Reports of Practical Oncology and Radiotherapy 25 (2020) 656–666

Available online at www.sciencedirect.com

Reports of Practical Oncology and Radiotherapy

journal homepage: http://www.elsevier.com/locate/rpor

Artificial intelligence in radiotherapy

Sarkar Siddique [a], James C.L. Chow [b,c,*]

[a] Department of Physics, Ryerson University, Toronto, ON M5B 2K3, Canada
[b] Radiation Medicine Program, Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1X6, Canada
[c] Department of Radiation Oncology, University of Toronto, Toronto, ON M5T 1P5, Canada

ARTICLE INFO

Article history:
Received 29 November 2019
Received in revised form 6 January 2020
Accepted 27 March 2020
Available online 6 May 2020

Keywords:
Artificial intelligence
Machine learning
Radiotherapy
Medical imaging

ABSTRACT

Artificial intelligence (AI) has already been implemented widely in the medical field in the recent years. This paper first reviews the background of AI and radiotherapy. Then it explores the basic concepts of different AI algorithms and machine learning methods, such as neural networks, that are available to us today and how they are being implemented in radiotherapy and diagnostic processes, such as medical imaging, treatment planning, patient simulation, quality assurance and radiation dose delivery. It also explores the ongoing research on AI methods that are to be implemented in radiotherapy in the future. The review shows very promising progress and future for AI to be widely used in various areas of radiotherapy. However, basing on various concerns such as availability and security of using big data, and further work on polishing and testing AI algorithms, it is found that we may not ready to use AI primarily in radiotherapy at the moment.

1. Introduction

1.1. Artificial intelligence

The idea of artificial intelligence (AI) is believed to be generated from the idea of robots. The idea becomes more and more prominent with more use of biosynthetic machines used in labour. AI can be defined as capability of a machine of imitating human intelligence. AI can be classified into two branches based on its application: virtual and physical. Physical component can be represented in medical devices, sophisticated robots (care bots) and limited mobility helper bots. The virtual component can be represented in machine learning. Machine learning is a mathematical algorithm that learns through experience.[1] For a physician, the two most important factors for patient care are knowledge and experience. Humans are limited in terms of learning by gathering large amount of data primarily due to time constraints. In the process of human learning, knowledge and experience are both required and gained along a lifetime career. Computer can use algorithms to gain far more experience and store data in significantly shorter amount of time than human. A radiologist will look at approximately 225,000 MRI/CT exams in 40 years, while AI can start with 225,000 scans to train itself and reach millions of scans within a very short period of time.[2] Nowadays, patients demand faster and more personalized patient care, which require physicians to interpret large amount of data and analyse it in a short period of time. Machine learning can aid in these situations by taking data analysis from the physician and provide more efficient, convenient and personalized clinical practice in a shorter time.[3] Massive amount of data is now available to train algorithms and modern computational hardware. These algorithms are being applied in many fields, such as drug discovery, medical diagnostics and imaging, remote patient care, risk management, hospital assistants and virtual assistance. Components that require a large amount of data analysis, such as DNA and RNA are expected to greatly benefit from these computational algorithms.[4] With the introduction of deep learning algorithm, the machine learning capabilities have advanced significantly in the recent years.[5]

1.2. Radiotherapy

Radiotherapy is an important component of cancer treatment and it is estimated that almost 50% of all cancer patients receive radiotherapy during their course of illness.[6] Radiotherapy can be classified into seven sections; imaging, treatment planning (TP), simulation, radiotherapy accessories, radiation delivery, radiotherapy verification, and patient monitoring.[7] Imaging process is the first step, where the physicians diagnose the patient for tumour; if there is a tumour present then important information related to the tumour is collected for later use. Imaging process provides the

* Corresponding author.
E-mail address: james.chow@rmp.uhn.ca (J.C.L. Chow).

https://doi.org/10.1016/j.rpor.2020.03.015
1507-1367/

physician with the estimated tumour volume, shape, location, surrounding organ that are at risk and other useful information which are very important for the radiation dose delivery.[8] More enhanced imaging modalities provide more accurate data, which makes the therapy sharper and more impactful.[9] In radiotherapy, distant metastasis is often a big problem and if not identified correctly in the imaging process, the tumour can re-grow and pose even higher threat in the future.[10] There are many advanced imaging modalities available to us now, they can very precisely take medical image with very little radiation exposure, such as computed tomography (CT), positron emission tomography (PET), and single photon emission tomography (SPECT).[11] There are also some non-radiative options available to us which can also take high resolution image with high accuracy, such as magnetic resonance imaging (MRI) and ultrasound.[12] Further advancement in these modalities have provided us with technology that can take four dimensional (4D) image, which accounts for patient movements.[13] During the TP and simulation process, various patient data such as tumour mass, height, weight, body mass index, previous exposure to radiation, and internal medical image of the patient is obtained and analysed to personalize the therapy by modifying different parameters to obtain the best outcome for the patient. Lot of risk, failure estimation and prediction go on in the TP process. The calculation of organ motion, therapy technique selection, beam intensity shaping, planning margin, and optimisation of treatment planning takes place in the TP process. Recently, TP process has improved significantly to keep up with the advances in the therapy modalities, such as three-dimensional conformal radiotherapy (3D-CRT) or intensity-modulated radiotherapy (IMRT), where dose to the surrounding organs-at-risk is reduced significantly while delivering high radiation dose to the target volume. Both modalities require very precise treatment setup, repeated monitoring and precise location of the target volume, which needs to account for the motion of the target due to internal organ motion and tumour regression.[14] The TP can be classified into many categories, such as knowledge based, expert based or artificial intelligence (AI) based.[15] After the patient simulation phase, treatment planning starts and the dose calculation takes place. Monte Carlo (MC) simulation is generally accepted as the gold standard for dose calculation in radiotherapy, since it provides real physical process involved in the interaction in general body or tissue. Some of the commonly used MC codes are the EGSnrc, PENELOPE, MCNP, and Geant4. These simulations are complicated, time consuming and require a large amount of computing and data storage power. These characteristics are not suitable for routine clinical use so the alternative solution is CloudMC, which is cheaper and undemanding in those aspects.[16] Radiotherapy accessories are often used to immobilize the patient to restrict their movement and have a more accurate estimation of the target volume location.[17] This immobilization procedure is usually carried out in the simulation phase. Radiotherapy process is the process where the radiation dose is delivered to the patient. Dose is energy deposited per unit mass. The main goal of the radiotherapy process is to kill tumour cells using ionizing radiation while sparing or minimizing the energy deposition to healthy cells as much as possible. There are many radiotherapy modalities available now, such as stereotactic body radiotherapy (SBRT), volumetic modulated arc therapy (VMAT), IMRT, proton therapy, electron therapy and brachytherapy.[18] After the treatment plan is created, the plan dosimetry will be verified by a patient specific quality assurance program to ensure an acceptable dose delivery accuracy. When the radiotherapy is performed, the patient is followed for a period of time (months to years) to observe side effects and outcomes.[19]

2. Method

Machine learning is the idea of computer learning to perform a task from studying a set of training examples. Machine learning is generally classified into two different categories: supervised learning and unsupervised learning. In supervised learning the training set contains the data and the correct output, where computer uses both the data and the output (label data) to predict the output of the future data. Input observations are known as the cause and output observations are the effect. The goal of supervised learning is to come up with a functional relationship from training data that generalizes testing data. This relationship can be in the form of equation, or numerical coefficient. There are various algorithms developed based on supervised learning, such as regression algorithm, classification and reinforcement learning.[20] In unsupervised learning, training set does not contain the solution, so the computer must find the solution on its own and use both the data and the derived solution to predict the outcome of the future data. The goal of unsupervised learning is to come up with the unknown variable behind the observation or find the relationship between samples. There are numerous algorithms that have been derived based on unsupervised learning, such as dimensionality reduction algorithms, clustering, blind source separation, and density estimation.[21,22] Semi-supervised learning is the combination of supervised and unsupervised learning methods. In semi-supervised learning, a training set contains data with solutions and data without solutions, and the method utilizes both labelled and unlabelled data to predict future outcome. This method reduces the need for labelled data, and labelling of data is often very expensive and not always possible. Some examples of semi-supervised learning are information recommendation system and semi-supervised classification.[23,24]

In-depth look at the various machine learning algorithms and their applications to radiotherapy and medical imaging is shown below:

2.1. Linear model for classification and regression

Linear model is based on a linear relationship between input and the output of the model. Linear model is generally used in applications that have computer-aided classification in research and development on radiotherapy.

