

Segmentation

- Semantic

Pre každý pixel povie kategóriu objektu

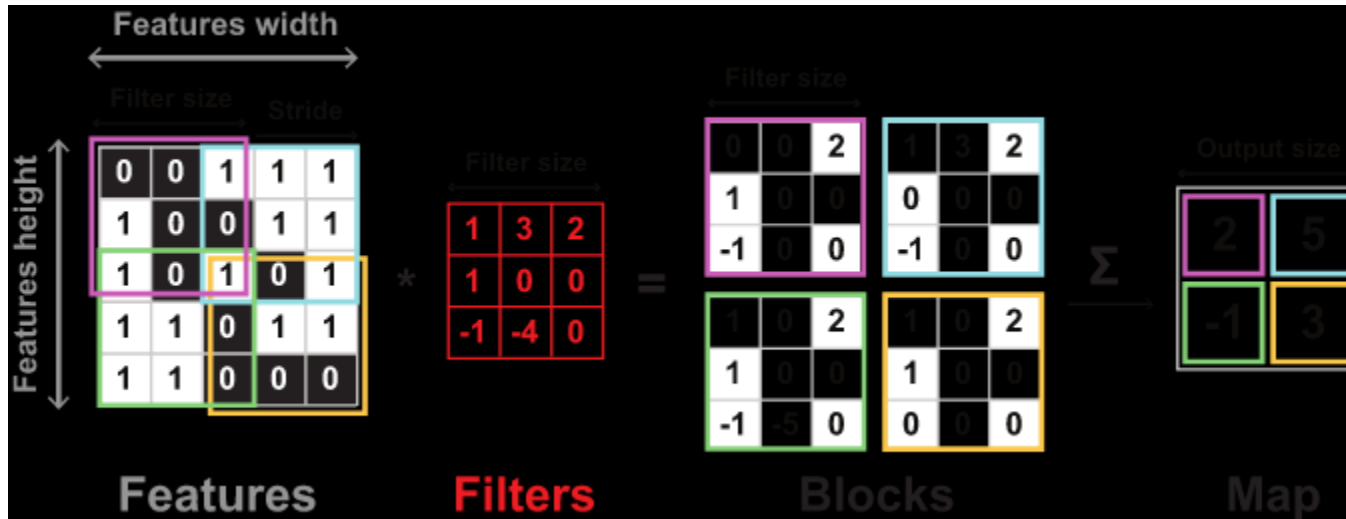
- Instance

Pre každý pixel povie kategóriu objektu + obdĺžnik každej jednej inštancie

