HIGH SCHOOL EDITION
Journal of Student Research



# The CNN: The Architecture Behind Artificial Intelligence Development

Yechan Lee

Dublin High School, USA

## ABSTRACT

The convolutional neural network (CNN) is a multilayer network architecture that is capable of training itself using an advanced algorithm to produce increasingly accurate results. The CNN is especially effective in spatial image recognition and is used in a multitude of fields, such as image recognition in the medical industry, and image segmentation in the security field. The ability of the CNN to show impressive results is seen in its multilayer composition. This multilayer network consists of the convolutional layer, the pooling layer, the activation layer, and the fully connected layer. Here, the convolutional layer, the pooling layer, and the activation layer have their own parameters which will influence the flow of input data throughout the CNN until it produces an ultimate output in the connected layer. As such, this paper will delve into the individual characteristics of each layer, and introduce its relationship with not only its immediate surrounding layers but with the entirety of the CNN. Additionally, this paper will stress important modules and parameters that can help improve the layers.

## Introduction

Since the 18th century, humankind has undergone four types of industrial revolutions. The first industrial revolution, which occurred in 1784, first broke the boundaries of the world through mechanization (Crafts, 1996; Groumpos, 2021). The second industrial revolution which followed in 1870 utilized electricity and the mechanized infrastructure built during the first industrial revolution to bring about mass production (Mohajan, 2019). One hundred years after the second industrial revolution, in 1969, automatons and computers entered the mainstream and led human society into the 3rd industrial revolution (The Strategic Foresight Initiative, 2013; Mohajan, 2021). Since the third industrial revolution, personal computers had slowly started to enter workplace desks and into homes, but it wasn't until the wide distribution of smartphones, personal devices which could be carried on a person and could connect people with everybody else in the world, that humanity would enter the fourth industrial revolution in 2011. An astute reader would have noticed that the time between each revolution is becoming shorter, and this is why people have already been hearing the rumblings of the fifth industrial revolution.

The fourth industrial revolution introduced computers and digitized personal libraries to society and improved the output and efficiency in various fields such as business, industry, and science (Tan and Shang-shu, 2017). However, the calculating method by which such prosperity was brought to society gave rise to the question, "Who is really in charge?" Computers were meant to help humans, not squeeze them of every ounce of energy in the name of efficiency. "Will the machine turn on its makers?" (Schmelzer, 2022). The looming fifth industrial revolution, which focuses extensively on artificial intelligence, consequently underlined the importance of cooperation between humans and artificial intelligence. Naturally, the interest in the mechanisms and algorithms brought about a new field where teaching artificial intelligence was the focus.

The fifth industrial revolution emphasizes the cooperation between the human and the machine and between efficiency and output (Aliet al, 2022; Callaghan, 2022; Nobleet al, 2022. While it seems to operate with the same parameters as the fourth industrial revolution, the fifth industrial revolution focuses on cooperation rather than replacement. As humanity entered the 21st century, it became exceedingly apparent that man and machine will not part ways. While initial conversations were about specific jobs being replaced by automatons, the reality was that technology enriched the life of an ordinary person. However, that is not to say that automated services did not replace human personnel at certain stations. The replacement of cashiers with self-calculating kiosk machines which can explain and process payment by themselves comes to mind (Lewis, 2018). Nevertheless, more often than people realize, they have been helped by the Algorithm more than once. The degree of how much artificial intelligence algorithms can help society will ultimately depend on the abilities of its makers.

As a result of the looming fifth industrial revolution, scientists, researchers, and even multinational companies have been actively participating in improving the machine algorithm. More specifically, these factions have been trying to make protocols which are deemed efficient; protocols which ease the burden of human initiation while providing stellar output results. As a result of this search, the scientific society reached back into time to pull out the term machine learning. Machine learning is one method by which humans try to improve artificial intelligence by adopting an off hands approach. The earliest machine learning program was a checker game program. This seemingly innocuous program was capable of learning from its human opponents by accumulating data and therefore anticipating its human opponent's actions. Its chess counterpart is an IBM machine known as Deep Blue, and it rose to infamy when it beat chess champion and grandmaster Garry Kasparov in a six-game match that lasted for several days (Newborn, 1997; Campbell, 2002). The story of Deep Blue was extraordinary at the time, as it was a revenge match for the machine. Learning from their loss to Garry Kasparov in 1996, the makers of Deep Blue upgraded its calculating power to process up to 200 million chess positions per second and in 1997, won against its world champion (IBM100, 2023).

In machine learning, the machine, or artificial intelligence, is capable of continuous learning without the need for human teachers. Higher-end algorithms have been developed for various fields such as text analysis to monitor child abuse, cell image classification, and video recommendations that reflect the user's preference (Deianaet al, 2015; Geetha & Sendhilkumar, 2023). The last example is one of the first ways an individual who does not specialize in computer engineering first encounters machine learning. Indeed, something as common as grammar suggestions and chatbots, are all products of machine learning.

In this manuscript, we will explain the concept and the purpose of machine learning. We will then consequently lead to the introduction of deep learning, a subset of machine learning that is currently the mainstream of artificial intelligence education and algorithm processing. Finally, we will introduce a comprehensive review of the main architecture of machine learning and deep learning paradigm, the central neural network (CNN).

