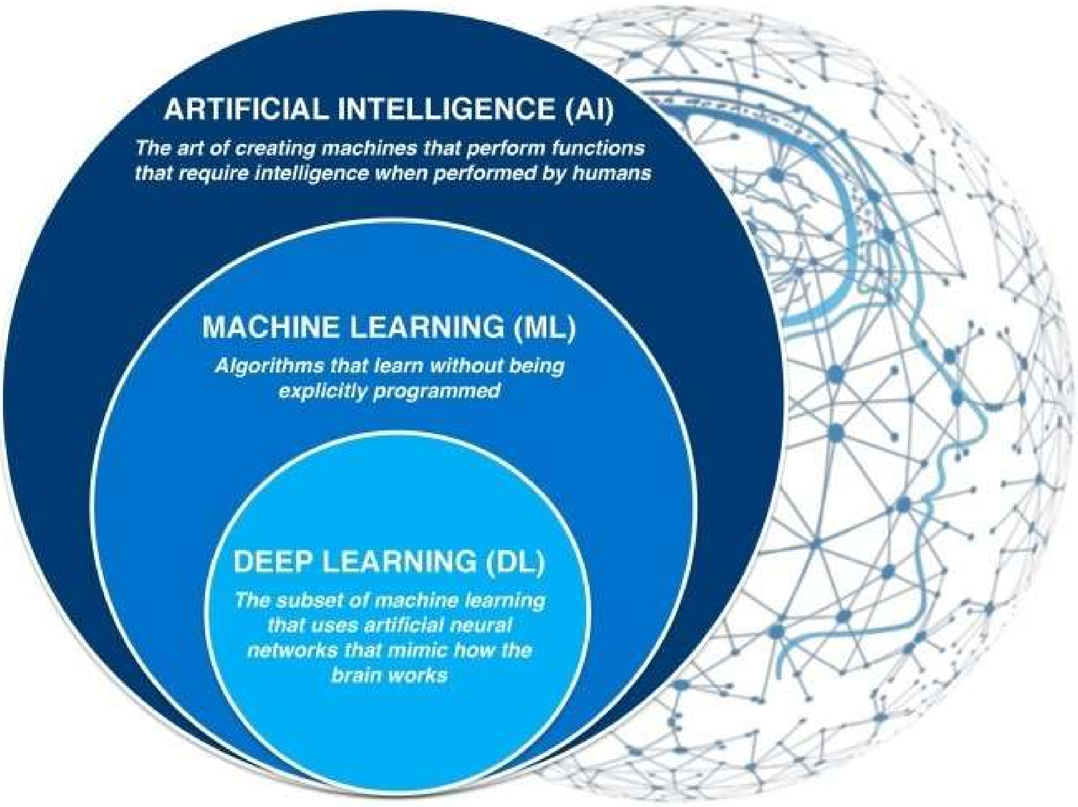
Artificial Neural Networks

Dictating Assistant / Neural Network Models (CNN)

