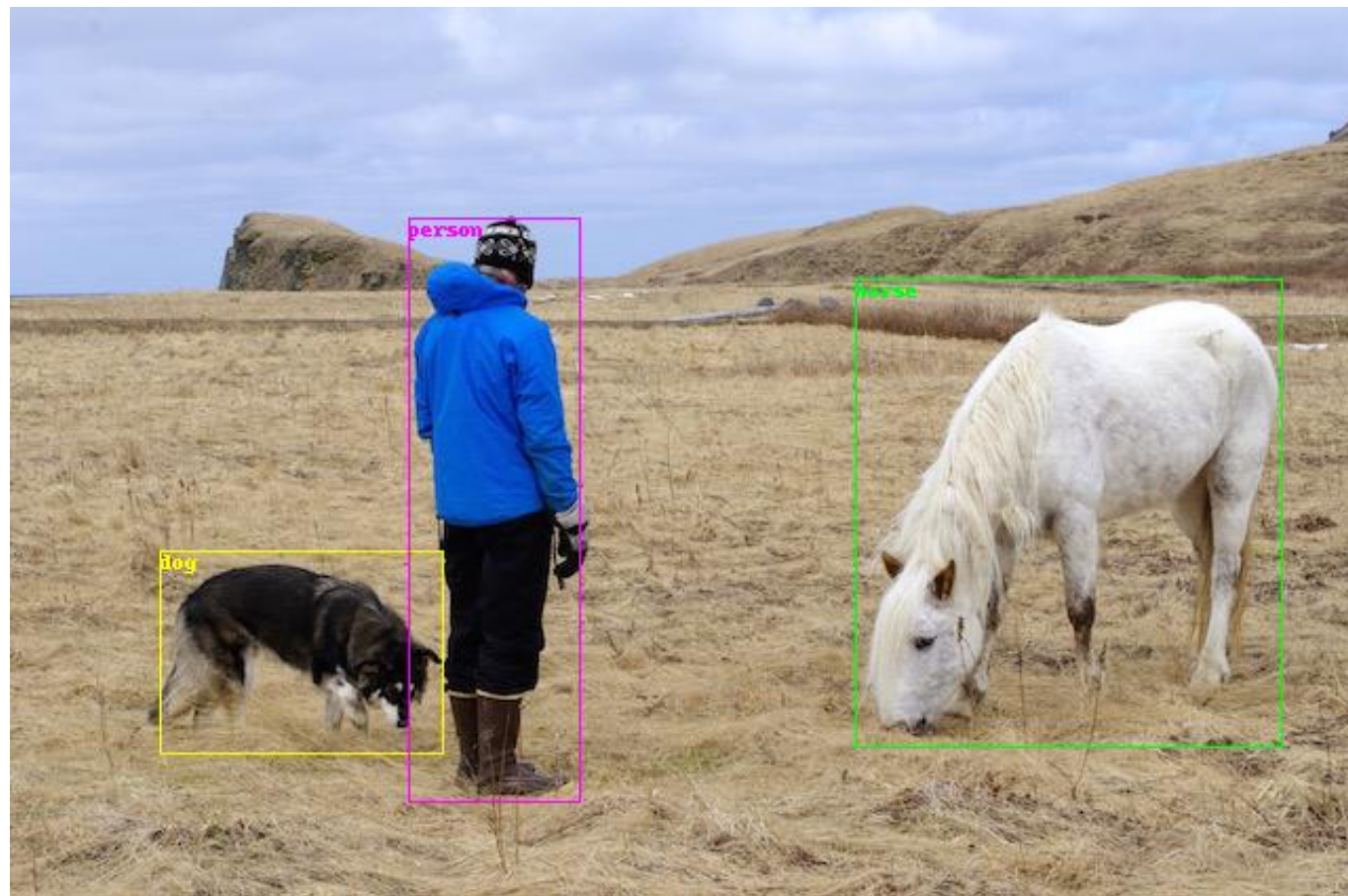


کارگاه یادگیری عمیق با پایتون: تشخیص اشیا

سید ناصر رضوی www.snrazavi.ir

۱۳۹۷



□ دسته‌بندی و مکان‌یابی.

- یک شی در تصویر
- شناسایی نقاط عطف
- تشخیص وضعیت قرارگیری

□ تشخیص اشیا.

- چند شی در تصویر
- پنجره لغزان
- شبکه‌های کاملاً کانولوشنی
- الگوریتم YOLO

تشخیص اشیا: اجرای نمایشی

۳



تشخیص اشیا

۴

□ دسته‌بندی تصویر و شناسایی اشیا.



دسته‌بندی تصویر



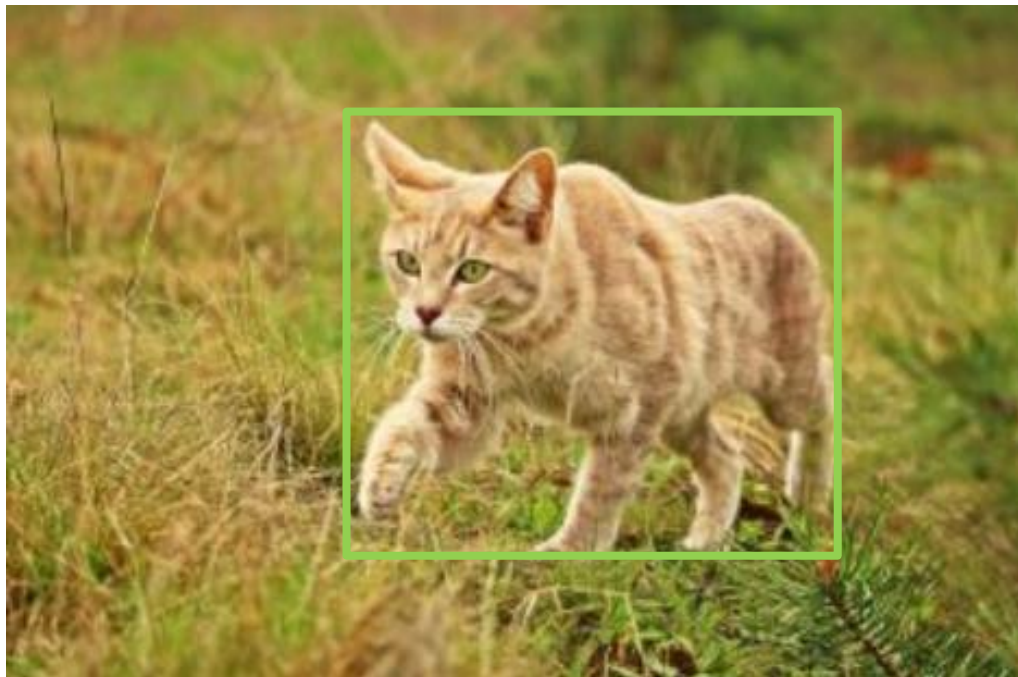
دسته‌بندی و مکان‌یابی



شناسایی اشیا

مکان‌یابی

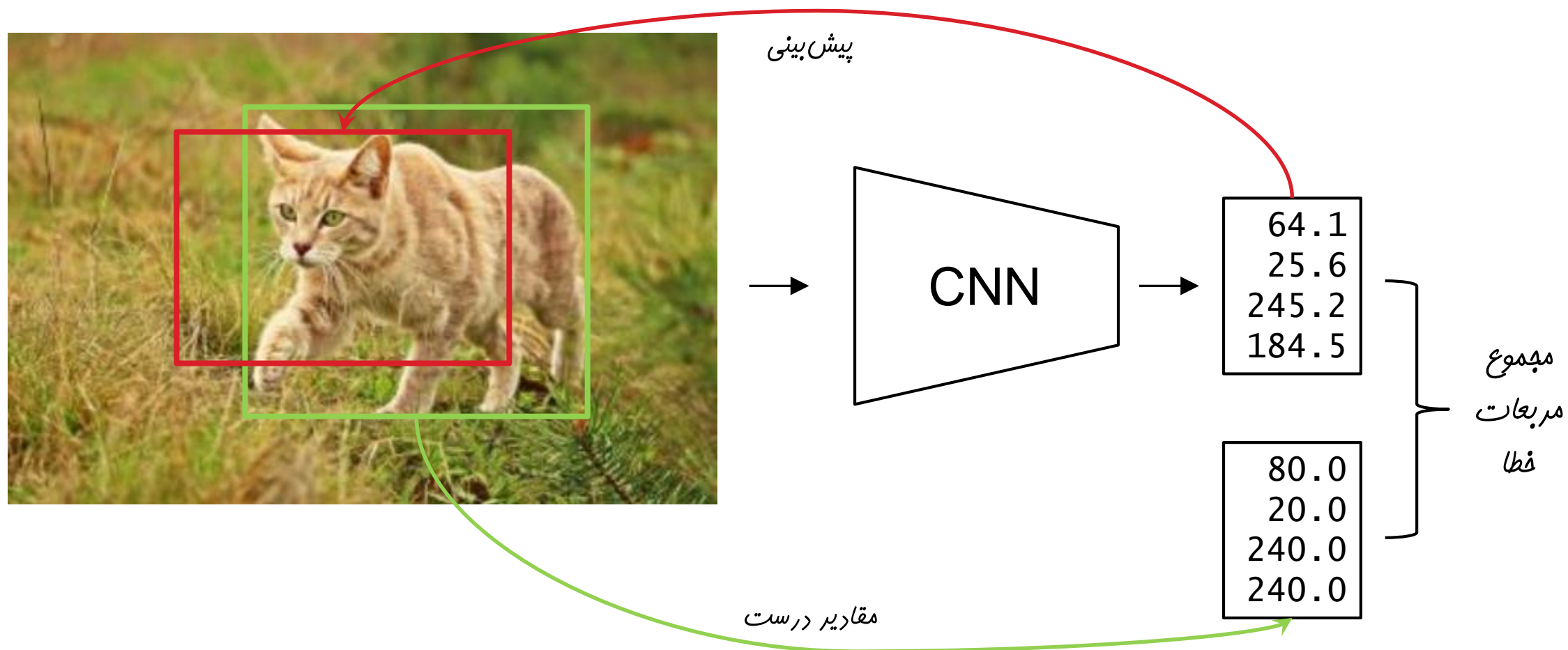
۵



- یک شی در هر تصویر
- تخمین مختصات جعبه محاطی
- ارزیابی از طریق نسبت اشتراک به اجتماع (IoU)

