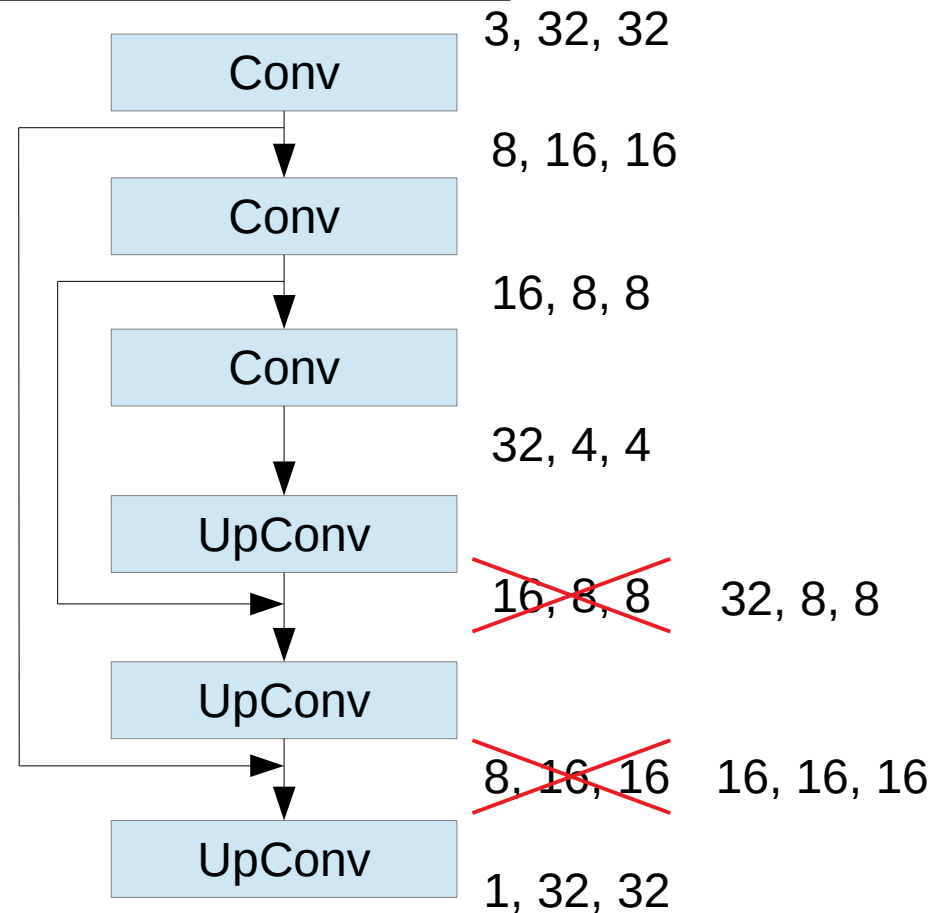
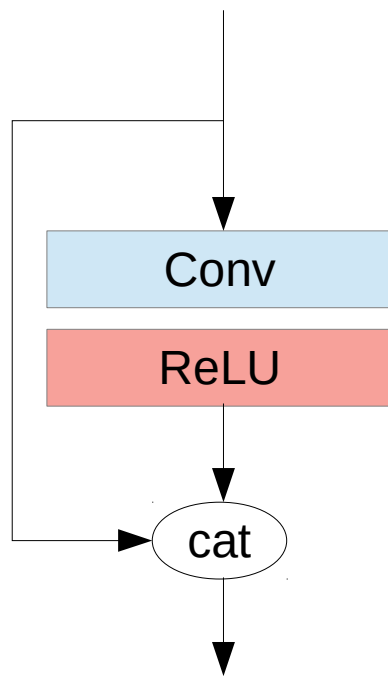
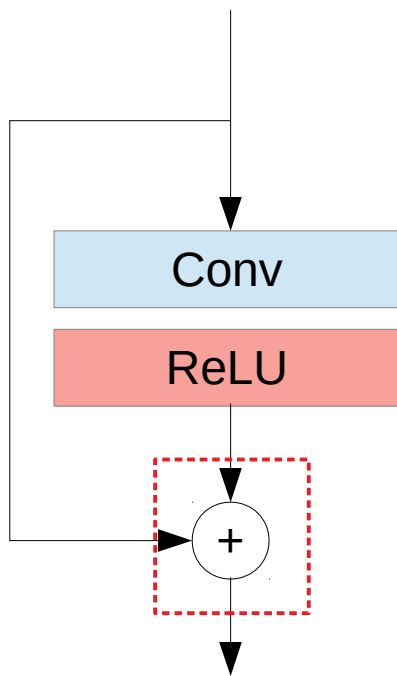




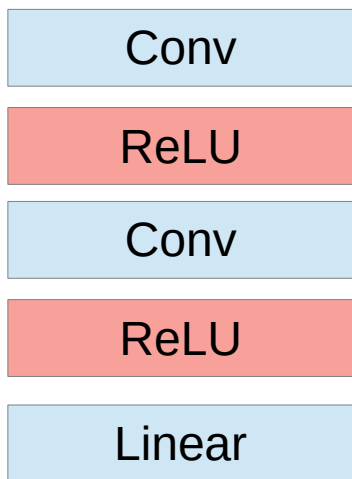
# Transfer Learning

# Aside: Skip Connections



# Transfer Learning

Small dataset:



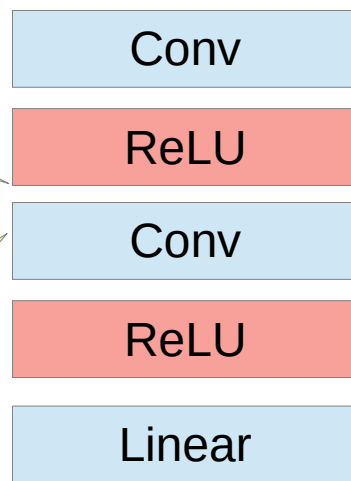
Overfits

Large dataset:



Different, related task

Pre-training

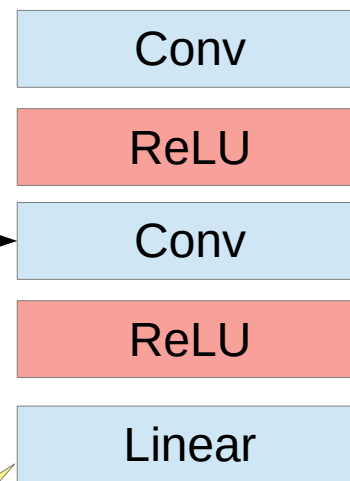


Overfits less

Original dataset:



Copy weights



Train just a little

Fine tuning

# Why does Transfer Learning Work?

Similar datasets – similar features



- Edges
- Colors
- Basic shapes
- Textures
- ...



- Edges
- Colors
- Basic shapes
- Textures
- ...

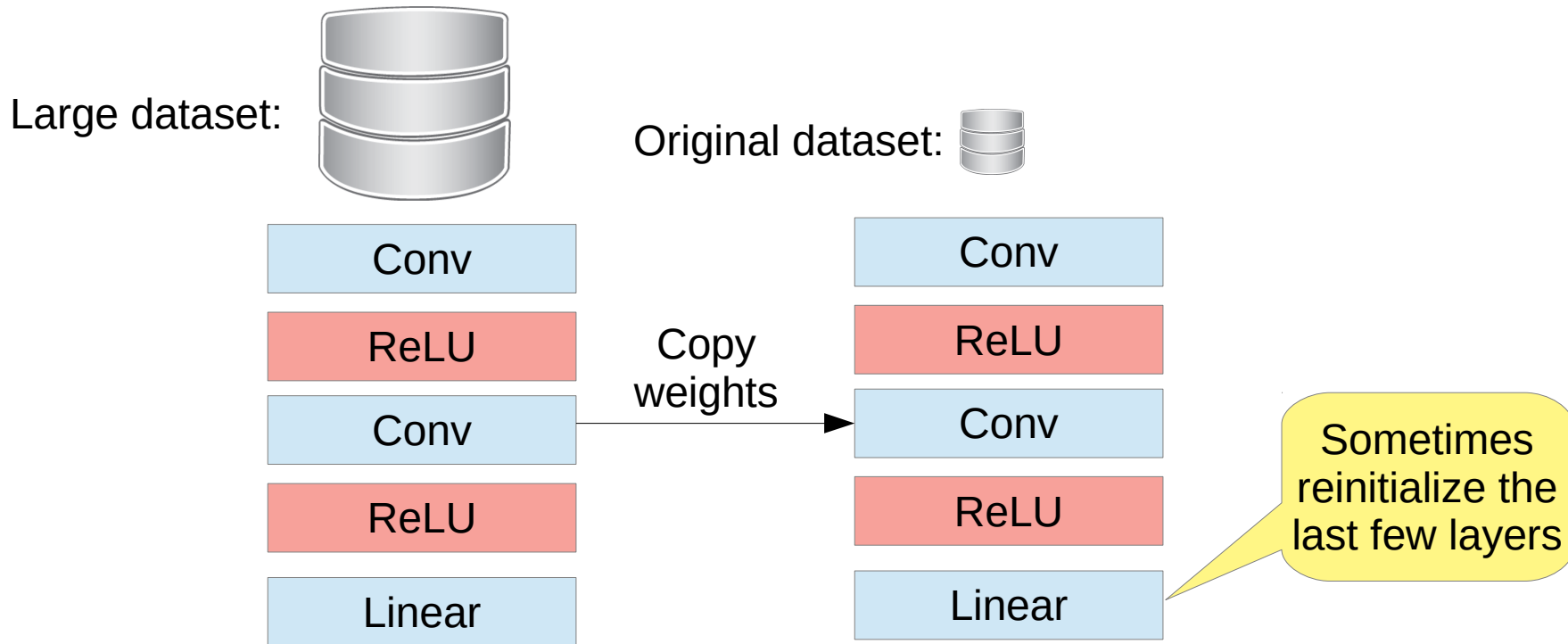
# Pre-training

- Vision: ImageNet
- Natural Language Processing
  - Self-Supervision
  - GPT3
- Model zoo
  - `torch.utils.model_zoo`

IMAGENET



# Fine Tuning



# When to use Transfer Learning

- Whenever you can
  - If a pre-trained model exists
  - Especially in early experiments
- In this class: don't use pretrained models
  - You can look at them for early experiments if you want
  - Your final models should be trained from scratch