

رسالة محمد

مبانی بینایی کامپیوتر

مدرس: محمدرضا محمدی

۱۴۰۱

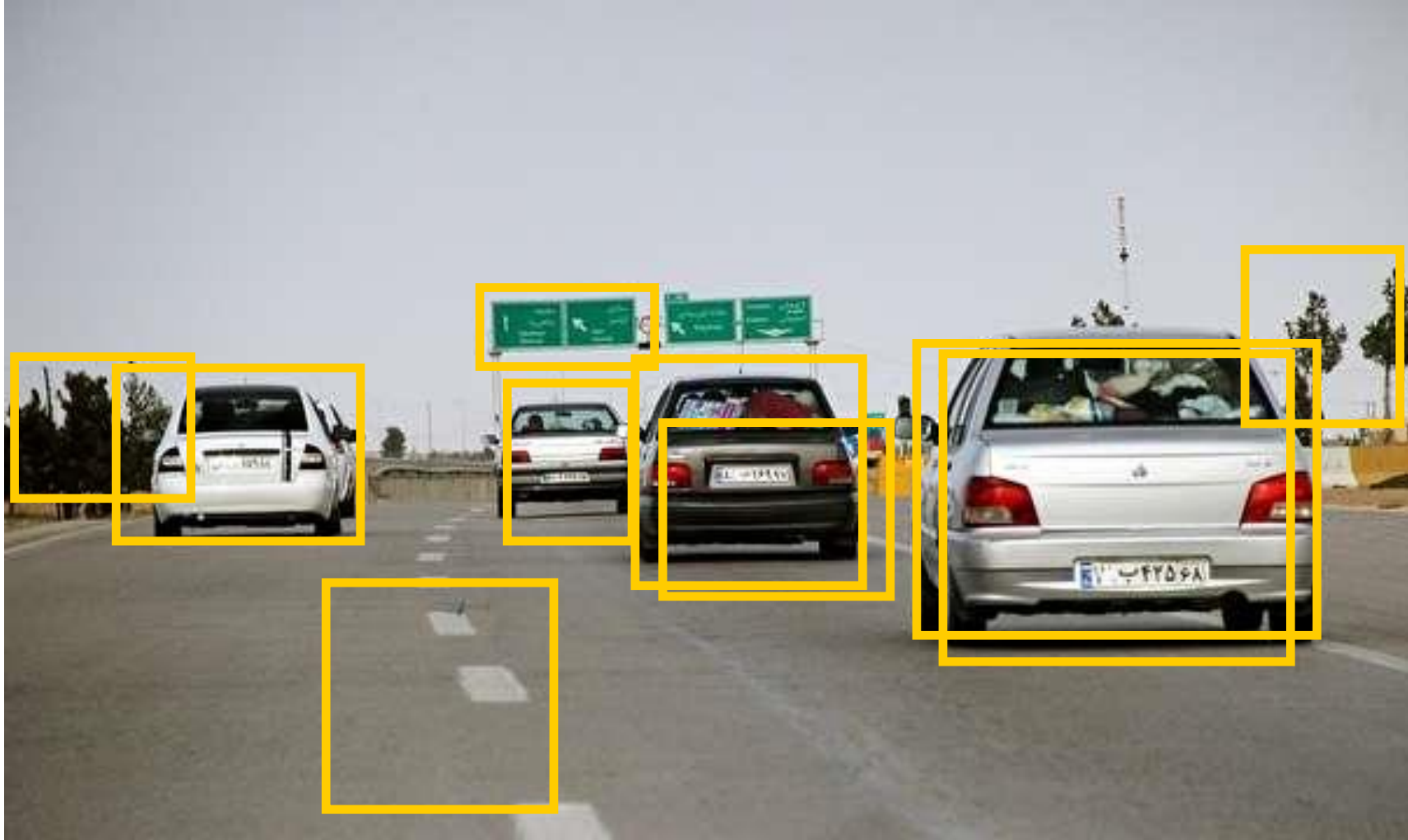
تشخيص اشیاء

Object Detection

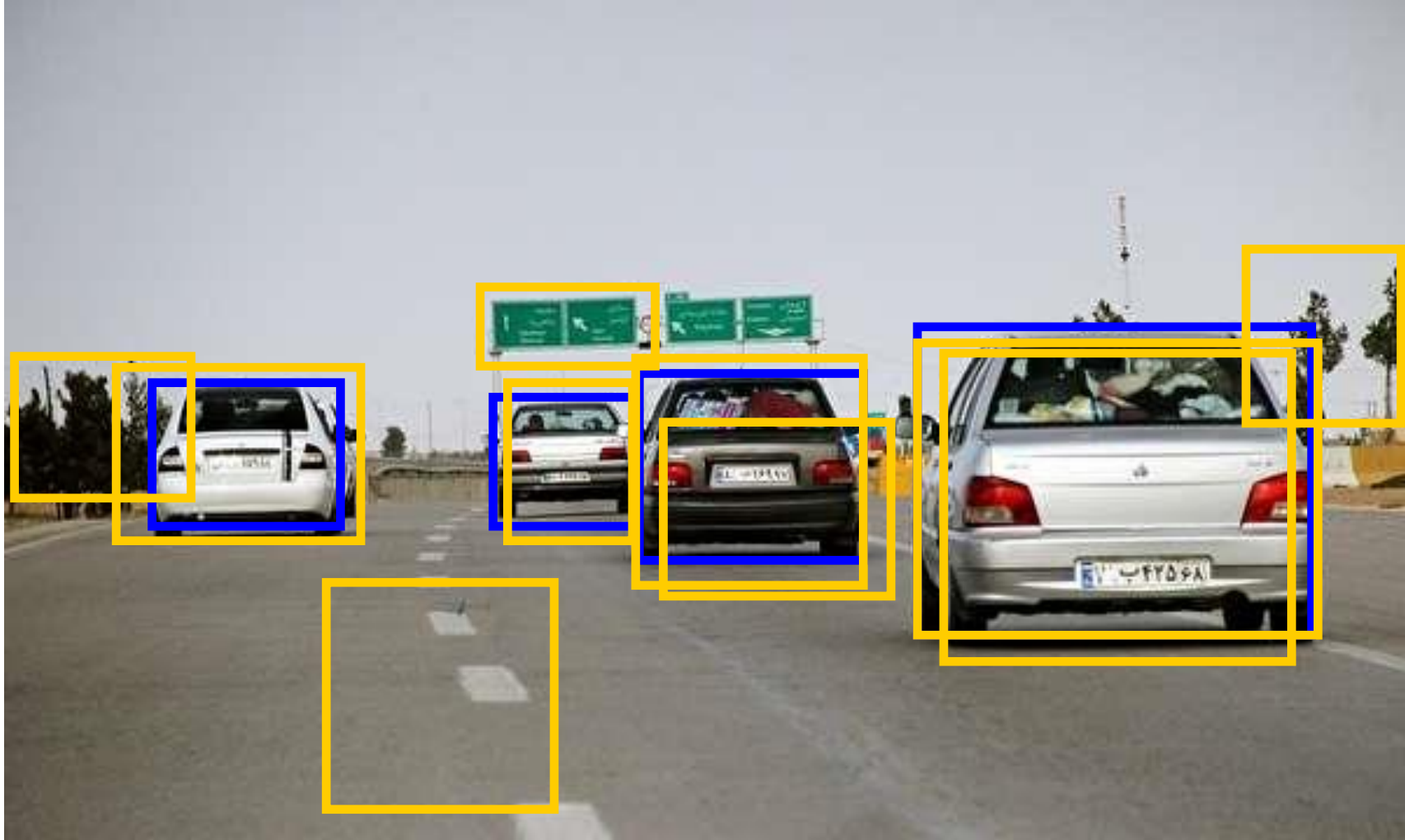
آموزش مدل

- ابتدا ناحیه‌های پیشنهادی از تصاویر 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

