



Medical image analysis based on deep learning approach

Muralikrishna Puttagunta¹ · S. Ravi¹ 

Received: 25 August 2020 / Revised: 28 November 2020 / Accepted: 10 February 2021 /

Published online: 6 April

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Medical imaging plays a significant role in different clinical applications such as medical procedures used for early detection, monitoring, diagnosis, and treatment evaluation of various medical conditions. Basics of the principles and implementations of artificial neural networks and deep learning are essential for understanding medical image analysis in computer vision. Deep Learning Approach (DLA) in medical image analysis emerges as a fast-growing research field. DLA has been widely used in medical imaging to detect the presence or absence of the disease. This paper presents the development of artificial neural networks, comprehensive analysis of DLA, which delivers promising medical imaging applications. Most of the DLA implementations concentrate on the X-ray images, computerized tomography, mammography images, and digital histopathology images. It provides a systematic review of the articles for classification, detection, and segmentation of medical images based on DLA. This review guides the researchers to think of appropriate changes in medical image analysis based on DLA.

Keywords Deep learning · Convolutional neural networks · Medical images · Segmentation · Classification · Detection

1 Introduction

In the health care system, there has been a dramatic increase in demand for medical image services, e.g. Radiography, endoscopy, Computed Tomography (CT), Mammography

✉ S. Ravi
sraivite@gmail.com

Muralikrishna Puttagunta
murali93940@gmail.com

¹ Department of Computer Science, School of Engineering and Technology, Pondicherry University, Pondicherry, India

Images (MG), Ultrasound images, Magnetic Resonance Imaging (MRI), Magnetic Resonance Angiography (MRA), Nuclear medicine imaging, Positron Emission Tomography (PET) and pathological tests. Besides, medical images can often be challenging to analyze and time-consuming process due to the shortage of radiologists.

Artificial Intelligence (AI) can address these problems. Machine Learning (ML) is an application of AI that can be able to function without being specifically programmed, that learn from data and make predictions or decisions based on past data. ML uses three learning approaches, namely, supervised learning, unsupervised learning, and semi-supervised learning. The ML techniques include the extraction of features and the selection of suitable features for a specific problem requires a domain expert. Deep learning (DL) techniques solve the problem of feature selection. DL is one part of ML, and DL can automatically extract essential features from raw input data [88]. The concept of DL algorithms was introduced from cognitive and information theories. In general, DL has two properties: (1) multiple processing layers that can learn distinct features of data through multiple levels of abstraction, and (2) unsupervised or supervised learning of feature presentations on each layer. A large number of recent review papers have highlighted the capabilities of advanced DLA in the medical field MRI [8], Radiology [96], Cardiology [11], and Neurology [155].

Different forms of DLA were borrowed from the field of computer vision and applied to specific medical image analysis. Recurrent Neural Networks (RNNs) and convolutional neural networks are examples of supervised DL algorithms. In medical image analysis, unsupervised learning algorithms have also been studied; These include Deep Belief Networks (DBNs), Restricted Boltzmann Machines (RBMs), Autoencoders, and Generative Adversarial Networks (GANs) [84]. DLA is generally applicable for detecting an abnormality and classify a specific type of disease. When DLA is applied to medical images, Convolutional Neural Networks (CNN) are ideally suited for classification, segmentation, object detection, registration, and other tasks [29, 44]. CNN is an artificial visual neural network structure used for medical image pattern recognition based on convolution operation. Deep learning (DL) applications in medical images are visualized in Fig. 1.

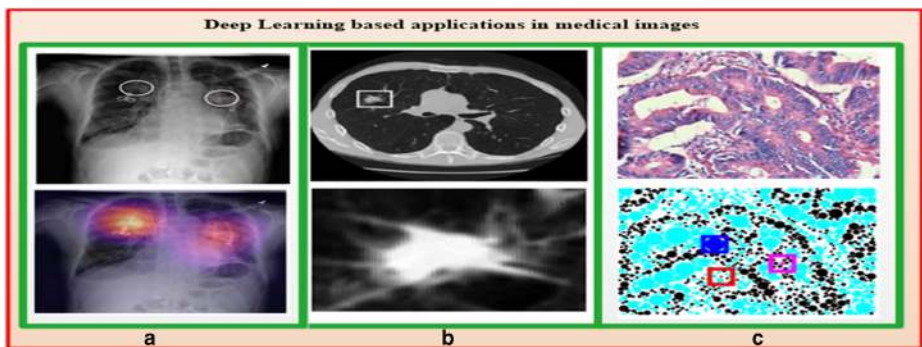


Fig. 1 a X-ray image with pulmonary masses [121] b CT image with lung nodule [82] c Digitized histopathological tissue image [132]

2 Neural networks

2.1 History of neural networks

The study of artificial neural networks and deep learning derives from the ability to create a computer system that simulates the human brain [33]. A neurophysiologist, Warren McCulloch, and a mathematician Walter Pitts [97] developed a primitive neural network based on what has been known as a biological structure in the early 1940s. In 1949, a book titled “Organization of Behavior” [100] was the first to describe the process of upgrading synaptic weights which is now referred to as the Hebbian Learning Rule. In 1958, Frank Rosenblatt’s [127] landmark paper defined the structure of the neural network called the perceptron for the binary classification task.

In 1962, Windrow [172] introduced a device called the Adaptive Linear Neuron (ADALINE) by implementing their designs in hardware. The limitations of perceptions were emphasized by Minski and Papert (1969) [98]. The concept of the backward propagation of errors for purposes of training is discussed in Werbose1974 [171]. In 1979, Fukushima [38] designed artificial neural networks called Neocognitron, with multiple pooling and convolution layers. One of the most important breakthroughs in deep learning occurred in 2006, when Hinton et al. [9] implemented the Deep Belief Network, with several layers of Restricted Boltzmann Machines, greedily teaching one layer at a time in an unsupervised fashion. In 1989, Yann LeCun [71] combined CNN with backpropagation to effectively perform the automated recognition of handwritten digits. Figure 2 shows important advancements in the history of neural networks that led to a deep learning era.

2.2 Artificial neural networks

Artificial Neural Networks (ANN) form the basis for most of the DLA. ANN is a computational model structure that has some performance characteristics similar to

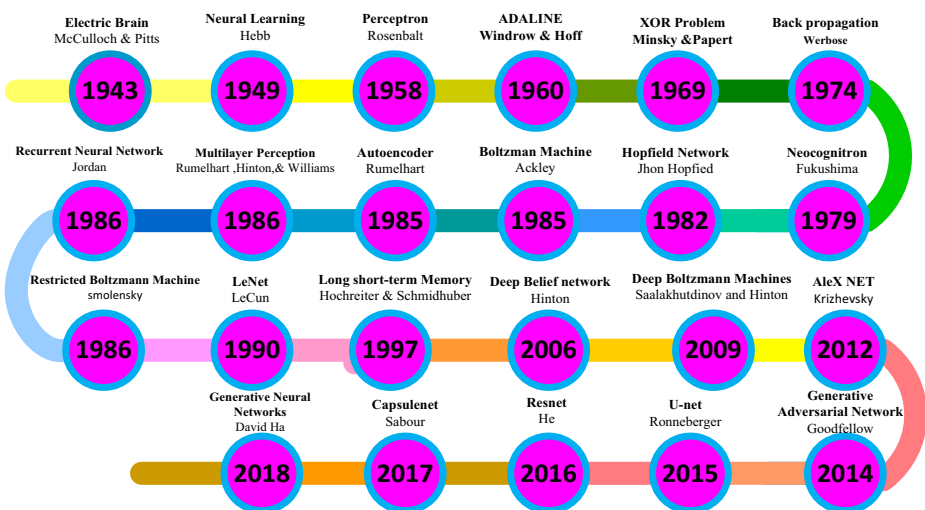


Fig. 2 Demonstrations of significant developments in the history of neural networks [33, 134]

biological neural networks. ANN comprises simple processing units called neurons or nodes that are interconnected by weighted links. A biological neuron can be described mathematically in Eq. (1). Figure 3 shows the simplest artificial neural model known as the perceptron.

$$y = f(wx^T + b) \tag{1}$$

Let $x \in \mathbb{R}^d$ be the vector of inputs (features), $w \in \mathbb{R}^d$ be the weight vector, $b \in \mathbb{R}$ be the bias, f be the non-linear activation (sigmoid) function and y be the scalar, the output of a node. For the construction of neural networks required parameters are i) The pattern of connections between the neurons (architecture) ii) The method of determining the weights on the connection (Training or learning) and iii) Activation function. Depending on the pattern of connection between nodes in the neural networks, they are mainly classified into two types, viz., a Feed-Forward Neural Network (FFNN) and Recurrent Neural Networks (RNN) [42]. For FFNN the requirement is that network has to be a directed acyclic graph. In RNN the connection between nodes forms a directed cycle. An FFNN consists of one or more layers in which each layer comprises one or more nodes. The network has one input layer and the last layer is called the output layer. The layers between the input and the output layer are considered as a hidden layer. FFNN is used for supervised [140] learning tasks like classification and regression.

2.3 Training a neural network with Backpropagation (BP)

In the neural networks, the learning process is modeled as an iterative process of optimization of the weights to minimize a loss function. Based on network performance, the weights are modified on a set of examples belonging to the training set. The necessary steps of the training procedure contain forward and backward phases. For Neural Network training, any of the activation functions in forwarding propagation is selected and BP training is used for changing weights. The BP algorithm helps multilayer

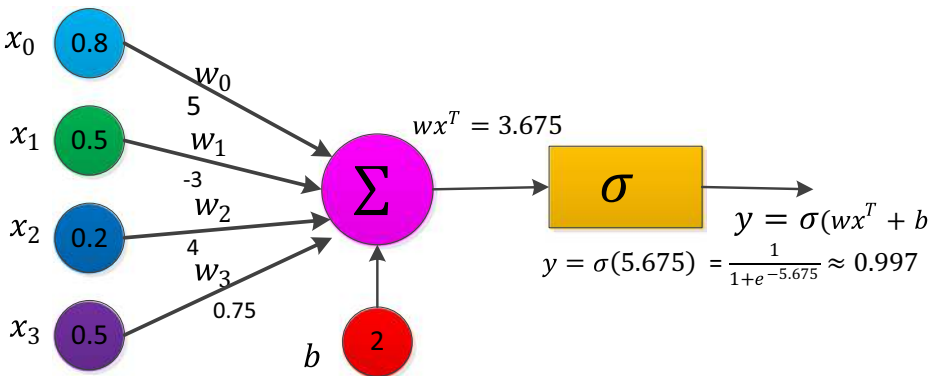


Fig. 3 Perceptron [77]

FFNN to learn input-output mappings from training samples [16]. Forward propagation and backpropagation are explained with the one hidden layer deep neural networks in the following algorithm.

The backpropagation algorithm is as follows for one hidden layer neural network

1. Initialize all weights to small random values.
2. While the stopping condition is false, do steps 3 through 10.
3. For each training pair $((x_1, y_1) \dots (x_n, y_n))$ do steps 4 through 9.

Feed-forward propagation:

4. Each input unit $(X_i, i = 1, 2, \dots, n)$ receives the input signal x_i and send this signal to all hidden units in the above layer.
5. Each hidden unit $(Z_j, j = 1, \dots, p)$ compute output using the below equation, and it transmits to the output unit (i.e.) $z_{j_in} = b_j + \sum_{i=1}^n w_{ij} x_i$ applies to an activation function $Z_j = f(Z_{j_in})$.
6. Compute the out signal for each output unit $(Y_k, k = 1, \dots, m)$.
 $y_{k_in} = b_k + \sum_{j=1}^p z_j w_{jk}$ and calculate activation $y_k = f(y_{k_in})$

Backpropagation

7. For input training pattern (x_1, x_2, \dots, x_n) corresponding output pattern (y_1, y_2, \dots, y_m) , let (t_1, t_2, \dots, t_m) be target pattern. For each output, the neuron computes network error δ_k
 At output-layer neurons $\delta_k = (t_k - y_k) f'(y_{k_in})$
8. For each hidden neuron, calculate its error information term δ_j while doing so, use δ_k of the output neurons as obtained in the previous step
 At Hidden layer neurons $\delta_j = f'(z_{j_in}) \sum_k \delta_k w_{jk}$
9. Update weights and biases using the following formulas where η is learning rate

Each output layer $(Y_k, k = 1, 2, \dots, m)$ updates its weights $(J = 0, 1, \dots, P)$ 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 $(Z_j, J = 1, 2, \dots, p)$ updates its weights $(i = 0, 1, \dots, n)$ biases:

$$w_{ij}(new) = w_{ij}(old) + \eta \delta_j x_i; b_j(old) = b_j(old) + \eta \delta_j$$

10. Test stopping condition

2.4 Activation function

The activation function is the mechanism by which artificial neurons process and transfers information [42]. There are various types of activation functions which can be used in neural networks based on the characteristic of the application. The activation functions are non-linear and continuously differentiable. Differentiability property is important mainly when training a neural network using the gradient descent method. Some widely used activation functions are listed in Table 1.

Table 1 Activation functions

Function name	Function equation	Function derivate
Sigmoid [86]	$f(x) = \frac{1}{1+e^{-x}}$	$f'(x) = f(x)(1-f(x))$
Hyperbolic tangent [87]	$f(x) = \tanh(x) = \frac{2}{1+e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$
Soft sign activation	$f(x) = \frac{x}{1+ x }$	$f'(x) = \frac{1}{(1+ x)^2}$
Rectified Linear Unit [68, 104] (ReLU)	$f(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$
Leaky Rectified Linear Unit [94] (leaky ReLU)	$f(x) = \begin{cases} \alpha x & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & x < 0 \\ 1 & x \geq 0 \end{cases}$
Parameterized Rectified Linear Unit(PReLU) [47]	PReLU is the same as leaky ReLU. The difference is α can be learned from training data via backpropagation	
Randomized Leaky Rectified Linear Unit [180]	$f(x) = \begin{cases} \alpha x & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} \alpha & x < 0 \\ 1 & x \geq 0 \end{cases}$
Soft plus [32]	$f(x) = \ln(1 + e^x)$	$f'(x) = \frac{1}{1+e^{-x}}$
Exponential Linear Unit (ELU) [24, 137]	$f(x) = \begin{cases} \alpha(e^x - 1) & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \alpha x & x < 0 \\ 1 & x \geq 0 \end{cases}$
Scaled exponential Linear Unit (SELU) [67]	$f(x) = \lambda \begin{cases} \alpha(e^x - 1) & x < 0 \\ x & x \geq 0 \end{cases}$	$f'(x) = \begin{cases} f(x) + \lambda \alpha x & x < 0 \\ \lambda & x \geq 0 \end{cases}$

3 Deep learning

Deep learning is a subset of the machine learning field which deals with the development of deep neural networks inspired by biological neural networks in the human brain.

