Master seminar: Solving localization problem in first person computer games with deep learning

Yauheni Selivonchyk[1],
Prof. Dr. Christian Bauckhage[2],
Prof. Dr. Stefan Wrobel[2],
Mr. Sc. Rafet Sifa[2]

[1]Institute of Computer Science, University of Bonn

[2]Fraunhofer IAIS

April 25, 2017

Outline

Recent progress in AI

Some of the recent advances in machine learning:

- Image classification
- Machine translation

Important attributes of it:

- Low cost and high processing abilities of modern computer chips
- Advances in machine learning algorithms
- Access to large labeled datasets

Importance of unsupervised learning for AI

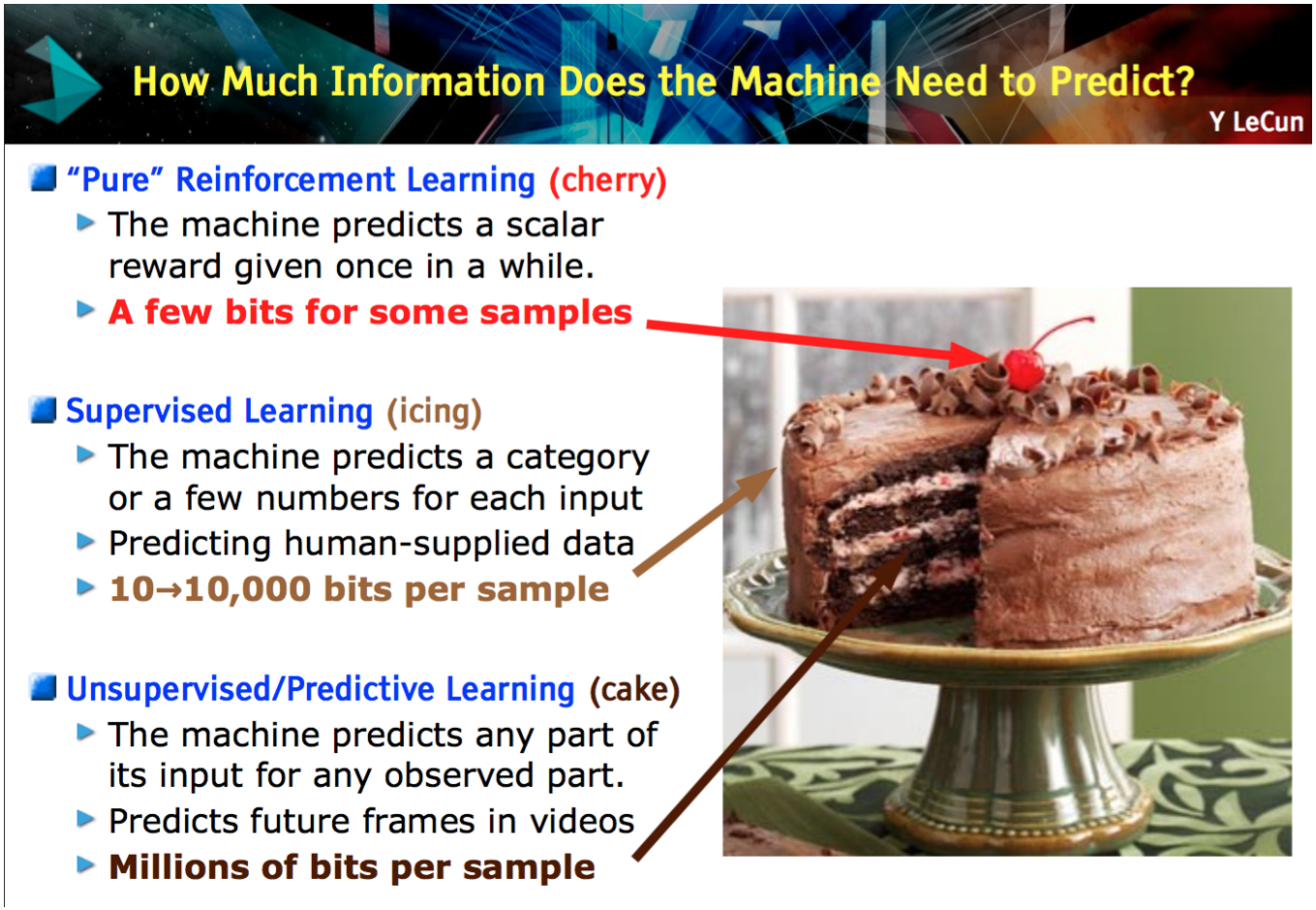


Figure: Slide from "Predictive learning" opening address given by Yann Lecun at NIPS2016.

