Watch, Ask, Learn, and Improve: a lifelong learning cycle for visual recognition

since 1558

Christoph Käding, Erik Rodner, Alexander Freytag, and Joachim Denzler
Computer Vision Group, Friedrich Schiller University Jena, Germany

Friedrich Schiller University Jena
Computer Vision Group

The WALI System for learning animal classifiers from YouTube data

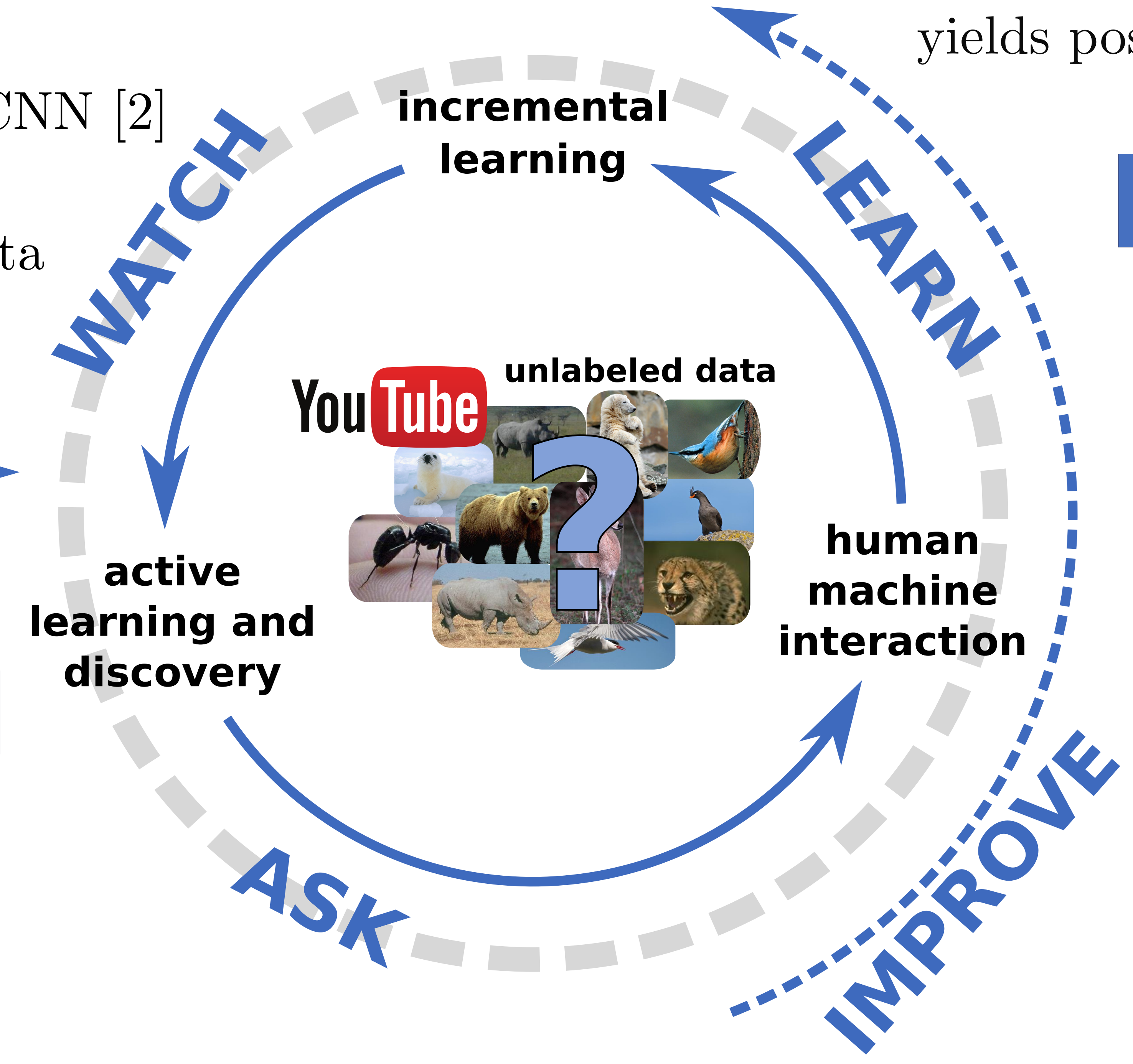
- **W**atch: download and watch YouTube videos autonomously
- **A**sk: actively select frames and ask human oracle for annotation
- **L**earn: incorporate new knowledge incrementally
- **I**mprove: the knowledge grows over time

Watch

- download and watch "animal documentary" videos from YouTube
- build long term memory out of 500 images which occurred during the last 5 hours of video
- normalized `relu7` features of BLVC AlexNet CNN [2] are used as feature representation
- every 10th frame is considered as unlabeled data

labeled training data

initial learning



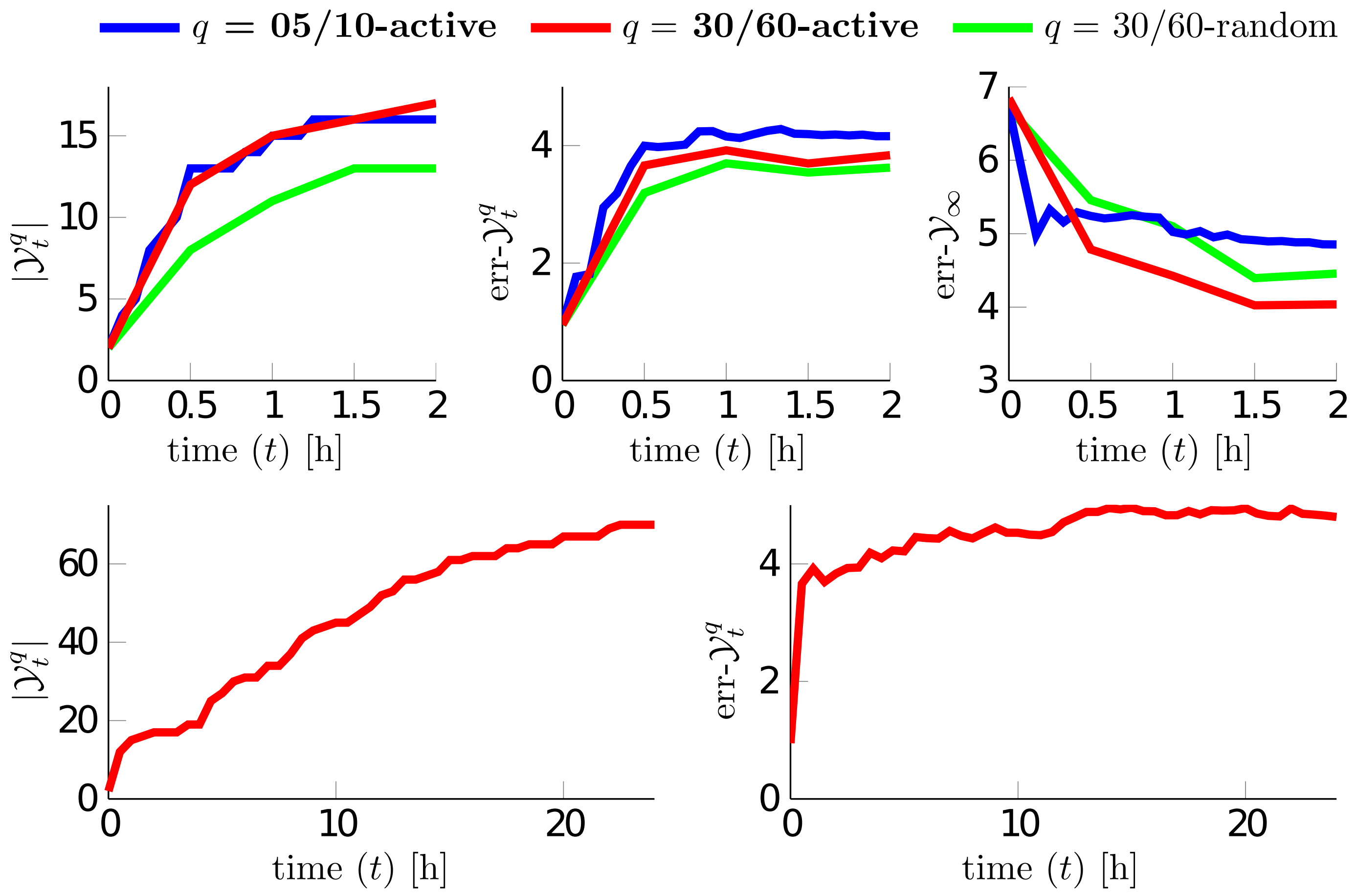
Ask

- one-vs-all classifier $\mathbf{w}_k^{\mathrm{T}}\mathbf{x}$ for each class $k \in \mathcal{Y}_t^q$
- active selection with 1-vs-2 strategy [3]:
$$\hat{k} = \operatorname{argmax}_{k \in \mathcal{Y}_t^q} \mathbf{w}_k^{\mathrm{T}}\mathbf{x}$$
$$q(\mathbf{x}) = \mathbf{w}_{\hat{k}}^{\mathrm{T}}\mathbf{x} - \operatorname{argmax}_{k \in \mathcal{Y}_t^q \setminus \{\hat{k}\}} \mathbf{w}_k^{\mathrm{T}}\mathbf{x}$$
- avoide inappropriate images via reject strategy [4]:
$$\tilde{q}(\mathbf{x}) = (1 - p(\text{rejection} \mid \mathbf{x})) \cdot q(\mathbf{x})$$

Learn

- update classifiers incrementally
- linear regression with quadratic loss function yields possibility for efficient updates [5]

Improve



- evaluation on corresponding ImageNet synsets [1]
- q = 05/10-active: watch 5 min and actively select 10 instances
- q = 30/60-active/random: watch 30 min and actively/randomly select 60 instances
- $|\mathcal{Y}_t^q|$: number of discovered classes
- err-$\mathcal{Y}_t^q$: hierarchical error [1] with respect to currently discovered classes $\mathcal{Y}_t^q$
- err-$\mathcal{Y}_\infty$: hierarchical error [1] regarding all classes $\mathcal{Y}_\infty$

[1] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet Large Scale Visual Recognition Challenge. *IJCV*, 2015. [2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *NIPS*, 2012. [3] Ajay J Joshi, Fatih Porikli, and Nikolaos Papanikolopoulos. Multi-class active learning for image classification. In *CVPR*, 2009. [4] Christoph Käding, Alexander Freytag, Erik Rodner, Paul Bodesheim, and Joachim Denzler. Active learning and discovery of object categories in the presence of unnameable instances. In *CVPR*, 2015. [5] Ronald L. Plackett. Some theorems in least squares. *Biometrika*, 1950.