

رسالة محمد

مبانی بینایی کامپیوتر

مدرس: محمدرضا محمدی

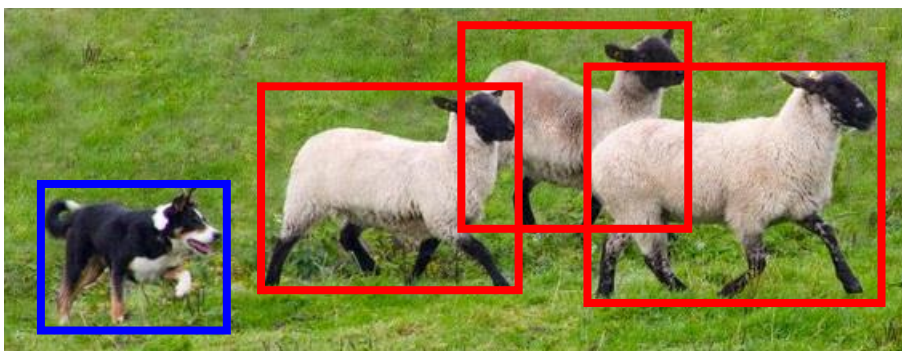
۱۴۰۱

ناحیه بندی معنایی

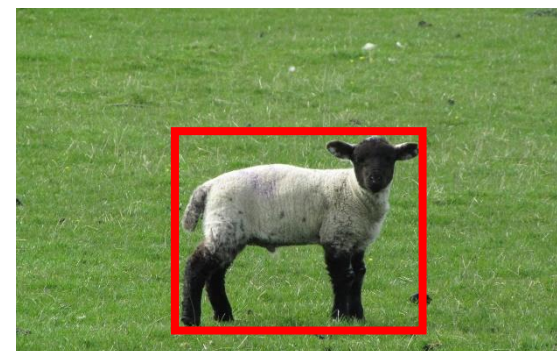
Semantic Segmentation

مسئله‌های بینایی کامپیوتر

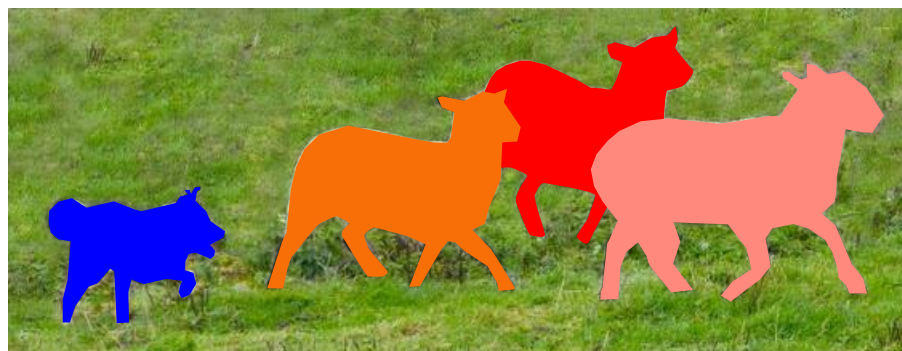
تشخیص اشیاء (Object Detection)



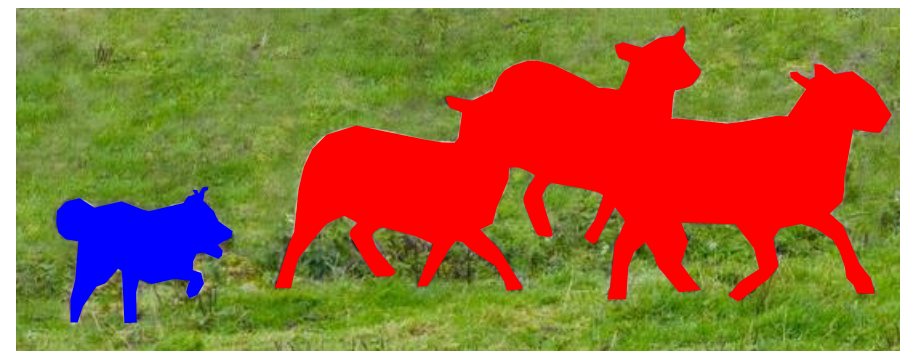
دسته‌بندی + مکان‌یابی



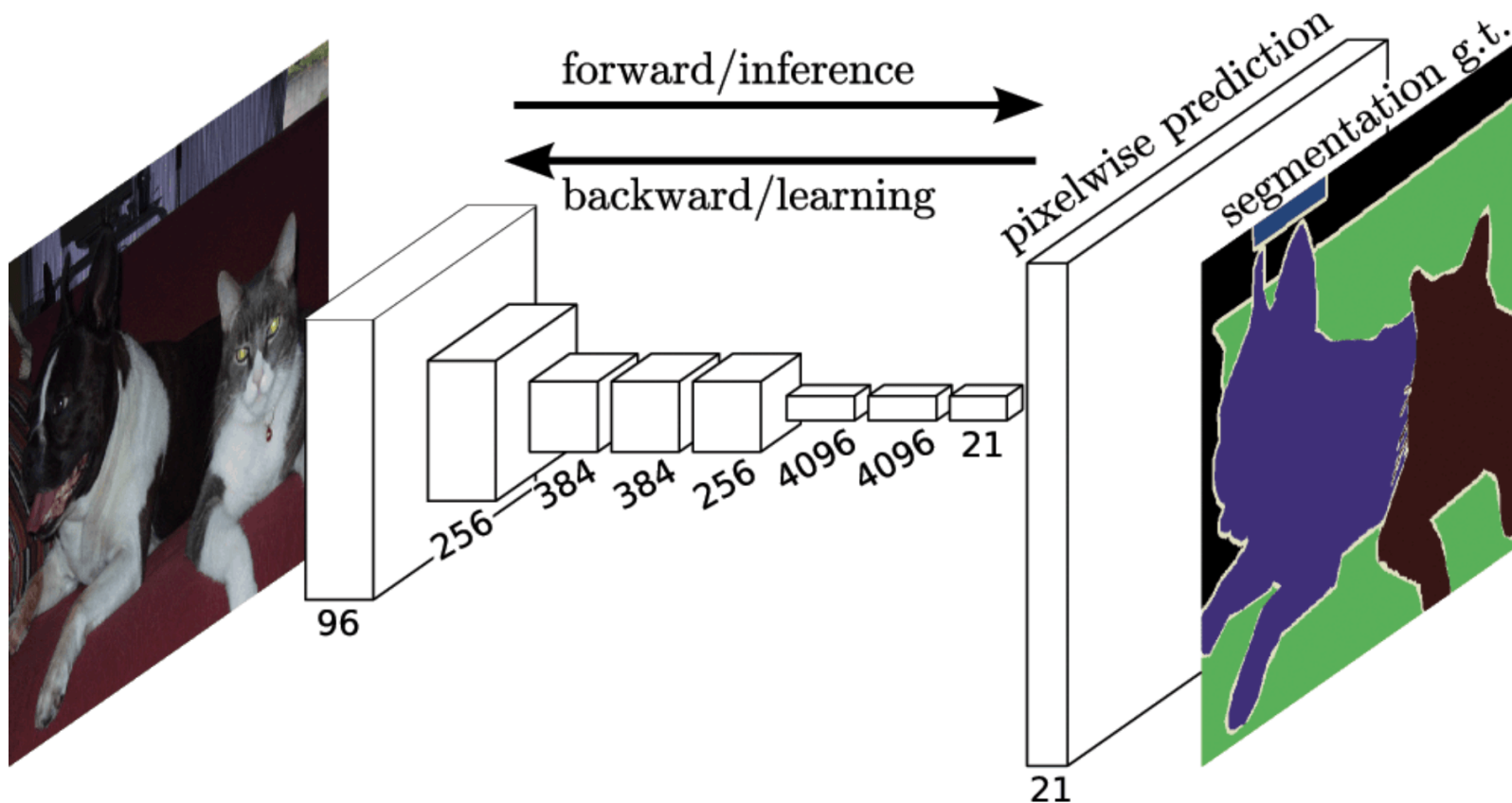
ناحیه‌بندی نمونه‌ها (Instance Segmentation)



