

Workshop on Interactive and Adaptive Learning in an Open World

Alexander Freytag¹, Vittorio Ferrari^{2,3}, Mario Fritz⁴, Uwe Franke⁵, Terrence Boulton⁶, Juergen Gall⁷, Walter Scheirer⁸, Angela Yao⁷, and Erik Rodner¹

¹ Carl Zeiss AG

² Google

³ University of Edinburgh

⁴ CISPA Helmholtz Center i.G.

⁵ Daimler AG

⁶ University of Colorado, Colorado Springs

⁷ University of Bonn, Germany

⁸ University of Notre Dame

Abstract. Next generation machine learning requires stepping away from classical batch learning towards interactive and adaptive learning. This is essential to cope with demanding machine learning applications we have already today. Our workshop at ECCV 2018 in Munich therefore served as a discussion forum for experts in this field and in the following we give a brief overview. **Please note that this discussion paper has not been peer-reviewed and only contains the subjective summary of the workshop organizers.**

Keywords: interactive learning, adaptive learning, open set, continuous learning

1 Scope of the workshop

Learning algorithms are the backbone of computer vision research and still focused on training from large amounts of already annotated data. The limitations we are currently observing in many applications are mostly due to the lack of annotations or changing data distributions over time. To overcome these barriers, the annotation and learning of models needs to be coupled strongly through human-machine interaction. Furthermore, models need to adapt as needed to handle either shifts or completely novel data. The goal of this workshop was to discuss and present the advances in technologies that support annotation, model learning through expert guidance, and continuous model adaptation.

The interactive and adaptive learning (IAL) workshop tried to bridge one of the gaps between results of basic AI research and their real-world applicability - availability of useful and easy-to-produce annotations and working solutions for efficient model adaptation. In consequence, the following topics have been central to the workshop:

- Online and incremental learning

- Interactive segmentation and detection to support annotation
- Transfer learning
- Active or self-taught
- Continuous / lifelong learning
- Open set learning
- Open domain learning
- Efficient fine-tuning of generic models

The topics are often seen as separated research fields, however, they should be considered jointly. While preparing this workshop, we once phrased this area as *Machine Didactics*, referring to the fact that we need to improve not only the training but also the teaching of models, which includes the way we collect and annotate data. For a lot of applications, it is simply unreasonable to assume that there is a clear division between annotation phase and the phase where a model is trained and tested with the data. In practice, this is always a continuous cycle of improvement and challenging the annotators with more data and further requirements. Currently, this process is still driven by manual work both from machine learning engineers as well as domain experts. The basic question for the future would be how this can be assisted by proper algorithms as well, such as active learning algorithms choosing the examples to annotate or bootstrapped feedback loops that allow experts to tune and check annotations rather than creating them from scratch with a lot of effort.

Another important aspect of the workshop stems from the fact that there are often unclear requirements of machine learning algorithms in the beginning. In reality, the number of classes that need to be differentiated in a classification task is simply often not defined or is likely to increase and change over time. This is referred to as an *open world* situation and is far more challenging than the standard ImageNet-like competition task most researchers focus on today. We were therefore very happy to have an associated challenge on open-set face recognition organized by Terry and Walter that was presented in detail during the workshop.

In addition to the aforementioned risks, a real-world machine learning application is likely to face changes of input conditions resulting from changing a sensor or the application field. Dealing with this problem requires training from only a few examples by transfer learning or learning of generic representations that allow for jump-starting learning for various tasks.

We invited extended abstract submissions related to the workshop scope and also compiled the list of invited speakers according to the fit of their main research interests to the workshop idea.

2 Invited speakers

2.1 Incremental learning: a critical view on the current state of affairs (Tinne Tuytelaars)

In the first invited talk of the workshop, Tinne Tuytelaars (KU Leuven) gave an overview on recent developments in the field of incremental learning. She high-

lighted current scenarios of incremental learning and argued that the majority of existing approaches is hardly comparable due to not-matching assumptions on the availability of tasks and data over time. She presented several approaches from her recent work [1–4] which tackle this issue and address the problem of catastrophic forgetting, e.g., by encouraging sparse representations to leave model capacity for subsequent tasks that are added over time.

2.2 Results and Evaluation of the Open-Face Challenge (Manuel Günther)

Manuel Günther (UCCD) presented the UnConstrained College Students (UCCS) dataset which is an Open-Face Challenge [5]. Subjects are photographed using a long-range high-resolution surveillance camera. Faces inside these images are of various poses, and varied levels of blurriness and occlusion. The challenge comes with a closed set recognition problem as well as an open set recognition problem. In addition, different attack scenarios are evaluated. More information about the challenge, the data, terms of usage, and recent results can be found on the challenge’s webpage at <http://vast.uccs.edu/Opensetface/>

During the discussion, the spent effort and the availability of the dataset was positively acknowledged. All participants further agreed on the difficulty of the task of re-identifying individuals based on single sub-images. Nonetheless, issues have been raised why the re-identification task is posed on single images, whereas the ground truth to validate the ids required entire video clips (which would also likely be the final application scenario).

