Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generating Natural Questions About an Image

[5]

Ryan Callihan & Sarah Taylor

Seminar für Sprachwissenschaft
Universität Tübingen

Janurary 12, 2018

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Outline

- Generating Questions from Pictures
  - The Authors' Objective
- Image Recognition using a CNN
- Datasets
  - MS COCO
  - Bing
  - Flicker
- Generative Models
  - Maximum Entropy Language Model
  - Long Short-Term Memory
  - Gated Recurrent Neural Network
- Retrieval Model
  - K-Nearest Neighbor
- Evaluation
- Discussion
  - Questions
- References

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Getting Started

Make groups!

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Challenge One - Setup

Please go to: and enter in the code:

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Challenge One

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Challenge One - Corpus Results

- Is this a religious ceremony?
- That looks very interesting, don't you think?
- What are they all gathered for?
- What are these people gathered for?
- Is this a satanic ritual?

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Challenge Two - Setup

Please go to: and enter in the code:

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Challenge Two

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Challenge Two - Corpus Results

- What are the people demonstrating about?
- What rally are they attending?
- What are these people protesting?
- What are they protesting?
- Who is in the gray jacket?

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

What is Image Recognition?

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Image recognition and neural networks

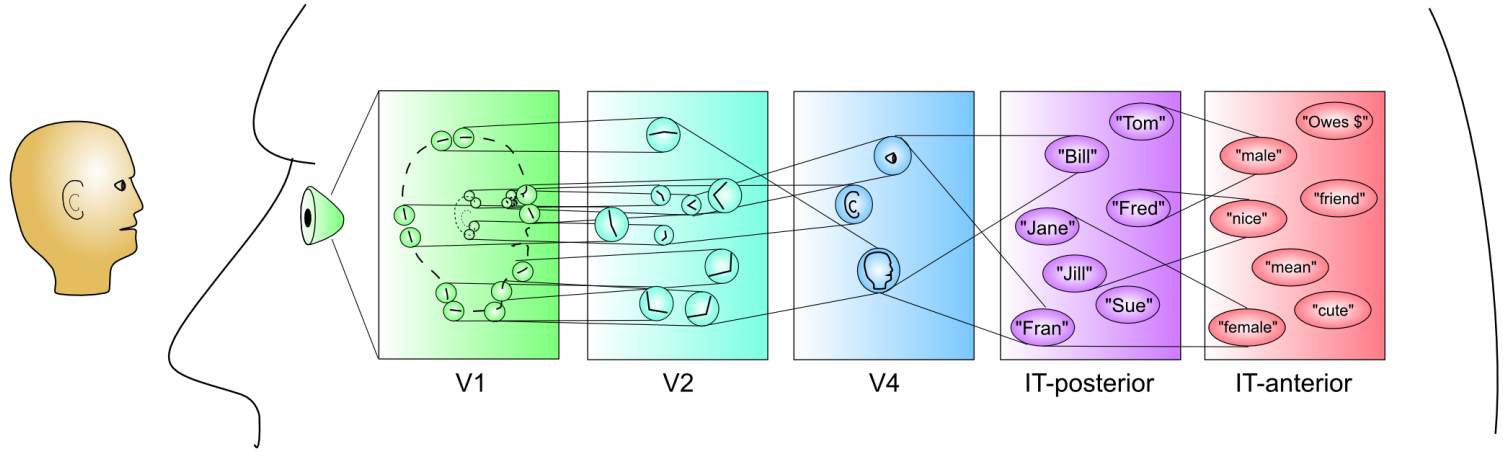


Figure: *Visual network representation. Image from grey.colorado.edu [6]*

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Convolutional neural network

