A Flexible Selection Scheme for Minimum-Effort Transfer Learning

Amélie Royer, Christoph H. Lampert
IST Austria
WACV 2020

Transfer Learning and Fine-tuning

Finetuning: - Pre-trained model on task T : X ⟶ Y

- Retrain last layer(s) for task T' : X ⟶ Y' with supervision

Hypothesis: Generality of early features for similar datasets

Example: ImageNet classification to Pascal detection

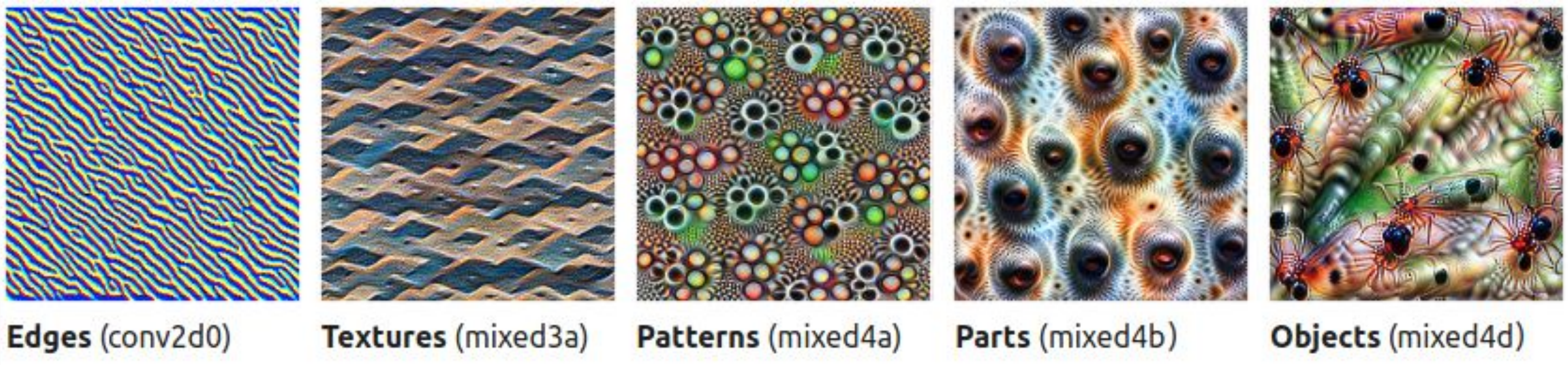
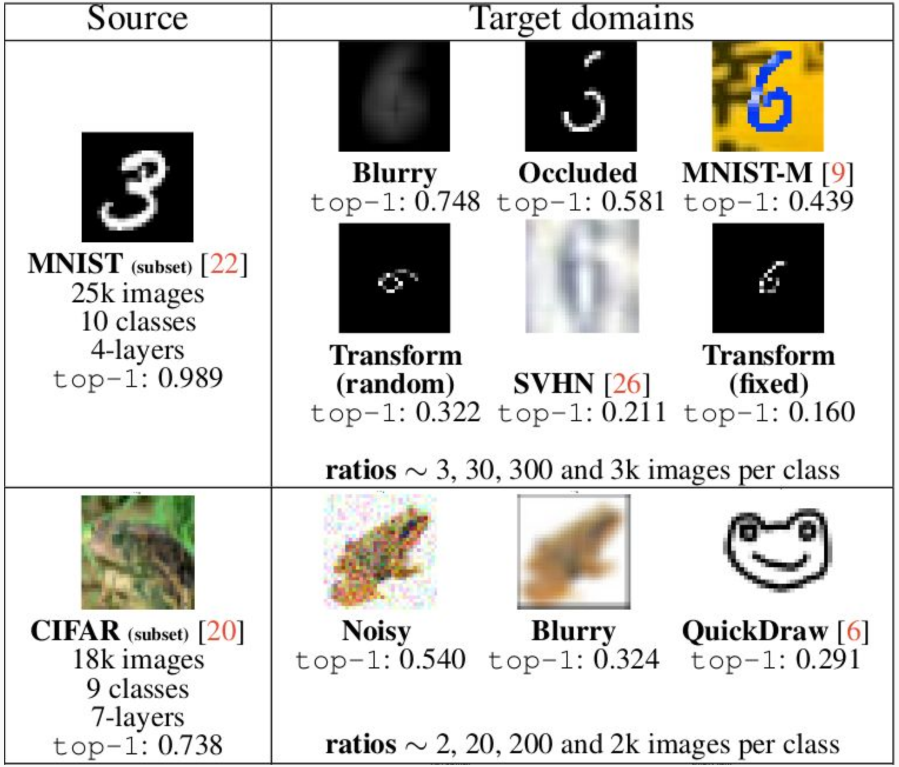


