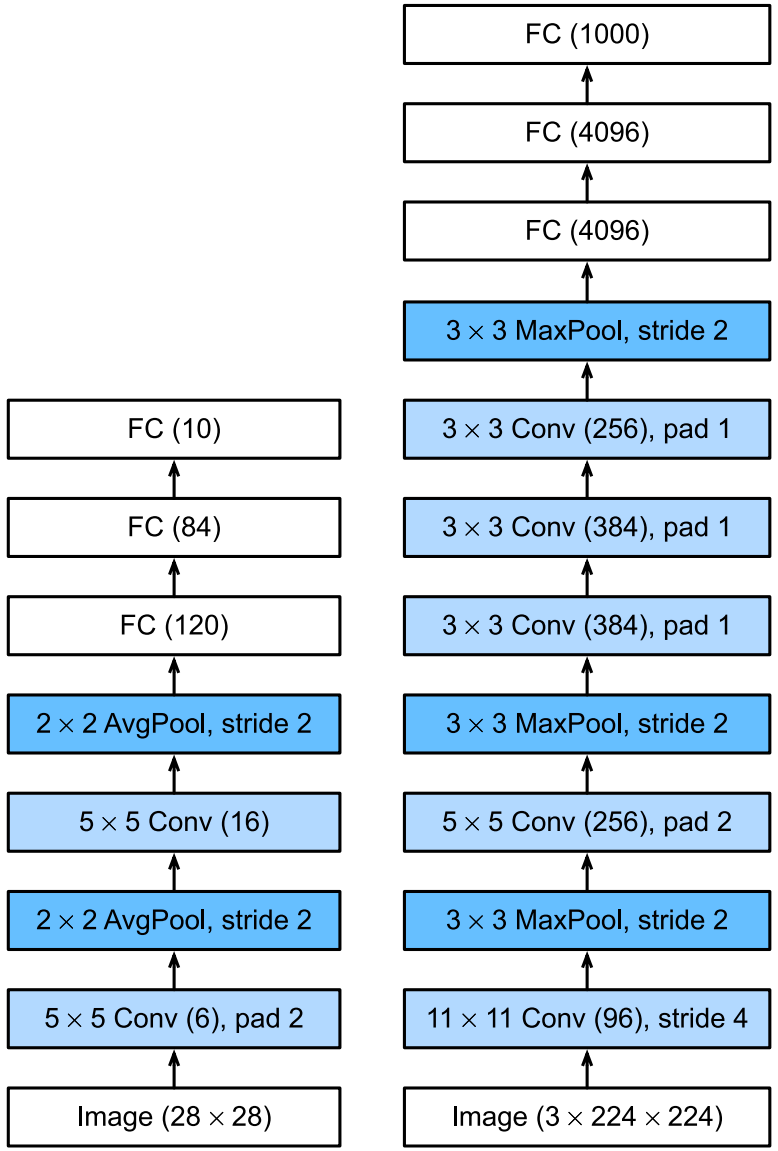
Introduction to Computer Vision

AlexNet

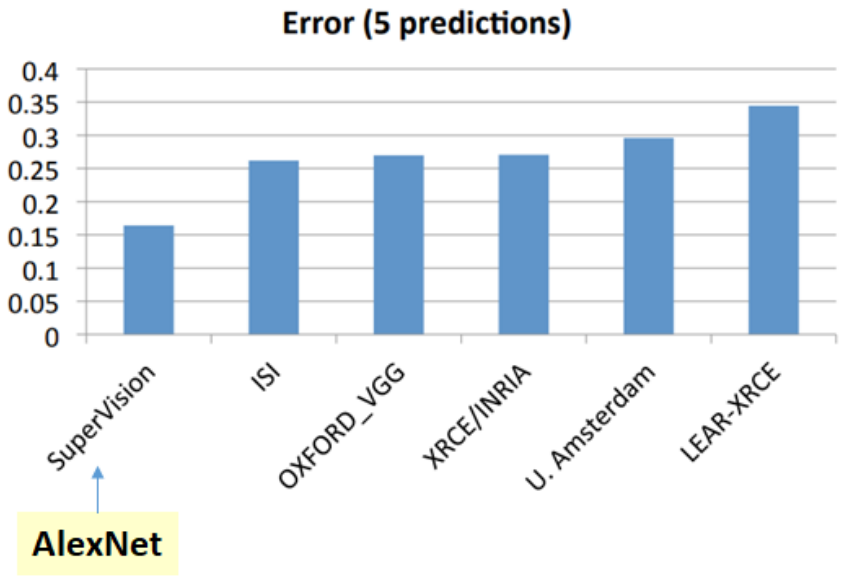
Cordonnier Xavier
2025/11/12



Context

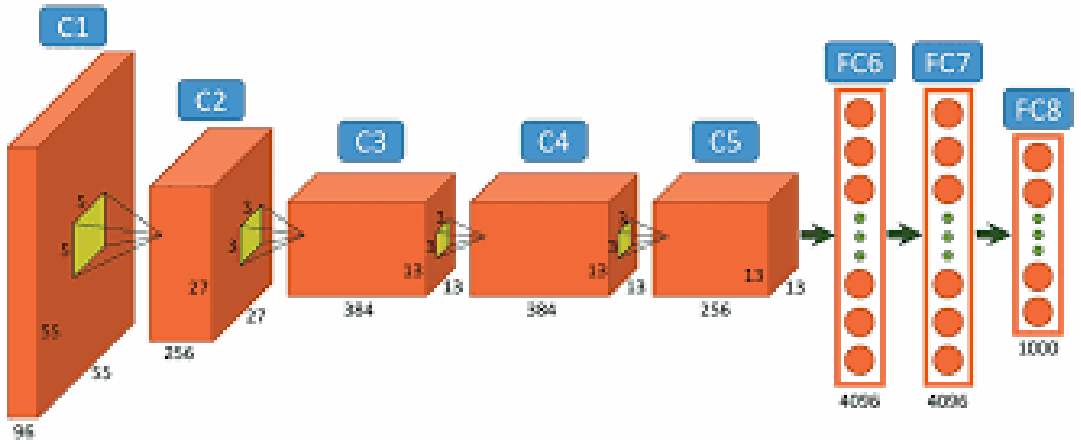
- New step in CNN usage
- Prove the potential of CNN over supervise methods
- Introduction of new solutions
- Use of 2 GPU as processing unit
- Win the 2012 ILSVRC

Ranking of the best results from each team



Technical description

- 8 layers network
- Max pooling
- ReLu and SoftMax activation function
- Image augmentation technics
- LRN and dropout normalization



Technical description

Feature map / Image size	Filter size	padding	Stride	pooling	Normalization
224*224*3	11*11	0	4	Yes	LRN
27*27*256	5*5	0	1	Yes	LRN
13*13*384	3*3	1	1	No	No
13*13*384	3*3	1	1	No	No
13*13*256	3*3	1	1	Yes	No

Technical description

Fully connected layer number	Number of neurons	Normalization
1	4096	Dropout 50%
2	4096	Dropout 50%
3	1000	No

Returns a 2d vector with the score for each of the 1000 classes

Technical description

Feature map	Pooling size	padding	Stride	Size after polling
55*55*96	3*3	0	2	27*27*96
27*27*256	3*3	0	2	13*13*256
13*13*256	3*3	0	2	6*6*256

Pooling windows are overlapping :

Observations

The last Convolutional layer :

6*6*256 = 9 216 and not 4 096

Around 60 millions parameters total