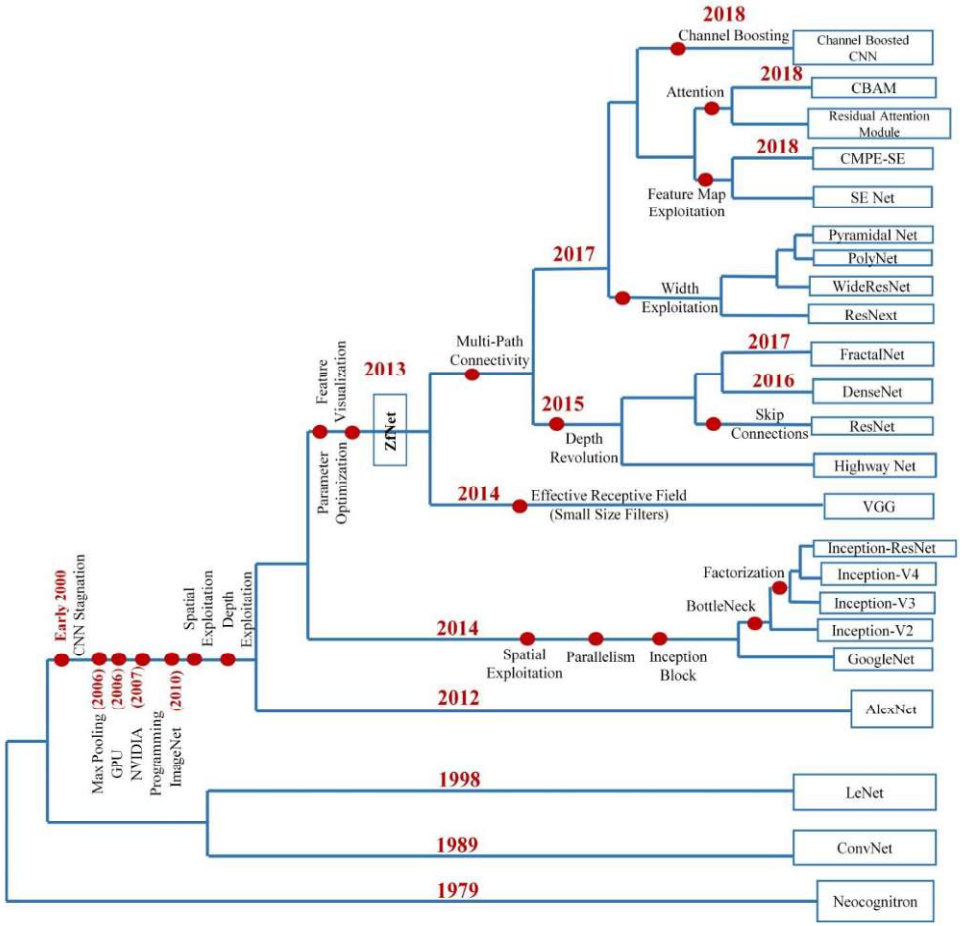
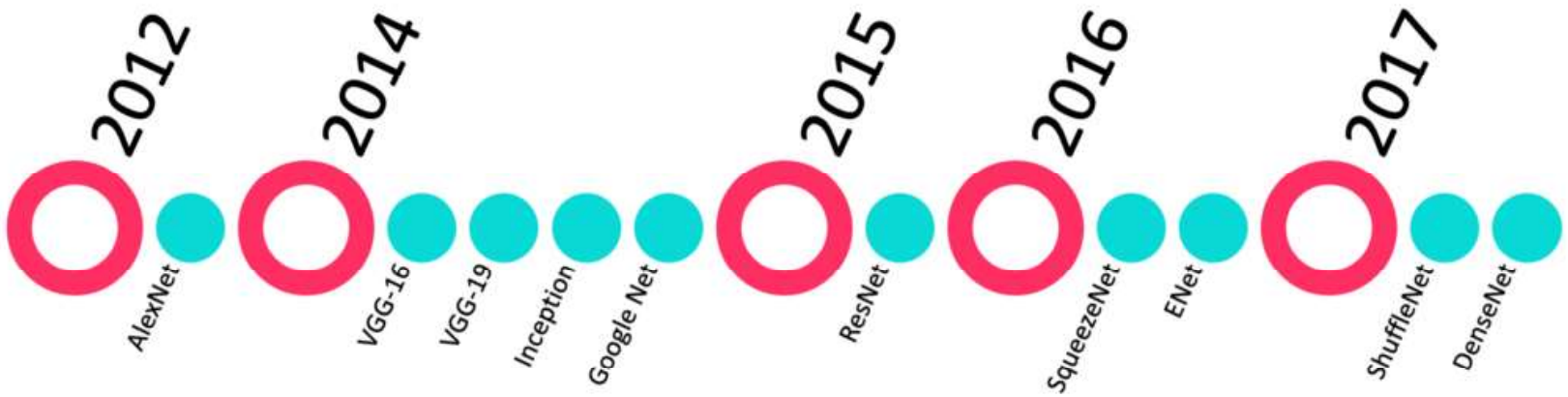
شبکه های عصبی معروف

