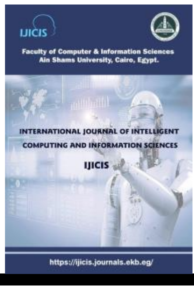
IJICIS, Vol.23, No.3, 59-72
DOI: 10.21608/ijicis.2023.211266.1274

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



Revolutionizing Medical Imaging through Deep Learning Techniques: An Overview

Salma Elgayar*
Computer Science Department, Faculty of Computer and Information Sciences, Ain shams university, Cairo, Egypt.
salma.elgayar@cis.asu.edu.eg

Safwat Hamad
Business Analytics Department School of Economic and Business Administration (SEBA), Saint Mary's College of California, Moraga CA-USA
Shh4@stmarys-ca.edu

El-Sayed M. El-Horbaty
Computer Science Department, Faculty of Computer and Information Sciences, Ain shams university, Cairo, Egypt.
shorbaty@cis.asu.ed.eg

Received 2023-05-15; Revised 2023-08-13; Accepted 2023-08-24

Abstract: *Medical imaging is a crucial tool for various clinical applications, including examination of medical issues, including early identification, monitoring, diagnosis, and therapy. To analyse medical images using computer vision, it is crucial to comprehend the ideas behind artificial neural networks and deep learning. In this particular article, we present recent principles of deep learning models and the different forms of widely used activation functions. Utilising Deep Learning Approach (DLA) has expanded quickly in the study of medical images, particularly in detecting either being present or not diseases. This study explores artificial neural networks development and provides a comprehensive analysis of DLA and potential uses for medical imaging. The majority of DLA implementations focus on pictures from digital histopathology, computer tomography, mammography, and X-rays.*
A thorough overview of the literature on the classification, detection, and segmentation of medical pictures using DLA is provided in the paper, which might help researchers decide what changes should be made to medical image analysis.

Keywords: *Convolutional neural network, Medical imagery, Deep learning, Classification, Segmentation, Detection.*

1. Introduction

The success of deep learning using medical image analysis has coincided with the growing prevalence of digital medical records. As evidenced by data from 2013 and 2019, the usage of electronic health records (EHRs) by both hospitals and physicians has significantly increased in recent years. However, despite this progress, there remains a shortage of qualified radiologists to analyze and interpret medical images such as nuclear medicine imaging, pathological tests, and positron emission tomography (PET). This process can be time-consuming and complex, but machine learning (ML) can offer a solution by using past data from a training dataset to make decisions without having programming instructions.

***Corresponding Author**: Salma Elgayar

Computer Science Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt

Email address: salma.elgayar@cis.asu.edu.eg

The field of machine learning comprises three primary learning methods: supervised machine learning, which involves algorithms like convolutional neural networks (CNNs) and Recurrent Neural Networks (RNNs); and unsupervised machine learning, which includes techniques like Deep Belief Networks (DBNs) and Restricted Boltzmann Machines.

Autoencoders, RBMs, GANs [1], and semi-supervised machine learning are all important techniques in the field of machine learning. These methods rely on domain experts to extract and select appropriate features for solving specific problems.
Feature selection problems can be solved using deep learning techniques, this enables for the feature extraction of input raw data [2]. The idea behind deep learning originated from cognitive and information theories. A part of artificial intelligence called "deep learning" has two main characteristics: first, it uses multiple processing layers to learn multiple abstracted features through these layers. Second, the learning feature approach of each layer is either supervised learning or unsupervised learning. Deep learning has been cited in many recent papers in the medical field, including MRI [3], cardiology, radiology, and neurology. Computer vision has influenced many forms of deep learning algorithms and has been applied to medical image analysis. These algorithms are particularly useful for detecting irregularities and identifying specific types of diseases. Convolutional neural networks (CNNs) are commonly used for tasks like object identification, registration, segmentation, classification, and others [4]. In medical pattern recognition, CNN is primarily used based on convolutional operation. Figure 1 displays visualized applications of some medical images [5].

