Information Fusion 66 (2021) 111–137

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Full length article

A survey on deep learning in medicine: Why, how and when?

Francesco Piccialli [a,*], Vittorio Di Somma [a], Fabio Giampaolo [a], Salvatore Cuomo [a], Giancarlo Fortino [b]

[a] University of Naples Federico II, Department of Mathematics and Applications "R. Caccioppoli", Italy
[b] Department of Informatics, Modeling, Electronics and Systems, University of Calabria, Rende (CS), Italy

ARTICLE INFO

Keywords:
Deep learning
Medicine
Artificial intelligence
Data science
Neural networks

ABSTRACT

New technologies are transforming medicine, and this revolution starts with data. Health data, clinical images, genome sequences, data on prescribed therapies and results obtained, data that each of us has helped to create. Although the first uses of artificial intelligence (AI) in medicine date back to the 1980s, it is only with the beginning of the new millennium that there has been an explosion of interest in this sector worldwide. We are therefore witnessing the exponential growth of health-related information with the result that traditional analysis techniques are not suitable for satisfactorily management of this vast amount of data. AI applications (especially Deep Learning), on the other hand, are naturally predisposed to cope with this explosion of data, as they always work better as the amount of training data increases, a phase necessary to build the optimal neural network for a given clinical problem. This paper proposes a comprehensive and in-depth study of Deep Learning methodologies and applications in medicine. An in-depth analysis of the literature is presented; how, where and why Deep Learning models are applied in medicine are discussed and reviewed. Finally, current challenges and future research directions are outlined and analysed.

1. Introduction

In the coming years, Artificial Intelligence (AI) will have an increasingly important role in the field of medicine, where it is already making a difference today. Medicine based on the observation of events has been, for many years, ever since the time of Hippocrates, the epistemological guiding criterion of the healthcare profession. This approach has evolved, with the progress of medicine, into what is termed Evidence-Based Medicine (EBM). Today, indeed, medicine based on signs which cannot be observed by any human doctor but can become evident with the use of Big Data and Deep Learning (DL) techniques has been developed. Such techniques are able to consider and process much more information than is possible for any human. State-of-the-art Deep Neural Networks (DNNs), also known as DL modes [1] have demonstrated remarkable results in image processing, classification and data analysis.

DL is increasingly attracting the interest of researchers in the medical and healthcare sectors, since, by using medical data, it is possible to increase the accuracy of medical applications. In particular, DL is rapidly replacing classic neural network techniques, named artificial neural networks (ANNs), whose goal is to mimic the **human brain**. This trend can be motivated by the following reasons. Firstly, DL can provide a better interpretation of a very complex phenomenon than classic statistical approaches if high-dimensional datasets are available, and the performance of a DL is directly proportional to the input size. This is a common scenario in medicine, where, as pointed out in [2], large amounts of data (about 15 to 20 TB) are collected and stored in optimized databases every day, also by using Cloud computing platforms [3]. Furthermore, DL is characterized by a high degree of flexibility. Medical data include different types of unstructured data, such as images, signals, genetic expressions and text data. Thanks to the complexity of their architectures, DL frameworks are able to benefit from this heterogeneity by achieving high levels of abstraction in data analysis. Finally, the high level of automation [4]. ML algorithms require a manual intervention to select the fundamental information from the input data and the corresponding transformation rules [5]. This is a crucial challenge because an experts' decision is needed and, therefore, there is a corresponding increase in the time and costs for a diagnosis [6]. However, DL can determine these elements by using large samples of examples. There are two main consequences of this facility. Firstly, there is a significant reduction in the cost and time of treatment and diagnosis, Secondly, the independence of the diagnosis means that patients can talk directly to data scientists and, by running

* Corresponding author.
E-mail addresses: francesco.piccialli@unina.it (F. Piccialli), vittorio.disomma@unina.it (V.D. Somma), fabio.giampaolo@consorzio-cini.it (F. Giampaolo), salvatore.cuomo@unina.it (S. Cuomo), giancarlo.fortino@unical.it (G. Fortino).

https://doi.org/10.1016/j.inffus.2020.09.006
Received 22 July 2020; Received in revised form 24 August 2020; Accepted 12 September 2020
Available online 15 September 2020

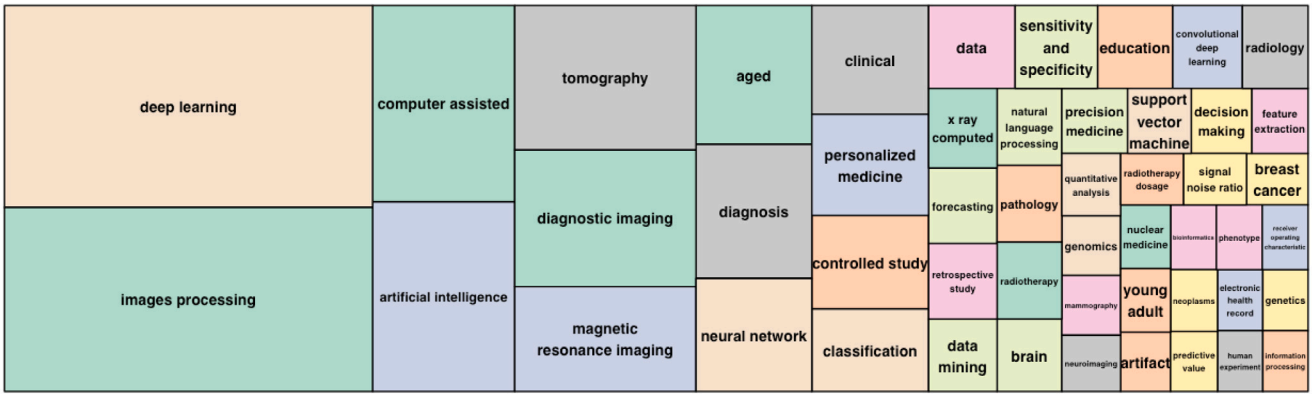

Fig. 1. A keyword-generated tree-map analysing a dataset of papers related to the application of DL methodologies in Medicine.

software, can understand the cause of their disease and obtain the best treatment.

The application of AI in real contexts can result in numerous potential advantages, such as the execution speed, potential reduction in costs, both direct and indirect, better diagnostic accuracy, greater clinical and operational efficiency ("algorithms don't sleep") and the possibility of providing access to the clinical information even for people who cannot otherwise benefit from this for geographical, political and economic reasons.

A great number of publications and surveys have addressed the use of DL in medicine, focusing on specific challenges or medical fields [7–10]. Nevertheless, most of these works are lacking in details, difficult to compare and do not provide the reader with a comprehensive overview of the applications of DL in the general medical area. Fig. 1 presents a keywords-generated tree-map extracted from a Scopus[1] dataset composed of papers related to "Deep Learning" and "Medicine" as input words. From 2016 until now, more than 1200 papers have been considered within this dataset. The tree-map has been generated by using the bibliometrix R-package,[2] an open-source tool for quantitative research in scientometrics and bibliometrics that includes all the main bibliometric methods of analysis. By analysing the tree-map (starting from the left side and considering the biggest squares), it is evident that DL in Medicine is mostly applied in the task of image processing, with a great focus on diagnostics. By continuing the analysis towards the right, some crucial keywords can be observed, such as "aged", "personalized medicine", and "classification". Summarizing the results of the keyword-based tree-map, it is possible to have an overview of the main medical fields, the principal tasks performed, and the most frequently used algorithms relating to DL in medicine during the last few years.

Starting from the above considerations, in this paper our aim is to provide an extensive analysis of DL applications in medicine, also categorizing DL models in relation to their applications in different medical areas. Afterwards, a comprehensive study of the state-of-the-art DL in medicine will be performed, taking into account existing DL surveys focused on specific medical fields. In summary, with this paper we aim to make the following contribution:

1. We will review the state-of-the-art in papers and surveys, especially of recent years, focusing on the application of DL in medicine, including all medical areas.
2. We will present a categorization of the DL models used and applied in medicine and give clear definitions of each, also providing an overview of hybrid DL architectures.

[1] http://www.scopus.com.
[2] https://www.bibliometrix.org/.

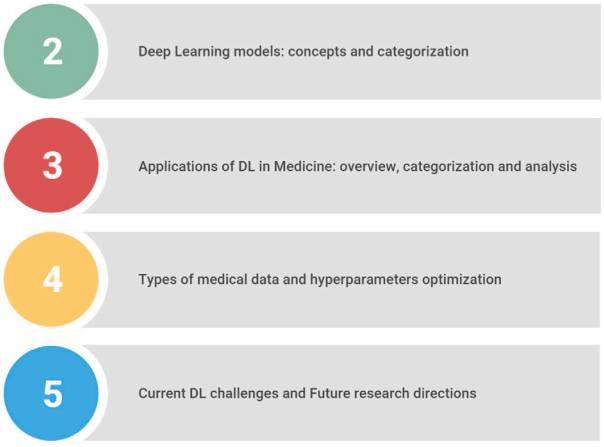

Fig. 2. An overview of the structure of this paper from Sections 2 to 6.

