

Introduction to DL & Applications

Subash Gandyer
Data Scientist, HealthChain

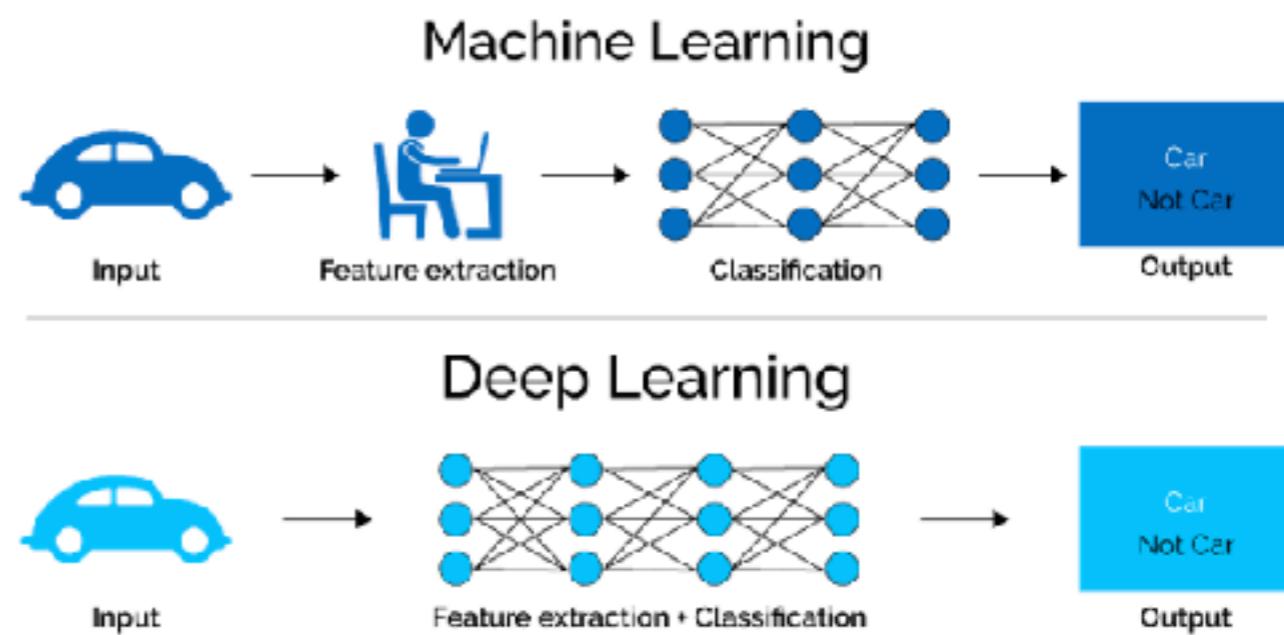
Agenda

- Deep Learning Concepts & Terminologies
- Deep Learning Applications
- Deep Learning Limitations

Concepts & Terminologies

What is Deep Learning

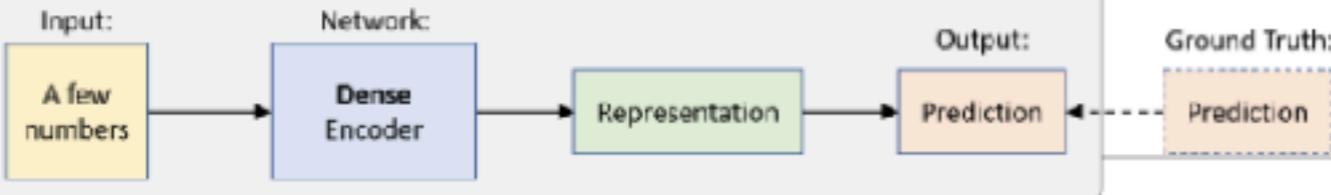
1. Deep Learning means using a neural network with several layers of nodes between input and output
2. The series of layers between input & output do feature identification and processing in a series of stages just as our brains seem to.



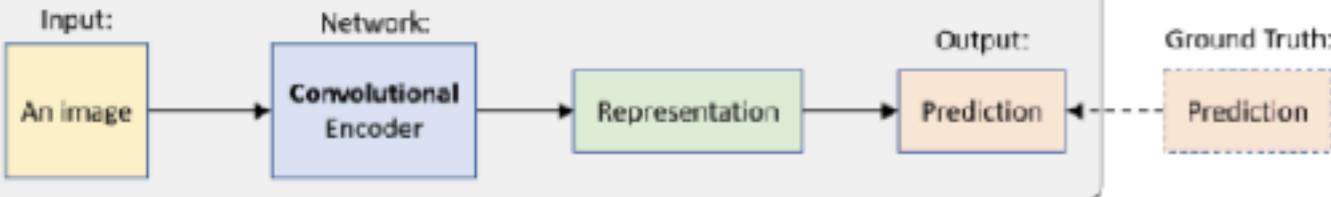
Deep Learning Architectures

Supervised Learning

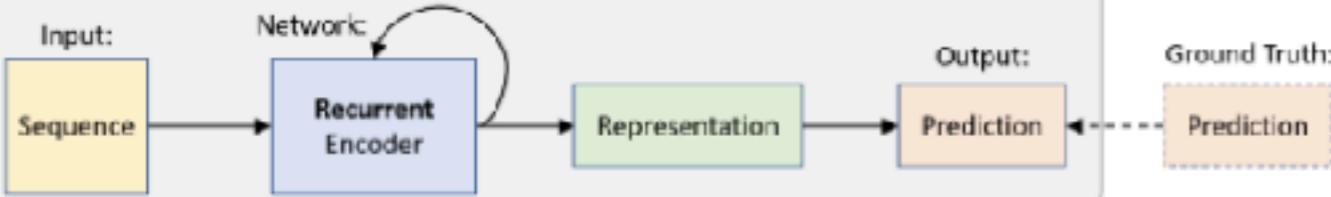
1. Feed Forward Neural Networks



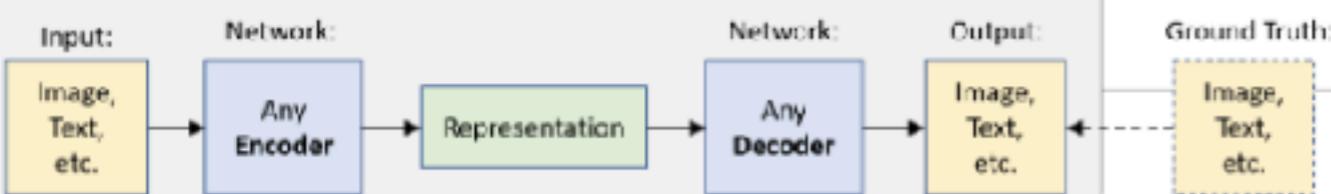
2. Convolutional Neural Networks



3. Recurrent Neural Networks



4. Encoder-Decoder Architectures

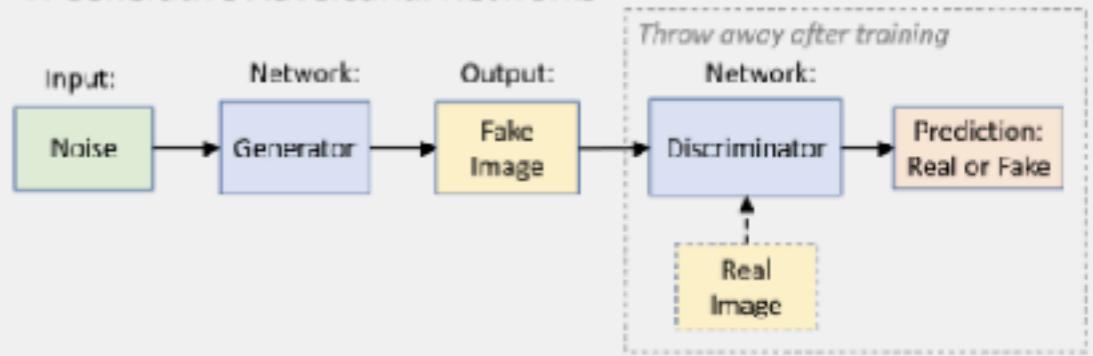


Unsupervised Learning

5. Autoencoder



6. Generative Adversarial Networks



Reinforcement Learning

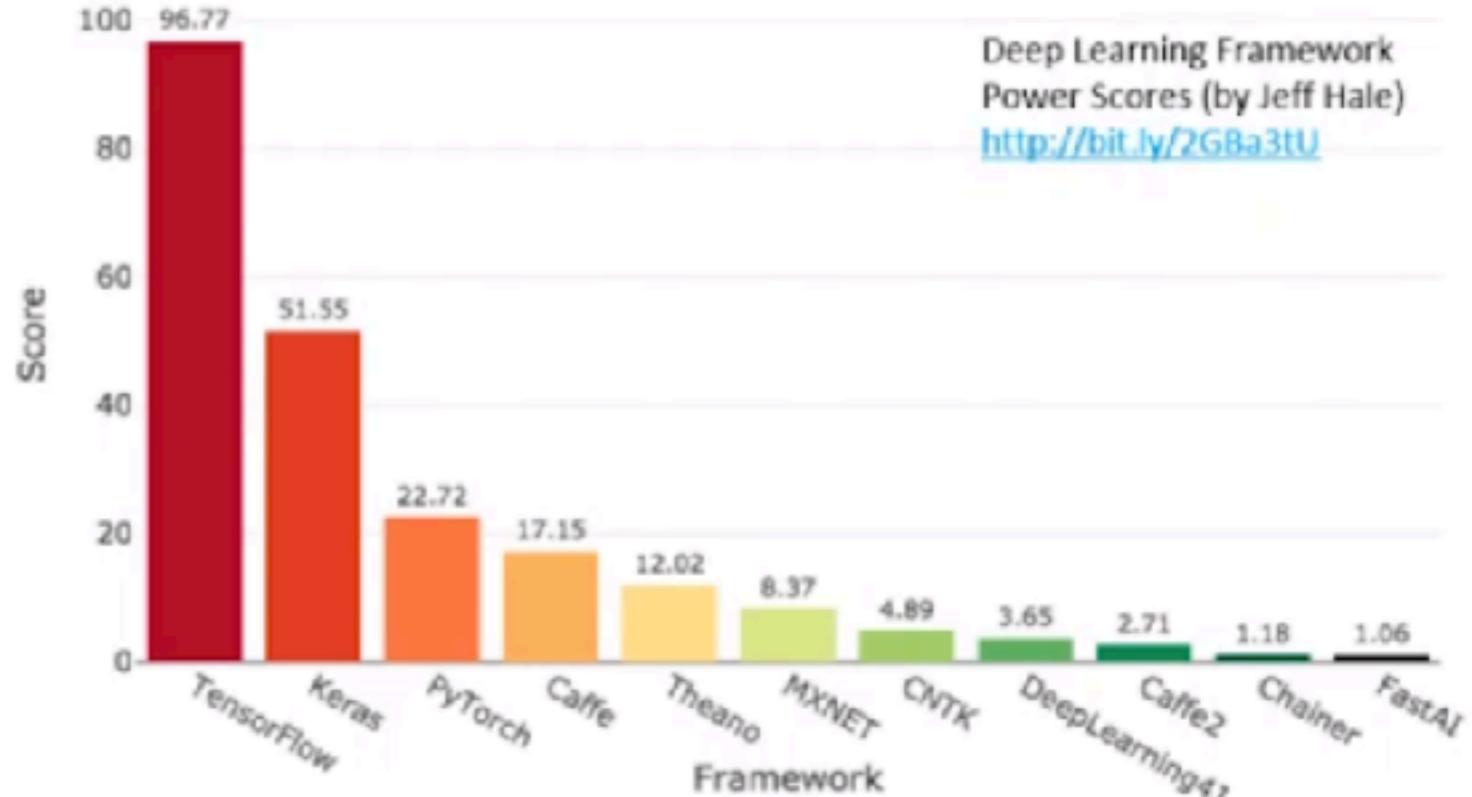
7. Networks for Actions, Values, Policies, and Models



Frameworks

Deep Learning Frameworks

- Pytorch (facebook)
- Tensorflow (google)
- MXNet (amazon)
- Chainer
- CNTK (microsoft)



Neuron

Weights

Bias

Activation Function

Activation Functions

- Sigmoid
- tanH
- ReLU
- Leaky ReLU
- Softmax

Datasets

Types of Datasets

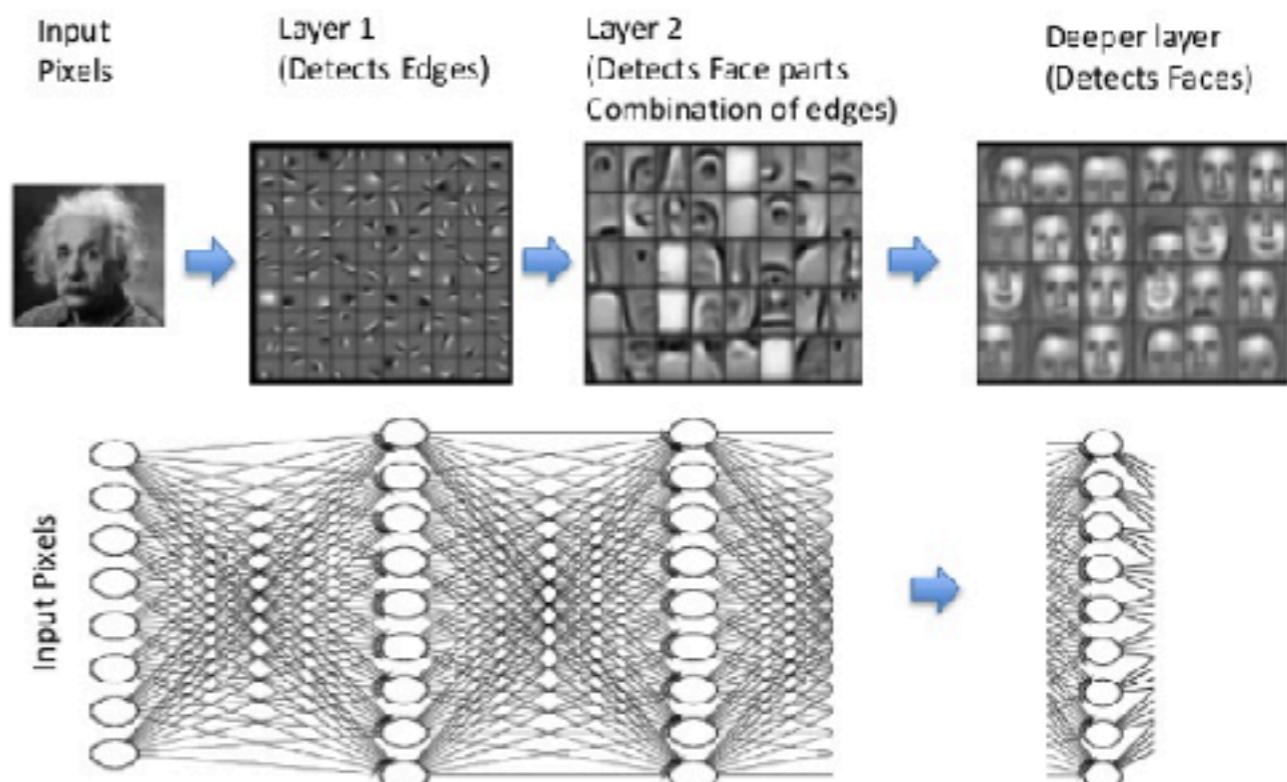
- **Training set** - we train the system on this data
- **Development set** - Used to tune parameters, select features etc
- **Test set** - used to evaluate the performance of the algorithm but not to make decisions about parameters

Neural Network

Feature Representation

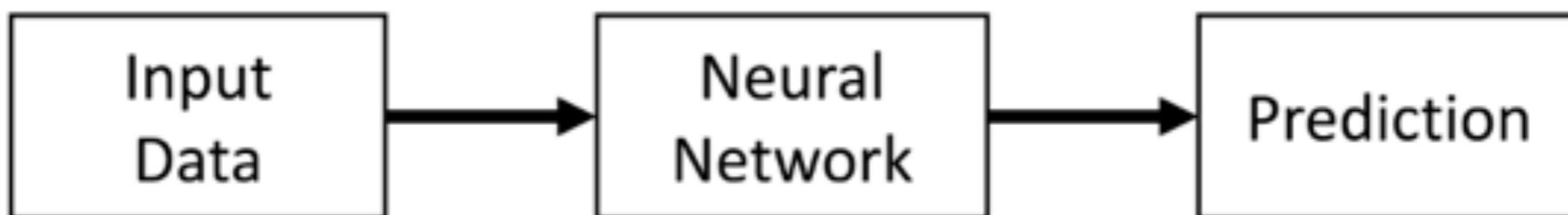
- Deep Networks Learn (re)usable features
- Usually deeper layers produce more dataset/task specific features

Feature Learning/Representation Learning
(Ex. Face Detection)

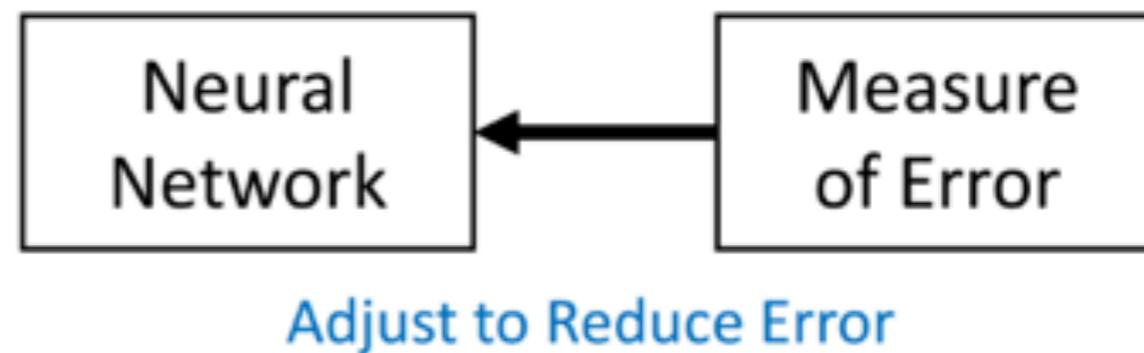


How Neural Networks Learn: Backpropagation

Forward Pass:



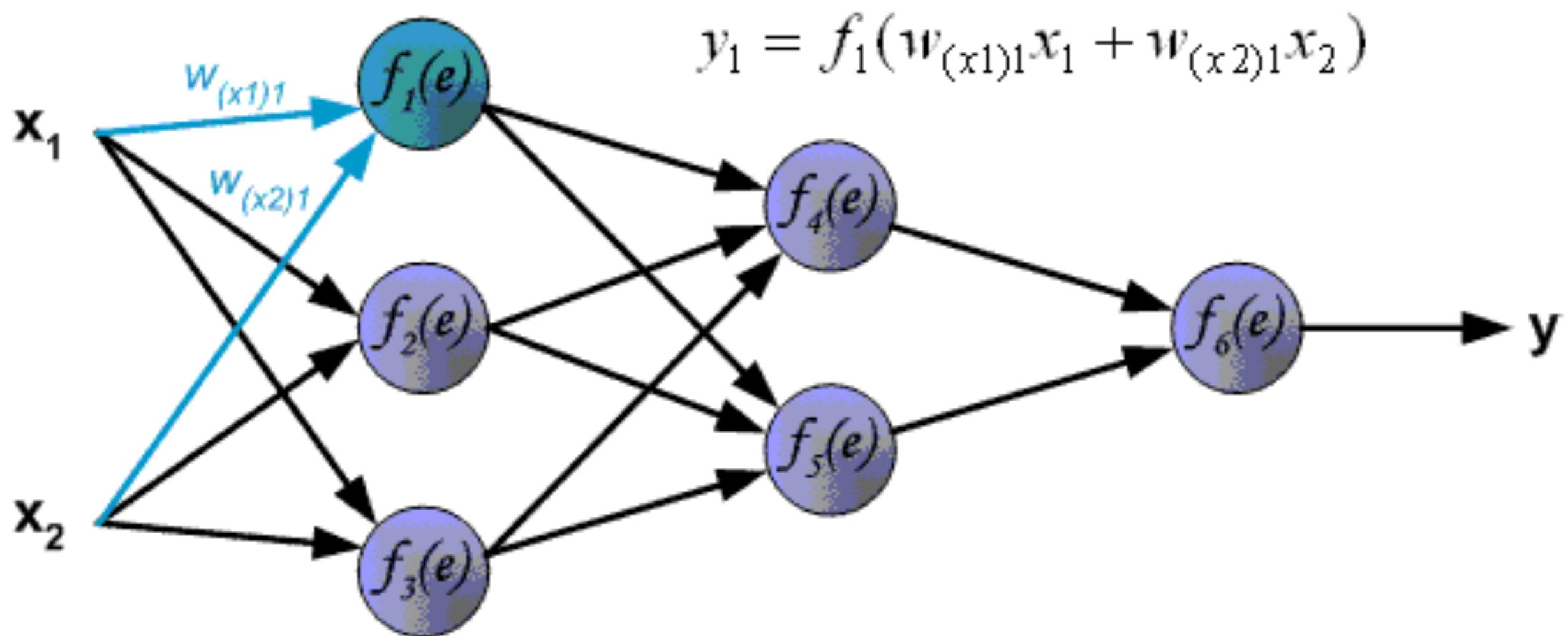
Backward Pass (aka Backpropagation):



Forward Propagation

Backward Propagation

FP



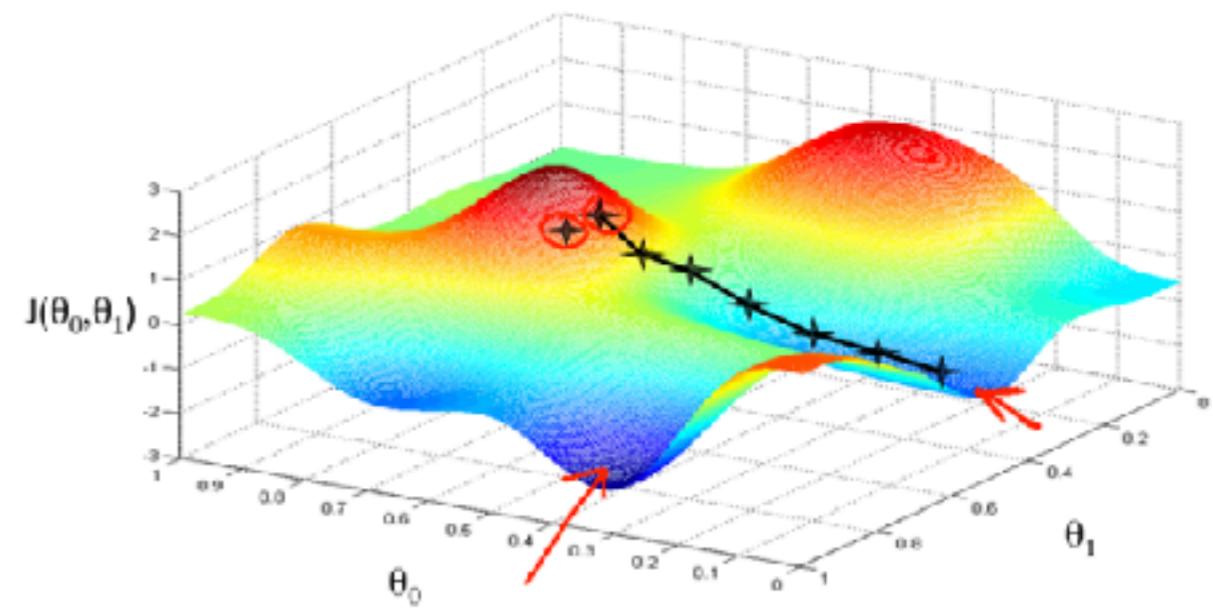
Loss Function

Cost Function

Gradient Descent

Gradient Descent (Updating the Weights)