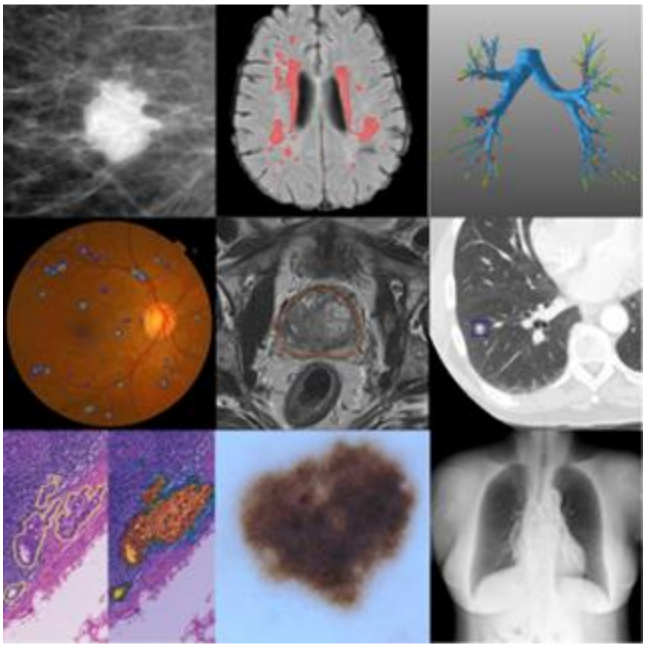


Fig.1 Deep learning usage has resulted in remarkable outcomes in various medical imaging applications, as illustrated in a collage showcasing some examples.

2. Neural Networks

2.1. History of Neural Networks

The field of Neural Network and Deep Learning is based on creating computer models that simulate the human brain. This idea was first explored by Warren S. McCulloch, a neurophysiologist, and Walter Pitts, a mathematician, in 1943. They created a basic neural network with an electric circuit that used threshold logic to mimic the human brain.

In 1949, Donald Hebb's book "Organization of Behavior" introduced the concept of updating synaptic weights in neural networks, which is now known as the Hebbian Learning Rule [6]. Frank Rosenblatt's 1958 paper on perceptron, the oldest neural network still in use, introduced the structure used to classify binary tasks**.**

Bernard Widrow and Marcian E. Hoff's 1960 paper [7] introduced the multiple Adaptive Linear Neuron (ADALINE) machine, which automatically adjusted its weights. In 1969, Minski and Papert highlighted the mathematical limitations of perceptron.

In 1974, Paul Werbose discussed the backpropagation concept for training [8]. In 1979, Fukushima designed a self-organizing neural network called Neocognitron, which had multiple convolution layers. Yann LeCun applied CNN and backpropagation for recognizing handwritten digits in 1989 [9].

Among of the most significant developments, Hinton et al. implemented the Deep Belief Network in deep learning in 2006. Restricted Boltzmann Machine for unsupervised learning, using greedy one layer at a time. Figure 2 summarizes the most important achievements in the development of neural networks, have led to the era of deep learning.

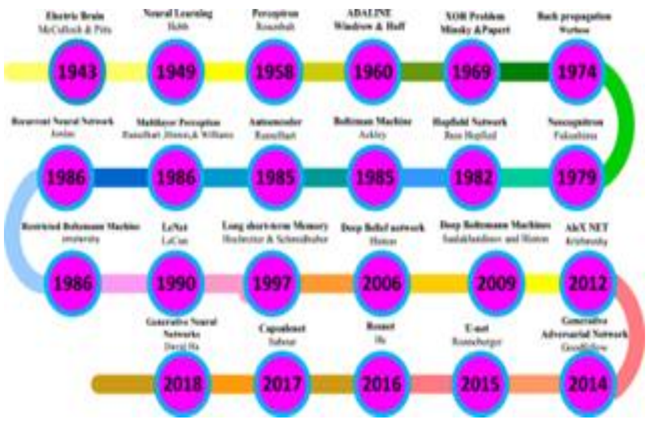


Fig.2 History of development of neural network [10].

2.2. Artificial Neural Networks

Deep learning algorithms are constructed using a foundational framework often referred to as artificial neural networks (ANN). These networks replicate the behavior of biological neural networks in humans, which is responsible for their exceptional performance. The ANN consists of interconnected neurons that carry out data processing, and these neurons are linked by weighted connections. The most basic model of a neuron can be expressed mathematically in Equation (1), and the perceptron is the most straightforward example of an artificial neural model, as depicted in Figure 3.

$$\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{w}\boldsymbol{x}^{\boldsymbol{T}} + \boldsymbol{b}) \tag{1}$$

Where $\boldsymbol{x} \in \boldsymbol{R}^{\boldsymbol{d}}$ the input feature vectors, $\boldsymbol{w} \in \boldsymbol{R}^{\boldsymbol{d}}$ is the interconnected weighted links vector, $\boldsymbol{b} \in \boldsymbol{R}$ the bias, The node's output is represented by a scalar value Y, and its activation function can be either linear or non-linear, denoted as f. The activation function f may involve certain numerical values enclosed in square brackets for specifying parameters or hyperparameters.

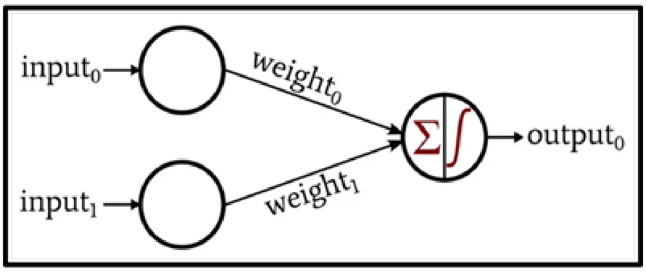


Fig 3. The perceptron [11]

We should define the following parameters for the construction of a neural network.