3. We will comprehensively classify medicine-related DL applications into macro-areas, also describing their sub-areas and the key aspects of the applied DL models.
4. We will analyse and discuss the recent and open challenges related to DL in medicine, also addressing future research directions, in order to provide the reader with a clear overview of the real-world scenario.

The rest of the paper is organized as follows, as depicted in Fig. 2. In Section 2, the various DL models are described and analysed in depth, with hybrid architectures also presented. In Section 3, a comprehensive overview and classification of DL applications in medicine is provided, including also a description of the main properties of the applied DL models. Section 4 presents a review of the kinds of medical data and hyperparameter optimization techniques. The current challenges in relation to the application of DL in medicine and future research directions are outlined in Section 5. Finally, in Section 6 our conclusions are presented.

2. Deep learning models

In this section we will provide a comprehensive overview of DL models applied in medicine. Starting from an in-depth study of the literature, we will present the main families of DL architectures: Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Autoencoders (AE), Generative Adversarial Networks (GAN), Deep Belief Networks (DBN) and Hybrid Architectures (HA) (see Fig. 3).

In the following subsections of this section, the various DL models will be discussed: Section 2.1 the CNNs; Section 2.2 the RNNs; Section 2.3 the AEs; Section 2.4 the GANs: Section 2.5 the DBNs; and

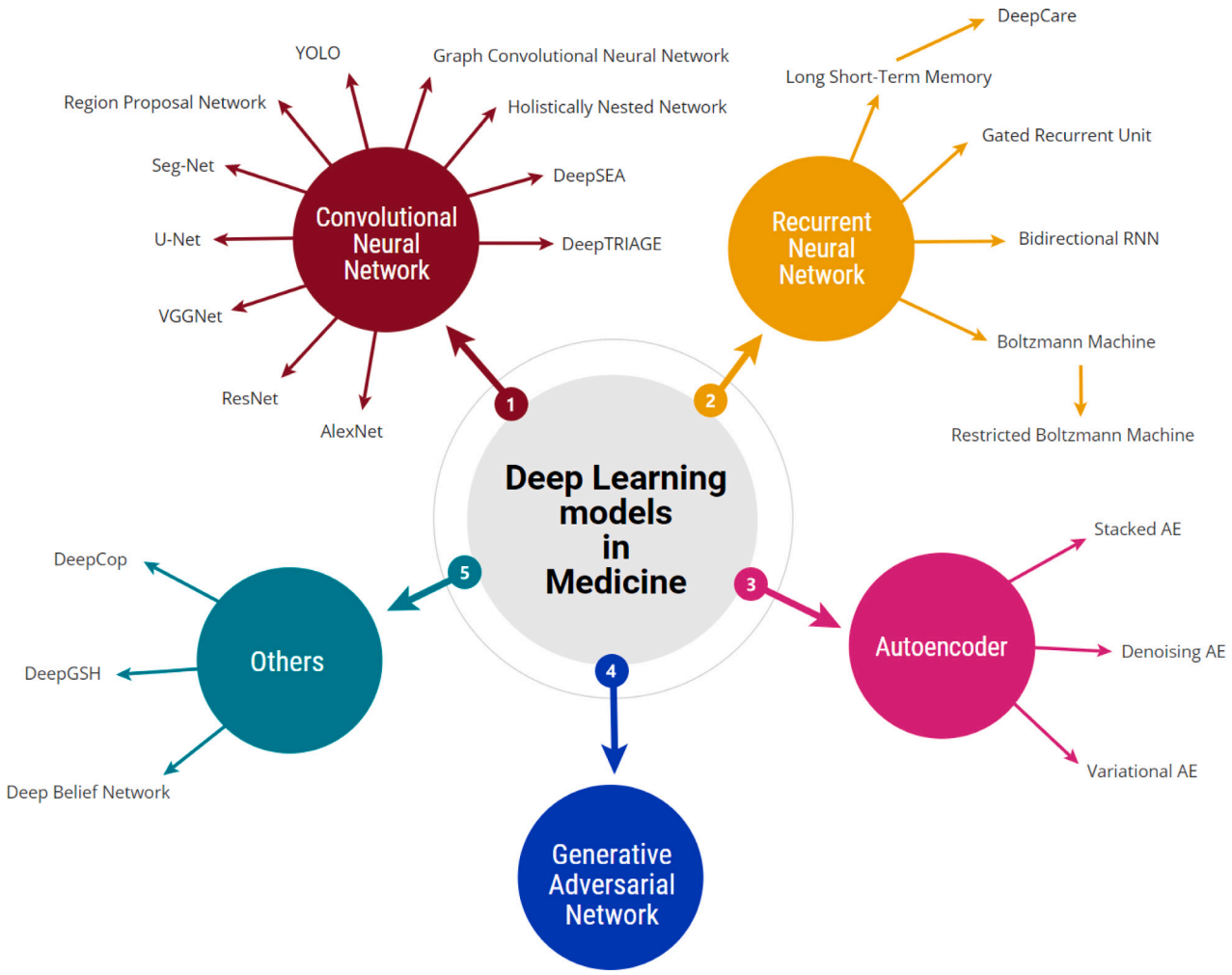

Fig. 3. A schematic representation of Deep Learning models used in Medicine.

Section 2.6 the HAs. For this study, we also make reference to and consider [1,11,12].

2.1. Convolutional neural networks

CNNs were introduced by LeCun in [13] as models for feature extraction tasks such that the spatial configuration of the data is preserved and the property of translation invariance is ensured. These characteristics make these models very popular in fields where an object shape is an important feature, such as image analysis [14–18], particularly in relation to the study of cancers and body lesions in the medical sector [8,19], and video analysis [20].

Concerning their basic architecture, this is composed of three kinds of layers, namely *convolutional layers*, *pooling layers* and *fully connected layers*.

As suggested by its name, a *convolutional layer* finds a real valued matrix, by computing a combination of the input values through a *discrete convolution* operation, whose goal is to extract specific features from the data. In formal terms, let $i(\cdot,\cdot,\cdot) : \mathbb{N}^3 \to \mathbb{R}$ be the input value, depending on its position in the space, $w_k : \mathbb{N}^3 \to \mathbb{R}$, $k = 1,\dots,K$ a set of $K \in \mathbb{N}$ weighted functions, named *kernels*; for every position $(x, y, z) \in \mathbb{N}^3$ of a generic pixel, the feature map $s_k(x, y, z)$ is given by

$$s_k(x,y,z) := \sum_m \sum_n \sum_p i(x-m, y-n, z-p) w_k(m,n,p).$$

A *pooling layer*, through the *pooling* operation, receives as input the feature maps and produces as output a particular statistical index (generally the maximum or the average) of the input itself. This operation is generally performed in order to reduce the dimensions of the inputs when high-dimensional structures are involved. Finally, the *fully connected layers* convert the pooling values into 1D-array and send them to the output layer.

Despite the great popularity of CNNs in image analysis, as discussed before, they suffer performance problems when high-resolution datasets are considered [21] and when a localization over great patches is required, especially in medical images [22]. These issues can be alleviated by using particular CNN architectures, such as AlexNet, VGGNet and ResNet, or specific strategies, such as in YOLO or U-net models, which will be described in the following.

2.1.1. AlexNet, VGGNet and ResNet

AlexNet was proposed by Krizhevsky and al. [21] for the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012. In terms of its structure, AlexNet is composed of 8 layers: the first 5 are convolutional layers and the other 3 are fully connected ones. The novel contribution of the AlexNet with respect to standard CNNs can be summarized as its: (i) *non-linear activation functions*; (ii) *dropout*; and (iii) *overlapping pooling*.

In relation to this first characteristic, in AlexNet the classical activation functions, such as softmax and hyperbolic tangent, are replaced by the ReLu function. This choice increases the accuracy of the network [21].

Concerning the second, this technique was first to alleviate the problem of overfitting of standard CNNs. The dropout decreases the

set of active neurons by selecting a random sample of neurons and eliminating it.

Finally, in terms of the third feature, in AlexNet the convolutional layers are followed by overlapping max-pooling layers. Differently to the classic approach followed in a standard CNN, in this variant the pixels over which the maximum is computed are overlapped according to a given stride. As shown in [21,23], the overlapping technique provides a higher accuracy than standard pooling operations.

This network was developed in order to improve the standard CNN performance in image analysis and, is usually applied in problems involving semantic segmentation and classification tasks with high-resolution data [24,25].

