Visual Content Generation via Machine Learning

› 113451a - Selected Topics of Artificial Intelligence WS19/20
Eröffnung - 25. April 2019
Get in touch

https://ai.hdm-stuttgart.de

mail: iaai@hdm-stuttgart.de

AI-Lab: Nobelstraße 10 - R143
About me

› Johannes Theodoridis

› PhD Student @ HdM & Uni Tübingen

› Audiovisuelle Medien @ HdM
› Computer Science and Media @ HdM
› Exchange @ KTH Stockholm

› You can do projects or your thesis with me 😊
› mail: theodoridis@hdm-stuttgart.de (or talk to me, Al-Lab R143)

(First slide: Left: A Neural Representation of Sketch Drawings [Ha and Eck 2017], Middle: Image-to-Image Translation with Conditional Adversarial Networks [Isola et al. 2017], Right: https://prisma-ai.com)
What do you do?
Deep Learning

What society thinks I do

What my friends think I do

What other computer scientists think I do

What mathematicians think I do

What I think I do

What I actually do

In [1]:
import keras
Using TensorFlow backend.
How to do Machine Learning

1. Collect Data
2. Create Dataset
3. Conduct Research

https://quickdraw.withgoogle.com
https://quickdraw.withgoogle.com/data
https://arxiv.org/abs/1704.03477
3. Conduct Research

A Neural Representation of Sketch Drawings
[Ha and Eck 2017]

https://arxiv.org/abs/1704.03477
How to do Machine Learning

1. Collect Data
2. Create Dataset
3. Conduct Research
4. Build more cool AI tools

https://quickdraw.withgoogle.com
https://quickdraw.withgoogle.com/data
https://arxiv.org/abs/1704.03477
https://magenta.tensorflow.org/assets/sketch_rnn_demo/index.html
Image Classification

Task: Classify an image based on the dominant object inside it.

Input: Image

Output: Class Label

The Neural Network learns a **mapping from pixels to class labels.**

- herd of sheep
- pack of wolves
Image Classification

**Task:** Classify an image based on the dominant object inside it.

**Input:**
X: Image

**Output:**
Y: Class Label

The Neural Network learns a (complicated) function, mapping from pixels to class labels.
Image Classification

Task: Classify an image based on the dominant object inside it.

Input: X: Image

Output: Y: Class Label

The Neural Network learns a (complicated) function, mapping from pixels to class labels.
Image Classification

**Task:** Classify an image based on the dominant object inside it.

**Input:**
X: Image

**Output:**
Y: Class Label

\[ f(x) = y \]

The Neural Network learns a (complicated) function, mapping from pixels to class labels.
Object Recognition

Task: Localize and classify all objects appearing in the image.

Input:
X: Image

Output:
Y: Labeled Bounding Boxes

Deep Neural Network

The Neural Network learns a mapping from pixels to Bounding Boxes.
Video Game Play

**Task:** Select the best Action in the current game situation (state).

**Input:**

X: Game Screen

**Output:**

Y: Next best Action

The Neural Network learns a **mapping from pixels to Actions.**
Visual Content Generation via Machine Learning

› 113451a - Selected Topics of Artificial Intelligence WS19/20
Inceptionism: Google DeepDream 2015

Left: Original photo by Zachi Evenor. Right: processed by Günther Noack, Software Engineer

https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html
Inceptionism: Google DeepDream 2015

Deep Neural Network

- Low-Level Features
- Mid-Level Features
- High-Level Features
- Trainable Classifier

Car

Banana
Inceptionism: Google DeepDream 2015

Understanding Neural Networks Through Deep Visualization
[Yosinski et al. – ICML 2015]
Inceptionism: Google DeepDream 2015

- Run the network in reverse.
- Keep the network fixed but gradually change the input picture to match a new object.
Inceptionism: Google DeepDream 2015
Artistic Style Transfer 2015

A Neural Algorithm of Artistic Style
[Gatys, Ecker, Bethge 2015]
https://deepart.io
Artistic Style Transfer 2015

- Start with a random noise image.
- Gradually force the image to incorporate features from the content and style sources.

Intuition: How good does the new image match the activations from the content and style images.

Reduce a combined Content + Style Loss
Artistic Style Transfer 2015

- Start with a random noise image.
- Gradually force the image to incorporate features from the content and style sources.

Intuition: How good does the new image match the activations from the content and style images.

Reduce a combined Content + Style Loss

Content Image
Translation Tasks

Text-to-Text

Translate from German (detected) → Translates into English →
Heute scheint die Sonne.
The sun is shining today.

