Applied Machine Learning with Big Data "EE 6973"

Topic:
Recurrent Neural Network
(RNN)

Paul Rad, Ph.D.

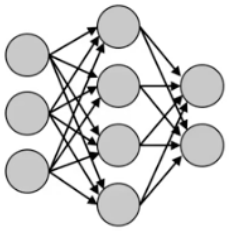
Chief Research Officer
UTSA Open Cloud Institute(OCI)
University of Texas at San Antonio

Outline

Sequences

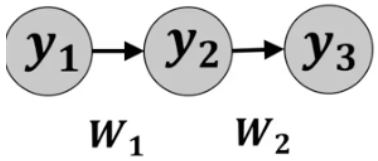
Recurrent Neural Network (RNN)

Deep Neural Network

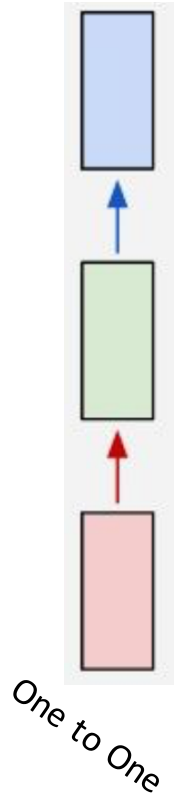


$$y_i = g\left(\sum_j W_{ij} x_j + b_i\right)$$

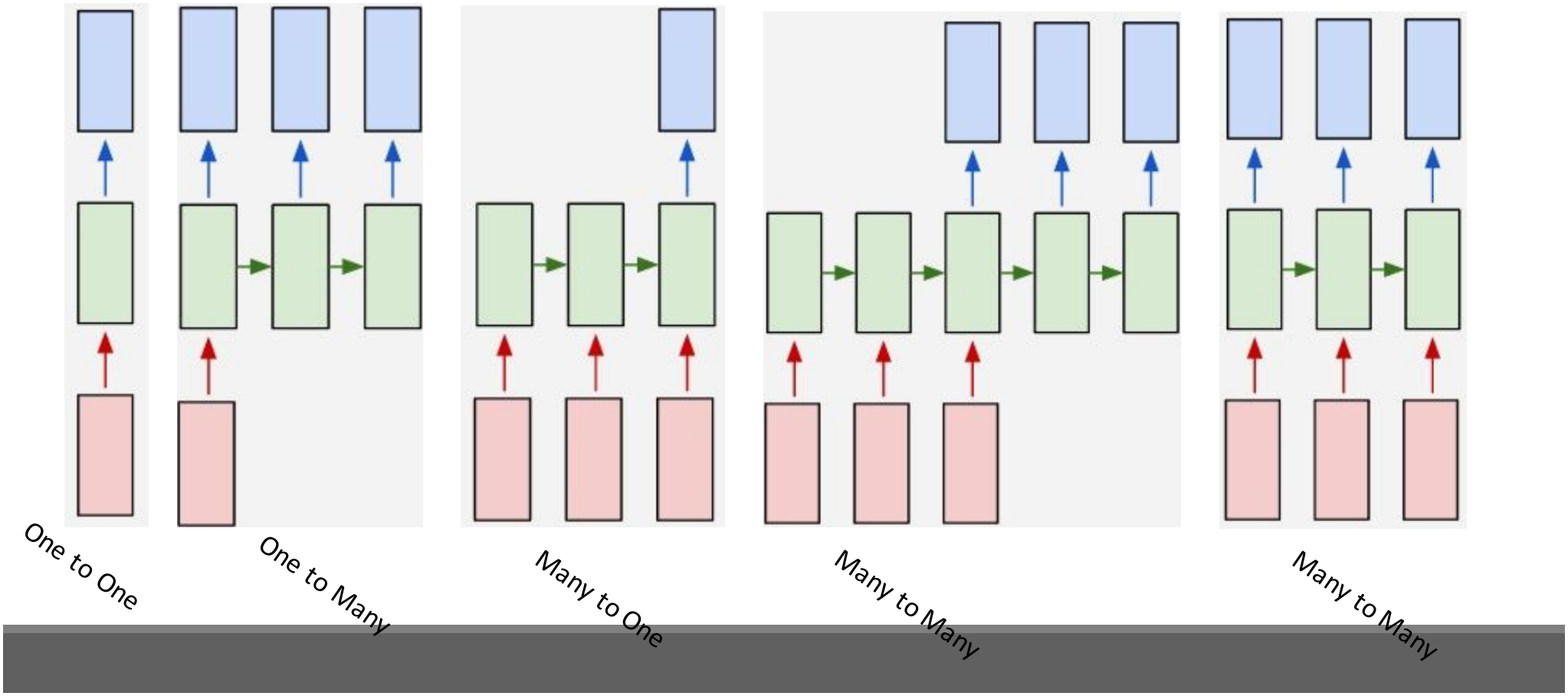
$$\frac{\partial C}{\partial W} = \frac{\partial C}{\partial g} \cdot \frac{\partial g}{\partial a} \cdot \frac{\partial a}{\partial W}$$



$$\boldsymbol{y}_k = g(W\boldsymbol{y}_{k-1} + \boldsymbol{b})$$



NN and RNN Models



Example

One to One = Vanilla Neural Networks

One to Many = Image Captioning

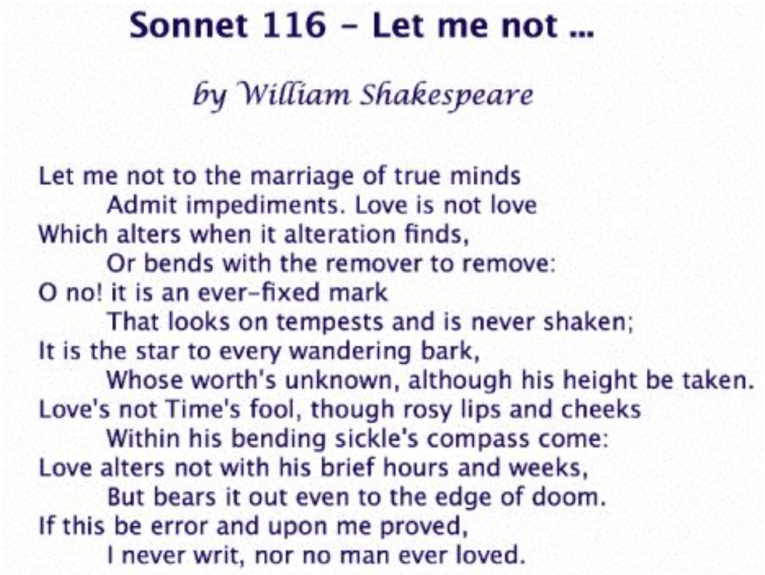
Many to One = Sentiment Classification

Many to Many = Machine Translation

Many to Many =Video Classification on Frame Level

"man in black shirt is playing guitar."

