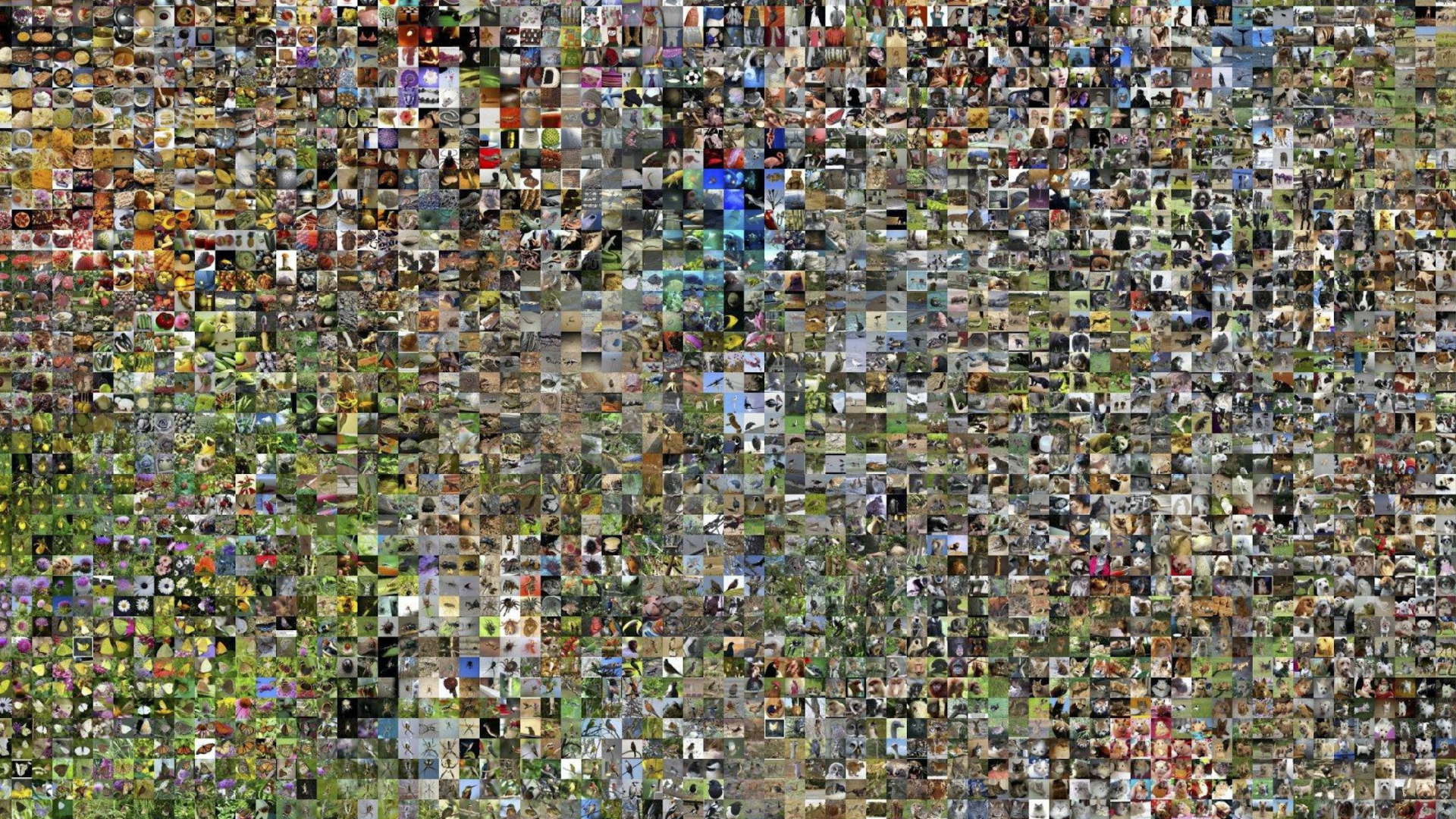
Deep Learning for ECE
EECE-580G

Transfer Learning

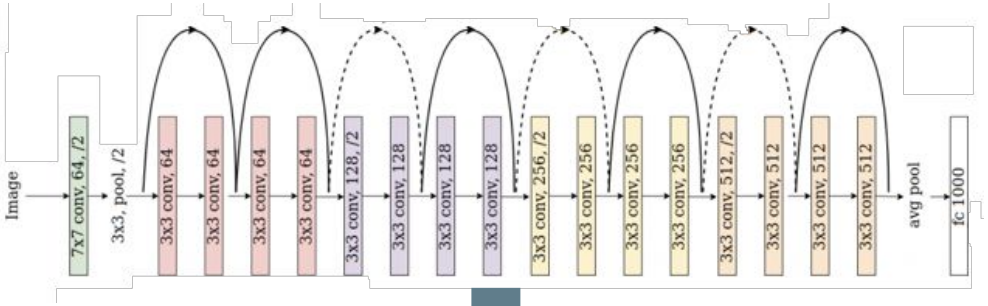


ImageNet

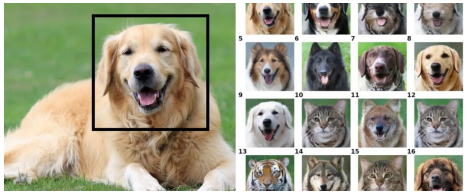
- > 1 Million images, 1,000 classes
- Hierarchical structure
- One of the most diverse image datasets
- Most computer vision research is benchmarked on ImageNet
- Benchmark
- Pretrained weights:
  - https://github.com/qubvel/classification_models
  - https://github.com/qubvel/efficientnet
  - https://www.tensorflow.org/api_docs/python/tf/keras/applications
  - ... github repos

