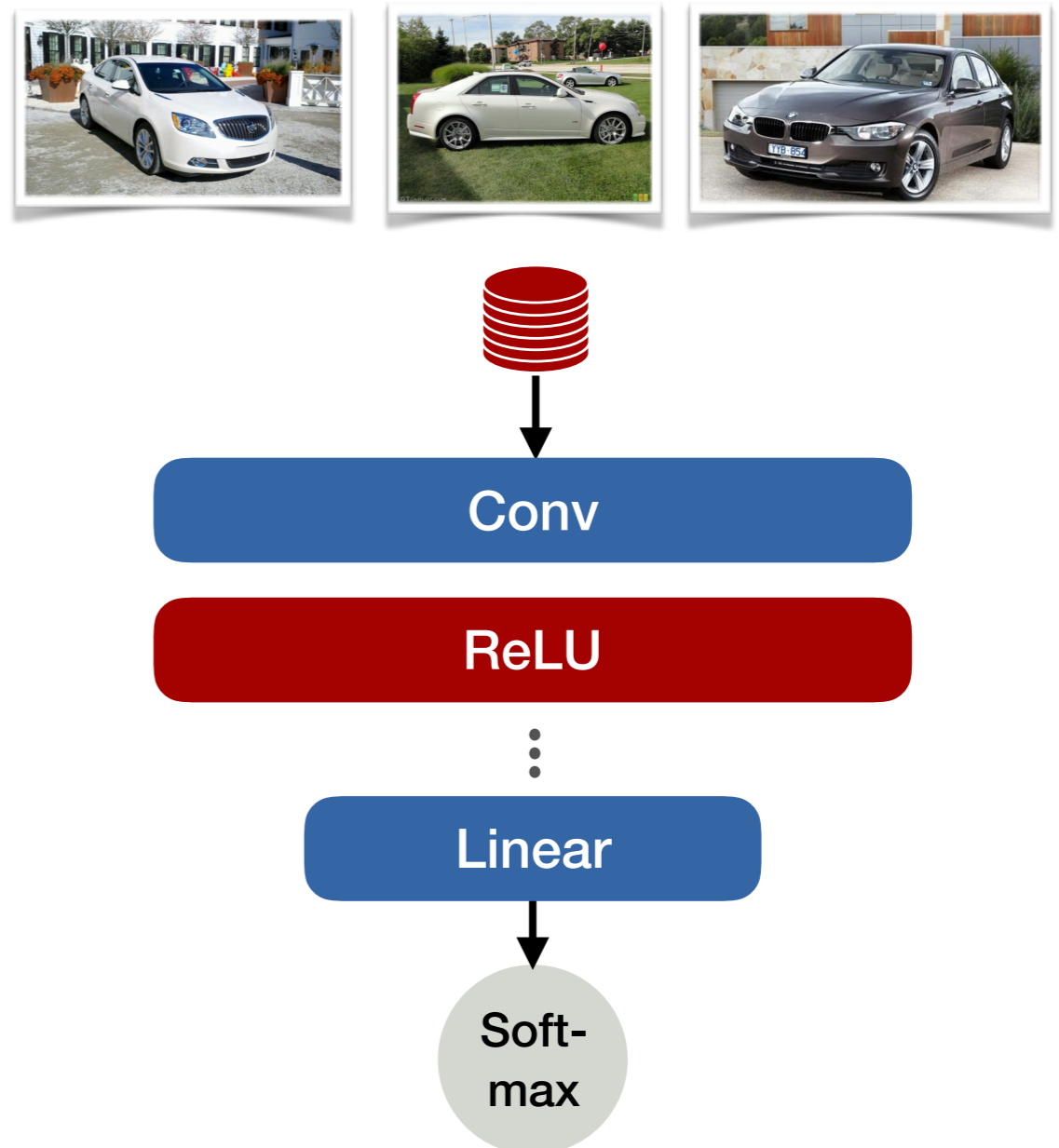


# Transfer learning

© 2019 Philipp Krähenbühl and Chao-Yuan Wu

