Continuous Learning for Long-term Ecological Monitoring

Paul Bodesheim

Computer Vision Group, Friedrich Schiller University Jena

CamTrap Ecology Meets AI Workshop 2022, September 28th

FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Computer Vision Group

AMMOD Automated Biodiversity Monitoring

Fine-grained recognition

- How work in this area started: bird species identification
- Distinction of highly similar species by small details
  $\Rightarrow$ Small (visual) differences between species (classes)
- *Wah et al. (2011): The Caltech-UCSD Birds-200-2011 Dataset.*
  $\Rightarrow$ Dataset paper with more than 2,500 citations by now

Fine-grained recognition

- How work in this area started: bird species identification
- Distinction of highly similar species by small details
  $\Rightarrow$ Small (visual) differences between species (classes)
- *Wah et al. (2011): The Caltech-UCSD Birds-200-2011 Dataset.*
  $\Rightarrow$ Dataset paper with more than 2,500 citations by now

- Part-based approaches (allow for attribution of decisions), history of work in our group:

Göring at al.: Nonparametric Part Transfer for Fine-grained Recognition. CVPR 2014.

Freytag et al.: Exemplar-specific Patch Features for Fine-grained Recognition. GCPR 2014.

Simon and Rodner: Neural Activation Constellations: Unsupervised Part Model Discovery with Convolutional Networks. ICCV 2015.

Rodner et al.: Fine-grained Recognition in the Noisy Wild: Sensitivity Analysis of Convolutional Neural Networks Approaches. BMVC 2016.

Simon et al.: Generalized orderless pooling performs implicit salient matching. ICCV 2017.

Korsch et al.: Classification-Specific Parts for Improving Fine-Grained Visual Categorization. GCPR 2019.

Simon et al.: The Whole Is More Than Its Parts? From Explicit to Implicit Pose Normalization. TPAMI 2020.

Korsch et al.: End-to-end Learning of Fisher Vector Encodings for Part Features in Fine-grained Recognition. GCPR 2021.

Individual identification: a special fine-grained scenario

Long species lists and corresponding datasets that have been considered in the past

Terrestrial:

- Amur Tiger
- Brown Bear
- Cheetah
- Elephants
- Great Apes
- Holstein-Friesian Cattle
- Panda
- Zebra
- ...

Aquatic:

- Common Dolphin
- Great White Shark
- Green Turtle
- Humpback Whale
- Manta Ray
- Ringed Seal
- ...

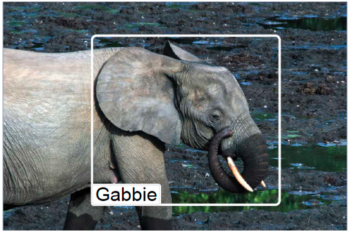
Insects:

- Bumblebee
- Fruit Fly
- ...

(Sorry if I missed your favorite species / the species you are working with!)

Our previous work on identifying individuals: elephants

- Elephant Listening Project, Cornell Lab of Ornithology
  ⇒ `https://elephantlisteningproject.org/`
- **Our elephant ID system is being used in the field!**
- *Körschens et al.: Towards Automatic Identification of Elephants in the Wild. AIWC Workshop 2018.*
- *Körschens and Denzler: ELPephants: A Fine-Grained Dataset for Elephant Re-Identification. ICCV Workshop 2019.*



Our previous work on identifying individuals: apes

Gorillas:

- **Identification system for field photographs** based on detecting and recognizing Gorilla faces
- Mbeli Bai study at the Nouabal-Ndoki National Park, Republic of Congo (Wildlife Conservation Society)
- *Brust et al.: Towards Automated Visual Monitoring of Individual Gorillas in the Wild. ICCV Workshop 2017.*

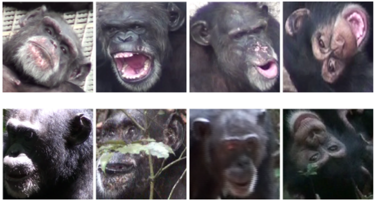
Our previous work on identifying individuals: apes

Gorillas:

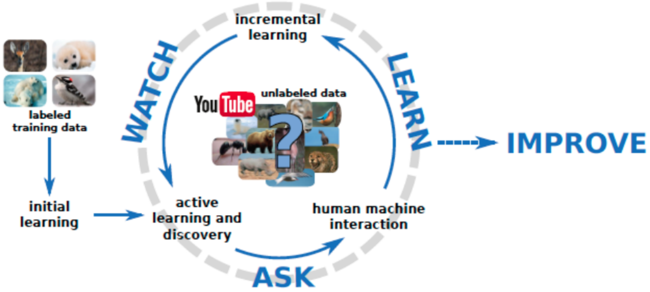
- **Identification system for field photographs** based on detecting and recognizing Gorilla faces
- Mbeli Bai study at the Nouabal-Ndoki National Park, Republic of Congo (Wildlife Conservation Society)
- *Brust et al.: Towards Automated Visual Monitoring of Individual Gorillas in the Wild. ICCV Workshop 2017.*

Chimpanzees:

- **Predicting IDs and attributes** (gender, age, age group)
- *Freytag et al.: Chimpanzee Faces in the Wild: Log-Euclidean CNNs for Predicting Identities and Attributes of Primates. GCPR 2016.*
- *Käding et al.: Active Learning for Regression Tasks with Expected Model Output Changes. BMVC 2018.*



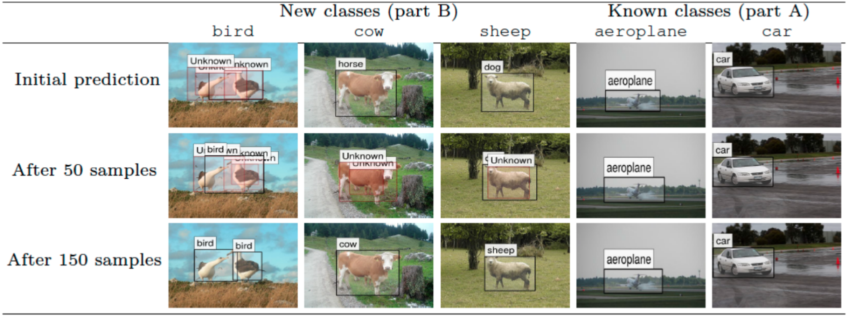
Active learning, e.g., via our WALI framework



- Machine learning model selects which samples are worth annotating (automatic selection)
- Making use of these newly labeled samples requires incremental model learning
- Always training from scratch is suboptimal / computationally expensive

Käding et al.: Watch, Ask, Learn, and Improve: A Lifelong Learning Cycle for Visual Recognition. European Symposium on Artificial Neural Networks (ESANN) 2016.

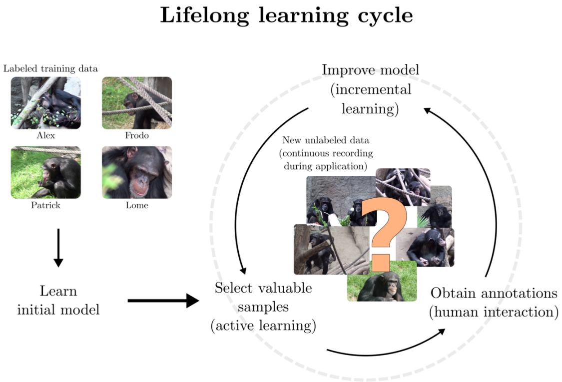
Incremental learning, e.g., of deep object detectors



Brust et al.: Active Learning for Deep Object Detection. VISAPP 2019.

