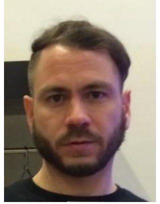
5th Vienna Deep Learning Meetup

22 September 2016 @ Automic Software

Automic™



Hosts: Josef Puchinger, Thomas Lidy, Jan Schlüter

Deep Learning

5th Vienna Deep Learning Meetup

Agenda:

- Welcome (Thomas Lidy)
- Deep Learning & The Future of Automation (Josef Puchinger, Automic Software)
- Going Deeper with GoogLeNet and CaffeJS (Christoph Körner)
- Latest News / Hot topics (Jan Schlüter, Alex Schindler)
- Open discussions

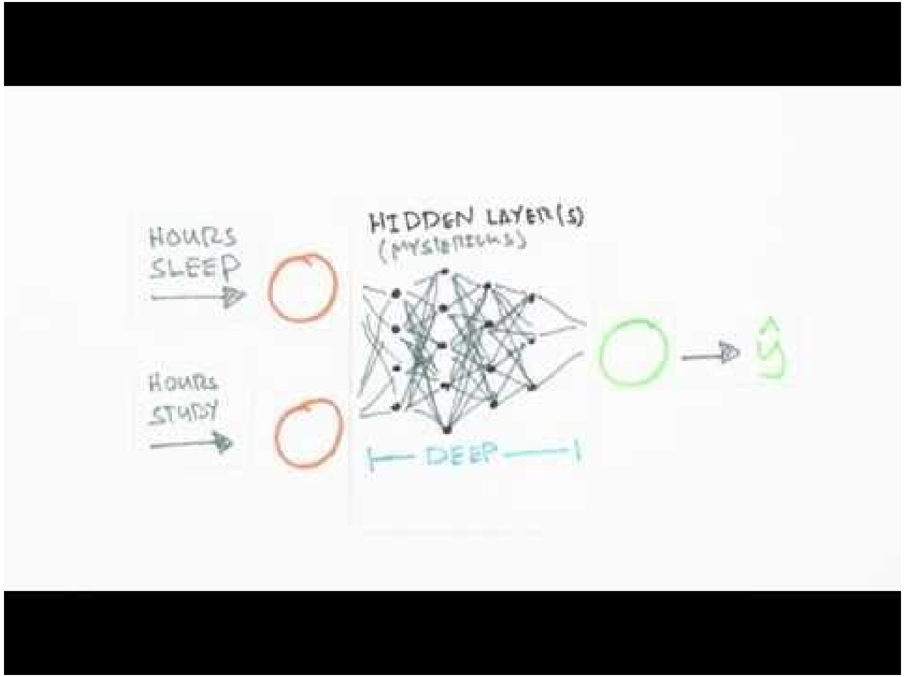
Latest News
Hot Topics

a 5-10 min block at every meetup to briefly present "trending topics"

Send us contributions (tom.lidy@gmail.com)
or come with slides to do a 5-10 min block yourself!

Neural Networks Demystified

Really nice introduction into Machine Learning / NN / DL / Data Preparation, Classification/Regression, etc.



https://www.youtube.com/watch?v=bxe2T-V8XRs

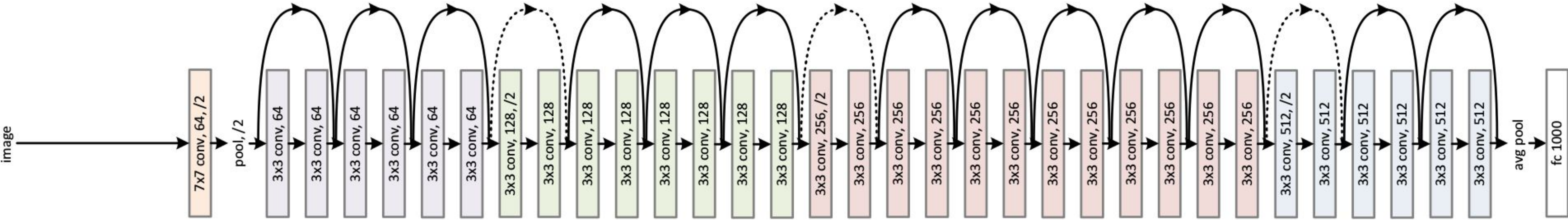
Apache Singa

- new open source **distributed** deep learning platform by Apache
- v 1.0 released on 8 Sep 2016
- flexible and scalable (“for big data analytics”)
- simple programming model
- making the distributed training process transparent to users
- training of DL models on a **GPU cluster** possible

http://singa.incubator.apache.org/

Residual Network improvements

Residual Network (Dec 2015):