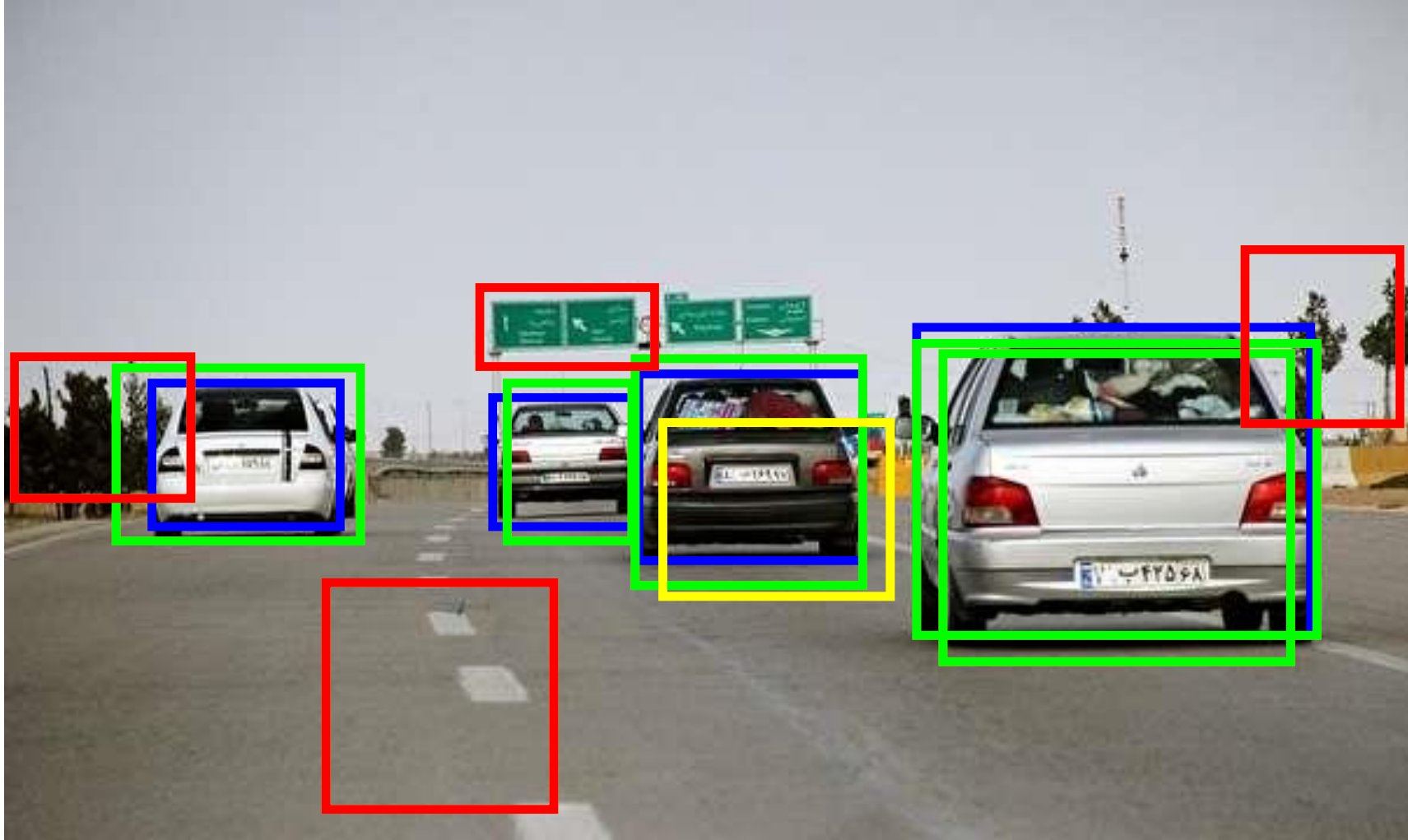
آموزش مدل



آموزش مدل

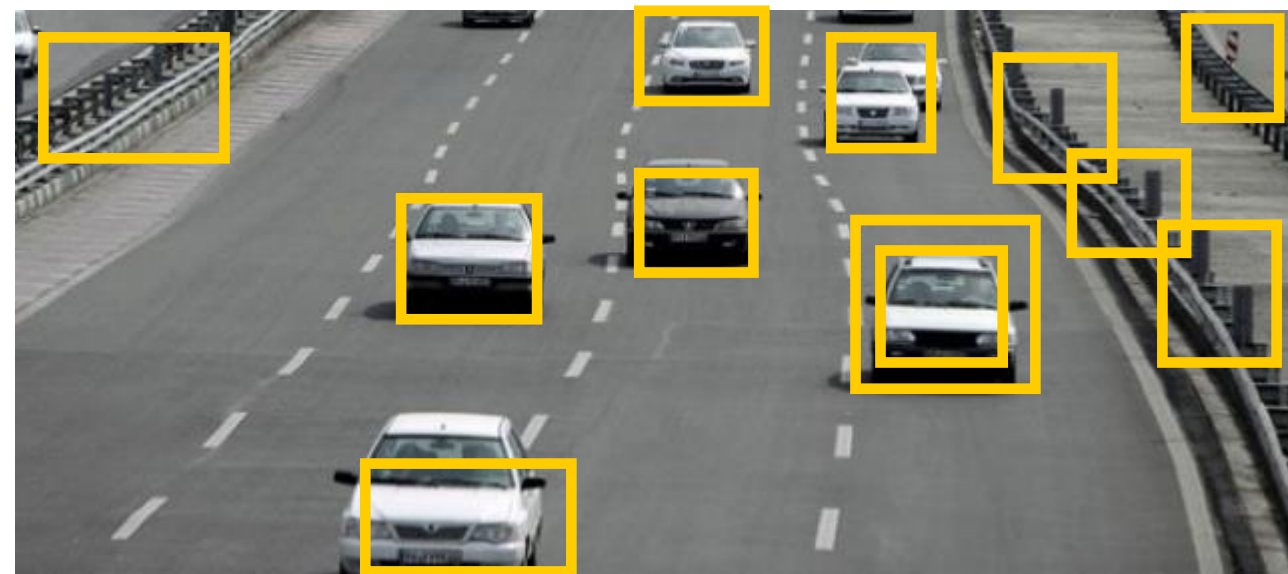


آموزش مدل



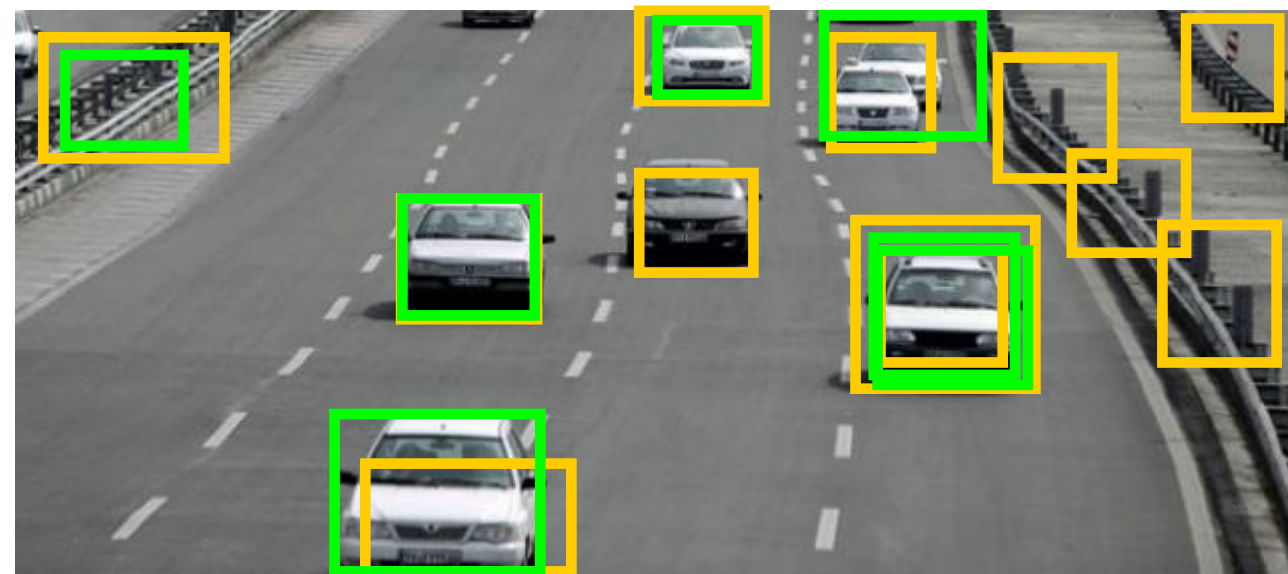
ارزیابی مدل

- ابتدا ناحیه‌های پیشنهادی از تصویر استخراج می‌شوند
- احتمالاً برخی از اشیاء مورد نظر در این مرحله حذف می‌شوند