Linear discriminant analysis (LDA) is a method to extract discriminant properties in pattern classification. LDA requires label data to learn a discriminant projection which significantly enlarges the distance between classes and reduce the distance within the class (distance refers to the order of the variable). LDA ultimately improves the classification accuracy. Various extensions of LDA have been developed to satisfy various special needs, such as two-dimensional linear discriminant analysis for small sample size which improves the performance and efficiency.[25]

Complex relationship often cannot be classified using LDA, especially when there is a quadratic relationship involved. Quadratic discriminant analysis (QDA) is used in such a case. QDA is very similar to LDA and an extension of LDA. QDA captures the relationship between independent and dependent variable and provides a more powerful discriminant tool. QDA is optimal when the predictor vector is normally distributed in both groups and each group has a different covariance structure. One of the drawbacks of QDA is that it estimates more parameters than LDA; however, it outperforms LDA in most cases.[26]

2.2. Artificial neutral network

Artificial neural network (ANN) has been derived from the fundamental idea of how our brain works. ANN is composed of nodes and interconnections, where nodes have limited computational power but they work as a triggering mechanism similar to how neurotransmitters work and accumulate to trigger a neuron; the density and complexity of the interconnector determines the network's computational power.[24] The relationship between the nodes and interconnectors (neurons) can be portrayed as an equation:

$$y_i = f_i(\sum_{j-1}^{n} w_{ij}x_j - \theta_i$$

where y_i is the output of node i, w_{ij} is the connection between the node i and j, x_j is the j^{th} input of the node, and θ_i is the threshold of the node. ANN can be divided into two classes according to their connectivity: feedforwarding and recurrent classes. In forward feeding nodes there is no connection from a node with a larger number to a node with a smaller number, otherwise it is a recurrent class. The order of the ANN is determined by the orders of the nodes in the ANN, for example higher order ANN contains a larger number of higher order nodes.[27] A few of the major advantages of ANN are that it can solve problems that are not linearly separable and can solve many types of problems. ANN can be used to predict the secondary structure of protein due to its ability to analyse a complex structure. There are many different types of neural network, such as Boltzmann machines and Kohonen networks.[28]

2.3. Kernels

Kernel method provides a powerful tool for data analysis by applying both the supervised and unsupervised method. In short, kernel corresponds to a dot product, generally in a high dimensional feature space. It measures the similarity, which allows us to construct algorithms in dot product space. For a given data set, $(x_1 + y_1), \ldots, (x_n, y_n)\, \varepsilon$ X x Y, where domain X are the inputs or predictor variables and Y are the targets or response variables, with given new input $x\, \varepsilon$ X, corresponding $y \varepsilon \{\pm 1\}$ needs to be predicted, then the kernel equals to:

$$K = X \times X \to R, \ \left(x, x'\right) K \ \left(x, x'\right)$$

$$K\left(x, x'\right) = \left(\theta\,(\mathrm{x}), \ \theta\,\left(\mathrm{x}'\right)\right)$$

whereθ maps into a dot product space (feature space), and K is the measure of kernel or similarity measure.[29] A set of supervised based kernel learning method is support vector machine (SVM) that can be used for classification and regression. SVM minimizes the empirical classification error and maximizes the geometric margin in the training set which gives a higher generalization ability on the new sample.[24] The commonly used kernel functions are polynomial, linear and Gaussian radial based. General kernel describes the point-to-point similarities, systematic and is known for its use in hyperspectral image classification.[30]

2.4. Probabilistic models

A probabilistic model allows us to predict a future event using present observation. The model uses probability distribution to represent all the uncertain unobserved quantities and explain how they relate to the data. Basic rules of probability theory are used to find the unobserved data from the observed data. This method of learning data from the probability theory is called Bayes rule. Bayes rule is the combination of the sum rule and product rule of the probability theory:

$$P\,(y|\,x) = \frac{P\,(x\,|y\,)\,P\,(y)}{\sum_{y\,\in\,\mathrm{Y}} P\,(x,\,y)}$$

where x and y are sets of observed and unknown quantities, $P(y)$ is probability of y occurring, $P(x,y)$ is combined probability of x and y occurring and $P(x|y)$ is the probability of x occurring conditioned on observing the value of y. Ultimately, the Bayes rule is,

$$P\,(m|D) = \frac{P\,(D|m)\,P\,(m)}{P\,(D)}$$

where m can be used to compare models and it is the condition term, and D is the observed data. Still, the graphical models are the most popular probabilistic models.[31]

2.5. Ensemble learning

Ensemble learning is another method of machine learning that works by training the learners by increasing the diversity of the ensemble classifier system. One of the advantages of ensemble learning is that the ensemble classifier can reach a high success ratio compared to a single classifier. That is why ensemble learning is widely used for image classification.[32] It is also a very effective mechanism to measure the uncertainty of space of statistical predictors. It can provide an accurate prediction by combining the estimates. The predictive models are trained using different subsets of the training data, this data is selected appropriately to ensure an appropriate amount of variability in predictive models. Bootstrap aggregating or bagging is a well-known selective method. It creates variability between predictors by sampling with replacement from the set of training data. One drawback of this method is the random selection of a subject in each training subset that can lead to large difference in predictor due to inter-subject difference.[33]

2.6. Cluster analysis

Clustering data is widely and very frequently used in radiotherapy and radiology. Natural data generally has an inherent clustering property and data samples belonging to the same cluster are similar or would have lower a distance than the samples from different cluster under distance matrices. Some of the commonly used clustering algorithms are hierarchical clustering, K-means clustering, DBSCAN, mixtures of Gaussians and normalized cut.

2.6.1. K-means algorithm

K-means algorithm uses parameter k, which divides n number of objects into clusters by having high similarities in clusters and low similarities between clusters, ultimately lowering the total distance between values in each cluster to the centroid. Centroid is an average value of a cluster. The reciprocating Euclidean distance is used to measure the similarity, where the lower the distance, the higher the similarities and vice versa. Euclidean distance can be calculated using the equation:[34]

$$d\left(x_j, c_j\right) = \sqrt{\sum_{j=1}^{n} \left(x_j - c_j\right)^2}.$$

2.6.2. DBSCAN

DBSCAN algorithm is a density-based clustering algorithm, where clusters are a dense region of the object and object of one cluster must be surrounded by objects of the same cluster. Some of the core concepts of DBSCAN are Eps-neighbour, which is a d-dimensional hypersphere with object p being its core and Eps being its radius. Other common concepts are core point, which is a given object P with more than minPts data points; and direct density-reachable, where object p is directly density-reachable, when p belongs to the Eps neighbourhood of a core point q given the values of Eps and minPts. Ultimately, DBSCAN checks Eps-neighbourhood of each data point in dimensional hyperspace and searches for the core point of the cluster and the objects which are directly reachable from the core point.[35]

2.6.3. Gaussian mixture model

Gaussian mixture model is used as a tool for model-based clustering and used in many fields such as pattern recognition, image analysis, data mining, and machine learning. Gaussian mixture model with k components can be modelled using:

$$f\left(y_i;\theta\right)=\sum_{j=1}^{k}\pi_j\theta^{p}\left(y_i;\mu_j\right),\sum j$$

where π_j are the positive weights subject to $\sum_{j=1}^{k}\pi_j=1$, and $\mu_j,\sum j,$ are the parameters of Gaussian components.[36] Volumetric analysis of brain necrosis in proton beam therapy is currently being studied using a mixture model.[37]

2.7. Dimensionality reduction and feature selection

To keep up with the latest technology in radiotherapy, AI is required to compute additional variables. Input of additional variables is problematic in AI-based machine learning classification problem due to a longer training time, more noise, larger number of training data required, overfitting and lower model interpretability. Dimensionality reduction reduces the number of input variables under consideration, which reduces most of the side effects of having a large number of initial variables. Dimensionality reduction can be divided into two categories: variable selection and feature extraction.[38]

2.7.1. Variable selection method

Variable selection or feature selection method works by figuring out and taking the best subset of original input variables. Feature selection method does not alter the original representation of the data. Variable selection method is preferred over feature extraction method for the diagnostic model because in the diagnostic model, variables are more useful when they are not transformed to non-interpretable features. Feature selection methods are classified into three categories: Filter, wrapper and embedded technique. Filter extracts the features from the data without any learning involved. Wrapper uses various learning techniques to evaluate which features are useful. Embedded techniques combine the feature selection step and the classifier construction. Some examples of commonly used variable selection methods in radiotherapy are the Pearson correlation coefficient based variable selection, minimum redundancy maximum relevance algorithm (mRMR), and the Markov blanket .[39]

2.7.1.1. Pearson correlation coefficient. Pearson correlation coefficient measures the strength of linear association between two variables.[40] It is one of the simplest approaches for selecting significant variables from independent input variables and target variables. Pearson correlation coefficient can be measured using the formula:

$$r=\frac{\sum(R_i-R_{av})(G_i-R_{av})}{\sqrt{\sum(R_i-R_{av})^2\sum(G_i-G_{av})^2}}$$

where R_i is the first variable, R_{av} is the arithmetic mean and G_i, G_{av} are corresponding second variable and arithmetic mean.[41]

2.7.1.2. The Markov blanket. The Makarov blanket (MB) was first named by Pear and represents a crucial part in the Bayesian network. MB of a variable consists of parent (direct cause), children (direct effects), and spouses (other direct causes of variable children) in a Bayesian network. One unique property of MB is that, given a target node, all other nodes are independent of the target node.[42] The MB of a variable gives a complete picture of local causal structure around the variable. By figuring out the MB of all variables in the data set, the MBs can be used as constraints to reduce the search spaces in a large scale Bayesian network.[43]

2.7.1.3. Minimum redundancy maximum relevance algorithm. Wang and Peng proposed minimum redundancy maximum relevance algorithm (mRMR), which is a variable selection approach. It tries to select features with high correlation in the output class and low correlation between the classes, and provides a measure of how potentially useful a candidate feature may be (scoring criterion) when used in a classifier model using the Battiti's proposed formula. It applies Peng's condition to avoid the need of a pre-defined threshold. The mRMR can be defined as:

$$f(x_i)=\mathrm{I}(\mathrm{X_i};\mathrm{Y})-\frac{1}{|S|}\sum_{\mathrm{X_j}\in\mathrm{S}}\mathrm{I}(\mathrm{X_i};\mathrm{X_j})$$

where f estimates the scoring criterion, $\mathrm{X_i}$ is the candidate feature, Y is the class attribute, and $\frac{1}{|S|}$ is the adjust of the subtraction comparability of the relevance and redundancy term.[44] In the case of a continuous feature, the correlation can be calculated using F-statistics and the correlation between classes (redundancy) can be calculated using Pearson's correlation. The features are selected one by one by applying a greedy search, which maximizes the objective function (function of relevance and redundancy). For the case of temporal data, the mRMR first needs the data to be flattened into a single matrix using preprocessing techniques. It also has a drawback due to the possibility of important data loss during the flattening process.[45]

2.7.2. Feature extraction method

Feature extraction methods are classified into two categories: linear feature extraction and non-linear feature extraction.

