



REVIEW

Deep Learning-Based Cancer Detection-Recent Developments, Trend and Challenges

Gulshan Kumar^{1,*} and Hamed Alqahtani²

¹Shaheed Bhagat Singh State University, Ferozepur, India

²King Khalid University, Abha, Saudi Arabia

*Corresponding Author: Gulshan Kumar. Email: gulshanahuja@gmail.com

Received: 23 July 2021 Accepted: 26 September 2021

ABSTRACT

Cancer is one of the most critical diseases that has caused several deaths in today's world. In most cases, doctors and practitioners are only able to diagnose cancer in its later stages. In the later stages, planning cancer treatment and increasing the patient's survival rate becomes a very challenging task. Therefore, it becomes the need of the hour to detect cancer in the early stages for appropriate treatment and surgery planning. Analysis and interpretation of medical images such as MRI and CT scans help doctors and practitioners diagnose many diseases, including cancer disease. However, manual interpretation of medical images is costly, time-consuming and biased. Nowadays, deep learning, a subset of artificial intelligence, is gaining increasing attention from practitioners in automatically analysing and interpreting medical images without their intervention. Deep learning methods have reported extraordinary results in different fields due to their ability to automatically extract intrinsic features from images without any dependence on manually extracted features. This study provides a comprehensive review of deep learning methods in cancer detection and diagnosis, mainly focusing on breast cancer, brain cancer, skin cancer, and prostate cancer. This study describes various deep learning models and steps for applying deep learning models in detecting cancer. Recent developments in cancer detection based on deep learning methods have been critically analysed and summarised to identify critical challenges in applying them for detecting cancer accurately in the early stages. Based on the identified challenges, we provide a few promising future research directions for fellow researchers in the field. The outcome of this study provides many clues for developing practical and accurate cancer detection systems for its early diagnosis and treatment planning.

KEYWORDS

Autoencoders (AEs); cancer detection; convolutional neural networks (CNNs); deep learning; generative adversarial models (GANs); machine learning

1 Introduction

Cancer is a disease that affects all kinds of life of patients, including business life, family life and social life. But, diagnosis and treatment of cancer is a long and challenging task in



comparison to other diseases [1]. Cancer is caused due to the formation of bad neoplasms along with normal division and reproduction of the cells in various organs and tissues [2].

Nowadays, cancer has been placed at the second position in diseases resulting in death. Cancer has many forms like skin cancer, breast cancer, leukaemia, prostate cancer, brain cancer etc. In 2018, approximately 9.6 million people suffered and died from cancer [3]. Figs. 1 and 2 present the top five most dangerous cancer spread in the US (men and women) and the number of deaths in the year 2019 [4].

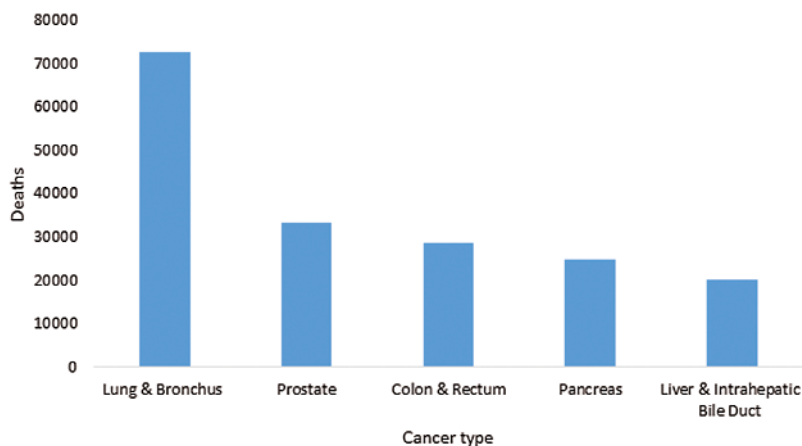


Figure 1: Top five most dangerous cancer types and deaths in men (in US) [4]

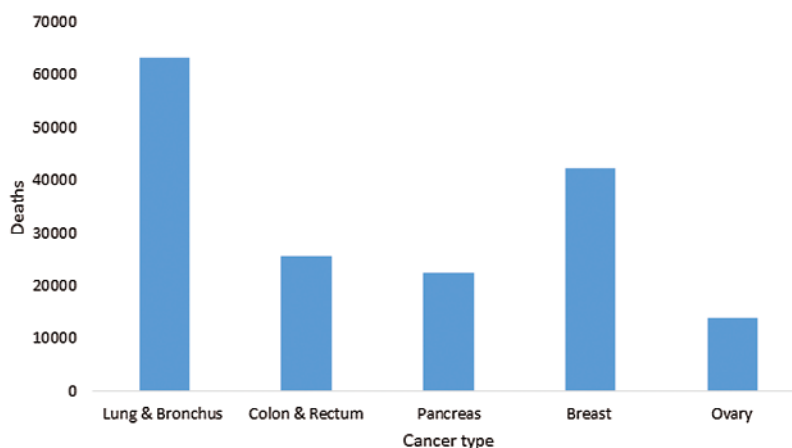


Figure 2: Top five most dangerous cancer types and deaths in women (in US) [4]

By 2030, it has been predicted that 13.1 million people will die due to cancer [5]. It has been analyzed from the cause of deaths that one in every ten persons (including both male and female) suffers from cancer disease. The primary cause of the increased number of deaths due to cancer is its diagnosis in later stages, which is challenging to treat. It has been observed that most of the patients suffering from cancer disease get well treated if diagnosed in the early stages [6].

Therefore, early diagnosis of cancer disease can help proper treatment and improve long term survival of the cancer patients [7].

Several medical imaging methods have been adopted for the early detection and diagnosis of cancer disease. These images can be used to early detect, monitor the situation and follow up during the treatment of the cancer patients [8]. Interpreting a large number of medical images by doctors is a challenging and time-consuming task. Manual interpretation of medical images can be biased and error-prone. To automate interpreting medical images for early detection of cancer, the computer-aided diagnosis has been started since 1980. Computer-aided diagnosis helps the doctor in detecting cancer disease in early stages by integrating medical images accurately and efficiently [9].

Numerous methods have been developed in recent years concerning the computer-aided diagnosis of cancer disease [2]. Artificial intelligence (AI) based systems that can imitate human intelligence are gaining increasing attention of doctors and practitioners for detecting cancer disease. AI-based systems, particularly deep learning systems, are learning systems widely used to solve real-life problems.

Deep learning systems help doctors and practitioners to detect cancer disease in early stages from medical images. Such systems require training of the machine learning methods for detecting similar patterns in new medical images. Deep learning systems can be divided into four categories based upon their learning strategy, supervised learning, semi-supervised learning, unsupervised learning and reinforced learning. The supervised learning method requires labelled data for training the machine learning algorithm. The trained model is further used to predict the label of unknown samples. The unsupervised learning method needs data without labels. It creates an implicit model based upon the training data to categorize unknown samples. The semi-supervised learning method uses labelled and unlabeled data for training the algorithm. The reinforced learning method learns from the environment through feedback. They improve their performance with experience. Deep learning systems have been successfully applied in many domains such as cybersecurity, industry production and diagnosing different diseases [10].

This paper contributes in the following ways:

- Describes the use of deep learning methods for detecting and diagnosing cancer disease.
- Explains and classifies the most popular deep learning methods for detecting cancer.
- Summarizes the recent developments in deep learning for detecting cancer disease.
- Identifies the challenges in applying deep learning for cancer detection.
- Provides many clues for the fellow researchers to conduct future research in this field.

The remaining part of this paper is organized as follows, [Section 2](#) highlights various phases being followed in detecting cancer using deep learning methods. [Section 3](#) classifies the existing deep learning architectures, followed by their details in [Section 4](#). [Section 5](#) provides a comprehensive review of deep learning methods for detecting cancer, particularly for breast cancer, brain cancer, skin cancer and prostate cancer. [Section 6](#) summarizes and discusses the key findings of this paper and highlights critical challenges in using deep learning methods for quick and accurate diagnosis of cancer disease. Finally, [Section 7](#) concludes the paper at the end.

2 Cancer Detection Process

Recent advancements in medical and information technologies have increased the success rate of disease diagnostic systems and hence medical treatment. Medical images such as MRI, CT and ultrasound play a crucial role in detecting many diseases in their early stages. The medical

images are very significant in detecting cancer in its early stage and plan the treatment accordingly. Detection of cancer from medical images using deep learning systems involve three phases, namely, pre-processing, segmentation and post-processing for better analysis of images to detect cancer disease [11]. Each phase in detecting cancer based on medical images involves applying different methods before training the deep-learning algorithm. The details are provided in the following subsections.

2.1 Pre-Processing Phase

Pre-processing is the first phase in the pipeline of detecting cancer based on medical images. Raw medical images may contain some noise. The presence of noise may affect the accuracy of analyzing the medical image for detecting cancer. Pre-processing phase removes the noise from raw medical images and improves the quality of the image. Pre-processing phase enables low contrast images between skin lesions, normal skin, irregular border, and other skin artifacts like skin lines, hairs, and black frames.

Several methods have been proposed for removing the different types of noise like Gaussian noise, Poisson noise, salt and pepper noise, and speckle noise. The most common noise removing methods include adaptive median filter, adaptive wiener filter, Gaussian filter, mean filter, and median filter.

The noise in the medical images is removed or adjusted during the pre-processing functions like colour correction, contrast adjustment, hair removal, image smoothing, localization, normalization, and vignetting effect removal. The appropriate combination of these pre-processing functions can result in more accurate detection of cancer using medical images.

Pre-processing methods include automatic colour equalization, black frame removal techniques, colour space transform, contrast enhancement, dull Razor, Gaussian filter, hair removal technique, Karhunen-Loe's transform [12], non-skin masking, and pseudo-random filter.

The pre-processing method helps to analyze the medical images for accurate detection of cancer disease. For instance, for detecting brain cancer, MRI images are transformed to grayscale images followed by contrast adjustment using smoothing function [13]. Sometimes, skull stripping is also helpful for extracting the brain tissues from the rest of the skull [14]. Lung cancer diagnosis involves transforming the CT scans into grayscale images, normalizing the image, and reducing the noise in images. The images may be converted into binary images for removing unwanted portions.

2.2 Image Segmentation Phase

Image segmentation is an essential phase in the pipeline of analyzing medical images. It involves dividing the medical image into different portions or regions of interest for extracting necessary information. This step differentiates the background of the image and the region of interest. Methods for segmenting the medical images can be divided into four categories described below:

- Model-based segmentation
- Pixel-based segmentation
- Region-based segmentation
- Threshold-based segmentation

Different methods have been developed for each category. For example, threshold-based segmentation methods include histogram-based threshold methods, local and global threshold

methods, maximum entropy, and Ostu's methods. Region-based methods include seeded region growing method and Watershed segmentation method. Pixel-based segmentation methods apply artificial neural networks, fuzzy c-means clustering, and Markov field method etc. Finally, model-based segmentation includes the employment of a parametric deformable model.

Several methods have been developed in different categories of image segmentation. The most well-known methods include active contours, adaptive thresholding, bootstrap learning [15], clustering and statistical region growing [16], contextual hypergraph [17], cooperative neural network segmentation, distributed and localized region identification, edge detection, fuzzy C-Means clustering, gradient flow vector, histogram thresholding, principal component transform, probabilistic modelling, region fused band and narrow band graph partition, sparse coding, and supervised learning. Multiple segmentation methods can be integrated to design hybrid segmentation methods to improve the accuracy of analysis results.

2.3 Post-Processing Phase

The post-processing phase in the pipeline of analyzing medical images for detecting cancer disease consists of grabbing relevant features after applying pre-processing and image segmentation steps. To accomplish the task of grabbing features, several methods have been proposed. The most common methods include border expansion, island removal, opening and closing operations, region merging, and smoothing. After applying the post-processing methods, the features are extracted from the selected region of the image to further analyze for detecting the disease. The common feature extraction methods include decision boundary features, Fourier power spectrum (FPS), Gaussian derivative kernels [18], grey level co-occurrence matrix (GLCM), principal component analysis (PCA), wavelet Packet Transform (WPT) [19]. Fig. 3 summarizes the phases in cancer detection with their methods.

