Karlsruhe Institute of Technology

Advanced Topics in Continual / Organic Machine Learning

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

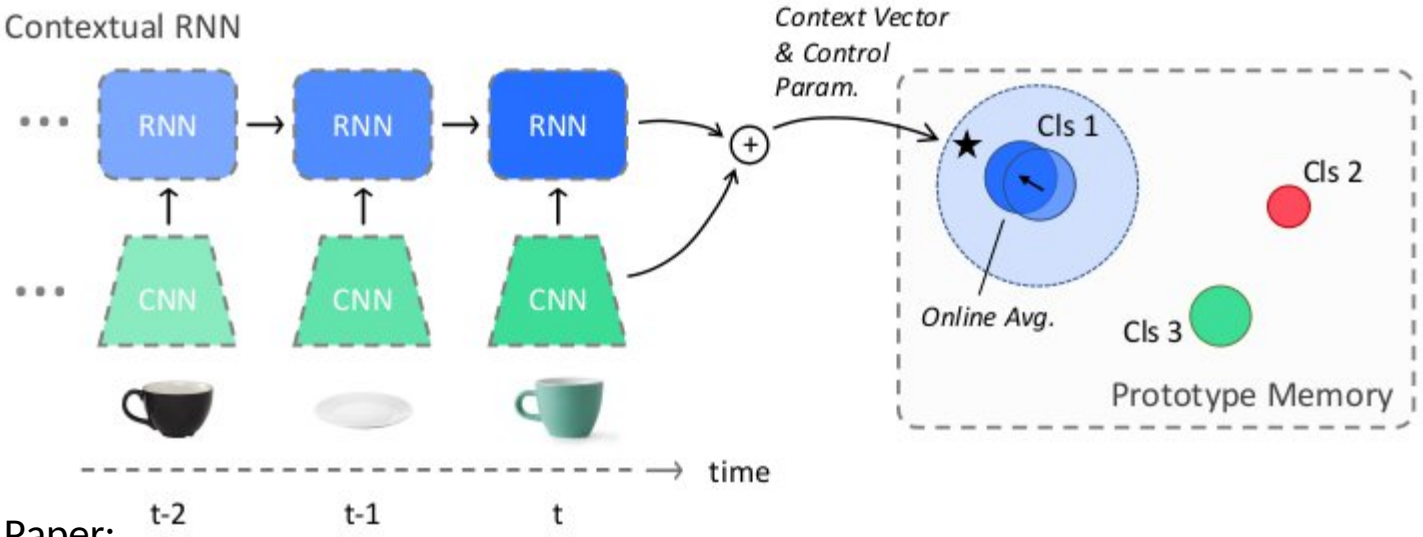
Course Organizer: F. Irem Eyiokur

Summer 23 Topics

"I'm still learning"
Michelangelo

Online Contextualized Few-shot Learning

- Few-shot incremental learning in online/continual setting.
- Classes are organized based in context
- To encode temporal contextual information → RNN
- Prototypical memory to remember prev. learned classes



Paper:

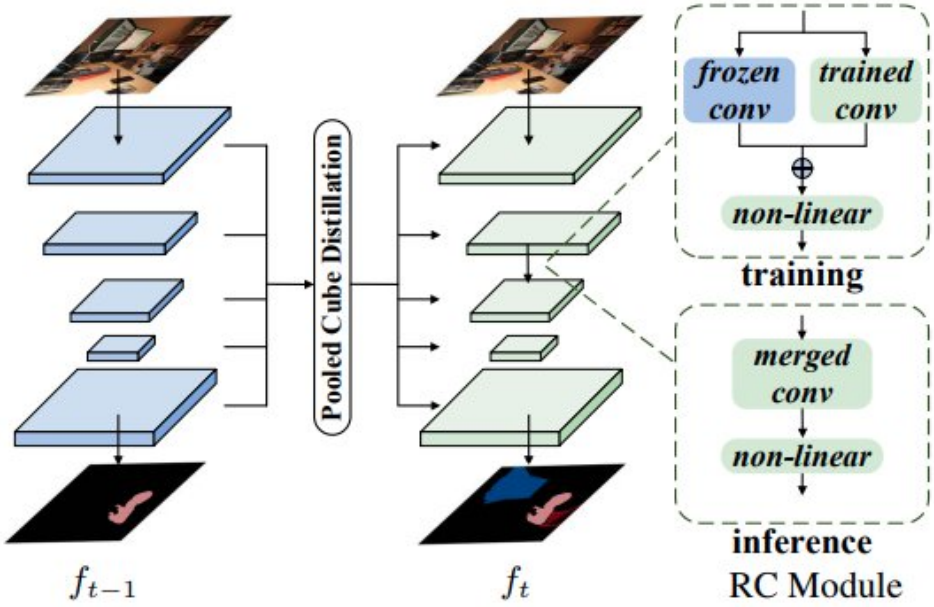
Ren, Mengye, et al. “Wandering within a world: Online contextualized few-shot learning”, ICLR 2021

