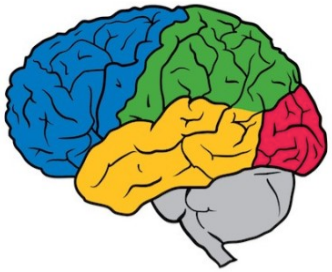
Efficient and Effective Models for Machine Reading Comprehension

Adams Wei Yu

Carnegie Mellon University

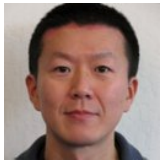


Collaborators


David Dohan


Thang Luong


Rui Zhao

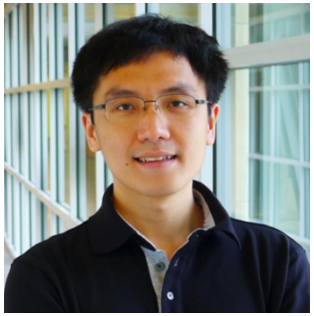

Kai Chen


Mohammad Norouzi


Hongrae Lee


Quoc Le

Bio

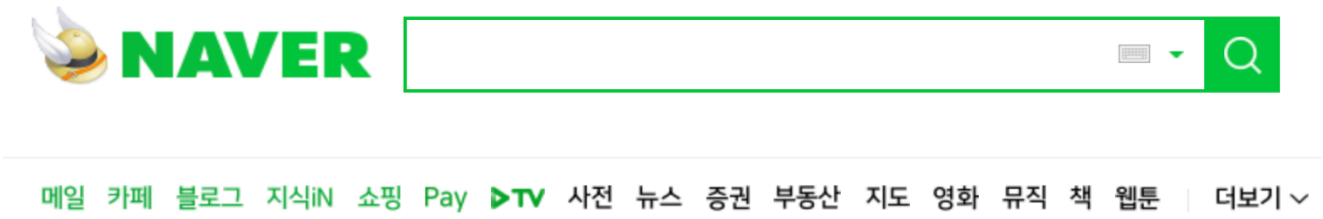


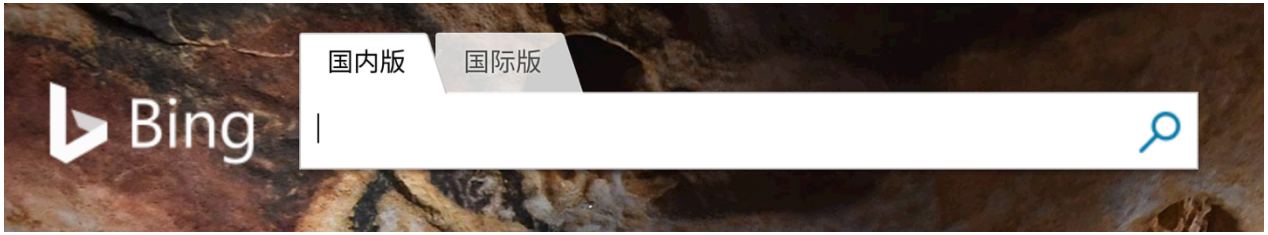
Adams Wei Yu

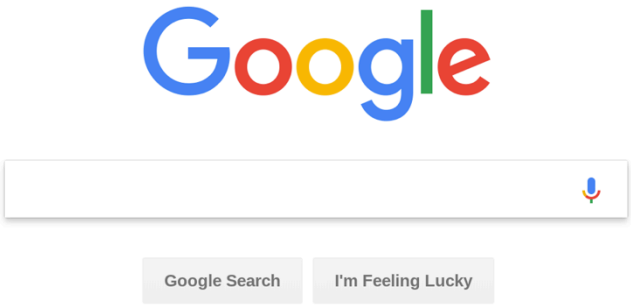
- Ph.D Candidate @ MLD, CMU
  - Advisor: Jaime Carbonell, Alex Smola
  - Research: **Efficient** AI, ML, NLP.
    - Algorithm: Large scale optimization
    - Model: Machine reading comprehension and question answering

This Talk: Question Answering

Question Answering is important







Most NLP tasks can be formulated as QA

- Translation
- Summarization
- Natural language inference
- Sentiment analysis
- Semantic role labeling
- Relation extraction
- Goal oriented dialogue
- Semantic parsing
- Pronoun resolution

Natural Language Decathlon

[McCann, 2018]

How good are the current QA systems?

who won the 2014 world cup?

All News Images Shopping Videos More Settings Tools

About 960,000,000 results (0.83 seconds)

2014 FIFA World Cup / Champion

Germany national football team

Gotze wonder goal crowns **Germany** champions. Mario Gotze scored a stunning extra-time goal to settle the 2014 FIFA World Cup Final in Germany's favour, crowning the Europeans as champions with a 1-0 victory over Argentina at the Maracana. Jul 13, 2014

2014 FIFA World Cup Brazil™ - Matches - Germany-Argentina - FIFA ...
www.fifa.com/worldcup/matches/round=255959/match=300186501/index.html

Concrete Answer

is germany still in the world cup?

All News Images Shopping Videos More Settings Tools

About 1,530,000,000 results (0.50 seconds)

Germany national football team

MATCHES NEWS STANDINGS PLAYERS

World Cup · Group F · Matchday 1 of 3

Germany 0
Mexico 1
FT Sun, 6/17
3:53

World Cup · Group F · Matchday 2 of 3

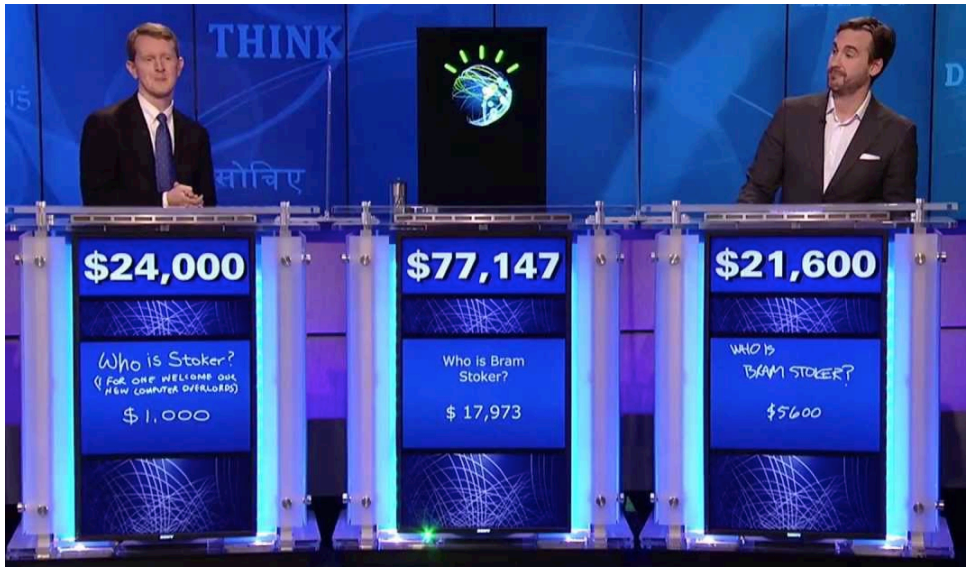
Germany 2
Sweden 1
FT Sat, 6/23
2:01

World Cup · Group F · Matchday 3 of 3

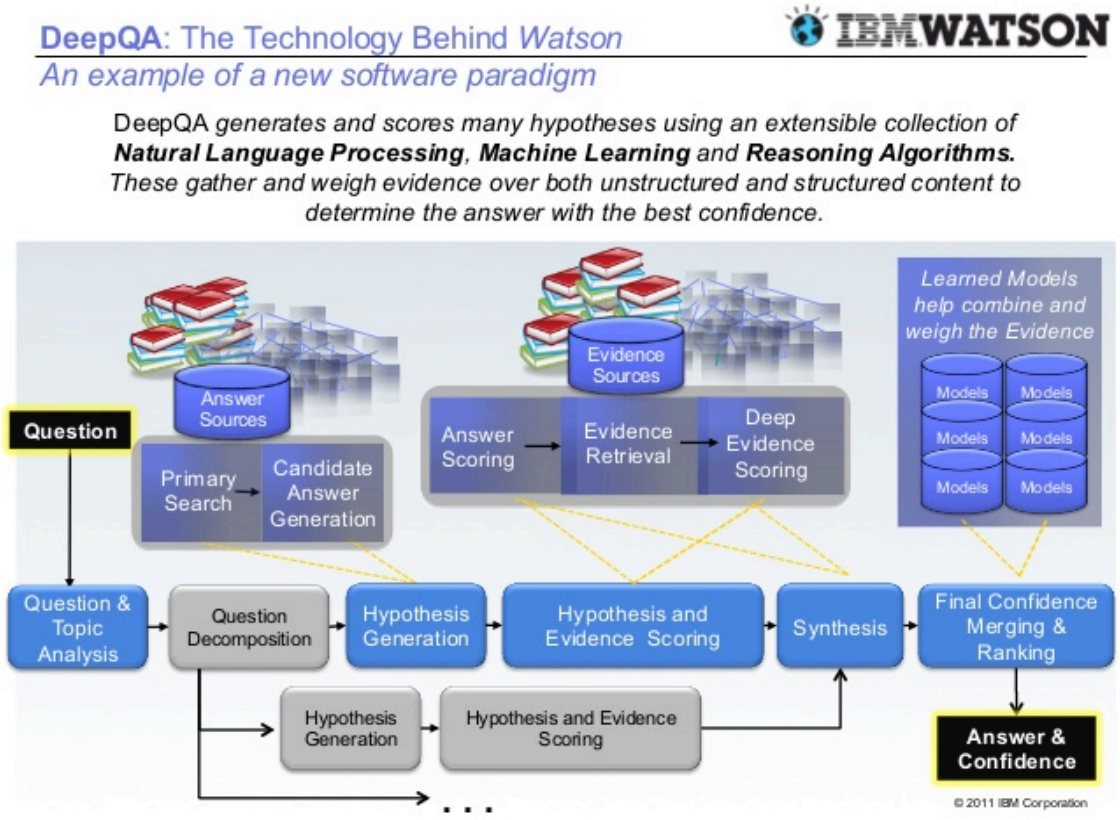
South Korea 2
Germany 0
FT Wed, 6/27
4:46

No clear answer

Early Success



Watson: complex multi-stage system



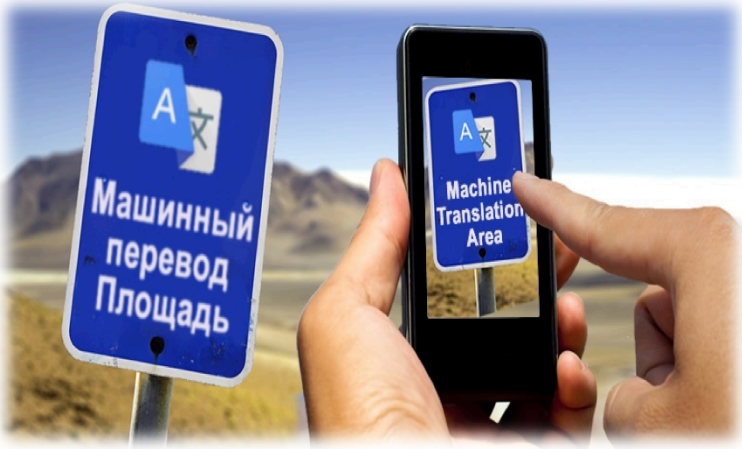
http://www.aaai.org/Magazine/Watson/watson.php

We Design End-to-End (Deep Learning) System

Deep Learning: Sparking the New Wave of AI

AlphaGo

Smart Assistant

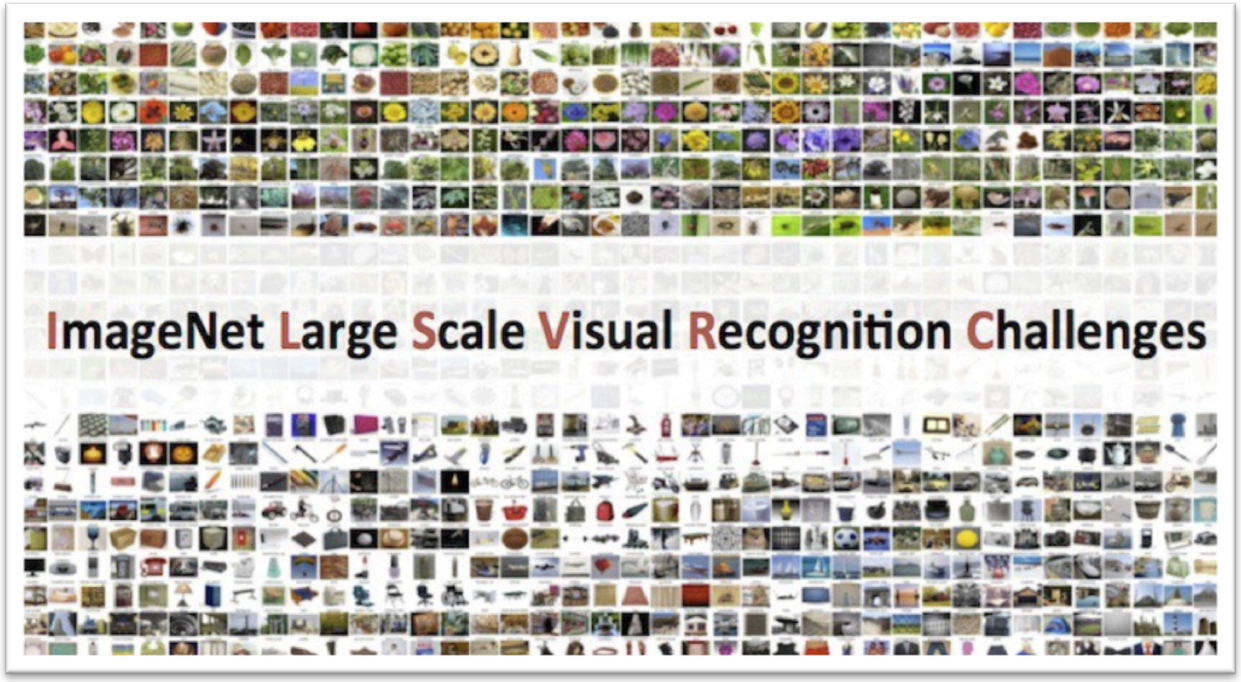


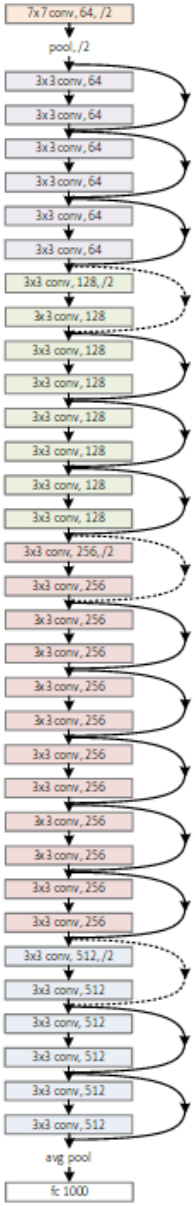
Machine Translation

Self-driving car

Bottleneck: Speed

ImageNet Classification

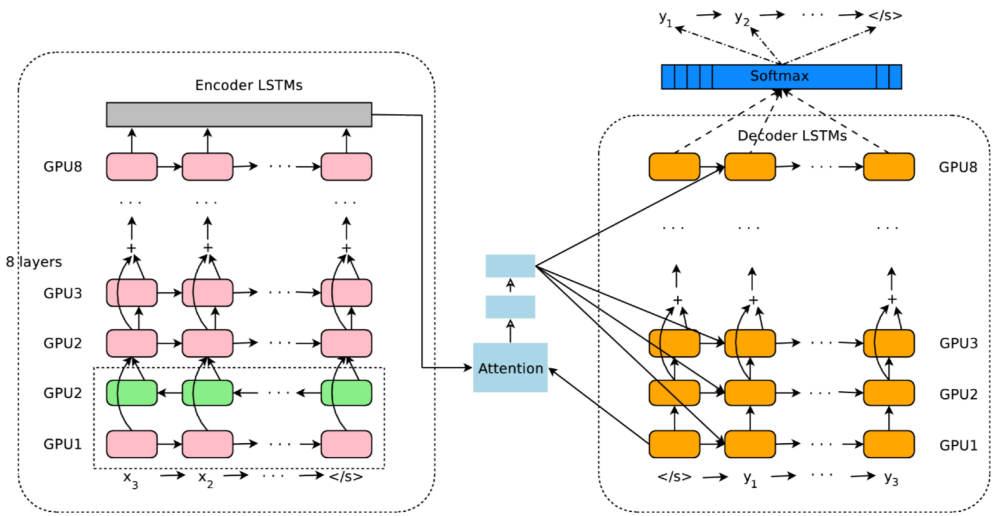




ResNet-152: 10 days to train

Bottleneck: Speed

Google Machine Translation



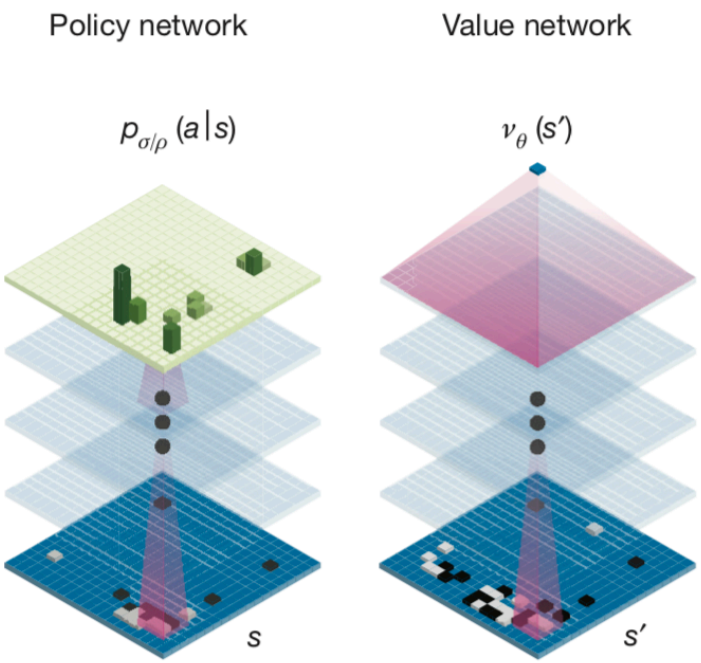
English → French

6 days to train

96 GPUs

Bottleneck: Speed

AlphaGo



4 weeks to train,
1920 CPUs and 280 GPUs

Also very slow for huge amount of texts

This review is from: Barefoot Running - The Movie: How to Run Light and Free by Getting in Touch with the Earth (NTSC/US Version) (DVD)

Wow! I just finished watching this and it is amazing. I'm completely blown away by the stunning cinematography. The movie was almost entirely filmed in Maui and I didn't want to take my eyes off the gorgeous colors on the screen. The beaches, trails, mountains and roads that Michael and Jessie run on throughout the movie, are absolutely breathtaking. There is even a bonus feature section that displays Michael's photography in a zen-like production of photographs that he captured during his runs.

