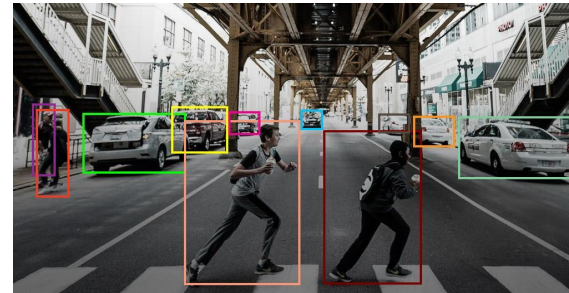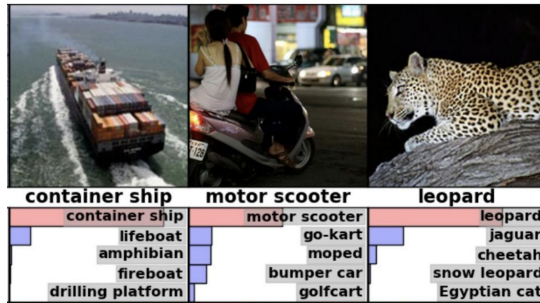# Seminar Vision Systems MA-INF 4208

12.02.2021
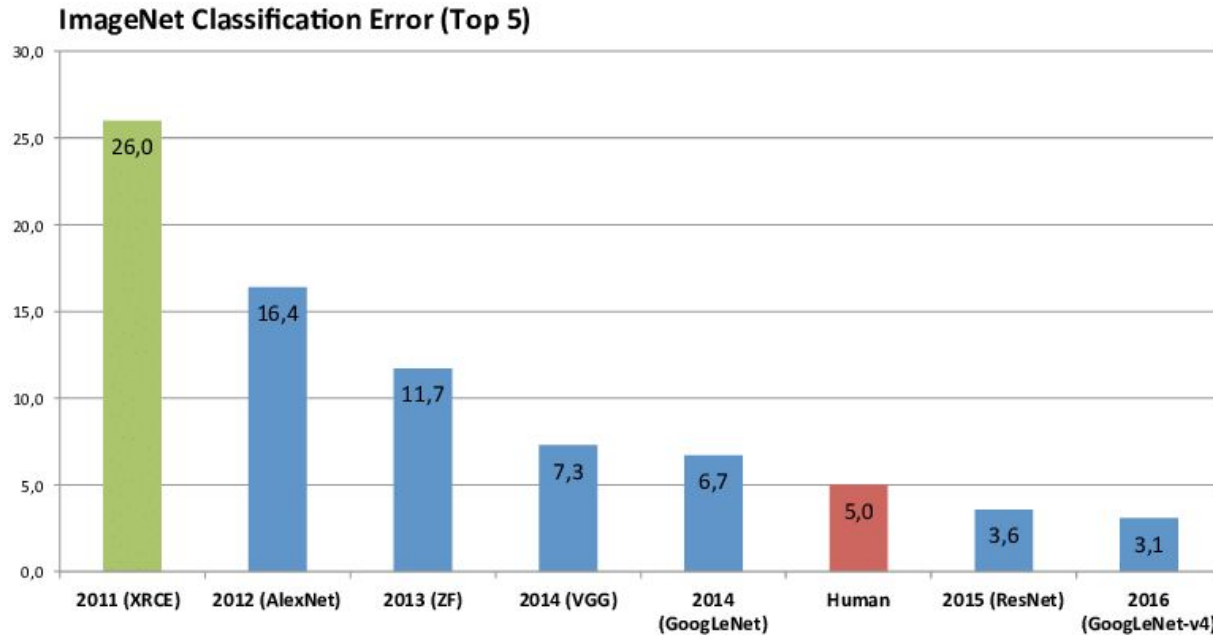
PROF. SVEN BEHNKE, <u>ANGEL VILLAR-CORRALES</u>, HAFEZ FARAZI

Contact: villar@ais.uni-bonn.de

# The Age of Deep Learning

# The Age of Deep Learning



ImageNet Classification Error (Top 5)

# The Age of Deep Learning

# In this seminar...

- Acquire/improve ability to:
  - deal with scientific publications (e.g., papers)
  - write a scientific report
  - present a scientific topic to an audience
  - engage technical topics

  ➡ Important skills for Master Thesis!

# In this seminar

- Discuss trending topics in deep learning and computer vision

- We will cover the following topics

  ○ Human Pose Estimation

  ○ Self-Supervised Learning

  ○ Understanding of Neural Networks

  ○ Domain Adaptation


  Paper List: https://www.ais.uni-bonn.de/WS2021/SeminarVision/PaperList.pdf
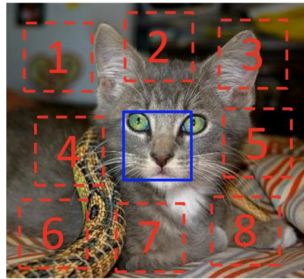
# Human Pose Estimation

- Predicting a pose skeleton for every person in an image

- Estimate location of body joints for each person

- Challenging task due to occlusion and variability

- Several applications
  - Action recognition
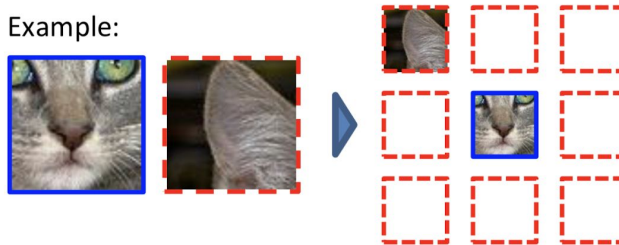  - Robot perception
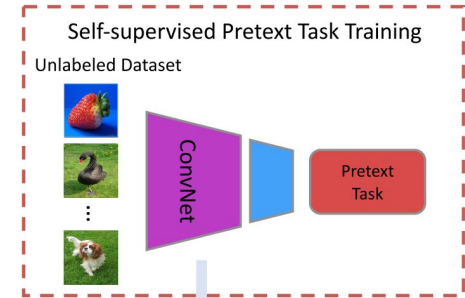  - Animation

