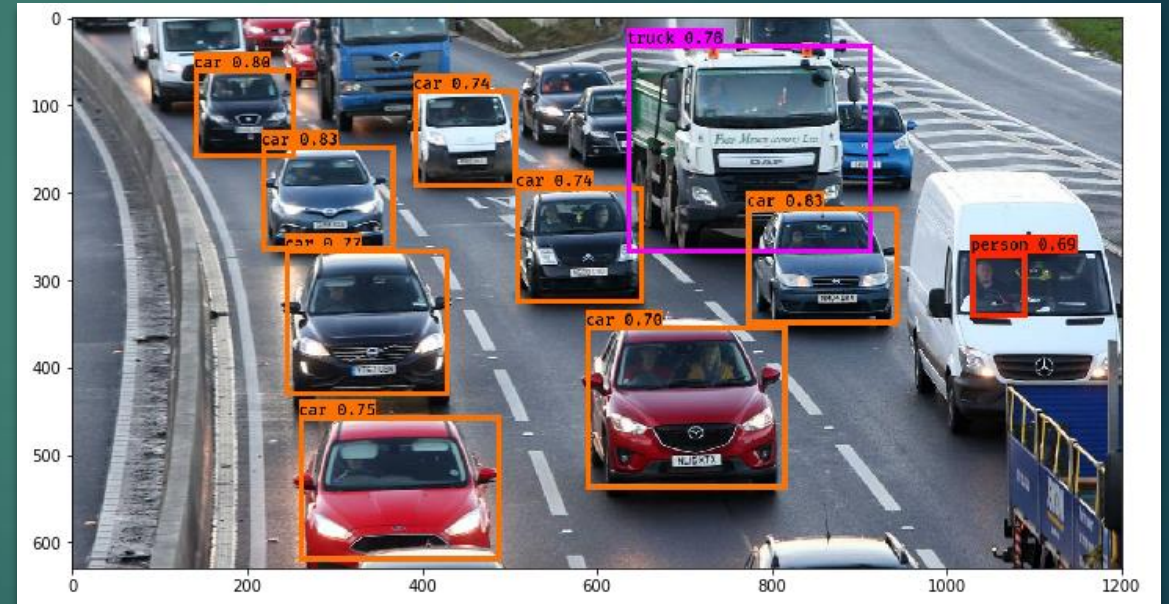
A large, stylized, light blue watermark of the word "YOLO" is centered in the background. A solid red vertical rectangle is positioned in the top right corner.

YOLO (You Only Look Once)

REAL-TIME OBJECT DETECTION

Principais vantagens

- ▶ Velocidade (45 frames por segundo)
- ▶ Código aberto
- ▶ Eficiência
- ▶ Realiza a detecção em uma única etapa

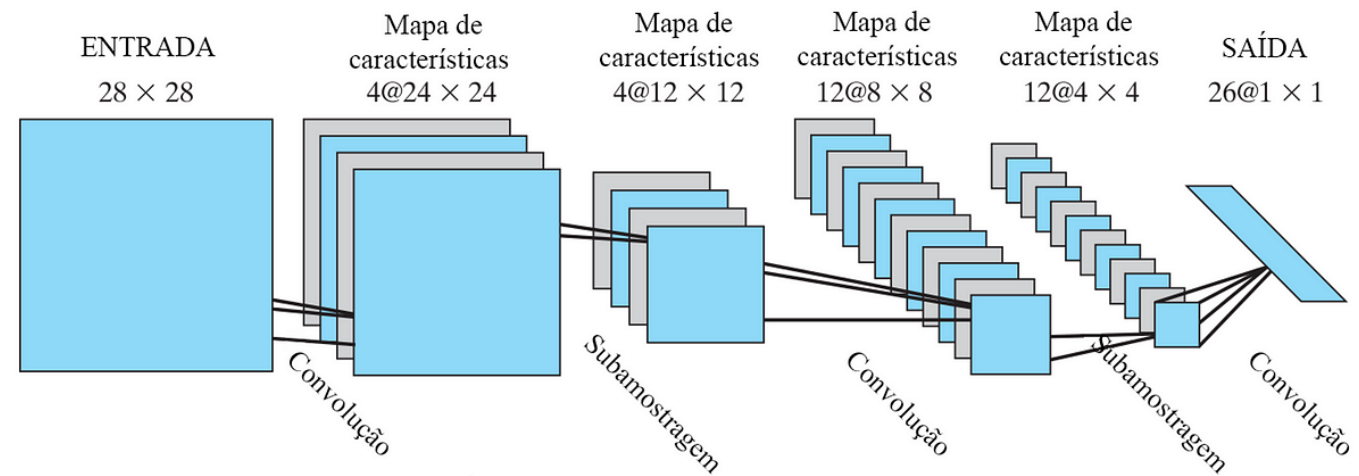


O que é o YOLO?