ناحیه‌بندی معنایی (Semantic Segmentation)

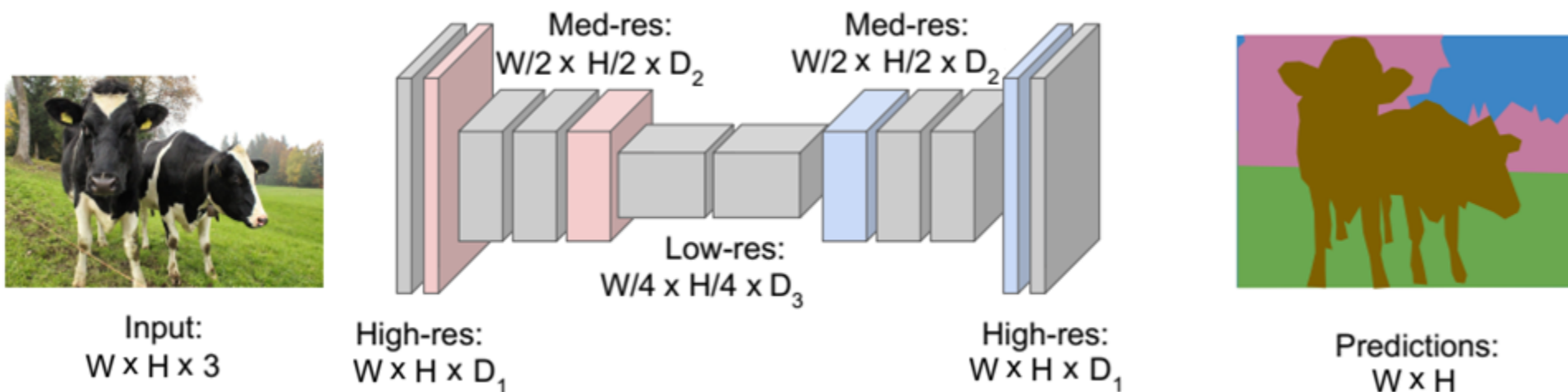


شبکه‌های کاملاً کانولوشنی



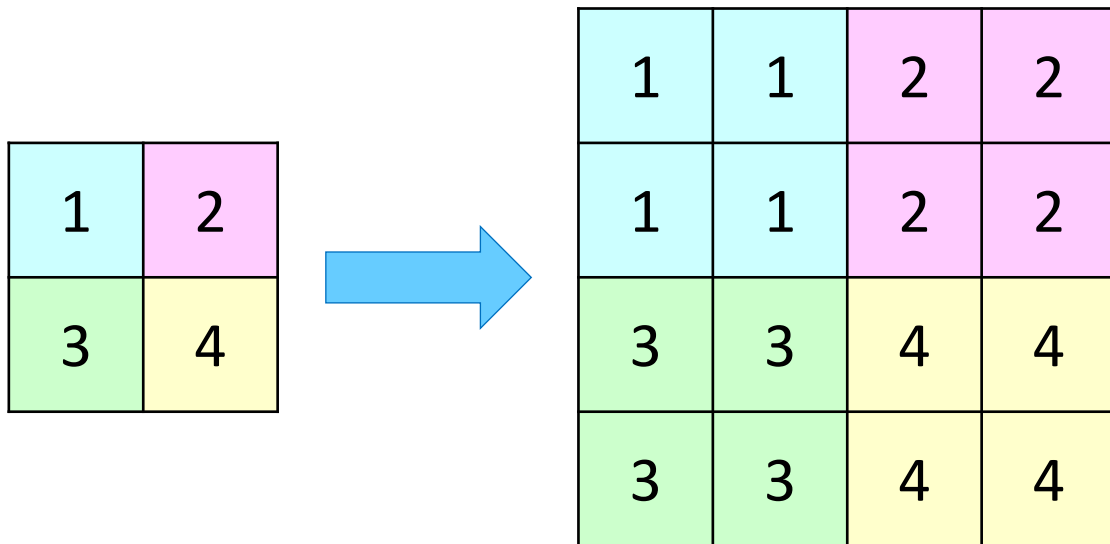
شبکه‌های کاملاً کانولوشنی

- می‌توان مشابه با شبکه‌های متداول، ابعاد مکانی را در طول شبکه کاهش داد و مجدد ابعاد مکانی را به صورت تدریجی افزایش داد
- برای کاهش ابعاد مکانی می‌توان از Pooling استفاده کرد
- چطور می‌توان ابعاد را افزایش داد؟

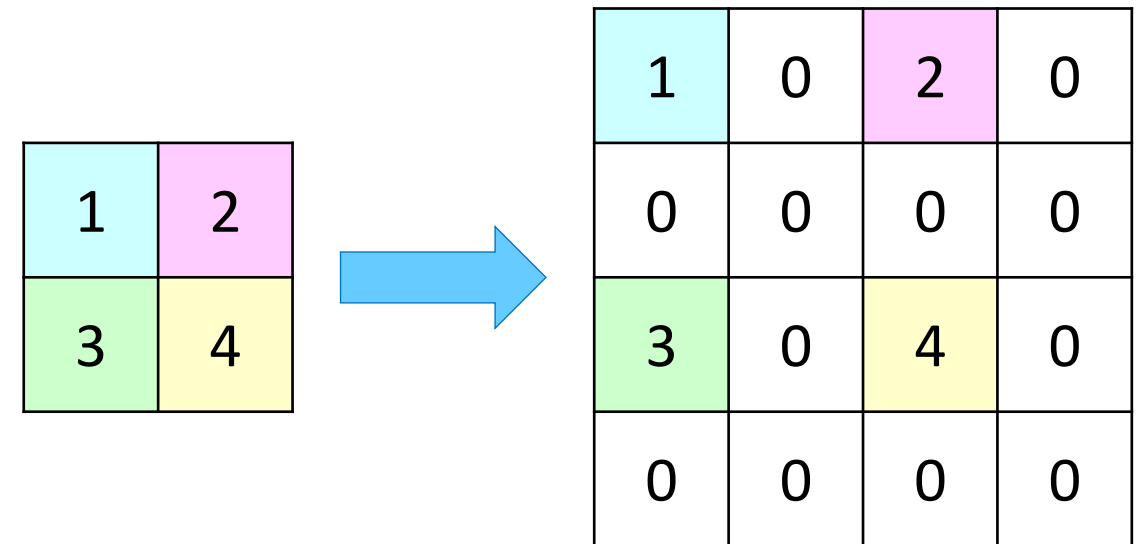


افزایش ابعاد (Unpooling)

