

Machine Learning - *Temporal* models

(69152) **RNNs, Transformers, ...**

Master in Robotics, Graphics and Computer Vision

Ana C. Murillo



Today

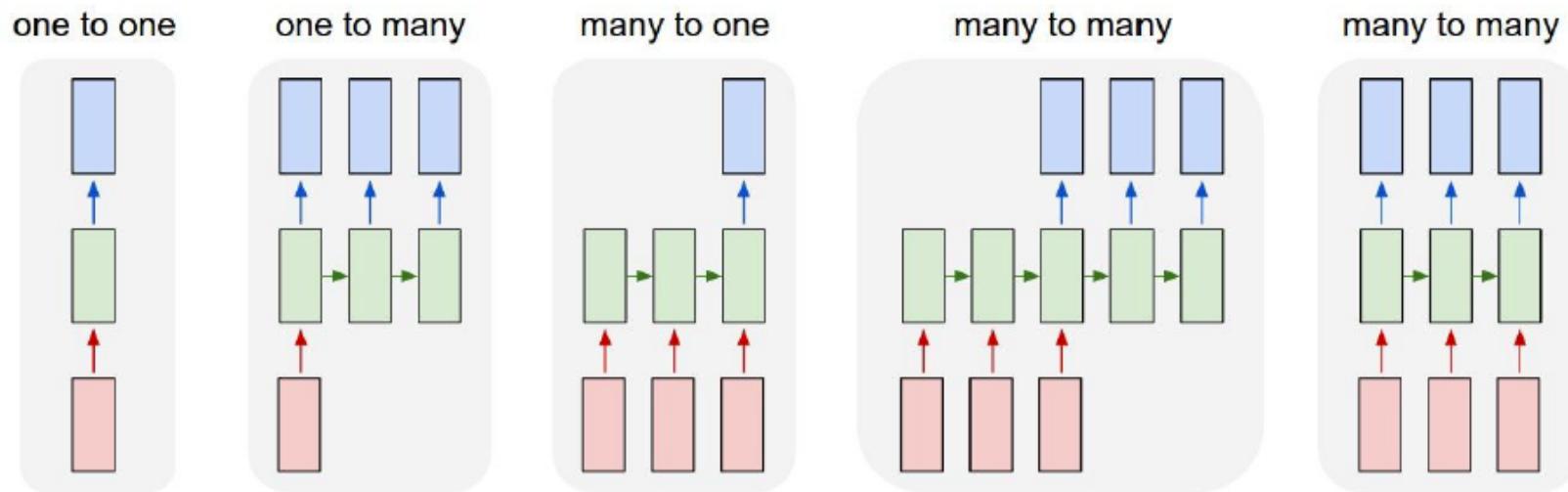
- RecursiveNN
- Resources:
 - <http://cs231n.stanford.edu>
 - <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
—> toy examples of RNN in python
 - Deep Learning book. Chapter 10.
www.deeplearningbook.org

Deep Learning & ...

- Time? Sequential Data?

Deep Learning & Sequential data

- Video and Speech/Text, etc? → **Recurrent NN**



Fei-Fei Li & Justin Johnson & Serena Yeung
<http://cs231n.stanford.edu>

- can handle **non-fixed length input/output**

Recurrent Neural Networks

- repeat (as many input/output/steps as needed)
- **share the weights** across iterations

one to one

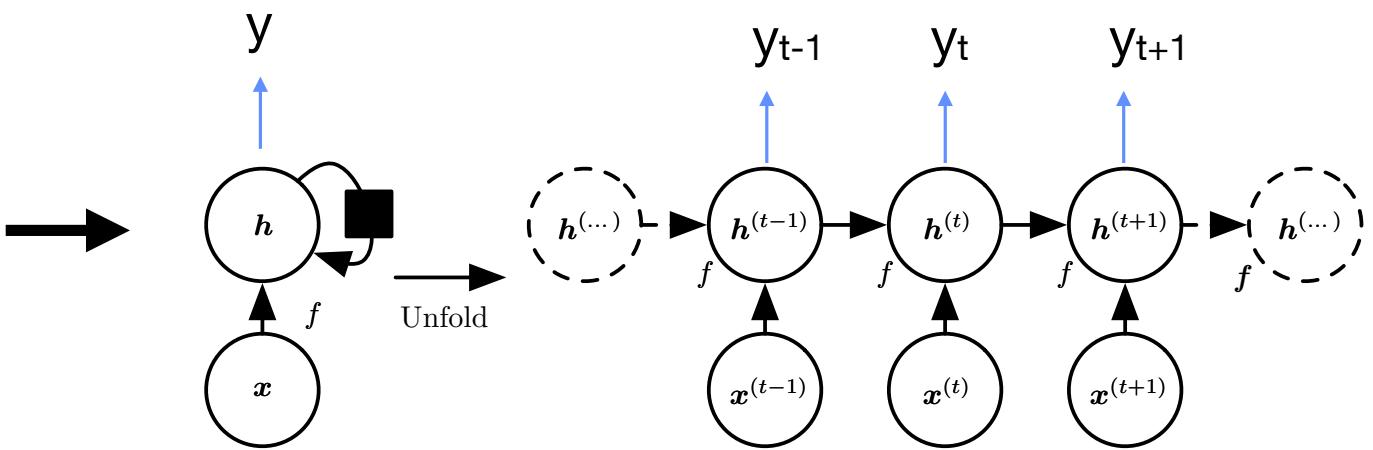
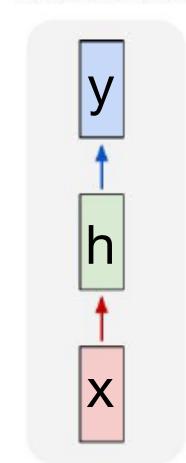