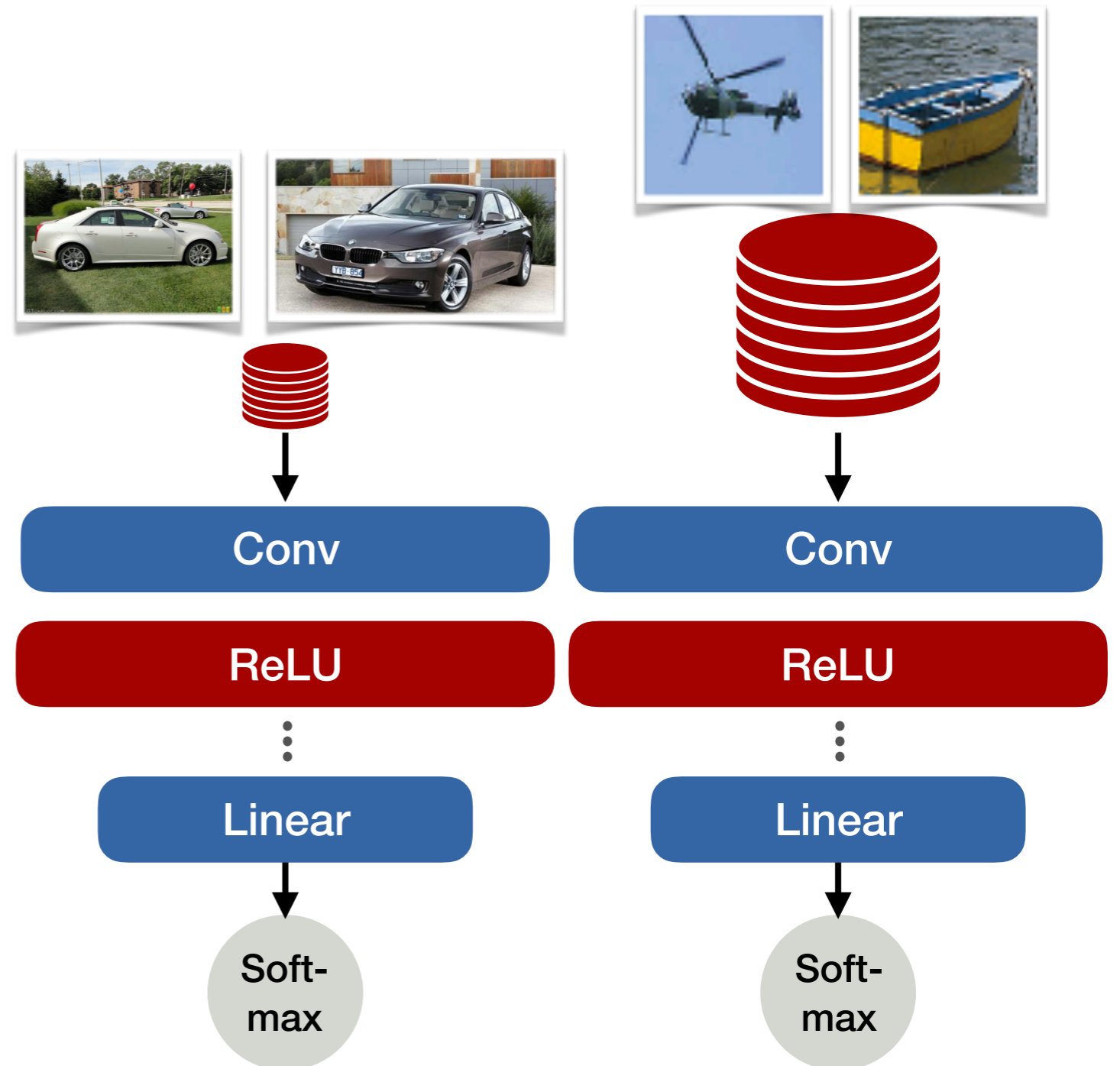
# Training on small datasets

- How to training
  - a large model
  - on a small dataset?