نزدیک ترین همسایه



Bed of Nails



افزایش ابعاد (Unpooling)

1	2
3	4



1	1.25	1.75	2
1.5	1.75	2.25	2.5
2.5	2.75	3.25	3.5
3	3.25	3.75	4



Ground Truth



$\frac{1}{4}$ Sized
Input



Bicubic



Super Resolution
Network

افزایش ابعاد (Max Unpooling)

1	2
3	4



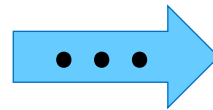
1	0	2	0
0	0	0	0
3	0	4	0
0	0	0	0

افزایش ابعاد (Max Unpooling)

0	5	2	1
1	0	6	2
7	1	2	0
1	0	0	8



5	6
7	8

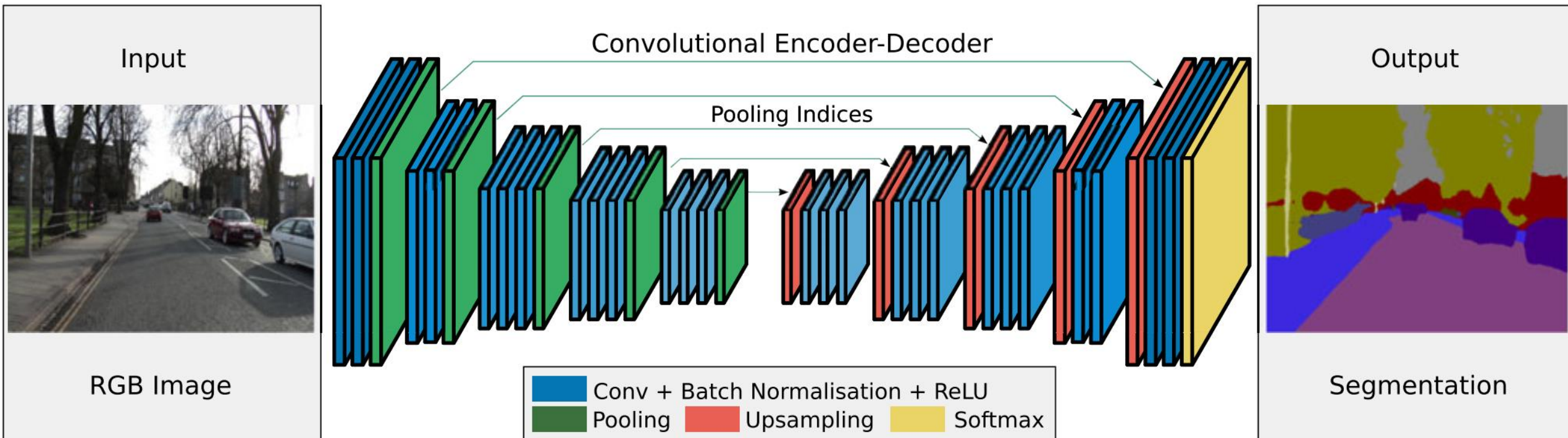


1	2
3	4

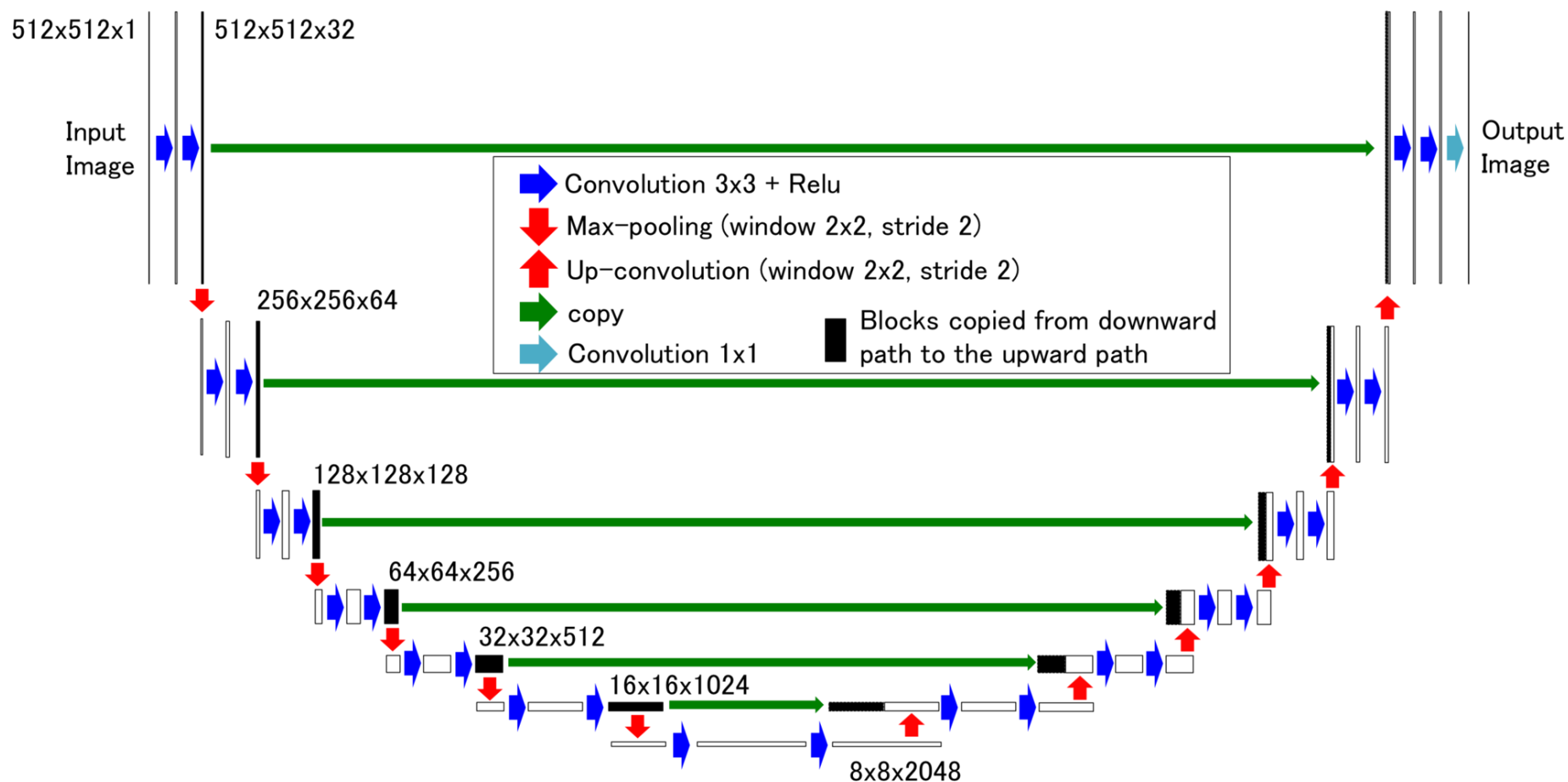


0	1	0	0
0	0	2	0
3	0	0	0
0	0	0	4

SegNet

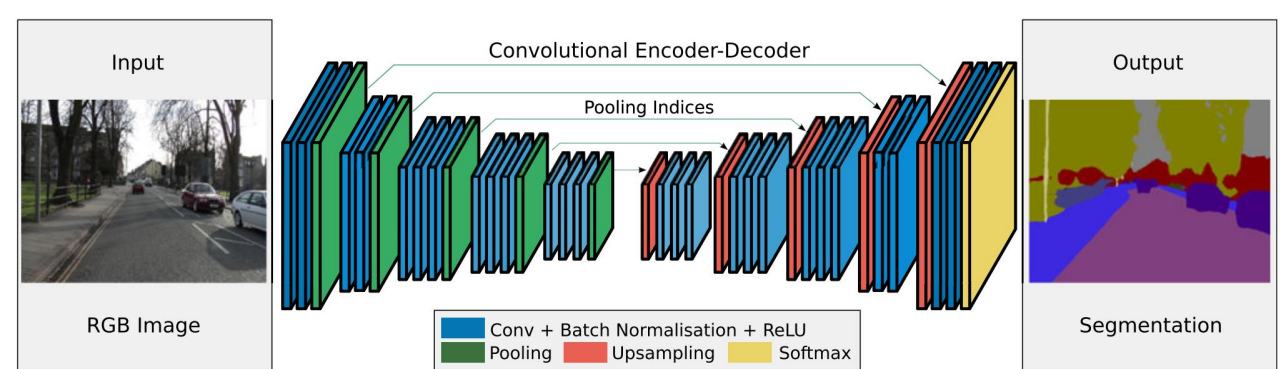
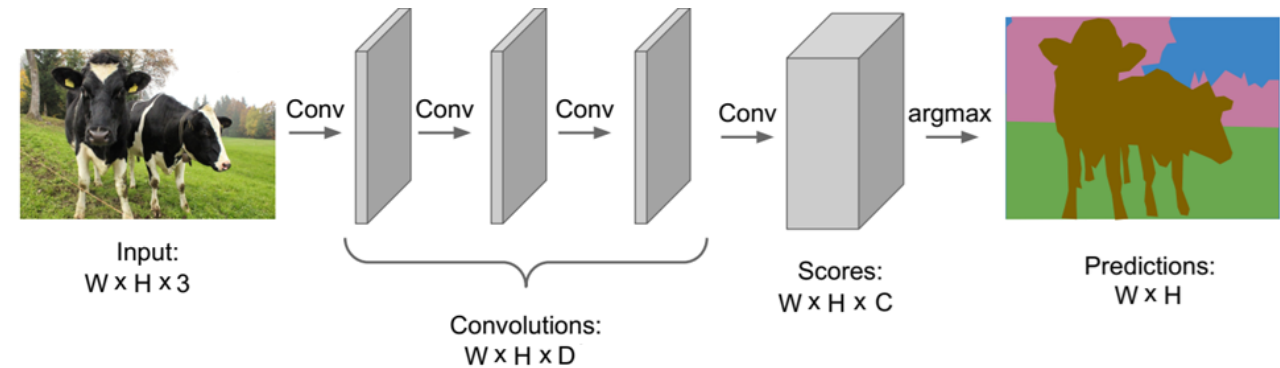


