

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

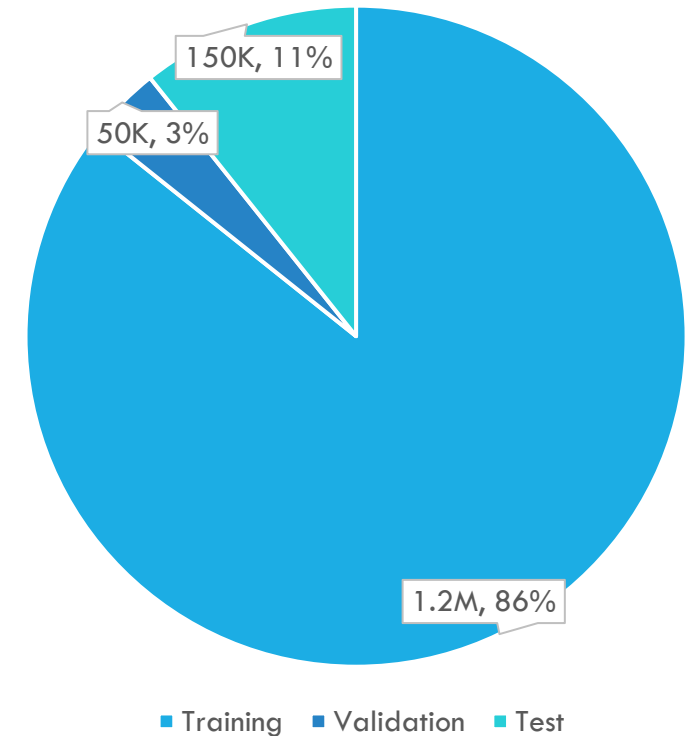
REDUCING OVERFITTING

- Data Augmentation
- Dropout

Test Results