Extension of above paper to unsupervised online learning:

Ren, Mengye, et al. “Online unsupervised learning of visual representations and categories”, arxiv preprint 2022

Representation Compensation Networks for Continual Semantic Segmentation

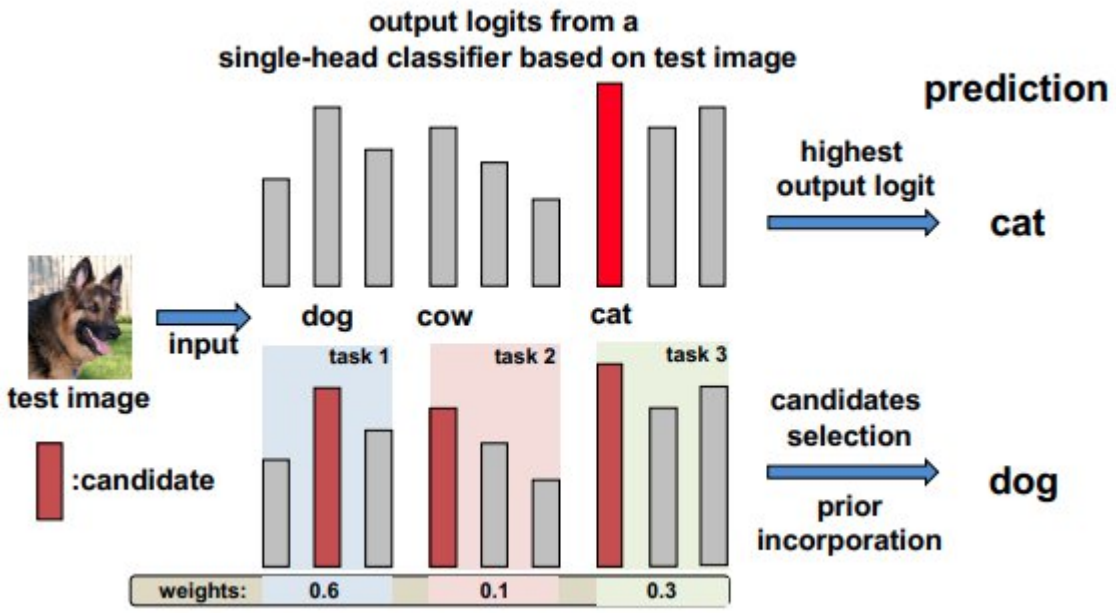
- Proposed a training framework for continual semantic segmentation while avoiding catastrophic forgetting
- Representation compensation (RC) module which include two branches for decouple frozen and trainable representations.
- Pooled Cube Distillation to enhance the plasticity and stability of the model
- During inference time, no increase in computation and memory cost



Paper:
Chang-Bin, Zhang, et al. “Representation Compensation Networks for Continual Semantic Segmentation”, CVPR 2022

