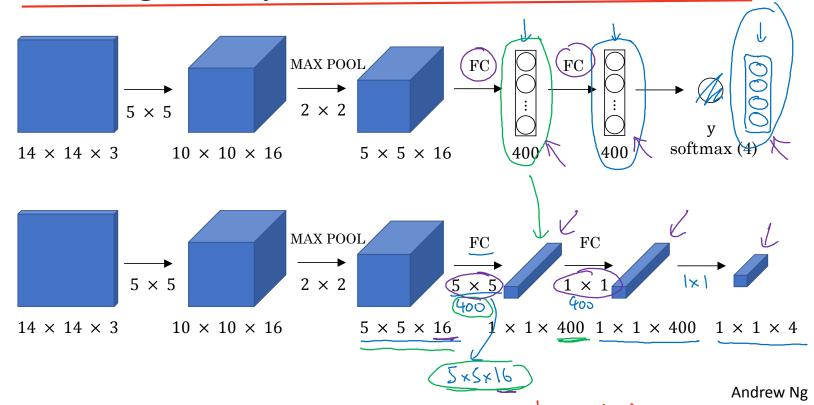


Object Detection

Convolutional implementation of sliding windows
- make all the predictions at the same time

Turning FC layer into convolutional layers



Convolution implementation of sliding windows MAX POOL FC 2×2 5×5 $5 \times 5 \times 16$ $1 \times 1 \times 400$ $1 \times 1 \times 400$ $1 \times 1 \times 4$ $14 \times 14 \times 3$ $10 \times 10 \times 16$ MAX POOL 2×2 $6 \times 6 \times 16$ $2 \times 2 \times 400$ $2 \times 2 \times 400$ $16 \times 16 \times 3$ $12 \times 12 \times 16$ MAX POOL

 $8 \times 8 \times 4$

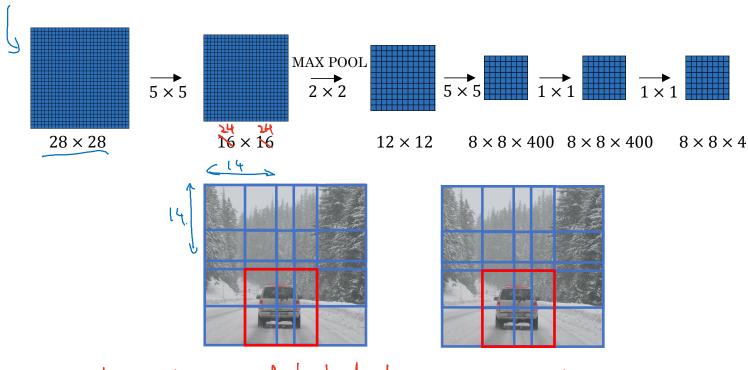
Andrew Ng

 $28 \times 28 \times 3$ $24 \times 24 \times 16$ $12 \times 12 \times 16$ $8 \times 8 \times 400$ $8 \times 8 \times 400$ [Sermanet et al., 2014, OverFeat: Integrated recognition, localization and detection using convolutional networks]

 2×2

 5×5

Convolution implementation of sliding windows



weakness: the position of the bounding boxes is not goly to be too accurate.

Andrew Ng