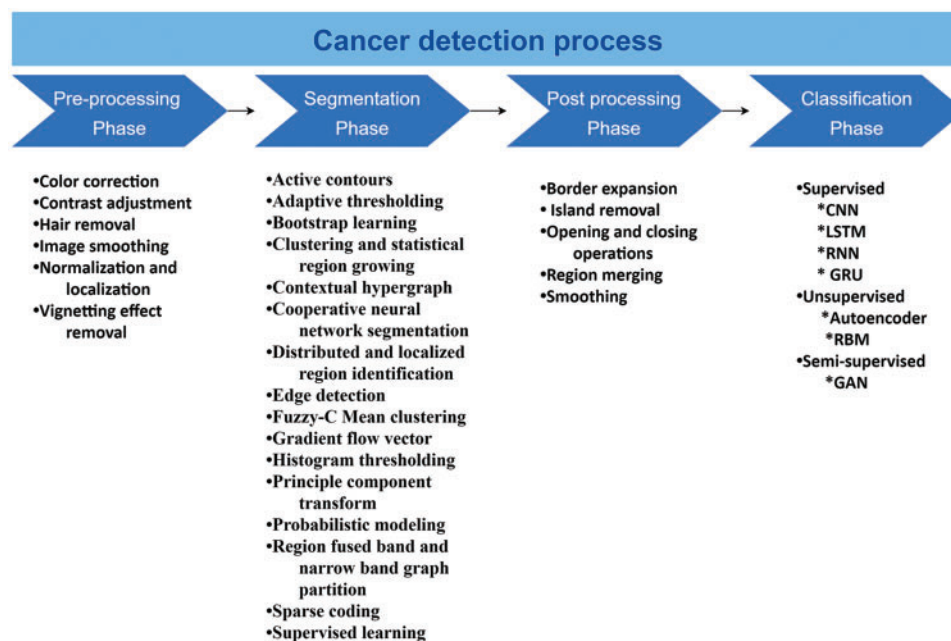


Figure 3: Summary of cancer detection phases and respective methods

2.4 Classification Phase

The classification phase involves the application of deep learning methods for classifying medical images into respective cancer categories based on the features extracted by the post-processing phase as described in [Sub-Section 2.3](#). Different types of deep learning methods have been developed based on different principles. Broadly, deep learning methods can be divided into four categories described in [Section 3](#). The classification phase involves training and testing the deep learning classifier for detecting cancer types based on the extracted features from images. The first phase trains a deep learning classifier using training dataset images. After appropriate training iterations, the trained deep learning model predicts the cancer type of unknown images based on the extracted features. The most commonly used deep learning methods for cancer detection are convolution neural network (CNN), long short term memory (LSTM), recurrent neural network (RNN), and gated recurrent units (GRU).

3 Classification of Deep Learning Models

Correct acquisition of medical images and their interpretation plays a significant role in accurately detecting and diagnosing cancer disease. There exist many image capturing devices with high-resolution such as CT scan, MRI scan and X-ray scans. After their pre-processing phase, the disease detection system involves extracting relevant features from these medical images and training the models from the extracted features. The trained model is further used to detect the disease from respective unknown medical images.

Due to considerable variation in medical images of different patients, the conventional machine learning method cannot provide authentic and accurate results. In recent years, deep learning methods have been successfully employed in different fields, specifically in analyzing medical images. These methods are beneficial and efficient in analyzing medical images to detect diseases, particularly cancer disease.

Deep learning methods are a subset of machine learning methods that allows approximating the outcome using the given data set and train the model as per the result. Deep learning methods involve neural networks with multiple layers of neurons like the input layer, multiple hidden layers, and output layer [20]. Due to the presence of multiple layers, the deep learning model gets trained more accurately. The deep learning models can be categorized based upon their learning strategies into four classes, namely, supervised learning, semi-supervised learning, unsupervised learning and reinforced learning models, presented in [Fig. 4](#).

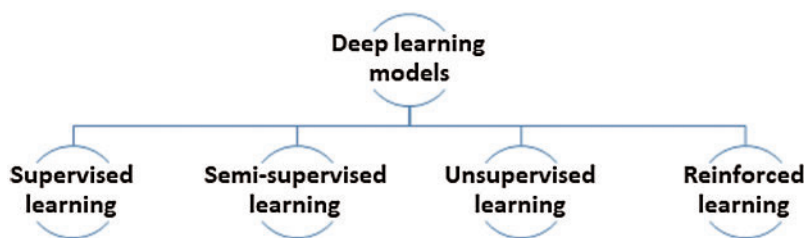


Figure 4: Classification of deep learning models

3.1 Supervised Deep Learning Models

This class of deep learning models require predefined labelled data for its training. During the training phase, the deep learning model needs to train with all possible input combinations

with the target class label. After training, the trained model is utilized to predict the label of unknown samples. The most commonly used deep learning models in the supervised learning category include CNN, LSTM, RNN, and GRU [21].

3.2 Un-Supervised Deep Learning Models

Supervised deep learning models do not require any labelled training data. These models analyze the intrinsic features of the data based on some relevant features to group similar data. These models are generally used for clustering and feature reduction purposes. The most commonly used unsupervised deep learning models include Auto Encoders (AE), Restricted Boltzmann Machines (RBM).

3.3 Semi-Supervised Deep Learning Models

A semi-supervised deep learning model partly uses labelled data and unlabeled data for its training. The most commonly used deep learning models in this category include RNN, LSTM, GRU and Generative Adversarial Networks (GAN).

3.4 Reinforced Deep Learning Models

Reinforced deep learning models work on taking appropriate actions to increase the rewards in a given environment. These models get trained for identification of actual behaviour by interacting with the environment [22].

4 Deep Learning Models

Several variants of deep learning architecture have been proposed in recent years. The most prominent deep learning models are described below.

4.1 Convolutional Neural Networks (CNN)

The convolutional neural network is a feed-forward neural network that is capable of processing input without any iteration cycles as presented in Fig. 5. each layer in the CNN has a different function.

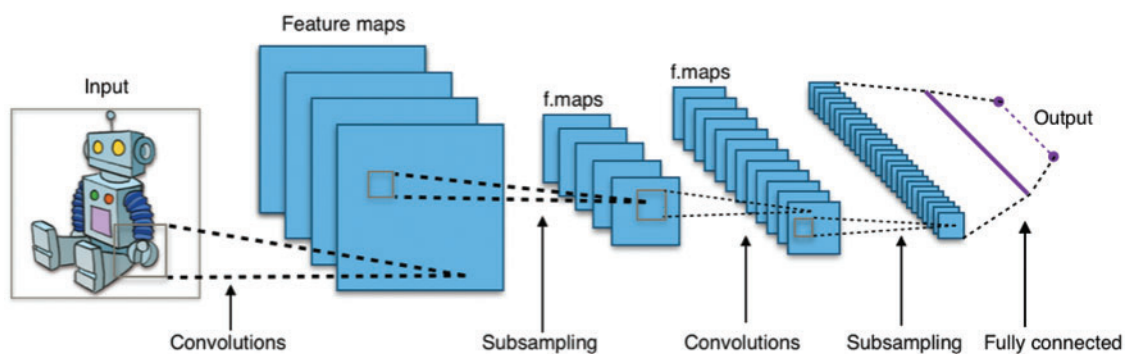


Figure 5: CNN architecture

Convolutional layer can be mathematically represented by Eq. (1).

$$G(X) = g_N(g_{N-1}(\dots(g_1(X)))) \quad (1)$$

Here, N gives hidden layer count, X represents input vector, and g represents a function to the layer N . A CNN model consists of a convolutional layer having a function g_N of multiple convolutional kernels ($h_1 \dots h_k$) denoting a linear function in k_{th} kernel as shown in Eq. (2).

$$h_k(x, y) = \sum_{s=-m}^m \sum_{t=-n}^n \sum_{v=-d}^d V_k(s, t, v) X(x-s, y-t, z-v) \quad (2)$$

Here, (x, y, z) defines the pixel position of input X , m gives height, n signifies width, and w shows the depth of the filter. V_k presents the weight of k_{th} kernel.

The sub-sampling layer of CNN is also called the pooling layer. This layer summarises surrounding pixels and computes the outcome at a given position with summarised features. This layer helps to reduce the features of the data. It also exhibits the invariance of translational and rotational transformations. Several methods have been proposed for the pooling layer [23], including max pooling, average pooling, etc.

Activation functions are used in activating the outcome of the layer based on a given input. Many activation functions have been proposed in the literature. The most commonly used activation functions include sigmoid function [24], tanh function, rectified linear unit (ReLU) function [25]. CNN models have been widely used for medical image analysis to detect the different diseases in different organs of the body.

4.2 Fully Convolutional Networks (FCNs)

Fully Convolutional Networks (FCNs) is designed using locally connected layers, such as convolution, pooling and upsampling layers as presented in Fig. 6 [26]. There is no dense layer used in this kind of architecture. It enables a reduction in the number of parameters and computation time. This network can work regardless of the original image size, without requiring any fixed number of units at any stage, given that all connections are local [27].

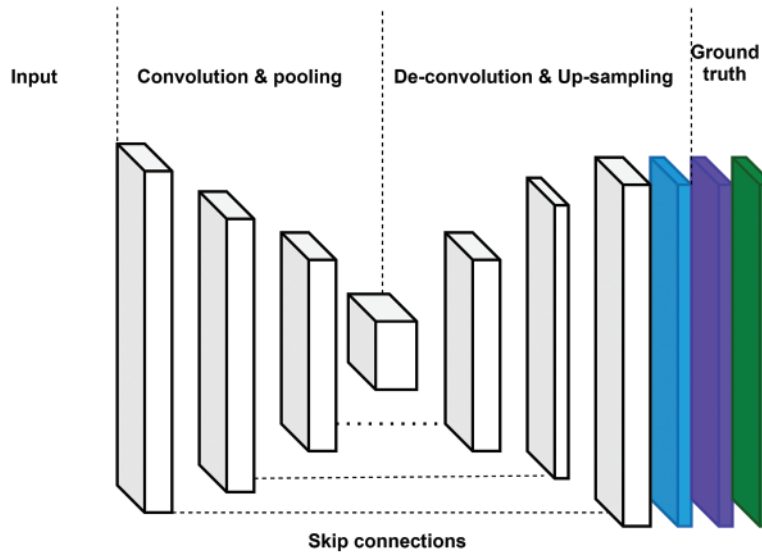


Figure 6: FCN architecture

4.3 U-Net Fully Convolutional Neural Network

This neural network has been developed for segmenting biomedical images. This network consists of two different parts known as in encoder and decoder. The encoder part extracts the context of the image using a sequence of convolutional and pooling layers. In contrast, the decoder part is responsible for the symmetric expanding path using transport convolutional and precise localization. This enables the network to take the input image of any size.

4.4 Generative Adversarial Networks (GANs)

A generative adversarial network was proposed by Goodfellow et al. [28] based on a two-player min-max game. It consists of two players. The first player is the generator and the second player is the discriminator. The generative network G transforms prior distribution p_z of the random noise $z \sim p_z$ to real looking images $G(z) \sim p_{fake}$ [29]. Discriminator network D classifies the fake image generated by G using a real training data set distribution $x \sim p_{real}$.

The argument of G is optimized as per the feedback received from D network, intending to make D network fool in the classification job. D generates better and more realistic fake images, whereas G trains itself to generate the actual images. GANs can map the random to a real distribution. It has been used in different domains, including detection, domain adaptation, reconstruction, and segmentation.

4.5 Recurrent Neural Networks (RNNs)

A recurrent neural network is an effective neural network for processing sequential or time-series based data [30]. This network computes a hidden vector sequence from the input values and generates a corresponding output sequence. In this network, the next state is a function of previous hidden States. This network is widely used for processing sequential data. RNNs learn from training data similar to the traditional feed-forward NNs. But, they have an additional memory element as presented in Fig. 7. The memory element considers the prior input that may impact the current input and output of RNN [31]. In contrast, feed-forward NNs assume independence between input and output. In contrast, the output of RNN depends upon current input and previous information in sequence.

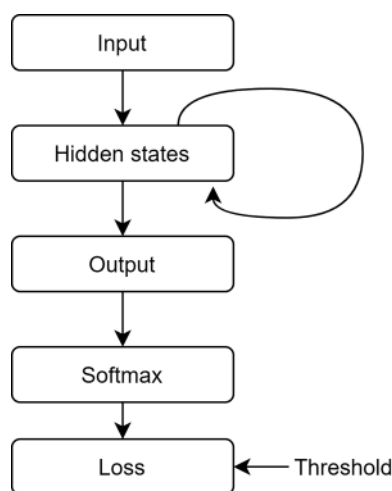


Figure 7: RNN architecture

RNNs can map different inputs to different outputs, unlike traditional feed-forward networks that map one input to one output. Input and output of RNN can vary in their length. According to the input and output length, different RNN types have been developed and used for different use cases, such as sentiment analysis, music generation, and machine translation. RNN can be one-to-one, one-to-many, many-to-one, and many-to-many types.

