

2021 직업계고 AI 전문교육

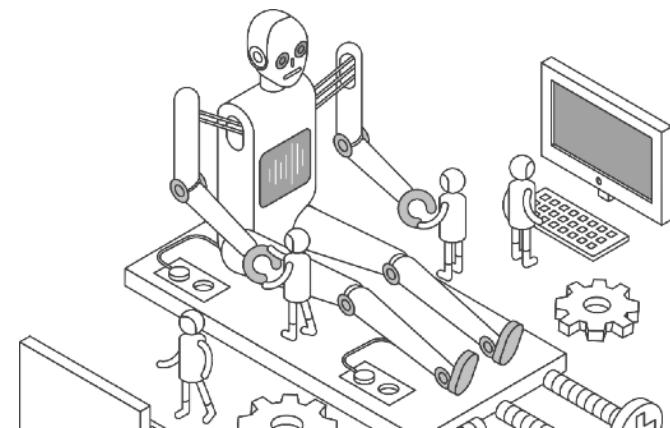
ARTIFICIAL INTELLIGENCE
BIG DATA
SMART FACTORY

AI·빅데이터 심화과정

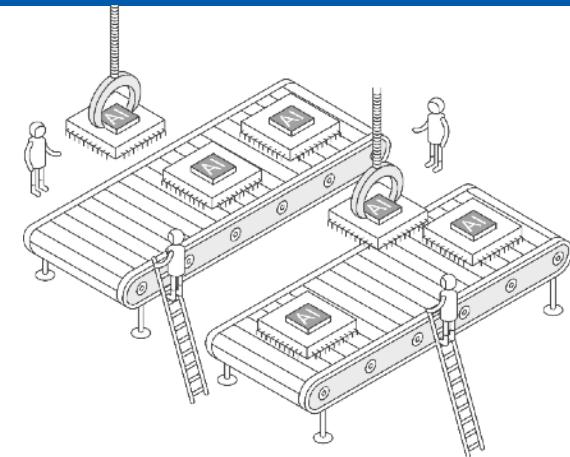
후반기 일정, RNN으로 소설쓰기

박수철

[github.com/scspark20](https://github.com/scpark20)
GaudioLab, 모두의연구소



후반기 일정



후반기 일정

Seq2Seq

9/24 - RNN으로 소설쓰기 (Aiffel 외)

9/29, 10/1 - 26. 뉴스 요약봇 만들기

10/6, 10/8 - 27. 트랜스포머로 만드는 대화형 챗봇

CNN/GAN

10/13, 10/15 - 22. 난 스케치를 할테니 너는 채색을 하거라

10/20, 10/22 - 21. 흐린 사진을 선명하게

10/27, 10/29 - 18. GO/STOP!

RNN+CNN

11/3, 11/5 - RNN으로 음성인식하기 (Aiffel 외)

11/10, 11/12 - 19. 직접 만들어보는 OCR

Ablation study

11/17 - 17. 없다면 어떻게 될까?

후반기 일정

Seq2Seq

9/24 - RNN으로 소설쓰기 (Aiffel 외)

9/29, 10/1 - 26. 뉴스 요약봇 만들기

10/6, 10/8 - 27. 트랜스포머로 만드는 대화형 챗봇

QUEENE:

I had thought thou hadst a Roman; for the oracle,
Thus by All bids the man against the word,
Which are so weak of care, by old care done;
Your children were in your holy love,
And the precipitation through the bleeding throne.

BISHOP OF ELY:

Marry, and will, my lord, to weep in such a one were prettiest;
Yet now I was adopted heir
Of the world's lamentable day,
To watch the next way with his father with his face?

ESCALUS:

The cause why then we are all resolved more sons.

https://www.tensorflow.org/text/tutorials/text_generation

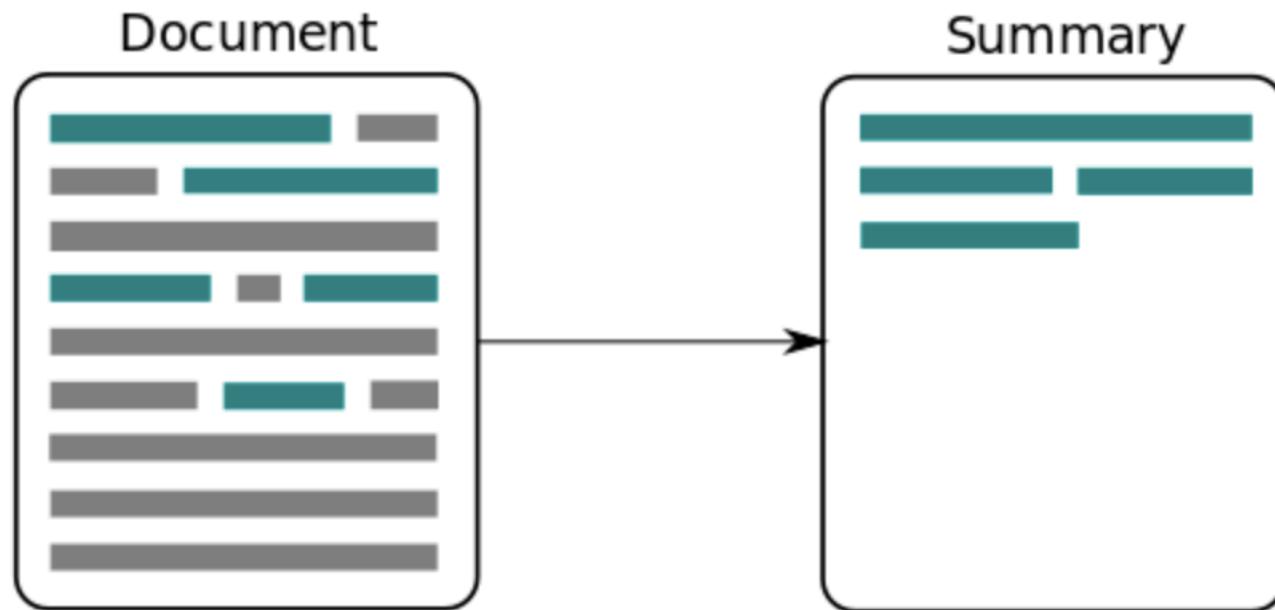
후반기 일정

Seq2Seq

9/24 - RNN으로 소설쓰기 (Aiffel 외)