Figure 10.2 - Deep Learning book. Ian Goodfellow

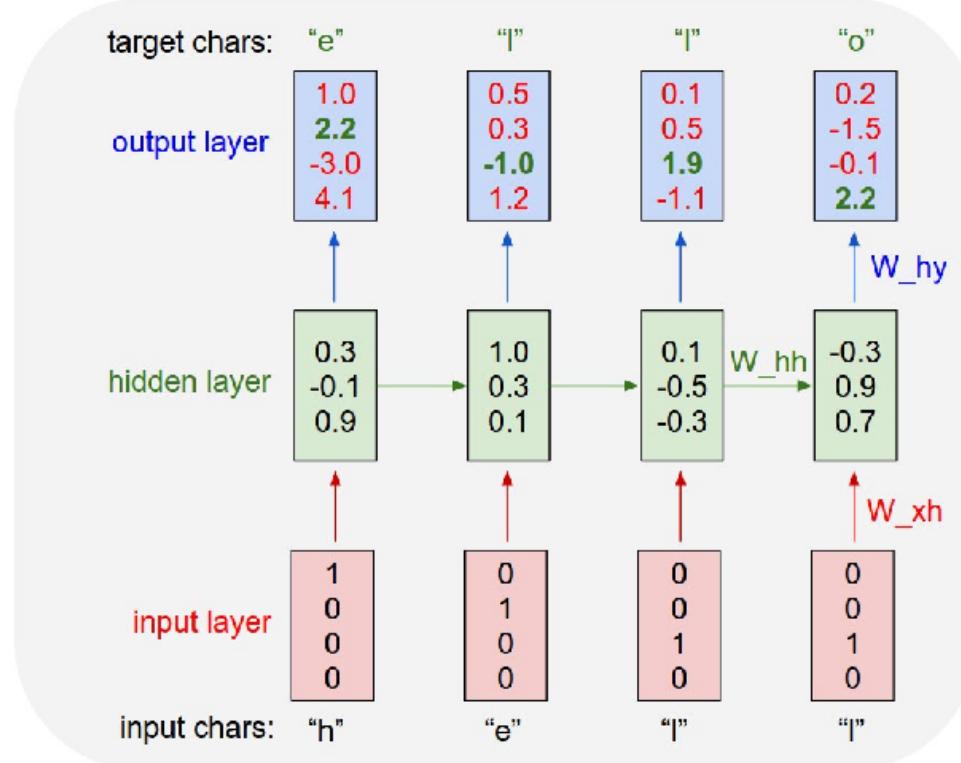
Recurrent Neural Networks

Main ingredients on a toy example:

- time: t
- hidden layer (s): h
- “vocabulary”: {h,e,l,o}
- weights: \mathbf{W}_{hh} , \mathbf{W}_{hh} , \mathbf{W}_{hy}