## Machine Learning

The terminology machine learning has been met with both curiosity and with trepidation. Curiosity from the notion of how far we can teach our artificial peers, and fear of what would happen if our new peers deem us a threat (Ford, 2020). This fear is not without reason, as the main principle behind machine learning is that artificial intelligence continues learning with minimal human intervention. However, this concept also means that the learning parameters of machine learning can be set by human restrictions (Wandelt, & Bailer-Jones, 2008; Theodoridis, 2015). In terms of efficiency, such human interaction is both a blessing and a curse. While the range of customization is broad, when the amount of input data increases the workload of the human programmer increases exponentially. Thus, feature engineering, a process where a machine learning model learns which variables to utilize for efficient data processing, was one of the first techniques studied.

This process and others that strive to improve the functionality of artificial intelligence have been under observation for several decades, but only recently have computer engineers been able to make strides in this field. An unseen growth in the computing power of processing units had heralded new techniques which had been previously thought to be impossible due to the limited technology of that time. Increased processing power allowed for greater and more complex algorithms capable of advanced functions, and artificial intelligence based on machine learning algorithms started to initially appear in biological and chemical scientific fields. Naturally, this was because the accumulation of data in these fields and the accumulation of data in machine learning initially came from human effort.

The cumulation of such techniques eventually points to an end goal: making an artificial intelligence model capable of imitating humans by acting like humans. To do so, an even more specialized version of machine learning known as deep learning was pioneered. The difference between deep learning and machine learning can be seen in the way they approach data sets and their limitations. As mentioned before, machine learning depends greatly on the human in charge of data inputs, and as such, is limited by the computer engineer themselves; only a small amount of data can be processed at a time, and the data processing method is linear (Carbone, 2022; Onyemaet al, 2022). On the other hand, deep learning was designed to circumvent the limitations of machine learning by applying a learning process that was very similar to how humans process data. As a result of this new data processing method, deep learning was capable of self-checking its own progress and learning from its own mistakes. The advantage of such a method minimized human interaction, but this also meant that a substantial dataset was required to initiate a deep learning algorithm, and a longer "training" period would be required. However, machine learning was not abandoned completely and it is currently still being employed in a number of different fields such as chemistry, linguistics, medical imaging, and even human paleontology, (Schmidtet al, 2019; Liet al, 2022; Varoquaux, & Cheplygina, 2022; Assaelet al, 2022).

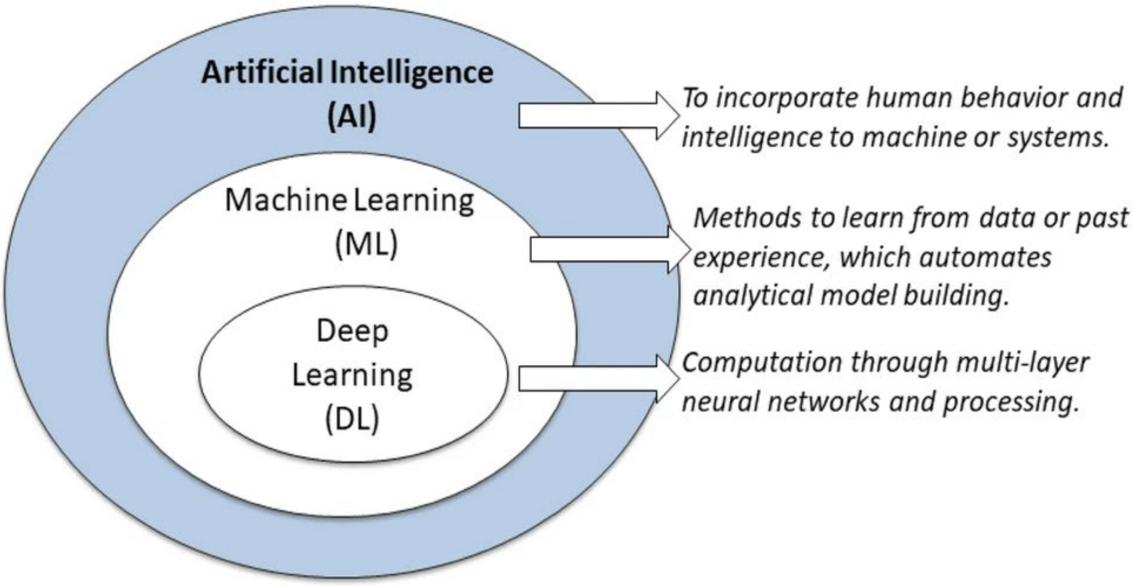


**Figure 1.** What is the difference between artificial intelligence, machine learning, and deep learning (Sarker, 2022)

## Deep Learning

The convolutional neural network (CNN) is the background on which the deep learning algorithm is built upon. However, while CNN is also used in machine learning, the deep learning version provides a better alternative due to how it utilizes transformations and graph technologies to modify its perimeters, and more importantly, it can "learn" in a multi-layer way (Alzubaidi et al., 2021). Add this to the fact that feature extraction, a process where an image is transformed into numbers while retaining the original information, is automatic and not manual, allowing the engineers to focus their efforts on other portions of the deep learning algorithm (Shaheenet al, 2016. In short, automated feature extraction uses algorithms to extract features from incoming signals and images without the need for a human counterpart, while manual feature extraction, like its namesake, requires a manual input; Hence, the former can be thought of as deep learning while the latter can be classified as machine learning.

