When you get a Deep Learning project idea and you can't find any papers on it

debojeet

YOU ARE NOT LOGGED IN! ⅔ of you!

CIS 522: 5T

Application domains overview & Final Project start
02/11/20

Penn
Engineering

What we will do today

- Overview over application domains
  - Computer Vision
  - Natural Language processing
  - Reinforcement learning
  - Other applications
  - Theory
- Start thinking about and working towards projects
- Time for a third of term Q&A

Feedback

HW1 had a mode of ~25h, a shorter mean, and a shortest solution time of just a few hours.

It was *really* hard without the prereqs

that was intended

Important: We will be curving for grades!

An architectures omission

RNNs

- A bit like convolution. Just over time
- We unroll time
- Use backpropagation on unrolled time

We will cover this in NLP lectures

But first, Re-asking the question question

Bad:

NLP

Vision

Good:

Quantify how much taking a Viking boatmaking course increases lifetime earnings.

Build an NLP system that runs through a deck of google slides, crosschecks them with r/sarcasm and then makes them funny

PollEV make better question.
We will revisit at end of lecture.

Overview of final project deliverables

There will be 40 projects with 3 to 4 students per team.

- Each of you will submit a short abstract of your project idea by Feb 18
  - Review literature before writing the abstract
  - Think of an interesting project topic and describe it
  - Outline methods to be used and expected outcome
- TAs and Konrad will select 40 projects from all submitted abstracts by Feb 20
- Students who submitted the selected abstracts work on that project

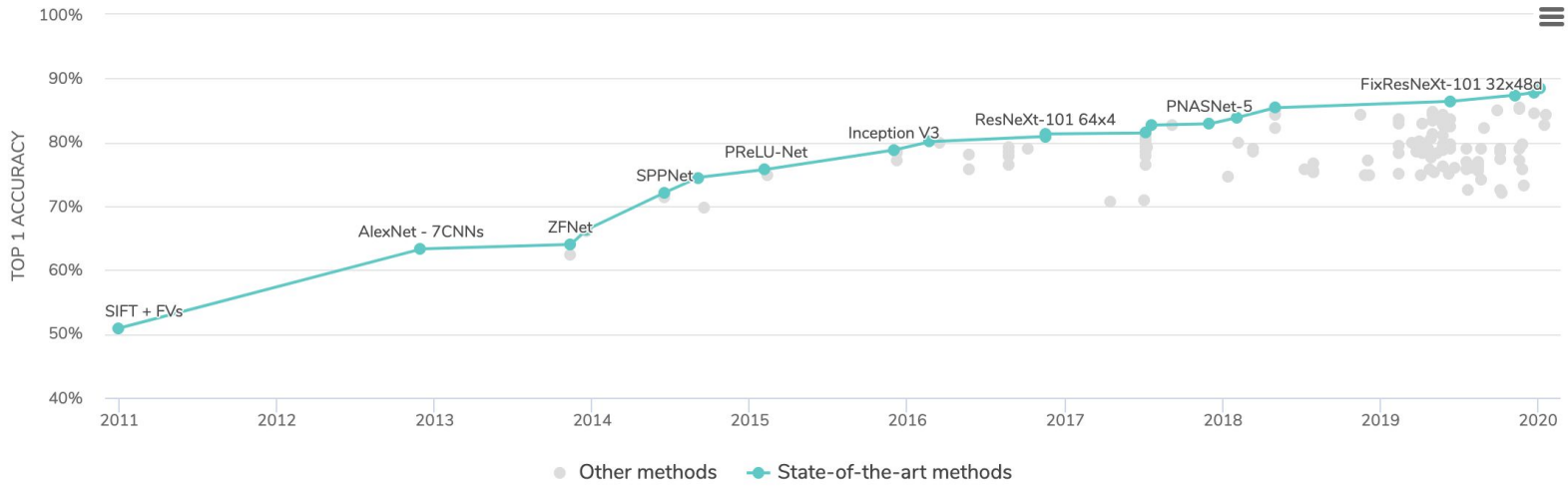
Overview of final project deliverables

- You will find a project, a team, and submit a longer proposal by Feb 28
  - Read the 40 selected abstracts and decide which you like
  - Use Piazza to find teammates
  - Project proposal will outline your planned milestones
- Milestone 1 will be due on March 20
- Milestone 2 will be due on April 10
- Milestone 3 and all code will be due on April 24
- Final project presentation fair will be on April 28

