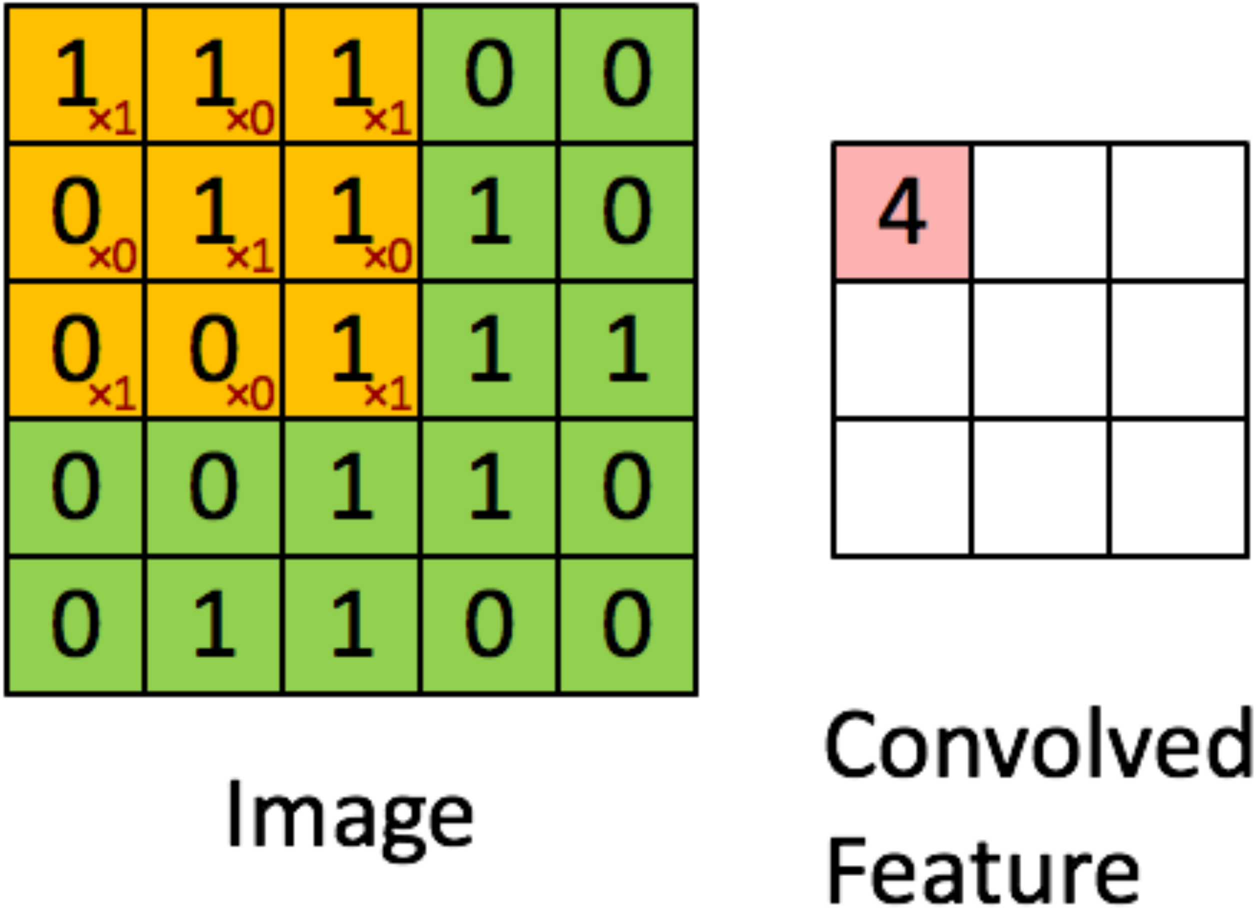
- Convolutional filtering is a general class of techniques used in signal processing for filtering continuous or discrete signals. It's based on the idea of a *finite impulse response filter* (FIR). In deep learning this was first applied to vision, wherein the convolutional filters are 2-dimensional patches (of varying sizes) that mimic the receptive fields in the human visual cortex.