The movie is designed as easy to follow, clear, concise chapters. It takes the viewer from the first step of shedding your shoes and putting your bare feet to the earth to running like a child again with a light foot and free spirit. I am not a runner yet this movie encourages me to take the journey of my first few years with out shoes. I plan to watch it over and over again while I practice what I've learned. This is an amazing resource guide put to life!

I think this movie is a terrific instructional video for experienced athletes too, who want to halt running injuries (that are discussed in detail in the movie) and add strength and speed to their prob[illegible]

This movie is filled with information such as (i) discussing "how to" go barefoot, (ii) what kind of shoe to purchase for the times when you must have a shoe, (iii) exercises to help your body heal before and between runs, and (iv) even ways of bringing the earth's energy into your body to help with overall health. It documents a serious accident that Michael had, which was the catalyst in getting him into barefoot running in the first place. Michael's story is incredibly inspirational and to see the strength of his muscles up close through the filming, after viewing the footage of what his body had been through prior to going barefoot, is fascinating.

Love it!!

Why Should We Care?

1. High turnaround time for experimentation.
2. Hard to train on large data.
3. Hard for real time applications.
4. High monetary cost.

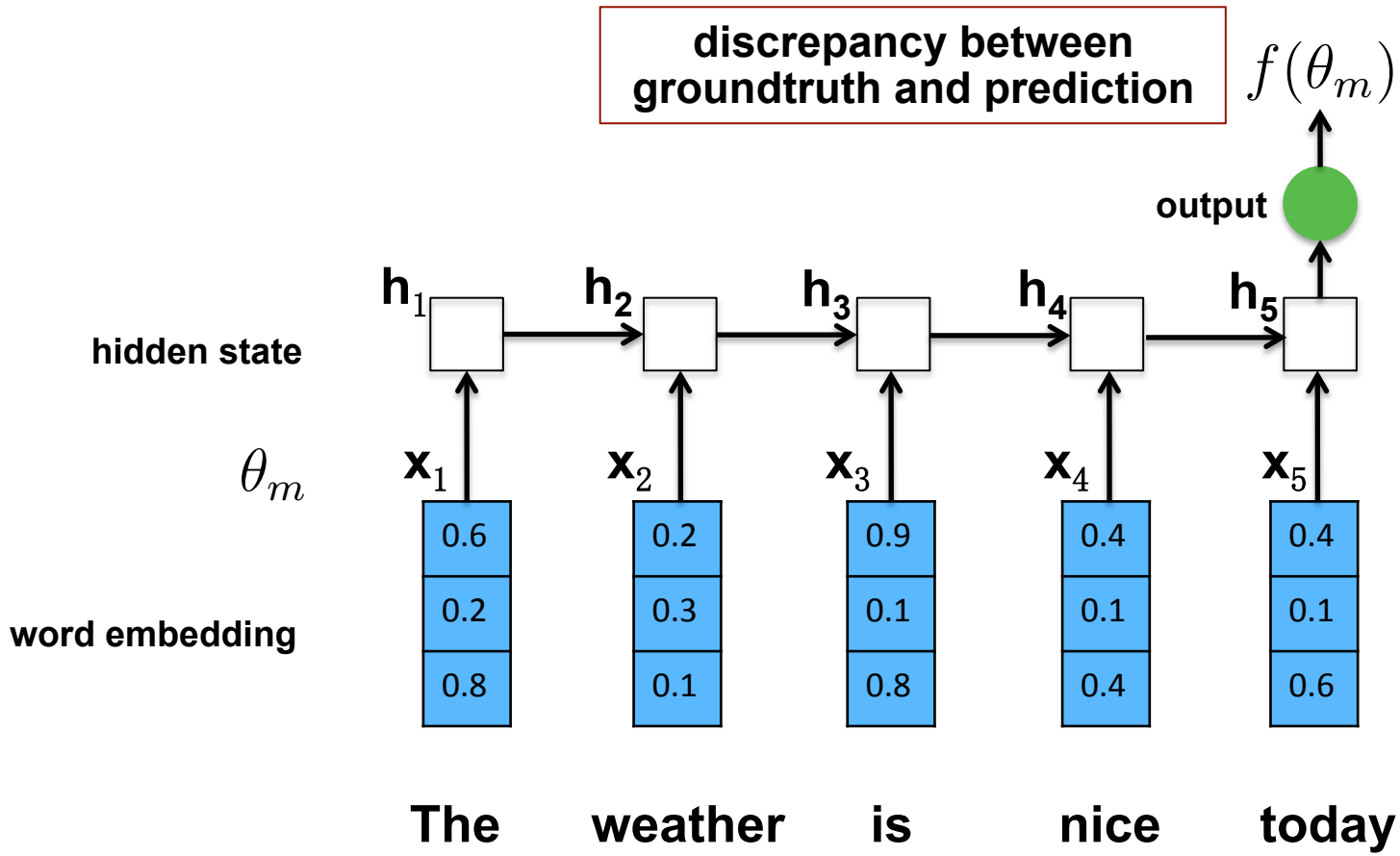
.....

This Talk: Efficient and Effective Deep Learning Models for Reading Comprehension

Roadmap

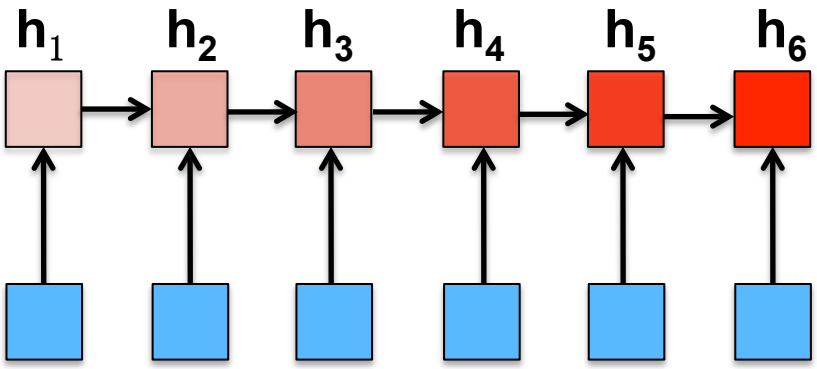
1. Skipping Irrelevant Information [ACL'17]
2. Discarding Recurrence [ICLR'18]
3. Future Work

Recurrent Neural Networks



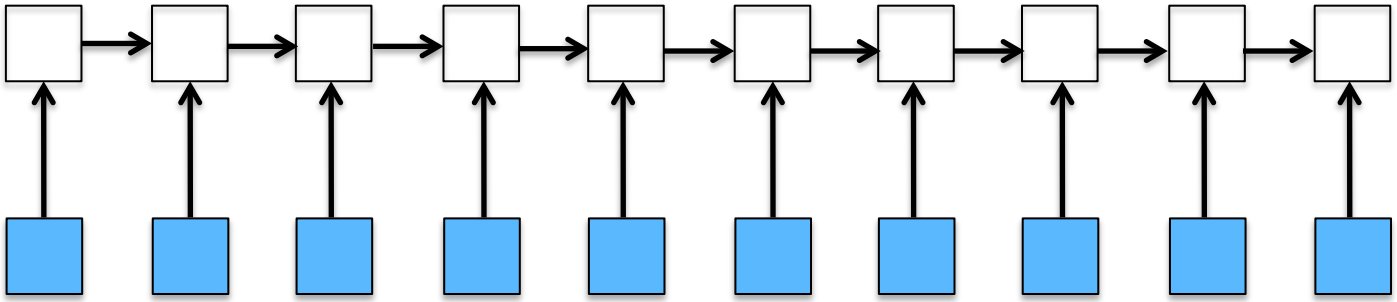
First Challenge: Hard to Capture Long Dependency

Being a long-time fan of Japanese film, I expected more than this. I can't really be bothered to write too much, as this movie is just so poor. The story might be the cutest romantic little something ever, pity I couldn't stand the awful acting, the mess they called pacing, and the standard "quirky" Japanese story. If you've noticed how many Japanese movies use characters, plots and twists that seem too "different", forcedly so, then steer clear of this movie. Seriously, a 12-year old could have told you how this movie was going to move along, and that's not a good thing in my book. Fans of "Beat" Takeshi: his part in this movie is not really more than a cameo, and unless you're a rabid fan, you don't need to suffer through this waste of film.



Second Challenge: Hard to Compute in Parallel

Strictly sequential



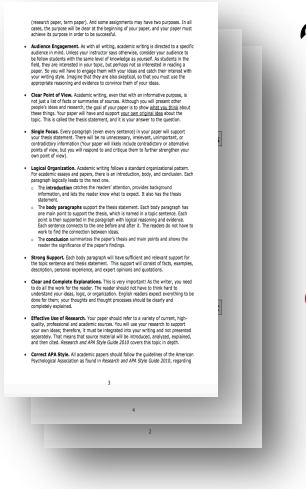
Reading Strategies

Q: What is dropout?

《Graphical Models》

《Statistics》

《Deep Learning》



Model

Algorithm

Regularization

Skim

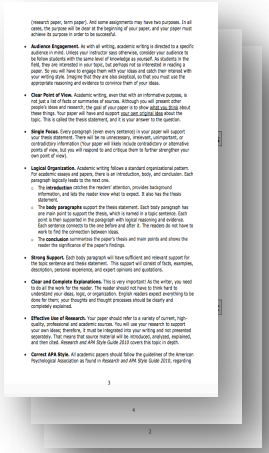
Reading Strategies

Q: What is dropout?

《Graphical Models》

《Statistics》

《Deep Learning》



Model

Algorithm

Regularization

Skim

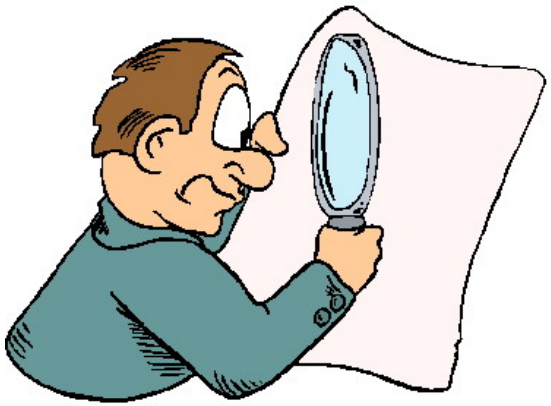
Inspect

Dropout Regularization For Neural Networks

Dropout is a regularization technique for neural network models proposed by Srivastava, et al. in their 2014 paper Dropout: A Simple Way to Prevent Neural Networks from Overfitting (download the PDF).

Dropout is a technique where randomly selected neurons are ignored during training. They are "dropped-out" randomly. This means that their contribution to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass.

As a neural network learns, neuron weights settle into their context within the network. Weights of neurons are tuned for specific features providing some specialization. Neighboring neurons become to rely on this specialization, which if taken too far can result in a fragile model too specialized to the training data. This reliant on context for a neuron during training is referred to complex co-adaptations.



Roadmap

1. Skipping Irrelevant Information [ACL'17]
2. Discarding Recurrence [ICLR'18]
3. Future Work

Learning to Skim Text

Yu et al.
ACL 2017

This review is from: **Barefoot Running - The Movie: How to Run Light and Free by Getting in Touch with the Earth (NTSC/US Version) (DVD)**

Wow! I just finished watching this and it is amazing. I'm completely blown away by the stunning cinematography. The movie was almost entirely filmed in Maui and I didn't want to take my eyes off the gorgeous colors on the screen. The beaches, trails, mountains and roads that Michael and Jessie run on throughout the movie, are absolutely breathtaking. There is even a bonus feature section that displays Michael's photography in a zen-like production of photographs that he captured during his runs.

The movie is designed as easy to follow, clear, concise chapters. It takes the viewer from the first step of shedding your shoes and putting your bare feet to the earth to running like a child again with a light foot and a free spirit. I am not a runner, yet this movie encourages me to take the journey of my first few yards without shoes. I plan to watch it over and over again while I practice what I've learned. This is an amazing resource guide put to life!

I think this movie is a terrific instructional video for experienced athletes too, who want to halt running injuries (that are discussed in detail in the movie) and find strength and speed that they probably never had.

This movie is filled with information such as (i) discussing "how to" go barefoot, (ii) what kind of shoe to purchase for the times when you must have a shoe, (iii) exercises to h[illegible] before [illegible] bringing th[illegible] a serious accident that Michael had, which was the catalyst in getting him into barefoot running in the first place. Michael's story is incredibly inspirational and to see the strength of his muscles up close through the filming, after viewing the footage of what his body had been through prior to going barefoot, is fascinating.

Love it!!

Texts are usually redundant …

Humans don't always read word by word

[Yu et al., ACL'17]

Sentiment Analysis

Positive or negative lengthy movie review?

Being a long-time fan of Japanese film, I expected more than this. I can't really be bothered to write too much, as this movie is just so poor. The story might be the cutest romantic little something ever, pity I couldn't stand the awful acting, the mess they called pacing, and the standard "quirky" Japanese story. If you've noticed how many Japanese movies use characters, plots and twists that seem too "different", forcedly so, then steer clear of this movie. Seriously, a 12-year old could have told you how this movie was going to move along, and that's not a good thing in my book. Fans of "Beat" Takeshi: his part in this movie is not really more than a cameo, and unless you're a rabid fan, you don't need to suffer through this waste of film.

[Yu et al., ACL'17]

Sentiment Analysis

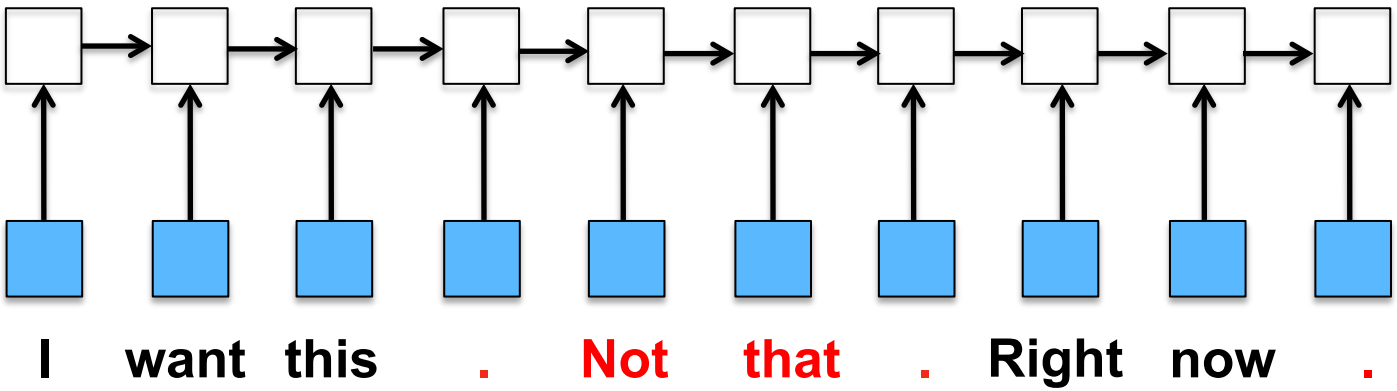
Positive or negative lengthy movie review?