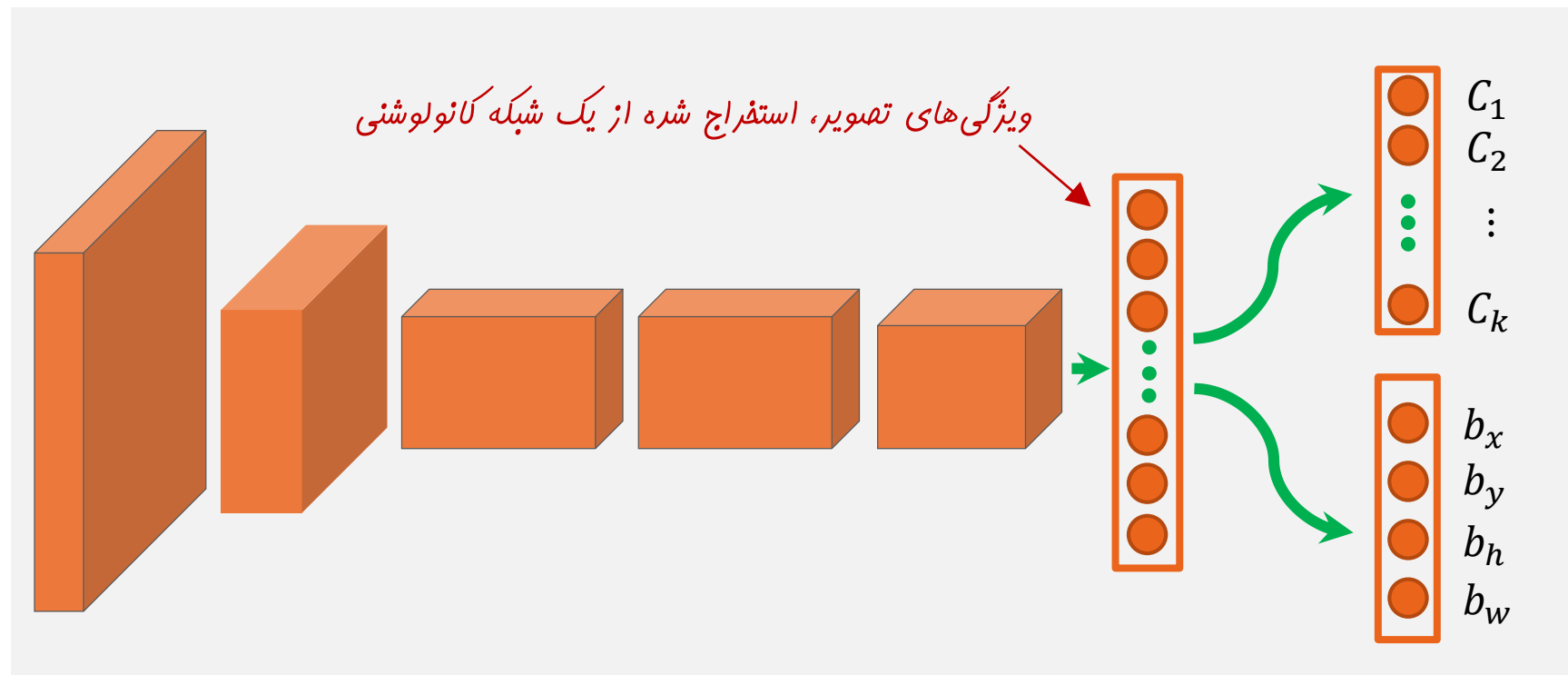
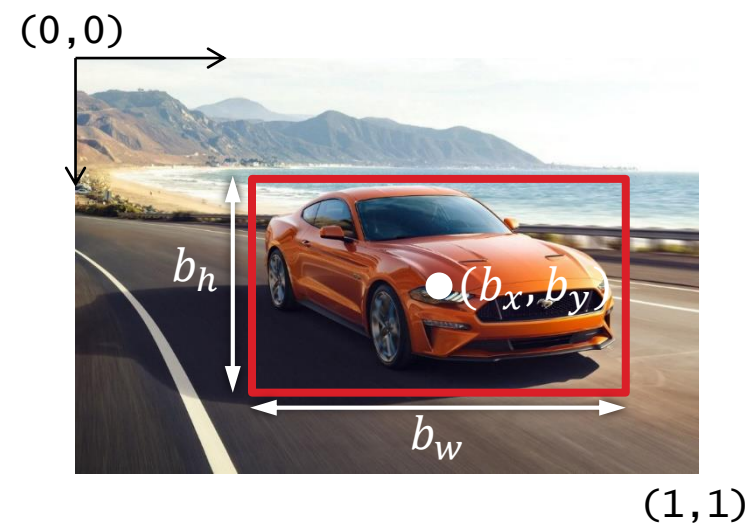
مکان‌یابی به عنوان رگرسیون

۶



دسته‌بندی و مکان‌یابی

۷



تابع هزینه: مجموع مربعات خطا

۸



$$\begin{bmatrix} 1 \\ b_x \\ b_y \\ b_h \\ b_w \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}$$

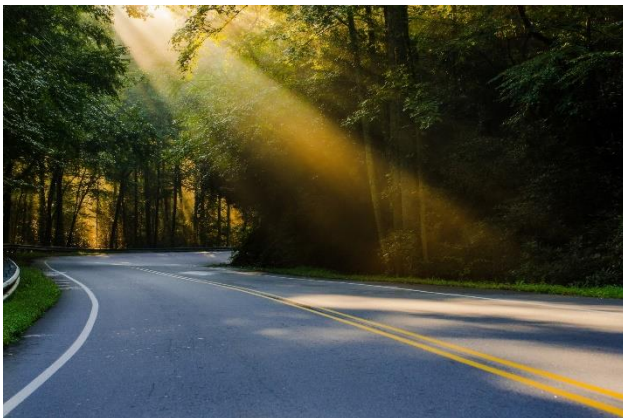
□ کلاس‌ها

۱ - رهگذر

۲ - خودرو

۳ - موتورسیکلت

۴ - پس‌زمینه



$$\begin{bmatrix} 0 \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \\ ? \end{bmatrix}$$

$$L(\hat{y}, y) = \begin{cases} \sum_{i=1}^f (\hat{y}_i - y_i)^2, & y_1 = 1 \\ (\hat{y}_1 - y_1)^2, & y_1 = 0 \end{cases}$$

تشخیص نقاط عطف: تشخیص احساسات

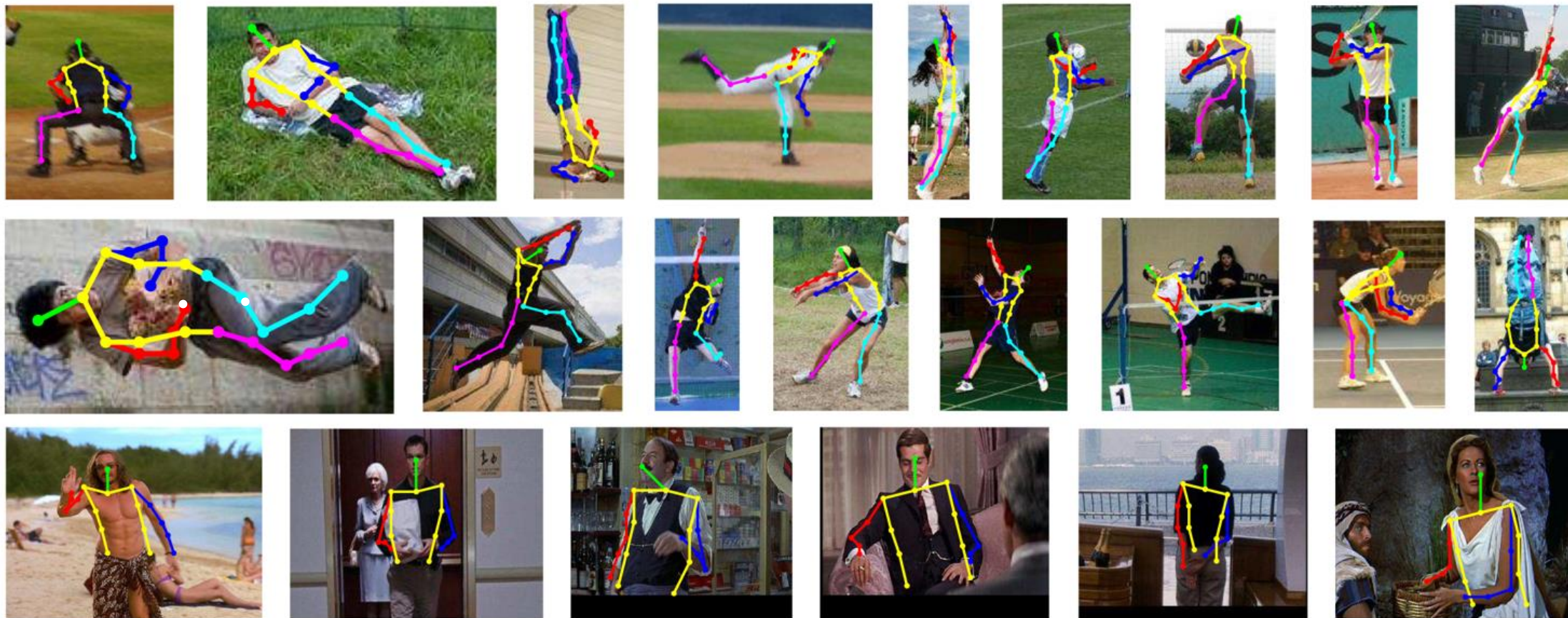
۹



Zhanpeng Zhang, et. al. Facial Landmark Detection by Deep Multi-task Learning, ECCV, 2014

تشخیص نقاط عطف: تشخیص وضعیت

۱۰

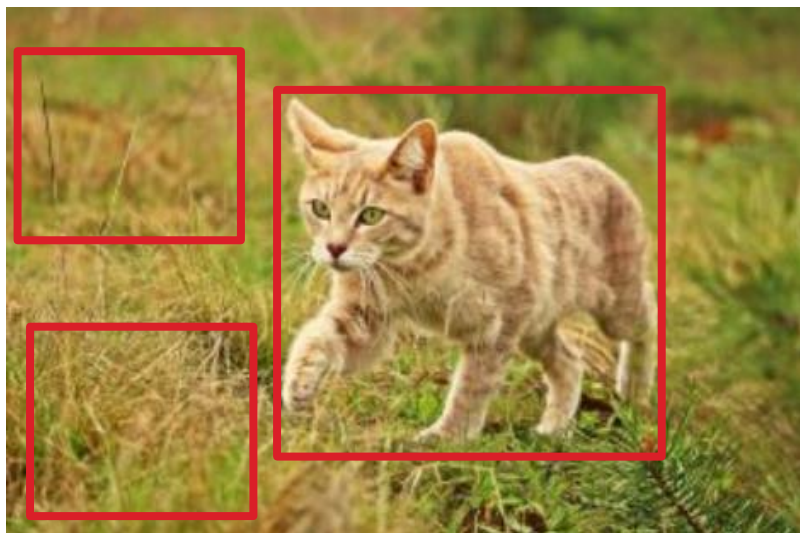


Xianjie Chen and Alan Yuille, Articulated Pose Estimation by a Graphical Model with Image Dependent Pairwise, NIPS, 2014.