- The point of GD (gradient descent) is to minimize the cost/loss function this means finding such combination of weights and biases that will give us the lowest possible loss
- To find the lowest error(deepest valley) in the cost function (with respect to one weight), we need to tweak the parameters of the model. How much do we tweak them though? Enter Calculus. Using calculus, we know that the slope of a function is the derivative of the function with respect to a value. This slope always points to the nearest valley

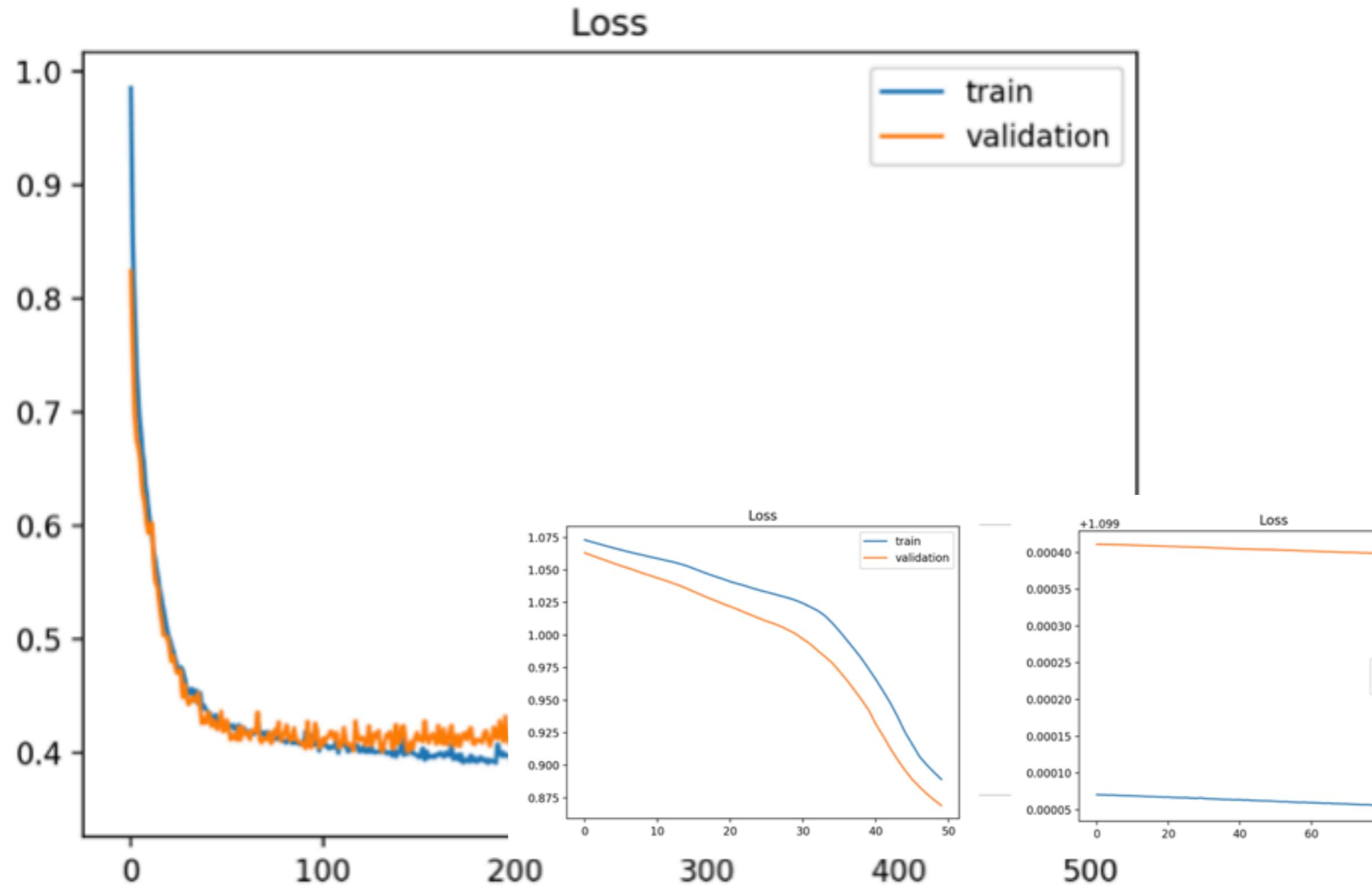


Gradient Descent Algorithms

- Batch Gradient Descent
- Stochastic Gradient Descent
- Mini-batch Gradient Descent

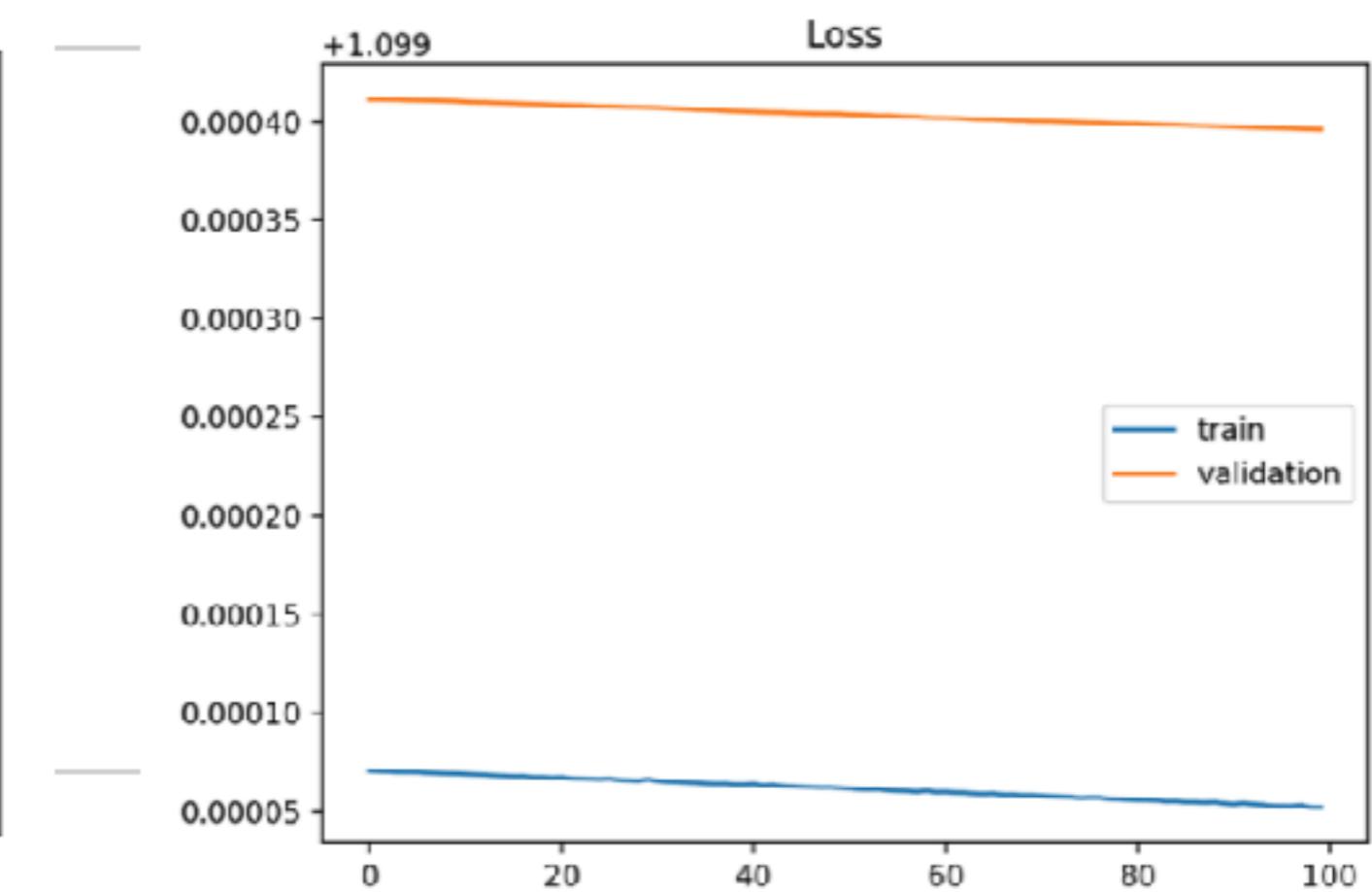
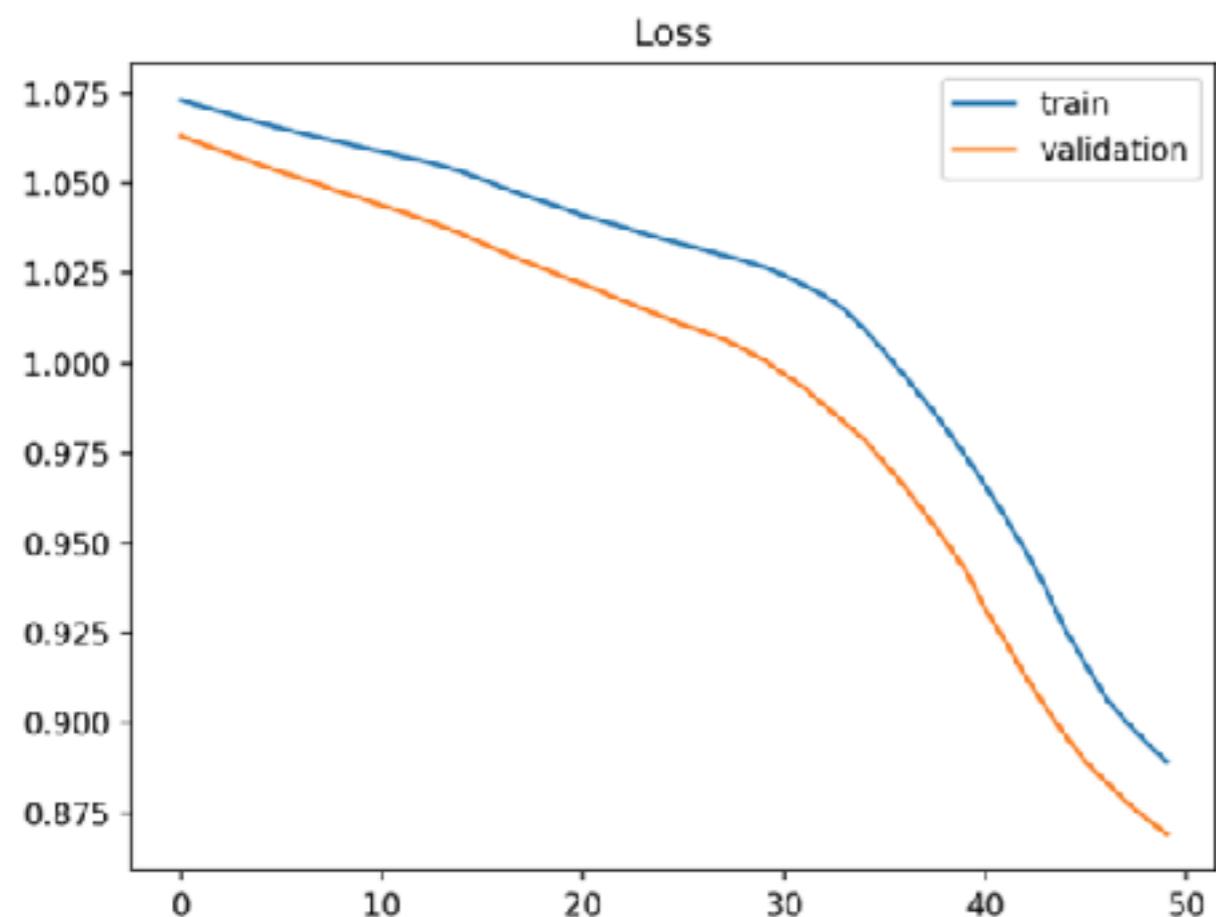
Overfitting

Overfitting



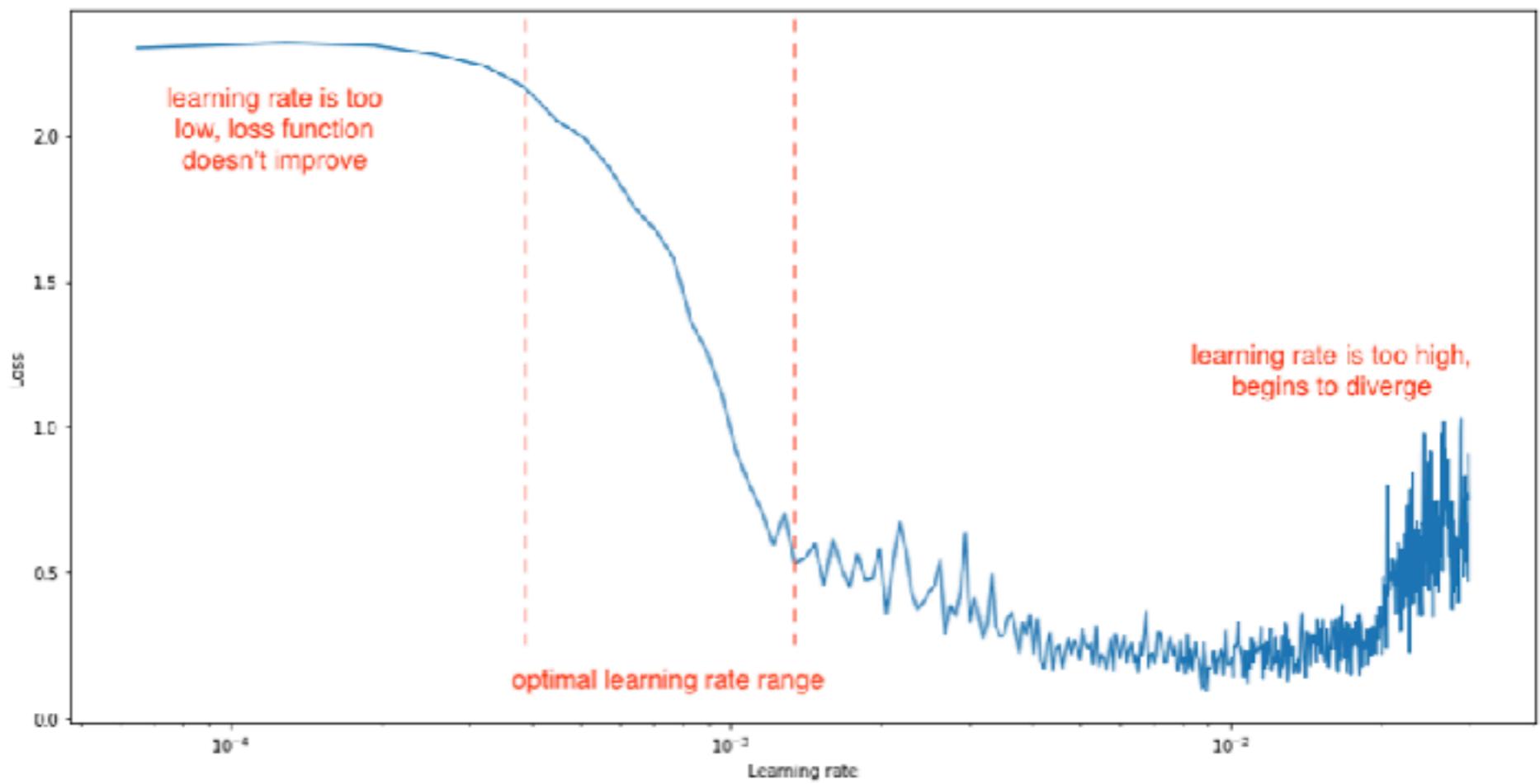
Underfitting

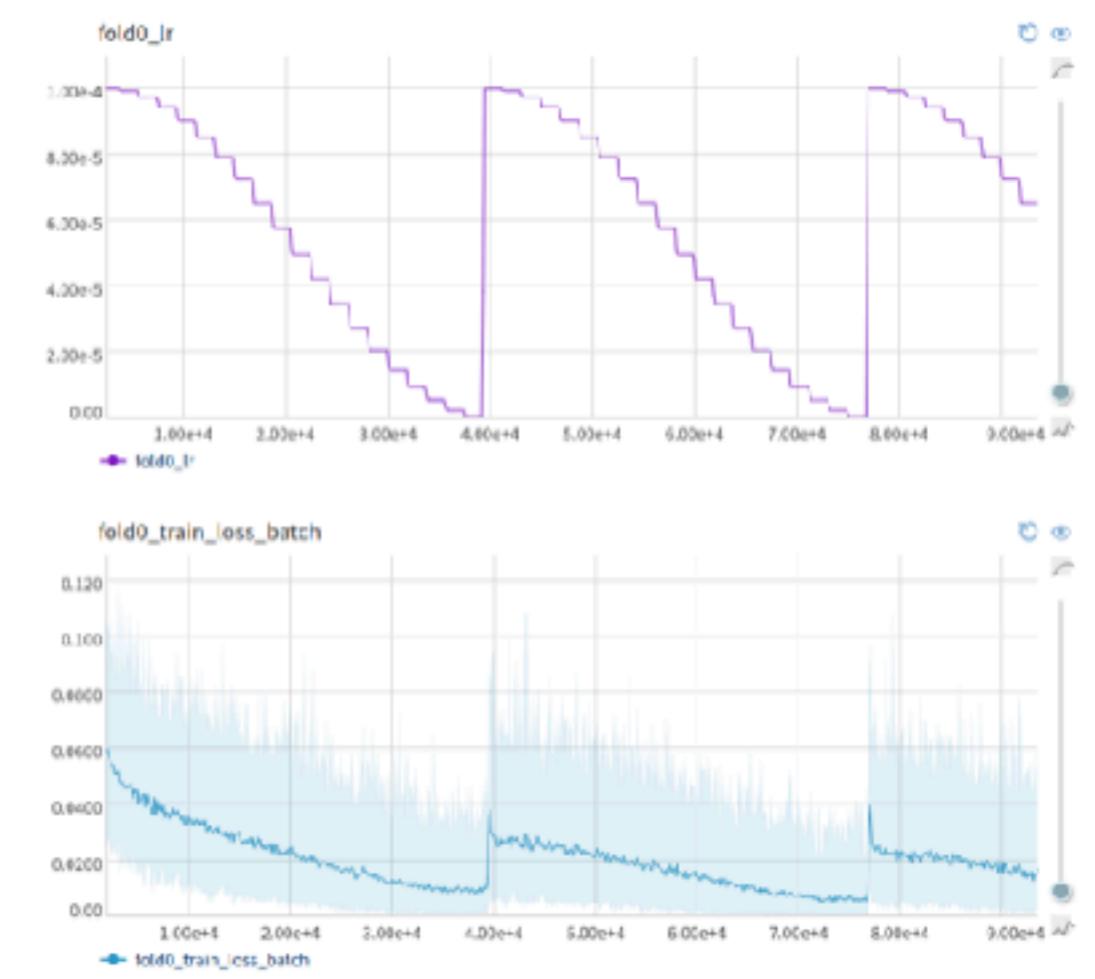
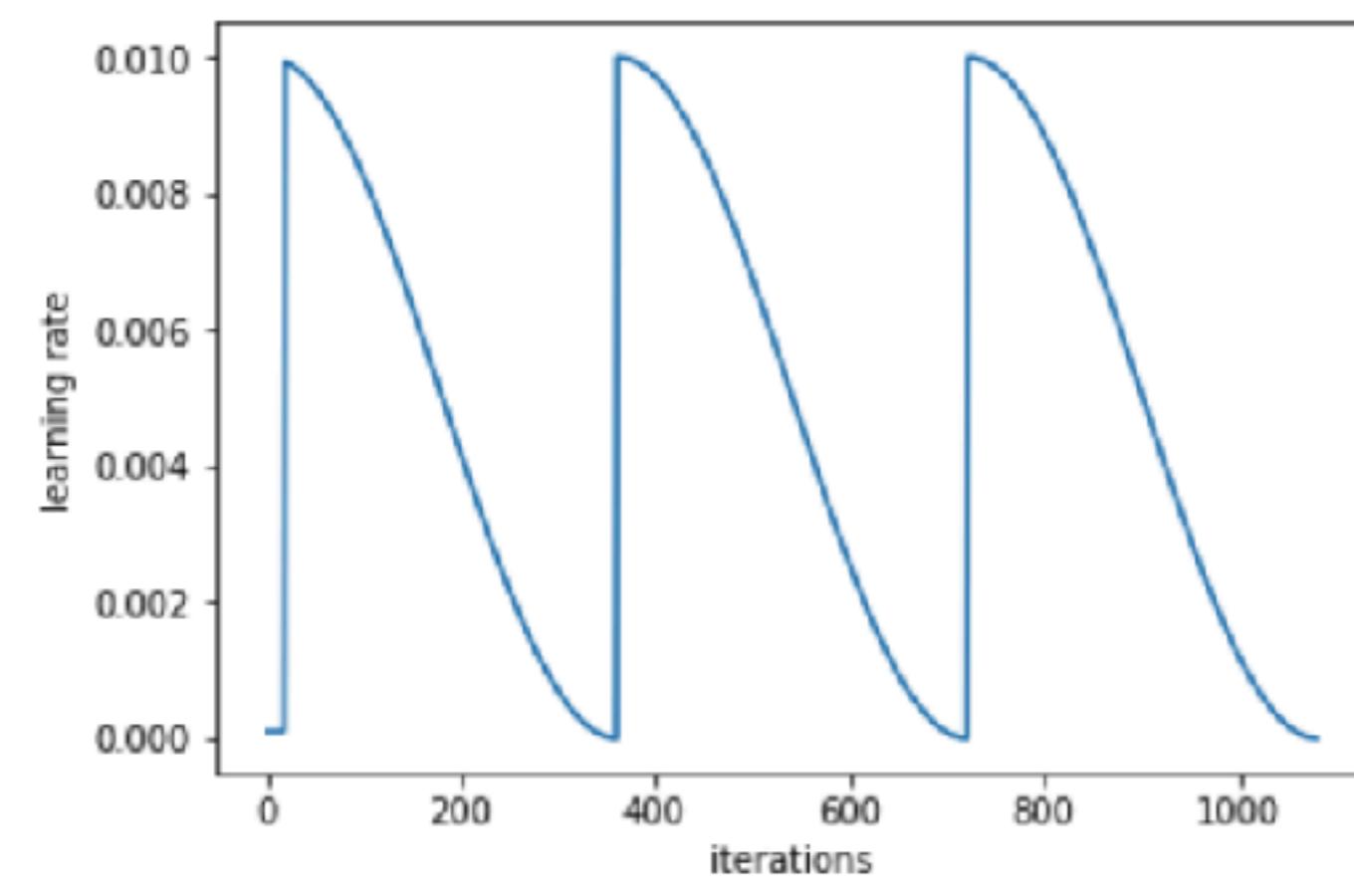
Underfitting



Learning Rate

Simple idea:
We are passing batches
through the network and
updating weights
starting with very low
learning rates up to very
high ones.
Observing this curve can
give us hints on what the
optimal LR might be.