3.1 Autoencoder

Autoencoder (AE) [128] is one of the deep learning models which exemplifies the principle of unsupervised representation learning as depicted in Fig. 4a. AE is useful when the input data have more number of unlabelled data compared to labeled data. AE encodes the input x into a lower-dimensional space z . The encoded representation is again decoded to an approximated representation x' of the input x through one hidden layer z .

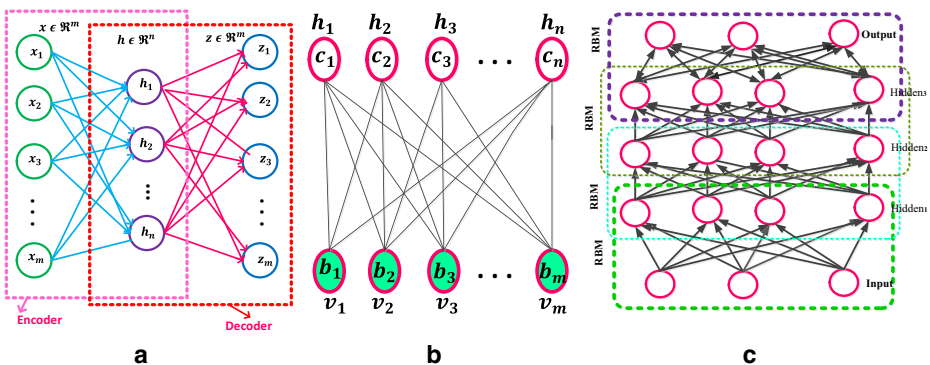


Fig. 4 a Autoencoder [187] b Restricted Boltzmann Machine with n hidden and m visible units [88] c Deep Belief Networks [88]

Basic AE consists of three main steps:

Encode: Convert input vector $x \in \mathfrak{R}^m$ into $h \in \mathfrak{R}^n$, the hidden layer by $h = f(wx + b)$ where $w \in \mathfrak{R}^{m \times n}$ and $b \in \mathfrak{R}^n$. m and n are dimensions of the input vector and converted hidden state. The dimension of the hidden layer h is to be smaller than x . f is an activate function.

Decode: Based on the above h , reconstruct input vector z by equation $z = f'(w'h + b')$ where $w' \in \mathfrak{R}^{n \times m}$ and $b' \in \mathfrak{R}^m$. The f' is the same as the above activation function.

Calculate square error: $L_{recons}(x, z) = \|x - z\|^2$, which is the reconstruction error cost function. Reconstruct error minimization is achieved by optimizing the cost function (2)

$$J(\theta) = \sum(x, z) \quad \theta = \{w, w', b, b'\} \tag{2}$$

Another unsupervised algorithm representation is known as Stacked Autoencoder (SAE). The SAE comprises stacks of autoencoder layers mounted on top of each other where the output of each layer was wired to the inputs of the next layer. A Denoising Autoencoder (DAE) was introduced by Vincent et al. [159]. The DAE is trained to reconstruct the input from random noise added input data. Variational autoencoder (VAE) [66] is modifying the encoder where the latent vector space is used to represent the images that follow a Gaussian distribution unit. There are two losses in this model; one is a mean squared error and the Kull back Leibler divergence loss that determines how close the latent variable matches the Gaussian distribution unit. Sparse autoencoder [106] and variational autoencoders have applications in unsupervised, semi-supervised learning, and segmentation.

3.2 Restricted Boltzmann machine

A Restricted Boltzmann machine [RBM] is a Markov Random Field (MRF) associated with the two-layer undirected probabilistic generative model, as shown in Fig. 4b. RBM contains visible units (input) v and hidden (output) units h . A significant feature of this model is that there is no direct contact between the two visible units or either of the two hidden units. In binary RBMs, the random variables (v, h) takes $(v, h) \in \{0, 1\}^{m+n}$. Like the general Boltzmann machine [50], the RBM is an energy-based model. The energy of the state $\{v, h\}$ is defined as (3)

$$E(v, h) = -\sum_{i=1}^n \sum_{j=1}^m w_{ij} h_i v_j - \sum_{j=1}^m b_j v_j - \sum_{i=1}^n c_i h_i \tag{3}$$

where v_j, h_i are the binary states of visible unit $j \in \{1, 2, \dots, m\}$ and hidden unit $i \in \{1, 2, \dots, n\}$, b_j, c_i are their biases of visible and hidden units, w_{ij} is the symmetric interaction term between the units v_j and h_i them. A joint probability of (v, h) is given by the Gibbs distribution in Eq. (4)

$$P(v, h) = \frac{1}{Z} e^{-E(v,h)} \tag{4}$$

Z is a “partition function” that can be given by summing over all possible pairs of visual v and hidden h (5).

$$Z = \sum_{v,h} e^{-E(v,h)} \quad (5)$$

A significant feature of the RBM model is that there is no direct contact between the two visible units or either of the two hidden units. In term of probability, conditional distributions $p(h|v)$ and $p(v|h)$ is computed as (6) $p(h|v) = \prod_{i=1}^n p(h_i|v)$

$$p(h|v) = \prod_{i=1}^n p(h_i|v), p(v|h) = \prod_{j=1}^m p(v_j|h) \quad (6)$$

For binary RBM condition distribution of visible and hidden are given by (7) and (8)

$$p(h_j = 1|v) = \sigma(b_j + \sum_{i=1}^m w_{ij}v_i) \quad (7)$$

$$p(v_i = 1|h) = \sigma(a_j + \sum_{j=1}^m w_{ij}h_j) \quad (8)$$

where $\sigma(\cdot)$ is a sigmoid function

RBM's parameters (w_{ij}, b_j, c_i) are efficiently calculated using the contrastive divergence learning method [150]. A batch version of k-step contrastive divergence learning (CD-k) can be discussed in the algorithm below [36]

Algorithm k-step contrastive divergence for RBM

Input: RBM ($V_1, V_2, \dots, V_m, H_1, H_2, \dots, H_n$), training batch s , learning rate ϵ

for $i = 1, 2, 3, \dots, n, j = 1, 2, \dots, m$.

 Initialize w with random values distributed uniformly in $[0, 1], \Delta w_{ij} = \Delta c_i = \Delta b_j = 0$,

 for $i = 1, 2, 3, \dots, n, j = 1, 2, \dots, m$.

 for all $v \in S \forall$ do, $v^{(0)} \leftarrow v$

 for $t = 0, \dots, K - 1$ do

 for $i = 1, 2, \dots, n$ do sample $h_i^{(t)} \sim p(h_i|v^{(t)})$;

 for $j = 1, 2, \dots, m$ do sample $v_j^{(t+1)} \sim p(v_j|h^{(t)})$;

 END for

 for $i = 1, 2, \dots, n, j = 1, 2, \dots, m$ do

$$\Delta w_{ij} \leftarrow \Delta w_{ij} + \epsilon \times [p(H_i = 1|v^{(0)}) \times v_j^{(0)} - p(H_i = 1|v^{(k)}) \times v_j^{(k)}]$$

$$\Delta b_j = \Delta b_j + \epsilon \times (v_j^{(0)} - v_j^{(k)})$$

$$\Delta c_i = \Delta c_i + \epsilon \times [p(H_i = 1|v^{(0)}) - p(H_i = 1|v^{(k)})]$$

$$w = w + \epsilon \times \Delta w_{ij}$$

 END for

 END for

Output: The RBM weight matrix w , gradient approximation $\Delta w_{ij}, \Delta c_i$ and Δb_j

3.3 Deep belief networks

The Deep Belief Networks (DBN) proposed by Hinton et al. [51] is a non-convolution model that can extract features and learn a deep hierarchical representation of training data. DBNs are generative models constructed by stacking multiple RBMs. DBN is a hybrid model, the first two layers are like RBM, and the rest of the layers form a directed generative model. A DBN has one visible layer v and a series of hidden layers $h^{(1)}, h^{(2)}, \dots, h^{(l)}$ as shown in Fig. 4c. The DBN model joint distribution between the observed units v and the l hidden layers $h^k (k = 1, \dots, l)$ as (9)

$$P(v, h^1, \dots, h^l) = \left(\prod_{k=0}^{l-2} P(h^{(k)} | h^{(k+1)}) \right) P(h^{(l-1)}, h^{(l)}) \tag{9}$$

where $v = h^{(0)}$, $P(h^k | h^{k+1})$ is a conditional distribution (10) for the layer k given the units of $k + 1$

$$P(h_i^{(k)} = 1 | h^{(k+1)}) = \sigma(b_i^{(k)} + W_{:,i}^{(k+1)} h^{(k+1)}) \forall i, \forall k \in 0, 1, \dots, l-2, \tag{10}$$

A DBN has l weight matrices: $W^{(1)}, \dots, W^{(l)}$ and $l + 1$ bias vectors: $b^{(0)}, \dots, b^{(l)}$ $P(h^{(l)}, h^{(l-1)})$ is the joint distribution of top-level RBM (11).

$$P(h^{(l)}, h^{(l-1)}) \propto e^{(b^{(l)} h^{(l)} + b^{(l-1)} h^{(l-1)} + h^{(l-1)} W^{(l)} h^{(l)})} \tag{11}$$

The probability distribution of DBN is given by Eq. (12)

$$P(v_i = 1 | h^{(1)}) = \sigma(b_i^{(0)} + W_{:,i}^{(1)} h^{(1)}) \forall i \tag{12}$$

3.4 Convolutional neural networks (CNN)

In neural networks, CNN is a unique family of deep learning models. CNN is a major artificial visual network for the identification of medical image patterns. The family of CNN primarily emerges from the information of the animal visual cortex [55, 116]. The major problem within a fully connected feed-forward neural network is that even for shallow architectures, the number of neurons may be very high, which makes them impractical to apply to image applications. The CNN is a method for reducing the number of parameters, allows a network to be deeper with fewer parameters.

CNN's are designed based on three architectural ideas that are shared weights, local receptive fields, and spatial sub-sampling [70]. The essential element of CNN is the handling of unstructured data through the convolution operation. Convolution of the input signal $x(t)$ with filter signal $h(t)$ creates an output signal $y(t)$ that may reveal more information than the input signal itself. 1D convolution of a discrete signals $x(t)$ and $h(t)$ is (13)

$$y(t) = x(t) * h(t) = \sum_{\tau=-\infty}^{\infty} x(\tau) h(t-\tau) \tag{13}$$

A digital image $x(n_1, n_2)$ is a 2-D discrete signal. The convolution of images $x(n_1, n_2)$ and $h(n_1, n_2)$ is (14)

$$y(n_1, n_2) = \sum_{k_1=0}^{M-1} \sum_{k_2=0}^{N-1} x(k_1, K_2) h(n_1 - k_1, n_2 - k_2) \quad (14)$$

where $0 \leq n_1 \leq M-1$, $0 \leq n_2 \leq N-1$.

The function of the convolution layer is to detect local features x^l from input feature maps x^{l-1} using kernels k^l by convolution operation ($*$) i.e. $x^{l-1} * k^l$. This convolution operation is repeated for every convolutional layer subject to non-linear transform (15)

$$x_n^{(l)} = f\left(\sum_m^{M^{l-1}} x_m^{(l-1)} * k_{mn}^{(l)} + b_m^{(l)}\right) \quad (15)$$

where $k_{mn}^{(l)}$ represents weights between feature map m at layer $l-1$ and feature map n at l . $x_m^{(l-1)}$ represents the m feature map of the layer $l-1$ and x_n^l is n feature map of the layer l . $b_m^{(l)}$ is the bias parameter. $f(\cdot)$ is the non-linear activation function. M^{l-1} denotes a set of feature maps. CNN significantly reduces the number of parameters compared with a fully connected neural network because of local connectivity and weight sharing. The depth, zero-padding, and stride are three hyperparameters for controlling the volume of the convolution layer output.

A pooling layer comes after the convolutional layer to subsample the feature maps. The goal of the pooling layers is to achieve spatial invariance by minimizing the spatial dimension of the feature maps for the next convolution layer. Max pooling and average pooling are commonly used two different pooling operations to achieve downsampling. Let the size of the pooling region M and each element in the pooling region is given as $x_j = (x_1, x_2, \dots, x_{M \times M})$, the output after pooling is given as x_i . Max pooling and average pooling are described in the following Eqs. (16) and (17).

$$x_i = \max_{1 \leq j \leq M \times M} (x_j) \quad (16)$$

$$x_i = \frac{1}{M \times M} \sum_{j=1}^{M \times M} x_j \quad (17)$$

The max-pooling method chooses the most superior invariant feature in a pooling region. The average pooling method selects the average of all the features in the pooling area. Thus, the max-pooling method holds texture information that can lead to faster convergence, average pooling method is called Keep background information [133]. Spatial pyramid pooling [48], stochastic pooling [175], Def-pooling [109], Multi activation pooling [189], and detailed preserving pooling [130] are different pooling techniques in the literature. A fully connected layer is used at the end of the CNN model. Fully connected layers perform like a traditional neural network [174]. The input to this layer is a vector of numbers (output of the pooling layer) and outputs an N-dimensional vector (N number of classes). After the pooling layers, the feature of previous layer maps is flattened and connected to fully connected layers.

The first successful seven-layered LeNet-5 CNN was developed by Yann LeCun in 1990 for handwritten digit recognition successfully. Krizhevsky et al. [68] proposed AlexNet is a deep convolutional neural network composed of 5 convolutional and 3 fully-connected layers. In AlexNet changed the sigmoid activation function to a ReLU activation function to make model training easier.