Lifelong learning concept

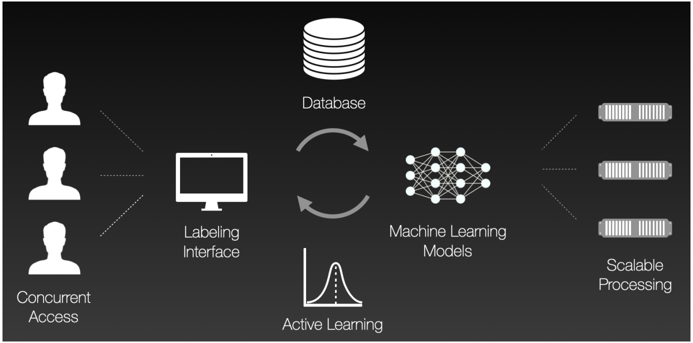
- Fixed pre-trained recognition models might quickly reach their limits when applied in long-term monitoring studies
- **Continuous learning with model adaptations** to new environments and unseen visual appearances of animals is required to improve recognition models over time
- **Exploit the continuous data stream of recordings during the application**
- The lifelong learning framework (with active learning and human-in-the-loop) offers a possible solution



Bodesheim et al.: Pre-trained models are not enough: active and lifelong learning is important for long-term visual monitoring of mammals in biodiversity research. Mammalian Biology 2022.

Related work on active learning by others

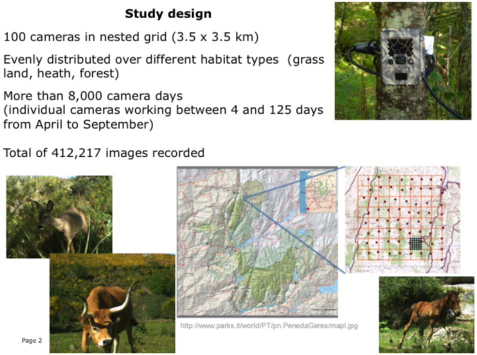
- *Norouzzadeh et al.: A deep active learning system for species identification and counting in camera trap images. Methods in Ecology and Evolution 2020.*
- *Kellenberger et al.: AIDE: Accelerating image-based ecological surveys with interactive machine learning. Methods in Ecology and Evolution 2020.*



AIDE workflow, image source: `https://github.com/microsoft/aerial_wildlife_detection`

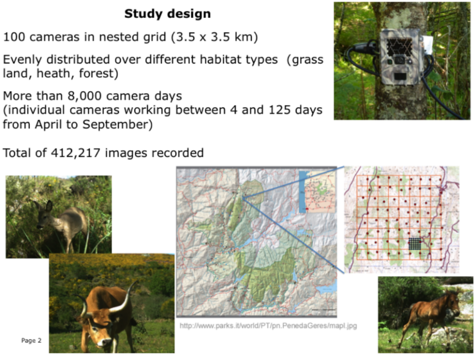
(Not to be confused with: *Dimitriadou et al.: AIDE: An Active Learning-Based Approach for Interactive Data Exploration. IEEE Transactions on Knowledge and Data Engineering 2016.*)

Our former monitoring task: herbivorous mammals in Portugal



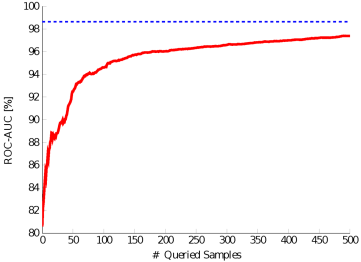
Joint work with Andrea Perino from iDiv (thanks for the figure)

Our former monitoring task: herbivorous mammals in Portugal



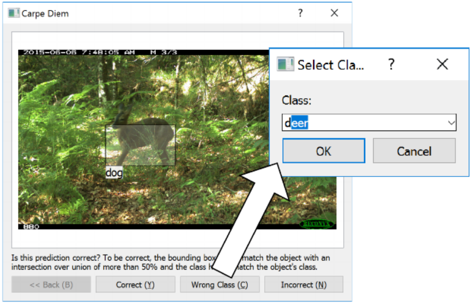
Joint work with Andrea Perino from iDiv (thanks for the figure)

First task: does the image contain an animal or not? (identify empty images)



Käding et al.: Large-scale Active Learning with Approximated Expected Model Output Changes. GCPR 2016.