DNNs Timeline



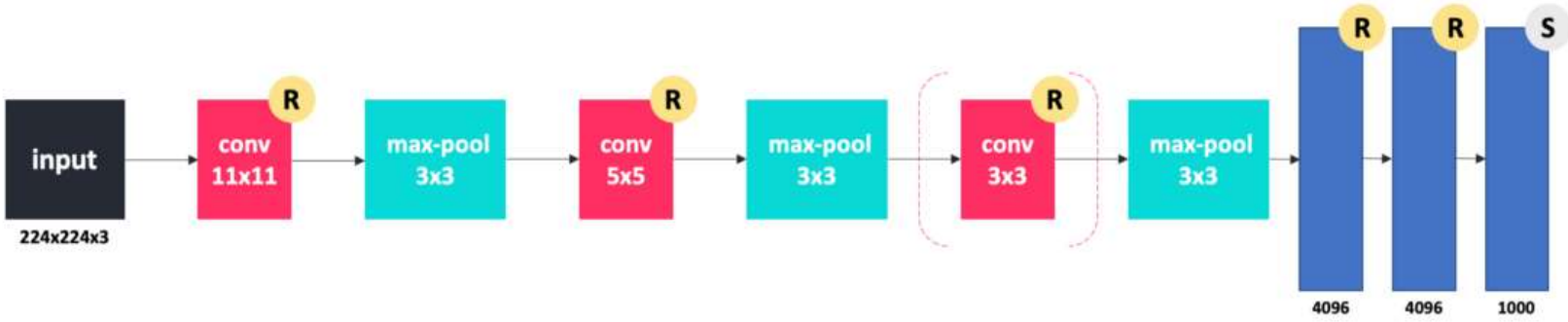
History

- **LeNet is the first successful application of CNNs and was developed by Yann Lecun in the 1990s that was used to read zip codes, digits, etc. The latest work is called LeNet-5 which a 5-layer CNN that reaches 99.2 % accuracy on insolated character recognition.**