It is necessary to choose project in area we have not covered yet

Computer Vision Applications

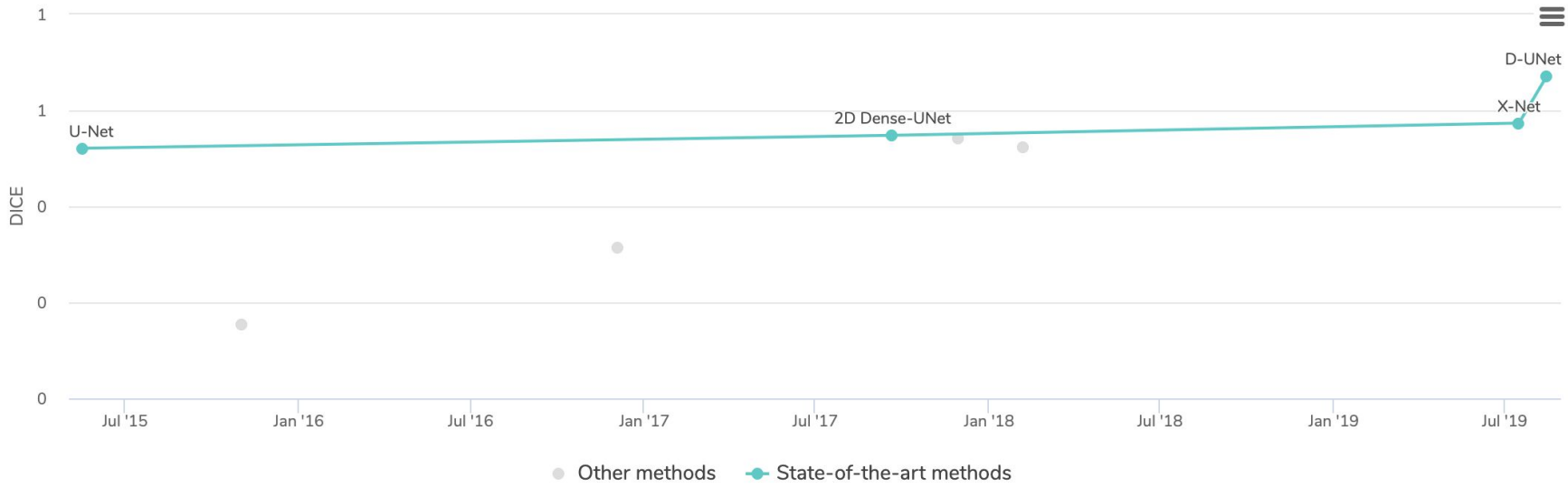
History of imagenet



Check out paperswithcode

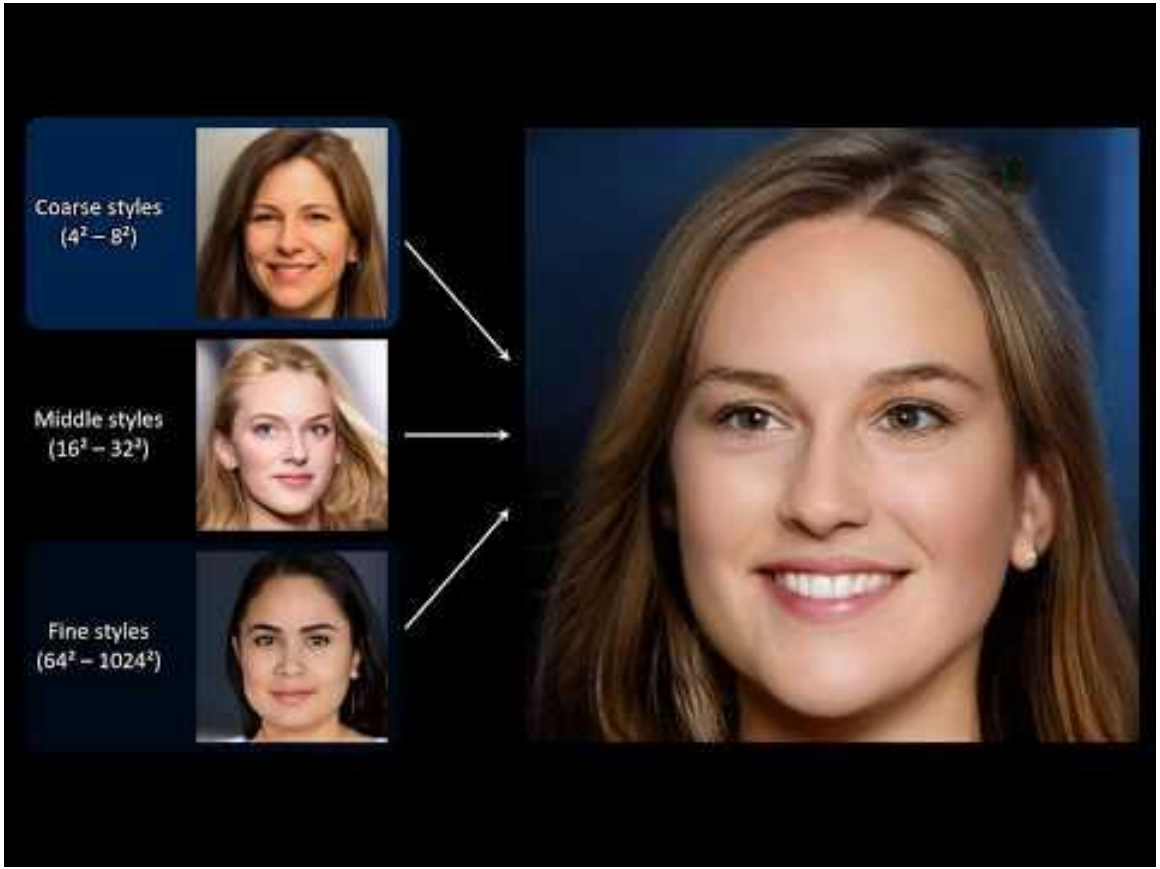
And lesion segmentation

Lesion Segmentation on Anatomical Tracings of Lesions After Stroke (ATLAS)

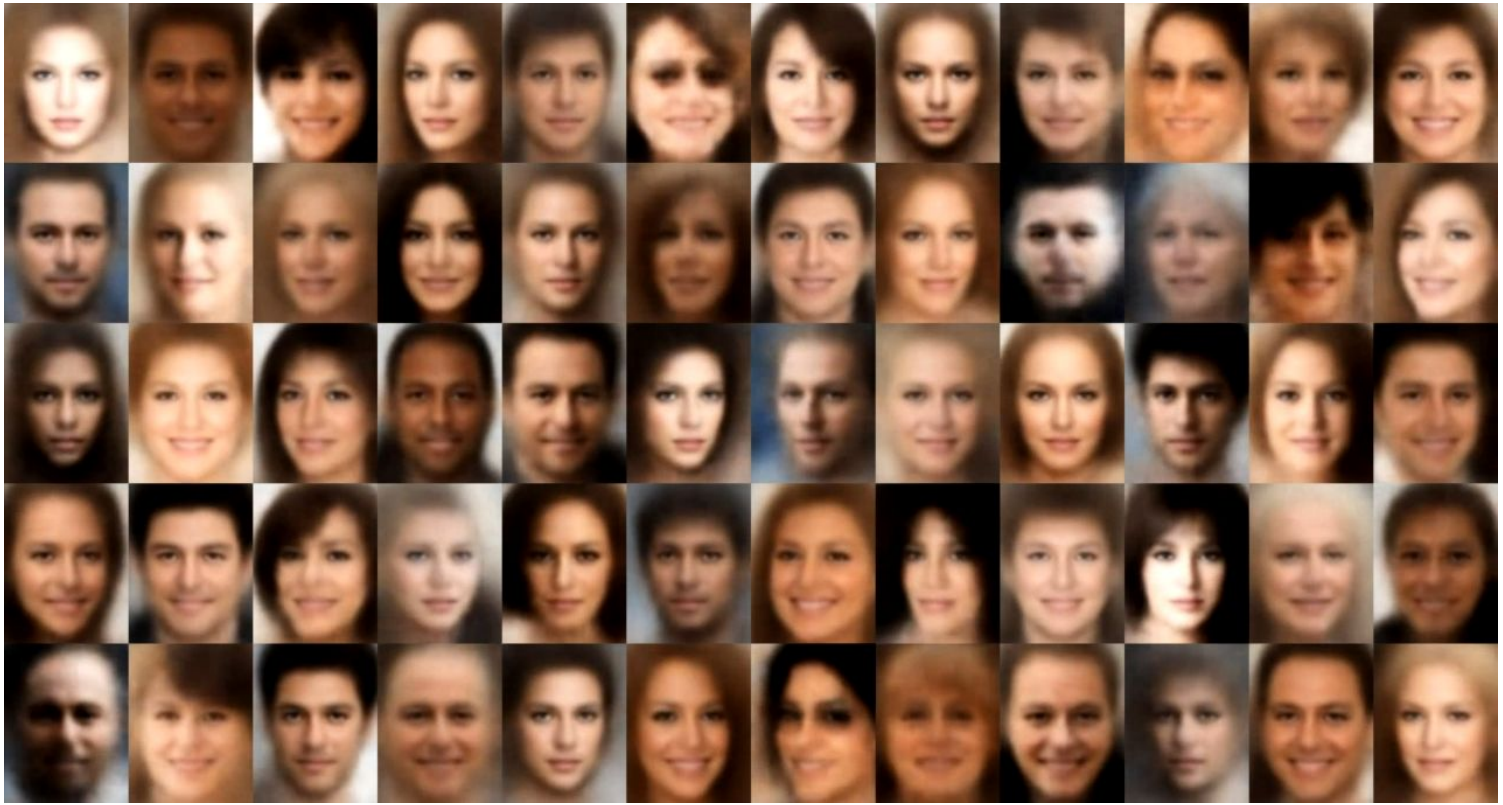


Pose Estimation

Generative Adversarial Networks

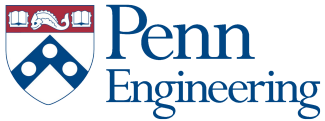


Variational Autoencoder ("Explainable AI")



Inverse Lego

Natural Language Processing Applications



Natural Language Processing

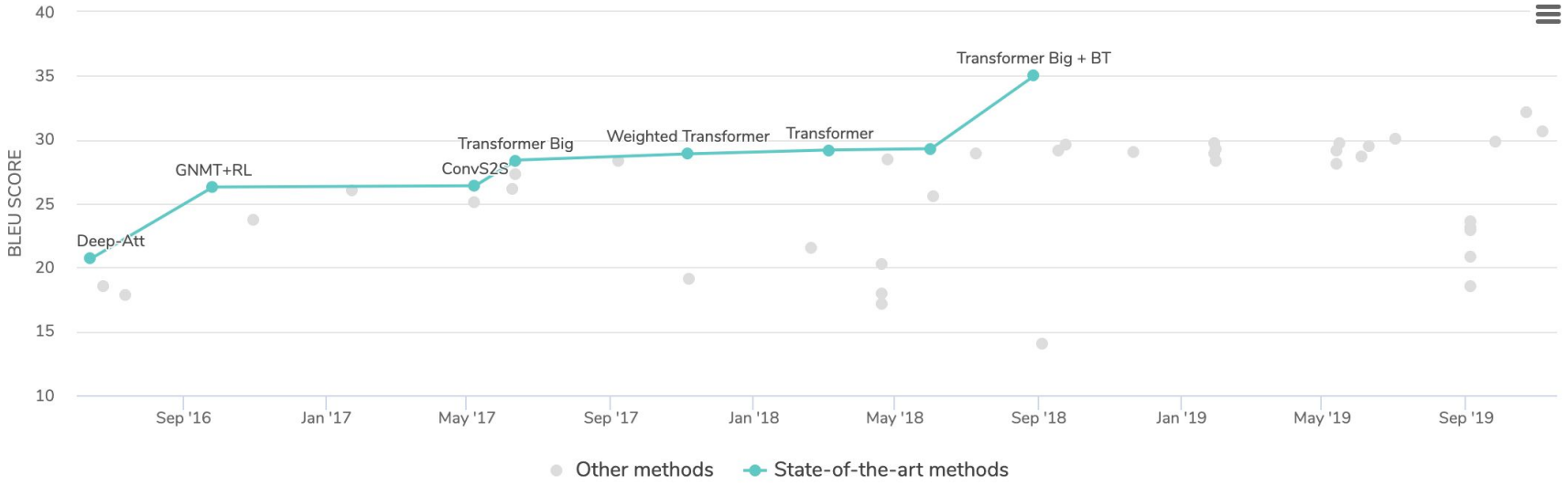
How can we computationally model language to perform useful tasks?