1- Neural network architecture single or multilayer.
2- The learning algorithm determines weights on the connections.
3- The kind of activation function used in a neural network depends on the type of connections between the nodes. Two common types of connections are Recurrent Neural Networks (RNN) and Feedforward Neural Networks (FFNN).

For supervised learning, classification, or regression, FFNN is employed[12]. The network consists of the input layer, the output layer, the final layer, and between the hidden levels. It may include one or more layers, each of which can have one or more nodes.
The primary distinction between FFNN and RNN is that the former uses a directed acyclic network, whilst the latter uses directed cycles to link its nodes.

3. Utilizing Backpropagation to Train a Neural Network (BB)

Neural network learning is based on updating connection weights between neurons in an iterative model to optimize loss function, updating weights are modified according to the network performance using the training dataset.
There are two phases for training forward and backward phase, forward activation function is selected while backward for calculating cost function and updating weights. Backpropagation used to learn input-output mapping using training samples in multilayer FNN.

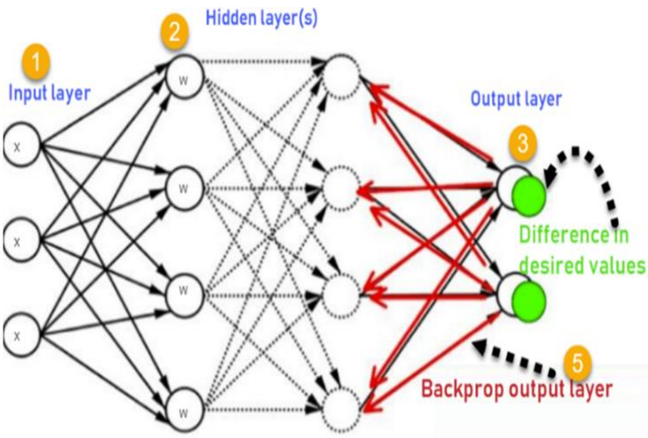


Figure 4 How Bp algorithm works.

Here is a sample of BB pseudocode explains one hidden layer NN.

1- Random weights initialization using small values.
2- Execute the steps repeatedly from 3 to 10 until the stopping condition becomes true.
3- For every set of training data consisting of an input and its corresponding desired output, iterate the feedforward propagation process from step 4 to step 9 inclusive.
4- Each of the input units (x_i, i = 1..., n) is responsible for receiving a specific input feature x_i and forwarding it to the subsequent hidden layer.
5- Each hidden neuron (y_j, j = 1..., p) calculate the net output using equation $Z_j = b_j + \sum_{i=1}^{n} w_{ij} * x_i$, Then apply it to the activation function $Z_j = f(Z_{j\_in})$.
6- Calculate the output signal for each output neuron, (y_k, k=1..., m), $Y_k = b_k + \sum_{j=1}^{p} w_{jk} * Z_j$, and calculate activation, $y_k = f(y_{k\_in})$.

Backpropagation propagation.

7- Given an input training pattern of n values (x_1, x_2…. x_n) and its matching output pattern of m values (y_1, y_2, …, y_m), we can define a target pattern (t_1, t_2, …. t_m). For each output value, a neuron in the network calculates the network error δ_k at output-layer neurons $\delta_k = (t_k - y_k)f'(y)$
8- Compute the error information term for every hidden neuron δ_j while doing so, use δ_k of the neurons at the output layer are determined by the neurons in the hidden layer from the previous step $\delta_j = f'(Z)\sum_{k}^{m} \delta_k * w_{jk}$

   To update the weights and biases in the output layer, use the following equations, where η represents the learning rate. For each output layer Y_k, where k ranges from 1 to m, update the weights (J = 0, 1, …P) and bias.
   $W_{jk}(new) = W_{jk}(old) + \eta*\delta_k * Z_j$,
   $b_k(new) = b_k(old) + \eta*\delta_k$
   Each hidden layer (Z_J, J = 1, 2, …p)
   updates its weights (i = 0, 1, …n) biases:
   $W_{ij}(new) = W_{ij}(old) + \eta*\delta* x_i$; $b_j(old) = b_j(old) + \eta*\delta_j$
9- Verify the halting criteria.

4. Activation Function

Artificial neurons utilize an approach known as activation function to receive and manipulate data. There are many types of activation functions based on the goal of the artificial neural network system. The optimal characteristic of the activation function is both non-linear and continuously differentiable. Non-linear to allow the NN to learn the non-linear relationship between training data, while the differentiability is important in backpropagation to optimize the new weights by minimizing the model's error using gradient descent cost function. Table 1 lists the most used activation functions.