2.7.2.1. Linear feature extraction (principal component analysis). Linear feature works when the data lies on a lower dimensional linear subspace. It projects the data on the subspace using matrix factorization. In 1901, Karl Pearson proposed the principal component analysis (PCA) method, which is also known as Karhunen–Loeve (K–L) method. It is one of the well-known dimensionality reduction algorithms. PCA transforms high-dimensional feature space into low-dimensional feature space using orthogonal transformation, where the dimension of the reduced feature space may be equal to or less than the original feature space. The transformation of the feature space carried out in a way that the highest variance lies in the first component, then the next highest variance lies in the second component and so on.[46] In PCA, the first computed variables are called principal components which are a linear combination of the original variables. First principal components have the largest possible variance. Second components are computed under the constraint of being orthogonal to the first components and have the largest possible inertia. The values of the new variables for the observations are called factor scores and can be computed as:

$$F = P\Delta = P\Delta Q^T Q = XQ$$

where F is the factor score or projection matrix, X is the data set, Q is the coefficient of linear combinations, Δ is the diagonal matrix of singular values.[47]

2.7.2.2. Nonlinear feature extraction. Nonlinear dimensionality reduction can be done using the linear dimensionality technique, but a few tricks need to be applied first. Low-dimensional surface needs to be mapped on a high-dimensional space so that a nonlinear relationship exists among them, which can easily be detected. Kernel function can be designed to create the same effect without the need to explicitly compute the lifting function. Another approach to solve the nonlinearity problem is to assume that the data line on an embedded nonlinear manifold has lower dimension than the raw data space and lies within it. The isomap map is a commonly used algorithm that constructs the manifold by joining neighbours in a form of map. The distance between the points is the geodesic distance on the output graph. Other manifold based algorithms that are used are Laplacian eigenmaps and locally linear embedding.[39]

2.8. Reinforcement learning

Reinforcement learning provides tools to optimize a sequence of decisions for long-term outcomes. Reinforcement learning algorithms generally take an input sequence of interactions known as histories between the decision maker and their environment. On each decision point, the algorithm chooses an action according to its policy and receives new observations and immediate outcomes (reward). It is used in many healthcare optimization such as antiretroviral therapy in HIV.[48] One of the most common reinforcement learning algorithms is Q-learning algorithm based on instantaneous strategy and irrelevant model. One of the main advantages of this algorithm is that it does not need prior knowledge. It interacts with the environment directly and obtains information from the state feedback in the process of control.[49] Current advances in AI and reinforcement learning provide a deep Q-network that can learn successful policies directly from the high dimensional sensory inputs using end-to-end reinforcement learning and can reach human level control.[50]

2.9. Multiple instance learning

Multiple instance learning (MIL) was first introduced by Dietterich in 1997 and it was mostly used for identification of proteins and content-based image retrieval. MIL deals with the uncertainty in the label. Several methods are available to solve MIL problems, some common examples are diverse density, kernel based SVM and K-nearest (K-NN) algorithm.[51] In MIL, a bag is used to represent an object and can be modelled as:

$$B_i = \left\{ x_{il} | l = 1, 2, \ldots\ldots, n_i \right\} \in R^d$$

where n_i is the number of feature vector in bag B_i, and R^d is the d-dimensional feature space. The training data set for MIL can be represented as:

$$T = \left\{ (B_i, y_i) | i = 1, 2, \ldots\ldots, N \right\}$$

where $y_i = (+1, -1)$ and it is the class label corresponding to each bag in the dataset. If the bag contains at least one positive instance, the bags are labeled +1, otherwise it is -1 or a negative labelled bag.[52]

2.10. Graph matching

Graphs are commonly used as an abstract representation of a complex structure. Graph matching is finding a correspondence between nodes of two graphs in a way that they look most similar when their vertices are labelled according to the correspondence. Nodes represent local features of the image, where edges correspond to a relational aspect between images. There are many models that approach graph matching in different ways, some examples of different methods are spectral methods, relaxation labelling, probabilistic approaches, semi-definite relaxations, graduated assignment, tree search, and replicated equations. Spectral method works by studying the similarities between the spectra of the adjacency or Laplacian matrices of the graphs and use that information to match them. Probabilistic and relaxation methods first define a probability distribution over mappings and optimize using discrete relaxation algorithms. Tree search runs sequential tests for compatibility of local parts of the graphs.[53] Integer-projected fixed-point algorithm is another commonly used graph matching algorithm used widely for image segmentation and has a high efficiency. Graph matching is an essential component of 2D and 3D feature matching and object recognition.[54]

2.11. Deep learning

Deep learning computational power has a high potential and it has already led to the rise of many new research companies in a start-up setting in the recent years, such as Deep Genomics. The deep learning method uses multiple processing layers to discover pattern in a large data set.[55] Deep learning is a branch of AI and a set of computational models composed of multiple layers of data processing. During the training phase, the system computes the error between the observed output and desired output and adjusts its internal parameters known as weights to reduce this error. The system also computes a gradient vector for each weight, which indicates the error deviation due to weight adjustment. The weight vector is adjusted in the opposite direction of the gradient vector. Generally, a method called stochastic gradient descent (SGD) is used to find an effective set of weights. In this process, the system is shown the input vector for a few examples and have it computed the output, error and the average gradient for those examples. The whole process is repeated multiple times with a small set of examples until the average of the object function stops decreasing.[56] Many deep learning-based networks have been developed and applied recently in radiotherapy for complex computation such as deep neural network and convolutional neural network.

2.11.1. Deep neural networks

Deep neural networks (DNNs) can be supervised or unsupervised. In machine learning, it is often necessary to reduce the complexity of the data and highlight the relevant pattern for the learning algorithm. In most cases, the AI function greatly depends on how effectively the learning algorithm function performs.[57] DNNs independently learn the order representation of the input data and require a large set of them. DNNs are a combination of multiple neural networks or multilayered ANNs. DNNs use the same principles of ANN to compute the data using nodes at first, then the output from the first layer becomes the input of the following layer and the process repeats, where final layer's output is the derived output for the system.[58] In unsupervised DNNs, the fundamentals are the same as unsupervised learning and ANN, so in the first layer ANN output is generated using unsupervised learning principle, then the outcome from the first layer serves as the input for the second layer. Unsupervised deep learning algorithms are more preferred over supervised DNNs due to the need of lower labelled data.

One drawback of unsupervised DNNs is that it is more complex to guarantee the learned representation will be meaningful.[59]

2.11.2. Convolutional neural networks

Convolutional neural networks (CNNs) are suitable to process data that comes in arrays, such as medical image. Typically, CNN consists of three types of layers: convolutional layer, pooling layer and fully connected layer. The role of convolutional layer is to detect similarities of the feature from the previous layer and learn feature representation of the input. Convolution layer number corresponds to the number of convolution kernels that are used to compute feature maps. In the feature map each neuron is connected to a neighbouring neuron on the previous layer. The role of the pooling layer is to achieve shift variance by reducing the resolution of the feature map. A pooling layer is generally placed between two convolutional layers, where each future map of the layer is connected to a corresponding feature map of the previous convolutional layer. The role of a fully connected layer is to perform high level reasoning by taking all neurons in the previous layer and connecting them to every single neuron of the current layer.[60]

3. Implementation of machine learning algorithms in radiotherapy

3.1. Medical imaging

Various machine learning methods are being implemented in modern medical imaging systems, such as computed tomography (CT) or magnetic resonance imaging (MRI). CT is a radiological imaging modality that can obtain volumetric, morphological information of the patient's anatomy through a medical image. A modern CT scanner can scan a patient from head to toe, with high resolution.[61] Another popular medical imaging system is MRI. MRI can differentiate soft tissue very well which makes it ideal to see inside a joint or ligament; however, MRI can be used to obtain image almost anywhere on the body, if it involves soft tissue density difference.[62] MRI obtains the image by aligning the spin moments of water protons (H) and can obtain better image than CT without the risk of radiation exposure; however, MRI cannot be used in the presence of metal such as a metal implant because it will interfere with the magnetic field of MRI.[63–65] Each of these imaging systems have its disadvantages, like radiation exposure, sensitivity to its surroundings and cost, whereas AI-based algorithm offers cheap non-radiative solutions and it has been widely researched in the recent years. AI-based algorithm does not come without any risk, they increase the risk of systematic errors with high consequences.[66] The accuracy of the radiation dose delivery to the target area is significantly improved with the advances in computing technology and medical imaging; which has led to the integration of MRI and linear accelerator that can provide real time MRI-guided radiotherapy.[67]

AI-based algorithms are primarily implemented in three aspects of medical imaging as shown below:

3.1.1. Image segmentation

Image segmentation plays a vital role in medical image analysis. There are a lot of segmentation methods available but none is universal. Some commonly used image segmentation methods are the snake model introduced by Kass et al. and the level set method (LMS). LMS can further be divided into two categories: region-based models and edge-based models. Edge-based model utilizes edge information, where region-based model utilizes region information to control the motion of the active contour. Region-based models are not sensitive to object with poorly defined boundaries; however, they are sensitive to inhomogeneity of image intensities (overlapping of intensity ranges), and parameter turning. Some other approaches to image segmentation are *K*-NN, SVM and extreme machine learning. They can analyse complex patterns but require post processing such as morphological operations.[68]

Segmentation can be applied to many structures, such as bones, organs, muscles and fractures. In recent research, tree-based segmentation method has been studied intensively for brain imaging. Neural networks and deep learning algorithms are also being used for the brain segmentation task. Together they provide a semi-automatic method to classify the brain MRI into cancer cells and healthy tissue. Early brain has lower tissue contrast, which makes it more difficult to segment than the adult brain. Fractional anisotropy image is more suitable for differentiating grey and white matter; while T2 weighted image is better at capturing cerebrospinal fluid. CNN method can combine these different modality's image data to enhance segmentation predominance. The neural network normally requires a convolutional layer to perform this task. Deep learning methods can automatically segment MRI brain images.[69] There are other segmentation methods available for brain MRI, such as intensity-based methods (clustering, thresholding reign growing, classification), atlas-based methods, surface-based methods (contours and surfaces and multiphase active contours) and hybrid segmentation. Most of these methods utilized previously explored algorithms and techniques, such as the Markov random field, *k*-point, regression and clustering.[70] New CNN-based interactive segmentation framework only needs the tumour core in one MR sequence for training and provides a more robust segmentation method than the other state-of-the-art regular CNN method. This method provides accurate result with fewer interaction and less time.[71]

Breast cancer is one of the commonly diagnosed cancers, responsible for 30% of all new cancer diagnosis in women. Ultrasound is normally used to diagnose breast cancer and improvement in segmentation of breast ultrasound images into functional tissues provides a better tumour localization, assessment of treatment response, and breast density measurement. Segmentation of ultrasound is very time consuming for radiologist and it is skill and experience dependent. Automated segmentation of ultrasound image will help mitigate those problems. Recent study shows convolutional CNN-based segmentation can segment the 3D image into four major tissues: skin, mass, fibro glandular tissue, and fatty tissue with high accuracy. This shows potential to provide segmentation in the future in clinical diagnosis of breast cancer.[72]

A fully automated whole-body segmentation for diagnostic CT has already been proposed. The segmentation method used random forest algorithms and explored its accuracy and limitation. Tissue segmentation of CT scans was done by training various data sets and applying them to neck, chest, pelvis and abdomen CT scans.[73] Another fully automated diagnostic system is also presented for abdomen disease that uses a feature extraction method to classify the disease, and uses genetic algorithms, SVM and ANN to classify the disease from CT images. Regions of interest segmented from the CT images were tumour, calculi, cyst, and normal liver cell.[74]

3.1.2. Medical image registration

Image registration is an application of machine learning. In the case when a patient is scanned by different or same modality multiple times from different positions, at different times or situations, then machine learning algorithms are used to combine the results to obtain a more accurate diagnosis. This process is often known as image fusion, matching or warping. The goal of this process is to find the optimal transformation that best aligns the structures of interest in the input images. Image registration is a very crucial step in image analysis. Few examples of commonly used image fusion modalities are CT or MRI with PET or SPECT. In the intensity-based registration method, the algorithm searches for geometric transformation iteratively so that when applied to the moving image it optimizes (maximizes or minimizes the similarity measure or cost function). Cost function is related to voxel intensity and computed in the overlapped region of the input image. In the feature-based registration method, the algorithm searches for optimal transformation after the features are established. Criterion based on geometrical, physical or statistical properties is used to match among features.[75] LDA is used to classify lung lesion using PET and CT radiomics features.[76]

3.1.3. Computer-aided detection (CAD) and diagnosis system

Automated image recognition has improved significantly in the recent years primarily due the availability of largescale data sets. CT lung node identification is an improved example in CAD. ImageNet consists of more than 1.2 million categorized natural images of over 1,000 classes, which play a big role in training this complex CNN. CAD-based CNN trained from ImageNet on thoraco-abdominal lymph node and interstitial lung disease shows promising results for clinical use.[77] CAD also shows a potential in detection of metastases, and colonic polyps.[78] Computer-aided detection and diagnosis (CADe and CADx) for colonoscopy use deep learning algorithms that are currently being studied. This AI has two principal roles in diagnosis: Polyp detection (CADe) and polyp characterization (CADx). CADe will decrease the polyp miss rate, which will lead to a better adenoma detection. CADx will improve the accuracy of colorectal polyp optical diagnosis.[79]

Prostate cancer is one of the leading causes of cancer death for men in western side of the world. Currently a transrectal ultrasound biopsy is performed to diagnose prostate cancer. One drawback of this diagnosis is that it has low specificity, so, recently, MRI has been increasingly used for prostate cancer. However, it requires a very expert radiologist team to detect prostate cancer from an MRI image with a lot of time. CAD can solve these problems. A new CAD system is being studied right now, which is composed of two stages. In the first stage, initial candidates are detected using multi-atlas based prostate segmentation, classification, local maxima detection and feature extraction. In the second stage, cancer probability of each filtered candidate is predicted using classification. The system was evaluated with 347 patients with MR-guided biopsy as the reference standard. The evaluation result showed that the system could potentially be used as a second reader or to improve the sensitivity of the radiologist. The system also shows a potential for first reader settings; however, more works need to be done for that to be operating effectively.[80]

Breast cancer has shown a significant survival rate if detected at early stage. In the United States alone nearly 40 million mammographic exams are performed per year. The mammographic image needs to be interpreted and analysed by one or more experienced readers, which is very time consuming and error-prone. CAD is being studied and employed widely as a second reader to increase efficiency. Computer does not suffer from concentration volume. They can also be trained with incredible number of samples, more than any radiologist can learn within a lifetime. A convolutional neural network-based CAD system can do pattern recognition and object detection with high accuracy. It was put to test against a manually designed feature set with 45,000 images and CNN outperforms at low sensitivity and shows comparable results at high sensitivity. The CNN network was also compared against a certified screening radiologist on a patch lever and showed no significant difference.[81]

3.2. Treatment planning

Low dose rate (LDR) prostate brachytherapy is one of the best treatment methods for prostate cancer. Several trials have proven that permanent implantation of low dose radioactive sources provides excellent result. The main goal of the therapy is to maximize the dose to the prostate while sparing normal tissue by implanting radioactive seeds in a 3D pattern. The implementation of those seeds and the precise location of the seeds are very important and require an expert brachytherapy team. Since error in systematic source placement will result in insufficient dose to the target, the TP process needs to be very carefully crafted and requires a lot of time. Computer optimized algorithms for TP in LDR brachytherapy was introduced to improve the efficiency and lighten the workload, some examples of the algorithms that are used are inverse planned simulated annealing (ISPA), and hybrid inverse planning optimization (HIOP). The CAD uses similarity matching and *K*-nearest neighbour methods, where they try to match the current situation to the most similar previous training data. CAD has proved to be able to make accurate treatment plans for LDR brachytherapy that are in par with those made by the expert treatment planer and radiologist.[82] Bayesian network and Makarov model-based decision aid were developed for intensity-modulated plan selection in a prostate cancer patient. The aid used Bayesian network in TP for local tumour control, regional spread, distant metastases, and normal tissue spearing; while Makarov method was used to calculate quality adjusted life expectancy. By combining the two outcomes the aid would provide a treatment decision.[83] *K*-mean clustering can be used as a classifier for adaptive radiotherapy in prostate cancer. The system compares delivered and planned radiotherapy in the patients and automatically identifies those that can benefit from adaptive treatment. The main goal of the system is to deal with dosimetry uncertainties due to the movement of hollow organs. *K*-means clustering algorithm is used to make an unsupervised predictive tool which detects incorrect setup due to stochastic physiological changes.[84] CNN-based automated method for predicting dosimetry eligibility of patients with prostate cancer undergoing IMRT has shown reasonably accurate results and showed its potential in the TP process.[85] In the study of Coates et al., QDA is used in prostate radiotherapy. QDA is used to predict various outcome parameters for both tumour control and radiotherapy induced normal tissue effect. Experimental results showed that QDA prediction matches expected outcome accurately for low dose and starts to deviate significantly over 6-8 Gy; however, most modern radiotherapy treatment regimens use 2-3 Gy. A modified QDA can be used to deal with other parameters that start to deviate from the expected result.[86]

