

Advanced Topics in Continual / Organic Machine Learning

Interactive Systems Lab (ISL)
Institute for Anthropomatics and Robotics (IAR)

Winter 20-21 Topics

"I'm still learning"
Michelangelo



Incremental Learning with Adversarial Networks

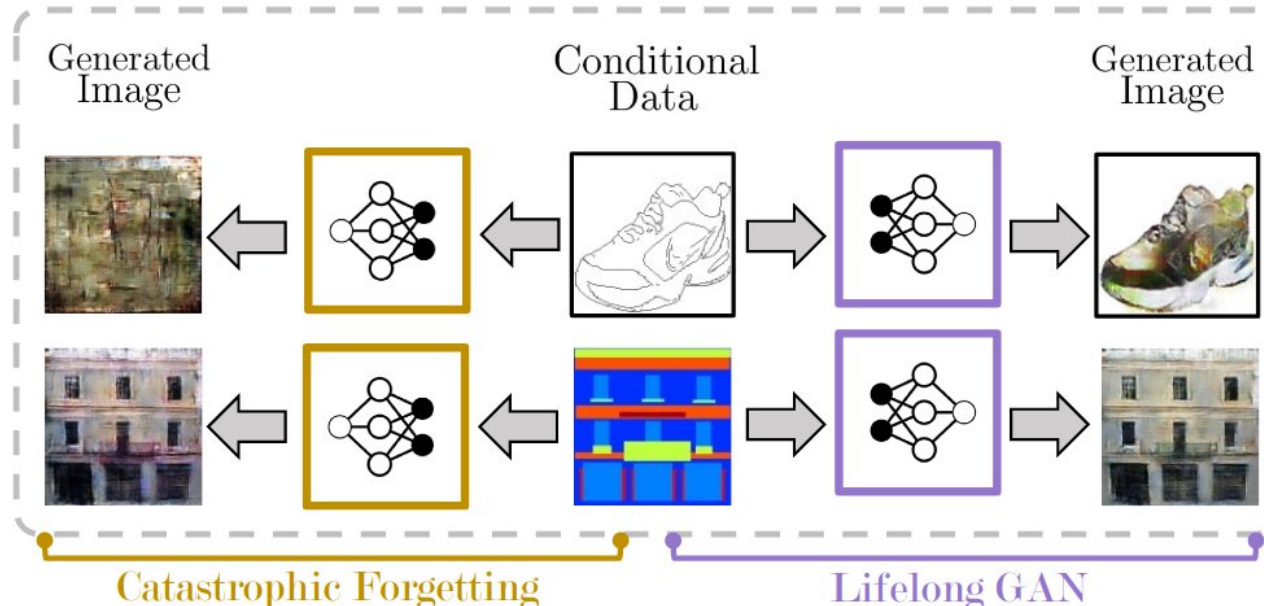
- GAN-based approach for incremental learning
- Proposed a new loss function
- Eliminate the unbalanced data distribution between old classes and new classes
- Generate historical data with GAN

Papers:

- Incremental Classifier Learning with Generative Adversarial Networks, Wu et al., 2018
- Incremental Learning Using Conditional Adversarial Networks, Xiang et al., 2019

Lifelong Learning with GANs

- GAN-based approach to avoid from catastrophic forgetting
- Task: Conditional image generation
- Knowledge distillation to transfer learned knowledge from old network to the new one

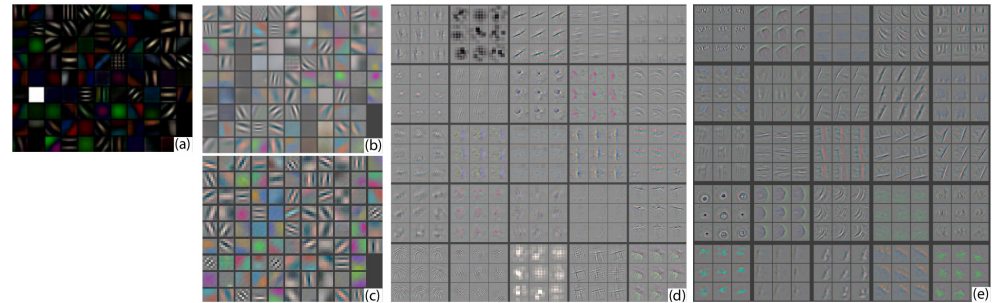


Papers:

- Lifelong GAN: Continual Learning for Conditional Image Generation, Zhai et al., 2019
- Piggyback GAN: Efficient Lifelong Learning for Image Conditioned Generation, Zhai et al., 2020