# Self-Supervised Learning

- Subcategory of unsupervised learning

- Use pretext task to train in a supervised fashion

- Hot-topic in deep learning community
  - Comparable to supervised pretraining
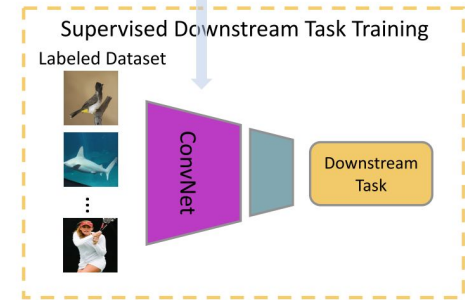  - No need for manual annotations



Example:

https://arxiv.org/pdf/1505.05192.pdf



Self-supervised Pretext Task Training
Unlabeled Dataset
ConvNet
Pretext Task

Knowledge Transfer

Supervised Downstream Task Training
Labeled Dataset
ConvNet
Downstream Task
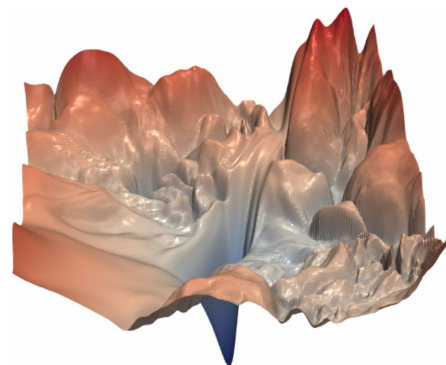
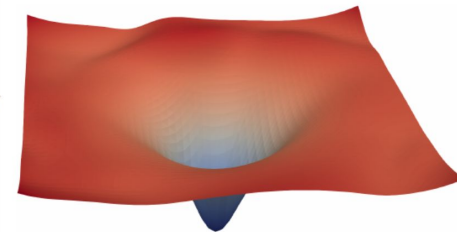https://vishal-keshav.github.io/publication/curriculum-learning

# Understanding CNNs

- Difficult to understand how the make predictions

- Easy to fool their decision making
  - Adversarial attacks

- Still a mathematical riddle
  - No convergence guarantees
  - Avoids bad local minima
  - Surprising generalization
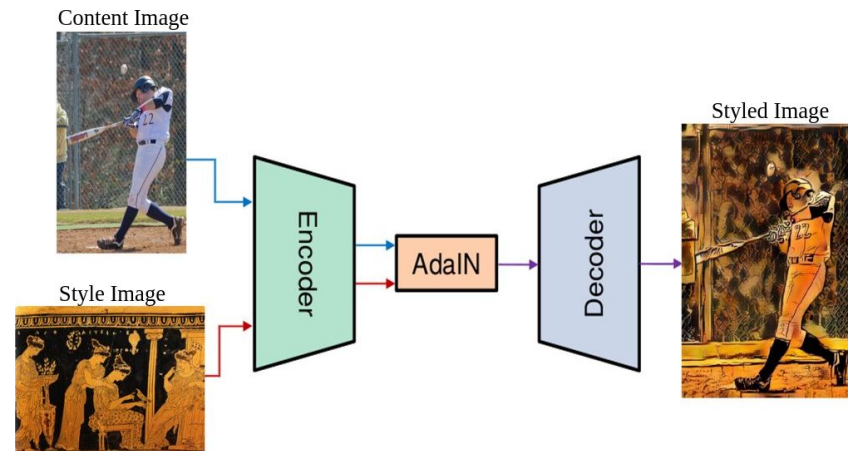


(a) without skip connections    (b) with skip connections

Loss surfaces

https://arxiv.org/pdf/1712.09913.pdf

# Domain Adaptation

- Models underperformed in unseen data domains

- Not possible to retrain
  - Lack of annotated data
  - Expensive & time consuming

- Bridge the gap between known and unseen domain
  - Style transfer
  - Feature alignment



Content Image

Style Image

Styled Image

Encoder

AdaIN

Decoder

https://arxiv.org/pdf/2012.05616.pdf

# Select your topic

- Use the Doodle to select a paper to review
  - Link on Seminar Website at 13:00
    - https://www.ais.uni-bonn.de/WS2021/4208_Sem_Vision_Systems
  - First come first serve

- Register in BASIS

- Send me an email at [villar@ais.uni-bonn.de](mailto:villar@ais.uni-bonn.de)
  - Your name
  - Your selected topic

# Deliverables

- Presentation (Thursday 25.03.21)
    - 30 min presentation
    - 15 min discussion

- Report (one week after presentation)
    - LaTex template
    - 8-10 pages
    - Brief but readable and informative
    - BibTex citations

→ Arrange a meeting with me ≈2 weeks before the presentation to check the preliminary materials for the presentation and report.

# Report

- Well structured:
    - Abstract
    - Introduction, methods, results, conclusion, ...
    - Tables and figures
    - Correct citations

- Your own scientific opinion:
    - What are the weak points of the paper?
    - What is missing?
    - Are comparisons fair and believable?
    - Possible future steps?

We don't want a copy of the paper!

# Grading

- 25%: Quality of the presentation slides

- 25%: Presentation skills and ability to answer questions

- 25%: Understanding of the concept

- 25%: Seminar report (8-10 pages, LaTEX)

# Questions?