Recent radiotherapy modalities such as photon-based VMAT require a lot of planning before dose delivery. The dose deposition in VMAT is very complex and an accurate prediction of the plan outcome allows radiation oncologists to make a better and more informed decision for therapy and saves a lot of time. New proposed machine learning algorithm can predict dose distribution for organs-at-risk and planning target volume. The algorithm's accuracy was validated on 69 plans for lung SBRT and 121 head-and-neck plans; this resulted in a mean error below 2.5 Gy. This shows a potential to be used as automated treatment plan in SBRT for lung and head-and-neck therapy.[87] In Cho et al., ANN shows outcome prediction capabilities for head-and-neck cancer. ANN combines relevant variables into a predictive model during training and analyses all possible correlation of variables. Out of 73 test subjects, 51 patients were used for the training set, 11 patients were used for the test set and the remaining 11 patients were used for the validation set. The result shows that for focal target control the accuracy for all combined sets is 90.4% and distant metastasis outcome accuracy is 91.8%, proving its viability as a prediction tool.[88] ANN also allows the prediction of survival of radiotherapy alone from uterine cervical cancer by evaluating important prognostic factors. ANN combines the histological grading of radiation effect from periodic biopsy examination, and the additional fundamental factors and provides an accurate prediction.[89] CNN-based rectal dose toxicity prediction model can serve as a practical powerful tool for rectal dose and induced toxicity analysis for high dose rate brachytherapy in cervical cancer.[90] ANN is also used to analyse prognostic factors related to radiation pneumonitis in patients with lung cancer.[91] DNNs can be used to quantitatively classify imaging data and integrate them with clinical risk factors to predict local failure after stereotactic body radiotherapy.[92]

Radiobiological effectiveness (RBE) is an important quantity to describe the effectiveness of cell killing by radiation. In the experiment of Friedrich et al., the relationship between RBE and linear energy transfer (LET) is analysed using the survival curves, which has been parameterized using linear quadratic (LQ) model. The RBE values derived from that model show the dependence of RBE on LET[24,93]. QDA can be used to predict the optimal position to efficiently spare the organs-at-risk in left-sided whole breast radiotherapy.[94] Linear regression algorithm is used to predict various parameters in TP, an example would be the linear regression model used to predict patient-specific skeletal spongiosa volume in molecular radiotherapy dosimetry.[95] Linear regression model is also used to study predictors of cardiac and lung dose to organs-at-risk in the deep inspiration breath hold method for left breast cancer treatment.[96]

A kernel-based optimization tool was developed to find the effective composition of carbon beams with input fluence and energies to deliver desired depth dose distribution over a spread-out Bragg peak region in carbon radiotherapy. The study shows it is an effective tool to use in the TP process which could ultimately substitute Monte Carlo (MC) simulation.[97]

3.3. Simulation

A hybrid framework of electron dose point kernel method was proposed to estimate the dose distribution around gold nanoparticles (GNP). GNP are used for dose enhancing in radiotherapy. This is a hybrid computational framework combined with the Geant4 MC simulator, which halves the calculation time when compared with the full Geant4 MC simulator ultimately lowering the cost and time efficiency.[98]

In prostate cancer, detection of intestinal gas is very important in image-guided radiotherapy. Miura et al. show that a deep convolutional neural network can detect intestinal gas in the pelvis region very effectively.[99]

3.4. Radiotherapy delivery

In microbeam radiotherapy (MRT), the treatment field is fractionated into arrays of a few tens of a micrometre wide planar beam with high peak doses that are separated by a low dose region. MRT has proven to spare normal tissue more efficiently than general radiotherapy. The dose calculation in MRT is based on MC simulations, which are time consuming. So, Debus et al. provided a kernel-based dose calculation algorithm which separates the photon and electron mediated energy transport and can calculate the valley and peak dose of MRT field within a few minutes. The peak dose value matched the MC simulation within 4% deviation and valley dose within 8%, except for the region close to the material interfaces.[100]

Kernel method provides an inexpensive computational solution to markerless tracking of respiration induced tumour motion in kilovoltage fluoroscopy image sequence in image-guided radiotherapy. The method first enhances the contrast of kilovoltage fluoroscopic image using histogram equalization, then the target tracking is formulated by maximizing the Bhattacharyya coefficient using the mean shift algorithm. The obtained result was compared with four clinical kilovoltage fluoroscopic image sequences and four conventional template matching methods. The kernel method proved superior to the conventional template matching method, showing comparable result to the fluoroscopic image sequence.[101] Markerless prostate localization strategy using DNNs to interpret projection x-ray images in image-guided radiotherapy has been investigated and the experimental result shows high accuracy and can be used for patient positioning and real-time target tracking.[102]

In IMRT the optimized beam angle typically clusters around in a distinct orientation, so a *K*-means algorithm is used to identify cluster centroids as irradiation angle of an IMRT treatment plan. The optimized beam angles provide better sparing of organs-at-risk in the case of pancreas and intracranial cancer.[103] In the development of radiation track structure clustering algorithm, a cluster-based analysis can be used to cluster the ionization events in each cell into two categories; simple or complex double-strand break, which can later be used to identify the RBE relationship.[104]

In carbon radiotherapy the energy deposition due to clustering fragments produced from the main beam can be estimated using the DBSCAN clustering algorithm. Each energy deposition cluster from each individual fragment can be individually estimated using this algorithm.[105] DBSCAN algorithm is also used to study the DNA cluster damage after the irradiation of fibroblast cell nucleus.[106]

Deep learning method is being studied to predict local control in non-small cell lung cancer, where not many label data available for machine learning, the experimental result showed that local control predictions were accurate and comparable to multi-layer perceptron. In the future, the deep learning method can be used to prospectively individualize dosing and guiding in altering systemic therapy.[107] DNNs are also being implemented in advanced radiotherapy modalities, such as stereotactic body radiotherapy to keep up with dose calculation, which can also transform the dose calculated in one algorithm to another with high speed and accuracy.[108] CNN is also being studied in integration with conventional CT to be a potential replacement for MRI only based prostate proton beam therapy.[109] CNN is also being studied to be implemented to track tumour boundary in MRI for lung cancer.[110]

3.5. Radiotherapy verification and patient monitoring

IMRT is heavily dependent on the accuracy and position of each radiation beam. Gamma analysis is the standard method for analysing the fidelity of IMRT. The gamma statistic is used to compare the measured dose distribution to the planned dose distribution. Gamma analysis does not correlate with many clinically relevant deviations in delivered dose and is insensitive to small errors in multi-leaf collimator positioning. A method was developed to detect specific errors using image features in gamma image. It treats the gamma distributions as an image and uses feature evaluation on the patient image to predict prognoses, response to therapy and other outcomes. The model studied using 186 IMRT beams from 23 patients, where half of them have head-and-neck cancer and the rest have rectal cancer, lung cancer, glioblastomas, and sarcomas.[111]

Bayesian network and Markov random field are two commonly used graph models. In their study, Kalet et al. used the Bayesian network model to detect error in radiotherapy TP. The probability of obtaining certain radiotherapy parameters were calculated using the network, set of initial clinical information and radiation oncology based clinical database system. A low probability in propagated network signals for potential error is flagged for investigation. The network performance was then compared with human experts and in the case of brain cancer the network outperforms human experts.[112] Bayesian network method can be used to detect external beam radiotherapy physician order errors. Chang et al. showed that using the Bayesian network method the average true and false positive rates of error detection were 98.72% and 1.99%, which are comparable to human error detection and can detect physician's order error accurately.[113]

Carrara et al. showed that ANN can be used to predict any selected acute or late toxicity endpoint after prostate radiotherapy.[114] LDA is used in various radiotherapy studies and applications such as identifying predictive genes using whole genome microarray data from prostate cancer patients to study cancer related fatigue[115]; and linking the connection between gut microorganism and chemotherapy induced diarrhoea from CapeOX regimen in resected stage III colorectal cancer patients.[116]

4. Current challenges and future works

The advances in deep learning algorithms have allowed AI to move forward an impressive distance and they are being implemented in various clinical tasks. They are statistically impressive but individually far from perfect due to their unreliability in certain cases and prone to make some mistakes that a human would not make. Many of the AI learning methods require large data sets and they are often not available or very expensive or protected by intellectual property right. The algorithm also needs to be extensively tested for accuracy before it can be implemented clinically and that is often time consuming and expensive; however, AI does show a great promise for the future. Another big challenge for implementing the AI method in clinical practice is legal and ethical reasons: if some reason AI fails to deliver the correct output who will take the responsibility for the mistake?[117] Many of the current CAD systems require high quality diagnostic image for testing and training, and they are often obtained from one institute due to cost and availability, which can jeopardize the CAD system's viability results. [118] In the future, it is expected that the deep learning method will be implemented with the CAD systems to further improve their accuracy and reliability.[119] AI in mammography has made great improvement and shows a very promising potential to have fully automated CAD system primarily for clinical use.[120]

5. Conclusion

It can be seen that AI has unlimited potential in radiotherapy; however, it is not completely tuned yet to be used widely by itself in clinical use. It is already being implemented in some diagnostic niche cases for a solo first reader; however, more works need to be done. In the future, with more research and development, AI is expected to take massive workload away from the radiation staff including radiotherapists, medical physicists and radiation oncologists in radiotherapy.

Conflict of interest

None declared.

Financial disclosure

None declared.