Transfer learning

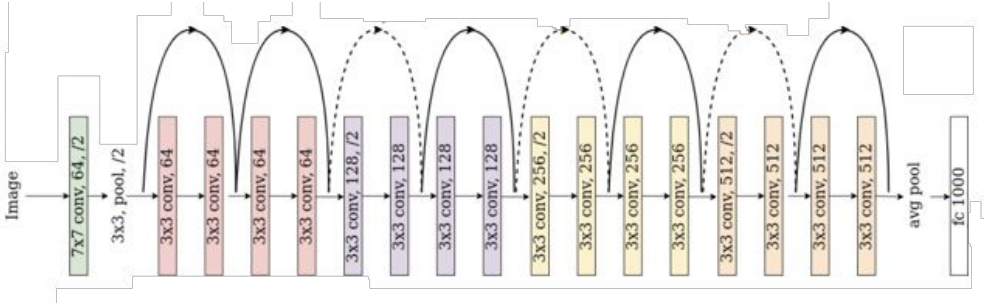
Train



Reuse weights

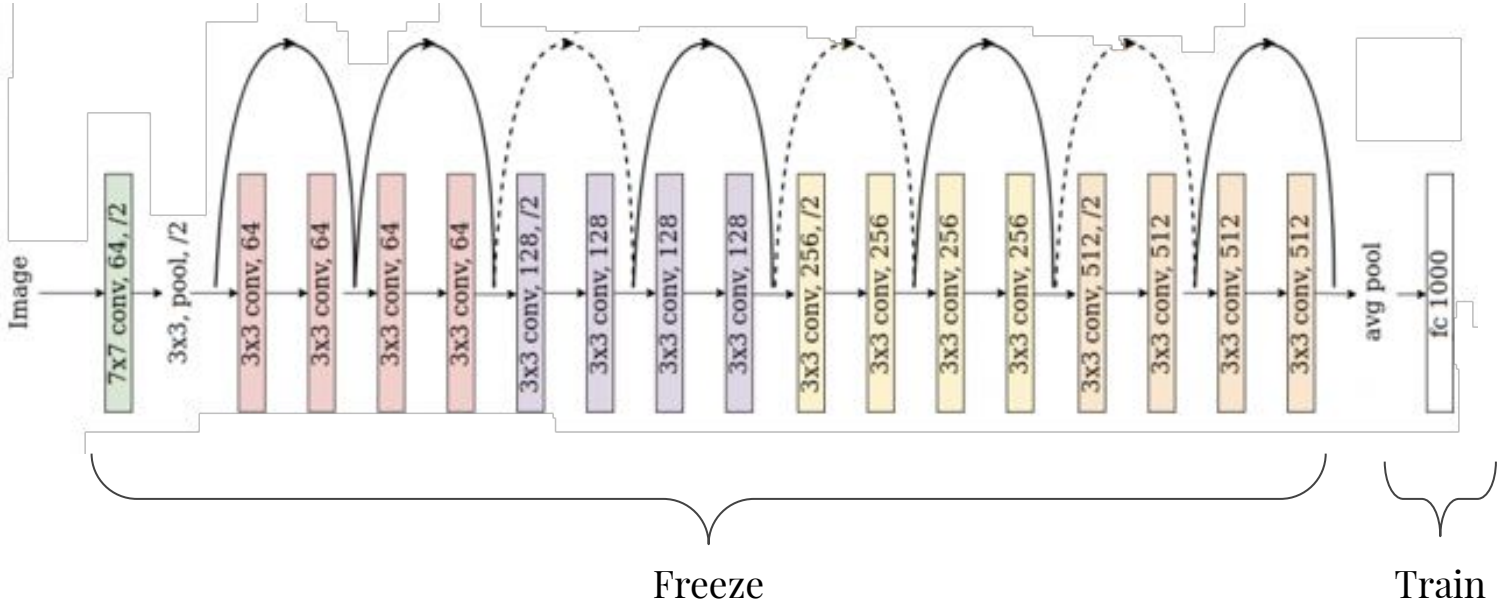