The architecture of the algorithms utilized during the automated feature extraction are multi-layered, and low-layered features as analyzed in the initial layers, and high-layered features are analyzed in the latter portion of the algorithm. The result of this advanced algorithm is minimalized human intervention and self-sufficient neural network training.

## Convolutional Neural Network (CNN) Architecture

The central neural network is the core pattern behind the advancement of artificial intelligence; more specifically, the CNN has been the basis upon which various deep learning algorithms have been built. The greatest advantage of the CNN is its ability to process spatial information. Spatial information is information related to the distance between two objects, be they locations, people, or objects. The increasing accuracy of Global Positioning System (GPS) locations on portable devices and crispier pictures of Magnetic Resonance Imaging (MRI) are all results of advancing CNN technology (Malleswaran, 2011; Gassenmaier, 2021). Society has been utilizing artificial intelligence in its daily work, but the average person does not understand the depth of the relationship between humans and machines. Other examples of CNN usage in everyday life include image recognition, cybersecurity protocols, and artificial assistants (Sharma and Jakovljevic 2018; Nedeljkovic & Jakovljevic).

The main structure of a CNN is composed of four layers: the convolutional layer, the pooling layer, the activation layer, and the fully connected layer (Taye, 2023. In the past, rudimentary CNNs were composed of just three layers, the convolutional layer, the pooling layer, and the fully connected layer, but recent improvements in the artificial intelligence algorithm have introduced the function of the activation layer, which will be explained in greater detail below. As mentioned before, the CNN was modeled after the human neuron system in the brain, and the layers mentioned above correspond to the layers in the visual cortex.

Aside from having similar characteristics to a human neuron system, the CNN of artificial learning has a major advantage over its fully connected network peers. The CNN has a special method by which it utilizes its 2D-input structure which assesses the image signals. By sharing weights and local connections, the CNN is capable of not only simplifying its learning process but also increasing the speed of its learning process. An astute biologist will recognize that this pattern is exactly how biological visual cortex cell work. A visual cortex cell does not take into account the whole picture. For example, if there is a picture of a cat, a human's biological visual cortex cells will not initially take the entirety of the cat into account. Rather the biological visual cortex cells assess a small portion of the cat, and then constantly recheck the input data.

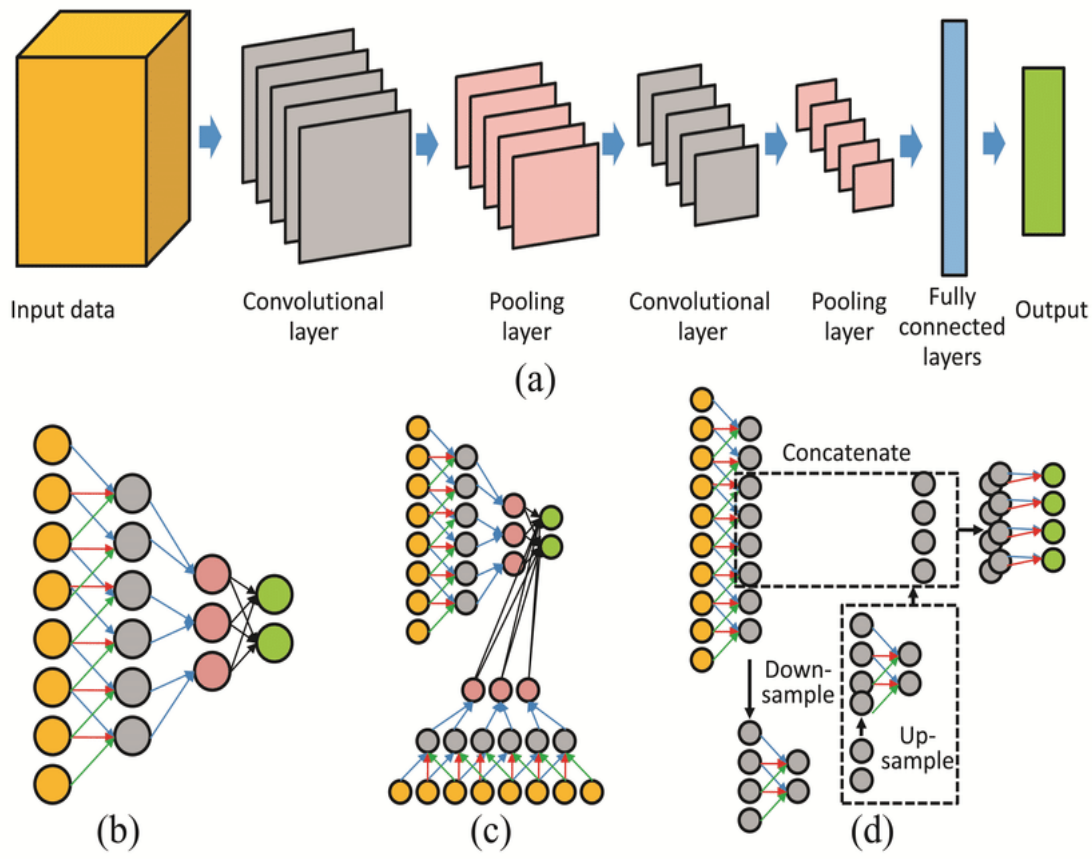


**Figure 2.** Examples of different CNN architecture (a) General CNN architecture; (b) Basic node placement for CNN architecture; (c) Node placement for double-stream CNN architecture; (d) Segmentation CNN nodeplacement (Monkam et al., 2019)

After a portion of the initial input layer has been analyzed, the image is further processed through multiple convolution layers. Although subsequent processes are not indicated in the image above, recent CNN have sub-sampling pooling layers after the convolution layer, and full connection layers after that.