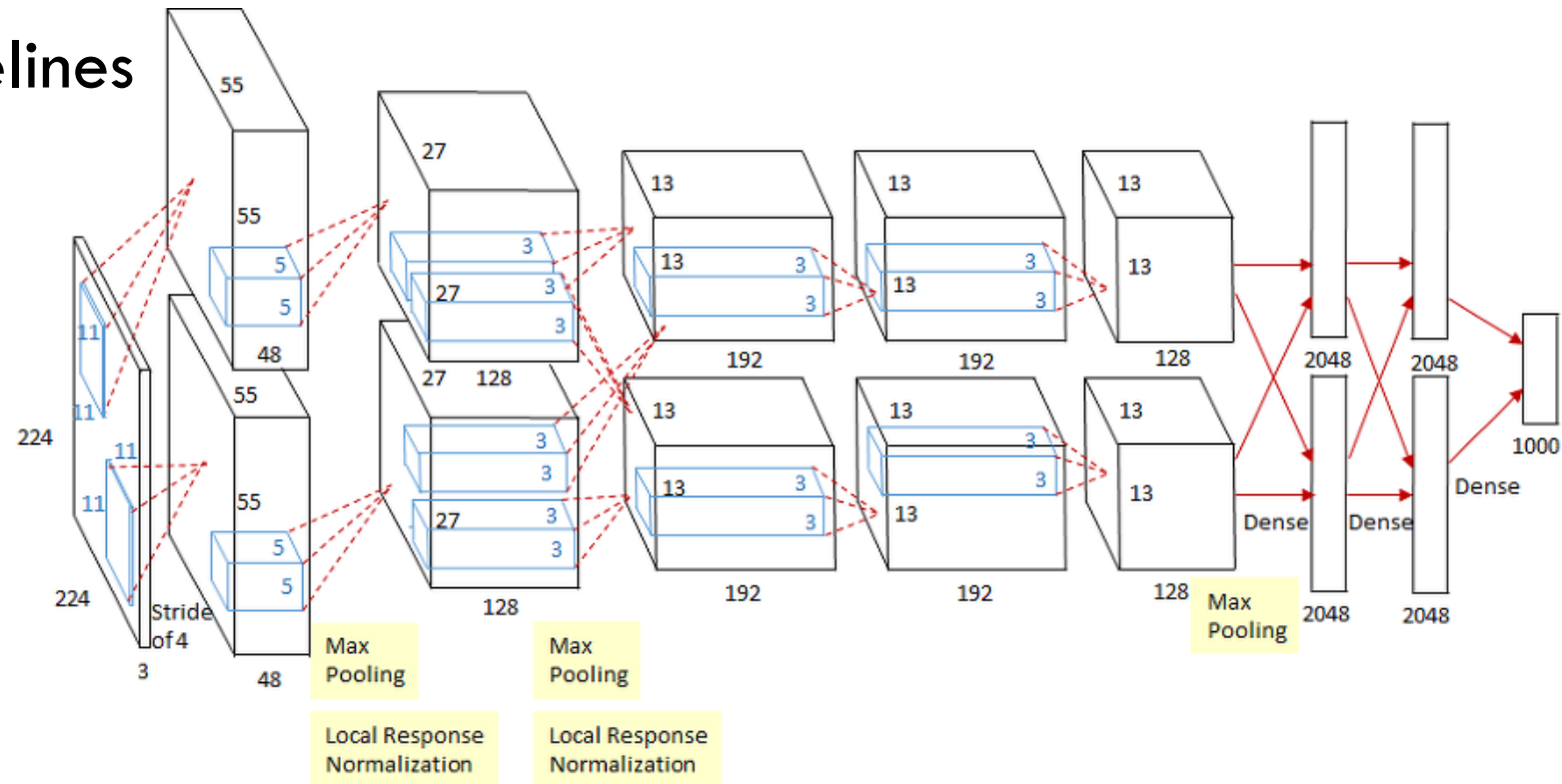
DATASET

- 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

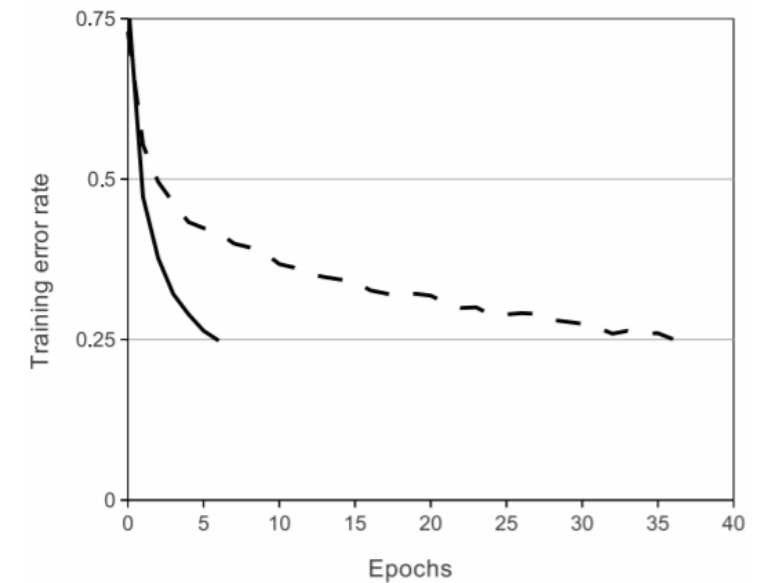
Why CNN?