É um método de detecção de objetos de passada única (single pass) que utiliza uma rede convolucional como extrator de características (features).

O que é uma rede convolucional?

São arquiteturas similares às Redes Neurais Artificiais, porém apresentam mais camadas e operações. Nessas redes, cada camada é responsável por extrair determinadas informações dos dados de entrada. A informação flui através de cada camada da rede, com a saída da camada anterior fornecendo a entrada para a camada seguinte da rede.

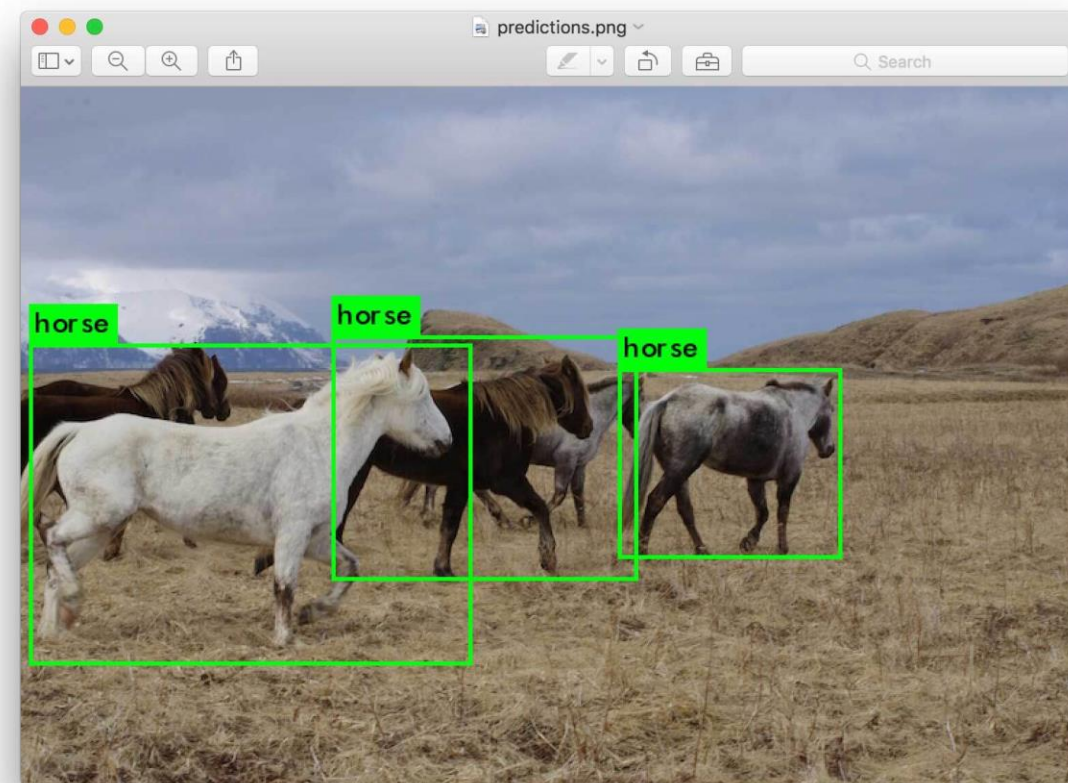


O que é o YOLO?

Diferente dos outros algoritmos de detecção de objetos, ele apenas precisa olhar pela imagem uma única vez para enviar para a rede neural.

Utiliza uma rede neural única usando as características da imagem inteira para detectar múltiplas caixas (Bounding Boxes), cada uma contendo um objeto.

Dimensões pré-definidas (âncoras) a de objetos anotados no conjunto de treinamento.