U-Net

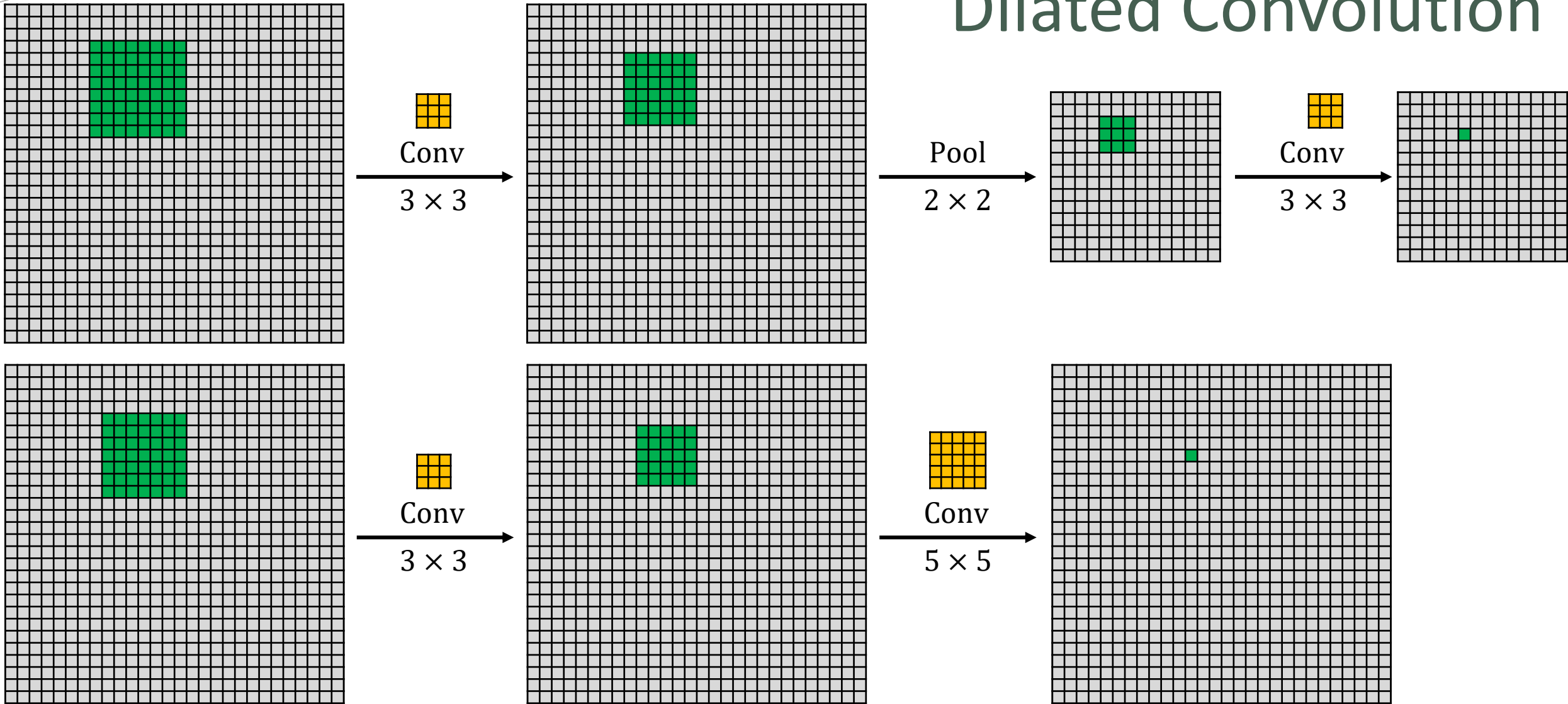


Dilated Convolution

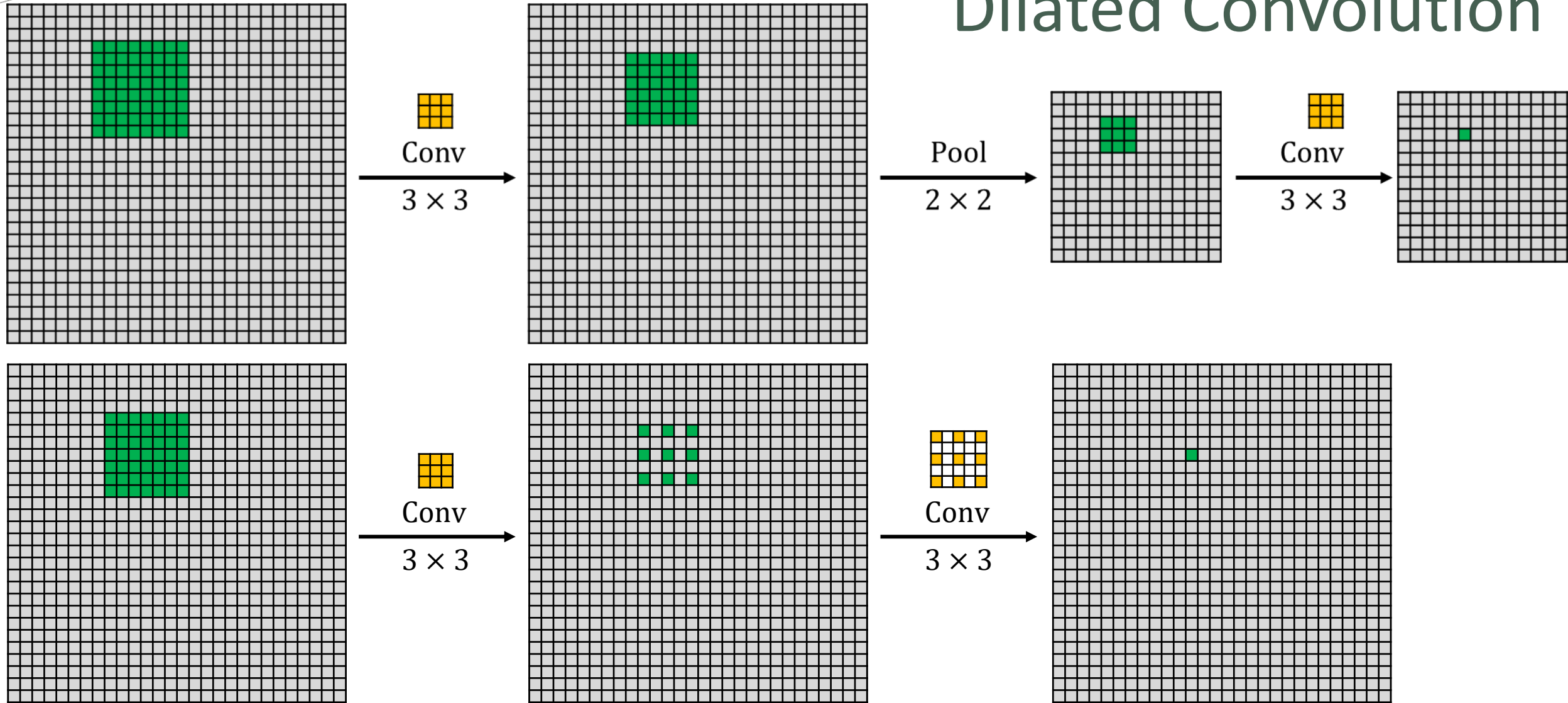
- یکی از دلایل کوچک کردن ابعاد مکانی ویژگی‌ها این است که برای داشتن میدان تاثیر بزرگ، نیازی با استفاده از فیلترهای بزرگ نداشته باشیم



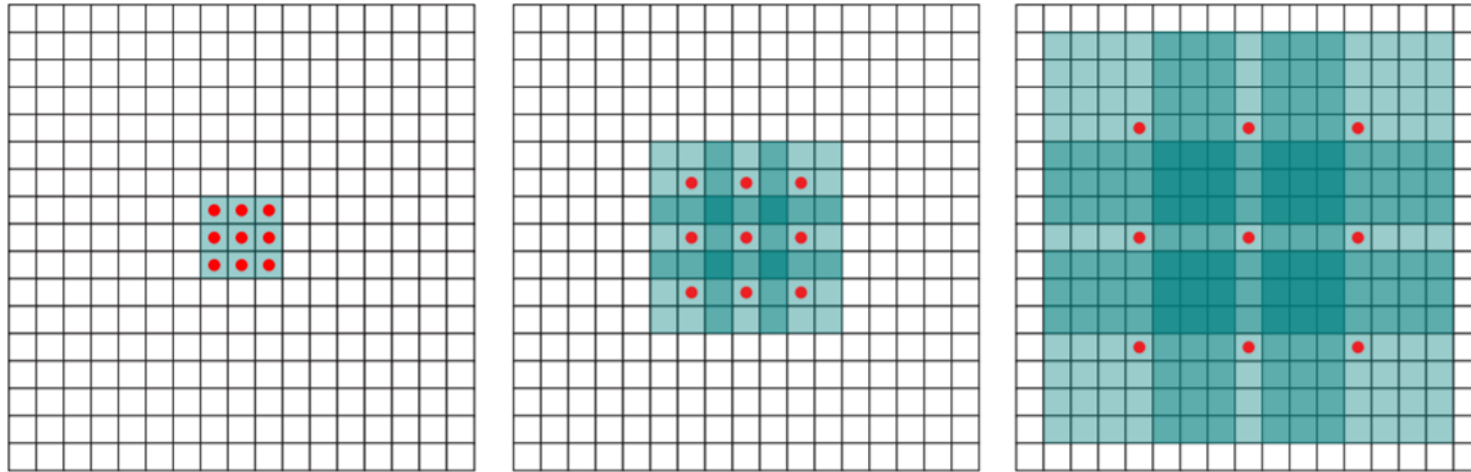
Dilated Convolution



Dilated Convolution



Dilated Convolution



Layer	1	2	3	4	5	6	7	8
Convolution	3×3	3×3	3×3	3×3	3×3	3×3	3×3	1×1
Dilation	1	1	2	4	8	16	1	1
Truncation	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Receptive field	3×3	5×5	9×9	17×17	33×33	65×65	67×67	67×67
Output channels								
Basic	C	C	C	C	C	C	C	C
Large	$2C$	$2C$	$4C$	$8C$	$16C$	$32C$	$32C$	C