Fine-tune

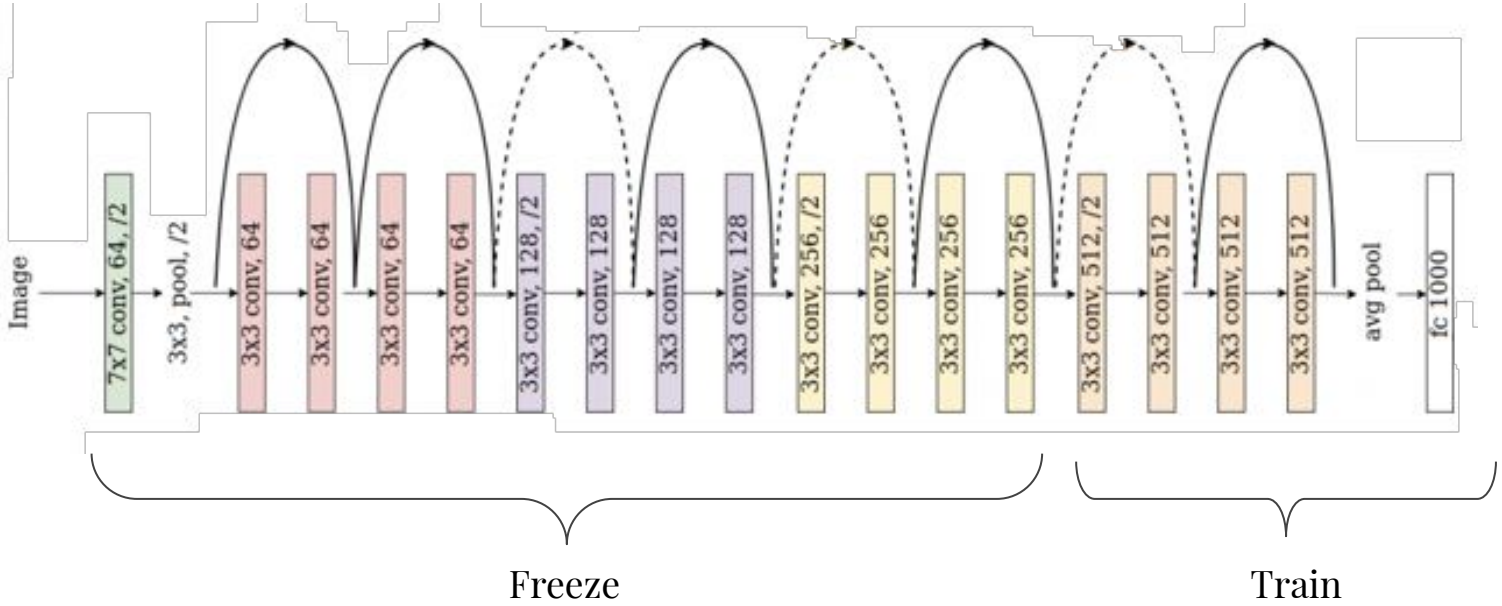


Multi-stage Transfer learning

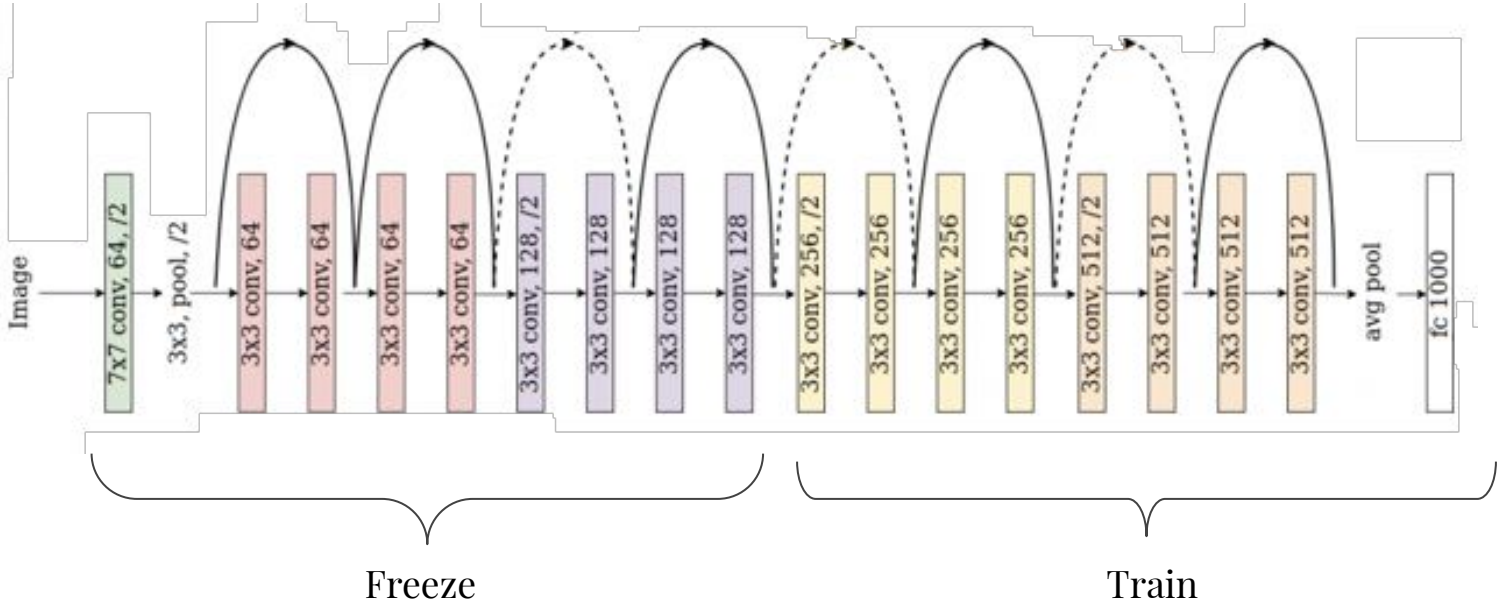
Howard, Jeremy, and Sebastian Ruder. "Universal language model fine-tuning for text classification." arXiv preprint arXiv:1801.06146 (2018).