- Panoptic

Pre každý pixel povie kategóriu a inštanciu objektu

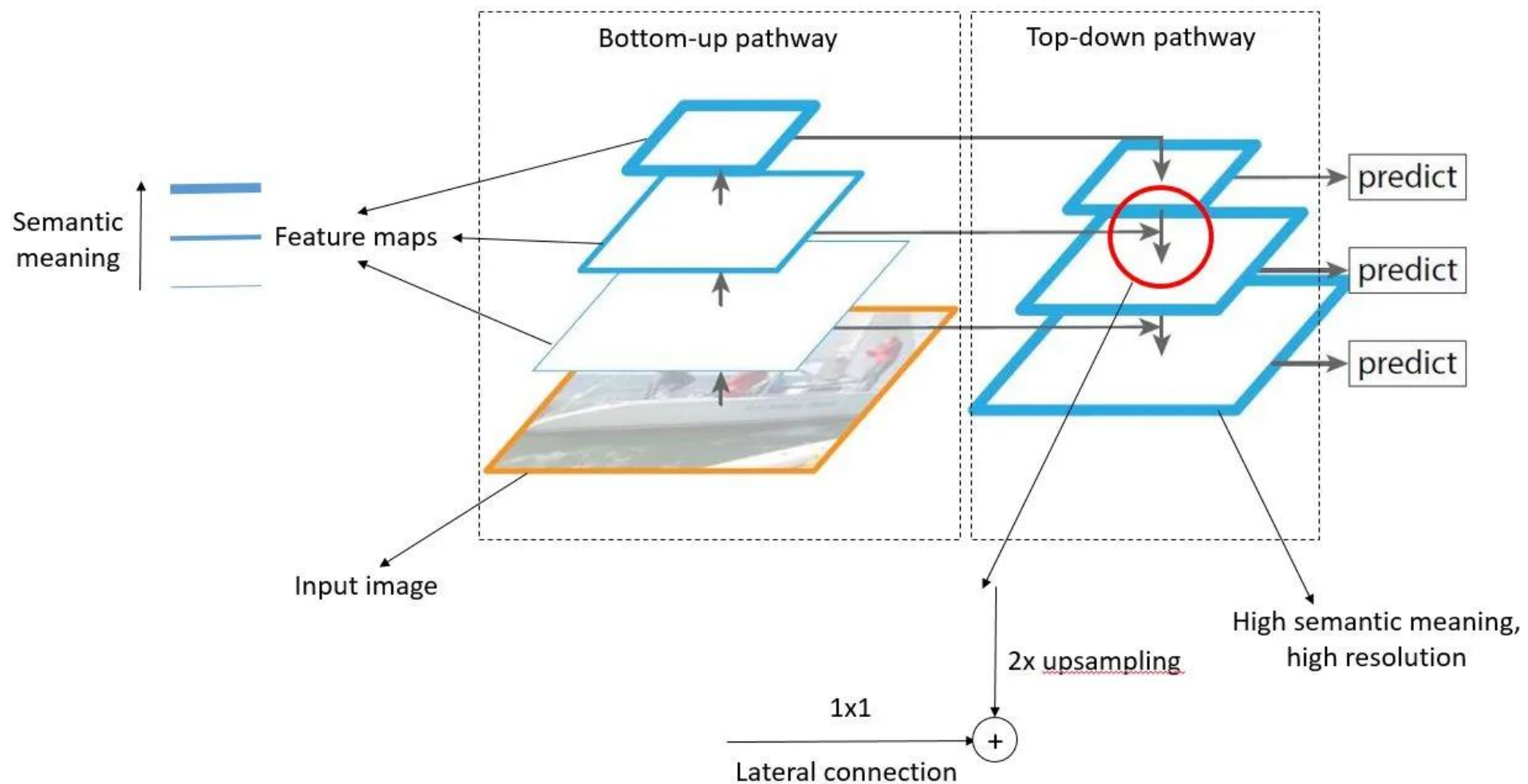
Max Pool vs. striding

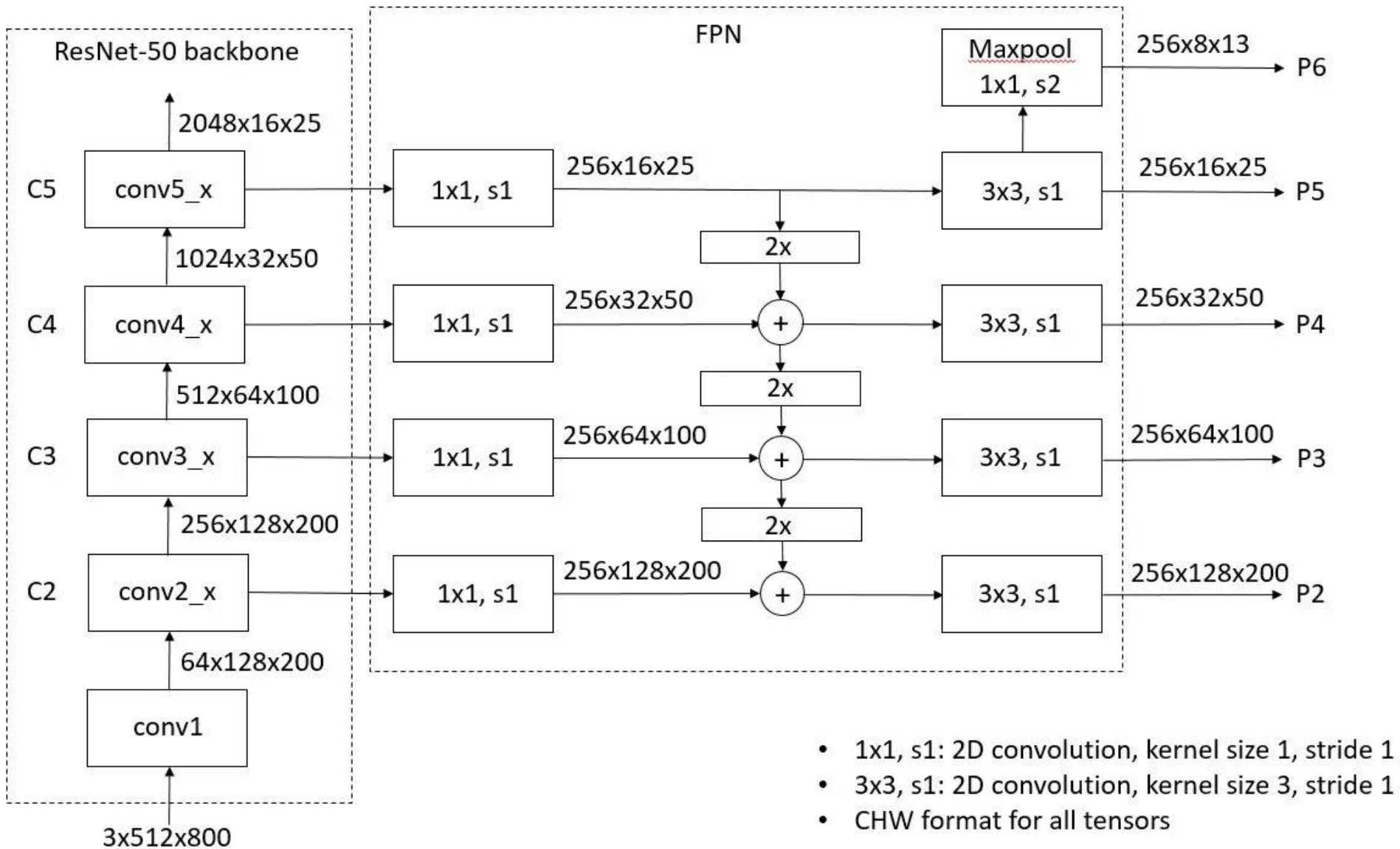