RNN Sample



Sonnet 116 – Let me not …

by William Shakespeare

Let me not to the marriage of true minds
Admit impediments. Love is not love
Which alters when it alteration finds,
Or bends with the remover to remove:
O no! it is an ever-fixed mark
That looks on tempests and is never shaken;
It is the star to every wandering bark,
Whose worth's unknown, although his height be taken.
Love's not Time's fool, though rosy lips and cheeks
Within his bending sickle's compass come:
Love alters not with his brief hours and weeks,
But bears it out even to the edge of doom.
If this be error and upon me proved,
I never writ, nor no man ever loved.

```
tyntd-iafhatawiaoihrdemot  lytdws  e ,tfti, astai f ogoh eoase rrranbyne 'nhthnee e
plia tklrgd t o idoe ns,smtt   h ne etie h,hregtrs nigtike,aoaenns lng
```

train more

```
"Tmont thithey" fomesscerliund
Keushey. Thom here
sheulke, anmerenith ol sivh I lalterthend Bleipile shuwy fil on aseterlome
coaniogennc Phe lism thond hon at. MeiDimorotion in ther thize."
```

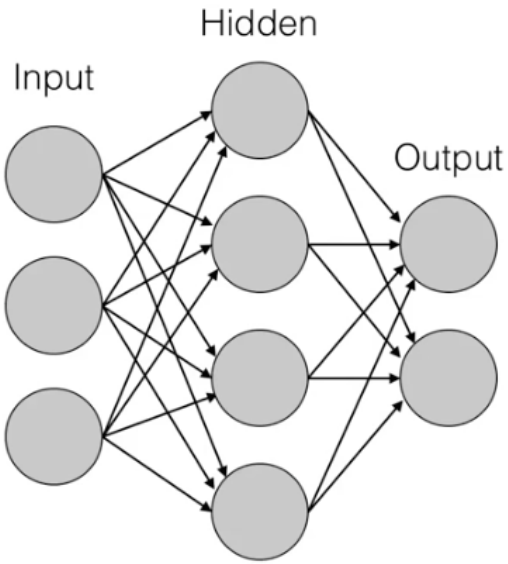
train more

```
Aftair fall unsuch that the hall for Prince Velzonski's that me of
her hearly, and behs to so arwage fiving were to it beloge, pavu say falling misfort
how, and Gogition is so overelical and ofter.
```

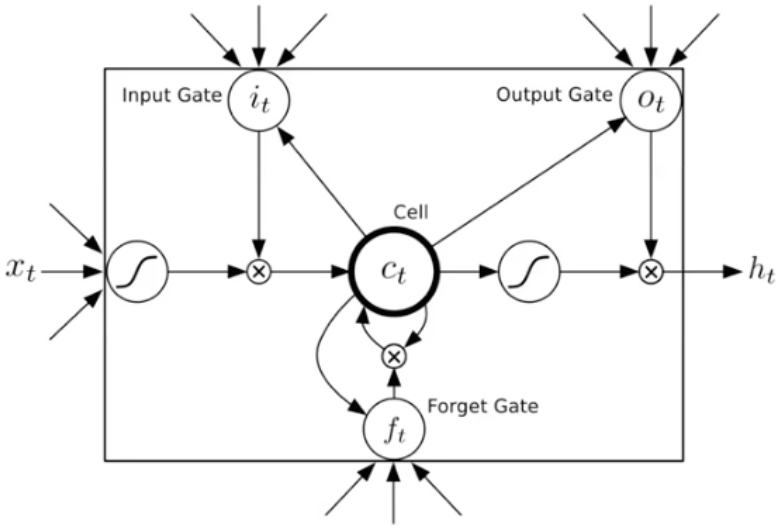
train more

```
"Why do what that day," replied Natasha, and wishing to himself the fact the
princess, Princess Mary was easier, fed in had oftened him.
Pierre aking his soul came to the packs and drove up his father-in-law women.
```

Why Recurrent Network

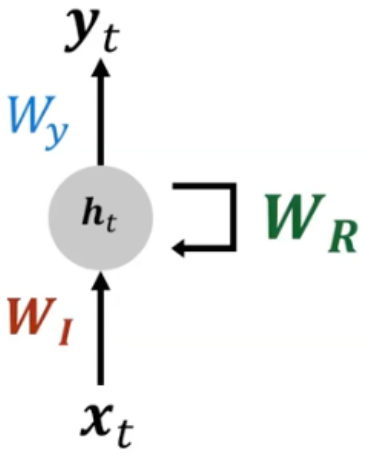


- Independence
- Fixed Length



- Temporal dependencies
- Variable sequence length

Recurrent Neural Network



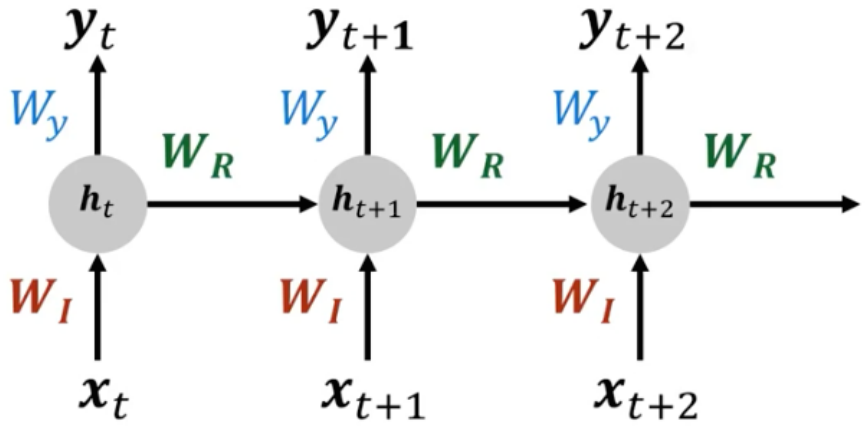
$$\boldsymbol{h}^{(t)} = g_h(W_{\mathrm{I}}\boldsymbol{x}^{(t)} + W_{\mathrm{R}}\boldsymbol{h}^{(t-1)} + \boldsymbol{b}_h)$$