$$h_t = \tanh(\mathbf{W}_{\text{hh}}h_{t-1} + \mathbf{W}_{\text{xh}}x_t)$$

$$y_t = \mathbf{W}_{\text{hy}}h_t$$



Andrej Karpathy
<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

Recurrent Neural Networks

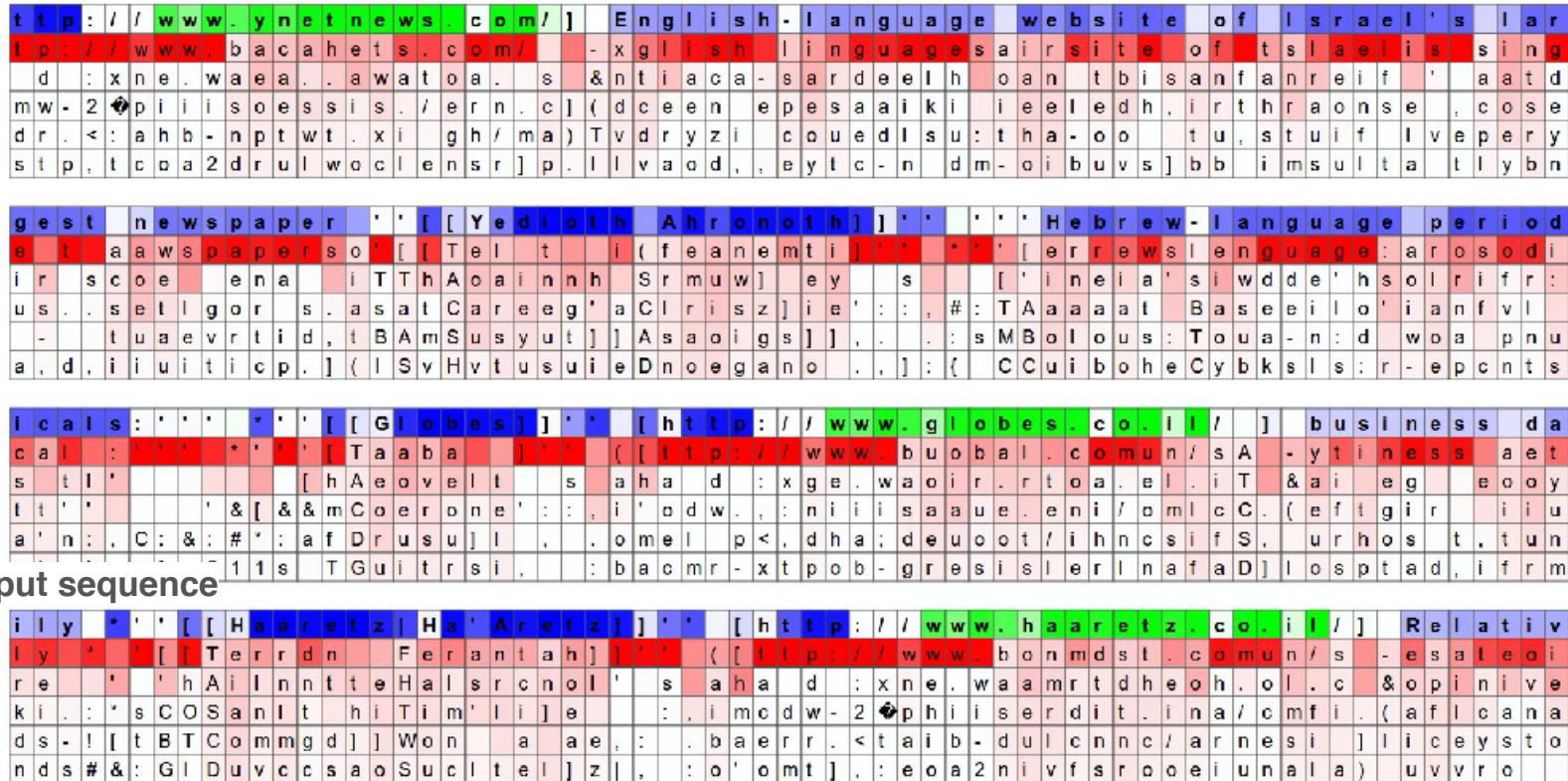
Interesting examples from a *vainilla* version

- text generator (Shakespeare, Wikipedia page, ...)
- code generator! (LaTex, C, ...)

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

Recurrent Neural Networks

Neuron that highlights URLs



<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

Recurrent Neural Networks

Neuron that highlights URLs

input sequence

t	t	p	:	/	/	w	w	w	.	y	n	e	t	n	e	w	s	.	c	o	m	/]	E	
t	p	:	/	/	w	w	w	.	b	a	c	a	h	e	t	s	.	c	o	m	/	-	x		
d	:	x	n	e	.	w	a	e	a	.	.	a	w	a	t	o	a	.	s	&	n				
m	w	-	2	?	p	i	i	i	s	o	e	s	s	i	s	.	/	e	r	n	.	c]	(d
d	r	.	<	:	a	h	b	-	n	p	t	w	t	.	x	i	g	h	/	m	a)	T	v	
s	t	p	,	t	c	o	a	2	d	r	u	l	w	o	c	l	e	n	s	r]	p	.	l	l

g	e	s	t		n	e	w	s	p	a	p	e	r	'	'	[[Y	e	d	i	o	th		
e	t	a	a	w	s	p	a	p	e	r	s	o	'	[[T	e	l	t	i	i				
i	r		s	c	o	e		e	n	a		i	T	T	h	A	o	a	i	n	n	h			
u	s	.	.	s	e	t	l	g	o	r	s	.	a	s	a	t	C	a	r	e	e	g	'	a	
-		t	u	a	e	v	r	t	i	d	,	t	B	A	m	S	u	s	y	u	t]]		
a	,	d	,	i	i	u	i	t	i	c	p	.]	(I	S	v	H	v	t	u	s	u	i	e

i	c	a	l	s	:	'	'	'	*	'	'	[[G	l	o	b	e	s]]	'	'	
c	a	l	:	'	'	'	*	'	'	'	[T	a	a	b	a]	'	'	(
c	a	l	,	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	l	