RNNs can use different types of activation functions like sigmoid, softmax, tanh and Relu expressed in Eqs. (3)–(6), respectively.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

$$f(x) = \frac{e^x}{\sum_{i=1}^n e^{x_i}} \quad (4)$$

$$f(x) = \frac{e^{-x} - e^{-x}}{e^{-x} + e^{-x}} \quad (5)$$

$$f(x) = \max(0, x) \quad (6)$$

4.6 Long Short-Term Memory (LSTM)

LSTM is a variant of the recurrent neural network [32]. LSTM addresses the problem of long-term dependencies. The long-term dependency occurs if the previous state influencing the current output is not in the recent past. This issue was not addressed by RNN accurately in predicting the current state. In order to address the long-term dependency issue, LSTMs have been designed with “cells” in the hidden layers of their architecture, consisting of three gates—an input gate, an output gate, and a forget gate [33]. Information flow is controlled using these gates for predicting the outcome of NN.

LSTM works in different steps. The first step decides the information be thrown from the cell state. This is decided by forget gate or sigmoid layer based on input x_t and cell state h_{t-1} . It generates number between 0 and 1 for each cell state h_{t-1} . 1 indicates to keep the value, and 0 indicate to throw the value. The second step decides the information for storing in cell state. It involves two parts: the input gate or sigmoid layer for updated values and the tanh layer, creating a new vector of values that can be added to the state. These two vectors are combined to update the state in the next step. Accordingly, old cell state h_{t-1} is updated to h_t considering forgettable information. The outcome of LSTM is based on filtered cell state, after passing through sigmoid layer, tanh layer, and multiplication with sigmoid layer as depicted in Fig. 8.

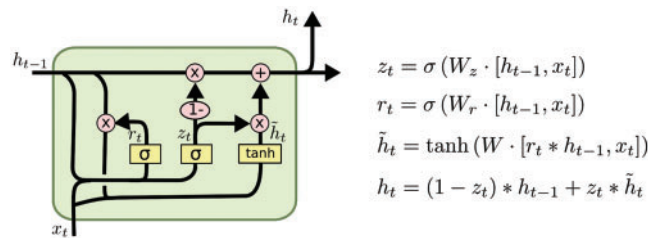


Figure 8: LSTM architecture

4.7 Restricted Boltzmann Machine (RBM)

RBM has been designed as a two-layer neural network [34,35]. It has generating capability. It can learn the probability distribution of input data. It has been widely used for classification, regression, feature reduction, collaborative filtering and modelling.

RBM is a particular case of the Boltzmann machine with restrictions on the connection between visible and hidden nodes as depicted in Fig. 9. It enables easy implementation of RBM relative to the Boltzmann machine. The two layers, visible and hidden layers, are connected by a fully bipartite graph. That implies each node of the visible layer is connected with each node of the hidden layer. But, there is no connection between nodes of the same layer. This restriction of no interconnection of nodes in the same layer enables an efficient training method in contrast to the Boltzmann machine. RBM is a Stochastic NN, where each neuron exhibits random behaviour upon activation. It differentiates RBM from autoencoder. It also consists layer of bias units for hidden and visible bias. The hidden bias of RBM produces activation on the forward pass, whereas visible bias enables RBM to reconstruct the input during backward pass. The reconstructed input is found different from the actual input as there is no connection between visible nodes; there is no way of transferring information among them.

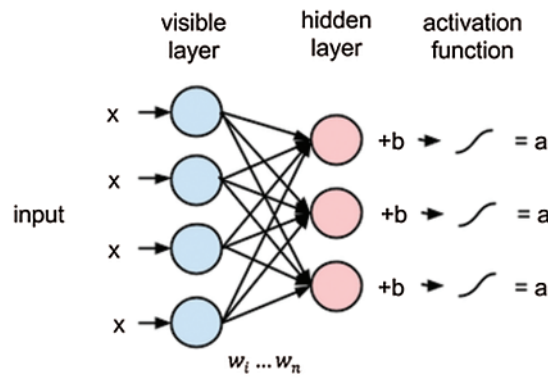


Figure 9: RBM architecture

4.8 Autoencoder (AE)

AE is unsupervised learning based deep neural network. This network learns input data to low dimensional feature space, as presented in Fig. 10.

The network consists of the input layer, hidden layer and output layer. The training process of AE consists of two phases, namely, encoding and decoding. In the first phase, the input is encoded as represented by Eq. (7).

$$J = \sigma(Y_{I,J} + B_{I,J}) \quad (7)$$

Here, σ represents an activation function. This representation J is decoded using a new weight matrix, as shown in Eq. (8).

$$\hat{I} = \hat{\sigma}(Y_{J,\hat{I}} + B_{J,\hat{I}}) \quad (8)$$

Here, $\hat{\sigma}$ represents an activation function. $Y_{J,\hat{I}}$ can be considered as the transpose of $Y_{I,J}$ or a new learnable matrix. This way, the autoencoder gets trained by minimizing the error. AE has many variants described in the following subsections.

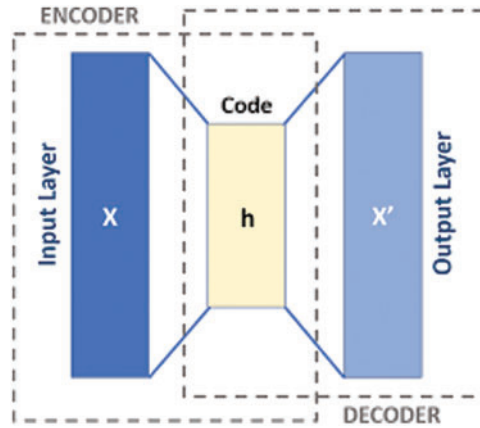


Figure 10: AE architecture

4.8.1 Stacked Autoencoders

The basic architecture of stacked AE consists of stacking of AEs in their hidden layers. The hidden layers are trained using an unsupervised learning strategy. It is followed by the fine-tuning phase using a supervised learning strategy. Stacked AE consists of three phases. The first atomic order gets trained based on input training data to create a feature vector in the first phase. In the second phase, that generated feature vector is passed to the next layer. This process is iterated to the last hidden layer. In the third phase, the back-propagation method is adopted to minimize the cost function after training the hidden layers. Accordingly, weights are updated to get an acceptable trained model.

4.8.2 Sparse Autoencoders (SAE)

SAE is a variant of autoencoder with sparsity involved in the hidden layers. In this case, the number of neurons in hidden layers is greater than that of the input layer. This network is trained using a greedy strategy when it is connected to the encoding network only. During the training process, the first hidden layer is trained as a stacked autoencoder separately. The output of this layer is fed to the next layer for training. In this network, features are captured using low level stacked autoencoder and later extracted features are passed to the higher levels of autoencoder for extracting deep features. In this way, sparse stacked autoencoder can extract 3D pictures from the raw training data.

4.8.3 Convolutional Autoencoders (CAE)

CAE involves learning the image features based on an end to end unsupervised learning strategy. This network is considered superior to the stacked autoencoder due to a relationship between image pixels. In this case, features are extracted after filters get trained. The extracted features are combined to form the input.

In the convolutional autoencoder, there is the same number of arguments for creating an activation map. This makes it suitable for medicinal images. By replacing the fully connected layer of a simple autoencoder with a convolutional layer, it becomes a convolutional autoencoder.

Except for decoding network, input and output layers have the same size, but it changes in the convolutional network.

4.9 Deep Belief Networks (DBN)

Deep belief network is a stacked restricted Boltzmann machine. It is a generative model as presented in Fig. 11.

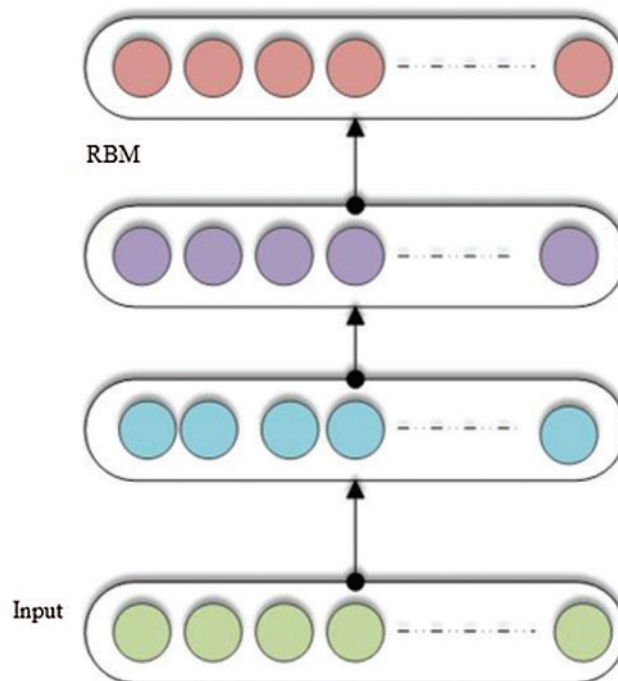


Figure 11: DBN architecture

The architecture of the deep belief network consists of two layers of restricted Boltzmann machine called visible layer or input layer and hidden layer. In this case, a restricted Boltzmann machine is optimized based on contractive diversion theorem by combining gradient Descent and Gibbs sampling using a greedy learning strategy.

4.10 Residual Network (ResNet)

He et al. [36] introduced a specific type of NN for solving complex problems. Generally, in tractional NN approaches, more layers are added to improve the accuracy of the resulting solutions. The main idea behind adding more layers is that these layers can learn more complex features of the problem and results in accurate solutions. But, it has been proved that adding more layers on top of NN results in an increased error % in top layers. This indicates that adding more layers on top of a network causes NNs' performance degradation in training and test datasets. The performance degradation can be linked to the optimization function, initialization of the network, and, more importantly, the vanishing gradient problem. This problem of training very deep networks has been alleviated with the introduction of ResNet or residual networks. Resnets consists of Residual Blocks as depicted in Fig. 12.

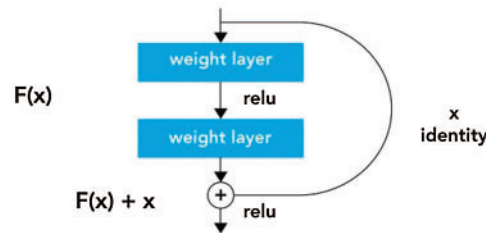


Figure 12: ResNet learning blocks

It can be seen from Fig. 12 that there exists a direct connection that enables skipping of some layers in between. Such connection is known as ‘skip connection’ and is the central concept of residual blocks. The output of the layer is different due to this skip connection. Using skip connection, residual block solves the problem of vanishing gradient in deep NN. It allows this alternate direct path for the gradient to flow through. Skip connection also helps the NN model learn the identity functions that ensure the higher layer performance will be better or the same as that of the lower layer. ResNet resulted in a considerable enhancement in the performance of NN with many layers in comparison to traditional NNs.

4.11 Xception

Chollet at Google Inc. introduced an extreme version of the Inception network called Xception [37]. Xception is a CNN architecture based entirely on depthwise separable convolution layers. The developer made a hypothesis that mapping of cross-channel correlations and spatial correlations in the feature maps of CNNs can be entirely decoupled. Accordingly, he proposed extreme Inception called Xception.

The proposed Xception architecture consists of 36 convolutional layers for extracting features, followed by a logistic regression layer. The 36 convolutional layers are structured into 14 modules, all of which have linear residual connections around them, except for the first and last modules. The Xception architecture is proposed as a linear stack of depthwise separable convolution layers with residual connections.

4.12 VGG Net

Simonyan et al. [38] proposed a CNN model called VGG Net. VGG Net resulted in an accuracy of 92.7% in the ImageNet dataset consisting of over 14 million images belonging to 1000 classes.

VGG Net architecture replaces large kernel-sized filters with 11 and 5 in the first and second layers. It reported a considerable improvement over AlexNet architecture, with multiple 3×3 kernel-sized filters one after another. It takes the input to CNN is a fixed-size 224×224 RGB image. The preprocessing step involves subtracting the mean RGB values computed on the training dataset for each pixel. Preprocessing step is followed by running of the image through a stack of CNN layers having filters of 3×3 . Its architecture also consists of 1×1 convolution filters that ensure the linear transformation of the input channels, followed by non-linear ones. The convolutional strides are assumed to be 1 pixel. The spatial padding of convolutional layer input is designed so that the spatial resolution is maintained after convolution. The padding process is followed by spatial pooling using five max-pooling layers over a 2×2 -pixel window, with stride 2.