$$\boldsymbol{y}^{(t)} = g_y(W_y\boldsymbol{h}^{(t)} + \boldsymbol{b}_y)$$

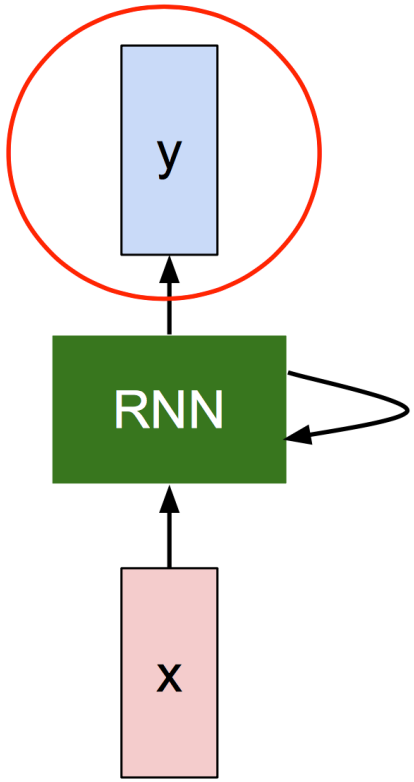
Unrolling a recurrent network into a feed forward network

$$\boldsymbol{h}^{(t)} = g_h(W_{\mathrm{I}}\boldsymbol{x}^{(t)} + W_{\mathrm{R}}\boldsymbol{h}^{(t-1)} + \boldsymbol{b}_h)$$

$$\boldsymbol{y}^{(t)} = g_y(W_y\boldsymbol{h}^{(t)} + \boldsymbol{b}_y)$$



Recurrent Neural Network



Recurrent Neural Network

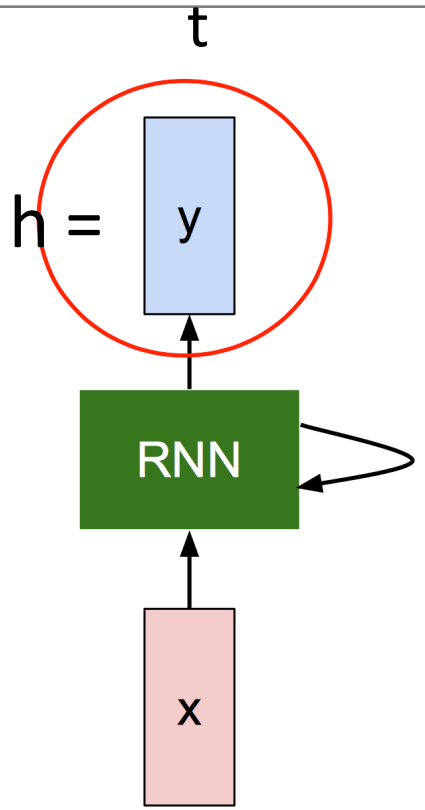
$$h_t = f_W(h_{t-1}, x_t)$$

new state

some function with parameters W

old state

input vector at some time step

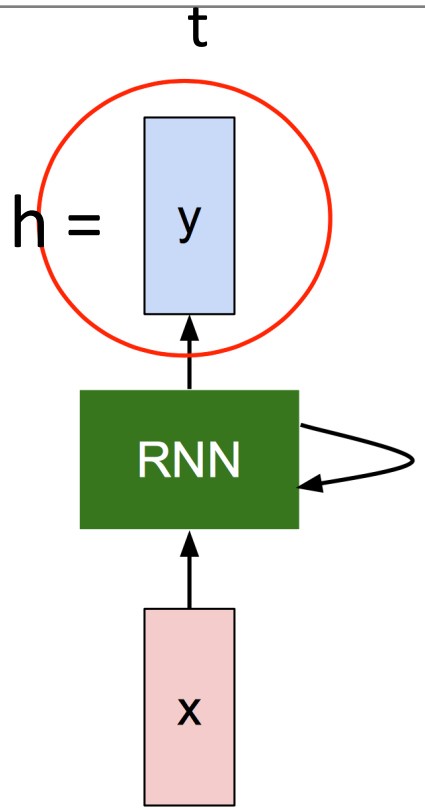


Vanilla Recurrent Neural Network

$$h_t = f_W(h_{t-1}, x_t)$$

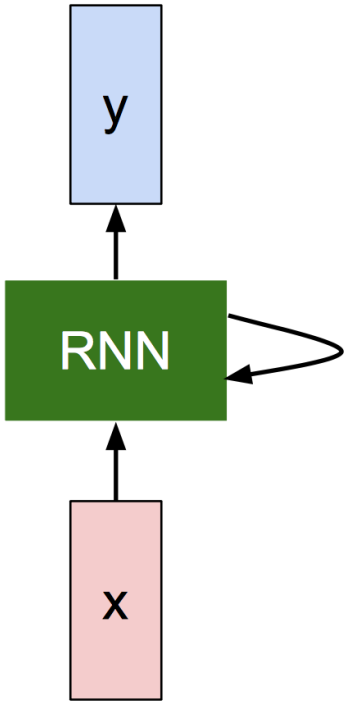
$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

$$y_t = W_{hy}h_t$$



$$h_t = f_W(h_{t-1}, x_t)$$

new state — some function with parameters W — old state — input vector at some time step



$$h_t = f_W(h_{t-1}, x_t)$$

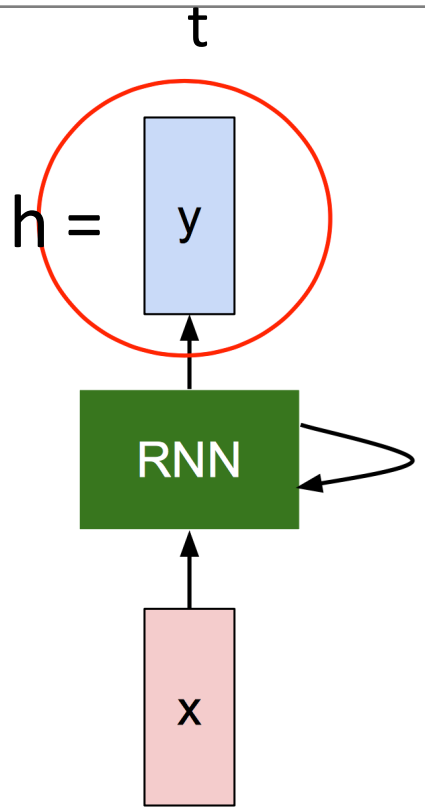
$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t)$$

$$y_t = W_{hy}h_t$$

Character-level language model example

Vocabulary:
[h,e,l,o]

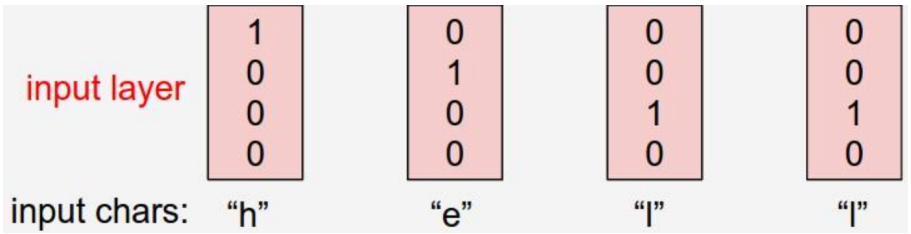
Example training sequence:
“hello”