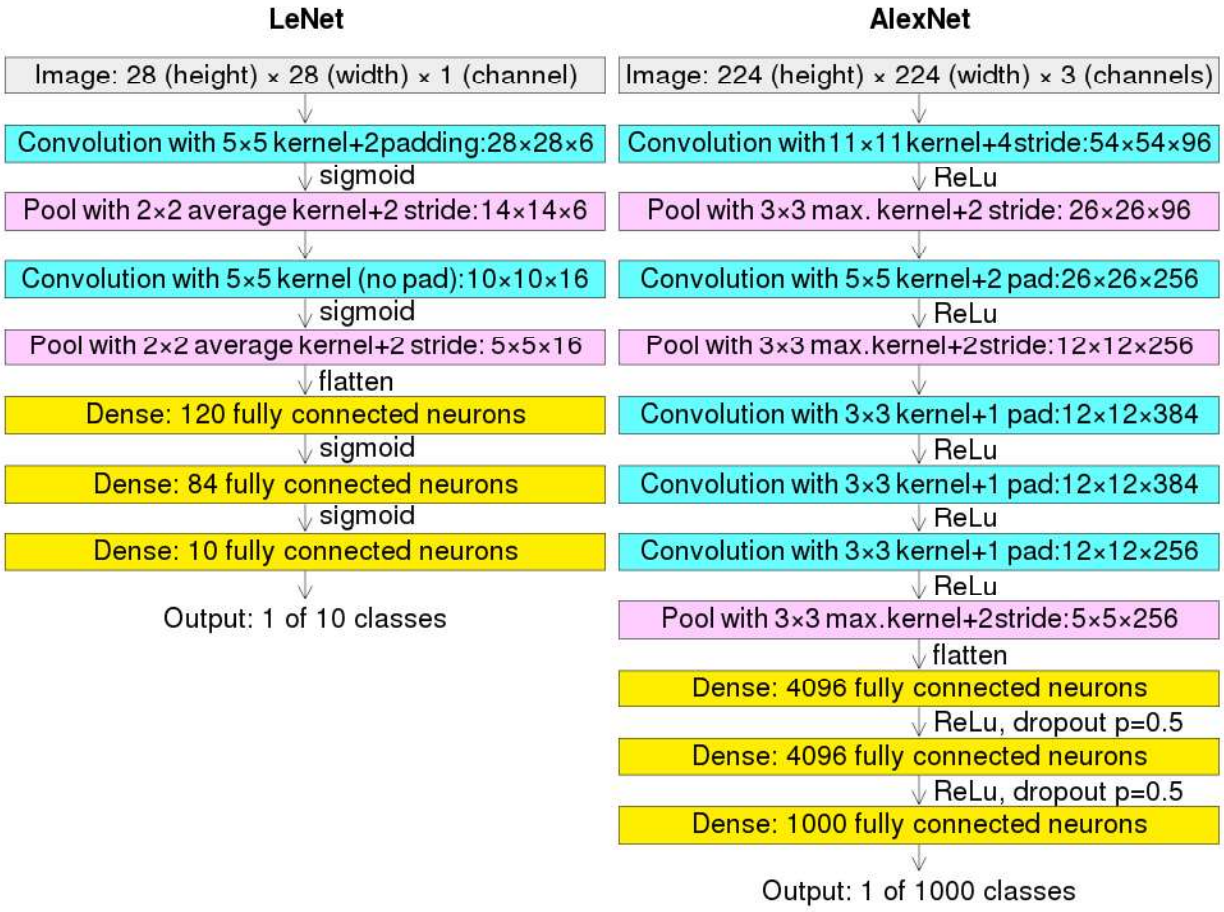
01. AlexNet (2012)

- **Alex Krizhevsky, Ilya Sutskever, and Geoff Hinton won the ImageNet Large Scale Visual Recognition Challenge with a test accuracy of 84.6%.**
- **The model significantly outperformed the second runner-up (top-5 error of 16% compared to runner-up with 26% error).**
- **The network consists of 5 convolutional layers and 3 fully connected layers.**



https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

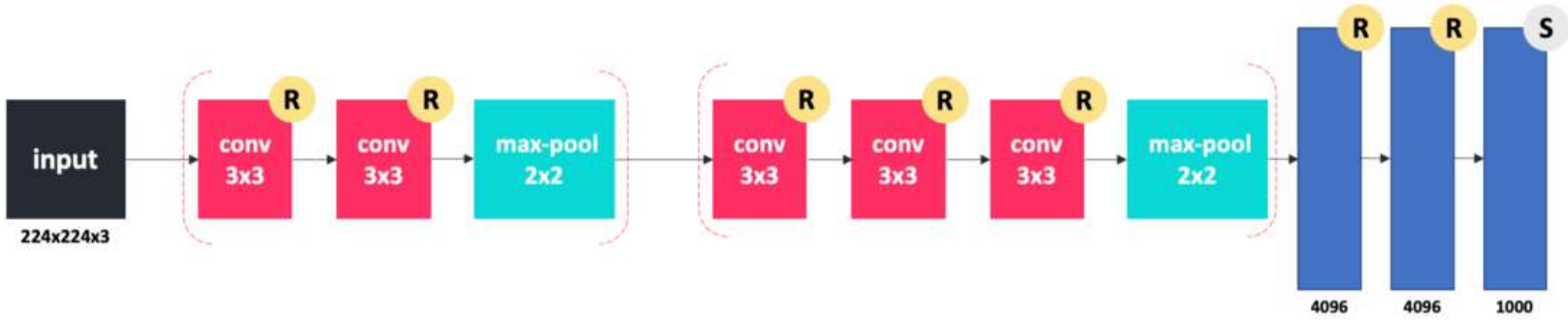
01. AlexNet (2012)