In a CNN model, an input layer has 3-dimensional aspects. It possesses a width and length of equal value and has a depth or height thickness. The main reason the dimensions of an input layer have to be set is that only a small portion of the input is initially sampled (Figure 2). The sample, otherwise known as a kernel, will have smaller dimensions than the input layer. For example, if the input layer has a dimension of L x L x H, where L is the width and H the height of the input layer, the kernel will have dimension M x M x R, where M is the width and height of the kernel and R is the depth of the kernel. In terms of magnitude, M is smaller than L and R is smaller or equal to H. The sudden dive into the technical terms may confuse the reader at this point, so it is best to summarize the advantages of the CNN before its multilayer structure is dissected. The CNN, and as a result deep learning, has three major advantages: 1) CNN has enhanced generalization capabilities due to the presence of a weight sharing feature. 2) The extracted features (output) of the CNN are highly accurate because the CNN architecture utilizes both the feature extraction layer and the classification layer. 3) As a result of generalization and accurate extraction values, the CNN can be deployed in larger environments (Alzubaidi et al., 2021).

## Convolution Layer

To understand the mechanics of a convolution layer, a computer engineer must understand what a kernel is in a CNN. Elementary speaking, the kernel is a filter that takes a certain portion of the input data and sharpens the edges of an image. In more technical terms, the kernel is a matrix that scans and isolates a particular area of the input data, performs a dot product of the area, and then releases the output data (Sun et al., 2016; Milosevic, 2020). During the data isolation, the kernel moves through utilizing a particular mechanism known as the stride or as a stride value. This stride value ultimately influences the output pixel, and thus, the term stride and pixel are sometimes used interchangeably. For example, a stride value of 5 implies that the pixel output is going to be a result of a filter which has been moved by 5 units; the greater the stride value the lower the resolution.

The relationship between a convolution layer and a kernel, henceforth referred to as a convolutional kernel layer, is similar to sampling a large portion of a large picture; a convolution layer can thus be said to be made of many convolutional kernel layers. These small numerous kernel layers are known as receptive fields and are utilized to identify patterns of an image from various points of view (Xu and Wang, 2022). In the end, the convolutional kernel layers all pool their processed information together to make a big picture.

It is during this kernel layer amassing process that the weight sharing aspect of the CNN really shines. Weights are the numbers within a particular filter or matrix. During weight sharing, the kernel is slid over an input tensor, and by doing so, the CNN can identify common themes and motifs, such as outlines and patterns. In essence, weight sharing allows the CNN to understand the input target even though the "point of view" may change when different samples are taken from the input layer (Prasad et al., 2012; Malallah et al., 2021).

In short, a kernel is a m x m matrix which is smaller than the input layer's n x n matrix. For example, if an input layer is a 4 x 4 matrix, the biggest kernel possible is a 3 x 3 matrix. Additionally, this 3 x 3 kernel is capable of having four positions that correspond to the corners of the matrix. This means the output of this kernel will be a 2 x 2 = 4 pixel. If this kernel was a 2 x 2 matrix, the number of unique positions will be 9 and the output will be a 3 x 3 = 9 pixel output. When a kernel moves from one unique position to another unique position, this process is the sliding process mentioned above. Through trial and error, the kernel will eventually produce a matrix that produces a certain result such as image sharpening, edge detection, and object identification.

In the convolution layer, an additional process known as padding often occurs to determine border dimensions. Padding layers can be thought of as the crust layer of a slice of bread, and depending on how the padding layers dimensions are fine tuned, this bread crust can be thin (1 x 1 dimension) or thick. When an engineer modifies both the padding parameters and the stride, the engineer is capable of fine tuning a filter before the CNN starts its learning process (Hashemi, 2019).

## Pooling Layer

The primary mechanism of the pooling layer lies in sub-sampling a previous convolution layer to produce a layer that has even smaller dimensions. However it is important to note that although the horizontal and vertical dimensions of the conventional layer shrinks after a "pooling" process, the depth layer, or the depth does not change. The purpose of this pooling process, which reduces the dimension of the input layer even further, is to eventually decrease the parameters to such a degree, the amount of computation required decreases (Jia, 2022; Zafar et al., 2022). It can be thought of as a summary of a book; an efficient way to understand as much as possible with minimal effort.

Although there are a number of pooling techniques, max pooling, average pooling, and global pooling are the most commonly used pooling techniques. Before the three major techniques are explained, it must be noted that pooling is a give-and-take process. Because the pooling process is a method which specializes in one particular aspect, other factors often suffer as a result.