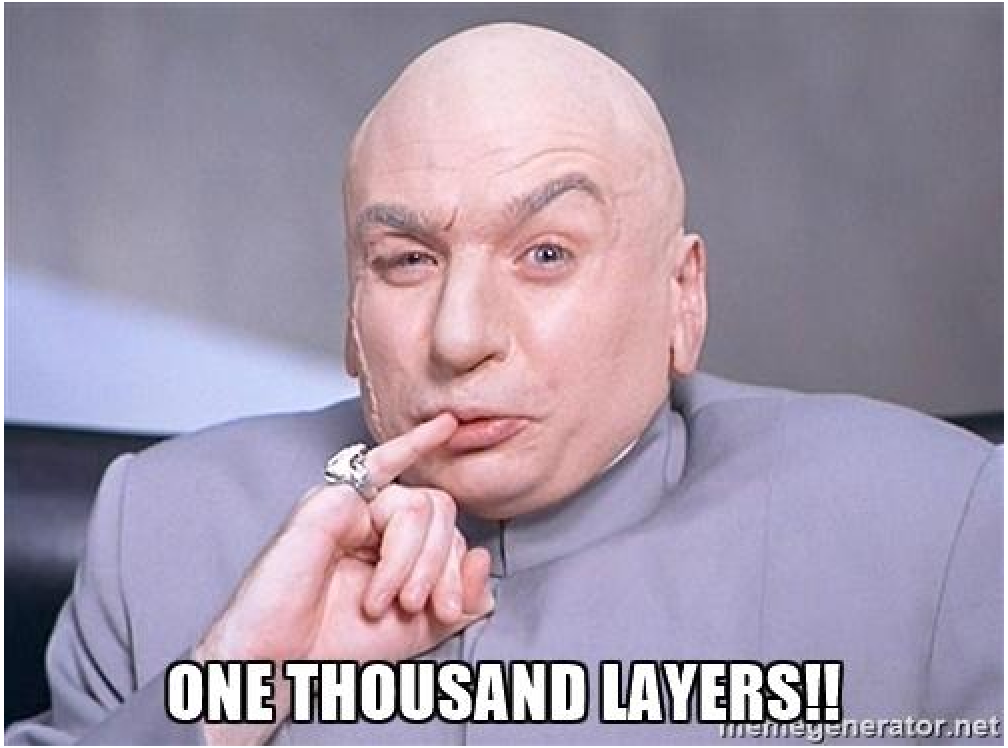


Allow deeper networks via skip connections.

Skip connections add input to output.
Thus, each layer learns to correct residual errors.

https://arxiv.org/abs/1512.03385, https://arxiv.org/abs/1603.05027

Residual Network improvements



Residual Network improvements

Wide ResNet (May 2016):

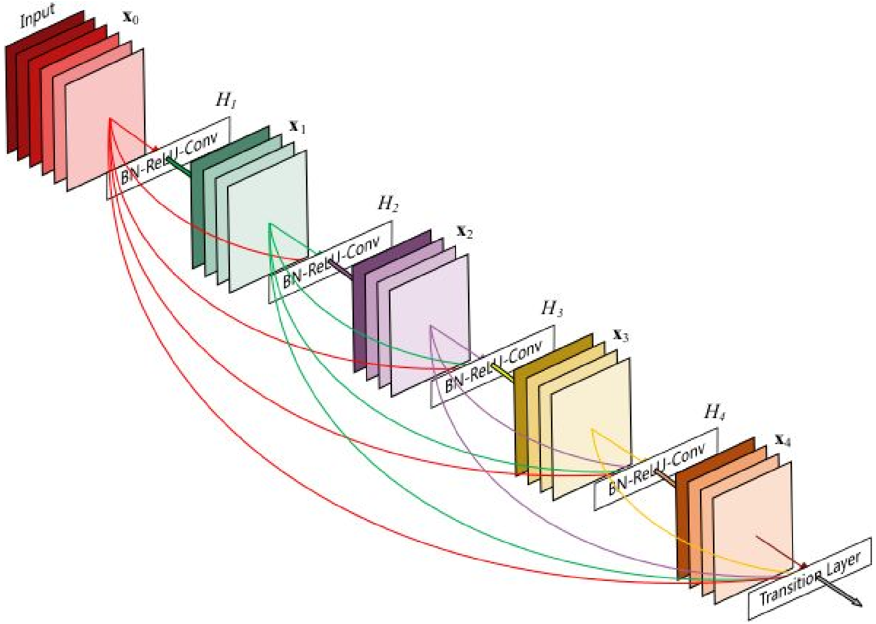
Make network wider (more filters), not deeper.