K. Simonyan and A. Zisserman invented the VGG-16 [143] which has 13 convolutional and 3 fully connected layers. The Visual Geometric Group (VGG) research group released a series of CNN starting from VGG-11, VGG-13, VGG-16, and VGG-19. The main intention of the VGG group to understand how the depth of convolutional networks affects the accuracy of the models of image classification and recognition. Compared to the maximum VGG19, which has 16 convolutional layers and 3 fully connected layers, the minimum VGG11 has 8 convolutional layers and 3 fully connected layers. The last three fully connected layers are the same as the various variations of VGG.

Szegedy et al. [151] proposed an image classification network consisting of 22 different layers, which is GoogleNet. The main idea behind GoogleNet is the introduction of inception layers. Each inception layer convolves the input layers partially using different filter sizes. Kaiming He et al. [49] proposed the ResNet architecture, which has 33 convolutional layers and one fully-connected layer. Many models introduced the principle of using multiple hidden layers and extremely deep neural networks, but then it was realized that such models suffered from the issue of vanishing or exploding gradients problem. For eliminating vanishing gradients' problem skip layers (shortcut connections) are introduced. DenseNet developed by Gao et al. [54] consists of several dense blocks and transition blocks, which are placed between two adjacent dense blocks. The dense block consists of three layers of batch normalization, followed by a ReLU and a 3×3 convolution operation. The transition blocks are made of Batch Normalization, 1×1 convolution, and average Pooling.

Compared to state-of-the-art handcrafted feature detectors, CNNs is an efficient technique for detecting features of an object and achieving good classification performance. There are drawbacks to CNNs, which are that unique relationships, size, perspective, and orientation of features are not taken into account. To overcome the loss of information in CNNs by pooling operation Capsule Networks (CapsNet) are used to obtain spatial information and most significant features [129]. The special type of neurons, called capsules, can detect efficiently distinct information. The capsule network consists of four main components that are matrix multiplication, Scalar weighting of the input, dynamic routing algorithm, and squashing function.

3.5 Recurrent neural networks (RNN)

RNN is a class of neural networks used for processing sequential information (deal with sequential data). The structure of the RNN shown in Fig. 5a is like an FFNN and the difference is that recurrent connections are introduced among hidden nodes. A generic RNN model at time t , the recurrent connection hidden unit h_t receives input activation from the present data x_t and the previous hidden state h_{t-1} . The output y_t is calculated given the hidden state h_t . It can be represented using the mathematical Eqs. (18) and (19) as

$$h_t = f(w_{hx}x_t + w_{hh}h_{t-1} + b_h) \quad (18)$$

$$y = \text{softmax}(w_{yh} + b_y) \quad (19)$$

Here f is a non-linear activation function, w_{hx} is the weight matrix between the input and hidden layers, w_{hh} is the matrix of recurrent weights between the hidden layers and itself

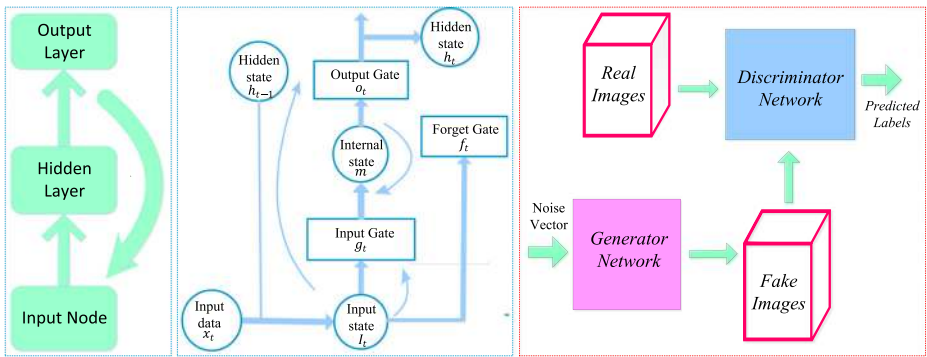


Fig. 5 **a** Recurrent Neural Networks [163] **b** Long Short-Term Memory [163] **c** Generative Adversarial Networks [64]

w_{yh} is the weight matrix between the hidden and output layer, and b_h and b_y are biases that allow each node to learn and offset. While the RNN is a simple and efficient model, in reality, it is, unfortunately, difficult to train properly. Real-Time Recurrent Learning (RTRL) algorithm [173] and Back Propagation Through Time (BPTT) [170] methods are used to train RNN. Training with these methods frequently fails because of vanishing (multiplication of many small values) or explode (multiplication of many large values) gradient problem [10, 112]. Hochreiter and Schmidhuber (1997) designed a new RNN model named Long Short Term Memory (LSTM) that overcome error backflow problems with the aid of a specially designed memory cell [52]. Figure 5b shows an LSTM cell which is typically configured by three gates: input gate g_t , forget gate f_t and output gate o_t , these gates add or remove information from the cell.

An LSTM can be represented with the following Eqs. (20) to (25)

$$\text{Input state } i_t = \sigma(w_{ix}x_t + w_{ih}h_{t-1} + b_i) \tag{20}$$

$$\text{Input gate } g_t = \phi(w_{gx}x_t + w_{gh}h_{t-1} + b_g) \tag{21}$$

$$\text{Forget gate } f_t = \sigma(w_{fx}x_t + w_{fh}h_{t-1} + b_f) \tag{22}$$

$$\text{Output gate } o_t = \sigma(w_{ox}x_t + w_{oh}h_{t-1} + b_o) \tag{23}$$

$$\text{Internal state } m_t = g_t \odot i_t + m_{t-1} \odot f_t \tag{24}$$

$$\text{Hidden state } h_t = m_t \odot o_t \tag{25}$$

3.6 Generative adversarial networks (GAN)

In the field of deep learning, one of the deep generative models are Generative Adversarial Networks (GANs) introduced by Good Fellow in [43]. GANs are neural networks that can generate synthetic images that closely imitate the original images. In GAN shown in Fig. 5c, there are two neural networks, namely generator, and discriminator, which are trained simultaneously. The generator G generates counterfeit data samples which aim to “fool” the discriminator D , while the discriminator attempts to correctly distinguish the true and false samples. In mathematical terms, D and G play a two player minimax game with the cost function of (26) [64].

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (26)$$

Where x represents the original image, z is a noise vector with random numbers. $p_{data}(x)$ and $p_z(z)$ are probability distributions of x and z , respectively. $D(x)$ represents the probability that x comes from the actual data $p_{data}(x)$ rather than the generated data. $1 - D(G(z))$ is the probability that it can be generated from $p_z(z)$. The expectation of x from the real data distribution p_{data} is expressed by $E_{x \sim p_{data}(x)}$ and the expectation of z sampled from noise is $E_{z \sim p_z(z)}$. The goal of the training is to maximize the loss function for the discriminator, while the training objective for the generator is to reduce the term $\log(1 - D(G(z)))$. The most utilization of GAN in the field of medical image analysis is data augmentation (generating new data) and image to image translation [107]. Trustability of the Generated Data, Unstable Training, and evaluation of generated data are three major drawbacks of GAN that might hinder their acceptance in the medical community [183].

3.7 U-net

Ronneberger et al. [126] proposed CNN based U-Net architecture for segmentation in biomedical image data. The architecture consists of a contracting path (left side) to capture context and an expansive symmetric path (right side) that enables precise localization. U-Net is a generalized DLA used for quantification tasks such as cell detection and shape measurement in medical image data [34].

3.8 Software frameworks

There are several software frameworks available for implementing DLA which are regularly updated as new approaches and ideas are created. DLA encapsulates many levels of mathematical principles based on probability, linear algebra, calculus, and numerical computation. Several deep learning frameworks exist such as Theano, TensorFlow, Caffé, CNTK, Torch, Neon, pylearn, etc. [138]. Globally, Python is probably the most commonly used programming language for DL. PyTorch and Tensorflow are the most widely used libraries for research in 2019. Table 2 shows the analysis of various Deep Learning Frameworks based on the core language and supported interface language.

Table 2 Comparison of various Deep Learning Frameworks

Framework	Core Language	Interface provided	Link
Caffe [61]	C ++	Python, MATLAB, C ++	http://caffe.berkeleyvision.org/
CNTK [186]	C ++	C ++, Python, Brain Script	https://github.com/Microsoft/CNTK
Chainer	–	Python	http://chainer.org/
DL4j	Java	Java, Python, Scala	https://deeplearning4j.org/
MXNet	C ++	Python, R, Scala, Perl, Julia, C ++, etc.	https://github.com/dmlc/mxnet
MatConvNet [158]	–	MATLAB	http://www.vlfeat.org/matconvnet/
Tensor Flow [1]	C ++	–	https://www.tensorflow.org/
Theano [6, 153]	Python	Python	http://deeplearning.net/software/theano/
Torch [25]	Lua	–	http://torch.ch/

4 Use of deep learning in medical imaging

4.1 X-ray image

Chest radiography is widely used in diagnosis to detect heart pathologies and lung diseases such as tuberculosis, atelectasis, consolidation, pleural effusion, pneumothorax, and hyper cardiac inflation. X-ray images are accessible, affordable, and less dose-effective compared to other imaging methods, and it is a powerful tool for mass screening [14]. Table 3 presents a description of the DL methods used for X-ray image analysis.

S. Hwang et al. [57] proposed the first deep CNN-based Tuberculosis screening system with a transfer learning technique. Rajaraman et al. [119] proposed modality-specific ensemble learning for the detection of abnormalities in chest X-rays (CXRs). These model predictions are combined using various ensemble techniques toward minimizing prediction variance. Class selective mapping of interest (CRM) is used for visualizing the abnormal regions in the CXR images. Loey et al. [90] proposed A GAN with deep transfer training for COVID-19 detection in CXR images. The GAN network was used to generate more CXR images due to the lack of the COVID-19 dataset. Waheed et al. [160] proposed a CovidGAN model based on the Auxiliary Classifier Generative Adversarial Network (ACGAN) to produce synthetic CXR images for COVID-19 detection. S. Rajaraman and S. Antani [120] introduced weakly labeled data augmentation for increasing training dataset to improve the COVID-19 detection performance in CXR images.

4.2 Computerized tomography (CT)

CT uses computers and rotary X-ray equipment to create cross-section images of the body. CT scans show the soft tissues, blood vessels, and bones in different parts of the body. CT is a high detection ability, reveals small lesions, and provides a more detailed assessment. CT examinations are frequently used for pulmonary nodule identification [93]. The detection of malignant pulmonary nodules is fundamental to the early diagnosis of lung cancer [102, 142]. Table 4 summarizes the latest deep learning developments in the study of CT image analysis.

Li et al. 2016 [74] proposed deep CNN for the detection of three types of nodules that are semisolid, solid, and ground-glass opacity. Balagourouchetty et al. [5] proposed GoogLeNet based an ensemble FCNet classifier for The liver lesion classification. For feature extraction,

Table 3 An overview of the DLA for the study of X-ray images

Reference	Dataset	Method	Application	Metrics
Lo et al., 1995 [89]	–	CNN	Two-layer CNN, each with 12 5×5 filters for lung nodule detection.	ROC
S.Hwang et al. 2016 [57]	KIT, MC, and Shenzhen	Deep CNN	Tuberculosis screening system with transfer learning technique	AUC
Rajpurkar et al. 2017 [122]	ChestX-ray14	CNN	Detects Pneumonia using CheXNet is a 121-layer CNN from a chest X-ray image.	F1 score
Lopes & Valiati 2017 [91]	Shenzhen and Montgomery	CNN	Comparative analysis of Pre-trained CNN as feature extractors for tuberculosis detection	Accuracy, ROC
Mittal et al. 2018 [99]	JSRT	LF-SegNet	Segmentation of lung field from CXR images using Fully convolutional encoder-decoder network	Accuracy
E.J.Hwang et al. 2019 [58]	57,481 CXR images	CNN	Deep learning-based automatic detection (DLAD) algorithm for tuberculosis detection on CXR	ROC
Souza et al. 2019 [148]	Montgomery	CNN	Segmentation of lungs in CXR for detection and diagnosis of pulmonary diseases using two CNN architecture	Dice coefficient
Hooda et al. [53]	Shenzhen, Montgomery, Belarus, JSRT	CNN	An ensemble of three pre-trained architectures ResNet, AlexNet, and GoogleNet for TB detection	Accuracy, ROC
Xu et al. 2019 [181]	chest X-ray 14	CNN, CXNet-m1	Design a hierarchical CNN structure for a new network CXNet-m1 to detect anomaly of chest X-ray images	Accuracy, F1-score, and AUC
Murphy et al. 2019 [103]	5565 CXR images		Deep learning-based CAD4TB software evaluation	ROC

Table 3 (continued)

Reference	Dataset	Method	Application	Metrics
Rajaraman and Antani 2020 [119]	RSNA, Pediatric pneumonia, and Indiana,	CNN	An ensemble of modality-specific deep learning models for Tuberculosis (TB) detection from CXR	Accuracy, AUC, CI
Capizzi et al. 2020 [15]	Open data set from radiologykey.com	PNN	The fuzzy system, combined with a neural network, can detect low-contrast nodules.	Accuracy
Abbas et al. 2020 [2]	196 X-ray images	CNN	Classification of COVID-19 CXR images using Decompose, Transfer, and Compose	Accuracy, SN, SP
Basu et al. 2020 [7]	225 COVID-19 CXR images	CNN	DETL (Domain Extension Transfer Learning) method for the screening of COVID-19 from CXR images	Accuracy
Wang & Wong 2020 [165]	13,975 X-ray images	CNN	A deep convolutional neural network COVID-Net design for the detection of COVID-19 cases	Accuracy, SN, PPV.
Ozturk et al. 2020 [110]	127 X-ray images	CNN	Deep learning-based DarkCovid net model to detect and classify COVID-19 cases from X-ray images	Accuracy.
Loey et al. 2020 [90]	306 X-ray images	AlexNet google Resnet18 CNN	A GAN with deep transfer learning for COVID-19 detection in limited CXR images.	Accuracy,
Apostolopoulos & Mpesiana 2020 [3]	1427 X-ray images	CNN	Transfer Learning-based CNN architectures to the detection of the Covid-19.	Accuracy, SN, SP

Table 4 A review of articles that use DL techniques for the analysis of the CT image