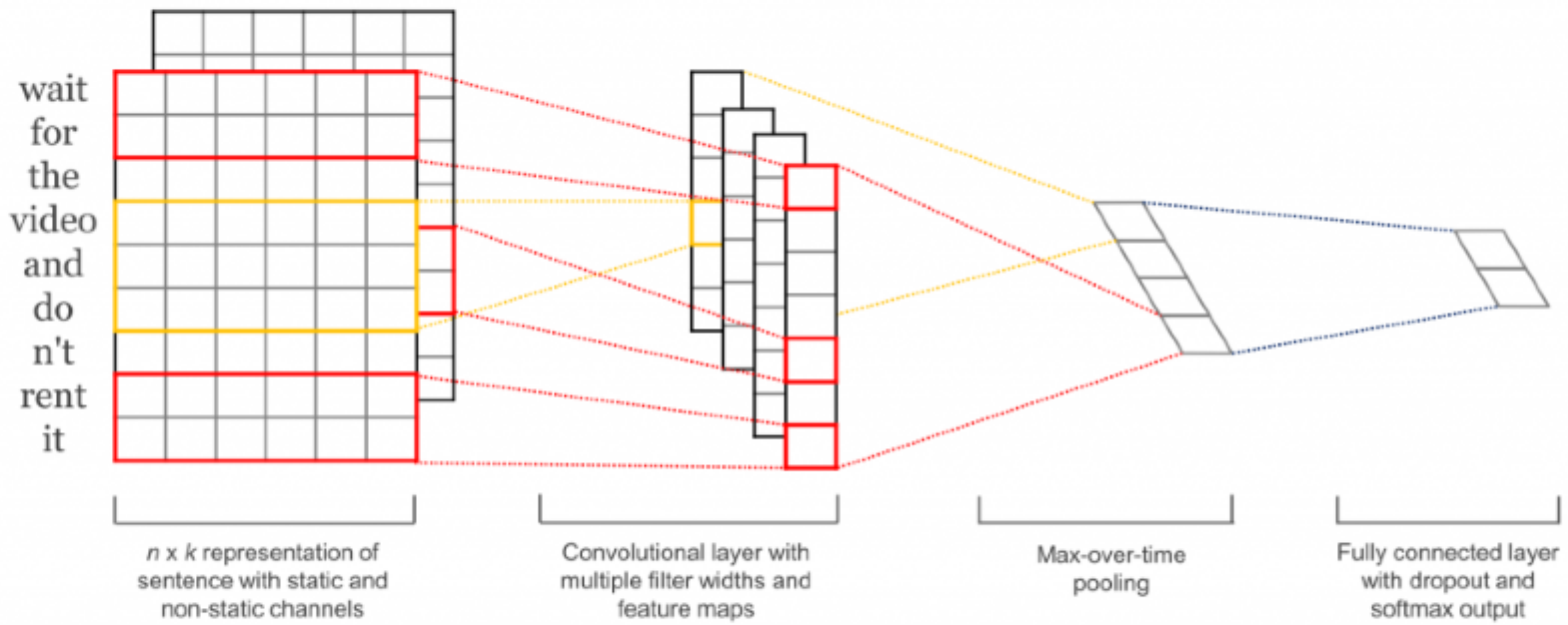
- In NLP, convolutions are performed over the sequence dimension in embedding space, followed by a summation (or similar operation) to collapse the convolved features along the sequence dimension (this is called *pooling)*, yielding a fixed size feature representation.

- Because of this pooling operation, convolutional filtering does not capture long range dependences well.