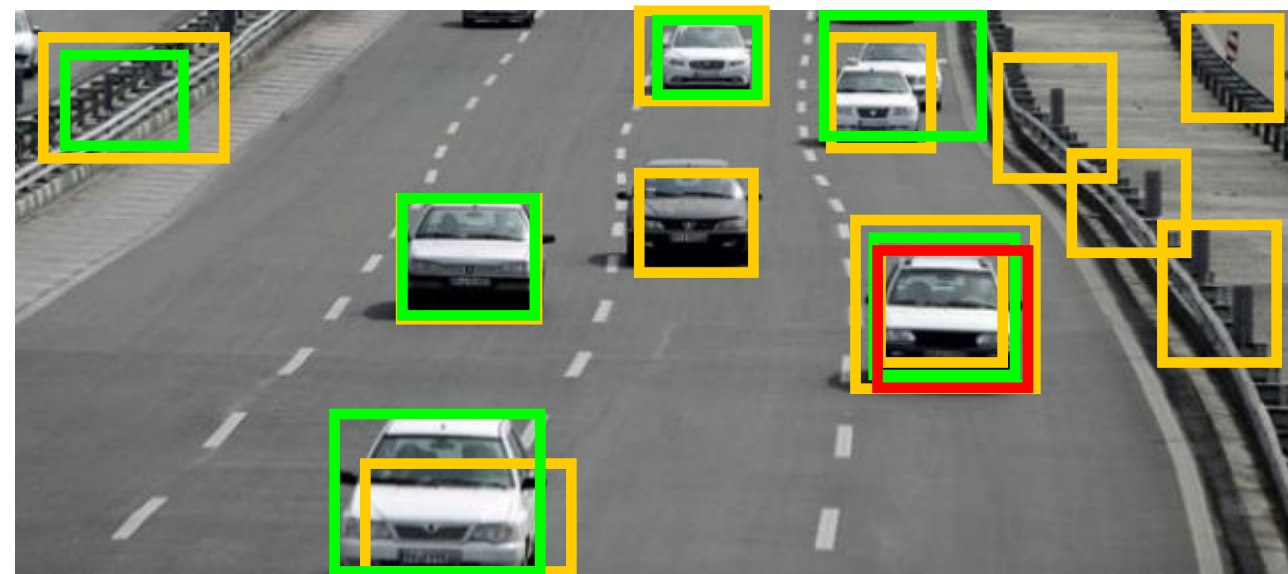
ارزیابی مدل

- ابتدا ناحیه‌های پیشنهادی از تصویر استخراج می‌شوند
 - احتمالا برخی از اشیاء مورد نظر در این مرحله حذف می‌شوند
- با استفاده از مدل آموزش دیده، برچسب هر ناحیه مشخص و مرز آن اصلاح می‌شود



ارزیابی مدل

- ابتدا ناحیه‌های پیشنهادی از تصویر استخراج می‌شوند
 - احتمالاً برخی از اشیاء مورد نظر در این مرحله حذف می‌شوند
- با استفاده از مدل آموزش دیده، برچسب هر ناحیه مشخص و مرز آن اصلاح می‌شود
- با استفاده از روش NMS ناحیه‌های دارای IOU بزرگ ترکیب می‌شوند



ارزیابی مدل

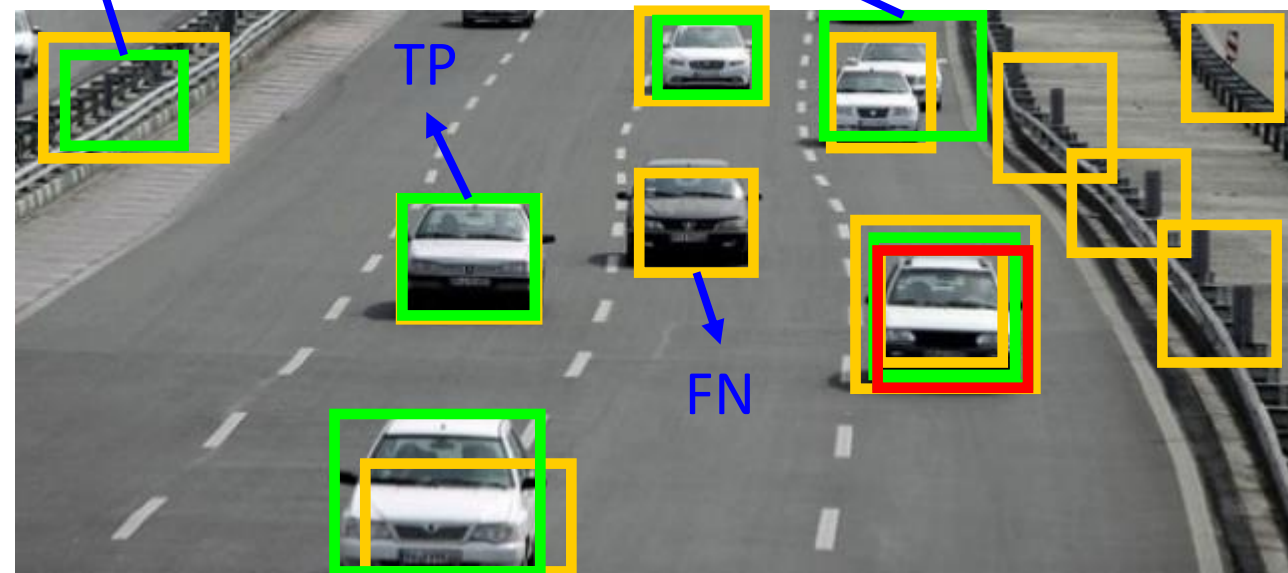
- ابتدا ناحیه‌های پیشنهادی از تصویر استخراج می‌شوند
 - احتمالاً برخی از اشیاء مورد نظر در این مرحله حذف می‌شوند
- با استفاده از مدل آموزش دیده، برچسب هر ناحیه مشخص و مرز آن اصلاح می‌شود
- با استفاده از روش NMS ناحیه‌های دارای IOU بزرگ ترکیب می‌شوند
- با مقایسه IOU ناحیه‌های بدست آمده با ناحیه‌های برچسب خورده، دقت الگوریتم محاسبه می‌شود