2.3 Recognition with unseen compositions and novel environments (Kristen Grauman)

Kristen Grauman (UT Austin) put emphasis on two aspects of open-ended learning: how to recognize unseen compositions of objects and operators as well as how to operate and navigate in unseen environments.

In her recent work [6], Kristen and her team show how operations such as slicing an apple, i.e., operations which transform objects, can be modeled as object-operator pairs and can be realized as operators applied to object representations. Appropriate embeddings are learned by optimizing a triplet-loss and additionally adding semantic regularizers, e.g., enforcing operators to be invertible which resembles undoing a transformation. In consequence, the notion of operators can also be generalized to new compositions of operator-object-pairs.

In the second part of her talk, Kristen focused on self-learning agents which are faced with environments that have been unseen at training time [7]. Based on a reinforcement learning approach, they proposed an additional reward for actions which reduce the estimated uncertainty about the agent’s environment. An interesting future direction is to combine this unsupervised exploration with active look-ahead strategies [8].

2.4 Interactive video segmentation: The DAVIS benchmark and first approaches (Jordi Pont-Tuset)

Jordi Pont-Tuset (Google AI) gave an overview of his work on video segmentation. In particular, he presented the DAVIS benchmark [9] and the video segmentation approach published in [10]. The latter only requires the annotation of a few key frames and allows propagating the region segmentation to the whole video. This work is one example for the focus of the workshop on reducing annotation efforts by interactive segmentation and in general assisting the annotator by propagating annotations in an intelligent manner. Especially for pixel-wise video segmentation, fully manual annotation often renders intractable. One of the key ideas of the underlying algorithm is to perform metric learning to phrase the segmentation as a retrieval problem on the pixel level later on.

2.5 Towards continual learning and interactive annotation (Christoph Lampert)

Christoph Lampert (IST Austria) presented recent results in the area of life-long learning and interactive annotation. In the first part of this talk, he reviewed iCARL, the Incremental Classifier and Representation Learning [11], which jointly learns appropriate embeddings and classification models upon the presence of newly added data. In continuous learning scenarios, it is further possible that unlabeled data is available and individuals tasks can be selected for annotation. How to select tasks such that information can be optimally transferred was shown in [12]. To assist in the annotation of new data, learnable bounding box dialogs for interactive annotation were presented in [13]. Finally, his work in [14] shows a simple yet powerful statistic test to detect if an incoming stream of data deviates from data a model has been trained on. By comparing distributions of model confidence scores, e.g., the maximum class score of deep convnets, the KS-test yields a probability if an entire batch of test samples stems from a different data distribution, e.g., induced by sensor drifts.

2.6 Elements of Continuous Learning for Wildlife Monitoring (Joachim Denzler)

Joachim Denzler (Univ. Jena) presented recent advances in continuous learning, especially focusing on active learning and anomaly detection. With the contributions of his group, he showed how application experts can be assisted in analyzing large-scale data using interactive machine learning tools, e.g., by spotting abnormal instances [15–17], interactively learning object classifiers [18, 19], regression models for animal age [20], or object detectors [21], and classifying large data collections from camera traps in a semi-automated fashion [22–24]. In summary, the recent tools and techniques already add large value to the application scientist’s work. Nonetheless, reliable and efficient interactive learning with deep neural networks remains an unsolved problem.

3 Extended Abstracts

Neal et al	Open Set Learning with Counterfactual Images
Gnther et al	Open-Set Recognition Challenge
Busto et al	Open Set Domain Adaptation for Image and Action Recognition
Dwivedi and Roig	Evaluation of plug and play modules for multi-domain learning
Jin et al	Unsupervised Hard Example Mining from Videos for Improved Object Detection
Osep et al	Towards Large-Scale Video Object Mining
Wang and Sharma	Unsupervised Representation Learning on Multispectral Imagery By Predicting Held-Out Bands
Sharma and Wang	Human-in-the-loop segmentation for improved segmentation and annotations
Bauermeister et al	Adaptive Network Architectures via Linear Splines
Rakelly et al	Few-Shot Segmentation Propagation with Guided Networks

4 Summary and Next Steps

The workshop successfully served as a venue for exchanging recent trends in the field of interactive and adaptive learning in an open world. The combination of invited speakers covering a broad technical spectrum as well as a short and informal poster session allowed for detailed discussions and for fostering connections.

The audience raised the strong interest in continuing the workshop within the next years. Of great benefit would be the continuation of a co-located challenge, especially in the area of open-set recognition.

