Generating Natural Questions About an Image

[?]

Ryan Callihan & Sarah Taylor

Seminar für Sprachwissenschaft
Universität Tübingen

Janurary 12, 2018

Outline

Getting Started

Make groups!

Challenge One - Setup

Please go to: and enter in the code:

Challenge One

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?

Challenge Two - Setup

Please go to: and enter in the code:

Challenge Two

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?

What is Image Recognition?

Image recognition and neural networks

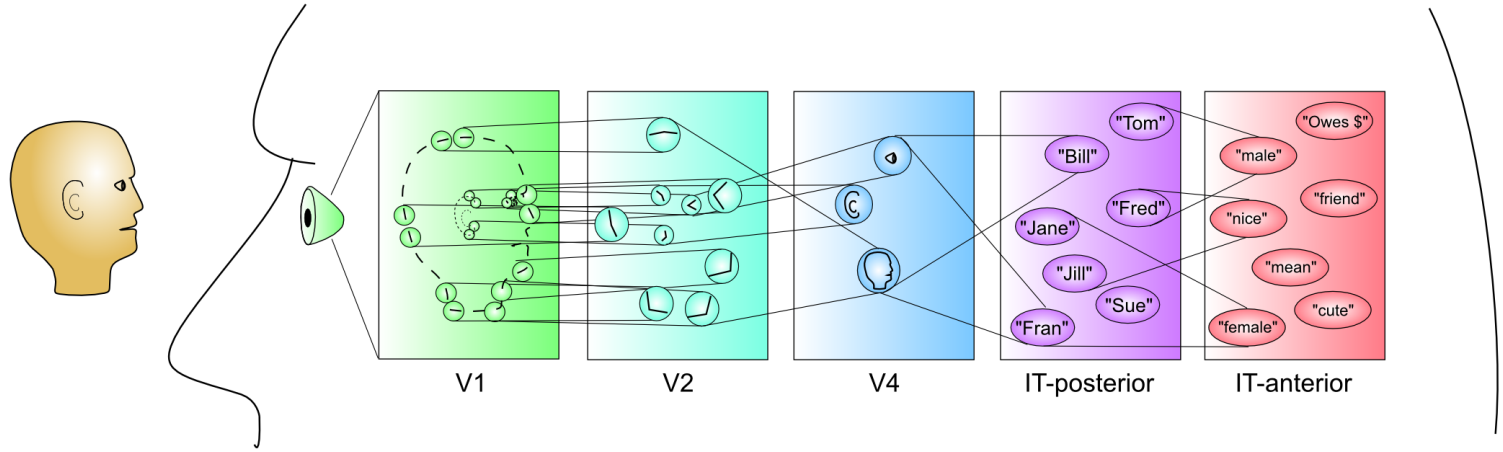


Figure: *Visual network representation. Image from grey.colorado.edu [?]*

Convolutional neural network

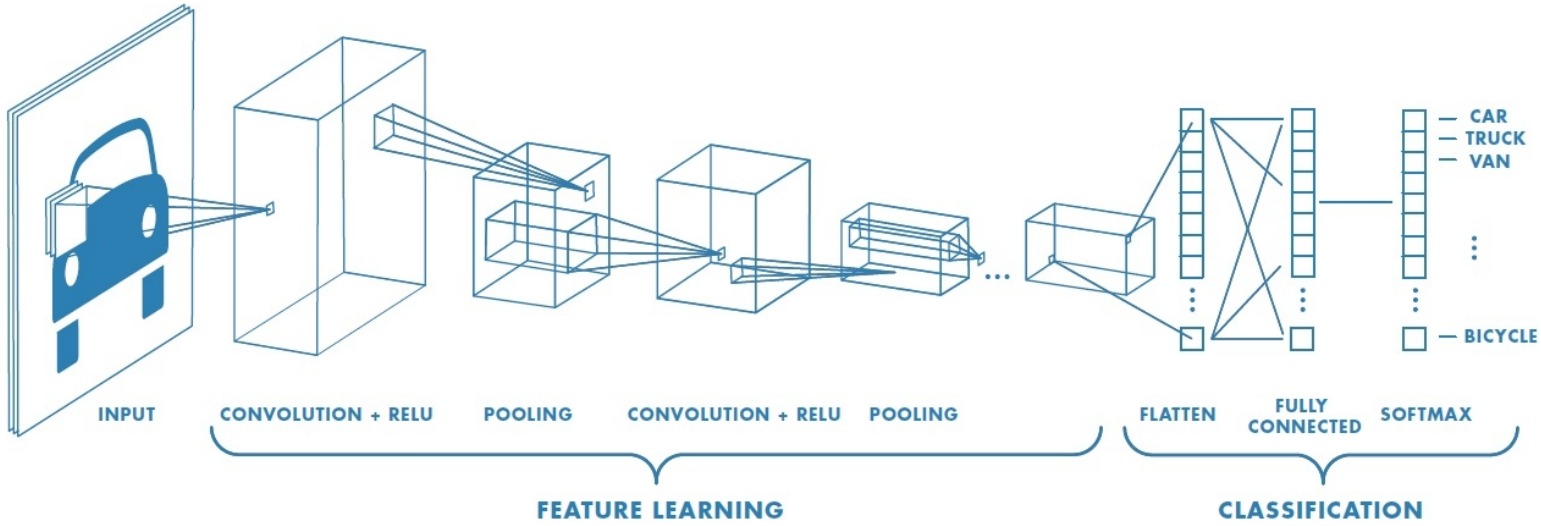


Figure: *CNN representation. Image from blog.floydhub.com/building-your-first-convnet [?]*