5. Deep Learning Architectures

Deep learning battleground of artificial intelligence, it is experiencing aggressive development because of using modern architectures and GPUs (graphical processing units), to accelerate their execution. Deep Learning is a significant subfield of machine learning that focuses on utilizing multiple layers of neural networks [3 or more] to replicate the complex processes of the human brain. Figure 5 shows the general architecture of neural networks.

Table 1 the most used activation functions.

Function Name	Equation Function	Derivative Function
Sigmoid [13]	$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)\big(1 - f(x)\big)$
Hyperbolic tangent [14]	$f(x) = tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
Soft sign activation.	$f(x) = \frac{x}{1+\lvert x\rvert}$	$f'(x) = \frac{x}{(1+\lvert x\rvert)^2}$
Rectified Linear Unit (ReLU) [15]	$\boldsymbol{f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}}$	$\boldsymbol{f'(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}}$
Leaky Rectified Linear Unit [16]	$\boldsymbol{f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases}}$	$\boldsymbol{f'(x) = \begin{cases} \alpha, & x < 0 \\ 1, & x \geq 0 \end{cases}}$
Parameterized Rectified Linear Unit(PReLU) [17]	Same equation as Leaky RelU.	Using backpropagation to learn α.
Randomized Leaky Rectified Linear Unit [18]	$f(x) = \begin{cases} \alpha x, & x < 0 \\ x, & x \geq 0 \end{cases}$	$\boldsymbol{f'(x) = \begin{cases} \alpha, & x < 0 \\ 1, & x \geq 0 \end{cases}}$
Soft plus [19]	$f(x) = len(1 + e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$
Exponential Linear Unit (ELU) [20]	$\boldsymbol{f(x) = \begin{cases} \alpha(e^x - 1), & x < 0 \\ x, & x \geq 0 \end{cases}}$	$f'(x) = \begin{cases} f(x) + \alpha x, & x < 0 \\ 1, & x \geq 0 \end{cases}$
Scaled exponential Linear Unit (SELU) [21]	$\boldsymbol{f(x) = \lambda \begin{cases} \alpha(e^x - 1), & x < 0 \\ x, & x \geq 0 \end{cases}}$	$f'(x) = \begin{cases} f(x) + \lambda \alpha x, & x < 0 \\ \lambda, & x \geq 0 \end{cases}$

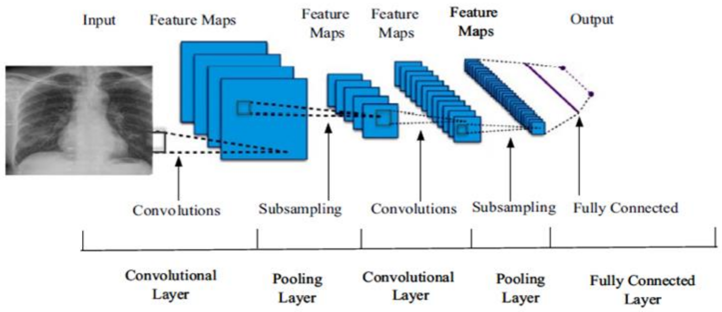


Figure 5 the general Design of NN.

Input layers, convolutional, pooling, activation function, and fully linked layers. are frequently seen in deep learning models, and output layers. However, there are certain architectures that may incorporate additional functional layers like normalization, dropout, cropping, and combination layers, as well as object detection and generative adversarial network layers.

The fundamental component of a successful network is the hidden layer(s), which are capable of representing complex data through their nodes and neurons. Access to only the input and output layers is provided, as the nodes in the hidden layer(s) are not known due to the absence of true values in the training dataset. The term "hidden layer" is used to refer to this fact. A neural network is required to have at least one hidden layer, and the optimal number of hidden units may be fewer than the inputs.

5.1. Deep Neural Network

Deep neural networks have a minimum of two layers in their architecture, allowing for the incorporation of nonlinear complexities. These networks are primarily used for classification and regression tasks. One of the major benefits of this model is its remarkable accuracy [22], which accounts for its widespread usage. However, a notable drawback is that during the training phase, numerous epochs are often necessary for the network to learn as errors from the backpropagation phase are transmitted back to the preceding layer, resulting in slow convergence.