Batch

Mini-Batch Size

Epoch



An Epoch represents one iteration over the entire dataset.

Batch



We cannot pass the entire dataset into the neural network at once. So, we divide the dataset into number of batches.

Iteration



If we have 10,000 images as data and a batch size of 200, then an epoch should contain $10,000/200 = 50$ iterations.

Mini-Batch size: Number of training instances the network evaluates per weight update step.

- Larger batch size = more computational speed
- Smaller batch size = (empirically) better generalization

Iteration

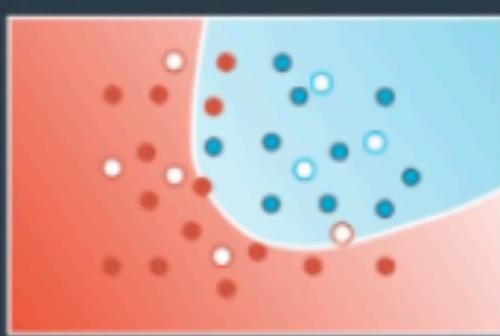
Epoch

Early Stopping

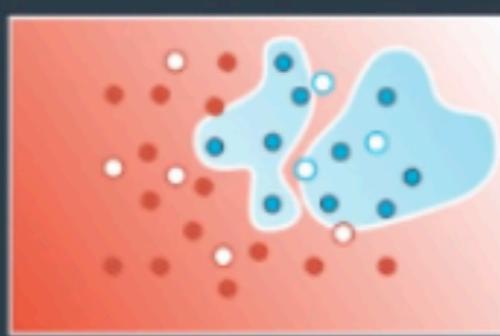
UNDERFITTING



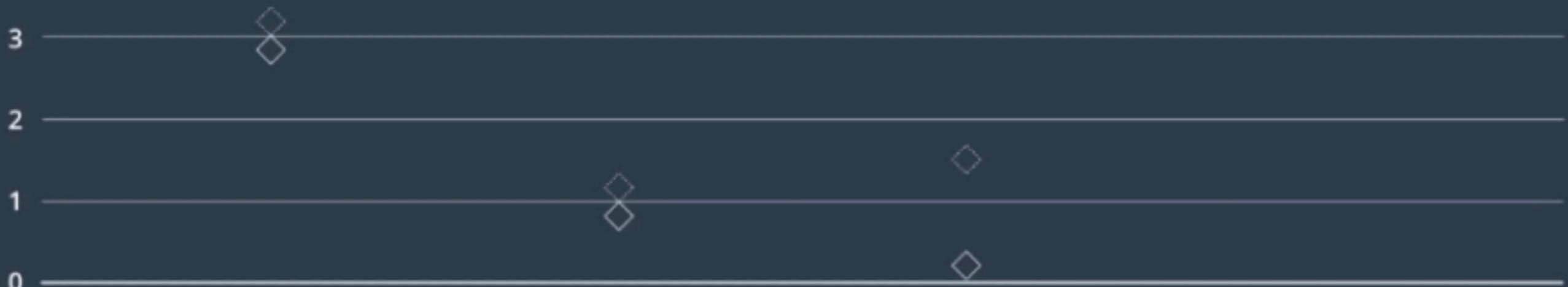
JUST RIGHT



OVERFITTING



OVERFITTING



EPOCH 1

Training Error: BIG
Testing Error: BIG

EPOCH 20

Training Error: SMALL
Testing Error: SMALL

EPOCH 100

Training Error: TINY
Testing Error: MEDIUM

EPOCH 600

Training Error: TINY
Testing Error: LARGE

Dropout

Augmentation

Vanishing Gradients

Exploding Gradients

Deep Learning Applications

Lip Syncing

Lip Syncing

A group from the University of Washington created a system that uses audio and synthesizes it with lip motion of a face in a video.



Synthesizing Obama: Learning Lip Sync from Audio

Supasorn Suwajanakorn

Steven M. Seitz

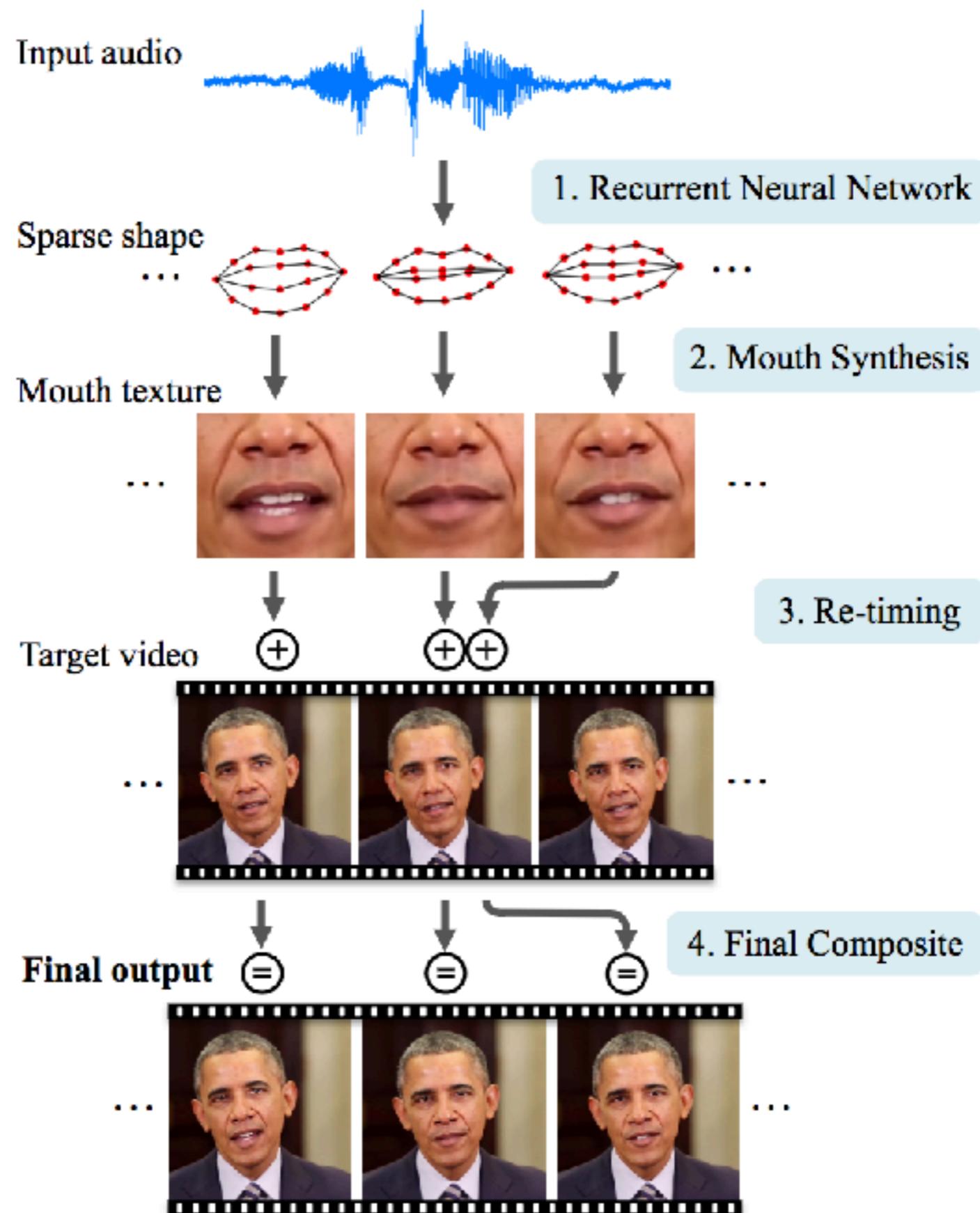
Ira Kemelmacher-Shlizerman

University of Washington

SIGGRAPH 2017

<http://grail.cs.washington.edu/projects/AudioToObama/>

Lip Sync Algorithm





Original Video for Input Audio



a) Our result



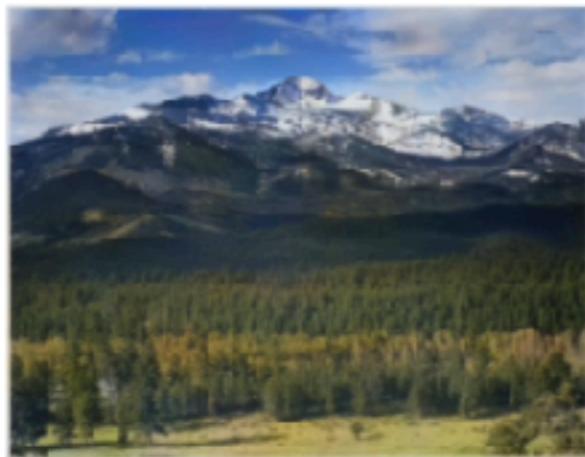
Original Video for Input Audio



b) Our result

Image Colorization

Image Colorization



Colorado National Park, 1941

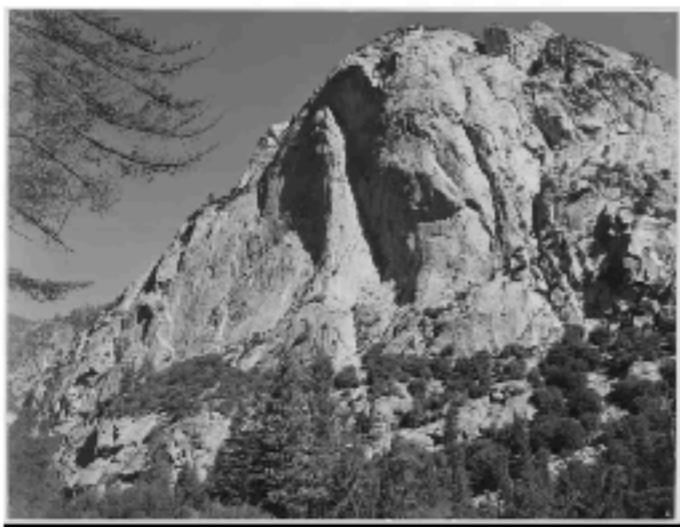
Textile Mill, June 1937

Berry Field, June 1909

Hamilton, 1936

Deep Learning network restores the color of old black & white photos.

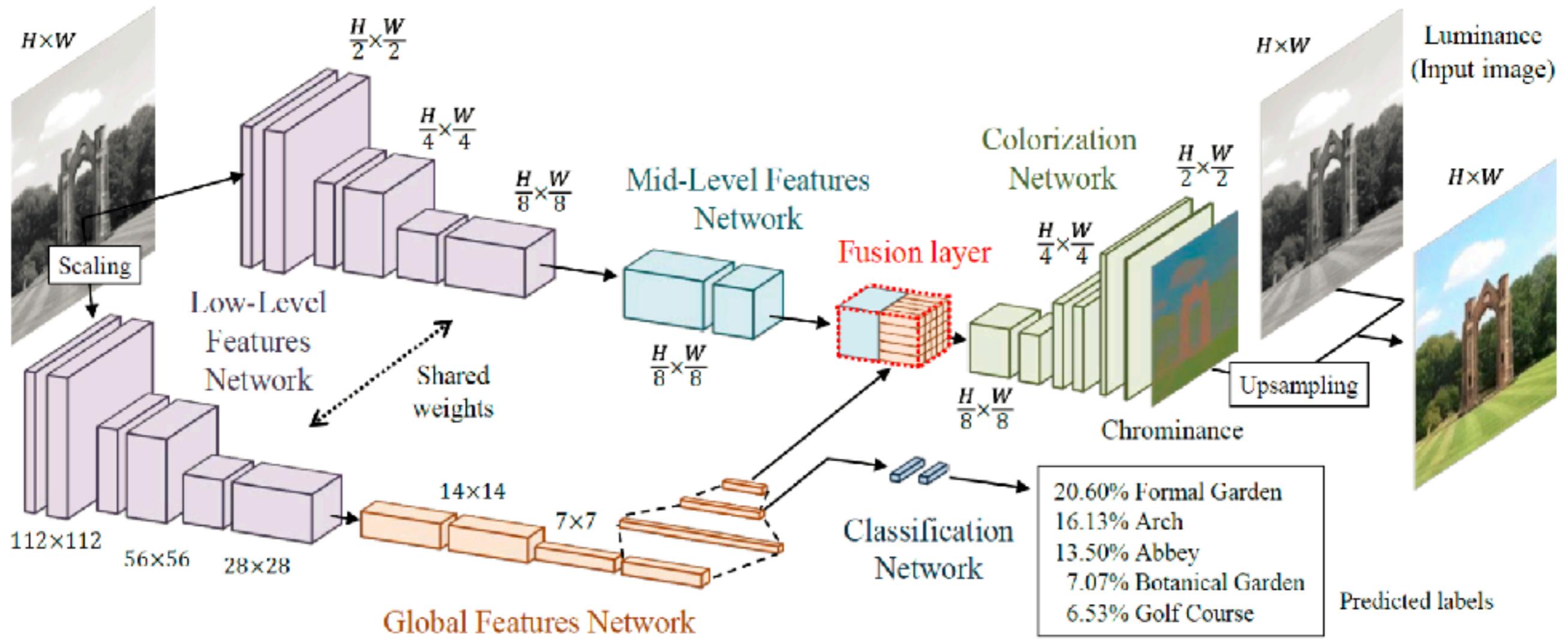
100 year old pictures...

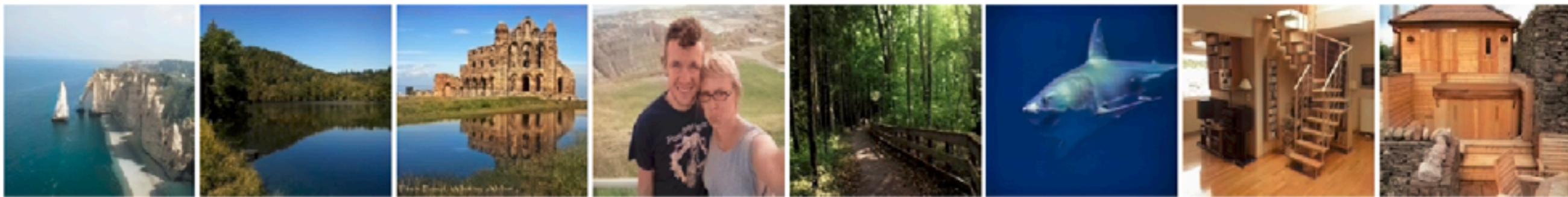


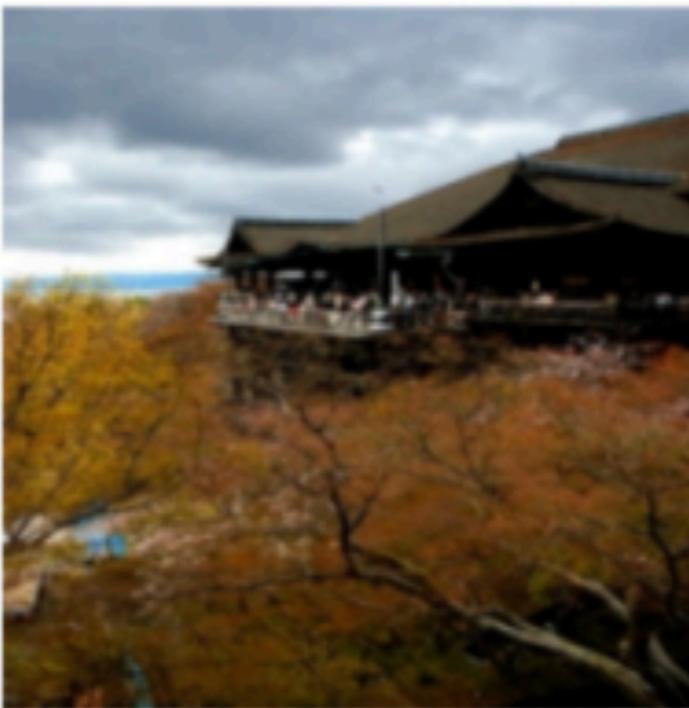
Code: https://github.com/satoshiiiizuka/siggraph2016_colorization

Paper: http://iizuka.cs.tsukuba.ac.jp/projects/colorization/data/colorization_sig2016.pdf

Colorization Algorithm



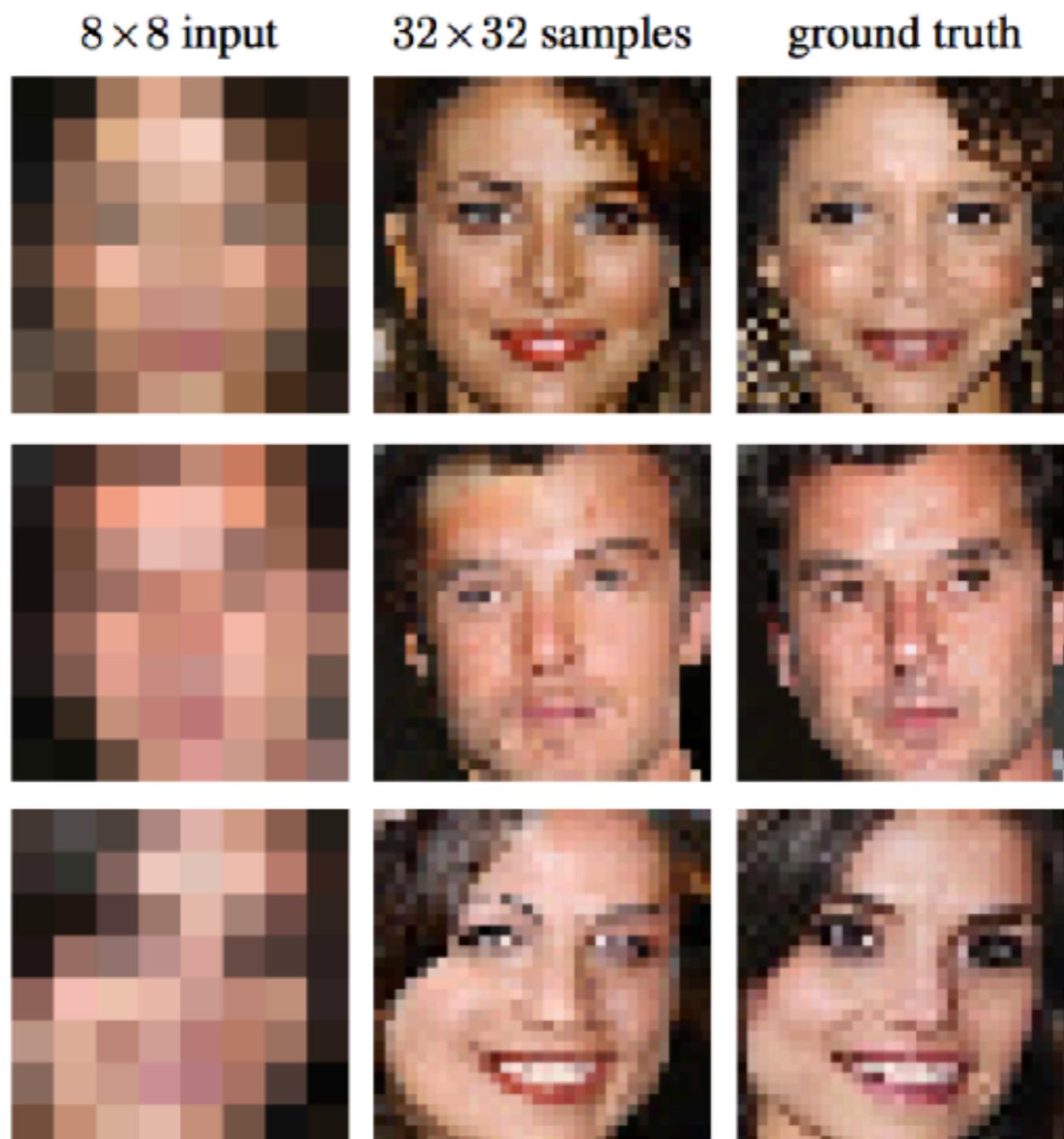




The computer restored the colors of these two B&W photos but made mistakes. Can you guess which photos are the real ones?