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

Recurrent Neural Networks

Neuron that highlights markdown env [[]]

*	*	[[Jerusalem Report]]	*	[http://www.jrep.com/]	Left-of-center En
*	[hToausal m aogurt]	*	(http://www.bsinioom/)	-iat af tenter (ng	
[' [Cassmene] Beaonds	s a[a d : xne. waaaoca. s &ato - nf hlsu m - ouc			
' s mFurnl s i aet allsa ' :	i' c dw- 2 t p i i soeg . er / . a] (oseswr - c iddr s [mt				
: * : Aq Denebiutn Cipre e . b1emr . 9 : ahb - npumughnmp) Teiretu : eoseodsald					
# T & Tf Si wrp e] aluve lru s : - mp rts < moa 2 deyshilr] c. Augl , 1p , larc : fae					
glish [[weekly newspaper]]	*	[[YNetNews]]	*	[http://www.ynetnews.c	
lish c[Caakly] cawspaper]	*	[hTaA at]	*	(http://www.bacanets.co	
iaci - lhSoip] i sec] enp] s . ' ' [Co * wess]	s a[a d : xne. wasa . awat o a				
ee na . pCci etnedlox] gicil s' [sAmFeSahon] t' :	, imomw - 2 t p i i soessis . / er				
syz . sf penn al ruel lrra . ' # ^ : oDuFreiuep . : b1edr . < ahb - nptwt . xi gh					
a dpeamAr bde orp itee] dts - T{ [BaAvTposwao , . . oacstp , tcoa 2 drulwoclens					
om/] English-language website of Israel's largest newspaper ' ' [[Yed					
m/] -xglish linguagesairsite of tslaelliss singe taawspaperso [[Tel					
. s &ntiaca - sardeelh can t bisan fanreif ' atdir score ena iTTAoai					
n. c] (dceen epesaaiiki ieledh , irthraonse , coseus . setlgor s . asat Care					
/ ma) Tvdryzi couedlsu : tha - oo tu , stuif lvep ery - tuaevrtid , tBAmSusy					
r] p . llvadd , , eytc - n dm - oibuv s] bb imsulta tlybna , d , iiuiticp -] (ISvHvtu					
loth Ahronoth] * * * Hebrew-language periodicals: * * * [[Globes]]					
t i(feanemt il) * * * [errewsle n gue : arosodical : * * * * * [Taaba]]					
nnh Srmuw] ey s ['ineia ' si wdd e ' hsolriffr : stl ']					
eg ' acIrisz] ie' : : , # : TAAAat Baseeil o 'ianfvl tt ']					
ut] Asaoigs] , . . : sMBo lous : Toua - n : d woa pnua ' n : , C : & : # : afDrusu] ,					
suiedDnoegano . .] : { CCui bohe Cybksls : r - epontsnk i <] : & 11s TGuitr si ,					

<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>

Recurrent Neural Networks

But ...

RNN vainilla version does not work that well in practice

Recurrent Neural Networks

- **LSTM networks** - Long short-term memory cells

- Very frequent type of RNN
- more sophisticated hidden function than basic RNN
- capable of learning long-term dependencies

$$h_t = \tanh \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right) \quad \rightarrow \quad \begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$
$$c_t = f \odot c_{t-1} + i \odot g$$
$$h_t = o \odot \tanh(c_t)$$

Hochreiter & Schmidhuber. "Long Short Term Memory". *Neural Computation*, 1997

Recurrent Neural Networks

- **LSTM networks** - Long short-term memory cells

- Very frequent type of RNN
- more sophisticated hidden function than basic RNN
- capable of learning long-term dependencies