- Natural Language Generation
  - 2019 is thought by some researchers to be the “year of the transformer”
- Natural Language Understanding (NLU)
- Speech Recognition

Speech Recognition

- In 1952 Bell made “Audrey” -- could recognize individually spoken digits, 10 years later could recognize up to 16 words
- By the 1970’s, SOTA was able recognize over **1000** words! A 3 year old’s vocabulary
- 1980’s saw development of Hidden Markov Models, now allowing recognition of thousands of words
- Google stagnated in 2001 at 80% word accuracy

Machine translation as benchmark



aidungeon

https://play.aidungeon.io/

Penn Engineering

Claim extraction for science of science

Multiplex genome engineering using CRISPR/Cas systems.

Functional elucidation of causal genetic variants and elements requires precise genome editing technologies. The type II prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats)/Cas adaptive immune system has been shown to facilitate RNA-guided site-specific DNA cleavage. We engineered two different type II CRISPR/Cas systems and demonstrate that Cas9 nucleases can be directed by short RNAs to induce precise cleavage at endogenous genomic loci in human and mouse cells. Cas9 can also be converted into a nicking enzyme to facilitate homology-directed repair with minimal mutagenic activity. Lastly, multiple guide sequences can be encoded into a single CRISPR array to enable simultaneous editing of several sites within the mammalian genome, demonstrating easy programmability and wide applicability of the RNA-guided nuclease technology.

Prediction

Functional elucidation of causal genetic variants and elements requires precise genome editing technologies. BACKGROUND

The type II prokaryotic CRISPR (clustered regularly interspaced short palindromic repeats)/Cas adaptive immune system has been shown to facilitate RNA-guided site-specific DNA cleavage. BACKGROUND

We engineered two different type II CRISPR/Cas systems and demonstrate that Cas9 nucleases can be directed by short RNAs to induce precise cleavage at endogenous genomic loci in human and mouse cells. RESULTS

Cas9 can also be converted into a nicking enzyme to facilitate homology-directed repair with minimal mutagenic activity. CONCLUSIONS

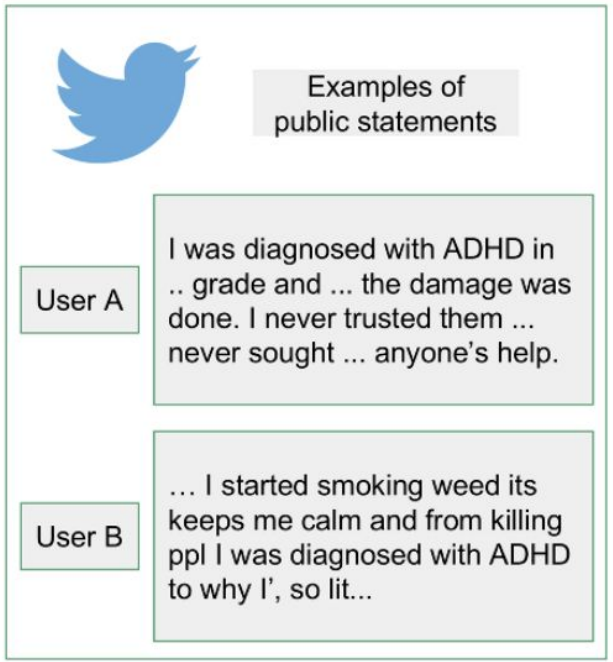
Lastly, multiple guide sequences can be encoded into a single CRISPR array to enable simultaneous editing of several sites within the mammalian genome, demonstrating easy programmability and wide applicability of the RNA-guided nuclease technology.
CONCLUSIONS CLAIM

Titipat Achakulvisut, Kording lab

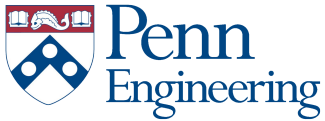
Do twitter users have ADHD?

Language of ADHD in Adults on Social Media

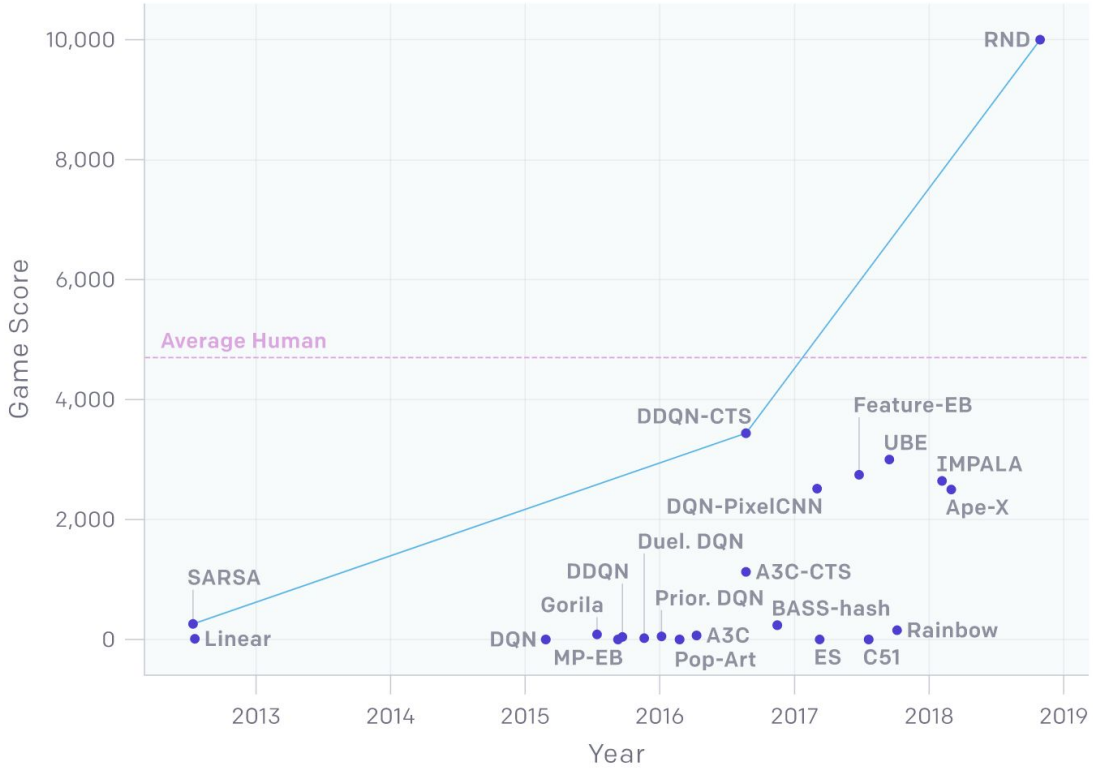
Sharath Chandra Guntuku[1], J. Russell Ramsay[1], Raina M. Merchant[1], and Lyle H. Ungar[1]