VGGNet was introduced by Simonyan and Zisserman in [26] in order to improve the accuracy of AlexNet by increasing the numbers of hidden layers, often referred to as the *depth* of the network, and introducing pre-training steps: with a greater number of smaller layers, the quantity of parameters decreases improving the performance of the model. This network has been applied in many fields, such as in medical image classification [27,28] and in image segmentation [29]. Despite the great popularity of VGGNet, there are several drawbacks related to this type of structure, the most important of which is that, if the depth becomes too high, the accuracy of the model starts to decrease. In order to alleviate this problem, in [30] a new CNN architecture, named *ResNet*, was proposed. Starting from the idea that the reduction in accuracy is related to the excessive number of direct mappings among the layers, ResNet is characterized by the following attributes: the decomposition of a mapping into the sum of a function, called a *residual mapping*, the identity function and the substitution of the original layer with a block of two hidden layers, called a *residual block*, where one layer learns from the residual, while the other learns from the identical function. In formal terms, given a learning map $H(x)$, the first layer of a residual block learns from the residual function $F(x)$, defined as

$$F(x) = H(x) - x,$$

while the second layer learns from the function $G(x) = x$.

Thanks to these improvements, ResNet has shown a high level of performance, especially in relation to image classification problems [31] and audio analysis tasks [32].

2.1.2. U-Net

U-Net, developed by Ronneberger [22], addresses the problem of images localization of a standard CNN. To overcome this issue, in U-Net the data features extraction is followed by a reconstruction step of the original dimension via an *upsampling* operation. Upsampling is a mathematical method applied to increase the sampling rate of a dataset. It can be divided in two steps, *expansion* and *interpolation*. In expansion, zero valued samples are added to the original data and ordered into a vector; in the interpolation step, a new dataset is generated by multiplying the output of the expansion step by a coefficients matrix, called a *filter*.

From a formal point of view, given the ordered input vector x and a coefficients matrix W, the final output y of the upsampling is given by:

$$y = Wx.$$

The U-Net is a particular type of *Encoder–Decoder (ED) model*: in the network, the features extraction is implemented by the *encoder*, which can be described as a standard CNN with max-pooling layers and no fully connected layer, while the features reconstruction is implemented by the *decoder*, which is composed of convolutional layers and upsampling layers, where the upsampling technique is performed. The main difference between the U-Net and an ED network is that, while in an ED the output of the encoder can belong to any space, in the U-Net this value must belong to the input space.

As previously mentioned, U-Net provides, in general, a better performance than a CNN in the image detection task. For this reason, this framework is widely applied in image segmentation problems, in particular in relation to the localization of cancer lesions [33–35].

A variant of U-Net is *SegNet* [36] whose goal is to reduce the space complexity of U-Net by making use of the max-pooling indices in the upsampling step.

2.1.3. Region-based CNN and YOLO

Region-based CNNs (RBCNN) were introduced by Girshick [37] to improve the performance of CNNs in detection problems. In this kind of problem a classical CNN is forced to work on many regions simultaneously, and this implies a high time complexity. In order to overcome this issue, a RBCNN determines the position of an object in an image by considering only some particular parts of an image, named *proposal regions*, generated by applying the following greedy algorithm: firstly, the original image is divided into multi parts, called *regions*; secondly, these regions are merged on the basis of some characteristics, as colour, size, texture; finally, the algorithm discards the regions with the lowest probability to contain the object to localize.

As mentioned at the beginning of this subsection, RBCNNs find application in object detection [38] and image segmentation [39]. However, they have a main drawback in that they can be very slow. In literature several DL models have been proposed to overcome this issue, and in the following subsection we will describe what is probably the most commonly used method, YOLO.

YOLO was introduced by Redmon et al. [40] to address the problem of the high time complexity of a RBCNN. Differently from standard RBCNNs, characterized by the selection of the proposal regions and their subsequent classification, this network divides the image into multi parts and deletes the ones with the lowest probability of enclosing an object (in this way, the number of the zones to analyse is decreased). Finally, the boxes with the highest probability of containing an object are merged in order to localize it with a higher precision. Thanks to its simplicity and to the fact that it can be easily extended to many image analysis tasks, YOLO has been gaining great popularity in the last few years, especially in relation to problems involving image detection [41].

2.1.4. Other CNNs

In this subsection we will briefly describe some particular CNN models, which have been developed to solve specific data analysis issues. First, we will consider a detection model, called *Holistically Nested Network*, consisting of two U-Nets, which follow an adaptive approach in the training step: this model is characterized by a lower space complexity than the classic U-Net [42].

DeepSea, developed by Zhou and Troyanskaya, is a CNN based model whose aim is to evaluate the non-coding variants effects in chromatin [43]. More precisely, this model determines these genetic effects by extracting the necessary features directly from DNA-sequences without implementing a pre-processing step and by combining different chromatin factors.

Another CNN based architecture, designed to identify several types of breast tumour by using genetic data, is *DeepTRIAGE* [44]. This model achieves a higher level of accuracy than other CNNs by using the information related to the heterogeneity of a class.

Finally, particular variants of CNNs, named *Graph Convolutional Networks* (GCNN), are networks where the hidden state does not depend only on the input, but also on the connections among the nodes, represented by the adjacency matrix. This kind of CNN has been applied in many fields, such as text classification [45] and medical data analysis [46].

2.2. Recurrent neural network

RNNs were developed by Rumelhart et al. [47] in order to exploit the correlations existing among the input data of a prediction problem. This property means that RNNs are frequently used in connection with problems where sequential data are processed, such as in relation to text analysis [48–50], particularly in medical electronic records in order to predict diseases [51,52], and speech recognition [53].

As pointed out at the beginning of this subsection, RNNs process data in which, at every step, the current conditions are affected by the historical conditions. This task is performed by introducing *internal memories* in the neural structure where information is stored and used to update the model in the subsequent steps. In addition, another consequence of the internal memory facility is the possibility of storing the parameters relative to the previous steps, which are shared among the neurons. Thanks to this property, known as parameter sharing, the space complexity is reduced on account of the absence of duplicates; moreover, networks of this type are able to work with data of different lengths due to the fact that information related to all the previous steps can be recovered, as shown in [54]

Starting from the previous considerations, a RNN can be formalized as follows: let h_t, y_t, x_t be respectively the hidden state, the output and the input; at the steps t, θ_t and σ_t the values of a set of parameters at t, f and g activation functions; at every step t, the values h_t and y_t are given by:

$$h_t = f(h_{t-1}, x_t, \theta_t) \tag{1}$$

$$y_t = g(h_t, \sigma_t). \tag{2}$$

The dependence on historical data provides more accurate results, but it poses two fundamental issues. Firstly, many applications, such as text analysis, require data related to many previous steps (*long-term dependencies*), while a basic RNN (1)–(2) is able to use only information about a short period in the past. Secondly, a RNN can suffer problems of instability because the network weights, which are proportional to the gradient of the activation function, could become null (this phenomenon is called *gradient vanishing*) or explode (*gradient exploding*) when dealing with heterogeneous data.

In order to overcome these issues, two particular RNN neural structures, called *long short-term memory* (LSTM) and *gated recurrent unit*, have been proposed in the literature. In the light of these considerations, in 2.2.1 we will discuss LSTMs and in 2.2.2 GRUs. Finally, in 2.2.3 two other common RNN architectures, *bidirectional recurrent neural network* (BRNN) and *Boltzmann machine* (BM), will be presented.

2.2.1. Long short-term memory

LSTMs were introduced by Hochreiter and Schmidhuber [55] in order to alleviate the problems of long-term dependency and gradient exploding and vanishing, as discussed in the previous subsection. This aspect makes them very commonly used in relation to speech recognition [56], path prediction [57] and medical diagnostics [58], where the authors propose a particular LSTM network, named *DeepCare*, whose goal is to predict possible clinical diseases by combining different kinds of medical information.

The basic architecture of a LSTM consists of three main components: the *forget gate*, the *input gate* and the *output gate*. This kind of structure gives to the cell a sort of internal memory, where information is stored to overcome the long-term dependency problem; moreover, the introduction of the forget gate, which controls which information in the state to discard, averts the problems of gradient vanishing and exploding, while the input and output gate monitor respectively the data flow which enters and goes out from the cell.

All the previous notions can be formalized as follows. Let σ be an activation function and, at every step t, x_t the input value, h_t the network state, s_t the cell state, $\theta_t^i, \theta_t^f, \theta_t^o$ network parameters. The forget gate determines the following weight f_t, measuring the relevance of the old information:

$$f_t = \sigma(h_{t-1}, x_t, c_{t-1}\theta_t^i).$$

The input gate computes the following weight f_i, measuring the relevance of the input data, given by:

$$f_i = \sigma(h_{t-1}, x_t, c_{t-1}\theta_t^f).$$

The cell updates its current state s_t using both f_t and f_i as:

$$s_t = f_t s_{t-1} + f_i x_t.$$

Finally, the output gate computes the following final output y_t and updates the network state h_t by applying the hyperbolic tangent function:

$$y_t = \sigma(h_{t-1}, x_t, c_t, \theta_t^o)$$
$$h_t = y_t \tanh(s_t).$$

Despite its ability to alleviate the problems relating to RNN, LSTM has a complex structure that increases the time complexity of the network. To address this issue, another kind of RNN cell, the *GRU*, has been developed, which will be presented in the next subsection.