Comparison of the LeNet and AlexNet convolution, pooling, and dense layers (AlexNet image size should be 227×227×3, instead of 224×224×3, so the math will come out right. The original paper said different numbers, but Andrej Karpathy, the head of computer vision at Tesla, said it should be 227×227×3 (he said Alex didn't describe why he put 224×224×3). The next convolution should be 11×11 with stride 4: 55×55×96 (instead of 54×54×96). It would be calculated, for example, as: [(input width 227 - kernel width 11) / stride 4] + 1 = [(227 - 11) / 4] + 1 = 55. Since the kernel output is the same length as width, its area is 55×55.)

02. VGG16 (2014)

- **Karen Simonyan and Andrew Zisserman of the Visual Geometry Group Lab of Oxford University.**
- **The model achieves 92.7% top-5 test accuracy in ImageNet.**
- **The 16 in VGG16 refers to it has a total of 16 layers that have weights.**



https://arxiv.org/abs/1409.1556

03. VGG19 (2014)

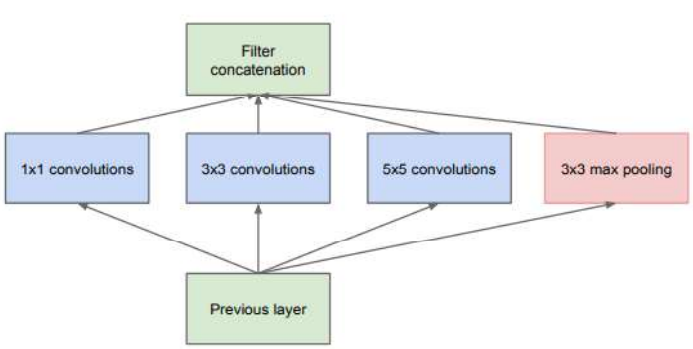
- **Karen Simonyan and Andrew Zisserman of the Visual Geometry Group Lab of Oxford University.**
- **The model trained on more than a million images from the Imagenet database with an accuracy of 92%.**
- **The 19 in VGG19 refers to it has a total of 19 layers that have weights.**

Model	top-5 classification error on ILSVRC-2012 (%)	
	validation set	test set
16-layer	7.5%	7.4%
19-layer	7.5%	7.3%
model fusion	7.1%	7.0%

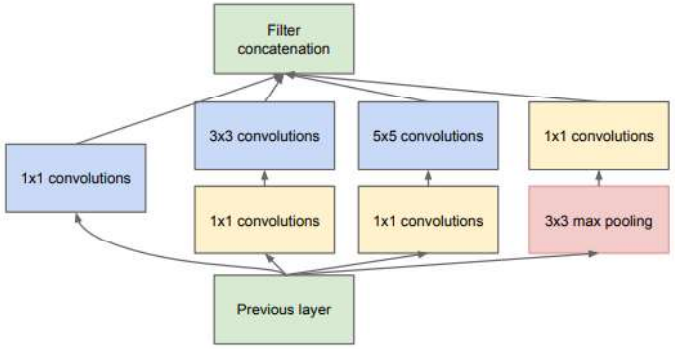
https://arxiv.org/abs/1409.1556

04. Inception and GoogLeNet (2014)

- **Christian Szegedy, Wei Liu and Yangqing Jia.**
- **GoogLeNet (or Inception v1) has 22 layers deep.**
- **With the accuracy of 93.3% this model won the 2014 ImageNet competition in both classification an detection task.**