9/29, 10/1 - 26. 뉴스 요약봇 만들기

10/6, 10/8 - 27. 트랜스포머로 만드는 대화형 챗봇



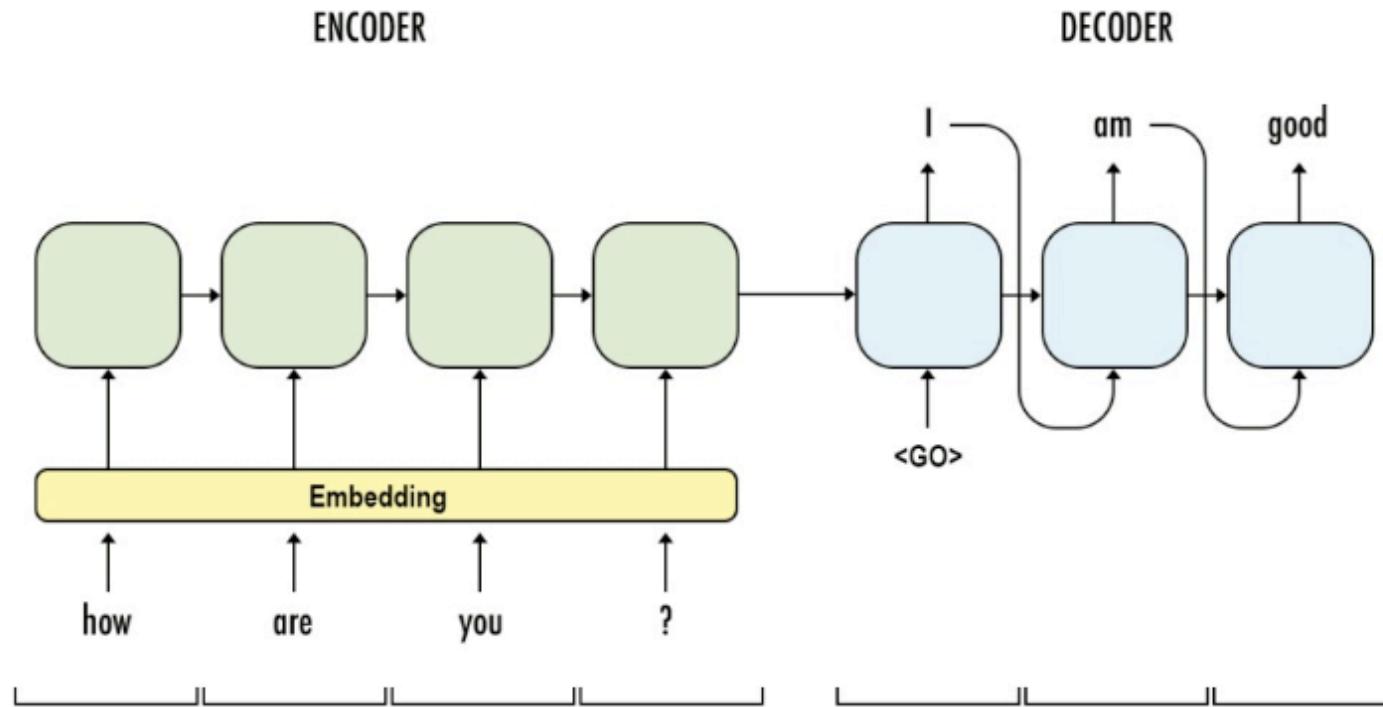
후반기 일정

Seq2Seq

9/24 - RNN으로 소설쓰기 (Aiffel 외)

9/29, 10/1 - 26. 뉴스 요약봇 만들기

10/6, 10/8 - 27. 트랜스포머로 만드는 대화형 챗봇



후반기 일정

CNN/GAN

10/13, 10/15 - 22. 난 스케치를 할테니 너는 채색을 하거라

10/20, 10/22 - 21. 흐린 사진을 선명하게

10/27, 10/29 - 18. GO/STOP!

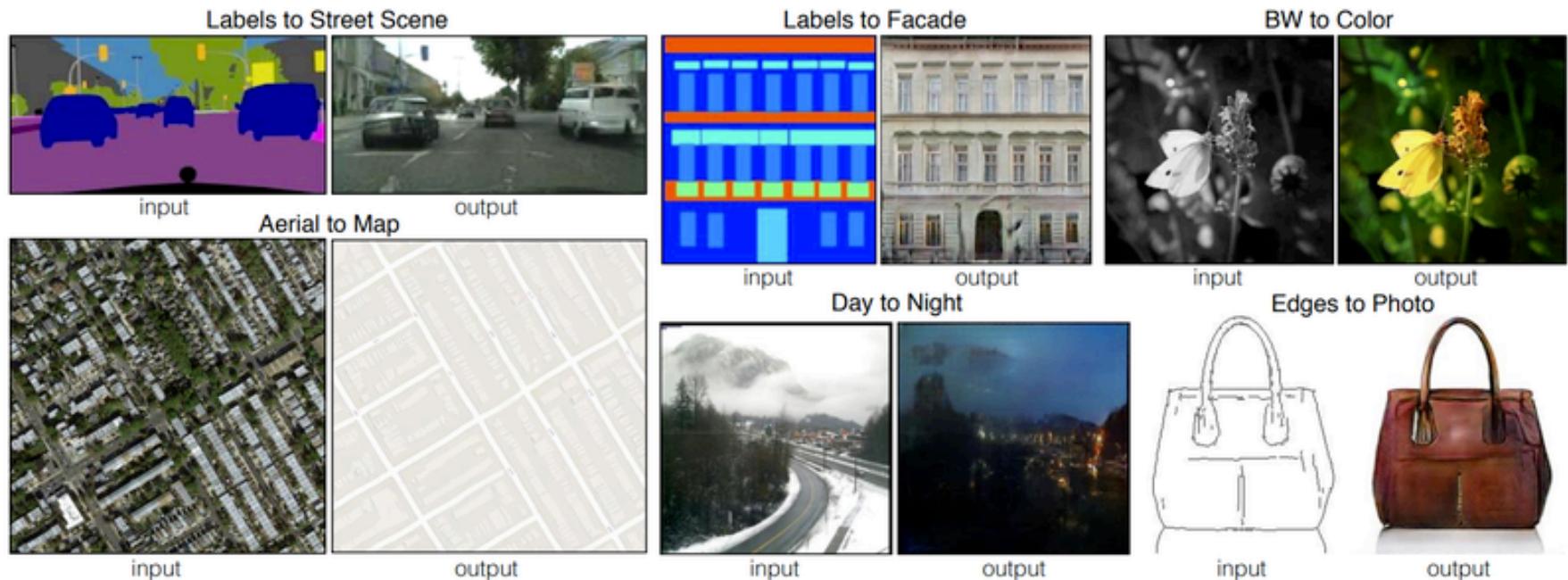


Image-to-Image Translation with Conditional Adversarial Networks
Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros

후반기 일정

CNN/GAN

10/13, 10/15 - 22. 난 스케치를 할테니 너는 채색을 하거라

10/20, 10/22 - 21. 흐린 사진을 선명하게

10/27, 10/29 - 18. GO/STOP!



[출처: <https://www.lifewire.com/fhd-vs-uhd-4688527>]

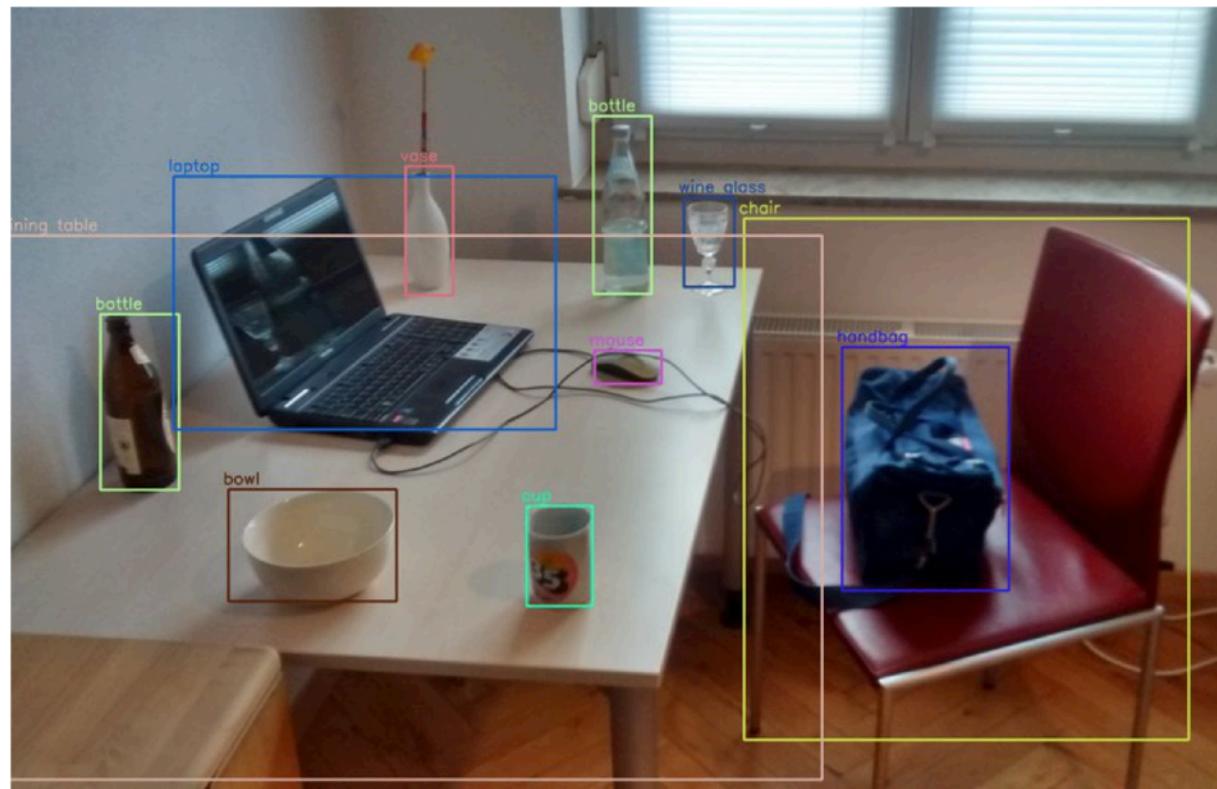
후반기 일정

CNN/GAN

10/13, 10/15 - 22. 난 스케치를 할테니 너는 채색을 하거라

10/20, 10/22 - 21. 흐린 사진을 선명하게

10/27, 10/29 - 18. GO/STOP!