5.2 Convolutional Neural Network (CNN)

The convolutional neural network (CNN) is a unique deep learning model that distinguishes itself from other neural networks. Its capacity to decrease the number of parameters is its main benefit, thereby enabling a deeper network with fewer parameters. CNN is especially valuable for visualizing and identifying medical image patterns and has become widely employed in image processing for healthcare professionals, following the exceptional performance of AlexNet in 2012 [23].
A significant challenge with fully connected feed-forward neural networks is that, even for shallow architectures, they can contain an impractically large number of neurons, making them unsuitable for image applications. Convolutional neural networks (CNNs), on the other hand, employ local connectivity and weight sharing, causing a significant decrease in the number of parameters compared to fully connected feed-forward neural networks. Additionally, CNNs contain three hyperparameters that regulate the output volume: depth, stride, and zero-padding.

- The depth of a convolutional neural network (CNN) is determined by the number of filters used, with each filter extracting a different feature from the input.
- The stride parameter dictates the number of pixels the filter is shifted by during convolution. When the stride is set to 1, the filter is shifted one pixel at a time.
- Zero-padding involves symmetrically adding zeros to the input matrix. This technique is commonly employed in the design of CNN layers to preserve input volume dimensions in the output volume. Square brackets indicate numerical values for each parameter.

Based on three key architectural ideas, CNNs are created: sparse interaction, shared weights, and local receptive fields. Sparse interactions involve using kernels smaller than the input size to detect only relevant features in the input image. This approach reduces the number of parameters, improves statistical efficiency, and minimizes memory usage.

Despite its benefits, training a CNN model can be challenging due to issues like overfitting, slow rate of coverage, exploding gradient, and class imbalance. Nevertheless, the key feature of CNNs is their ability to handle unstructured data through convolutional operations.
Convolution operation is processed on the input signal x(t) with filter h(t) which result in an output signal y(t) that may show more features than the input signal.
Here is the equation (2) of 1-D convolution of discrete signals x(t) and h(t).

$$\boldsymbol{Y}(\boldsymbol{t}) = \boldsymbol{x}(\boldsymbol{t}) * \boldsymbol{h}(\boldsymbol{t}) = \sum_{\boldsymbol{t}=-\infty}^{\infty} \boldsymbol{x}(\mathrm{T})\boldsymbol{h}(\boldsymbol{t}-\mathrm{T}) \tag{2}$$

Using the convolution function (3) on every convolutional layer to detect local non-linear features x_l from the previous input feature x_{l-1} by applying kernels k_l window, using the (*) convolution operation.

$$x_n^{(l)} = f(\sum_{m}^{M-1} x_m^{(l-1)} * K_{mn}^{(l)} + b_m^{(l)}) \tag{3}$$

where k_{mn} represents weights between two layers(l) with feature map n, and the previous layer(l-1) with feature map n.
$(x)_m{}^{l-1}$ represents (m) feature of the layer (l – 1) and
$(x)_n{}^{l}$ represent (n) feature of the layer (l)
$(b)_m{}^{l}$ is the bias parameter, f(.) is the non-linearly activation function, M_{l-1} denotes a set of feature maps.
The pooling layer plays a crucial role in subsampling the feature map following the convolutional layer. Its primary objective is to reduce the number of learnable parameters while minimizing the computational workload in the network through the reduction of the feature map's dimensions. Max pooling and average pooling are the two most used styles of pooling layers. The first option chooses the biggest element on the feature map inside the filter's area, resulting in a new feature map with only the essential features, as shown in equation (4).

$$x_i = ma_1 \leq j \ \leq M \times M(x_j) \tag{4}$$

The function of the average pooling layer is to calculate the mean value of the elements within the specific region of the feature map that is encompassed by the filter [24]. This results in the output layer of the pooling layer containing the average features that exist within the patch, as per the equation (5).

$$x_i = \frac{1}{M \times M} \sum_{j=1}^{M \times M} x_j \tag{5}$$

The output x_i after a pooling layer can be obtained by dividing the input into a pooling region of size M and computing a function on the elements x_j within that region. Different pooling techniques, including stochastic pooling , detailed preserving pooling , Spatial pyramid pooling , Multi activation pooling, and Def pooling, exist for this purpose. In a fully connected CNN model, the output of the last pooling layer serves as the input to a fully connected layer, which functions like a traditional neural network with the output representing the number of classes to be classified.

CNNs have undergone several developments since their inception. LeNet-5, a successful CNN for handwritten digit recognition, was created by Yann LeCunn in 1990 and consisted of seven layers. AlexNet was a deep convolutional neural network with five convolutional and three fully connected layers that was developed by Krizhevsky et al. [15]. To facilitate training, the ReLU activation function was employed rather than the sigmoid activation function. K. Simoyan and A. Zisserman's VGG-16 [25] contained 13 convolutional layers and three fully linked layers. In order to investigate the correlation between the depth of convolutional networks and the precision of image classification models, the Visual Geometric Group (VGG) created a series of CNNs, including VGG-11, VGG-13, VGG-16, and VGG-19.