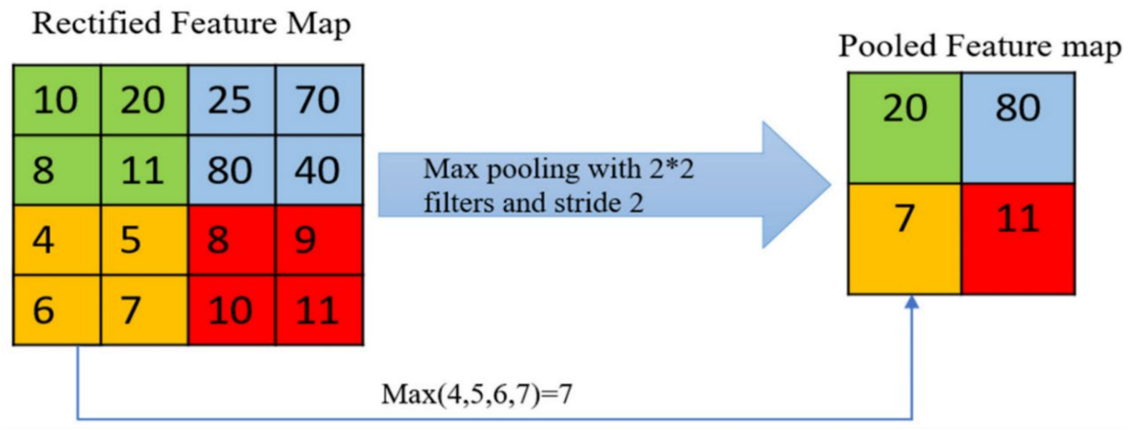


**Figure 3.** Example of Max Pooling (Zafaret al, 2022)

Max pooling is a pooling technique where the most prominent aspect of a region, or the greatest input values, is selected (Gholamalinezhad & Khosravi, 2020). As max pooling focuses on the most salient portions of an image, it is usually used to procure distinct edges. Average pooling is when the averages of a region are sub-sampled. While average pooling appears to be inferior to max pooling due to its lack of focus on specific qualities, average pooling is arguably capable of retaining the most information out of the three pooling methods. Due to its ability to retain the most information out of an input feature map, average pooling is mainly used for object detection (Özdemir, 2023). An example of average pooling is the "find all images of a fire hydrant" question many people encountered when they were logging into a particular site. The final pooling method, the global pooling method produces a single value from the pooling region. While max pooling and average pooling still retained a simple 3D structure of l x l x d, where l is the horizontal and vertical length and d is the depth, the global pooling method produces a result of 1 x 1 x d. Depending on whether the global pooling method was a global max pooling method or a global average pooling method, the resulting single value will differ greatly. As a result of this extreme simplification process, the global pooling method sometimes shows the greatest disadvantages of the pooling layer process: information loss and over-smoothing. This will affect the final classification process.

To summarize, the pooling layer is a process which simplifies or focuses on the convolutional layer that came immediately before it. If the convolutional seeks to attain specific details from the input layer, the pooling layer simplifies the data obtained in the convolutional layer. As such, sometimes there are multiple consecutive conventional layer-pooling layer complexes. The reason behind this is threefold: 1) To select the most important features deemed by the engineer, 2) to achieve translation invariance and discard the possibility of positional variation on the output, and 3) to reduce the required computing power by reducing the spatial dimensions.

## Activation Functions

The previous layers have been mostly regions where information is analyzed, processed, and summarized into certain values. It had been previously mentioned that the CNN architecture was influenced by the neural system of living beings. The activation function is the part where the neurons of the CNN system are activated. The activation function's goal is to act as a transfer function and produce an output that corresponds to the input values measured beforehand (Dubey et al., 2022).

The activation function can be either a linear function or a non-linear function, but non-linear functions are the main workhorse that drives the advancement of computer science forward. Most recent advancements in the artificial intelligence field that can easily be identified in everyday life, such as language processing, are because of these non-linear functions (Hao et al., 2020).

Simple activation functions have two types of transfer functions. They can be classified as either a linear function or a step function. The linear function is the simplest activation function available, and here, the output values are the same as its input values, and the step function can be seen as a binary function, where the output is either a 0 or a 1. As a result of its simplicity, the step function loses an incredible amount of data during its transfer process.

In contrast to the simple activation functions, the non-linear functions are both complex and numerous in number. First and foremost, the sigmoidal function is the easiest to utilize amongst the non-linear functions due to its output range of 0~1. This range implies that, unlike the step function which produces a binary output value as a response, the sigmoidal function is capable of producing a more complex and detailed result than simple activation functions. However, it is not without its disadvantages as the sigmoidal shape itself is a limitation to the output values. When the input values lean toward extreme values, the sigmoid functions give an output value that is close to zero. Additionally, the exponential nature of the sigmoidal function applies a significant amount of processing stress to the neural network and ultimately, hampers the computing speed of the algorithm.

The next non-linear function, the tanh function, addresses a critical factor of the sigmoidal function. Not only does the tanh have an output range that is twice that of the sigmoidal function (-1 ~ 1), the range of the tanh function centers the 0 value, a feature that was impossible in the sigmoidal function. As a result of these two characteristics, the tanh function is capable of shepherding input values through multiple hidden levels. Nevertheless, because the tanh function is also an exponential function, it requires significant computing power.