Image-to-Text

Image caption: A woman is throwing a frisbee in a park.

Text-to-Image

The bird is blue with white and has a very short beak.

Image-to-Image

Real-Time User-Guided Image Colorization with Learned Deep Priors
[Zhang et al. 2017]

Show, Attend and Tell: Neural Image Caption Generation with Visual Attention
[Xu et al. 2015]

StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks
[Zhang et al. 2017]

DeepL, Google Translate etc.
Image-to-Image Translation

Colorful Image Colorization
[Zhang, Isola, Efros 2016]

Real-Time User-Guided Image Colorization with Learned Deep Priors
[Zhang et al. 2017]

Style2Paints
https://github.com/lllyasviel/style2paint/tree/master/V3

AI-Powered Automatic Colorization
https://paintschainer.preferred.tech/
Image-to-Image Translation

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks
[Zhu and Park et al. 2017]
› **Task:**
  › Synthesize realistic looking images, art, …

› **Problem:**
  › We don’t know how to formalize our objective. What is a good measure of “realistic”?

› **Idea:**
  › We as humans can easily assess what looks realistic and what’s not. Why not train a Deep Neural Network that can do the same?
Generative Adversarial Networks (GAN) [Goodfellow et al. 2014]

The Generator network will try to generate fake images that fool the discriminator. The Discriminator network will try to distinguish between a real and a generated image.
Welcome to the GAN Zoo

 › Generate bedrooms - 2016

Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks
[Radford et al. ICLR 2016]
Welcome to the GAN Zoo

› Generate bedrooms, buildings, cats - 2017

StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks
[Zhang et al. 2017]
Welcome to the GAN Zoo

- Generate celebrities 2018

High resolution: 1024 x 1024 pixel

Progressive Growing of GANs for Improved Quality, Stability, and Variation
[Harras et al. 2018]

IntroVAE: Introspective Variational Autoencoders for Photographic Image Synthesis
[Huang et al. 2018]
GAN Progress – 2014 to SOTA

All images here are generated (fake)!

A Style-Based Generator Architecture for Generative Adversarial Networks
(Karras, Laine, Aila 2018) [https://github.com/NVlabs/stylegan]
More GANs

Generate Anime Characters

https://make.girls.moe/
More GANs

- Image-to-Image translation

Image-to-Image Translation with Conditional Adversarial Networks [Isola et al. 2017]

Edges2cats

Image-to-Image Demo
https://affinelayer.com/pixsrv/
Vision

Interactive drawing applications

Generative Visual Manipulation on the Natural Image Manifold
[Zhu et al. 2016]

Suggestive Drawing among humans and artificial intelligences.
[Master Thesis Harvard Graduate School of Design 2017]
nono.ma
Vision

› Interactive drawing applications

Suggestive Drawing among humans and artificial intelligences.
[Master Thesis Harvard Graduate School of Design 2017]
nono.ma

NVIDIA Research - SPADE
nvlabs.github.io/SPADE/
What do you think?

Do we have creative AI?
What are we missing?

These questions are easy for us because we have an abstract model of the world.
Ok, let’s do something more “intelligent” :D

Creative drawing game (10min)

1. Team up with your neighbour.
2. Silently! Think about a random real animal.
3. If you are both ready, tell your partner what kind of animal you have imagined.
4. Both of you, draw a combination of those two animals.

- Drawing skill doesn’t matter!!!
- Be creative 😊

If you like, send me a pic of your drawings!
- theodoridis@hdm-stuttgarg.de
What did just happen?

You created something truly “new”.

Have you ever seen such animals in real life?

Was it hard to recombine the two animals?

Humans can generate and imagine new ideas because they learn a compositional and flexible hierarchy of abstract concepts.

Compositionality and concept learning are key for better Visual Content Generation.

Left: Top row shows a combination of penguin and snail. Bottom row shows a combination of jellyfish and butterfly.

Right: Lion-man of the Hohlenstein-Stadel cave in Germany (≈ 35,000 years old). The ivory figurine with the head of a lion and the body of a human is one of the earliest examples of art and the human cognitive ability to imagine things that do not really exist. (Image: Dagmar Hollmann / Wikimedia Commons / CC BY-SA 4.0)
› If you think this is interesting or you want to participate in research or you already have a cool idea!
› Come to me 😊

› creative AI projects
› bachelor thesis etc.

› we also go to Hackathons 😊

› ai.hdm-stuttgart.de
› theodoridis@hdm-stuttgart.de
› AI-Lab (R143)

› Thank you!