شناسایی اشیاء: پنجره لغزان

۱۱

ایجاد مجموعه آموزشی



1

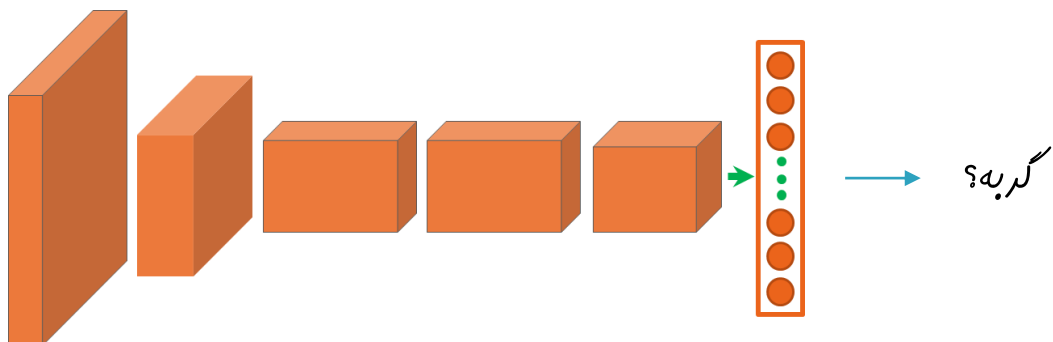


0



0

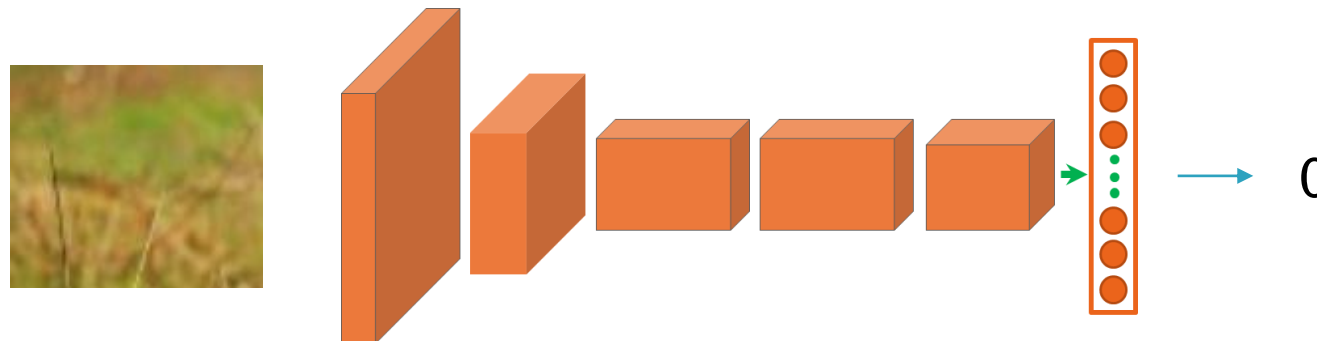
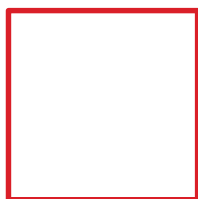
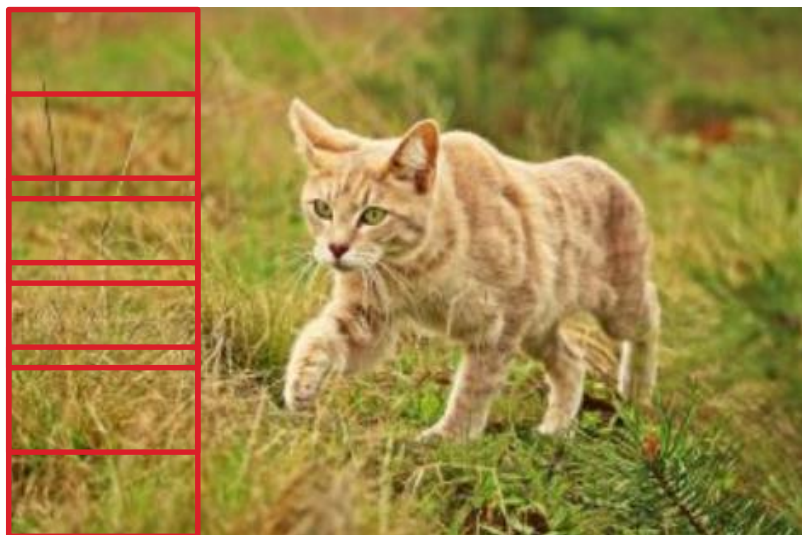
آموزش شبکه کانولوشنی



شناسایی اشیاء: پنجره لغزان

۱۲

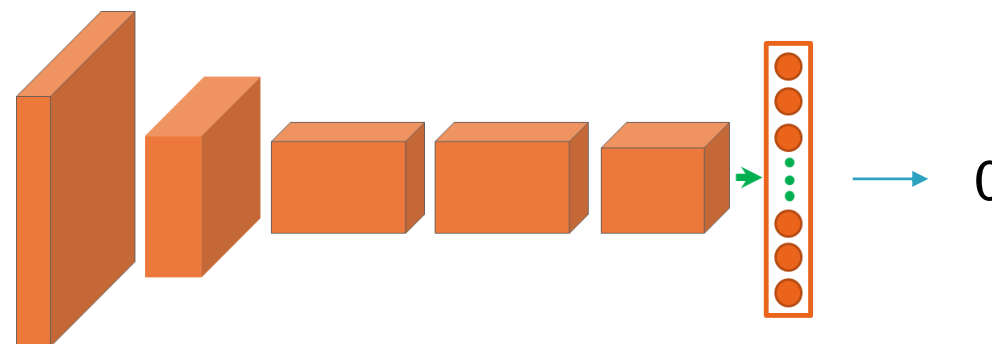
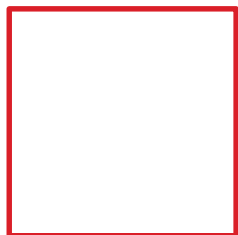
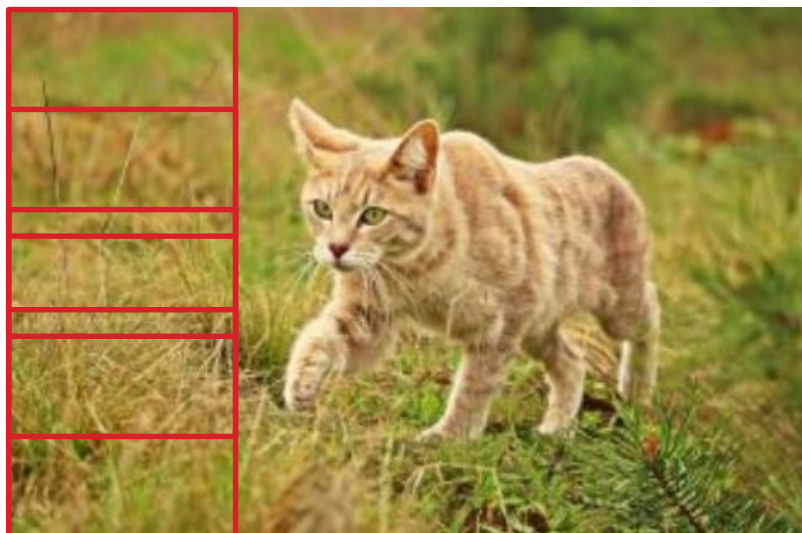
شناسایی اشیاء با پنجره لغزان



شناسایی اشیاء: پنجره لغزان

۱۳

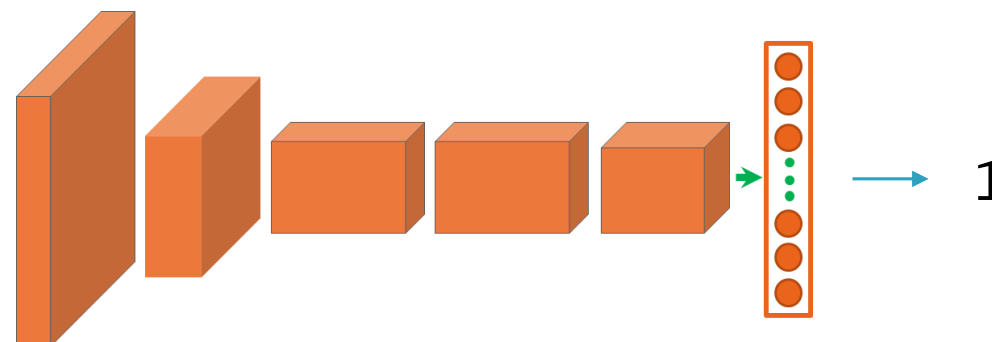
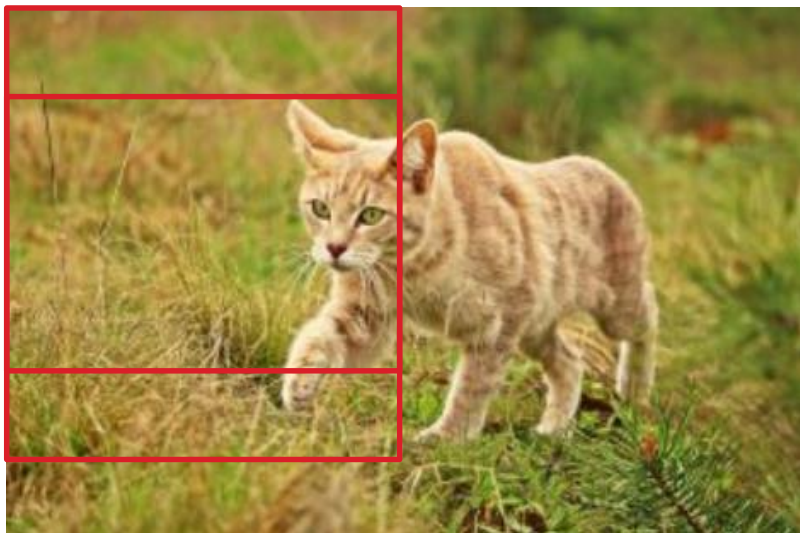
شناسایی اشیاء با پنجره لغزان



شناسایی اشیاء: پنجره لغزان

۱۴

شناسایی اشیاء با پنجره لغزان

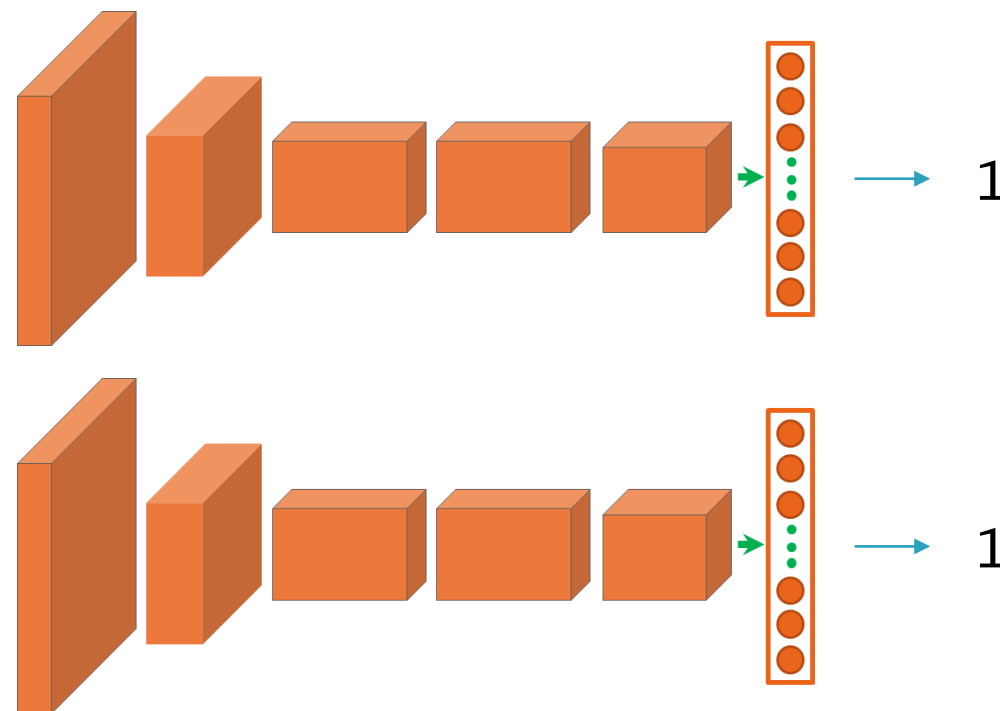
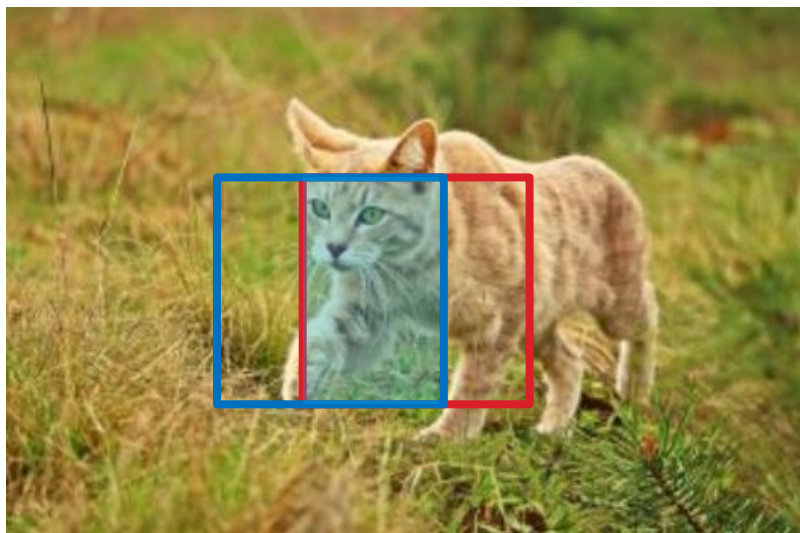


هزینه محاسباتی بسیار بالا!