(a) Inception module, naïve version

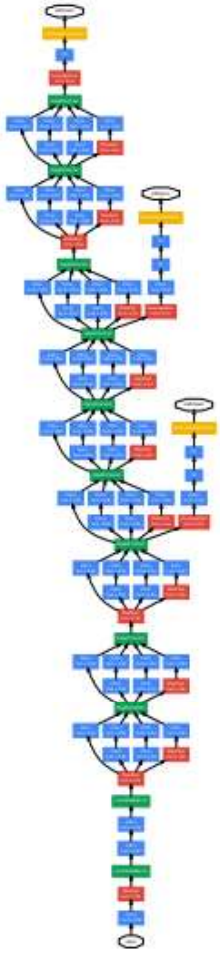


(b) Inception module with dimension reductions

https://arxiv.org/pdf/1409.4842.pdf

04. Inception and GoogLeNet (2014)

type	patch size/ stride	output size	depth	#1×1	#3×3 reduce	#3×3	#5×5 reduce	#5×5	pool proj	params	ops
convolution	7×7/2	112×112×64	1							2.7K	34M
max pool	3×3/2	56×56×64	0								
convolution	3×3/1	56×56×192	2		64	192				112K	360M
max pool	3×3/2	28×28×192	0								
inception (3a)		28×28×256	2	64	96	128	16	32	32	159K	128M
inception (3b)		28×28×480	2	128	128	192	32	96	64	380K	304M
max pool	3×3/2	14×14×480	0								
inception (4a)		14×14×512	2	192	96	208	16	48	64	364K	73M
inception (4b)		14×14×512	2	160	112	224	24	64	64	437K	88M
inception (4c)		14×14×512	2	128	128	256	24	64	64	463K	100M
inception (4d)		14×14×528	2	112	144	288	32	64	64	580K	119M
inception (4e)		14×14×832	2	256	160	320	32	128	128	840K	170M
max pool	3×3/2	7×7×832	0								
inception (5a)		7×7×832	2	256	160	320	32	128	128	1072K	54M
inception (5b)		7×7×1024	2	384	192	384	48	128	128	1388K	71M
avg pool	7×7/1	1×1×1024	0								
dropout (40%)		1×1×1024	0								
linear		1×1×1000	1							1000K	1M
softmax		1×1×1000	0								



https://arxiv.org/pdf/1409.4842.pdf

05. ResNet (2016)

- **The network has been created and acquainted by Microsoft.**
- **With 96.4% accuracy this model won the 2016 ImageNet competition.**
- **It is well-known due to its depth (to 152 layers) and the introduction of residual blocks.**

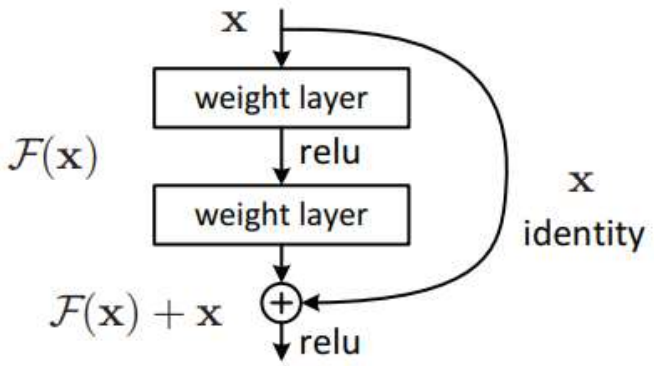
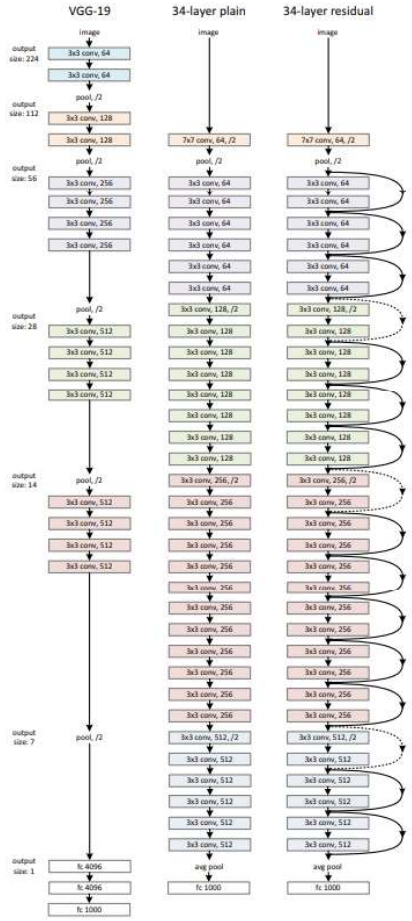


