Convolutional Neural Networks DeepLearning AI

Practical Advice for Using ConvNets

Alec Dewulf

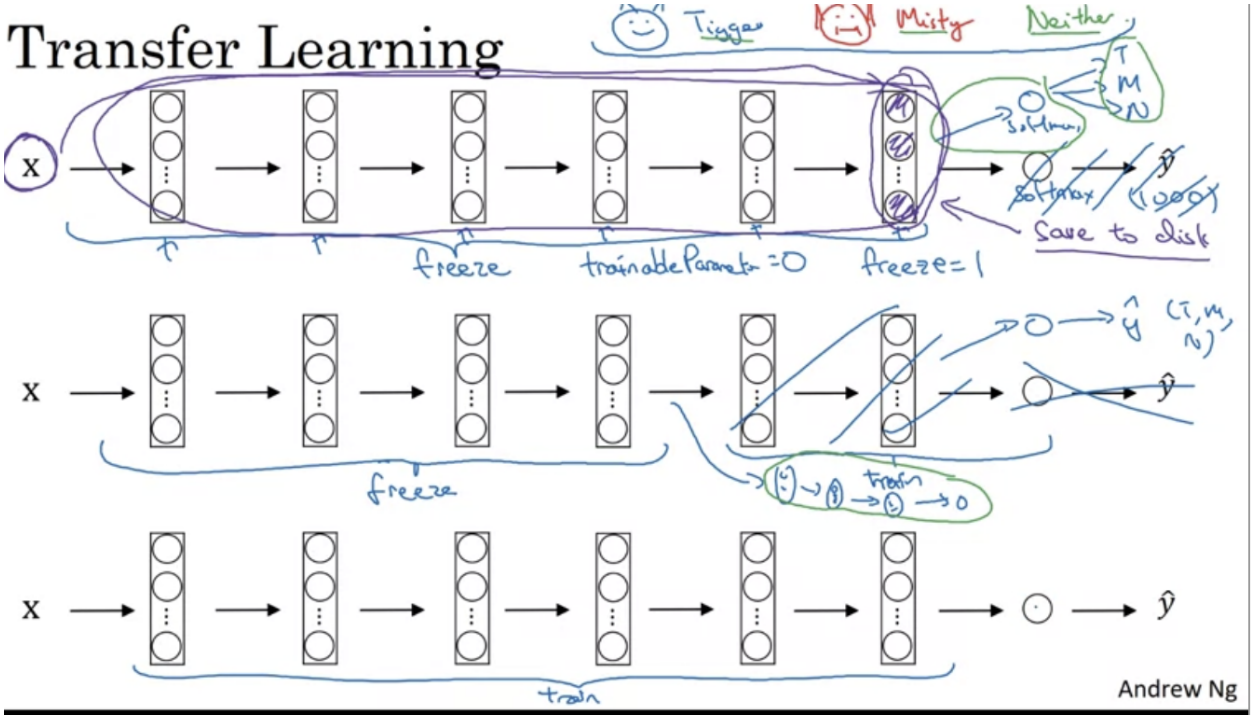
Transfer Learning

Rather than training the weights from scratch, you can download the weights someone else has already trained.

You can specify freeze parameters which prevent weights from being edited in certain layers of the network.

You can for example, freeze everything but the softmax to apply a complex function learned by the network to your own specific task.

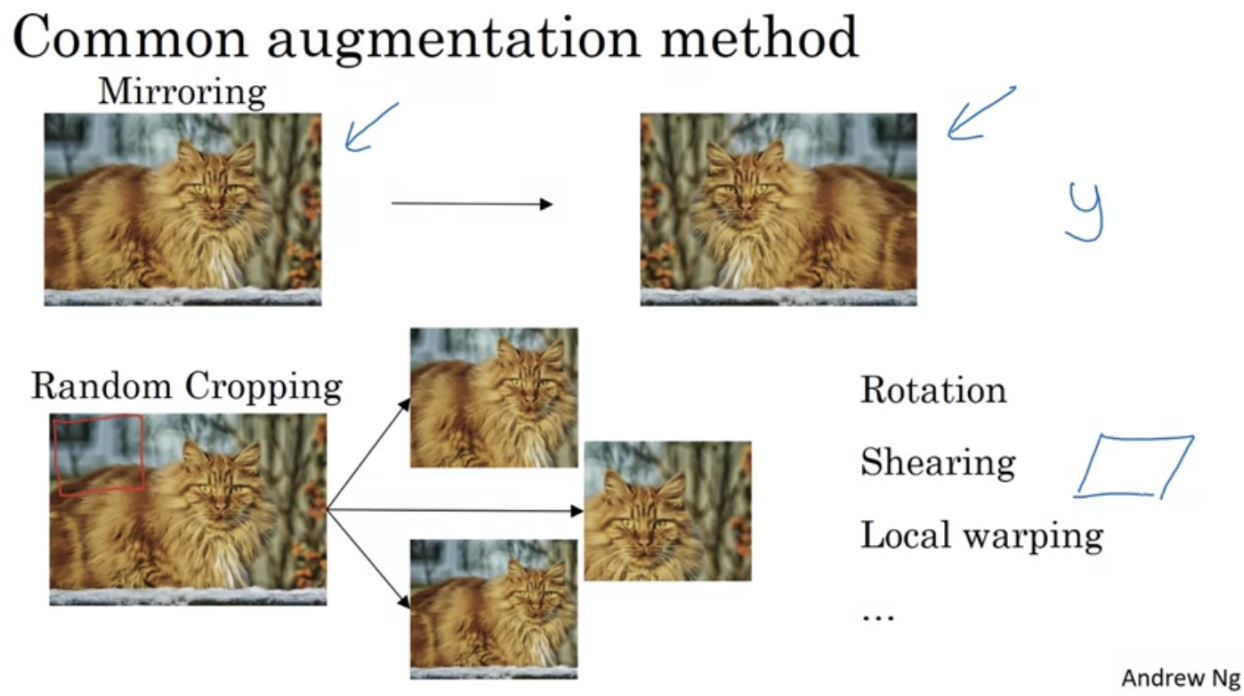
The more data you have, the less layers you freeze. This allows you to learn more specific functions.