References

1. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. 2017;69:S36–S40, http://dx.doi.org/10.1016/j.metabol.2017.01.011.
2. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019;28(2):73–81, http://dx.doi.org/10.1080/13645706.2019.1575882.
3. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. *J Am Coll Cardiol*. 2018;71(23):2668–2679, http://dx.doi.org/10.1016/j.jacc.2018.03.521.
4. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500–510, http://dx.doi.org/10.1038/s41568-018-0016-5.
5. Steinberg MH. Clinical trials in sickle cell disease: adopting the combination chemotherapy paradigm. *Am J Hematol*. 2008;83(1):1–3, http://dx.doi.org/10.1002/ajh.21033.
6. Baskar R, Lee KA, Yeo R, Yeoh KW. Cancer and radiation therapy: current advances and future directions. *Int J Med Sci*. 2012;9(3):193–199, http://dx.doi.org/10.7150/ijms.3635.
7. Khoo VS. Radiotherapeutic techniques for prostate cancer, dose escalation and brachytherapy. *Clin Oncol*. 2005;17(7):560–571, http://dx.doi.org/10.1016/J.CLON.2005.07.006.
8. Zaidi H, Vees H, Wissmeyer M. Molecular PET/CT imaging-guided radiation therapy treatment planning. *Acad Radiol*. 2009;16(9):1108–1133, http://dx.doi.org/10.1016/j.acra.2009.02.014.
9. Moody A, Brody H, Grayson M, et al. *OUTLOOK*; 2012:115.
10. De Felice F, Piccioli A, Musio D, Tombolini V. The role of radiation therapy in bone metastases management. *Oncotarget*. 2017;8(15):25691–25699, http://dx.doi.org/10.18632/oncotarget.14823.
11. Deshmukh P, Levy MS. Effective radiation dose in coronary imaging modalities: back to Basics. *Catheter Cardiovasc Interv*. 2015;85(7):1182–1183, http://dx.doi.org/10.1002/ccd.26013.
12. Weissleder R, Nahrendorf M. Advancing biomedical imaging. *Proc Natl Acad Sci USA*. 2015;112(47):14424–14428, http://dx.doi.org/10.1073/pnas.1508524112.
13. Li G, Citrin D, Camphausen K, et al. Advances in 4D medical imaging and 4D radiation therapy. *Technol Cancer Res Treat*. 2008;7(1):67–81, http://dx.doi.org/10.1177/153303460800700109.
14. Van De Bunt L, Van Der Heide UA, Ketelaars M, De Kort GAP, Jürgenliemk-Schulz IM. Conventional, conformal, and intensity-modulated radiation therapy treatment planning of external beam radiotherapy for cervical cancer: the impact of tumor regression. *Int J Radiat Oncol Biol Phys*. 2006;64(1):189–196, http://dx.doi.org/10.1016/j.ijrobp.2005.04.025.
15. Nwankwo O, Mekdash H, Sihono DSK, Wenz F, Glatting G. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: validation of a KBRT algorithm for prostate cancer treatment planning. *Radiat Oncol*. 2015;10(1):1–5, http://dx.doi.org/10.1186/s13014-015-0416-6.
16. Miras H, Jiménez R, Perales Á, et al. Monte Carlo verification of radiotherapy treatments with CloudMC. *Radiat Oncol*. 2018;13(1):1–9, http://dx.doi.org/10.1186/s13014-018-1051-9.
17. Wroe AJ, Bush DA, Schulte RW, Slater JD. Clinical immobilization techniques for proton therapy. *Technol Cancer Res Treat*. 2014;14(1):71–79, http://dx.doi.org/10.7785/tcrt.2012.500398.
18. Shaverdian N, Verruttipong D, Wang PC, et al. Exploring value from the patient's perspective between modern radiation therapy modalities for localized prostate cancer. *Int J Radiat Oncol Biol Phys*. 2017;97(3):516–525, http://dx.doi.org/10.1016/j.ijrobp.2016.11.007.
19. McDonald DG, Jacqmin DJ, Mart CJ, et al. Validation of a modern second-check dosimetry system using a novel verification phantom. *J Appl Clin Med Phys*. 2017;18(1):170–177, http://dx.doi.org/10.1002/acm2.12025.
20. Bzdok D, Krzywinski M, Altman N. Points of significance: machine learning: supervised methods. *Nat Methods*. 2018;15(1):5–6, http://dx.doi.org/10.1038/nmeth.4551.
21. Lopez C, Tucker S, Salameh T, Tucker C. An unsupervised machine learning method for discovering patient clusters based on genetic signatures. *J Biomed Inform*. 2018;85(2017):30–39, http://dx.doi.org/10.1016/j.jbi.2018.07.004.
22. Amari S. Machine learning. *Appl Math Sci*. 2016;194:231–278, http://dx.doi.org/10.1007/978-4-431-55978-8_11.
23. Hoi SCH, Jin R, Zhu J, Lyu MR. Semisupervised SVM batch mode active learning with applications to image retrieval. *ACM Trans Inf Syst*. 2009;27(3):1–29, http://dx.doi.org/10.1145/1508850.1508854.
24. Wang S, Summers RM. Machine learning and radiology. *Med Image Anal*. 2012;16(5):933–951, http://dx.doi.org/10.1016/j.media.2012.02.005.
25. Wen J, Fang X, Cui J, et al. Robust sparse linear discriminant analysis. *IEEE Trans Circuits Syst Video Technol*. 2019;29(2):390–403, http://dx.doi.org/10.1109/TCSVT.2018.2799214.
26. Berrendero JR, Cárcamo J. *Linear Components of Quadratic Classifiers*. Berlin, Heidelberg: Springer; 2019:13, http://dx.doi.org/10.1007/s11634-018-0321-6.
27. Yao X. Evolving artificial neural networks. *Proc IEEE*. 1999;87(9):1423–1447, http://dx.doi.org/10.1109/5.784219.
28. Krogh A. What are artificial neural networks? *Nat Biotechnol*. 2008;26(2):195–197, http://dx.doi.org/10.1038/nbt1386.
29. Hofmann T, Schölkopf B, Smola AJ. Kernel methods in machine learning. *Ann Stat*. 2008;36(3):1171–1220 http://www.jstor.org/stable/25464664.
30. Peng J, Zhou Y, Chen CLP. Region-kernel-based support vector machines for hyperspectral image classification. *IEEE Trans Geosci Remote Sens*. 2015;53(9):4810–4824, http://dx.doi.org/10.1109/TGRS.2015.2410991.
31. Ghahramani Z. Probabilistic machine learning and artificial intelligence. *Nature*. 2015;521(7553):452–459, http://dx.doi.org/10.1038/nature14541.
32. Ergul U, Bilgin G. Multiple-instance ensemble learning for hyperspectral images. *J Appl Remote Sens*. 2017;11(4):1, http://dx.doi.org/10.1117/1.jrs.11.045009.
33. Simpson I.J.A., Woolrich M.W., Andersson J.R., Groves A.R., Schnabel J., Uncertain Registration 2013;32(4):748-756.
34. Dharshinni NP, Azmi F, Fawwaz I, Husein AM, Siregar SD. Analysis of accuracy K-means and apriori algorithms for patient data clusters. *J Phys Conf Ser*. 2019;1230:012020, http://dx.doi.org/10.1088/1742-6596/1230/1/012020.