Online Continual Learning Via Candidates Voting

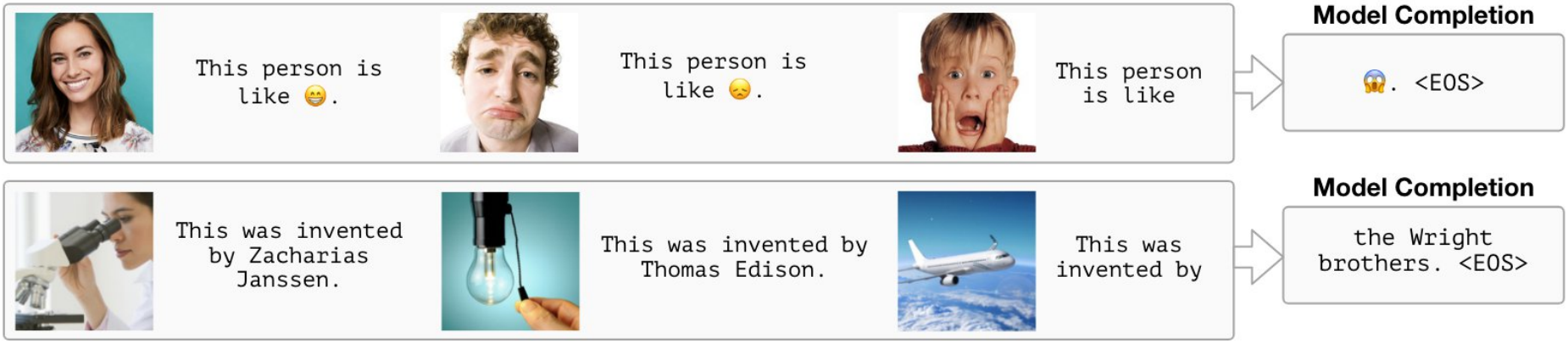
- Online continual learning to learn sequence of new tasks from data stream
- Selecting candidates from each learned task and then incorporating the corresponding weights instead of single-head classifier
- No need for original data



Paper:
Jiangpeng, He, et al. “Online Continual Learning Via Candidates Voting”, WACV 2022

Multimodal Few-Shot-Learning

- Huge language models (LMs) like GPT-3 show zero/few-shot learning capabilities although only trained to predict next word
- Transfer this to multimodal setting
- From simple frozen LM to advanced adapter layers



Papers:

- Tsimpoukelli et al., “Multimodal Few-Shot Learning with Frozen Language Models,” (NIPS, 2021)
- Alayrac et al., “Flamingo: A Visual Language Model for Few-Shot Learning” (arXiv, 2022)
- Yang et al., “Zero-Shot Video Question Answering via Frozen Bidirectional Language Models” (arXiv, 2022)
- Driess et al., “PaLM-E: An Embodied Multimodal Language Model” (arXiv, March 6, 2023)

Augmented Language Models – Survey

- Huge Language Models do not observe and act in the world
- Augmentation allows usage of tools, APIs etc.
- From calculator for solving math task...
- ... to using search engine
- ... to LLM writing code to control a robot

Question: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?
Answer: Roger started with 5 balls.
`tennis_balls = 5`
2 cans of 3 tennis balls each is
`bought_balls = 2 * 3`
tennis balls. The answer is
`answer = tennis_balls * bought_balls`

Question: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?
Answer:
<LM>

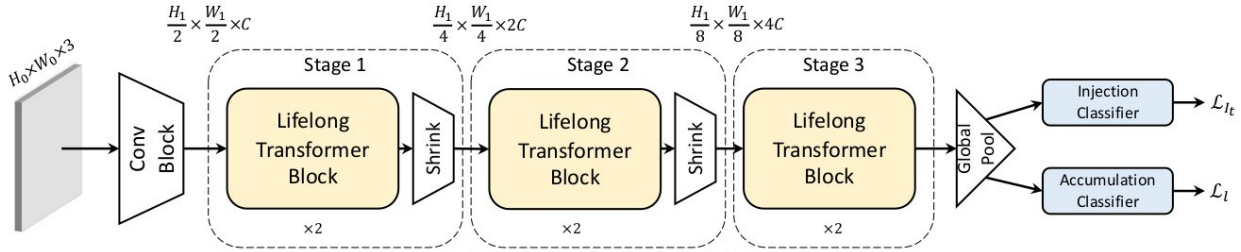
Figure 6: An example of few-shot PAL (Gao et al., 2022) prompt.