VGG Net is the most popular deep learning architecture used for benchmarking on any particular task. It has been made open source and can be used out of the box for various applications.

The deep learning architectures mentioned above are summarized in Fig. 13.

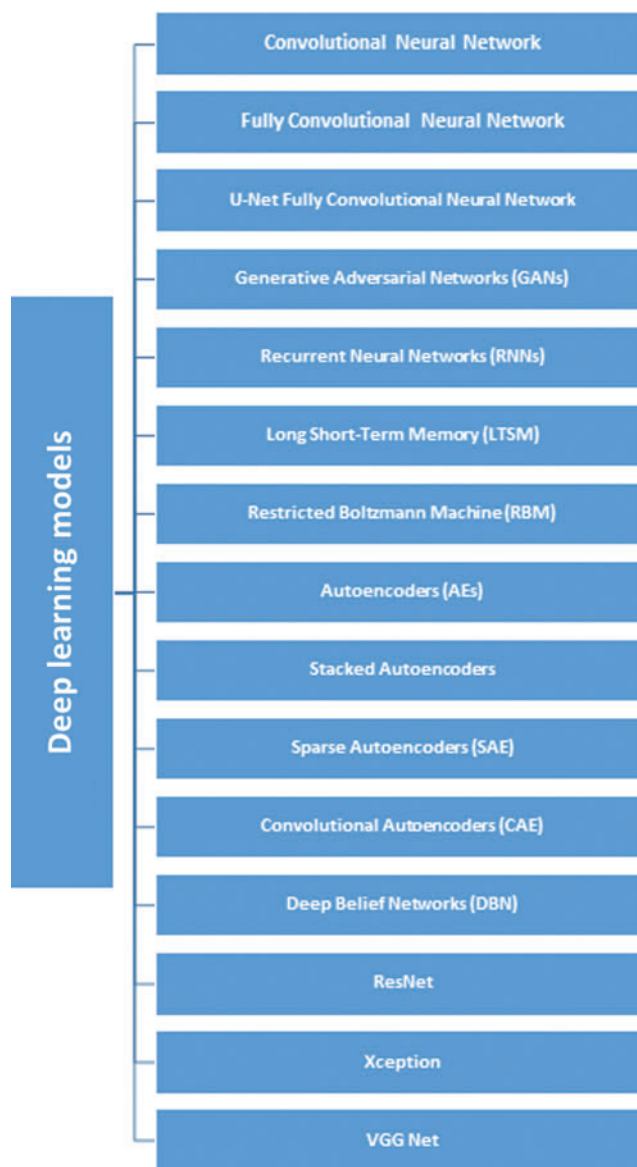


Figure 13: Deep learning architectures

5 Deep Learning Methods in Cancer Detection

Deep learning is a subset of machine learning algorithms that are highly capable of analyzing digital images. In the recent past, deep learning methods have reported extraordinary accurate results compared to conventional machine learning algorithms. Deep learning models

have minimized the use of manually extracted features and image processing methods by using multiple layered architectures [39]. The models have significant characteristics of extracting deeper-level images by transforming them into two consecutive layers. Deep learning models enable an analysis of different shapes, patterns, density functions and colours of the images using different feature maps. Deep learning models have been successfully employed in several natural language processing, image processing, and time series analysis [15]. Besides, deep learning has also become the most popular method for analyzing medical images to detect different diseases. The researchers have successfully applied deep learning models for detecting the different kinds of cancers based upon medical images such as CT scans, MRI scans and ultrasound scans.

This section provides a comprehensive review of state of the art in deep learning methods used for detecting cancer disease. We focus on four kinds of cancer: breast, brain, skin, and prostate, using deep learning methods.

5.1 Breast Cancer

Many researchers have applied deep learning methods for detecting and diagnosing breast cancer in recent years. For example, Albayrak et al. [40] proposed a deep learning-based approach that extracts features from the histopathology images for detecting breast cancer. The extracted features are supplied to support vector machine for its training and classification purpose to detect best cancer images. Spanhol et al. [41] used AlexNet to construct a CNN for classification of normal and abnormal metosis using breast best histopathological images. The researchers used different deep learning models to extract intrinsic features from the images and classify the images.

Similarly, Chen et al. [42] proposed the detection of mitosis using breast histology images. They proposed to extract mitosis patient features from images using a pre-trained fully CNN. It is followed by the application of CaffeNet model for the classification of mitosis samples. Further, three networks having a fully connected layer with different settings have been employed to generate multiple probability scores. The authors proposed to take an average of multiple scores to obtain the final output.

Albarqouni et al. [43] also trained deep CNN model for expert crowd annotations concerning biomedical context. The author proposed a multi-scale CNN architecture that combines crowd and annotations in a particular way. So that, for each softmax layer, there is an aggression layer for aggregating the predicted results from multiple candidates. The authors of [44] suggested the use of a stacked sparse autoencoder based approach for classifying nuclei using histopathological images of the breast. They suggested optimizing the model using a greedy approach that considers one hidden layer at a time for training. The output of the previously hidden layer is supplied to the next layer as an input. They demonstrated the validity of their proposed approach using a mammographic image dataset.

Wichakam et al. [45] proposed a hybrid approach by integrating the support vector machine and the CNN model to detect mass on digital mammograms. They suggested using Mammogram patches to train the CNN model for extracting high-level features from the last fully connected layer. The extracted features are supplied to support vector machines for training and classification purposes.

Some researchers focused on transfer learning methods to address the issue of insufficient training data. Wichakam et al. [45] proposed the use of a pre-trained CNN model for detecting mass in the given mammogram. However, in the case of limited training data, there can be an

over-fitting problem. To avoid the over-fitting problem, Swiderski et al. [46] used the statistical self-similarity method and non-negative matrix factorization to increase the training data.

Ertosun et al. [47] presented an approach to detect the presence of a mass in a mammogram followed by the location of the mass in the images. It requires learning of the features from mammogram images in multiple scales. So, Kallenberg et al. [48] suggested a model on tag convolutional autoencoder for learning features from mammograms. They validated the robustness of the model by introducing sparsity regularization in their model. Similarly, Dhungel et al. [49] used a structured support vector machine for combining many functions, including a mixture model of deep belief networks before the location method for segmenting the mass in mammograms. They also suggested a hybrid model of deep learning model and random forest classifier for detecting mass in mammograms [50]. Kim et al. [51] developed a 3D multi-view approach that learns bilateral features using digital breast tomosynthesis. Similarly, CNN is also applied by many researchers using different breast cancer datasets in [52–54].

The review of deep learning methods for detecting breast cancer is summarized in Table 1.

Table 1: Summary of deep learning based breast cancer detection methods (in chronological order)

Study	DL model	Target disease	Image type	Dataset	Results
Xing et al. [52]	CNN	Nucleus segmentation	Histopathology	-	Precision = 0.71, Recall = 0.88, F-Measure = 0.78
Ertosun et al. [47]	CNN	Mass segmentation	Mammographic	DDSM	Accuracy = 0.85
Dhungel et al. [49]	DBN	Mass segmentation	Mammographic	DDSM + INbreast	Dice index = 0.88
Dhungel et al. [50]	Hybrid (DBN + CNN)	Mass detection	Mammographic	DDSM + INbreast	TPR = 0.75
Xu et al. [44]	SSAE	Nuclei classification	Histopathology	Synthetic	Accuracy = 0.76
Wichakam et al. [53]	CNN	Mass detection	Mammographic	INbreast	Accuracy = 0.98
Chen et al. [42]	Hybrid (FCN + CNN)	Mitosis detection	Histopathology	MITOSATYPIA-12, MITOSATYPIA-14	Rank = 1
Albarqouni et al. [43]	CNN	Mitosis detection	Histopathology	MITOSATYPIA-13	AUC = 0.76
Albayrak et al. [40]	CNN	Mitosis detection	Histopathology	MITOSATYPIA-14	Accuracy = 0.96
Kim et al. [51]	CNN	Latent bilateral feature representation learning	Tomosynthesis	Synthetic	AUC = 0.847

(Continued)

Table 1 (continued).

Study	DL model	Target disease	Image type	Dataset	Results
Kallenberg et al. [48]	SSAE	Breast density segmentation & risk scoring	Mammographic	Synthetic	AUC = 0.59
Spanhol et al. [41]	CNN	Breast cancer classification	Histopathology	BreaKHis	Accuracy = 0.89
Suzuki et al. [54]	CNN	Mass detection	Mammographic	DDSM	DR = 0.899
Swiderski et al. [46]	CNN	Lesion recognition	Mammographic	DDSM	Accuracy = 0.858

5.2 Brain Cancer

Brain cancer is one of the most challenging tasks in detecting cancer. It is an uncontrolled development, and that can develop in any portion of the brain. The most challenging task in detecting brain cancer is to separate the affected portion of the brain from the healthy part. Several research efforts have been made in detecting brain cancer.

Guo et al. [55] proposed two algorithms based on two dimensional CNN and three dimensional CNN. The proposed algorithms work on two-dimensional slice brain images and three-dimensional brain images extracting prominent features. The integration of these two algorithms reported better performance than kaze feature and scale-invariant features Swift (SIFT).

Similarly, Pereira et al. [56] suggested an image segmentation approach using CNN that segments MRI images automatically. In their research, the authors analyzed intensity normalization as well as other functions for detecting brain tumors.

Zhao et al. [57] advocated a hybrid approach using a conditional random field and fully connected convolutional neural network for segmenting the brain cancer images. In the first phase, the authors proposed to train a conditional random field and fully connected convolutional neural network using image patches. Finally, they fine-tuned their system using image slices directly.

Kamnitsas et al. [58] suggested joining adjacent image patches into one pass based on a dense training method of the CNN model. They reduced false positives based on three dimensional fully connected random field. Similarly, researchers also used CNN for detecting brain cancer in [59,60].

The review of deep learning methods for detecting brain cancer is summarized in Table 2.

Table 2: Summary of deep learning based brain cancer detection methods (in chronological order)

Study	DL model	Target disease	Image type	Dataset	Results
Zhao et al. [57]	CNN	Brain tumor Segmentation, Cancer detection	MRI	BraTS	Accuracy = 0.77

(Continued)

Table 2 (continued).

Study	DL model	Target disease	Image type	Dataset	Results
Liu et al. [59]	CNN	Feature learning of brain tumor	MRI	Synthetic	Accuracy = 0.95
Pereira et al. [56]	CNN	Brain tumor segmentation	MRI	BraTS	Sensitivity = 0.86
Kamnitsas et al. [58]	CNN	Brain tumor segmentation	MRI	BraTS	Precision = 0.68
Ahmed et al. [60]	CNN	Classification of brain tumor	MRI	Synthetic	Accuracy = 0.81

5.3 Lung Cancer

Deep learning methods have been successfully applied for detecting lung cancer and the survival time of the patients. For example, Zhu et al. [61] proposed a deep convolutional neural network for predicting the survival time of lung cancer patients using their pathological images. Similarly, Paul et al. [62] use a pre-trained CNN model to detect lung cancer based on CT scan images. Han et al. [63] suggested a hybrid approach of CNN and DBN using an end-to-end learning strategy for protecting lung cancer from the images. The authors use Tut medicinal CT Scan images for pulmonary load classification.

On similar lines, Hussein et al. [64] employed the CNN model on 3D CT Scan images using an end to end at learning strategy. They suggested extracting two-dimensional patches from three dimensional CT Scan images rather than creating a map of two-dimensional model.

The review of deep learning methods for detecting lung cancer can be summarized in Table 3.

Table 3: Summary of deep learning based lung cancer detection methods (in chronological order)

Study	DL model	Target disease	Image type	Dataset	Results
Hua et al. [65]	DBN & CNN	Nodule classification	CT scan slices	LIDC-IDRI	Sensitivity = 0.73, Specificity = 0.82
Setio et al. [66]	CNN	Pulmonary nodules detection	Volume CT scan	LIDC-IDRI, ANODE09, DLCST	Sensitivity = 0.90
Paul et al. [62]	CNN	Survival prediction	Volume CT scan	Synthetic	Accuracy = 0.825
Zhu et al. [61]	CNN	Survival analysis	Histopathology	Synthetic	Concordance index = 0.629

(Continued)

Table 3 (continued).