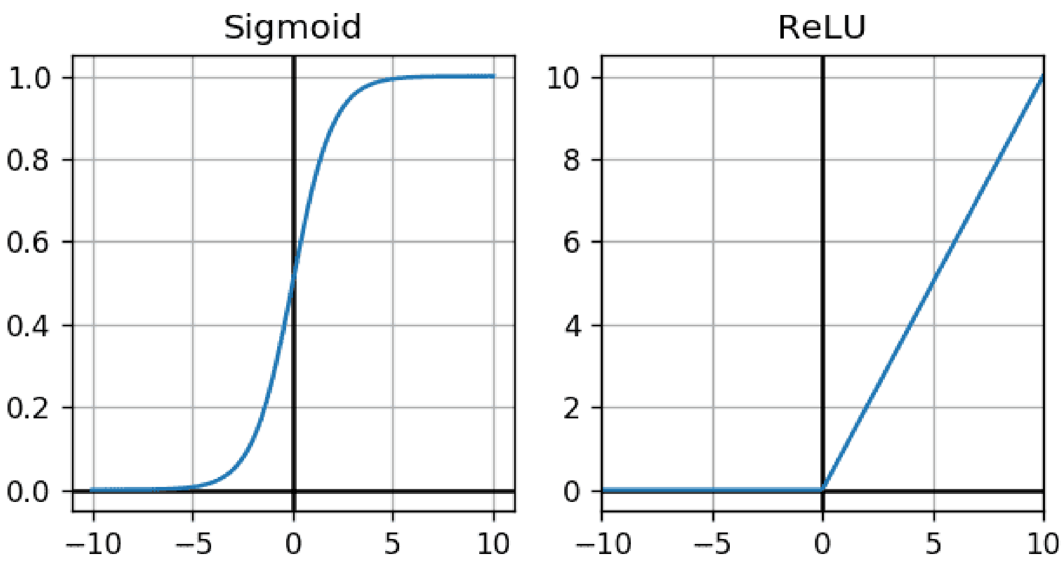


**Figure 4**. Side-by-side comparison of a Sigmoid function and a ReLU function. (Thakur, 2023) https://dhruvs.space/posts/ml-basics-issue-4/

While the exponential functions mentioned above both acknowledge and fix the limitations of the simple activation function, they have their own disadvantages. As such, a general function known as the Rectified Linear Unit (ReLU) function was made and then split into other ReLU-based functions. The ReLU function is capable of overcoming the gradients that formed during the small and large input values in the exponential functions. Additionally, it simplifies the computing process by streamlining the output values (Islam et al., 2021). Negative input values are set to 0 while positive values retain their original value. Even so, the ReLU itself has a disadvantage known as “dead neurons”. Dead neurons are neurons which do not correct and fix themselves during the artificial intelligence learning process. As such, the variations of the ReLU seek to address this dead neuron problem (Mahima et al., 2023). Due to the sheer number of ReLU variations, only two major ReLU variants will be introduced.

The Parametric ReLU is and ReLU variant which sacrifices economic computing processing for more precision. The Parametric ReLU tries to minimize the number of dead neurons by paying closer attention to the negative values which were treated as 0 by the base ReLU function. However, it is important to note that the parameters applied to the negative value for the Parametric ReLU function are also the main source of error. As such, the engineer in charge of the Parametric ReLU will have to pay specific attention to its build.

The Gaussian Error Linear Unit (GELU) function is the last activation function that will be introduced. As its name pimples the GELU function utilizes the Gaussian distribution during the processing of the input data, and consequently, the output of the GELU function ultimately depends on the sign of the input value and its weight.

### Fully Connected Layer

The fully connected layer was named due to its relationship with the layer that immediately precedes it. Whether the preceding layer is a convolution layer or a pooling layer, each neuron of the fully connected layer is connected to every neuron in the preceding layer. The fully connected layer is also the last layer in a CNN architecture, and thus, its output will be the final result.

## Discussion

To prospective computer scientists, understanding the main structure of the CNN architecture is the first step to eventually producing a CNN architecture of their own. The next step in building the CNN architecture will include the fine-tuning of various parameters mentioned for each layer. For example, the dataset optimization techniques, such as data dropout or data augmentation will be the first variable that will need to be addressed. Additionally, choosing or designing the algorithm will require expert knowledge of not only the optimizer itself but also on other enhancers which will be needed to, for the loss of a better word, optimize the CNN learning process.

Finally, as stated at the beginning of this manuscript, machine learning and deep learning require a significant initial data output; the more input data the more specific and correct the output. Once the initial batch of data is acquired, a teacher-student approach may be required for initial program training and transfer learning. Additional hyperparameters will be required to eliminate the unwanted noise in the algorithm, but this topic and others mentioned in the discussion portion of the manuscript will require their own review papers and to continue on is to do injustice to the professionals in this field.

The CNN architecture had been the golden standard upon which recent artificial intelligence, machine learning, and deep learning had their foundations built. In the past, artificial neural networks (ANN) and recurrent neural networks (RNN) were used to develop algorithms, but eventually, they fell out of favor as the input variety started to emphasize different characteristics. That is not to say, however, that the ANN and RNN are not used in contemporary machines. ANNs are still used to solve complex problems and RNNs are best used to analyze temporal data but when pure efficiency is considered, CNNs provide the best option while also specializing in spatial image analysis.