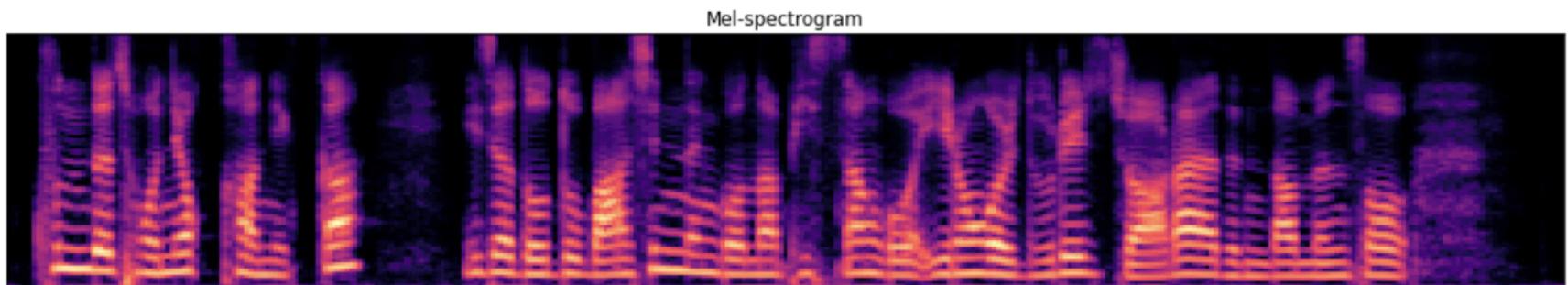
[출처: https://ko.wikipedia.org/wiki/객체_탐지]

후반기 일정

RNN+CNN

11/3, 11/5 - RNN으로 음성인식하기 (Aiffel 외)

11.10, 11/12 - 19. 직접 만들어보는 OCR



Ground Truth

군대 전역하니까 이제 너도 맛있는 거나 비싼 거 이런 거 누릴 줄 알아야 된다면서

Prediction

군대 전역하니까 이제 너도 맛있는 거나 비싼 거 이런 걸 눌리 주 알아야 된다면서

[https://github.com/scspark20/aispeech_flip/blob/main/asr/ctc_lstm_kspon.ipynb](https://github.com/scpark20/aispeech_flip/blob/main/asr/ctc_lstm_kspon.ipynb)

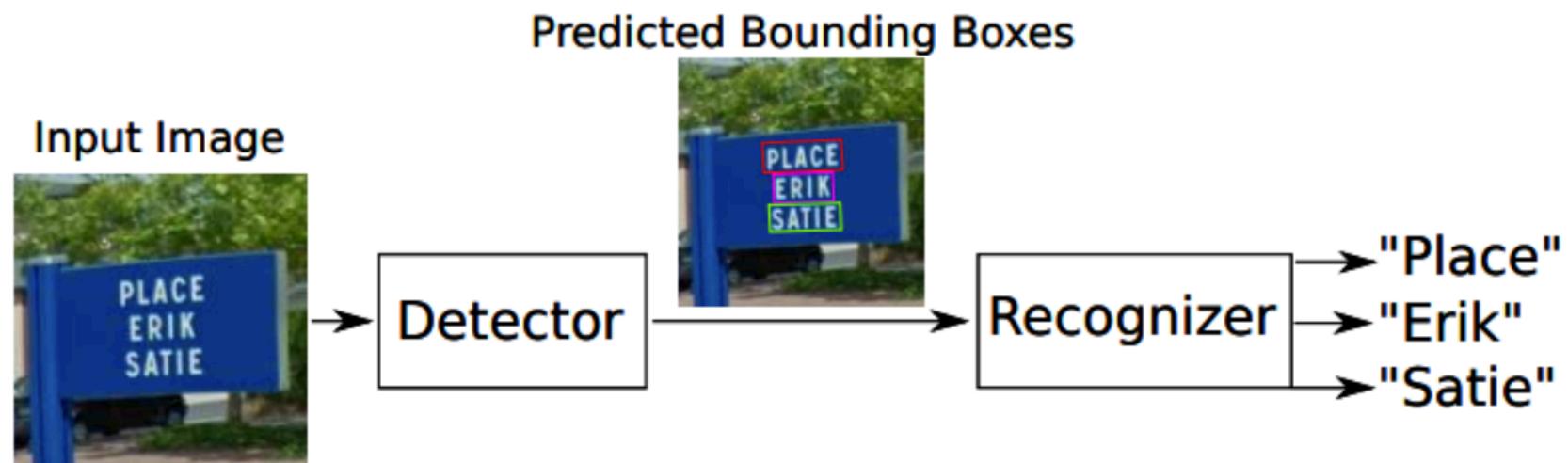
후반기 일정

RNN+CNN

11/3, 11/5 - RNN으로 음성인식하기 (Aiffel 외)

11.10, 11/12 - 19. 직접 만들어보는 OCR

OCR = Text detection + Text recognition



[출처 : <https://arxiv.org/pdf/1707.08831.pdf>]

Ablation study

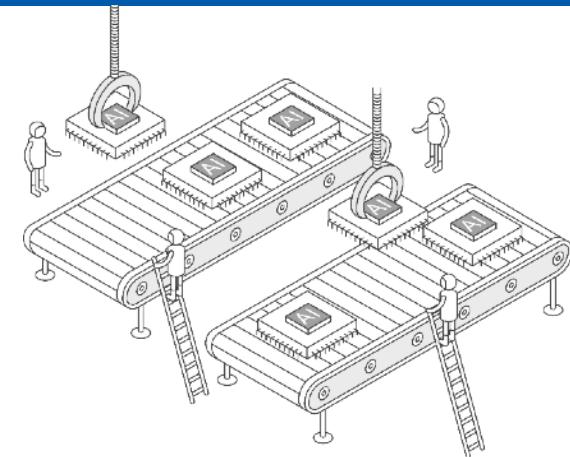
11/17 - 17. 없다면 어떻게 될까?

	Number of Parameters (millions)		Top-5 Error Rate (%)
VGG-11	133		10.4
VGG-11 (LRN)	133		10.5
VGG-13	133		9.9
VGG-16 (Conv1)	134		9.4
VGG-16	138		8.8
VGG-19	144		9.0

The diagram illustrates the architecture of six variants of the VGG model. Each variant starts with an input image and follows a similar pattern of layers: Conv3-64, Conv3-64, Max pool, followed by either LRN or another Max pool layer, and then a series of Conv3-128 and Conv3-256 layers with Max pool layers in between. The final layers consist of Conv3-512, Max pool, and then a fully connected layer (FC-4096, FC-4096, FC-1000, Soft-max). The 'VGG-11' and 'VGG-11 (LRN)' variants differ in the presence of the LRN layer. The 'VGG-13' variant has one less Max pool layer than the others. The 'VGG-16 (Conv1)' and 'VGG-16' variants differ in the first few layers; 'VGG-16 (Conv1)' uses Conv3-512, Conv3-512, Conv1-512, and Max pool, while 'VGG-16' uses Conv3-512, Conv3-512, Conv3-512, and Max pool. The 'VGG-19' variant adds an additional Max pool layer before the final FC-4096 layer.

모두 모두 파이!!!

