Lesson 2

- Interesting Links
- Successful applications

Books and Tutorials

- Deep Learning, MIT Press. By Y.Bengio, I.Goodfellow and A.Courville.
- Dive into Deep Learning. An interactive book with code, theory and discussions
- Tensorflow Tutorials By Tensorflow org.
- Keras Blog. By F.Chollet.
- Deep Learning Tutorial. LISA lab. University of Montreal.
- Towards Data Science. A Medium publication sharing concepts, ideas, and codes.
- ... so many others

Code and Datasets

The State of the Art site! (papers with code)

- labeled natural images: ImageNet (@Stanford Vision Lab)

  $\approx$ 15M high res color images covering 22K object classes
  ground truth for discrimination, segmentation, borders
- faces
  - CelebA (many facial attributes: hear color, beard, mustaches, age, glasses, ...)
  - Labeled Faces in the Wild (detection/recognition)

Some Dataset repositories

Tensor flow dataset   Kaggle Datasets
Amazon Datasets   Biomedical challenges . . .

Computational facilities

Training may be expensive.

Some example:

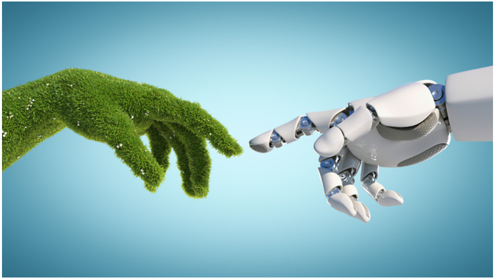
- the hyper-realistic Generative Adversarial Network for face generation by Tesla takes 4 days 8 Tesla V100 GPUs
- training of BERT, a well known generative model for NLP, takes about 96 hours on 64 TPU2 chips.

Major companies offer free computational resources on their clouds:

- Colab, by Google.
- Kaggle
- Amazon Web Services (AWS)
- ...

Green AI

The growing consumption of computational resources is raising social concerns. People is aiming to a more Green AI



Emphasis on efficiency as well as performance.
See this article for a discussion of evaluation metrics.

Next argument

Examples of
successful applications

Examples of successful applications

- **Image Processing**
  - Image Classification and Detection
  - Image Segmentation, Scene understanding
  - Style transfer
  - Deep dreams and Inceptionism
- **Natural Language Processing**
  - Speech Recognition
  - Text processing (translation, summarization, generation, . . .)
- **Generative modeling** (GANs, VAEs, Cycle Gans)
- **Deep Reinforcement Learning**
  - Robot navigation and autonomous driving
  - Model-free learning

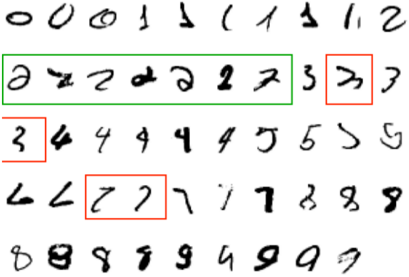
Next argument

Image Processing

MNIST

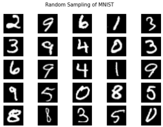
Modified National Institute of Standards and Technology database

- grayscale images of handwritten digits, 28×28 pixels each
- 60,000 training images and 10,000 testing images



MNIST

A comparison of different techniques



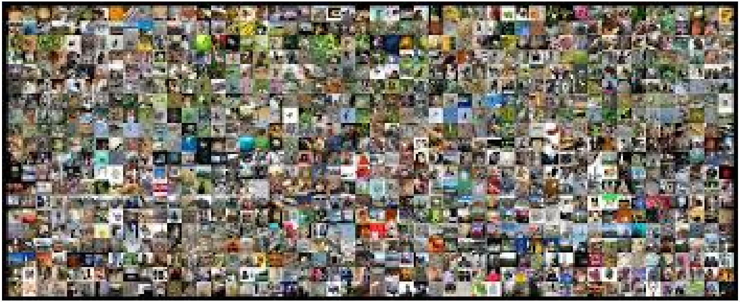
Classifier	**Error rate**
Linear classifier	7.6
K-Nearest Neighbors	0.52
SVM	0.56
Shallow neural network	1.6
Deep neural network	0.35
Convolutional neural network	0.21

See LeCun's page the mnist database for more data.