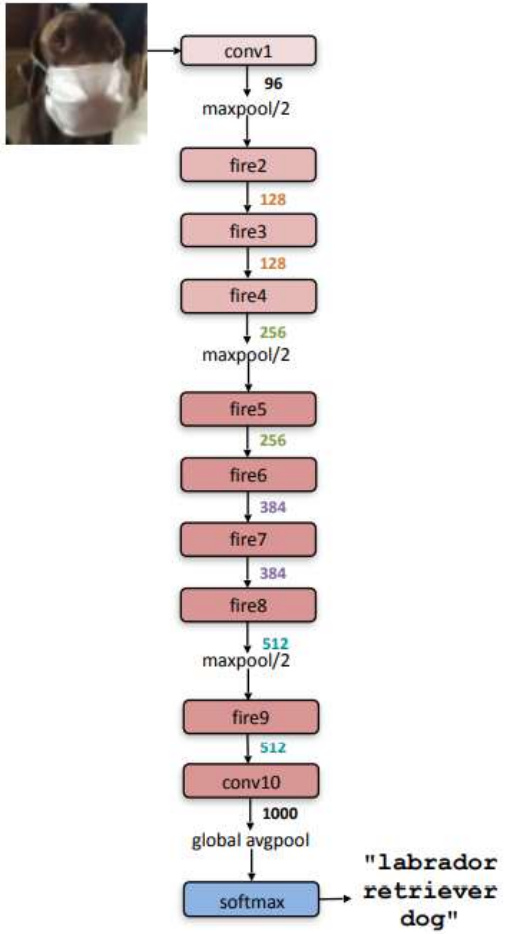
Figure 2. Residual learning: a building block.



https://arxiv.org/abs/1512.03385

06. Squeeze Net (2016)

- **Forrest N. Iandola, Song Han, Matthew W. Moskewicz.**
- **Smaller DNNs require less communication across servers during distributed training.**
- **Smaller DNNs require less bandwidth to export a new model from the cloud to an autonomous car.**
- **Smaller DNNs are more feasible to deploy on FPGAs and other hardware with limited memory.**
- **SqueezeNet can be three times faster and 510 times smaller.**



https://arxiv.org/pdf/1602.07360.pdf

07. DenseNet (2017)

- **Gao Huang, Zhuang Liu, and their team.**
- **It received the best paper award at CVPR Conference.**
- **With traditional convolutional networks with n layers have n connections but DensetNet has n(n+1)/2 connections in total becaues of feed-forward fashion.**

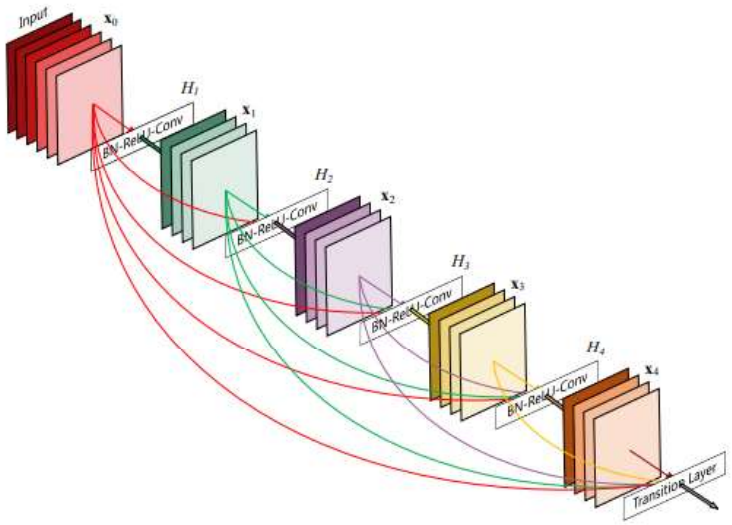


Figure 1: A 5-layer dense block with a growth rate of $k = 4$. Each layer takes all preceding feature-maps as input.

https://openaccess.thecvf.com/content_cvpr_2017/papers/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.pdf