$$h_t = \tanh \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right)$$

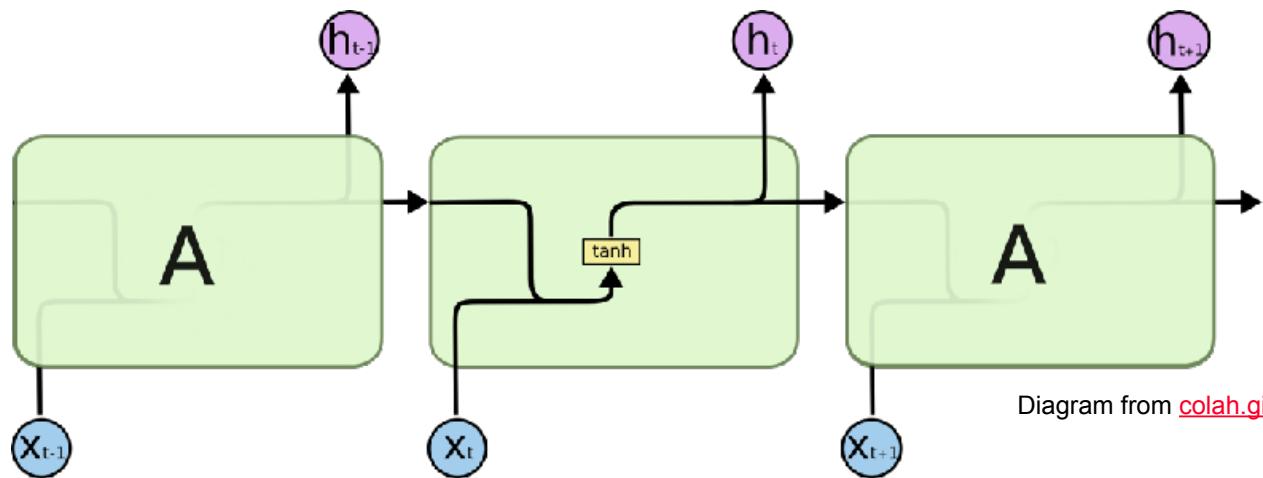


Diagram from colah.github.io

Recurrent Neural Networks

- **LSTM networks** - Long short-term memory cells
 - Very frequent type of RNN
 - more sophisticated hidden function than basic RNN
 - capable of learning long-term dependencies

$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$
$$c_t = f \odot c_{t-1} + i \odot g$$
$$h_t = o \odot \tanh(c_t)$$

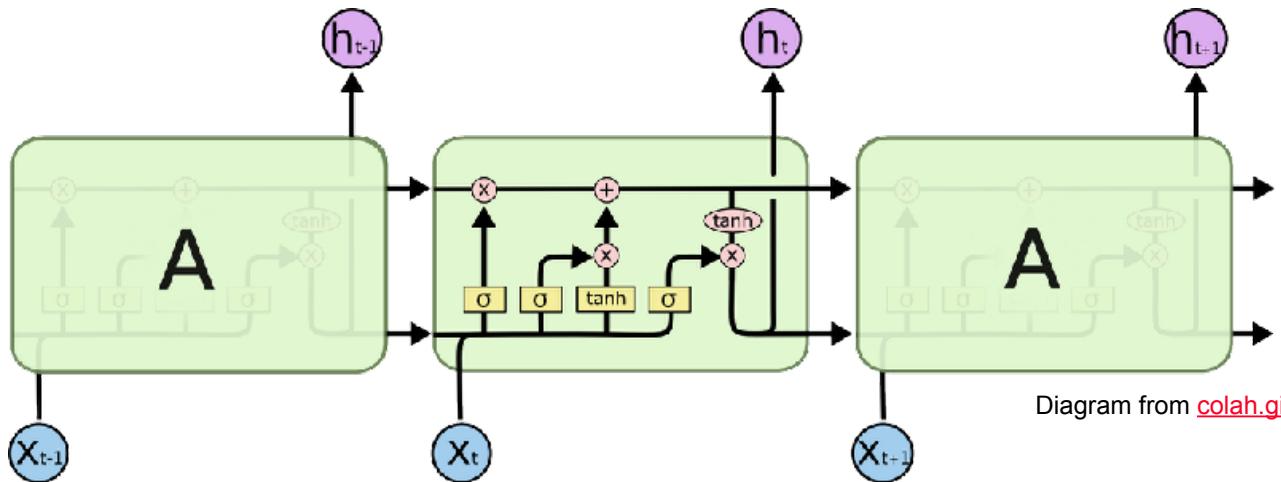


Diagram from colah.github.io

Recurrent Neural Networks

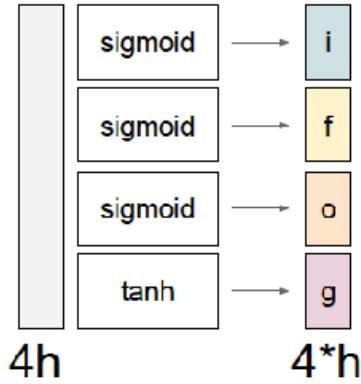
- **LSTM networks** - Long short-term memory cells

- **f: forget gate.**
Erase cell or not?

- **i: input gate.**
Write cell or not?