# Solution: Pre-training / fine-tuning

- Train model on large dataset (Pre-training)
  - on related task
- Copy model
- Continue training on small dataset (Fine-tuning)



# Pre-training

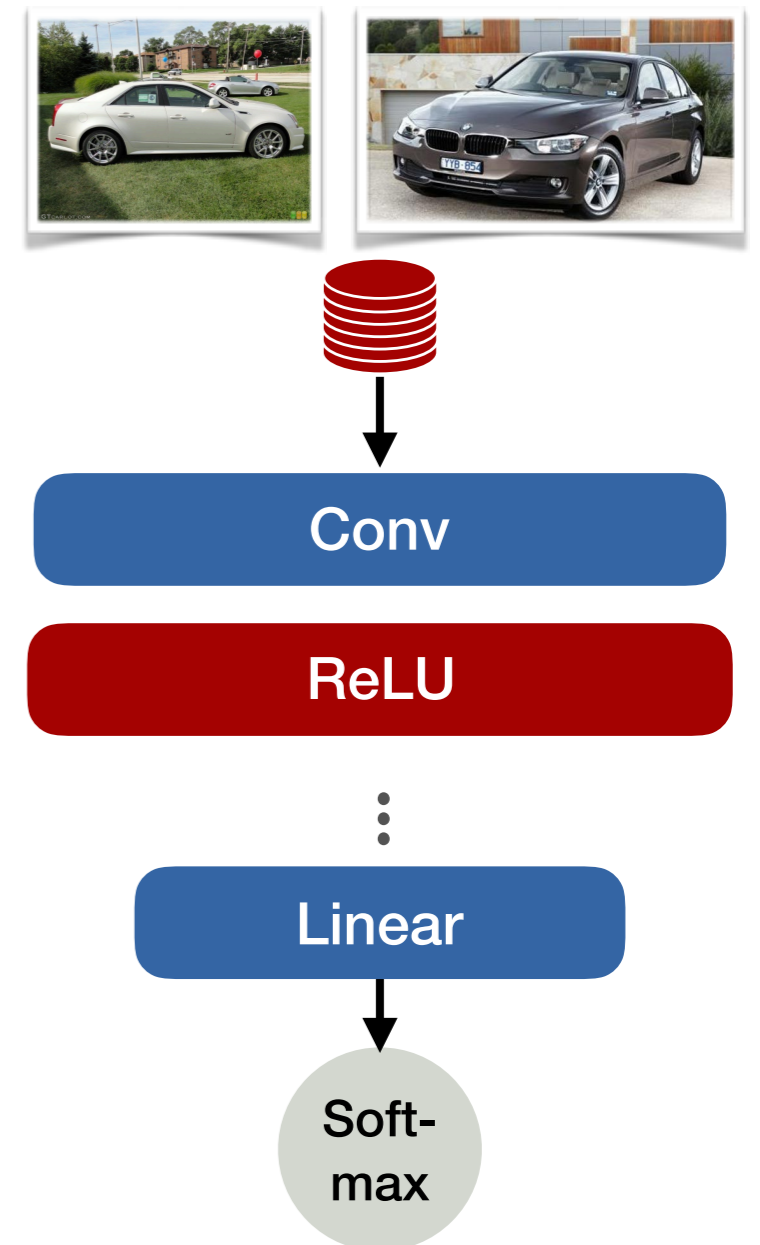
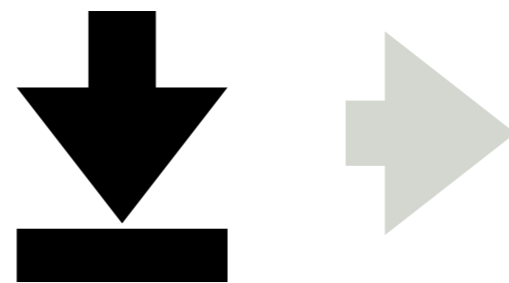
- Computer vision
  - Supervised
    - e.g. ImageNet
- Natural Language Processing
  - Self-supervised
    - Unlabeled text