ImageNet

ImageNet (@Stanford Vision Lab)

- high resolution color images covering 22K object classes
- over 15 million labeled images from the web



ImageNet competition

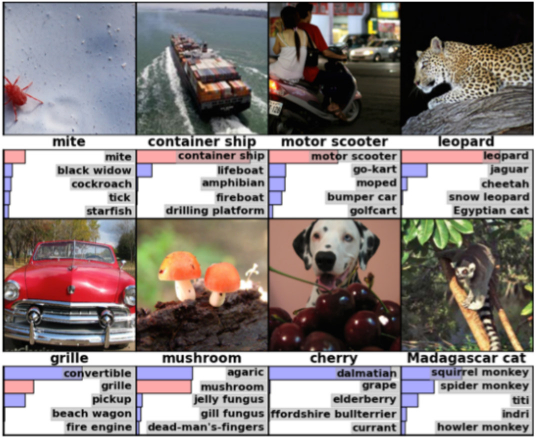
Annual competition of image classification: 2010-2017.

- 1.2 Million images, covering 1K different categories
- make five guesses about image label, ordered by confidence

EntleBucher

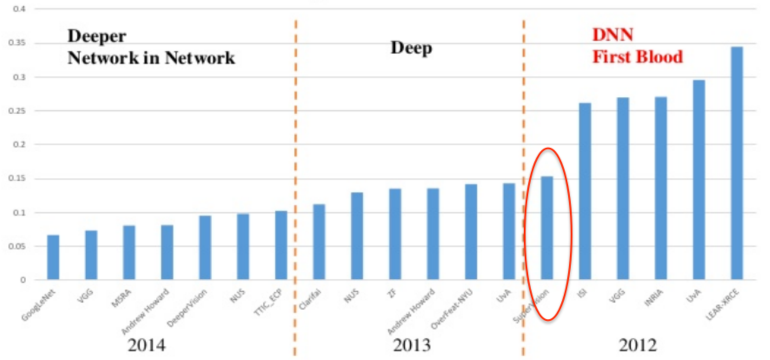
Appenzeller

ImageNet samples



ImageNet results

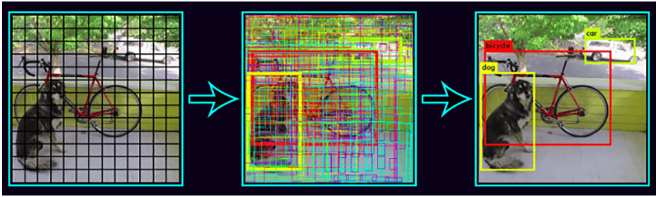
ImageNet Classification error throughout years and groups



Li Fei-Fei: ImageNet Large Scale Visual Recognition Challenge, 2014 http://image-net.org/

Image Detection

YOLO: Real-Time Object Detection

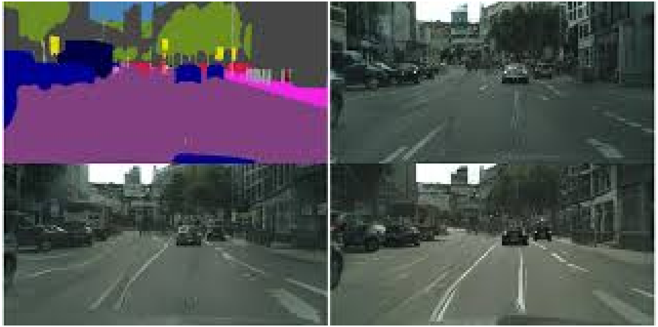


You only look once (YOLO) is a state-of-the-art, real-time object detection system. On a Pascal Titan X it processes images at 30 FPS and has a mAP of 57.9% on COCO test-dev.

First release in 2016, now at version 7.

Image Segmentation - Scene understanding

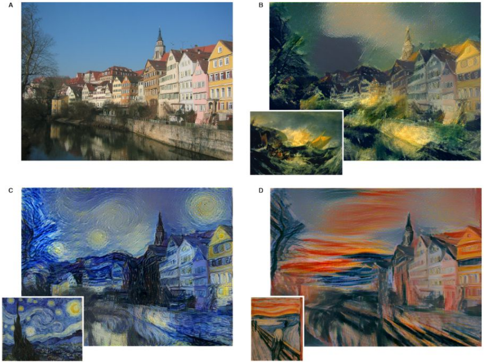
Video-to-Video Synthesis



Mimicking style

A neural algorithm of artistic style

L.A. Gatys, A.S. Ecker, M. Bethge



Change the style of an image, preserving the content.