Reference	Dataset	Method	Application	Metrics
Van Ginneken 2015 [157]	LIDC (865 CT scans)	CNN	Nodule detects in chest CT with pre-trained CNN models from orthogonal patches around the candidate	FROC
Li et al. 2016 [74]	LIDC database.	CNN	Nodule classification with 2D CNN that processes small patches around a nodule	SN, FP/exam Accuracy
Setio et al. 2016 [136]	LIDC-IDRI, ANODE09	Multi-view Conv Net	CNN-based algorithms for pulmonary nodule detection with 9-patches per candidate.	Sensitivity FROC
Shin et al. 2016 [141]	ILD dataset	CNN	Interstitial lung disease (ILD) classification and Lymph node (LN) detection using transfer learning-based CNNs	AUC
Qiang, Yan et al. 2017 [117]	Independent dataset	Deep SDAE-ELM	Discriminative features of nodules in CT and PET images are combined using the fusion method for classification of nodules	SN,SP,AUC,
Onishi Y et al. 2019 [108]	Independent dataset	CNN	CNN trained by Wasserstein GAN for pulmonary nodule classification	SN, SP, AUC Accuracy
Li et al. .2018 [75]	2017 LiTS, 3DIRCADb dataset	H-Dense Unet	H-Dense UNet for tumor and liver segmentation from CT volume	DICE
Pezeshk et al. 2018 [114]	LIDC	3DFCN and 3DCNN	3DFCN is used for nodule candidate generation and 3D CNN for reducing the false-positive rate	FROC
Balagourouchetty et.al 2019 [5]	634 liver CT images	GoogLeNet based FCNet Classifier	The liver lesion classification using GoogLeNet based ensemble FCNet classifier	Accuracy, ROC
Y.Wang et a2019 [166]	Independent dataset	Faster RCNN and ResNet	Intelligent Imaging Layout System (IILS) for the detection and classification of pulmonary nodules	SN, SP AUC Accuracy
Pang et al. 2020 [111]	Shandong Provincial Hospital	CNN (DenseNet)	Classification of lung cancer type from CT images using the DenseNet network.	Accuracy
Masood et al. 2020 [95]	LIDC	mRFCN	Lung nodule classification and detection using mRFCN based automated decision support system	SN, SP, AUC, Accuracy
Zhao and Zeng 2019 [190]	KiTS19 challenge	3D-UNet	Multi-scale supervised 3D U-Net to simultaneously segment kidney and kidney tumors from CT images	DICE, Recall Accuracy Precision
Fan et al. 2020 [35]	COVID-19 infection dataset	Inf-Net	COVID-19 lung CT infection segmentation network	DICE, SN, SP MAE
Li et al. 2020 [79]	4356 Chest CT images	COVNet	COVID-19 detection neural network (COVNet) used for the recognition of COVID-19 from volumetric chest CT exams	AUC, SN, SP

AUC: area under ROC curve; FROC: Area under the Free-Response ROC Curve; SN: sensitivity; SP: specificity; MAE: mean absolute error LIDC: Lung Image Database Consortium; LIDC-IDRI: Lung Image Database Consortium-Image Database Resource Initiative.

basic GoogLeNet architecture is modified with three modifications. Masood et al. [95] proposed the multidimensional Region-based Fully Convolutional Network (mRFCN) for lung nodule detection/classification and achieved a classification accuracy of 97.91%. In lung nodule detection, the feature work is the detection of micronodules (less than 3 mm) without loss of sensitivity and accuracy. Zhao and Zeng 2019 [190] proposed DLA based on supervised MSS U-Net and 3DU-Net to automatically segment kidneys and kidney tumors from CT images. In the present pandemic situation, Fan et al. [35] and Li et al. [79] used deep learning-based techniques for COVID-19 detection from CT images.

4.3 Mammograph (MG)

Breast cancer is one of the world's leading causes of death among women with cancer. MG is a reliable tool and the most common modality for early detection of breast cancer. MG is a low-dose x-ray imaging method used to visualize the breast structure for the detection of breast diseases [40]. Detection of breast cancer on mammography screening is a difficult task in image classification because the tumors constitute a small part of the actual breast image. For analyzing breast lesions from MG, three steps are involved that are detection, segmentation, and classification [139].

The automatic classification and detection of masses at an early stage in MG is still a hot subject of research. Over the past decade, DLA has shown some significant overcome in breast cancer detection and classification problem. Table 5 summarizes the latest DLA developments in the study of mammogram image analysis.

Fonseca et al. [37] proposed a breast composition classification according to the ACR standard based on CNN for feature extraction. Wang et al. [161] proposed twelve-layer CNN to detect Breast arterial calcifications (BACs) in mammograms image for risk assessment of coronary artery disease. Ribli et al. [124] developed a CAD system based on Faster R-CNN for detection and classification of benign and malignant lesions on a mammogram image without any human involvement. Wu et al. [176] present a deep CNN trained and evaluated on over 1,000,000 mammogram images for breast cancer screening exam classification. Conant et al. [26] developed a Deep CNN based AI system to detect calcified lesions and soft-tissue in digital breast tomosynthesis (DBT) images. Kang et al. [62] introduced Fuzzy completely connected layer (FFCL) architecture, which focused primarily on fused fuzzy rules with traditional CNN for semantic BI-RADS scoring. The proposed FFCL framework achieved superior results in BI-RADS scoring for both triple and multi-class classifications.

4.4 Histopathology

Histopathology is the field of study of human tissue in the sliding glass using a microscope to identify different diseases such as kidney cancer, lung cancer, breast cancer, and so on. The staining is used in histopathology for visualization and highlight a specific part of the tissue [45]. For example, Hematoxylin and Eosin (H&E) staining tissue gives a dark purple color to the nucleus and pink color to other structures. H&E stain plays a key role in the diagnosis of different pathologies, cancer diagnosis, and grading over the last century. The recent imaging modality is digital pathology

Deep learning is emerging as an effective method in the analysis of histopathology images, including nucleus detection, image classification, cell segmentation, tissue segmentation, etc.

Table 5 Summary of DLA for MG image analysis

Reference	Dataset	Method	Application	Metrics
Sahiner et al. 1996 [131]	Manually extracted ROIs from 168 mammograms	CNN	CNN for classification of masses and normal tissue on MG.	ROC, TP, FP
Fonseca et al. 2015 [37]	–	CNN	CNN for feature extraction in combing with an SVM as a classifier for breast density estimation	Accuracy
Huych et al. .2016 [56]	607 Digital MG images(219 breast lesions)	CNN	Pre-trained CNN models (MG-CNN) for mass classification	AUC
Wang et al. .2017 [161]	840 standard screening FFDMs	Deep CNN	Detection of cardiovascular disease based on vessel calcification	FROC
Geras et al. 2017 [41]	Screening mammograms images 129, 208	MV-CNN	Multi-view deep CNN for breast cancer screening and image resolution on the prediction accuracy	Accuracy, ROC, TP, FP
Zhang et al. 2017 [188]	3000 MG images	CNN	Data augmentation and transfer learning methods with a CNN for classification	ROC
Wu et al. 2017 [177]	200,000 Breast cancer screening exams	DCN	Deep CNN for breast density classification	AUC
Kyono et al. 2018 [69]	Private dataset of 8162 patients	MAMMO-CNN	MAMMO is a novel multi-view CNN with multi-task learning (MTL) a clinical decision support system capable of triaging MG	Accuracy
Lehman et al. [72]	41,479 Mammogram images	ResNet-18	Deep learning-based CNN for mammographic breast density classification	Accuracy
Kim et al. 2018 [65]	29,107 Digital MG (24,765 normal cases and 4339 cancer cases)	DIB-MG	DIB-MG is weakly supervised learning. DIB-MG learns radiologic features without any human annotations.	SN, SP, Accuracy
Ribli et al. 2018 [124]	DDSM (2620), INbreast (115), Private database	Faster <i>R</i> -CNN, VGG16	CNN detects and classifies malignant or benign lesions on MG images	AU
Chougrad et al. 2018 [23]	MIAS,DDSM, INbreast, BCDR	VGG16, ResNet50, Inceptionv3	Transfer learning and fine-tuning strategy based CNN to classify MG mass lesions	AUC, Accuracy
Karthik et al. 2018 [63]	WBCD	DNN-RFS	Deep neural network (DNN) as a classifier model for breast cancer data	Accuracy, Precision, SP, SN, F-score
Cai et al. 2019 [13]	990 MG images, 540 Malignant masses, and 450 benign lesions	DCNN	Deep CNN for microcalcification discrimination for breast cancer screening	Accuracy, Precision, SP, AUC, SN
Wu et al. 2019 [176]	1000 000 images	DCNN	CNN-based breast cancer screening classifier	AUC
Conant et al. .2019 [26]	12,000 cases, including 4000	DCNN	Deep CNN based system detected soft tissue and	AUC

Table 5 (continued)

Reference	Dataset	Method	Application	Metrics
Rodriguez-Ruiz et al. 2019 [125]	biopsy-proven cancers 9000 Cancer cases and 180,000 normal cases	DCNN	calcific lesions in the DBT images CNN based CAD system	AUC
Ionescu et al. 2019 [59]	Radiologists Private data set	CNN	Breast density estimation and risk scoring	

MIAS: Mammographic Image Analysis Society dataset; DDSM: Digital Database for Screening Mammography; BI-RADS: Breast Imaging Reporting and Data System; WBCD: Wisconsin Breast Cancer Dataset; DIB-MG: data-driven imaging biomarker in mammography. FFDMs: Full-Field Digital Mammograms; MAMMO: Man and Machine Mammography Oracle; FROC: Free response receiver operating characteristic analysis; SN: sensitivity; SP: specificity.

[178]. Tables 6 and 7 summarize the latest deep learning developments in pathology. In the study of digital pathology image analysis, the latest development is the introduction of whole slide imaging (WSI). WSI allows digitizing glass slides with stained tissue sections at high resolution. Dimitriou et al. [30] reviewed challenges for the analysis of multi-gigabyte WSI images for building deep learning models. A. Serag et al. [135] discuss different public “Grand Challenges” that have innovations using DLA in computational pathology.

4.5 Other images

Endoscopy is the insertion of a long nonsurgical solid tube directly into the body for the visual examination of an internal organ or tissue in detail. Endoscopy is beneficial in studying several systems inside the human body, such as the gastrointestinal tract, the respiratory tract, the urinary tract, and the female reproductive tract [60, 101]. Du et al. [31] reviewed the Applications of Deep Learning in the Analysis of Gastrointestinal Endoscopy Images. A revolutionary device for direct, painless, and non-invasive inspection of the gastrointestinal (GI) tract for detecting and diagnosing GI diseases (ulcer, bleeding) is Wireless capsule endoscopy (WCE). Soffer et al. [145] performed a systematic analysis of the existing literature on the implementation of deep learning in the WCE. The first deep learning-based framework was proposed by He et al. [46] for the detection of hookworm in WCE images. Two CNN networks integrated (edge extraction and classification of hookworm) to detect hookworm. Since tubular structures are crucial elements for hookworm detection, the edge extraction network was used for tubular region detection. Yoon et al. [185] developed a CNN model for early gastric cancer (EGC) identification and prediction of invasion depth. The depth of tumor invasion in early gastric cancer (EGC) is a significant factor in deciding the method of treatment. For the classification of endoscopic images as EGC or non-EGC, the authors employed a VGG-16 model. Nakagawa et al. [105] applied DL technique based on CNN to enhance the diagnostic assessment of oesophageal wall invasion using endoscopy. J.choi et al. [22] express the feature aspects of DL in endoscopy.

Positron Emission Tomography (PET) is a nuclear imaging tool that is generally used by the injection of particular radioactive tracers to visualize molecular-level activities within tissues. T. Wang et al. [168] reviewed applications of machine learning in PET attenuation correction (PET AC) and low-count PET reconstruction. The authors discussed the advantages

Table 6 Summary of articles using DLA for digital pathology image - Organ segmentation

Reference	Staining/ Image modality	Method	Application	Dataset	Metrics
Ronneberger et al. 2015 [126]	EM	U-net architecture with deformation augmentation	Segmentation of neuronal structures, cell segmentation	ISBI cell tracking challenge 2014 and 2015	Warping, Rand, Pixel Error
Song et al. 2016 [147]	Pap, H & E	Multi-scale CNN model	Segmentation of cervical cells in Pap smear images	ISBI 2015 Challenge, Shenzhen University (SZU) Dataset	Dice Coefficient
Xing et al. 2016 [179]	IHC H & E,	CNN and sparse shape model	Nuclei segmentation	Private set containing brain tumor (31), pancreatic NET (22), breast cancer (35) images	–
Chen et al. 2017 [19]	H & E	Multi-task learning framework with contour-aware FCN model for instance segmentation	Deep contour-aware CNN Segmentation of colon glands	GLAS challenge (165 images), MICCAI2015 nucleus segmentation challenge (33 images)	Dice coefficient
Van Eycke et al. (2018) [156]	H & E	Integration of DCAN, UNet, and ResNet models	Segmentation of glandular epithelium in H & E and IHC staining images	GLAS challenge (165 images) and a private set containing colorectal tissue microscopy images	F1-score, object dice coefficient
Liang et al. 2018 [81]	H & E	Patch-based FCN+iterative learning approach	first-time deep learning applied to the gastric tumor segmentation	2017 China Big Data and AI challenge (1900 images)	Mean IoU, mean accuracy
Qu et al. 2019 [118]	H & E	FCN trained with perceptual loss	Jointly classifies and segments various types of nuclei from histopathology images	40 tissue images of lung adenocarcinoma (private set)	F1score, Dice coefficient accuracy,
Pinckaers and Litjens 2019 [115]	H & E	Incorporating NODE in U - Net to allow an adaptive receptive field	Segmentation of colon glands	GLAS challenge (165) images	Object Dice, F1 score
Gadermayr et al. 2019 [39]	Stain agnostic	CycleGAN + UNet segmentation	Multi-Domain Unsupervised Segmentation of object-of interest in WSIs	23 PAS, 6 AFOG, 6 CoI3 & 6 CD31 WSIs	F1 score
Sun et al. 2019 [149]	H & E	Multi-scale modules and specific convolutional operations	Deep learning architecture for gastric cancer segmentation	500 pathological images of gastric areas, with cancerous regions	–