**WIKIPEDIA**  
The Free Encyclopedia

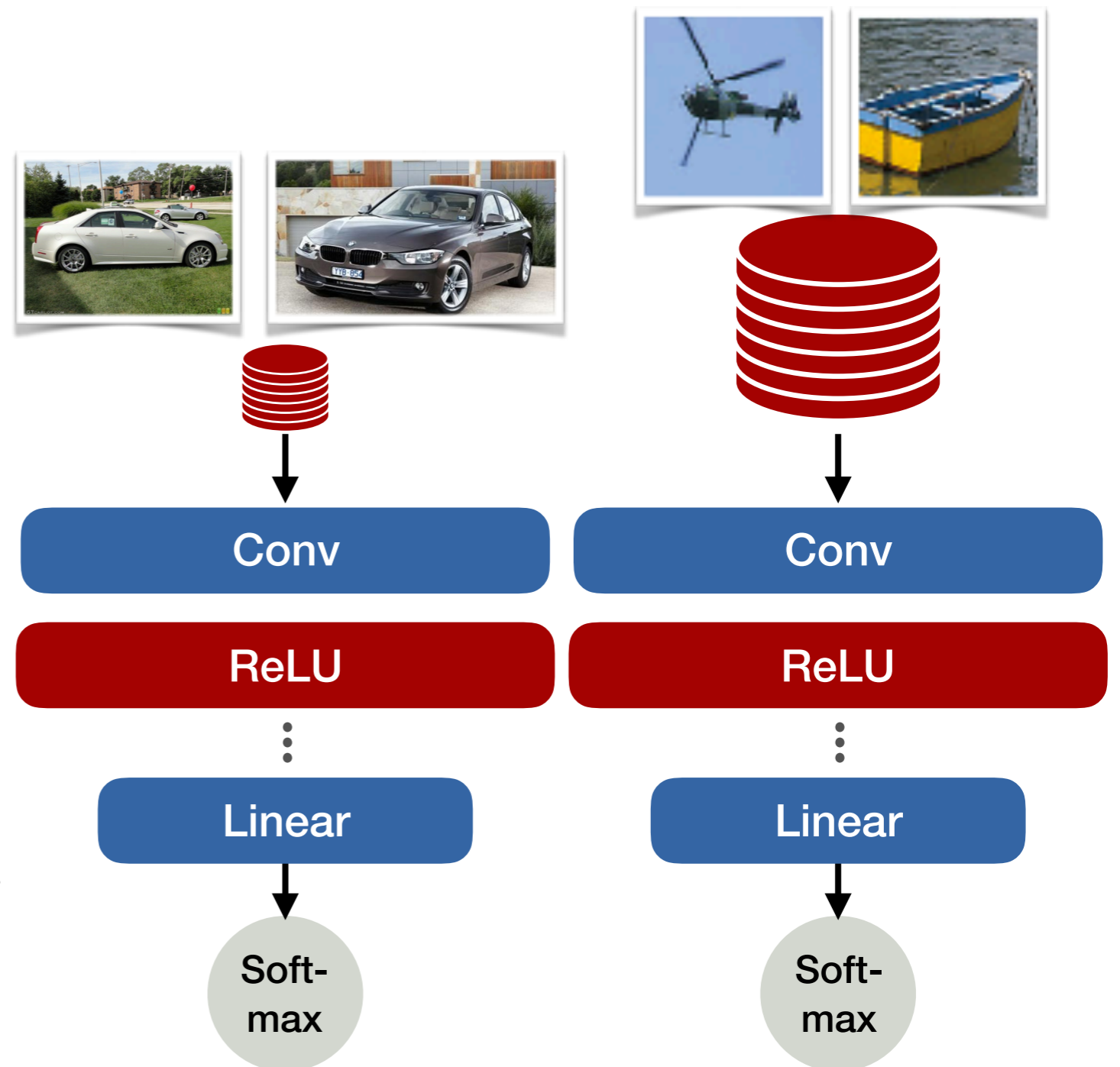
# Pre-training / fine-tuning in practice

- Download a pre-trained model
  - Model-zoo
  - Directly for author
- Run a few training iterations on small dataset



# Why does transfer learning work?

- Similar inputs
  - e.g. Images, Text, ...
  - Transfer between tasks
- Good initialization
  - Learned weights are tuned well



# When to use transfer learning?

- Whenever possible
  - In early experiments
  - Large pre-trained model exists