Negative

- Being a long-time fan of Japanese film, I expected more than this. I can't really be bothered to write too much, as this movie is just so poor. The story might be the cutest romantic little something ever, pity I couldn't stand the awful acting, the mess they called pacing, and the standard "quirky" Japanese story. If you've noticed how many Japanese movies use characters, plots and twists that seem too "different", forcedly so, then steer clear of this movie. Seriously, a 12-year old could have told you how this movie was going to move along, and that's not a good thing in my book. Fans of "Beat" Takeshi: his part in this movie is not really more than a cameo, and unless you're a rabid fan, you don't need to suffer through this waste of film. 2/10

[Yu et al., ACL'17]

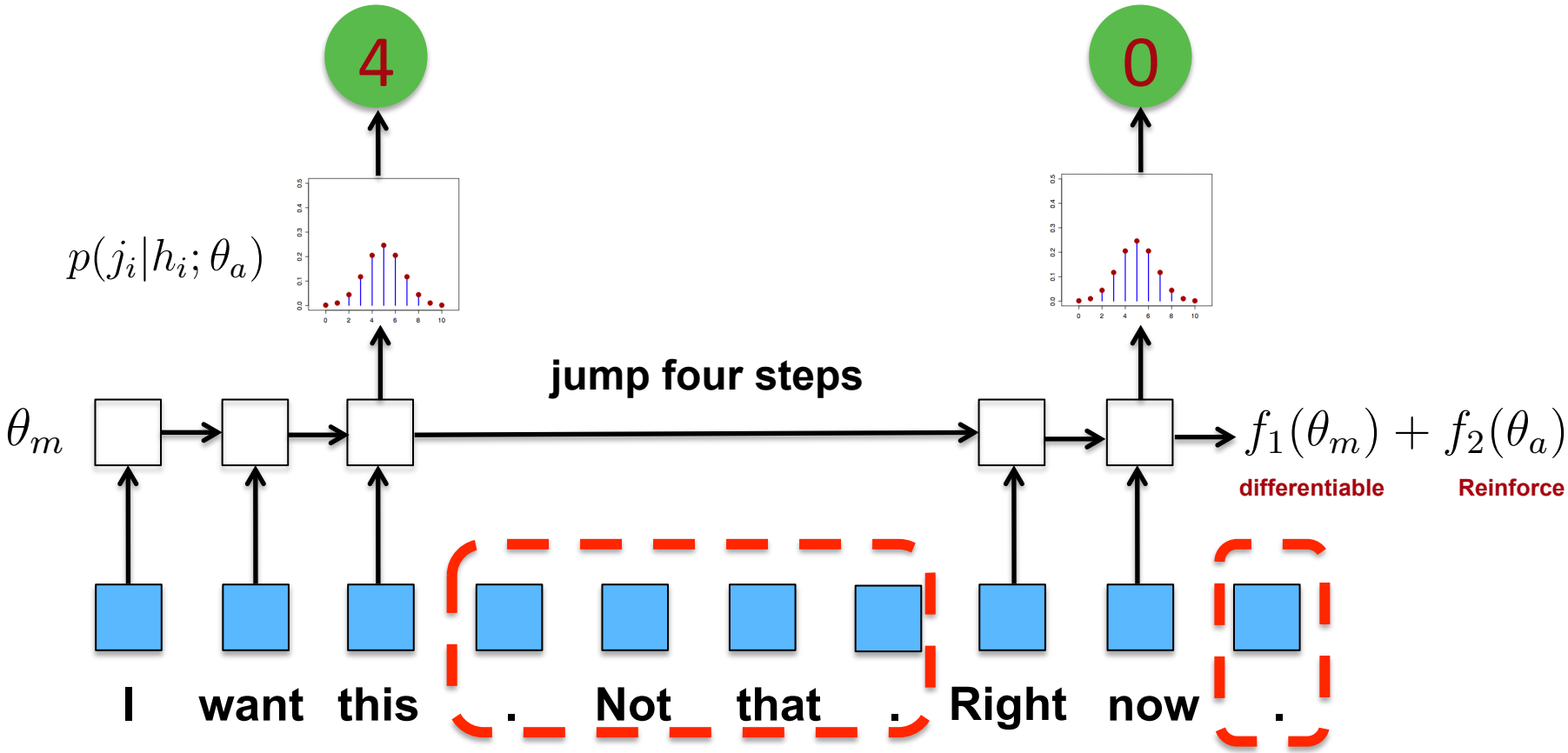
Learning to Skim Text



[Yu et al., ACL'17]

Learning to Skim Text

LSTM-Jump



[Yu et al., ACL'17]

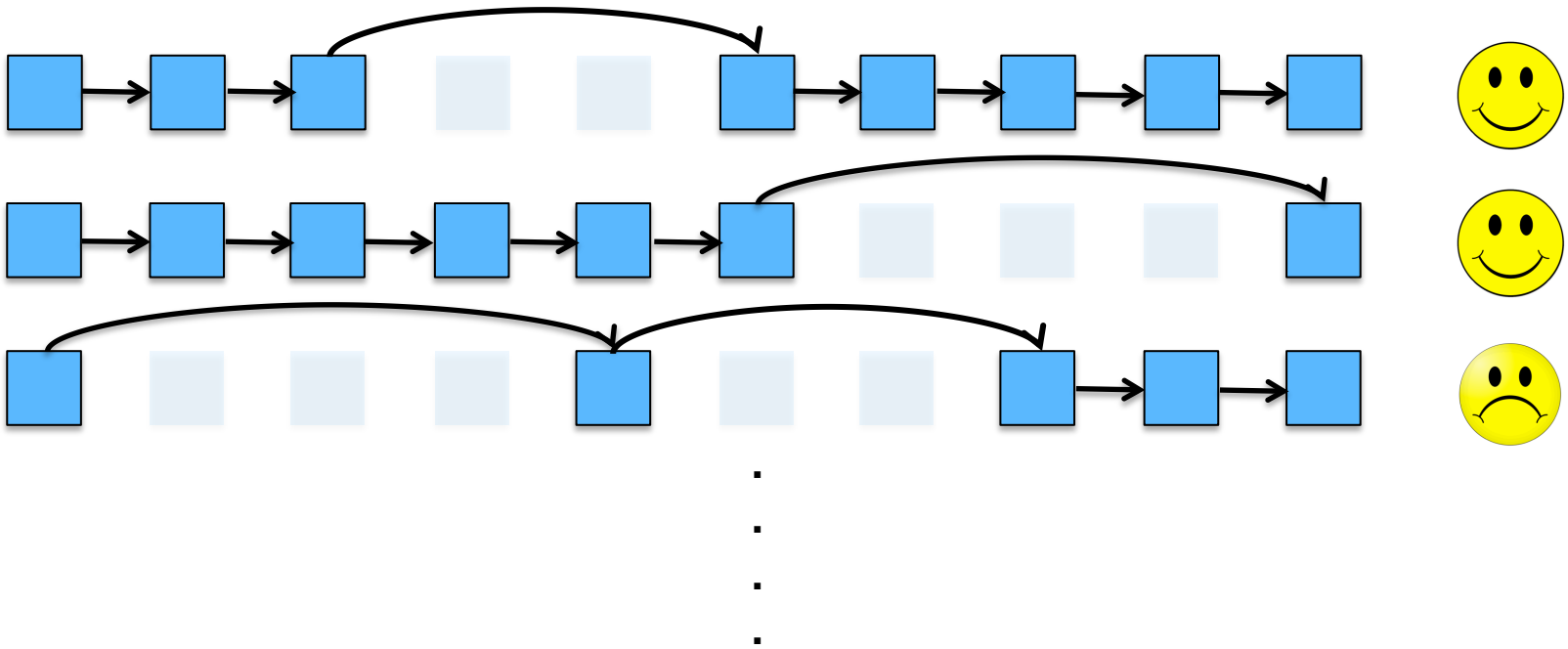
Learning to Skim Text

$$\min_{\theta_m,\ \theta_a} f_1(\theta_m) + \boxed{f_2(\theta_a)}$$

differentiable

gradient?

policy gradient: REINFORCE



$$f_2(\theta_a) = -\mathbb{E}_{p(N;\theta_a)}[R]$$

$$R = \begin{cases} 1 \\ -1 \end{cases}$$

[Yu et al., ACL'17]

Number prediction

Input: $\{x_t\}_{t=0}^{T}$

Output: x_{x_0}

1. Input: 4, 5, 1, 7, 6, 2 Output: 6

2. Input: 2, 4, 9, 4, 5, 6 Output: 9

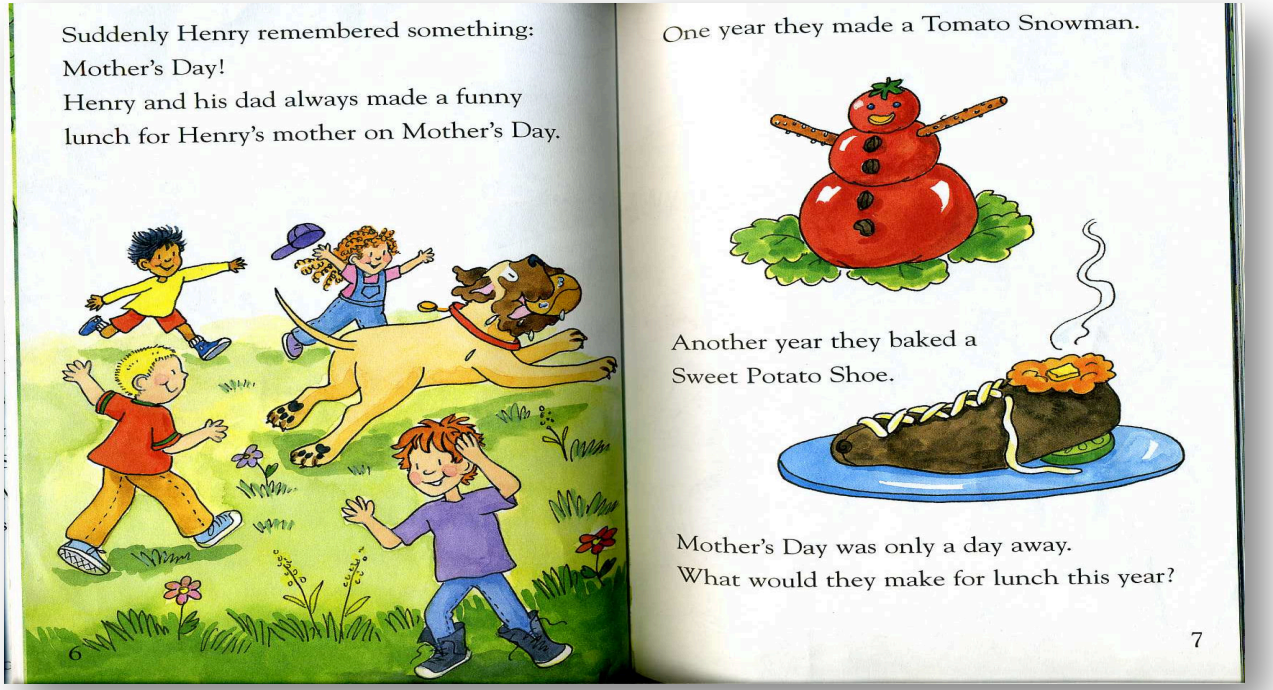
	Accuracy	Token read	Test Time	Speedup
LSTM-Jump	**90%**	**3.0**	**19 s**	**66x**
LSTM	80%	1,000	1,250 s	-

Length = 1,000

[Yu et al., ACL'17]

QA on Children's Book Test

- Context: 20 contiguous sentences
- Query: A sentence with a word removed.
- Task: Fill the blank from 10 candidates.



[Yu et al., ACL'17]

Query: Yes, I call (__) a nuisance.

Candidates:

1. Christmas
2. boys
3. day
4. dinner
5. half
6. interest
7. rest
8. stockings
9. things
10. uncles

1. But to you and me it would have looked just as it did to Cousin Myra – a very discontented and unbecoming scowl.

2. “I’m awfully glad to see you, Cousin Myra, ”explained Frank carefully.

3. “But Christmas is just a bore – a regular bore.”

4. That was what Uncle Edgar called things that didn’t interest him, so that Frank felt pretty sure of.

5. Nevertheless, he wondered uncomfortably what made Cousin Myra smile so queerly.

6. “Why, how dreadful!”

7. She said brightly.

8. “I thought all boys and girls looked upon Christmas as the very best time in the year.”

9. “We don’t, ”said Frank gloomily.

10. “It’s just the same old thing year in and year out.

11. We know just exactly what is going to happen.

12. We even know pretty well what presents we are going to get.

13. And Christmas Day itself is always the same.

14. We’ll get up in the morning , and our stockings will be full of things.

15. Then there ’s dinner.

16. It ’s always so poky.

17. And all the uncles and aunts come to dinner – just the same old crowd, every year.

18. Aunt Desda always says, ‘Why, Frankie, how you have grown!’

19. She knows I hate to be called Frankie.

20. And after dinner they’ll sit round and talk the rest of the day, and that’s all.

[Yu et al., ACL’17]

Query: Yes, I call (__) a nuisance.

Candidates:

1. Christmas
2. boys
3. day
4. dinner
5. half
6. interest
7. rest
8. stockings
9. things
10. uncles

1. But to you and me it would have looked just as it did to Cousin Myra – a very discontented and unbecoming scowl.

2. "I'm awfully glad to see you, Cousin Myra, "explained Frank carefully.

3. "But Christmas is just a bore – a regular bore."

4. That was what Uncle Edgar called things that didn't interest him, so that Frank felt pretty sure of.

5. Nevertheless, he wondered uncomfortably what made Cousin Myra smile so queerly.

6. "Why, how dreadful!"

7. She said brightly.

8. "I thought all boys and girls looked upon Christmas as the very best time in the year."

9. "We don't, "said Frank gloomily.

10. "It's just the same old thing year in and year out.

11. We know just exactly what is going to happen.

12. We even know pretty well what presents we are going to get.

13. And Christmas Day itself is always the same.

14. We'll get up in the morning , and our stockings will be full of things.

15. Then there 's dinner.

16. It 's always so poky.

17. And all the uncles and aunts come to dinner – just the same old crowd, every year.

18. Aunt Desda always says, 'Why, Frankie, how you have grown!'

19. She knows I hate to be called Frankie.

20. And after dinner they'll sit round and talk the rest of the day, and that's all.

[Yu et al., ACL'17]

Query: Yes, I call (__) a nuisance.

Candidates:

1. Christmas
2. boys
3. day
4. dinner
5. half
6. interest
7. rest
8. stockings
9. things
10. uncles

1. But to you and me it would have looked just as it did to Cousin Myra – a very discontented and unbecoming scowl.

2. "I'm awfully glad to see you, Cousin Myra, "explained Frank carefully.

3. "But Christmas is just a bore – a regular bore."

4. That was what Uncle Edgar called things that didn't interest him, so that Frank felt pretty sure of.

5. Nevertheless, he wondered uncomfortably what made Cousin Myra smile so queerly.

6. "Why, how dreadful!"

7. She said brightly.

8. "I thought all boys and girls looked upon Christmas as the very best time in the year."

9. "We don't, "said Frank gloomily.

10. "It's just the same old thing year in and year out.

11. We know just exactly what is going to happen.

12. We even know pretty well what presents we are going to get.

13. And Christmas Day itself is always the same.

14. We'll get up in the morning , and our stockings will be full of things.

15. Then there 's dinner.

16. It 's always so poky.

17. And all the uncles and aunts come to dinner – just the same old crowd, every year.

18. Aunt Desda always says, 'Why, Frankie, how you have grown!'

19. She knows I hate to be called Frankie.

20. And after dinner they'll sit round and talk the rest of the day, and that's all.

[Yu et al., ACL'17]

Query: Yes, I call (__) a nuisance.

Candidates:

1. Christmas
2. boys
3. day
4. dinner
5. half
6. interest
7. rest
8. stockings
9. things
10. uncles

1. But to you and me it would have looked just as it did to Cousin Myra – a very discontented and unbecoming scowl.

2. "I'm awfully glad to see you, Cousin Myra, "explained Frank carefully.

3. "But Christmas is just a bore – a regular bore."

4. That was what Uncle Edgar called things that didn't interest him, so that Frank felt pretty sure of.

5. Nevertheless, he wondered uncomfortably what made Cousin Myra smile so queerly.

6. "Why, how dreadful!"

7. She said brightly.

8. "I thought all boys and girls looked upon Christmas as the very best time in the year."

9. "We don't, "said Frank gloomily.

10. "It's just the same old thing year in and year out.