Study	DL model	Target disease	Image type	Dataset	Results
Tajbakhsh et al. [67]	MTANN & CNN	Lung nodule detection	CT scan slices	Synthetic	AUC = 0.8806
Kim et al. [68]	SSAE	Pulmonary nodule classification	CT scan slices	Synthetic	Accuracy = 0.955
Dou et al. [69]	CNN	Pulmonary nodules detection	Volume CT scan	LIDC-IDRI	a Sensitivity = 0.677
Hirayama et al. [70]	CNN	Extraction of ground glass opacity (GGO) candidate region	CT scan slices	LIDC	TPR = 0.93
Wang et al. [71]	CNN	Lung nodule classification	CT scan slices	JSRT	Sensitivity = 0.69, Specificity = 0.96
Hussein et al. [64]	CNN	Nodule characterization	Volume CT scan	LIDC-IDRI	Accuracy = 0.92
Shen et al. [72]	CNN	Lung nodule malignancy suspiciousness classification	Volume CT scan	LIDC-IDRI	Accuracy = 0.87

5.4 Skin Cancer

Skin cancer or melanoma is one of the most challenging jobs in detecting cancer. Several methods have been proposed for the accurate detection of skin cancer. For instance, Pomponiu et al. [73] proposed a pre-trained CNN and AlexNet model for extracting features from skin images. They used a k-nearest neighbour classifier for classifying the lesion. The proposed model reported an accuracy of 93.62% based on 399 skin images.

Esteva et al. [74] also used a pre-trained CNN model using 129,450 images. They used a mixture of skin lesion images and dermatoscopic images. Their research work performed two sets of classification experiments, namely benign nevi vs. malignant melanoma and benign seborrheic keratosis vs. keratinocytes carcinomas. They used transfer learning for extracting features and classification of diseases.

Han et al. [63] conducted a set of experiments for the classification of twelve kinds of skin diseases based upon 19,398 images using the ResNet model.

Masood et al. [75] suggested a self-advised semi-supervised model detecting melanoma based on two self-advised support vector machines and a deep belief network. They conducted their experiments using three different datasets combined with two kernels, namely, polynomial kernel and radial basis function kernel. They fine-tuned the model using the exponential loss function to maximize the separation between label data.

In another approach, researchers [76] integrated manually extracted features and deep learning features for training the SVM classifier models to predict the probability score of samples. The outcome is computed based upon the highest score of the samples.

Sabbaghi et al. [77] integrated a deep neural network to classify medical images and improved the accuracy of the bag of features.

Demyanov et al. [78] applied stochastic gradient descent for training the CNN model. The trained model was used for detecting regular globules and network patterns. Residual blocks are also used for replacing convolutional layers of the fully connected convolutional neural network [79]. The resultant network is called a fully convolutional residual network (FCRN) and is applied in the classification of unknown samples. Similarly, CNN model was also used for detecting melanoma using pre-processed medical images in [80,81].

Some researchers used ABCDE method for detecting cancer using medical images [79,82]. They used different pre-processing methods like histogram analysis, segmentation techniques and contour tracer.

Some researchers applied clustering techniques in combination with classification techniques to analyze skin cancer [83,84]. They used fuzzy C-mean clustering, K-means clustering in combination with K-nearest neighbour classifier and support vector machine.

He et al. [85] trained a CNN model consisting of eight layers using the back-propagation method based upon 900 images. The trained CNN model achieved 91.92% accuracy on the training data set and 89.5% accuracy on the test dataset.

Pham et al. [86] used data augmentation methods for improving the accuracy of the CNN model. In their work, they attempted to address the issue of the limited availability of training data. By using data augmentation, the authors reported an accuracy of 89.2% based on 600 medical images. They also highlighted the impact of using data augmentation methods on the performance of the CNN model in their paper.

Zhang et al. [87] conducted a set of experiments for detecting four types of cutaneous diseases using deep learning techniques. They summarized classification and diagnosis features in the form of a hierarchical structure. They reported an accuracy of 87.25% for detecting the diseases with a probability error of 2.24%.

Vesal et al. [88] developed a convolutional neural network called SkinNet. SkinNet is used for segmenting and detecting skin cancer diseases. The proposed system is a modified version of U-net CNN model. The authors reported the values of 85.1%, 93% and 76.67% for dice coefficient, sensitivity and Jaccard index, respectively. The authors also suggested a multi-task conventional neural network in combination with segmentation and joint detection framework in [89]. The proposed system is called the faster region-based CNN model. They used region proposals and bounding boxes for localizing the lesion. Bounding blocks were defined by using the softmax function.

Horie et al. [90] developed a CNN model to detect oesophageal cancer, adenocarcinoma, and SCC (Squamous cell carcinoma). They trained their proposed model using 8,428 medical images of 384 patients in Japan. The trained model was tested based upon 1,118 images collected from 47 patients. They reported accuracy and sensitivity of 98% both.

Gomez-Martin et al. [91] analyzed dermoscopic, clinical, and confocal parameters to detect flat leg lesions pink shaded in elders. An accuracy of 49.1%, specificity of 73.4% and sensitivity of 68.7% have been reported using the clinical diagnosis process. Whereas 59.6%, 85% and 67.6%

accuracy, sensitivity and specificity respectively is reported using dermoscopy. Confocal microscopy resulted in 85.1%, 97.5% and 88.2% accuracy, sensitivity and specificity, respectively.

The review of deep learning methods for detecting skin cancer is summarized in [Table 4](#).

Table 4: Summary of deep learning based skin cancer detection methods (in chronological order)

Study	DL model	Target disease	Image type	Dataset	Results
Masood et al. [75]	DBN	Skin lesion classification	Dermoscopy	Synthetic	Accuracy = 0.89
Sabbaghi et al. [77]	SSAE	Melanomas classification	Dermoscopy	Synthetic	Accuracy = 0.95
Pomponiu et al. [73]	CNN	Skin mole lesion classification	Dermoscopy	DermIS, DermQuest	Accuracy = 0.96
Sabouri et al. [81]	CNN	Lesion border detection	Clinical photography	DermIS, DermQuest	Accuracy = 0.867
Demyanov et al. [78]	CNN	Dermoscopy patterns classification	Dermoscopy	ISIC	Accuracy = 0.88
Majtner et al. [76]	CNN	Skin lesion classification	Dermoscopy	ISIC	Accuracy = 0.826
Yu et al. [79]	Hybrid (FCN+ CNN)	Melanoma recognition	Dermoscopy	ISIC	Accuracy = 0.855
Nasr-Esfahani et al. [80]	CNN	Melanoma detection	Clinical photography	MED-NODE	Accuracy = 0.81
Esteva et al. [74]	CNN	Dermatologist-level skin cancer classification	Dermoscopy	Open-access online dataset, Unpublished clinical dataset	AUC = 0.96
Mahbod et al. [92]	CNN	Skin lesion classification	Dermoscopy	ISIC	AUC = 0.975

5.5 Prostate Cancer

Prostate cancer is highly diagnosable in males. It is considered as the third-highest cause of death in males [55]. It requires successful segmentation of the medical images for its timely diagnosis and radiotherapy. Guo et al. [55] suggested a combined model of deep learning and sparse patch matching for successful segmentation of prostate medical images. They suggested the use of the sparse stacked autoencoder method for extracting the features from MRI images.

Similarly, Yan et al. [93] also proposed a prostate cancer detection method using stacked autoencoder as a classifier. They suggested using a sparse text autoencoder to improve the extracted features in the supervised fine-tuning method.

Shen et al. [72] developed a multivariate convolutional neural network raising the issue of variable nodule size. This network is capable of generating multi-scale features. For this purpose, this network replaces the max-pooling layer with a multi-crop pooling layer in the CNN model. The authors suggested using the randomized leaky rectified linear unit for doing non-linear transformation in their research.

Yu et al. [79] developed a volumetric convolutional neural network segmenting three dimensional MRI images of prostate patients to detect cancer. The authors suggested an extension of their research by using the residual block that enables the volume to volume prediction.

Ma et al. [94] proposed an approach based on CNN for using the region of interest and detecting prostate cancer based on image patches. They computed the final results based on a multi-atlas label function.

The review of deep learning methods for detecting prostate cancer is summarized in Table 5.

Table 5: Summary of deep learning based prostate cancer detection methods (in chronological order)

Study	DL model	Target disease	Image type	Dataset	Results
Guo et al. [55]	SSAE	Deformable prostate segmentation	MRI	Synthetic	Precision = 0.87
Zhao et al. [57]	CNN	Prostate segmentation	MRI	BraTS	Accuracy = 0.77
Milletari et al. [95]	CNN	Prostate segmentation	3D MRI	PROMISE12	Dice loss = 0.869
Yu et al. [79]	FCN	Prostate segmentation	3D MRI	PROMISE12	Accuracy = 0.855
Yan et al. [93]	SSAE	Prostate recognition	MRI	PROMISE12	DSC = 0.89, HD = 3.8
Kallen et al. [96]	CNN	Gleason grading	Histopathology	Synthetic	Accuracy = 0.89
Gummeson et al. [97]	CNN	Gleason grading	Histopathology	Synthetic	Error rate = 7.3%
Kwak et al. [98]	CNN	Lumen-based prostate cancer detection	Histopathology	Synthetic	AUC = 0.95
Ma et al. [94]	CNN	Prostate segmentation	CT scan slices	Synthetic	Dice similarity coefficient = 86.80%

(Continued)

Table 5 (continued).

Study	DL model	Target disease	Image type	Dataset	Results
Tian et al. [99]	FCN	Prostate segmentation	MRI	Synthetic	Dice similarity coefficient = 85.3%
Cheng et al. [100]	HNN	Prostate segmentation	MRI	Synthetic	Dice similarity coefficient = 88.70%

5.6 Other Types of Cancer

Many doctors and practitioners also applied deep learning methods for detecting cancer of different forms like bladder cancer, cervical cancer, etc. For instance, Song et al. [101] suggested the use of multi-scale conventional neural networks in association with the partitioning method for segmenting by cervical cytoplasm. Their proposed system helps to perform segmentation and graph partitioning to improve the segmentation results. Similarly, Xu et al. [102] also used deep learning-based CNN model for diagnosing cervical dysplasia diagnosis. They fine-tuned their CNN model using cervigram dataset. They proposed joining the extracted features of images with multi-model clinical features to train the classifier. The trained classifier is further used to detect unknown samples of cancer.

Many researchers used deep learning methods to detect bladder cancer. For instance, deep learning was used the CNN model for pleural segmentation based on CT urographical images in [103]. In this research, the trained CNN model was used for computing probability map of patterns around the bladder in CT urographical images.

Liver cancer is also considered the primary form of cancer. Deep learning methods can help to diagnose liver cancer using laparoscopic videos. For example, Gibson et al. [104] developed an automatic segmentation system for laparoscopic videos to diagnose liver cancer. Similarly, Li et al. [105] used the CNN model for the automatic segmentation process to detect liver tumors using CT scan.

The review of deep learning methods for detecting cancer is summarized in Table 6.

Table 6: Summary of deep learning based other types of cancer detection methods (in chronological order)

Study	DL model	Target disease	Image type	Dataset	Results
Song et al. [101]	CNN	Cervical cytoplasm and nuclei segmentation	Histogram pathology	Synthetic	Correct split = 90.2%

(Continued)

Table 6 (continued).

Study	DL model	Target disease	Image type	Dataset	Results
Li et al. [105]	CNN	Liver tumor segmentation	CT slices	Synthetic	Dice similarity coefficient = 0.81
Mao et al. [106]	CNN	Circulating tumor-cell detection	Histogram pathology	Synthetic	F Score = 0.97
Song et al. [107]	CNN	Cervical cytoplasm segmentation	Histogram pathology	ISBI, SZU	F Score = 0.94
BenTaieb et al. [108]	CNN	Colon adeno-carcinoma glands segmentation	Histogram pathology	Warwick-QU	Accuracy = 0.89
Cha et al. [103]	CNN	Bladder cancer treatment response assessment	CT slices	Synthetic	Jaccard index = 0.762
Gibson et al. [104]	CNN	Liver segmentation on Laparoscopic videos	Laparoscopy	Synthetic	Dice scores = 0.95
Gordon et al. [109]	CNN	Inner/outer bladder wall segmentation	CT slices	Synthetic	Average volume intersection = $90.0 \pm 8.7\%$

6 Summary and Discussion

In this study, we reviewed 59 papers published for detecting cancer, focusing on brain cancer, breast cancer, skin cancer, prostate cancer, and other types of cancer diseases. The summary of papers reviewed category wise is presented in Table 7.

It is evident from the comprehensive review mentioned above that deep learning methods have reported the best performance for detecting different types of cancer. CNN model has reported the best performance in all architectures. Many researchers have focused on using pre-trained CNN models for detecting cancer diseases.