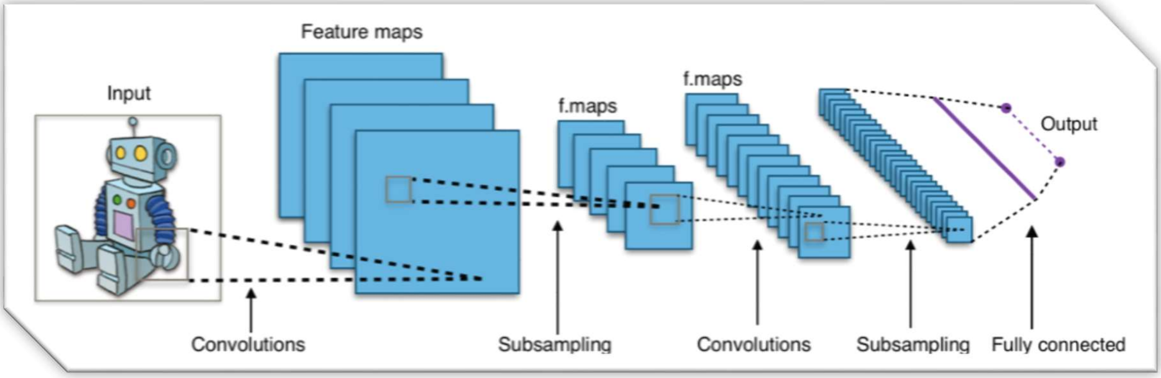
Name: Mert DUMANLI

Student ID: 160315002 – NORMAL


ARTIFICIAL INTELLIGENCE (AI)
The art of creating machines that perform functions that require intelligence when performed by humans
MACHINE LEARNING (ML)
Algorithms that learn without being explicitly programmed
DEEP LEARNING (DL)
The subset of machine learning that uses artificial neural networks that mimic how the brain works

Convolutional Neural Networks

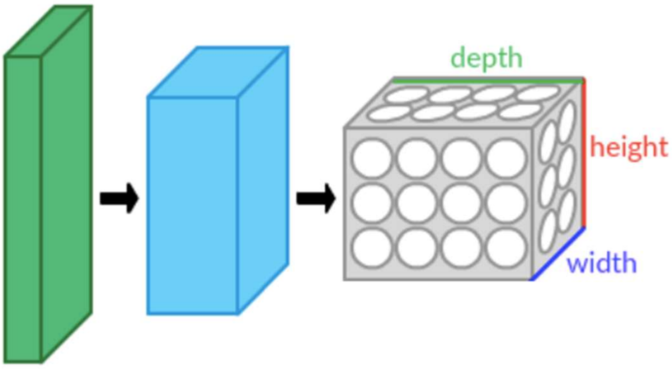
In deep learning, a convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analysing visual imagery. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics. They have applications in image and video recognition, recommender systems, image classification, medical image analysis, and natural language processing.



CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "fully-connectedness" of these networks makes them prone to overfitting data. Typical ways of regularization include adding some form of magnitude measurement of weights to the loss function. CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble more complex patterns using smaller and simpler patterns. Therefore, on the scale of connectedness and complexity, CNNs are on the lower extreme.

Convolutional networks were inspired by biological processes in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field.

CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage.



Definition

The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. Convolution is a specialized kind of linear operation. Convolutional networks are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers

Design

A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of a series of convolutional layers that convolve with a multiplication or other dot product. The activation function is commonly a RELU layer, and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers, referred to as hidden layers because their inputs and outputs are masked by the activation function and final convolution. The final convolution, in turn, often involves backpropagation in order to more accurately weight the end product.

Though the layers are colloquially referred to as convolutions, this is only by convention. Mathematically, it is technically a sliding dot product or cross-correlation. This has significance for the indices in the matrix, in that it affects how weight is determined at a specific index point.

Convolutional

When programming a CNN, the input is a tensor with shape (number of images) x (image width) x (image height) x (image depth). Then after passing through a convolutional layer, the image becomes abstracted to a feature map, with shape (number of images) x (feature map width) x (feature map height) x (feature map channels). A convolutional layer within a neural network should have the following attributes:

- *Convolutional kernels defined by a width and height (hyper-parameters).*
- *The number of input channels and output channels (hyper-parameter).*
- *The depth of the Convolution filter (the input channels) must be equal to the number channels (depth) of the input feature map.*

Convolutional layers convolve the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a specific stimulus. Each convolutional neuron processes data only for its receptive field. Although fully connected feedforward neural networks can be used to learn features as well as classify data, it is not practical to apply this architecture to images. A very high number of neurons would be necessary, even in a shallow (opposite of deep) architecture, due to the very large input sizes associated with images, where each pixel is a relevant variable. For instance, a fully connected layer for a (small) image of size 100 x 100 has 10,000 weights for each neuron in the second layer. The convolution operation brings a solution to this problem as it reduces the number of free parameters, allowing the network to be deeper with fewer parameters. For instance, regardless of image size, tiling regions of size 5 x 5, each with the same shared weights, requires only 25 learnable parameters. In this way, it resolves the vanishing or exploding gradients problem in training traditional multi-layer neural networks with many layers by using backpropagation.