Table 7 Summary of articles using DLA for digital pathology image - Detection and classification of disease

Reference	Staining/ image modality	Method	Application	Data set
Xu et al. 2016 [182]	H&E	Stacked sparse autoencoders	Nucleus detection from Breast Cancer Histopathology Images	537 H&E images from Case Western Reserve University
Coudray et al. (2018) [27]	H&E	Patch-based Inception-V3 model	Lung cancer histopathology images classify them into LUAD, LUSC, or normal lung tissue	FFPE sections (140 s) Frozen sections (98 s), and lung biopsies (102 s)
Song et al. 2018 [146]	H&E	Deep autoencoder	Simultaneous detection and classification of cells in bone marrow histology images	–
Yi et al. 2018 [184]	H&E	FCN	Microvessel prediction in H&E Stained Pathology Images	Lung adenocarcinoma (ADC) patients images 38
Bulten and Litjens 2018 [12]	H&E, IHC	Self-clustering Convolutional adverse Arial Autoencoders	Classification of the pros take into tumor vs non-tumor	94 registered WSIs from Radboud University Medical Center
Valkonen et al. 2019 [154]	ER, PR, Ki-67	Fine-tuning partially pre-trained CNN network	Recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67	Digital Pan CK (152 – invasive breast cancer images)
Wei et al. 2019 [169]	H&E	ResNet-18 based patch classifier	Classification of histologic subtypes on lung adenocarcinoma	143 WSIs private set
Wang et al. (2019) [167]	H & E	Patch-based FCN and context-aware block selection + feature aggregation strategy	Lung cancer image classification	Private (939 WSIs), TCGA (500 WSIs)
Li et al. 2019 [80]	H & E	FCN trained with a concentric loss on weakly annotated centroid label	Mitosis detection in breast histopathology images	ICPR12 (50 images), ICPR14 (1696 images), AMIDA13 (606 images), TUPAC16 (107 images)
Tabibu et al. .2019 [152]	H & E	Pre-trained Res Net based patch classifier	Classification of Renal Cell Carcinoma subtypes and survival prediction	TCGA(2, 093WSI)
Lin et al. 2019 [83]	H & E	Fast Scan Net: FCN based model	Automatic detection of breast cancer metastases from whole-slide image	2016 Camelyon Grand Challenge (400 WSI)

NODE: Neural Ordinary Differential Equations; IoU: mean Intersection over Union coefficient

of deep learning over machine learning in the applications of PET images. AJ reader et al. [123] reviewed the reconstruction of PET images that can be used in deep learning either directly or as a part of traditional reconstruction methods.

5 Discussion

The primary purpose of this paper is to review numerous publications in the field of deep learning applications in medical images. Classification, detection, and segmentation are

essential tasks in medical image processing [144]. For specific deep learning tasks in medical applications, the training of deep neural networks needs a lot of labeled data. But in the medical field, at least thousands of labeled data is not available. This issue is alleviated by a technique called transfer learning. Two transfer learning approaches are popular and widely applied that are fixed feature extractors and fine-tuning a pre-trained network. In the classification process, the deep learning models are used to classify images into two or more classes. In the detection process, Deep learning models have the function of identifying tumors and organs in medical images. In the segmentation task, deep learning models try to segment the region of interest in medical images for processing.

5.1 Segmentation

For medical image segmentation, deep learning has been widely used, and several articles have been published documenting the progress of deep learning in the area. Segmentation of breast tissue using deep learning alone has been successfully implemented [104]. Xing et al. [179] used CNN to acquire the initial shape of the nucleus and then isolate the actual nucleus using a deformable pattern. Qu et al. [118] suggested a deep learning approach that could segment the individual nucleus and classify it as a tumor, lymphocyte, and stroma nuclei. Pinckaers and Litjens [115] show on a colon gland segmentation dataset (GlaS) that these Neural Ordinary Differential Equations (NODE) can be used within the U-Net framework to get better segmentation results. Sun 2019 [149] developed a deep learning architecture for gastric cancer segmentation that shows the advantage of utilizing multi-scale modules and specific convolution operations together. Figure 6 shows U-Net is the most usually used network for segmentation (Fig. 6).

5.2 Detection

The main challenge posed by methods of detection of lesions is that they can give rise to multiple false positives while lacking a good proportion of true positive ones. For tuberculosis detection using deep learning methods applied in [53, 57, 58, 91, 119]. Pulmonary nodule detection using deep learning has been successfully applied in [82, 108, 136, 157].

Shin et al. [141] discussed the effect of CNN pre-trained architectures and transfer learning on the identification of enlarged thoracoabdominal lymph nodes and the diagnosis of interstitial lung disease on CT scans, and considered transfer learning to be helpful, given the fact that natural images vary from medical images. Litjens et al. [85] introduced CNN for the

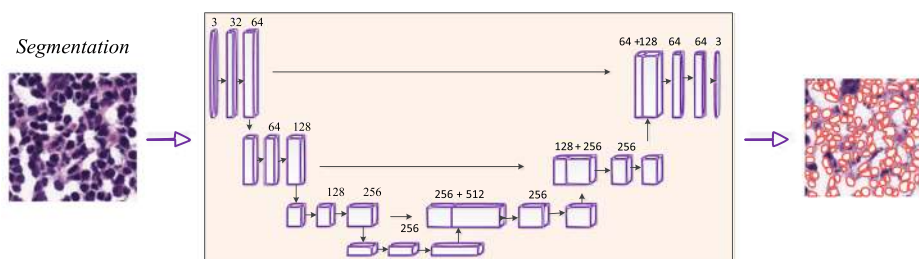


Fig. 6 U-Net architecture for segmentation, comprising encoder (downsampling) and decoder (upsampling) sections [135]

identification of Prostate cancer in biopsy specimens and breast cancer metastasis identification in sentinel lymph nodes. The CNN has four convolution layers for feature extraction and three classification layers. Riddle et al. [124] proposed the Faster R-CNN model for the detection of mammography lesions and classified these lesions into benign and malignant, which finished second in the Digital Mammography DREAM Challenge. Figure 7 shows VGG architecture for detection.

An object detection framework named Clustering CNN (CLU-CNNs) was proposed by Z. Li et al. [76] for medical images. CLU-CNNs used Agglomerative Nesting Clustering Filtering (ANCF) and BN-IN Net to avoid much computation cost facing medical images. Image saliency detection aims at locating the most eye-catching regions in a given scene [21, 78]. The goal of image saliency detection is to locate a given scene in the most eye-catching regions. In different applications, it also acts as a pre-processing tool including video saliency detection [17, 18], object recognition, and object tracking [20]. Saliency maps are a commonly used tool for determining which areas are most important to the prediction of a trained CNN on the input image [92]. NT Arun et al. [4] evaluated the performance of several popular saliency methods on the RSNA Pneumonia Detection dataset and was found that GradCAM was sensitive to the model parameters and model architecture.

5.3 Classification

In classification tasks, deep learning techniques based on CNN have seen several advancements. The success of CNN in image classification has led researchers to investigate its usefulness as a diagnostic method for identifying and characterizing pulmonary nodules in CT images. The classification of lung nodules using deep learning [74, 108, 117, 141] has also been successfully implemented.

Breast parenchymal density is an important indicator of the risk of breast cancer. The DL algorithms used for density assessment can significantly reduce the burden of the radiologist. Breast density classification using DL has been successfully implemented [37, 59, 72, 177]. Ionescu et al. [59] introduced a CNN-based method to predict Visual Analog Score (VAS) for breast density estimation. Figure 8 shows AlexNet architecture for classification.

Alcoholism or alcohol use disorder (AUD) has effects on the brain. The structure of the brain was observed using the Neuroimaging approach. S.H.Wang et al. [162] proposed a 10-layer CNN for alcohol use disorder (AUD) problem using dropout, batch normalization, and PReLU techniques. The authors proposed a 10 layer CNN model that has obtained a sensitivity of 97.73, a specificity of 97.69, and an accuracy of 97.71. Cerebral micro-bleeding (CMB) are small chronic brain hemorrhages that can result in cognitive impairment, long-term disability, and neurologic dysfunction. Therefore, early-stage identification of CMBs for prompt treatment is essential. S. Wang et al. [164] proposed the

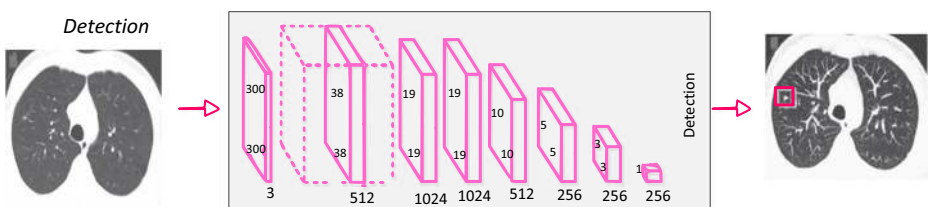


Fig. 7 CNN architecture for detection [144]

transfer learning-based DenseNet to detect Cerebral micro-bleedings (CMBs). DenseNet based model attained an accuracy of 97.71% (Fig. 8).

5.4 Limitations and challenges

The application of deep learning algorithms to medical imaging is fascinating, but many challenges are pulling down the progress. One of the limitations to the adoption of DL in medical image analysis is the inconsistency in the data itself (resolution, contrast, signal-to-noise), typically caused by procedures in clinical practice [113]. The non-standardized acquisition of medical images is another limitation in medical image analysis. The need for comprehensive medical image annotations limits the applicability of deep learning in medical image analysis. The major challenge is limited data and compared to other datasets, the sharing of medical data is incredibly complicated. Medical data privacy is both a sociological and a technological issue that needs to be discussed from both viewpoints. For building DLA a large amount of annotated data is required. Annotating medical images is another major challenge. Labeling medical images require radiologists' domain knowledge. Therefore, it is time-consuming to annotate adequate medical data. Semi-supervised learning could be implemented to make combined use of the existing labeled data and vast unlabelled data to alleviate the issue of "limited labeled data". Another way to resolve the issue of "data scarcity" is to develop few-shot learning algorithms using a considerably smaller amount of data. Despite the successes of DL technology, there are many restrictions and obstacles in the medical field. Whether it is possible to reduce medical costs, increase medical efficiency, and improve the satisfaction of patients using DL in the medical field cannot be adequately checked. However, in clinical trials, it is necessary to demonstrate the efficacy of deep learning methods and to develop guidelines for the medical image analysis applications of deep learning.

6 Conclusion and future directions

Medical imaging is a place of origin of the information necessary for clinical decisions. This paper discusses the new algorithms and strategies in the area of deep learning. In this brief introduction to DLA in medical image analysis, there are two objectives. The first one is an introduction to the field of deep learning and the associated theory. The second is to provide a general overview of the medical image analysis using DLA. It began with the history of neural networks since 1940 and ended with breakthroughs in medical applications in recent DL algorithms. Several supervised and unsupervised DL algorithms are first discussed, including auto-encoders, recurrent, CNN, and restricted Boltzmann machines. Several optimization techniques and frameworks in this area include Caffe, TensorFlow, Theano, and PyTorch are discussed. After that, the most successful DL methods were reviewed in various medical

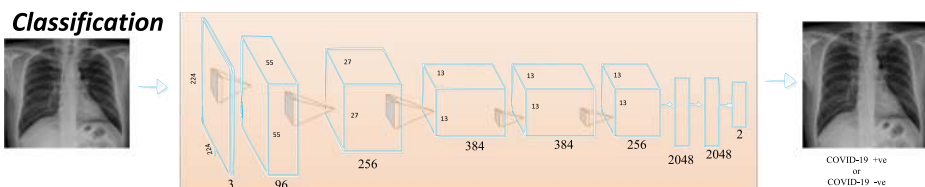


Fig. 8 CNN architecture for classification [144]

image applications, including classification, detection, and segmentation. Applications of the RBM network is rarely published in the medical image analysis literature. In classification and detection, CNN-based models have achieved good results and are most commonly used. Several existing solutions to medical challenges are available. However, there are still several issues in medical image processing that need to be addressed with deep learning. Many of the current DL implementations are supervised algorithms, while deep learning is slowly moving to unsupervised and semi-supervised learning to manage real-world data without manual human labels.

DLA can support clinical decisions for next-generation radiologists. DLA can automate radiologist workflow and facilitate decision-making for inexperienced radiologists. DLA is intended to aid physicians by automatically identifying and classifying lesions to provide a more precise diagnosis. DLA can help physicians to minimize medical errors and increase medical efficiency in the processing of medical image analysis. DL-based automated diagnostic results using medical images for patient treatment are widely used in the next few decades. Therefore, physicians and scientists should seek the best ways to provide better care to the patient with the help of DLA. The potential future research for medical image analysis is the designing of deep neural network architectures using deep learning. The enhancement of the design of network structures has a direct impact on medical image analysis. Manual design of DL Model structure requires rich knowledge; hence Neural Network Search will probably replace the manual design [73]. A meaningful feature research direction is also the design of various activation functions. Radiation therapy is crucial for cancer treatment. Different medical imaging modalities are playing a critical role in treatment planning. Radiomics was defined as the extraction of high throughput features from medical images [28]. In the future, Deep-learning analysis of radiomics will be a promising tool in clinical research for clinical diagnosis, drug development, and treatment selection for cancer patients. Due to limited annotated medical data, unsupervised, weakly supervised, and reinforcement learning methods are the emerging research areas in DL for medical image analysis. Overall, deep learning, a new and fast-growing field, offers various obstacles as well as opportunities and solutions for a range of medical image applications.