Papers:

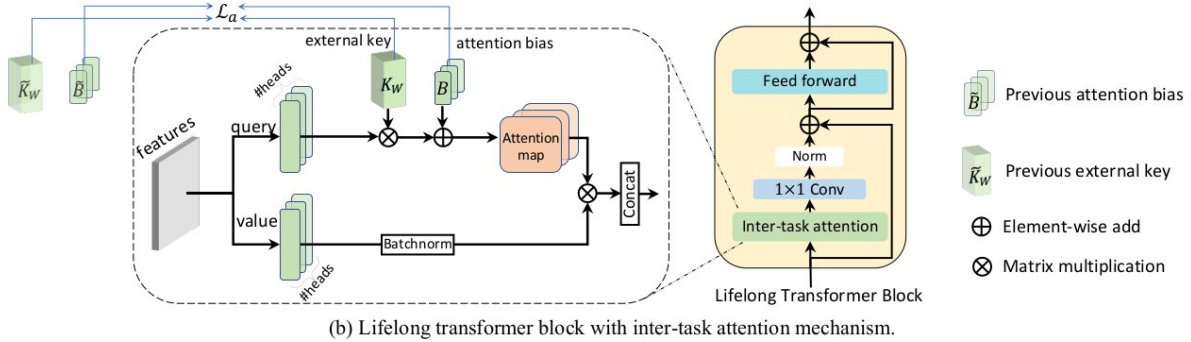
- Mialon et al., “Augmented Language Models: A Survey” (arXiv, February 15, 2023)
- Liang et al., “Code as Policies: Language Model Programs for Embodied Control,” in *Workshop on Language and Robotics at CoRL 2022*

Continual Learning with Lifelong Vision Transformer

- Training with sequential data => Existing methods are based on CNN
- Novel attention-based framework Lifelong Vision Transformer (LVT)
- Inter-task attention mechanism: Implicitly absorbs information from the previous tasks
- Slows down the drift of important attention between previous and current tasks.



(a) Overall architecture of LVT.

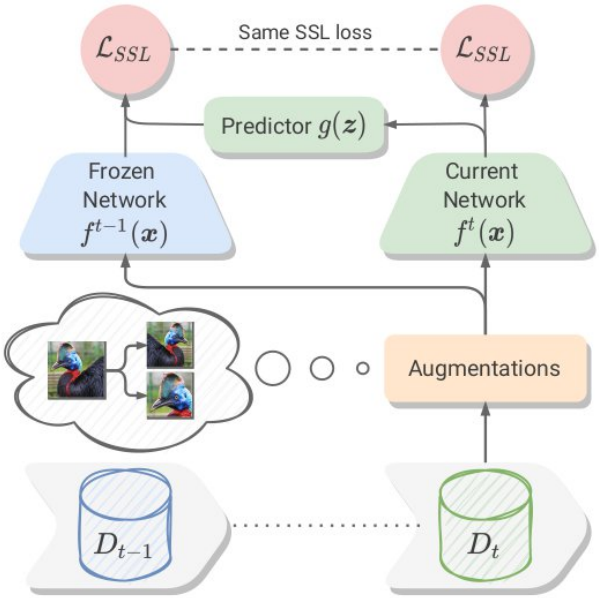


(b) Lifelong transformer block with inter-task attention mechanism.

Paper: Wang, Zhen et al. “Continual Learning with Lifelong Vision Transformer”, CVPR 2022

Self-Supervised Models are Continual Learners

- **Task:** Self-supervised continual visual representation learning
- **Method:**
  - Benefit from self-supervised loss function
  - Convert it into distillation mechanism for continual learning by using a predictor network
  - Predictor network: Maps the current state of the representations to their past state.