Pooling

Convolutional networks may include local or global pooling layers to streamline the underlying computation. Pooling layers reduce the dimensions of the data by combining the outputs of neuron clusters at one layer into a single neuron in the next layer. Local pooling combines small clusters, typically 2 x 2. Global pooling acts on all the neurons of the convolutional layer. In addition, pooling may compute a max or an average. Max pooling uses the maximum value from each of a cluster of neurons at the prior layer. Average pooling uses the average value from each of a cluster of neurons at the prior layer.

Fully Connected

Fully connected layers connect every neuron in one layer to every neuron in another layer. It is in principle the same as the traditional multi-layer perceptron neural network (MLP). The flattened matrix goes through a fully connected layer to classify the images.

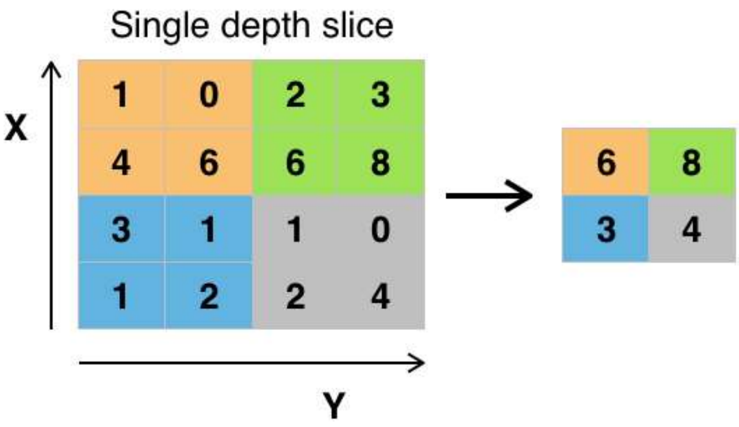
Receptive Field

In neural networks, each neuron receives input from some number of locations in the previous layer. In a fully connected layer, each neuron receives input from every element of the previous layer. In a convolutional layer, neurons receive input from only a restricted subarea of the previous layer. Typically the subarea is of a square shape (e.g., size 5 by 5). The input area of a neuron is called its receptive field. So, in a fully connected layer, the receptive field is the entire previous layer. In a convolutional layer, the receptive area is smaller than the entire previous layer.

Weights

Each neuron in a neural network computes an output value by applying a specific function to the input values coming from the receptive field in the previous layer. The function that is applied to the input values is determined by a vector of weights and a bias (typically real numbers). Learning, in a neural network, progresses by making iterative adjustments to these biases and weights.

The vector of weights and the bias are called filters and represent particular features of the input (e.g., a particular shape). A distinguishing feature of CNNs is that many neurons can share the same filter. This reduces memory footprint because a single bias and a single vector of weights are used across all receptive fields sharing that filter, as opposed to each receptive field having its own bias and vector weighting.