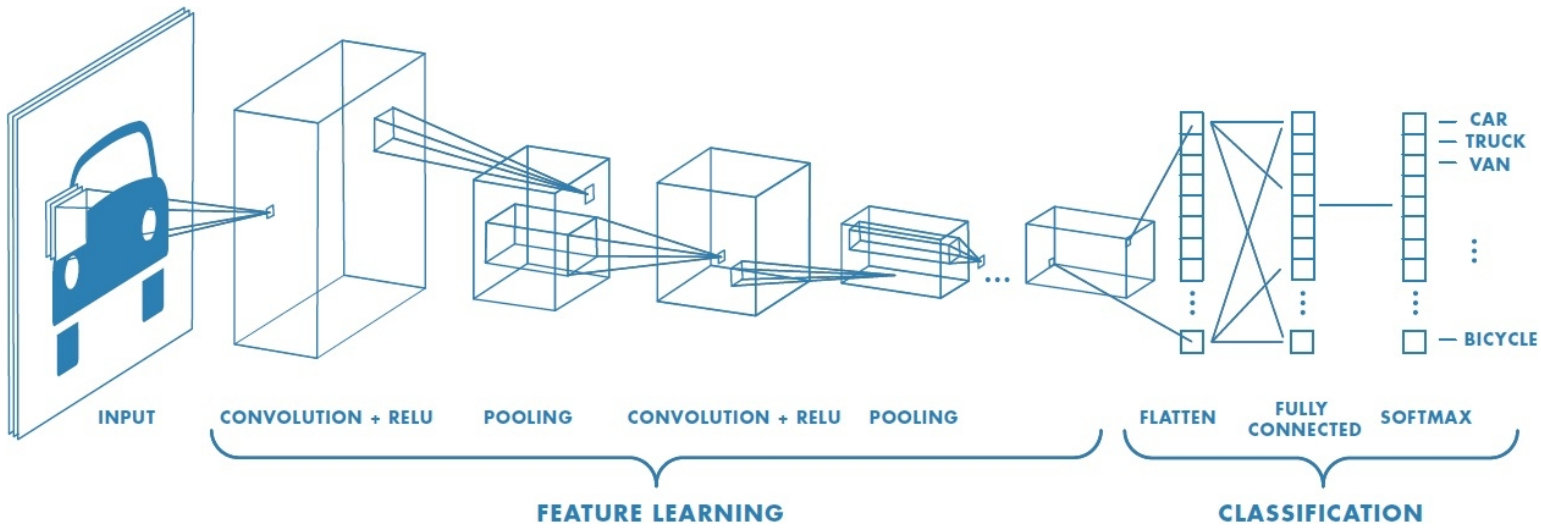


Figure: *CNN representation. Image from blog.floydhub.com/building-your-first-convnet [8]*