Recent advances in Unsupervised learning

Some of the recent influential models:

- Word embedding (T. Mikolow, 2013)
- Variational autoencoders (D. Kingma, 2013)
- Generative Adversarial Networks (I. Goodfellow, 2014)

Our goal

- Sufficient amount of training data
- Computational feasibility of the problem

Localization

Localization as a task of extracting, tracking or predicting object's position in some environment from available sensory data.

Types of data:

- Visual data: images or video sequences
- Depth map
- Information about position/direction of the sensors
- etc.

Localization example: Tracking

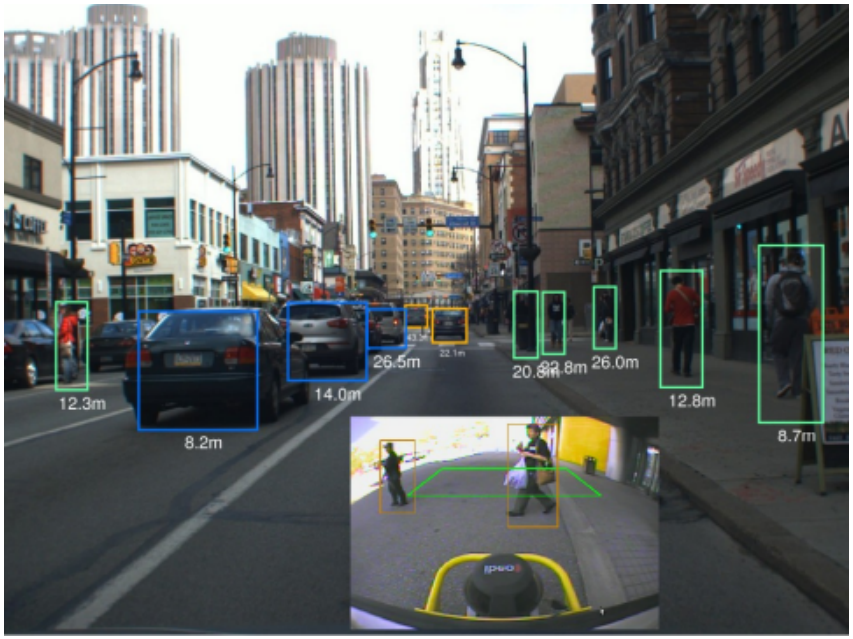


Figure: Pedestrian tracking visualization [1].

[1] H. Cho et. al. "Real-Time Pedestrian and Vehicle Detection for Automotive Active Safety Systems"

Localization example: SLAM

Figure: Example solution of SLAM problem on PC3 dataset (courtesy of University of Michigan).