11. We know just exactly what is going to happen.

12. We even know pretty well what presents we are going to get.

13. And Christmas Day itself is always the same.

14. We'll get up in the morning , and our stockings will be full of things.

15. Then there 's dinner.

16. It 's always so poky.

17. And all the uncles and aunts come to dinner – just the same old crowd, every year.

18. Aunt Desda always says, 'Why, Frankie, how you have grown!'

19. She knows I hate to be called Frankie.

20. And after dinner they'll sit round and talk the rest of the day, and that's all.

[Yu et al., ACL'17]

Result

Children's Book Test

	Accuracy	Infer Time	Speedup
LSTM-Jump	**45.2%**	**20 s**	**6x**
LSTM	43.8%	124 s	-

Other Classification Problems

Dataset	Rotten Tomatoes	IMDB	AG News
LSTM-Jump	**2x**	**2.5x**	**1.3x**

Similar or better accuracy as LSTM

[Yu et al., ACL'17]

So Far

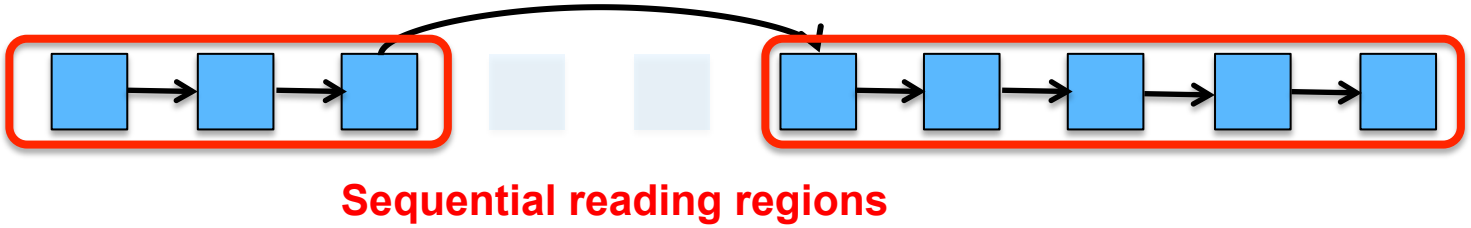
RNN / LSTM

LSTM-Jump

1. Skips redundant info

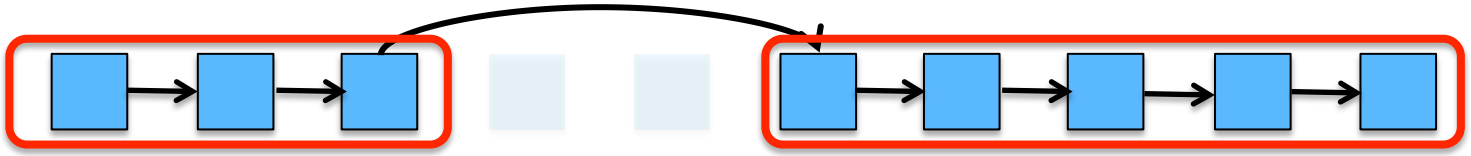
2. Works well on simple tasks

1. Still reads sequentially between jumps



Hard to Compute in Parallel

1. Still reads sequentially between jumps



Sequential reading regions

2. Might not handle complicated tasks

Roadmap

1. Skipping Irrelevant Information [ACL'17]
2. Discarding Recurrence [ICLR'18]
3. Future Work

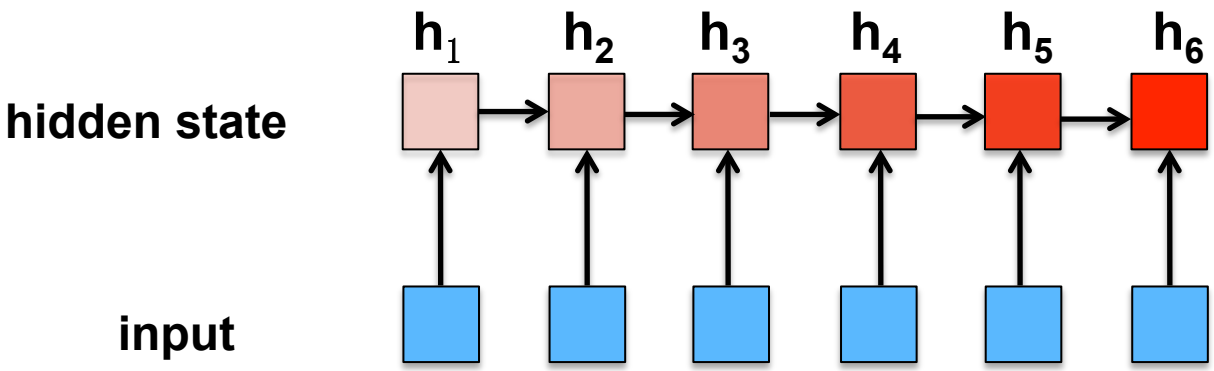
QANet: Combining Local Convolution and Global Self-Attention for Reading Comprehension

Yu et al.
ICLR 2018

What do RNNs capture?

2. global interaction

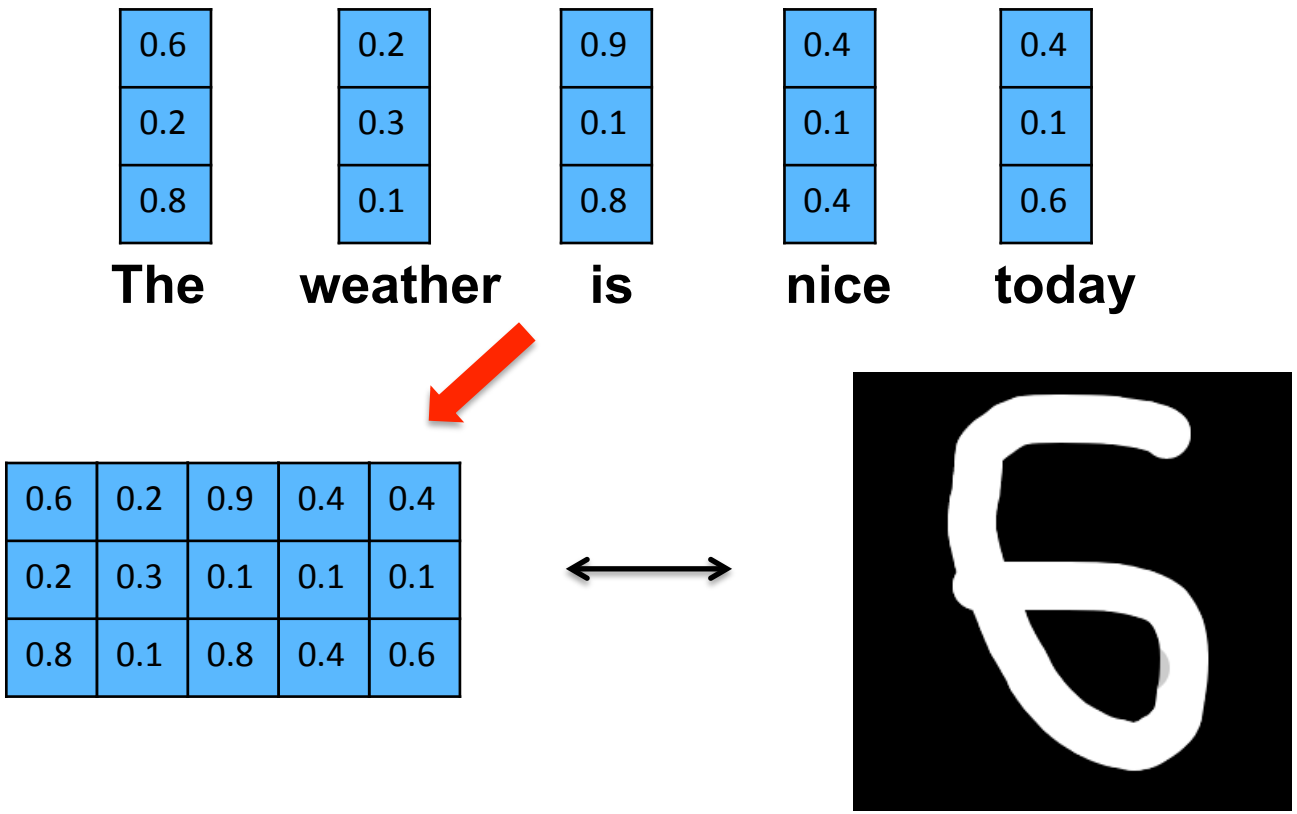
1. local context



3. Temporal info

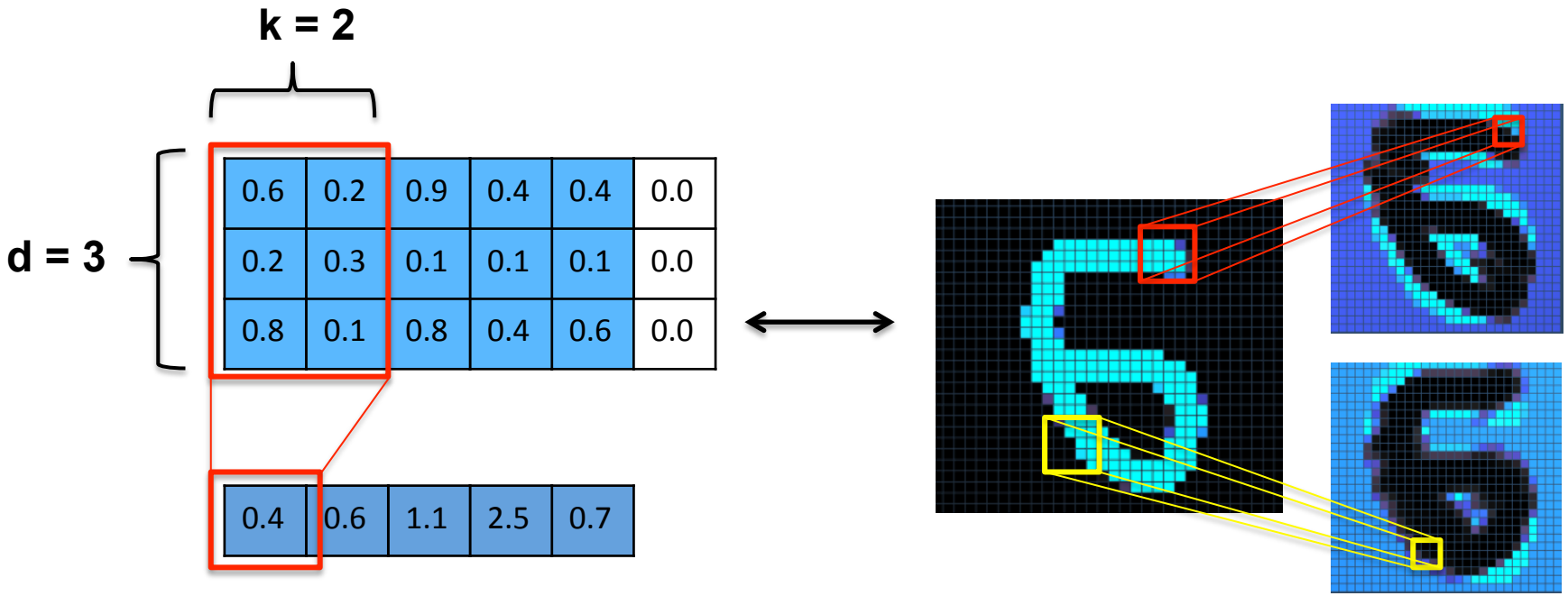
Efficient substitution?

Convolution: Capturing Local Context



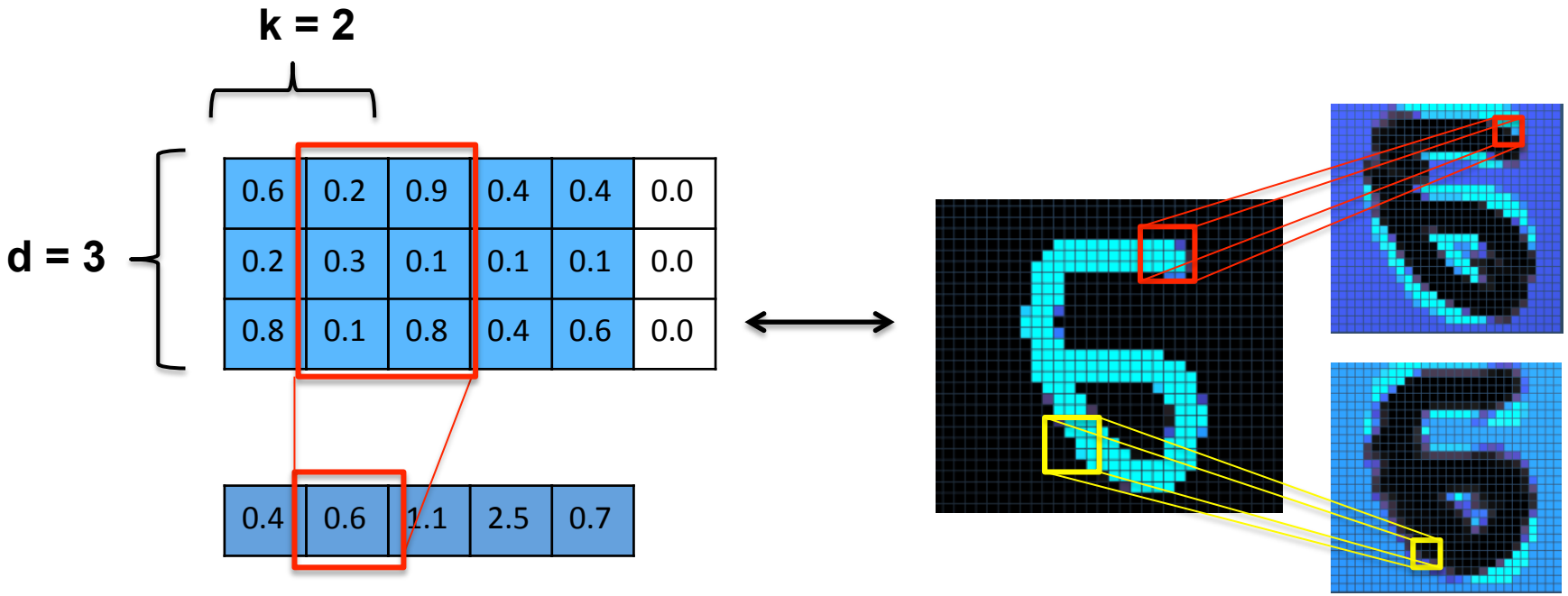
[Kim, EMNLP'14]; [Johnson and Zhang, NAACL'15]; [Conneau, EACL'17]; [Gehring et al., ICML'17]

Convolution: Capturing Local Context



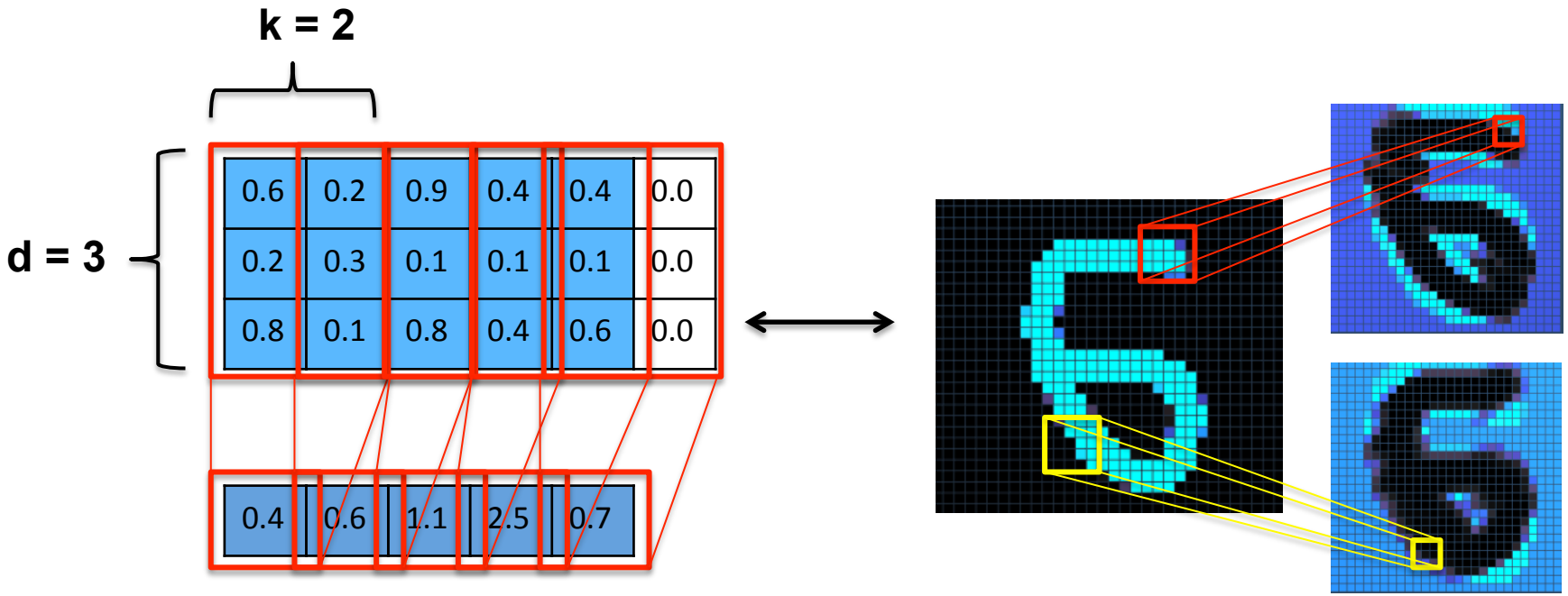
[Kim, EMNLP'14]; [Johnson and Zhang, NAACL'15]; [Conneau, EACL'17]; [Gehring et al., ICML'17]