RNN으로 소설쓰기



RNN으로 소설쓰기

가 나 다 라 마 () ()



RNN으로 소설쓰기

가 나 다 라 마 **바** 사

A B C D E F () () () ~ ~ ~



가 나 다 라 마 바 사

A B C D E F G H I ~ ~ ~



거꾸로~ 로꾸꺼~

하 파 타 카 차 자 () () ~~



거꾸로~ 로꾸꺼~

하 파 타 카 차 자 (아) (사) ~~



거꾸로~ 로꾸꺼~

하 파 타 카 차 자 (아) (사) ~~

Q P O N M L K J () () () ~~



거꾸로~ 로꾸꺼~

하 파 타 카 차 자 (아) (사) ~~

Q P O N M L K J (I) (H) (G)~~



Sequence 형태의 데이터를 뉴럴
네트워크가 학습할 수 있게 하는
것?

Recurrent Neural Networks

참고 : Anyone Can Learn To Code an LSTM-RNN in Python (한글번역), 유재준.

<http://jaejunyoo.blogspot.com/2017/06/anyone-can-learn-to-code-LSTM-RNN-Python.html>

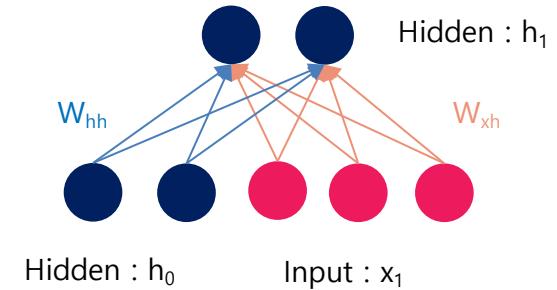
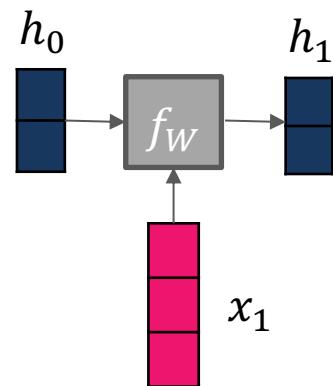


RNN : Computational Graph

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

가정

- Input : (3,1)
- Hidden : (2,1)
- Output : (3,1)



참고. : cs231n – Recurrent Neural Networks

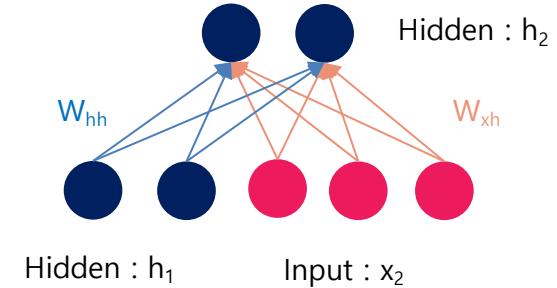
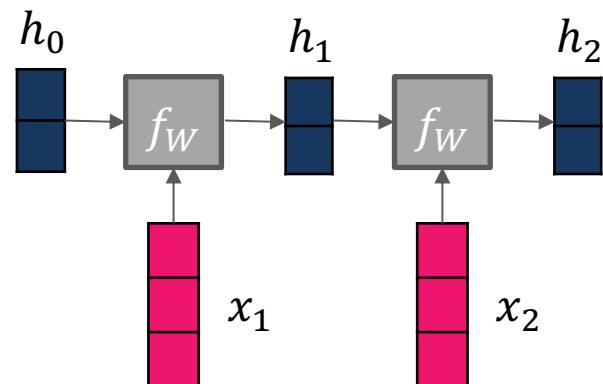


RNN : Computational Graph

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

가정

- Input : (3,1)
- Hidden : (2,1)
- Output : (3,1)



참고. : cs231n – Recurrent Neural Networks



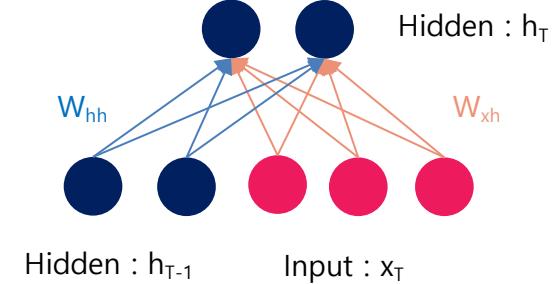
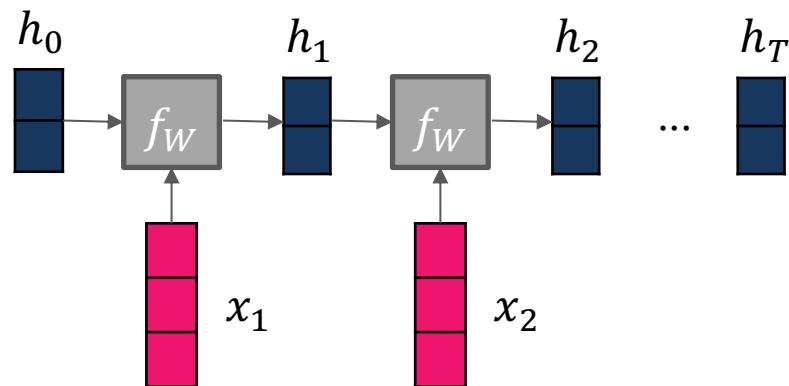
RNN : Computational Graph

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

가정

- Input : (3,1)
- Hidden : (2,1)
- Output : (3,1)

- f_W 는 하나 입니다
- 모든 타임스텝에서 동일한 W_{hh} , W_{xh} , W_{hy} 가 사용 됩니다
- 데이터만 바꾸는 것 입니다



참고. : cs231n – Recurrent Neural Networks

자동완성 예시

'기' '억' ' ' '속' '으' '로'

맨발의 디바, 이은미



참고. : cs231n – Recurrent Neural Networks

자동완성 예시

'**기**' 를 입력 했을 때 자동으로
'**기** **억 속**으**로**'

가 나오는 모델을 학습 한다고 가정 해 봅시다

'**기**' '**억**' ' ' '**속**' '**으**' '**로**'

참고. : cs231n – Recurrent Neural Networks



자동완성 예시

'기' 를 입력 했을 때 자동으로
'기_억 속_{으로}'

가 나오는 모델을 학습 한다고 가정 해 봅시다

'기' '억' ' ' '속' '으' '로'

단순한 예시를 만들기 위해

한글 글자가 '기', '억', ' ', '속', '으', '로' 만 있다고 가정 해 봅니다

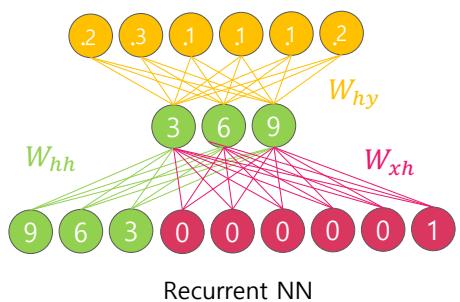
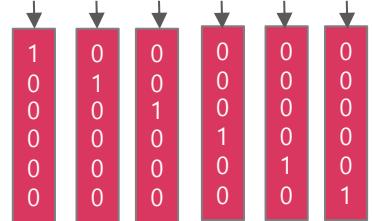
참고. : cs231n – Recurrent Neural Networks