## References

Ali, S. H., Al-Sultan, H. A., & Al Rubaie, M. T. (2022). Fifth industrial revolution. *International Journal of Business, Management and Economics*, 3(3), 196–212. https://doi.org/10.47747/ijbme.v3i3.694

Assael, Y., Sommerschield, T., Shillingford, B., Bordbar, M., Pavlopoulos, J., Chatzipanagiotou, M., Androutsopoulos, I., Prag, J., & de Freitas, N. (2022). Restoring and attributing ancient texts using deep neural networks. *Nature,* 603(7900), 280–283. https://doi.org/10.1038/s41586-022-04448-z

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of Deep Learning: Concepts, CNN Architectures, challenges, applications, Future Directions. *Journal of Big Data*, *8*(1). https://doi.org/10.1186/s40537-021-00444-8

Callaghan, C. (2022). The Fifth Industrial Revolution: An Unfolding Knowledge Productivity Revolution. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4307753

Campbell, M., Hoane, A. J., & Hsu, F (2002) Deep Blue, *Artificial Intelligence,* 134(2), 57-83,, https://doi.org/10.1016/S0004-3702(01)00129-1.

Carbone, M. R. (2022). When not to use machine learning: A perspective on potential and limitations. *MRS Bulletin*, *47*(9), 968–974. https://doi.org/10.1557/s43577-022-00417-z

Crafts, N. F. R (1996). The First Industrial Revolution: A Guided Tour for Growth Economists. T*he American Economic Review,* 86(2), 197–201. http://www.jstor.org/stable/2118122

Deiana, A. M., Tran, N., Agar, J., Blott, M., Di Guglielmo, G., Duarte, J., Harris, P., Hauck, S., Liu, M., Neubauer, M. S., Ngadiuba, J., Ogrenci-Memik, S., Pierini, M., Aarrestad, T., Bahr, S., Becker, J., Berthold, A.-S., Bonventre, R. J., Bravo, T. E. M., … Weng, O. (2021, October 25). *Applications and techniques for fast machine learning in science*. arXiv.org. https://arxiv.org/abs/2110.13041

Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in Deep learning: A comprehensive survey and benchmark. *Neurocomputing*, *503*, 92–108. https://doi.org/10.1016/j.neucom.2022.06.111

Ford, P. (2020, April 2). *Our fear of Artificial Intelligence*. MIT Technology Review. https://www.technologyreview.com/2015/02/11/169210/our-fear-of-artificial-intelligence/

Geetha, T. V., & Sendhilkumar, S. (2023). Machine learning applications. *Machine Learning*, 295–315. https://doi.org/10.1201/9781003290100-12

Gassenmaier, S., Küstner, T., Nickel, D., Herrmann, J., Hoffmann, R., Almansour, H., Afat, S., Nikolaou, K., & Othman, A. E. (2021). Deep learning applications in Magnetic Resonance Imaging: Has the future become present? *Diagnostics*, *11*(12), 2181. https://doi.org/10.3390/diagnostics11122181

Gholamalinezhad, H., & Khosravi, H. (2020, September 16). Pooling methods in Deep Neural Networks, a review. arXiv.org. https://arxiv.org/abs/2009.07485

Groumpos, P. P. (2021). A critical historical and scientific overview of all industrial revolutions. *IFAC-PapersOnLine,* 54(13), 464–471. https://doi.org/10.1016/j.ifacol.2021.10.492

IBM100 - Deep Blue. (2023). https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/

Hao, W., Yizhou, W., Yaqin, L., & Zhili, S. (2020). The role of activation function in CNN. *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*. https://doi.org/10.1109/itca52113.2020.00096

Hashemi, M. (2019). Enlarging smaller images before inputting into convolutional neural network: Zero-padding vs. interpolation. *Journal of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0263-7

Islam, M. A., Wimmer, H., & Rebman, C. M. (2021). Examining sigmoid vs relu activation functions in deep learning. *Interdisciplinary Research in Technology and Management*, 432–437. https://doi.org/10.1201/9781003202240-68

Lewis, B. (2018, February 8). *Predictive maintenance for more resilient self-service vending*. https://www.insight.tech/retail/predictive-maintenance-for-more-resilient-self-service-vending

Li, Y., Choi, D., Chung, J., Kushman, N., Schrittwieser, J., Leblond, R., Eccles, T., Keeling, J., Gimeno, F., Dal Lago, A., Hubert, T., Choy, P., de Masson d'Autume, C., Babuschkin, I., Chen, X., Huang, P.-S., Welbl, J., Gowal, S., Cherepanov, A., … Vinyals, O. (2022). Competition-level code generation with AlphaCode. *Science*, 378(6624), 1092–1097. https://doi.org/10.1126/science.abq1158

Mahima, R., Maheswari, M., Roshana, S., Priyanka, E., Mohanan, N., & Nandhini, N. (2023). A comparative analysis of the most commonly used activation functions in deep neural network. *2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)*. https://doi.org/10.1109/icesc57686.2023.10193390

Malallah, H., Zeebaree, S. R., Zebari, R. R., Sadeeq, M. A., Ageed, Z. S., Ibrahim, I. M., Yasin, H. M., & Merceedi, K. J. (2021). A comprehensive study of kernel (issues and concepts) in different operating systems. *Asian Journal of Research in Computer Science*, 16–31. https://doi.org/10.9734/ajrcos/2021/v8i330201

Malleswaran, M., Vaidehi, V., & Deborah, S. A. (2011). CNN based GPS/INS data integration using new dynamic learning algorithm. *2011 International Conference on Recent Trends in Information Technology (ICRTIT)*. https://doi.org/10.1109/icrtit.2011.5972270

Milosevic, N. (2020). Convolutions and Convolutional Neural Networks. *Introduction to Convolutional Neural Networks*. https://doi.org/10.1007/978-1-4842-5648-0_12

Mohajan, H. (2019, October 21). The Second Industrial Revolution has brought modern social and economic developments. *Munich Personal RePEc Archive*. https://mpra.ub.uni-muenchen.de/98209/

Mohajan, H. (2021). Third Industrial Revolution Brings Global Development. 7. 239-251.