Convolutional filtering

14

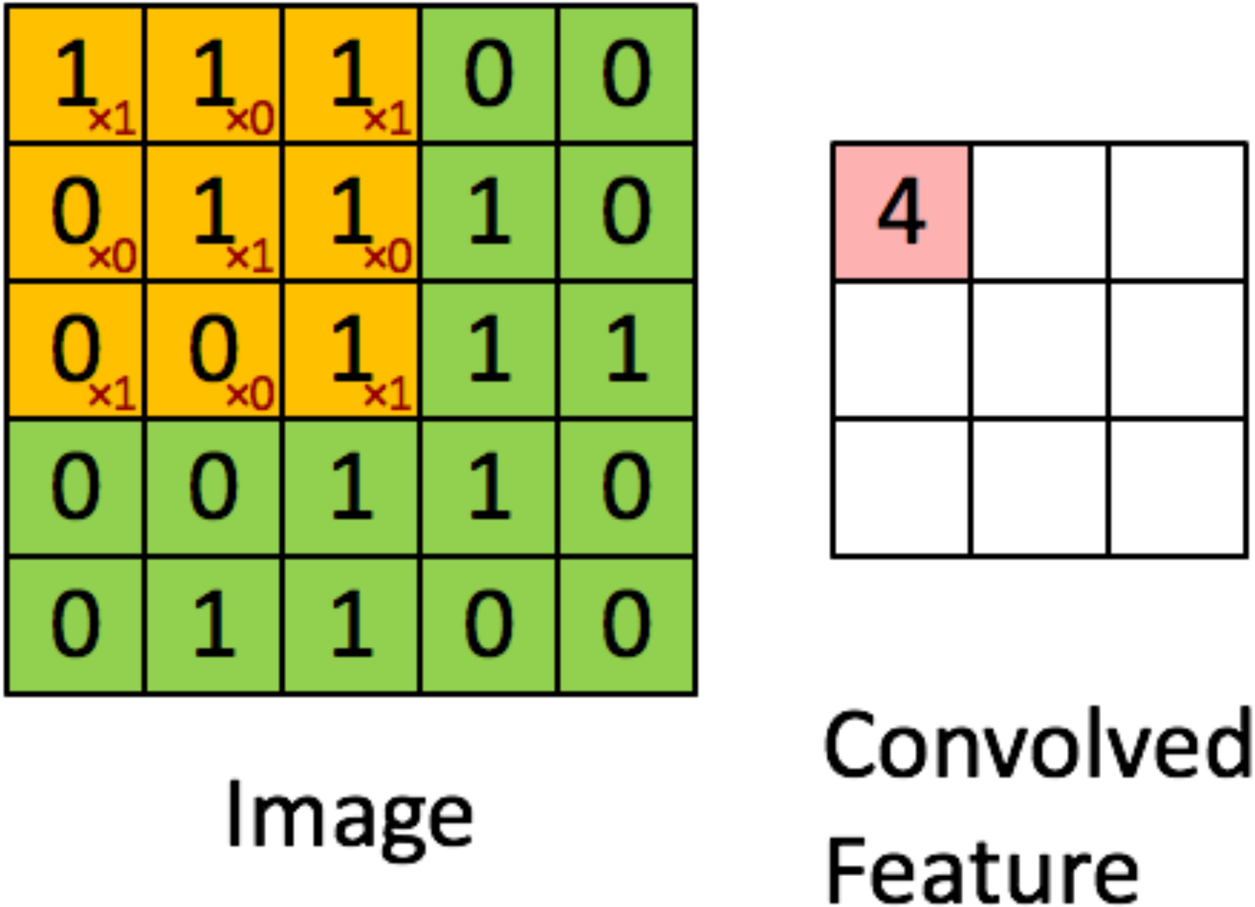

1 ×1 1 ×0 1 ×1 0 0
0 ×0 1 ×1 1 ×0 1 0
0 ×1 0 ×0 1 ×1 1 1
0 0 1 1 0
0 1 1 0 0
4
Image
Convolved Feature


wait
for
the
video
and
do
n't
rent
it
n x k representation of sentence with static and non-static channels
Convolutional layer with multiple filter widths and feature maps
Max-over-time pooling
Fully connected layer with dropout and softmax output

Convolutions applied to images

Convolutions applied to text

Illustrations taken from Stanford UFLDL Wiki


1 ×1 1 ×0 1 ×1 0 0
0 ×0 1 ×1 1 ×0 1 0
0 ×1 0 ×0 1 ×1 1 1
0 0 1 1 0
0 1 1 0 0
4
Image
Convolved Feature