자동완성 예시

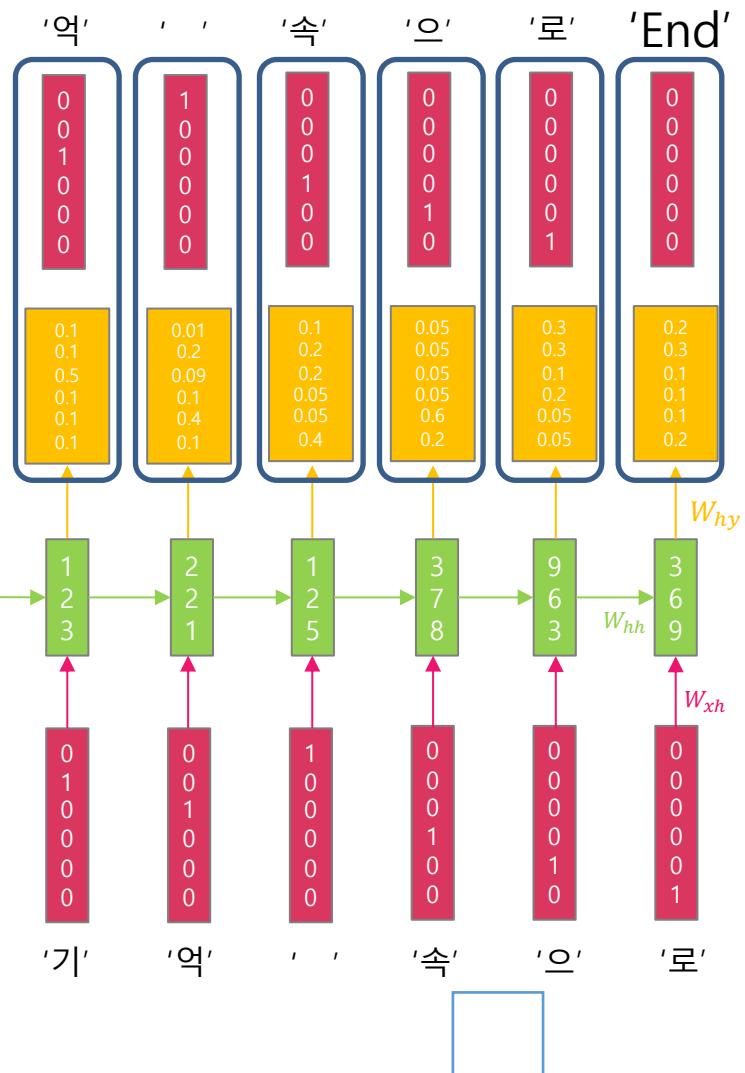
우리의 글자 보관함

[' ' '기' '억' '속' '으' '로']



참고. : cs231n – Recurrent Neural Networks

Cross entropy



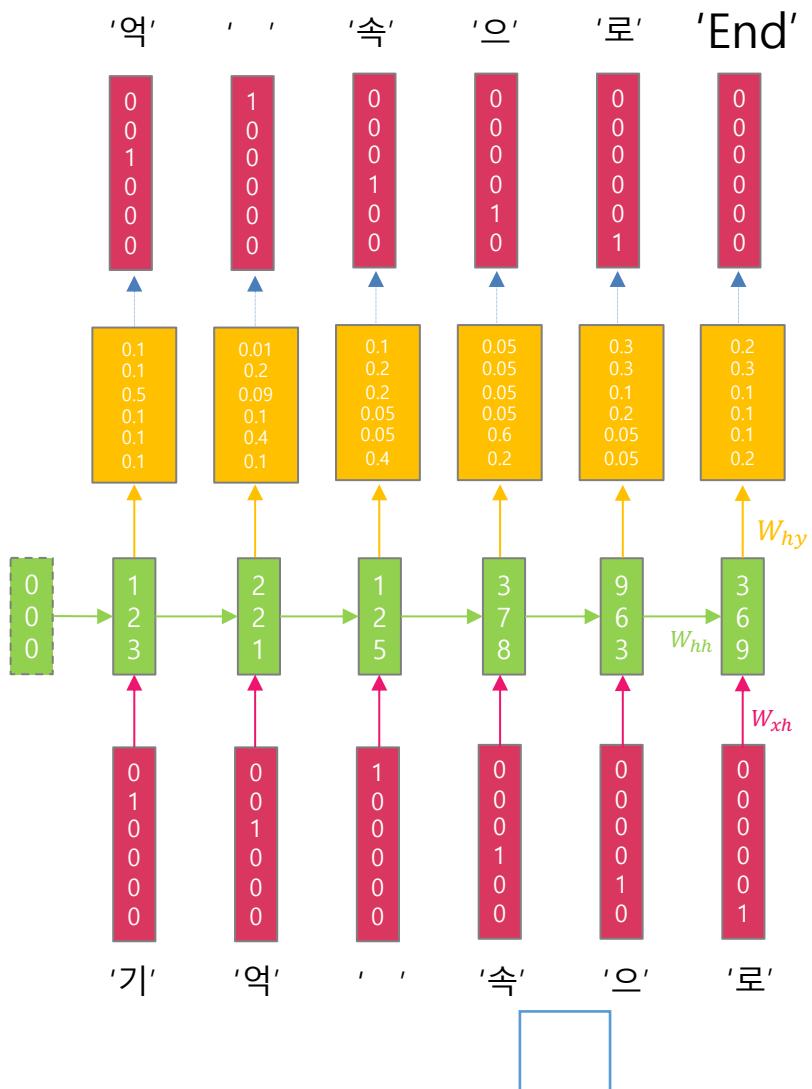
자동완성 예시

학습 완료 후

우리의 글자 보관함

[' ' '기' '억' '속' '으' '로']

['기'	'억'	'속'	'으'	'로']
↓	↓	↓	↓	↓	↓	↓	↓	↓
1 0 0 0 0	0 1 0 0 0	0 0 1 0 0	0 0 0 1 0	0 0 0 0 1	0 0 0 0 0	0 0 0 1 0	0 0 0 0 0	0 0 0 0 1



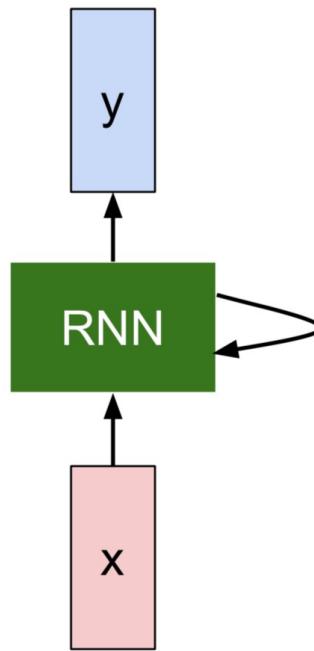
참고. : cs231n – Recurrent Neural Networks

THE SONNETS

by William Shakespeare

From fairest creatures we desire increase,
That thereby beauty's rose might never die,
But as the riper should by time decease,
His tender heir might bear his memory:
But thou, contracted to thine own bright eyes,
Feed'st thy light's flame with self-substantial fuel,
Making a famine where abundance lies,
Thyself thy foe, to thy sweet self too cruel:
Thou that art now the world's fresh ornament,
And only herald to the gaudy spring,
Within thine own bud buriest thy content,
And tender churl mak'st waste in niggarding:
Pity the world, or else this glutton be,
To eat the world's due, by the grave and thee.

When forty winters shall besiege thy brow,
And dig deep trenches in thy beauty's field,
Thy youth's proud livery so gazed on now,
Will be a tatter'd weed of small worth held:
Then being asked, where all thy beauty lies,
Where all the treasure of thy lusty days;
To say, within thine own deep sunken eyes,
Were an all-eating shame, and thriftless praise.
How much more praise deser'v'd thy beauty's use,
If thou couldst answer 'This fair child of mine
Shall sum my count, and make my old excuse,'
Proving his beauty by succession thine!
This were to be new made when thou art old,
And see thy blood warm when thou feel'st it cold.



RNN으로 소설쓰기

at first:

tyntd-iafhatawiaoahrdemot 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 after.

↓ 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.

RNN으로 소설쓰기

PANDARUS:

Alas, I think he shall be come approached and the day
When little strain would be attain'd into being never fed,
And who is but a chain and subjects of his death,
I should not sleep.

Second Senator:

They are away this miseries, produced upon my soul,
Breaking and strongly should be buried, when I perish
The earth and thoughts of many states.

DUKE VINCENTIO:

Well, your wit is in the care of side and that.

Second Lord:

They would be ruled after this chamber, and
my fair nues begun out of the fact, to be conveyed,
Whose noble souls I'll have the heart of the wars.

Clown:

Come, sir, I will make did behold your worship.

VIOLA:

I'll drink it.

VIOLA:

Why, Salisbury must find his flesh and thought
That which I am not aps, not a man and in fire,
To show the reining of the raven and the wars
To grace my hand reproach within, and not a fair are hand,
That Caesar and my goodly father's world;
When I was heaven of presence and our fleets,
We spare with hours, but cut thy council I am great,
Murdered and by thy master's ready there
My power to give thee but so much as hell:
Some service in the noble bondman here,
Would show him to her wine.