35. Xie N, Miao Z, Wang J, Zhang Q. Adaptive DBSCAN algorithm based on sample density gradient. *J Phys Conf Ser*. 2019;1229(1), http://dx.doi.org/10.1088/1742-6596/1229/1/012056.
36. Viroli C, McLachlan GJ. Deep Gaussian mixture models. *Stat Comput*. 2019;29(1):43–51, http://dx.doi.org/10.1007/s11222-017-9793-z.
37. Niyazi M, Niemierko A, Paganetti H, et al. Volumetric and actuarial analysis of brain necrosis in proton therapy using a novel mixture cure model. *Radiother Oncol*. 2019, http://dx.doi.org/10.1016/j.radonc.2019.09.008.
38. Lee SD, Lee JH, Choi YG, You HC, Kang JH, Jun CH. Machine learning models based on the dimensionality reduction of standard automated perimetry data for glaucoma diagnosis. *Artif Intell Med*. 2019;94(2016):110–116, http://dx.doi.org/10.1016/j.artmed.2019.02.006.
39. Hira ZM, Gillies DF. A review of feature selection and feature extraction methods applied on microarray data. *Adv Bioinformatics*. 2015;2015(1), http://dx.doi.org/10.1155/2015/198363.
40. Williams S. Pearson's correlation coefficient. *N Z Med J*. 1996;109(1015):38, http://dx.doi.org/10.1136/bmj.e4483.
41. Adler J, Parmryd I. Quantifying colocalization by correlation: the pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytom Part A*. 2010;77(8):733–742, http://dx.doi.org/10.1002/cyto.a.20896.
42. Gao T, Ji Q. Efficient markov blanket discovery and its application. *IEEE Trans Cybern*. 2017;47(5):1169–1179, http://dx.doi.org/10.1109/TCYB.2016.2539338.
43. Yu K, Liu L, Li J, Chen H. Mining markov blankets without causal sufficiency. *IEEE Trans Neural Networks Learn Syst*. 2018;29(12):6333–6347, http://dx.doi.org/10.1109/TNNLS.2018.2828982.
44. Che J, Yang Y, Li L, Bai X, Zhang S, Deng C. Maximum relevance minimum common redundancy feature selection for nonlinear data. *Inf Sci (NY)*. 2017;409-410:68–86, http://dx.doi.org/10.1016/j.ins.2017.05.013.
45. Radovic M, Ghalwash M, Filipovic N, Obradovic Z. Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. *BMC Bioinformatics*. 2017;18(1):1–15, http://dx.doi.org/10.1186/s12859-016-1423-9.
46. Bharti KK, Singh PK. Hybrid dimension reduction by integrating feature selection with feature extraction method for text clustering. *Expert Syst Appl*. 2015;42(6):3105–3114, http://dx.doi.org/10.1016/j.eswa.2014.11.038.
47. Rasmus Broa AKS. Principal component analysis Tutorial Review. *Anal Methods*. 2014;6:1–16, http://dx.doi.org/10.1002/wics.101.
48. Gottesman O, Johansson F, Komorowski M, et al. Guidelines for reinforcement learning in healthcare. *Nat Med*. 2019;25(1):16–18, http://dx.doi.org/10.1038/s41591-018-0310-5.
49. Wang L, Liu Y, Zhai X. Design of reinforce learning control algorithm and verified in inverted pendulum. *Chinese Control Conf CCC*. 2015:3164–3168, http://dx.doi.org/10.1109/ChiCC.2015.7260128.
50. Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. *Nature*. 2015;518(7540):529–533, http://dx.doi.org/10.1038/nature14236.
51. Poursaeidi MH, Kundakcioglu OE. Robust support vector machines for multiple instance learning. *Ann Oper Res*. 2014;216(1):205–227, http://dx.doi.org/10.1007/s10479-012-1241-z.
52. Tomar D, Agarwal S. Multiple instance learning based on twin support vector machine. *Adv Intell Syst Comput*. 2017;553:497–507, http://dx.doi.org/10.1007/978-981-10-3770-2_46.
53. Caetano TS, McAuley JJ, Cheng L, Le QV, Smola AJ. Learning graph matching. *IEEE Trans Pattern Anal Mach Intell*. 2009;31(6):1048–1058, http://dx.doi.org/10.1109/TPAMI.2009.28.
54. Leordeanu M, Sukthankar R, Hebert M. Unsupervised learning for graph matching. *Int J Comput Vis*. 2012;96(1):28–45, http://dx.doi.org/10.1007/s11263-011-0442-2.
55. Goodfellow I. Deep learning(working version). *Nat Publ Gr*. 2011;13(1):35, http://dx.doi.org/10.1038/nmeth.3707.
56. Lecun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444, http://dx.doi.org/10.1038/nature14539.
57. Schmidhuber J. Deep learning in neural networks: an overview. *Neural Networks*. 2015;61:85–117, http://dx.doi.org/10.1016/j.neunet.2014.09.003.
58. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med*. 2018;98(May):126–146, http://dx.doi.org/10.1016/j.compbiomed.2018.05.018.
59. Sun Y, Yen GG, Yi Z. Evolving unsupervised deep neural networks for learning meaningful representations. *IEEE Trans Evol Comput*. 2019;23(1):89–103, http://dx.doi.org/10.1109/TEVC.2018.2808689.
60. Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks. *Pattern Recognit*. 2018;77:354–377, http://dx.doi.org/10.1016/j.patcog.2017.10.013.
61. Heismann BJ, Schmidt BT, Flohr TG. *Spectral Computed Tomography*. SPIE; 2012, http://dx.doi.org/10.1117/3.977546.
62. Kilcoyne RF, Richardson ML, Porter BA, Olson DO, Greenlee TK, Lanzer W. Magnetic resonance imaging of soft tissue masses. *Clin Orthop Relat Res*. 1988;52(228):13–19, http://dx.doi.org/10.1053/j.ro.2017.04.003.
63. Ahrens ET, Bulte JWM. Tracking immune cells in vivo using magnetic resonance imaging. *Nat Rev Immunol*. 2013;13(10):755–763, http://dx.doi.org/10.1038/nri3531.
64. Koch KM, Hargreaves BA, Pauly KB, Chen W, Gold GE, King KF. Magnetic resonance imaging near metal implants. *J Magn Reson Imaging*. 2010;32(4):773–787, http://dx.doi.org/10.1002/jmri.22313.
65. Mushtaq N, To K, Gooding C, Khan W. Radiological imaging evaluation of the failing total hip replacement. *Front Surg*. 2019;6(June):1–13, http://dx.doi.org/10.3389/fsurg.2019.00035.
66. Geis JR, Brady AP, Wu CC, et al. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. *Can Assoc Radiol J*. 2019:1–6, http://dx.doi.org/10.1016/j.carj.2019.08.010.
67. Hall WA, Paulson ES, van der Heide UA, et al. The transformation of radiation oncology using real-time magnetic resonance guidance: a review. *Eur J Cancer*. 2019;122:42–52, http://dx.doi.org/10.1016/j.ejca.2019.07.021.
68. Pratondo A, Chui CK, Ong SH. Integrating machine learning with region-based active contour models in medical image segmentation. *J Vis Commun Image Represent*. 2017;43(1):1–9, http://dx.doi.org/10.1016/j.jvcir.2016.11.019.
69. Zhang Z, Sejdić E. Radiological images and machine learning: trends, perspectives, and prospects. *Comput Biol Med*. 2019;108(2018):354–370, http://dx.doi.org/10.1016/j.compbiomed.2019.02.017.
70. Despotović I, Goossens B, Philips W. MRI segmentation of the human brain: challenges, methods, and applications. *Comput Math Methods Med*. 2015;2015, http://dx.doi.org/10.1155/2015/450341.
71. Wang G, Li W, Zuluaga MA, et al. Interactive medical image segmentation using deep learning with image-specific fine tuning. *IEEE Trans Med Imaging*. 2018;37(7):1562–1573, http://dx.doi.org/10.1109/TMI.2018.2791721.
72. Xu Y, Wang Y, Yuan J, Cheng Q, Wang X, Carson PL. Medical breast ultrasound image segmentation by machine learning. *Ultrasonics*. 2019;91(2018):1–9, http://dx.doi.org/10.1016/j.ultras.2018.07.006.
73. Polan DF, Brady SL, Kaufman RA. Tissue segmentation of computed tomography images using a random forest algorithm: a feasibility study. *Phys Med Biol*. 2016;61(17):6553–6569, http://dx.doi.org/10.1088/0031-9155/61/17/6553.
74. Sethi G, Saini BS. Computer aided diagnosis system for abdomen diseases in computed tomography images. *Biocybern Biomed Eng*. 2016;36(1):42–55, http://dx.doi.org/10.1016/j.bbe.2015.10.008.
75. Oliveira FPM, Tavares JMRS, Oliveira FPM, Manuel J, Medical RST. Computer methods in biomechanics and biomedical engineering medical image registration: a review. *Comput Methods Biomech Biomed Engin*. 2016;17(2):73–93, http://dx.doi.org/10.1080/10255842.2012.670855.
76. Kirienko M, Cozzi L, Rossi A, et al. Ability of FDG PET and CT radiomics features to differentiate between primary and metastatic lung lesions. *Eur J Nucl Med Mol Imaging*. 2018;45(10):1649–1660, http://dx.doi.org/10.1007/s00259-018-3987-2.
77. Lu L, Shin H, Roth HR, et al. Deep convolutional neural networks for computer-aided detection : CNN architectures. *IEEE Trans Med Imaging*. 2016;35(5):1285–1298, http://dx.doi.org/10.1109/TMI.2016.2528162.
78. Roth HR, Lu L, Liu J, et al. Improving computer-aided detection using convolutional neural networks and random view aggregation. *IEEE Trans Med Imaging*. 2016;35(5):1170–1181, http://dx.doi.org/10.1109/TMI.2015.2482920.
79. Vinsard DG, Mori Y, Misawa M, et al. Quality assurance of computer-aided detection and diagnosis in colonoscopy. *Gastrointest Endosc*. 2019;90(1):55–63, http://dx.doi.org/10.1016/j.gie.2019.03.019.
80. Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-aided detection of prostate cancer in MRI. *IEEE Trans Med Imaging*. 2014;33(5):1083–1092, http://dx.doi.org/10.1109/TMI.2014.2303821.
81. Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal*. 2017;35:303–312, http://dx.doi.org/10.1016/j.media.2016.07.007.
82. Nicolae A, Morton G, Chung H, et al. Evaluation of a machine-learning algorithm for treatment planning in prostate low-dose-rate brachytherapy. *Int J Radiat Oncol Biol Phys*. 2017;97(4):822–829, http://dx.doi.org/10.1016/j.ijrobp.2016.11.036.
83. Smith WP, Doctor J, Meyer J, Kalet IJ, Phillips MH. A decision aid for intensity-modulated radiation-therapy plan selection in prostate cancer based on a prognostic Bayesian network and a Markov model. *Artif Intell Med*. 2009;46(2):119–130, http://dx.doi.org/10.1016/j.artmed.2008.12.002.
84. Guidi G, Maffei N, Vecchi C, et al. Expert system classifier for adaptive radiation therapy in prostate cancer. *Australas Phys Eng Sci Med*. 2017;40(2):337–348, http://dx.doi.org/10.1007/s13246-017-0535-5.
85. Kajikawa T, Kadoya N, Ito K, et al. Automated prediction of dosimetric eligibility of patients with prostate cancer undergoing intensity-modulated radiation therapy using a convolutional neural network. *Radiol Phys Technol*. 2018;11(3):320–327, http://dx.doi.org/10.1007/s12194-018-0472-3.
86. Coates J, Souhami L, El Naqa I. Big data analytics for prostate radiotherapy. *Front Oncol*. 2016;6(Jun):1–17, http://dx.doi.org/10.3389/fonc.2016.00149.
87. Valdes G, Wojtowicz L, Pattison AJ, et al. OC-0253: machine learning-based enables data-driven radiotherapy treatment planning decision support. *Radiother Oncol*. 2017;123:S127–S128, http://dx.doi.org/10.1016/s0167-8140(17)30696-5.
88. Cho DD, Wernicke AG, Nori D, Chao KC, Chang J, Parashar B. Predicting radiation therapy outcome for head and neck cancer patients using artificial neural network (ANN). *Int J Radiat Oncol*. 2014;90(1):S852, http://dx.doi.org/10.1016/j.ijrobp.2014.05.2442.
89. Ochi T, Murase K, Fujii T, Kawamura M, Ikezoe J. Survival prediction using artificial neural networks in patients with uterine cervical cancer treated by radiation therapy alone. *Int J Clin Oncol*. 2002;7(5):294–300, http://dx.doi.org/10.1007/s101470200043.
90. Zhen X, Chen J, Zhong Z, et al. Deep convolutional neural networks with transfer learning for rectum toxicity prediction in combined brachytherapy and external beam radiation therapy for cervical cancer. *Int J Radiat Oncol*. 2017;99(2):S168, http://dx.doi.org/10.1016/j.ijrobp.2017.06.386.