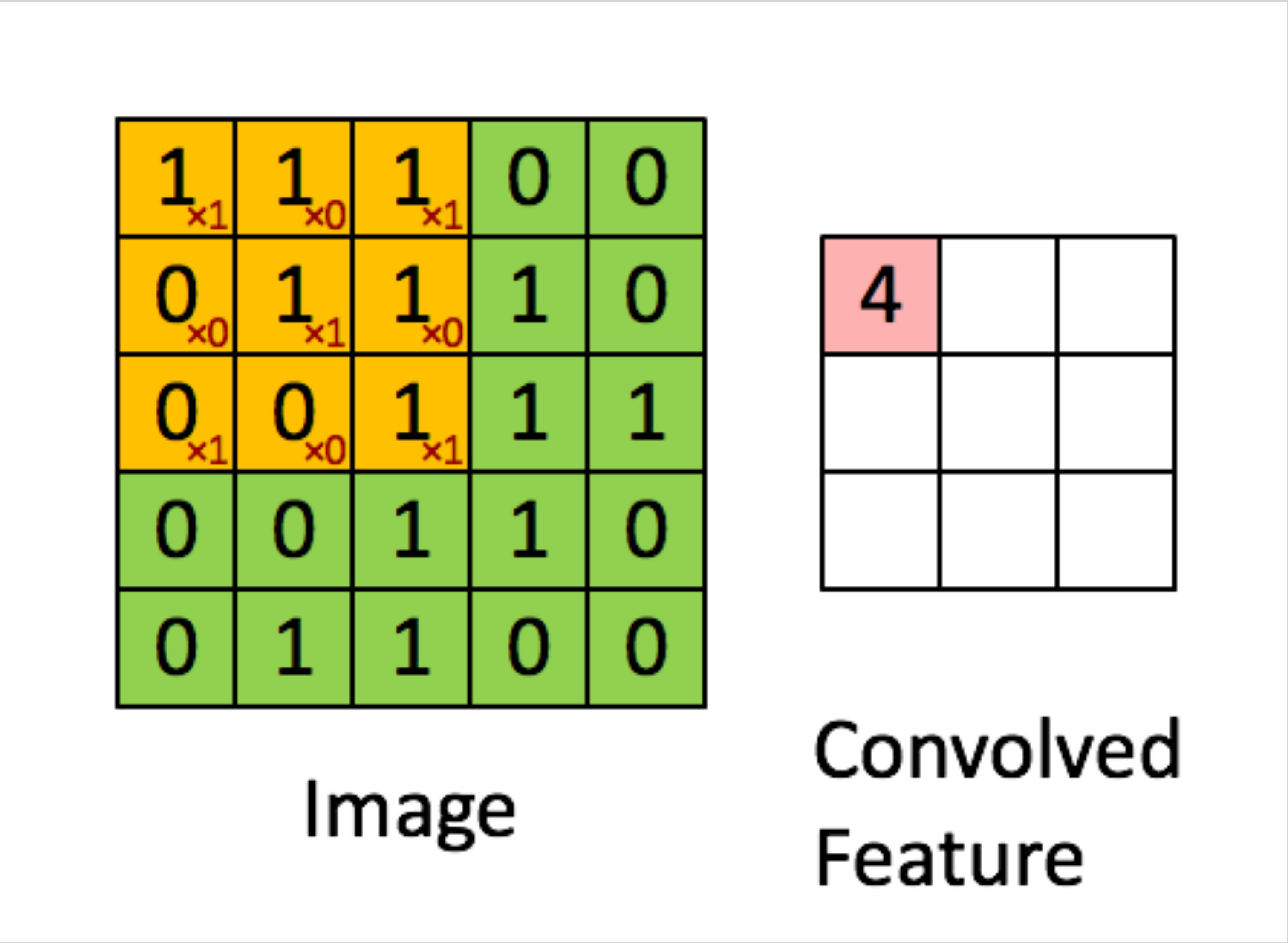
Convolutional filtering

- Convolutional filtering is a general class of techniques used in signal processing for filtering continuous or discrete signals. It's based on the idea of a *finite impulse response filter* (FIR). In deep learning this was first applied to vision, wherein the convolutional filters are 2-dimensional patches (of varying sizes) that mimic the receptive fields in the human visual cortex.

- In NLP, convolutions are performed over the sequence dimension in embedding space, followed by a summation (or similar operation) to collapse the convolved features along the sequence dimension (this is called *pooling)*, yielding a fixed size feature representation.