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

CNN visualized

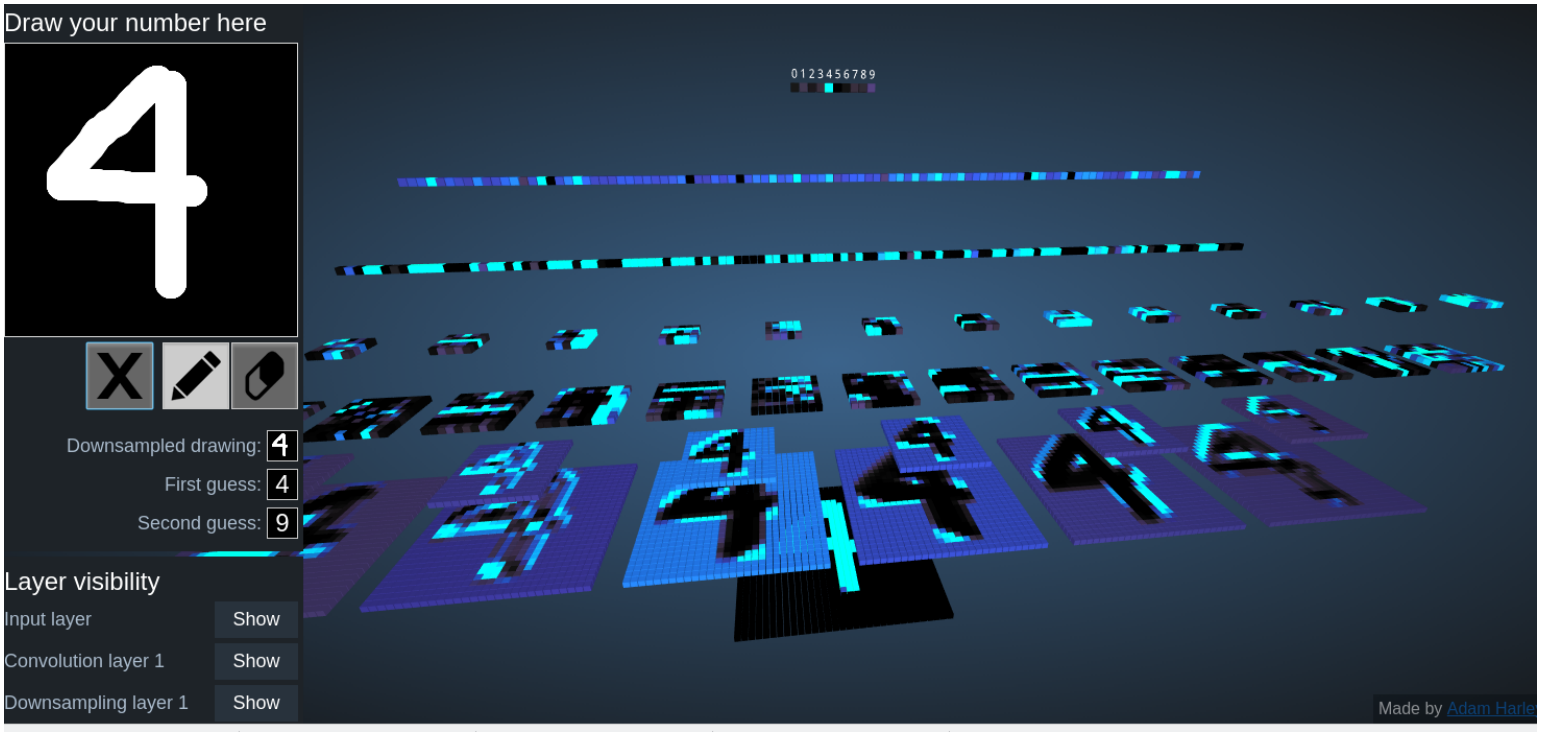


Figure: *CNN visualized. Image from scs.ryerson.ca/ [4]*

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generative Models

Caption Bot [2]

captionbot.ai was used throughout this paper to automatically generate captions. It is a Microsoft project based on the Computer Vision API, Emotion API, and Bing Image Search API.