In the case of breast cancer detection, histopathology and mammography remain in the focus of doctors and practitioners (refer to Table 1). Most researchers used MRI scans for detecting brain cancer (refer to Table 2). Volumetric CT scans are most helpful in diagnosing lung cancer (refer to Table 3). Dermoscopy and clinical photography are found be suitable for diagnosing skin cancer.

Table 7: Summary of deep learning based cancer detection methods

Type of cancer	Number of papers	Studies
Brain	5	Ahmed et al. [60], Liu et al. [59], Kamnitsas et al. [58], Pereira et al. [56], Zhao et al. [57]
Breast	14	Xing et al. [52], Spanhol et al. [41], Ertosun et al. [47], Swiderski et al. [46], Suzuki et al. [54], Dhungel et al. [50], Dhungel et al. [49], Wichakam et al. [53], Chen et al. [42], Albarqouni et al. [43], Albayrak et al. [40], Kallenberg et al. [48], Kim et al. [68], Xu et al. [44]
Lung	11	Wang et al. [71], Hirayama et al. [70], Hussein et al. [64], Dou et al. [69], Shen et al. [72], Hua et al. [65], Setio et al. [66], Kim et al. [51], Paul et al. [62], Tajbakhsh et al. [67], Zhu et al. [61]
Prostate	10	Milletari et al. [95], Yu et al. [79], Yan et al. [93], Cheng et al. [100], Gummesson et al. [97], Guo et al. [55], Kallenberg et al. [48], Kwak et al. [98], Ma et al. [94], Tian et al. [99]
Skin	10	Pomponiu et al. [73], Sabouri et al. [81], Demyanov et al. [78], Majtner et al. [76], Yu et al. [79], Mahbod et al. [92], Nasr-Esfahani et al. [80], Esteva et al. [74], Masood et al. [75], Sabbaghi et al. [77]
Others	9	BenTaieb et al. [108], Cha et al. [103], Gibson et al. [104], Gordon et al. [109], Li et al. [105], Mao et al. [106], Song et al. [101], Song et al. [107], Zhao et al. [57]

Several benchmark datasets have been developed for validating the deep learning models. Table 8 presents the most commonly used datasets in detecting cancer.

Table 8: Summary of benchmark datasets for cancer detection methods (in alphabetical order)

Benchmark dataset	Medical images	Usage
ANODE09 [110]	CT scans	Detection of pulmonary nodules in thoracic CT scans
ASU-Mayo Clinic Colonoscopy Video (c) Database [111]	Colonoscopy videos	Polyp detection in colonoscopy videos.
BraTS [112]	Four MRI sequences: T1, T1c, T2, and FLAIR	Brain tumor images

(Continued)

Table 8 (continued).

Benchmark dataset	Medical images	Usage
BreaKHis [113]	Breast tumor images	Benign and malignant breast tumor recognition
CRCHistoPhenoty pes [114]	H&E stained histology images of colorectal adenocarcinomas	Detection of nuclei in routine colon cancer histology images
DDSM [115]	Digitized film screen mammograms	Detection and diagnosis of mammary lesions
DermIS-BioGPS [116]	Dermoscopic images	Skin lesion detection
DermQuest [117]	Dermoscopic images	Skin lesion detection
DLCST [118]	CT screening of smokers	Lung cancer detection and diagnosis
INbreast [118]	Mammograms	Detection and diagnosis of mammary lesions
ISIC [119]	Dermoscopic images	Skin lesion detection
JSRT [120]	Chest X-ray images	Lung nodule detection
LIDC-IDRI [121]	Diagnostic and lung cancer screening thoracic computed tomography (CT) scans	Lung cancer detection and diagnosis
MED-NODE [122]	Dermoscopic images	MED-NODE system
MITOS-ATYPIA [123]	Breast cancer biopsy images	Nuclear atypia scoring
PROMISE12 [124]	Prostate MRI	MRI of the prostate
TCIA collections [125]	Cancer-related imaging	Advanced medical images
Warwick-QU [126]	Images of Hematoxylin and Eosin (H&E) stained slides	Segmentation of glands

Out of 59 papers reviewed in this work, 16 researchers have validated their work using synthetic or unpublished datasets. Most brain cancer-detecting approaches used BraTS dataset for validating their proposals. DDSM and INbreast datasets are the most commonly used datasets for detecting breast cancer. Skin cancer-detecting approaches have been validated using DermIS, DermQuest and ISIC datasets. The researchers working for lung cancer detection preferred to use LIDC-IDRI, ANODE09, and DLCST datasets for demonstrating the performance of their deep learning models.

Despite successful employment and improved performance of deep learning models for detecting cancer, there are many challenges in implementing deep learning to diagnose cancer in the early stages accurately.

One of the significant challenges is the lack of benchmark datasets for validating the novel deep learning models. The deep learning methods need a large quantity of quality data for developing a trained model. Many hospitals and organizations are reluctant to share the confidential image data of the patients. However, the picture archiving and communication society (PACS) has made several efforts to provide medical images of several patients to the research community for validating new deep learning models. Many researchers have also used their collected medical image data from different cancer Research hospitals and organizations to validate their model.

To address the issue, Esteva et al. [74] provided several medical images for validating deep learning models. Their data set contains 127,463 and 1,942 medical images for training and testing purposes of deep learning models in cancer detection. Besides, many datasets are available online in the form of raw images, researchers have to explore the ground truth at their level before using the available datasets to validate deep learning models.

Many researchers advocated using data augmentation methods to address the lack of availability of limited data for validating deep learning models. The most commonly used data augmentation method includes rotation, filtering and cropping to increase the number of medical image data for training and testing purposes of the deep learning models. However, this may cause the over-fitting problem. Some researchers use the pre-trained models and transfer learning methods to avoid the over-fitting problem.

It has been observed that low-quality images like low contrast and low signal to noise ratio of images can be a significant cause to deteriorate the performance of deep learning models. Low-quality medical images are another primary concern that impacts the performance of deep learning models.

Some researchers have argued to maintain the performance of the deep learning models, especially in brain cancer, about training data collected from multiple sources. It has been observed that the performance of the deep learning model deteriorates on using data from different institutions.

Several institutions have provided medical image data for validating deep learning models in detecting cancer. These datasets can help researchers in the field to develop new models and validate them. Table 8 presents the most commonly used benchmark datasets in the field of cancer detection.

A significant issue that arises in the training data is the unequal distribution of data in benchmark datasets. Deep learning models trained on imbalanced data leads to biased results towards majority classes. Most of the researchers ignored biased cancer detection results towards majority classes in their research.

Another problem in implementing the convolutional neural network is the object's size to be detected in the image. As the object's size in the target image may vary in different images, deep learning models must be trained to learn size variation in different scales. To address the issue of multi-scale features, Shen et al. [72] proposed to replace the average pooling function with a multi-crop pooling function.

One general and notable point in the recent deep learning developments is that it has resulted in tremendous growth in the number of parameters of deep learning models. This indicates that the models become increasingly complex in context to memory and computing resource consumption [127]. The increased requirement of computational resources has also increased power consumption. The increased complexity of deep learning models requires more input-output operations. The input or output operations are more expensive in comparison to arithmetic operations, which has led to more power consumption [128].

It has been observed that most of the hardware platforms are unable to cope up with the exponential growth of complexity and size of deep learning models [128] due to the increased power consumption of these models. Besides, there is an increased demand for installing machine learning models into limited resource devices [129,130]. The increased size of deep learning models in terms of their number of parameters also impacts security and efficiency. The heavy sized

machine-learning model has limited applicability in the domains having a limited bandwidth channel.

In order to solve the problems mentioned above, compression algorithms are considered one of the promising directions in deep learning algorithms. Compression algorithms can compress the size of machine learning models and make them practical for solving complex problems. Application of compression algorithms in developing machine learning models have practical benefits of the requirement of reduced computing resources, communication overhead, and memory overhead [131–133].

7 Conclusion

Cancer disease is one of the most severely affecting diseases in human life. It is well established that cancer diagnosis and its treatment is a long and challenging procedure. Several research efforts have been made to address the long and challenging procedure accurately and diagnose cancer in the early stages.

In recent years, deep learning, a subset of artificial intelligence, has increased researchers and practitioners' attention in detecting cancer using medical images like CT scan, MRI scans, ultrasound, etc. Deep learning methods have been successfully implemented in classification problems of different domains such as computer vision, cybersecurity, natural language processing and many more.

Recently deep learning methods have been applied to detect cancer and its types in different organs of the human body by analyzing medical images. Deep learning models have provided extraordinary accurate results in diagnosing cancer in the early stages and help increase the patients' survival time.

In this study, we introduced the cancer detection problem, phases of the cancer detection process, deep learning models and their classification, and applications of deep learning methods in detecting different types of cancer. We highlighted recent developments in detecting cancer by focusing on four types: brain cancer, breast cancer, skin cancer, and prostate cancer. We compared different deep learning models with their pros & cons and described their applications in detecting cancer in detail.

Recent developments in cancer detection are summarized and analyzed to identify challenges in implementing deep learning methods for cancer detection. The identified challenges provide future research directions to fellow researchers. The outcome of this study is helpful for researchers to develop new deep learning models for cancer disease diagnosis in early stages accurately and treatment planning accordingly.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Basarslan, M. S., Kayaalp, F. (2021). Performance evaluation of classification algorithms on diagnosis of breast cancer and skin disease. *Deep learning for cancer diagnosis*, pp. 27–35. Singapore: Springer.
2. Kose, U., Alzubi, J. (2021). *Deep learning for cancer diagnosis*. Singapore: Springer.

3. Kaya, U., Yılmaz, A., Dikmen, Y. (2019). Sağlık alanında kullanılan derin öğrenme yöntemleri. *Avrupa Bilim ve Teknoloji Dergisi*, (16), 792–808. DOI 10.31590/ejosat.573248. <https://dergipark.org.tr/en/download/article-file/773628>.
4. Center, J. S. C. (2020). Top five most dangerous cancers in men and women. <https://www.unitypoint.org/desmoines/services-cancer-article.aspx?id=c9f17977-9947-4b66-9c0f-15076e987a5d>.
5. WHO (2020). Cancer. <https://www.who.int/>.
6. Sencard.com.tr (2019). Importance of early diagnosis in cancer.
7. Hu, Z., Tang, J., Wang, Z., Zhang, K., Zhang, L. et al. (2018). Deep learning for image-based cancer detection and diagnosis—A survey. *Pattern Recognition*, 83, 134–149. DOI 10.1016/j.patcog.2018.05.014.
8. Fass, L. (2008). Imaging and cancer: A review. *Molecular Oncology*, 2(2), 115–152. DOI 10.1016/j.molonc.2008.04.001.
9. Doi, K. (2007). Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31(4–5), 198–211. DOI 10.1016/j.compmedimag.2007.02.002.
10. Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8–17. DOI 10.1016/j.csbj.2014.11.005.
11. Kumar, S., Fred, A. L., Padmanabhan, P., Gulyas, B., Kumar, H. A. et al. (2021). Deep learning algorithms in medical image processing for cancer diagnosis: Overview, challenges and future. *Deep learning for cancer diagnosis*, vol. 908, pp. 37–66. Singapore: Springer. DOI 10.1007/978-981-15-6321-8_3.
12. Messadi, M., Bessaid, A., Taleb-Ahmed, A. (2009). Extraction of specific parameters for skin tumour classification. *Journal of Medical Engineering & Technology*, 33(4), 288–295. DOI 10.1080/03091900802451315.
13. Reddy, B. V., Reddy, P. B., Kumar, P. S., Reddy, S. S. (2014). Developing an approach to brain MRI image preprocessing for tumor detection. *International Journal of Research*, 1(6), 725–731.
14. Zacharaki, E. I., Wang, S., Chawla, S., Soo Yoo, D., Wolf, R. et al. (2009). Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 62(6), 1609–1618. DOI 10.1002/mrm.22147.
15. Tong, N., Lu, H., Ruan, X., Yang, M. H. (2015). Salient object detection via bootstrap learning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, Massachusetts, USA.
16. Emre Celebi, M., Kingravi, H. A., Iyatomi, H., Alp Aslandogan, Y., Stoecker, W. V. et al. (2008). Border detection in dermoscopy images using statistical region merging. *Skin Research and Technology*, 14(3), 347–353. DOI 10.1111/j.1600-0846.2008.00301.x.
17. Li, X., Li, Y., Shen, C., Dick, A., van den Hengel, A. (2013). Contextual hypergraph modeling for salient object detection. *Proceedings of the IEEE International Conference on Computer Vision*. Sydney, Australia.
18. Zhou, H., Chen, M., Rehg, J. M. (2009). Dermoscopic interest point detector and descriptor. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 1318–1321. Boston, MA, USA.
19. Sikorski, J. (2004). Identification of malignant melanoma by wavelet analysis. *Proceedings of Student/Faculty Research Day*, CSIS, Pace University.
20. Uyulan, Ç., Ergüzel, T. T., Tarhan, N. (2019). Elektroensefalografi tabanlı sinyallerin analizinde derin öğrenme algoritmalarının kullanılması. *The Journal of Neurobehavioral Sciences*, 108, 108–124. DOI 10.5455/JNBS.1553607558.
21. Beşer, F., Kizrak, M. A., Bolat, B., Yildirim, T. (2018). Recognition of sign language using capsule networks. *26th Signal Processing and Communications Applications Conference*, pp. 1–4. Izmir, Turkey.
22. Prassanna, J., Rahim, R., Bagyalakshmi, K., Manikandan, R., Patan, R. (2020). Effective use of deep learning and image processing for cancer diagnosis. *Deep learning for cancer diagnosis*, pp. 147–168. Singapore: Springer.
23. Lee, C. Y., Gallagher, P. W., Tu, Z. (2016). Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree. *Artificial Intelligence and Statistics*, 464–472.