Image from "Feature Visualization", C. Olah et al, Distill, 2017

How to handle visually dissimilar domains ?



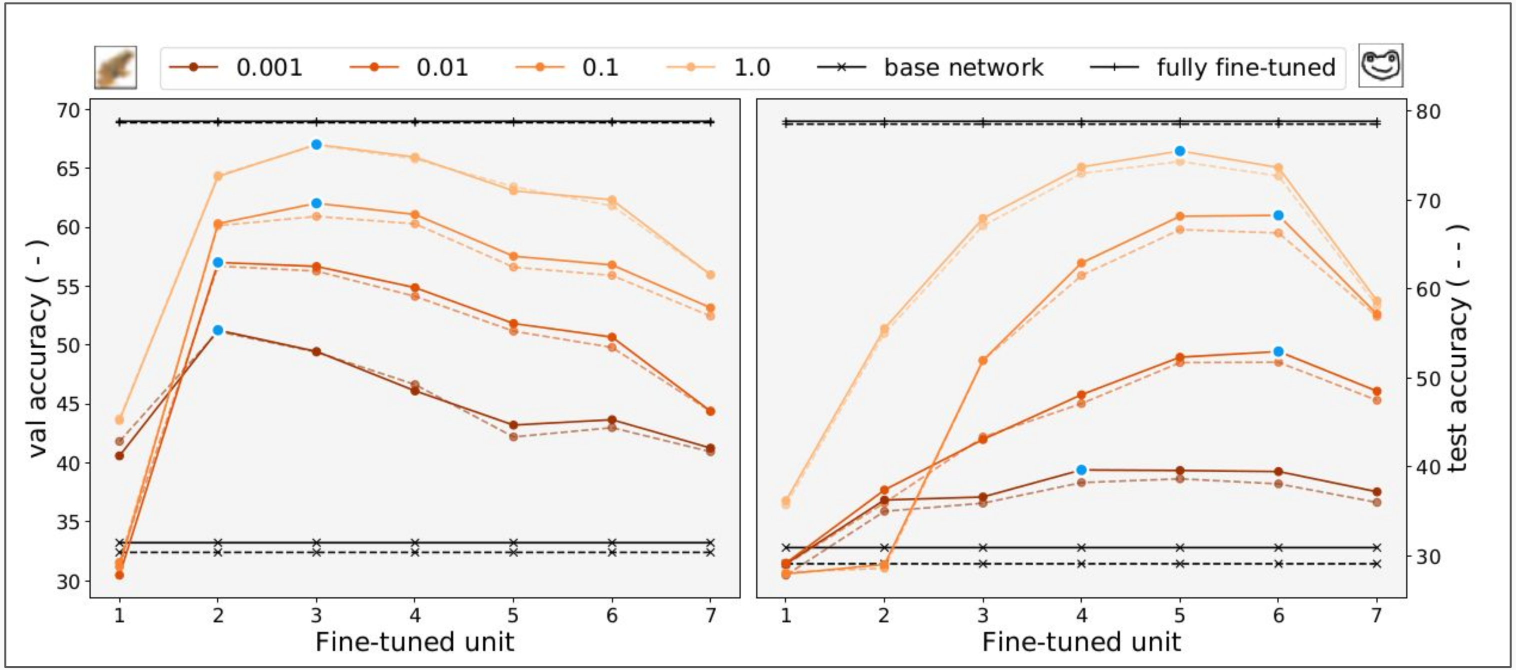
Assumption of similar early features break.
<u>but</u> fine-tuning the whole architecture is not always an option
(e.g., lack of data, risk of overfitting)