References

1. Aljundi, R., Chakravarty, P., Tuytelaars, T.: Expert gate: Lifelong learning with a network of experts. In: CVPR. (2017) 7120–7129
2. Aljundi, R., Babiloni, F., Elhoseiny, M., Rohrbach, M., Tuytelaars, T.: Memory aware synapses: Learning what (not) to forget. In: European Conference on Computer Vision (ECCV). (2018)
3. Rannen, A., Aljundi, R., Blaschko, M.B., Tuytelaars, T.: Encoder based lifelong learning. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2017)
4. Aljundi, R., Rohrbach, M., Tuytelaars, T.: Selfless sequential learning. arXiv preprint arXiv:1806.05421 (2017)
5. Günther, M., Cruz, S., Rudd, E.M., Boulton, T.E.: Toward open-set face recognition. In: IEEE Conference on Computer Vision and Pattern Recognition – Workshops (CVPRW), IEEE (2017) 573–582
6. Nagarajan, T., Grauman, K.: Attributes as operators. In: European Conference on Computer Vision (ECCV). (2018)
7. Jayaraman, D., Grauman, K.: Learning to look around: Intelligently exploring unseen environments for unknown tasks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2018)
8. Ramakrishnan, S., Grauman, K.: Sidekick policy learning for active visual exploration. In: European Conference on Computer Vision (ECCV). (2018)
9. Caelles, S., Montes, A., Maninis, K.K., Chen, Y., Van Gool, L., Perazzi, F., Pont-Tuset, J.: The 2018 davis challenge on video object segmentation. arXiv preprint arXiv:1803.00557 (2018)
10. Chen, Y., Pont-Tuset, J., Montes, A., Van Gool, L.: Blazingly fast video object segmentation with pixel-wise metric learning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018) 1189–1198
11. Rebuffi, S.A., Kolesnikov, A., Sperl, G., Lampert, C.H.: icarl: Incremental classifier and representation learning. In: CVPR. (2017)
12. Pentina, A., Lampert, C.H.: Multi-task learning with labeled and unlabeled tasks. In: International Conference on Machine Learning (ICML). (2017)
13. Konyushkova, K., Uijlings, J., Lampert, C.H., Ferrari, V.: Learning intelligent dialogs for bounding box annotation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2018)
14. Sun, R., Lampert, C.H.: Ks(conf): A light-weight test if a convnet operates outside of its specifications. In: German Conference on Pattern Recognition (GCPR). (2018)
15. Barz, B., Rodner, E., Garcia, Y.G., Denzler, J.: Detecting regions of maximal divergence for spatio-temporal anomaly detection. IEEE Transactions on Pattern Analysis and Machine Intelligence (2018)
16. Garcia, Y.G., Shadaydeh, M., Mahecha, M., Denzler, J.: Extreme anomaly event detection in biosphere using linear regression and a spatiotemporal mrf model. Natural Hazards (2018) 1–19
17. Schultheiss, A., Kding, C., Freytag, A., Denzler, J.: Finding the unknown: Novelty detection with extreme value signatures of deep neural activations. In: German Conference on Pattern Recognition (GCPR). (2017) 226–238
18. Kding, C., Freytag, A., Rodner, E., Bodesheim, P., Denzler, J.: Active learning and discovery of object categories in the presence of unnameable instances. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2015) 4343–4352

19. Kding, C., Rodner, E., Freytag, A., Denzler, J.: Watch, ask, learn, and improve: A lifelong learning cycle for visual recognition. In: European Symposium on Artificial Neural Networks (ESANN). (2016) 381–386
20. Kding, C., Rodner, E., Freytag, A., Mothes, O., Barz, B., Denzler, J.: Active learning for regression tasks with expected model output changes. In: British Machine Vision Conference (BMVC). (2018)
21. Brust, C.A., Kding, C., Denzler, J.: Active learning for deep object detection. In: arXiv preprint arXiv:1809.09875. (2018)
22. Krschens, M., Barz, B., Denzler, J.: Towards automatic identification of elephants in the wild. In: AI for Wildlife Conservation Workshop (AIWC). (2018)
23. Brust, C.A., Burghardt, T., Groenenberg, M., Kding, C., Khl, H., Manguette, M.L., Denzler, J.: Towards automated visual monitoring of individual gorillas in the wild. In: ICCV Workshop on Visual Wildlife Monitoring (ICCV-WS). (2017) 2820–2830
24. Freytag, A., Rodner, E., Simon, M., Loos, A., Khl, H., Denzler, J.: Chimpanzee faces in the wild: Log-euclidean cnns for predicting identities and attributes of primates. In: German Conference on Pattern Recognition (GCPR). (2016) 51–63