پنجره لغزان: کاهش هزینه‌های محاسباتی

۱۵

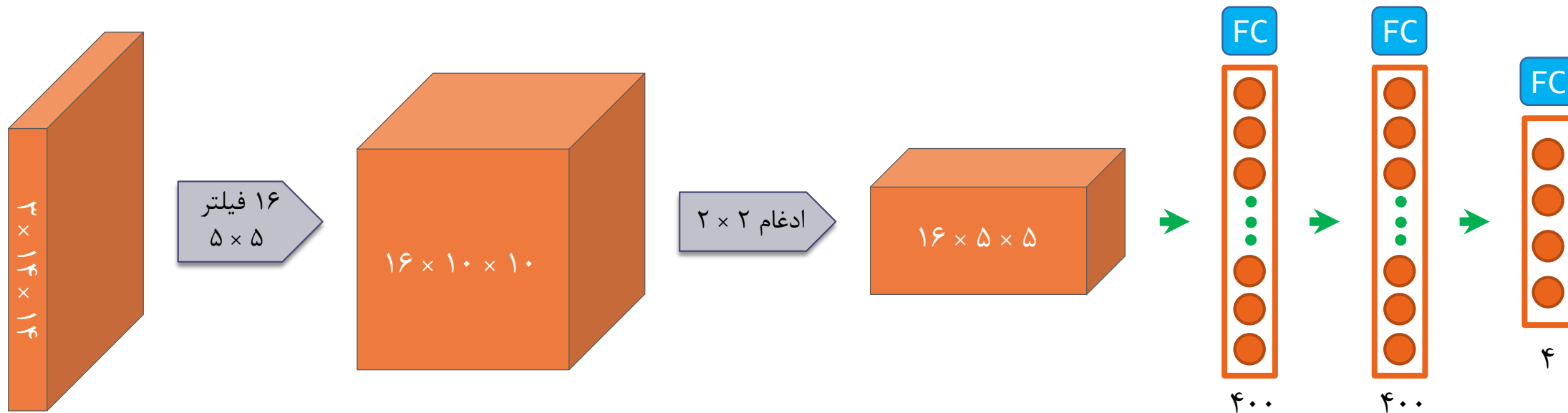
شناسایی اشیاء با پنجره لغزان



پنجره لغزان: پیاده‌سازی کانولوشنی

۱۶

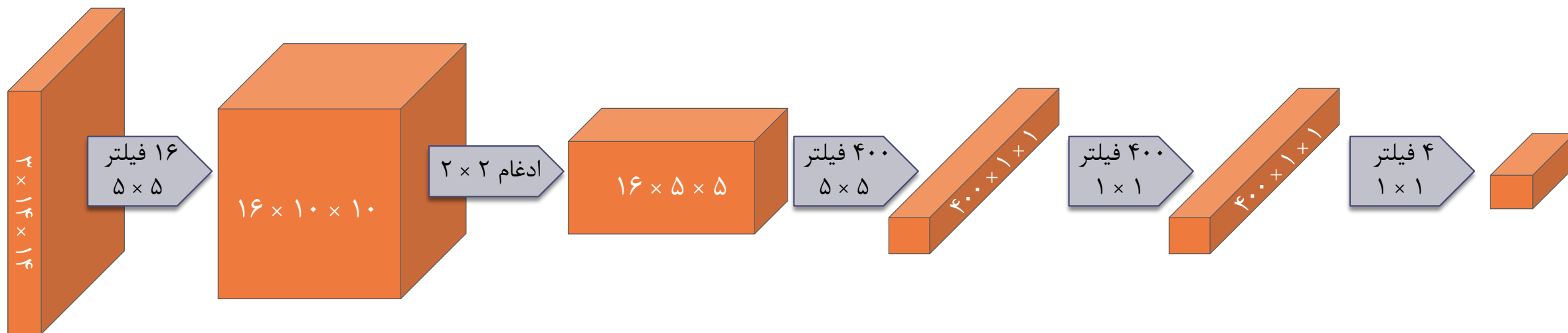
□ تبدیل لایه‌های کاملاً متصل به لایه‌های کانولوشنی.



پنجره لغزان: پیاده‌سازی کانولوشنی

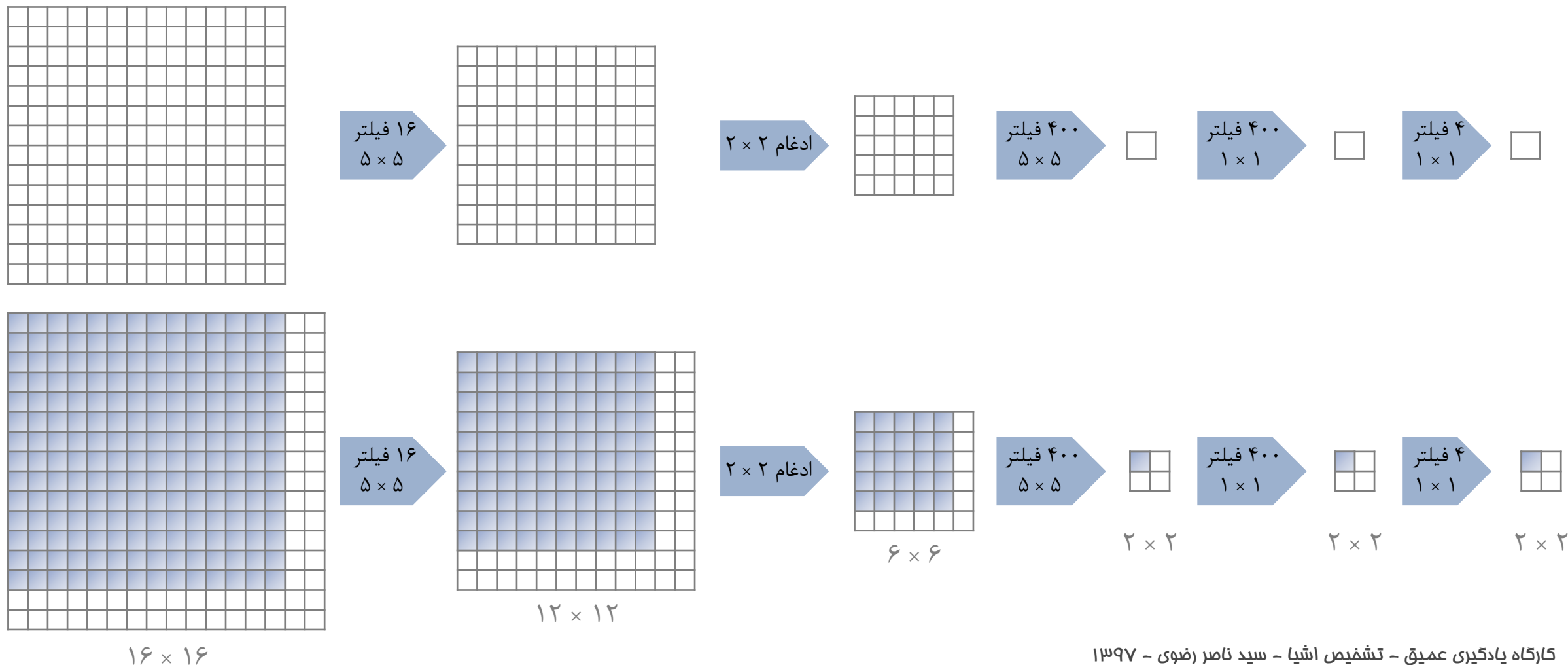
۱۷

□ تبدیل لایه‌های کاملاً متصل به لایه‌های کانولوشنی.