$$O = ((W - K + 2P) / S) + 1$$

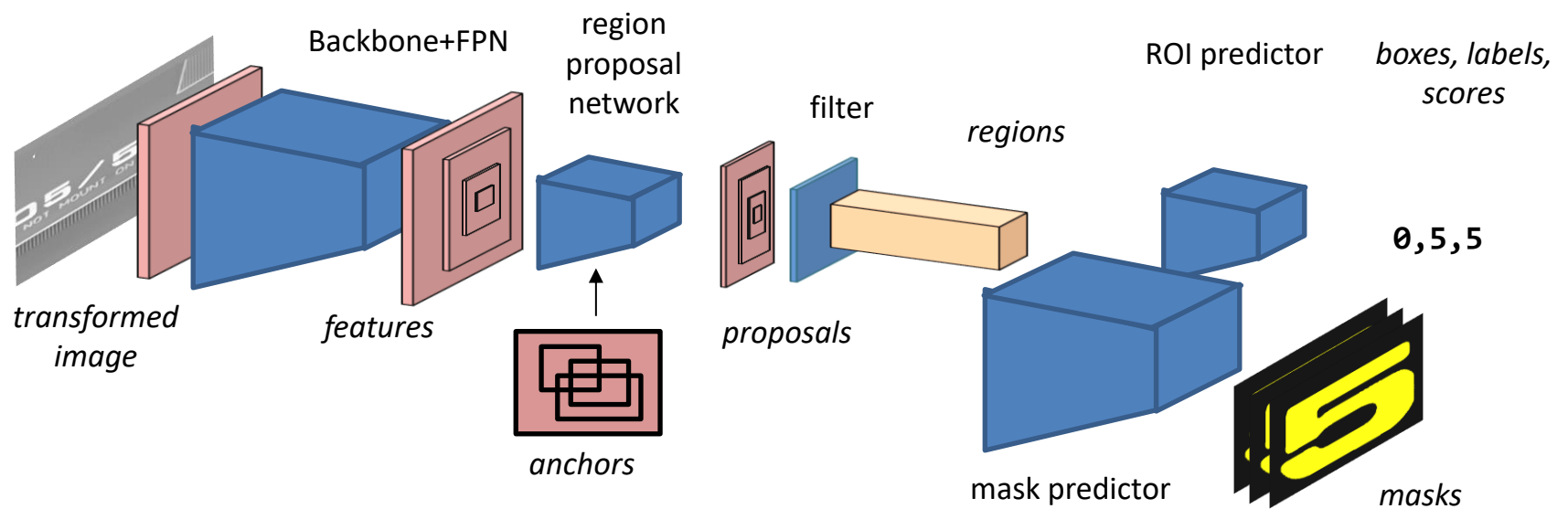
- O is the output size
- W is the input size (width or height)
- K is the kernel size
- P is the padding
- S is the stride