FP

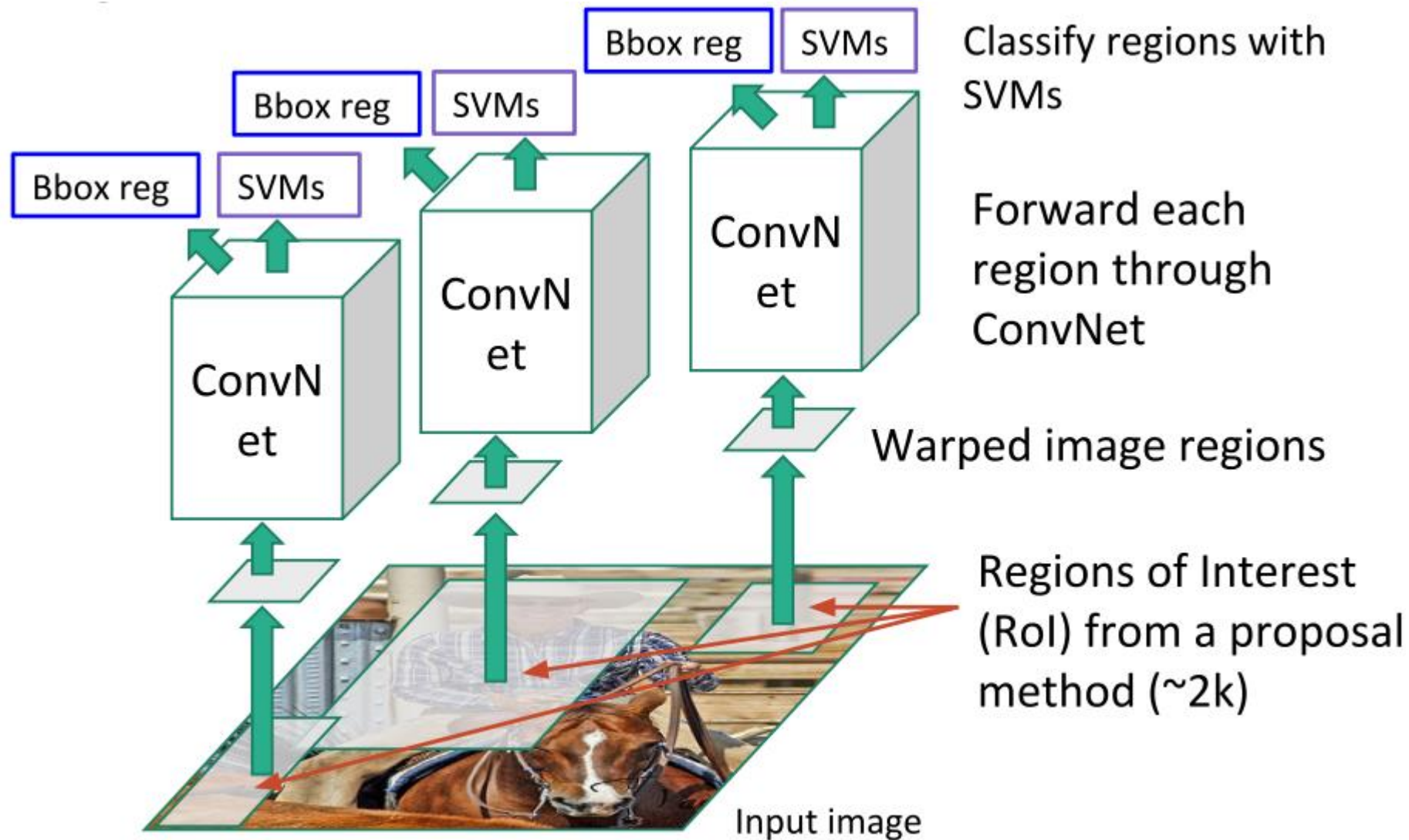
FP & FN

TP

FN



R-CNN



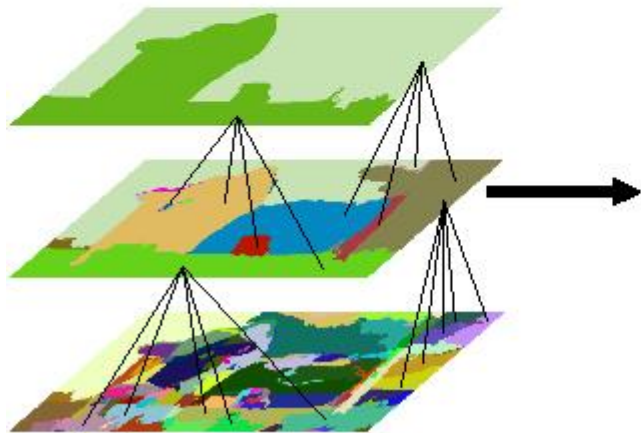
R-CNN



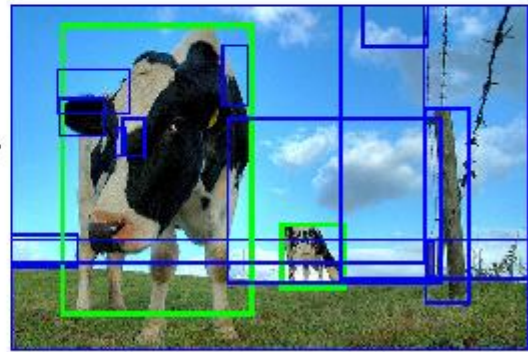
(a)



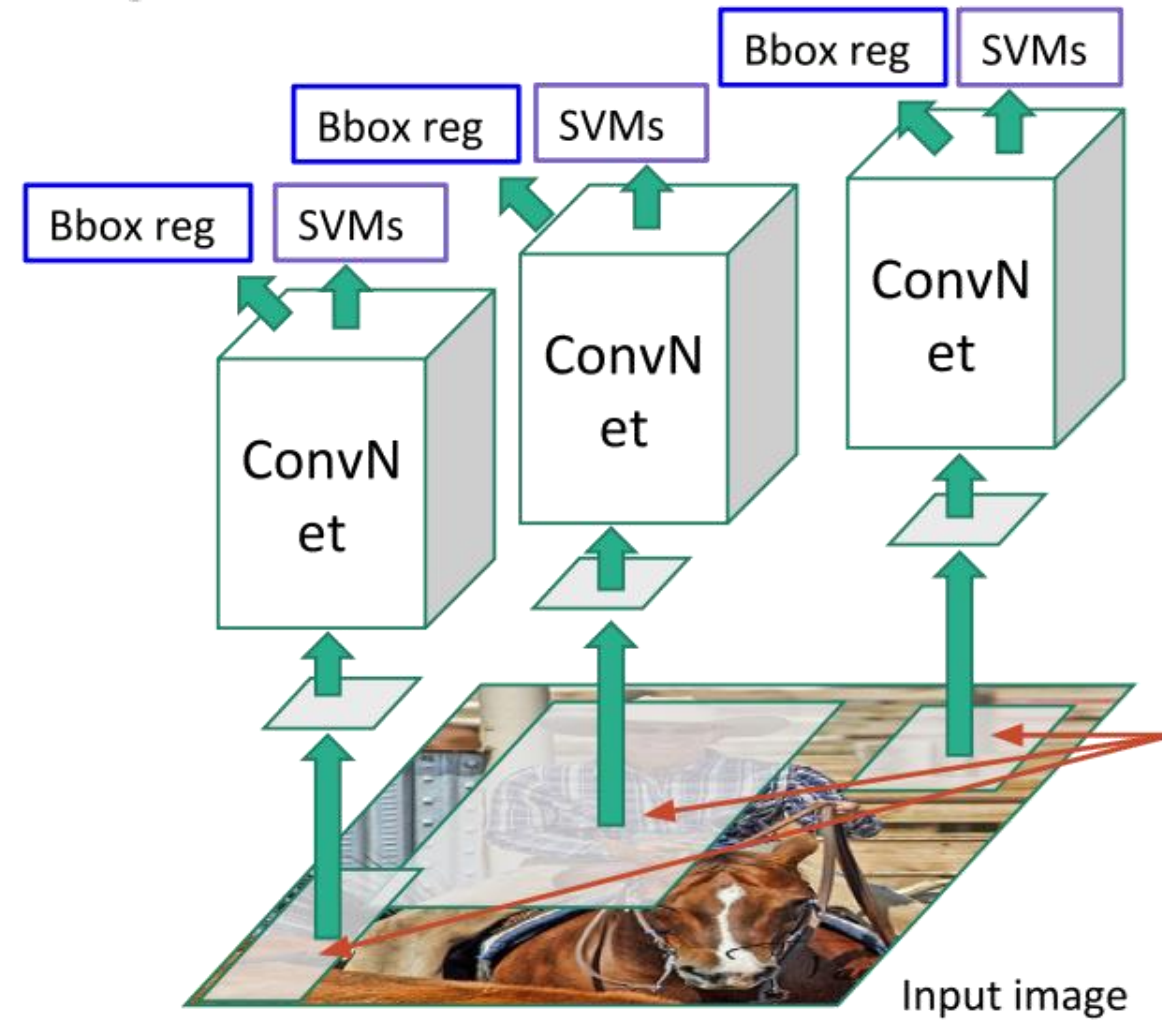
(b)



