

# James D. McCaffrey

Software Research, Development, Testing, and Education

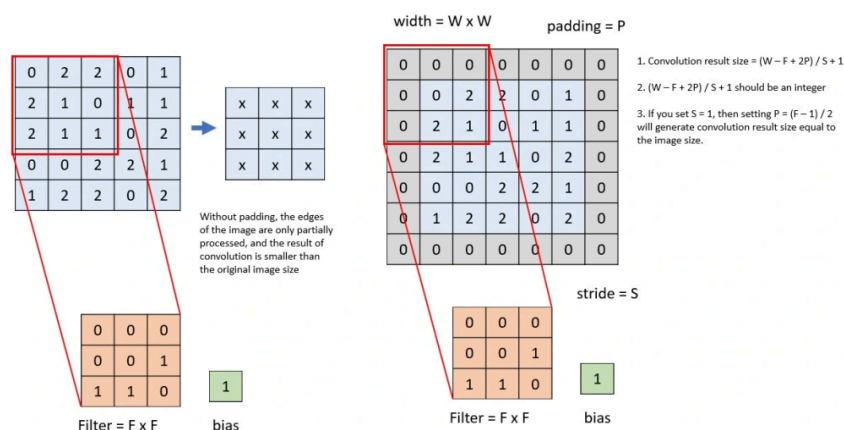


## Convolution Image Size, Filter Size, Padding and Stride

A convolutional neural network (CNN) applies a filter to an image in a very tricky way. When you use a CNN you have to be aware of the relationship between the image size, the filter size, the size of the padding around the image, and the distance the filter moves (the stride) during convolution.

Without image padding, the pixels on the edge of the image are only partially processed (which may be OK), and the result of convolution will be smaller than the original image size (usually not good).

I'll cut to the chase and give the key facts.



Suppose an image has size  $W \times W$ , the filter has size  $F \times F$ , the padding is  $P$ , and the stride is  $S$ . Then:

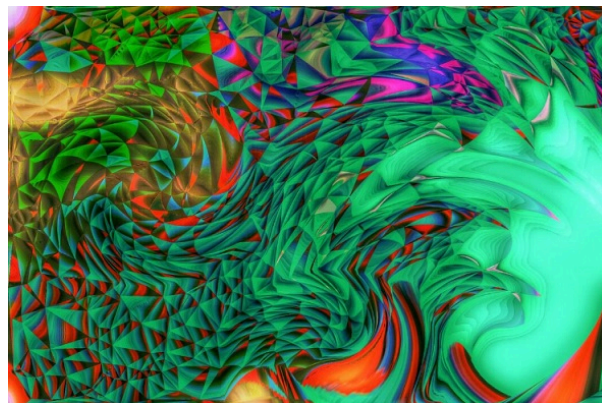
1. The result size of a convolution will be  $(W - F + 2P) / S + 1$ . For example, if an image is  $100 \times 100$ , a filter is  $6 \times 6$ , the padding is 7, and the stride is 4, the result of convolution will be  $(100 - 6 + (2)(7)) / 4 + 1 = 27$ .

$$/ 4 + 1 = 28 \times 28.$$

2. Therefore, the quantity  $(W - F + 2P) / S + 1$  should be an integer, and so  $(W - F + 2P)$  should be evenly divisible by  $S$ . This will never be a problem if  $S = 1$  but could be a problem if  $S$  is greater than 1.

3. If you set  $S = 1$  (very common), then by setting  $P = (F - 1) / 2$  the result size of convolution will be the same as the image size (which is usually what you want). If  $S$  is greater than 1, then you need to adjust  $P$  and/or  $F$  if you want to retain the original image size.

When I sat down to write this blog post it was my intention to explain exactly where these relation equations come from. But I quickly realized that providing a full explanation would take a couple pages of text (at least). So, instead I'll just say that when you work with convolution, you can't just use any values for your filter, padding, and stride.



*"Convolutional Synchronicity", Sheila Robbins.*

May 30, 2018

---

[« Previous](#)

[Next »](#)



PGT-ART on June 1, 2018 at 7:50 am

It seems not so complex, so without math symbols.

Be aware with convolution operations that the tile centre has a border, apply the border to the

original image (usually by mirroring the border). (ea copy original image to larger image).  
Or leave the border a background value or color (if it needs to future-less)

Be aware that its not ideal to train your net with border containing tiles, since those borders are made up (mirrored data), so they are not real data. Corner cases are often less optimal in graphics functions.

If you want to make Neural Network image denoiser, then get into contact with Blender Developers.

Blender is a popular opensource 3D editor.

[View Full Site](#)

[Blog at WordPress.com.](#)