References

1. Abadi M *et al.* (2016) TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, [Online]. Available: <http://arxiv.org/abs/1603.04467>.
2. Abbas A, Abdelsamea MM, Gaber MM (2020) Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, pp. 1–9, [Online]. Available: <http://arxiv.org/abs/2003.13815>.
3. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks, *Phys Eng Sci Med*, no. 0123456789, pp. 1–6, DOI: <https://doi.org/10.1007/s13246-020-00865-4>.
4. Arun NT *et al.* (2020) Assessing the validity of saliency maps for abnormality localization in medical imaging, pp. 1–5, [Online]. Available: <http://arxiv.org/abs/2006.00063>.
5. L. Balagourouchetty, J. K. Pragatheeswaran, B. Pottakkat, and R. G, “GoogLeNet based ensemble FCNet classifier for focal liver lesion diagnosis,” *IEEE J Biomed Heal Inf*, vol. 2194, no. c, pp. 1–1, 2019, DOI: <https://doi.org/10.1109/jbhi.2019.2942774>, 1694.
6. Bastien F *et al.* (2012) Theano: new features and speed improvements, pp. 1–10, [Online]. Available: <http://arxiv.org/abs/1211.5590>.
7. Basu S, Mitra S, Saha N (2020) Deep Learning for Screening COVID-19 using Chest X-Ray Images, pp. 1–6, [Online]. Available: <http://arxiv.org/abs/2004.10507>.

8. Bauer S, Wiest R, Nolte LP, Reyes M (2013) A survey of MRI-based medical image analysis for brain tumor studies. *Phys Med Biol* 58(13):1–44. <https://doi.org/10.1088/0031-9155/58/13/R97>
9. Bengio Y, Lamblin P, Popovici D, Larochelle H (2006) Greedy layer-wise training of deep networks. In: *The 19th International Conference on Neural Information Processing Systems(NIPS'06)*, pp 153–160. <https://doi.org/10.5555/2976456.2976476>
10. Bengio Y, Simard P, Palo F (1994) Learning long -term dependencies with gradient descent is difficult. *IEEE Trans Neural Netw* 5(2):157–166
11. Bizopoulos P, Koutsouris D (2019) Deep learning in cardiology. *IEEE Rev Biomed Eng* 12(c):168–193. <https://doi.org/10.1109/RBME.2018.2885714>
12. Bulten W, Litjens G (2018) Unsupervised Prostate Cancer Detection on H&E using Convolutional Adversarial Autoencoders, [Online]. Available: <http://arxiv.org/abs/1804.07098>.
13. Cai H *et al.* (2019) Breast Microcalcification Diagnosis Using Deep Convolutional Neural Network from Digital Mammograms, *Comput Math Methods Med*, vol. 2019, DOI: <https://doi.org/10.1155/2019/2717454>.
14. Candemir S, Rajaraman S, Thoma G, Antani S (2018) Deep learning for grading cardiomegaly severity in chest x-rays : an investigation. In: *2018 IEEE Life Sciences Conference (LSC)*, pp 109–113. <https://doi.org/10.1109/LSC.2018.8572113>
15. Capizzi G, Lo Sciuto G, Napoli C, Polap D (2020) Small Lung Nodules Detection based on Fuzzy-Logic and Probabilistic Neural Network with Bio-inspired Reinforcement Learning, *IEEE Trans Fuzzy Syst*, vol. PP, no. XX, p. 1. <https://doi.org/10.1109/TFUZZ.2019.2952831>.
16. Chen DS, Jain RC (1994) A robust back propagation learning algorithm for function approximation. *IEEE Trans. Neural Networks* 5(3):467–479. <https://doi.org/10.1109/72.286917>
17. Chen C, Li S, Qin H, Pan Z, Yang G (2018) Bilevel feature learning for video saliency detection. *IEEE Trans Multimed* 20(12):3324–3336. <https://doi.org/10.1109/TMM.2018.2839523>
18. Chen C, Li S, Wang Y, Qin H, Hao A (2017) Video saliency detection via spatial-temporal fusion and low-rank coherency diffusion. *IEEE Trans Image Process* 26(7):3156–3170. <https://doi.org/10.1109/TIP.2017.2670143>
19. Chen H, Qi X, Yu L, Dou Q, Qin J, Heng PA (2017) DCAN: deep contour-aware networks for object instance segmentation from histology images. *Med Image Anal* 36:135–146. <https://doi.org/10.1016/j.media.2016.11.004>
20. Chen C, Wang G, Peng C, Zhang X, Qin H (2020) Improved robust video saliency detection based on long-term spatial-temporal information. *IEEE Trans Image Process* 29:1090–1100. <https://doi.org/10.1109/TIP.2019.2934350>
21. Chen C, Wei J, Peng C, Zhang W, Qin H (2020) Improved saliency detection in RGB-D images using two-phase depth estimation and selective deep fusion. *IEEE Trans Image Process* 29:4296–4307. <https://doi.org/10.1109/TIP.2020.2968250>
22. Choi J, Shin K, Jung J, Bae HJ, Kim DH, Byeon JS, Kim N (2020) Convolutional neural network technology in endoscopic imaging: artificial intelligence for endoscopy. *Clin Endosc* 53(2):117–126. <https://doi.org/10.5946/ce.2020.054>
23. Chougrad H, Zouaki H, Alheyane O (2018) Deep convolutional neural networks for breast cancer screening. *Comput Methods Prog Biomed* 157:19–30. <https://doi.org/10.1016/j.cmpb.2018.01.011>
24. Clevert DA, Unterthiner T, Hochreiter S (2016) Fast and accurate deep network learning by exponential linear units (ELUs). In: *4th International Conference on Learning Representations, ICLR 2016*, pp 1–14
25. Collobert R, Kavukcuoglu K, Farabet C (2011) Torch7: A matlab-like environment for machine learning, *BigLearn, NIPS Work*, pp. 1–6, [Online]. Available: http://infoscience.epfl.ch/record/192376/files/Collobert_NIPSWORKSHOP_2011.pdf.
26. Conant EF *et al* (2019) Improving Accuracy and Efficiency with Concurrent Use of Artificial Intelligence for Digital Breast Tomosynthesis. *Radiol Artif Intell* 1(4):e180096. <https://doi.org/10.1148/ryai.2019180096>
27. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, Moreira AL, Razavian N, Tsirigos A (2018) Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med* 24(10):1559–1567. <https://doi.org/10.1038/s41591-018-0177-5>
28. Dercele L, Henry T, Carré A, Paragios N, Deutsch E, Robert C (2020) Reinventing radiation therapy with machine learning and imaging bio-markers (radiomics): State-of-the-art, challenges and perspectives, *Methods*, no. May, pp. 0–1, DOI: <https://doi.org/10.1016/j.ymeth.2020.07.003>.
29. Dhillon A, Verma GK (2019) Convolutional neural network: a review of models, methodologies, and applications to object detection *Prog Artif Intell*, no. 0123456789, DOI: <https://doi.org/10.1007/s13748-019-00203-0>.
30. Dimitriou N, Arandjelović O, Caie PD (2019) Deep Learning for Whole Slide Image Analysis: An Overview. *Front Med* 6(November):1–7. <https://doi.org/10.3389/fmed.2019.00264>

31. Du W et al (2019) Review on the applications of deep learning in the analysis of gastrointestinal endoscopy images. *IEEE Access* 7:142053–142069. <https://doi.org/10.1109/ACCESS.2019.2944676>
32. Dugas C, Bengio Y, Bélisle F, Nadeau C, Garcia R (2000) Incorporating second-order functional knowledge for better option pricing. In: 13th International Conference on Neural Information Processing Systems (NIPS'00), pp 451–457. <https://doi.org/10.5555/3008751.3008817>
33. Eberhart RC, Dobbins RW (1990) Early neural network development history: the age of Camelot. *IEEE Eng Med Biol Mag* 9(3):15–18. <https://doi.org/10.1109/51.59207>
34. Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, Böhm A, Deubner J, Jäckel Z, Seiwald K, Dovzhenko A, Tietz O, Dal Bosco C, Walsh S, Saltukoglu D, Tay TL, Prinz M, Palme K, Simons M, Diester I, Brox T, Ronneberger O (2019) U-net: deep learning for cell counting, detection, and morphometry. *Nat Methods* 16(1):67–70. <https://doi.org/10.1038/s41592-018-0261-2>
35. Fan D-P et al. (2020) Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Scans, pp. 1–10, [Online]. Available: <http://arxiv.org/abs/2004.14133>.
36. Fischer A, Igel C (2014) Training restricted Boltzmann machines: an introduction. *Pattern Recogn* 47(1): 25–39. <https://doi.org/10.1016/j.patcog.2013.05.025>
37. Fonseca P et al (2015) Automatic breast density classification using a convolutional neural network architecture search procedure. *Med Imaging 2015 Comput Diagnosis* 9414(c):941428. <https://doi.org/10.1117/12.2081576>
38. Fukushima K (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern* 36(4):193–202. <https://doi.org/10.1007/BF00344251>
39. Gadermayr M, Gupta L, Appel V, Boor P, Klinkhammer BM, Merhof D (2019) Generative adversarial networks for facilitating stain-independent supervised and unsupervised segmentation: a study on kidney histology. *IEEE Trans Med Imaging* 38(10):2293–2302. <https://doi.org/10.1109/TMI.2019.2899364>
40. Gardezi SJS, Elazab A, Lei B, Wang T (2019) Breast cancer detection and diagnosis using mammographic data: systematic review. *J Med Internet Res* 21(7):1–22. <https://doi.org/10.2196/14464>
41. Geras KJ et al. (2017) High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional Neural Networks, pp. 1–9, [Online]. Available: <http://arxiv.org/abs/1703.07047>.
42. Goodfellow I, Bengio Y, Courville A (2016) “Deep learning,” DOI: <https://doi.org/10.1038/nmeth.3707>
43. Goodfellow IJ et al (2014) Generative adversarial nets. *Adv Neural Inf Process Syst* 3(January): 2672–2680
44. Greenspan H, Van Ginneken B, Summers RM (2016) Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. *IEEE Trans Med Imaging* 35(5):1153–1159. <https://doi.org/10.1109/TMI.2016.2553401>
45. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B (2009) Histopathological image analysis: a review. *IEEE Rev Biomed Eng* 2:147–171. <https://doi.org/10.1109/RBME.2009.2034865>
46. He JY, Wu X, Jiang YG, Peng Q, Jain R (2018) Hookworm detection in wireless capsule endoscopy images with deep learning. *IEEE Trans Image Process* 27(5):2379–2392. <https://doi.org/10.1109/TIP.2018.2801119>
47. He K, Zhang X, Ren S., Sun J (2015) Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proc IEEE Int Conf Comput Vis*, vol. 2015 Inter, pp 1026–1034, DOI: <https://doi.org/10.1109/ICCV.2015.123>.
48. He K, Zhang X, Ren S, Sun J (2015) Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans Pattern Anal Mach Intell* 37(9):1904–1916
49. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 2016(Decem):770–778. <https://doi.org/10.1109/CVPR.2016.90>
50. Hinton G (2014) Boltzmann Machines, *Encycl Mach Learn Data Min*, no. 1, pp. 1–7, DOI: https://doi.org/10.1007/978-1-4899-7502-7_31-1.
51. Hinton GE, Osindero S, Teh Y-W (2006) A fast learning algorithm for deep belief nets. *Neural Comput* 18:1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
52. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
53. Hooda R, Mittal A, Sofat S (2019) Automated TB classification using ensemble of deep architectures. *Multimed Tools Appl* 78(22):31515–31532. <https://doi.org/10.1007/s11042-019-07984-5>
54. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. *Proc - 30th IEEE Conf Comput Vis Pattern Recognition, CVPR 2017* 2017(Janua):2261–2269. <https://doi.org/10.1109/CVPR.2017.243>
55. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J Physiol* 160(1):106–154. <https://doi.org/10.1113/jphysiol.1962.sp006837>

56. Huynh BQ, Li H, Giger ML (2016) Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. *J Med Imaging* 3(3):034501. <https://doi.org/10.1117/1.jmi.3.3.034501>
57. Hwang S, Kim H-E, Jeong J, Kim H-J (2016) A novel approach for tuberculosis screening based on deep convolutional neural networks. *Med Imaging 2016 Comput Diagnosis* 9785:97852W. <https://doi.org/10.1117/12.2216198>
58. Hwang EJ, Park S, Jin KN, Kim JI, Choi SY, Lee JH, Goo JM, Aum J, Yim JJ, Park CM, Deep Learning-Based Automatic Detection Algorithm Development and Evaluation Group, Kim DH, Woo W, Choi C, Hwang IP, Song YS, Lim L, Kim K, Wi JY, Oh SS, Kang MJ (2019) Development and validation of a deep learning–based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs. *Clin Infect Dis* 69(5):739–747. <https://doi.org/10.1093/cid/ciy967>
59. Ionescu GV et al (2019) Prediction of reader estimates of mammographic density using convolutional neural networks. *J Med Imaging* 6(03):1. <https://doi.org/10.1117/1.jmi.6.3.031405>
60. Jani KK, Srivastava R (2019) A survey on medical image analysis in capsule endoscopy. *Curr Med Imaging Rev* 15(7):622–636. <https://doi.org/10.2174/1573405614666181102152434>
61. Jia Y et al. (2014) Caffe: Convolutional architecture for fast feature embedding.” *MM 2014 – Proc 2014 ACM Conf Multimed* , pp. 675–678, DOI: <https://doi.org/10.1145/2647868.2654889>.
62. Kang C, Yu X, Wang SH, Guttery DS, Pandey HM, Tian Y, Zhang YD (2020) A heuristic neural network structure relying on fuzzy logic for images scoring. *IEEE Trans Fuzzy Syst* 6706(c):1–1. <https://doi.org/10.1109/tfuzz.2020.2966163>
63. S. Karthik, R. Srinivasa Perumal, and P. V. S. S. R. Chandra Mouli, “Breast cancer classification using deep neural networks,” *Knowl Comput Its Appl Knowl Manip Process Tech Vol. 1*, pp. 227–241, 2018, DOI: https://doi.org/10.1007/978-981-10-6680-1_12
64. Kazemina S et al. (2020) GANs for Medical Image Analysis,” *Artif Intell Med*, p. 104262, DOI: <https://doi.org/10.1016/j.jece.2020.104262>.
65. Kim EK, Kim HE, Han K, Kang BJ, Sohn YM, Woo OH, Lee CW (2018) Applying data-driven imaging biomarker in mammography for breast Cancer screening: preliminary study. *Sci Rep* 8(1):1–8. <https://doi.org/10.1038/s41598-018-21215-1>
66. Kingma DP, Welling M Auto-encoding variational bayes. In: 2nd International Conference on Learning, ICLR 2014, vol 2014, pp 1–14
67. Klambauer G, Unterthiner T, Mayr A, Hochreiter S (2017) Self-normalizing neural networks. *Adv Neural Inf Process Syst* 2017(Decem):972–981
68. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: *The 25th International Conference on Neural Information Processing Systems*, pp 1097–1105. <https://doi.org/10.1145/3065386>
69. Kyono T, Gilbert FJ, van der Schaar M (2018) MAMMO: A Deep Learning Solution for Facilitating Radiologist-Machine Collaboration in Breast Cancer Diagnosis, pp. 1–18, [Online]. Available: <http://arxiv.org/abs/1811.02661>.
70. LeCun Y, Bengio Y (1998) Convolutional networks for images, speech, and time-series. In: *The handbook of brain theory and neural networks*, pp 255–258
71. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, Jackel LD (1989) Backpropagation applied to digit recognition. *Neural Comput* 1(4):541–551
72. Lehman CD, Yala A, Schuster T, Dontchos B, Bahl M, Swanson K, Barzilay R (2019) Mammographic breast density assessment using deep learning: clinical implementation. *Radiology* 290(1):52–58. <https://doi.org/10.1148/radiol.2018180694>
73. Lei T, Wang R, Wan Y, Du X, Meng H, Nandi AK (2020) Medical Image Segmentation Using Deep Learning: A survey, vol. 171, pp. 17–31, DOI: https://doi.org/10.1007/978-3-030-32606-7_2.
74. Li W, Cao P, Zhao D, Wang J (2016) Pulmonary Nodule Classification with Deep Convolutional Neural Networks on Computed Tomography Images, *Comput Math Methods Med*, vol. 2016, DOI: <https://doi.org/10.1155/2016/6215085>.
75. Li X, Chen H, Qi X, Dou Q, Fu CW, Heng PA (2018) H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans Med Imaging* 37(12):2663–2674. <https://doi.org/10.1109/TMI.2018.2845918>
76. Li Z, Dong M, Wen S, Hu X, Zhou P, Zeng Z (2019) CLU-CNNs: Object detection for medical images. *Neurocomputing* 350(May):53–59. <https://doi.org/10.1016/j.neucom.2019.04.028>
77. Li Y, Huang C, Ding L, Li Z, Pan Y, Gao X (2019) Deep learning in bioinformatics: introduction, application, and perspective in the big data era. *Methods* 166:4–21. <https://doi.org/10.1016/j.ymeth.2019.04.008>