Paper:

- Fini, Enrico, et al. “Self-Supervised Models are Continual Learners” CVPR 2022

Online Continual Learning of End-to-End Speech Recognition Models

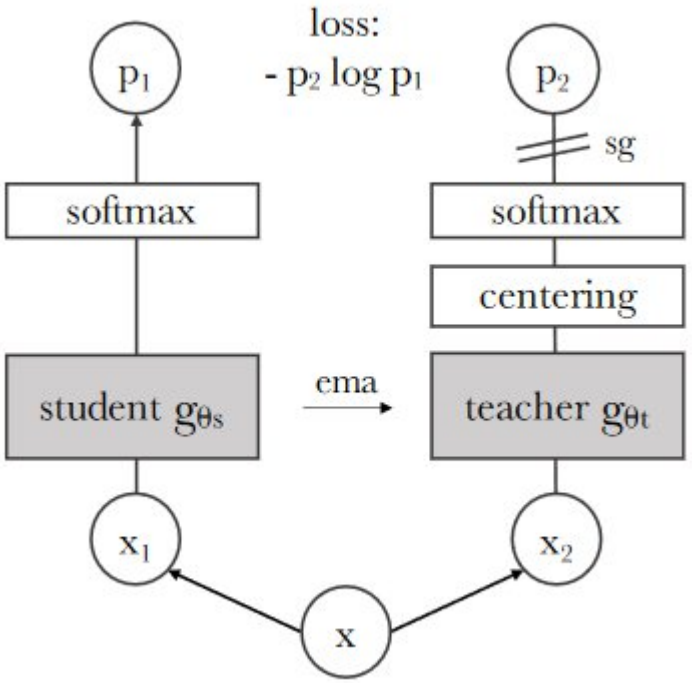
- Task: Explain Online Gradient Episodic Memory (GEM) method and the application to speech recognition

- Yang, Muqiao, Ian Lane, and Shinji Watanabe. "Online Continual Learning of End-to-End Speech Recognition Models." arXiv preprint arXiv:2207.05071 (2022).

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

- Emerging properties in self-supervised vision transformers (Caron et al. 2021)



Automatic Prompt Design

- **Review methods for automatic prompt design for pretrained language models (PLM)**
- Prompt design has big impact on the performance of the language model
- Idea: Automatically optimize or search for good prompt designs
- Recent approaches include AutoPrompt[1] (2020), Prefix-Tuning[2] (2021), P-tuning[3] (2021), Prompt Tuning[4] (2022), Automatic Prompt Engineer[5] (2022), Auto-CoT[6] (2023)