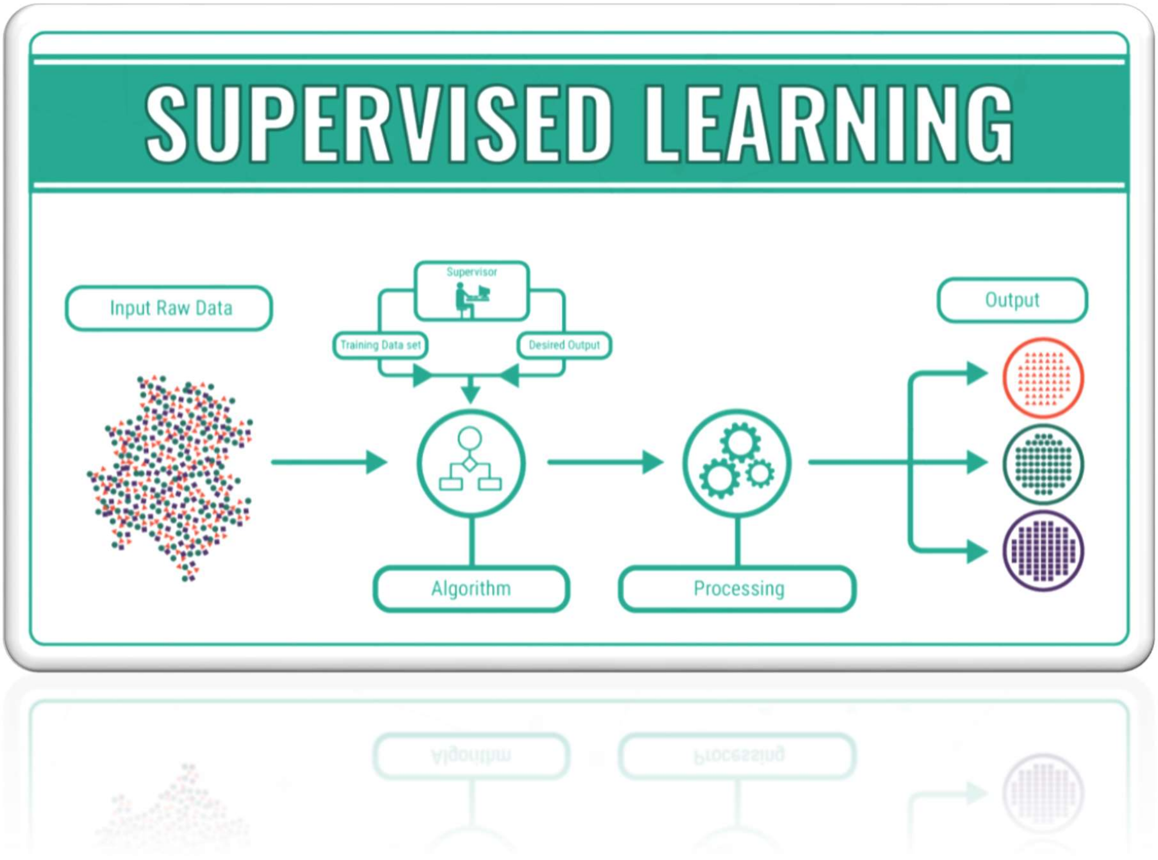


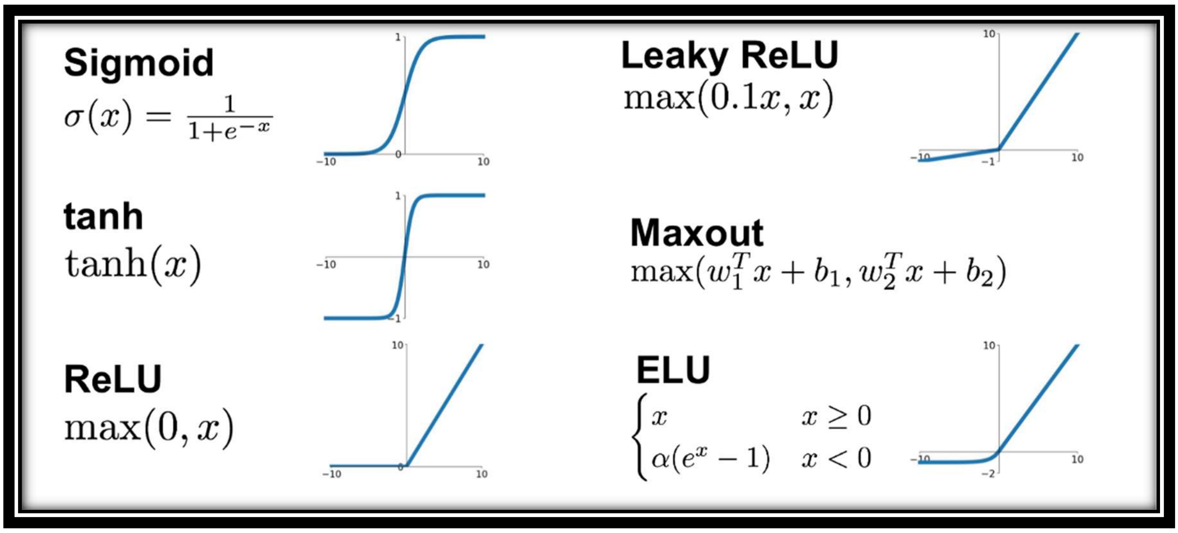
1	0	2	3
4	6	6	8
3	1	1	0
1	2	2	4

6	8
3	4

Supervised Learning

Supervised learning uses a set of paired inputs and desired outputs. The learning task is to produce the desired output for each input. In this case the cost function is related to eliminating incorrect deductions. A commonly used cost is the mean-squared error, which tries to minimize the average squared error between the network's output and the desired output. Tasks suited for supervised learning are pattern recognition (also known as classification) and regression (also known as function approximation). Supervised learning is also applicable to sequential data (e.g., for hand writing, speech and gesture recognition). This can be thought of as learning with a "teacher", in the form of a function that provides continuous feedback on the quality of solutions obtained thus far.





References

- *https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python?utm_source=adwords_ppc&utm_campaignid=1455363063&utm_adgroupid=65083631748&utm_device=c&utm_keyword=&utm_matchtype=b&utm_network=g&utm_adpostion=1t1&utm_creative=278443377095&utm_targetid=aud-392016246653:dsa-473406581915&utm_loc_interest_ms=&utm_loc_physical_ms=21070&gclid=Cj0KCQiAmZDxBRDIARIsABnkbYRGNsWivXSbTguZ-d5f_TGPOIWVPu5ULqAHhZEo8dQaoILaaOdj0pcaAuqCEALw_wcB*
- *https://tr.coredump.biz/questions/33885051/using-spyder-python-to-open-npy-file*
- *https://python-forum.io/Thread-ValueError-Found-input-variables-with-inconsistent-numbers-of-samples-5-6*