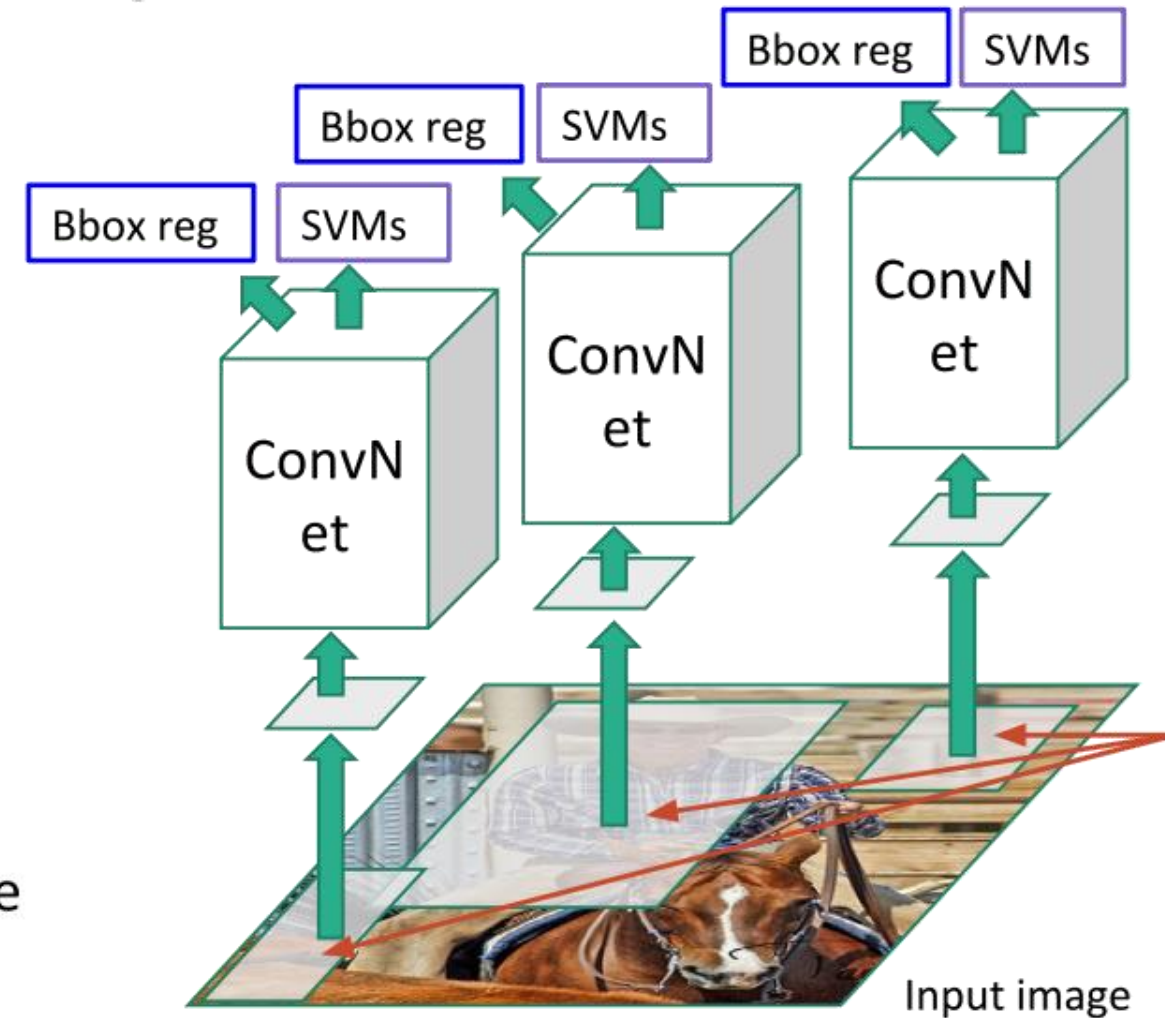
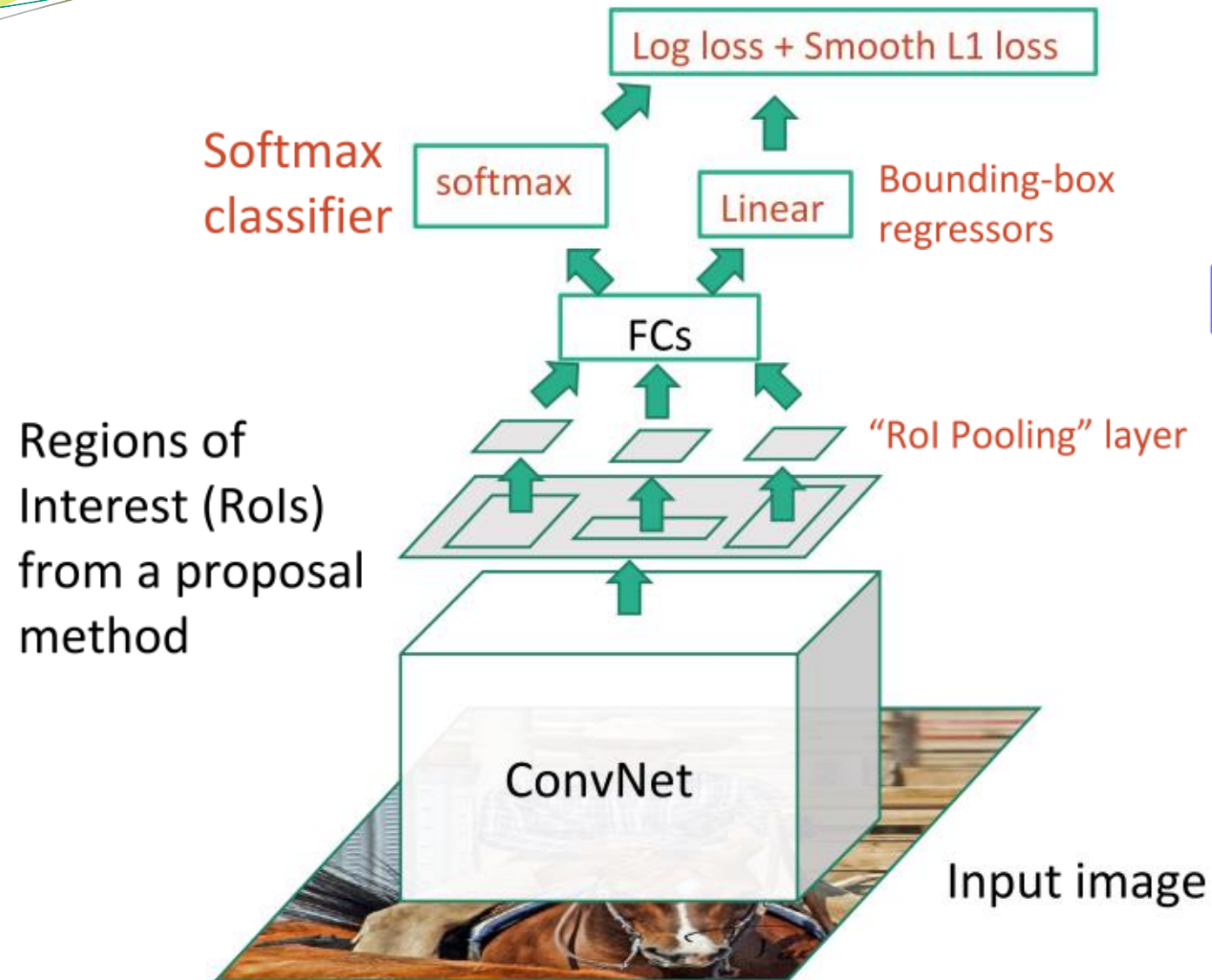
(c)