91. Ju E, Kim KH, Chang KH, et al. Artificial neural network-based variable importance analysis of prognostic factors related to radiation pneumonitis in patients with lung cancer: preliminary study. *J Korean Phys Soc*. 2019;75(4):277–282, http://dx.doi.org/10.3938/jkps.75.277.
92. Aneja S, Shaham U, Kumar RJ, et al. Deep neural network to predict local failure following stereotactic body radiation therapy: integrating imaging and clinical data to predict outcomes. *Int J Radiat Oncol*. 2017;99(2):S47, http://dx.doi.org/10.1016/j.ijrobp.2017.06.120.
93. Friedrich T, Scholz U, Elsässer T, Durante M, Scholz M. Systematic analysis of RBE and related quantities using a database of cell survival experiments with ion beam irradiation. *J Radiat Res*. 2013;54(3):494–514, http://dx.doi.org/10.1093/jrr/rrs114.
94. Lin H, Liu T, Shi C, et al. Feasibility study of individualized optimal positioning selection for left-sided whole breast radiotherapy: DIBH or prone. *J Appl Clin Med Phys*. 2018;19(2):218–229, http://dx.doi.org/10.1002/acm2.12283.
95. Brindle JM, Trindade AA, Shah AP, et al. Linear regression model for predicting patient-specific total skeletal spongiosa volume for use in molecular radiotherapy dosimetry. *J Nucl Med*. 2006;47(11):1875–1883.
96. Cao N, Kalet AM, Young LA, et al. Predictors of cardiac and lung dose sparing in DIBH for left breast treatment. *Phys Medica*. 2019;67:27–33, http://dx.doi.org/10.1016/j.ejmp.2019.09.240.
97. Waligórski MPR, Grzanka L, Korcyl M, Olko P. A TPS kernel for calculating survival vs. depth: distributions in a carbon radiotherapy beam, based on Katz's cellular track structure theory. *Radiat Prot Dosimetry*. 2015;166(1-4):347–350, http://dx.doi.org/10.1093/rpd/ncv202.
98. Jayarathna S, Manohar N, Ahmed MF, Krishnan S, Cho SH. Evaluation of dose point kernel rescaling methods for nanoscale dose estimation around gold nanoparticles using Geant4 Monte Carlo simulations. *Sci Rep*. 2019;9(1):3583, http://dx.doi.org/10.1038/s41598-019-40166-9.
99. Miura H, Ozawa S, Doi Y, et al. Automatic gas detection in prostate cancer patients during image-guided radiation therapy using a deep convolutional neural network. *Phys Medica*. 2019;64(1):24–28, http://dx.doi.org/10.1016/j.ejmp.2019.06.009.
100. Debus C, Oelfke U, Bartzsch S. A point kernel algorithm for microbeam radiation therapy. *Phys Med Biol*. 2017;62(21):8341–8359, http://dx.doi.org/10.1088/1361-6560/aa8d63.
101. Zhang X, Homma N, Ichiji K, et al. A kernel-based method for markerless tumor tracking in kV fluoroscopic images. *Phys Med Biol*. 2014;59(17):4897–4911, http://dx.doi.org/10.1088/0031-9155/59/17/4897.
102. Zhao W, Han B, Yang Y, et al. Incorporating imaging information from deep neural network layers into image guided radiation therapy (IGRT). *Radiother Oncol*. 2019;140:167–174, http://dx.doi.org/10.1016/j.radonc.2019.06.027.
103. Bangert M, Oelfke U. Spherical cluster analysis for beam angle optimization in intensity-modulated radiation therapy treatment planning. *Phys Med Biol*. 2010;55(19):6023–6037, http://dx.doi.org/10.1088/0031-9155/55/19/025.
104. Douglass M, Bezak E, Penfold S. Development of a radiation track structure clustering algorithm for the prediction of DNA DSB yields and radiation induced cell death in Eukaryotic cells. *Phys Med Biol*. 2015;60(8):3217–3236, http://dx.doi.org/10.1088/0031-9155/60/8/3217.
105. Francis Z, Seif E, Incerti S, et al. Carbon ion fragmentation effects on the nanometric level behind the Bragg peak depth. *Phys Med Biol*. 2014;59(24):7691–7702, http://dx.doi.org/10.1088/0031-9155/59/24/7691.
106. Meylan S, Incerti S, Karamitros M, et al. Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci Rep*. 2017;7(1):1–16, http://dx.doi.org/10.1038/s41598-017-11851-4.
107. Cui S, Luo Y, Jolly S, Ten Haken RK, El Naqa I. Prediction of local control in non-small cell lung cancer patients after radiation therapy by composite deep learning neural networks. *Int J Radiat Oncol*. 2018;102(3):S4–S5, http://dx.doi.org/10.1016/j.ijrobp.2018.06.107.
108. Dong P, Xing L. DoseNet: A Deep neural network for accurate dosimetric transformation between different spatial resolutions and/or different dose calculation algorithms for precision radiation therapy. *Int J Radiat Oncol*. 2019;105(1):S255–S256, http://dx.doi.org/10.1016/j.ijrobp.2019.06.2471.
109. Yang X, Liu Y, Lei Y, et al. MRI-based proton radiotherapy for prostate cancer using deep convolutional neural networks. *Int J Radiat Oncol*. 2019;105(1):S200, http://dx.doi.org/10.1016/j.ijrobp.2019.06.263.
110. Tahmasebi N, Boulanger P, Noga M, Punithakumar K. A fully convolutional deep neural network for lung tumor boundary tracking in MRI. *Conf Proc. Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Annu Conf*. 2018;2018:5906–5909, http://dx.doi.org/10.1109/EMBC.2018.8513607.
111. Wootton LS, Nyflot MJ, Chaovalitwongse WA, Ford E. Error detection in intensity-modulated radiation therapy quality assurance using radiomic analysis of gamma distributions. *Int J Radiat Oncol Biol Phys*. 2018;102(1):219–228, http://dx.doi.org/10.1016/j.ijrobp.2018.05.033.
112. Kalet AM, Gennari JH, Ford EC, Phillips MH. Bayesian network models for error detection in radiotherapy plans. *Phys Med Biol*. 2015;60(7):2735–2749, http://dx.doi.org/10.1088/0031-9155/60/7/2735.
113. Chang X, Li HH, Kalet AM, Yang D. Development and validation of a bayesian network method to detect external beam radiation therapy physician order errors. *Int J Radiat Oncol Biol Phys*. 2019;105(2):423–431, http://dx.doi.org/10.1016/j.ijrobp.2019.05.034.
114. Carrara M, Rancati T, Fiorino C, et al. A method to develop reliable "ready to use" graphical tools based on artificial neural networks for the prediction of toxicities after high dose prostate radiation therapy. *Int J Radiat Oncol*. 2015;93(3):E610–E611, http://dx.doi.org/10.1016/j.ijrobp.2015.07.2106.
115. Feng LR, Fernández-Martínez JL, Zaal KJM, Deandrés-Galiana EJ, Wolff BS, Saligan LN. MGluR5 mediates post-radiotherapy fatigue development in cancer patients. *Transl Psychiatry*. 2018;8(1), http://dx.doi.org/10.1038/s41398-018-0161-3.
116. Fei Z, Lijuan Y, Xi Y, et al. Gut microbiome associated with chemotherapy-induced diarrhea from the CapeOX regimen as adjuvant chemotherapy in resected stage III colorectal cancer. *Gut Pathog*. 2019;11(1):1–10, http://dx.doi.org/10.1186/s13099-019-0299-4.
117. McBee MP, Awan OA, Colucci AT, et al. Deep learning in radiology. *Acad Radiol*. 2018;25(11):1472–1480, http://dx.doi.org/10.1016/j.acra.2018.02.018.
118. Längkvist M, Jendeberg J, Thunberg P, Loutfi A, Lidén M. Computer aided detection of ureteral stones in thin slice computed tomography volumes using convolutional neural networks. *Comput Biol Med*. 2018;97(February):153–160, http://dx.doi.org/10.1016/j.compbiomed.2018.04.021.
119. Acharya UR, Hagiwara Y, Deshpande SN, et al. Characterization of focal EEG signals: A review. *Futur Gener Comput Syst*. 2019;91:290–299, http://dx.doi.org/10.1016/j.future.2018.08.044.
120. Fuchsjäger M. Is the future of breast imaging with AI? *Eur Radiol*. 2019:4822–4824, http://dx.doi.org/10.1007/s00330-019-06286-6.