Feature Pyramid Network (FPN)

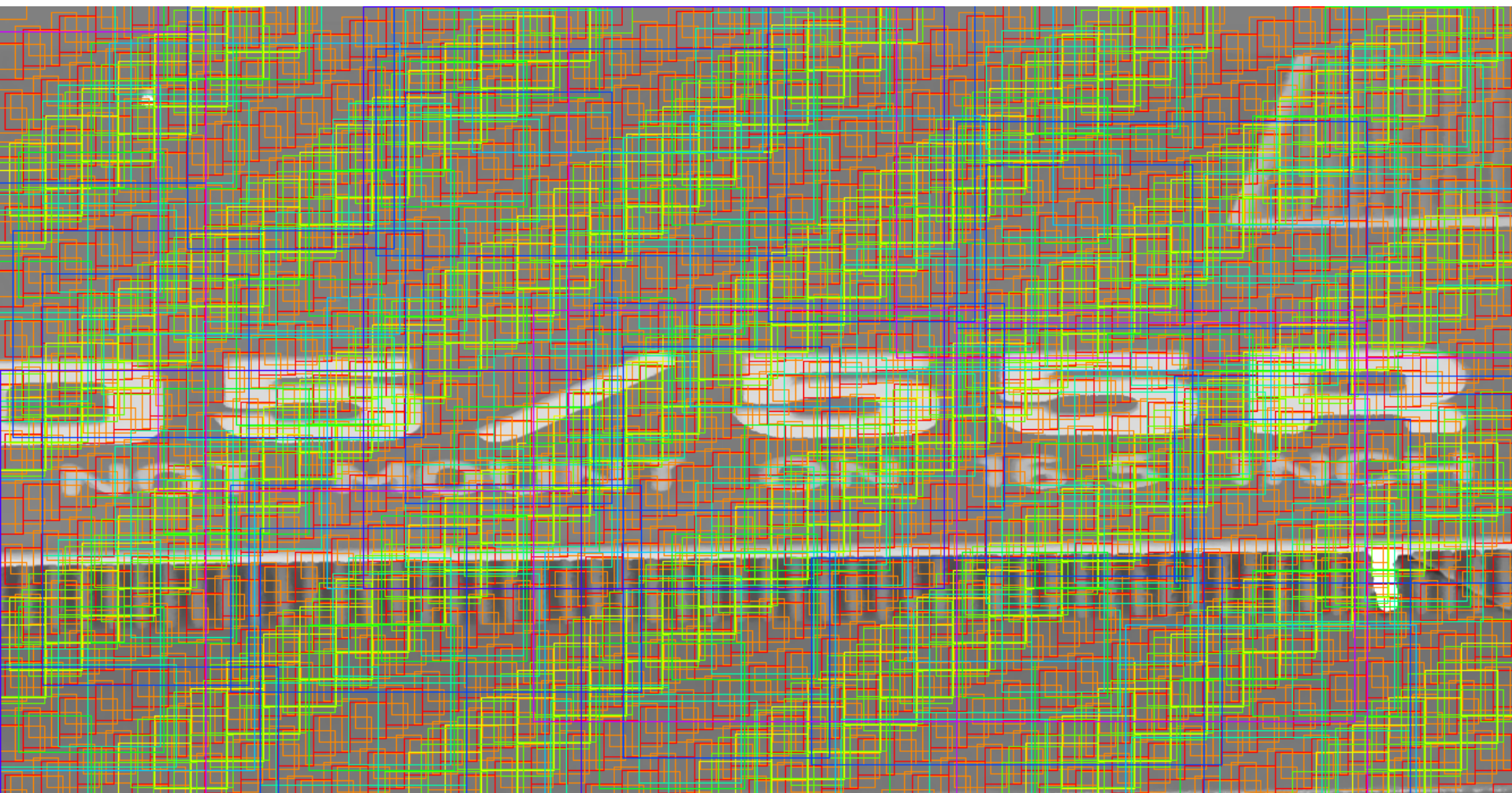




■ MaskRCNN