Wide ResNet of 16 layers > original ResNet of 1001 layers.

https://arxiv.org/abs/1605.07146

Residual Network improvements

DenseNet (Aug 2016):



Similar to ResNet, but skip connections *concatenate* input to output instead of adding it.

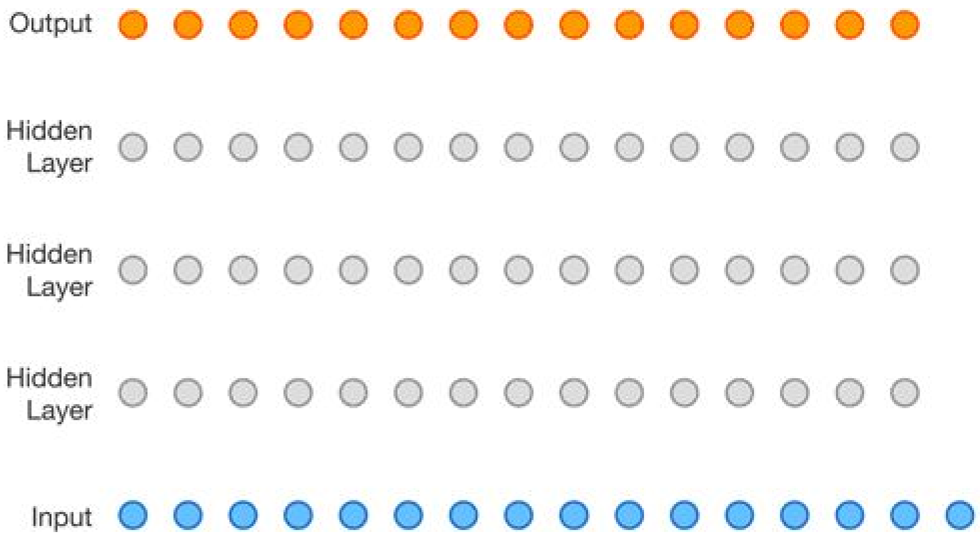
Each layer can access the outputs of all previous layers, and adds a few feature maps to the pool.

SOTA on CIFAR10/100, SVHN

https://arxiv.org/abs/1608.06993

WaveNet

ConvNet that can **predict next step of time sequence**, using a clever architecture for processing a large temporal context (about 3000-6000 past time steps)



https://deepmind.com/blog/wavenet-generative-model-raw-audio/

WaveNet

Trained to **generate** samples of **raw audio** (200-300ms context)

Without side inputs:
example A example B

With side inputs (phones, speaker id):
example A example B example C

Without side inputs, trained on piano YouTube videos:
example A example B

https://deepmind.com/blog/wavenet-generative-model-raw-audio/

Deep Learning

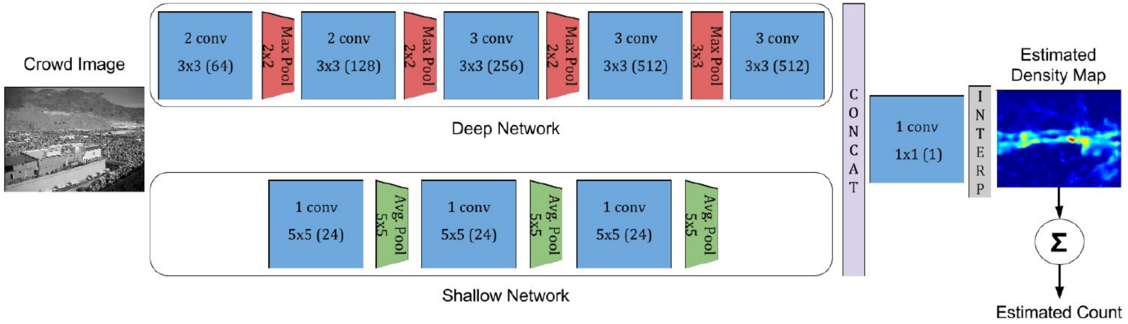
CrowdNet: A Deep Convolutional Network for Dense Crowd Counting