(d)



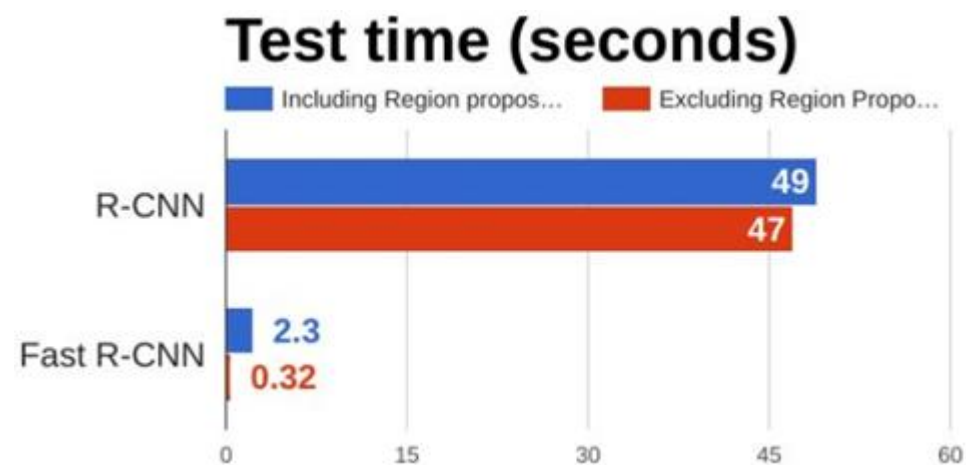
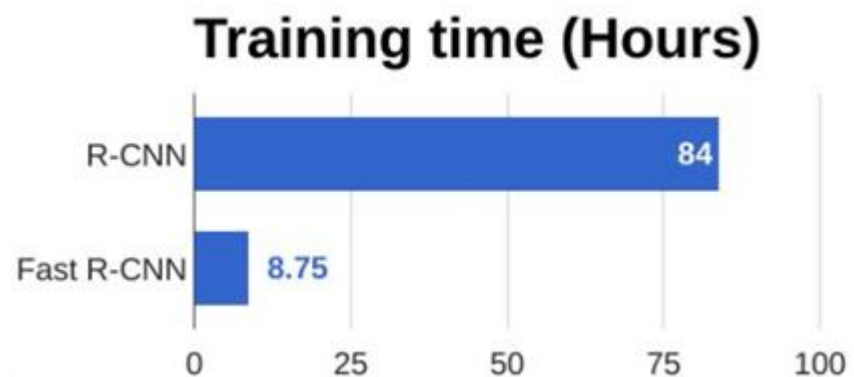
Fast R-CNN



ROI Pooling

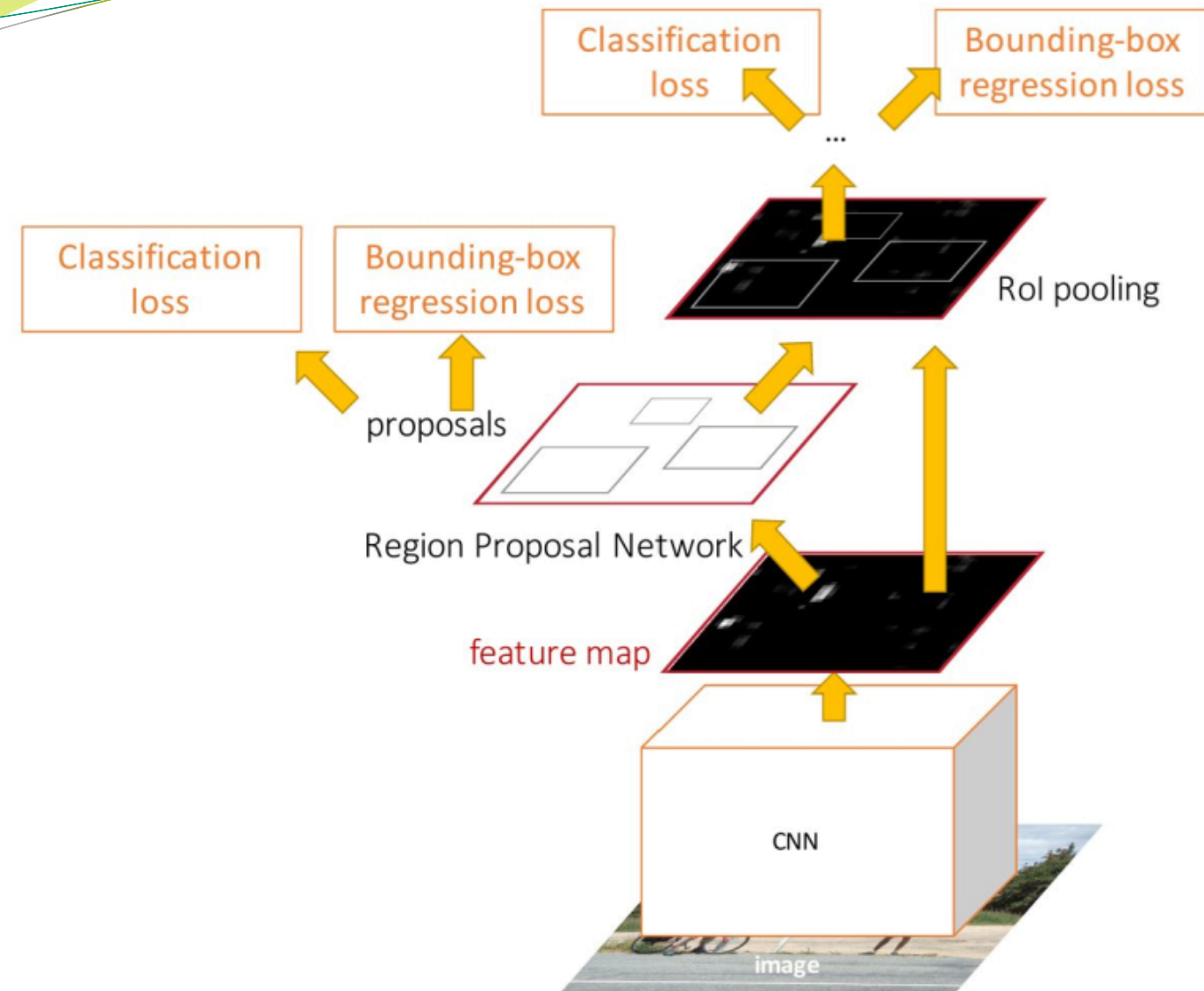


مقایسه زمانی

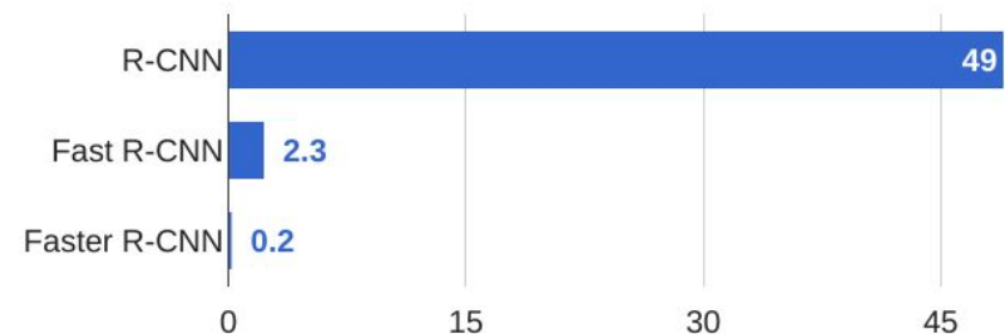


Faster R-CNN

- ناحیه‌های پیشنهادی هم توسط بخشی از شبکه (RPN) تولید می‌شوند
- آموزش شبکه end-to-end انجام می‌شود