Deep dreams

Source: Google Inceptionism

Visit Deep dreams generator

Many videos on youtube (e.g. this)

Next argument

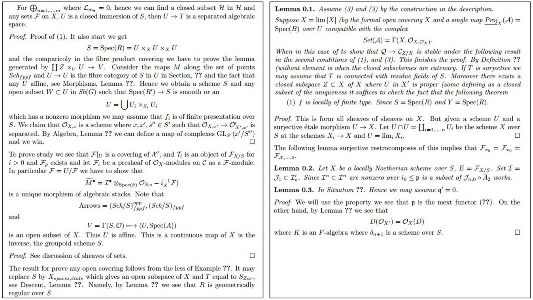
Natural Language Processing

Language Modeling

Predict the next character in a document (self-supervised)

First attempts with RNN (LSTM).

See Andrej Karpathy's blog The Unreasonable Effectiveness of Recurrent Neural Networks (old but still inspiring)



For $\bigoplus_{n=1,\dots,m}$ where $\mathcal{L}_{m_\bullet} = 0$, hence we can find a closed subset $\mathcal{H}$ in $\mathcal{H}$ and any sets $\mathcal{F}$ on X, U is a closed immersion of S, then $U \to T$ is a separated algebraic space.

Proof. Proof of (1). It also start we get

$$S = \mathrm{Spec}(R) = U \times_X U \times_X U$$

and the comparicoly in the fibre product covering we have to prove the lemma generated by $\coprod Z \times_U U \to V$. Consider the maps M along the set of points Sch_{fppf} and $U \to U$ is the fibre category of S in U in Section, **??** and the fact that any U affine, see Morphisms, Lemma **??**. Hence we obtain a scheme S and any open subset $W \subset U$ in $Sh(G)$ such that $\mathrm{Spec}(R') \to S$ is smooth or an

$$U = \bigcup U_i \times_{S_i} U_i$$

which has a nonzero morphism we may assume that f_i is of finite presentation over S. We claim that $\mathcal{O}_{X,x}$ is a scheme where $x, x', s'' \in S'$ such that $\mathcal{O}_{X,x'} \to \mathcal{O}'_{X',x'}$ is separated. By Algebra, Lemma **??** we can define a map of complexes $\mathrm{GL}_{S'}(x'/S'')$ and we win. □

To prove study we see that $\mathcal{F}|_U$ is a covering of $\mathcal{X}'$, and $\mathcal{T}_i$ is an object of $\mathcal{F}_{X/S}$ for $i > 0$ and $\mathcal{F}_p$ exists and let $\mathcal{F}_i$ be a presheaf of $\mathcal{O}_X$-modules on $\mathcal{C}$ as a $\mathcal{F}$-module. In particular $\mathcal{F} = U/\mathcal{F}$ we have to show that

$$\widetilde{M}^\bullet = \mathcal{I}^\bullet \otimes_{\mathrm{Spec}(k)} \mathcal{O}_{S,s} - i_X^{-1}\mathcal{F})$$

is a unique morphism of algebraic stacks. Note that

$$\mathrm{Arrows} = (Sch/S)^{opp}_{fppf}, (Sch/S)_{fppf}$$

and

$$V = \Gamma(S, \mathcal{O}) \longmapsto (U, \mathrm{Spec}(A))$$

is an open subset of X. Thus U is affine. This is a continuous map of X is the inverse, the groupoid scheme S.

Proof. See discussion of sheaves of sets. □

The result for prove any open covering follows from the less of Example **??**. It may replace S by $X_{spaces,étale}$ which gives an open subspace of X and T equal to S_{Zar}, see Descent, Lemma **??**. Namely, by Lemma **??** we see that R is geometrically regular over S.

Lemma 0.1. *Assume (3) and (3) by the construction in the description.*

Suppose $X = \lim |X|$ *(by the formal open covering* X *and a single map* $\underline{Proj}_X(\mathcal{A}) = \mathrm{Spec}(B)$ *over* U *compatible with the complex*

$$Set(\mathcal{A}) = \Gamma(X, \mathcal{O}_{X,\mathcal{O}_X}).$$

When in this case of to show that $\mathcal{Q} \to \mathcal{C}_{Z/X}$ *is stable under the following result in the second conditions of (1), and (3). This finishes the proof. By Definition* **??** *(without element is when the closed subschemes are catenary. If* T *is surjective we may assume that* T *is connected with residue fields of* S*. Moreover there exists a closed subspace* $Z \subset X$ *of* X *where* U *in* X' *is proper (some defining as a closed subset of the uniqueness it suffices to check the fact that the following theorem*

(1) f *is locally of finite type. Since* $S = \mathrm{Spec}(R)$ *and* $Y = \mathrm{Spec}(R)$.