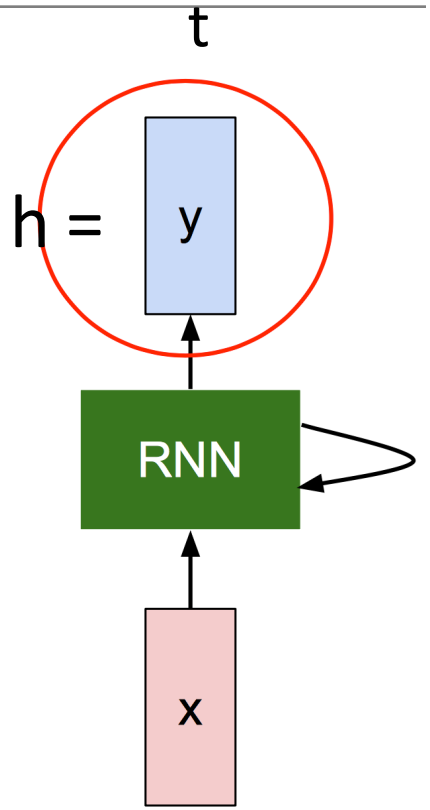


Character-level language model example

Vocabulary:
[h,e,l,o]

Example training sequence:
“hello”





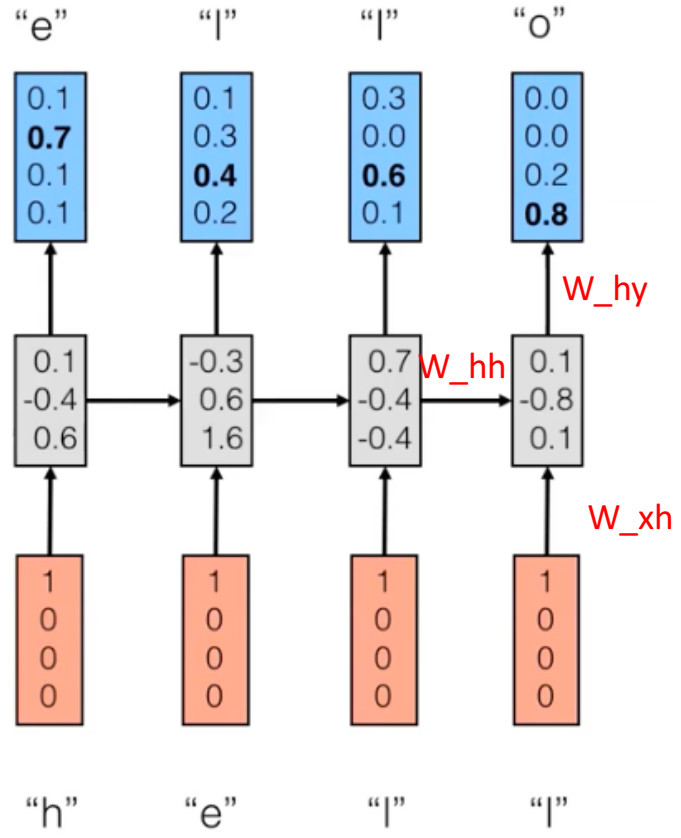
$$h_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$$

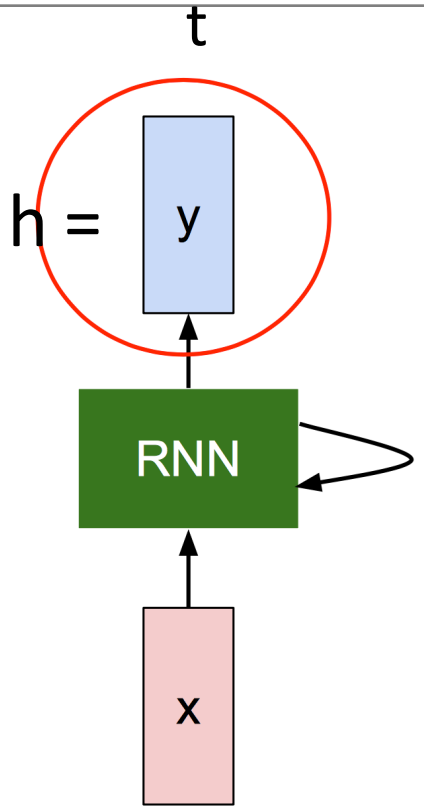
$$y_t = W_{hy} h_t$$

Character-level language model example

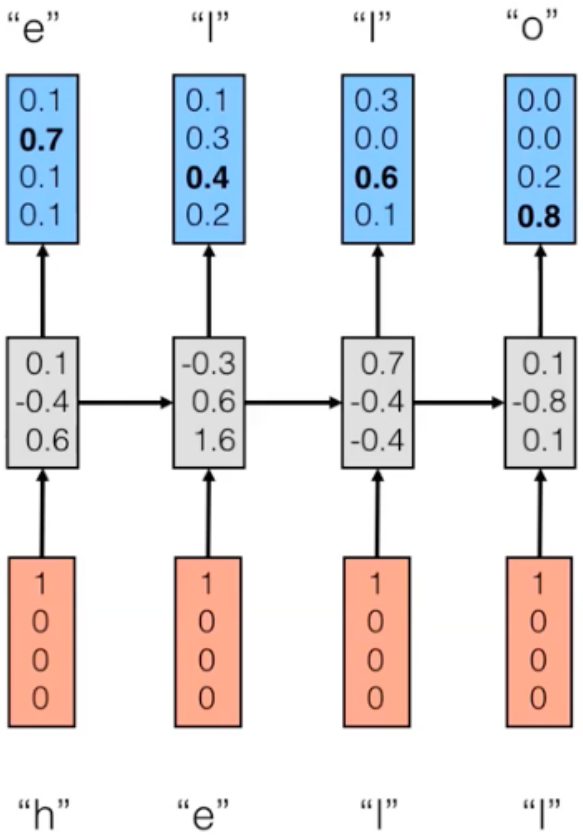
Vocabulary:
[h,e,l,o]

Example training sequence:
“hello”





