Data vs. hand-engineering

Little data ←————————————————————————————————→ Lots of data

More hand-engineering ("hacks")

↑ Transfer learning

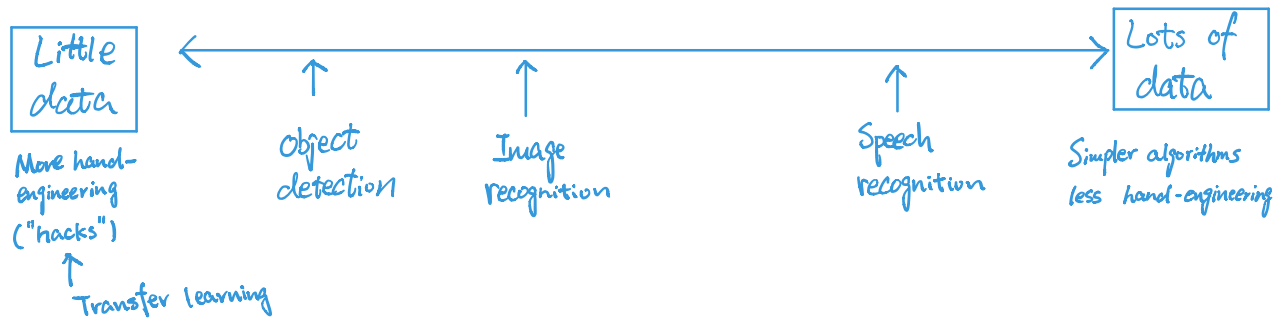↑ Object detection   ↑ Image recognition   ↑ Speech recognition

Simpler algorithms less hand-engineering

Two sources of knowledge
- Labeled data (x, y)
- Hand engineering features/network architecture/other components

Tips for doing well on benchmark/winning competitions   (Useless in pratice)
- Ensembling   3-15 networks   — $\hat{y}$
  ‣ Train several networks independently and average their outputs

- Multi-crop at test time
  ‣ Run classifier on multiple versions of test images and average results

Use open source code
- Use architectures of networks published in the literature

- Use open source implementation if possible

- Use pretrained models and fine-tune on your dataset