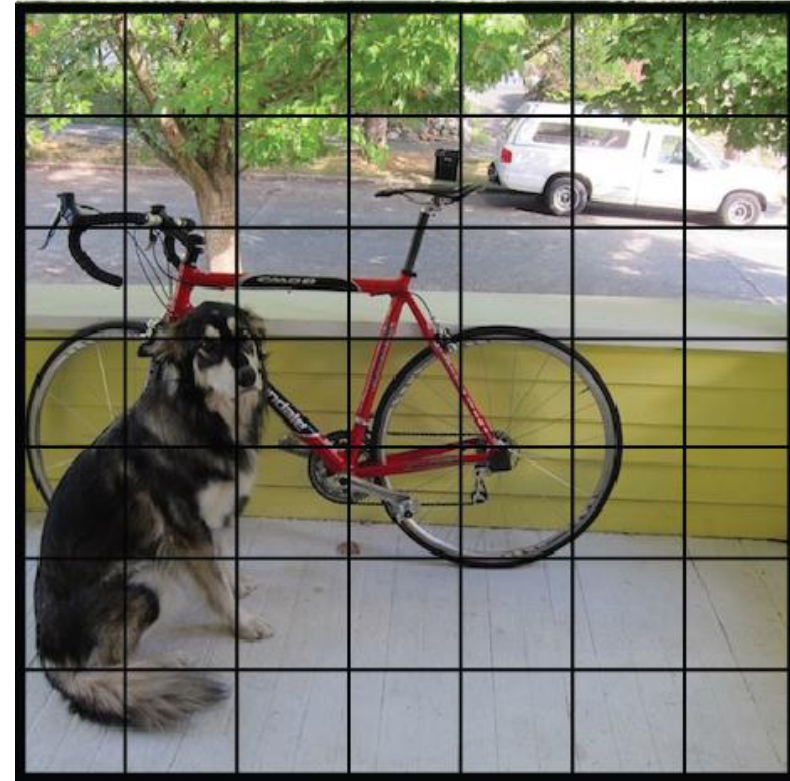


Etapas

Ele se baseia na ideia de segmentar a imagem em imagens menores

Cada uma dessas células faz a predição de B caixas delimitadoras e prevê pontuações de confiança para cada caixa (0.0 a 1.0), para caso tenha mais de um objeto naquela célula. B é configurável.

A célula na qual reside o centro de um objeto, por exemplo o centro da bicicleta, é a célula responsável por detectar esse objeto.