- 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%

Overlapping pooling:

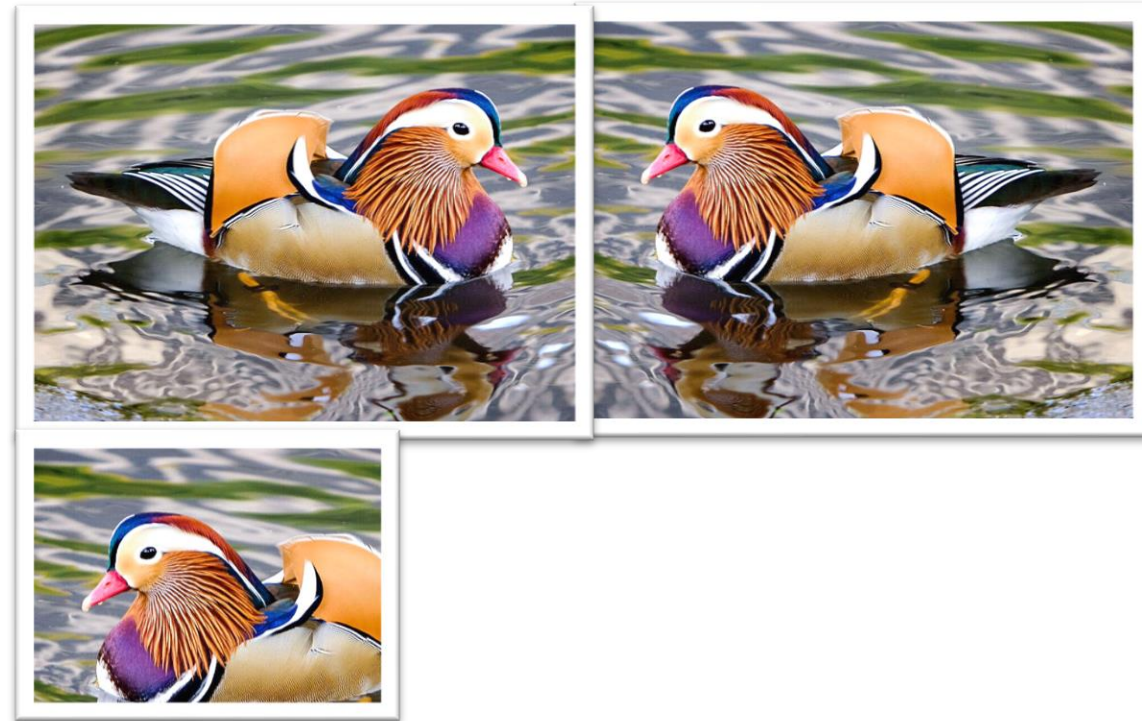
- Reduces top-1 and top-5 error rates by 1.7% and 1.2%
- It is slightly more difficult to overfit

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

REDUCING OVERFITTING

Data Augmentation:

- Random Horizontal Flipping
 - 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 (%)
<i>Sparse coding</i> ²	47.1	28.2
<i>SIFT + FVs</i> 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, %)
<i>SIFT + FVs</i> 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

			
mite	container ship	motor scooter	leopard
<div> <div></div> <div>mite</div> <div>black widow</div> <div>cockroach</div> <div>tick</div> <div>starfish</div> </div>	<div> <div></div> <div>container ship</div> <div>lifeboat</div> <div>amphibian</div> <div>fireboat</div> <div>drilling platform</div> </div>	<div> <div></div> <div>motor scooter</div> <div>go-kart</div> <div>moped</div> <div>bumper car</div> <div>golfcart</div> </div>	<div> <div></div> <div>leopard</div> <div>jaguar</div> <div>cheetah</div> <div>snow leopard</div> <div>Egyptian cat</div> </div>
			
grille	mushroom	cherry	Madagascar cat
<div> <div></div> <div>convertible</div> <div>grille</div> <div>pickup</div> <div>beach wagon</div> <div>fire engine</div> </div>	<div> <div></div> <div>agaric</div> <div>mushroom</div> <div>jelly fungus</div> <div>gill fungus</div> <div>dead-man's-fingers</div> </div>	<div> <div></div> <div>dalmatian</div> <div>grape</div> <div>elderberry</div> <div>ffordshire bullterrier</div> <div>currant</div> </div>	<div> <div></div> <div>squirrel monkey</div> <div>spider monkey</div> <div>titi</div> <div>indri</div> <div>howler monkey</div> </div>