Reinforcement Learning Applications

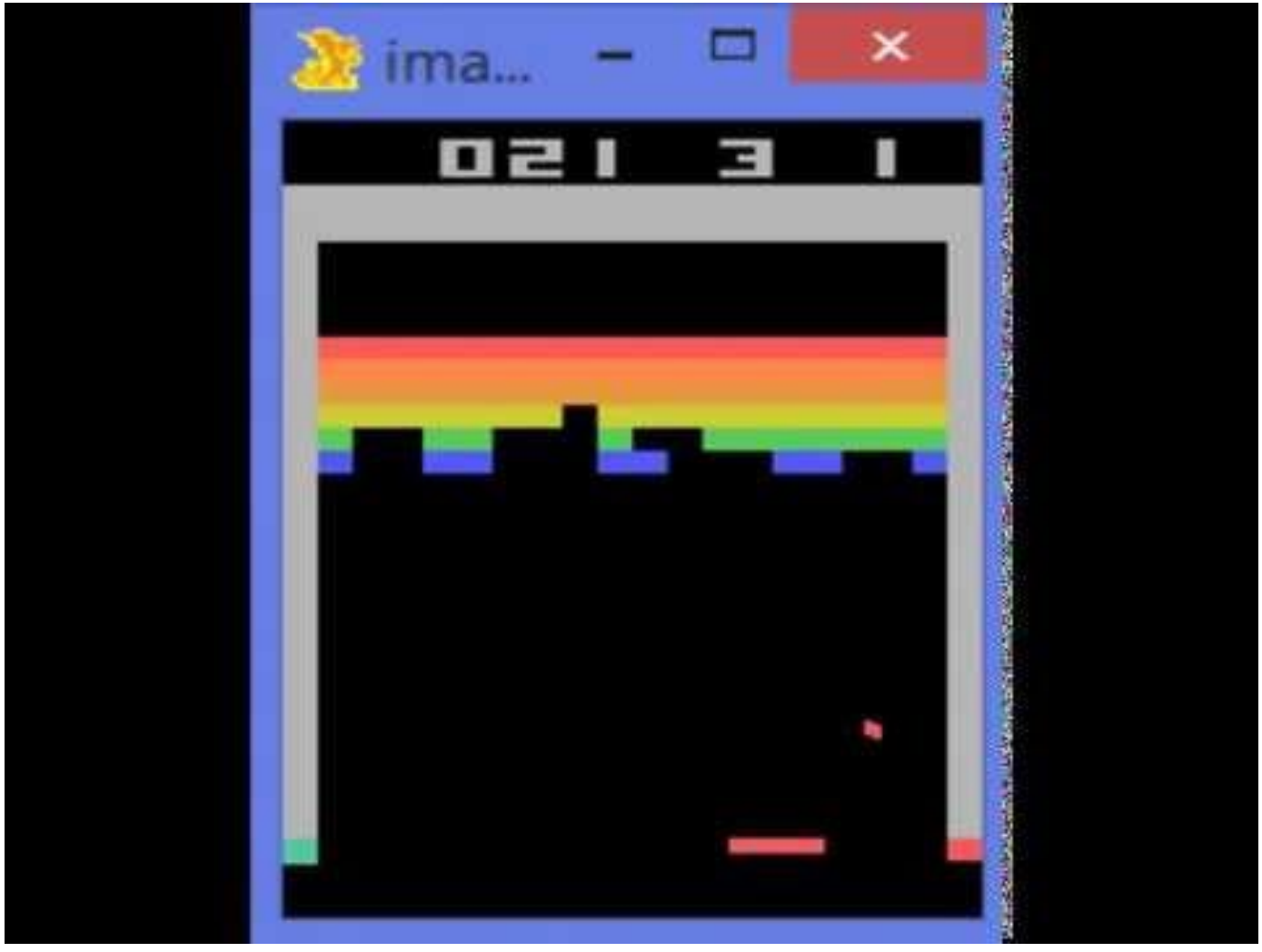


Progress at Montezuma's revenge



OpenAI

Atari Game playing



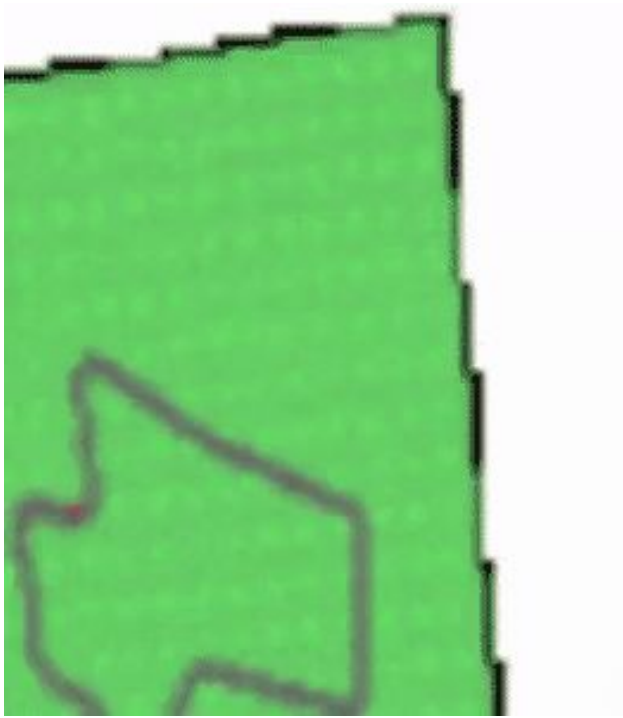
Deepmind

World Model

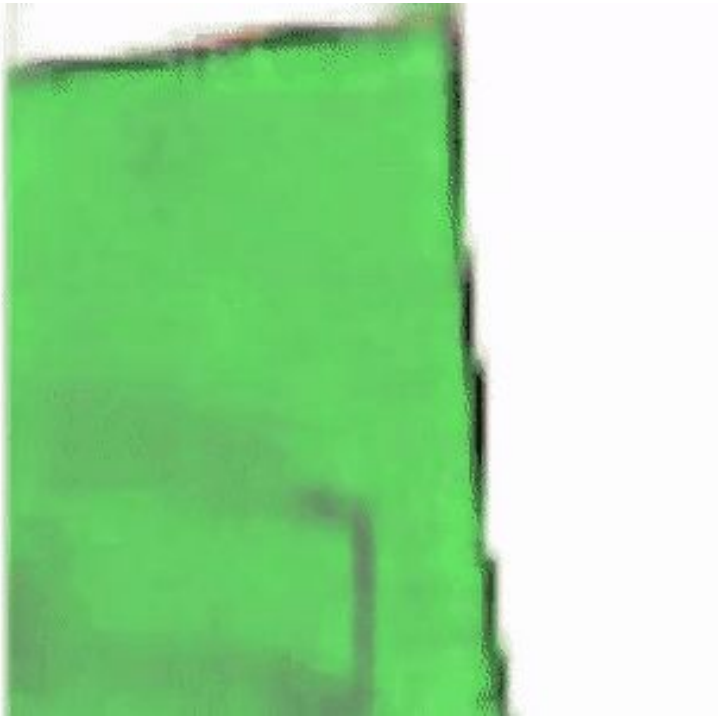
- Learn by dreaming!
- Mix of Generative and Reinforcement learning
- Train by generating new data through VAE's (Create a virtual training space a.k.a train inside your dreams)
- More information at (https://worldmodels.github.io)