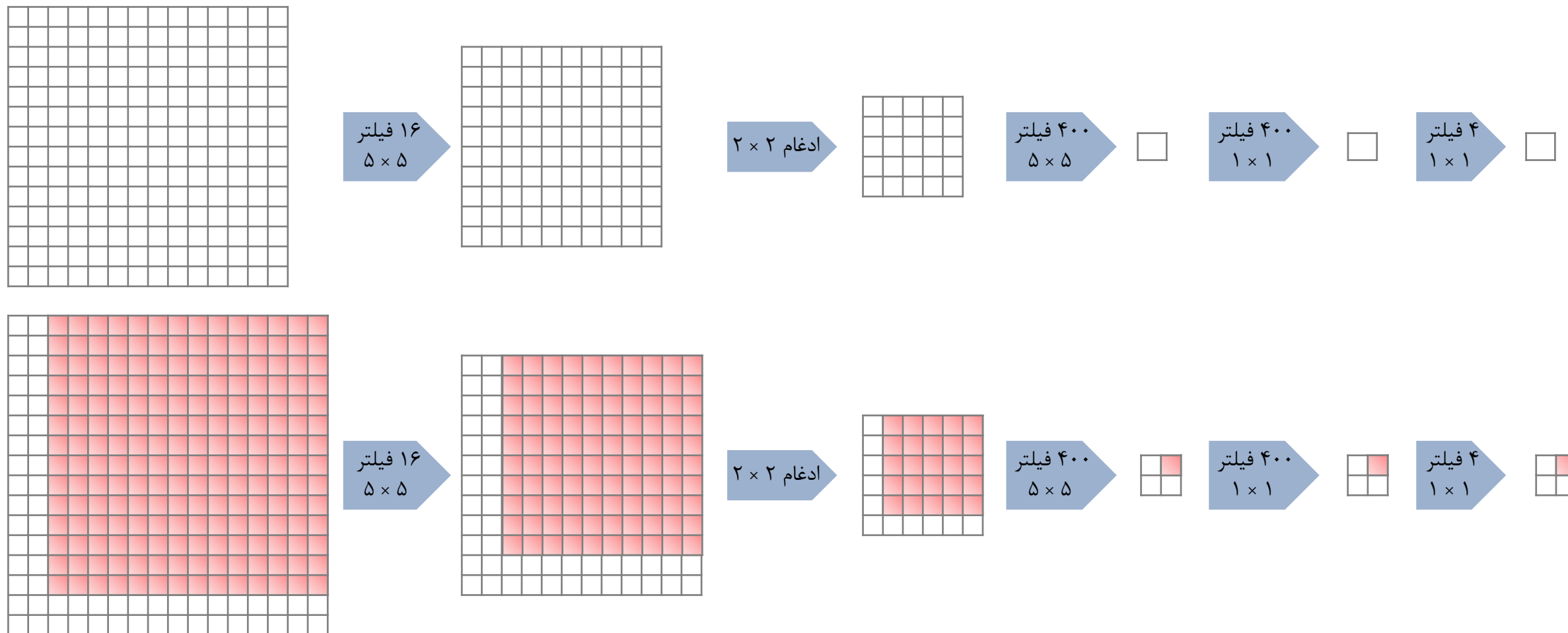
پنجره لغزان: پیاده‌سازی کانولوشنی

۱۸



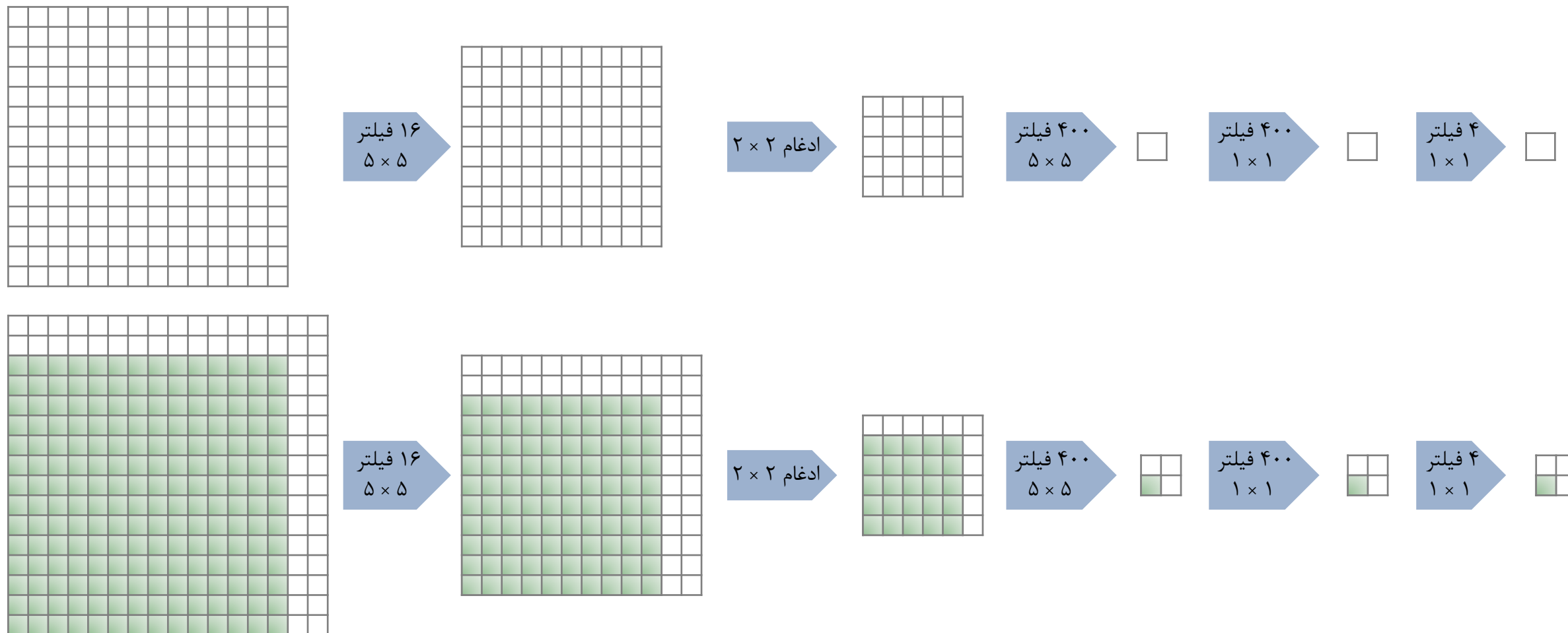
پنجره لغزان: پیاده‌سازی کانولوشنی

۱۹



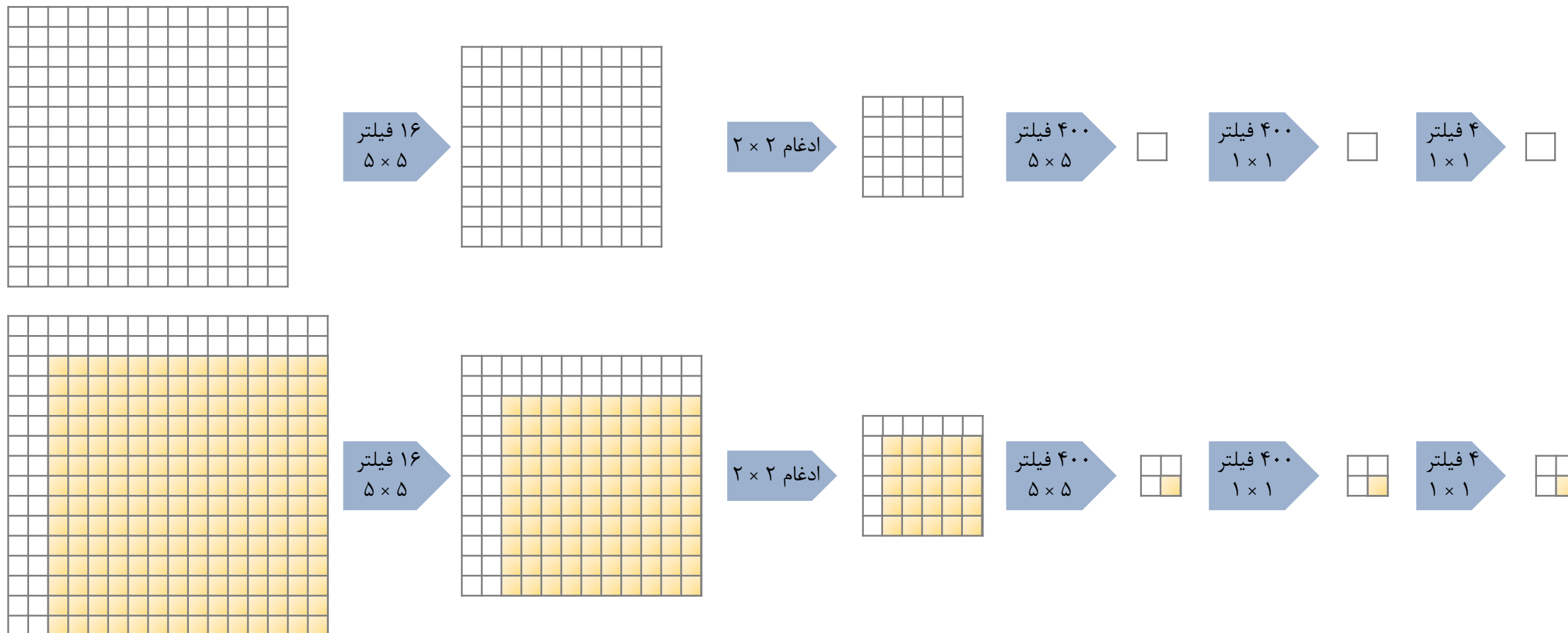
پنجره لغزان: پیاده‌سازی کانولوشنی

۲۰



پنجره لغزان: پیاده‌سازی کانولوشنی

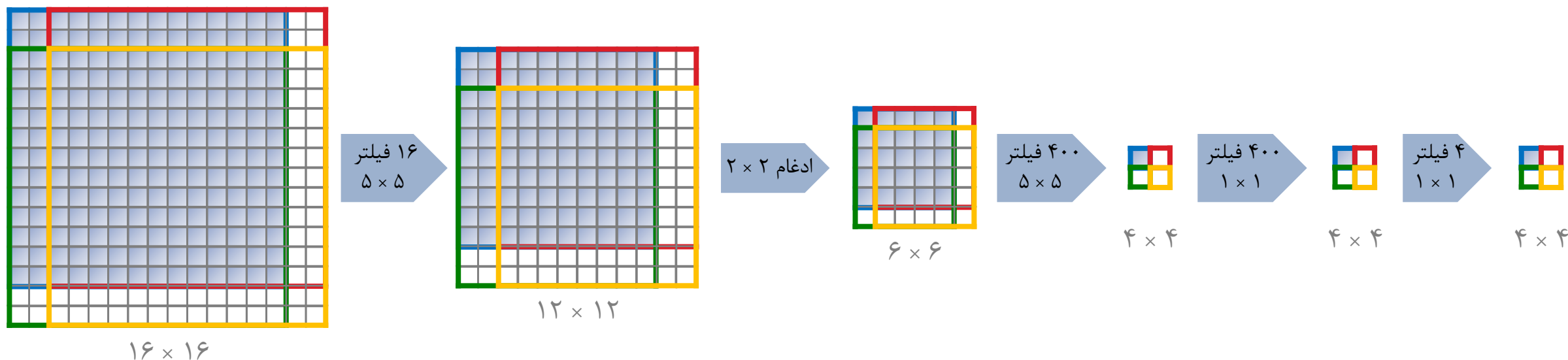
۲۱



پنجره لغزان: پیاده‌سازی کانولوشنی

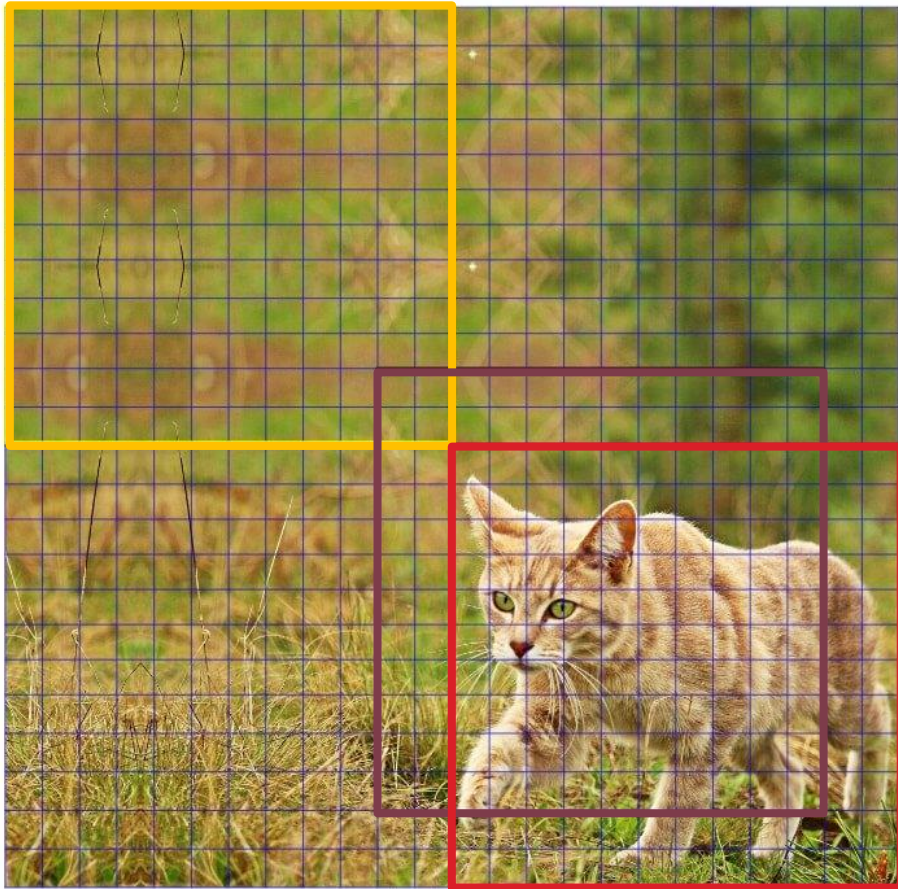
۲۲

انجام محاسبات مربوط به پنجره لغزان به صورت موازی و همزمان!

