

# Intuition Report 4

Ramtin Mojtahedi Saffari - 20307293

---

Assessed link 1: [What is Convolutional Neural Network](#)

## Self-reflection

- **Short Summary**

The link provides a great interactive tool to visualizing and manually setting a convolutional neural network. Also, the user can personalize images that wants to be given to the network. The interactive tool can clearly visualize the concepts of input layers, convolutional layers, padding, kernel size, stride, activations functions, pooling layers, and flatten layer.

- **Hypothesis and Expectation**

I hypothesized that that by increasing the padding, the size of the output image is increased. Also, when we increase the stride, the resolution of the output image will be decreased.

- **What I Achieved**

There should be a tradeoff between the size of the padding and kernel size that affects the stride. When we increase the input size, the kernel size and the padding

size should be increased. In this way as the stride size being increased, the resolution of the output image becomes decreased. In addition, when the padding size is being increased with a fixed size of input image, the kernel and stride is being decreased with respect to the padding size. In this sense, the size of the output image increased.

- **What I Learned**

The adjustment of the hyperparameters in a CNN is very important. I found that adding padding is very helpful in prevention of shrinking size of the image in every convolutional process. Also, it guarantees that the edge pixels of an image being enough considered in the process [1]. On the other hand, padding can increase the computational costs of the network.

Another important hyperparameter in the network is stride. The stride's size affects the CNN output image because nearby pixels are strongly correlated (especially in the lower layers), it makes sense to minimize the output size by subsampling (pooling) the filter response. The greater the distance between two pixels, the less associated they are. As a result, a large step in the pooling layer results in significant information loss [2]. This translates into the initial hypothesize that larger stride will decrease the resolution of the output image. Therefore, making tradeoff between these hyperparameters is significant to both keep the quality information of the image as well as optimizing the performance of the network.

### **Suggestion and Filling the Gaps**

I found the interactive link very interesting and helpful in learning-by-practice the basic concepts in a neural network. I recommend that authors can expand the

visualization for the features such as what is proposed in [3]. Overall, I found the link extremely helpful.

---

**Assessed link 2:** [Draw the Number](#)

## **Self-reflection**

- **Short Summary**

The website provides interactive tool for drawing our customized number inside the box and then predicts top two guesses for the number. The proposed tool predicts the drawn number using CNN network that its layers are visually shown to the user in order to get a better sense of the network's functionality as well as what is happening in each layer of the CNN.

- **Hypothesis**

I hypothesized that in this handwritten drawing of a number, the direction and length of the drawings are important factors that affect the outcome of the prediction.

- **Testing the hypothesize and what I have learned**

The provided interactive tool can predict the drawn digit with a high accuracy in most of the cases. After several experiments, I understand that the direction of the drawn digit and the length of the digit's parts affect the outcome of the prediction. As an example, when the number '4' is being drawn while its bottom part deviated to the left or right rather than being perpendicular, the network predicts it as '9'. Particularly when the deviation becomes larger. This happens despite the fact that

the top shape of '4' is clearly different with '9'. Also, the network can predict the digits between 0-9 and if we input two numbers, it considers the last drawn digit. The other challenge that I found is when there is a discretion in the drawn number. This causes large error in the predicted number and network mostly predicted based on the angle of the drawn number and their length. Since 1998, one of the well-known datasets that is used for training handwritten digits is called Modified National Institute of Standards and Technology (MNIST) with 58,527 digits written by 500 different writers [4]. Probably this dataset has been used for training of this network. There are a number of methods to increase the level of accuracy and performance of the proposed network using this dataset. One of the methods is changing the learning rate of the network. Also, we can use batch normalization in the network. After convolutional and fully connected layers, Batch normalization can be utilized. It has the effect of altering the layer's output distribution, especially by standardizing the outputs. This has the effect of speeding and stabilizing the learning process. the other improvement method is increasing the depth of the proposed model such as number of convolutional layers followed by a max pooling layer [5]. These methods can improve the model's performance and prediction outcome of the network.

### **Suggestion and Filling the Gaps**

Similar to what I mentioned before, I found the interactive link very interesting and helpful in learning-by-practice the basic concepts in a convolutional neural network for prediction of handwritten digit. I suggest that the developed add some explanations to each of the proposed interactive tools. For example, they can use the materials in the reference [6] to explain the used layers and parameters in the proposed interactive tool and network.

## Self-evaluation:

In this intuition report, I have gone through an in-depth analysis of two of the provided links. In my assessment, I have completely considered the required expectation for deep exploration, including proposing a hypothesis and what I expected, reporting on what I achieved and explored, in-depth discussion of what I have learned, and providing gaps and recommendations to fill them. Considering the quality and assessment level, I deserve to get the full mark (4 points) for this intuition report.

In advance, thank you very much for your time and consideration of this report.

Best regards,

Ramtin

## References

- [1] DeepAI. (2020, June 25). *Padding (Machine Learning)*. <https://deepai.org/machine-learning-glossary-and-terms/padding>
- [2] P. (2020, February 5). *Understanding of Convolutional Neural Network (CNN) — Deep Learning*. Medium. <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>
- [3] Olah, C. (2019, August 28). *Feature Visualization*. Distill. <https://distill.pub/2017/feature-visualization/>
- [4] Wikipedia contributors. (2021, August 20). *MNIST database*. Wikipedia. [https://en.wikipedia.org/wiki/MNIST\\_database](https://en.wikipedia.org/wiki/MNIST_database)
- [5] Brownlee, J. (2020, August 24). *How to Develop a CNN for MNIST Handwritten Digit Classification*. Machine Learning Mastery. <https://machinelearningmastery.com/how-to->

[develop-a-convolutional-neural-network-from-scratch-for-mnist-handwritten-digit-classification/](#)

- [6] Madiraju, N. (2020, July 15). *Digit recognizer using CNN - Towards Data Science*. Medium.  
<https://towardsdatascience.com/digit-recognizer-using-cnn-55c65ca7f9e5>