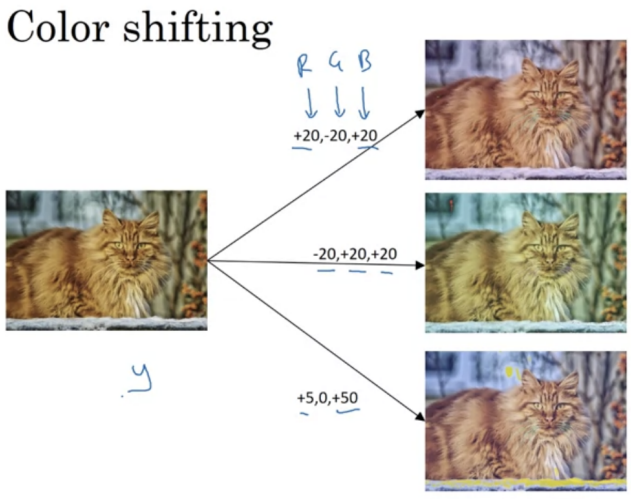
If you have a lot of data, you can use the weights you find online as initial parameters which may decrease train time.

Data Augmentation

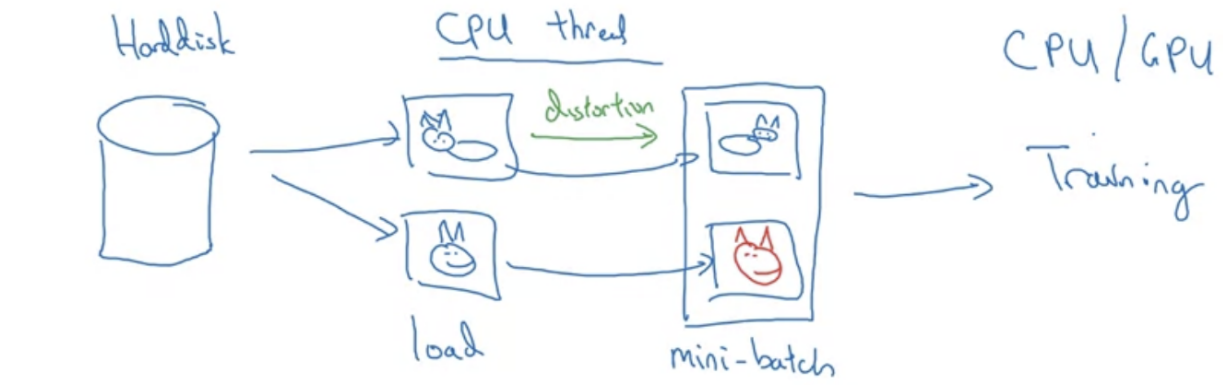
One of the most common augmentation methods is mirroring. Another is *random cropping* which takes random crops of the image. Random cropping is likely more dangerous because the key features may not be contained in the crop.



You can also do color shifting by adding some values to the channels in the image.



This makes the network more robust to slight color changes which could be caused by shadows or additional sunlight for example.



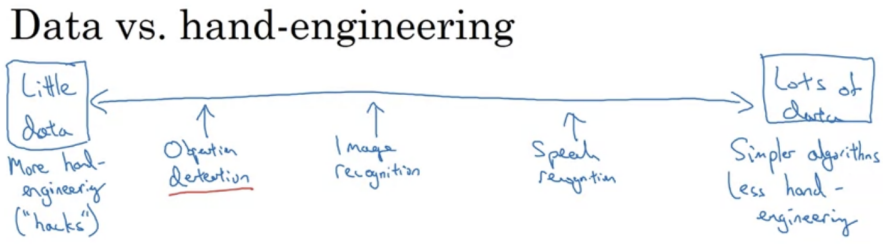
You can use one or multiple threads for loading the data and implementing distortions and then passing the result to some other thread that does the training.

These processes also have hyperparameters that describe things like the values to shift the color.

The State of Computer Vision

CV engineers still wish they had more data. There is also less data for object detection which is more complicated.

When you have a lot of data, you tend to have people getting away with using simpler networks.



Two sources of knowledge

- Labeled data (x, y)
- Hand engineered features/network architecture/other components

Andrew Ng

Ensembling is training several networks independently and averaging their outputs. This might help you do 1% or 2% better on a benchmark dataset. You might run the data through 3-15 networks on average. This is 3-15 times slower and for that reason rarely used in production

Multi-crop at test time: run a classifier on multiple versions of test images and average the results.

Tips for doing well on benchmarks/winning competitions

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

10-crop

Andrew Ng

The 10-crop is a specific method of generating 10 different images from a single image by random cropping.

Use open source code

- Use architectures of networks published in the literature
- Use open source implementations if possible
- Use pretrained models and fine-tune on your dataset