Convolution: Capturing Local Context



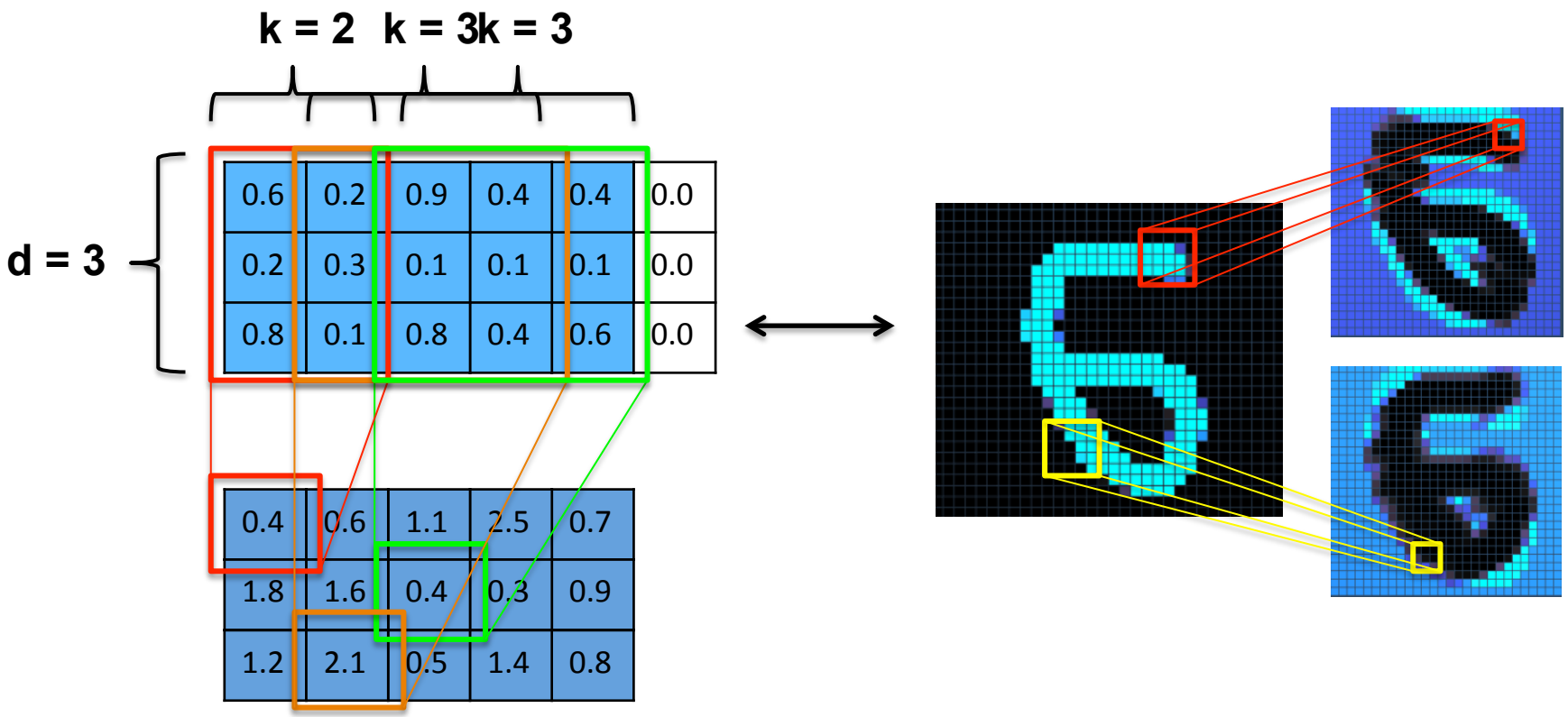
[Kim, EMNLP'14]; [Johnson and Zhang, NAACL'15]; [Conneau, EACL'17]; [Gehring et al., ICML'17]

Convolution: Capturing Local Context



[Kim, EMNLP'14]; [Johnson and Zhang, NAACL'15]; [Conneau, EACL'17]; [Gehring et al., ICML'17]

Convolution: Capturing Local Context

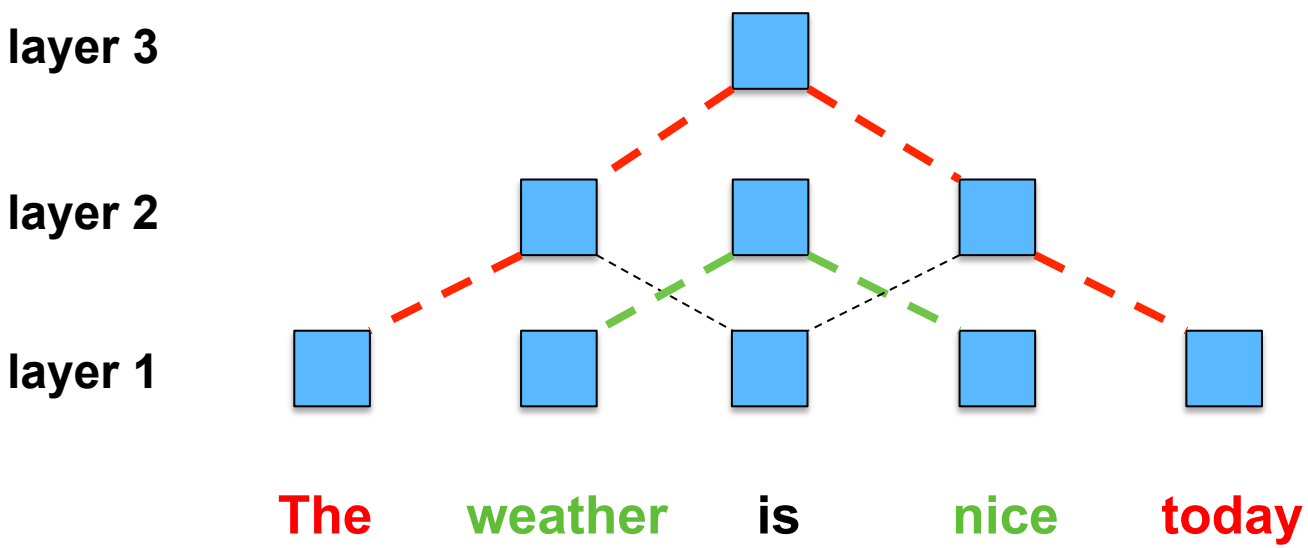


k-gram features

Fully parallel!

[Kim, EMNLP'14]; [Johnson and Zhang, NAACL'15]; [Conneau, EACL'17]; [Gehring et al., ICML'17]

How about Global Interaction?



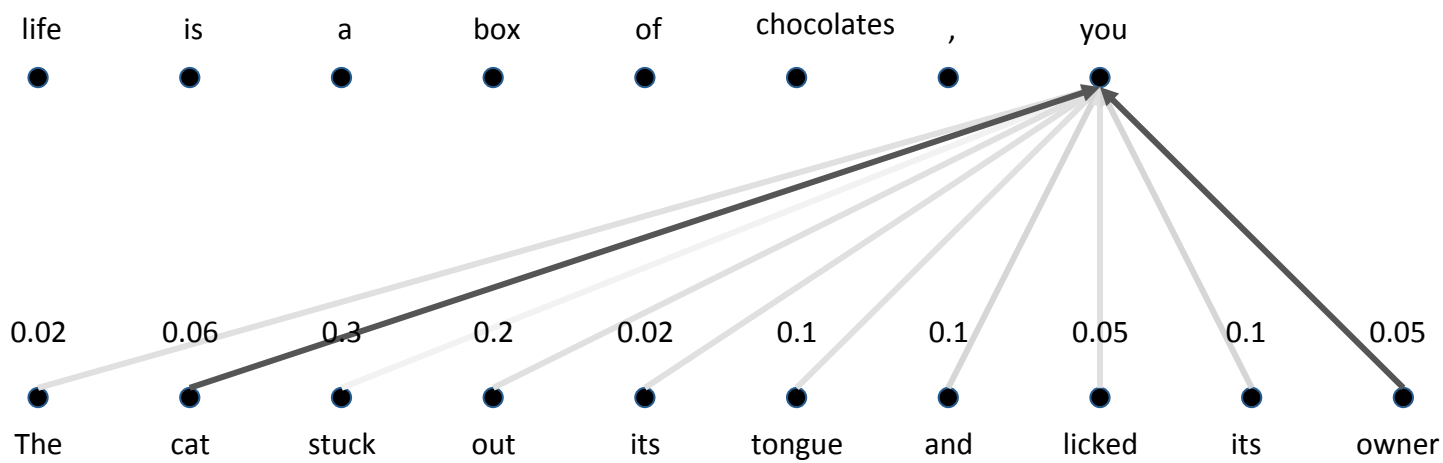
N: Seq length.
k: Filter size.

1. May need $O(\log_k N)$ layers
2. Interaction may become weaker

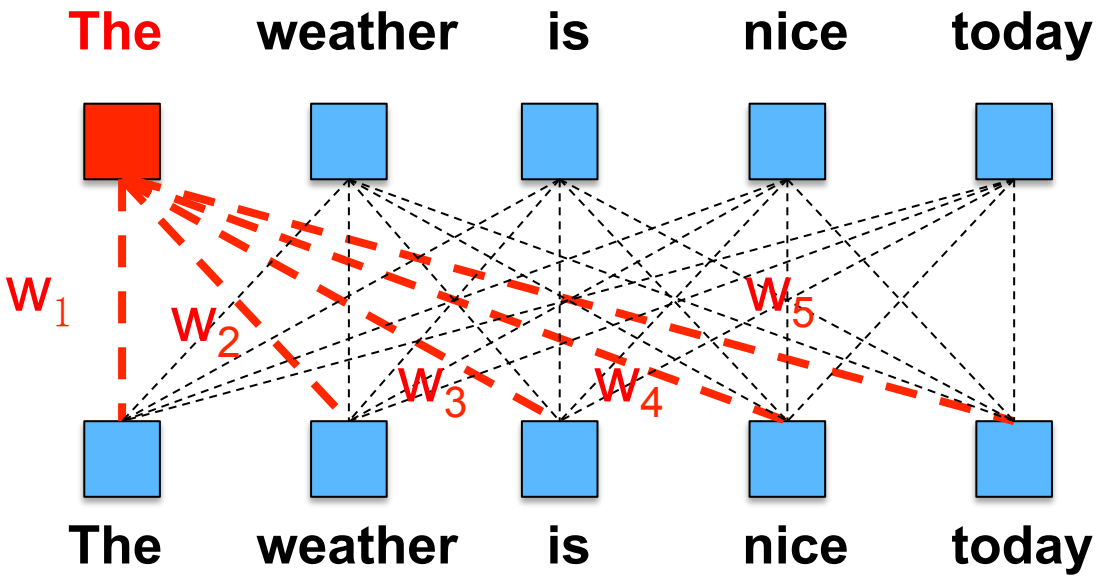
[Yu et al., ICLR'18]

Self-Attention: Global Interaction

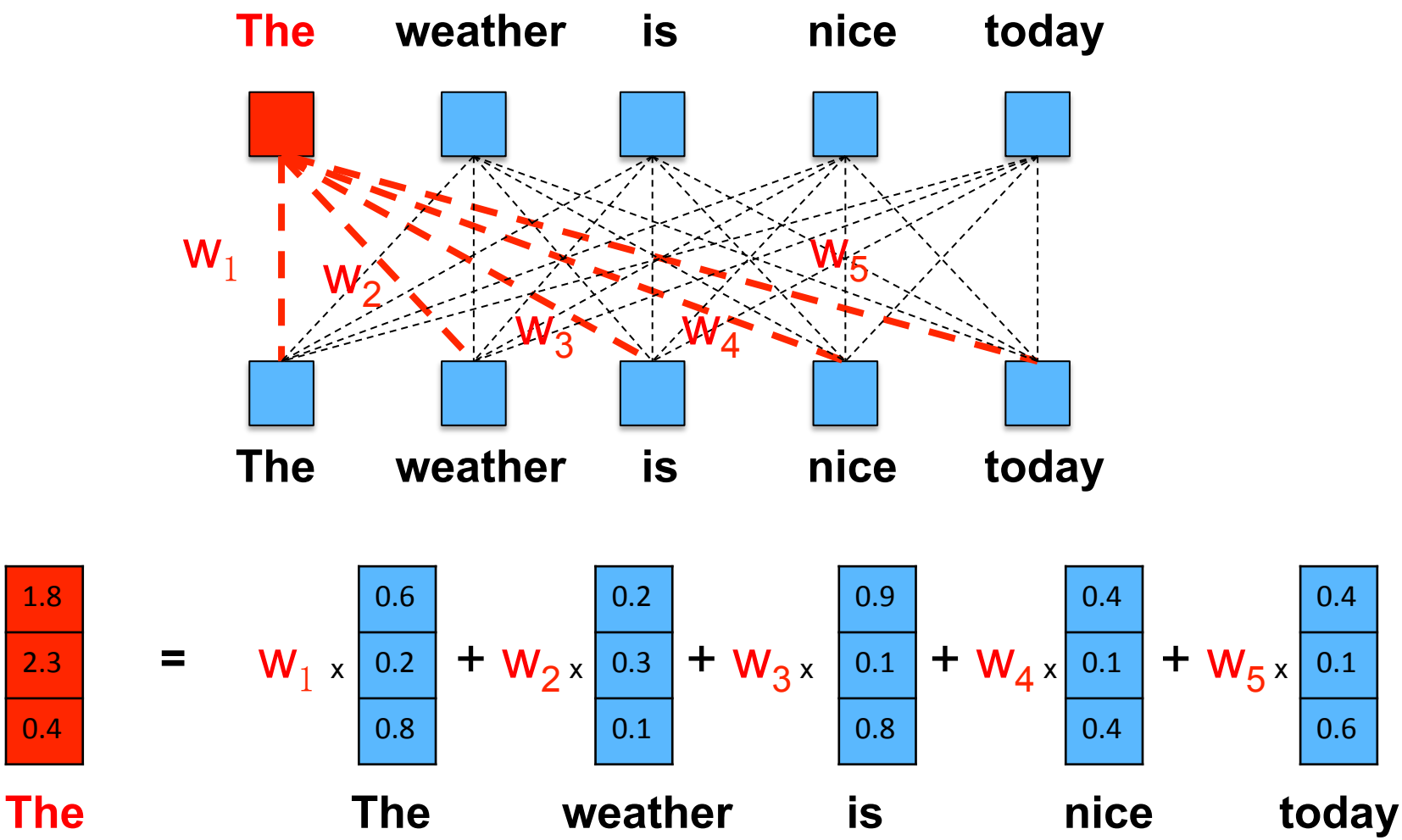
Attention: interaction of two seqs with weighted average



[Bahdanau et al., ICLR'15]; [Luong et al., EMNLP'15]

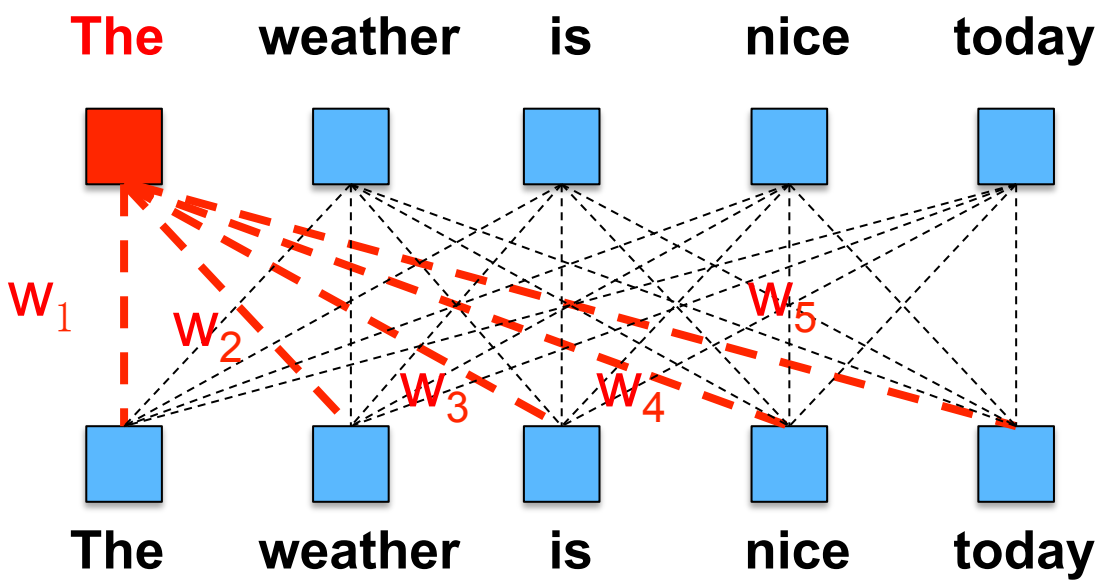


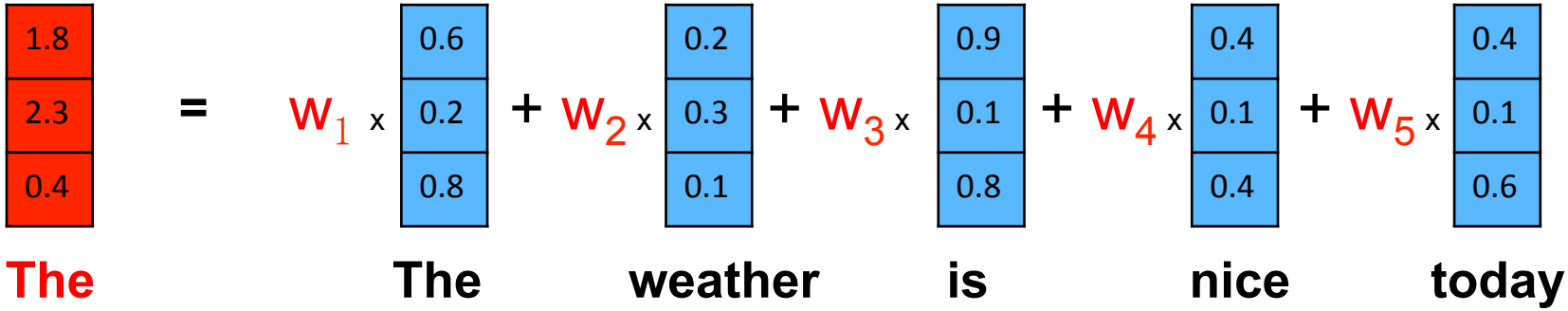
[Vaswani et al., NIPS'17]