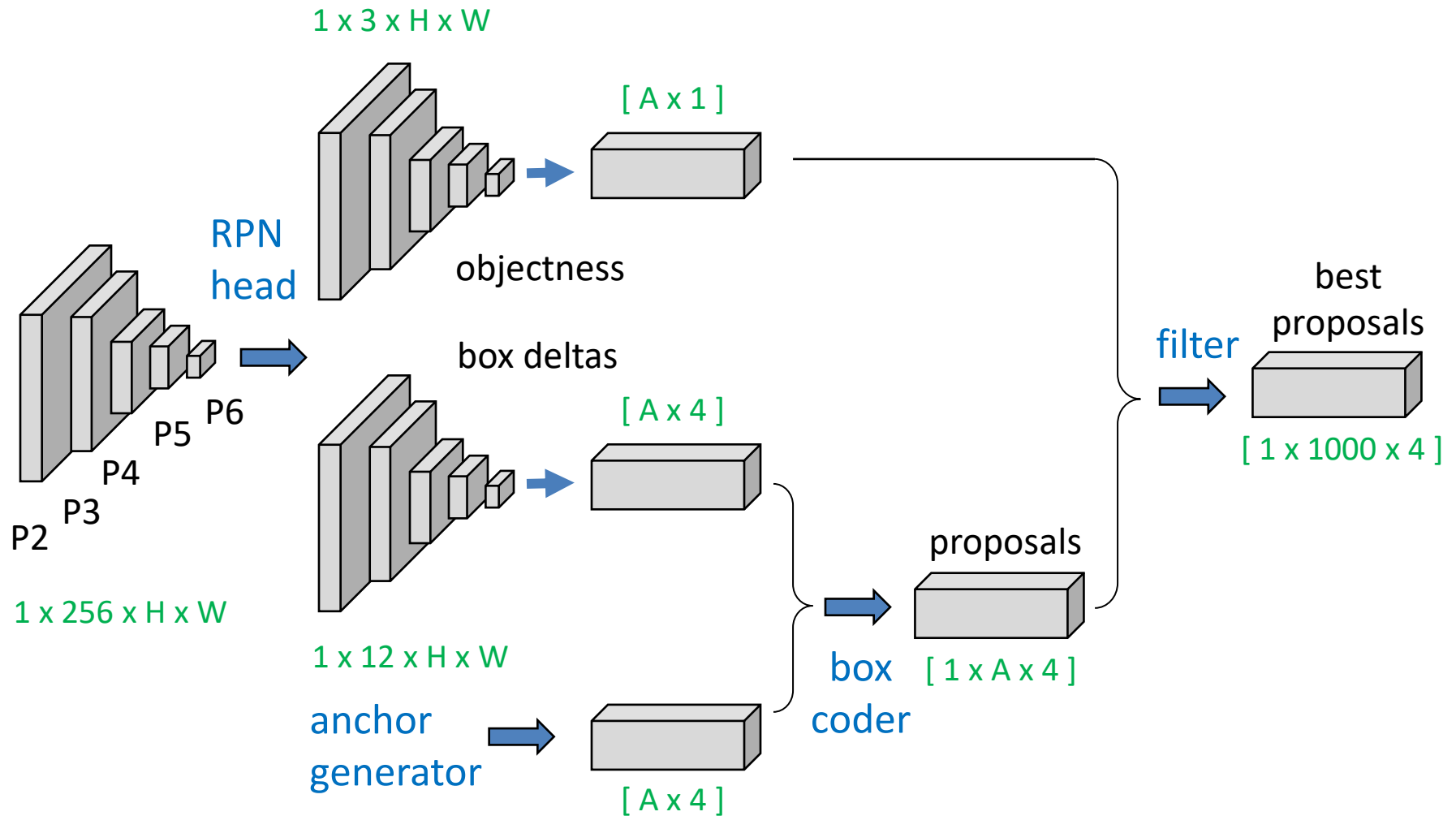


Anchors



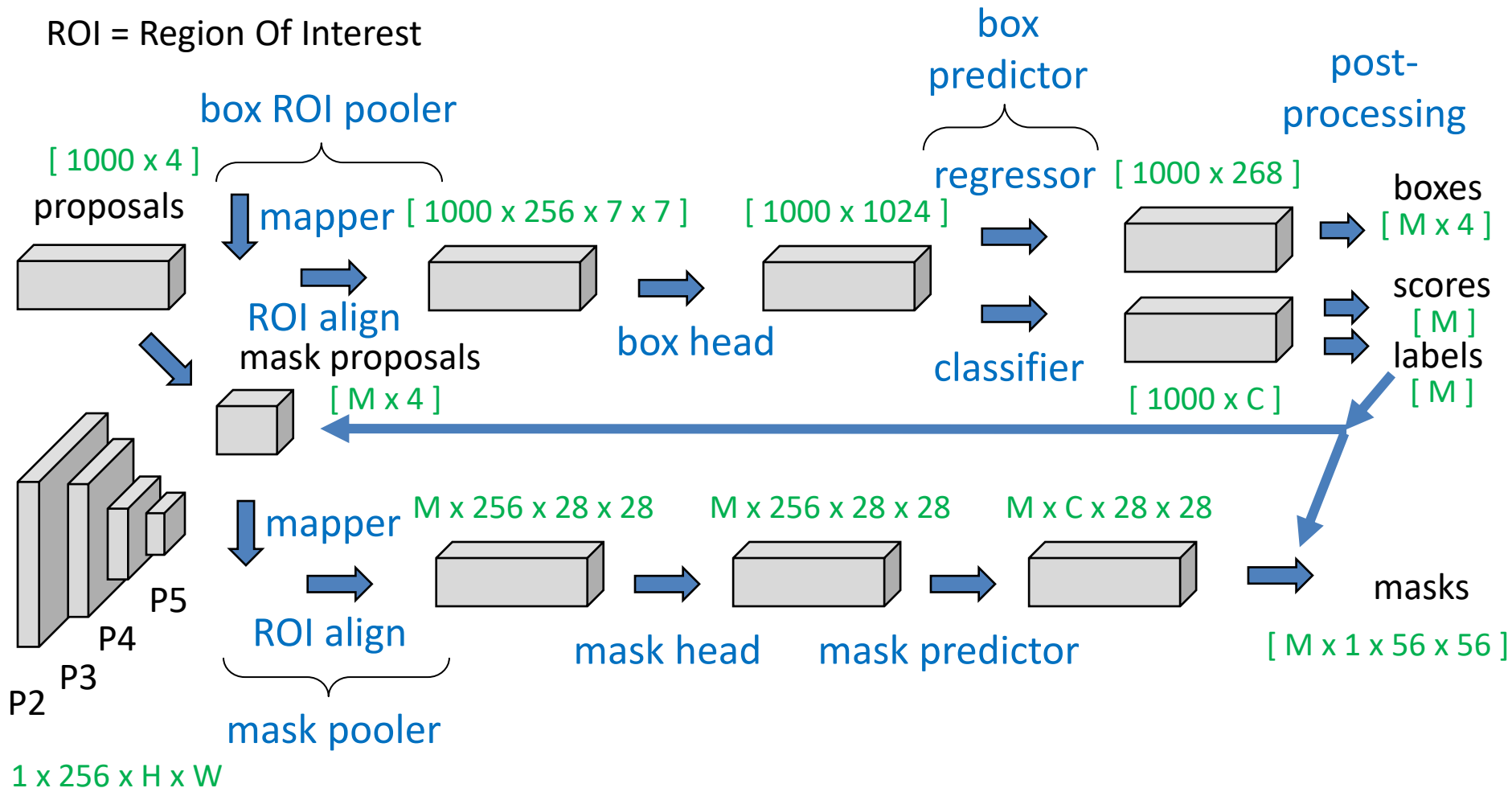
(na obrázku je každá stá kotva, modré sú najväčšie, červené najmenšie)