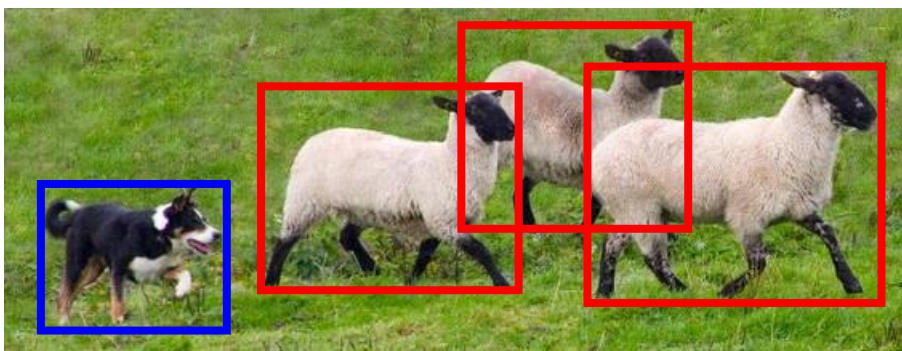


تشخيص اشیاء

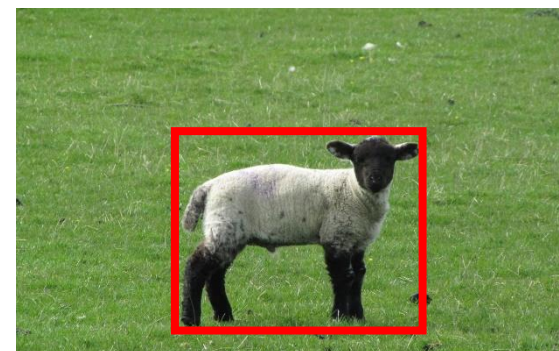
Object Detection

مسئله‌های بینایی کامپیوتر

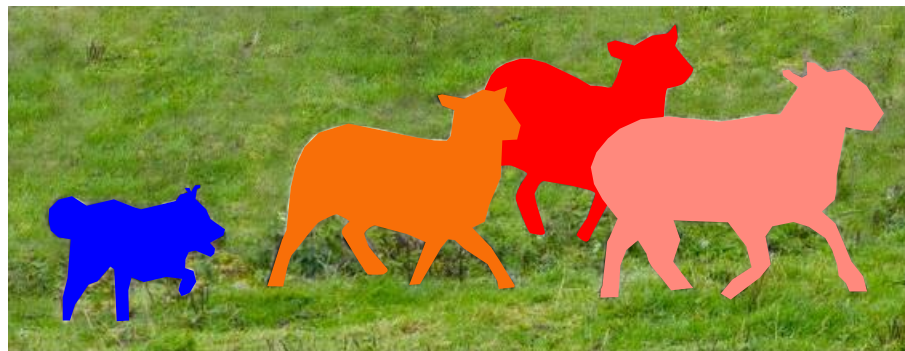
تشخیص اشیاء (Object Detection)



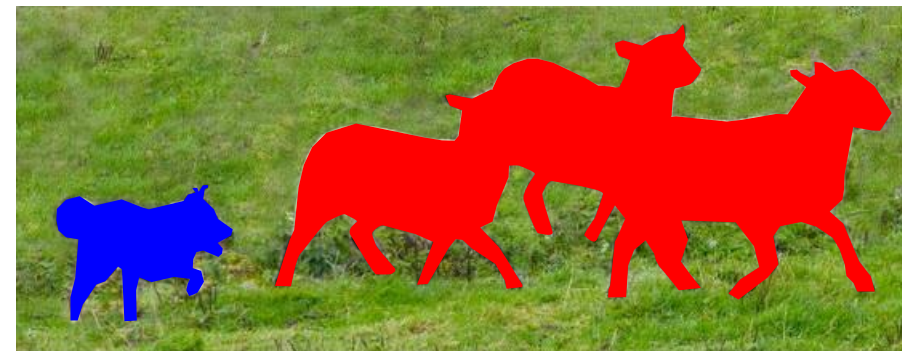
دسته‌بندی + مکان‌یابی



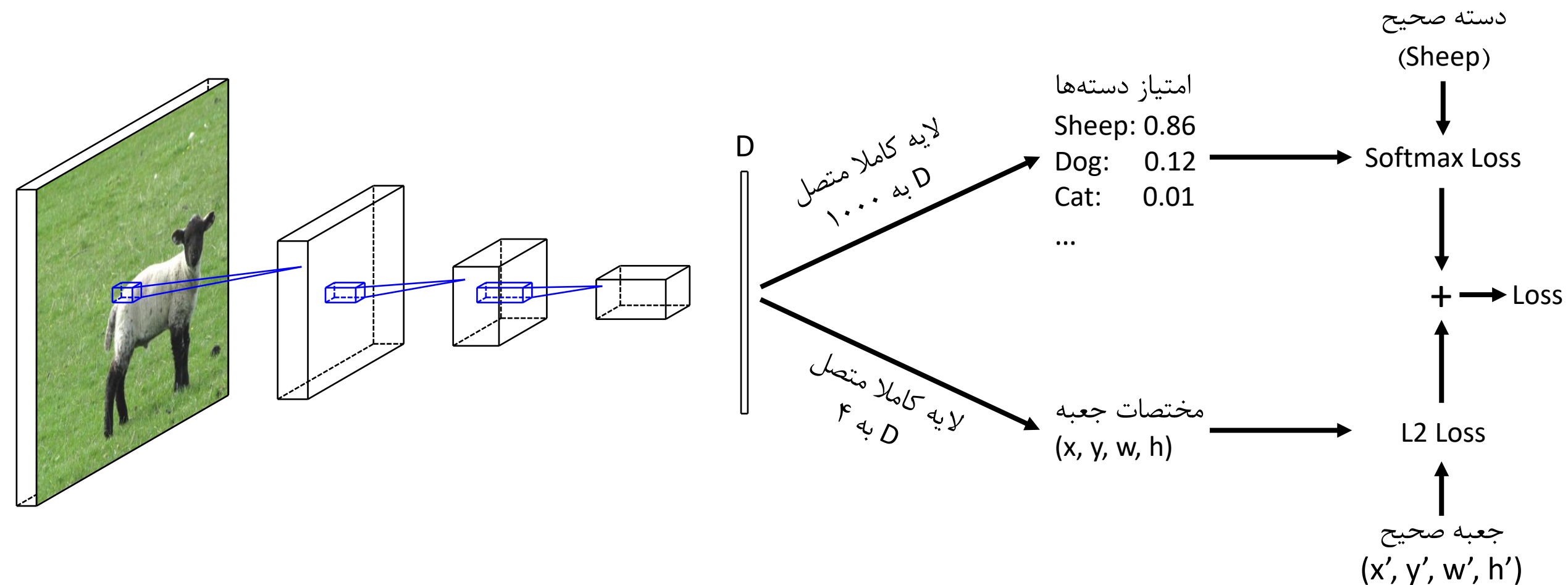
ناحیه‌بندی نمونه‌ها (Instance Segmentation)