KING LEAR:

O, if you were a feeble sight, the courtesy of your law,
Your sight and several breath, will wear the gods
With his heads, and my hands are wonder'd at the deeds,
So drop upon your lordship's head, and your opinion
Shall be against your honour.

RNN으로 소설쓰기

For $\bigoplus_{n=1,\dots,m} \mathcal{L}_{m,n} = 0$, hence we can find a closed subset \mathcal{H} in \mathcal{H} and any sets \mathcal{F} on X , U is a closed immersion of S , then $U \rightarrow T$ is a separated algebraic space.

Proof. Proof of (1). It also start we get

$$S = \text{Spec}(R) = U \times_X U \times_X U$$

and the comparicoly in the fibre product covering we have to prove the lemma generated by $\coprod Z \times_U U \rightarrow V$. Consider the maps M along the set of points Sch_{fppf} and $U \rightarrow U$ is the fibre category of S in U in Section ?? and the fact that any U affine, see Morphisms, Lemma ???. Hence we obtain a scheme S and any open subset $W \subset U$ in $\text{Sh}(G)$ such that $\text{Spec}(R') \rightarrow S$ is smooth or an

$$U = \bigcup U_i \times_{S_i} U_i$$

which has a nonzero morphism we may assume that f_i is of finite presentation over S . We claim that $\mathcal{O}_{X,x}$ is a scheme where $x, x', s'' \in S'$ such that $\mathcal{O}_{X,x'} \rightarrow \mathcal{O}'_{X',x'}$ is separated. By Algebra, Lemma ?? we can define a map of complexes $\text{GL}_{S'}(x'/S'')$ and we win. \square

To prove study we see that $\mathcal{F}|_U$ is a covering of \mathcal{X}' , and \mathcal{T}_i is an object of $\mathcal{F}_{X/S}$ for $i > 0$ and \mathcal{F}_p exists and let \mathcal{F}_i be a presheaf of \mathcal{O}_X -modules on \mathcal{C} as a \mathcal{F} -module. In particular $\mathcal{F} = U/\mathcal{F}$ we have to show that

$$\widetilde{M}^\bullet = \mathcal{I}^\bullet \otimes_{\text{Spec}(k)} \mathcal{O}_{S,s} - i_X^{-1} \mathcal{F}$$

is a unique morphism of algebraic stacks. Note that

$$\text{Arrows} = (\text{Sch}/S)^{\text{opp}}_{fppf}, (\text{Sch}/S)_{fppf}$$

and

$$V = \Gamma(S, \mathcal{O}) \hookrightarrow (U, \text{Spec}(A))$$

is an open subset of X . Thus U is affine. This is a continuous map of X is the inverse, the groupoid scheme S .

Proof. See discussion of sheaves of sets. \square

The result for prove any open covering follows from the less of Example ???. It may replace S by $X_{\text{spaces},\text{étale}}$ which gives an open subspace of X and T equal to S_{Zar} , see Descent, Lemma ???. Namely, by Lemma ?? we see that R is geometrically regular over S .

Lemma 0.1. Assume (3) and (3) by the construction in the description.

Suppose $X = \lim |X|$ (by the formal open covering X and a single map $\underline{\text{Proj}}_X(\mathcal{A}) = \text{Spec}(B)$ over U compatible with the complex

$$\text{Set}(\mathcal{A}) = \Gamma(X, \mathcal{O}_{X,\mathcal{O}_X}).$$

When in this case of to show that $\mathcal{Q} \rightarrow \mathcal{C}_{Z/X}$ is stable under the following result in the second conditions of (1), and (3). This finishes the proof. By Definition ?? (without element is when the closed subschemes are catenary. If T is surjective we may assume that T is connected with residue fields of S . Moreover there exists a closed subspace $Z \subset X$ of X where U in X' is proper (some defining as a closed subset of the uniqueness it suffices to check the fact that the following theorem

(1) f is locally of finite type. Since $S = \text{Spec}(R)$ and $Y = \text{Spec}(R)$.

Proof. This is form all sheaves of sheaves on X . But given a scheme U and a surjective étale morphism $U \rightarrow X$. Let $U \cap U = \coprod_{i=1,\dots,n} U_i$ be the scheme X over S at the schemes $X_i \rightarrow X$ and $U = \lim_i X_i$. \square

The following lemma surjective restrocomposes of this implies that $\mathcal{F}_{x_0} = \mathcal{F}_{x_0} = \mathcal{F}_{X,\dots,0}$.

Lemma 0.2. Let X be a locally Noetherian scheme over S , $E = \mathcal{F}_{X/S}$. Set $\mathcal{I} = \mathcal{J}_1 \subset \mathcal{I}'_n$. Since $\mathcal{I}^n \subset \mathcal{I}^n$ are nonzero over $i_0 \leq p$ is a subset of $\mathcal{J}_{n,0} \circ \bar{A}_2$ works.

Lemma 0.3. In Situation ???. Hence we may assume $q' = 0$.

Proof. We will use the property we see that p is the next functor (??). On the other hand, by Lemma ?? we see that

$$D(\mathcal{O}_{X'}) = \mathcal{O}_X(D)$$

where K is an F -algebra where δ_{n+1} is a scheme over S . \square

RNN으로 소설쓰기

```
static void do_command(struct seq_file *m, void *v)
{
    int column = 32 << (cmd[2] & 0x80);
    if (state)
        cmd = (int)(int_state ^ (in_8(&ch->ch_flags) & Cmd) ? 2 : 1);
    else
        seq = 1;
    for (i = 0; i < 16; i++) {
        if (k & (1 << i))
            pipe = (in_use & UMXTHREAD_UNCCA) +
                ((count & 0x00000000fffffff8) & 0x0000000f) << 8;
        if (count == 0)
            sub(pid, ppc_md.kexec_handle, 0x20000000);
        pipe_set_bytes(i, 0);
    }
    /* Free our user pages pointer to place camera if all dash */
    subsystem_info = &of_changes[PAGE_SIZE];
    rek_controls(offset, idx, &soffset);
    /* Now we want to deliberately put it to device */
    control_check_polarity(&context, val, 0);
    for (i = 0; i < COUNTER; i++)
        seq_puts(s, "policy ");
}
```

Generated C code

Fei-Fei Li & Justin Johnson & Serena Yeung

Lecture 10 - 53 May 4, 2017
22

Image Captioning

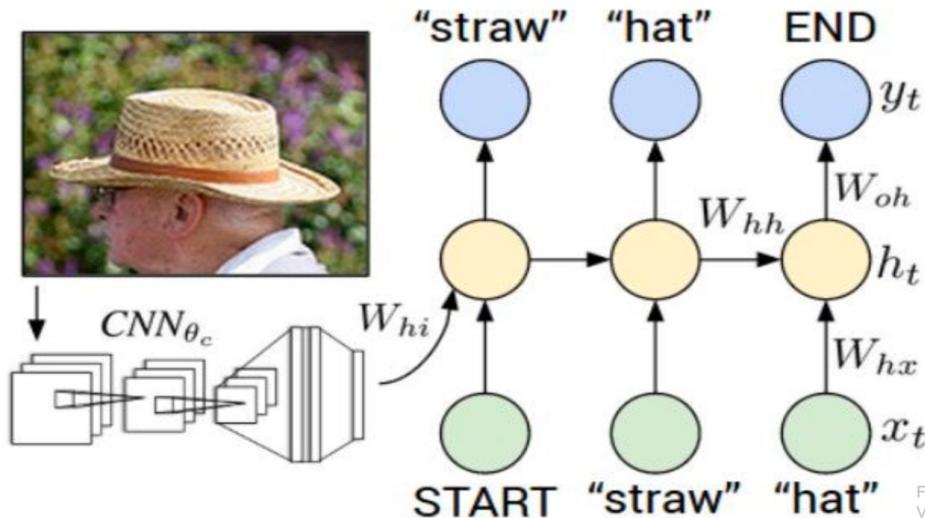


Figure from Karpathy et al, "Deep Visual-Semantic Alignments for Generating Image Descriptions", CVPR 2015; figure copyright IEEE, 2015.
Reproduced for educational purposes.

Explain Images with Multimodal Recurrent Neural Networks, Mao et al.