Yet, similar semantics (output space)

Can we learn to mend the representation by tuning one intermediate unit (layer of block of layers) in the architecture ?

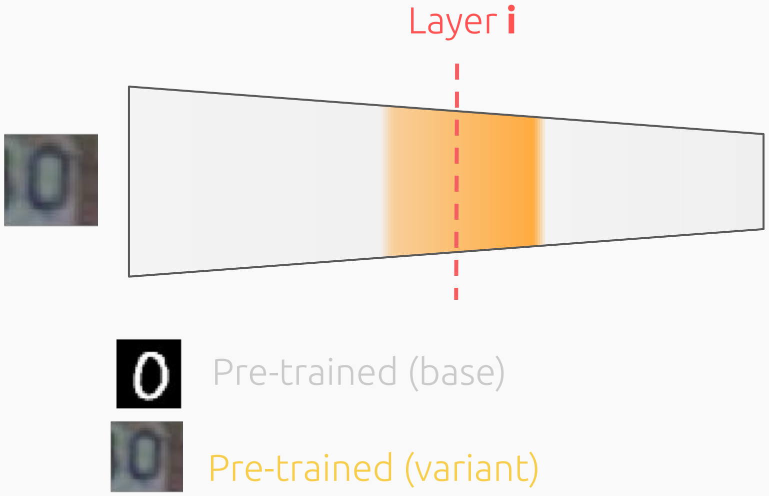
Flextuning: A model selection criterion

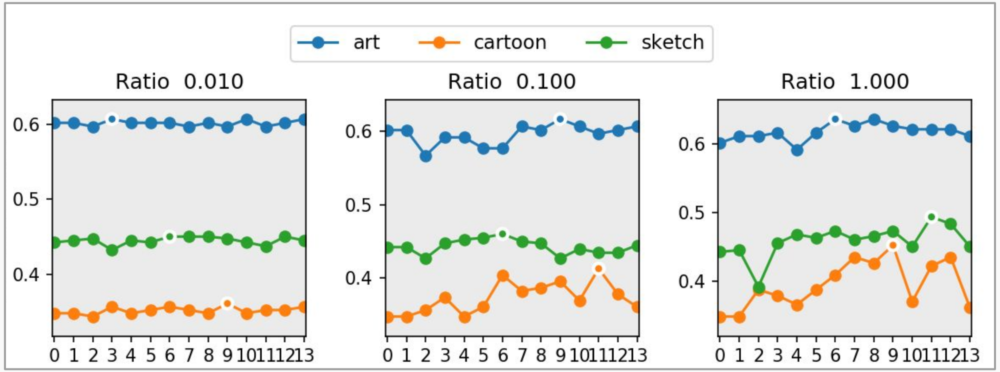
Evaluate best unit to tune based on accuracy on the validation set



Flextuning: Faster variants

Approximate how each unit will impact the end model accuracy by network “surgery” after several (fast-flex) or one (faster-flex) training epoch