Super Resolution

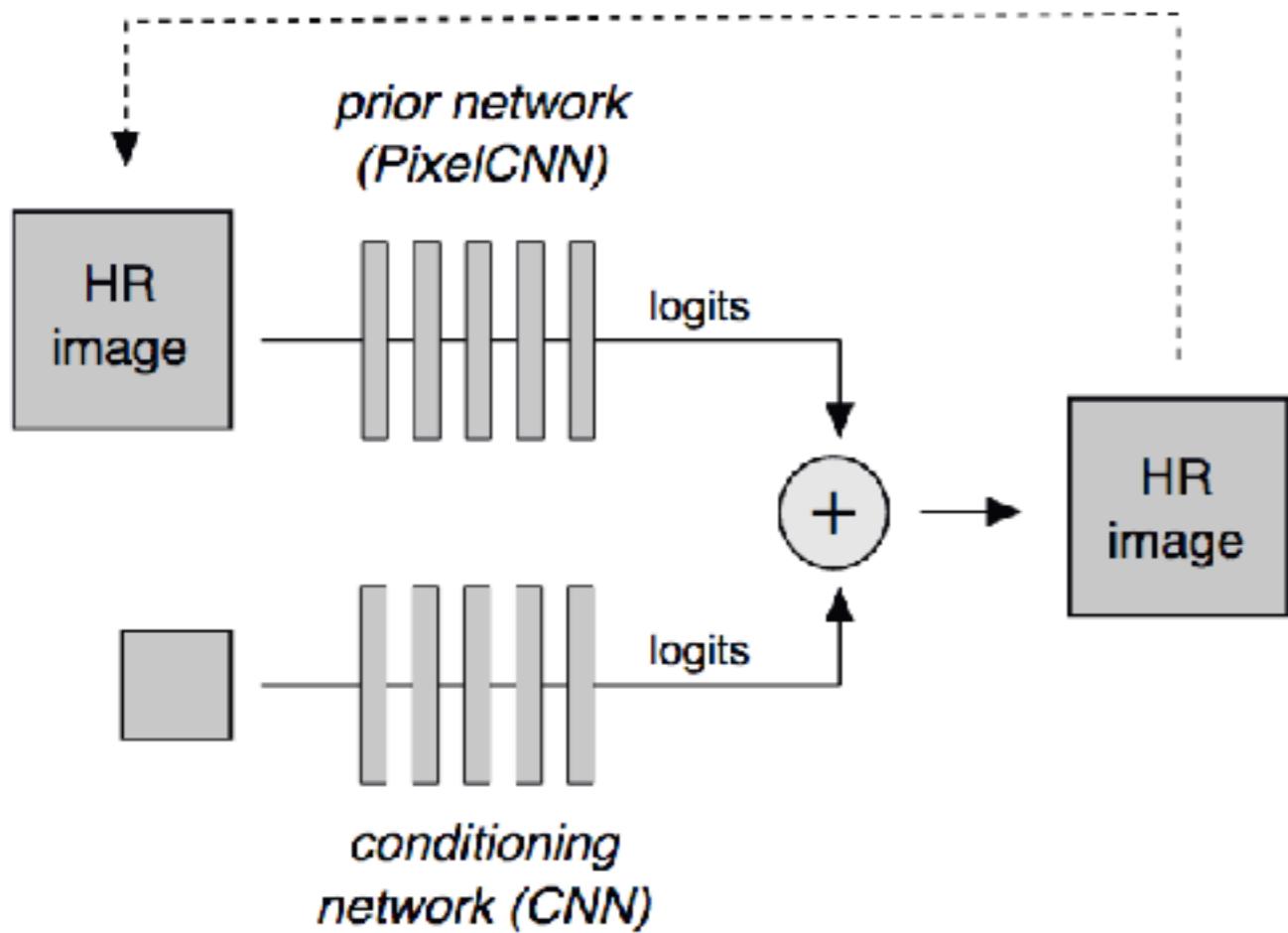
Super Resolution



8x8 pixel photos were inputted into a Deep Learning network which tried to guess what the original face looked like. As you can see it was fairly close (the correct answer is under "ground truth").



Algorithm



Pose Estimator

Pose Estimator



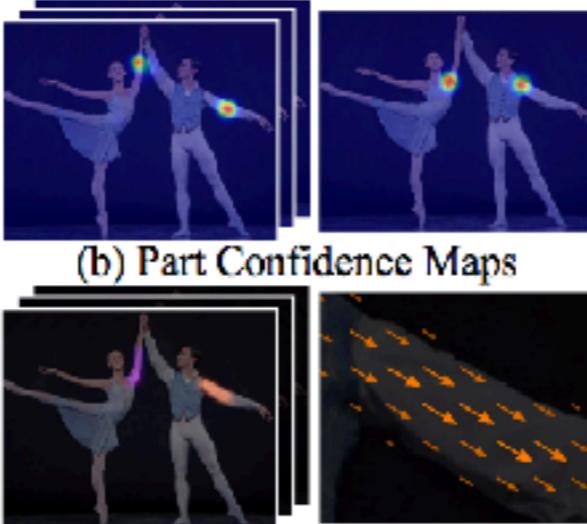
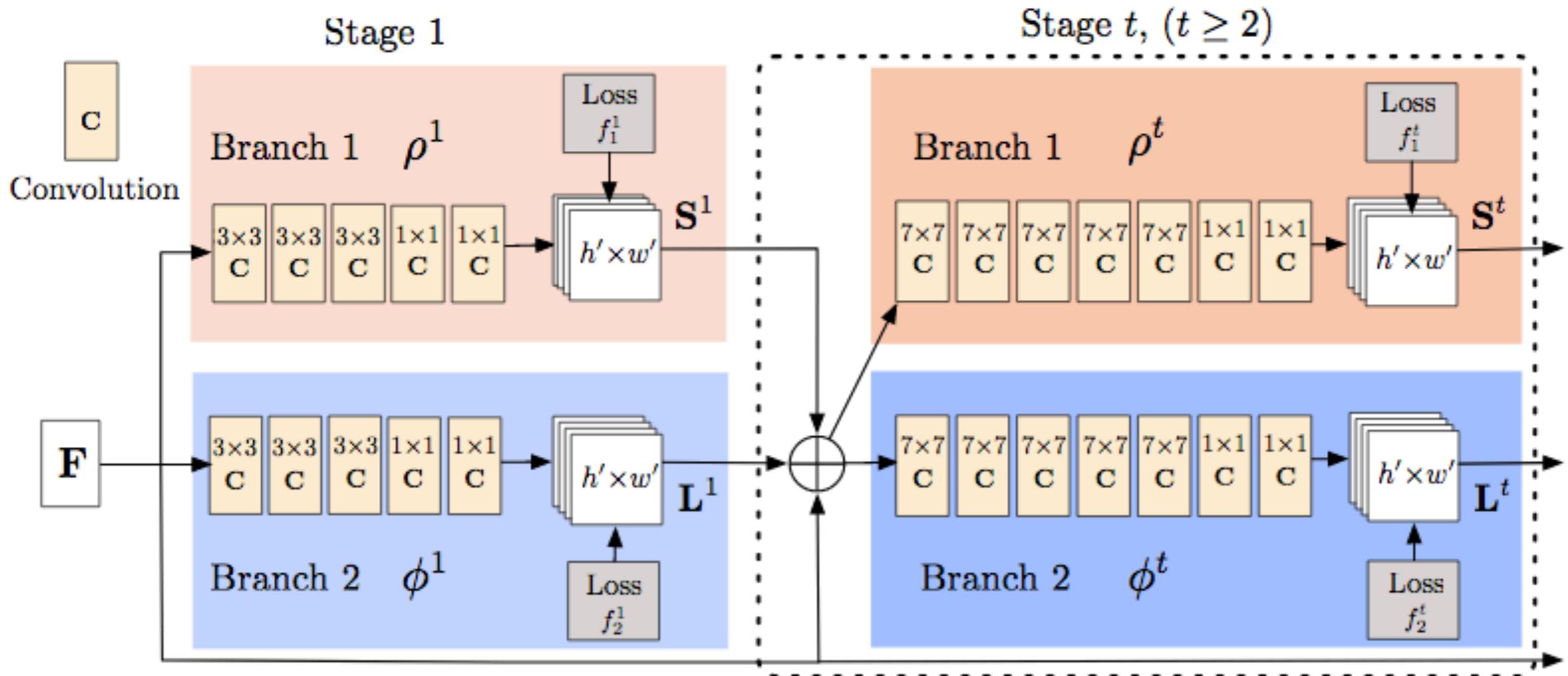
Figure 1. **Top:** Multi-person pose estimation. Body parts belonging to the same person are linked. **Bottom left:** Part Affinity Fields (PAFs) corresponding to the limb connecting right elbow and right wrist. The color encodes orientation. **Bottom right:** A zoomed in view of the predicted PAFs. At each pixel in the field, a 2D vector encodes the position and orientation of the limbs.

Real-time Multi-Person 2D Pose Estimation Using Part Affinity Fields

Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh

Carnegie Mellon University

Algorithm



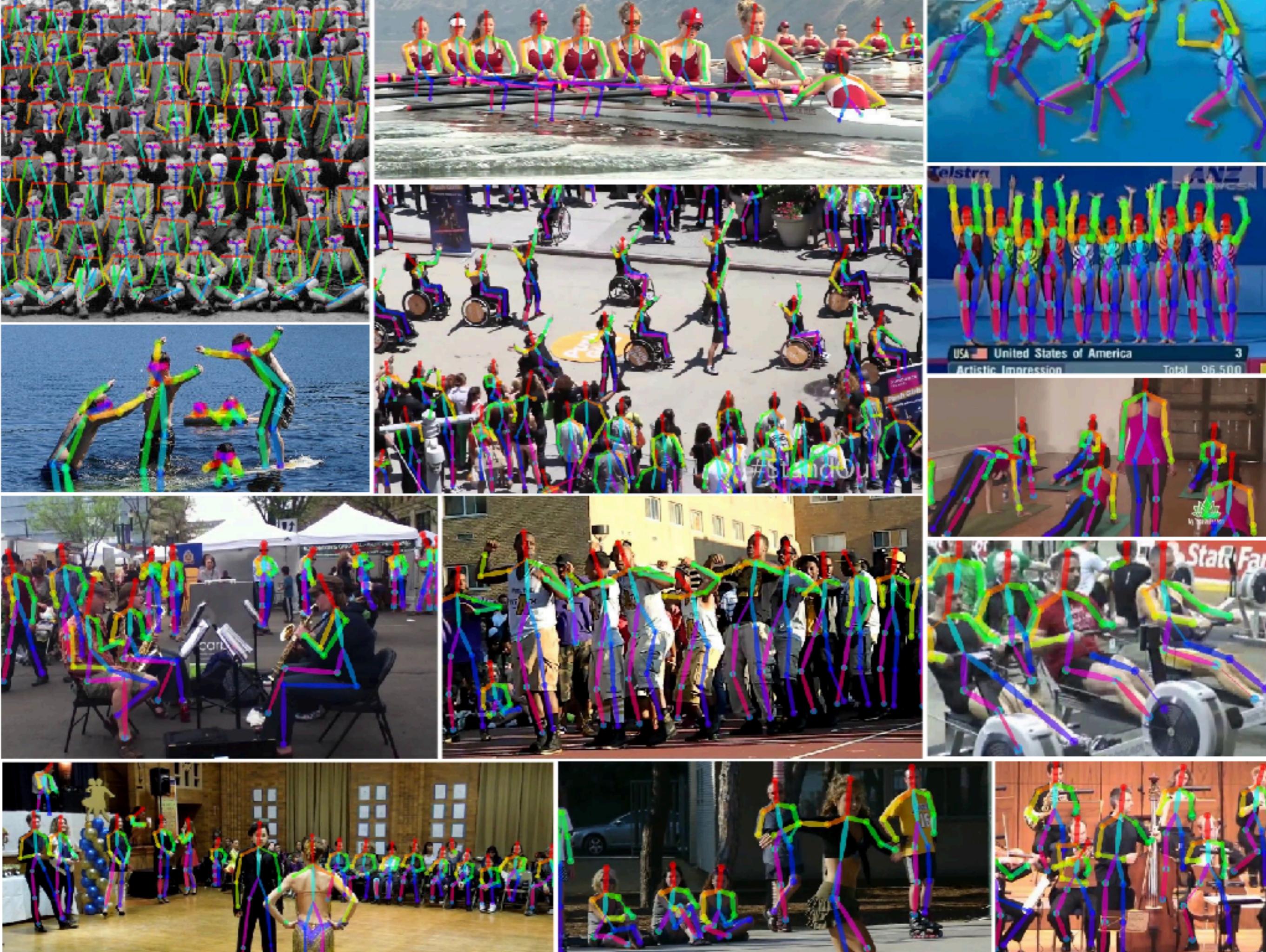
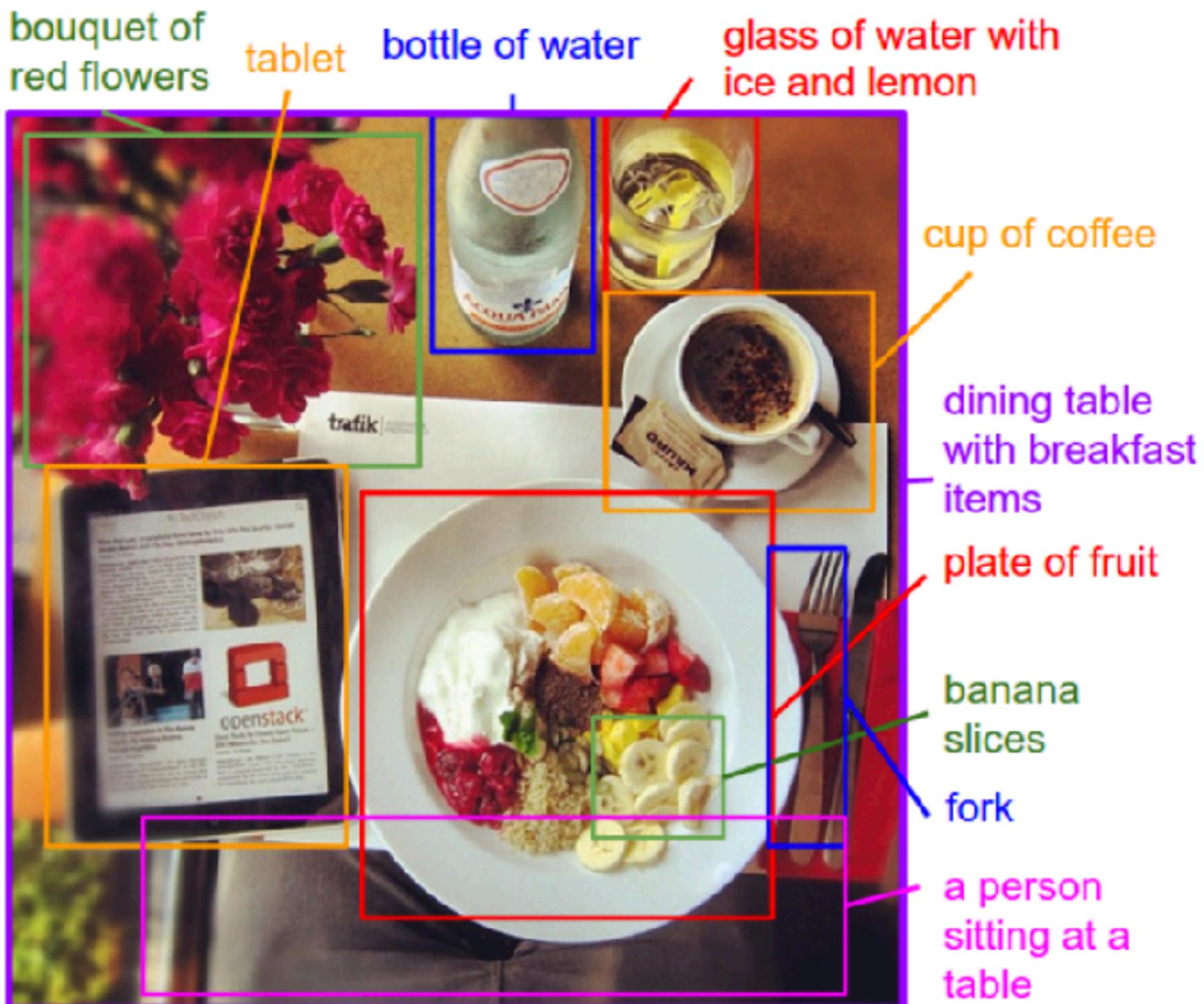


Image Descriptor

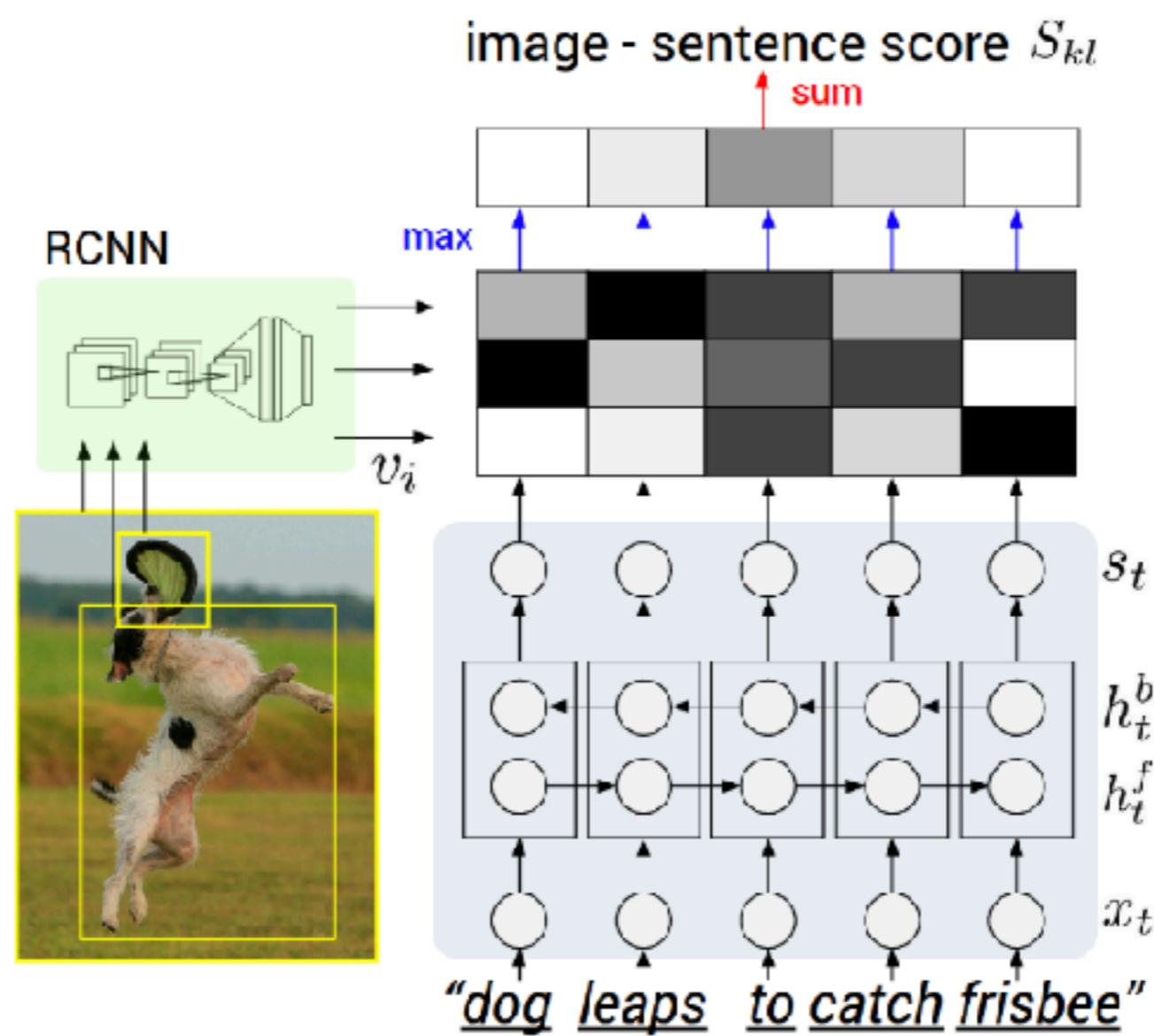
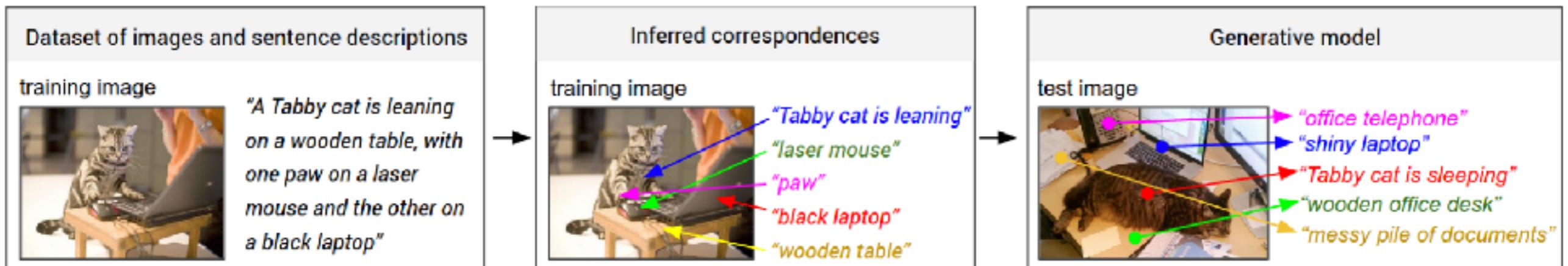
Image Descriptor



A Deep Learning network automatically segments an image and writes a short sentence describing each segment with a proper English grammar.



Algorithm



Results



man in black shirt is playing guitar.

construction worker in orange safety vest is working on road.

two young girls are playing with lego toy.

boy is doing backflip on wakeboard.

Demo

<https://cs.stanford.edu/people/karpathy/deepimagesent/generationdemo/>

Deep Warp

Deep Warp

- Maximum image dimension: 400px. Will be resized if exceeds the limit.
- Minimum image dimensions (after resize): 200x25px.

News:

- 24.02.17: Check out an article (in Russian) about the technology! It comes with a [Telegram bot](#).
- 11.03.17: [New!](#) An article (in English) on [the Verge](#).
- 11.03.17: [New!](#) The online demo should now be working on [iOS](#) and in the [Safari](#) browser.

Choose file

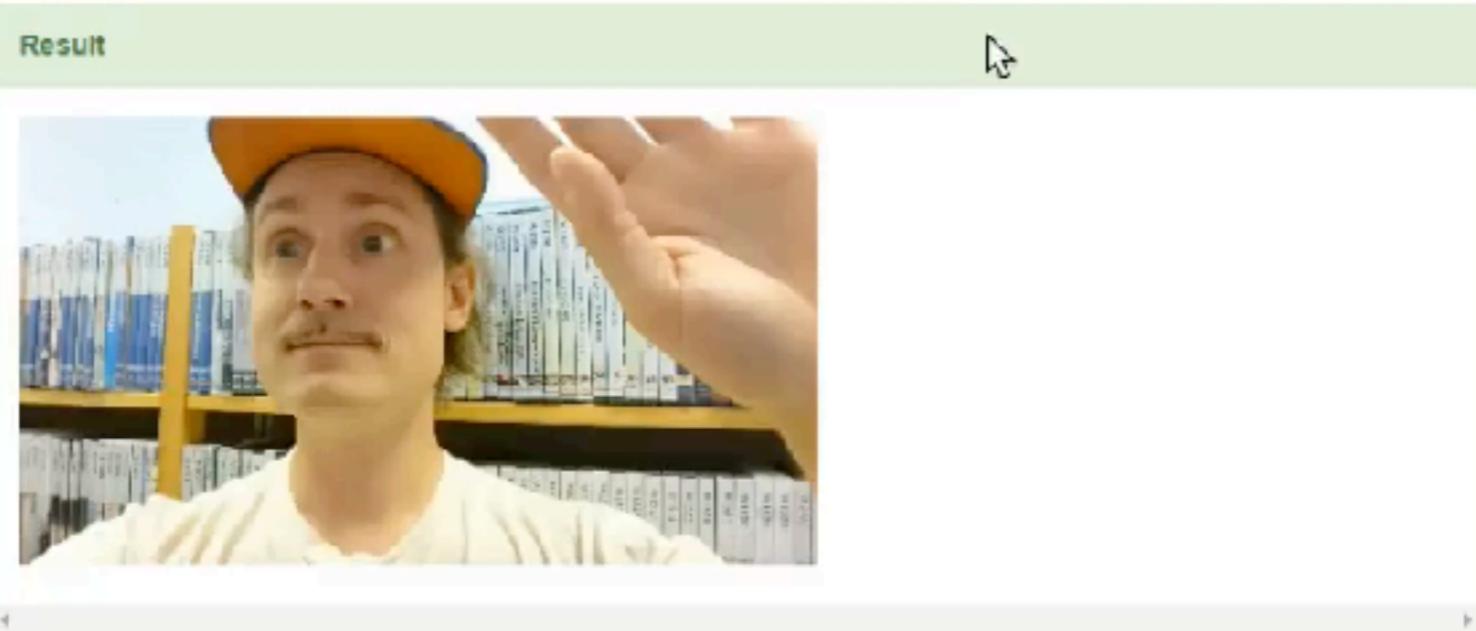
DSC_0932.JPG

Let's roll!