GoogleNet, proposed by Szegedy et al. in their paper [26], introduced the Inception neural network, which comprises of 22 distinct layers for image classification. One of the essential elements of this network is the use of inception layers that utilize varying filter sizes to perform convolution on the input layer.

5.3 Recurrent Neural Network (RNN)

Traditional neural networks treat each input and output as independent, while recurrent neural networks (RNNs) take into account the previous output as input to the current step. RNNs are particularly useful when predicting the next word in a sentence, as they need to remember the previous words to generate the next one. They are a type of neural network that is designed to handle sequential data.

At time t, the RNN model utilizes a recurrent connection hidden unit h_t, which receives input activation from the present data x_t as well as the previous hidden state h_{t-1}. The output y_t is then computed based on the hidden state h_t, which can be expressed mathematically using the following equations (6) and (7), enclosed in square brackets:

$$\boldsymbol{h_t = f\ (w_{hx}x_t + w_{hh}h_{t-1} + b_h)} \tag{6}$$

$$\boldsymbol{y = softmax(w_{yh} + b_y)} \tag{7}$$

(f) is a non-linearly activation function, (w_{hx}) is a matrix weight between the input and hidden layers, (w_{hh}) is a recurrent weights matrix between the hidden layers and itself, (w_{yh}) is a matrix between the hidden and output layer weights, and (b_h)and (b_y) are biases.

5.4 Autoencoder

The Autoencoder is a unique neural network architecture type that ensures the input and output layers contain identical values. This architecture typically comprises three stages: the encoder, the hidden layer, and the decoder, as illustrated in Figure (6).

The primary objective of the encoder's job is to reduce the input data's dimension. referred to as code, which the decoder then decodes to create an approximation of the original input data. This unsupervised approach uses regular feed-forward propagation, where the backpropagation technique computes the gradient of the loss function.

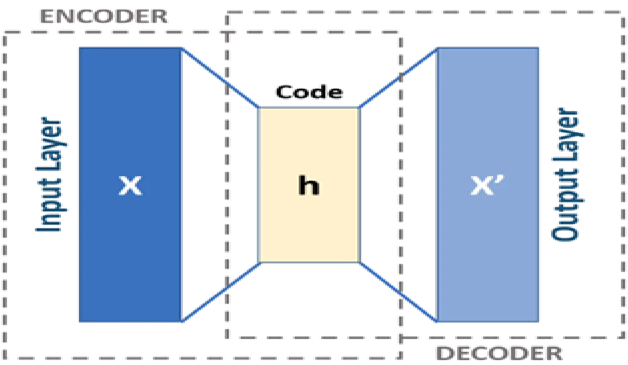


Figure 6.Autoencoder

6. Deep Learning in Medical Imaging

Deep Learning has revolutionized the field of medical imaging by enabling more accurate and efficient image analysis, interpretation, and diagnosis. With the accessibility of massive medical imaging datasets, Deep Learning models have been created to carry out various tasks such as image segmentation, detection, classification, and registration. Convolutional Neural Networks (CNNs) are commonly used in medical imaging applications due to their ability to learn features from images automatically.

6.1 X-ray Image

Chest X-rays are a commonly utilized diagnostic tool for identifying heart and lung ailments. They generate an image encompassing the heart, lungs, chest and spinal bones, blood vessels, and airways. They can detect a range of internal conditions, including heart-related lung problems such as atelectasis, pleural effusion, consolidation, pneumothorax, and hyper cardiac inflation. X-rays Easy to use, affordable and allows higher quality care for mass screening.
As a result of limited availability of COVID-19 dataset, researchers utilized Generative Adversarial Networks (GAN) to generate additional CXR images. Loey et al. [27] proposed a GAN-based approach for deep learning-based COVID-19 detection using CXR images. Meanwhile, Waheed et al. [28] developed CovidGAN, a model that utilizes manufactured CXR pictures are created using the Auxiliary Classifier Generative Adversarial Network (ACGAN) in order to identify COVID-19.

6.2 Computerized tomography (CT)

Medical technique using computerized x-ray imaging is employed to obtain three-dimensional images of a patient by directing a narrow beam of x-rays around the body to generate cross-sectional slices. This procedure aids in the easier identification of tumors or abnormalities in comparison to traditional x-ray imaging of basic structures. CT scans are primarily utilized for detecting pulmonary nodules, especially for identifying them in the early stages. Detecting malignant pulmonary nodules in the early stages is critical for diagnosing lung cancer. Furthermore, the early diagnosis of breast cancer is also vital.

6.3 Mammograph

A mammogram is a breast X-ray imaging that utilizes little radiation. Medical professionals utilize mammograms to detect early signs of breast cancer by visualizing the breast's structure [29]. Breast cancer is one of the primary causes of cancer-related deaths among women worldwide. Therefore, the detection and classification of breast masses at an early stage using mammography remains a hot topic of research. However, the task of using mammography for breast cancer detection is challenging because tumors only represent a small portion of the breast image.
Fortunately, the Deep Learning Approach has made significant advancements in breast cancer detection and classification problems. Analyzing breast lesions using mammography involves three steps: detection, segmentation, and classification.