Tools and systems for human interaction: label correction



- *Käding et al.: Large-scale Active Learning with Approximated Expected Model Output Changes. GCPR 2016.*
- *Brust et al.: Active and Incremental Learning with Weak Supervision. KI 2020.*
- *Brust et al.: Carpe Diem: A Lifelong Learning Tool for Automated Wildlife Surveillance. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.*

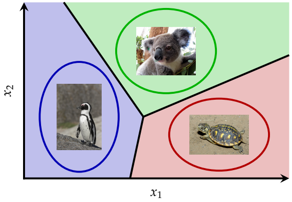
Tools and systems for human interaction: identification

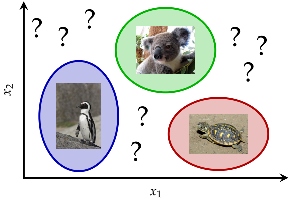
- *Käding et al.: Active Learning and Discovery of Object Categories in the Presence of Unnameable Instances. CVPR 2015.*
- *Käding et al.: Large-scale Active Learning with Approximated Expected Model Output Changes. GCPR 2016.*
- *Brust et al.: Active and Incremental Learning with Weak Supervision. KI 2020.*
- *Brust et al.: Carpe Diem: A Lifelong Learning Tool for Automated Wildlife Surveillance. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.*

Tools and systems for human interaction: discovery

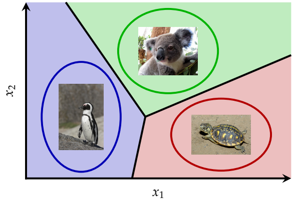
- *Käding et al.: Active Learning and Discovery of Object Categories in the Presence of Unnameable Instances. CVPR 2015.*
- *Käding et al.: Large-scale Active Learning with Approximated Expected Model Output Changes. GCPR 2016.*
- *Brust et al.: Active and Incremental Learning with Weak Supervision. KI 2020.*
- *Brust et al.: Carpe Diem: A Lifelong Learning Tool for Automated Wildlife Surveillance. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.*

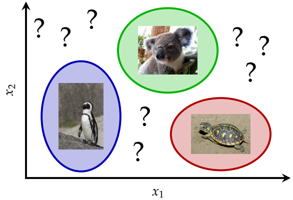
Closed-world vs. open-set recognition: novelty detection

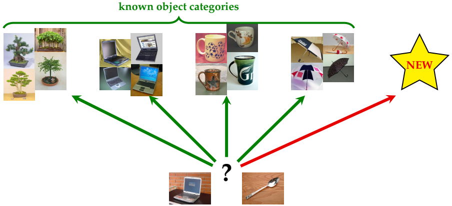


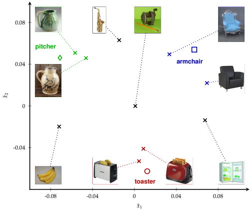


Closed-world vs. open-set recognition: novelty detection



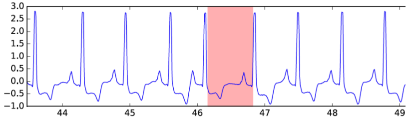






- *Bodesheim et al.: Kernel Null Space Methods for Novelty Detection. CVPR 2013.*
- *Bodesheim et al.: Local Novelty Detection in Multi-class Recognition Problems. WACV 2015.*
- *Schultheiss et al.: Finding the Unknown: Novelty Detection with Extreme Value Signatures of Deep Neural Activations. GCPR 2017.*

Anomaly detection in videos and time series data

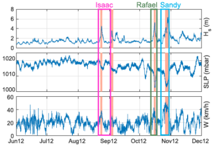


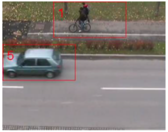
Preliminaries

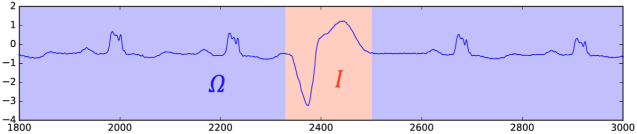
- Given: multivariate time-series $(x_t)_{t=1}^{n},\ x_t \in \mathbb{R}^D$
- $I = \{t \mid t_1 \leq t < t_2\},\ \Omega = \{1, \ldots, n\} \setminus I$