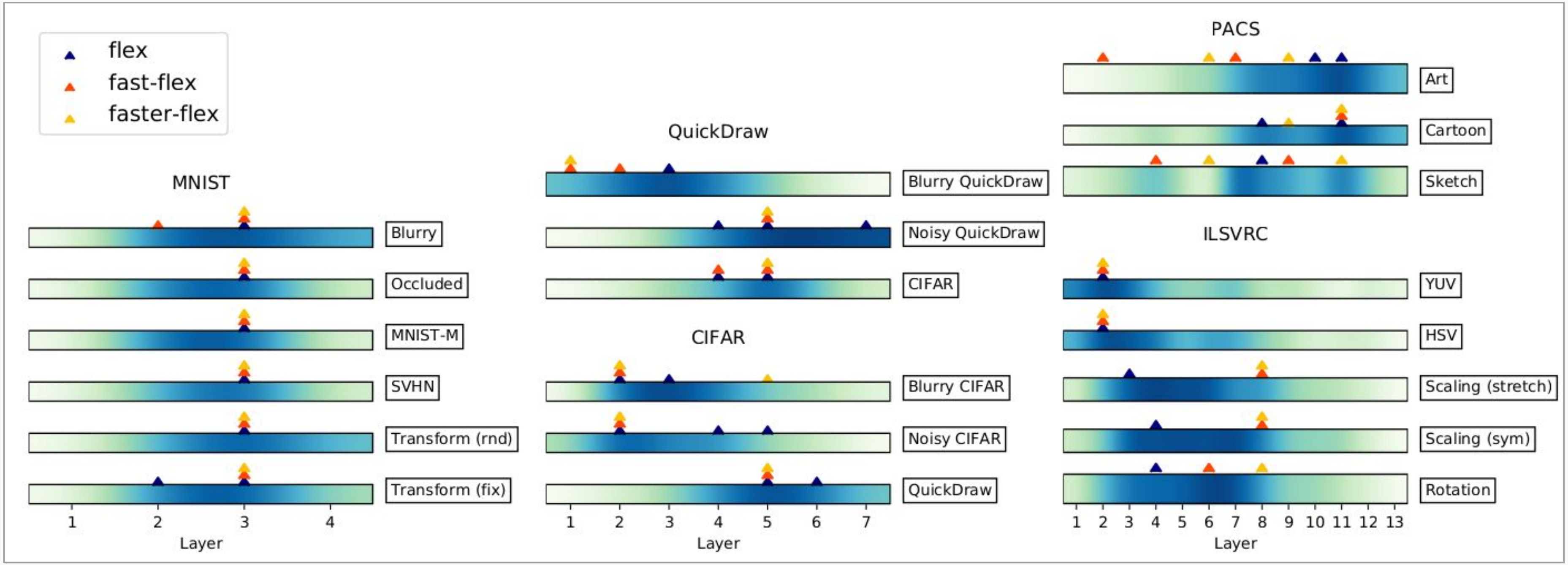




Qualitative results

Agreement across the different criteria

Selection pattern dependent on the domain shift



Quantitative results

Experiments across different visual **domain shifts** and **data scarcity** scenarios.

- On large architectures, less prone to overfitting.

- On local distorsions, improved accuracies on the target set by selecting early units

ILSVRC	flex			ft-		
	flex	fast	faster	fc	ss	all
YUV (0.84)	**0.893**	**0.893**	**0.893**	0.835	0.699	0.808
HSV (0.38)	**0.856**	**0.856**	**0.856**	0.533	0.646	0.687
Scaling (stretch) (0.44)	**0.724**	**0.696**	**0.696**	0.502	0.584	0.653
Scaling (sym.) (0.52)	**0.770**	**0.757**	**0.757**	0.663	0.650	0.716
Rotation (0.74)	**0.826**	**0.832**	**0.812**	0.667	0.652	0.771

ILSVRC	flex			ft-		
	flex	fast	faster	fc	ss	all
ratio: 200 images per class						
Art (0.53)	**0.906**	**0.906**	0.823	0.791	0.887	0.746
Cartoon (0.32)	**0.958**	0.956	0.952	0.868	0.956	0.925
Sketch (0.14)	**0.924**	**0.924**	0.890	0.767	0.916	0.875