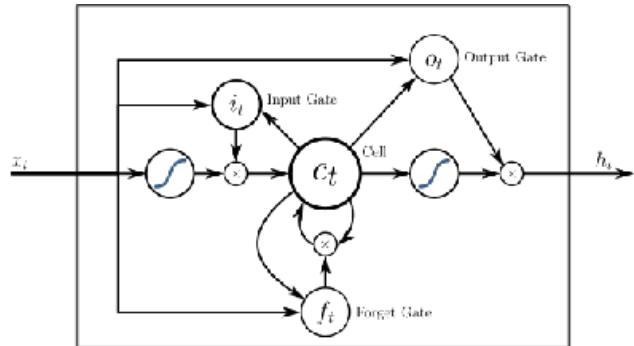
- **g: Cell Gate.**
How much to write to cell

- **o: Output gate.**
How much to reveal cell value



Fei-Fei Li & Justin Johnson & Serena Yeung
<http://cs231n.stanford.edu>

$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$
$$c_t = f \odot c_{t-1} + i \odot g$$
$$h_t = o \odot \tanh(c_t)$$

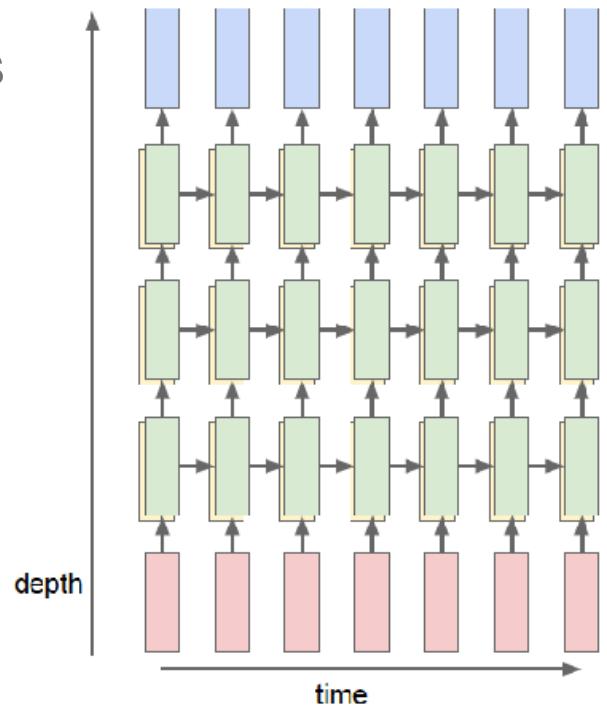


Wikipedia

Recurrent Neural Networks

- Many successful examples and variations:

- Many variations of LSTM cells
- Multi-layer RNN



Fei-Fei Li & Justin Johnson & Serena Yeung

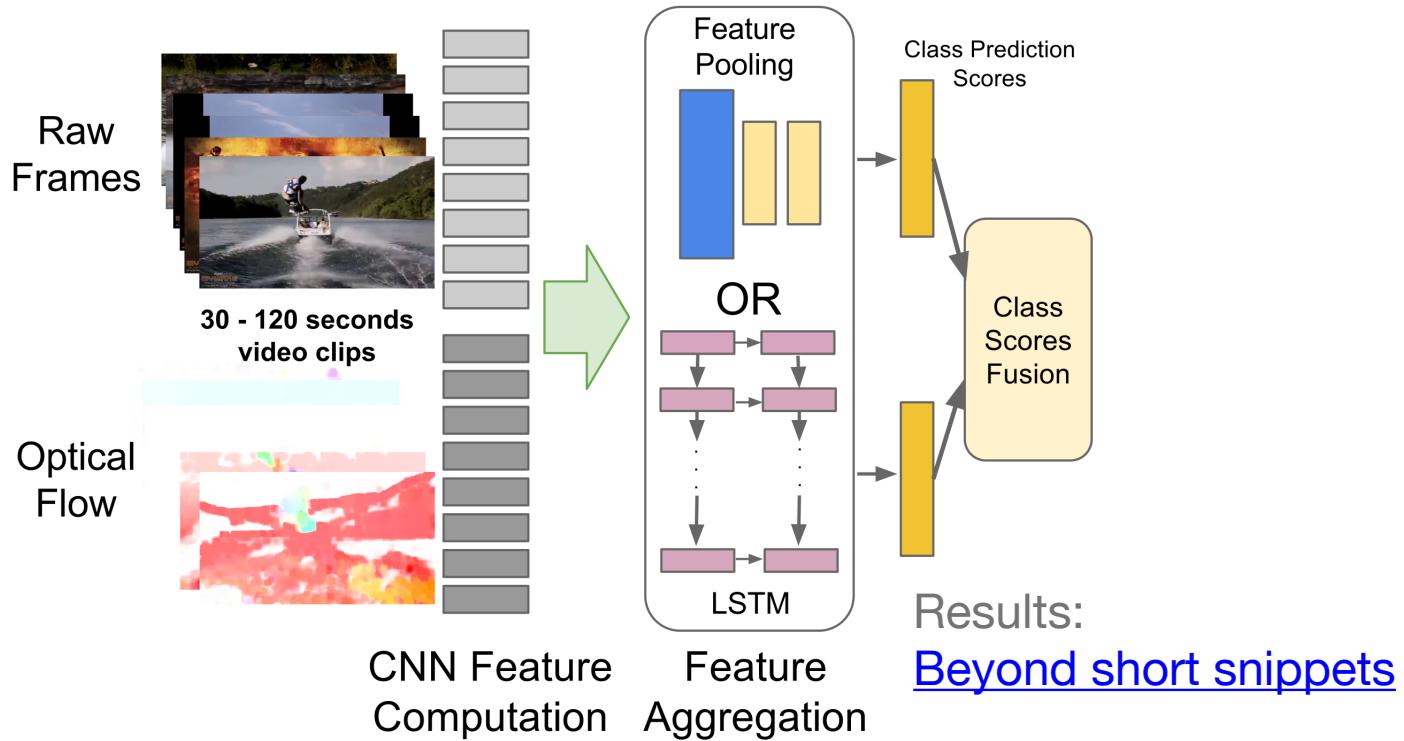
<http://cs231n.stanford.edu>

Recurrent Neural Networks

- Besides “pure” RNN examples
- plenty of combinations of RNN with others

Recursive NN examples

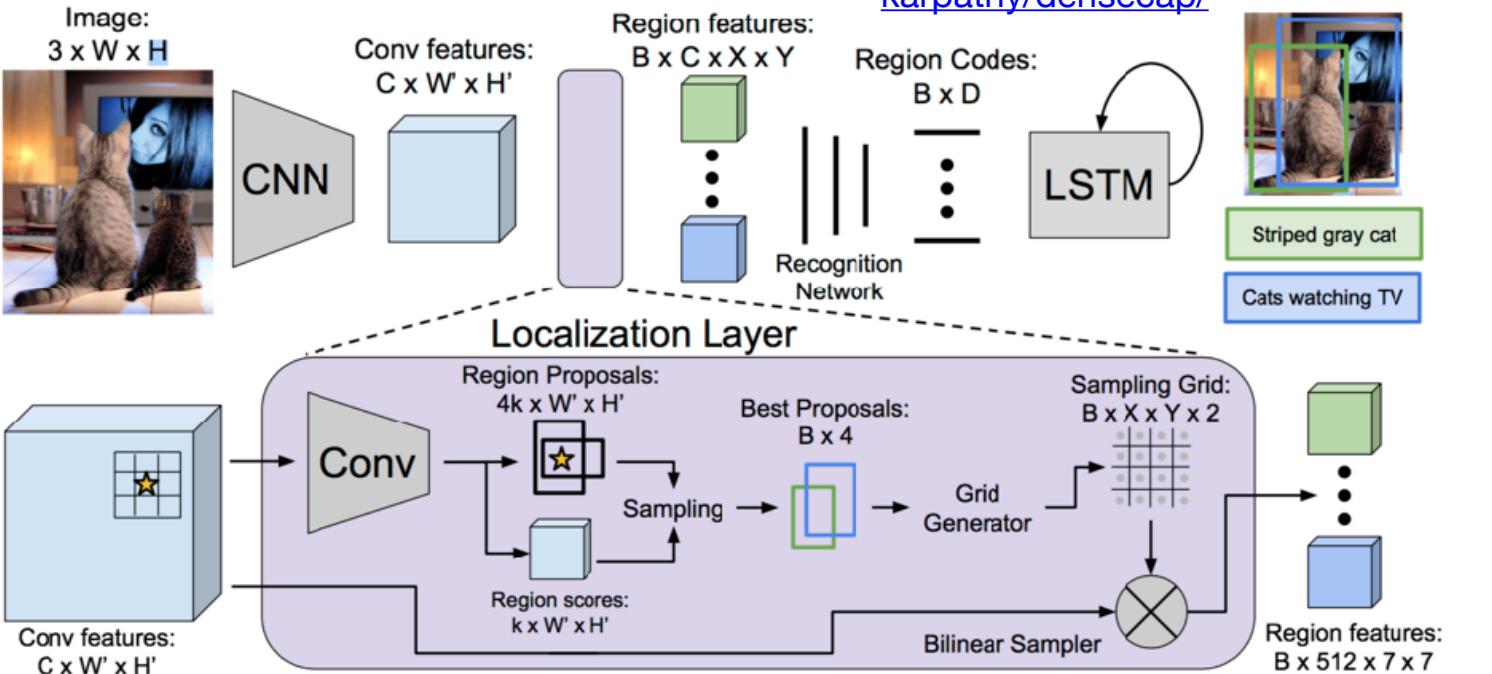
- **Action Classification.** Video and Flow (temporal data)



Beyond Short Snippets: Deep Networks for Video Classification.
Joe Yue-Hei Ng et al. Google. CVPR2015.