Etapas

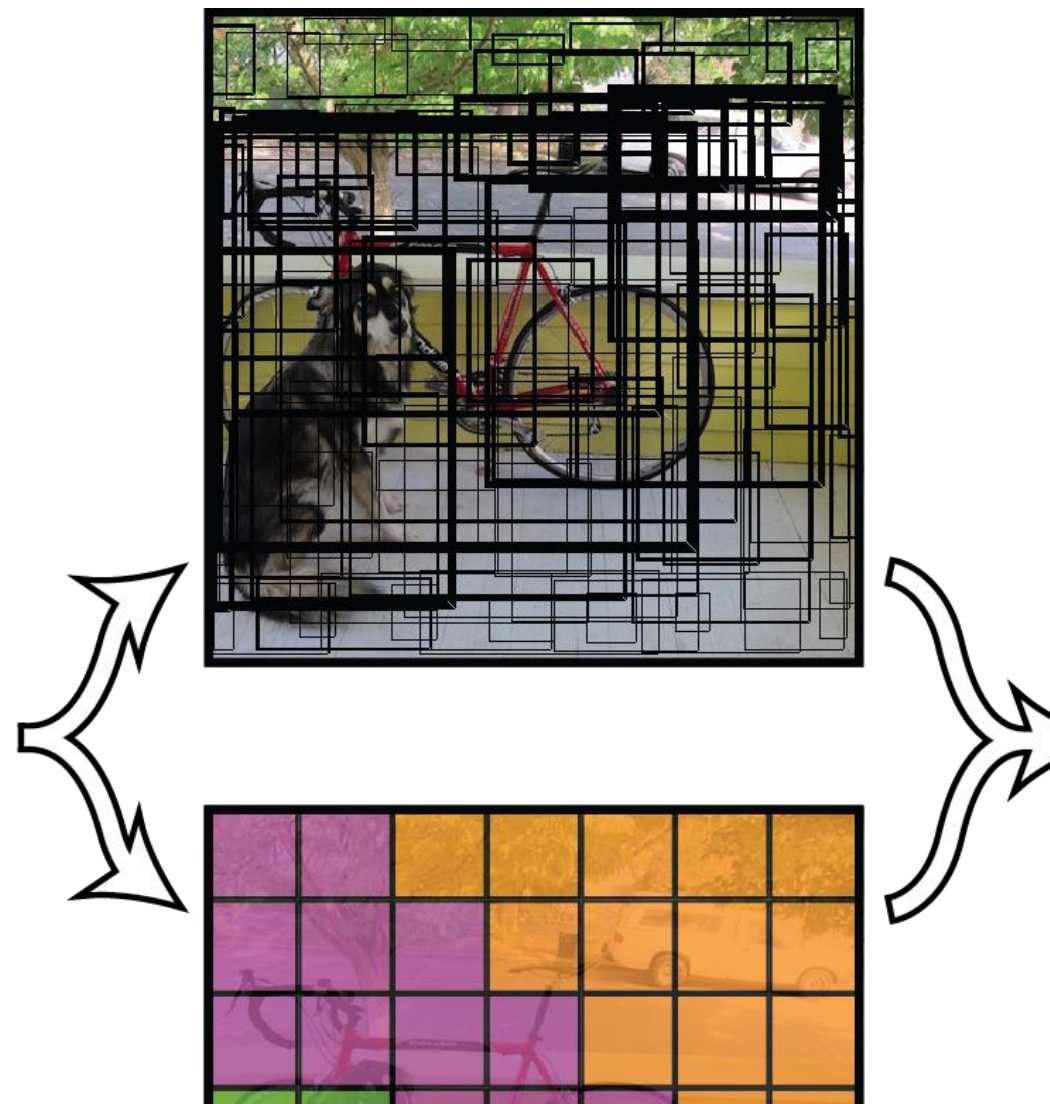
Retorna uma pontuação de confiança: quanto de certeza o algoritmo tem que aquela célula contém um objeto e a precisão da caixa delimitadora.

Cada caixa consistem em 5 previsões: x , y , w , h e confiança.

As coordenadas (x,y) representam a localização do centro da caixa delimitadora.

A largura e altura (w,h) são frações relativas ao tamanho total da imagem

Confiança representa a IoU (fração entre 0 e 1)

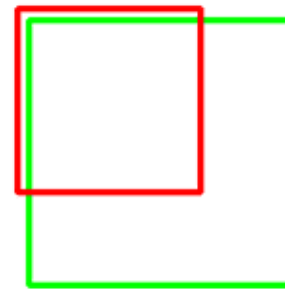


Etapas - IoU

Ajuda a avaliar a precisão do modelo. Pode descobrir quão bem a caixa prevista se sobrepõe à caixa de verdade.

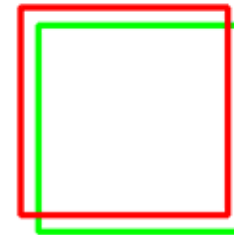
Ajuda a remover caixas duplicadas. Para isso, classifica todas as previsões / objetos em uma ordem decrescente de sua confiança. Se duas caixas apontar para o mesmo objeto, a IoU será muito alta, próxima de 1. Ele escolhe aquela com maior confiança e rejeita a outra. Da mesma forma o contrário, se for muito baixa, possivelmente significa que apontam para objetos diferentes da mesma classe (como cães ou gatos diferentes na mesma imagem).