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

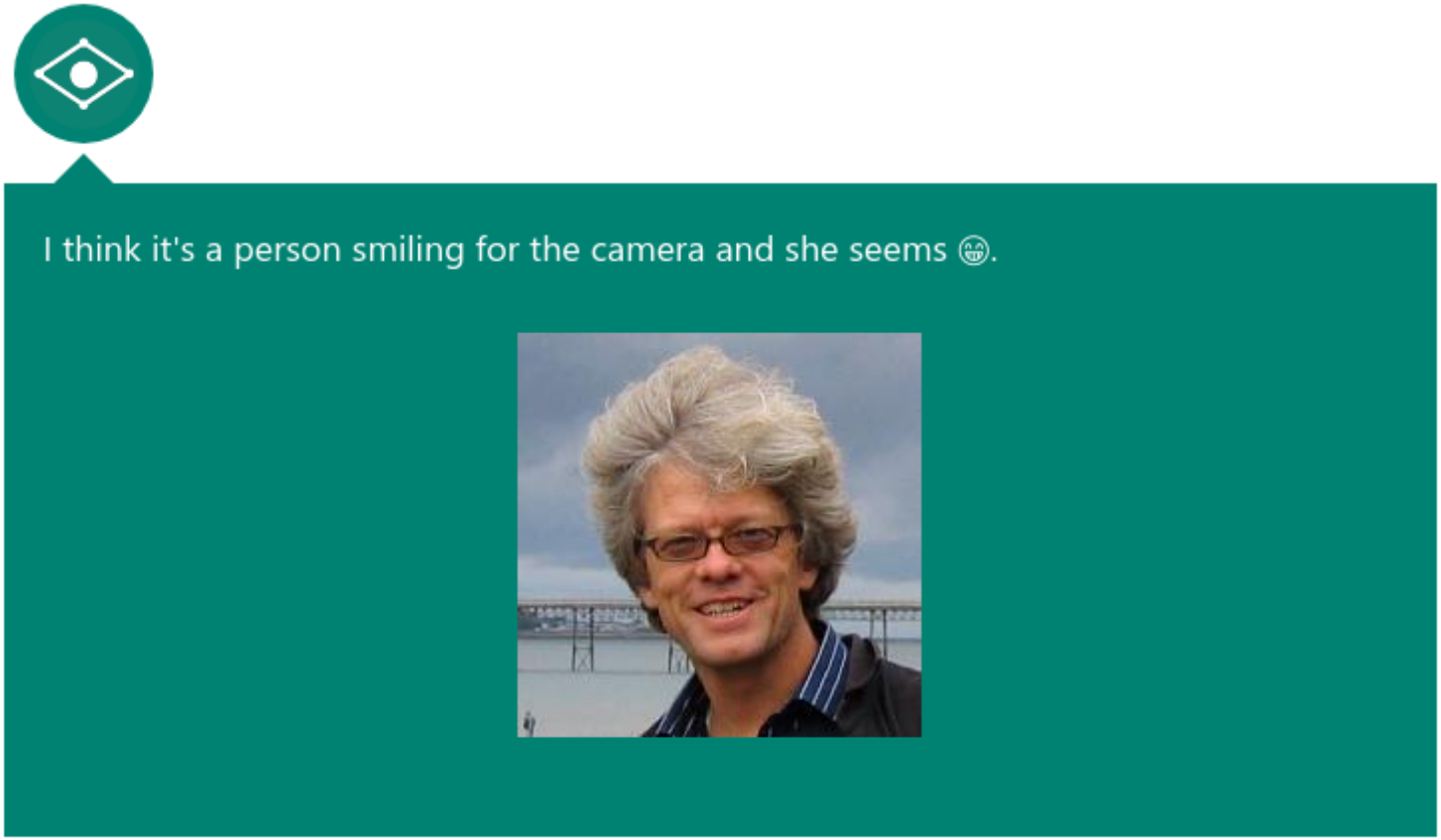
Discussion

Questions

References

References

Caption Bot example



Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

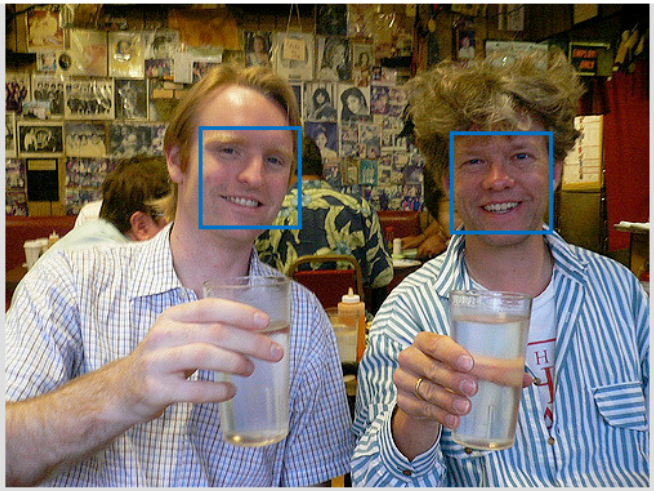
Discussion

Questions

References

References

Computer Vision API example



FEATURE NAME:	VALUE
Description	{ "tags": ["person", "man", "indoor", "table", "sitting", "holding", "food", "woman", "glasses", "people", "posing", "drinking", "wine", "restaurant", "plate", "smiling", "pizza", "phone", "young", "standing", "store", "group", "white"], "captions": [{ "text": "a man sitting at a table in a restaurant", "confidence": 0.9105153 }] }
Tags	[{ "name": "person", "confidence": 0.999498367 }, { "name": "man", "confidence": 0.928230047 }, { "name": "indoor", "confidence": 0.8648256 }, { "name": "restaurant", "confidence": 0.193121776 }]
Image format	"Jpeg"

- Description "tags": ["person", "man", "indoor", "table", "sitting", "holding", "food", "woman", "glasses", "people", "posing", "drinking", "wine", "restaurant", "plate", "smiling", "pizza", "phone", "young", "standing", "store", "group", "white"], "captions": ["text": "a man sitting at a table in a restaurant", "confidence": 0.9105153]
- Faces ["age": 25, "gender": "Male", "faceRectangle": "top": 94, "left": 149, "width": 79, "height": 79 , "age": 33, "gender": "Male", "faceRectangle": "top": 97, "left": 343, "width": 79, "height": 79]