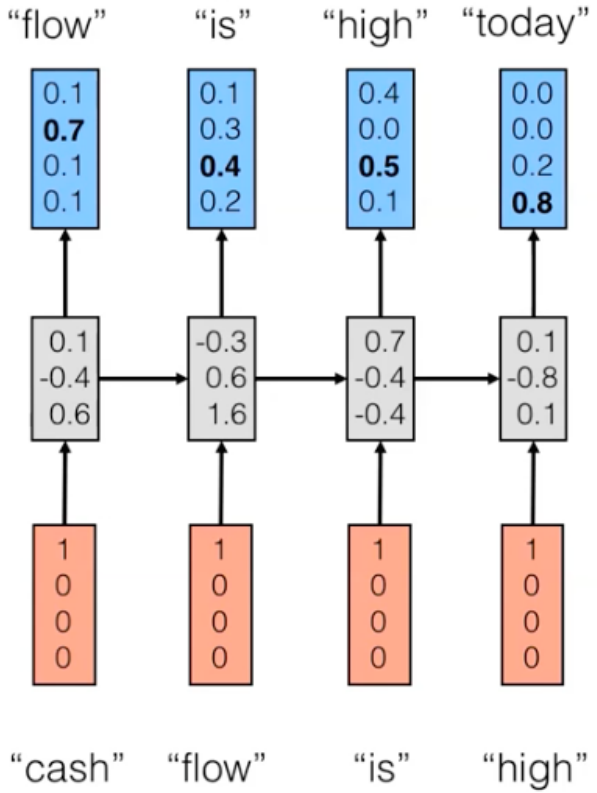
RNN



Learned a language model!

$$P(c_t|\{c_{t-1}, c_{t-2}, \dots, c_0\})$$

RNN – Use Case

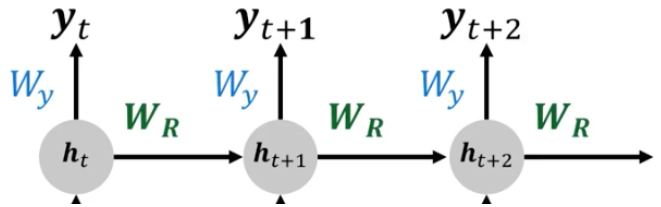


Learned a language model!

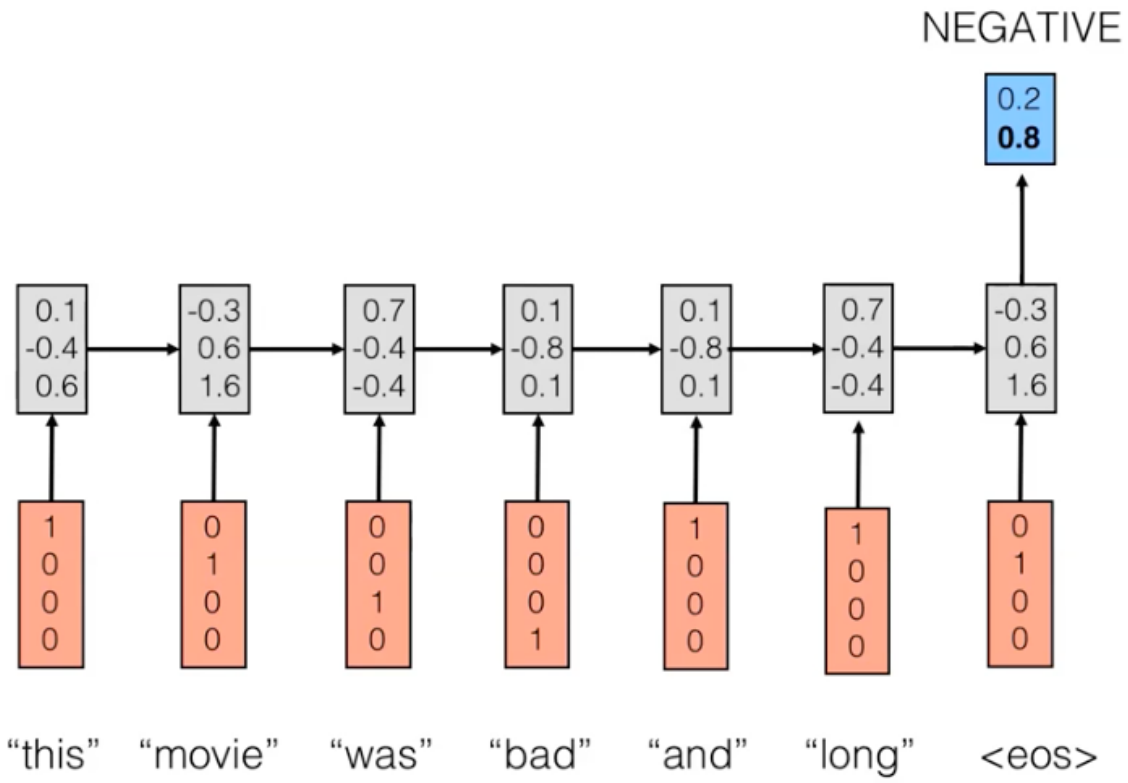
$$P(c_t|\{c_{t-1}, c_{t-2}, \dots, c_0\})$$

$$\boldsymbol{h}^{(t)} = g_h(W_{\mathrm{I}}\boldsymbol{x}^{(t)} + W_{\mathrm{R}}\boldsymbol{h}^{(t-1)} + \boldsymbol{b}_h)$$

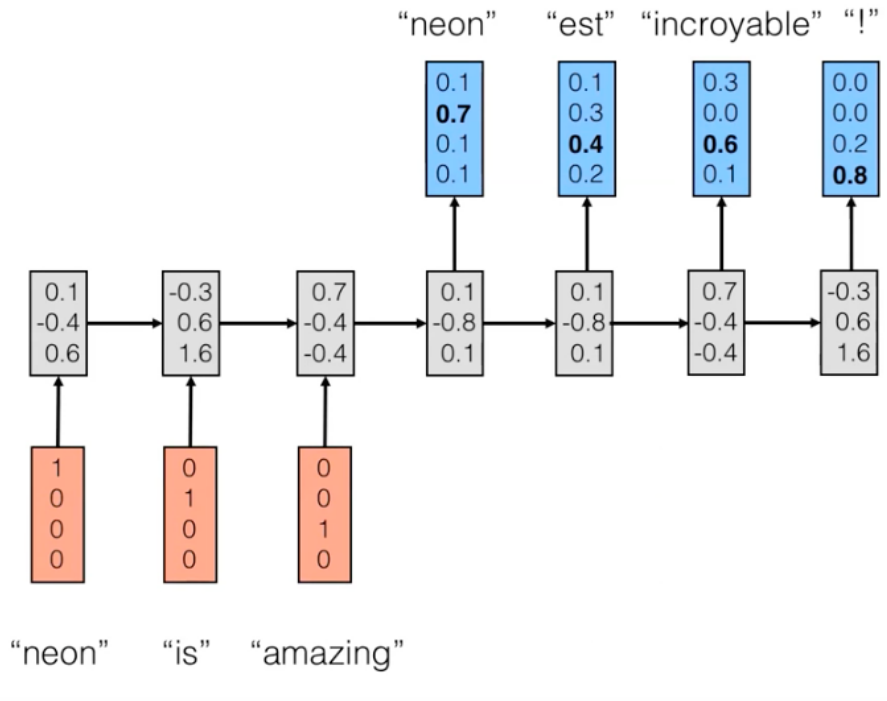
$$\boldsymbol{y}^{(t)} = g_y(W_y\boldsymbol{h}^{(t)} + \boldsymbol{b}_y)$$