24. Nwankpa, C., Ijomah, W., Gachagan, A., Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. arXiv preprint arXiv:1811.03378.
25. Nair, V., Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. *ICML*.
26. Long, J., Shelhamer, E., Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440. Boston, Massachusetts.
27. Noh, H., Hong, S., Han, B. (2015). Learning deconvolution network for semantic segmentation. *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1520–1528. Cambridge, MA, USA.
28. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680. DOI 10.1145/3422622.
29. Baur, C., Albarqouni, S., Navab, N. (2018). Melanogans: High resolution skin lesion synthesis with gans. arXiv preprint arXiv:1804.04338.
30. Kaur, K., Mittal, S. (2020). Classification of mammography image with CNN-RNN based semantic features and extra tree classifier approach using LSTM. *Materials Today: Proceedings*. DOI 10.1016/j.matpr.2020.09.619.
31. Zheng, Y., Yang, C., Wang, H. (2020). Enhancing breast cancer detection with recurrent neural network. *Mobile Multimedia Image Processing, Security, and Applications*, 11399. DOI 10.1117/12.2558817.
32. Budak, Ü., Cömert, Z., Rashid, Z. N., Şengür, A., Çıbuk, M. (2019). Computer-aided diagnosis system combining fcn and bi-ISTM model for efficient breast cancer detection from histopathological images. *Applied Soft Computing*, 85, 105765. DOI 10.1016/j.asoc.2019.105765.
33. Gao, R., Huo, Y., Bao, S., Tang, Y., Antic, S. L. et al. (2019). Distanced LSTM: Time-distanced gates in long short-term memory models for lung cancer detection. *International Workshop on Machine Learning in Medical Imaging*, pp. 310–318. Shenzhen, China, Springer.
34. Anandaraj, A. P. S., Gomathy, V., Punitha, A. A. A., Kumari, D. A., Rani, S. S. et al. (2021). Internet of Medical Things (IOMT) enabled skin lesion detection and classification using optimal segmentation and restricted boltzmann machines. *Cognitive internet of medical things for smart healthcare*, pp. 195–209. Singapore: Springer.
35. Mathappan, N., Soundariya, R., Natarajan, A., Gopalan, S. K. (2020). Bio-medical analysis of breast cancer risk detection based on deep neural network. *International Journal of Medical Engineering and Informatics*, 12(6), 529–541. DOI 10.1504/IJMEI.2020.111027.
36. He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778. San Juan, PR, USA.
37. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1251–1258. San Juan, PR, USA.
38. Simonyan, K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
39. Zhu, W., Xiang, X., Tran, T. D., Hager, G. D., Xie, X. (2018). Adversarial deep structured nets for mass segmentation from mammograms. *IEEE 15th International Symposium on Biomedical Imaging*, pp. 847–850. Washington DC, USA.
40. Albayrak, A., Bilgin, G. (2016). Mitosis detection using convolutional neural network based features. *2016 IEEE 17th International Symposium on Computational Intelligence and Informatics*, pp. 335–340. Budapest, Hungary.
41. Spanhol, F. A., Oliveira, L. S., Petitjean, C., Heutte, L. (2016). Breast cancer histopathological image classification using convolutional neural networks. *2016 International Joint Conference on Neural Networks*, pp. 2560–2567. Vancouver, Canada.
42. Chen, H., Qi, X., Yu, L., Heng, P. A. (2016). DCAN: Deep contour-aware networks for accurate gland segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2487–2496. San Francisco, CA, USA.
43. Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S. et al. (2016). Aggnet: Deep learning from crowds for mitosis detection in breast cancer histology images. *IEEE Transactions on Medical Imaging*, 35(5), 1313–1321. DOI 10.1109/TMI.2016.2528120.

44. Xu, B., Wang, N., Chen, T., Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. arXiv preprint arXiv:1505.00853.
45. Wichakam, I., Vateekul, P. (2016). Combining deep convolutional networks and SVMs for mass detection on digital mammograms. *8th International Conference on Knowledge and Smart Technology*, pp. 239–244. Chiang Mai, Thailand.
46. Swiderski, B., Kurek, J., Osowski, S., Kruk, M., Barhoumi, W. (2017). Deep learning and non-negative matrix factorization in recognition of mammograms. *Eighth International Conference on Graphic and Image Processing*, vol. 10225, pp. 102250B. Tokyo, Japan.
47. Ertosun, M. G., Rubin, D. L. (2015). Probabilistic visual search for masses within mammography images using deep learning. *IEEE International Conference on Bioinformatics and Biomedicine*, pp. 1310–1315. Washington, DC, USA.
48. Kallenberg, M., Petersen, K., Nielsen, M., Ng, A. Y., Diao, P. et al. (2016). Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE Transactions on Medical Imaging*, 35(5), 1322–1331. DOI 10.1109/TMI.2016.2532122.
49. Dhungel, N., Carneiro, G., Bradley, A. P. (2015). Deep structured learning for mass segmentation from mammograms. *2015 IEEE International Conference on Image Processing*, pp. 2950–2954. Québec City, Canada.
50. Dhungel, N., Carneiro, G., Bradley, A. P. (2015). Automated mass detection in mammograms using cascaded deep learning and random forests. *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Adelaide, Australia, pp. 1–8.
51. Kim, D. H., Kim, S. T., Ro, Y. M. (2016). Latent feature representation with 3-D multi-view deep convolutional neural network for bilateral analysis in digital breast tomosynthesis. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 927–931. Shanghai, China.
52. Xing, F., Xie, Y., Yang, L. (2015). An automatic learning-based framework for robust nucleus segmentation. *IEEE Transactions on Medical Imaging*, 35(2), 550–566. DOI 10.1109/TMI.2015.2481436.
53. Wichakam, I., Vateekul, P. (2016). Combining deep convolutional networks and SVMs for mass detection on digital mammograms. *8th International Conference on Knowledge and Smart Technology*, pp. 239–244. Chiang Mai, Thailand.
54. Suzuki, S., Zhang, X., Homma, N., Ichiji, K., Sugita, N. et al. (2016). Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis. *55th Annual Conference of the Society of Instrument and Control Engineers of Japan*, pp. 1382–1386. Tsukuba, Japan.
55. Guo, Y., Gao, Y., Shen, D. (2015). Deformable MR prostate segmentation via deep feature learning and sparse patch matching. *IEEE Transactions on Medical Imaging*, 35(4), 1077–1089. DOI 10.1109/TMI.2015.2508280.
56. Pereira, S., Pinto, A., Alves, V., Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. *IEEE Transactions on Medical Imaging*, 35(5), 1240–1251. DOI 10.1109/TMI.2016.2538465.
57. Zhao, L., Jia, K. (2015). Deep feature learning with discrimination mechanism for brain tumor segmentation and diagnosis. *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 306–309. Adelaide, Australia.
58. Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D. et al. (2017). Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical Image Analysis*, 36, 61–78. DOI 10.1016/j.media.2016.10.004.
59. Liu, R., Hall, L. O., Goldgof, D. B., Zhou, M., Gatenby, R. A. et al. (2016). Exploring deep features from brain tumor magnetic resonance images via transfer learning. *International Joint Conference on Neural Networks*, pp. 235–242. Vancouver, Canada.
60. Ahmed, K. B., Hall, L. O., Goldgof, D. B., Liu, R., Gatenby, R. A. (2017). Fine-tuning convolutional deep features for MRI based brain tumor classification. *Medical Imaging 2017: Computer-Aided Diagnosis*, 10134, 101342E. DOI 10.1117/12.2253982.

61. Zhu, X., Yao, J., Huang, J. (2016). Deep convolutional neural network for survival analysis with pathological images. *IEEE International Conference on Bioinformatics and Biomedicine*, pp. 544–547. Shenzhen, China.
62. Paul, R., Hawkins, S. H., Hall, L. O., Goldgof, D. B., Gillies, R. J. (2016). Combining deep neural network and traditional image features to improve survival prediction accuracy for lung cancer patients from diagnostic CT. *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 2570–2575. Budapest, Hungary.
63. Han, F., Wang, H., Zhang, G., Han, H., Song, B. et al. (2015). Texture feature analysis for computer-aided diagnosis on pulmonary nodules. *Journal of Digital Imaging*, 28(1), 99–115. DOI 10.1007/s10278-014-9718-8.
64. Hussein, S., Gillies, R., Cao, K., Song, Q., Bagci, U. (2017). Tumornet: Lung nodule characterization using multi-view convolutional neural network with Gaussian process. *IEEE 14th International Symposium on Biomedical Imaging*, pp. 1007–1010. Melbourne, Australia.
65. Hua, K. L., Hsu, C. H., Hidayati, S. C., Cheng, W. H., Chen, Y. J. (2015). Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and Therapy*, 8, 2015–2022. DOI 10.2147/OTT.S80733.
66. Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C. et al. (2016). Pulmonary nodule detection in CT images: False positive reduction using multi-view convolutional networks. *IEEE Transactions on Medical Imaging*, 35(5), 1160–1169. DOI 10.1109/TMI.42.
67. Tajbakhsh, N., Suzuki, K. (2017). Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: Mtnns vs. CNNs. *Pattern Recognition*, 63, 476–486. DOI 10.1016/j.patcog.2016.09.029.
68. Kim, B. C., Sung, Y. S., Suk, H. I. (2016). Deep feature learning for pulmonary nodule classification in a lung CT. *2016 4th International Winter Conference on Brain-Computer Interface*, pp. 1–3. Yongpyong Resort, Korea.
69. Dou, Q., Chen, H., Yu, L., Qin, J., Heng, P. A. (2016). Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection. *IEEE Transactions on Biomedical Engineering*, 64(7), 1558–1567. DOI 10.1109/TBME.2016.2613502.
70. Hirayama, K., Tan, J. K., Kim, H. (2016). Extraction of ggo candidate regions from the lidc database using deep learning. *16th International Conference on Control, Automation and Systems*, pp. 1–8. HICO, Gyeongju, Korea.
71. Wang, C., Elazab, A., Wu, J., Hu, Q. (2017). Lung nodule classification using deep feature fusion in chest radiography. *Computerized Medical Imaging and Graphics*, 57, 10–18. DOI 10.1016/j.compmedimag.2016.11.004.
72. Shen, W., Zhou, M., Yang, F., Yu, D., Dong, D. et al. (2017). Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification. *Pattern Recognition*, 61, 663–673. DOI 10.1016/j.patcog.2016.05.029.
73. Pomponiu, V., Nejati, H., Cheung, N. M. (2016). Deepmole: Deep neural networks for skin mole lesion classification. *IEEE International Conference on Image Processing*, pp. 2623–2627. Phoenix, Arizona.
74. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M. et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. DOI 10.1038/nature21056.
75. Masood, A., Al-Jumaily, A., Anam, K. (2015). Self-supervised learning model for skin cancer diagnosis. *7th International IEEE/EMBS Conference on Neural Engineering*, pp. 1012–1015. Montpellier, France.
76. Majtner, T., Yildirim-Yayilgan, S., Hardeberg, J. Y. (2016). Combining deep learning and hand-crafted features for skin lesion classification. *Sixth International Conference on Image Processing Theory, Tools and Applications*, pp. 1–6. Oulu, Finland.
77. Sabbaghi, S., Aldeen, M., Garnavi, R. (2016). A deep bag-of-features model for the classification of melanomas in dermoscopy images. *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1369–1372. Orlando, USA.
78. Demyanov, S., Chakravorty, R., Abedini, M., Halpern, A., Garnavi, R. (2016). Classification of dermoscopy patterns using deep convolutional neural networks. *IEEE 13th International Symposium on Biomedical Imaging*, pp. 364–368. Prague, Czech Republic.