2.2.2. Gated recurrent unit

GRUs, created by Kyunghyun Cho et al. in 2014 [59], are a class of RNNs whose goal is to improve the performances of LSTM networks when large amounts of data are involved. The idea underlying a GRU is to simplify the internal structure of the LSTM cell in order to decrease the time complexity of the network operations.

With respect to LSTM neurons, GRU cells have a simplified gate structure: only a *reset gate* and an *update gate* are present in this case. The former helps the model to determine how much of the past information (from the previous time steps) needs to be passed along to the future, while the second is used to manage how much of the past information can be forgotten. It is worth emphasizing that the absence of the output layer implies that both the output and the state of a cell coincide.

In formal terms, for every step t, let u_t and r_t be respectively the result of the update and the reset step, x_t the output, h_t the current hidden value and θ_{1t}^u, θ_{2t}^u, θ_{1t}^r, θ_{2t}^r, θ_t^1, θ_t^2, σ an activation function. A GRU can be summarized as follows:

1. Current hidden state update:

$$h_t = u_{t-1}h_{t-1} + (1 - u_{t-1})\sigma(\theta_{1t}x_{t-1} + \theta_{2t}r_{t-1}h_{t-1}).$$

2. Update and reset steps:

$$u_t = \sigma(\theta_{1t}^u x_t + \theta_{2t}^u h_t)$$
$$r_t = \sigma(\theta_{1t}^r x_t + \theta_{2t}^r h_t).$$

As discussed before, GRUs are an appropriate alternative to LSTMs when large data samples are involved. Moreover, there is no criterion which lets us choose between them a priori, but this can be determined only a posteriori [1]. Thanks to this interchangeability, GRU networks have the same application fields as LSTMs, in particular in speech recognition [60].

2.2.3. Bidirectional recurrent neural network and Boltzmann machine

BRNNs, introduced by Schuster and Paliwal [61], are a family of RNNs whose goal is to improve the accuracy of the results by increasing the input samples. More precisely, the hidden state is updated by using the past information, as in a classical RNN, and by using the information related to future instants. From a technical point of view, this task is implemented by dividing the neurons into two parts, one taking the past information, the other gathering the data from the following instants.

BRNNs have been applied in many fields, in particular in handwriting and speech recognition where they are used to detect the missing parts of a sentence from a knowledge of the other words [62,63].

BM models, introduced by Hinton et al. [64], are a family of stochastic RNNs where every neuron has the following characteristics: (i) it is modelled as a random variable; (ii) it is linked to those of the other hidden layers and these links are bidirectional (we can say that the connection is full); and (iii) its state takes only the values 0 and 1. The idea underlying a BM is to generalize inference methods to DL in order to outperform standard statistical approaches (such as maximum likelihood estimation and Monte Carlo methods) when large datasets are involved. In short, the goal of a BM is to infer the probability distribution of a high dimensional sample.

In the following we formalize a BM. Let i and j be two generic neurons, z_i the input of i, to which a bias b_i is associated, w_{ij} the weight between i and j, and s_j the state of the neuron, defined as

$$s_j = \begin{cases} 1 & \text{if i and j are connected} \\ 0 & \text{otherwise.} \end{cases}$$

The input x_i of i is equal to

$$x_i = b_i + \sum_j s_j w_{ij}.$$

As in the Hopfield network, the hidden state v_i is expressed by a function E, named the *energy function*, defined as

$$E(v_i) = -\left(\sum_{i<j} w_{ij} s_i s_j + \sum_i b_i v_i \right).$$

Given all the possible states v_j, the probability $P(v)$ associated with the state v_i is:

$$P(v) = \frac{exp(-E(v_i))}{\sum_j exp(-E(v_j))}.$$

Thanks to their easy implementation and their ability to replicate many probability distributions, BMs are often used as pre-training models in all applications where *a priori* knowledge of data is required, such as *computer vision* and *image classification*. In particular, in computer vision, BMs can be combined with other generative models (see Subsection 2.4.1) in relation to object localization problems, [65,66]. In image classification, BMs can be joined with Gaussian models to find the geometric structure of an image, with reference to, for example, an investigation into the presence of a tumour [67].

The main drawbacks of BMs are related to their full connection between the neurons. This feature makes this kind of structure very slow and not efficient when the size of the data grows exponentially [68]. A very common solution to alleviate this problem is to relax the connections among the neurons of the same layer (in this case the DL model is named a *Restricted BM*) or to assume a unidirectional connection among the neurons, as proposed in [1].

2.3. Auto encoder

AEs, developed by Rumellhart et al. [69], are DL models whose aim is to reduce the dimension of the data. For this reason, they are used as feature extractors in all the applications where very noisy and complex data are involved, such as medical image analysis [70,71], natural language processing [72] and video analysis [73].

In a generic AE network, every hidden layer is composed of two parts: an *encoder* and a *decoder*. The encoder generates a reduction of the input by mapping it into a vector, called a *code*, which belongs to a dimension space lower than the original one, called a *feature space*. The decoder produces a new representation of the original input by reverse mapping the code into the vector of the input space with the minimum distance from the original input.

All these concepts can be formalized as follows: let x be a generic input, X the input space, and Y the feature space. The decoder h is a function defined as

$$h : X \rightarrow Y$$
$$x \rightarrow y$$

the decoder $\tilde{h}$ is a function defined as

$$\tilde{h} : Y \rightarrow X$$
$$y \rightarrow x'$$

selected in order to minimize a given loss function L:

$$\min_{x'} L(x, x') \quad s.t. \quad x' = \tilde{h}(h(x)) \tag{3}$$

Despite their great ability to simplify very complex data, AEs have a drawback, namely the impossibility of extracting specific features when the number of the layers is too high, as shown in [1]. This issue is mitigated by constructing particular AEs, called *regularized AEs*, whose main idea is to decrease the number of active neurons. In more details, through the use of a loss function very sensitive to the input, a regularized AE is able to determine the neurons playing a less significant role in the feature extraction and make them inactive. According to different strategies, two main categories of regularized AEs can be identified: *sparse AE* (SAE) and *denoising AE* (DAE).

In a SAE, the loss function is modified by adding a *penalty term* depending on the code. From a mathematical point of view, defined $\Omega(h)$ the penalty related to the code $h(x)$, Problem (3) can be rewritten as

$$\min_{x'} L(x, x') + \Omega(h) \quad s.t. \quad x' = \tilde{h}(h(x)).$$

Differently from SAEs, a DAE does not add a penalty to the loss function, but substitutes the original input with its noisy version. Formally, given an input x, in a DAE the new value $\hat{x}$, defined as

$$\hat{x} = x + \varepsilon,$$

with ε random variable, is considered as input; then the following minimization problem is solved:

$$\min_{x''} L(x, x'') \quad s.t. \quad x'' = \tilde{h}(h(\hat{x}))$$

Additional variants of AE that can be found in the literature are *variational AEs* (VAEs). In a VAE, the encoder is represented by the probability density function of the input in the feature space and, after the encoding stage, a sampling of the new data using the *PDF* is added. Differently to DAE and SAE, a VAE is not a regularized AE, but belongs to the class of *generative models*. This class of models are ML frameworks that sample new data from a density probability function, estimated by using the available data input;. As will be detailed in the next sections, they are widely used as pre-training models to deal with the overfitting problem.

2.4. Generative adversarial network

GANs, introduced by Goodfellow [74], are a family of generative models whose goal is to generate new data from the estimated input data distribution (see Part 2.3.1). As generative models, GANs are used as pre-training models to alleviate the overfitting problem in many fields, such as image analysis [75,76] or natural language processing [75].

The hidden layers of a GAN structure are composed of two parts: (i) the *generator*, which estimates the data probability density from the actual input values and exploits this to sample new data (*fake data*); and (ii) the *discriminator*, which tries to recognize the fake data and separate them from the real data. Briefly, a GAN is based on a two-player zero-sum game: at every step, the discriminator receives data from the generator and discards the fake data and, consequently, the generator is forced to improve the data quality; the game ends when the Nash equilibrium is obtained, i.e. the discriminator is no longer able to distinguish real from fake data.

The previous discussion can be formalized as follows. Let x be a generic input, $g(x)$ the sample produced by the generator, $d(x)$ the

probability that $g(x)$ is real, $\mathbb{E}_{[\cdot]}$ the expectation of a random variable, p_d the distribution of the real data, p_m the distribution of the fake data, $v(g,d)$ the function defined as:

$$v(g,d) = \mathbb{E}_{x\approx p_d}[\ln d(x)] + \mathbb{E}_{x\approx p_m}[\ln(1-d(x))].$$

A GAN determines the solution of the following *max–min* problem:

$$\max_d \min_g v(g,d). \tag{4}$$

Despite their ability to generate high-quality data and to replicate very complex probability distributions, GANs are characterized by two main issues [77]: (i) after several iterations, a GAN can produce many duplicates, which do not make any contribution to the knowledge of the problem (this phenomenon is known as the *mode collapse problem*); and (ii) a GAN can become very unstable since the existence of the solution of (4) is not guaranteed. Possible solutions to overcome these problems are described in [78].