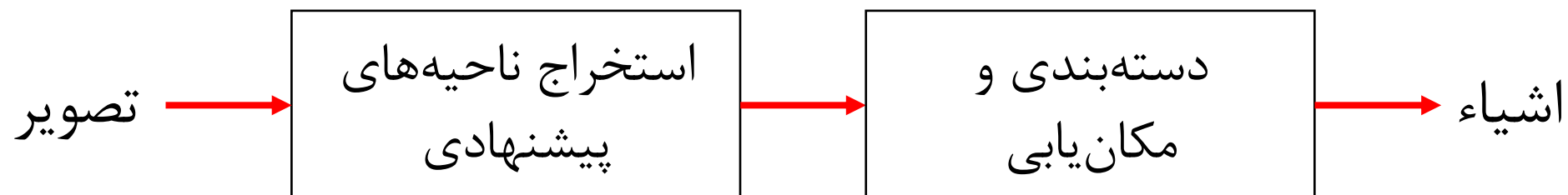
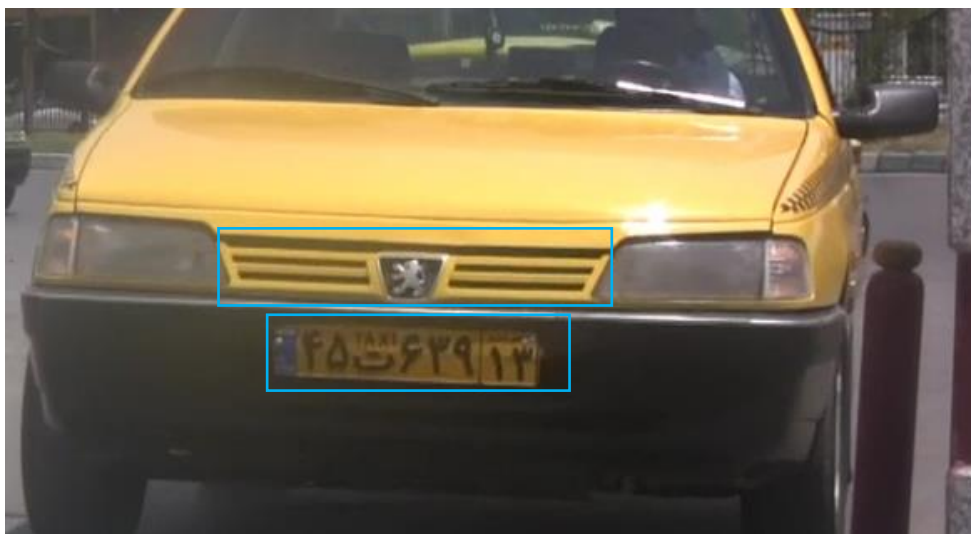
ناحیه‌بندی معنایی (Semantic Segmentation)



دسته‌بندی و مکان‌یابی



تشخیص اشیاء



پنجره لغزان

- ساده‌ترین روش برای استخراج ناحیه‌های پیشنهادی استفاده از پنجره لغزان است
- در این روش، بخش‌های مختلف تصویر توسط یک پنجره جستجو می‌شوند
- علاوه بر این، نیاز است تا از پنجره‌های دارای ابعاد و نسبت‌های مختلف استفاده شود



تولید ناحیه‌های پیشنهادی

- تعداد ناحیه‌هایی که در روش‌های مبتنی بر پنجره لغزان بررسی می‌شوند بسیار زیاد است و نمی‌توان از بسیاری از روش‌های استخراج ویژگی و دسته‌بندی متداول استفاده کرد
- روش‌های مختلفی توسعه یافته‌اند تا تعداد محدودی ناحیه پیشنهادی تولید کنند
- مانند ناحیه‌بندی تصویر
- بسیاری از روش‌های تولید ناحیه‌های پیشنهادی خاص منظوره هستند