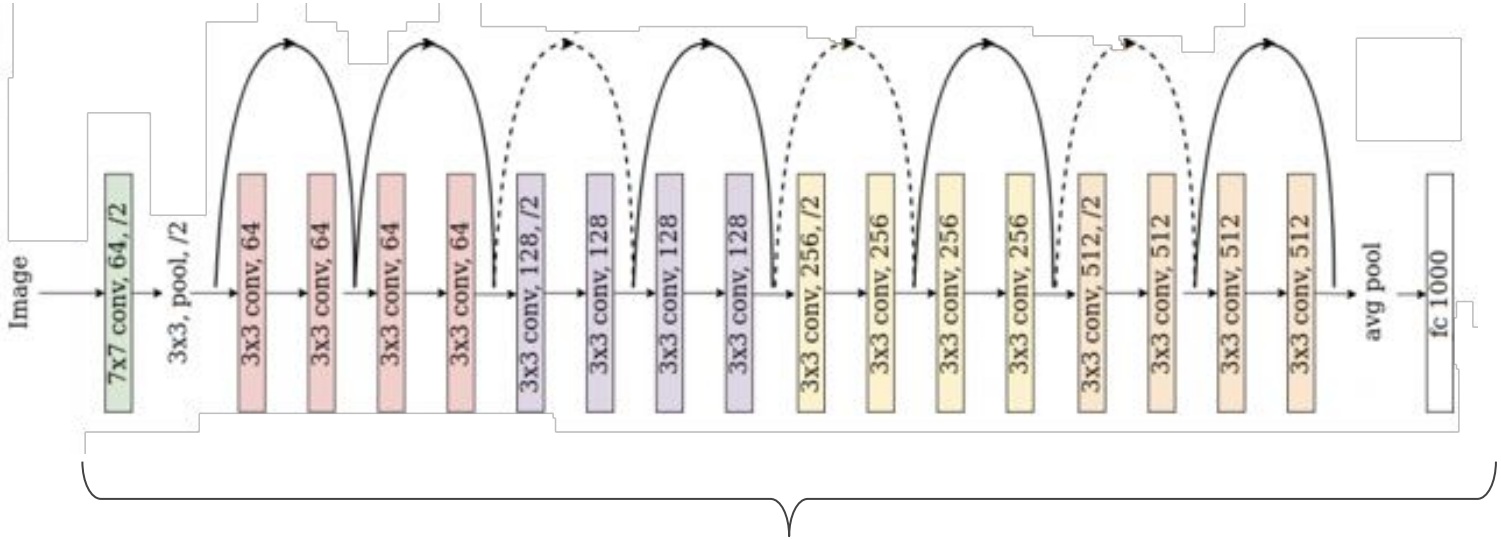
Multi-stage Transfer learning



Multi-stage Transfer learning

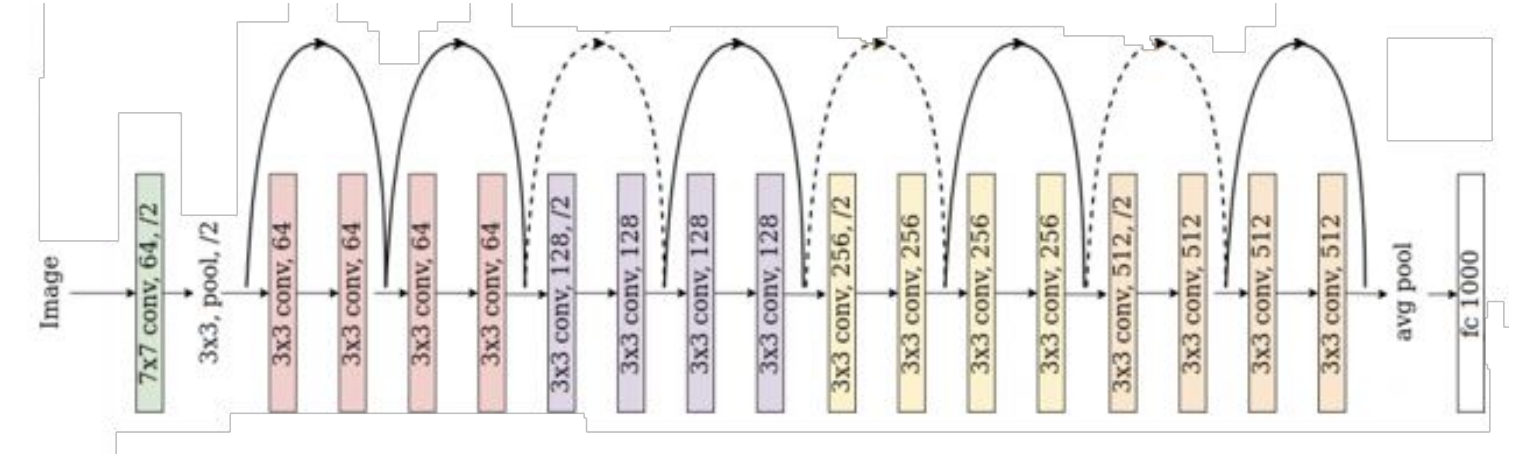


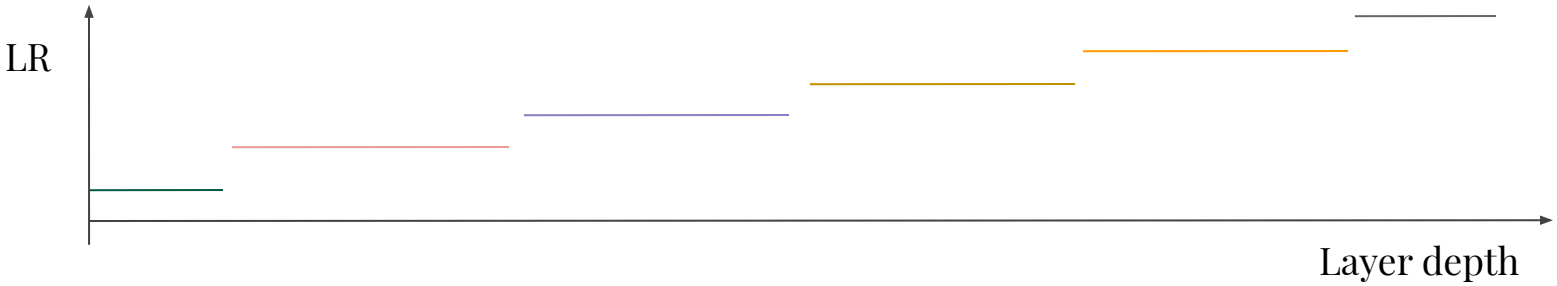
Multi-stage Transfer learning



Differential (discriminative) Learning Rate