Lokesh Boominathan
Video Analytics Lab
Indian Institute of Science
Bangalore, INDIA - 560012
boominathanlokesh@gmail.com

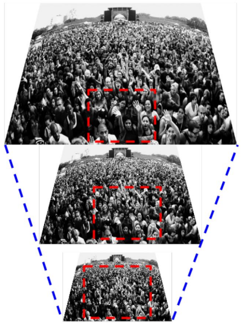
Srinivas S S Kruthiventi
Video Analytics Lab
Indian Institute of Science
Bangalore, INDIA - 560012
kssaisrinivas@gmail.com

R. Venkatesh Babu
Video Analytics Lab
Indian Institute of Science
Bangalore, INDIA - 560012
venky@cds.iisc.ac.in

Network Architecture



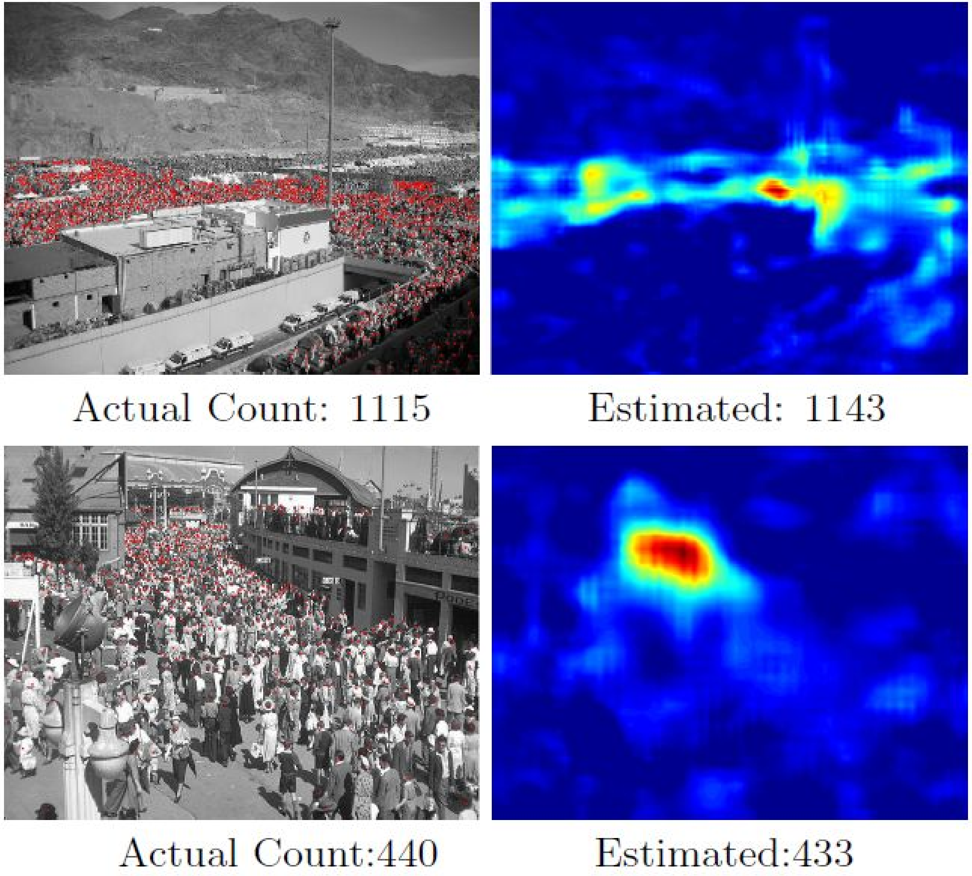
Scale Invariance



http://arxiv.org/abs/1608.06197

Deep Learning

CrowdNet: A Deep Convolutional Network for Dense Crowd Counting



CrowdNet: A Deep Convolutional Network for Dense Crowd Counting

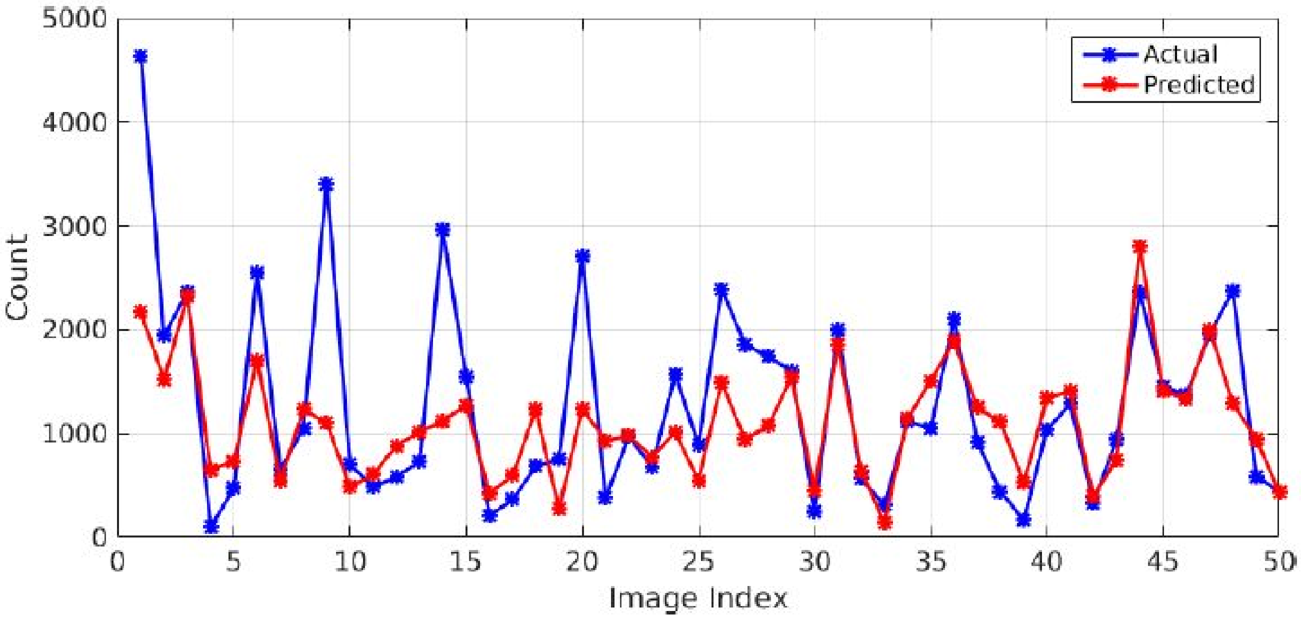


Figure 4: Actual count vs. Predicted Count for each of the 50 images in the UCF_CC_50 dataset.

Deep Learning

