



Algoritma Mask R-CNN

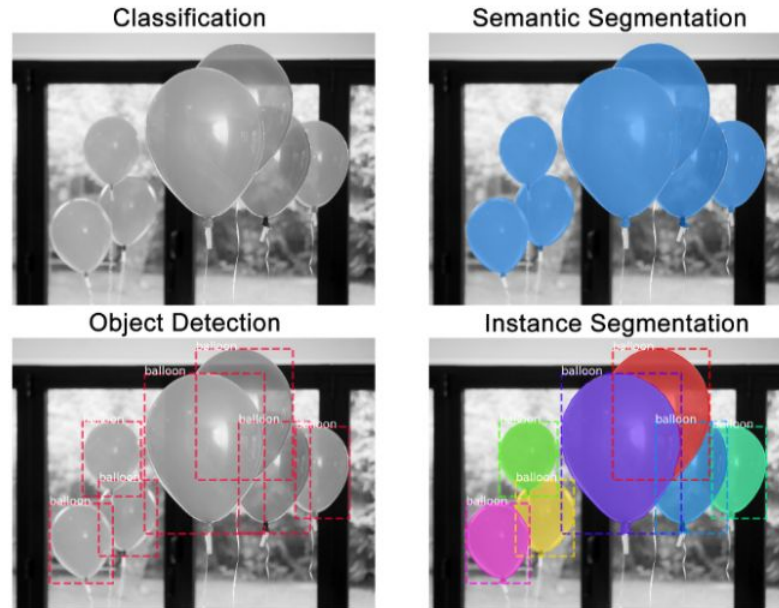
Kaiming He et al.



Pengertian Mask R-CNN

- Mask R-CNN adalah *deep neural network* yang bertujuan untuk memecahkan masalah *instance segmentation* dalam *machine learning* atau *computer vision*. Dengan kata lain, algoritma ini dapat memisahkan objek yang berbeda dalam gambar atau video.
- Apa itu *instance segmentation*? Segmentasi instance adalah tugas mengidentifikasi outlines objek pada tingkat piksel.

Tugas-tugas dalam instance segmentation



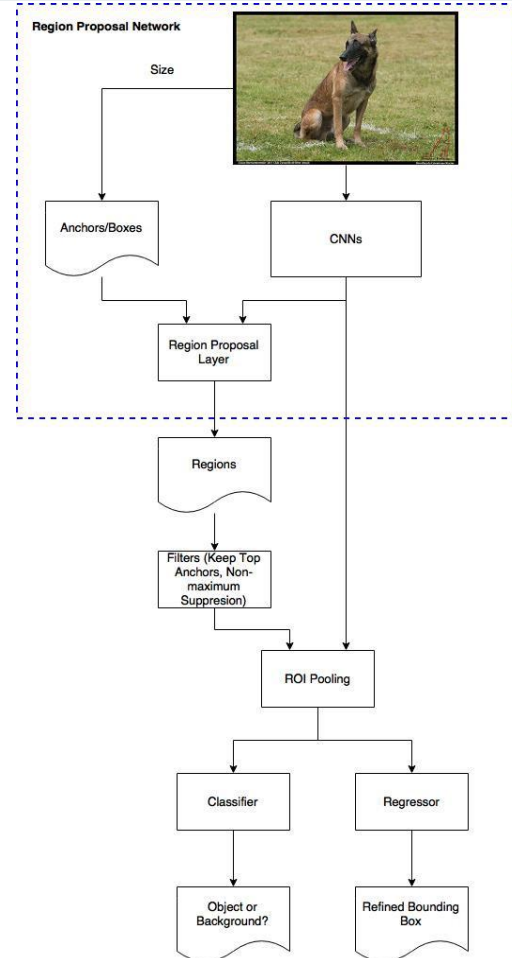


Arsitektur Mask R-CNN

Mask R-CNN (jaringan saraf konvolusional regional) adalah kerangka kerja dua tahap: tahap pertama memindai gambar dan menghasilkan proposal tentang wilayah di mana mungkin ada objek berdasarkan gambar masukan. Kedua, memprediksi kelas objek, menyempurnakan kotak pembatas, dan menghasilkan mask dalam level piksel objek berdasarkan proposal tahap pertama. Kedua tahapan tersebut terhubung ke struktur backbone.