Howard, Jeremy, and Sebastian Ruder. "Universal language model fine-tuning for text classification." arXiv preprint arXiv:1801.06146 (2018).





Progressive resizing

64x64

Train

128x128

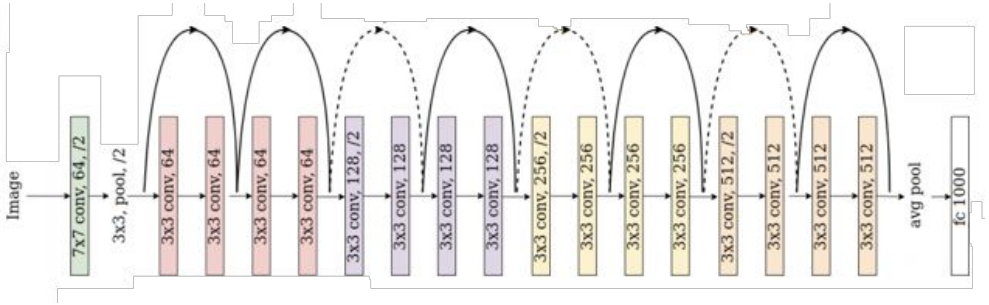
Fine-tune

Jeremy Howard. Fastai - progressive resizing. https://www.fast.ai/2018/04/30/dawnbench-fastai/, 2018.