79. Yu, L., Chen, H., Dou, Q., Qin, J., Heng, P. A. (2016). Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, 36(4), 994–1004. DOI 10.1109/TMI.2016.2642839.
80. Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S. M. R., Jafari, M. H. et al. (2016). Melanoma detection by analysis of clinical images using convolutional neural network. *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1373–1376. Orlando, USA.
81. Sabouri, P., GholamHosseini, H. (2016). Lesion border detection using deep learning. *IEEE Congress on Evolutionary Computation*, pp. 1416–1421. Vancouver, Canada.
82. Chandrhasa, M., Vadigeri, V., Salecha, D. (2016). Detection of skin cancer using image processing techniques. *International Journal of Modern Trends in Engineering and Research*, 3(5), 111–114.
83. Mehta, P., Shah, B. (2016). Review on techniques and steps of computer aided skin cancer diagnosis. *Procedia Computer Science*, 85, 309–316. DOI 10.1016/j.procs.2016.05.238.
84. Sumithra, R., Suhil, M., Guru, D. (2015). Segmentation and classification of skin lesions for disease diagnosis. *Procedia Computer Science*, 45, 76–85. DOI 10.1016/j.procs.2015.03.090.
85. He, J., Dong, Q., Yi, S. (2018). Prediction of skin cancer based on convolutional neural network. *International Conference on Mechatronics and Intelligent Robotics*, pp. 1223–1229. Changchun, China.
86. Pham, T. C., Luong, C. M., Visani, M., Hoang, V. D. (2018). Deep CNN and data augmentation for skin lesion classification. *Asian Conference on Intelligent Information and Database Systems*, pp. 573–582. Dong Hoi City, Vietnam.
87. Zhang, X., Wang, S., Liu, J., Tao, C. (2018). Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge. *BMC Medical Informatics and Decision Making*, 18(2), 59. DOI 10.1186/s12911-018-0631-9.
88. Vesal, S., Ravikumar, N., Maier, A. (2018). Skinnet: A deep learning framework for skin lesion segmentation. *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings*, pp. 1–3. Sydney, Australia.
89. Vesal, S., Patil, S. M., Ravikumar, N., Maier, A. K. (2018). A multi-task framework for skin lesion detection and segmentation. In: *OR 2.0 Context-Aware operating theaters, computer assisted robotic endoscopy, clinical image-based procedures, and skin image analysis*, pp. 285–293. Springer.
90. Horie, Y., Yoshio, T., Aoyama, K., Yoshimizu, S., Horiuchi, Y. et al. (2019). Diagnostic outcomes of esophageal cancer by artificial intelligence using convolutional neural networks. *Gastrointestinal Endoscopy*, 89(1), 25–32. DOI 10.1016/j.gie.2018.07.037.
91. Gómez-Martín, I., Moreno, S., Duran, X., Pujol, R. M., Segura, S. (2019). Diagnostic accuracy of non-melanocytic pink flat skin lesions on the legs: Dermoscopic and reflectance confocal microscopy evaluation. *Acta Dermato-Venereologica*, 99(1–2), 33–40. DOI 10.2340/00015555-3029.
92. Mahbod, A., Schaefer, G., Wang, C., Ecker, R., Ellinge, I. (2019). Skin lesion classification using hybrid deep neural networks. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1229–1233. Brighton, UK.
93. Yan, K., Li, C., Wang, X., Li, A., Yuan, Y. et al. (2016). Automatic prostate segmentation on MR images with deep network and graph model. *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 635–638. Florida, USA.
94. Ma, L., Guo, R., Zhang, G., Tade, F., Schuster, D. M. et al. (2017). Automatic segmentation of the prostate on CT images using deep learning and multi-atlas fusion. *Medical Imaging 2017: Image Processing*, 10133, 101332O. DOI 10.1117/12.2255755.
95. Milletari, F., Navab, N., Ahmadi, S. A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. *Fourth International Conference on 3D Vision*, pp. 565–571. Stanford, CA, USA.
96. Kallen, H., Molin, J., Heyden, A., Lundström, C., Åström, K. (2016). Towards grading gleason score using generically trained deep convolutional neural networks. *IEEE 13th International Symposium on Biomedical Imaging*, pp. 1163–1167. Prague, Czech Republic.

97. Gummeson, A., Arvidsson, I., Ohlsson, M., Overgaard, N. C., Krzyzanowska, A. et al. (2017). Automatic gleason grading of h and e stained microscopic prostate images using deep convolutional neural networks. *Medical Imaging 2017: Digital Pathology*, 10140, 101400S. DOI 10.1117/12.2253620.
98. Kwak, J. T., Hewitt, S. M. (2017). Lumen-based detection of prostate cancer via convolutional neural networks. *Medical Imaging 2017: Digital Pathology*, 10140, 1014008. DOI 10.1117/12.2253513.
99. Tian, Z., Liu, L., Fei, B. (2017). Deep convolutional neural network for prostate MR segmentation. *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, 10135, 101351L. DOI 10.1007/s11548-018-1841-4.
100. Cheng, R., Roth, H. R., Lay, N., Lu, L., Turkbey, B. et al. (2017). Automatic MR prostate segmentation by deep learning with holistically-nested networks. *Medical Imaging 2017: Image Processing*, vol. 10133, pp. 101332H. International Society for Optics and Photonics.
101. Song, Y., Zhang, L., Chen, S., Ni, D., Lei, B. et al. (2015). Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. *IEEE Transactions on Biomedical Engineering*, 62(10), 2421–2433. DOI 10.1109/TBME.2015.2430895.
102. Xu, T., Zhang, H., Huang, X., Zhang, S., Metaxas, D. N. (2016). Multimodal deep learning for cervical dysplasia diagnosis. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 115–123. Athens, Greece.
103. Cha, K. H., Hadjiiski, L., Samala, R. K., Chan, H. P., Caoili, E. M. et al. (2016). Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets. *Medical Physics*, 43(4), 1882–1896. DOI 10.1118/1.4944498.
104. Gibson, E., Robu, M. R., Thompson, S., Edwards, P. E., Schneider, C. et al. (2017). Deep residual networks for automatic segmentation of laparoscopic videos of the liver. *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, 10135, 101351M. DOI 10.1117/12.2255975.
105. Wen, L., Jia, F., Hu, Q. et al. (2015). Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. *Journal of Computer and Communications*, 3(11), 146. DOI 10.4236/jcc.2015.311023.
106. Mao, Y., Yin, Z., Schober, J. (2016). A deep convolutional neural network trained on representative samples for circulating tumor cell detection. *IEEE Winter Conference on Applications of Computer Vision*, pp. 1–6. Lake Placid, NY, USA.
107. Song, Y., Cheng, J. Z., Ni, D., Chen, S., Lei, B. et al. (2016). Segmenting overlapping cervical cell in pap smear images. *IEEE 13th International Symposium on Biomedical Imaging*, pp. 1159–1162. Prague, Czech Republic.
108. BenTaieb, A., Kawahara, J., Hamarneh, G. (2016). Multi-loss convolutional networks for gland analysis in microscopy. *IEEE 13th International Symposium on Biomedical Imaging*, pp. 642–645. Prague, Czech Republic.
109. Gordon, M., Hadjiiski, L., Cha, K., Chan, H. P., Samala, R. et al. (2017). Segmentation of inner and outer bladder wall using deep-learning convolutional neural network in CT urography. *Medical Imaging 2017: Computer-Aided Diagnosis*, 10134, 1013402. DOI 10.1117/12.2255528.
110. van Ginneken, B., Armato III, S. G., de Hoop, B., van Amelsvoort-van de Vorst, S., Duindam, T. et al. (2010). Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study. *Medical Image Analysis*, 14(6), 707–722. DOI 10.1016/j.media.2010.05.005.
111. Tajbakhsh, N., Gurudu, S. R., Liang, J. (2015). Automated polyp detection in colonoscopy videos using shape and context information. *IEEE Transactions on Medical Imaging*, 35(2), 630–644. DOI 10.1109/TMI.2015.2487997.
112. Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K. et al. (2014). The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10), 1993–2024. DOI 10.1109/TMI.2014.2377694.
113. Spanhol, F. A., Oliveira, L. S., Petitjean, C., Heutte, L. (2015). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462. DOI 10.1109/TBME.2015.2496264.

114. Sirinukunwattana, K., Raza, S. E. A., Tsang, Y. W., Snead, D. R., Cree, I. A. et al. (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE Transactions on Medical Imaging*, 35(5), 1196–1206. DOI 10.1109/TMI.2016.2525803.
115. Bowyer, K., Kopans, D., Kegelmeyer, W., Moore, R., Sallam, M. et al. (1996). The digital database for screening mammography. *Third International Workshop on Digital Mammography*, 58, 27.
116. Dey, T. K. (2006). *Curve and surface reconstruction: Algorithms with mathematical analysis*. Cambridge University Press. DOI 10.1017/CBO9780511546860.
117. DermQuest (2020). Online medical resource. <http://www.dermquest.com>.
118. Pedersen, J. H., Ashraf, H., Dirksen, A., Bach, K., Hansen, H. et al. (2009). The danish randomized lung cancer CT screening trial-overall design and results of the prevalence round. *Journal of Thoracic Oncology*, 4(5), 608–614. DOI 10.1097/JTO.0b013e3181a0d98f.
119. Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M. et al. (2016). Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC). arXiv preprint arXiv:1605.01397.
120. Shiraishi, J., Katsuragawa, S., Ikezoe, J., Matsumoto, T., Kobayashi, T. et al. (2000). Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules. *American Journal of Roentgenology*, 174(1), 71–74. DOI 10.2214/ajr.174.1.1740071.
121. Armato III, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R. et al. (2011). The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, 38(2), 915–931. DOI 10.1118/1.3528204.
122. Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F. et al. (2015). Med-node: A computer-assisted melanoma diagnosis system using non-dermoscopic images. *Expert Systems with Applications*, 42(19), 6578–6585. DOI 10.1016/j.eswa.2015.04.034.
123. Veta, M., van Diest, P. J., Willems, S. M., Wang, H., Madabhushi, A. et al. (2015). Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Medical Image Analysis*, 20(1), 237–248. DOI 10.1016/j.media.2014.11.010.
124. Litjens, G., Toth, R., van de Ven, W., Hoeks, C., Kerkstra, S. et al. (2014). Evaluation of prostate segmentation algorithms for MRI: The promise12 challenge. *Medical Image Analysis*, 18(2), 359–373. DOI 10.1016/j.media.2013.12.002.
125. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J. et al. (2013). The cancer imaging archive (tcia): Maintaining and operating a public information repository. *Journal of Digital Imaging*, 26(6), 1045–1057. DOI 10.1007/s10278-013-9622-7.
126. Sirinukunwattana, K., Pluim, J. P., Chen, H., Qi, X., Heng, P. A. et al. (2017). Gland segmentation in colon histology images: The glas challenge contest. *Medical Image Analysis*, 35, 489–502. DOI 10.1016/j.media.2016.08.008.
127. Xu, X., Ding, Y., Hu, S. X., Niemier, M., Cong, J. et al. (2018). Scaling for edge inference of deep neural networks. *Nature Electronics*, 1(4), 216–222. DOI 10.1038/s41928-018-0059-3.
128. Horowitz, M. (2014). 1.1 computing's energy problem (and what we can do about it). *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, pp. 10–14. San Francisco, CA, USA.
129. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W. et al. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.
130. Ota, K., Dao, M. S., Mezaris, V., Natale, F. G. D. (2017). Deep learning for mobile multimedia: A survey. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 13(3s), 1–22. DOI 10.1145/3092831.
131. Cheng, Y., Wang, D., Zhou, P., Zhang, T. (2017). A survey of model compression and acceleration for deep neural networks. arXiv preprint arXiv:1710.09282.

132. Cheng, Y., Wang, D., Zhou, P., Zhang, T. (2018). Model compression and acceleration for deep neural networks: The principles, progress, and challenges. *IEEE Signal Processing Magazine*, 35(1), 126–136. DOI 10.1109/MSP.2017.2765695.
133. Wiedemann, S., Müller, K. R., Samek, W. (2019). Compact and computationally efficient representation of deep neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 31(3), 772–785. DOI 10.1109/TNNLS.5962385.