2.5. Deep belief network

DBNs, created by Hinton [79], are a class of IGM generative models (see subsection 2.3.1). As discussed in the introduction to this section, DBNs estimate the probability distribution of the data using a set of binary variables, named *latent variables*. In particular, these latent variables are assumed to follow an *exponential distribution*.

Given an input x, a set of latent variables h_i, a normalization constant Z and two sets of biases a_i and b_i, the joint probability $\pi(\cdot,\cdot)$ for the latent variables and input can be defined as:

$$\pi(x,h) = \frac{1}{Z}\prod_i exp[-E(x,h_i)],$$

where E is the so-called *energy function*, describing the hidden state, given by

$$E(x,h_i) = -\left(\sum_i a_i x + \sum_i b_i h_i\right).$$

A DBN is the result of a sequential connection of several BMs, but it cannot be classified as a RNN. This point can be explained by the fact that, in a BM, the recursion holds because of the connection among the neurons of the same layer, while in a DBM every neuron can interact only with those of the previous and next layers.

As generative models, DBNs are widely applied as pre-training models in *image recognition* and *speech recognition*. In the first class of problem, DBNs are often used in classification tasks to detect lesions in medical diagnostics and, more generally, in video recognition to identify the presence of people [80]. In relation to speech recognition, for example, DBNs have been used to understand the missing words in a sentence [81]. Moreover, an interesting application of this kind of network applied on physiological signals to recognize human emotion can be found in [82].

2.6. Hybrid architectures

As discussed in this section, all DL models in medicine can be classified in five main categories, with the choice of model depending on the type of data involved in the problem to solve. Besides this, many applicative scenarios require different types of data as input. In order to address problems of this kind, a strategy can be found in *hybrid architectures* (HA), sequential combinations of different DL models whose aim is to benefit from the characteristics of every single model. In the following subsection, three HAs very commonly applied in medical problems HAs are discussed: *CNN+RNN*, *AE+CNN* and *GAN+CNN*.

CNN+RNN structures are obtained by combining, as the name suggests, CNNs, which are used for their feature extraction capabilities, and RNNs [83]. Since the output of a CNN is generally a $3D$-value and a RNN works with $2D$-data, an extra layer, called the *reshape layer*, is generally added between the CNN and the RNN in order to convert the output of the CNN into a matrix. CNN+RNN networks have been successfully applied in text analysis to identify missing words [84] and in image analysis to increase, for example, the speed of storing magnetic resonance images [85,86]. Variants of CNN+RNNs are obtained by replacing the standard RNN component with a LSTM one [87,88].

The AE+CNN architecture is a sequential combination of an AE, used as a pre-training model when data with high levels of noisiness are involved, and a CNN, which is used as a feature extractor model. AE+CNNs are applied, for example, in images analysis to classify noisy medical images [89] and in the reconstruction of medical images [71, 90].

A GAN+CNN network is obtained as sequential combination of a GAN, used as a pre-training model to moderate the overfitting problem, and a CNN, which is used as a feature extractor. This HA is often used in image analysis problems [91,92].

3. Applications fields of DL in medicine

DL achieves a remarkable accuracy and quality of its results thanks to its multi-layer architecture, which is able to obtain a high level of abstraction by working with large data samples. For this reason, it is gaining a great popularity in all fields where the process of information extraction from data involves various problems, such as the medical sector [93–95].

In detail, the analysis of medical data encounters three main issues, which are summarized in the following discussion. Medical data can be of different types, from images to text values, and a specific technique is required for each. This poses several difficulties when different kinds of data must be combined to obtain, for example, a medical diagnosis [2]. A second problem is that traditional ML algorithms have shown bad performances in the analysis of unstructured information, especially in the case of medical data such as clinical images and text notes, which is considered one of the main medical tools in the prediction and prevention of human disease. Finally, many medical data, such as genetic expressions and biosignals, suffer from high levels of noisiness and variability, which make the process of knowledge extraction very difficult.

Starting from an in-depth study of the main papers in Scopus, this section covers the main DL applications in medicine, which have been classified into the following four macro-areas on the basis of the kind of data involved in the analysis process: *Clinical Images*, *Biosignals*, *Biomedicine* and *Electronic Health Records* (EHR). According to this classification, this subsection is organized as follows: Section 3.1 addresses Clinical Images; Section 3.2 Biosignals; Section 3.3 Biomedicine; and Section 3.4 EHR (see Fig. 4).

3.1. Clinical images

Clinical images are the main tool used by doctors in diagnosis and disease prediction, a fact which explains the necessity of using statistical models for the analysis of such data. For this reason DL can be considered as the new paradigm in clinical image analysis, as confirmed by the large number of papers relating to this topic [7,96–100].

In the present subsection the principal applications of DL in clinical image analysis are, for the sake of simplicity, categorized into six main groups [101,102]: (i) *classification*, which is the process of identifying a particular feature in an image; (ii) *detection*, which is the process of determining the location of multiple features; (iii) *segmentation*, which is the process of dividing an image into multiple parts;

(iv) *reconstruction*, which is the process of reproducing the features in their original conditions; (v) *registration*, which is the process of merging two different images into one; (vi) *dose estimation*, which is the process of estimating the right dose quantity of a substance.

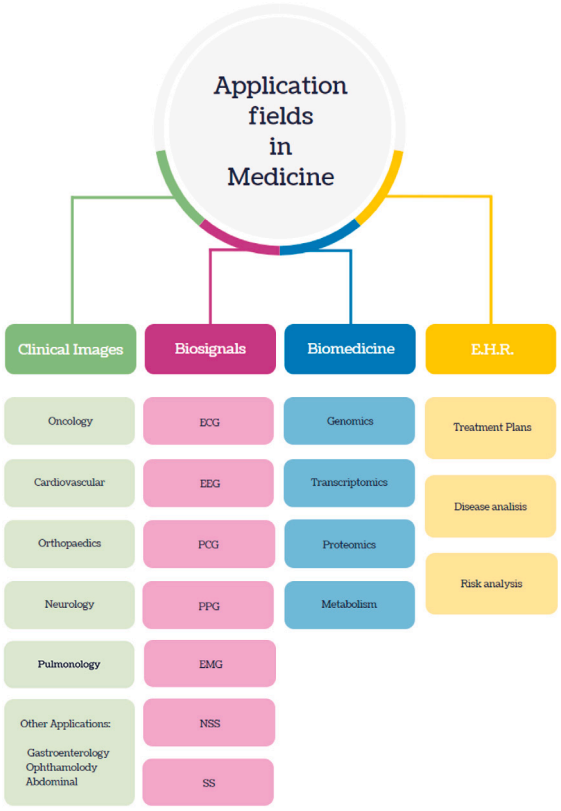

Fig. 4. A structured representation of the applications fields of DL in Medicine. It is possible to observe, from a medical point of view, the four macro-areas and the related sub-areas.

An in-depth exploration of the literature reveals that the DL models usually applied in problems involving images are *CNNs*, *AEs* and *GANs*. CNNs are widely used in this field due to their ability to preserve the spatial structure of data as well as their proven skills as features extractors (in particular, U-Net is often applied on segmentation tasks due to its ability to reproduce the initial dimensions of data features). AEs are very common as data reduction models when multi-features and complex images are involved in the analysis process, while GANs are mainly found as pre-training models for the generation of new input instances whose goal is to manage *overfitting* (for more details, see Section 5).

In the following subsection the applications and contributions of DL will be discussed on the basis of the type of medical disease analysed by the images. In detail, this subsection has been structured as follows: in 3.1.1 we will discuss the contribution of DL to *Oncology*; in 3.1.2 its contribution to *Cardiovascular* diseases; in 3.1.3 its contribution to *Orthopaedics*; in 3.1.4 its contribution to *Neurology*, in 3.1.5 its contribution to *Pulmonology* and finally, in 3.1.6 its contribution to ***Other Applications***, which includes all the other medical sectors.

3.1.1. Oncology

Oncology is in absolute terms the main DL application field in medicine, as confirmed by the large number of papers in relation to this topic found in public databases (Scopus, Pubmed) [103–107]. This trend can be explained by the role played by cancer as a cause of death globally and by the lack of knowledge about many mechanical aspects of a tumour.

According to the categorization provided above, in this subsection we regroup the contributions of DL models in the medical field under examination by type of application.

In classification and detection tasks, CNNs have been widely used in the identification of many kinds of tumours [108–110]. In particular, CNNs have shown optimal results in terms of accuracy and quality in the estimation of the cancer typology (benign, malignant or equivocal) [111,112], in the analysis of the cancer sub-types [113,114] and in the analysis of tumour stratification [115] and grading [116,117]. An interesting application can be found in [118] where, with the aim of analysing rectal cancer, Fu et al. introduce the transfer learning approach, which consists in the use of information deriving from a similar problem as a priori knowledge. Non-standard CNNs have been proposed to perform particular classification tasks, such as ResNet to recognize the lesions of a breast cancer [119], VGGNet to determine the level of lymph-vascular space invasion in cervical tumour [120] and a fully connected CNN (*DenseNet*) to compute the amount of lymph node metastasis in gastric tumours [121]. Moreover, as pointed out in Section 2.3, CNNs are often combined with an AE in order to reduce the noisiness level of very complex images; an example of this HA applied to cancer image analysis can be found in [101,122,123].