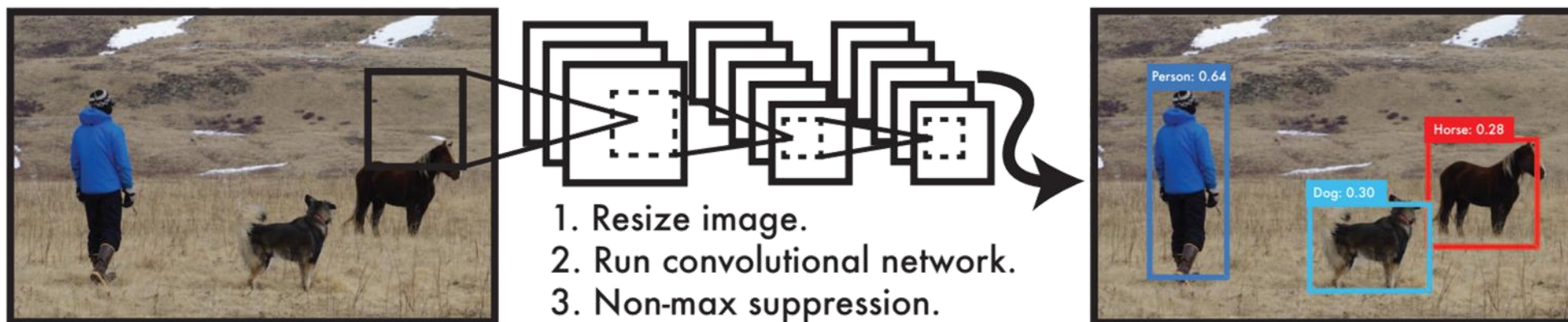


R-CNN Test-Time Speed

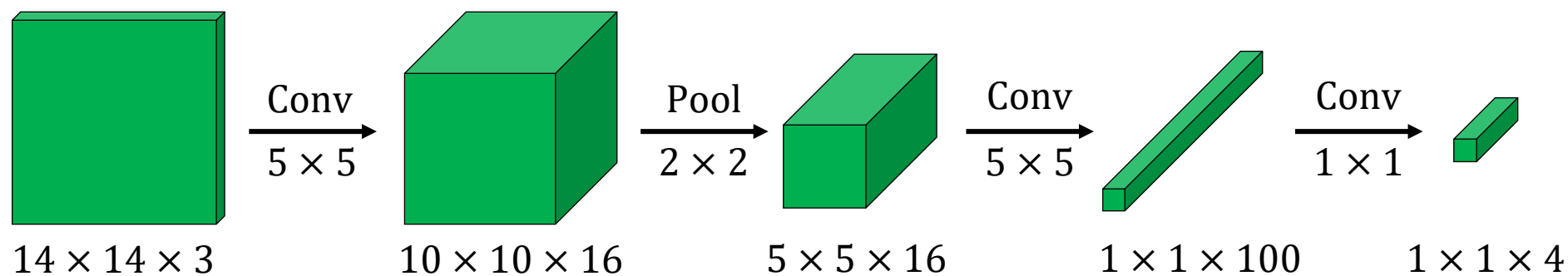
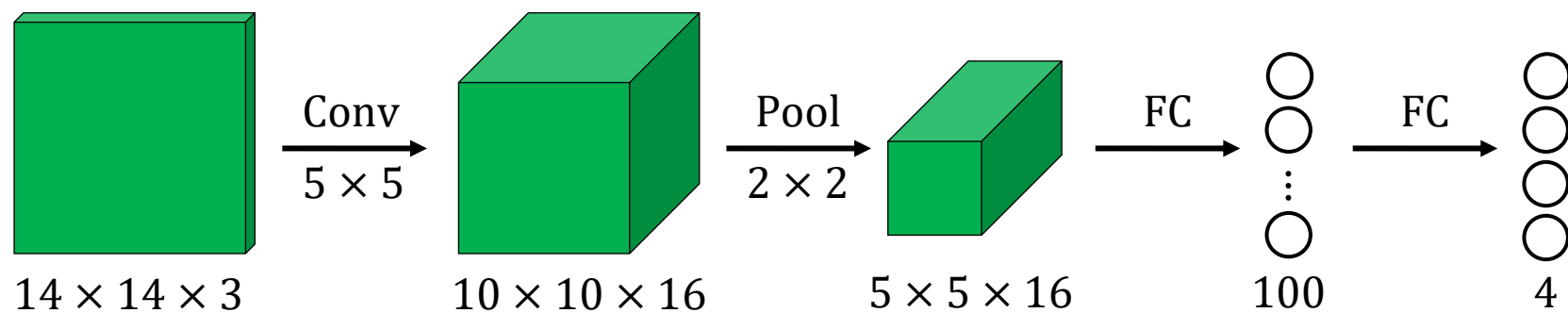


تشخیص اشیاء یک مرحله‌ای

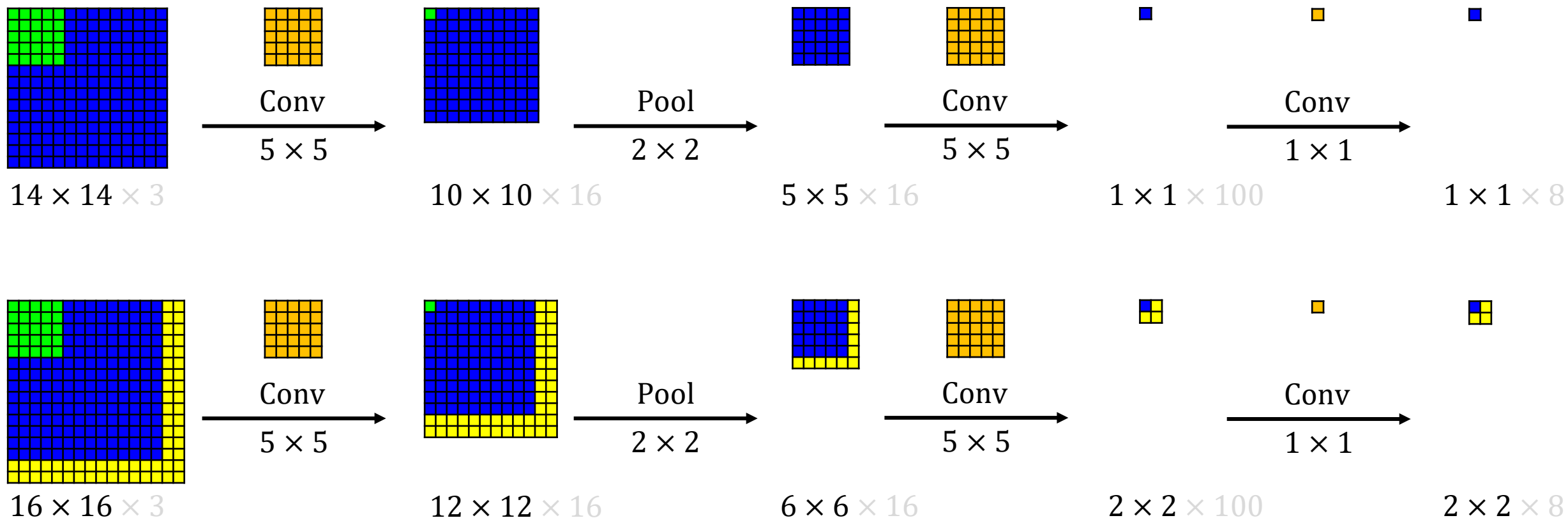
- برخی از الگوریتم‌های جدیدتر مانند YOLO، SSD و RetinaNet بدون بخش مستقیمی برای تولید ناحیه‌های پیشنهادی توسعه یافته‌اند
- این الگوریتم‌ها مبتنی بر پنجره لغزان هستند
 - به نوعی، توسعه روش دسته‌بندی و مکان‌یابی هستند