Faster R-CNN

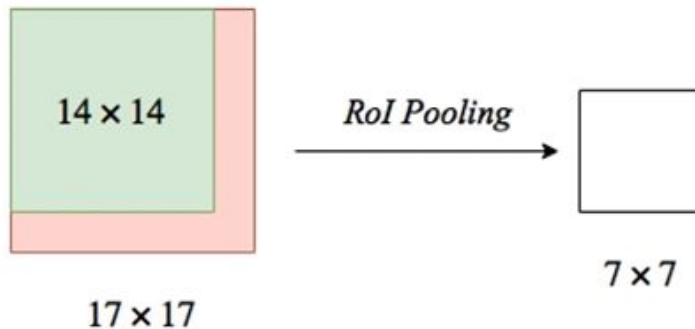
Untuk tahap pertama yaitu deteksi daerah objek menggunakan arsitektur yang mirip dengan faster R-CNN



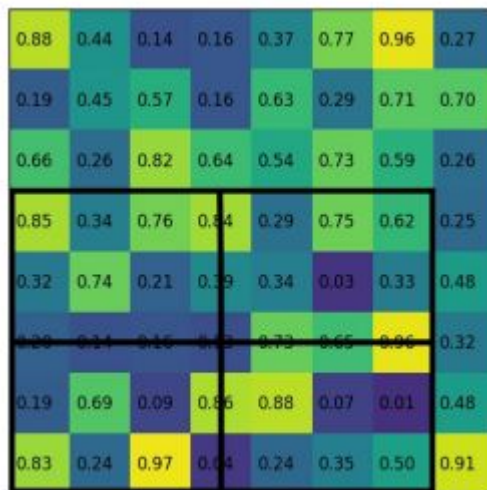
Rol Pooling

Rol Pooling: Stride is quantized.

$$stride = \frac{17}{7} = 2.42 \quad stride_{RoIPool} = [2.42] = 2$$



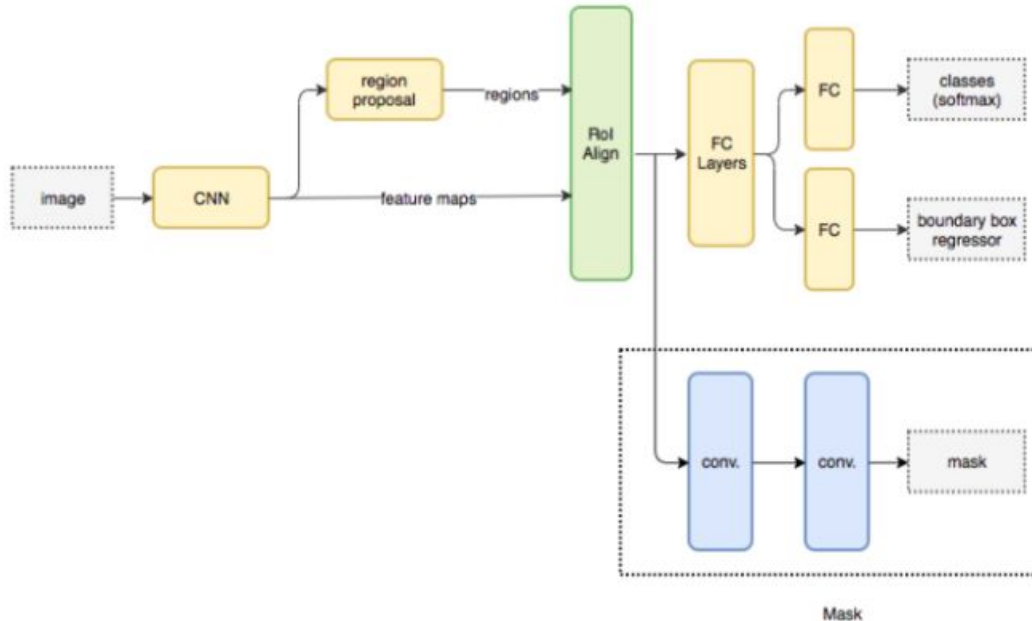
RoI Align



	align?	bilinear?	agg.	AP	AP ₅₀	AP ₇₅
<i>RoIPool</i> [12]			max	26.9	48.8	26.4
<i>RoIWarp</i> [10]		✓	max	27.2	49.2	27.1
			ave	27.1	48.9	27.1
<i>RoIAlign</i>	✓	✓	max	30.2	51.0	31.8
			ave	30.3	51.2	31.5

(c) **RoIAlign** (ResNet-50-C4): Mask results with various RoI layers. Our RoIAlign layer improves AP by ~3 points and AP₇₅ by ~5 points. Using proper alignment is the only factor that contributes to the large gap between RoI layers.

Arsitektur Mask R-CNN





Refrensi

- Kaiming He, Georgia Gkioxari, Piotr Dollar, Ross Girshick. 2017. Masker R-CNN. ` Riset AI Facebook (FAIR).
- <https://medium.com/@tibastar/mask-r-cnn-d69aa596761f>
- <https://medium.com/@smallfishbigsea/faster-r-cnn-explained-864d4fb7e3f8>