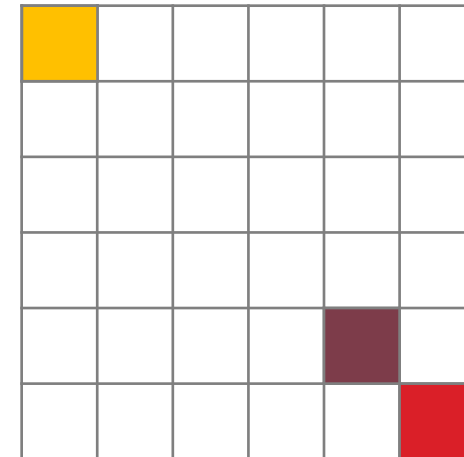


پنجره لغزان: پیاده‌سازی کانولوشنی

۲۳



CNN



احتمال وجود اشیا
در هر یک از پنجره‌ها

اشتراک بر روی اجتماع

۲۴

□ معیاری به منظور محاسبه میزان همپوشانی دو جعبه محاطی.

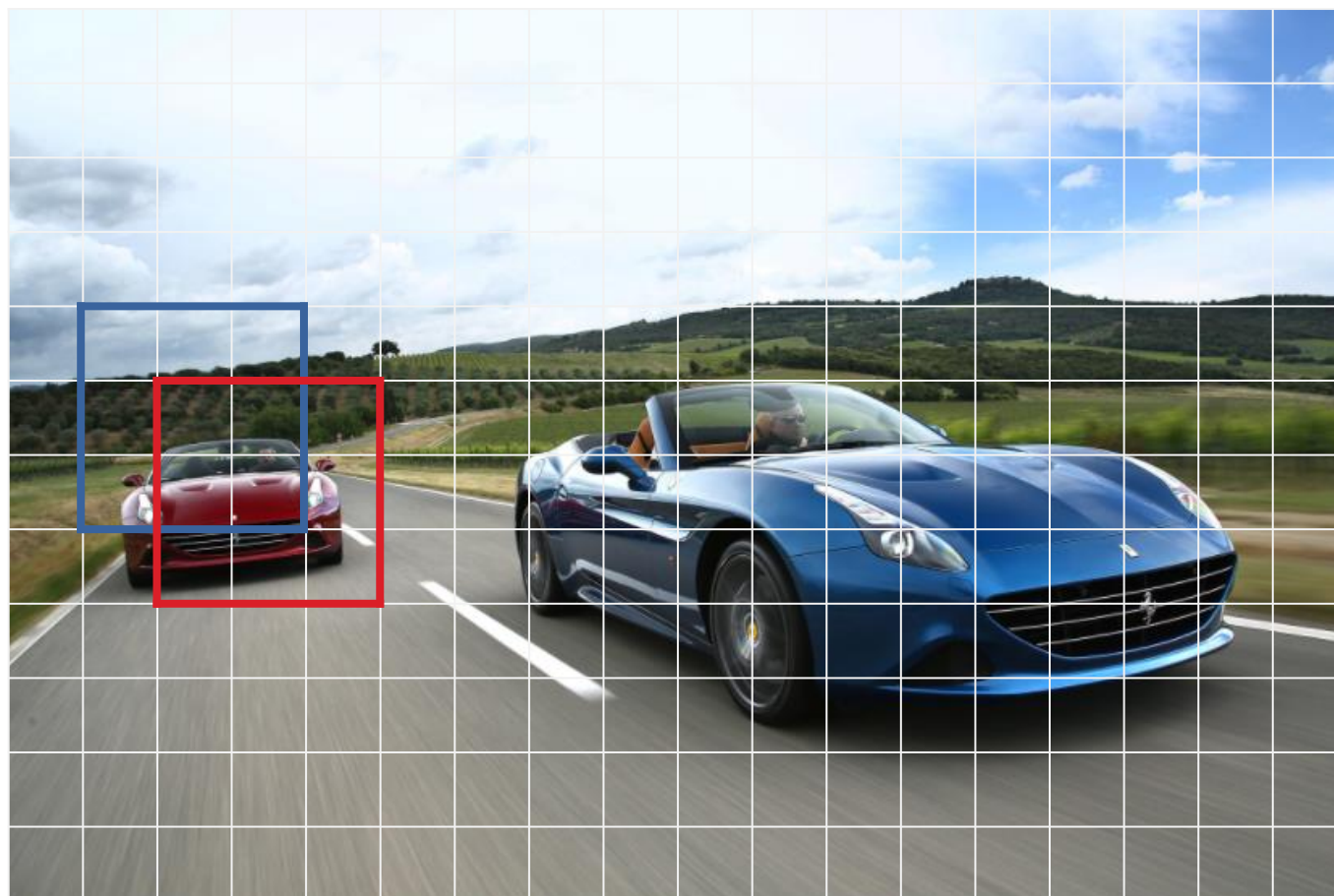


$$IoU = \frac{|B_1 \cap B_2|}{|B_1 \cup B_2|} \geq 0.5$$

سرکوب غیر بیشینه

۲۵

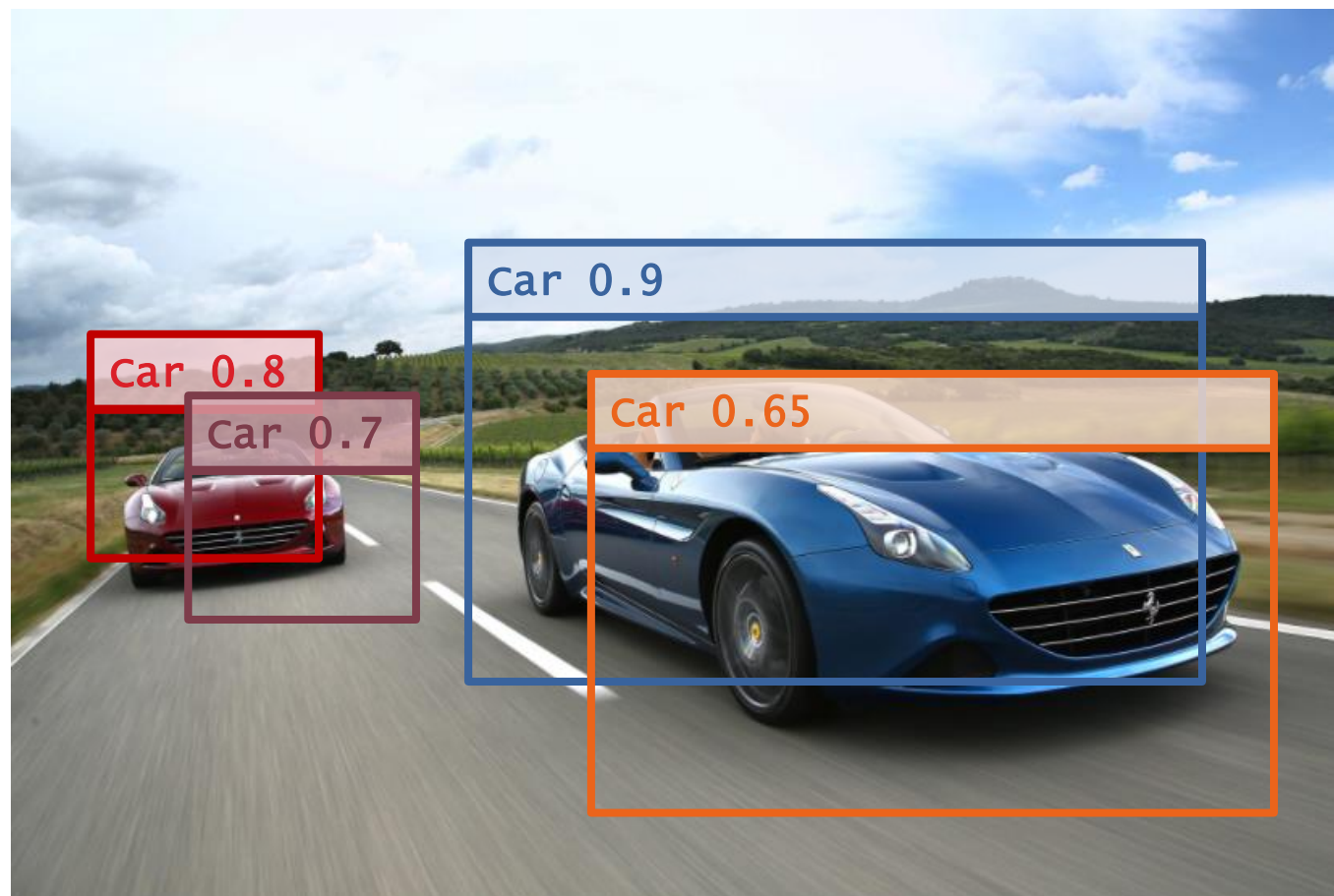
□ اطمینان از این که هر شی، بیش از یک بار شناسایی نمی شود.



سرکوب غیر بیشینه

۲۶

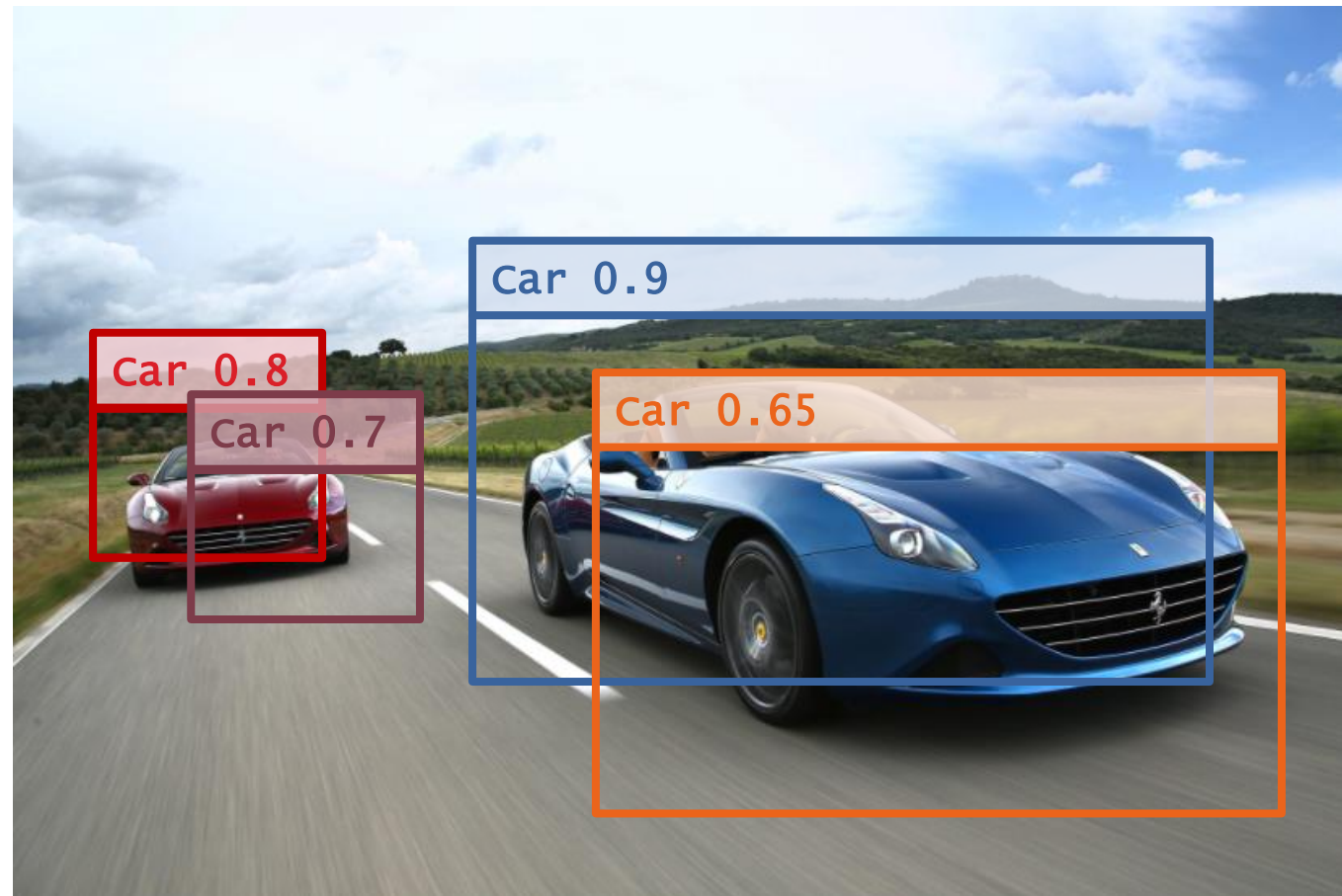
□ اطمینان از این که هر شی، بیش از یک بار شناسایی نمی شود.