World Model

Original



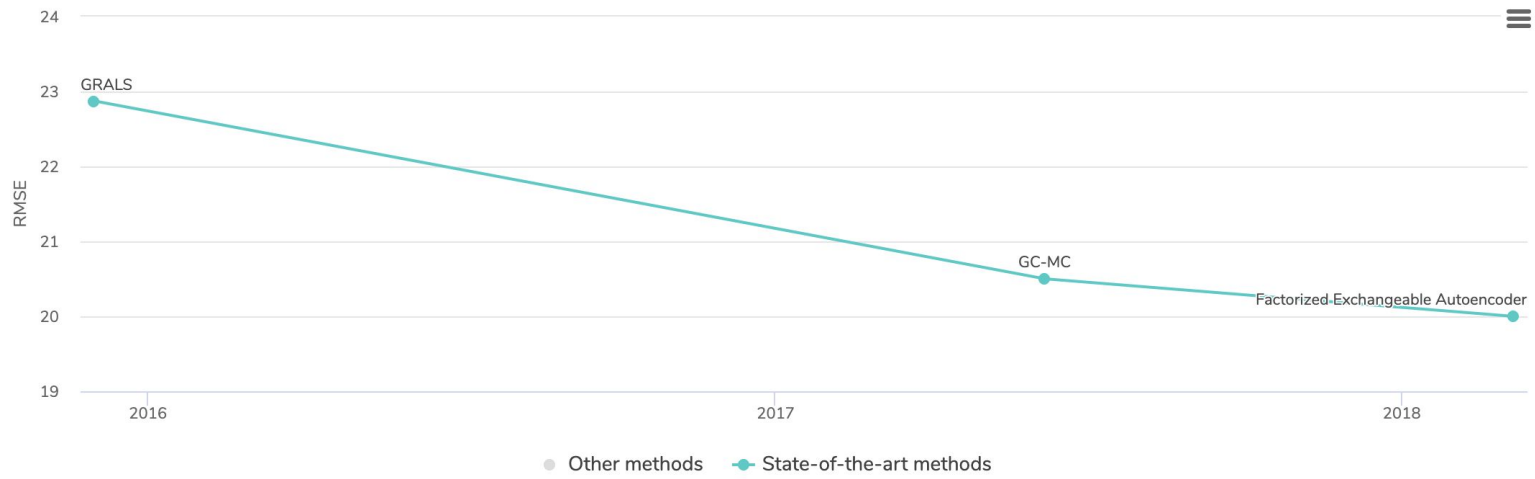
Dream



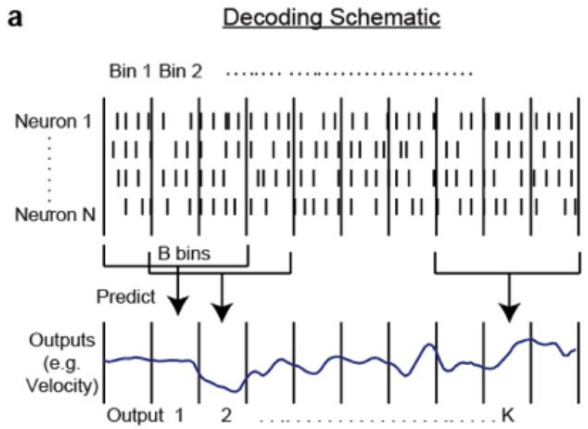
Other Applications

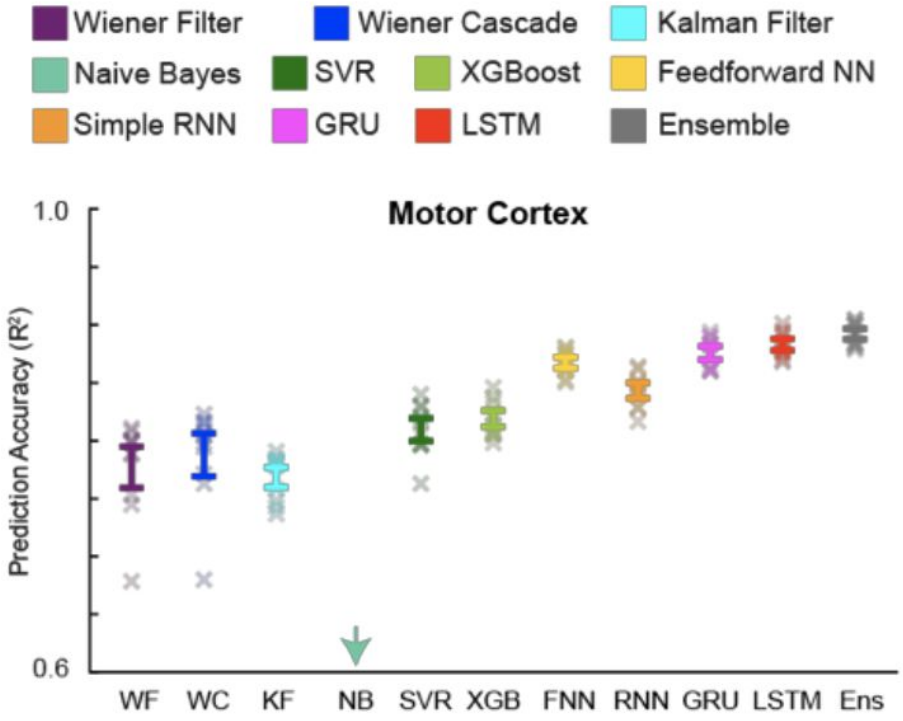
Penn Engineering

Recommendation Systems on YahooMusic



Neural decoding





Glaser … Kording, in press

Make this sound great

Theoretical Problems

The notorious “why does DL generalize question”

Neural Tangent Kernels

Rich vs Kernel Regimes

How many regions do ReLU networks have

Does SGD regularize DL systems?

Inverse DL

Can I give you DL system and you tell me how it works?

- Activation function?
- Loss function?
- Training data?
- Optimizer?

Toward good projects

Penn Engineering

Must be doable

Not too much compute

There must be an available dataset

Code to do similar things must exist

Should teach the muscles you want to have

Complicated

But not too complicated

In an area that you may envision working in in future

Must answer a real question/ solve a problem

If granny does not understand question it is not good

If the answer can not either be a product or a few sentences it is not a good question

Better if new

You can replicate a paper you admire

But better (and more points) if a new idea

- Something like 10% are for good (and new) ideas

Two examples

How Deep Learning Beat the SAT Reading Section	Dewang Sultania, Leo Murri, Sam Oshay, Peyman Norouzi	Learned to answer multiple-choice reading comprehension questions from the RACE dataset using logistic regression, CNNs, BERT, and Dual Co-Matching Networks. Interpreted the deep models by checking lexical, syntactic, and semantic data hypotheses.
Fill in the ____: Investigating the Machine Reading Comprehension Problem with Deep Learning	Brandon Lin, Yonah Mann, Rohan Menezes	Learned cloze-style question-answering: i.e. how to answer fill-in-the-blank queries after reading a related article. Used logistic regression, bidirectional LSTMs, and attentioned LSTMs to model CNN and Daily Mail articles.

PollEV questions again

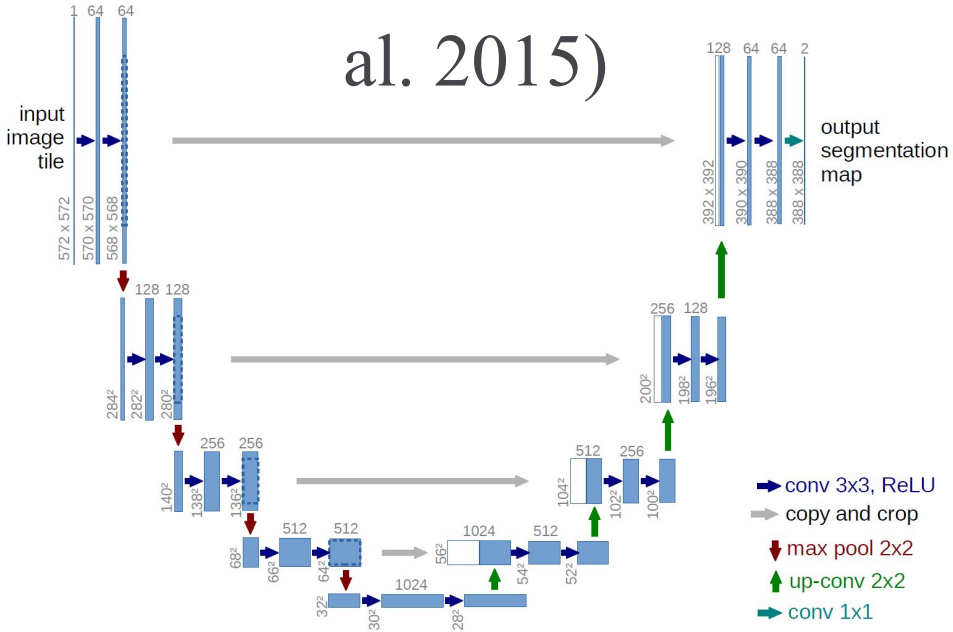
What we have done today

- Overview over application domains and how DL is great
  - Computer Vision
  - Natural Language processing
  - Reinforcement learning
  - Other applications
  - Theory
- Start thinking about and working towards projects
- Probably had a Q&A session

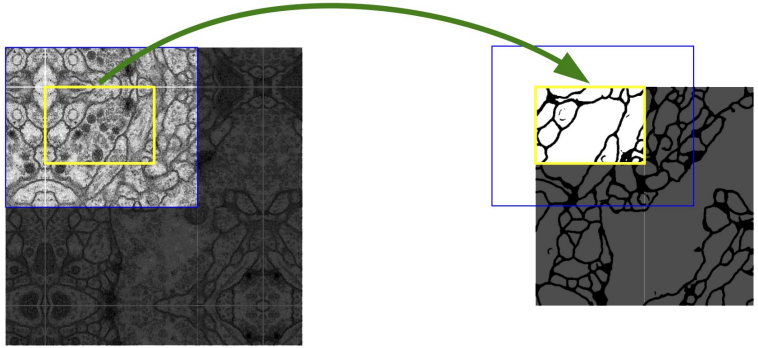
PollEV feedback

Image segmentation

U-Net (Ronneberger et al. 2015)



Segmenting biological cell membranes



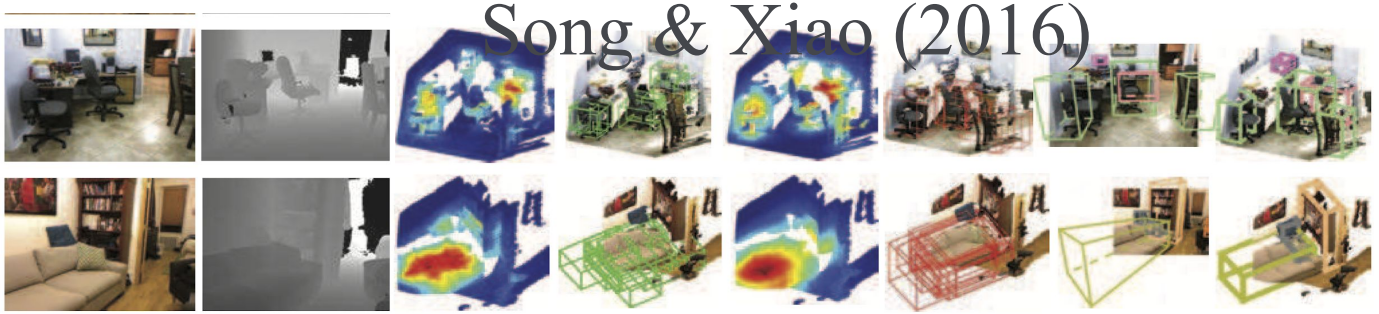
Insights of U-Net

- "Up-convolution"
  - Fixes the problem of shrinking images with CNNs
  - One example of how to make "fully convolutional" nets: pixels to pixels
  - "Up-convolution" is just upsampling, then convolution
  - Allows for refinement of the upsample by learned weights
  - Goes along with decreasing the number of feature channels
  - Not the same as "de-convolution".
- Connections across the "U" in the architecture.