In segmentation tasks, U-Net has been widely applied thanks to its ability to reconstruct the original dimensions of the features extracted during the learning process [124–126]. In [127] Boers et al. develop a U-Net to improve the quality of the colour contrast in the analysis of pancreatic tumours, while in [128] the same network architecture is proposed in order to study the impact of dimensional variations on the DL model performance in relation to thyroid melanomas. Moreover, also GANs have been used in image segmentation as a pre-training model with a U-Net to alleviate overfitting, as in [91], or as DL models with the aim of increasing the accuracy level of the results, as shown in [92].

In image registration, CNNs have shown very good results in terms of both time complexity and image accuracy, as in the analysis of the lung melanoma [129,130].

Finally, in the dose estimation task, using a set of initial values, GAN has been applied to estimate the optimal dose in a treatment plan [131] and in dosimetric analysis in computer tomography [132].

3.1.2. Cardiovascular

The cardiovascular region has always been a central topic in medicine since cardiovascular diseases, such as heart attack, are the number 1 cause of death globally [133], taking an estimated 17.9 million lives each year.[3] Many experts in this field are turning their attention to DL in order to obtain a new perspective in the analysis and prediction of the main diseases of this region [133–138]. In the following we summarize some of the principal results relating to image analysis applications in haematology, cardiology and angiology, addressing classification, segmentation and reconstruction tasks, and describe the contribution of DL models in this field.

Concerning the classification task, CNNs [139] provide good results in the study of heart atriums, which are improved in [140] by replacing a standard CNN with ResNet. In [141] a CNN is applied in the analysis of particular heart characteristics by combining images and echocardiograms [141]. In [142] the authors propose a HA composed of a GAN and a ResNet in the classification of blood cell images.

In the detection task, CNNs have been applied in the localization of potential arterial calcifications [139] and in the analysis of thick blood smears in order to detect malaria [143] (in this paper, the authors make use of one of the particular CNNs discussed in Section 2.1, YOLO).

In the segmentation task, a U-Net has been implemented in order to identify endoleaks after aortic aneurysm repair [144] (the authors introduce new measures of accuracy to estimate the performances of the model). In [139] Wong et al. propose two DL models for atrium segmentation: a DBN, which overcomes the problem of poor image quality, and a hybrid CNN+AE structure, which improves the accuracy of the traditional models.

[3] https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

Finally, in the context of reconstruction tasks, CNNs have been applied to improve the Signal-to-Noise Ratio in the study of arterial diseases [145,146]; moreover, in [147] Vu et al. present a GAN model which monitors the blood vessels in order to prevent the possible onset of cardiovascular disease.

3.1.3. Orthopaedics

DL in orthopaedic research is experiencing a wide diffusion in the diagnosis of the main bones diseases, such as osteoarthritis [42,148–152]. Concerning this medical field, the principal DL tasks applied on clinical images prove to be classification and segmentation, both discussed in the following paragraphs.

In the area of image classification problems, two of the main challenges are linked to the evaluation of bones age [153] and to the impact of gender in the analysis of skeletal muscles [154]. As in case of oncology, CNNs are the principal class of DL model applied in this field. In [153] the authors propose a VGGNet to estimate the bone age, in [154] the authors develop a multi-step CNN in which the first stage is implemented through an AlexNet, and the subsequent stages through VGGNet.

As regards image segmentation, this task applies mainly to the analysis of the cartilage structure and the development of new biomarkers for the treatment of common orthopaedics diseases. As pointed out also in the introduction of this section and in the case of oncology, this task is very commonly addressed with U-Net architectures, which, for example, have been used to find new biomarkers for osteoarthritis in [149,150], to reconstruct the original cartilage fracture in [148,155] and to study the structure of irregular cartilages in [42]. Finally, in [151] Gaj et al. propose a hybrid model, obtained by the combination of a GAN in the pre-training step and a U-Net in the features extraction step for the study of the osteoarthritis.

3.1.4. Neurology

Neurology is probably one of the medical fields in which clinical images are most widely used both for research purposes in relation to the brain and for diagnostic procedures linked to neurological diseases. In this context DL applications have shown very promising results in the analysis of such images, since they are usually characterized by a high degree of complexity [156] which can be managed by complex DL structures. The following discussion is mainly focused on the three principal tasks that can be found in the literature with regard to the study of neurological images, namely classification, segmentation and reconstruction.

In the classification task, the central issue consists in the risk estimation of developing neurological diseases. In this case CNNs are one of the most widely used models, as shown, for example, in [157,158], where this kind of neural network achieves a high level of accuracy in relation to the prevention of Alzheimer's disease.

In conventional non-quantitative magnetic resonance imaging, the image contrast is consistent within the images, but the absolute intensity can vary arbitrarily between scans. For the quantitative analysis of intensity data, the images are typically normalized to a consistent reference. The most convenient reference is a tissue that is always present in the image, and is unlikely to be affected by pathological processes. In multiple sclerosis neuroimaging, both the white and grey matter are affected, so normalization techniques that depend on the brain tissue may introduce bias or remove biological changes of interest.

In the segmentation task, Brown et al. [159] propose a robust calibration procedure for the contrast intensity data of MRI images based on a segmentation operation through a U-net. In [160] the problem of neural diseases is studied by means of a GAN model, which generates new brain images in order to increase the accuracy of the medical diagnosis.

Finally, in image reconstruction, CNNs have provided good results in magnetic resonance analysis [161,162]. In [163] these results have been improved by introducing a transfer learning approach. In particular, as mentioned in the previous applications, in [162] a CNN is combined with a generative model (in this case a DAE) as a pre-training model to overcome the overfitting.

3.1.5. Pulmonology

Pulmonology is a medical field where DL is outperforming the more classical statistical methods presented in the literature in terms of accuracy and efficiency [164]. As regards this discipline, the main DL tasks are classification and segmentation.

In the classification task, CNNs provide good results in the study of lung diseases [165,166], in particular in the diagnosis of chronic obstructive pulmonary (COP) diseases. In [167] Tang et al. propose a ResNet architecture in the analysis of (COP) diseases, whose goal is to improve the accuracy of standard CNNs. In [168] the COP diseases are studied by using a hybrid architecture, obtained by combining a standard CNN and a VGGNet. An alternative to CNNs has been presented in [169], where the authors analyse patients' exacerbation frequency to predict the presence of COP diseases by means of a DBN.

In the segmentation task, as in the case of the DL applications treated in the previous subsections, U-Net is the most DL common model. In [170] Park et al. develop a U-Net architecture for the lung segmentation by making use of images extracted from different kinds of datasets.

3.1.6. Other applications

In this subsection the main DL models used in *gastroenterology*, *ophthalmology* and *abdominal analysis* are discussed, these areas considered as new emerging fields where Neural Network techniques are applied.

As a starting point, the principal DL tasks in gastroenterology are explored [171]. CNNs prove to be the most frequently applied methodologies as regards classification tasks, since they provide, in general, very good results, for example in the analysis of the celiac diseases [172], in the identification of the non-alcoholic fatty liver diseases [173] and in the estimation of the characteristics of a normal pancreas [174]. Among detection tasks, an interesting reference can be found in [175], where a combination of a U-Net and a ResNet has been implemented to measure the neoplasia levels. Finally, regarding the segmentation task, a CNN has been applied to determine the right position of a catheter [176].

Also in ophthalmology CNNs have been found to be the most popular DL models in many classification tasks, such as, for example, the diagnosis of cataracts [177,178], the identification of the effects of diabetes on the eyes [179] and the examination of blood vessels in the retina [180]. An alternative to CNNs has been proposed in [181], where Zhang et al. developed a classification and segmentation framework by combining a CNN with a LSTM in order to prevent possible visual abnormalities in children.

Segmentation and reconstruction tasks prove to be the principal areas of DL algorithm application in abdominal analysis: in [182] the authors address the problem of measuring the adipose tissue using a CNN; in [183,184] CNN architectures are implemented for image reconstruction, where a multi-image is summarized into a single image obtaining a substantial error reduction; in [185] the classification of abdominal images is performed by using different CNN models (AlexNet, VGGNet) and a transfer approach is applied in order to improve the accuracy of the results; finally, in [186], Zollner et al. present a HA, obtained as a combination of a CNN and a GAN, in order to improve the accuracy and alleviate the overfitting problem. To summarize the above discussion, in gastroenterology, ophthalmology and the abdominal analysis field, CNNs are the most popular frameworks in most classification, segmentation and registration tasks, particularly in relation to disease diagnosis. The accuracy and quality of CNNs are sometimes improved by particular CNNs (AlexNet, U-Net, VGNet) or hybrid models, whose principal goal is to reduce the overfitting problem.

