Sequence to sequence models

Basic models

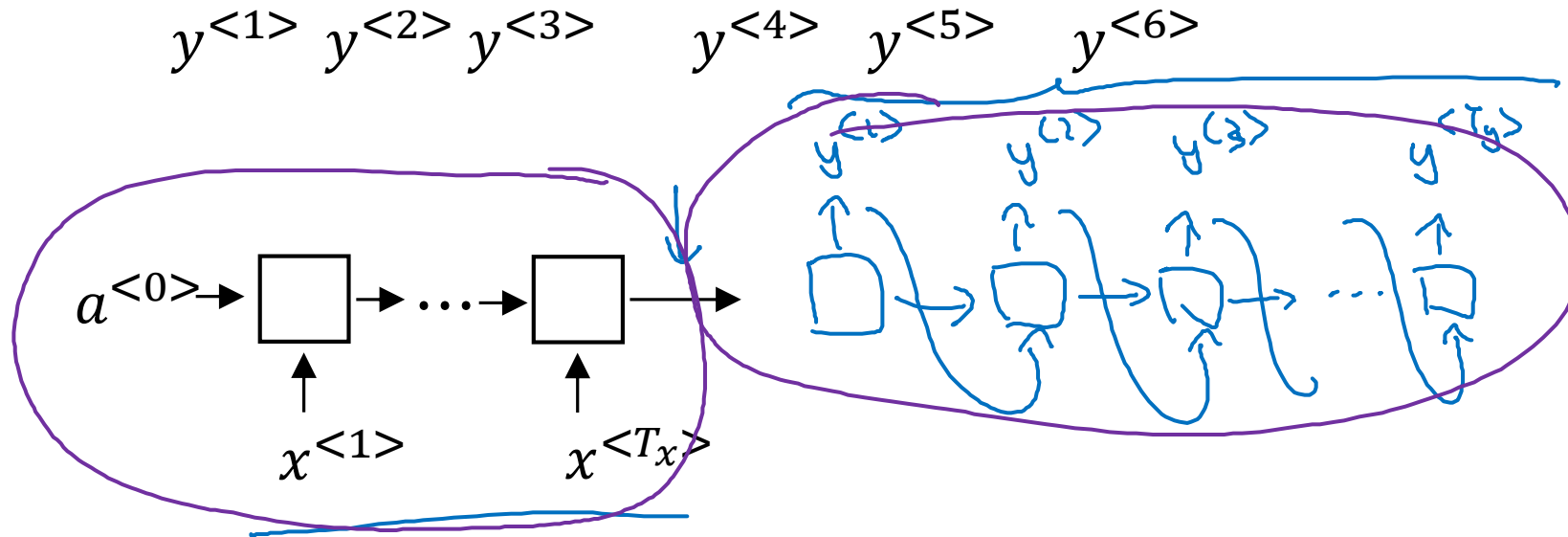deeplearning.ai

Sequence to sequence model

$x^{<1>}$ $x^{<2>}$ $x^{<3>}$ $x^{<4>}$ $x^{<5>}$

Jane visite l'Afrique en septembre

$\longrightarrow$ Jane is visiting Africa in September.