Proof. This is form all sheaves of sheaves on X. But given a scheme U and a surjective étale morphism $U \to X$. Let $U \cap U = \coprod_{i=1,\dots,n} U_i$ be the scheme X over S at the schemes $X_i \to X$ and $U = \lim_i X_i$. □

The following lemma surjective restrocomposes of this implies that $\mathcal{F}_{x_0} = \mathcal{F}_{x_0} = \mathcal{F}_{\mathcal{X},\dots,0}$.

Lemma 0.2. *Let* X *be a locally Noetherian scheme over* S, $E = \mathcal{F}_{X/S}$*. Set* $\mathcal{I} = \mathcal{J}_1 \subset \mathcal{I}'_n$*. Since* $\mathcal{I}^n \subset \mathcal{I}^n$ *are nonzero over* $i_0 \leq \mathfrak{p}$ *is a subset of* $\mathcal{J}_{n,0} \circ \overline{A}_2$ *works.*

Lemma 0.3. *In Situation* **??**. *Hence we may assume* $\mathfrak{q}' = 0$.

Proof. We will use the property we see that $\mathfrak{p}$ is the mext functor (**??**). On the other hand, by Lemma **??** we see that

$$D(\mathcal{O}_{X'}) = \mathcal{O}_X(D)$$

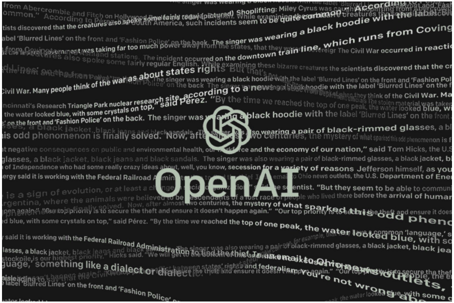
where K is an F-algebra where δ_{n+1} is a scheme over S. □

Examples of fake algebraic documents generated by a RNN.

Transformers

RNNs have been replaced by **Transformers**, based on a mechanism called **attention**

See Bert, Albert, GPT, . . .



GPT2 is a huge model, with 1.2 billion parameters, trained over 8 million web pages.

Other applications in NLP

- Sentiment analysis. Classify a document according to its “polarity”
- Machine Translation
- Text summarization/completion
- Text Generation: a truly generative task
- Speech recognition
- Dialog Systems - Chatboxes

Next argument

Generative Modeling

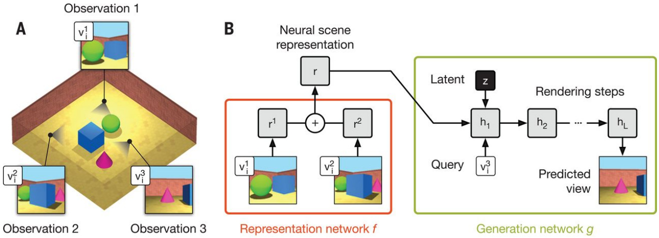
Generative Modeling

Goal: Generate new samples similar to training data.

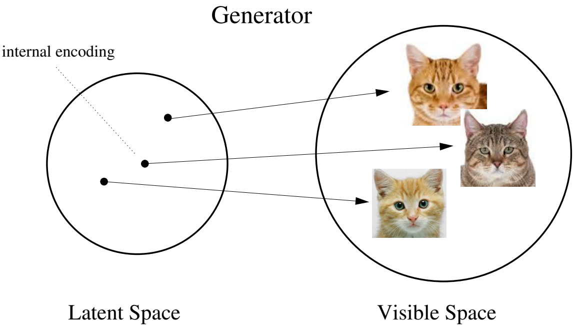
Face generation video by Nvidia

Scene representation and rendering

Neural scene representation and rendering (VAE)
Work published on Science (June 2018)