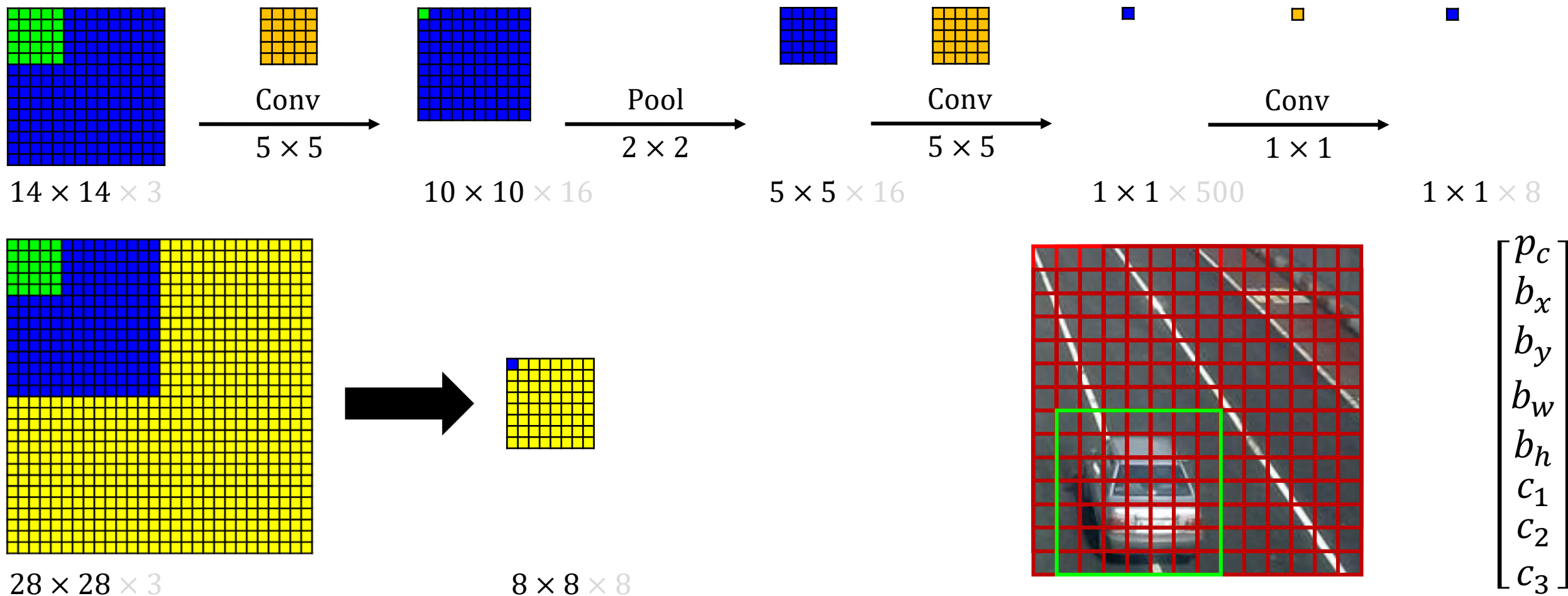
تبدیل لایه‌های Conv به FC



پیاده‌سازی کانولوشنی پنجره لغزان



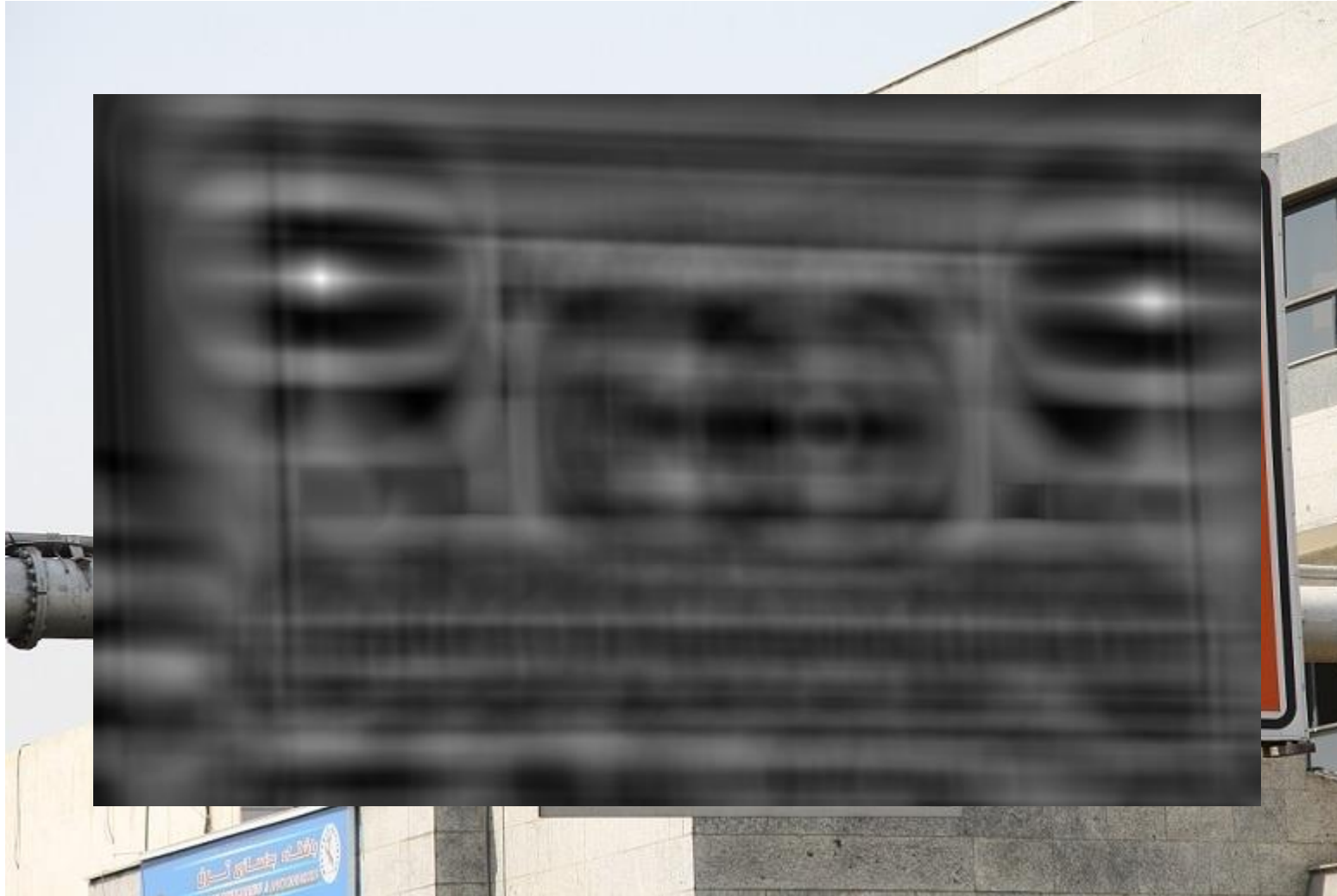
پیاده‌سازی کانولوشنی پنجره لغزان



تطبیق کلیشه

- روش تطبیق کلیشه ساده‌ترین روش برای یافتن یک شیء در تصویر است
- در این روش، یک کلیشه (template) از شیء مورد نظر ساخته شده و مشابهت (یا فاصله) هر ناحیه از تصویر نسبت به آن سنجیده می‌شود

تطبيق کلیشه



تطبيق کلیشه



- method

- TM_SQDIFF
- TM_SQDIFF_NORMED
- TM_CCORR
- TM_CCORR_NORMED
- TM_CCOEFF
- TM_CCOEFF_NORMED

$$R(x, y) = \sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2$$

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}}$$

$$R(x, y) = \sum_{x', y'} (T(x', y') \cdot I(x + x', y + y'))$$

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') \cdot I(x + x', y + y'))}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}}$$

$$R(x, y) = \sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y'))$$

$$R(x, y) = \frac{\sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x', y'} T'(x', y')^2 \cdot \sum_{x', y'} I'(x + x', y + y')^2}}$$

$$T'(x', y') = T(x', y') - 1/(w \cdot h) \cdot \sum_{x'', y''} T(x'', y'')$$

$$I'(x + x', y + y') = I(x + x', y + y') - 1/(w \cdot h) \cdot \sum_{x'', y''} I(x + x'', y + y'')$$