Recursive NN examples

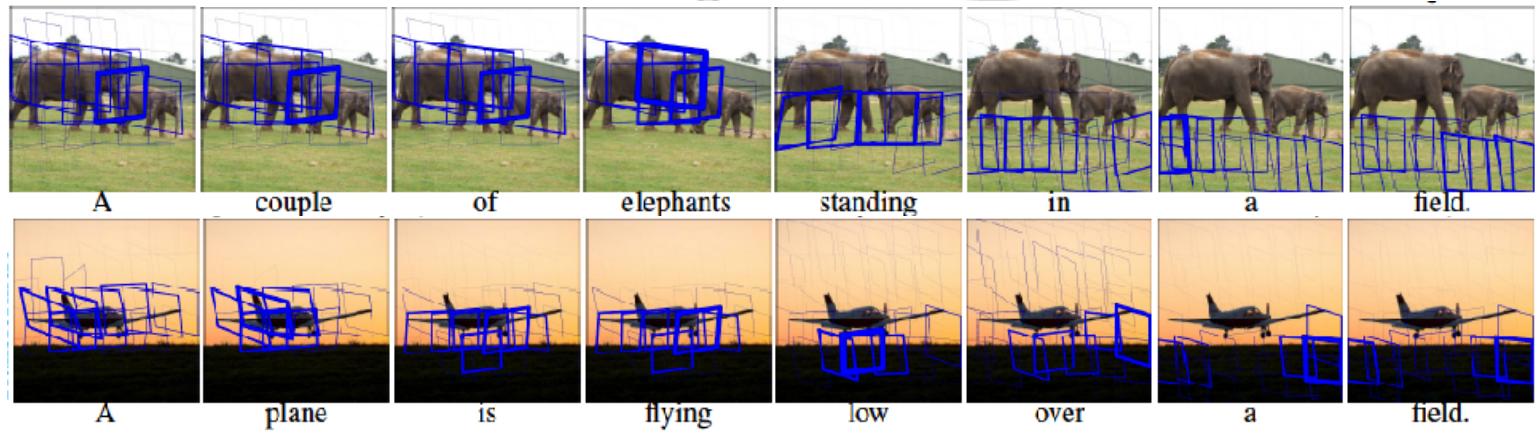
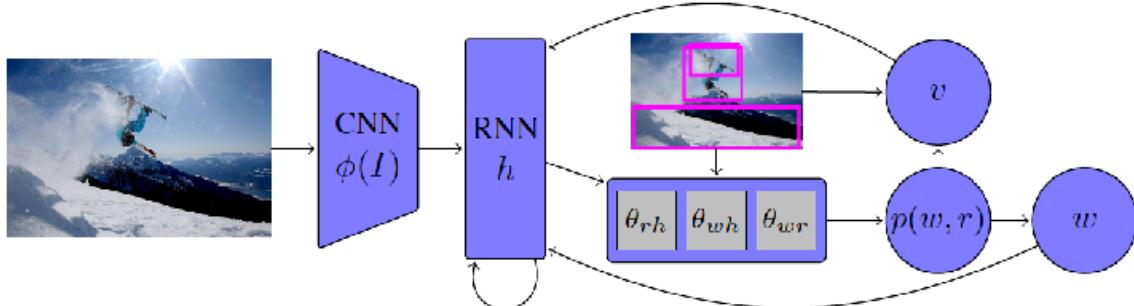
- **Multimodal problems.** Image and Text. Dense Captions (other types of sequential data)



DenseCap: Fully Convolutional Localization Networks for Dense Captioning.
Johnson, Justin and Karpathy, Andrej and Fei-Fei, Li. CVPR 2016

Recursive NN examples

- **Multimodal problems.** Image and Text. Captioning & Attention (other types of sequential data)



Areas of Attention for Image Captioning. Marco Pedersoli, Thomas Lucas, Cordelia Schmid, Jakob Verbeek; ICCV 2017

RNN & CuDNN example

Exercise in class: we will implement a simple RNN.

Build a text file of similar size to this one (200K words):

[https://storage.googleapis.com/download.tensorflow.org/
data/shakespeare.txt](https://storage.googleapis.com/download.tensorflow.org/data/shakespeare.txt)

of certain type of text/speech

e.g., speech from someone; lyrics of certain type;
programming language; ... some more examples here:
<http://karpathy.github.io/2015/05/21/rnn-effectiveness/>)

RNN & CuDNN example/exercise

- Use your own text file of similar size to this one (200K words):
<https://storage.googleapis.com/download.tensorflow.org/data/shakespeare.txt> of certain type of text/speech (e.g., speech from someone; lyrics of certain type; programming language; ... some more examples here: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>)
- **Run Colab example** to implement a simple RNN using your text style data:
https://colab.research.google.com/drive/1Qjcip3iErC_UyKEL62DwO_S2GBWKrxMo
(extracted from https://www.tensorflow.org/tutorials/text/text_generation)
 1. Read and understand the current example
 2. Re-run with your own data

Next

- RNN exercise
- **QUESTIONNAIRES** - [https://encuestas.unizar.es/
english-questionnaires](https://encuestas.unizar.es/english-questionnaires)
- Other *temporal* models
 - TCN
 - Transformers