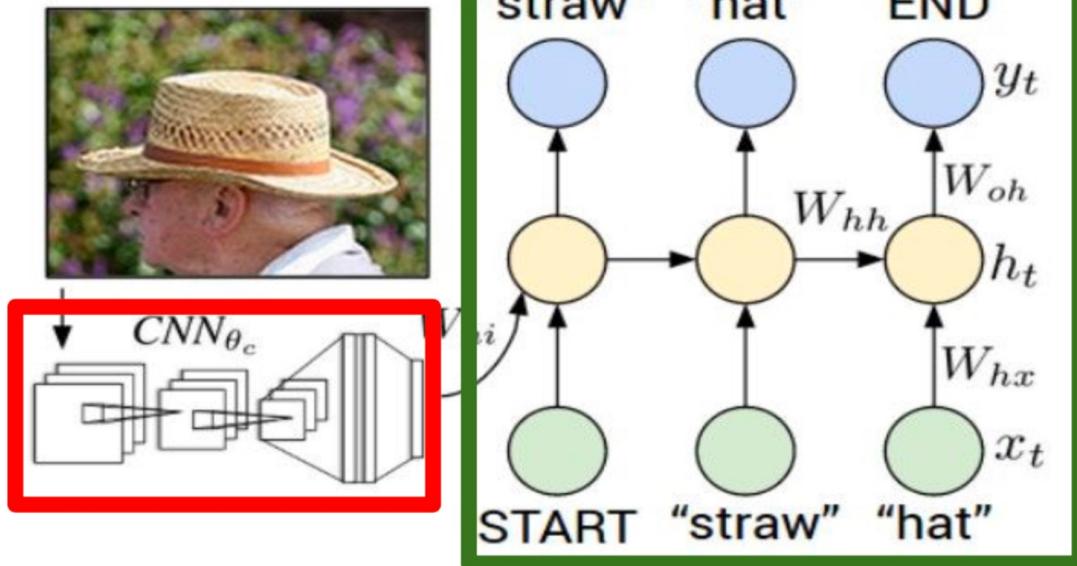
Deep Visual-Semantic Alignments for Generating Image Descriptions, Karpathy and Fei-Fei

Show and Tell: A Neural Image Caption Generator, Vinyals et al.

Long-term Recurrent Convolutional Networks for Visual Recognition and Description, Donahue et al.

Learning a Recurrent Visual Representation for Image Caption Generation, Chen and Zitnick

Recurrent Neural Network



Convolutional Neural Network

RNN으로 소설쓰기

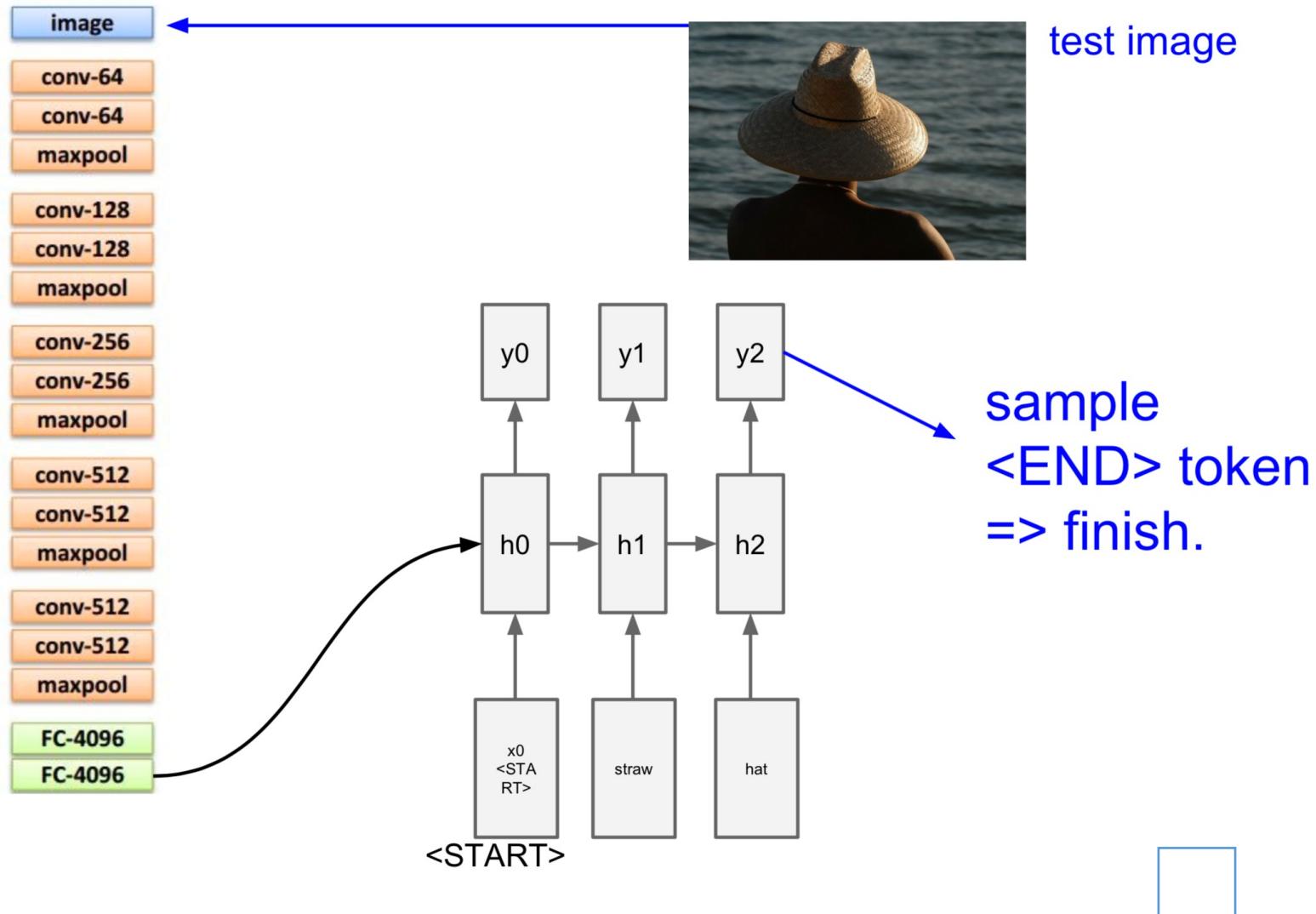


Image Captioning: Example Results



A cat sitting on a suitcase on the floor



A cat is sitting on a tree branch



A dog is running in the grass with a frisbee



A white teddy bear sitting in the grass



Two people walking on the beach with surfboards



A tennis player in action on the court



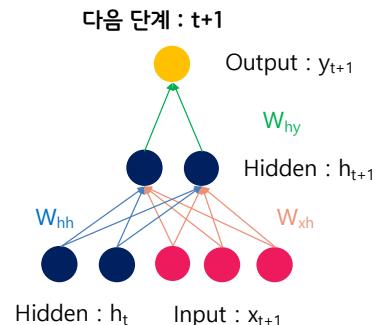
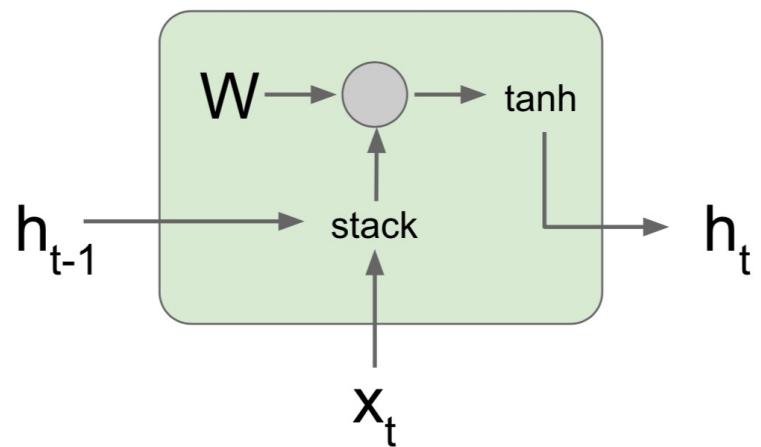
Two giraffes standing in a grassy field