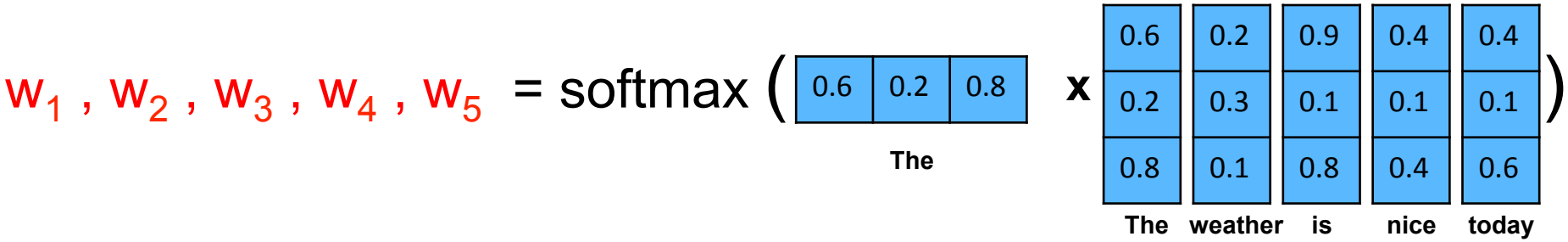


The weather is nice today

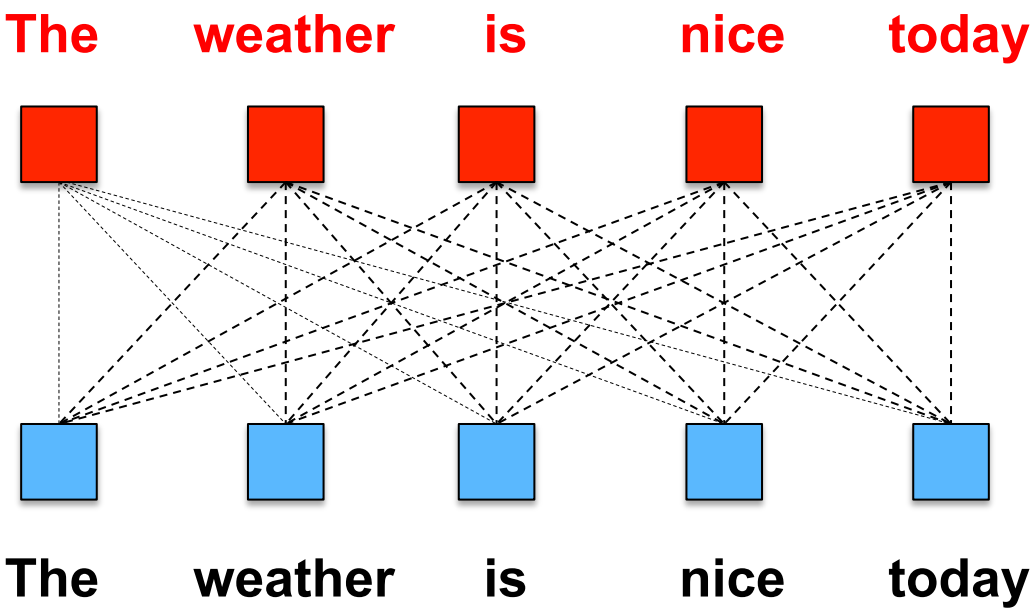
[Vaswani et al., NIPS'17]







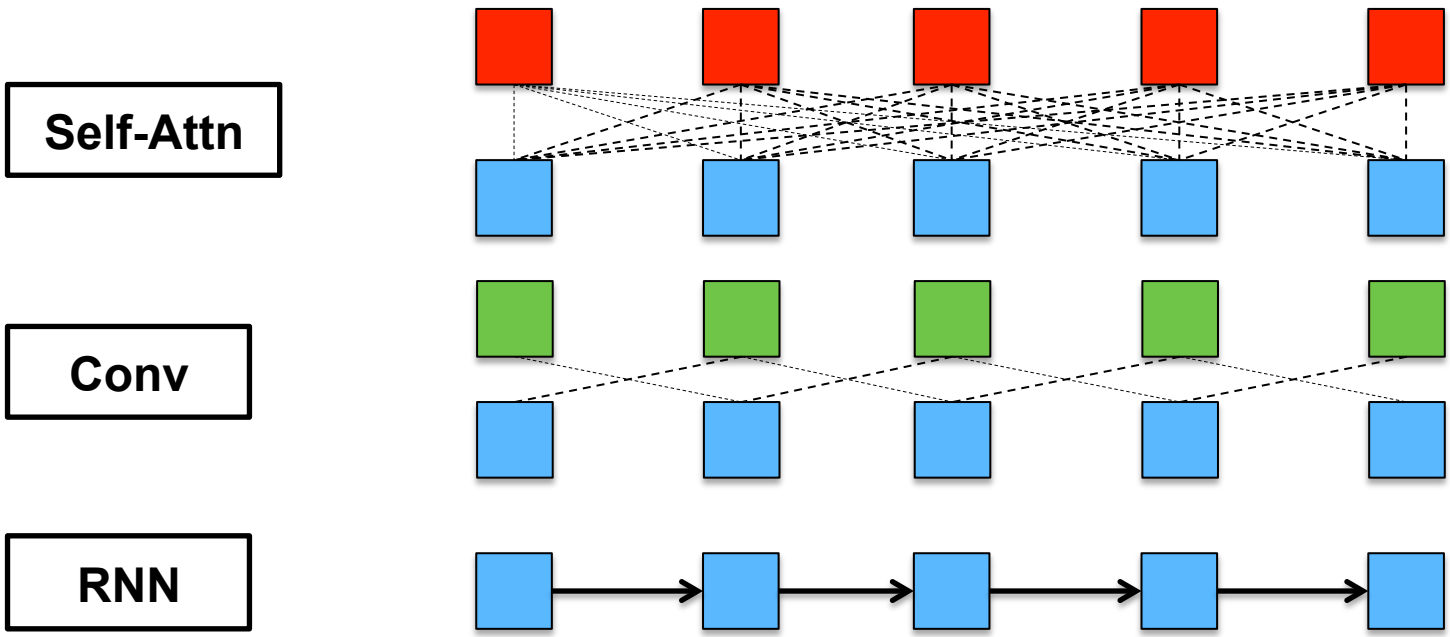
[Vaswani et al., NIPS'17]



Fully parallel!

[Vaswani et al., NIPS'17]

Complexity

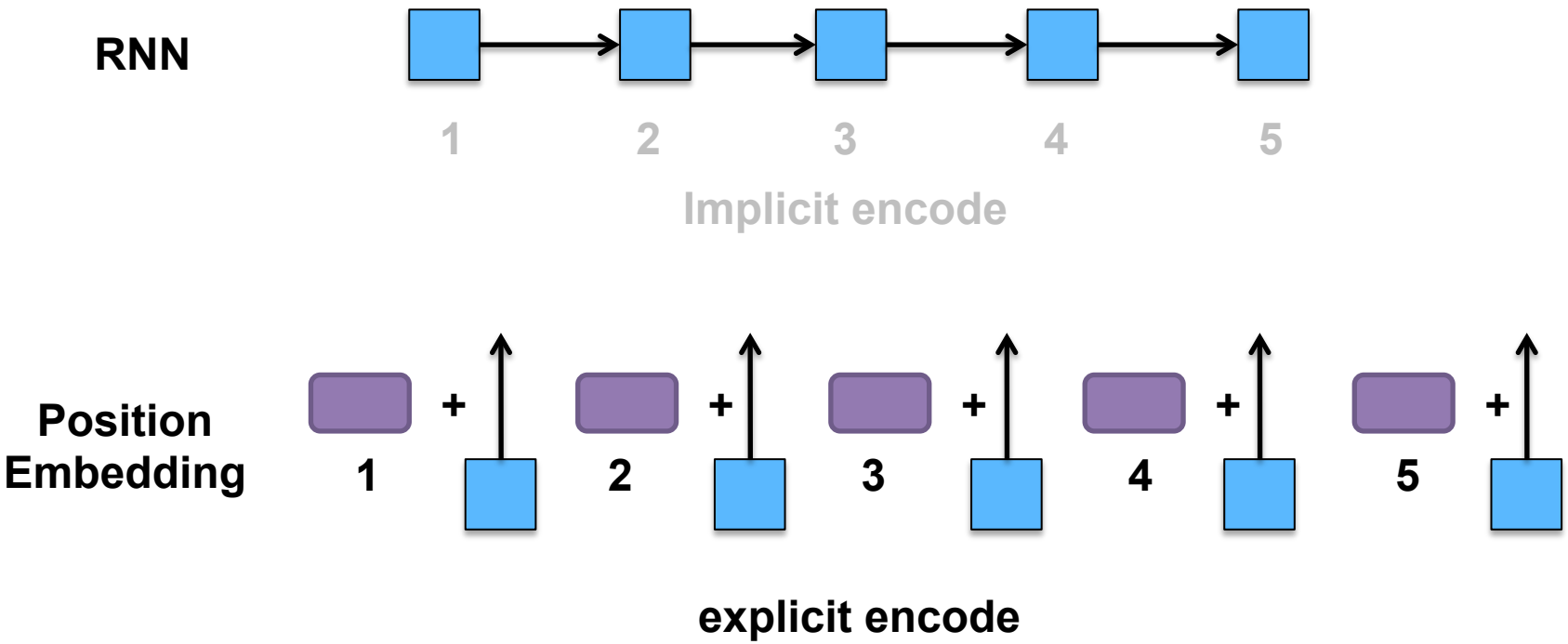


N: Seq length.
d: Dim. (N > d)
k: Filter size.

	Per Unit	Total Per Layer	Sequential Op (Path Memory)
Self-Attn	$O(Nd)$	$O(N^2d)$	$O(1)$
Conv	$O(kd^2)$	$O(kNd^2)$	$O(1)$
RNN	$O(d^2)$	$O(Nd^2)$	$O(N)$

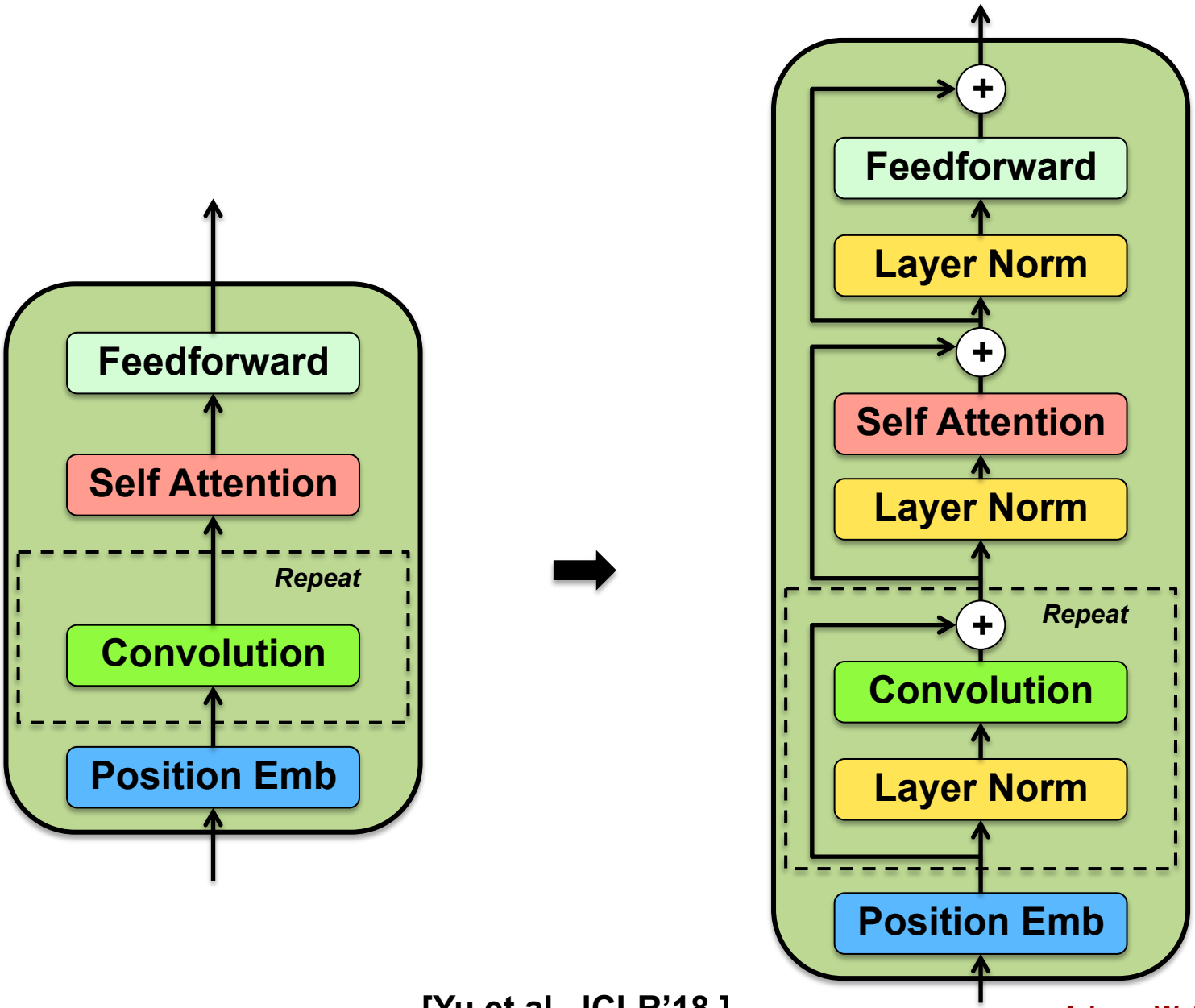
[Yu et al., ICLR'18]

Explicitly Encode Temporal Info



[Vaswani et al., NIPS'17]

QANet Encoder Block



[Yu et al., ICLR'18]

We don’t have recurrence anymore

[Yu et al., ICLR’18]

Lots of QA datasets Available

SQuAD
The Stanford Question Answering Dataset

SQuAD2.0
The Stanford Question Answering Dataset

QuAC
Question Answering in Context

HotpotQA
A Dataset for Diverse, Explainable Multi-hop Question Answering

CoQA
A Conversational Question Answering Challenge

RACE

Narrative QA

MS MARCO

TriviaQA

Stanford Question Answer Dataset (SQuAD)

Data: Crowdsourced 100k question-answer pairs on 500 Wikipedia articles.

Passage: In education, teachers facilitate student learning, often in a school or academy or perhaps in another environment such as outdoors. A teacher who teaches on an individual basis may be described as a tutor.

Question: **What is the role of teachers in education?**

Groundtruth: **facilitate student learning**

Prediction 1: **facilitate student learning** — **EM = 1, F1 = 1**

Prediction 2: **student learning** — **EM = 0, F1 = 0.8**

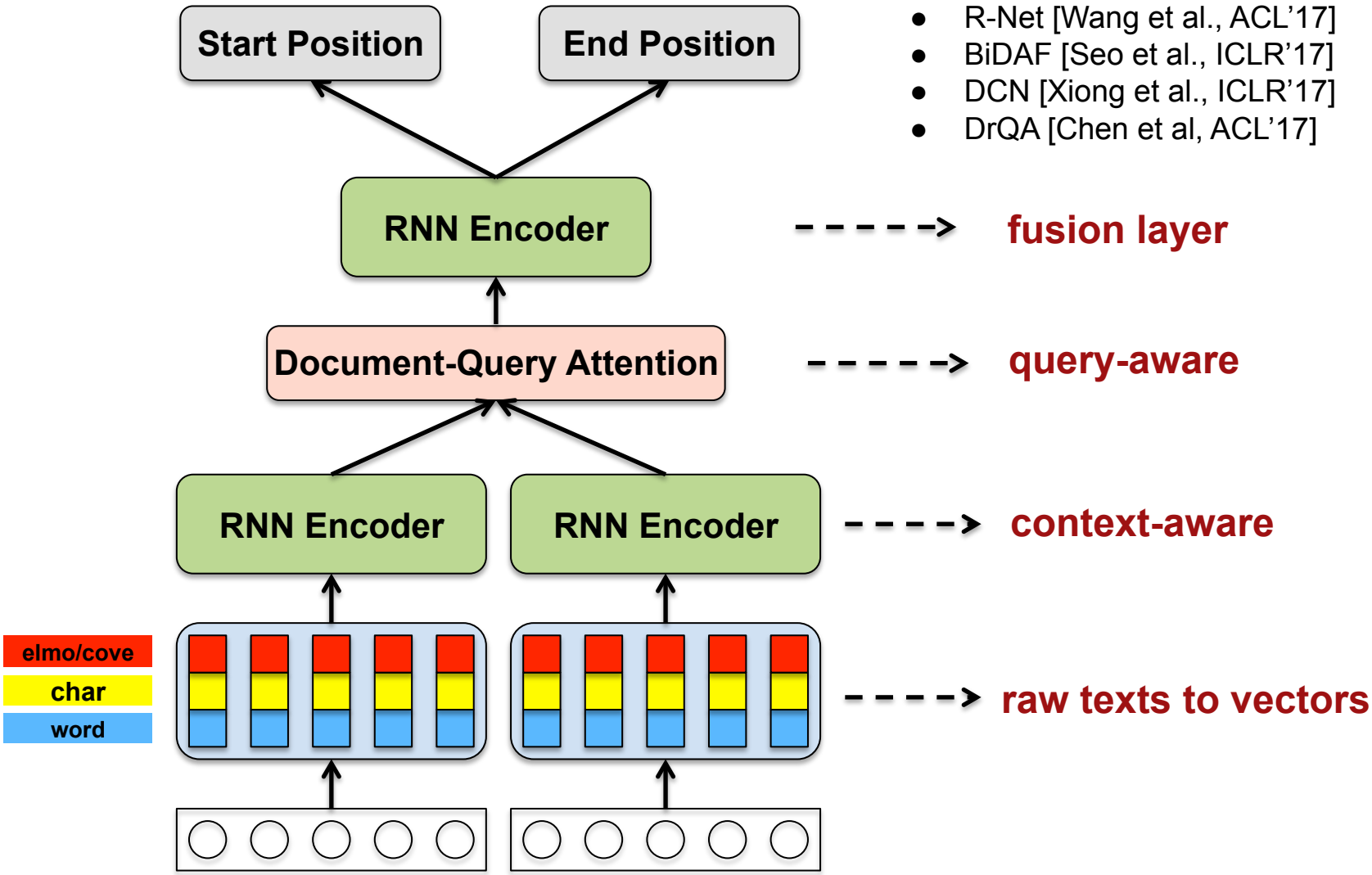
Prediction 3: **teachers facilitate student learning** — **EM = 0, F1 = 0.86**