This subsection has examined the main DL applications in clinical image analysis, which have been classified, on the basis of the medical discipline, into six classes: oncology, cardiovascular diseases, orthopaedics, neurology, pulmonology and **other applications**. For each of these classes, we have highlighted the contributions found in the literature in relation to five principal tasks, namely classification,

detection, segmentation, reconstruction and registration. Subsequently, the main DL models used in this field, i.e. CNNs, AEs, GANs, have been discussed, and some relevant applications for every task have been indicated. We will conclude this section by highlighting that CNNs, and in general CNN-based architectures, are widely used in every task thanks to their ability to preserve the spatial structure of the objects and to extract specific features from the data, characteristics that prove to be essential when dealing with images. On the other hand, AEs are mainly used in pre-training steps especially to reduce the noise of multi-feature clinical images, while GANs are mainly implemented in HAs, combined with CNNs, in order to manage the problems resulting from overfitting issues. These results are summarized in Table 1, which is structured as follows: (i) the first column, headed *Area*, reports the areas into which the macro areas have been divided, which correspond to the parts into which this subsection has been divided; (ii) the second column, headed *Task*, contains the DL clinical image tasks analysed in this subsection; (iii) the third column, headed *Contribution*, reports the principal contributions of the paper analysed in this subsection; (iv) the fourth column, headed *Model*, reports the main DL models associated with the DL tasks in relation to clinical images.

3.2. Biosignals

With the word biosignals we denote the set of all the electrical, mechanical and thermal signals produced by a human body [194], measured by doctors to improve the accuracy of their diagnoses and studied to prevent several diseases. However, even if biosignals are considered as a great source of information, their analysis is usually difficult due to their high degree of noisiness and non-stationarity. Nevertheless, DL has shown promising results in addressing these issues thanks to the ability of Neural Networks to extract information from very complex and noisy data [195–197]. Starting from these considerations, this subsection covers the main DL applications to biosignals, which have been divided on the basis of the nature of the signal into the following groups: (i) *electrocardiograms* (ECG); (ii) *encephalograms* (EEG); and (iii) *other biosignals*, which includes *phonocardiograms* (PCG), *photoplethysmograms* (PPG), *electromyograms* (EMG), *spectroscopy signals* (SS) and *nanopore sequencing* (NSS).

For every task, the most common DL models and contributions have been highlighted: CNNs are widely used in biosignals analysis thanks to their ability as feature extractors on very complex data. In addition, two other common DL models in this field are *RNNs* and *AEs*: RNNs usually provide good results since, as discussed in Section 2.2, the presence of the internal memory allows them to benefit from the correlations existing among biosignal data; AEs, on the other hand, are usually applied as methodologies whose goal is to reduce the noisiness level of the data. In detail, this subsection is organized as follows: in 3.2.1 the contributions of DL to ECGs are discussed; in 3.2.2 the DL applications to EEGs are described; and, finally, in 3.2.3 the DL frameworks used in problems involving *other biosignal* types are examined.

3.2.1. ECG

ECGs measure the electrical signals produced by a human heart. As discussed in the introduction to this subsection, DL provides promising results in the analysis of ECG data. The main DL applications in relation to ECGs are discussed in this Subsection, according to an organization based on the nature of the disease investigated: (i) *atrial fibrillation*, which is a particular abnormal heart rhythm; (ii) *cardiac arrhythmia*, which contains all other heart irregularities apart from atrial from the atrial fibrillation; and (iii) *sleep apnoea-glucose*, which relates to the analysis of the sleep apnoea and glucose.

In recognizing atrial fibrillation, the most common models that can be found in the literature are CNNs and RNNs: two remarkable examples can be found in [198], where Baalman et al. propose a RNN to separate different kinds of signals to identify atrial fibrillation, and in [199], where a CNN classification model is developed in order to study atrial fibrillation by using a particular database, which contains only low-dimensional and not-noisy data.

Also for identification of cardiac arrhythmia, as for atrial fibrillation, CNNs and RNNs are the most common DL models. As an example, in [200] the authors classify this heart problem by the use of a LSTM, while in [201] the LSTM is replaced by a CNN framework for the same purpose. Moreover, this kind of analysis can be shown to benefit from combining these two types of Neural Network, as reported in [202], where cardiac arrhythmias are studied by using a hybrid model, obtained as a combination of a CNN and a RNN, where the CNN extracts the features which are successively processed by the RNN.

Finally, CNNs have been successfully used in the classification of obstructive sleep apnoea and hypopnoea, as reported in [203], while in [204] a hybrid model, resulting from the combination of a CNN and a RNN, is constructed to provide an optimal classification of blood glucose levels.

From the application cases presented above, it emerges that CNNs and RNNs can be identified as reference models in dealing with biosignals, thanks to their ability in extracting features from data and in exploiting the internal dependencies of signals, respectively. Moreover, combining these characteristics in hybrid structures usually leads to enhanced results, often even better than those provided by just one of these applications individually.

3.2.2. EEG

EEGs relate to the set of all the brain's electrical signals. This type of biosignal is mainly used in the diagnosis and treatment of neurological diseases, and also in many other medical practices, such as rehabilitation [205]. As in relation to ECGs, the high variability of this kind of signal, as well as the potential presence of noise in the data, makes its analysis very challenging. Nevertheless, the use of DL in this field has proven to be effective in overcoming these issues. As is the case for the other types of signals, for EEGs we will identify some critical areas of investigation in which DL methodologies have been successfully applied in literature: (i) *sleep phase monitoring*, in which the signals under examination are collected while the patient sleeps, and (ii) *epilepsy*, where neural networks exploit signals connected to this neurological disorder which causes unprovoked, recurrent seizures.

In relation to the first area, an important experience is provided by Tsinalis et al. in [206] where a SAE is designed in order to provide a tool for automatic sleep stage scoring; this network, extracting the features by reducing the nosiness level, has been proven to outperform a classic AE. An alternative solution has been proposed in [206], where the authors combine a feature extraction CNN model with a RNN.

In the study of epilepsy signals, due to the importance of the sequentiality of data, RNNs result to be the most popular DL models [195,207]. An application an this type of data can be found in [207], where the authors combine the "memory features" of an LSTM with the ability of a CNN in recognizing important patterns, proposing a hybrid CNN+LSTM model.

3.2.3. Other biosignals applications

ECGs and EEGs are probably the two most significant biosignals on which DL techniques are applied with successful outcomes, but they are not the only ones. Good results have been obtained also in the case of many other kinds of biosignal, such as PCGs, PPGs, EMGs, SSs and NSSs. In the following paragraphs DL applications on these type of data are summarized and some interesting results are reported.

PCGs are heart signals obtained during a cardiac cycle. Thanks to their ability in the automatic extraction, CNNs are the most popular DL model applied on this type of signal, as, for example, in order to predict coronary artery disease as in [208], to analyse heart abnormalities as in [209] and to classify heart sounds as reported in [210].

With PPGs we consider the set of all the human signals linked to pulsations of blood. As in ECGs, there are two DL models which are commonly applied for the study of this kind of data: LSTM neural

Table 1
For each medical area this table lists the DL task, the obtained contributions and the most frequently used DL models.