Image super-resolution through deep learning

- Upscale 16x16 images by factor 4

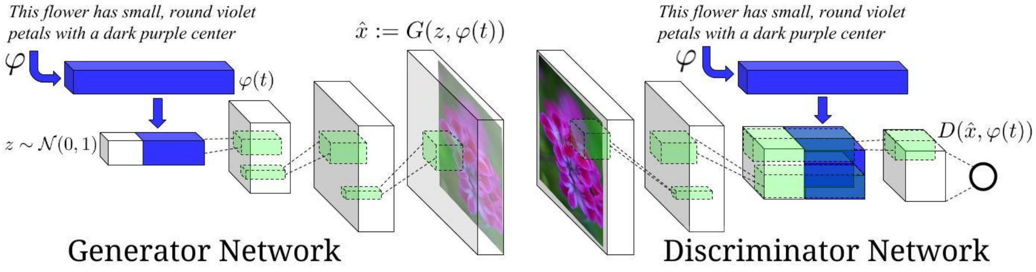
- Deep Convolutional Generative Adversarial Network (DCGAN)
  - Image as input instead of gaussian noise
  - Loss function measures difference between input and scaled version
  - Generator uses Residual Network (ResNet) modules

- Code on Github
  - https://github.com/david-gpu/srez

Deep Learning

Generative Adversarial Text to Image Synthesis

Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran REEDSCOT[1], AKATA[2], XCYAN[1], LLAJAN[1]
Bernt Schiele, Honglak Lee SCHIELE[2],HONGLAK[1]
[1] University of Michigan, Ann Arbor, MI, USA (UMICH.EDU)
[2] Max Planck Institute for Informatics, Saarbrücken, Germany (MPI-INF.MPG.DE)



- Paper: http://arxiv.org/abs/1605.05396
- Github: https://github.com/paarthneekhara/text-to-image