سرکوب غیر بیشینه

۲۷

□ اطمینان از این که هر شی، بیش از یک بار شناسایی نمی شود.



سرکوب غیر بیشینه

۲۸

□ اطمینان از این که هر شی، بیش از یک بار شناسایی نمی شود.



سرکوب غیر بیشینه: الگوریتم

۲۹

$$y = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \end{bmatrix}$$

□ مرحله اول: حذف تمام پیش‌بینی‌ها با احتمال کمتر از ۰/۶

□ مرحله دوم: تا زمانی که پیش‌بینی (جعبه محاطی) پردازش نشده‌ای وجود دارد:

□ یک جعبه با بیشترین احتمال را انتخاب و آن را به عنوان خروجی در نظر بگیر.

□ تمامی جعبه‌هایی را که هم‌پوشانی زیادی با جعبه انتخاب شده دارند، حذف کن.


$$IoU \geq 0.5$$

همپوشانی اشیا و جعبه‌های لنگر

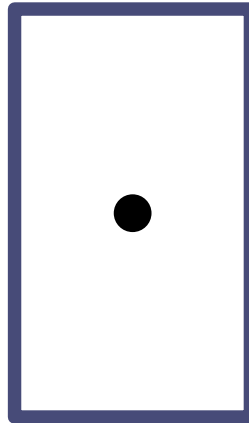
۳۰



$$y = \begin{bmatrix} y_{a1} \\ y_{a2} \end{bmatrix}$$

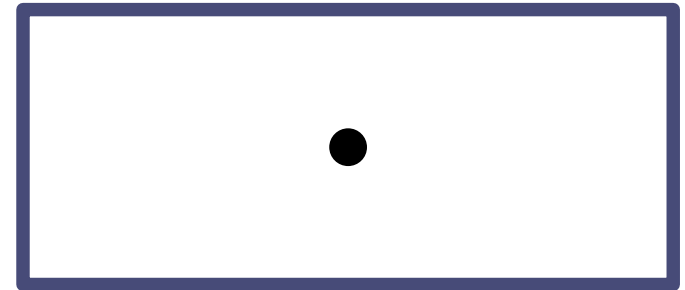
به ازای هر خانه

جعبه لنگر ۱



$$y_{a1} = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}$$

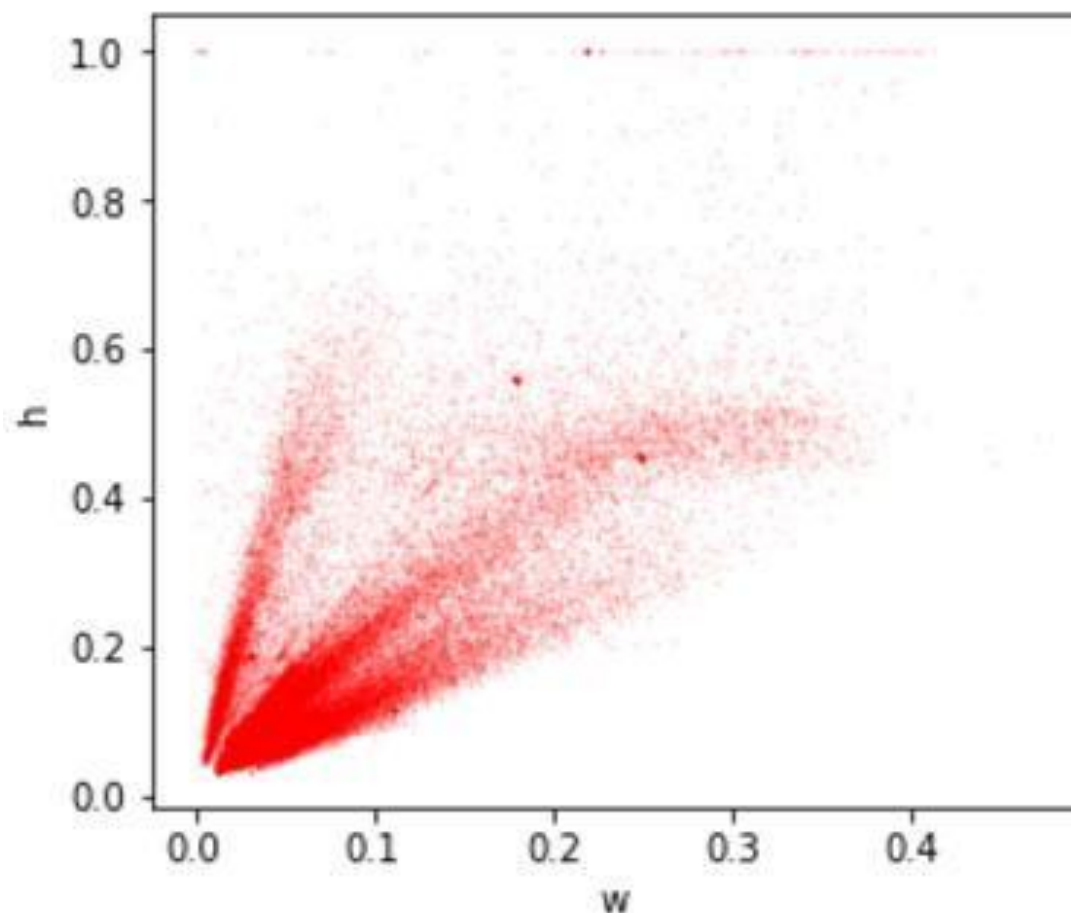
جعبه لنگر ۲



$$y_{a2} = \begin{bmatrix} p_c \\ b_x \\ b_y \\ b_h \\ b_w \\ C_1 \\ C_2 \\ C_3 \end{bmatrix}$$

جعبه‌های لنگر

۳۱



□ انتخاب جعبه‌های لنگر

□ طراحی جعبه‌های لنگر به صورت دستی

□ استفاده از الگوریتم خوشه‌بندی [۵ خوشه]

$$D(B_1, B_2) = 1 - IoU(B_1, B_2)$$

معیار فاصله

الگوریتم YOLO

۳۲

□ ایجاد مجموعه آموزشی.

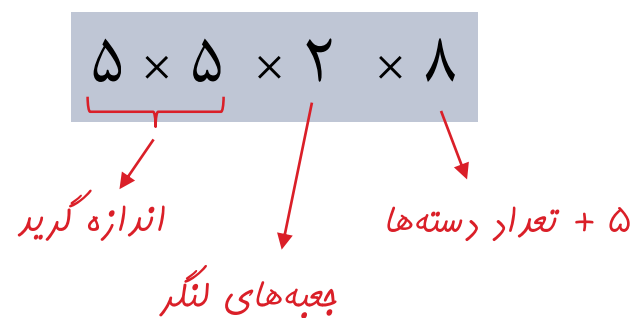
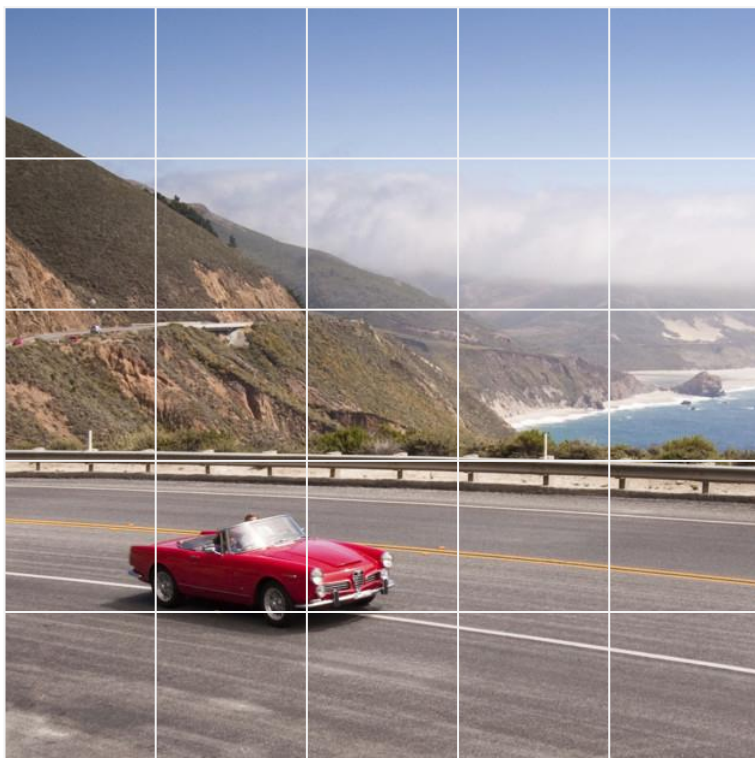
□ دسته‌ها.

۱ - رهگذر

۲ - خودرو

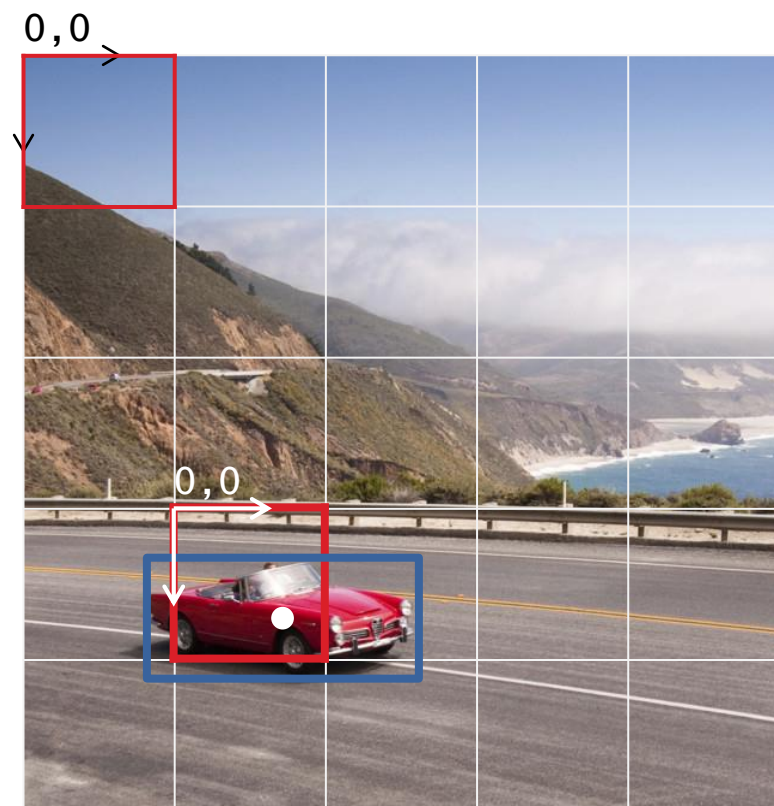
۳ - موتورسیکلت

□ خروجی.