78. Li Y, Li S, Chen C, Hao A, Qin H (2020) A Plug-and-play Scheme to Adapt Image Saliency Deep Model for Video Data, *IEEE Trans Circuits Syst Video Technol*, no. Xx, pp. 1–1, DOI: <https://doi.org/10.1109/tcsvt.2020.3023080>.
79. Li L, Qin L, Yin Y, Wang X et al (2019) Artificial Intelligence Distinguishes COVID-19 from Community Acquired Pneumonia on Chest CT. *Radiology* 2020:1–5. <https://doi.org/10.1007/s10489-020-01714-3>
80. Li C, Wang X, Liu W, Latecki LJ, Wang B, Huang J (2019) Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Med Image Anal* 53:165–178. <https://doi.org/10.1016/j.media.2019.01.013>
81. Liang Q, Nan Y, Coppola G, Zou K, Sun W, Zhang D, Wang Y, Yu G (2019) Weakly supervised biomedical image segmentation by reiterative learning. *IEEE J Biomed Heal Inf* 23(3):1205–1214. <https://doi.org/10.1109/JBHI.2018.2850040>
82. Liao F, Liang M, Li Z, Hu X, Song S (2019) Evaluate the malignancy of pulmonary nodules using the 3-D deep leaky Noisy-OR network. *IEEE Trans Neural Netw Learn Syst* 30(11):3484–3495. <https://doi.org/10.1109/TNNLS.2019.2892409>
83. Lin H, Chen H, Graham S, Dou Q, Rajpoot N, Heng PA (2019) Fast ScanNet: fast and dense analysis of multi-Gigapixel whole-slide images for Cancer metastasis detection. *IEEE Trans Med Imaging* 38(8): 1948–1958. <https://doi.org/10.1109/TMI.2019.2891305>
84. Lijtens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42(1995):60–88. <https://doi.org/10.1016/j.media.2017.07.005>
85. Lijtens G et al (2016) Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep* 6(January):1–11. <https://doi.org/10.1038/srep26286>
86. Little WA (1974) The existence of persistent states in the brain. *Math Biosci* 19(1–2):101–120. [https://doi.org/10.1016/0025-5564\(74\)90031-5](https://doi.org/10.1016/0025-5564(74)90031-5)
87. Little WA, Shaw GL (1978) Analytic study of the memory storage capacity of a neural network. *Math Biosci* 39(3–4):281–290. [https://doi.org/10.1016/0025-5564\(78\)90058-5](https://doi.org/10.1016/0025-5564(78)90058-5)
88. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE (2017) A survey of deep neural network architectures and their applications. *Neurocomputing* 234(November 2016):11–26. <https://doi.org/10.1016/j.neucom.2016.12.038>
89. Lo SLJLMFMCMSMC, Lo SCB, Lou SLA, Chien MV, Mun SK (1995) Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Trans Med Imaging* 14(4):711–718. <https://doi.org/10.1109/42.476112>
90. Loey M, Smarandache F, Khalifa NEM (2020) Within the lack of chest COVID-19 X-ray dataset: A novel detection model based on GAN and deep transfer learning, *Symmetry (Basel)*, vol. 12, no. 4, DOI: <https://doi.org/10.3390/SYM12040651>.
91. Lopes UK, Valiati JF (2017) Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. *Comput Biol Med* 89(August):135–143. <https://doi.org/10.1016/j.compbiomed.2017.08.001>
92. Ma G, Li S, Chen C, Hao A, Qin H (2020) Stage-wise salient object detection in 360 omnidirectional image via object-level Semantical saliency ranking. *IEEE Trans Vis Comput Graph* 26:3535–3545. <https://doi.org/10.1109/tvcg.2020.3023636>
93. Ma J, Song Y, Tian X, Hua Y, Zhang R, Wu J (2020) Survey on deep learning for pulmonary medical imaging. *Front Med* 14(4):450–469. <https://doi.org/10.1007/s11684-019-0726-4>
94. Maas AL, Hannun AY, Ng AY (2013) Rectifier nonlinearities improve neural network acoustic models. In: *The 30th International Conference on Machine Learning*, vol 30
95. Masood A, Sheng B, Yang P, Li P, Li H, Kim J, Feng DD (2020) Automated decision support system for lung Cancer detection and classification via enhanced RFCN with multilayer fusion RPN. *IEEE Trans Ind Inf* 3203(c):1–1. <https://doi.org/10.1109/tii.2020.2972918> 7801
96. Mazurowski MA, Buda M, Saha A, Bashir MR (2019) Deep learning in radiology: an overview of the concepts and a survey of the state of the art with a focus on MRI. *J Magn Reson Imaging* 49(4):939–954. <https://doi.org/10.1002/jmri.26534>
97. McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys* 5:115–133. <https://doi.org/10.1007/BF02478259>
98. Minsky M, Papert S (1969) *Perceptrons: an introduction to computational geometry*, vol 522. MIT Press, Cambridge MA, pp 20–522. [https://doi.org/10.1016/S0019-9958\(70\)90409-2](https://doi.org/10.1016/S0019-9958(70)90409-2)
99. Mittal A, Hooda R, Sofat S (2018) LF-SegNet : a fully convolutional encoder – decoder network for segmenting lung fields from chest, *Wirel Pers Commun*, DOI: <https://doi.org/10.1007/s11277-018-5702-9>
100. Morris RGM, Hebb DO (1949) *The Organization of Behavior*, Wiley: New York; 1949,” *Brain Res Bull*, vol. 50, no. 5–6, p. 437, DOI: [https://doi.org/10.1016/S0361-9230\(99\)00182-3](https://doi.org/10.1016/S0361-9230(99)00182-3).

101. Münzer B, Schoeffmann K, Böszörményi L (2018) Content-based processing and analysis of endoscopic images and videos: a survey. *Multimed Tools Appl* 77(1):1323–1362. <https://doi.org/10.1007/s11042-016-4219-z>
102. Murphy A, Skalski M, Gaillard F (2018) The utilisation of convolutional neural networks in detecting pulmonary nodules: a review. *Br J Radiol* 91(1090):1–6. <https://doi.org/10.1259/bjr.20180028>
103. Murphy K *et al.* (2019) Computer aided detection of tuberculosis on chest radiographs: An evaluation of the CAD4TB v6 system, pp. 1–11, [Online]. Available: <http://arxiv.org/abs/1903.03349>.
104. Nair V, Hinton GE (2010) Rectified linear units improve restricted Boltzmann machines. *Proc 27th Int Conf Mach Learn (ICML-10)*, 807–814 33(5):807–814
105. Nakagawa K, Ishihara R, Aoyama K, Ohmori M (2019) Classification for invasion depth of esophageal squamous cell carcinoma using a deep neural network compared with experienced endoscopists. *Gastrointest Endosc* 90(3):407–414. <https://doi.org/10.1016/j.gie.2019.04.245>
106. Ng A (2011) Sparse autoencoder. *CS294A Lect. Notes* 72:1–19
107. Nie D, Trullo R, Lian J, Wang L, Petitjean C, Ruan S, Wang Q, Shen D (2018) Medical image synthesis with deep convolutional adversarial networks. *IEEE Trans Biomed Eng* 65(12):2720–2730. <https://doi.org/10.1109/TBME.2018.2814538>
108. Onishi Y *et al.* (2019) Automated Pulmonary Nodule Classification in Computed Tomography Images Using a Deep Convolutional Neural Network Trained by Generative Adversarial Networks, *Biomed Res Int*, vol. 2019, DOI: <https://doi.org/10.1155/2019/6051939>.
109. Ouyang W *et al.* (2015) DeepID-Net: Deformable deep convolutional neural networks for object detection. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit* 07–12(June):2403–2412. <https://doi.org/10.1109/CVPR.2015.7298854>
110. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 121(April):103792. <https://doi.org/10.1016/j.combiomed.2020.103792>
111. Pang S, Zhang Y, Ding M, Wang X, Xie X (2020) A deep model for lung Cancer type identification by densely connected convolutional networks and adaptive boosting. *IEEE Access* 8:4799–4805. <https://doi.org/10.1109/ACCESS.2019.2962862>
112. Pascanu R, Mikolov T, Bengio Y (2013) On the difficulty of training recurrent neural networks. *30th Int Conf Mach Learn ICML 2013(PART 3)*:2347–2355
113. Perone CS, Cohen-Adad J (2019) Promises and limitations of deep learning for medical image segmentation. *J Med Artif Intell* 2:1–1. <https://doi.org/10.21037/jmai.2019.01.01>
114. Pezeshk A, Hamidian S, Petrick N, Sahiner B (2018) 3D convolutional neural networks for automatic detection of pulmonary nodules in chest CT. *IEEE J Biomed Heal Inf PP(c)*:1. <https://doi.org/10.1109/JBHI.2018.2879449>
115. Pinckaers H, Litjens G (2019) Neural Ordinary Differential Equations for Semantic Segmentation of Individual Colon Glands, no. *NeurIPS*, [Online]. Available: <http://arxiv.org/abs/1910.10470>.
116. Poggio T, Serre T (2013) Models of visual cortex. *Scholarpedia* 8(4):3516. <https://doi.org/10.4249/scholarpedia.3516>
117. Qiang Y, Ge L, Zhao X, Zhang X, Tang X (2017) Pulmonary nodule diagnosis using dual-modal supervised autoencoder based on extreme learning machine. *Expert Syst* 34(6):1–12. <https://doi.org/10.1111/exsy.12224>
118. Qu H *et al.* (2019) Joint Segmentation and fine -grained classification of nuclei in histopathology images. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp 900–904. <https://doi.org/10.1109/ISBI.2019.8759457>
119. Rajaraman S, Antani SK (2020) Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs. *IEEE Access* 8:27318–27326. <https://doi.org/10.1109/ACCESS.2020.2971257>
120. Rajaraman S, Antani S (2020) Weakly labeled data augmentation for deep learning: a study on COVID-19 detection in chest X-rays. *Diagnostics* 10(6):1–17. <https://doi.org/10.3390/diagnostics10060358>
121. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz CP, Patel BN, Yeom KW, Shpanskaya K, Blankenberg FG, Seekins J, Amrhein TJ, Mong DA, Halabi SS, Zucker EJ, Ng AY, Lungren MP (2018) Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med* 15(11):1–17. <https://doi.org/10.1371/journal.pmed.1002686>
122. Rajpurkar P *et al.* (2017) CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning, pp. 3–9, [Online]. Available: <http://arxiv.org/abs/1711.05225>.
123. Reader AJ, Corda G, Mehranian A, da Costa-Luis C, Ellis S, Schnabel JA (2020) Deep learning for PET image reconstruction. *IEEE Trans Radiat Plasma Med Sci* 7311(1):1–1. <https://doi.org/10.1109/trpms.2020.3014786> 25
124. Ribli D, Horváth A, Unger Z, Pollner P, Csabai I (2018) Detecting and classifying lesions in mammograms with deep learning. *Sci Rep* 8(1):1–7. <https://doi.org/10.1038/s41598-018-22437-z>