Monkam, P., Qi, S., Ma, H., Gao, W., Yao, Y., & Qian, W. (2019). Detection and classification of pulmonary nodules using Convolutional Neural Networks: A survey. IEEE Access, 7, 78075–78091. https://doi.org/10.1109/access.2019.2920980

Nedeljkovic, D., & Jakovljevic, Z. (2022). CNN based method for the development of cyber-attacks detection algorithms in industrial control systems. *Computers & Security*, *114*, 102585. https://doi.org/10.1016/j.cose.2021.102585

Newborn, M. (1997). Kasparov versus Deep Blue. https://doi.org/10.1007/978-1-4612-2260-6

Noble, S. M., Mende, M., Grewal, D., & Parasuraman, A. (2022). The Fifth Industrial Revolution: How Harmonious Human–machine collaboration is triggering a retail and service [r]evolution. *Journal of Retailing,* 98(2), 199–208. https://doi.org/10.1016/j.jretai.2022.04.003

Onyema, E. M., Almuzaini, K. K., Onu, F. U., Verma, D., Gregory, U. S., Puttaramaiah, M., & Afriyie, R. K. (2022). Prospects and challenges of using machine learning for academic forecasting. *Computational Intelligence and Neuroscience*, *2022*, 1–7. https://doi.org/10.1155/2022/5624475

Özdemir, C. (2023). Avg-topk: A new pooling method for Convolutional Neural Networks. *Expert Systems with Applications*, *223*, 119892. https://doi.org/10.1016/j.eswa.2023.119892

Prasad, P. S., & R. Upadhyay, A. (2012). Design of hybrid kernel and the performance improvement of the operating system. *International Journal of Engineering and Technology*, *4*(2), 162–165. https://doi.org/10.7763/ijet.2012.v4.340

Sarker, I.H. AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. S*n Computer Science*. 3, 158 (2022). https://doi.org/10.1007/s42979-022-01043-x

Shaheen, F., Verma, B., & Asafuddoula M. (2016). Impact of Automatic Feature Extraction in Deep Learning Architecture, *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Gold Coast, QLD, Australia, 2016, pp. 1-8, doi: 10.1109/DICTA.2016.7797053.

Sharma, N., & Jakovljevic, Z. (2018, June 8). *An analysis of convolutional neural networks for Image Classification*. Procedia Computer Science. https://www.sciencedirect.com/science/article/pii/S1877050918309335

Schmelzer, R. (2022, October 12). *Should we be afraid of ai?*. Forbes. https://www.forbes.com/sites/cognitiveworld/2019/10/31/should-we-be-afraid-of-ai/?sh=344b91414331

Schmidt, J., Marques, M. R., Botti, S., & Marques, M. A. (2019). Recent advances and applications of machine learning in solid-state materials science. *Npj Computational Materials*, 5(1). https://doi.org/10.1038/s41524-019-0221-0

Tan, T.-B., & Shang-su, W. (2017). The Fourth Industrial Revolution Explained. In PUBLIC POLICY IMPLICATIONS OF THE FOURTH INDUSTRIAL REVOLUTION FOR SINGAPORE (pp. 5–7). S. Rajaratnam School of International Studies. http://www.jstor.org/stable/resrep17650.5

Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions. *Computation*, *11*(3), 52. https://doi.org/10.3390/computation11030052

Thakur, D. (2019, September 4). *ML Basics #4: Replace negatives with Zeros!*. dhruvs space RSS. https://dhruvs.space/posts/ml-basics-issue-4/

Theodoridis, S. (2015). Parameter learning. *Machine Learning*, 327–402. https://doi.org/10.1016/b978-0-12-801522-3.00008-2

The Strategic Foresight Initiative. (2013). The Third Industrial Revolution. In Envisioning 2030: US Strategy for the Coming Technology Revolution (pp. 15–22). Atlantic Council. http://www.jstor.org/stable/resrep03584.8

Varoquaux, G., & Cheplygina, V. (2022). Machine Learning for Medical Imaging: Methodological Failures and recommendations for the future. *Npj Digital Medicine,* 5(1). https://doi.org/10.1038/s41746-022-00592-y

Wandelt, B. D., & Bailer-Jones, C. A. L. (2008). Precision parameter estimation and machine learning. *AIP Conference Proceedings*. https://doi.org/10.1063/1.3059073

Xu, C., & Wang, H. (2022). Research on a convolution kernel initialization method for speeding up the convergence of CNN. *Applied Sciences*, *12*(2), 633. https://doi.org/10.3390/app12020633

Zafar, A., Aamir, M., Mohd Nawi, N., Arshad, A., Riaz, S., Alruban, A., Dutta, A. K., & Almotairi, S. (2022). A comparison of pooling methods for Convolutional Neural Networks. Applied Sciences, 12(17), 8643. https://doi.org/10.3390/app12178643