Localization example: surgery

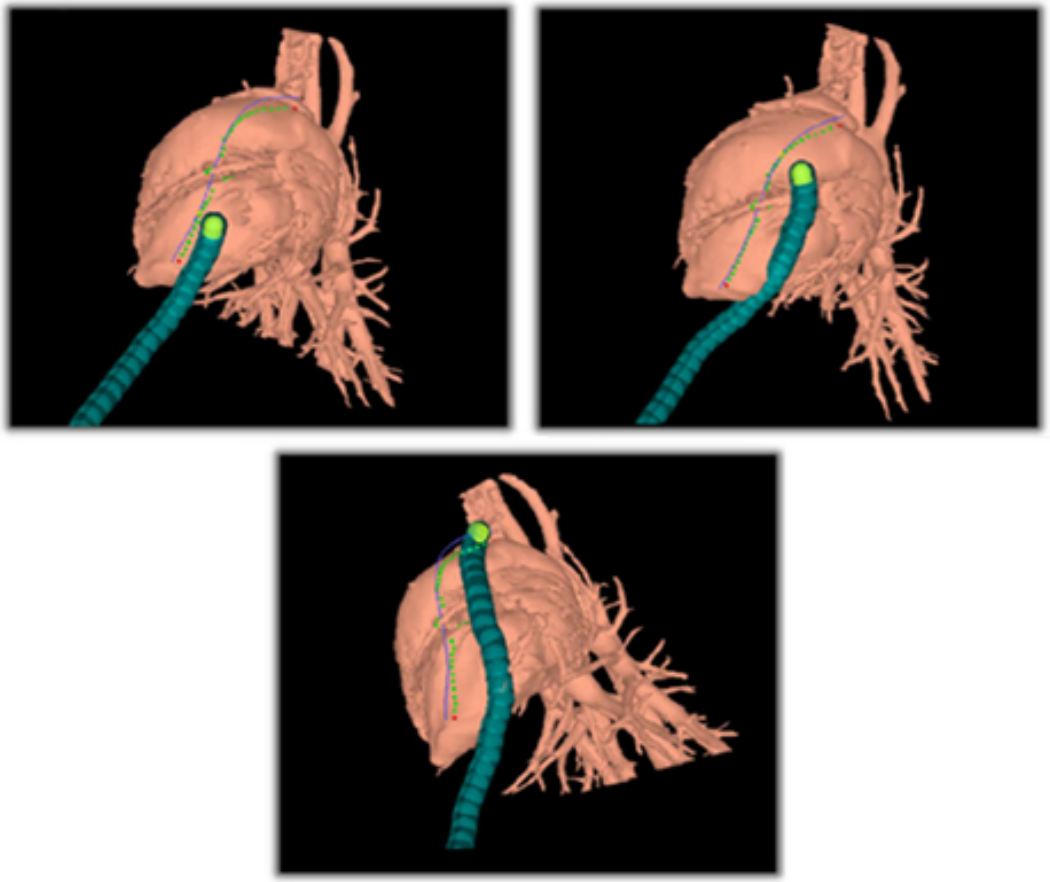


Figure: Mapping the position of a tool in minimally invasive surgery [http://biorobotics.ri.cmu.edu/research/medicalSLAM.html].

Motivation. Continuied

Goal of this work: reconstruction of the actors trajectory in first-person shooter (games) from visual data.

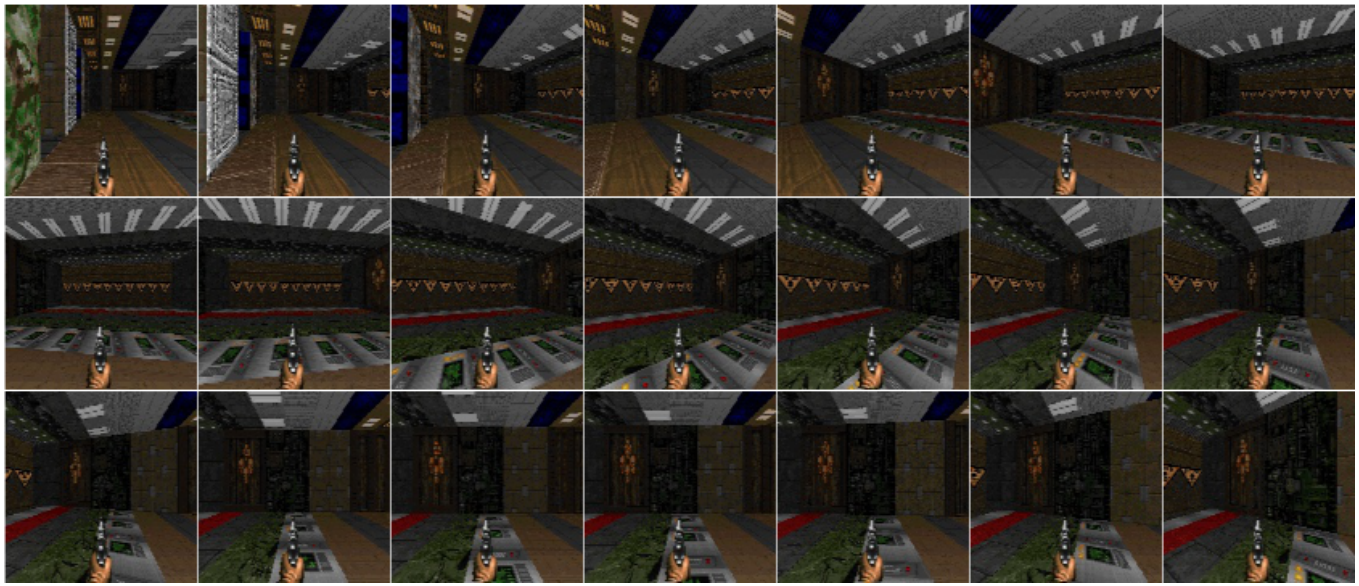
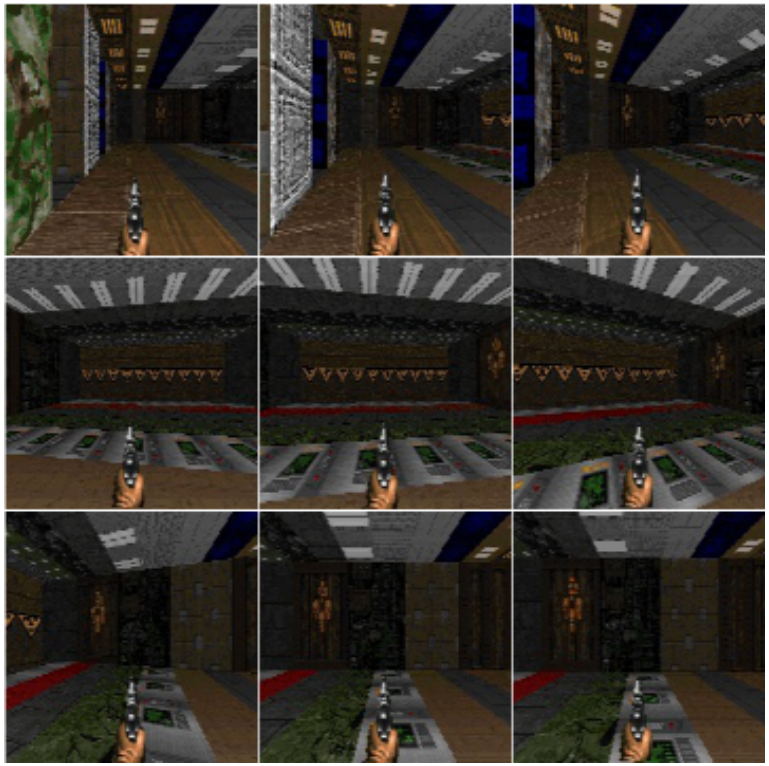