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures

The Authors' Objective

Image Recognition using a CNN

Datasets

MS COCO

Bing

Flicker

Generative Models

Maximum Entropy Language Model

Long Short-Term Memory

Gated Recurrent Neural Network

Retrieval Model

K-Nearest Neighbor

Evaluation

Discussion

Questions

References

References

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Generative models

Three generative models were used in this study

- Maximum entropy language model (MELM)
- RNN with long short-term memory cells (LSTM)
- RNN with gated recurrent units (GRU)

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Model input

The last fully connected layer of the image recognition CNN was used as input

It was a 4096-dimentional vector

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Maximum Entropy Language Model

- Generated word probabilities using a CNN

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Maximum Entropy Language Model

- Generated word probabilities using a CNN
- MELM using candidate word probabilities

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Maximum Entropy Language Model

- Generated word probabilities using a CNN
- MELM using candidate word probabilities
- Relevant questions but some grammatical errors. Warrents futher research

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

RNN with Long Short-Term Memory

- Given a caption, model generates a question

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

RNN with Long Short-Term Memory

- Given a caption, model generates a question
- Used gold standard caption-question dataset for training

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

RNN with Long Short-Term Memory

- Given a caption, model generates a question
- Used gold standard caption-question dataset for training
- Consists of 2 RNNs
  - An encoder which processes the caption
  - A decoder which generates the question

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

RNN with Long Short-Term Memory

- Given a caption, model generates a question
- Used gold standard caption-question dataset for training
- Consists of 2 RNNs
  - An encoder which processes the caption
  - A decoder which generates the question
- Output was incoherent

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Gated Recurrent Neural Network

- Uses final fully connected layer from the image recognition CNN (Captionbot I assume) into a 500-dimentional vector

Code based on this model can be found here: github.com/JamesChuanggg/VQG-tensorflow

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Gated Recurrent Neural Network

- Uses final fully connected layer from the image recognition CNN (Captionbot I assume) into a 500-dimentional vector
- Generates a single token at a time until the end of sentence token (aka "?")

Code based on this model can be found here: github.com/JamesChuanggg/VQG-tensorflow

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Gated Recurrent Neural Network

- Uses final fully connected layer from the image recognition CNN (Captionbot I assume) into a 500-dimentional vector
- Generates a single token at a time until the end of sentence token (aka "?")
- "Where is this?" Is the top generated question for 29.3% of the images. More interesting and meaningful quesions generally ranked lower

Code based on this model can be found here: github.com/JamesChuanggg/VQG-tensorflow

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Gated Recurrent Neural Network

- Uses final fully connected layer from the image recognition CNN (Captionbot I assume) into a 500-dimentional vector
- Generates a single token at a time until the end of sentence token (aka "?")
- "Where is this?" Is the top generated question for 29.3% of the images. More interesting and meaningful quesions generally ranked lower

Code based on this model can be found here: github.com/JamesChuanggg/VQG-tensorflow

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Gated Recurrent Neural Network

- Uses final fully connected layer from the image recognition CNN (Captionbot I assume) into a 500-dimentional vector
- Generates a single token at a time until the end of sentence token (aka "?")
- "Where is this?" Is the top generated question for 29.3% of the images. More interesting and meaningful quesions generally ranked lower
- They added a constraint that which rejects questions with less that 6 tokens

Code based on this model can be found here: github.com/JamesChuanggg/VQG-tensorflow

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Retrieval Model

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

K-Nearest Neighbor

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Evaluation

Two methods of evaluation were used

- Human evaluation using AMT on a 3 point scale

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Evaluation

Two methods of evaluation were used

- Human evaluation using AMT on a 3 point scale
- Automatic evaluation
  - BLEU [7]
  - $\Delta BLEU$ [3]
  - METEOR [1]

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Human Evaluation

Using AMT, 3 people evaluated each question on a 3 point scale from 1 (worst) - 3 (best)

