

## Exercício 21 - Aula 28

Aluna: Williana Luzia Sousa Leite

Matrícula: 390239

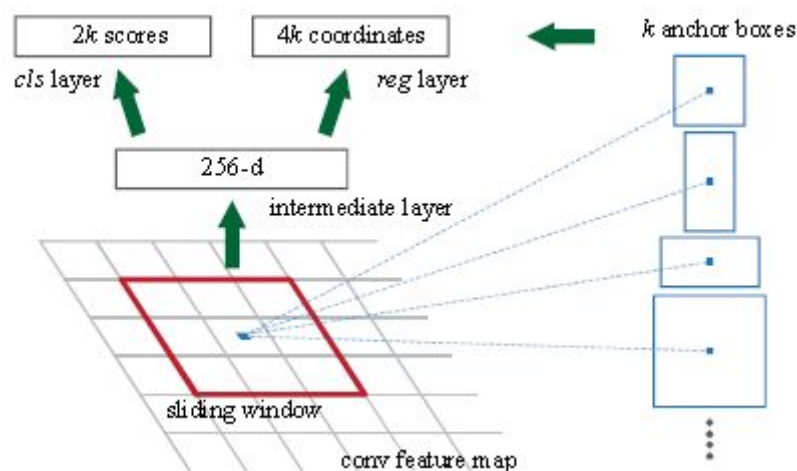
Departamento: UFC - Campus Russas

Exercício:

Pesquise e descreva como é feito o recorte da imagem para detecção de objetos nos seguintes métodos: YOLO (You Only Look Once), SSD (Single Shot Detector) e Faster-RCNN.

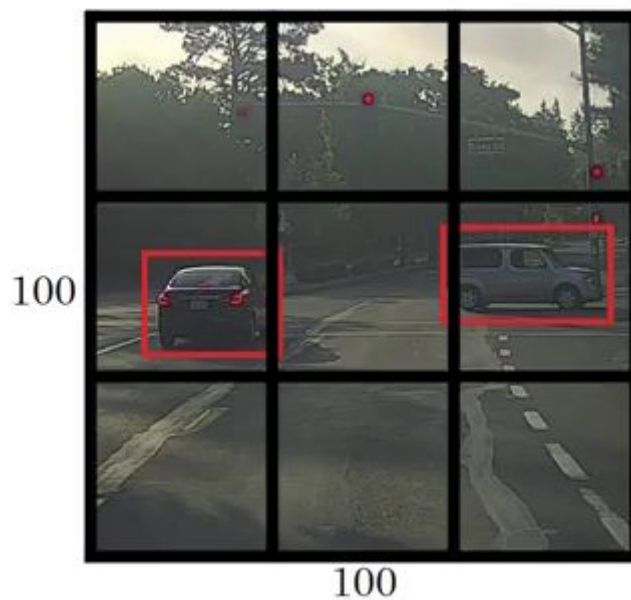
Resposta:

- Faster-RCNN
  - A proposta de regiões do Faster-RCNN é realizada através da rede Region Proposal Networks(RPN). O RPN recebe como entrada uma imagem e tem como saída um conjunto de propostas. Para encontrar essas propostas o algoritmo primeiro passa a imagem para uma rede de extração de características que tem como resultado um mapa de recursos, logo após uma janela deslizante percorre cada ponto do mapa gerando  $k$  propostas de âncoras, com a configuração padrão de 3 tamanhos com 3 proporções diferentes. Este processo é melhor compreendido analisando a imagem desenvolvida pelos autores do projeto, como apresentada a seguir.

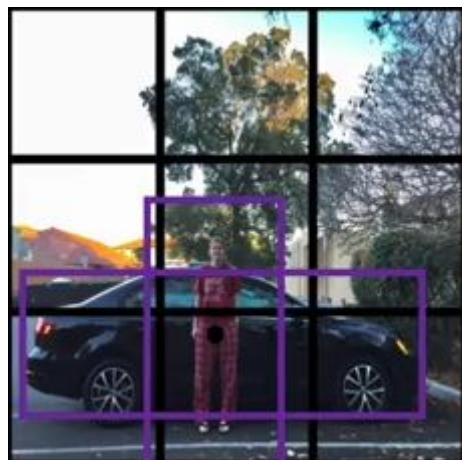


- YOLO

- De acordo com este [link](#), o yolo recebe a imagem de entrada e a divide em grade, logo após realiza classificação e localização para cada grade, como apresentado no exemplo a seguir utilizando uma grid de 3x3:



- Para solucionar o problema de mais de um objeto na mesma grid, o YOLO também utiliza o conceito de âncoras, como apresentado no exemplo a seguir:



- SSD

- Segundo este [link](#), o SSD não usa uma rede de proposta de região delegada. Em vez disso, ele utiliza um método muito simples. Ele

calcula as pontuações de localização e classe usando pequenos filtros de convolução. Depois de extrair os mapas de recursos, o SSD aplica filtros de convolução  $3 \times 3$  para cada célula para fazer previsões. (Esses filtros calculam os resultados da mesma forma que os filtros regulares da CNN.) Cada filtro gera 25 canais: 21 pontuações para cada classe mais uma caixa de limite. Como apresentado no exemplo a seguir:

