Fine-grained Flower and Fungi Classification at CMP

Milan Šulc[1], Lukáš Picek[2], Jiří Matas[1]

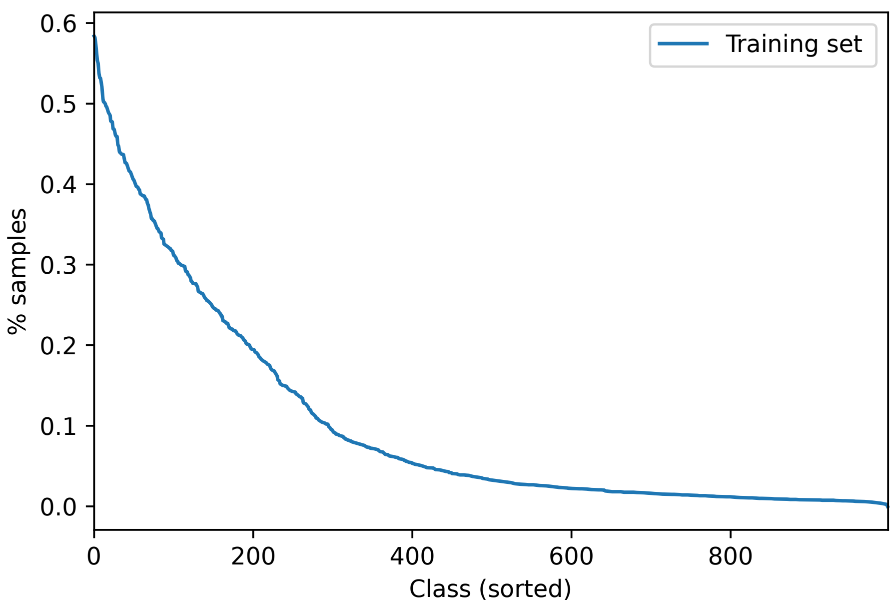
1 Visual Recognition Group, Center for Machine Perception, Dept. of Cybernetics, FEE Czech Technical University in Prague

2 Dept. of Cybernetics, FAS University of West Bohemia

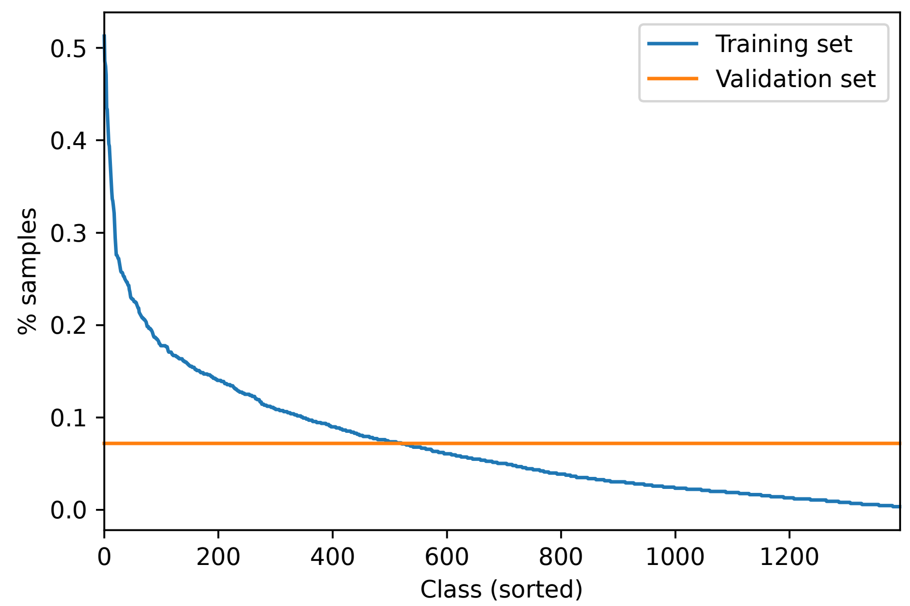
Label Distributions

FGVCx Flower and Fungi Classification datasets available for training follow a "long-tail distribution" of classes, which may not correspond with the test-time distribution.

FGVCx Flowers



FGVCx Fungi



Label Distributions

We recently observed a similar problem in the LifeCLEF plant identification challenge: majority of training data comes from the web, while test images come from a different source.

Can we compensate for this imbalance?

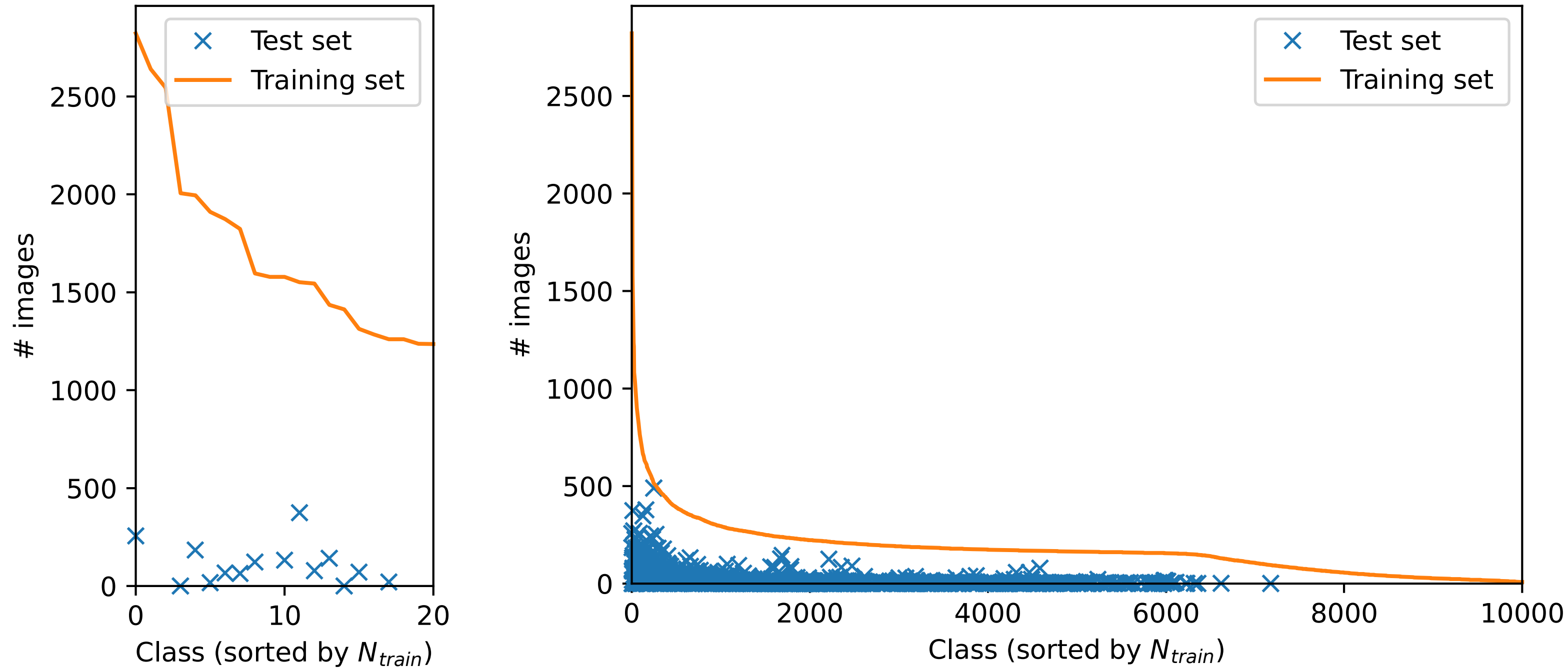


Figure: PlantCLEF 2017 label distribution in the “trusted” training set.

[1] Improving CNN classifiers by estimating test-time priors. Milan Šulc and Jiří Matas. arXiv:1805.08235 [cs.CV], 2018.

CNN Outputs as Posterior Estimates

Training neural networks (*f* with parameters θ) by cross-entropy loss minimization means training it to estimate the posterior probabilities:

$$\theta^* = \arg\min_\theta L_{\text{CE}} = \arg\max_\theta \sum_i \sum_k c_{ik} \log f(c_k|\mathbf{x}_i, \theta)$$

where $$c_{ik} = \begin{cases} 1 \text{ if } k = y_i \\ 0 \text{ otherwise} \end{cases}$$

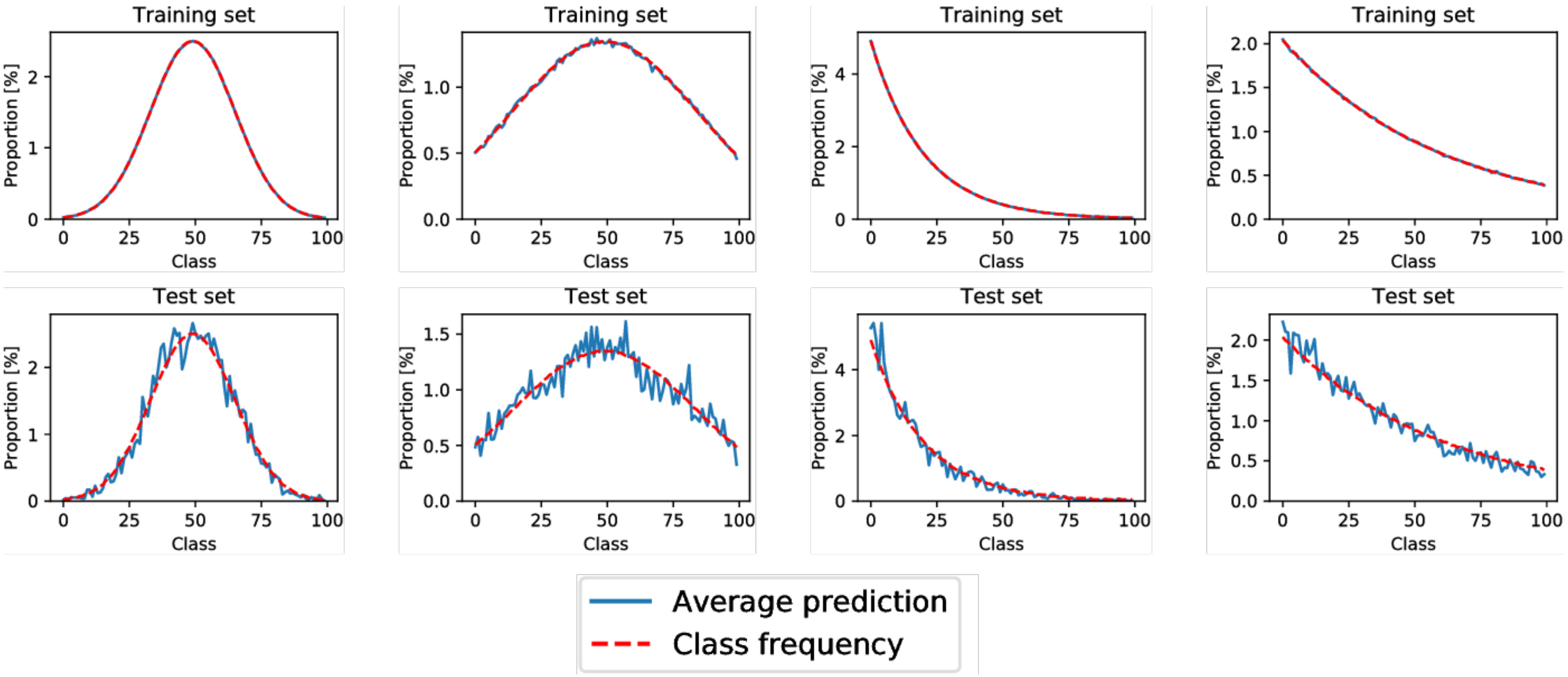
Then: $$\min_\theta L_{\text{CE}} = \min_\theta - \sum_i \log f(c_{y_i}|\mathbf{x}_i, \theta) =$$

$$= \max_\theta \prod_i f(c_{y_i}|\mathbf{x}_i, \theta) \approx \prod_i p(c_{y_i}|\mathbf{x}_i)$$

CNN Outputs as Posterior Estimates

Experiment on selected subsets of CIFAR-100 with different class priors:
How well do the posterior estimates marginalize over dataset samples?

$$\frac{1}{N}\sum_{i=1}^{N} f_{\mathrm{CNN}}(c_k|\mathbf{x}_i) \approx \frac{N_k}{N}$$



[1] Improving CNN classifiers by estimating test-time priors. Milan Šulc and Jiří Matas. arXiv:1805.08235 [cs.CV], 2018.

Adjusting Estimates to New Priors

Assuming that the probability density function $p(\mathbf{x}_i|c_k)$ remains unchanged:

$$p(\mathbf{x}_i|c_k) = \frac{p(c_k|\mathbf{x}_i)p(\mathbf{x}_i)}{p(c_k)} = p_e(\mathbf{x}_i|c_k) = \frac{p_e(c_k|\mathbf{x}_i)p_e(\mathbf{x}_i)}{p_e(c_k)}$$

The mutual relation of the posteriors is:

$$p_e(c_k|\mathbf{x}_i) = p(c_k|\mathbf{x}_i)\frac{p_e(c_k)p(\mathbf{x}_i)}{p(c_k)p_e(\mathbf{x}_i)} = \frac{p(c_k|\mathbf{x}_i)\dfrac{p_e(c_k)}{p(c_k)}}{\sum_{j=1}^{K} p(c_j|\mathbf{x}_i)\dfrac{p_e(c_j)}{p(c_j)}} \boxed{\propto p(c_k|\mathbf{x}_i)\frac{p_e(c_k)}{p(c_k)}}$$

When Test Set Priors Are Unknown

How to estimate the test-set priors?

Saerens et al. [1] proposed a simple EM procedure to maximize the likelihood *L(**x**$_0$,**x**$_1$,**x**$_2$,...)*:

$$p_e^{(s)}(c_k|\mathbf{x}_i) = \frac{p(c_k|\mathbf{x}_i)\dfrac{p_e^{(s)}(c_k)}{p(c_k)}}{\sum_{j=1}^{K} p(c_j|\mathbf{x}_i)\dfrac{p_e^{(s)}(c_j)}{p(c_j)}}$$

$$p_e^{(s+1)}(c_k) = \frac{1}{N}\sum_{i=1}^{N} p_e^{(s)}(c_k|\mathbf{x}_i)$$

This procedure is equivalent [2] to fixed-point-iteration minimization of the KL divergence between $p_e(\mathbf{x})$ and $q_e(\mathbf{x}) = \sum_{k=1}^{K} P_k p(\mathbf{x}|c_k)$.

[1] Adjusting the outputs of a classifier to new a priori probabilities: a simple procedure. Marco Saerens, Patrice Latinne, and Christine Decaestecker. Neural computation 14.1 (2002): 21-41.

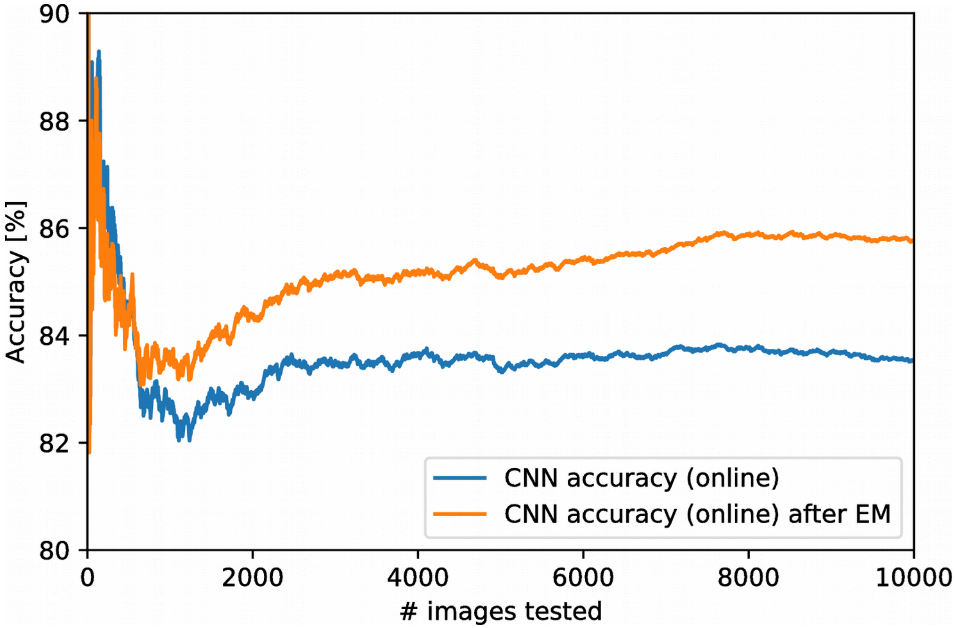
[2] Semi-supervised learning of class balance under class-prior change by distribution matching. Marthinus Christoffel Du Plessis and Masashi Sugiyama. Neural Networks, 50:110–119, 2014.

Test Set Prior Estimation in LifeCLEF

Preliminary experiments (using the 2017 test set for validation):

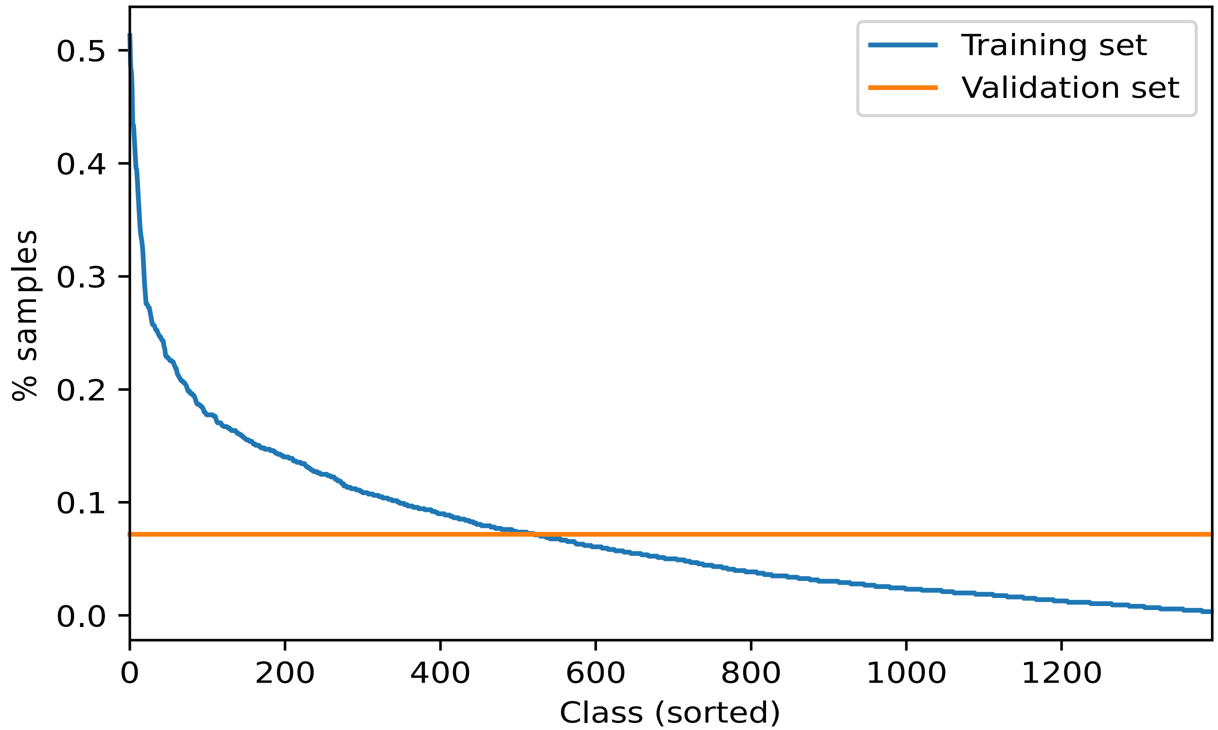
- When the whole test set is available:
  Inception-ResNet-v2: **82.9% → 85.8%**
  Inception-v4: **82.8% → 86.3%**

- On-line [1] after each new test image:



[1] Improving CNN classifiers by estimating test-time priors. Milan Šulc and Jiří Matas. arXiv:1805.08235 [cs.CV], 2018.

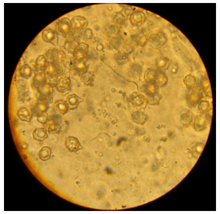
When New Priors Are Known



FGVCx Fungi 2018





When New Priors Are Known

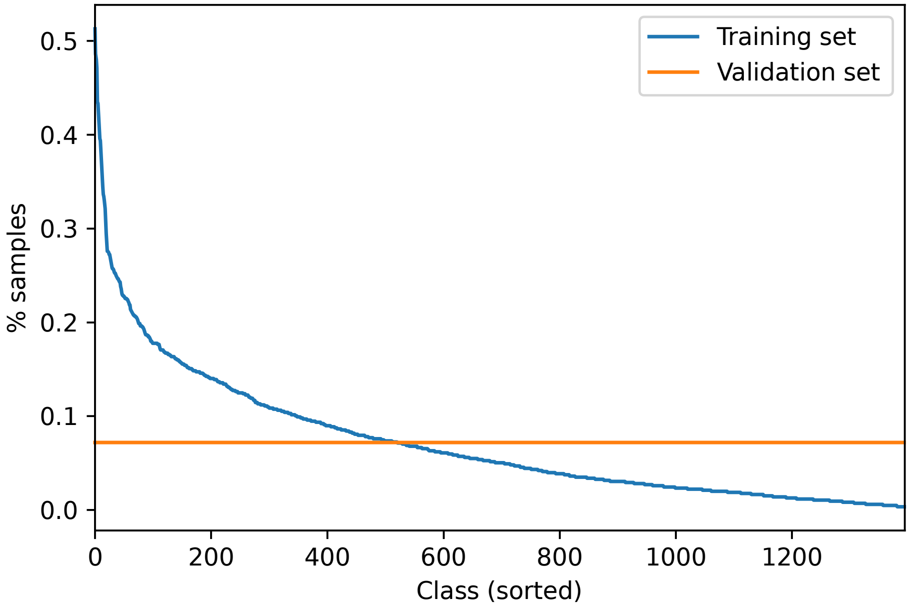
Note: in the iNaturalist 2017 challenge, the winning GMV submission [1] approached the change in priors as follows:

“To compensate for the imbalanced training data, the models were further **fine-tuned on** the 90% subset of the **validation** data that has a more balanced distribution.”

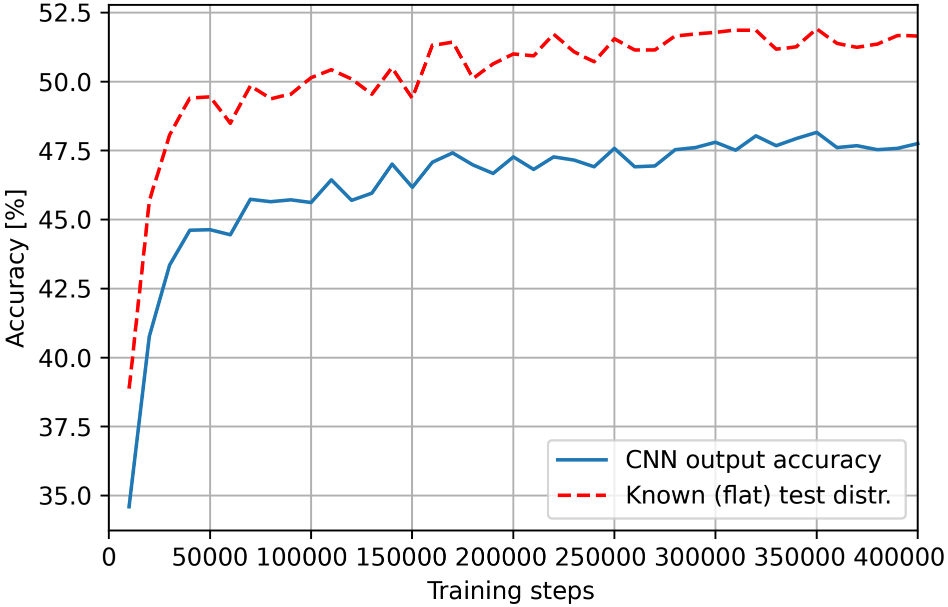
We, instead, only use the **validation set statistics** – i.e. uniform class distribution in this case.

[1] The iNaturalist Species Classification and Detection Dataset-Supplementary Material. Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. Reptilia 32, no. 400: 5426.

When New Priors Are Known



FGVCx Fungi 2018



Inception v4

[1] Improving CNN classifiers by estimating test-time priors. Milan Šulc and Jiří Matas. arXiv:1805.08235 [cs.CV], 2018.

Tricks used in both challenges

Predictions re-weighted simply assuming uniform class priors.

Moving average of trained variables (exponential decay).

Training time augmentation:
- Random crops
- Color distortions

Test-time data augmentation:
14× per image :   7 crops ×2 (mirror)

Final Ensembles

FGVCx Fungi: 6 nets (averaged)

2x Inception-v4 299x299 initialized from ImageNet and LifeCLEF ckpts
2x Inception-v4 598x598 initialized from ImageNet and LifeCLEF ckpts
2x Inception-ResNet-v2 299x299 from ImageNet and LifeCLEF ckpts

FGVCx Flowers: 5 nets (modus)

3x Inception-v4 299x299 initialized from ImageNet, LifeCLEF, iNaturalist ckpts
1x Inception-v4 598x598 initialized from LifeCLEF ckpt
1x Inception-ResNet-v2 299x299 initialized from LifeCLEF ckpt

Leaderboard

FGVCx Fungi

#	△pub	Team Name	Kernel	Team Members	Score
1	—	CMP			0.21197
2	—	digitalspecialists			0.23188
3	—	Val An			0.25091
4	—	Deep Learning Analytics			0.28341
5	—	Invincibles			0.28751

FGVCx Flowers

#	△pub	Team Name	Kernel	Team Members	Score
1	—	CMP			0.07599
2	▲ 2	fadivugibs			0.08177
3	▼ 1	DLUT_VLG (Dalian University ...			0.08242
4	▼ 1	yen			0.08396
5	—	xiaoxiao			0.09579

Discussion

- Standard CNN classifiers (and their ensembles) achieve best results in plant and fungi recognition.
  - Future work: Learning from Ensembles?

- Important to take into account change in class prior distribution [1]
  - New priors can be estimated on-line, as new test-samples appear.

- Q & A

sulcmila@cmp.felk.cvut.cz

[1] Improving CNN classifiers by estimating test-time priors. Milan Šulc and Jiří Matas. arXiv:1805.08235 [cs.CV], 2018.