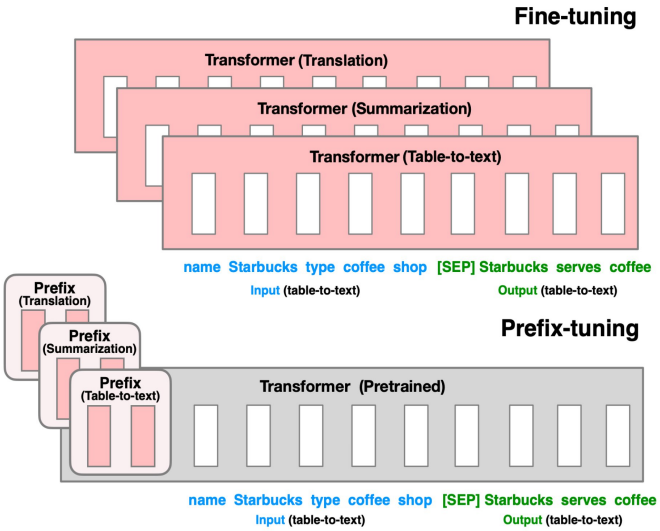


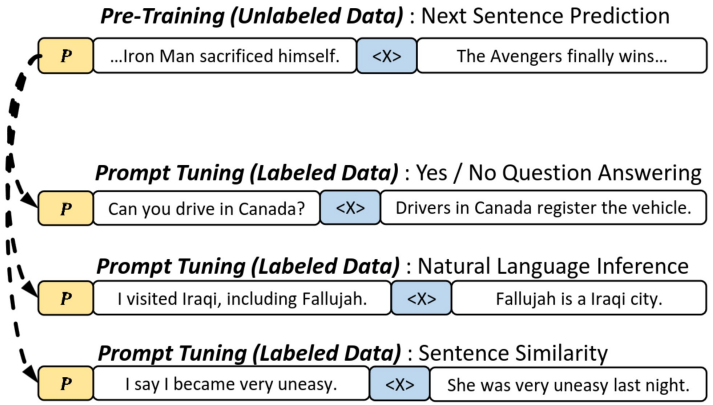
Illustration of Prefix-Tuning from [2]

Papers (Suggestions):

- [1] *AutoPrompt: Eliciting Knowledge from Language Models with ..., Shin et al. (2020)*
- [2] *GPT Understands, Too, Liu et al. (2021)*
- [3] *Prefix-Tuning: Optimizing Continuous Prompts for Generation, Li et al. (2021)*
- [4] *The Power of Scale for Parameter-Efficient Prompt Tuning, Lester et al. (2022)*
- [5] *Large Language Models Are Human-Level Prompt Engineers, Zhou et al. (2022)*
- [6] *Automatic Chain of Thought Prompting in Large Language Models, Zhang et al. (2023)*
- [7] *Automatic Prompt Augmentation and Selection with Chain-of-Thought ..., Shum et al. (2023)*

PPT: Pre-trained Prompt Tuning for Few-shot Learning

- pre-trained large language models (PLLMs) have large general domain knowledge
- fine-tuning often too time consuming, hard prompt tuning is a good alternative:
  sentence: I like eating apples
  sentiment: positive
  sentence: The film was horrible
  sentiment:
- alternative: soft prompts that are automatically generated (with backpropagation algorithm)
- in some few-shot learning scenarios prompt tuning is worse than fine-tuning a smaller model
- one solution: adding soft prompts into the pre-training stage to improve performance for these scenarios



Paper: Yuxian Gu, Xu Han, Zhiyuan Liu, and Minlie Huang, “PPT: Pre-trained Prompt Tuning for Few-shot Learning”, Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL), 2022

Variable Few Shot Class Incremental and Open World Learning

- Previous work: Fixed-Few-Shot Class Incremental Learning (number of ways and shots are known, but this is not true in real-world cases)
- Number of ways: how many new classes are added
- Number of shots: how many examples per class are provided
- Variable-Few-Shot Class Incremental Learning (Up-to N-Ways, Up-to K-Shots class incremental learning)
- Detect unknown classes
- Main paper assumes labels for unknown classes, possible solution in real-world: ask user for labels (organic learning like a child)

Paper: Ahmad, Dhamija, Jafarzadeh, Cruz, Rabinowitz, Li, Boult., “Variable Few Shot Class Incremental and Open World Learning”, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022

Sleep prevents catastrophic forgetting in spiking neural networks

- Spiking Neural Networks (SNN): Spiking neural networks (SNNs) are a type of artificial neural network that model the behavior of biological neurons more closely than traditional neural networks.

- In an SNN, the neurons communicate with each other by sending brief electrical impulses, or "spikes," rather than continuously varying signals. These spikes can be thought of as digital events that occur when a neuron's input exceeds a certain threshold.

- One advantage of SNNs is that they can use a sparse coding scheme, meaning that only a subset of the neurons are active at any given time. This can lead to more efficient use of computational resources and may better reflect the way that the brain processes information.

Sleep has been hypothesized to play an important role in memory and learning by enabling spontaneous reactivation of previously learned memory patterns.

SNN used with reinforcement learning to show during new task training sleep-like activity optimizes the network's memory representation in synaptic weight space to prevent forgetting old memories.

Paper: Golden, Ryan et al. "Sleep prevents catastrophic forgetting in spiking neural networks by forming a joint synaptic weight representation", PLOS COMPUTATIONAL BIOLOGY, 2022