IoU: 0.4034



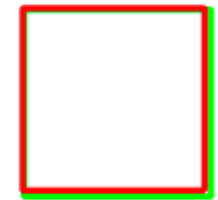
Poor

IoU: 0.7330



Good

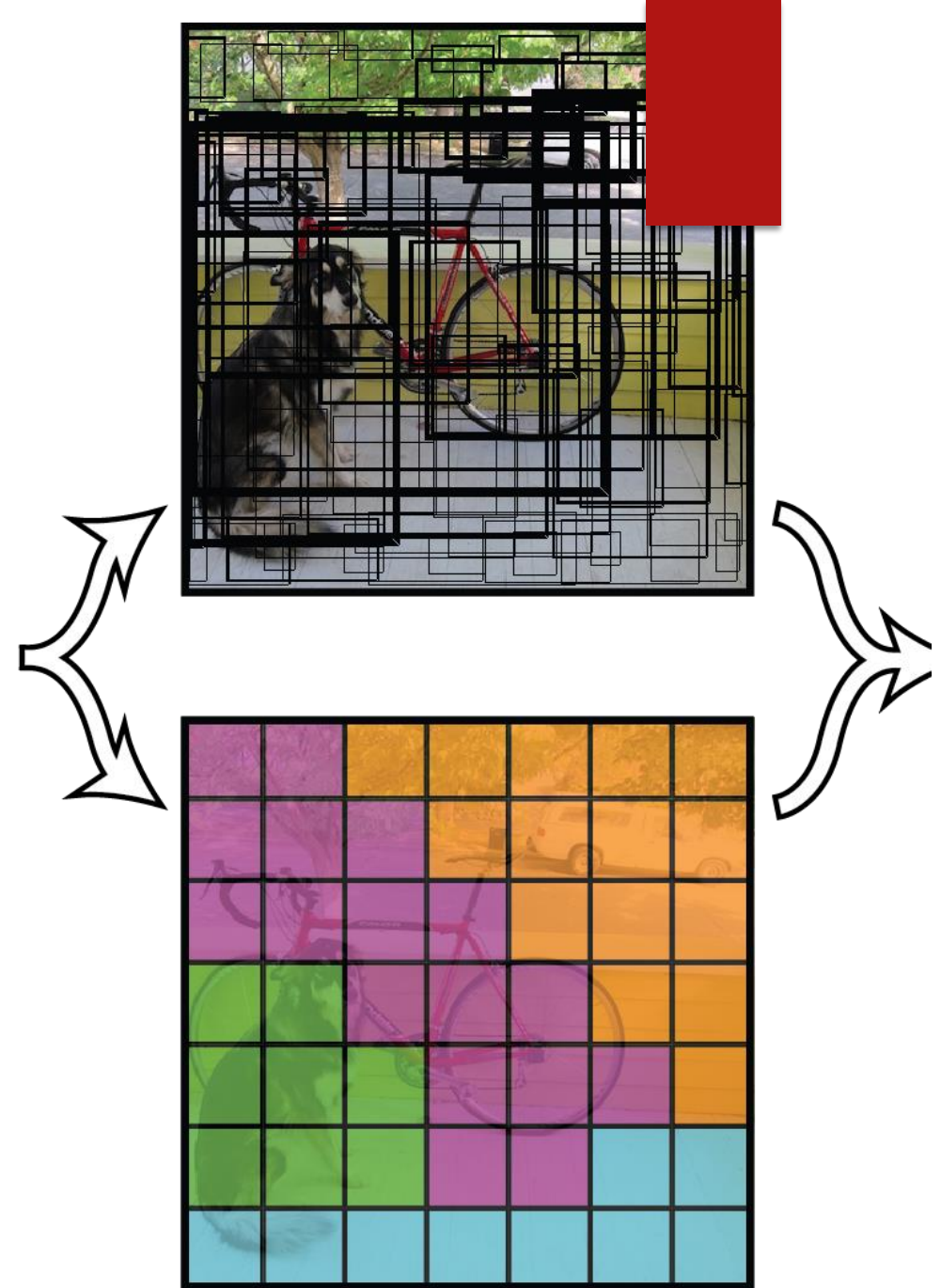
IoU: 0.9264



Excellent

Etapas

Cada célula da grade também prevê a classe condicional C probabilidades, Pr (Classe | Objeto). Essas probabilidades são condicionais na célula da grade que contém um objeto. Prevê um conjunto de probabilidade de classe por célula da grade, independentemente do número de caixas B .



Etapas

No momento do teste, multiplica a probabilidade da classe condicional e as previsões de confiança de cada caixa individual, ou seja:

$$\Pr(\text{Classe} | \text{Objeto}) * \Pr(\text{Objeto}) * \text{IoU} = \Pr(\text{Classe}) * \text{IoU}$$

O que nos dá pontuações de confiança específicas da classe para cada caixa. Essas pontuações codificam a probabilidade dessa classe aparecendo na caixa e quão bem a caixa prevista se encaixa no objeto