RNN – Use Case


NEGATIVE
0.2
0.8
0.1
-0.4
0.6
-0.3
0.6
1.6
0.7
-0.4
-0.4
0.1
-0.8
0.1
0.1
-0.8
0.1
0.7
-0.4
-0.4
-0.3
0.6
1.6
1 0 0 0
0 1 0 0
0 0 1 0
0 0 0 1
1 0 0 0
1 0 0 0
0 1 0 0
“this”
“movie”
“was”
“bad”
“and”
“long”
<eos>

RNN- Use Case


"neon"
"est"
"incroyable"
"!"
0.1 -0.4 0.6
-0.3 0.6 1.6
0.7 -0.4 -0.4
0.1 -0.8 0.1
0.1 -0.8 0.1
0.7 -0.4 -0.4
-0.3 0.6 1.6
0.1 0.7 0.1 0.1
0.1 0.3 0.4 0.2
0.3 0.0 0.6 0.1
0.0 0.0 0.2 0.8
1 0 0 0
0 1 0 0
0 0 1 0
"neon"
"is"
"amazing"

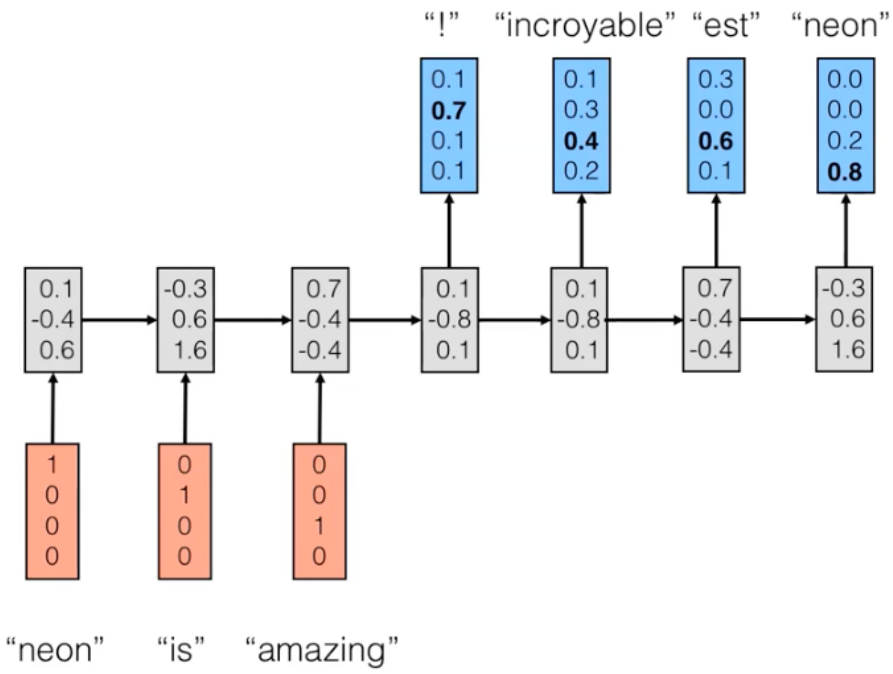

"!"
"incroyable"
"est"
"neon"
0.1 -0.4 0.6
-0.3 0.6 1.6
0.7 -0.4 -0.4
0.1 -0.8 0.1
0.1 -0.8 0.1
0.7 -0.4 -0.4
-0.3 0.6 1.6
0.1 0.7 0.1 0.1
0.1 0.3 0.4 0.2
0.3 0.0 0.6 0.1
0.0 0.0 0.2 0.8
1 0 0 0
0 1 0 0
0 0 1 0
"neon"
"is"
"amazing"