Generative Adversarial Text to Image Synthesis

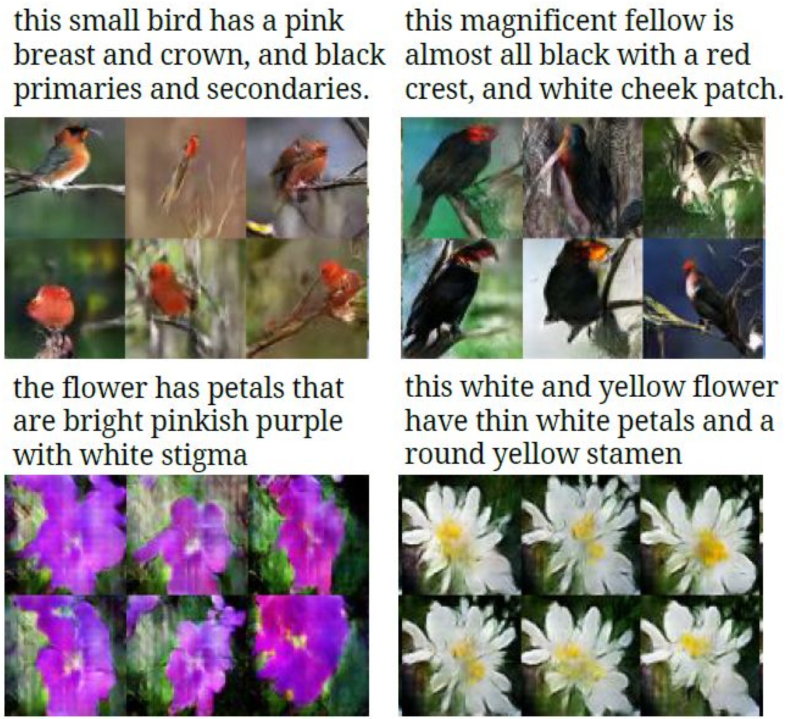
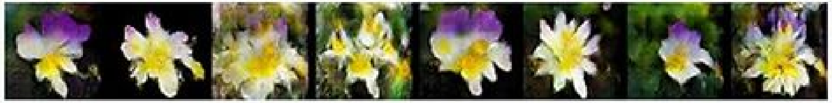
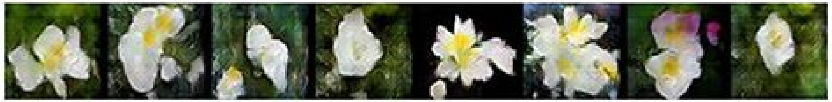
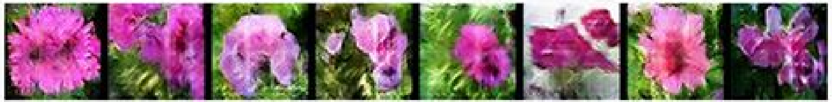


Figure 1. Examples of generated images from text descriptions. Left: captions are from zero-shot (held out) categories, unseen text. Right: captions are from the training set.

Caption	Generated Images
the flower shown has yellow anther red pistil and bright red petals	
this flower has petals that are yellow, white and purple and has dark lines	
the petals on this flower are white with a yellow center	
this flower has a lot of small round pink petals.	
this flower is orange in color, and has petals that are ruffled and rounded.	
the flower has yellow petals and the center of it is brown	