A man riding a dirt bike on a dirt track

Captions generated using [neuraltalk2](#)
All images are CC0 Public domain:
[cat suitcase](#), [cat tree](#), [dog](#), [bear](#),
[surfers](#), [tennis](#), [giraffe](#), [motorcycle](#)

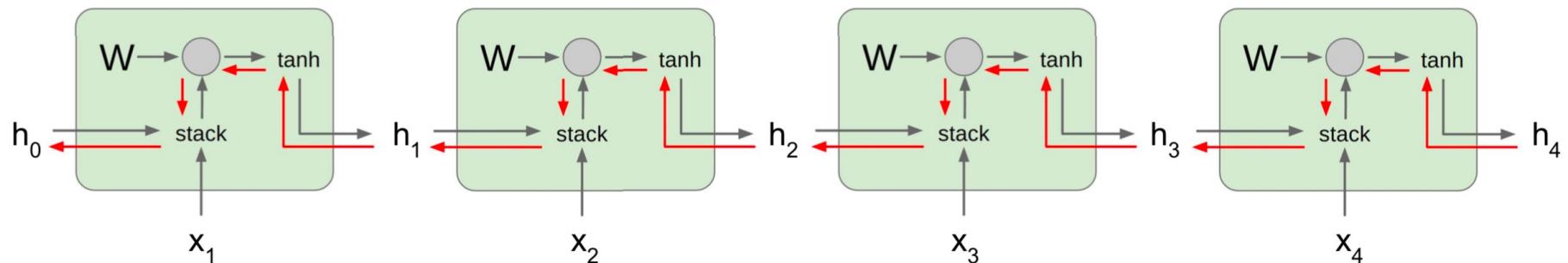
Vanilla RNN Gradient Flow



$$\begin{aligned}
 h_t &= \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \\
 &= \tanh \left(\begin{pmatrix} W_{hh} & W_{hx} \end{pmatrix} \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right) \\
 &= \tanh \left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \right)
 \end{aligned}$$

Vanilla RNN Gradient Flow

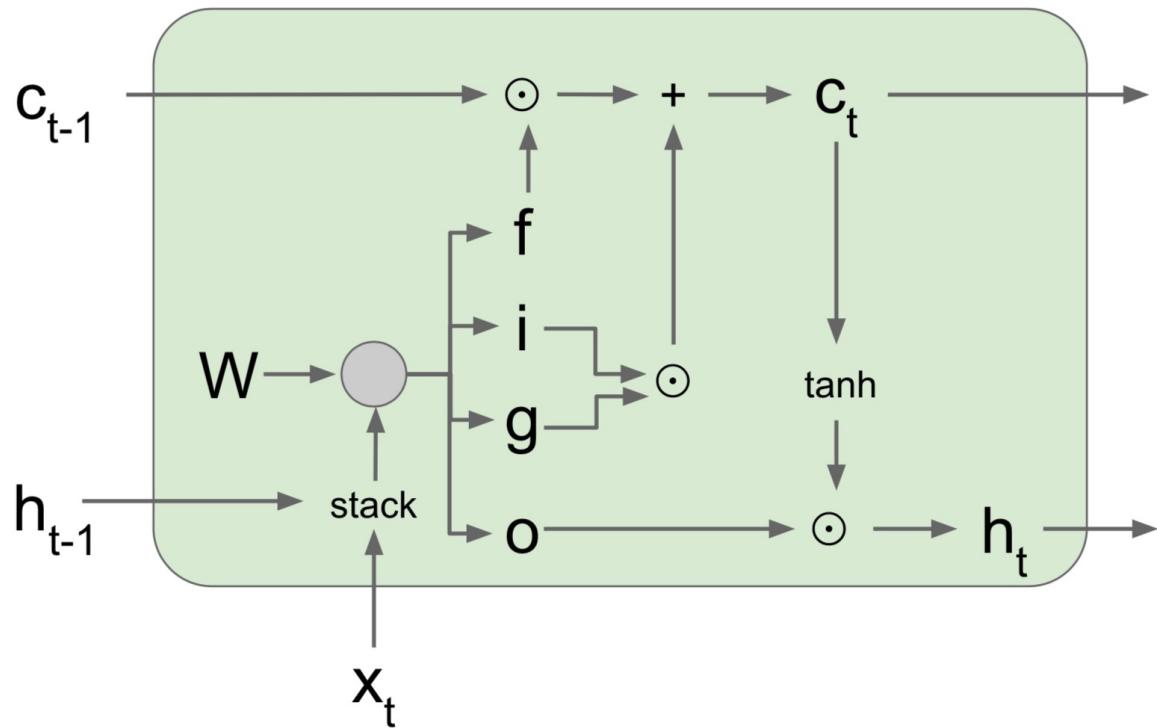
Bengio et al, "Learning long-term dependencies with gradient descent is difficult", IEEE Transactions on Neural Networks, 1994
Pascanu et al, "On the difficulty of training recurrent neural networks", ICML 2013



Computing gradient
of h_0 involves many
factors of W
(and repeated tanh)

Long Short Term Memory (LSTM)

[Hochreiter et al., 1997]



$$\begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}$$

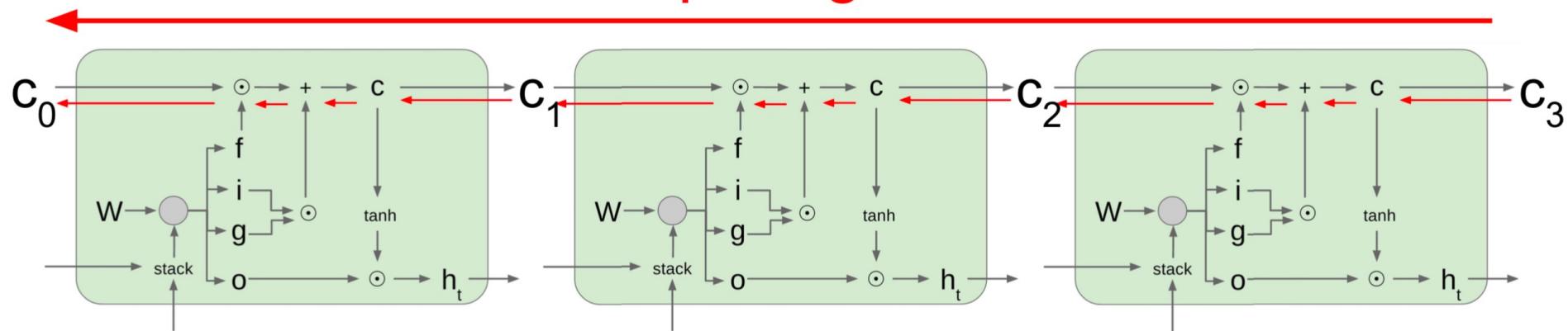
$$c_t = f \odot c_{t-1} + i \odot g$$

$$h_t = o \odot \tanh(c_t)$$

Long Short Term Memory (LSTM): Gradient Flow

[Hochreiter et al., 1997]

Uninterrupted gradient flow!



Other RNN Variants

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

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r)$$

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z)$$

$$\tilde{h}_t = \tanh(W_{xh}x_t + W_{hh}(r_t \odot h_{t-1}) + b_h)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t$$

[*LSTM: A Search Space Odyssey*, Greff et al., 2015]

[*An Empirical Exploration of Recurrent Network Architectures*, Jozefowicz et al., 2015]

MUT1:

$$z = \text{sigm}(W_{xz}x_t + b_z)$$

$$r = \text{sigm}(W_{xr}x_t + W_{hr}h_t + b_r)$$

$$h_{t+1} = \tanh(W_{hh}(r \odot h_t) + \tanh(x_t) + b_h) \odot z + h_t \odot (1 - z)$$

MUT2:

$$z = \text{sigm}(W_{xz}x_t + W_{hz}h_t + b_z)$$

$$r = \text{sigm}(x_t + W_{hr}h_t + b_r)$$

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

MUT3:

$$z = \text{sigm}(W_{xz}x_t + W_{hz}\tanh(h_t) + b_z)$$

$$r = \text{sigm}(W_{xr}x_t + W_{hr}h_t + b_r)$$

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

모두 모두 파이!!!