3D convolution

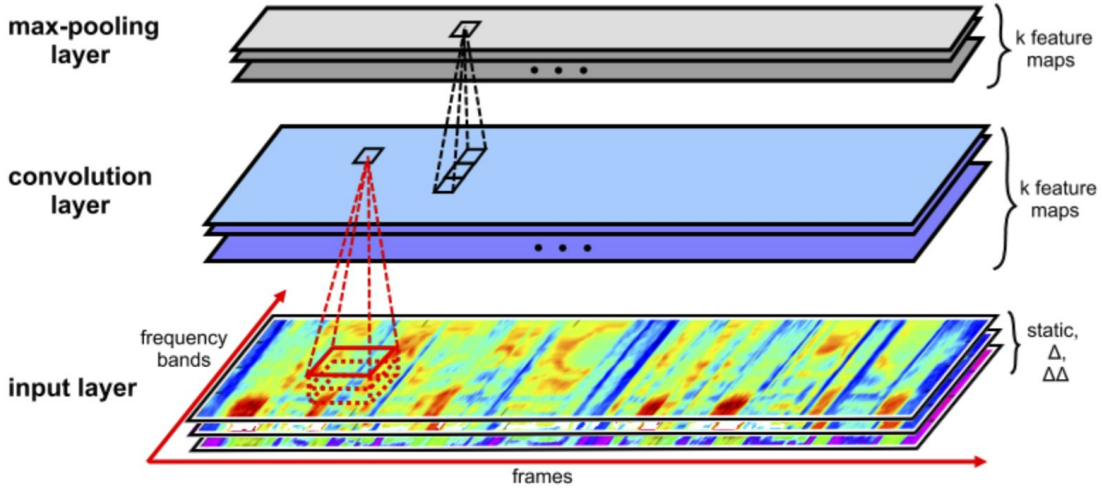
Motion in videos - Tran et al. (2015)

Objects from images w/ depths - Song & Xiao (2016)



Speech recognition

Zhang et al. 2017



- CNN competitive with RNNs (e.g. LSTMs)
- 10 layers, 3x5 conv, 3x1 pooling - deep enough for

Myth: CNNs are for computer vision, and RNNs are for NLP.

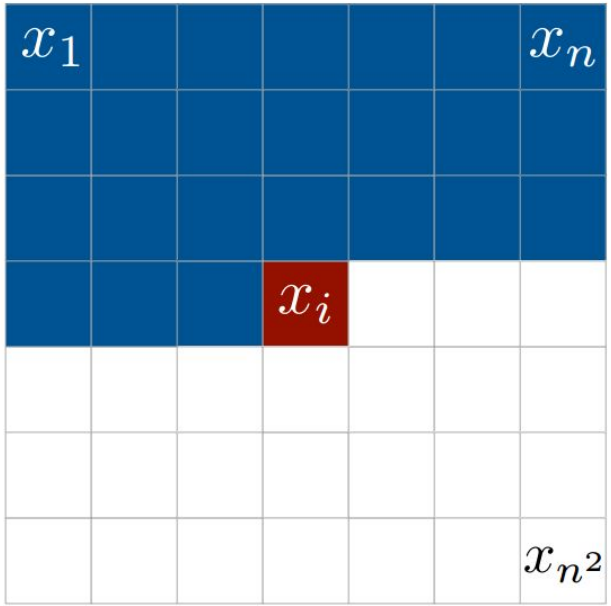
CNNs are SOTA in many NLP tasks

3. Text Classification

Research Paper	Datasets	Metric	Source Code	Year
Learning Structured Text Representations	Yelp	Accuracy: 68.6	• Tensorflow	2017
Attentive Convolution	Yelp	Accuracy: 67.36	• Theano	2017

Source:
https://github.com/bentrevett/pytorch-sentiment-analysis/blob/master/4%20-%20Convolutional%20Sentiment%20Analysis.ipynb

Pixel RNN - RNN's in Computer Vision!



Predicts next pixel by looking at all the previous pixels generated

Image Source: https://towardsdatascience.com/auto-regressive-generative-models-pixelrnn-pixelcnn-32d192911173

Pixel RNN - RNN's in Computer Vision!

Figure 1. Image completions sampled from a PixelRNN.

Image Source: https://towardsdatascience.com/auto-regressive-generative-models-pixelrnn-pixelcnn-32d192911173

Object Detection

Bounding-Box Object Detection

Object detection

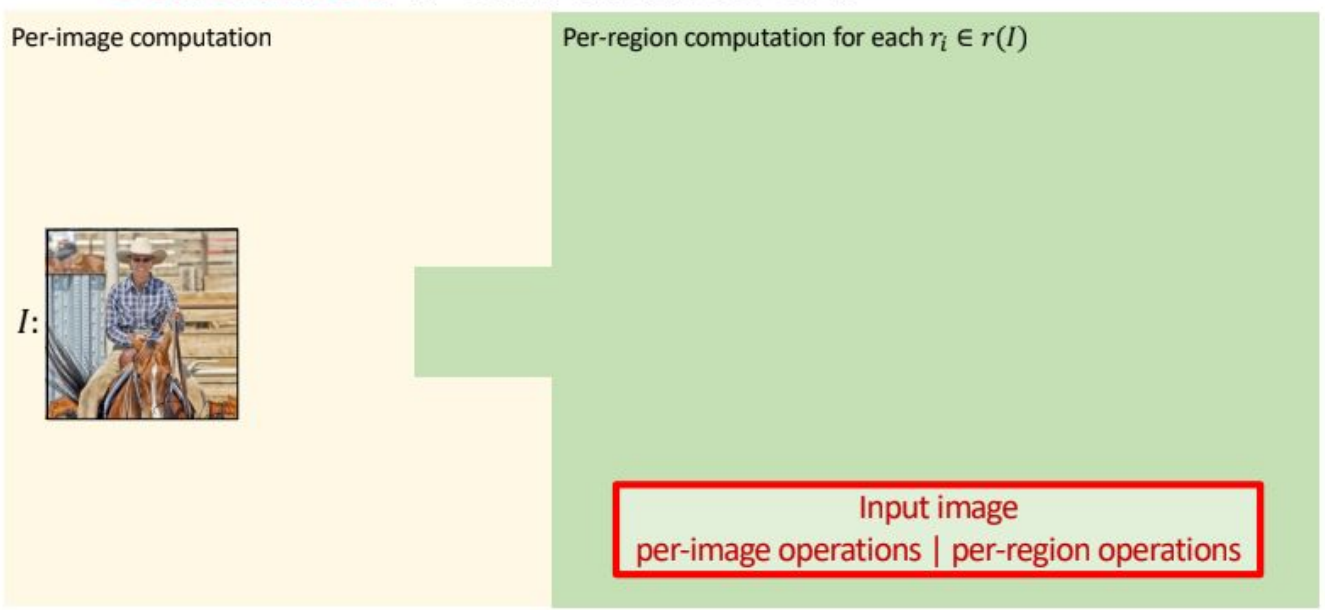
Object Detection

Detection Beyond Bounding Boxes

Mask R-CNN
[He, Gkioxari, Dollár, Girshick]