Anomaly Score: Kullback-Leibler Divergence

$$KL(p_I, p_\Omega) = \int p_I(x_t) \cdot \log \frac{p_I(x_t)}{p_\Omega(x_t)} dx_t$$





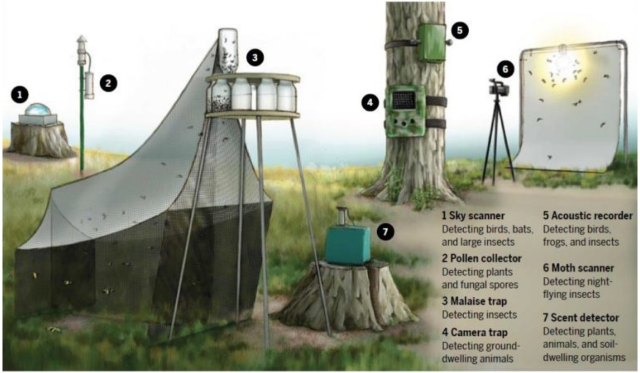


Barz et al.: Detecting Regions of Maximal Divergence for Spatio-Temporal Anomaly Detection. PAMI 2018.

`https://cvjena.github.io/libmaxdiv` and `https://github.com/cvjena/libmaxdiv`

The idea of a “weather station for biodiversity”

AMMOD project: **A**utomated **M**ultisensor Stations For **M**onitoring **O**f Bio**D**iversity (`https://ammod.de`)



Graphic by V.ALTOUNIAN/SCIENCE. From „Where have all the insects gone?“ by Gretchen Vogel, SCIENCE, May 10, 2017 (doi:10.1126/science.aal1160).

- **Visual monitoring / Camera traps**
- Smellscapes / Scent detector
- Metabarcoding / Pollen collector and Malaise traps
- Acoustic monitoring / Sound recordings
- Self-sustaining stations / Automatized data transfer to central cloud storage
- *Wägele et al.: Towards a multisensor station for automated biodiversity monitoring. Basic and Applied Ecology 2022.*

visAMMOD teams

- TU Munich, Department of Computer Science
  - Bernd Radig, Franziska Schmickler, Ludwig Kürzinger
  - Hardware development of moth scanner and stereo camera setup
- University of Bonn, Institute of Computer Science 4: Security and Networked Systems
  - Volker Steinhage, Timm Haucke, Morris Klasen, Frank Schindler
  - Depth estimation, 3D reconstruction and analysis
- **Friedrich Schiller University Jena, Computer Vision Group**
  - **Paul Bodesheim, Joachim Denzler, Daphne Auer, Julia Böhlke, Dimitri Korsch**
  - **Detection and species identification, moths (light traps) and wildlife (camera traps)**

Associated partners:

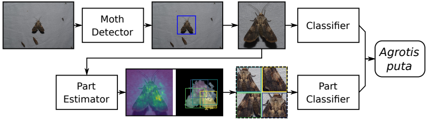
- *Tilo Burghardt (University of Bristol)*
- *Christian Fiderer and Marco Heurich (Bavarian Forest National Park)*
- *Gunnar Brehm (Phyletic Museum Jena)*

The AMMOD moth scanner prototype in Jena

- Joint work with Gunnar Brehm (Phyletic Museum Jena) who built the setup, UV light source to attract insects
- Continuous monitoring since June 2021, images taken at regular intervals during night, e.g., every 2 minutes
- June to October 2021: 95 nights, 27,455 images (201 GB), detection & **recognition down to species level!**
- *Korsch et al.: Automated Visual Monitoring of Nocturnal Insects with Light-based Camera Traps. CVPR Workshop on Fine-grained Visual Classification 2022.*

Two approaches for moth species identification

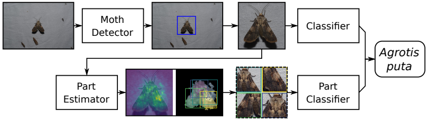
Part-based classification:



- *Korsch et al.: Classification-Specific Parts for Improving Fine-Grained Visual Categorization. GCPR 2019.*
- *Korsch et al.: Deep Learning Pipeline for Automated Visual Moth Monitoring: Insect Localization and Species Classification. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.*

Two approaches for moth species identification

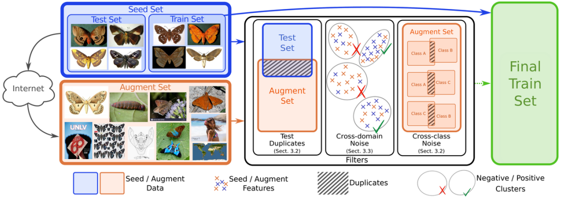
Part-based classification:



- *Korsch et al.: Classification-Specific Parts for Improving Fine-Grained Visual Categorization. GCPR 2019.*
- *Korsch et al.: Deep Learning Pipeline for Automated Visual Moth Monitoring: Insect Localization and Species Classification. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.*

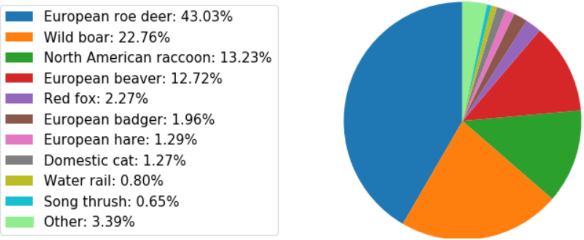
Webly supervised learning:

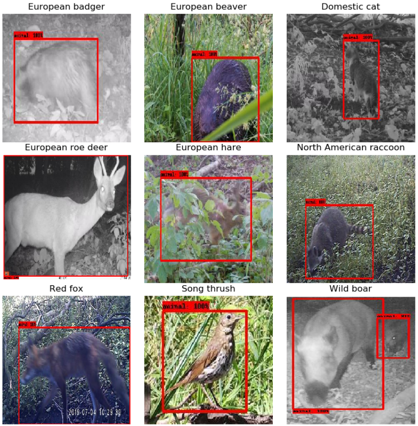
- *Böhlke et al.: Lightweight Filtering of Noisy Web Data: Augmenting Fine-grained Datasets with Selected Internet Images. VISAPP 2021.*
- *Böhlke et al.: Exploiting Web Images for Moth Species Classification. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.*



Wildlife camera traps: Brandenburg video dataset

- Joint work with Tilo Burghardt (University of Bristol) and his student George Ioannou
- 1.5 TB of *video data* from camera traps in Brandenburg
- 72.3% of the videos are empty (do not contain an animal), species in remaining videos:





Video dataset provided by Hjalmar Kühl

Wildlife camera traps: Bavarian Forest National Park

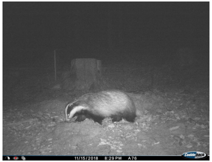

Red deer


Red fox


Red squirrel


Wild boar


Two wild boars


European badger

Challenges:

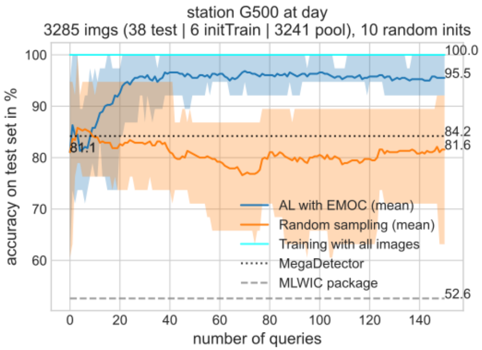
- Daytime vs. nighttime
- Small vs. large animals
- Occlusion and truncation
- Scene / background clutter

Approach:

- Filter empty images (binary task: empty or not)
- Species classification in a lifelong learning scenario

Image datasets from different camera site networks (hundreds of locations, forest and trail sites) provided by Christian Fiderer and Marco Heurich $\Rightarrow$ many images do not contain an animal (often more than 50%)

Active learning for filtering empty images (without animals)



- Region-specific models, e.g., one per station
- Distinction between daytime and nighttime possible
- Active learning approach gradually improves recognition performance with minimal annotation efforts
- **Less than 5% of the images need to be annotated to achieve 95.5% accuracy!**

Auer et al.: Minimizing the Annotation Effort for Detecting Wildlife in Camera Trap Images with Active Learning. Computer Science for Biodiversity Workshop (CS4Biodiversity) 2021.