Deep Learning

Learning Temporal Transformations From Time-Lapse Videos

Yipin Zhou   Tamara L. Berg

University of North Carolina at Chapel Hill
{yipin,tlberg}@cs.unc.edu

Ground Truth

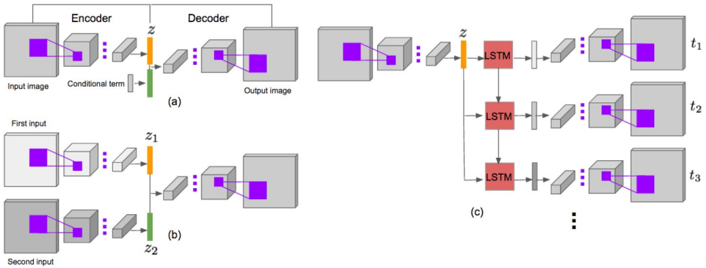
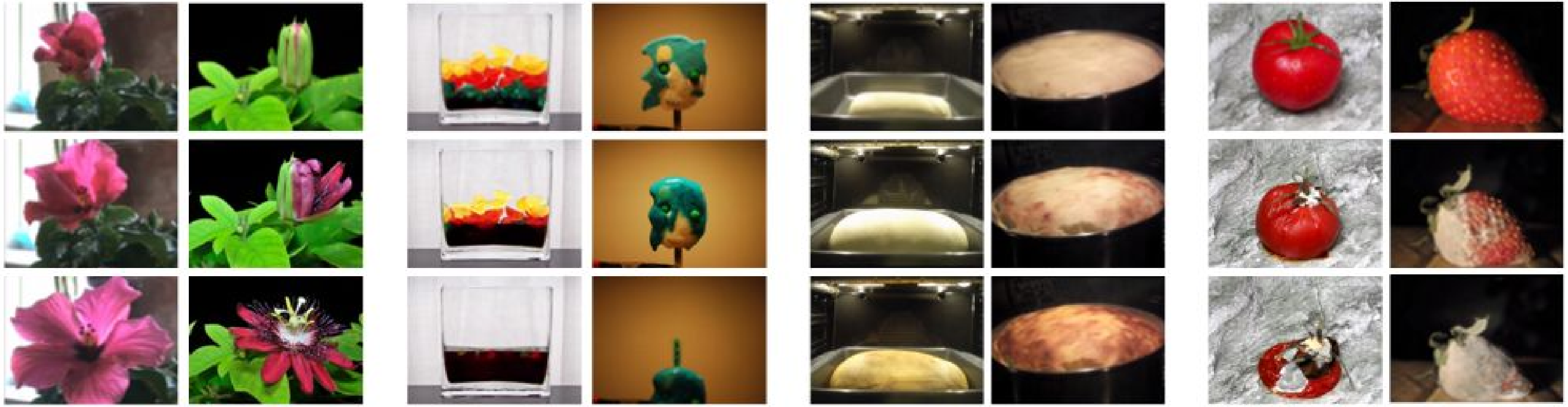


Fig. 2. Model architectures of three generation tasks: (a) Pairwise generator; (b) Two stack generator; (c) Recurrent generator.

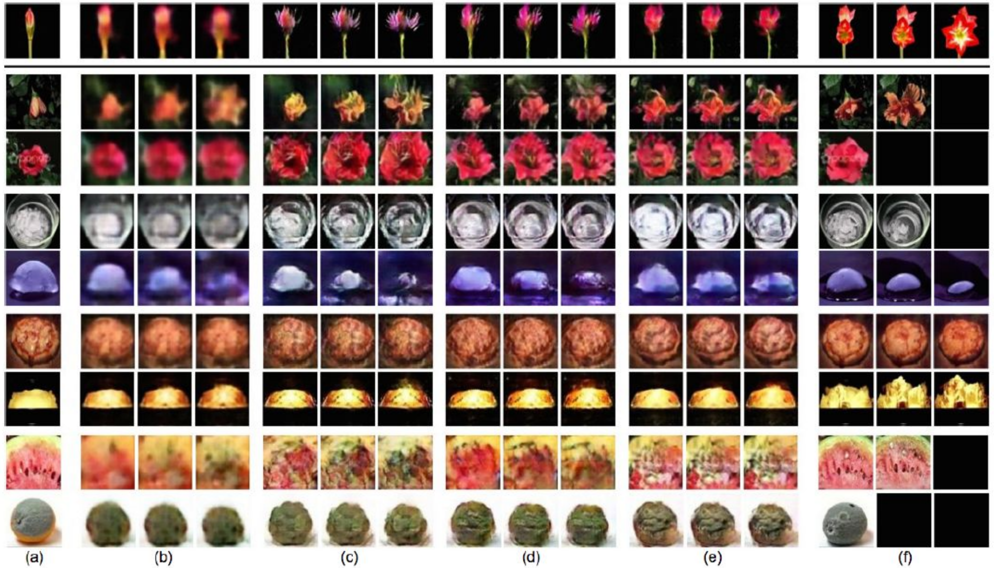
Learning Temporal Transformations From Time-Lapse Videos



Blooming Melting Baking Rotting

Deep Learning

Learning Temporal Transformations From Time-Lapse Videos



Deep Learning

Pre-Trained Models now available in Lasagne and Keras

Available models

Models for image classification with weights trained on Im

- VGG16
- VGG19
- ResNet50
- InceptionV3

Examples

Classify ImageNet classes with ResNet50

```
from keras.applications.resnet50 import ResNet50
from keras.preprocessing import image
from keras.applications.resnet50 import preprocess_i

model = ResNet50(weights='imagenet')

img_path = 'elephant.jpg'
img = image.load_img(img_path, target_size=(224, 224
x = image.img_to_array(img)
x = np.expand_dims(x, axis=0)
x = preprocess_input(x)

preds = model.predict(x)
print('Predicted:', decode_predictions(preds))
# print: [[u'n02504458', u'African_elephant']]
```

Lasagne recipes: examples, IPython notebooks, ...