الگوریتم YOLO

۳۳



خانه ۱

$p_c = 0$
 $b_x = 0$
 $b_y = 0$
 $b_h = 0$
 $b_w = 0$
 $C_1 = 0$
 $C_2 = 0$
 $C_3 = 0$
 $p_c = 0$
 $b_x = 0$
 $b_y = 0$
 $b_h = 0$
 $b_w = 0$
 $C_1 = 0$
 $C_2 = 0$
 $C_3 = 0$

خانه ۱۷

$p_c = 0$
 $b_x = 0$
 $b_y = 0$
 $b_h = 0$
 $b_w = 0$
 $C_1 = 0$
 $C_2 = 0$
 $C_3 = 0$
 $p_c = 1$
 $b_x = 0.7$
 $b_y = 0.7$
 $b_h = 0.8$
 $b_w = 1.9$
 $C_1 = 0$
 $C_2 = 1$
 $C_3 = 0$

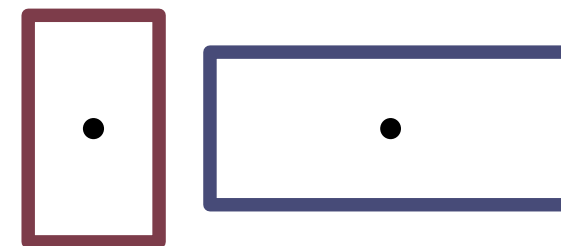
□ ایجاد مجموعه آموزشی.

□ دسته‌ها.

۱ - رهگذر

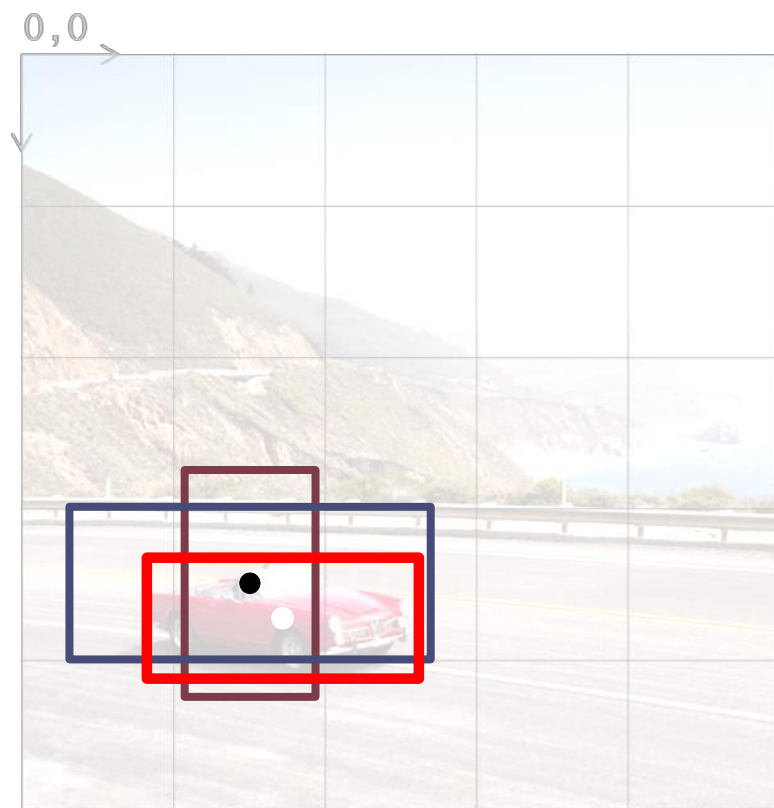
۲ - خودرو

۳ - موتورسیکلت



الگوریتم YOLO

۳۴



$19 \times 19 \times 5 \times 1$

خانه ۱

$p_c = 0$
 $b_x = 0$
 $b_y = 0$
 $b_h = 0$
 $b_w = 0$
 $C_1 = 0$
 $C_2 = 0$
 $C_3 = 0$
 $p_c = 0$
 $b_x = 0$
 $b_y = 0$
 $b_h = 0$
 $b_w = 0$
 $C_1 = 0$
 $C_2 = 0$
 $C_3 = 0$

خانه ۱۷

$p_c = 0$
 $b_x = 0$
 $b_y = 0$
 $b_h = 0$
 $b_w = 0$
 $C_1 = 0$
 $C_2 = 0$
 $C_3 = 0$
 $p_c = 1$
 $b_x = 0.7$
 $b_y = 0.7$
 $b_h = 0.8$
 $b_w = 1.9$
 $C_1 = 0$
 $C_2 = 1$
 $C_3 = 0$

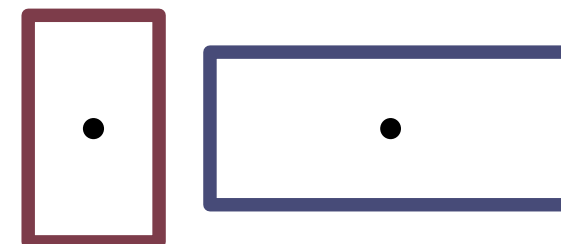
□ ایجاد مجموعه آموزشی.

□ دسته‌ها.

۱ - رهگذر

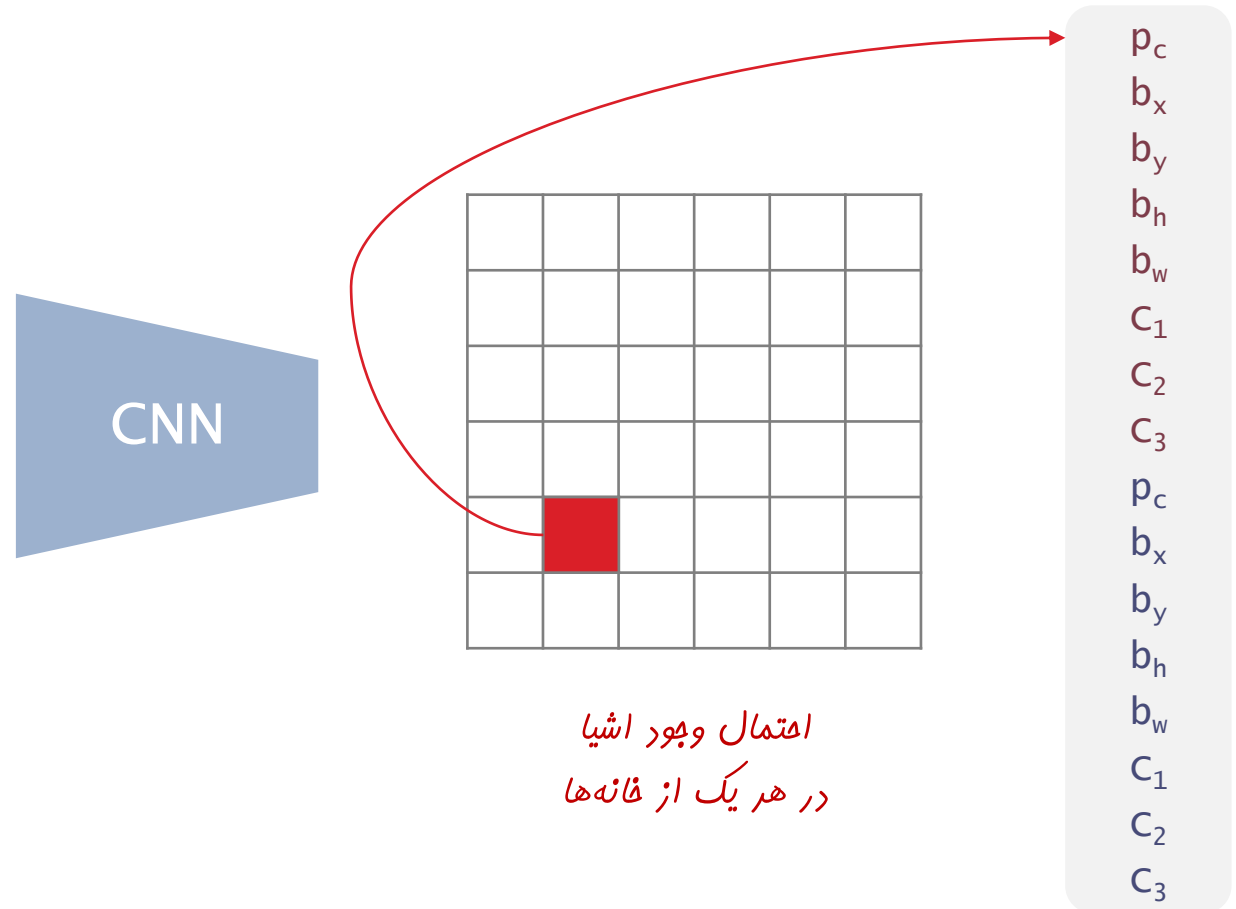
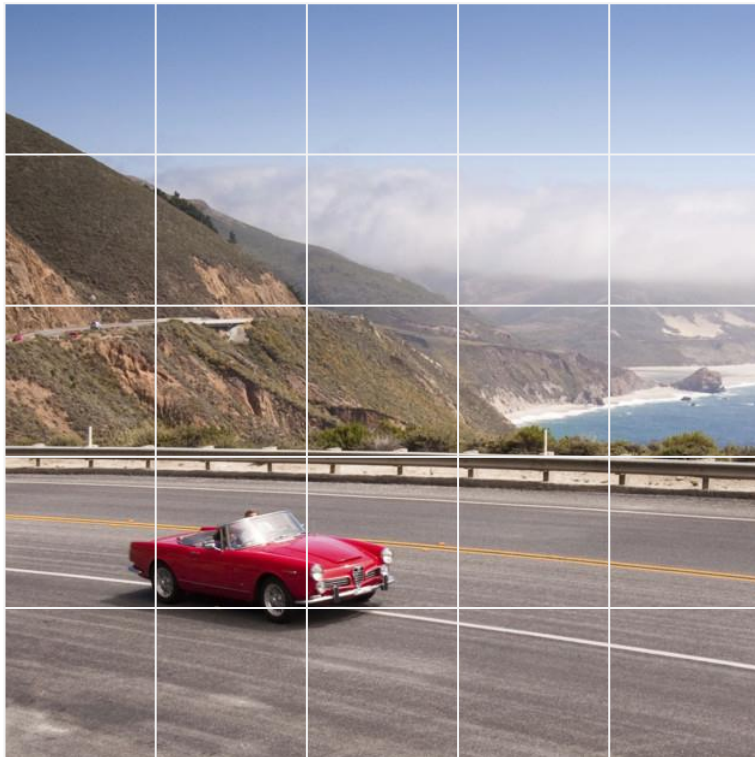
۲ - خودرو

۳ - موتورسیکلت



الگوریتم YOLO: پیش‌بینی

۳۵



الگوریتم YOLO: سرکوب غیر بیشینه

۳۶

□ پیش‌بینی‌هایی را که احتمال کمی دارند، حذف کن.

□ برای هر دسته [ره‌گذر، خودرو، موتورسیکلت]:

□ تا زمانی که پیش‌بینی دیگری وجود دارد:

- پیش‌بینی با بیشترین احتمال را انتخاب کن.
- این پیش‌بینی را به عنوان خروجی در نظر بگیر.
- تمام پیش‌بینی‌هایی را که هم‌پوشانی بالایی با پیش‌بینی انتخاب شده دارند، حذف کن.