Progressive resizing

256x256

Fine-tune

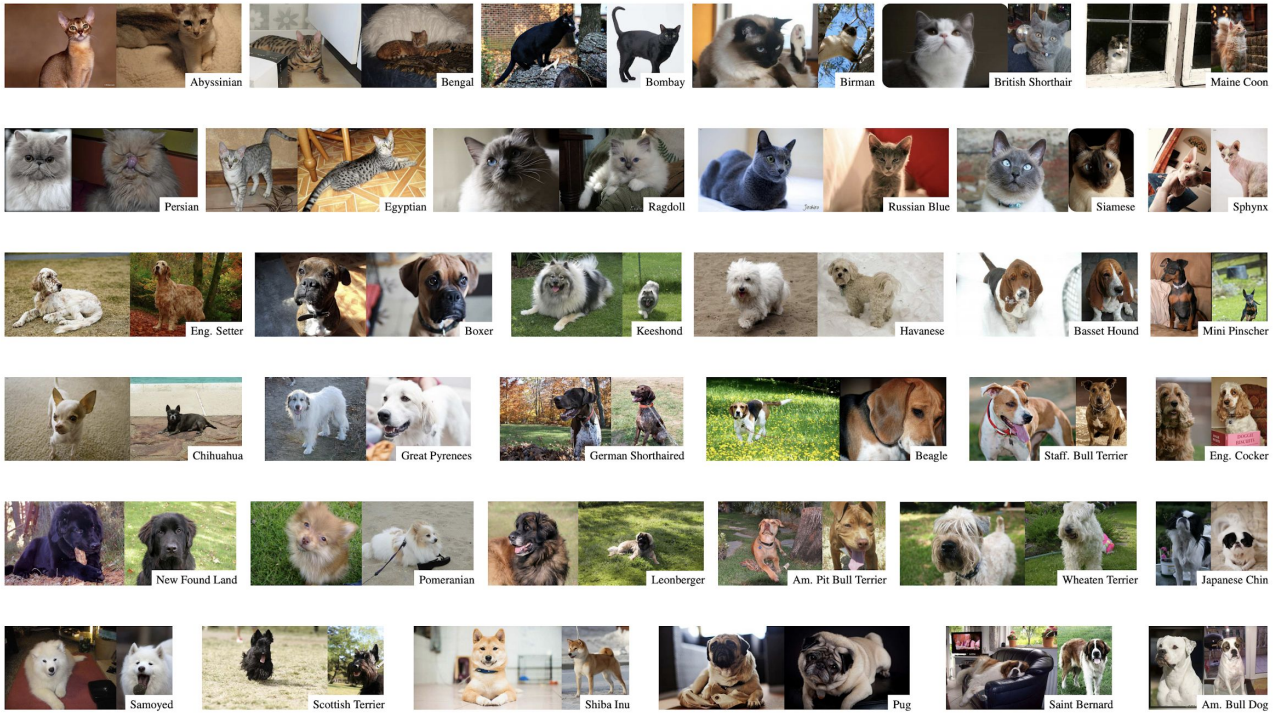


...

Example

- Oxford pets dataset
- Many classes, not enough samples per class
- This is usually the case in DL
- Need data augmentation!

>> Colab



Source: kaggle

Conclusion

In practice

- Transfer learning is very effective = usually the **norm in Deep Learning**
- Maybe not very useful when large dataset
- Still very heuristic and manual procedures…
- **Progressive unfreezing** + Some **LR** tricks (differential/schedules/…) seem to work well

Research

- Need more research related to Transfer Learning
- Works even when the task is very different from ImageNet classification! (e.g. forensics, steganalysis, etc.)

End