86 commits | 3 branches | 0 releases

Branch: master | New pull request

f0k committed on GitHub Merge pull request #60 from joelmoniz/recipe/rnn_word_gen ...

examples	Merge pull request #60 from jo
modelzoo	added header to resnet50.py; a
papers	Fixed typo in README
stale	some directory structure
tutorials	add tutorial links
utils	minor changes, added docume
.gitignore	reverse .gitignore to default
LICENSE	Initial commit
README.md	Update README.md

README.md

- c3d.py
- caffe_reference.py
- cifar10_nin.py
- googlenet.py
- inception_v3.py
- resnet50.py
- vgg16.py
- vgg19.py
- vgg_cnn_s.py

Deep Learning

Pretrained models
5-min Quick-Test

red_fox
dhole
kit_fox
grey_fox
Arctic_fox
red_wolf
coyote
timber_wolf
white_wolf
lion

monastery
palace
fountain
church
bell_cote
castle
triumphal_arch
analog_clock
obelisk
dome

Using Python and Keras/Theano

Deep Learning

DL for Sound Recognition

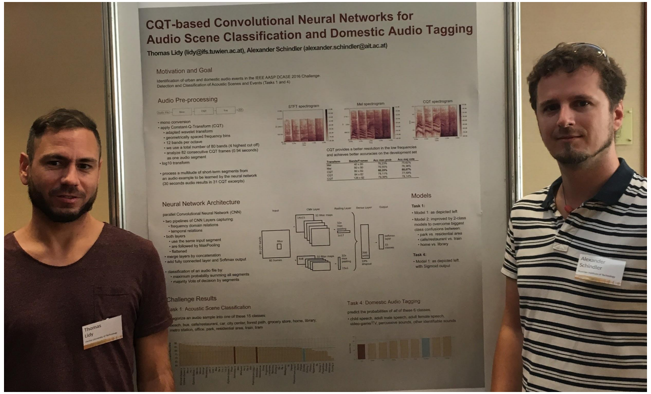
Recent IEEE AASP DCASE Workshop in Budapest

- Urban audio scene classification (bus, metro, tram, park, …)
- Audio event detection in a stream (keyboard, door, phone …)
- Domestic audio tagging (child, male, female, TV, game, household)

-> many used CNN/DNN

- JKU Linz winner of Audio scene task
- TU Wien winner of Domestic audio task

Data + algorithm descriptions available:
http://www.cs.tut.fi/sgn/arg/dcase2016/



Announcements

Deep Learning

Call for Meetup Topics / Talks

Looking for:

- Industry Applications of DL
- Real World Use Cases of DL
- Interesting novel scientific results
- Niche applications
- Latest News / Hot topics

Get in touch if you want to contribute something!

tom.lidy@gmail.com

Mon Style

Job Offer in DL Startup

Mon Style is a tech startup that uses **machine-learning** to deliver highly **accurate results to online shoppers**. Their mission is to revolutionize the way we discover, interact and shop products.

Job Profile:

- Full time position based in Vienna
- Flexible working hours and home office
- Knowledge of algorithms and data structures, Python
- Knowledge of machine learning (Classification, Regression and Clustering)
- Experience using deep learning frameworks like Caffe/Torch/Theano/TensorFlow + numpy, pandas etc.

Application: office@mon-style.com

BitCoin Analytics

- Funded MSC Thesis
- Analyzing Bitcoin transactions
  - Fraud Detection
  - Anomaly Detection
  - Large Scale Data Analysis

Contact: Alexander.Schindler@ait.ac.at and Bernhard.Haslhofer@ait.ac.at

Waves Music Hackday

Register! http://www.wavescentraleurope.com/waves-music-hackday

netidee Open Source Community Camp

https://www.whataventure.com/events/netidee/2016/

Thanks a lot to Automic Software for hosting us!