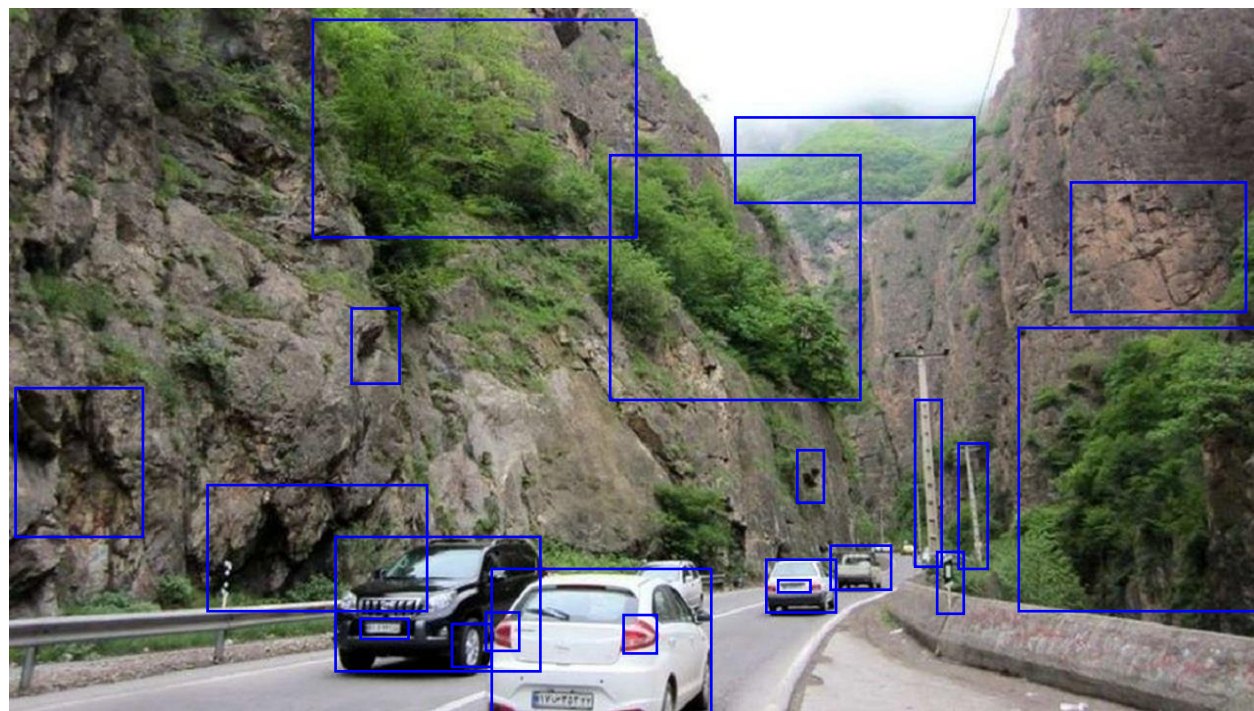


تولید ناحیه‌های پیشنهادی

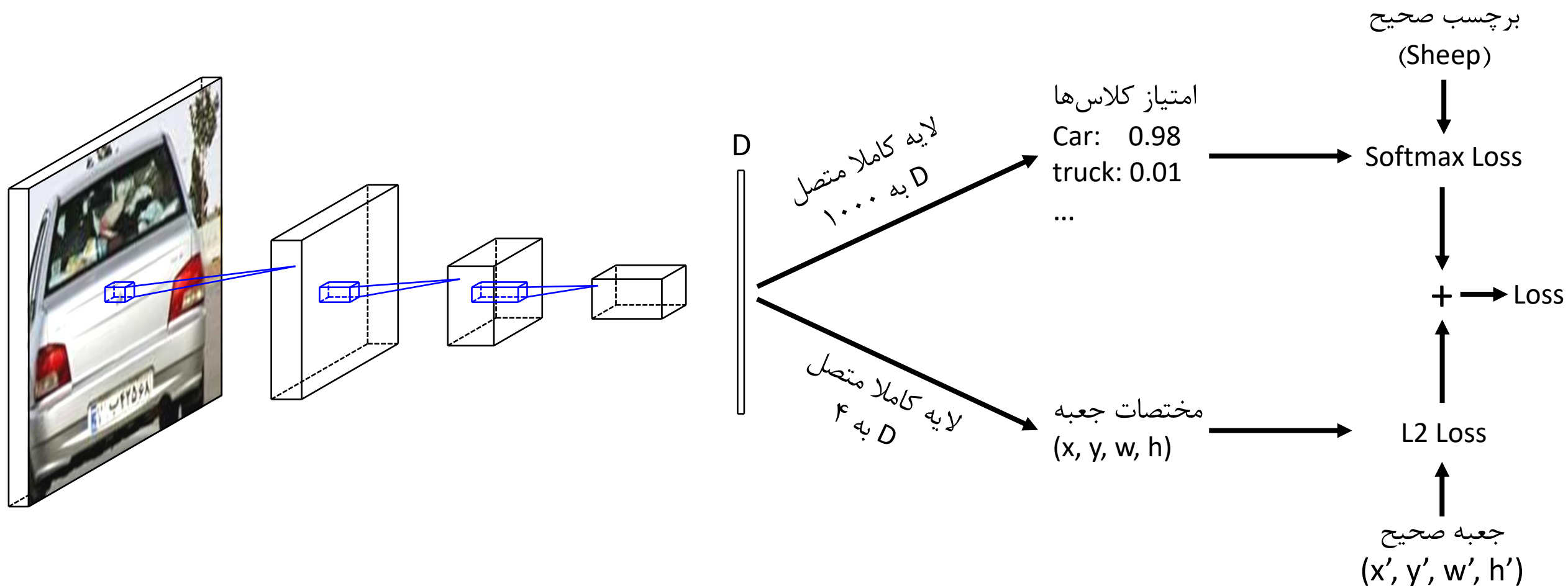
- با وجود آنکه عمده روش‌های تولید ناحیه‌های پیشنهادی خاص منظوره هستند، در سال‌های اخیر الگوریتم‌هایی توسعه یافته‌اند که بتوانند ناحیه‌های عمومی مناسبی از تصویر استخراج کنند

- تعداد ناحیه‌های پیشنهادی باید تا حد امکان کم باشد و در عین حال اشیاء مورد نظر باید تا حد امکان انتخاب شوند و البته سریع باشد

- از جمله این الگوریتم‌ها می‌توان به Edge Boxes، BING و Selective Search اشاره کرد



دسته‌بندی و مکان‌یابی ناحیه‌های پیشنهادی

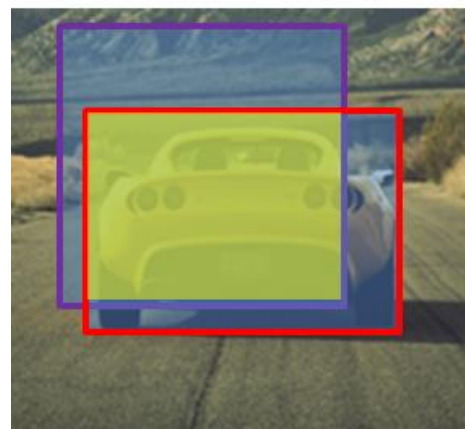


ارزیابی ناحیه‌ها

- متداول‌ترین معیار برای مقایسه دو ناحیه، اشتراک به اجتماع است (IoU)