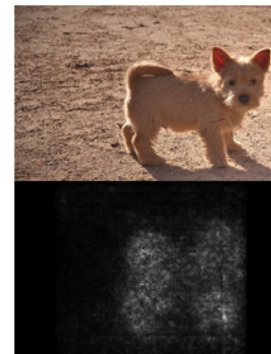
Understanding neural networks' internals

- End-to-end (E2E) models
 - Straight forward training
 - State-of-the-art performance
 - Blackbox → hard to analyse

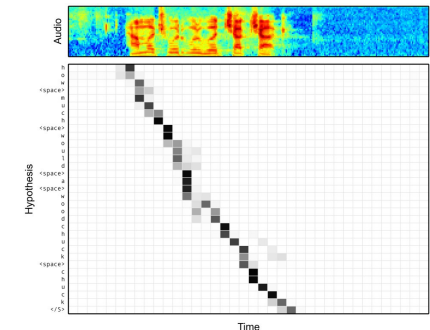


1: „Visualizing and Understanding Convolutional Networks”, Zeiler et Fergus, 2013

- Analysis methods for images
 - Filter visualisation (1)
 - Saliency map plots (2)
 - Display attention-alignments (3)



2: „Deep Inside Convolutional Networks Visualising Image Classification Models and Saliency Maps”, Simonyan et al., 2014



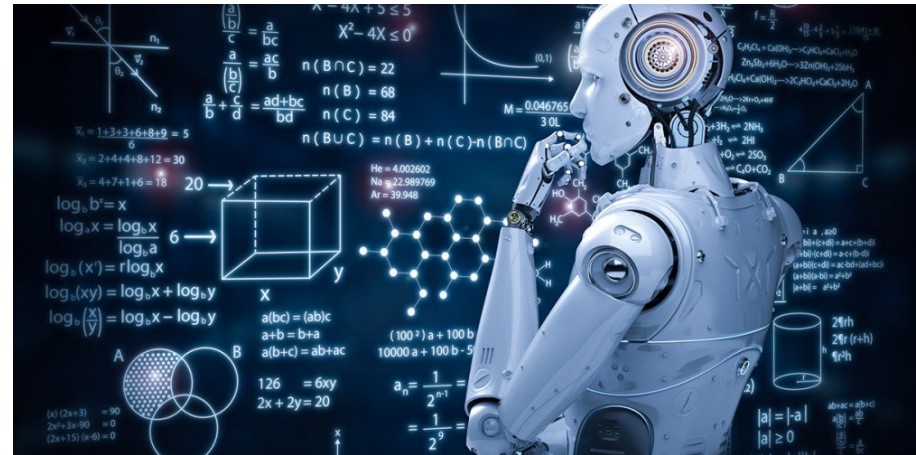
3: „Listen, attend and spell: A neural network”, Chan et al., 2015

- Method for acoustic signals (speech)?

„What does a network layer hear? Analyzing hidden representations of end-to-end ASR through speech synthesis“, Li et al., 2019

Self-learning with multiple hypotheses

- Supervised learning
→ Labelled data?
- Unsupervised learning:
Knowledge Distillation (KD)
→ Gigantic teacher model?
- Self-training
- Multi-label learning („Multitask learning“, Caruana, 1997)
- Application: Automatic speech recognition (ASR)



Source: <https://www.roboticsbusinessreview.com/wp-content/uploads/2018/07/SmarterRobots-iStock-1000x500.jpg>

*„Sequence-level self-learning with multi-task learning framework“,
Kumatani et al., 2020*

Self-Supervised Learning

- A machine learning system that uses supervised learning techniques (e.g. NNs) to learn from automatically labelled data
 - No labelled data needed
 - Explain methods/applications
- Deep Clustering for Unsupervised Learning of Visual Features (Caron et al. 2018)

Meta-Learning - Learning to learn

- An approach to learn how to learn
 - Meta-Learning
- Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, Finn et al. 2017

Never ending Learning

- a computer system that learns over time to read the web
- since January 2010, the computer system called NELL (Never-Ending Language Learner) is learning
- extracts facts from text found in hundreds of millions of web pages (e.g., `playsInstrument(George_Harrison, guitar)`)
- attempts to improve its reading competence

Papers:

- Never-ending learning, Mitchell et al., 2018
- Learning without forgetting, Li et al, 2017