Image Captioning

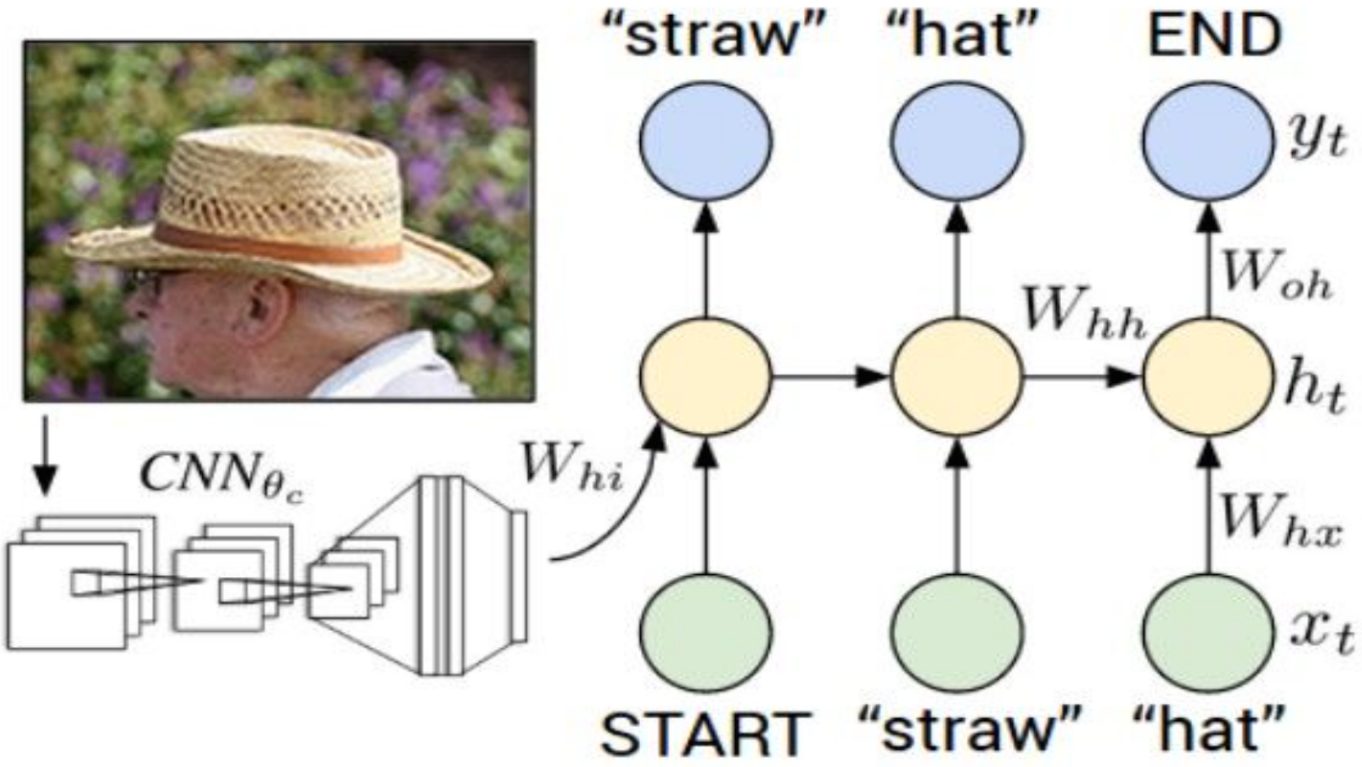
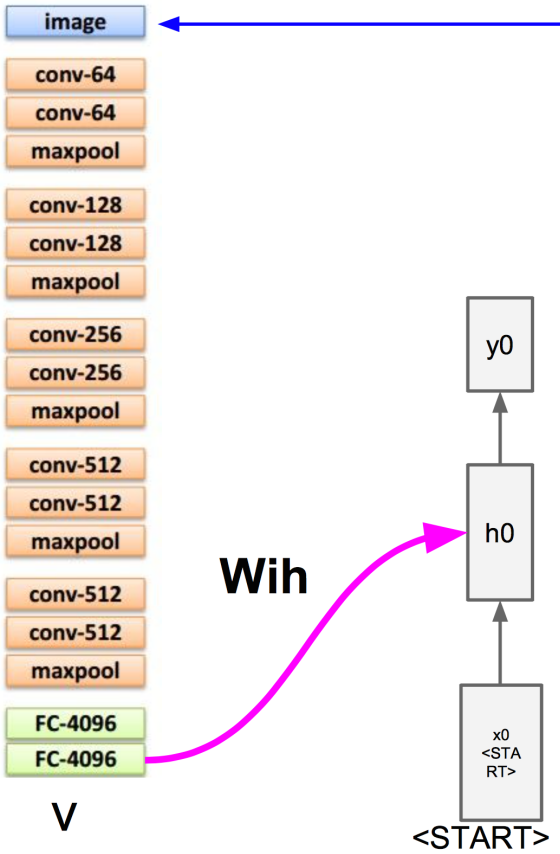


Image Captioning



test image

before:

h = tanh(Wxh * x + Whh * h)

now:

h = tanh(Wxh * x + Whh * h + **Wih * v**)

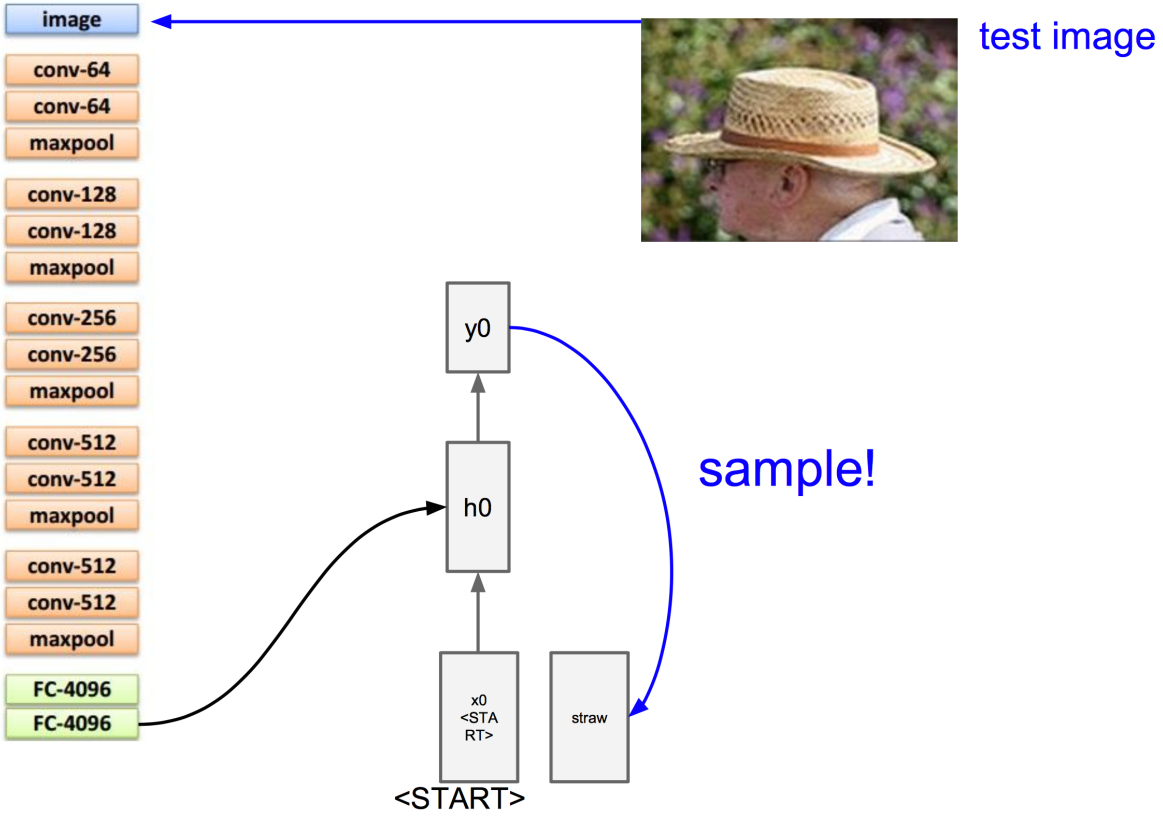


image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
maxpool
conv-512
conv-512
maxpool
conv-512
conv-512
maxpool
FC-4096
FC-4096
test image
y0
h0
x0
<STA
RT>
straw
<START>
sample!