CNN visualized

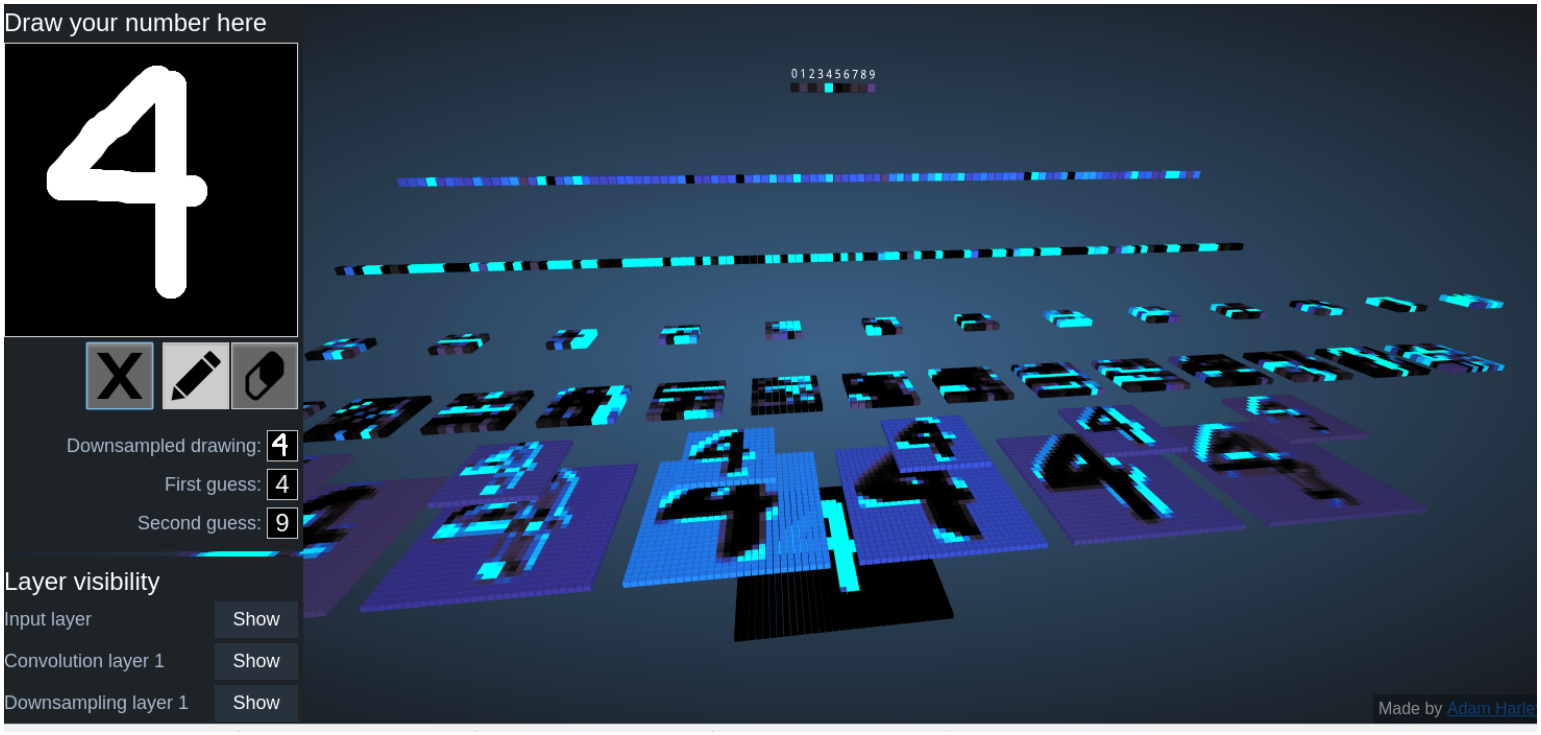


Figure: *CNN visualized. Image from scs.ryerson.ca/ [?]*

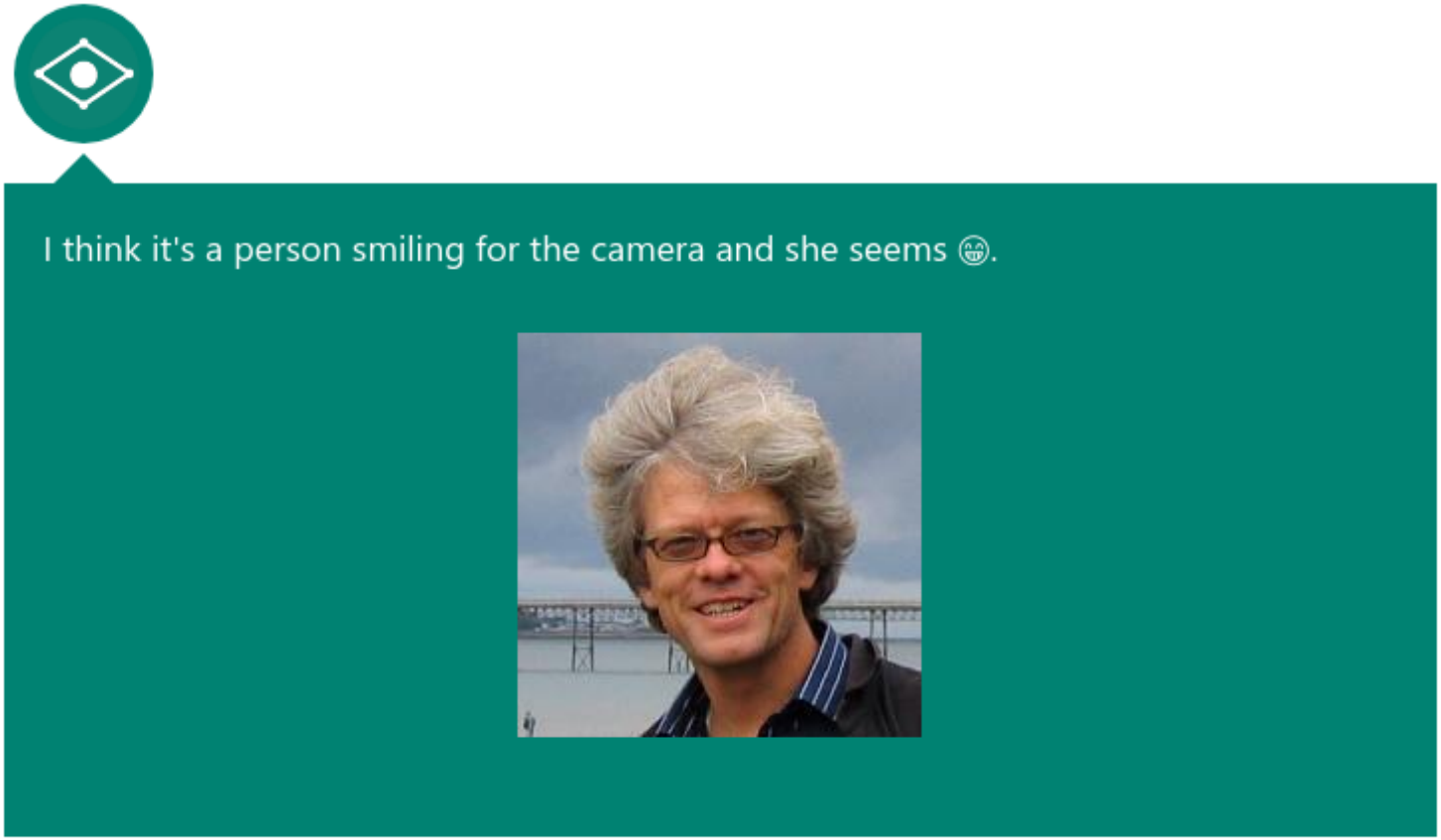
Why Does It Work Well for Images?

Generative Models

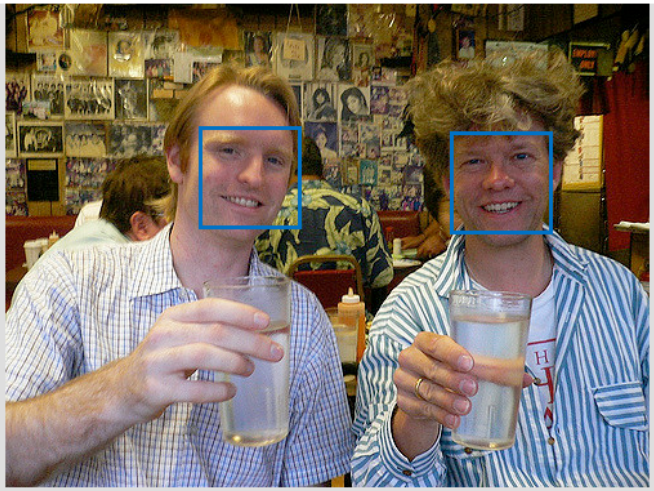
Caption Bot [**?**]

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.

Caption Bot example



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]

MELM

LSTM

GRNN

Retrieval Model

K-Nearest Neighbor

BLEU

Evaluation

BLEU & METEOR

Discussion - Authors' Thoughts

Questions

References I