- *https://github.com/quantopian/pyfolio/issues/486*
- *https://stackoverflow.com/questions/43477496/found-input-variables-with-inconsistent-numbers-of-samples-489-1890-kneighbo*
- *https://datascience.stackexchange.com/questions/20199/train-test-split-error-found-input-variables-with-inconsistent-numbers-of-sam*
- *https://medium.com/@sddkal/keras-ile-ses-s%C4%B1n%C4%B1fland%C4%B1rma-ecs-50-ad8d3829f993*
- *https://stackoverflow.com/questions/52283840/i-cant-install-pyaudio-on-my-python-how-to-do-it*
- *https://ai.googleblog.com/2017/08/launching-speech-commands-dataset.html*
- *https://towardsdatascience.com/logistic-regression-classifier-8583e0c3cf9*
- *https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/*
- *https://medium.com/turkce/keras-ile-derin-%C3%B6%C4%9Frenmeye-giri%C5%9F-40e13c249ea8*
- *https://github.com/keras-team/keras-applications*
- *https://pathmind.com/wiki/neural-network*
- *https://www.wikizeroo.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQXJ0aWZpY2lhbF9uZXVyYWxfbmV0d29yaw*
- *https://towardsdatascience.com/understanding-neural-networks-19020b758230*
- *https://datafloq.com/read/machine-learning-explained-understanding-learning/4478?source=post_page-----f1009cad907c----------------------*
- *https://www.quora.com/Why-isnt-Logistic-regression-called-Logistic-Classification*
- *https://keras.io/*
- *https://medium.com/turkce/keras-ile-derin-%C3%B6%C4%9Frenmeye-giri%C5%9F-40e13c249ea8*
- *https://github.com/ozgung/nn-tutorial/blob/master/notebooks/keras-tutorial.ipynb*
- *http://aisdatasets.informatik.uni-freiburg.de/freiburg_groceries_dataset/*