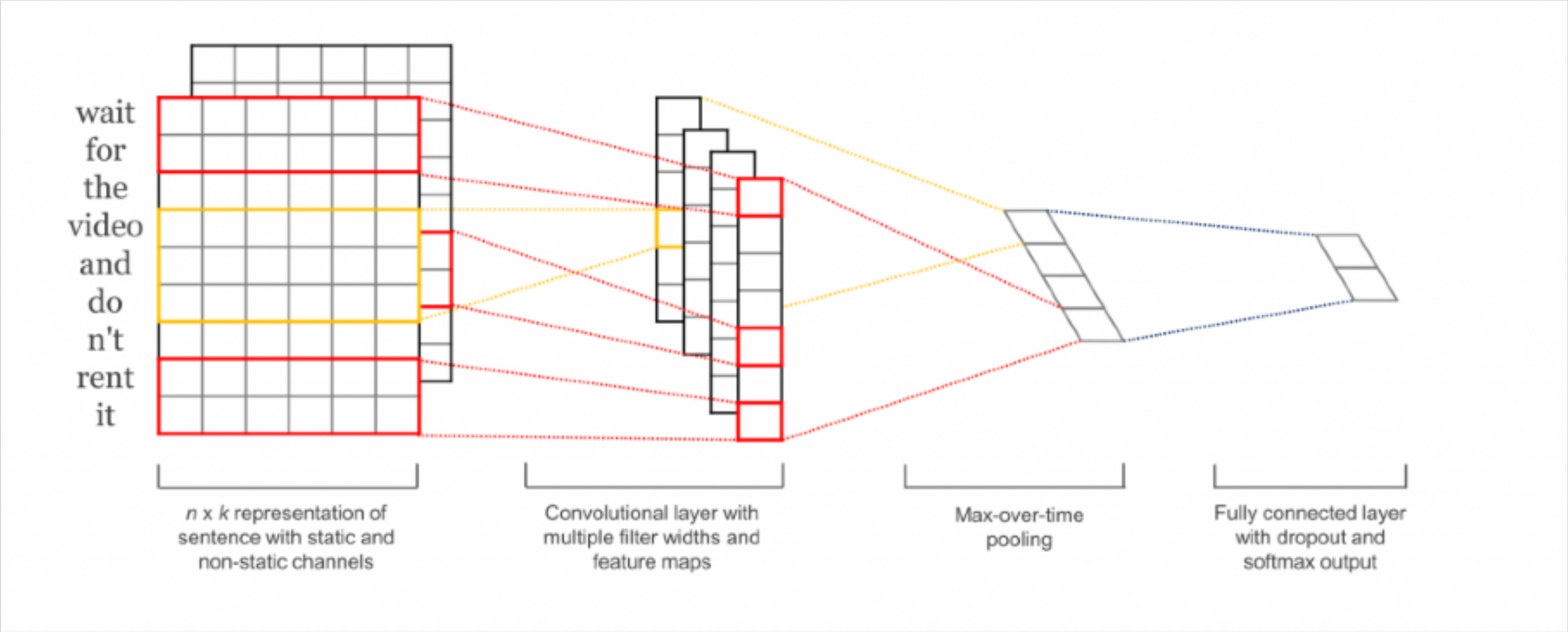
- Because of this pooling operation, convolutional filtering does not capture long range dependences well.

Convolutions applied to images



Illustrations taken from Stanford UFLDL Wiki

Convolutions applied to text



Illustrations taken from Stanford UFLDL Wiki

Recurrent connections

- Recurrent neural networks are based on the idea of an *infinite impulse response filter* (IRR), whereby the feature representation at the t^{th} sequence position, $\mathbf{h}^{(t)}$, is a function of both $\mathbf{x}^{(t)}$ and $\mathbf{h}^{(t-1)}$. This *hidden state* can then be used in a variety of of ways, for example in language modeling a word/token is predicted at each sequence step, whereas for a text classification task, only the feature layer at the last step, $\mathbf{h}^{(T)}$, is used to predict the output.

- In theory recurrent connections enable us to maximally capture context. In practice, training these networks becomes increasingly difficult for long sequences due to a phenomenon called vanishing gradients.

- The popular Long-Short Term Memory (LSTM) cell block is based on this idea!