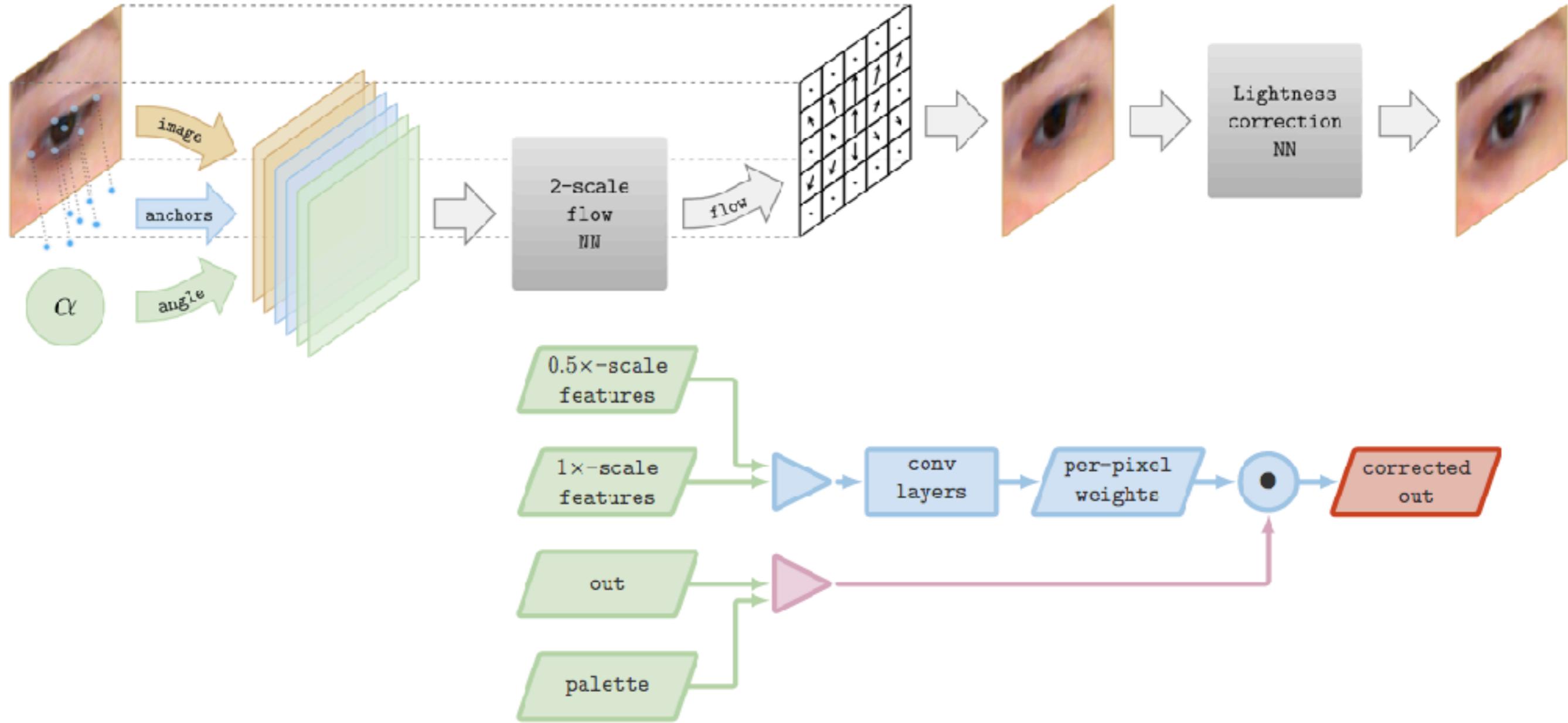
Let's scroll!

Let's cross!

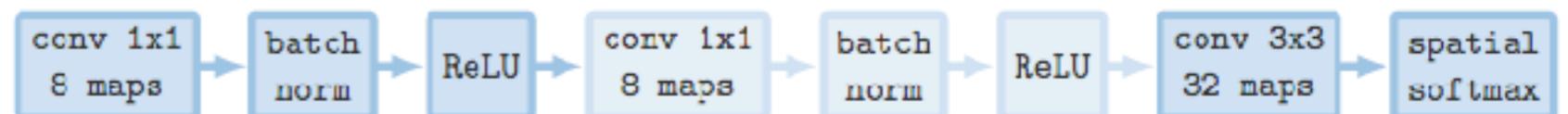
Let's shift!



Algorithm



(a) Architecture.



(b) Convolutional layers.

Demo

<http://sites.skoltech.ru/compvision/projects/deepwarp/>

Behaviour Analysis

Deep Glint

DeepGlint Haomu Behavior Analysis System

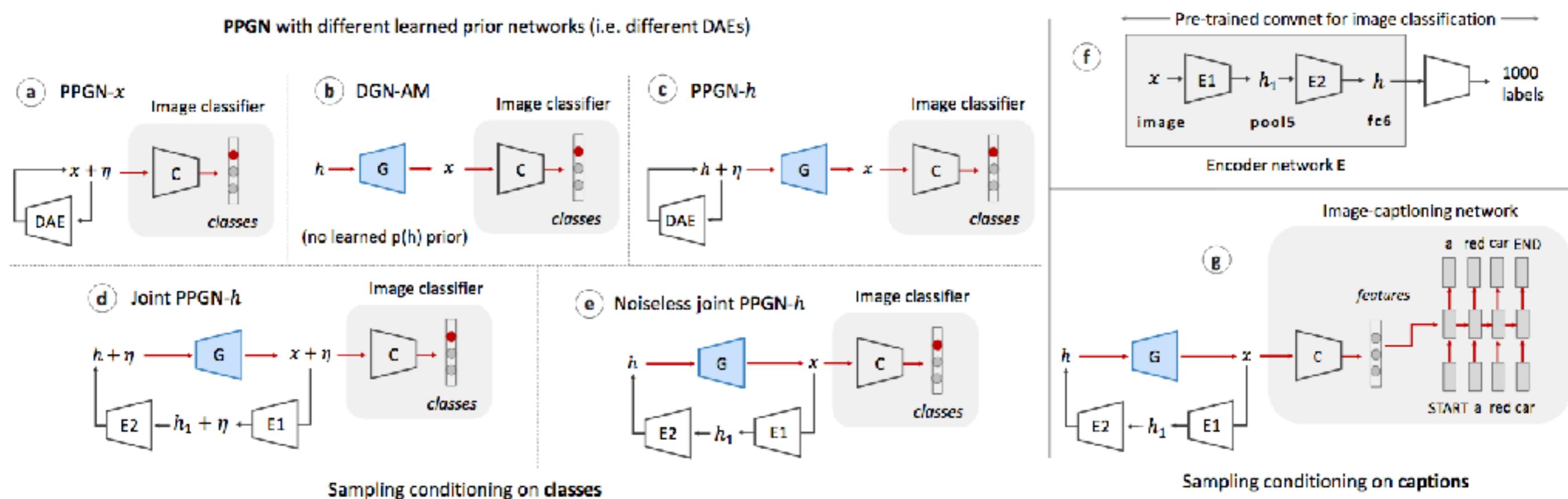
DEEPCGLINT
格 灵 深 瞳

Photo Creator



volcano

Algorithm



Results

(a) Real: top 9



(b) DGN-AM [37]



(c) Real: random 9



(d) PPGN (this)



cardoon



cock

Results



broccoli

Paper

<https://arxiv.org/pdf/1612.00005.pdf>

Magazine

http://www.nature.com/polopoly_fs/1.21398!/menu/main/topColumns/topLeftColumn/pdf/542016a.pdf

Translation



MÖRK
Mörk → Dark



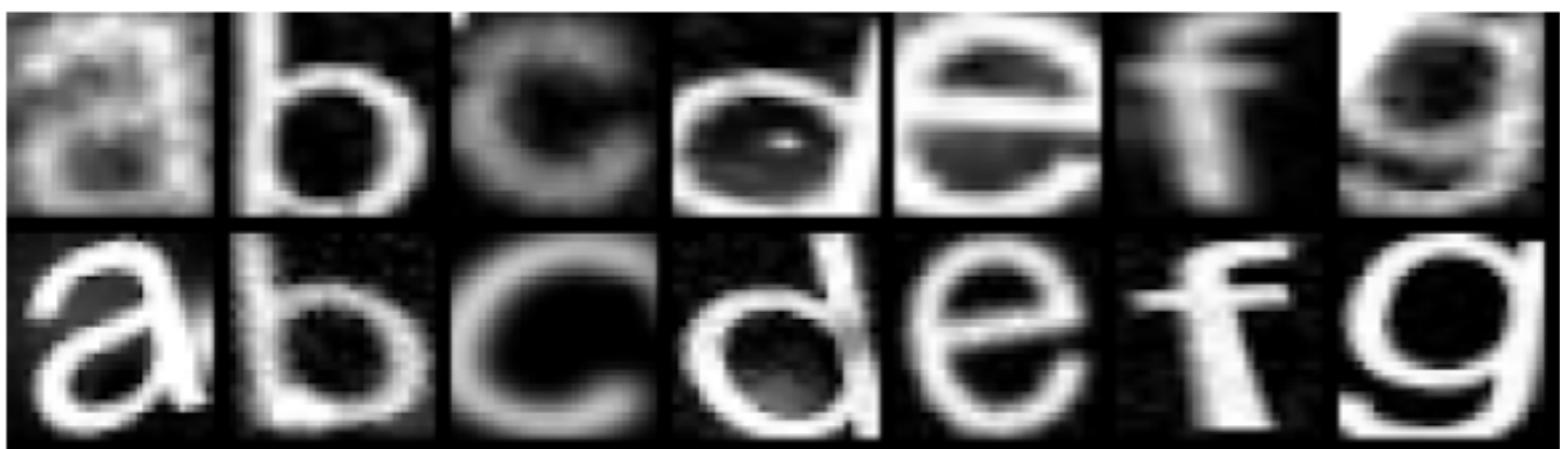
A photo of a packaged food taken by the phone, and Google Translate "reads" the text and replaces it with a text in English in real-time.

Algorithm

- Object Detection - Locate letters
- Object Classification - Identify letters (OCR)
- Translation Dictionary LookUp
- Imposing Translated Text on Actual Text

<https://ai.googleblog.com/2015/07/how-google-translate-squeezes-deep.html>

Training with Dirty Real-life letters



Saving Whales

Strata+ Hadoop

WORLD

#StrataHadoop
strataconf.com



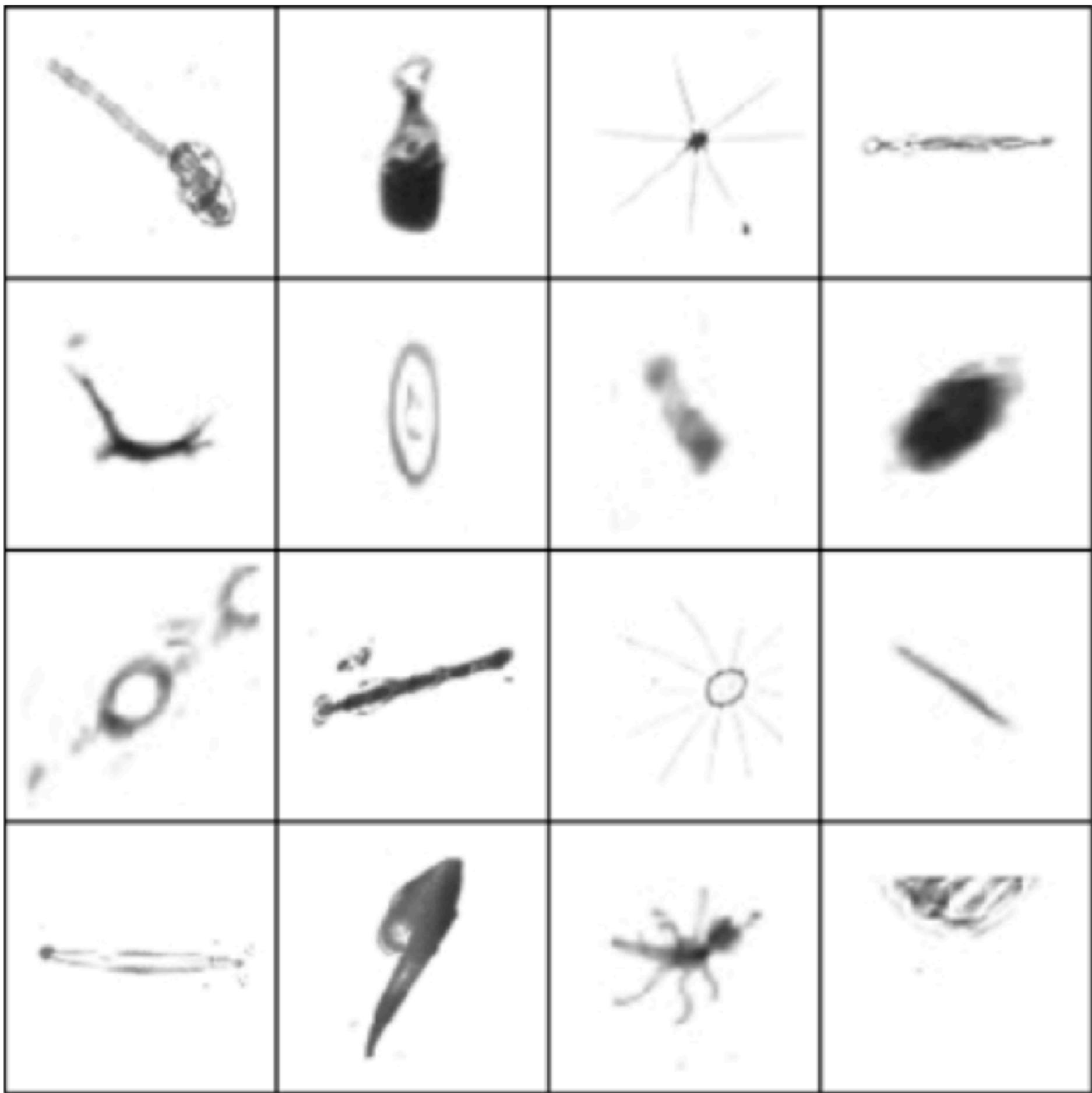
Saving whales with deep learning

The Leadenhall Building
122 Leadenhall Street, London, EC3V

Piotr Niedźwiedź
deepsense.io

Sponsored by *deepsense.io*

Plankton Classification



<https://benanne.github.io/2015/03/17/plankton.html>

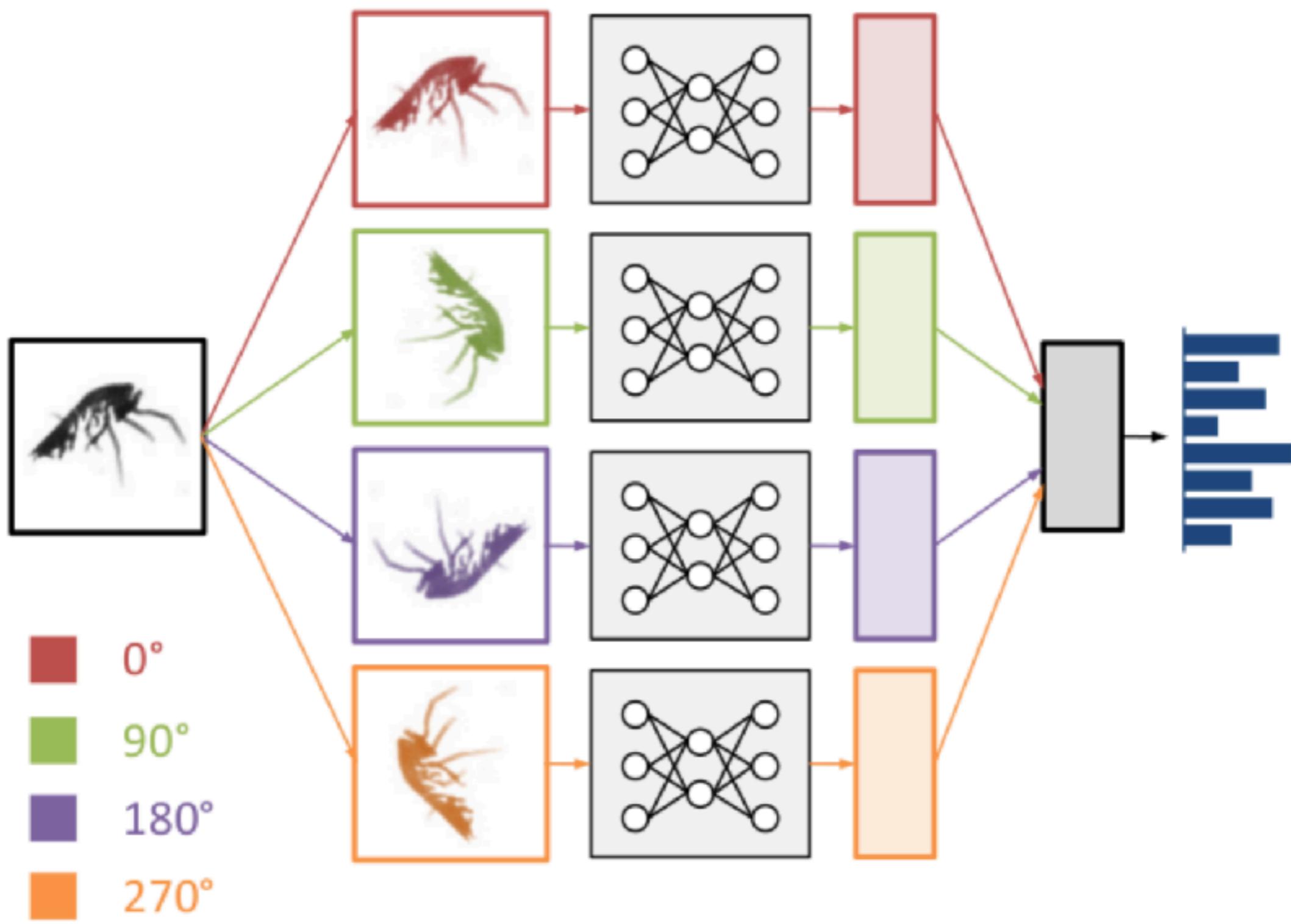
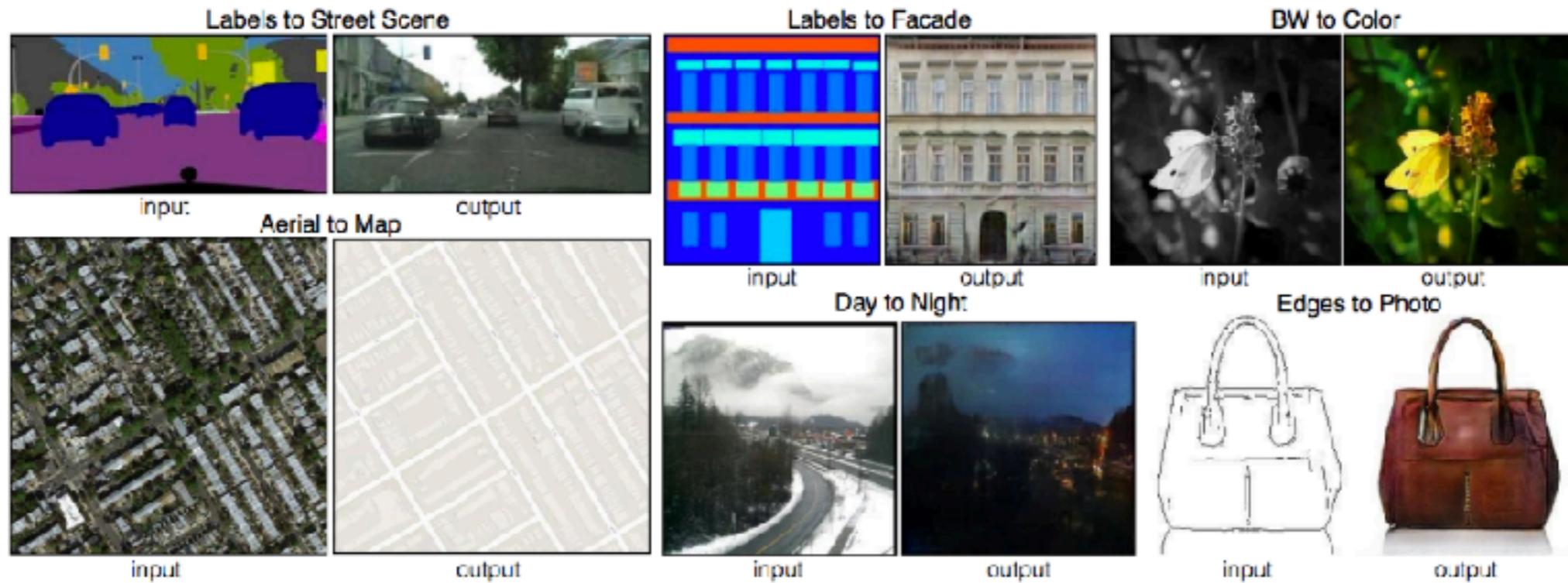


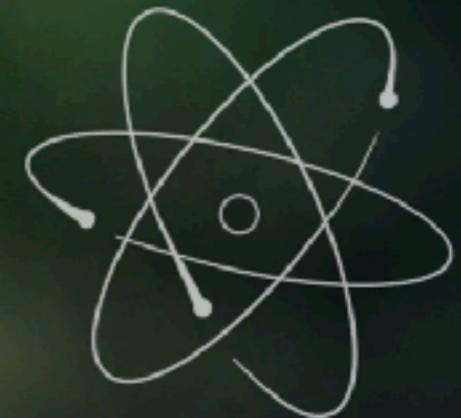
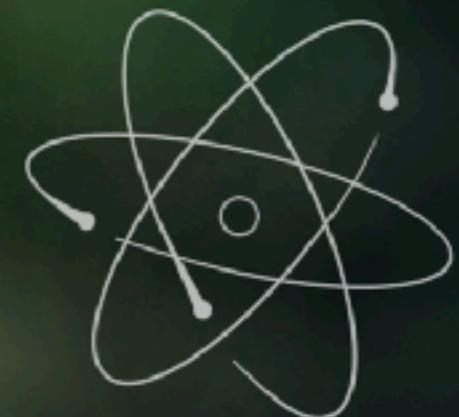
Image Creator



A deep learning network generates new photos. For example, photos taken at daytime were converted to photos that seem to be taken at night.



