



Take Charge of Your Privacy

Discover new ways to manage y
profile and privacy settings. Op
Facebook.

Facebook®

Vis

Calculate output size of Convolution

Machine Learning (ML)

cnn

2022 → Internship at**OpenGenus**
2023 → Internship at FAANG
2024 → Full-time High Paying Job

Apply now:
→ internship.opengenus.org

Taking this
Remote Internship
will change the whole
equation of your career



Facebook Live पर और अधिक खोजें

Ad Facebook®

[Get this book -> Problems on Array: For Interviews and Competitive Programming](#)

दिवाली Facebook पर है

दिवाली की अपनी मनपसंद परंपराओं के करके अपनी फीड को जगमगा दें। Facebook पर जाएँ।

Facebook®

[साइट](#)

Convolution is the most important operation in Machine Learning models where more than 70% of computational time is spent. The input data has specific dimensions and we can use the values to calculate the size of the output.

In short, the answer is as follows:

- **Output height** = $(\text{Input height} + \text{padding height top} + \text{padding height bottom} - \text{kernel height}) / (\text{stride height}) + 1$
- **Output width** = $(\text{Output width} + \text{padding width right} + \text{padding width left} - \text{kernel width}) / (\text{stride width}) + 1$

Up

The concerned parameters taken by Convolution are:



- Kernel height and Kernel width
- Stride height and stride width
- Padding height top, Padding height bottom
- Padding width right, Padding width left

We will understand each of the parameters and then, understand how it can be used to calculate the size of the output.

Input dimension

Ideally, the input in Machine Learning models are 4 dimensions (4D) where the dimensions are:

- Height (H)
- Width (W)
- Number of batches (N)
- Number of channels (C)

Height and width of the input is the height and width of the input images.

Number of batches (N) or the batch size is the number of 2D images processed together or passed into the Machine Learning model. When batch size is 1, the inference process is known as latency and when the batch size is greater than 1, it is known as throughput.

Common batch size (N) is 1 and 256.

Number of channels (C) is the number of channels in the input image. Images are defined by 3 dimensions: height, width and channels. Channels determine the features in the images. In colored images, it is usually 3 for RGB or HSV. This parameter can be of any dimension.

Note that the number of batches (N) and number of channels (C) is same for input and output as the parameters that transform the input in Convolution works on the 2D input and keeps the other dimensions preserved.

Kernel

Up

The kernel is ideally of 2 dimensions: height (KH) and width (KW). In some cases, it has a third dimension: channels (KC).

The 2D kernel applies on each 2D images. Each C (third dimension) 2D kernels are applied on



Stride

Stride is the value by which the kernel slides over the input data. By default, it is 1. Strides are of two types:

- Stride along width (SW)
- Stride along height (SH)

Common stride values are 2 and 3.

Padding

Padding is the amount of default data added to the sides of the input to maintain the size of the output. The values in the padding area is usually 0. There are four types of padding:

- Padding along height on top (P_H1)
- Padding along height on bottom (P_H2)
- Padding along width on left (P_W1)
- Padding along width on right (P_W2)

This modifies the dimensions of the input data as follows:

- Height: Height + Padding along height on top + Padding along height on bottom = $H + P_H1 + P_H2$
- Width: Width + Padding along width on left + Padding along width on right = $W + P_W1 + P_W2$

Calculation of output size

Input dimensions:

- Height (H)
- Width (W)
- Number of batches (N)
- Number of channels (C)

Up



- Kernel width (KW)
- Kernel channels (KC)
- Stride height (SH)
- Stride width (SW)
- Padding: P_H1, P_H2, P_W1, P_W2

The input image dimension: $N \times H \times W \times C$

The modified height and width of input due to padding is as follows:

- Height modified $H1 = H + P_H1 + P_H2$
- Width modified $W1 = W + P_W1 + P_W2$

On applying the kernel ($KH \times KW$), the modified height and width will be as follows:

- $H2 = H1 - KH + 1 = H + P_H1 + P_H2 - KH + 1$
- $W2 = W1 - KW + 1 = W + P_W1 + P_W2 - KW + 1$

The kernel channel (KC) modifies the output channel to KC as for each kernel channel, all input channels are added together.

On considering strides, we get the following height and width:

- $H3 = (H1 - KH) / SH + 1$
- $W3 = (W1 - KW) / SW + 1$

This makes the dimensions of the output as follows:

- Number of batches = N
- Number of channels = KC
- Height = $(\text{Input Height} + \text{Padding height (top and bottom)} - \text{Kernel Height}) / \text{Stride Height} + 1$
- Width = $(\text{Input Width} + \text{Padding width (left and right)} - \text{Kernel Width}) / \text{Stride Width} + 1$

Example

Up

Let us take some example values of input data and calculate the dimension of the output accordingly.

Input: $[N \ H \ W \ C] = [100 \ 255 \ 255 \ 64]$



Padding: $[P_H1, P_H2, P_W1, P_W2] = [2\ 2\ 2\ 2]$

Output Height = $(255 + 2 + 2 - 7) / 3 + 1 = 85$

Output Width = $(255 + 2 + 2 - 7) / 4 + 1 = 64$

Output channels = 128

Output batch size = 100

Hence, the output size is: $[N\ H\ W\ C] = 100 \times 85 \times 64 \times 128$

With this article at [OpenGenus](#), you must have the complete idea of computing the output size of convolution. Enjoy.

Learn more:

- [Convolution Layer](#) by **Surya Pratap Singh** at **OpenGenus**
- [Convolutional Neural Network \(CNN\) questions](#) by **Leandro Baruch** at **OpenGenus**
- [Building a Convolution Neural Network \(CNN\) for handwritten digit recognition in Python using Keras](#) by **Piyush Mishra** at **OpenGenus**
- [Convolutional Neural Networks \(CNN\)](#) by **Piyush Mishra** and **Junaid N Z** at **OpenGenus**
- [Machine Learning topics](#) at **OpenGenus**

Control Your Profile Details

Ad Facebook®



Connect to the World

Start connecting and sharing with people you know. See the latest in your community.

Facebook®

Visit



Ue Kiao, PhD

Ue Kiao is a Technical Author and Software Developer with B. Sc in Computer Science at National Taiwan University and PhD in Algorithms at Tokyo Institute of Technology | Researcher at TaoBao

Read More

Improved & Reviewed by:

— OpenGenus IQ: Computing Expertise & Legacy —

Machine Learning (ML)

ReLU6 in ML

Radial Basis Function Neural Network

Delta Rule in Neural Network

ALGORITHMS

Partition a set into two subsets such that sum of each subset is same

An analysis and explanation of the solution for the partition problem, in which a set of numbers is to be divided into two sub arrays such that their sum is equal by using dynamic programming. Up



[See all 498 posts →](#)

[REETIKA POOSARLA](#)

SOFTWARE ENGINEERING

Complex Numbers in Python

In this article, we have explored how to work with Complex Numbers in Python which is supported in Python inherently as a fundamental datatype. We have covered the different operations as well.



[LYNDI CASTREJON](#)