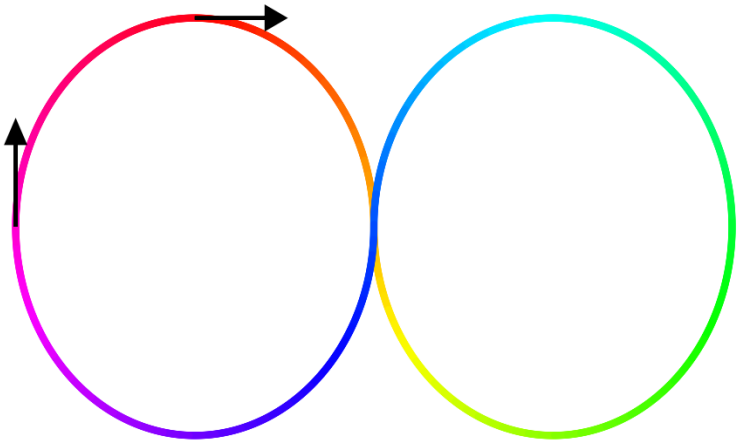
Figure: Example visual data.

Autoencoder model

Input video information:

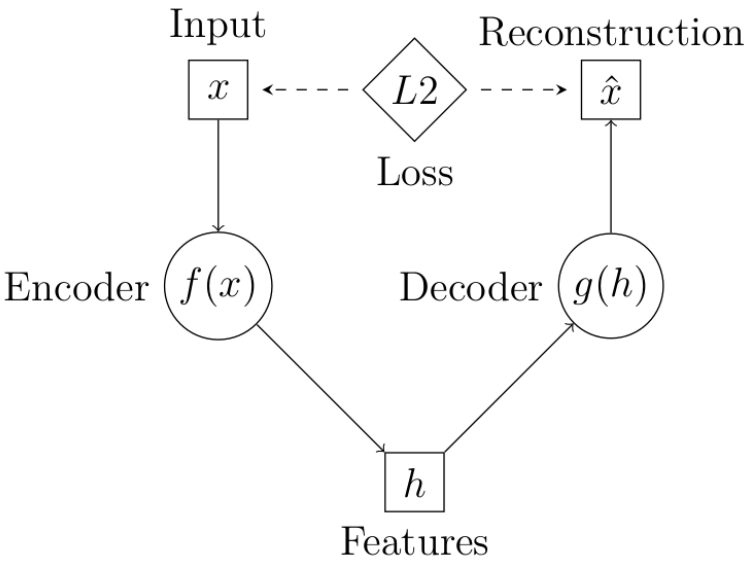


Corresponding players path:



Autoencoder model

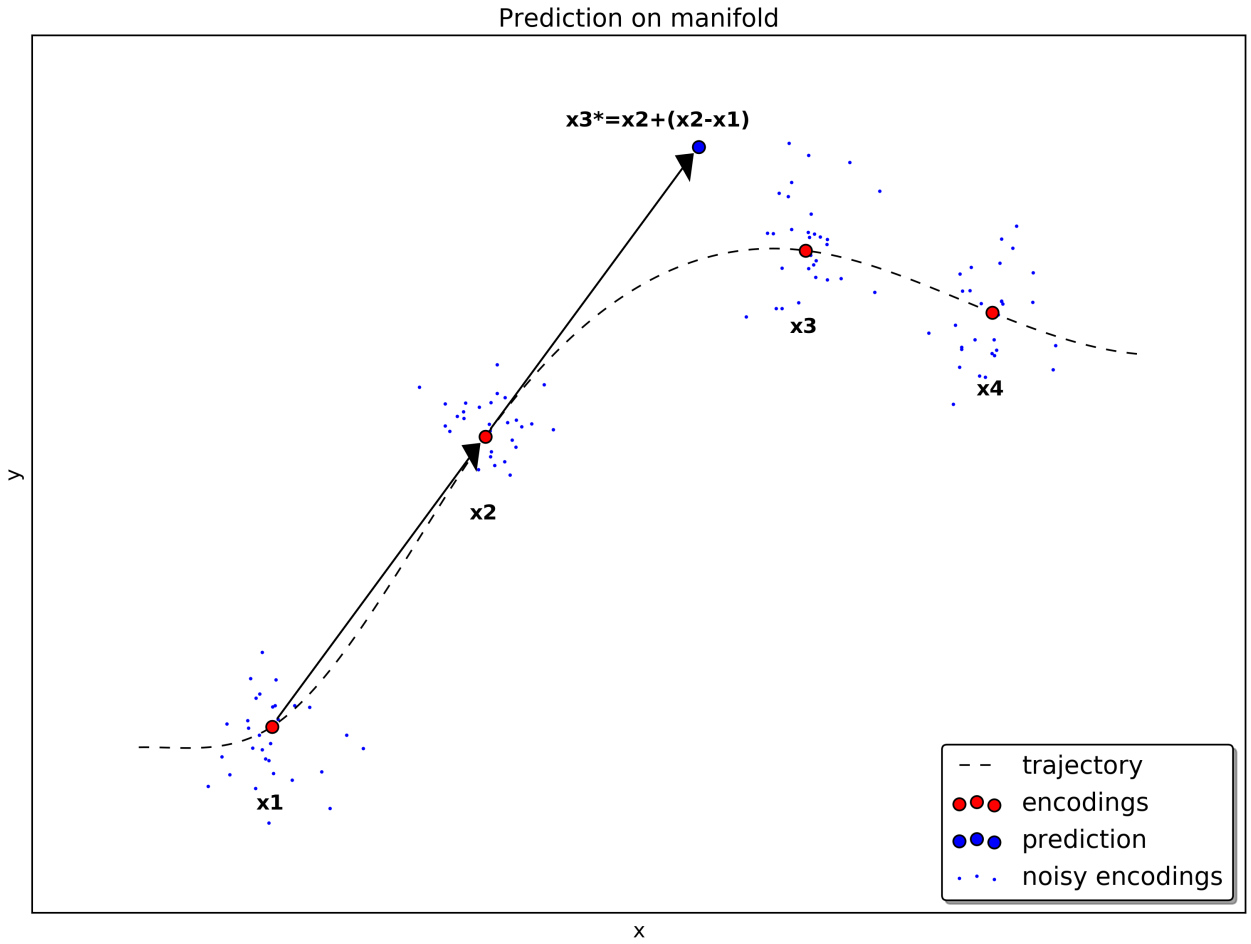
Autoencoders learn to project the input x into some embedding space $h \in H$ and simultaneously reconstruct the original information $\hat{x}$.



- $f : \mathbb{R}^N \rightarrow \mathbb{R}^M$
- $g : \mathbb{R}^M \rightarrow \mathbb{R}^N$

Predictive regularization

Try to estimate positional encoding of the next frame using last two frames of the video.



Metrics

Data collection

Data collection

Results

Results

Definition

Summary

- The **first main message** of your talk in one or two lines.

- Outlook
  - Something you haven't solved.
  - Something else you haven't solved.