Draw the outline of a bag or a shoe and the Deep Learning network will color it for you. The results are pretty creative.



TWO MINUTE PAPERS

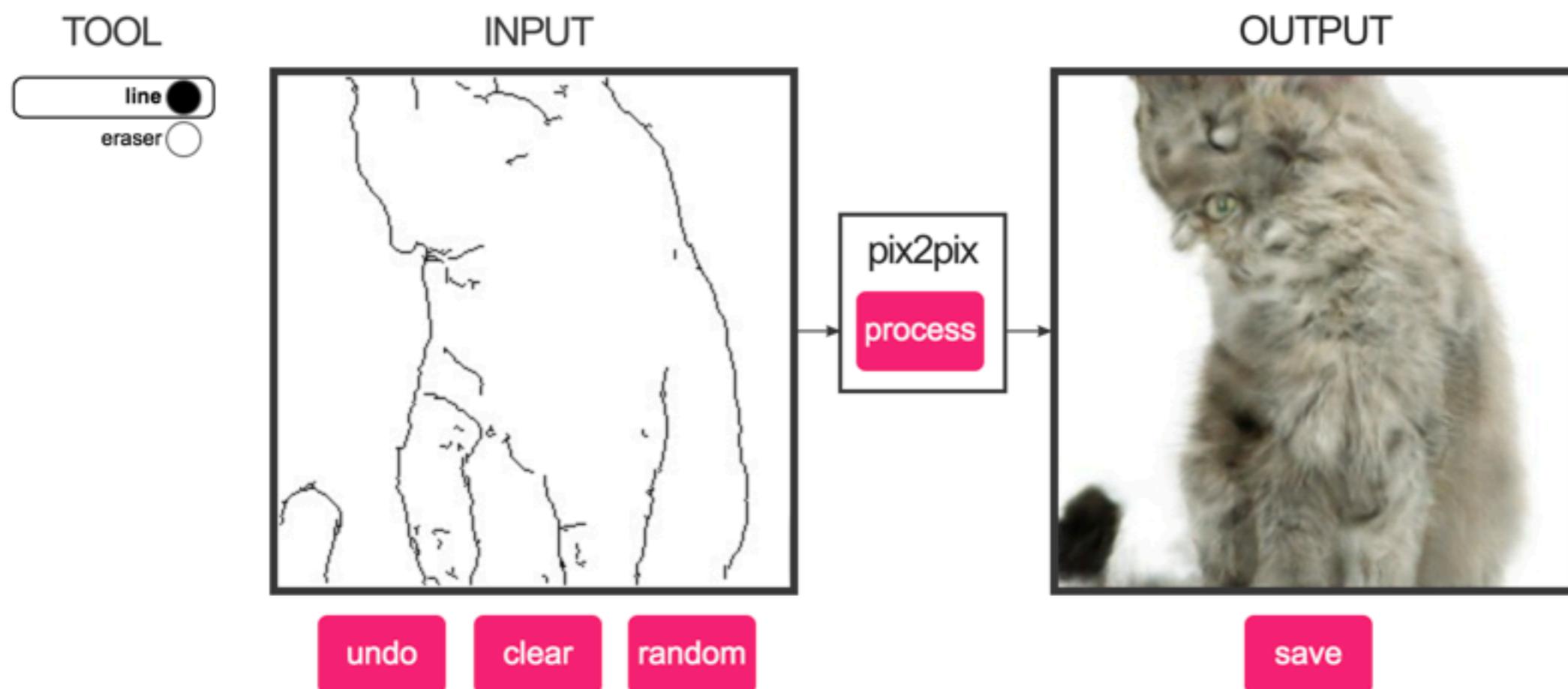
WITH KÁROLY ZSOLNAI-FEHÉR (KZF)

AI MAKES STUNNING PHOTOS FROM YOUR DRAWINGS

Disclaimer: I was not part of this research project, I am merely providing commentary on this work.

Demo

edges2cats



<https://affinelayer.com/pixsrv/>

Text Identifier

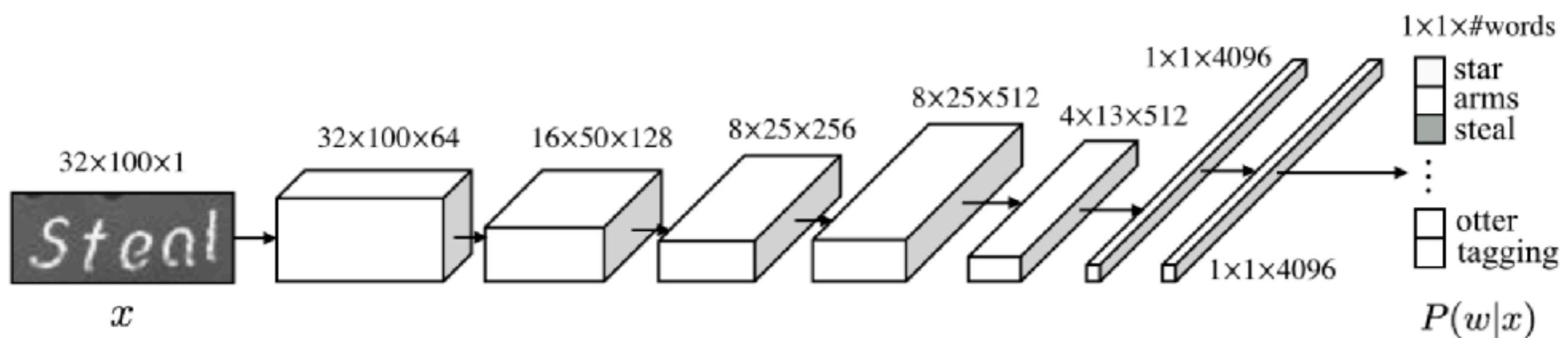
Search for something

Text to search...

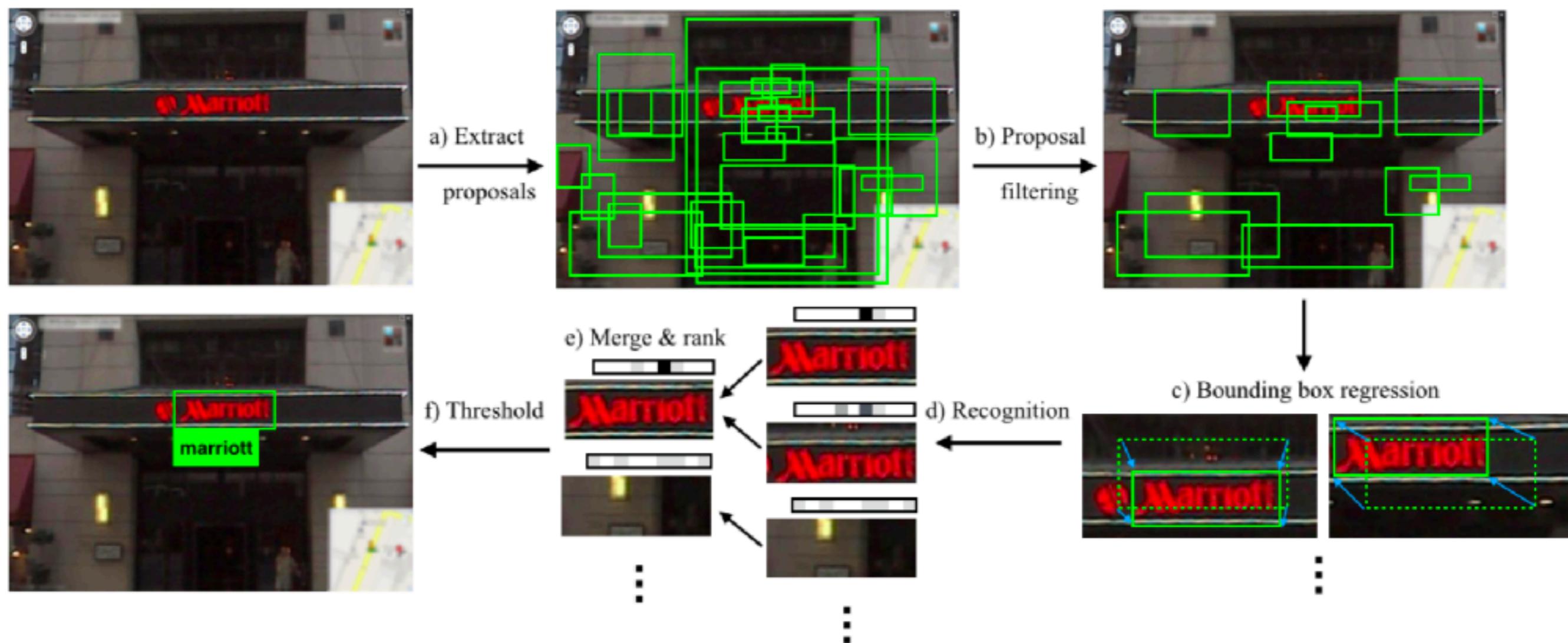
HEALTH LING GROUP WHICH MANCHESTER PLAN
ATTACK MURDER THURSDAY THAN REPORT
WORK NIGHT BRIEF SPORTS ABOUT SPORTSMAN
WEDNESDAY HOSPITAL BUCHENWALD SCOTLAND
CENTRAL SUNDAY ENGLAND OTHER SOUTH GOVERNMENT
ROYAL MINISTER PLYMOUTH
SAID YORK BIRMINGHAM OBAMA SAYS TIME BEIJING DEATHS
IRELAND HOUR WESTMINSTER STATE HIGH OFFICE FIRST SEND AGAINST
RELAYS SPORT BEEN FROM LASS TERMS
INTERACTIVE BRITISH BANK LIVE HAVE PEOPLE SWINE LABOR
ANDY NEXT EXPENSES WORLD POLICE PUBLIC LAND
COMING FRIDAY OVERAFTER NEWSSES FINAL MOUTH
PICTURE TEXT FOUND

<http://zeus.robots.ox.ac.uk/textsearch/#/search/>

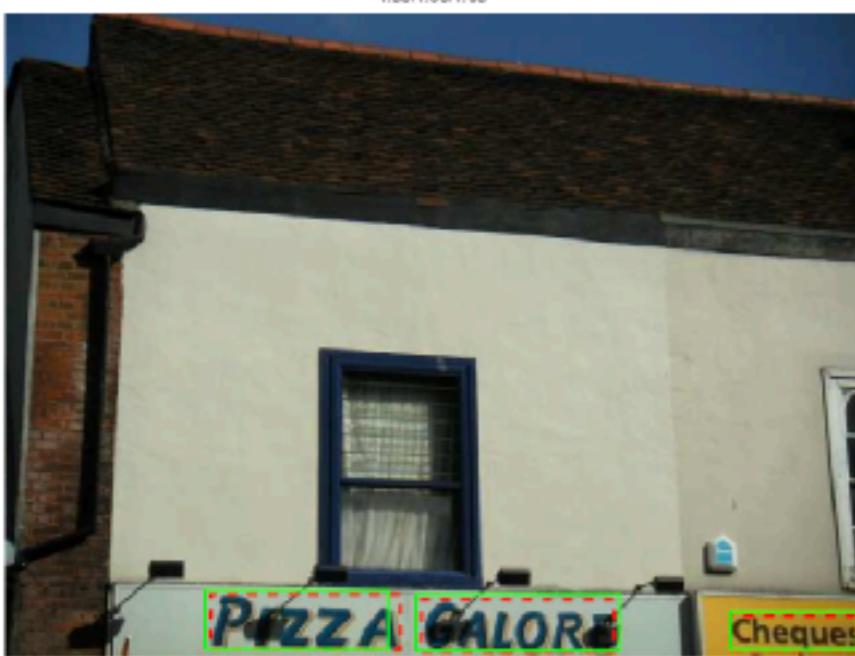
Algorithm



Algorithm



Results



SunRoof

✓ Analysis complete. Your roof has:

 **1,531 hours of usable sunlight per year**
Based on day-to-day analysis of weather patterns

 **758 sq feet available for solar panels**
Based on 3D modeling of your roof and nearby trees

\$5,000 savings
Estimated net savings for your roof with a 20-year lease

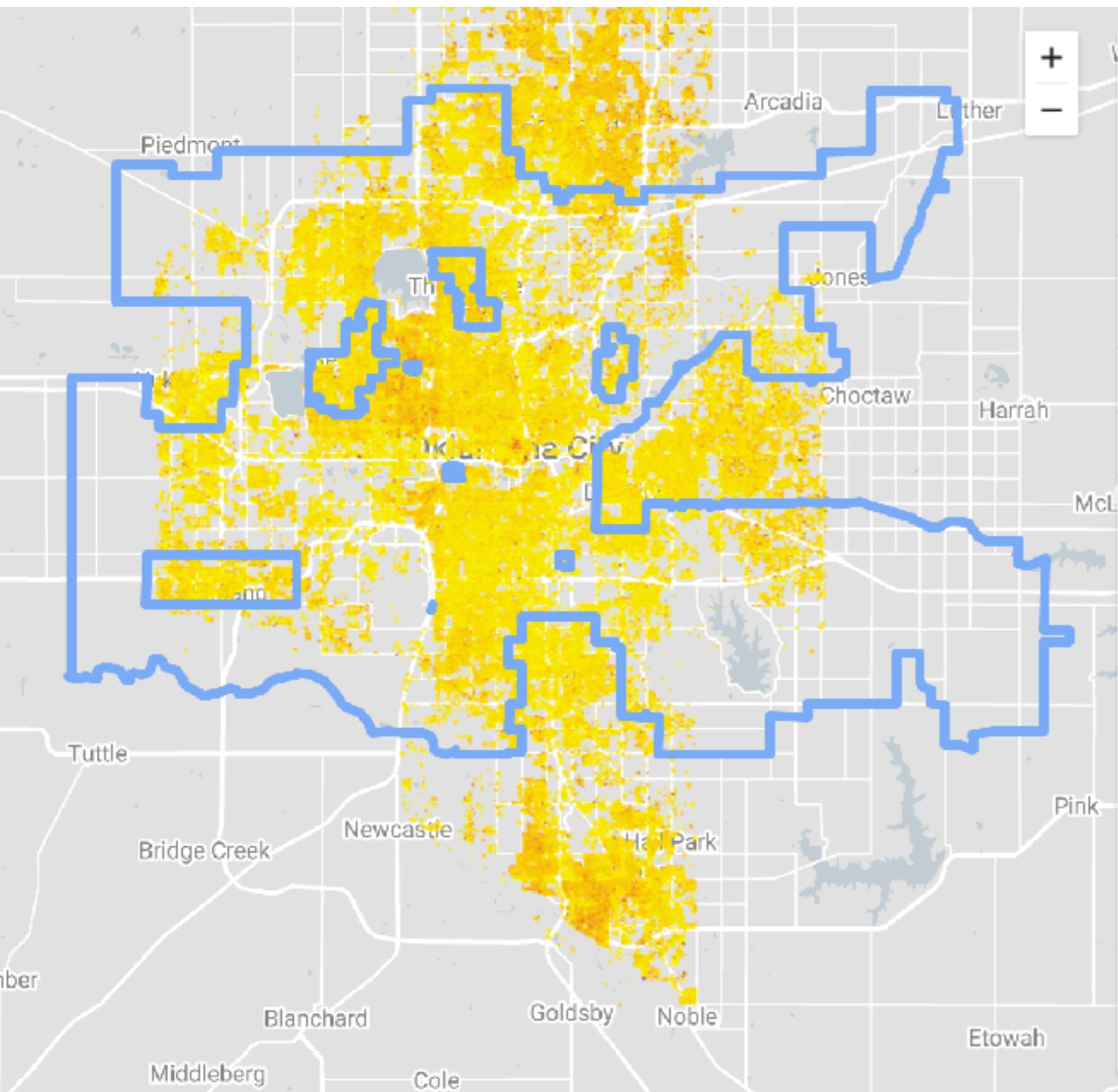
[FINE-TUNE ESTIMATE](#) [SEE SOLAR PROVIDERS](#)

Wrong roof? Drag the marker to the right one.



Google Sunroof predicts how much money you can save if you install solar panels. It uses Deep Learning to analyze the Google Earth aerial images.

<https://www.google.com/get/sunroof#/p=0>



Estimated rooftop solar potential of Oklahoma City, OK

Last updated: 11/2018

Sunlight on rooftops

Shady Sunny

Existing solar arrays

Median household income

Buildings

96%
solar-viable

274
existing solar installations

Based on 90% data coverage over buildings in this geographic area. All estimates are based on buildings viable for solar panels. Included panels receive at least 75% of the maximum annual sun in the county. For Oklahoma City, the average value of the threshold is 1,155 kWh/kW. Read about Project Sunroof's methodology for defining solar viability below.

[READ METHODOLOGY](#)