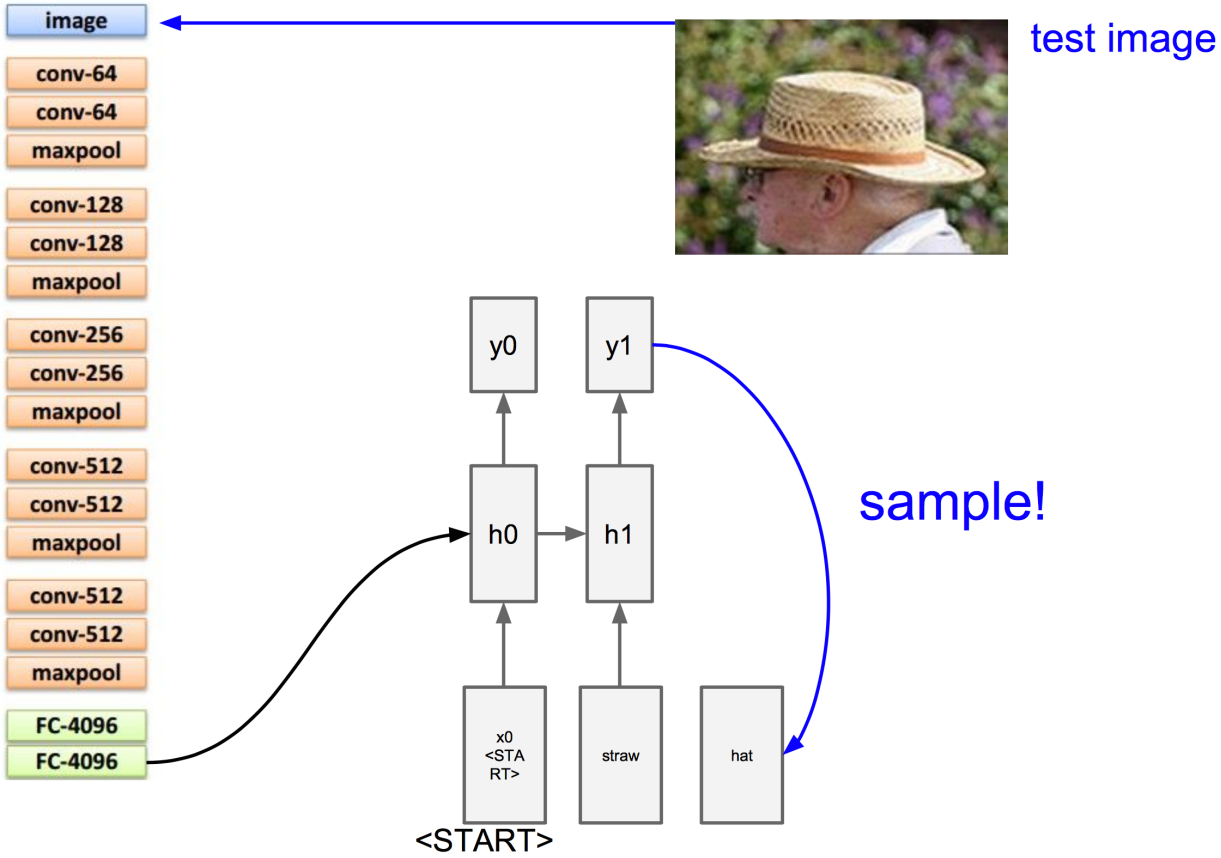
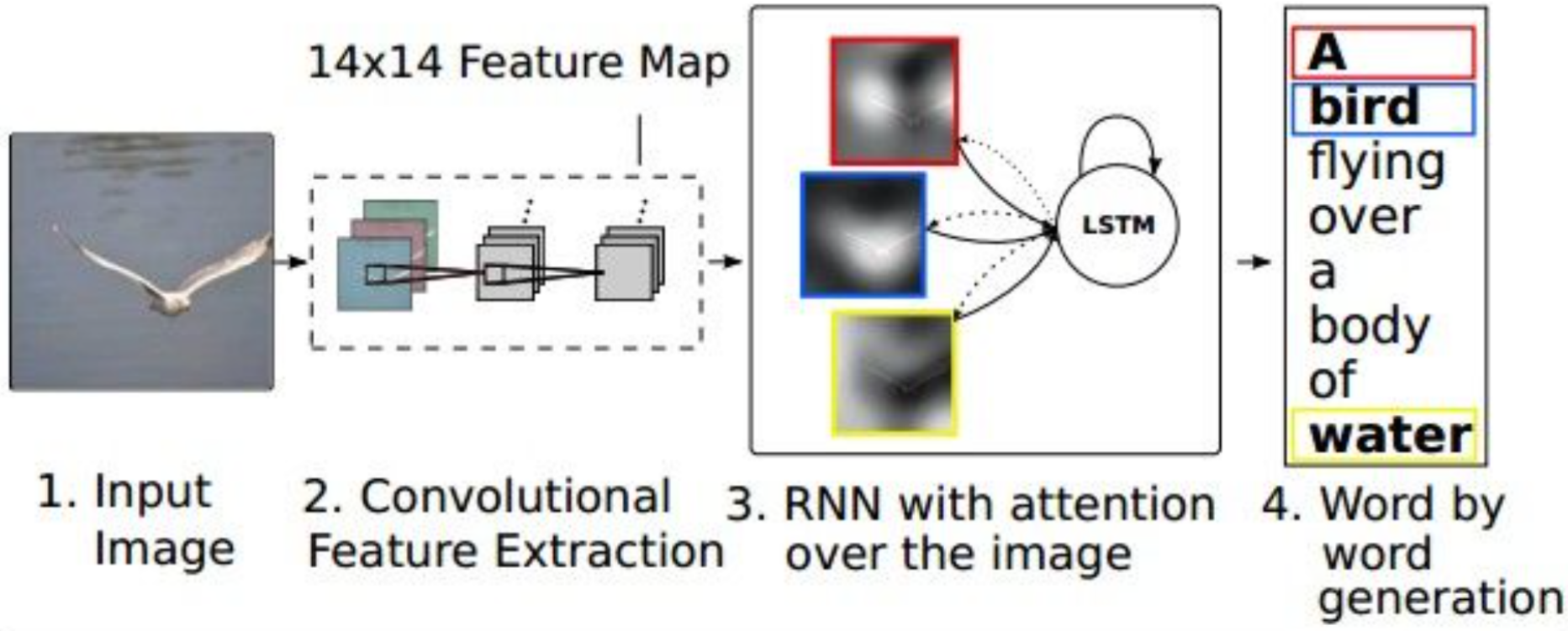


image
conv-64
conv-64
maxpool
conv-128
conv-128
maxpool
conv-256
conv-256
maxpool
conv-512
conv-512
maxpool
conv-512
conv-512
maxpool
FC-4096
FC-4096
test image
y0
y1
h0
h1
x0
<STA
RT>
straw
hat
<START>
sample!


14x14 Feature Map
LSTM
A
bird
flying
over
a
body
of
water
1. Input Image
2. Convolutional Feature Extraction
3. RNN with attention over the image
4. Word by word generation