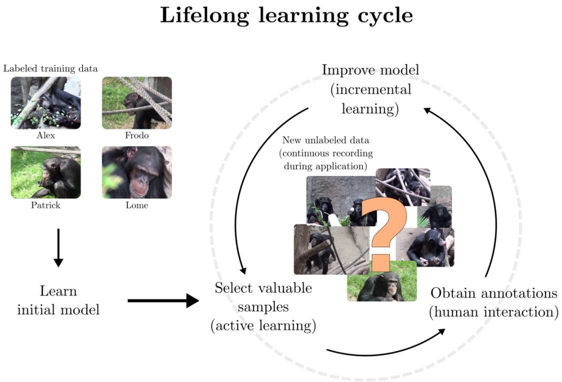
Continuous learning with neural networks: rehearsal learning

- Continuously incoming data (streams, experiences) that can be used to update/improve the species classifier
- Train the neural network for further epochs when new data arrives (fine-tuning), different strategies are possible:

1. **Naive approach:** only use the new data for additional training steps
   $\Rightarrow$ Overfitting to new data, catastrophic forgetting, concept drift
2. **Cumulative approach:** use the new data and all the previously seen data
   $\Rightarrow$ Computational costs explode over time due to steadily increasing number of samples
3. **Rehearsal learning:** use the new data and a subset of the previously seen data
   $\Rightarrow$ Trade-off, requires strategies for selecting the subset

Again the lifelong learning concept

- Fixed pre-trained recognition models might quickly reach their limits when applied in long-term monitoring studies
- **Continuous learning with model adaptations** to new environments and unseen visual appearances of animals is required to improve recognition models over time
- **Exploit the continuous data stream of recordings during the application**
- The lifelong learning framework (with active learning and human-in-the-loop) offers a possible solution



Bodesheim et al.: Pre-trained models are not enough: active and lifelong learning is important for long-term visual monitoring of mammals in biodiversity research. Mammalian Biology 2022.

Summary

- Fine-grained species recognition and individual identification are challenging tasks for humans *and* machines $\Rightarrow$ Team up, **human-in-the-loop approaches**
- Incorporate **expert knowledge and feedback**, e.g., via active learning
- Continuous adaptations are required for successful long-term monitoring (changing environments, new individuals, ...) $\Rightarrow$ **Lifelong learning**, e.g., within the AMMOD project

Summary

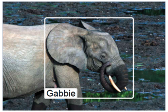
- Fine-grained species recognition and individual identification are challenging tasks for humans *and* machines $\Rightarrow$ Team up, **human-in-the-loop approaches**
- Incorporate **expert knowledge and feedback**, e.g., via active learning
- Continuous adaptations are required for successful long-term monitoring (changing environments, new individuals, ...) $\Rightarrow$ **Lifelong learning**, e.g., within the AMMOD project
- **Systems are required** that can be used by ecologists (e.g., to provide feedback and to derive ecological metrics), examples from our group:
  - Monitoring system for herbivorous mammals in Portugal (iDiv, Leipzig)
  - EIS: Elephant Identification System (Cornell Lab of Ornithology, Elephant Listening Project)
  - Identification and age estimation of gorillas (WCS, Mbeli Bai Study)
  - AMMOD biodiversity monitoring system (in development, currently several sites in Germany)

Thank you for your attention!

Contact:

Paul Bodesheim (`paul.bodesheim@uni-jena.de`)

Computer Vision Group: `https://www.inf-cv.uni-jena.de`
The AMMOD project: `https://ammod.de`

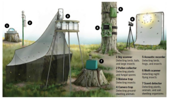


Acknowledgments for project funding:

SPONSORED BY THE

Federal Ministry of Education and Research

FONA
Research for Sustainability



Lifelong learning cycle