Voice Generator



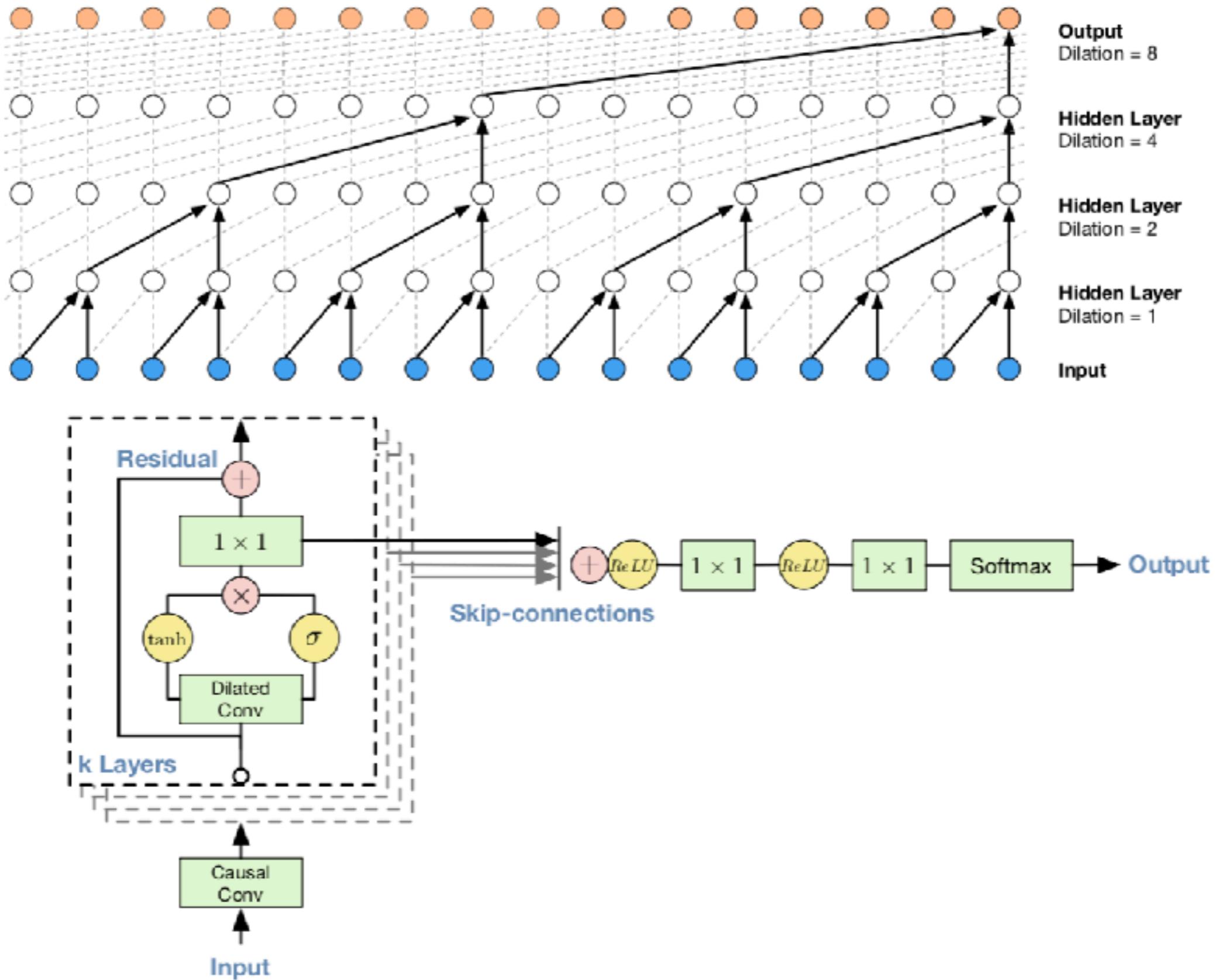
Live Demo



Baidu Research

Andrew Ng

Wavenet



Music Composer

tion

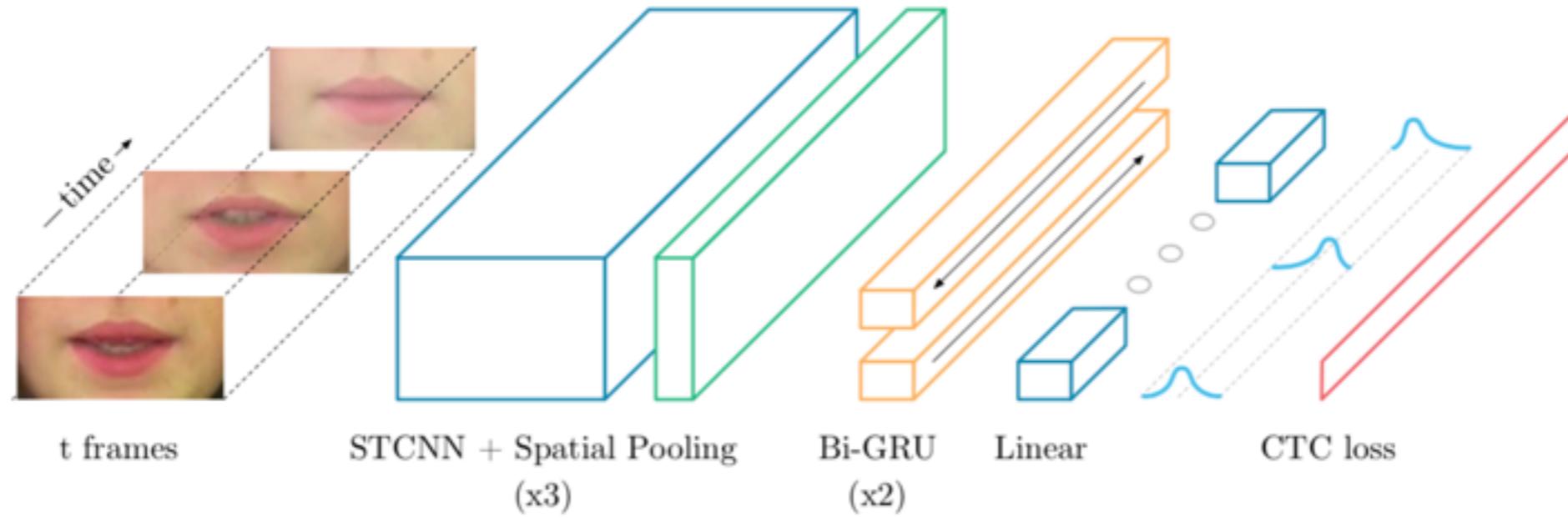
with

Visual Sound Creator

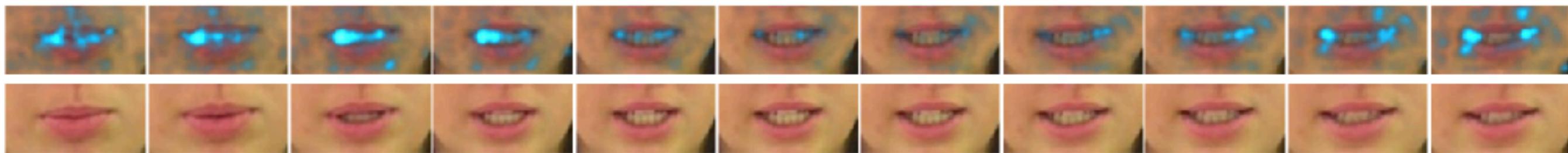
Lip Reading



Algorithm



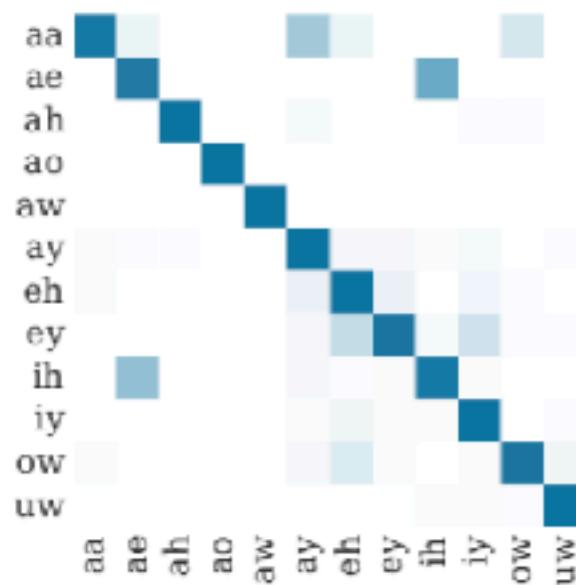
p l e e a a s s u u e



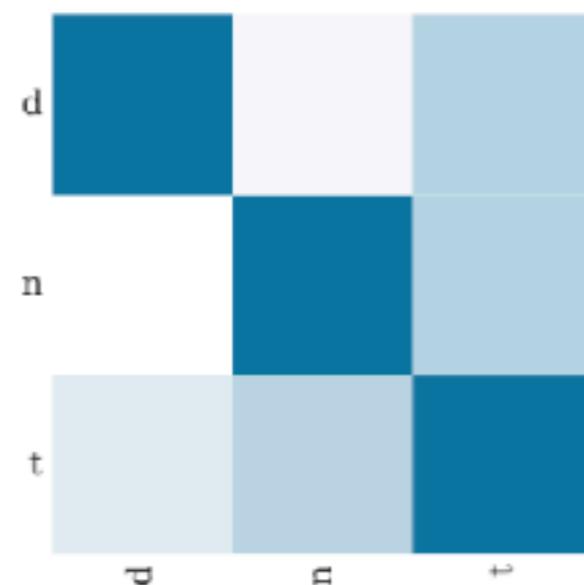
l u u u a a u u y



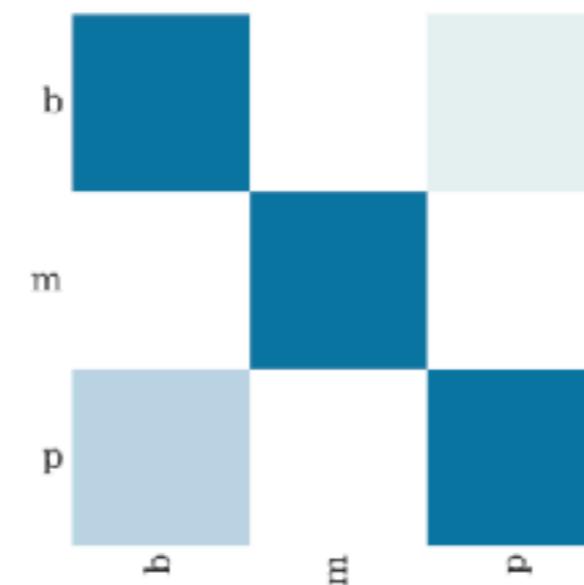
1 2 3 4 5 6 7 8 9 10 11



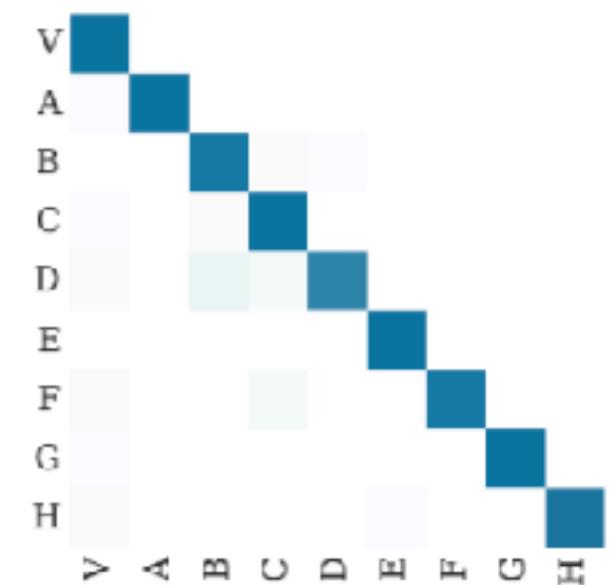
(a) Lip-rounding vowels



(b) Alveolar



(c) Bilabial



(d) Viseme Categories

Style Transfer

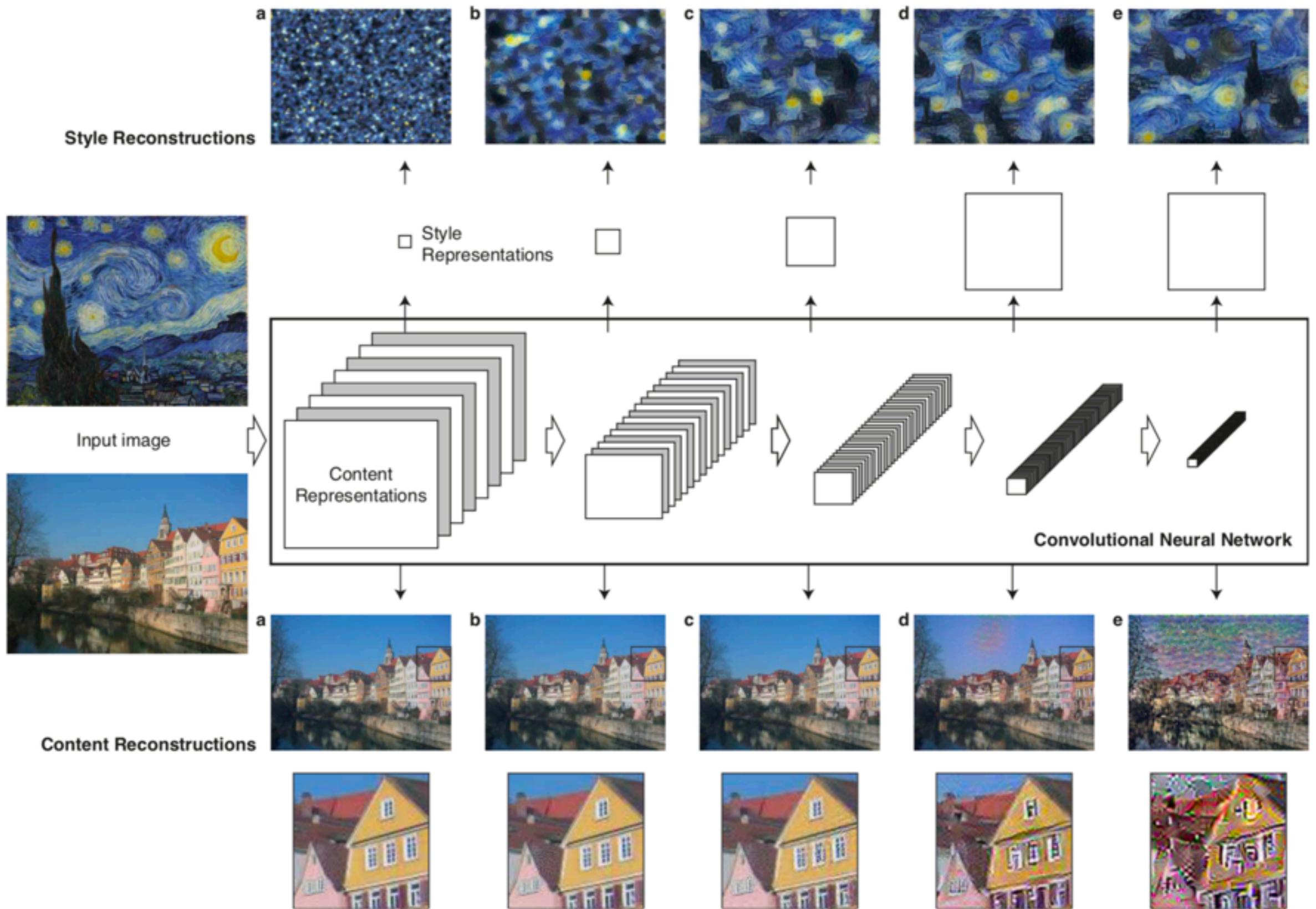


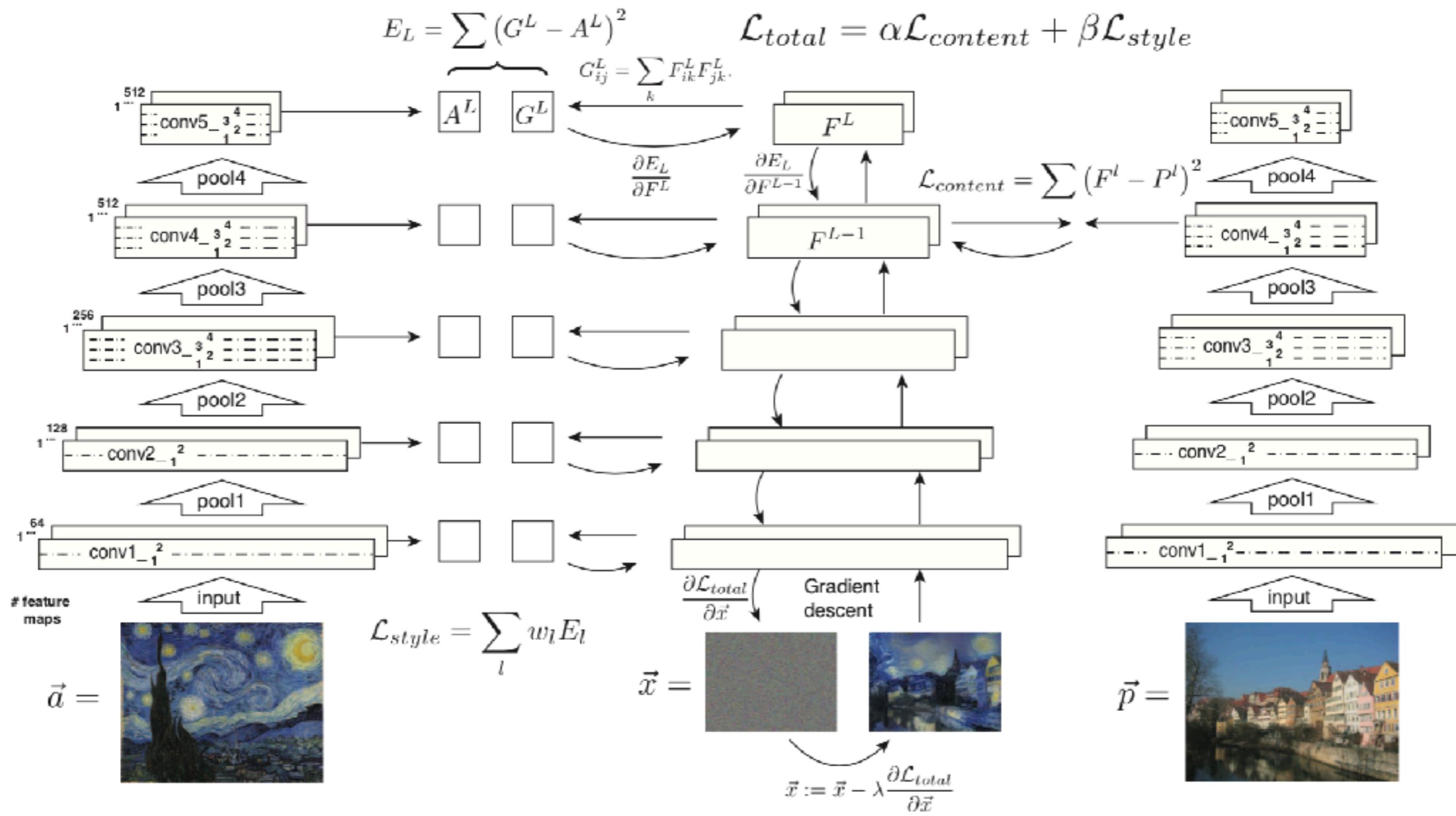
Style transfer is a technique where a Deep Learning network can transfer artistic styles from known pieces of art to new images

Demo

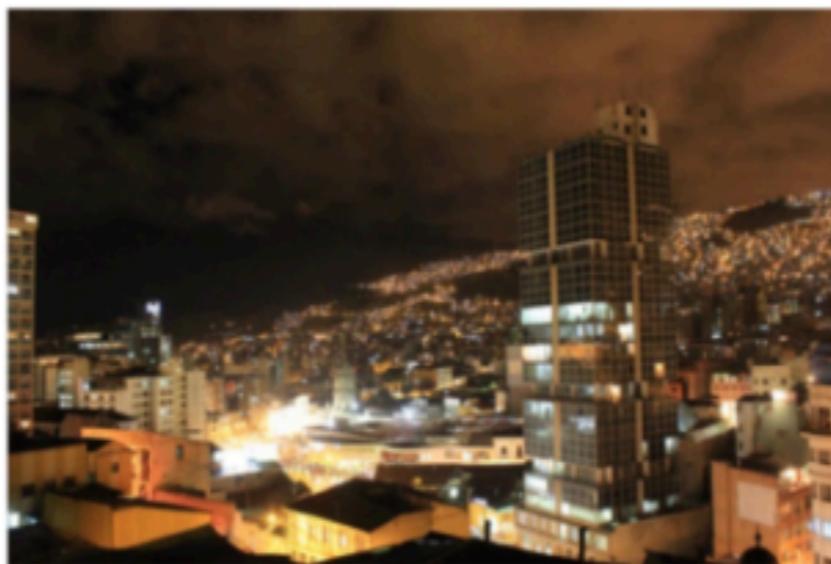
<http://genekogan.com/works/style-transfer/>

Algorithm





Results



Text Generator

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.

This text reads like Shakespeare, but was actually written by a Deep Learning (LSTM) network who read a lot of Shakespeare.

Naturalism and decision for the majority of Arab countries' capitalide was grounded by the Irish language by [[John Clair]], [[An Imperial Japanese Revolt]], associated with Guangzham's sovereignty. His generals were the powerful ruler of the Portugal in the [[Protestant Immineners]], which could be said to be directly in Cantonese Communication, which followed a ceremony and set inspired prison, training. The emperor travelled back to [[Antioch, Perth, October 25|21]] to note, the Kingdom of Costa Rica, unsuccessful fashioned the [[Thrales]], [[Cynth's Dajoard]], known in western [[Scotland]], near Italy to the conquest of India with the conflict. Copyright was the succession of independence in the slop of Syrian influence that was a famous German movement based on a more popular servicious, non-doctrinal and sexual power post. Many governments recognize the military housing of the [[Civil Liberalization and Infantry Resolution 265 National Party in Hungary]], that is sympathetic to be to the [[Punjab Resolution]] (PJS)[<http://www.humah.yahoo.com/guardian.cfm/7754800786d17551963s89.htm> Official economics Adjoint for the Nazism, Montgomery was swear to advance to the resources for those Socialism's rule, was starting to signing a major tripad of aid exile.]

Deep Learning networks can be trained on various types of text. Here the network generates a Wikipedia article from scratch.

LSTM

https://www.wikiwand.com/en/Long_short-term_memory

Math Text

<http://cs.stanford.edu/people/jcjohns/fake-math/4.pdf>

```
static void start_ctrl_regs_pc_filter(struct function *feature,
                                      struct fuse_ctrl *p_ctrl,
                                      unsigned int cur_ctrl, unsigned int dfl_sched_ok)
{
    struct fuse_ctrl *ctrl;
    int enabled;

    ctrl = firmware->fsr;
    false = false;

    if ((media_entities->type & FUSE_TYPE_CCR) == MFC_CONFIG) {
        if (fieldmode == FLIP_CONTROL) {
            ctrl_reg = FLUSH_CTRL;
            fimc->get_time = 1;
        }
    }

    rc = ctrl_get_field32(&ctrl->lock, FUNCTION_RV,
                         TYPE_FLDSCR_NACK);
    if (rc != 0) {
        pr_err("failed to retrieve t1pc for frame toggle register.\n");
        goto rdev_init_fail;
    }

    fname = t1pci_int;
    ctrl_reg = FIELD26(RDS, rc);
    reg |= field(fieldmode, 8);

    if (n < sizeof(ctrl))
        return -EINVAL;

    return;
}

error_out:
    return ret;
}

int firmware_calc_n_frames(struct nfc_hci_ctlr *fc, struct firmware *fw)
{
    struct firmware *fw = &libcfs_dev->ctrl_ri;

    strlcpy(five_taps, "1", len);
    first_firmware_status = 0; /* match as follows: remove the states */

    ctrl->state = STATUS_HCA_TRANSFER;
    ctrl->state = FIP_STAT_PROXIMITY;
    ctrl->poll_count = 0;
```

