Süleyman Alperen GÜLER
201101027

IMAGENET CLASSIFICATION WITH DEEP CONVOLUTIONAL NEURAL NETWORKS

Dataset

Architecture

- Why CNN
- ReLu Nonlinearities
- Multi GPU Training
- Local Response Normalization
- Overlapping Pooling

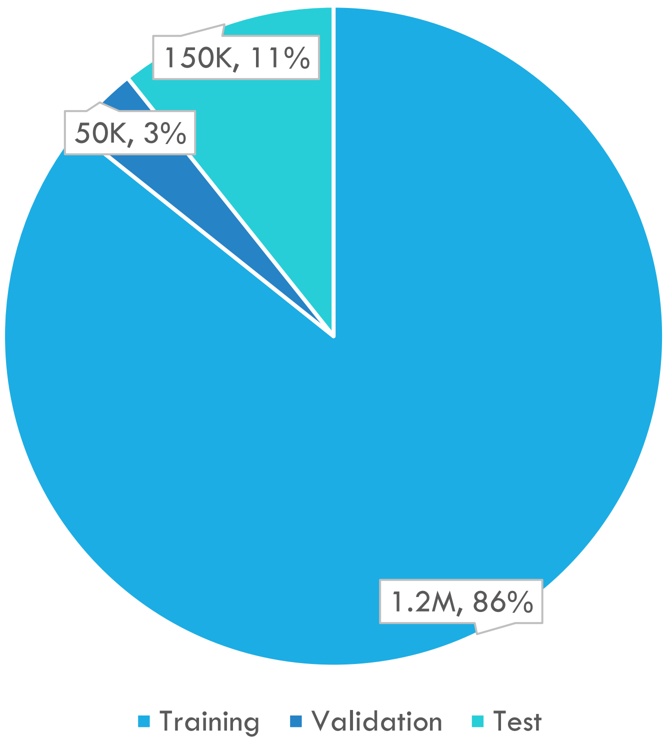
REDUCING OVERFITTING

- Data Augmentation
- Dropout

Test Results

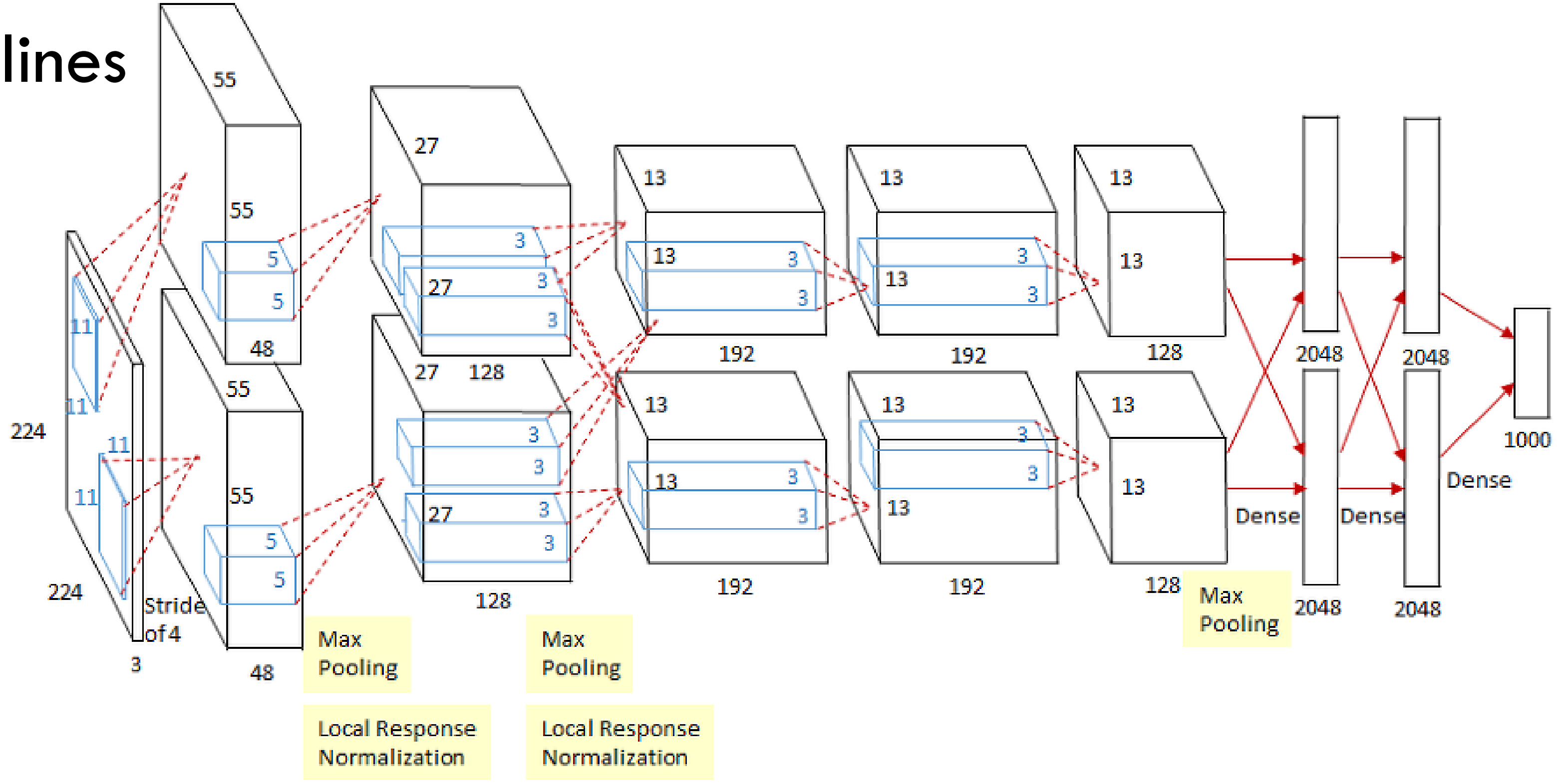
DATASET

- Original dataset has over 15 million images, which labelled by human labellers
- Images down-sampled to a fixed resolution of 256×256



ARCHITECTURE

- Network was so large, didn't fit into one GPU
- It has 2 different pipelines



ARCHITECTURE

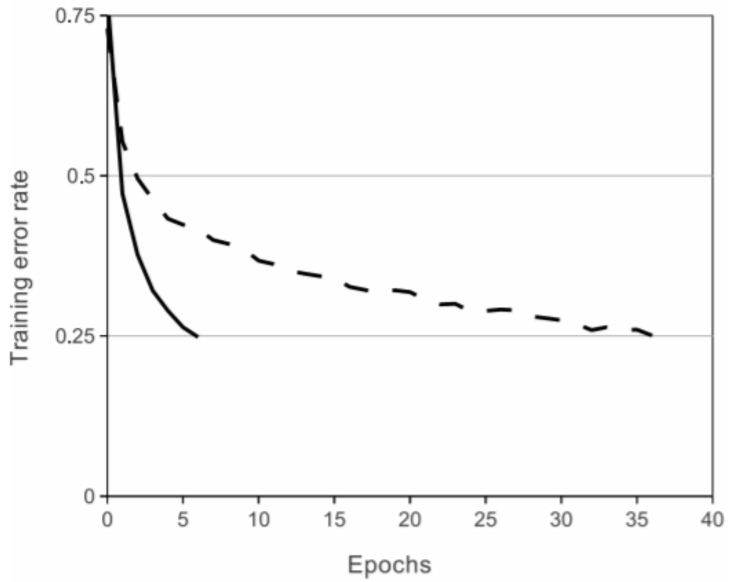
Why CNN?

- Has lots of capacity
- The problem needs lots of prior knowledge
- However expensive to apply in large-scale high-resolution images

ARCHITECTURE

ReLu Nonlinearity:

- At the time tanh and sigmoid functions were popular
- In terms of training time tanh and sigmoid functions much slower than relu



ARCHITECTURE

Multi GPU Training:

- Distributed the model in GPUs
- GPUs communicate only in certain layers
- Reduced top-1 error by 1.7% and top-5 by 1.2%

ARCHITECTURE

Local Response Normalization:

- k, n, α, and β are hyperparameters
- Reduces top-1 and top-5 error rates by 1.4% and 1.2%

$$b_{x,y}^{i} = a_{x,y}^{i} \Big/ \left(k + \alpha \sum_{j=\max(0,\, i-n/2)}^{\min(N-1,\, i+n/2)} \left(a_{x,y}^{j}\right)^{2} \right)^{\beta}$$

Overlapping pooling:

- Reduces top-1 and top-5 error rates by 1.7% and 1.2%
- It is slightly more difficult to overfit

REDUCING OVERFITTING

Data Augmentation:

- Random Horizontal Fliping
- Random Cropping
- Translation
- Color Jittering

- These applications reduces error by 1%

REDUCING OVERFITTING

Dropout:

- Sets 0 the output of each hidden neuron with probability of 0.5
- At test time, multiply all neuron outputs by 0.5
- This application roughly doubles number of iterations

TEST RESULTS

Table 1. Comparison of results on ILSVRC-2010 test set.

Model	Top-1 (%)	Top-5 (%)
Sparse coding[2]	*47.1*	*28.2*
SIFT + FVs[29]	*45.7*	*25.7*
CNN	**37.5**	**17.0**

Table 2. Comparison of error rates on ILSVRC-2012 validation and test sets.

Model	Top-1 (val, %)	Top-5 (val, %)	Top-5 (test, %)
SIFT + FVs[6]	–	–	*26.2*
1 CNN	40.7	18.2	–
5 CNNs	38.1	16.4	**16.4**
1 CNN*	39.0	16.6	–
7 CNNs*	36.7	15.4	**15.3**

TEST RESULTS