125. Rodríguez-Ruiz A, Krupinski E, Mordang JJ, Schilling K, Heywang-Köbrunner SH, Sechopoulos I, Mann RM (2019) Detection of breast cancer with mammography: effect of an artificial intelligence support system. *Radiology* 290(3):1–10. <https://doi.org/10.1148/radiol.2018181371>
126. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 9351: 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
127. Rosenblatt F (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Rev* 65(6):386–408. <https://doi.org/10.1037/h0042519>
128. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323(9):533–536
129. Sabour S, Frosst N, Hinton GE (2017) Dynamic routing between capsules. *Adv Neural Inf Process Syst* 2017-Decem(Nips):3857–3867
130. Saeedan F, Weber N, Goesele M, Roth S (2018) Detail-Preserving Pooling in Deep Networks,” *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, no. June, pp. 9108–9116, DOI: <https://doi.org/10.1109/CVPR.2018.00949>.
131. Sahiner B, Heang-Ping Chan, Petrick N, Datong Wei, Helvie MA, Adler DD, Goodsitt MM (1996) Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE Trans Med Imaging* 15(5):598–610. <https://doi.org/10.1109/42.538937>
132. Sari CT, Gunduz-Demir C (2019) Unsupervised feature extraction via deep learning for Histopathological classification of Colon tissue images. *IEEE Trans Med Imaging* 38(5):1139–1149. <https://doi.org/10.1109/TMI.2018.2879369>
133. Scherer D, Müller A, Behnke S (2010) Evaluation of pooling operations in convolutional architectures for object recognition. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 6354 LNCS(PART 3):92–101. https://doi.org/10.1007/978-3-642-15825-4_10
134. Schmidhuber J (2015) Deep learning in neural networks: an overview. *Neural Netw* 61:85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
135. Serag A et al (2019) Translational AI and Deep Learning in Diagnostic Pathology. *Front Med* 6(October): 1–15. <https://doi.org/10.3389/fmed.2019.00185>
136. Setio AAA, Ciompi F, Litjens G, Gerke P, Jacobs C, van Riel SJ, Wille MMW, Naqibullah M, Sanchez CI, van Ginneken B (2016) Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. *IEEE Trans Med Imaging* 35(5):1160–1169. <https://doi.org/10.1109/TMI.2016.2536809>
137. Shah A, Kadam E, Shah H, Shinde S, Shingade S (2016) Deep residual networks with exponential linear unit. *ACM Int Conf Proceeding Ser* 21–24(Sept):59–65. <https://doi.org/10.1145/2983402.2983406>
138. Shatnawi A, Al-Bdour G, Al-Qurran R, Al-Ayyoub M (2018) A comparative study of open source deep learning frameworks. 2018 9th Int Conf Inf Commun Syst ICICS 2018 2018-Janua:72–77. <https://doi.org/10.1109/IACS.2018.8355444>
139. Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W (2019) Deep learning to improve breast Cancer detection on screening mammography. *Sci Rep* 9(1):1–13. <https://doi.org/10.1038/s41598-019-48995-4>
140. Shickel B, Tighe PJ, Bihorac A, Rashidi P (2017) Deep EHR : A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record, vol. 2194, no. c, pp. 1–17, DOI: <https://doi.org/10.1109/JBHI.2017.2767063>.
141. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging* 35(5):1285–1298. <https://doi.org/10.1109/TMI.2016.2528162>
142. Siegel RL, Miller KD, Jemal A (2019) Cancer statistics, 2019. *CA Cancer J Clin* 69(1):7–34. <https://doi.org/10.3322/caac.21551>
143. Simonyan K, Zisserman (2015) A Very deep convolutional networks for large-scale image recognition. In: 3rd International Conference on Learning Representations, ICLR 2015, pp 1–14
144. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E (2019) Convolutional neural networks for radiologic images: a Radiologist’s guide. *Radiology* 290(3):590–606. <https://doi.org/10.1148/radiol.2018180547>
145. Soffer S, Klang E, Shimon O, Nachmias N, Eliakim R (2020) Deep learning for wireless capsule endoscopy : a systematic review and meta-analysis. *Gastrointest Endosc* 92(4):831–839.e8. <https://doi.org/10.1016/j.gie.2020.04.039>
146. Song TH, Sanchez V, Eidaly H, Rajpoot NM (2019) Simultaneous cell detection and classification in bone marrow histology images. *IEEE J Biomed Heal Inf* 23(4):1469–1476. <https://doi.org/10.1109/JBHI.2018.2878945>

147. Song Y, Tan EL, Jiang X, Cheng JZ, Ni D, Chen S, Lei B, Wang T (2017) Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE Trans Med Imaging* 36(1):288–300. <https://doi.org/10.1109/TMI.2016.2606380>
148. Souza JC, Bandeira Diniz JO, Ferreira JL, França da Silva GL, Corrêa Silva A, de Paiva AC (2019) An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. *Comput Methods Prog Biomed* 177:285–296. <https://doi.org/10.1016/j.cmpb.2019.06.005>
149. Sun M, Zhang G, Dang H, Qi X, Zhou X, Chang Q (2019) Accurate gastric Cancer segmentation in digital pathology images using deformable convolution and multi-scale embedding networks. *IEEE Access* 7: 75530–75541. <https://doi.org/10.1109/ACCESS.2019.2918800>
150. Swersky K, Chen B, Marlin B, de Freitas N (2010) A tutorial on stochastic approximation algorithms for training restricted Boltzmann machines and deep belief nets,” 2010 Inf Theory Appl Work ITA 2010, Conf Proc, pp. 80–89, DOI: <https://doi.org/10.1109/ITA.2010.5454138>.
151. Szegedy C, Reed S, Sermanet P, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: The IEEE conference on Computer Vision and Pattern Recognition (CVPR), pp 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>
152. Tabibu S, Vinod PK, Jawahar CV (2019) Pan-renal cell carcinoma classification and survival prediction from histopathology images using deep learning. *Sci Rep* 9(1):1–9. <https://doi.org/10.1038/s41598-019-46718-3>
153. Theano Development Team *et al.* (2016) Theano: A Python framework for fast computation of mathematical expressions, pp. 1–19. [Online]. Available: <http://arxiv.org/abs/1605.02688>.
154. Valkonen M, Isola J, Ylilinen O, Muhonen V, Saxlin A, Tolonen T, Nykter M, Ruusuvaara P (2020) Cytokeratin-supervised deep learning for automatic recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67. *IEEE Trans Med Imaging* 39(2):534–542. <https://doi.org/10.1109/TMI.2019.2933656>
155. Valliani AAA, Ranti D, Oermann EK (2019) Deep learning and neurology: a systematic review. *Neurol Ther* 8(2):351–365. <https://doi.org/10.1007/s40120-019-00153-8>
156. Van Eycke YR, Balsat C, Verset L, Debeir O, Salmon I, Decaestecker C (2018) Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise IHC biomarker quantification: a deep learning approach. *Med Image Anal* 49:35–45. <https://doi.org/10.1016/j.media.2018.07.004>
157. van Ginneken B, Setio AAA, Jacobs C, Ciompi F (2015) Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. In: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp 286–289. <https://doi.org/10.1109/ISBI.2015.7163869>
158. Vedaldi A, Lenc K (2015) MatConvNet: Convolutional neural networks for MATLAB, MM 2015 – Proc 2015 ACM Multimedia Conf, pp. 689–692, DOI: <https://doi.org/10.1145/2733373.2807412>.
159. Vincent P, Larochelle H, Larochelle H, Bengio Y, Manzagol PA (2010) Stacked denoising autoencoders: learning useful representations in a deep network with a local Denoising criterion. *J Mach Learn Res* 11: 3371–3408
160. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR (2020) CovidGAN: data augmentation using auxiliary classifier GAN for improved Covid-19 detection. *IEEE Access* 8:91916–91923. <https://doi.org/10.1109/ACCESS.2020.2994762>
161. Wang J, Ding H, Bidgoli FA, Zhou B, Iribarren C, Molloy S, Baldi P (2017) Detecting cardiovascular disease from mammograms with deep learning. *IEEE Trans Med Imaging* 36(5):1172–1181. <https://doi.org/10.1109/TMI.2017.2655486>
162. Wang SH, Muhammad K, Hong J, Sangaiah AK, Zhang YD (2020) Alcoholism identification via convolutional neural network based on parametric ReLU, dropout, and batch normalization. *Neural Comput & Applic* 32(3):665–680. <https://doi.org/10.1007/s00521-018-3924-0>
163. Wang H, Raj B (2017) On the Origin of Deep Learning,” pp. 1–72, [Online]. Available: <http://arxiv.org/abs/1702.07800>.
164. Wang S, Tang C, Sun J, Zhang Y (2019) Cerebral micro-bleeding detection based on densely connected neural network. *Front Neurosci* 13(MAY):1–11. <https://doi.org/10.3389/fnins.2019.00422>
165. Wang L, Wong A (2020) COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images, pp. 1–12, [Online]. Available: <http://arxiv.org/abs/2003.09871>.
166. Wang Y, Yan F, Lu X, Zheng G, Zhang X, Wang C, Zhou K, Zhang Y, Li H, Zhao Q, Zhu H, Chen F, Gao C, Qing Z, Ye J, Li A, Xin X, Li D, Wang H, Yu H, Cao L, Zhao C, Deng R, Tan L, Chen Y, Yuan L, Zhou Z, Yang W, Shao M, Dou X, Zhou N, Zhou F, Zhu Y, Lu G, Zhang B (2019) IILS: intelligent imaging layout system for automatic imaging report standardization and intra-interdisciplinary clinical workflow optimization. *EBioMedicine* 44:162–181. <https://doi.org/10.1016/j.ebiom.2019.05.040>
167. Wang X *et al* (2019) Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis. *IEEE Trans Cybern PP*:1–13. <https://doi.org/10.1109/tcyb.2019.2935141>

168. Wang T et al (2020) Machine learning in quantitative PET: A review of attenuation correction and low-count image reconstruction methods. *Phys Medica* 76(March):294–306. <https://doi.org/10.1016/j.ejmp.2020.07.028>
169. Wei JW, Tafe LJ, Linnik YA, Vaickus LJ, Tomita N, Hassanpour S (2019) Pathologist-level classification of histologic patterns on resected lung adenocarcinoma slides with deep neural networks. *Sci Rep* 9(1):1–8. <https://doi.org/10.1038/s41598-019-40041-7>
170. Werbos PJ (1990) Backpropagation through time: what it does and how to do it. *Proc IEEE* 78(10):1550–1560. <https://doi.org/10.1109/5.58337>
171. Werboose J (1974) Beyond regression: new tools for prediction and analysis in the behavioral
172. Widrow B, Hoff ME (1962) Associative Storage and Retrieval of Digital Information in Networks of Adaptive ‘Neurons. *Biol Prototypes Synth Syst*:160–160. https://doi.org/10.1007/978-1-4684-1716-6_25
173. Williams RJ, David Z (1995) Gradient-based learning algorithms for recurrent networks and their computational complexity. In: *Back-propagation: theory, architectures and applications*. L. Erlbaum Associates Inc, pp 433–486
174. Wu J (2017) Convolutional Neural Networks. *Med Imaging Inf Sci* 34(2):109–111. <https://doi.org/10.11318/mii.34.109>
175. Wu H, Gu X (2015) Max-pooling dropout for regularization of convolutional neural networks. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* 9489:46–54. https://doi.org/10.1007/978-3-319-26532-2_6
176. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, Jastrzebski S, Fevry T, Katsnelson J, Kim E, Wolfson S, Parikh U, Gaddam S, Lin LLY, Ho K, Weinstein JD, Reig B, Gao Y, Toth H, Pysarenko K, Lewin A, Lee J, Airola K, Mema E, Chung S, Hwang E, Samreen N, Kim SG, Heacock L, Moy L, Cho K, Geras KJ (2019) Deep neural networks improve radiologists’ performance in breast Cancer screening. *IEEE Trans Med Imaging* 39:1–1. <https://doi.org/10.1109/tmi.2019.2945514> 1194
177. Wu N et al (2018) Breast density classification with deep convolutional neural networks. *ICASSP, IEEE Int Conf Acoust Speech Signal Process - Proc 2018-April*:6682–6686. <https://doi.org/10.1109/ICASSP.2018.8462671>
178. Xing F, Xie Y, Su H, Liu F, Yang L (2018) Deep learning in microscopy image analysis: a survey. *IEEE Trans Neural Netw Learn Syst* 29(10):4550–4568. <https://doi.org/10.1109/TNNLS.2017.2766168>
179. Xing F, Xie Y, Yang L (2016) An automatic learning-based framework for robust nucleus segmentation. *IEEE Trans Med Imaging* 35(2):550–566. <https://doi.org/10.1109/TMI.2015.2481436>
180. Xu B, Wang N, Chen T, Li M (2015) Empirical Evaluation of Rectified Activations in Convolutional Network . [Online]. Available: <http://arxiv.org/abs/1505.00853>.
181. Xu S, Wu H, Bie R (2019) CXNet-m1: anomaly detection on chest X-rays with image-based deep learning. *IEEE Access* 7(c):4466–4477. <https://doi.org/10.1109/ACCESS.2018.2885997>
182. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J, Madabhushi A (2016) Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Trans Med Imaging* 35(1):119–130. <https://doi.org/10.1109/TMI.2015.2458702>
183. Yi X, Walia E, Babyn P (2019) Generative adversarial network in medical imaging: A review,” *Med Image Anal*, vol. 58, DOI: <https://doi.org/10.1016/j.media.2019.101552>.
184. Yi F, Yang L, Wang S, Guo L, Huang C, Xie Y, Xiao G (2018) Microvessel prediction in H&E Stained Pathology Images using fully convolutional neural networks. *BMC Bioinform* 19(1):1–9. <https://doi.org/10.1186/s12859-018-2055-z>
185. Yoon HJ et al (2019) A Lesion-Based Convolutional Neural Network Improves Endoscopic Detection and Depth Prediction of Early Gastric Cancer. *J Clin Med* 8(9):1310. <https://doi.org/10.3390/jcm8091310>
186. Yu D et al (2014) An Introduction to Computational Networks and the Computational Network Toolkit. *INTERSPEECH, Microsoft Research*
187. Zhang S, Zhang S, Wang B, Habetler TG (2020) Deep learning algorithms for bearing fault diagnostics - a comprehensive review. *IEEE Access* 8:29857–29881. <https://doi.org/10.1109/ACCESS.2020.2972859>
188. Zhang X et al (2017) Whole mammogram image classification with convolutional neural networks. *Proc - 2017 IEEE Int Conf Bioinforma Biomed BIBM 2017 2017-Janua(Cc)*:700–704. <https://doi.org/10.1109/BIBM.2017.8217738>
189. Zhao Q, Lyu S, Zhang B, Feng W (2018) Multiactivation pooling method in convolutional neural networks for image recognition. *Wirel Commun Mob Comput* 2018:1–16. <https://doi.org/10.1155/2018/8196906>
190. Zhao W, Zeng Z (2019) Multi Scale Supervised 3D U-Net for Kidney and Tumor Segmentation., DOI: <https://doi.org/10.24926/548719.007>.