**Convolutional Neural Network Architecture: Forging Pathways to the Future**

<https://missinglink.ai/guides/convolutional-neural-networks/convolutional-neural-network-architecture-forging-pathways-future/>

Advances in AI and deep learning have enabled the rapid evolution in the fields of computer vision and image analysis. This is all made possible by the emergence and progress of Convolutional Neural Networks (CNNs).

## BASIC CONVOLUTIONAL NETWORK ARCHITECTURE

CNN architecture is inspired by the organization and functionality of the visual cortex and designed to mimic the connectivity pattern of neurons within the human brain.

The neurons within a CNN are split into a three-dimensional structure, with each set of neurons analyzing a small region or feature of the image. In other words, each group of neurons specializes in identifying one part of the image. CNNs use the predictions from the layers to produce a final output that presents a vector of probability scores to represent the likelihood that a specific feature belongs to a certain class.

## POPULAR CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES

The architecture of a CNN is a key factor in determining its performance and efficiency. The way in which the layers are structured, which elements are used in each layer and how they are designed will often affect the speed and accuracy with which it can perform various tasks.

### THE IMAGENET CHALLENGE

The ImageNet project is a visual database designed for use in the research of visual object recognition software. The ImageNet project has more than 14 million images specifically designed for training CNN in object detection, one million of which also provide bounding boxes for the use of networks such as [YOLO](https://missinglink.ai/guides/neural-network-concepts/yolo-deep-learning-dont-think-twice/).

Since 2010, the project hosts an annual contest called the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). The contenders of the contest build software programs that attempt to correctly detect and classify objects and scenes within the given images. Currently, the challenge uses a cut down list of a thousand separate classes.

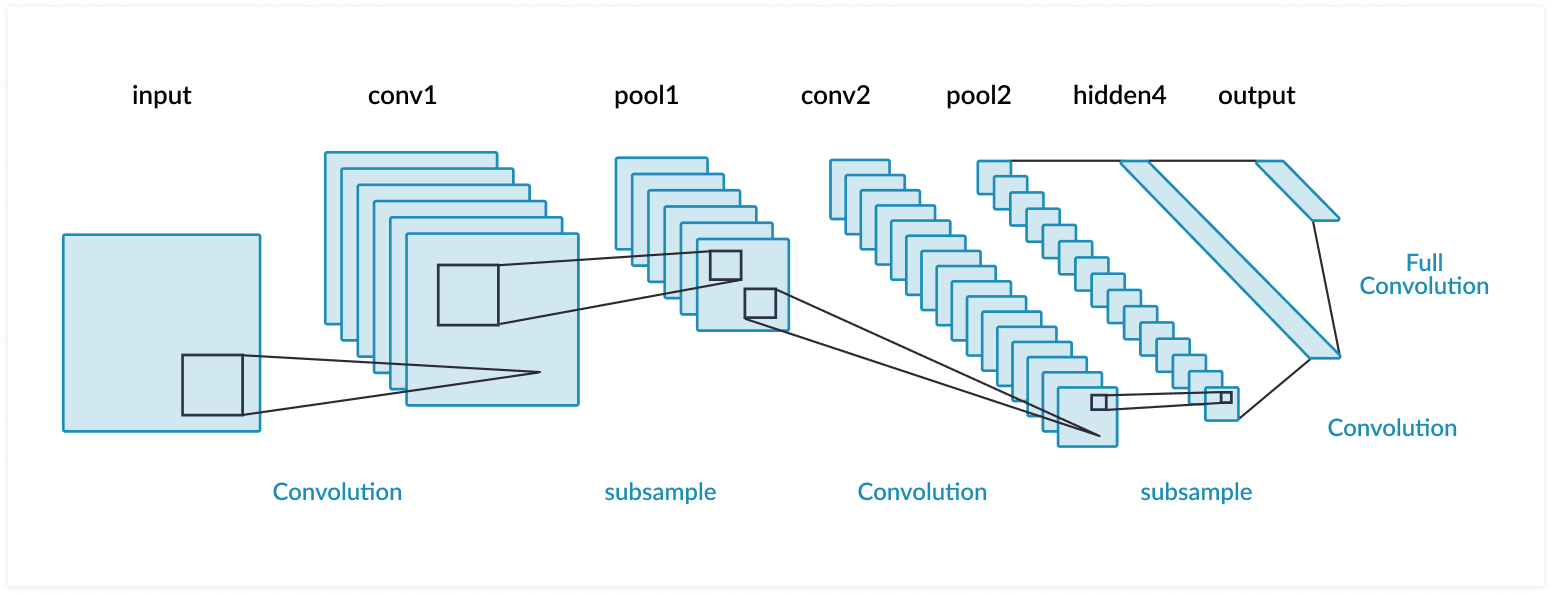
When the annual ILSVRC competition began, a good classification rate was 25%, the first major leap in performance was achieved by a network called AlexNet in 2012, which dropped the classification rate by 10%. Over the next years, the error rates dropped to lower percentages and finally exceeded human capabilities.

### POPULAR ARCHITECTURES – THE DIFFERENCES IN ARCHITECTURE AND THEIR ADVANTAGES

There are many popular CNN architectures, many of them gained recognition by achieving good results at the ILSVRC.

### LeNET-5 (1998)

This 7-layer CNN classified digits, digitized 32×32 pixel greyscale input images. it was used by several banks to recognize the hand-written numbers on checks.