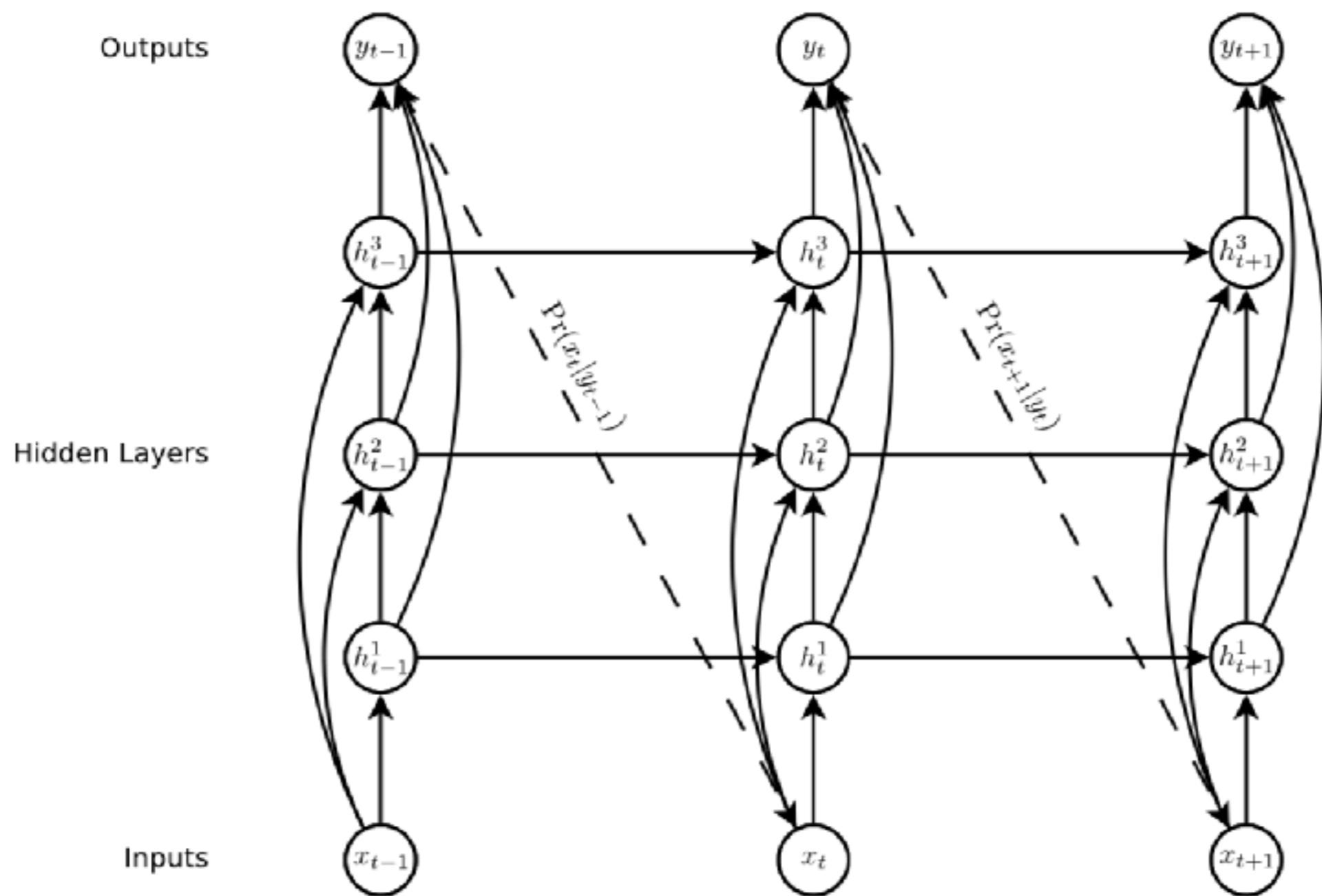
Handwriting Generator

recurrent neural network handwriting generation demo

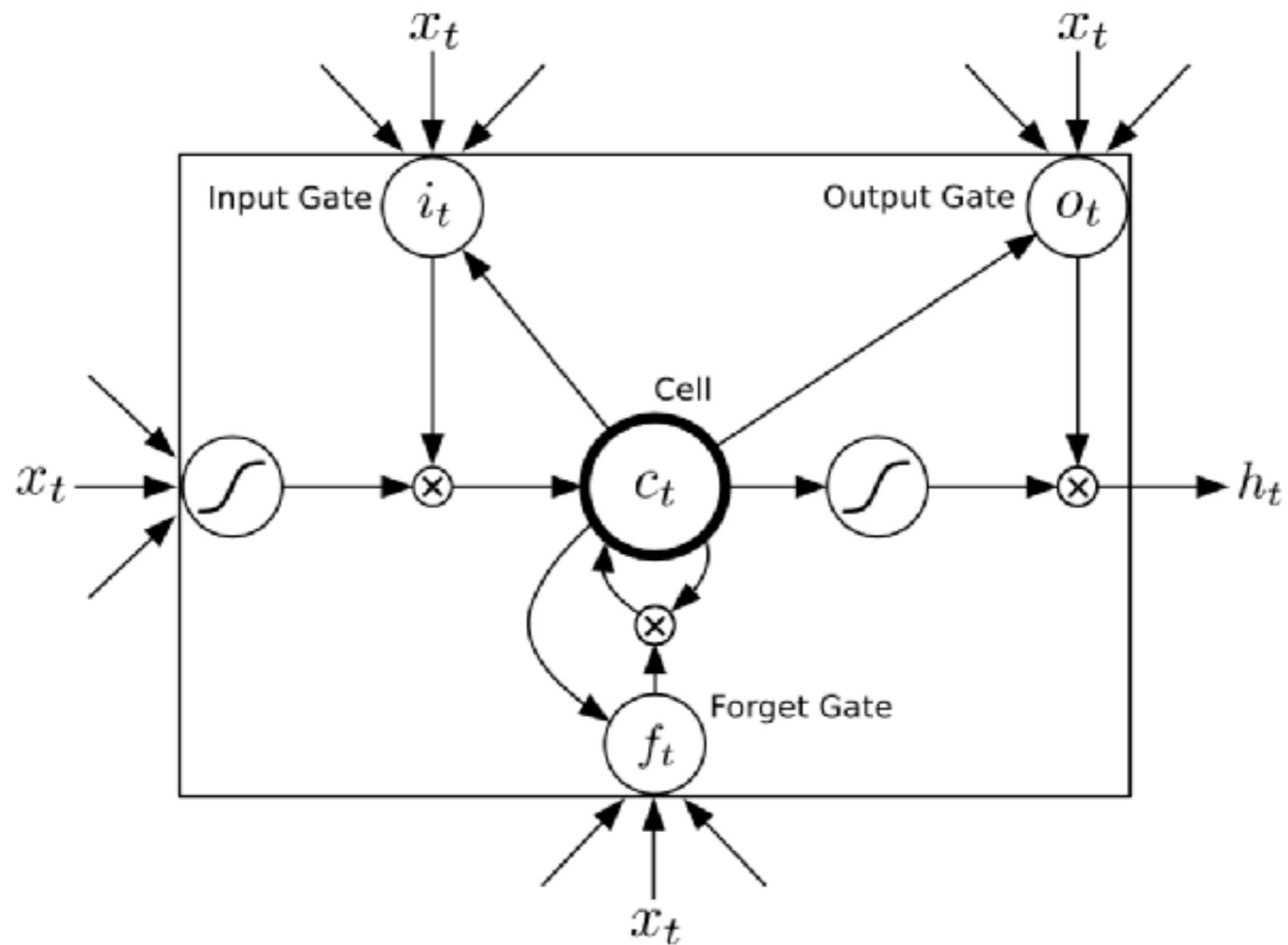
This is a handwritten text created by a computer, and not a hand! Tap on it to generate your own texts on the blog of Alex Graves.

Demo

<http://www.cs.toronto.edu/~graves/handwriting.html>



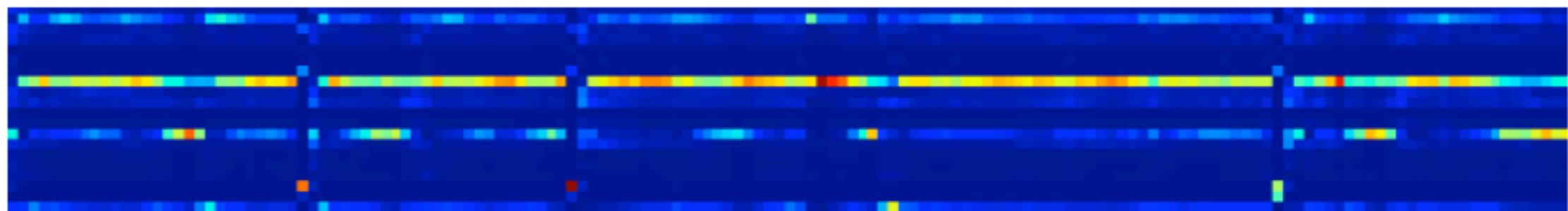
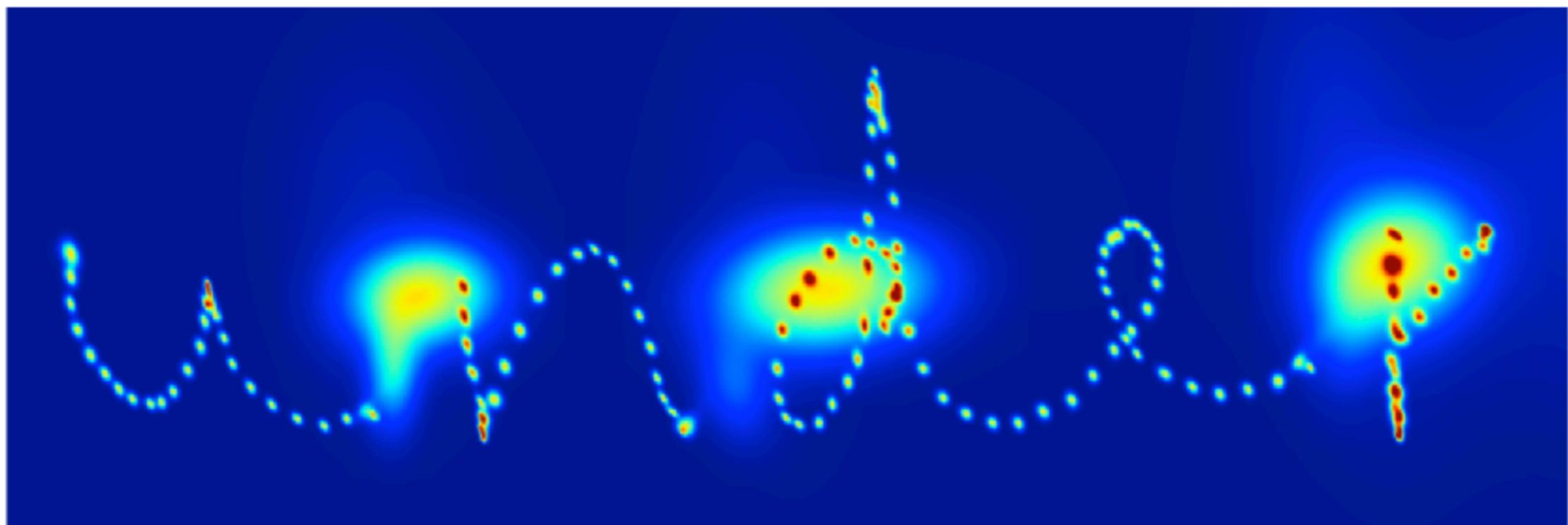
LSTM Cell



Samples

would find the bus safe and sound
As for Clark, unless it was a
cancer at the age of fifty-five
Editorial. Dilemma of
the tides in the affairs of men;

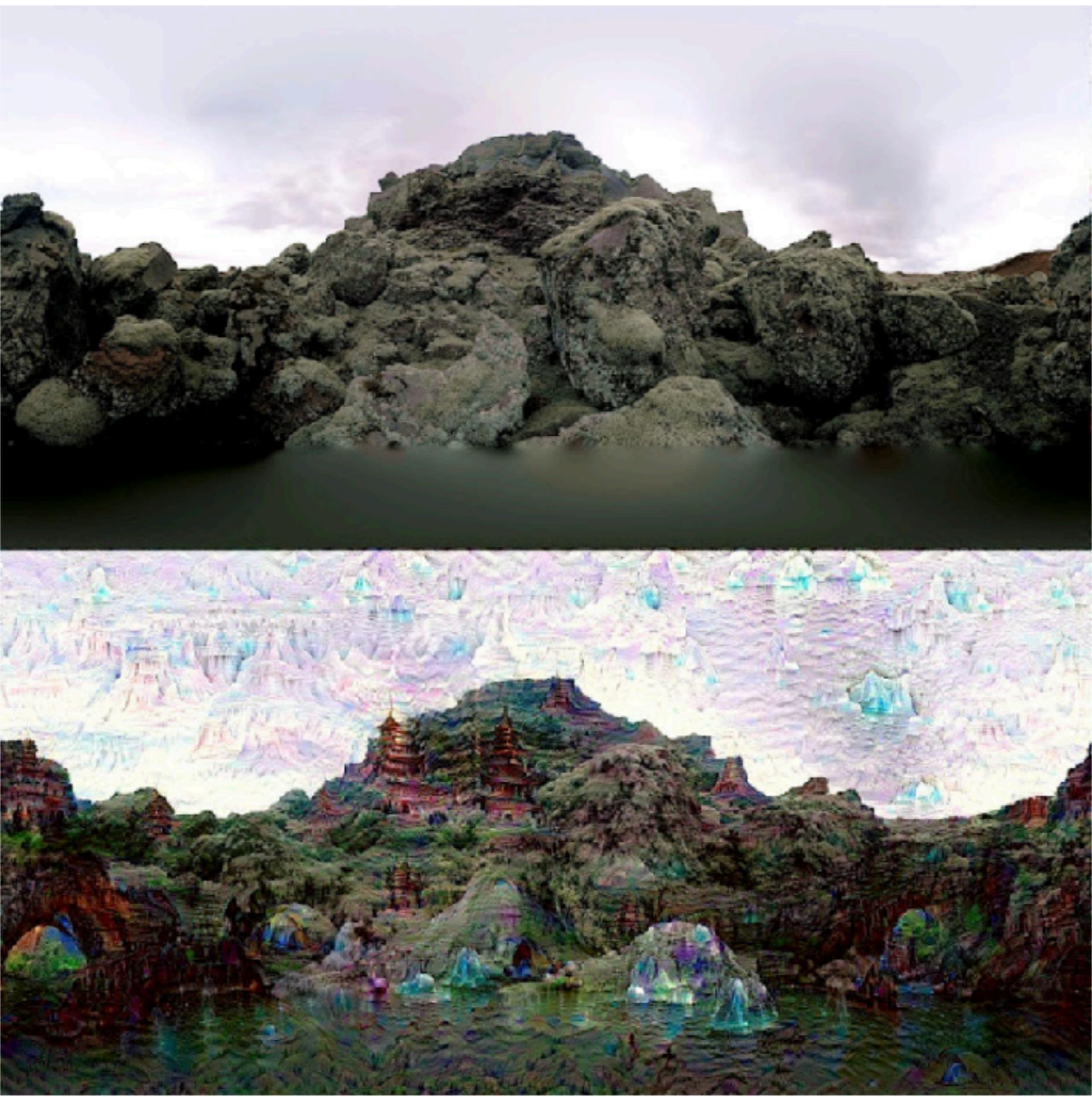
Weight Density Maps



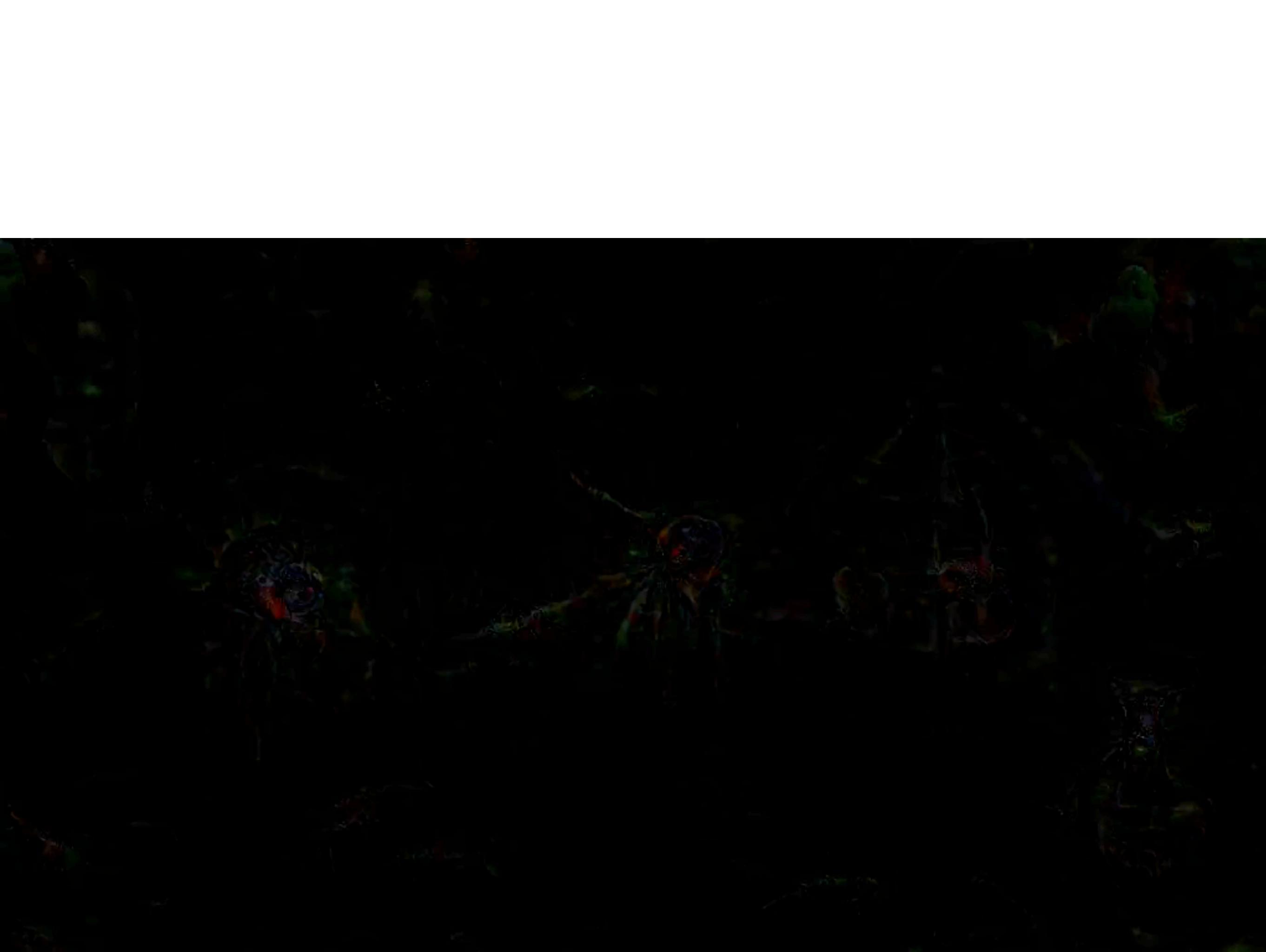
urn ay under Goucage there will
egg med anche. ' beskies the
Anaine Cenele of hye Woodit's
see Boung a. me accretions fa
purne h' mistaen b'w lured
bopes & cold winefs wine cura
heist. Y Geesh the gather w
spycle satet Jomw In doing Te a
nver & highe eame. Tend., madp

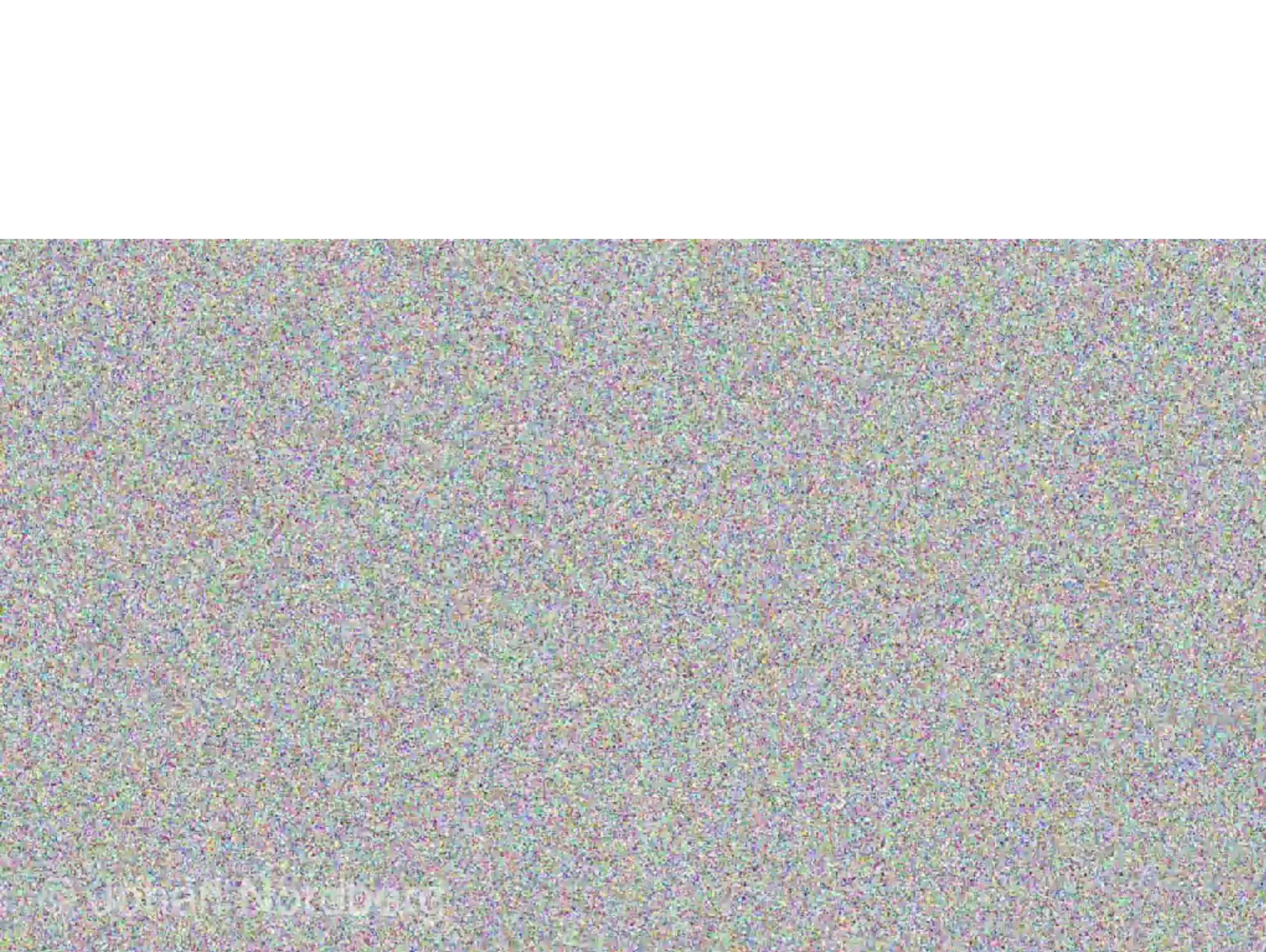
more of national temperament

Deep Dream



The computer hallucinates imaginary buildings and structures on top of this hill.





Deep Code

neural complete!

Choose a model

neural_token

Add

```
# generated code follows below  
from keras.layers.core import Dense, Activation  
from keras.models import Sequential  
model = Sequential()
```

Neural Complete is a Deep Learning network that can write Deep Learning code!

NLP Medication Parser

Sig Parser

Q Amoxicillin 250mg/5ml suspension (amoxicillin) X

AMOXICILLIN 250MG/5ML SUSPENSION (AMOXICILLIN) [Save as Favourite](#)

Instructions

t 250mg tid x10d [Print](#)

Take 3 times a day for 10 days

Intelligence

Response Code

200

Response Headers

```
{  
    "access-control-allow-origin": "*",
    "date": "Tue, 18 Jun 2019 17:25:41 GMT",
    "server": "Werkzeug/0.14.1 Python/3.6.6",
    "content-length": "507",
    "x-context": "{ 'message': 'Prescribed DIN is not valid', 'type': 'DinNotFoundError', 'level': 'critical' }",
    "content-type": "application/json"
}
```

```
"",
"patientInstruction": "take 1 tablet of 500mg Cephalexin",
"text": "cephalexin t1t"
```

Response Code

200

Response Headers

```
{  
    "access-control-allow-origin": "*",
    "date": "Wed, 19 Jun 2019 16:34:07 GMT",
    "server": "Werkzeug/0.14.1 Python/3.6.6",
    "content-length": "2261",
    "x-context": "{ 'message': 'Conflict between prescribed DIN Ceftin/cefuroxime axetil and specified drugname cephalexin' }",
    "content-type": "application/json"
}
```

Intelligence

Precision Verification:

```
{  
  "x-context": [  
    {  
      "message": "No exact strength match is found",  
      "type": "StrengthNotFoundError",  
      "level": "critical"  
    },  
    {  
      "message": "Closest drug strength found for 300mg is 2",  
      "type": "ClosestStrengthFoundError",  
      "level": "critical"  
    }  
  ]  
}
```

t2t oxyno bid x30d

x

Precision Verification:

```
[  
  {  
    "message": "Common Drug strength is provided",  
    "type": "MissingStrengthDataError",  
    "level": "critical"  
  }  
]
```

Formulary:

```
{  
  "details": "No details found for din 2372568",  
  "din": "2372568"  
}
```

Precision Intelligence:

```
{  
  "medicationRequest": {  
    "dispenseRequest": {  
      "expectedSupplyDuration": {  
        "unit": "d",  
        "value": 30  
      },  
      "status": "active"  
    }  
  }  
}
```

Limitations