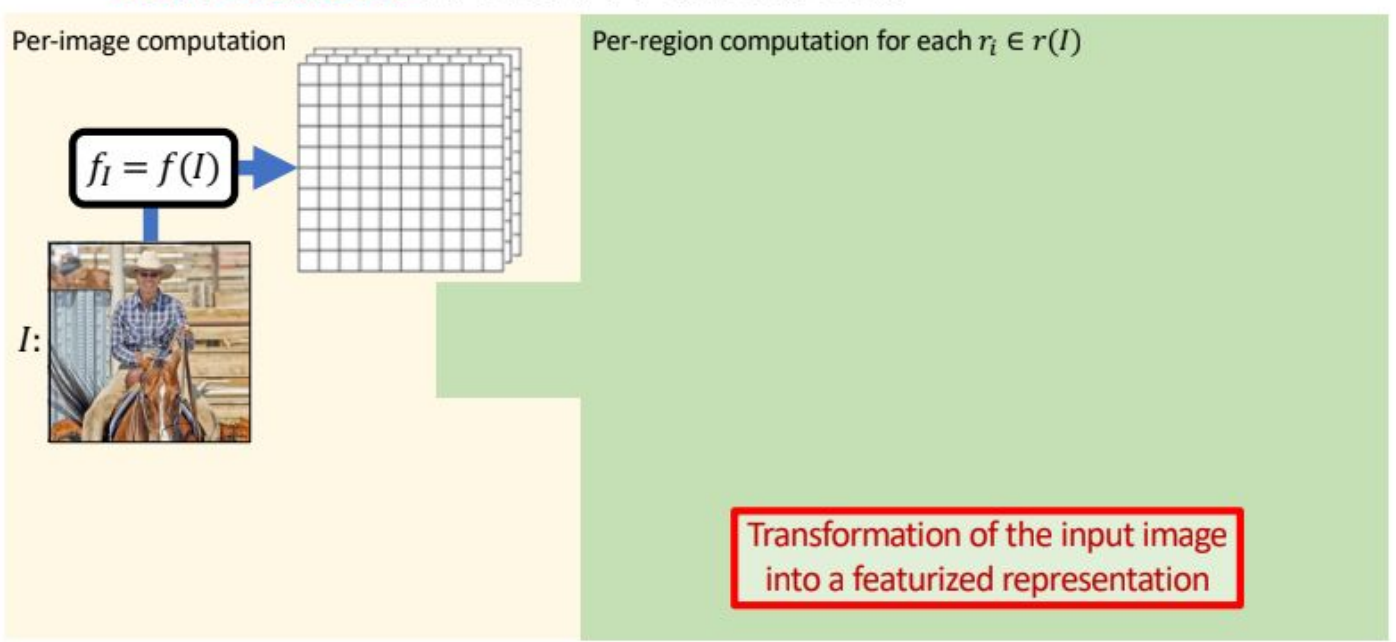
- Detects the objects.
- Does Semantic Segmentation for each object.

Generalized R-CNN framework for Object Detection



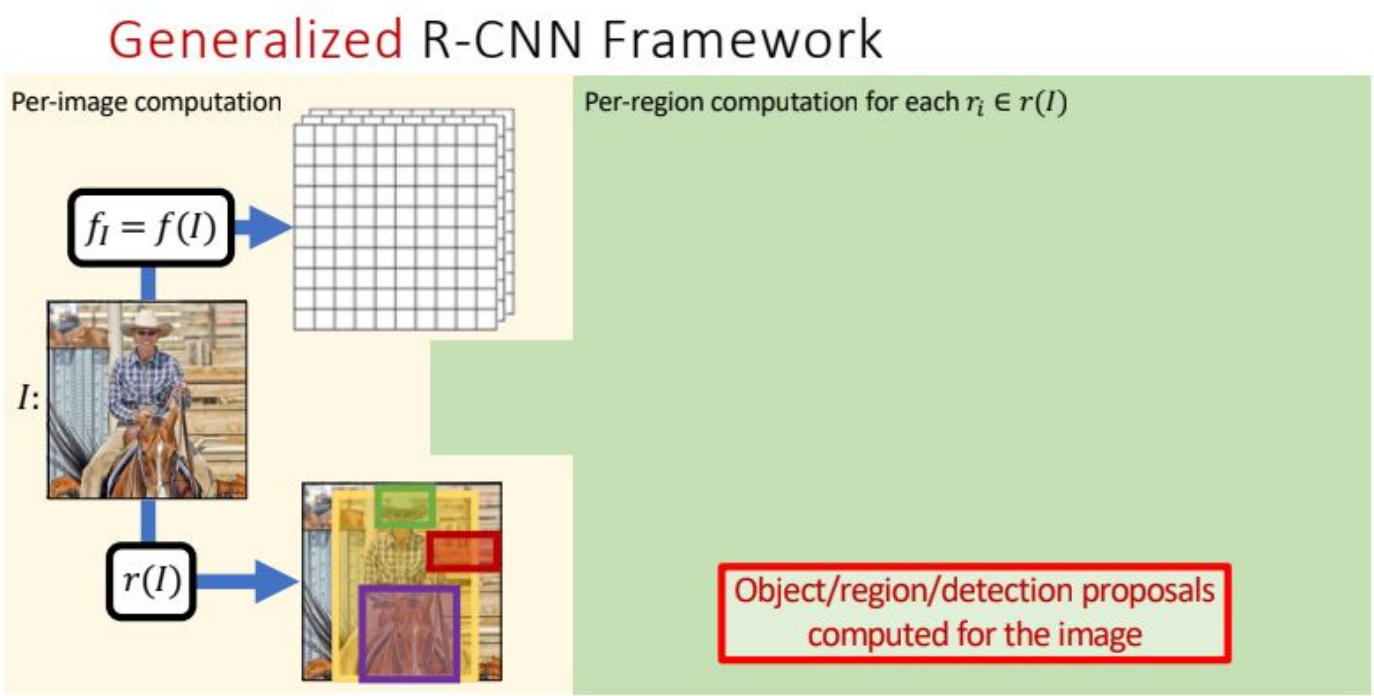
https://drive.google.com/file/d/1Z1i8mk8AqdtI_Q8avqtrzLQnLTG7QW6T/view

Generalized R-CNN framework for Object Detection



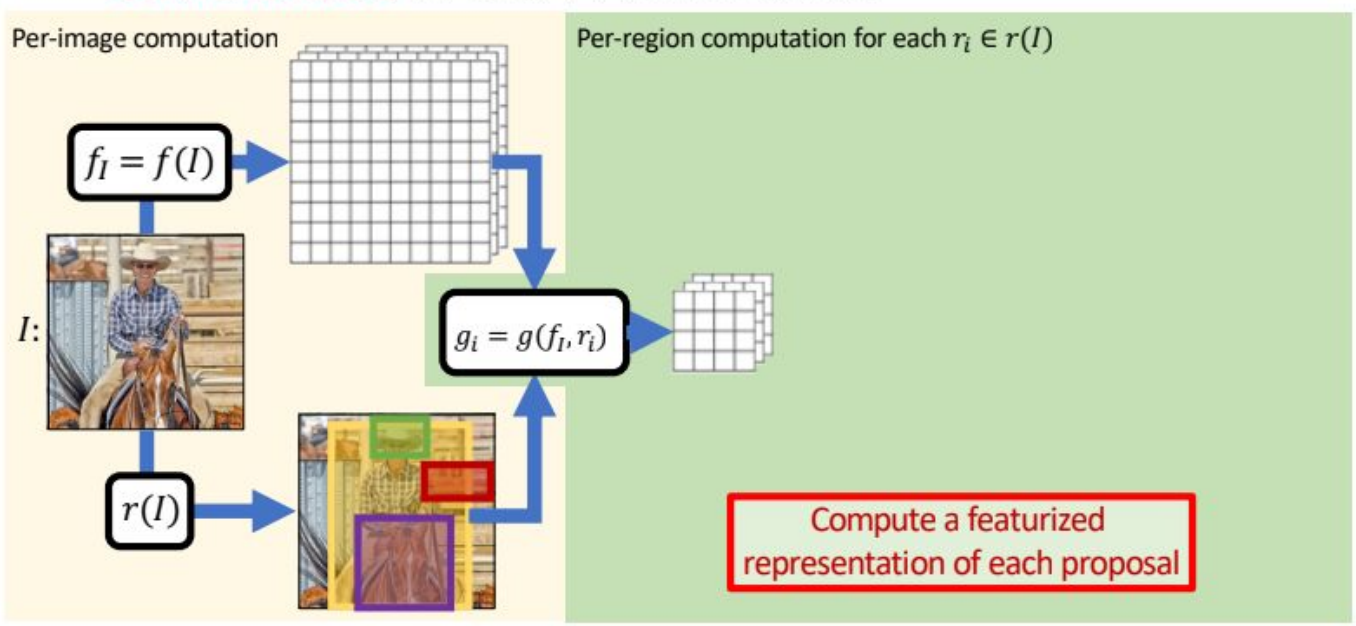
https://drive.google.com/file/d/1Z1i8mk8AqdtI_Q8avqtrzLQnLTG7QW6T/view