6.4. Histopathology

The study of tissue under a microscope is known as histopathological samples to identify various diseases, cancers of the kidney, the lung, the breast, and others by studying their characteristic features. Staining techniques are utilized in histopathology to enhance visualization and emphasize cellular structures within the tissue. Hematoxylin and Eosin (H&E) staining is a widely used method that imparts different hues to the nucleus and other structures, enabling the diagnosis and grading of various pathologies, including cancer, over the past century.

Digital pathology is a cutting-edge imaging modality that has gained prominence in recent times. The effectiveness of deep learning methods has been proving to be highly efficient in analyzing histopathological images. These techniques are useful for tasks such as detecting nuclei, classifying images, segmenting cells, and tissues, among others.

6.5. Evaluation metrics

The most popular metrics for measuring how well NNs perform on medical pictures are covered in this section. Multiple performance measures are employed to evaluate and analyze the NN's performance in both detection and classification tasks. One popular method for evaluating detection performance is Intersection over Union (IoU), which measures the degree of overlap between the predicted bounding box and the ground truth. Equation (8) demonstrates how IoU can be calculated.

$$IoU = \frac{A \cap B}{A \cup B} \quad \text{The ground truth box is represented by B, while A is the anticipated bounding box.} \qquad (8)$$

Additionally, both anomaly detection and classification are evaluated using sensitivity, specificity, and accuracy. It's important to also consider the confusion matrix, which displays the count of True Positive, False Positive, True Negative, and False Negative are the four types of results. Table 2 demonstrates the layout of the confusion matrix, with the following components:

- True Positive (TP) signifies the instances where the system accurately identified or categorized the cases as positive.
- False Positive (FP) refers to the count of negative cases that were inaccurately identified or categorized as positive.
- True Negative (TN) refers to negative cases that are accurately identified as negative in detection or classification.
- False Negative (FN) represents cases that were not accurately identified or classified.

Table 2 fundamental structure of the confusion matrix

Table 2 fundamental structure of the confusion matrix

		Predicted	
		+	-
Actual	+	**TP**	**FP**
	-	**FN**	**TN**

Sensitivity, Recall, or True Positive Rate (TPR) measures how likely it is for a true positive to be correctly classified as a positive case. By using equation (9), this can be calculated. Specificity, also known as True Negative Rate (TNR), is the likelihood that, in the absence of any abnormalities in the image, a true negative will be correctly identified as such. Equation (10) provides the formula for calculating specificity.

$$\text{Sensitivity} = TP / (TP + FN) \quad (9)$$
$$\text{Specificity} = TN / (TN + FP) \quad (10)$$

Accuracy (Acc), which gauges the proportion of accurate predictions to all predictions, can be used to evaluate a system's performance across all classes, as shown in Equation (11). Receiver Operating Characteristics (ROC) is a crucial performance metric for such systems in addition to these evaluation techniques. The True Positive Rate (TPR) and False Positive Rate (FPR) trade-off at various classification thresholds is represented by the ROC curve. TPR is on the y-axis of the ROC curve and FPR is on the x-axis. How effectively the system distinguishes between positive and negative classes is shown by the area under the ROC curve (AUC).

$$Acc = (TP + TN) / (TP + TN + FP + FN) \quad (11)$$

The F1-Score is another performance evaluation measure employed for assessing a model's binary classification performance, which involves classifying cases as either positive or negative. The F1-Score is computed based on the model's precision and recall.

Finally, the effectiveness of object detection models can also be evaluated using the mAP (mean Average Precision) metric. Calculating the Average Precision (AP) for each class and then averaging the AP values across all classes yields the mAP.

7. Conclusion

This article explores recent advances in deep learning algorithms and strategies as they relate to medical image analysis, a crucial source of information for clinical decision-making. The paper has two main goals: to give a general overview of the field using DLA and to introduce an overview of the application of deep learning to medical picture processing. Several supervised and unsupervised DL algorithms are covered in the article, including auto-encoders, recurrent, and CNN, after providing a brief history of neural networks.

DLA has the potential to significantly reduce medical errors and improve efficiency in medical image analysis, leading to more accurate auto-diagnostic outcomes utilizing medical imagery. As such, physicians and scientists should prioritize finding excellent practices for leverage DLA to offer superior care for patients. One promising area for future research is the development of deep neural network architectures for medical image analysis. The design of network structures directly impacts the quality and accuracy of medical image analysis, making this a critical area for further investigation.