[Yu et al., ICLR'18]

General Framework

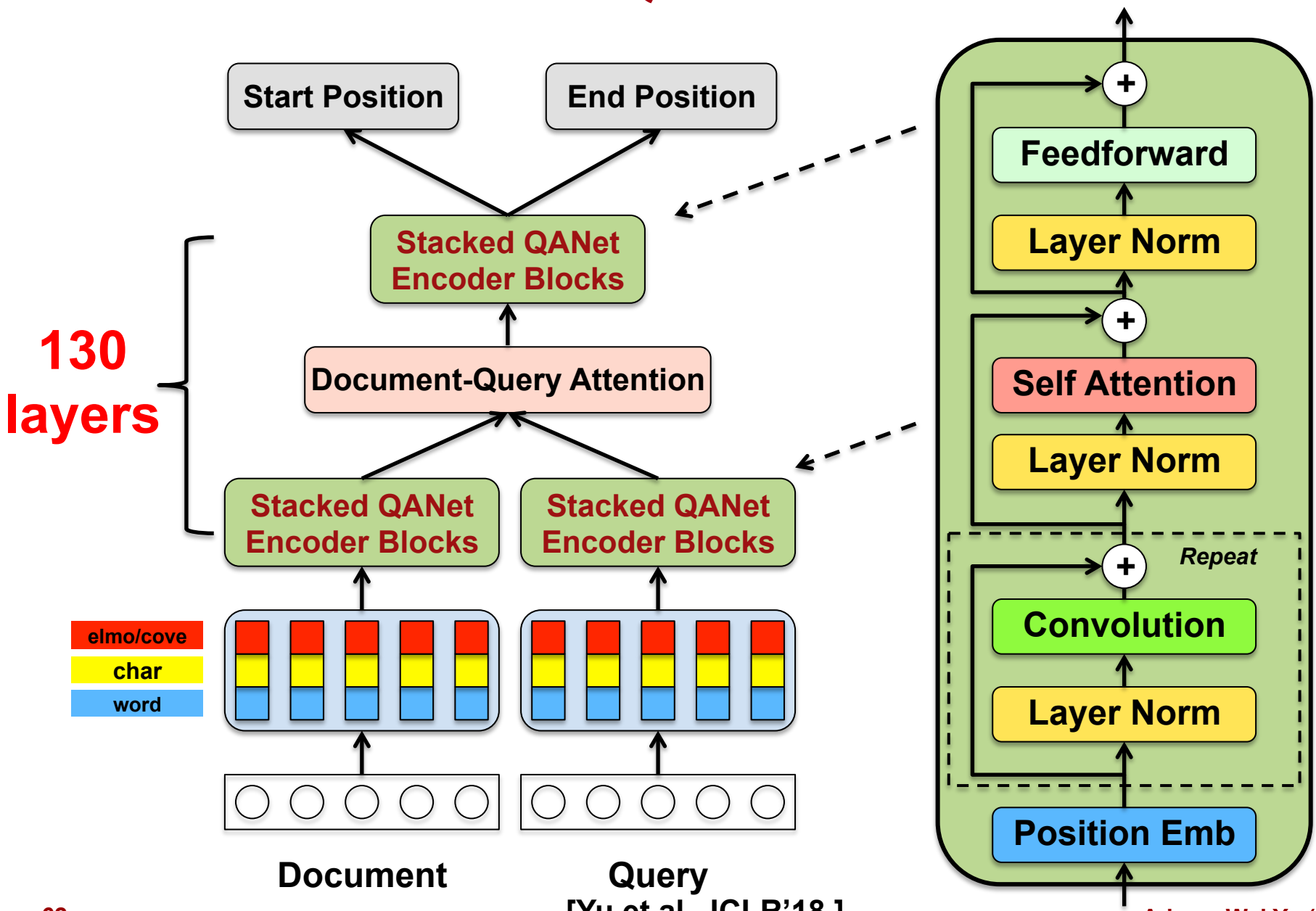
Similar architectures:

- R-Net [Wang et al., ACL'17]
- BiDAF [Seo et al., ICLR'17]
- DCN [Xiong et al., ICLR'17]
- DrQA [Chen et al, ACL'17]



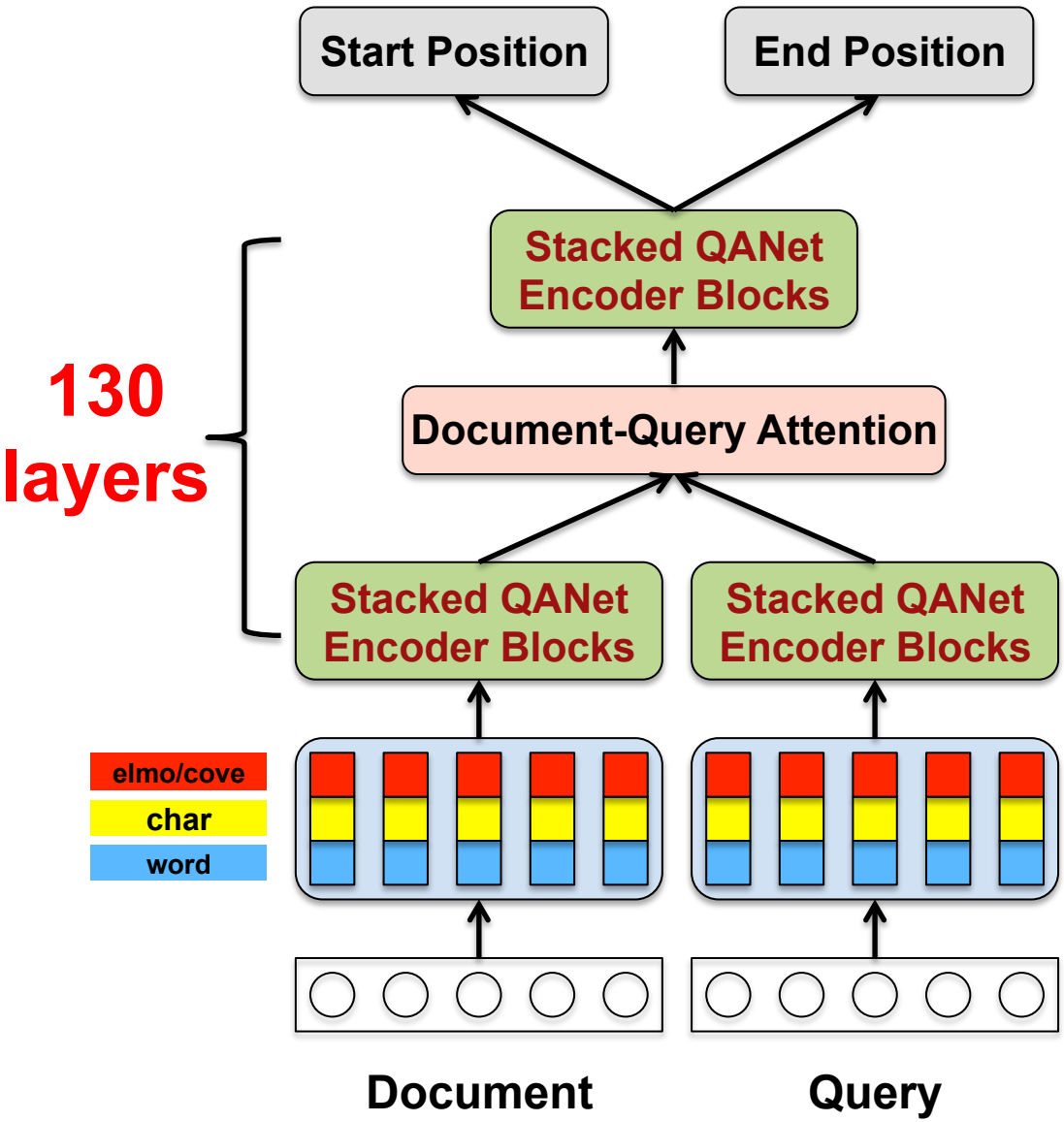
[Yu et al., ICLR'18]

QANet



[Yu et al., ICLR'18]

QANet



[Yu et al., ICLR'18]

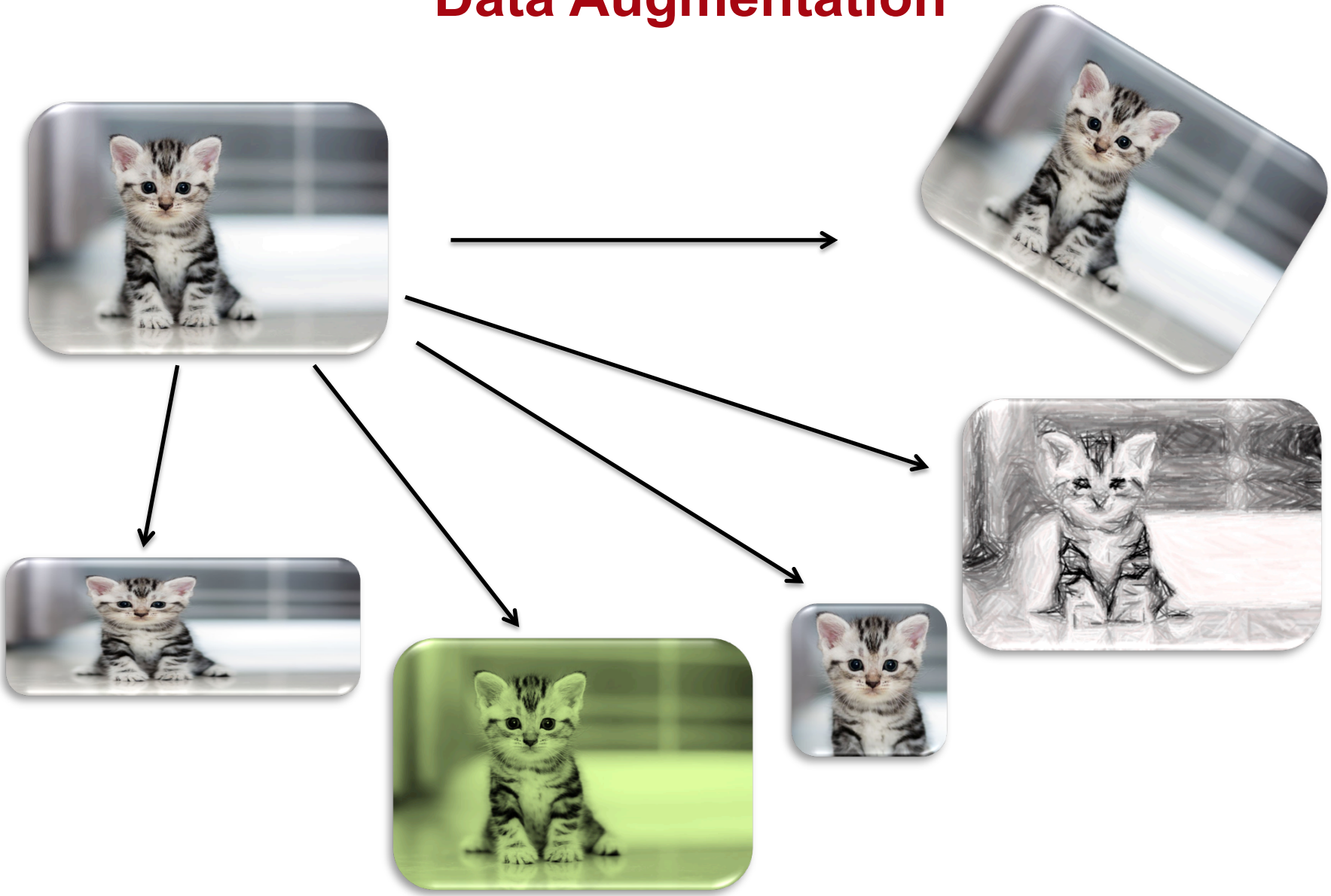
1. Leverage existing techniques in image classification for free

stochastic depth,
residual connection,
squeeze and excitation,
.....