Generalized R-CNN framework for Object Detection



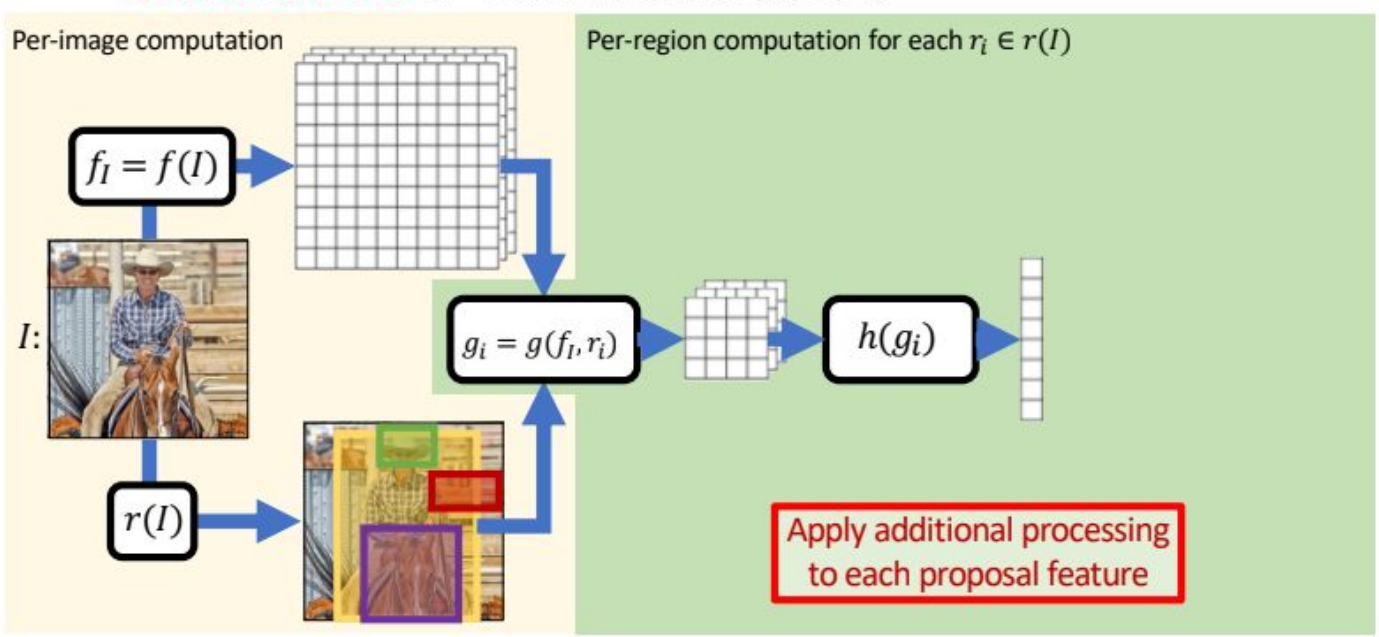
https://drive.google.com/file/d/1Z1i8mk8AqdtI_Q8avqtrzLQnLTG7QW6T/view

Generalized R-CNN framework for Object Detection



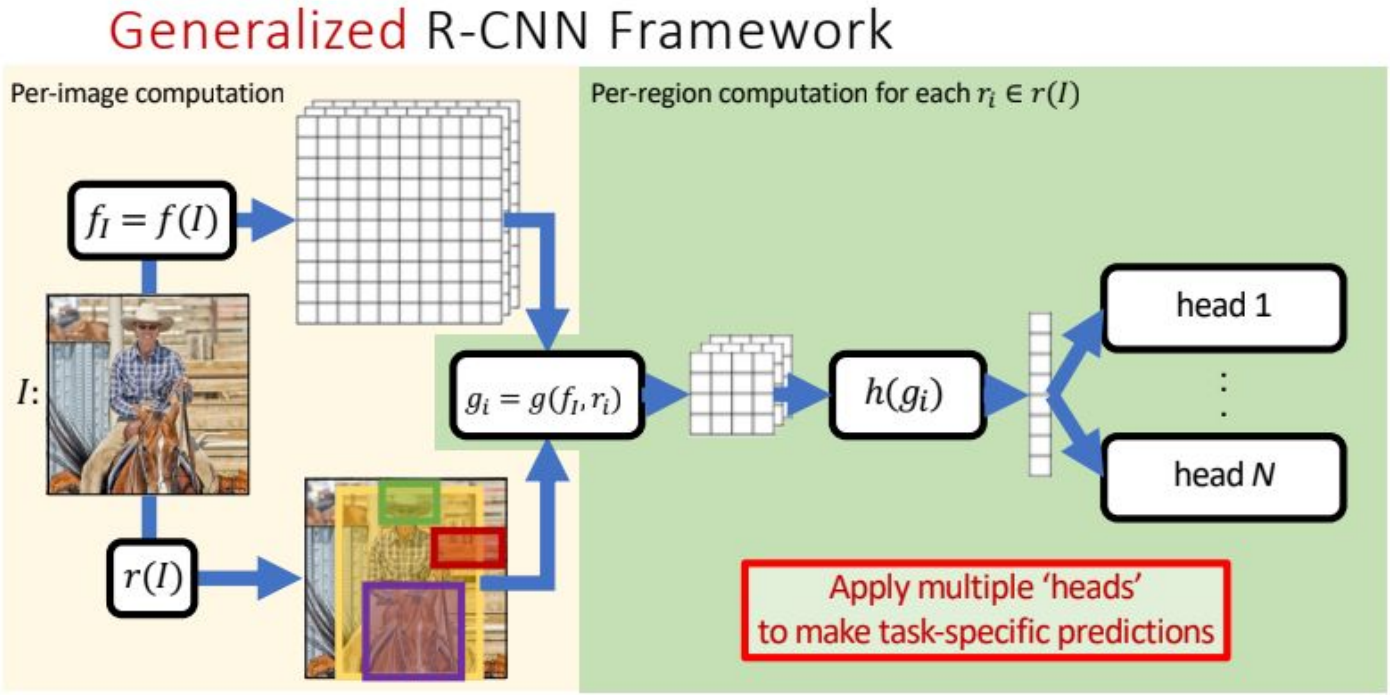
https://drive.google.com/file/d/1Z1i8mk8AqdtI_Q8avqtrzLQnLTG7QW6T/view

Generalized R-CNN framework for Object Detection



https://drive.google.com/file/d/1Z1i8mk8AqdtI_Q8avqtrzLQnLTG7QW6T/view

Generalized R-CNN framework for Object Detection

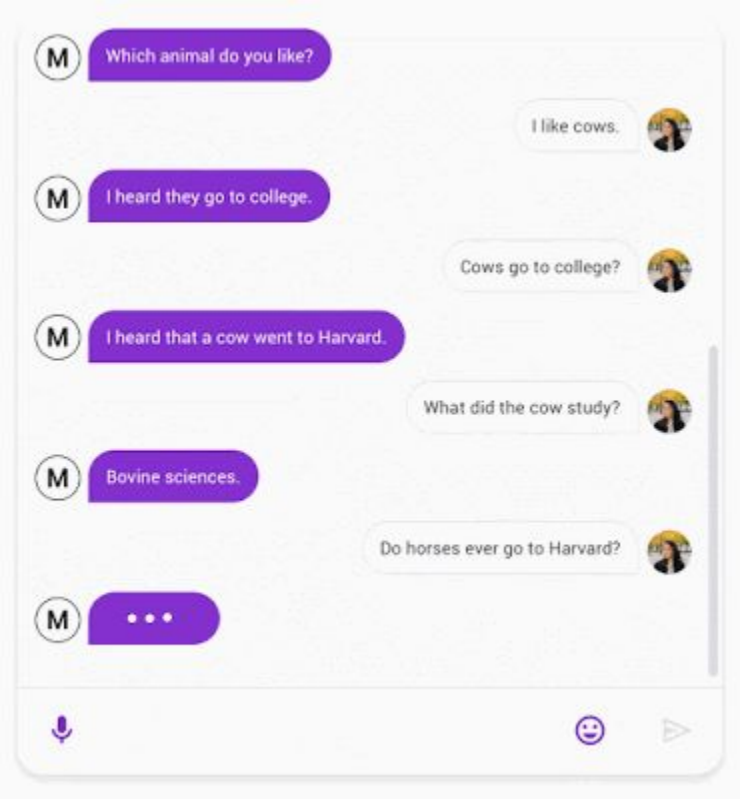


https://drive.google.com/file/d/1Z1i8mk8AqdtI_Q8avqtrzLQnLTG7QW6T/view

Deep Learning Domains

Penn Engineering

Transformers



Speech Recognition -- Now