Region Proposal Network (RPN)



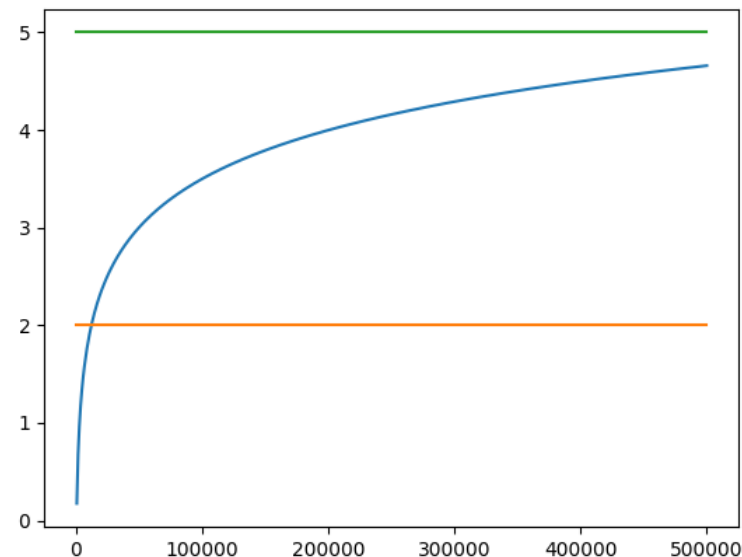
ROI Head

ROI = Region Of Interest



Mapper

- asociuje box $(x1,y1,x2,y2)$ s P_L podľa heuristiky:



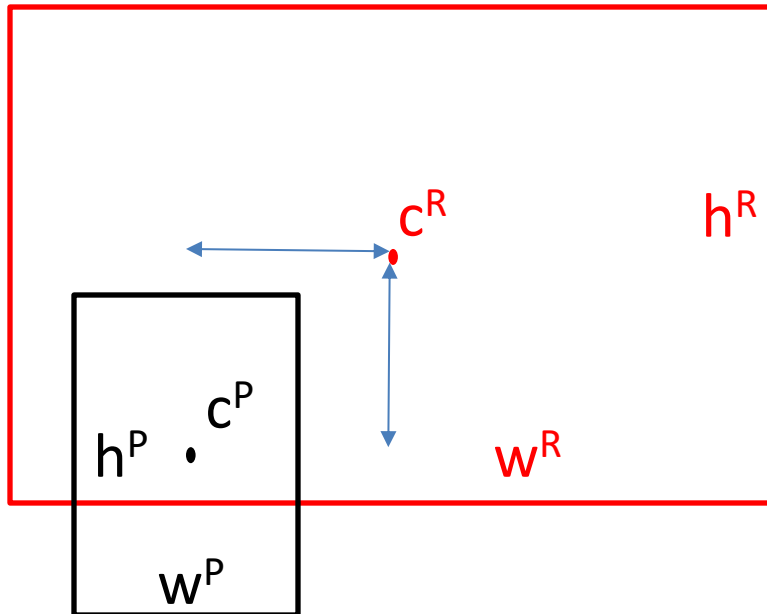
$$area = (x2 - x1)(y2 - y1)$$

$$L = \left[\min \left\{ 5, \max \left\{ 2, 3 + \log_2 \frac{\sqrt{area}}{224} + \epsilon \right\} \right\} \right]$$

224x224 bolo kanonické rozlíšenie obrázkov datasetu pre ImageNet

Box encoding

Reference box (a placed anchor)



Proposal box
to encode

$$\left(\frac{c_x^R - c_x^P}{w^P}, \frac{c_y^R - c_y^P}{h^P}, \ln \frac{w^P}{w^R}, \ln \frac{h^P}{h^R} \right)$$