Intersection over union (IoU)

$$= \frac{\text{size of } \text{yellow box}}{\text{size of } \text{blue box}}$$

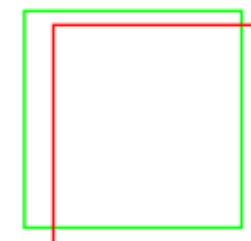


IoU=0.92



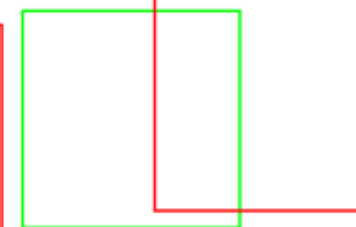
Excellent

IoU=0.71

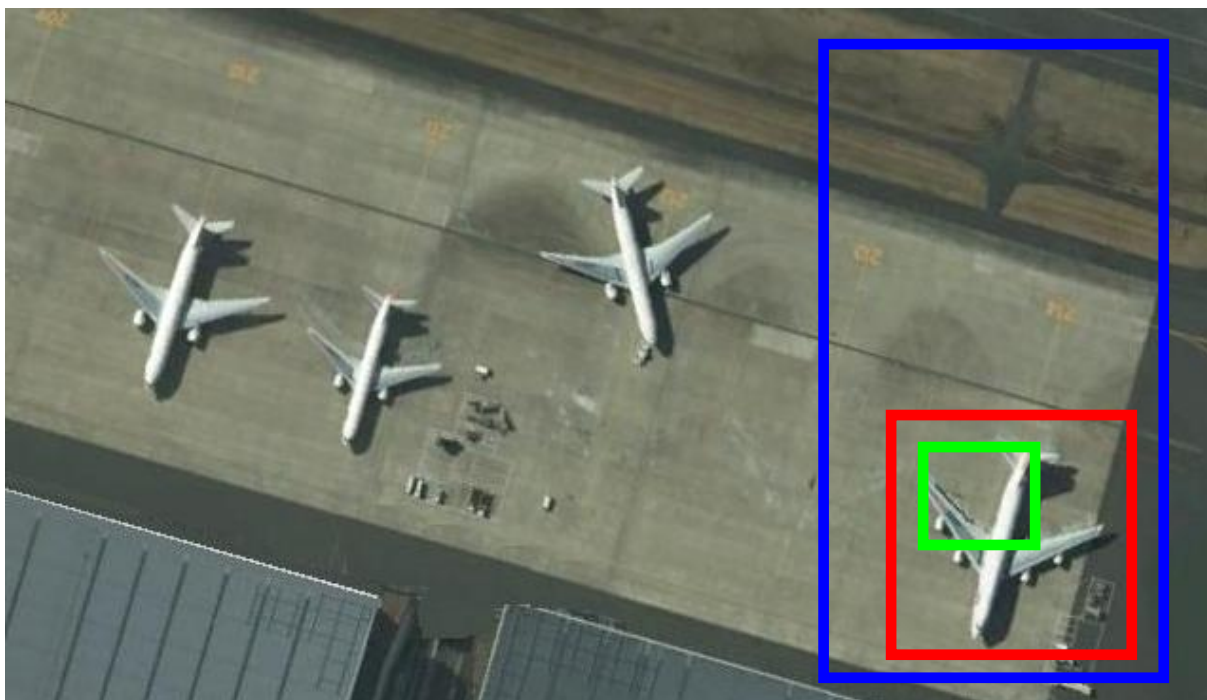


Good

IoU=0.39



Poor



آموزش مدل

- ابتدا ناحیه‌های پیشنهادی از تصویر استخراج می‌شوند
- ناحیه‌هایی که دارای IOU بزرگی با یکی از ناحیه‌های برچسب خورده هستند به عنوان اشیاء مثبت به دسته‌بند وارد می‌شوند (و جعبه آنها در آموزش رگرسیون استفاده می‌شود)
- ناحیه‌هایی که با هیچ ناحیه برچسب‌خورده‌ای دارای اشتراک نبوده یا دارای IOU کوچکی باشند به عنوان اشیاء منفی به دسته‌بند وارد می‌شوند
- باقی ناحیه‌ها وارد فاز آموزش نمی‌شوند