Pretraining techniques

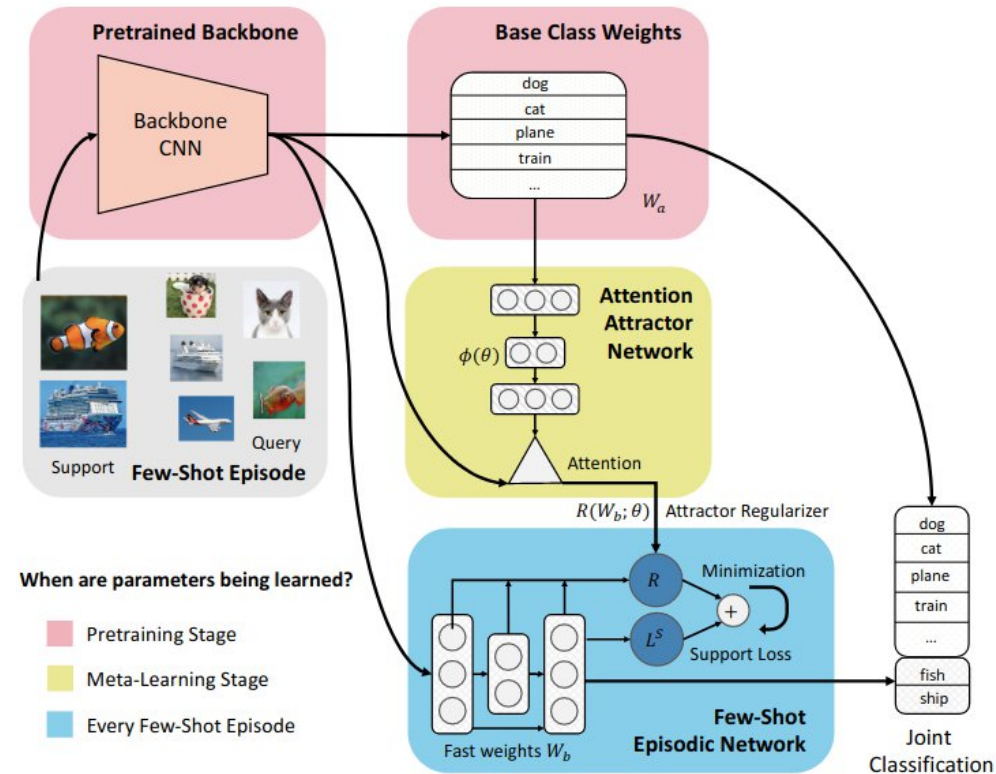
- to do organic learning, a base system is needed
- if not many data is available, a pretrained system can boost the performance
- it is trained on general data and can be finetuned to special data

Papers:

- Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer, Raffel al., 2019
- Pre-training via Paraphrasing, Lewis et al, 2020

Incremental few-shot learning

- Adaptation with few data belonging to novel classes
- Using pretrained backbone model
- Meta learning based incremental training episodes
- Avoid from catastrophic forgetting

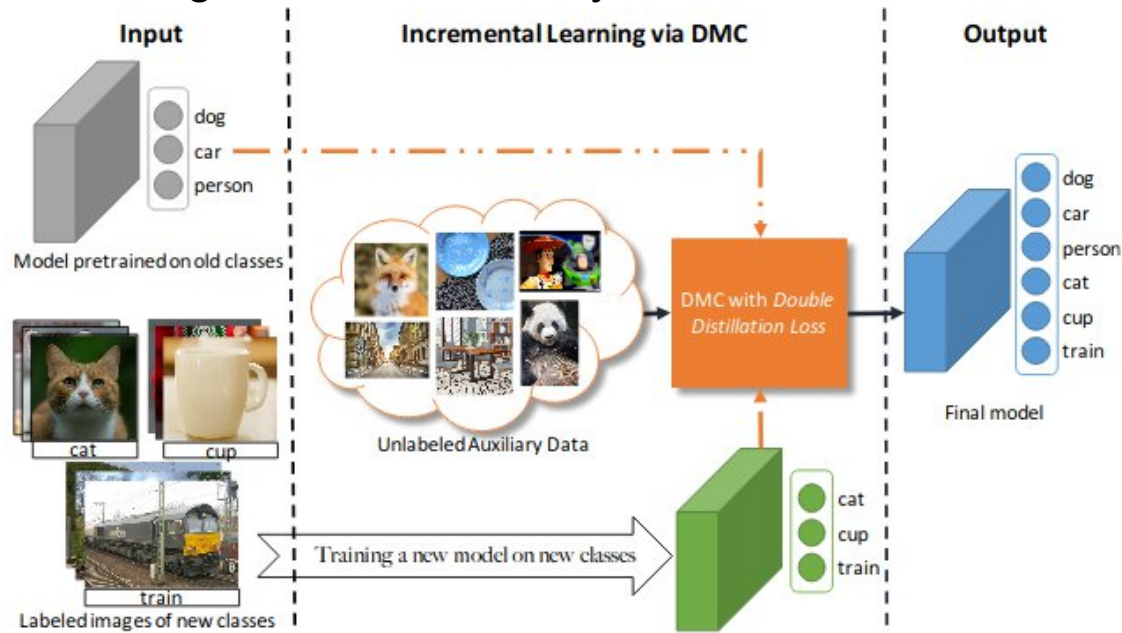


Papers:

- Incremental Few-Shot Learning with Attention Attractor Networks, Ren et al., NeurIPS 2019
- XtarNet: Learning to Extract Task-Adaptive Representation for Incremental Few-Shot Learning, Yoon et al., arXiv, 2020 (Accepted for ICML)

Class Incremental Learning

- Avoid from biased model towards old classes and new classes
- Novel objective function
- Solve the data imbalance problem
- Solve the increasing number of visually similar classes

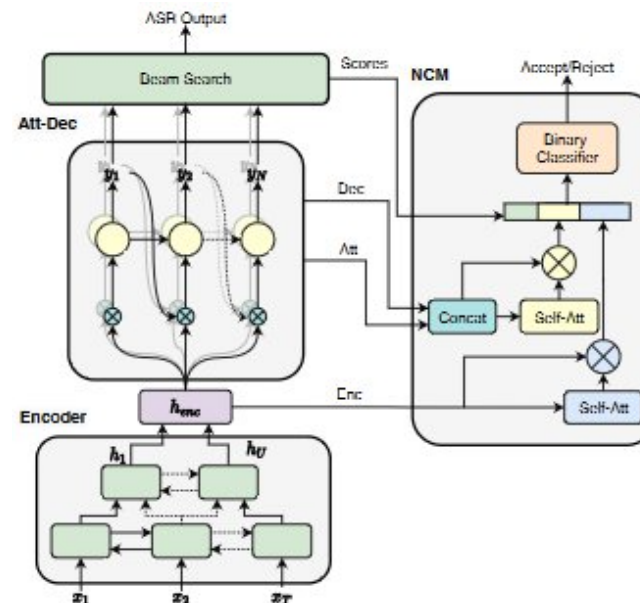


Papers:

- Class-incremental Learning via Deep Model Consolidation, Zhang et al., 2020
- Large Scale Incremental Learning, Wu et al., 2019

Neural Confidence Measure

- Define binary classifier
- Approach for solving poorly performing confidence estimations
- Evaluation: Encoder, Decoder, Attention features
- Significant computation saved

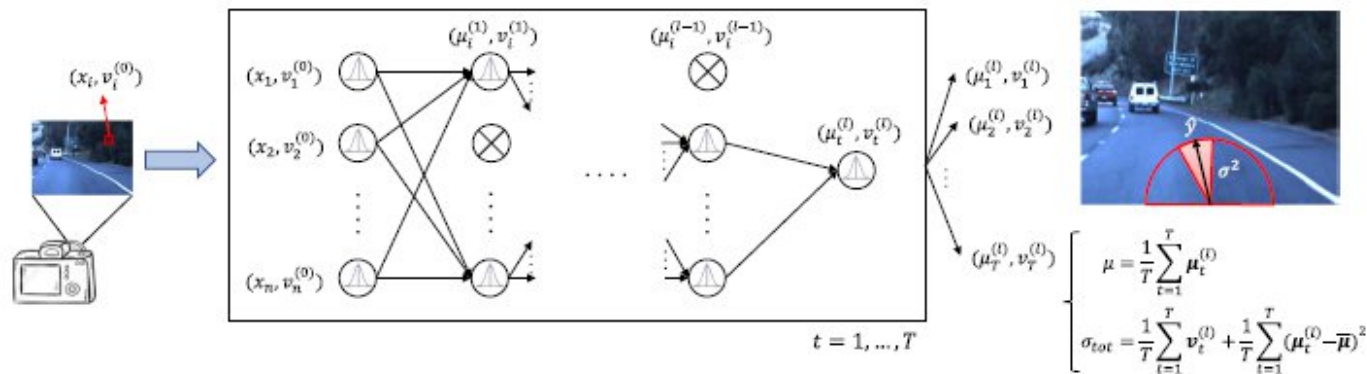


Papers:

- Kumar, Ankur, et al. "Utterance Confidence Measure for End-to-End Speech Recognition with Applications to Distributed Speech Recognition Scenarios", 2020
- Li, Qiuja, et al. "Confidence Estimation for Attention-based Sequence-to-sequence Models for Speech Recognition", 2020

General Framework Uncertainty Estimation

- Input sample: out of training or noise corrupted
- Approaches the problem of ignoring prior knowledge about the training data underestimation of uncertainty
- Novel framework based on Bayesian networks and Monte-Carlo sampling
- General framework: on top of already existing architecture



Papers:

- Loquercio, Antonio, et al. "A General Framework for Uncertainty Estimation in Deep Learning", 2020

Differentiable neural computers

- New model architecture
- Based on
 - Von Neumann architecture
 - Human working memory
- Memory based
- "Hybrid computing using a neural network with dynamic external memory", Graves et al., 2016