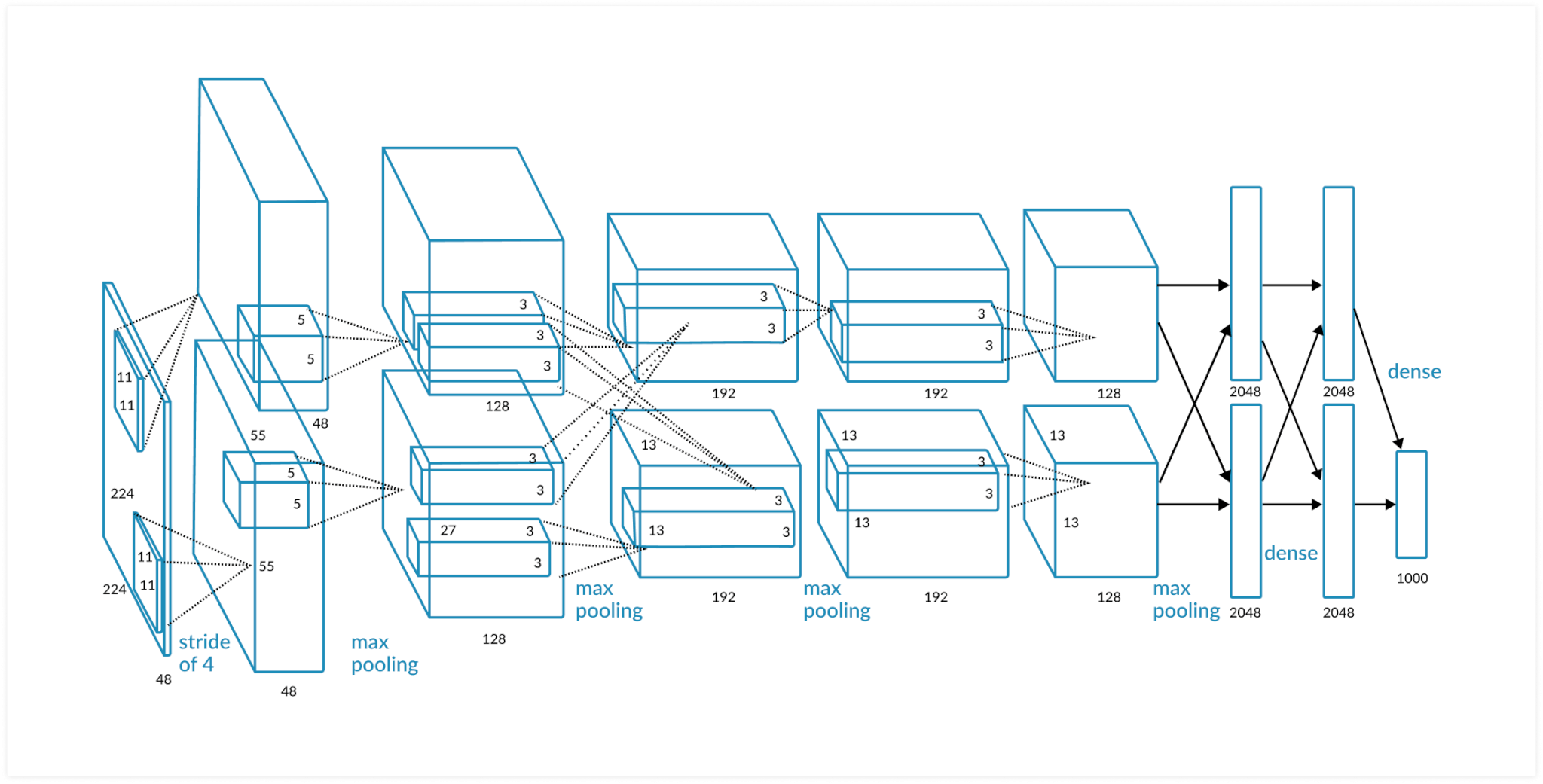


### ALEXNET (2012)

AlexNet is designed by SuperVision group, with a similar architecture to LeNet, but deeper━it has more filters per layer as well as stacked convolutional layers. It is composed of five convolutional layers followed by three fully connected layers.

One of the most significant differences between AlexNet and other object detection algorithms is the use of ReLU for the non-linear part instead of Sigmond function or Tanh like traditional neural networks. AlexNet leverages ReLU’s faster training to make their algorithm faster.

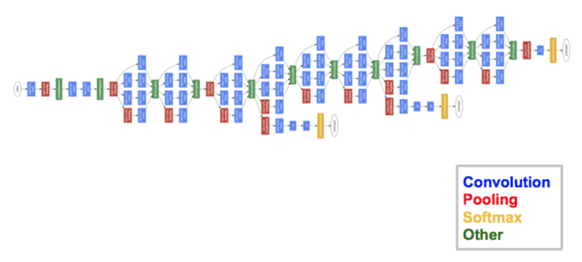
The creators of AlexNet split their network into two pipelines because they used two Nvidia Geforce GTX 580 [Graphics Processing Units](https://missinglink.ai/guides/deep-learning-frameworks/complete-guide-deep-learning-gpus/) (GPUs) to train their CNN.



### GOOGLENET (2014)

Built with a CNN inspired by LetNet, the GoogleNet network, which is also named Inception V1, was made by a team at Google. GoogleNet was the winner of ILSVRC 2014 and achieved a top-5 error rate of less than 7%, which is close to the level of human performance.

GoogleNet architecture consisted of a 22 layer deep CNN used a module based on small convolutions, called “inception module”, which used batch normalization, RMSprop and image to reduce the number of parameters from 60 million like in AlexNet to only 4 million.



### VGGNET (2014)

VGGNet, the runner-up at the ILSVRC 2014, consisted of 16 convolutional layers. Similar to AlexNet, it used only 3×3 convolutions but added more filters. VGGNet trained on 4 GPUs for more than two weeks to achieve its performance.

The problem with VGGNet is that it consists of 138 million parameters, 34.5 times more than GoogleNet, which makes it challenging to run.