References

1. Litjens, Geert et al. “A survey on deep learning in medical image analysis.” Medical image analysis vol. 42 (2017).
2. S.Pang and X.Yang, “Deep Convolutional Extreme Learning Machine, & Its Application in Handwritten digit classification, computational intelligence & neuroscience”, (2016).
3. Bauer, Stefan et al. “A survey of MRI-based medical image analysis for brain tumor studies.” Physics in medicine and biology vol. 58,13 (2013).
4. Anamika D. and Gyanendra V. “Convolutional neural network: a review of models, methodologies and applications to object detection.” Progress in Artificial Intelligence, (2019).

5. Geert Litjens, Thijs Kooi, B. Babak Ehteshami, Arnaud AA. Setio, Francesco C., Mohsen Gh., Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez "A survey on deep learning in medical image analysis.", Med Image Anal, December 2017.
6. Richard K., Wulfram G., and J. Leo Hemmen, "Hebbian learning and spiking neurons", Published 1 April 1999
7. Widrow, Marcian Hoff and Bernard. Associative Storage and Retrieval of Digital Information in Networks of Adaptive-Neurons. (1962).
8. Werbos, Paul, Beyond Regression: New Tools for Prediction and Analysis in Behavioral Science. Thesis (Ph. D.),Appl. Math. Harvard University, (1974).
9. Yann LeCun, Backpropagation Applied to Handwritten Zip Code Recognition, in *Neural Computation*, Dec. 1989,
10. Eberhart and R. W. Dobbins, "Early neural network development history: the age of Camelot," IEEE Engineering in Medicine & Biology Magazine, 1990 Sept.
11. Yu Li, Chao Huang, Lizhong Ding, Zhongxiao Li,Yijie Pan, Xin Gao "Deep learning in bioinformatics: Introduction, application, and perspective in the big data era. Methods." 2019 Aug.
12. Goodfellow, Ian & Pouget-Abadie, Jean & Mirza, Mehdi & Xu, Bing & Warde-Farley, David & Ozair, Sherjil & Courville, Aaron & Bengio, Y. Generative Adversarial Networks. Advances in Neural Information Processing Systems, (2014).
13. Little, William A. "The existence of persistent states in the brain." *Bellman Prize in Mathematical Biosciences"* (1974).
14. Little, William A. and Gordon L. Shaw. "Analytic study of the memory storage capacity of a neural network." *Bellman Prize in Mathematical Biosciences* (1978).
15. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems", (2012).
16. Maas, A.L, Hannun, A.Y. and Ng, A.Y, "Rectifier Nonlinearities Improve Neural Network Acoustic Models." Proceedings of the 30th International Conference on Machine Learning, (2013).
17. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification." IEEE International Conference on Computer Vision, (2015).
18. Bing Xu, Naiyan Wang, Tianqi Chen, Mu Li, "Empirical Evaluation of Rectified Activations in Convolutional Network", (2015).
19. Charles Dugas, Yoshua Bengio, François Bélisle, Claude Nadeau, René Garcia, "Incorporating Second-Order Functional Knowledge for Better Option Pricing" (2000).
20. Djork-Arné Clevert & Thomas Unterthiner, & Hochreiter, Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). Under Review of ICLR-2016 . Sep. (2015).
21. Günter Klambauer, Thomas Unterthiner, Andreas Mayr, Sepp Hochreiter, Self-Normalizing Neural Networks, (2017).
22. Phi Vu Tran, A Fully Convolutional Neural Network for Cardiac Segmentation in Short-Axis MRI.
23. Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, Dimitris Metaxas (2016). "Multimodal Deep Learning for Cervical Dysplasia Diagnosis."
24. Scherer, Dominik & Müller, Andreas & Behnke, Sven – 2010, "Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition."
25. Karen Simonyan, Andrew Zisserman – 2015, "Very Deep Convolutional Networks for Large-Scale Image Recognition." The 3rd International Conference on Learning Representations.
26. Christian Szegedy & Wei Liu & Yangqing Jia & Pierre Sermanet & Scott Reed & Dragomir Anguelov & Dumitru Erhan & Vincent Vanhoucke & Andrew Rabinovich – 2014, "Going Deeper with Convolutions."

27. Nour Eldeen M Khalifa, Florentin Smarandache, Gunasekaran Manogaran, Mohamed Loey - 2020. "Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. Symmetry."
28. Abdul Waheed, Muskan Goyal, Deepak Gupta, Ashish Khanna, Fadi Al-Turjman, Placido Rogerio Pinheiro-2002. "CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection."
29. Wang, Tianfu & Gardezi, Jamal & Elazab, Ahmed & Lei, Baiying. -2019. "Breast Cancer Detection & Diagnosis Using Mammographic Data, Systematic Review." Journal of Medical Internet Research.