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective

Image Recognition using a CNN

Datasets
MS COCO
Bing
Flicker

Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network

Retrieval Model
K-Nearest Neighbor

Evaluation

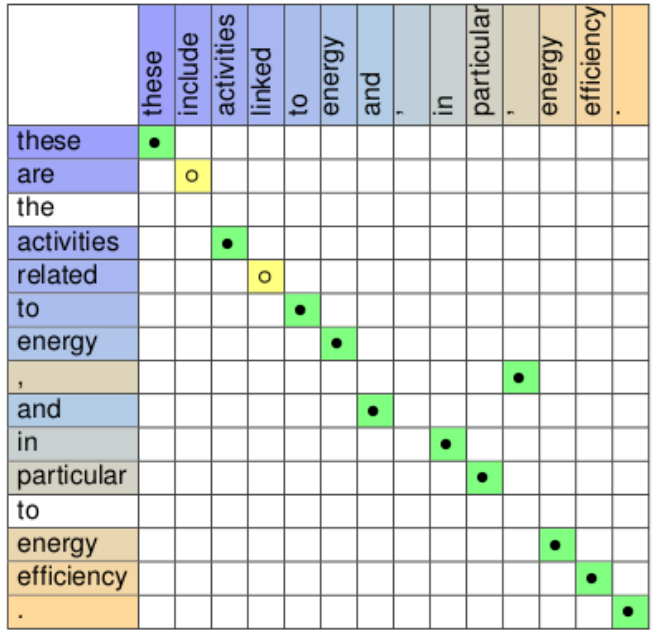
Discussion
Questions

References

References

BLEU, $\Delta BLEU$, & METEOR

Similar concept to Levenshtein or Needleman-Wunsch Distance



Segment 2022

P:	0.897
R:	0.907
Frag:	0.514
Score:	0.440

Figure: *Taken from cs.cmu.edu*

Correlation between human and automatic metrics

	METEOR	*BLEU*	Δ*BLEU*
r	0.916 (4.8e-27)	0.915 (4.6e-27)	0.915 (5.8e-27)
ρ	0.628 (1.5e-08)	0.67 (7.0e-10)	0.702 (5.0e-11)
τ	0.476 (1.6e-08)	0.51 (7.9e-10)	0.557 (3.5e-11)

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Evaluation discussion

With your group, take a look at the evaluation results table and:

- Decide which model and data set worked the best and which worked the worst?

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Evaluation discussion

With your group, take a look at the evaluation results table and:

- Decide which model and data set worked the best and which worked the worst?
- Why do these differences exist?

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Discussion - Authors' Thoughts

Generating Natural Questions About an Image

Ryan Callihan & Sarah Taylor

Generating Questions from Pictures
The Authors' Objective
Image Recognition using a CNN
Datasets
MS COCO
Bing
Flicker
Generative Models
Maximum Entropy Language Model
Long Short-Term Memory
Gated Recurrent Neural Network
Retrieval Model
K-Nearest Neighbor
Evaluation
Discussion
Questions
References
References

Questions

References I

[1] Denkowski, M. and Lavie, A. (2014). Meteor universal: Language specific translation evaluation for any target language. In *Proceedings of the EACL 2014 Workshop on Statistical Machine Translation*.

[2] Fang, H., Gupta, S., Iandola, F., Srivastava, R. K., Deng, L., Dollár, P., Gao, J., He, X., Mitchell, M., Platt, J. C., et al. (2015). From captions to visual concepts and back. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1473–1482.

[3] Galley, M., Brockett, C., Sordoni, A., Ji, Y., Auli, M., Quirk, C., Mitchell, M., Gao, J., and Dolan, B. (2015). deltableu: A discriminative metric for generation tasks with intrinsically diverse targets. *CoRR*, abs/1506.06863.

[4] Harley, A. W. (2015). An interactive node-link visualization of convolutional neural networks. In *ISVC*, pages 867–877.

[5] Mostafazadeh, N., Misra, I., Devlin, J., Mitchell, M., He, X., and Vanderwende, L. (2016). Generating natural questions about an image. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1802–1813, Berlin, Germany. Association for Computational Linguistics.

References II

[6] O'Reilly, R. C., Munakata, Y., Frank, M., Hazy, T., et al. (2012). *Computational cognitive neuroscience*. PediaPress.

[7] Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 311–318. Association for Computational Linguistics.

[8] Wadhwa, S. (2017). Building your first convnet.