2. Fully Feedforward – Fast

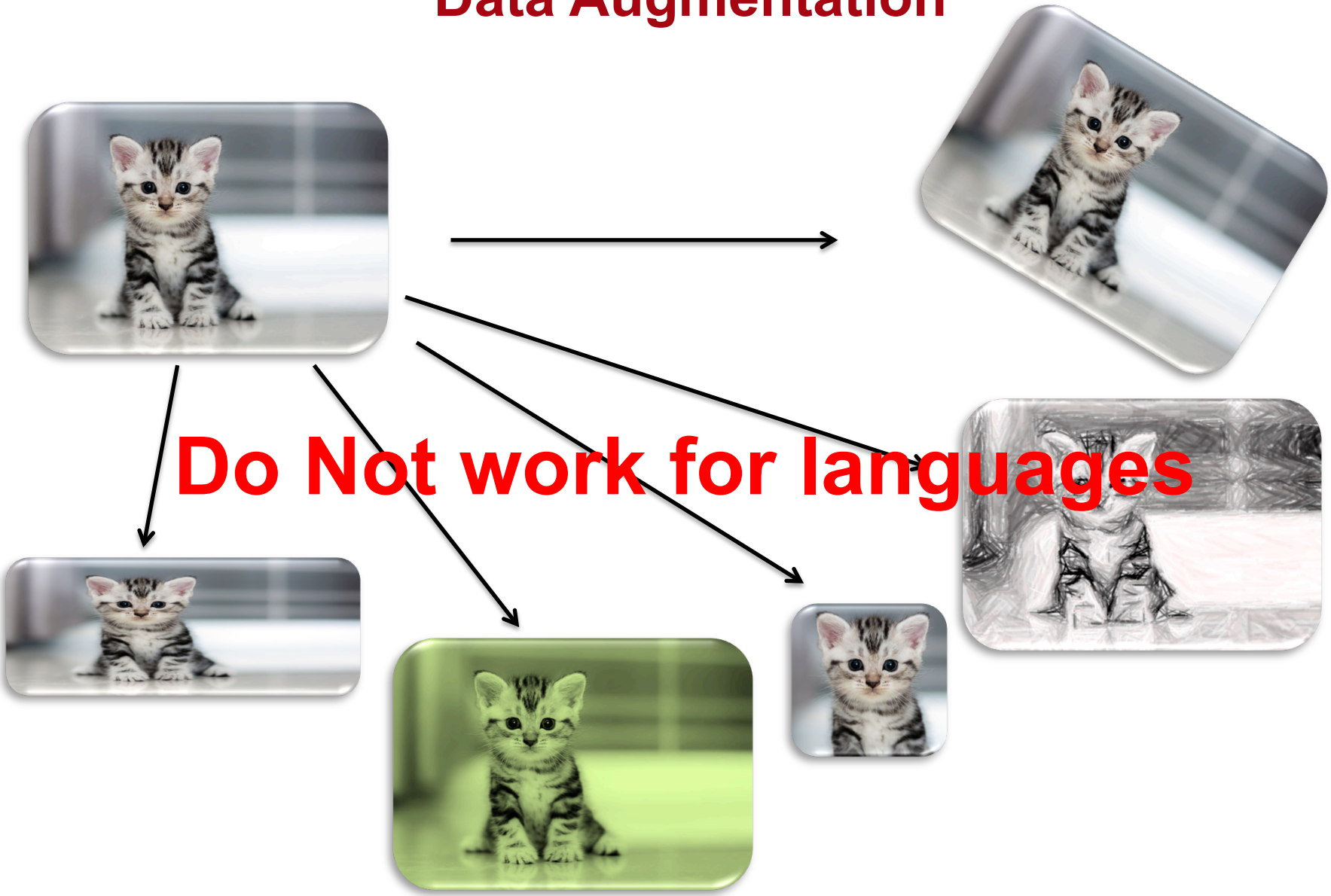
train with more data

Data Augmentation



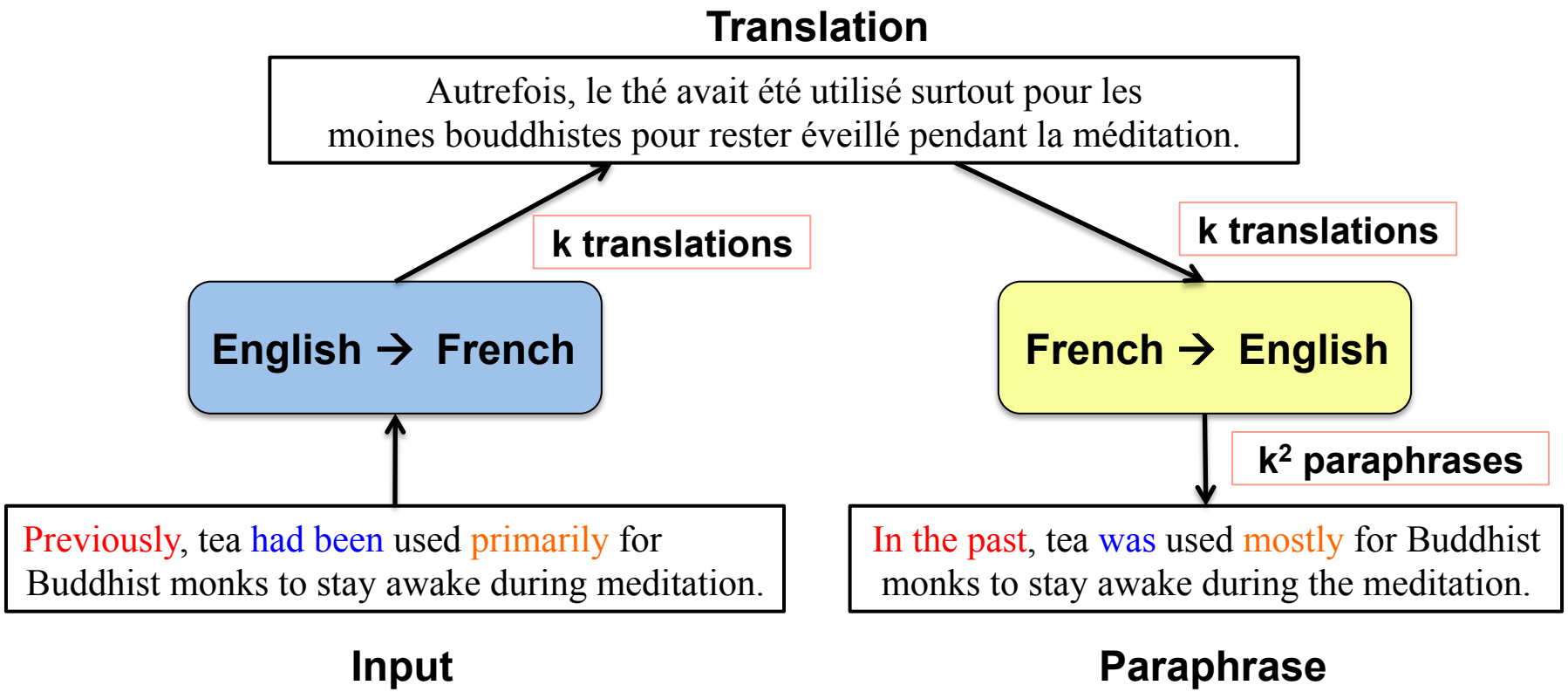
[Yu et al., ICLR'18]

Data Augmentation



[Yu et al., ICLR'18]

Data Augmentation via Cyclic-translation



English → German

English → Czech

English → Spanish

[Yu et al., ICLR'18]

About the SQuAD

- Starts from Aug, 2016.
- More than 50 teams.
- Hundreds of submissions.

Industry:

Academia:

[Yu et al., ICLR'18]

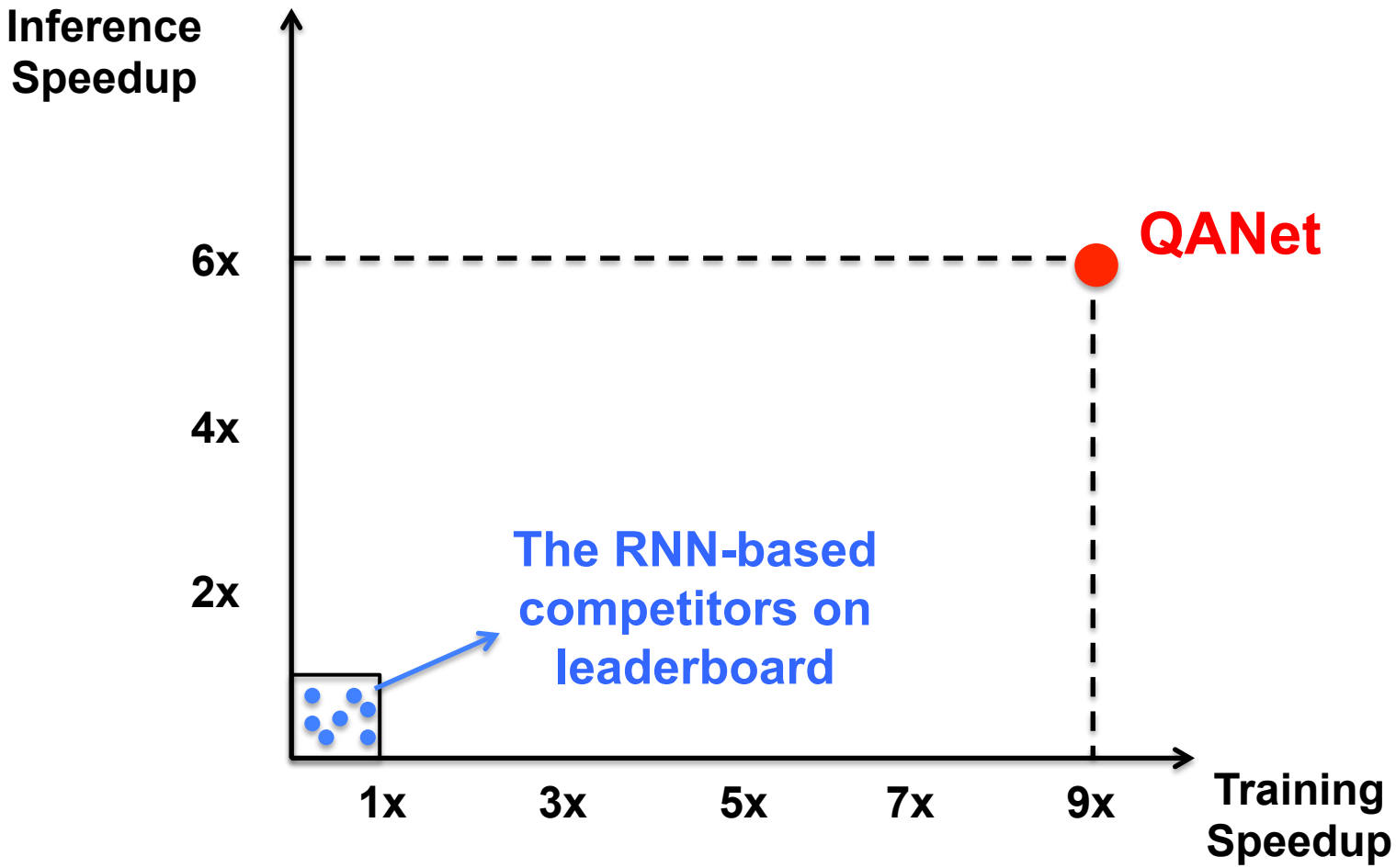
Home

Rank	Model	EM	F1
	Human Performance *Stanford University* (Rajpurkar et al. '16)	82.304	91.221
1 Jul 12, 2018	QANet (ensemble) *Google Brain & CMU*	**84.454**	**90.490**
2 Jul 09, 2018	r-net (ensemble) *Microsoft Research Asia*	84.003	90.147
3 Jun 21, 2018	MARS (ensemble) *YUANFUDAO research NLP*	83.982	89.796
4 Mar 20, 2018	QANet (ensemble) *Google Brain & CMU*	83.877	89.737
5 Jun 21, 2018	MARS (single model) *YUANFUDAO research NLP*	83.122	89.224
6 Mar 07, 2018	QANet (ensemble) *Google Brain & CMU*	82.744	89.045
7 May 09, 2018	MARS (single model) *YUANFUDAO research NLP*	82.587	88.880
7 Feb 20, 2018	Reinforced Mnemonic Reader + A2D (ensemble model) *Microsoft Research Asia & NUDT*	82.849	88.764
7 Jun 20, 2018	QANet (single) *Google Brain & CMU*	82.471	89.306

#1 on SQuAD (Mar-Aug 2018)

Rank	Model	EM	F1
	Human Performance *Stanford University* (Rajpurkar et al. '16)	82.304	91.221
1 Oct 05, 2018	BERT (ensemble) *Google AI Language* https://arxiv.org/abs/1810.04805	**87.433**	**93.160**
2 Oct 05, 2018	BERT (single model) *Google AI Language* https://arxiv.org/abs/1810.04805	85.083	91.835
2 Sep 09, 2018	nlnet (ensemble) *Microsoft Research Asia*	85.356	91.202
2 Sep 26, 2018	nlnet (ensemble) *Microsoft Research Asia*	85.954	91.677
3 Jul 11, 2018	QANet (ensemble) *Google Brain & CMU*	84.454	90.490
4 Jul 08, 2018	r-net (ensemble) *Microsoft Research Asia*	84.003	90.147
5 Mar 19, 2018	QANet (ensemble) *Google Brain & CMU*	83.877	89.737

Speedup



Fast and Accurate!

[Yu et al., ICLR'18]

Stanford Dawnbench Competition:

Model	Time to 75 F1	Hardware
QANet	**0:45:56**	1 TPUv2
BiDAF	7:38:10	1 K80/61 GB/4 CPU
BiDAF	7:51:22	1 P100/512 GB/56 CPU
BiDAF	8:43:40	1 K80/30 GB/8 CPU
BiDAF	10:50:22	60 GB/16 CPU

An End-to-End Deep Learning Benchmark and Competition

10x speedup

[Yu et al., ICLR'18]

Failure cases

Question

Who else did Tesla make the acquaintance of in 1886?

Context

… In late 1886 Tesla met Alfred S. Brown , a Western Union superintendent , and New York attorney **Charles F. Peck** … Together in April 1887 they formed the Tesla Electric Company with an agreement that profits from generated patents would go ⅓ to Tesla , ⅓ to Peck and Brown , and ⅓ to fund development .

Ground truths

Charles F. Peck

Predictions from 14 models

model 0: Alfred S. Brown
model 1: Alfred S. Brown
model 2: Alfred S. Brown
model 3: Alfred S. Brown
model 4: Alfred S. Brown
model 5: Alfred S. Brown
model 6: Alfred S. Brown
model 7: Alfred S. Brown
model 8: Alfred S. Brown
model 9: Alfred S. Brown
model 10: Alfred S. Brown
model 11: Alfred S. Brown, a Western Union superintendent, and New York attorney Charles F. Peck
model 12: Alfred S. Brown
model 13: Peck and Brown

Failure cases

Question

Where by mass is oxygen a major part?

Context

… Oxygen constitutes 49.2 % of the Earth ' s crust by mass and is the major component of the **world' s oceans** (**88.8 %** by mass) . Oxygen gas is the second most common component of the Earth ' s atmosphere , taking up 20.8 % of its volume and 23.1 % of its mass (some 1015 tonnes) ...

Ground truths

world's oceans

Predictions from 14 models

model 0: 88.8%
model 1: 88.8%
model 2: world's oceans
model 3: 88.8%
model 4: the world's oceans
model 5: 88.8%
model 6: 88.8%
model 7: 88.8%
model 8: world's oceans
model 9: world's oceans
model 10: 88.8%
model 11: 88.8%
model 12: world's oceans
model 13: world's oceans

QA is far from Solved!

Roadmap

1. Skipping Irrelevant Information [ACL'17]
2. Discarding Recurrence [ICLR'18]
3. Future Work

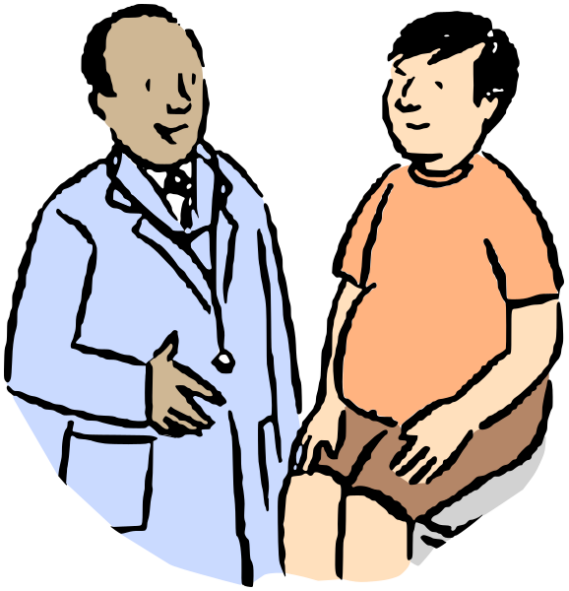
Future Work

1. Downstream Applications
2. AutoML for QA.
3. Real Human Level Comprehension

Downstream Applications

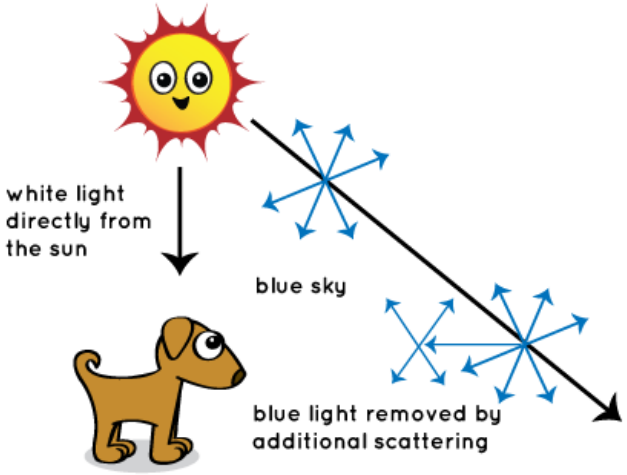
1. Medical QA

- Input: doctor-patient dialog.
- Question: what is the syndrome of the patient?

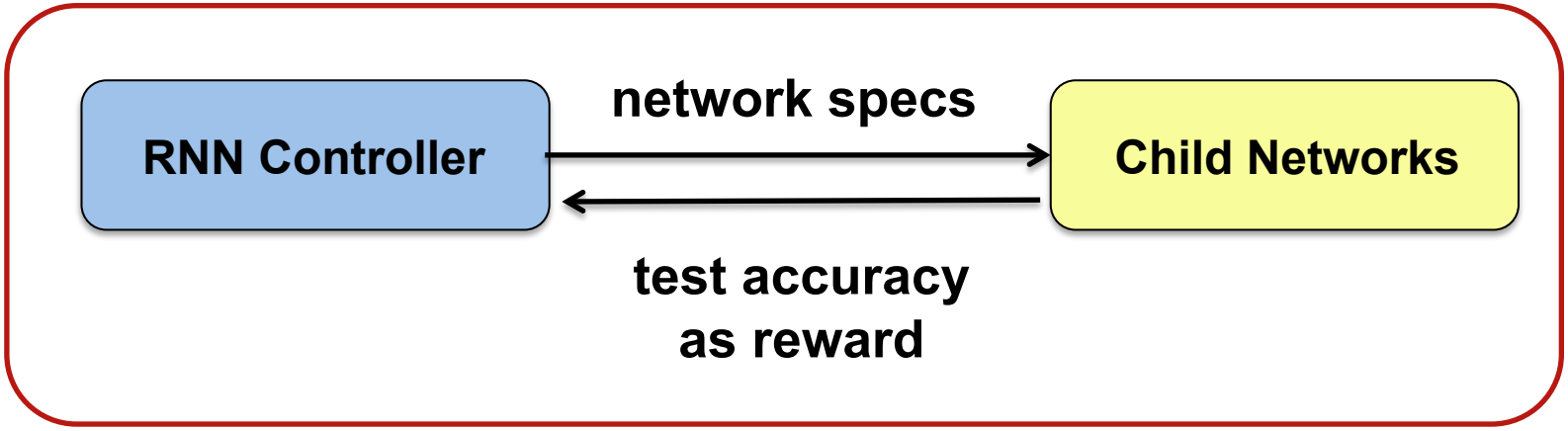


2. Web QA

- Query: Why is the sky blue?



AutoML (Neural Architecture Search)



Reinforcement Learning

Human Level Comprehension

US·MLE
United States
Medical
Licensing
Examination

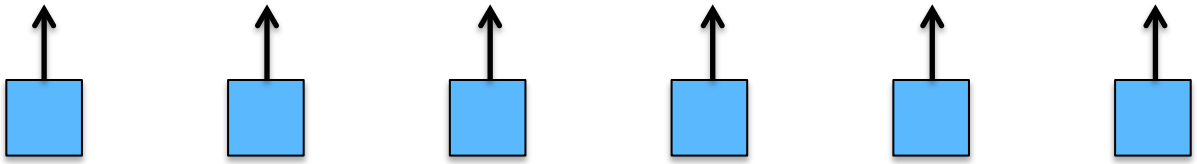
1. Extremely long texts (Many books)
2. Complicated reasoning (e.g., Medical License Exam)

Summary

RNN / LSTM

LSTM-Jump

QANet



Thank you!