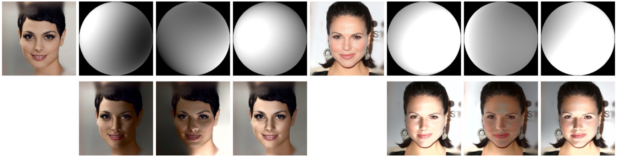


Latent space

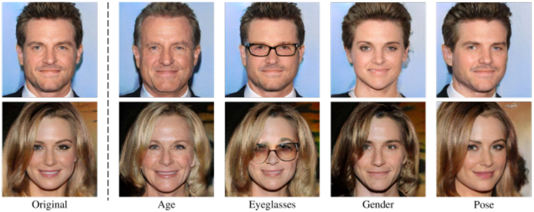


Suggested reading:
Comparing the latent space of generative models

Conditional generation

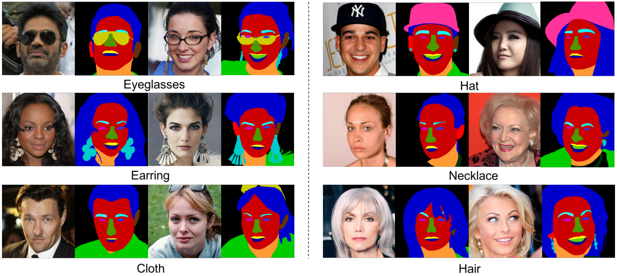


Deep Single Image Portrait Relighting



Interpreting the Latent Space of GANs for Semantic Face Editing

Conditional generation



Eyeglasses

Hat

Earring

Necklace

Cloth

Hair

MaskGAN: Towards Diverse and Interactive Facial Image Manipulation

Dall·E - OpenAI

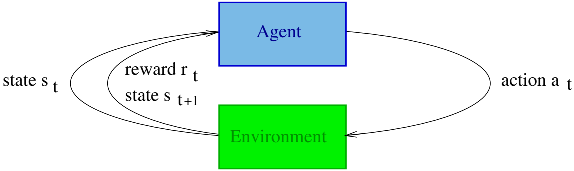
Dall·E is a new AI system that can create realistic images and art from a description in natural language.

Next argument

Reinforcement Learning

Reinforcement Learning

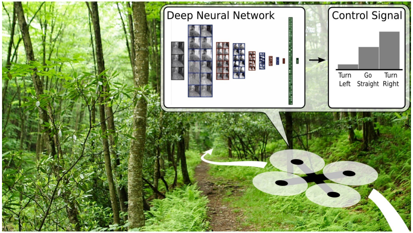
Problems involving an agent interacting with an environment, which provides numeric rewards



Goal: learn how to take actions in order to maximize the future cumulative reward.

Robot navigation

Quadcopter Navigation in the Forest using Deep Neural Networks



Robotics and Perception Group, University of Zurich, Switzerland & Institute for Artificial Intelligence (IDSIA), Lugano Switzerland

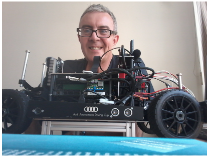
Based on Imitation Learning

Autonomous driving

Develop intelligent, fully automatic driving functions for vehicles.

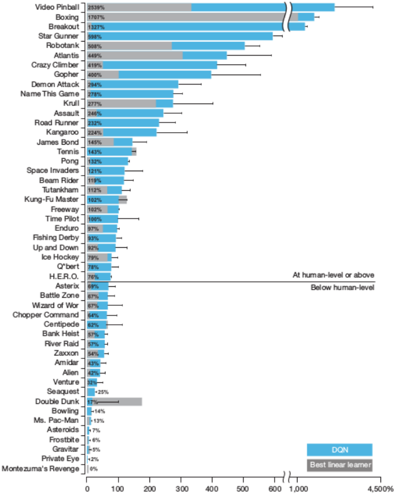
Merging of signals collected by different sensors (camera, lidar, sonar, dots). Needs to accurately evaluate distances and speeds.

Turn observations into actions.



Several competitions around. We took part to the 2018 Audi Autonomous Driving Cup

Game Simulation



Google DeepMind's system playing Atari games (2013)

The same network architecture was applied to all games

End-to-end training starting from screen frames

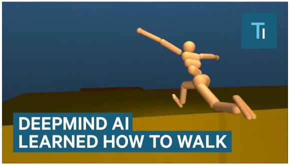
Works well for reactive games; problems with planning...

but see An investigation of Model-Free planning (ICML 2019)

Open AI-gym

OpenAI Gym is a Gym is a toolkit for developing and comparing reinforcement learning algorithms (DQN, A3C, A2C, Acer, PPO, . . .).

It offers many learning scenarios, from walking to playing games like Pong or Pinball, as well as other classical physical “equilibrium” problems.

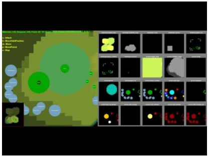


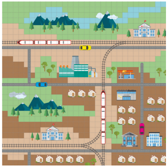
Multi agent DRL

It requires interaction and cooperation of multiple agents.

Examples:

StarCraft II: a RL environment based on the game StarCraft II. The environment consists of three sub-components: a Linux StarCraft II binary, the StarCraft II API providing programmatic control over the game, and a python wrapper over the API called PyC2.





Flatland: a train rescheduling problem on a complex grid world environment.

Flatland is organized every year by AIcrowd in collaboration with the Swiss Federal Railways, SBB