$y^{<1>}$ $y^{<2>}$ $y^{<3>}$ $y^{<4>}$ $y^{<5>}$ $y^{<6>}$

$a^{<0>} \rightarrow \square \rightarrow \cdots \rightarrow \square \rightarrow$

$x^{<1>}$ $x^{<T_x>}$
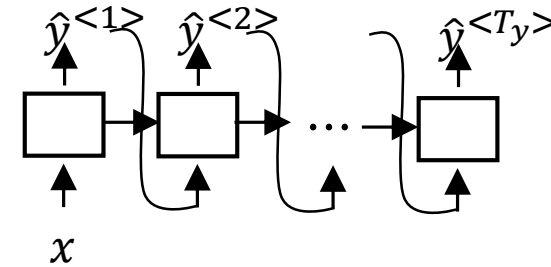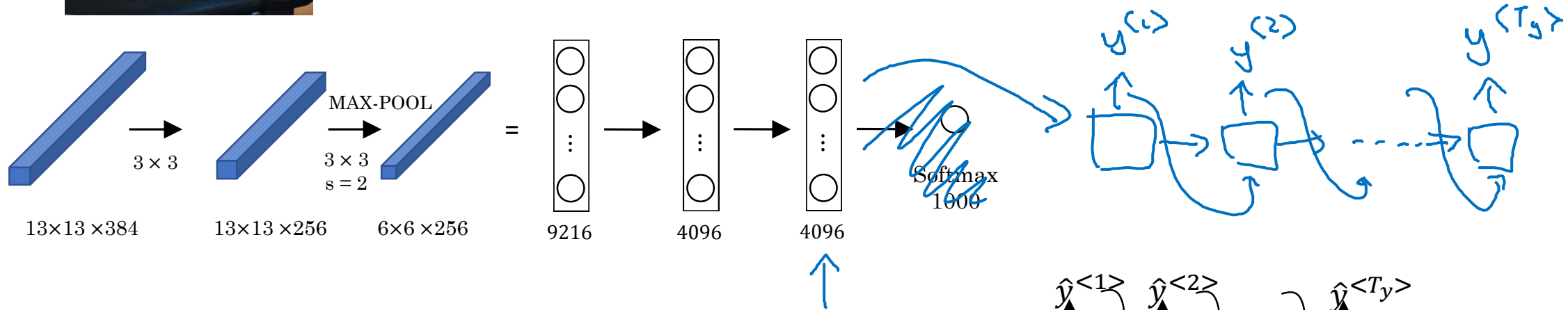
[Sutskever et al., 2014. Sequence to sequence learning with neural networks]

[Cho et al., 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation]

Andrew Ng
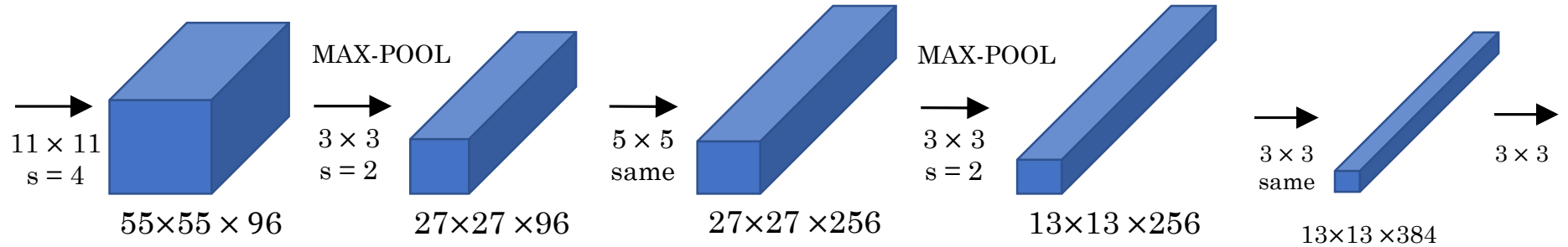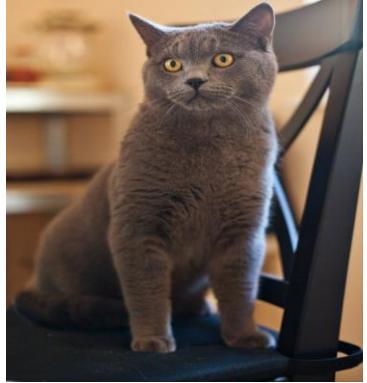
Image captioning

$y^{<1>}$ $y^{<2>}$ $y^{<3>}$ $y^{<4>}$ $y^{<5>}$ $y^{<6>}$

A cat sitting   on   a   chair

11 × 11
s = 4

55×55 × 96

MAX-POOL

3 × 3
s = 2

27×27 ×96

5 × 5
same

27×27 ×256

MAX-POOL

3 × 3
s = 2

13×13 ×256

3 × 3
same

13×13 ×384

3 × 3

13×13 ×384

3 × 3

13×13 ×256

MAX-POOL

3 × 3
s = 2

6×6 ×256

=

9216

4096

4096

Softmax
1000

$y^{(1)}$ $y^{(2)}$ $y^{<T_y>}$

$\hat{y}^{<1>}$ $\hat{y}^{<2>}$ $\hat{y}^{<T_y>}$

$\cdots$

x

[Mao et. al., 2014. Deep captioning with multimodal recurrent neural networks]
[Vinyals et. al., 2014. Show and tell: Neural image caption generator]
[Karpathy and Li, 2015. Deep visual-semantic alignments for generating image descriptions]

Andrew Ng