Area	Task	Contribution	Model
Oncology	Classification	Benefit from unlabelled data for lung tumour stratification	DBN [115]
		Introduction of a transfer learning approach in rectal cancer prediction	CNN [118]
		Identification of bladder tumour sub-types from histopathological images	ResNet [114]
		Improvement in breast tumour estimation by considering a large set of risk factors	CNN [187]
		Estimation of the cancer grade	CNN [116]
		Estimation of the cancer type	CNN [111,112], ResNet [119]
	Detection	Optimal localization of lung cancer sub-types	CNN [113]
	Segmentation	Analysis of colour contrast and parameter variability issues in pancreatic tumour	U-Net [127]
		Impact of dimension variations on DL model performance in thyroid melanomas	U-Net [128]
		Limitation of the overfitting problem in bone cancer	CNN [188], GAN+U-Net [91]
		Improvement in image accuracy in lung and prostate cancer	U-Net [124,125], GAN [92]
	Registration	Optimized DL model in terms of time complexity and accuracy in lung melanoma estimation	CNN [129,130]
	Dose estimation	Estimation of the right substance doses	GAN [131,132]
Cardiovascular	Classification	Limitation of overfitting	GAN [142], ResNet [140]
		Analysis of the particular characteristics of the heart by using echocardiograms	ResNet [141]
	Detection	Low-cost object detection for malaria	YOLO [143]
	Segmentation	DL model for multi-step integration and registration error reduction in atrial fibrillation analysis	CNN+LSTM [189]
		Accuracy in the analysis of irregular pelvic hematoma images	U-Net [190]
		Improvement in aortic disease analysis with the introduction of new accuracy measures	U-Net [144]
		Introduction of the transfer learning approach in atrium study	U-Net [140]
	Reconstruction	Improvement in the Signal-to-Noise Ratio Multi-data integration	CNN [145]
		Improvement in image quality at high levels in the study of coronary diseases	CNN [146]
Orthopaedics	Classification	Improvement in bone image quality	U-Net [153]
		Analysis of the impact of gender on skeletal muscles	CNN [154]
	Segmentation	Analysis of the impact of the image quality in osteoarthritis	U-Net [149], RCNN [152]
		Introduction of transfer learning and attention mechanism in the study of the knees	VGGNet+U-Net [150]
		Improvement in image accuracy of the cartilage	U-Net [148], HNN [42], U-Net+GAN [151], RCNN
		Combination of the region-based approach with U-Net for bone diseases	RCC+U-Net [191]
	Reconstruction	Application of CNNs to computed tomography for chest digital images	CNN [192]
Neurology	Classification	Automatic estimation of brain diseases risk	AlexNet [158], CNN [193]
	Segmentation	Limitation of overfitting in White Matter analysis	GAN [160]
		Colour quality improvement in orbital analysis	U-Net [159]
	Reconstruction	Introduction of a DAE as a priori model for noise density in magnetic resonance	DAE [162]
		Analysis of perturbation effects	CNN [157]
		Introduction of transfer learning into magnetic resonance	CNN [163]
Pulmonology	Classification	Improvement of accuracy and efficiency in COP diseases	ResNet [167], VGGNet+CNN [168], DBN [169]
		Analysis of interstitial lung diseases	CNN [165]
	Segmentation	Segmentation of lung lob using different types of datasets	U-Net [170]
Gastroenterology	Classification	Estimation of the normal levels of the pancreas	CNN [172,174]
	Detection	Improvement in image accuracy in neoplasia analysis	ResNet [175]
	Segmentation	Analysis of image effects in neoplasia and catheter detection	U-Net [175], RNN [176]
Ophthalmology	Classification	Improvement in image quality	CNN [179], CNN+LSTM [181]
Abdominal	Classification	Improvement in accuracy in abdominal ultrasounds	CNN [185]
	Reconstruction	Limitation of overfitting	CNN+GAN [186]

networks, where the sequential nature of the data is of critical importance, as in the blood parameter quantitative estimation presented in [211], and CNNs, where feature extraction is a necessary task, as in the evaluation of blood glucose levels [212] or in the monitoring of the pulse [213].

EMGs measure the set of electrical signals of the skeletal muscles, where CNNs have been successfully applied in the analysis of the conditions of the muscles [214] and AEs have been used in gesture recognition problems [215].

Finally, SS and NSS consist in the set of signals coming from a brain spectroscopy and the set of signals produced by genetic molecules (DNA or RNA), respectively. In such cases, we can observe that the CNN approach is the principal technique through which these biosignals are studied as can be appreciated, for example, in [216,217].

The discussion presented in this subsection has investigated the main DL applications to biosignals, understood as the electrical and mechanical signals produced by the human body. Due to the intrinsic complexity of this kind of data, different strategies can be found in the literature, but all of them focusing on three main characteristics: when it is necessary to take into account sequential dependencies among data, the most common approach consists in adopting a LSTM network, thanks to its ability to store information from the past; when the extraction of data features is necessary, a CNN proves to be the most suitable kind of network; finally, in a case where the data present a high level of noise, the most common strategy is to apply an AE network. The findings of this subsection have been inserted into Table 2, which is structured as Table 1.

Table 2
For each medical area this table lists the DL tasks, the obtained contributions and the main used DL models.

Area	Task	Contribution	Model
ECG	Arrhythmia	Classification and detection of arrhythmia levels	RNN [198], LSTM [200], CNN [218], CNN+RNN [219]
	Glucose	Evaluation of low glucose levels	CNN [204]
	Ventricular	Data relative to different races	CNN [201]
	Atrial	Choice of low-dimensional datasets	CNN [199]
		Use of ECG with different lengths	CNN [220]
	Foetal	Restriction of data only one ECG channel	AE [221]
	Apnoea	Classification of obstructive sleep apnoea and hypopnoea	CNN [203]
EEG	Epilepsy	Multimodal approach	CNN+LSTM [222]
		Increase in the SNR	CNN [223]
	Emotions	Improvement in performance in heterogeneous samples	DL [222].
PCG	Heart abnormalities	Optimal identification of heart errors	CNN [208]
PPG	Blood analysis	Use of signals correlations	LSTM [211], CNN [212,213]
EMG	Muscle conditions	Improvements in signals accuracy	CNN [214], AE [215]
NNS, SS	Biosignals analysis	Accurate analysis of biosignals	CNN [216,217]

3.3. Biomedicine

Biomedicine is the set of disciplines which study the molecules involved in biological processes [224,225]. In recent years these disciplines are gaining popularity in the medical sector due to the role played by biomedical data, also named *omics data*, in the treatment of particular diseases, such as cancer and Alzheimer's disease [226–229]. Unfortunately, omics data generally experience problems on account of their high degree of heterogeneity and variability, related to the complexity of the genetic and/or molecular structures, which complicate the process of the analysis of this kind of information. In this context, DL offers convenient strategies in order to address such problems, therefore gaining an increasing popularity as a statistical model both to describe significant correlations among data and to extract information from them. In addition, the use of DL in biomedicine is suggested by the similarity between omics expressions and other kinds of data, such as images and texts, where DL has provided optimal results (see Section 2) [224].

Based on these premises, this part of the work will investigate the main applications and contributions of DL in this area under discussion. For the sake of clarity, the DL tasks in biomedicine have been organized into the following main groups: (i) *omics structure analysis*, which groups the applications related to the identification of particular omics regions and to the simplification of their gene expressions; (ii) *disease prediction*, in which the contributions of DL to phenotype analysis are described; and (iii) *drug discovery*, including the DL applications related to the analysis of new drugs.

These principal tasks will then be discussed for each of the main areas of application in the field of biomedicine, identified on the basis of processed data, an organization which also provides the structure of the following subsections: *Genomics*, which is presented in 3.3.1, *Transcriptomics* in 3.3.2, *Proteomics* in 3.3.3 and ***Metabolomics*** in 3.3.4.

3.3.1. Genomics

Genomics is the discipline which studies the human gene, with particular attention focused on the structure and properties of DNA. As emphasized in Section 3.3, genomics data can present a high level of heterogeneity which often proves impossible to analyse with classic statistical methods or even Machine Learning strategies. Thanks to its capability of integrating and dealing with different and complex data types, DL has been successfully applied in particular on two main group of tasks, namely *DNA structure analysis* and *disease prediction*, which are discussed in the following paragraphs.

In the DNA structure analysis task, RNNs have been applied to find the missing values of DNA expressions by exploiting the dependencies among the available molecules, as reported in [230]. Moreover, in [230] a GCNN was developed in order to identify particular DNA regions on the basis of the spatial configuration of the molecules. The problem of feature extraction from very complex data is instead addressed in [231], where Alzubaidi et al. demonstrate the feasibility of implementing a simplified approach to DNA structure analysis by means of an AE.

Regarding diseases prediction, AEs constitute the main class of model capable of reducing the high level of nosiness of omics data. As remarkable examples, in [231] a SAE model is applied to improve the accuracy of tumour prediction using molecules with a sparse topological structure, while in [232,233] the authors propose an AE model whose goal is to improve the accuracy of cancer prediction by integrating multi type omics data.

Another interesting application of the task under examination relates to the identification of cancer subtypes in order to personalize patient treatment. Such a type of problem is addressed in [44], where the authors achieve a high level of accuracy in the analysis of breast cancer subtypes by applying the Deep Triage model (see Section 2.5).

Briefly speaking, the most frequently used DL models related to DNA structure analysis and disease prediction prove to be RNNs, due to their ability to exploit molecule dependencies, AEs, used as knowledge extractors by isolating specific data features, and CNNs, which find application in collecting information from data by exploiting the molecular configuration. Moreover, in the literature, combinations of these structures through HAs can be found when it is necessary to combine their respective characteristics, as for example in the work of Guo et al. [234], where a CNN+RNN is designed to perform a parallel feature extraction in order to predict chromatin accessibility.

3.3.2. Transcriptomics

Transcriptomics is a discipline which studies the RNA transcripts generated by the human genome. In the following paragraphs the main DL tasks in this field are discussed, namely *RNA structure analysis*, *disease prediction* and *drug discovery*. In these types of problem, thanks to its ability to achieve high levels of abstraction in data feature extraction, DL has given promising results in terms of efficiency and flexibility [10].

In the case of RNA structure analysis, CNNs are usually applied in the simplification of RNA-structure [235] and in the classification of the RNA components. RNNs and AEs represent alternative models to CNNs, used, as example, in [236] by Deng et al. where a RNN which exploits the correlations among the molecules to classify the RNA sequences is proposed, or in [235,237], where an AE is used to extract data features from omics data by reducing the data dimensionality and sparsity.

Concerning disease prediction, CNNs have proven their ability in providing accurate results in phenotype analysis with the aim of classifying cancer types [238] and in recognizing the RNA variations related to the presence of various diseases, such as autism [239]. In addition, in [240] the authors propose an AE as an alternative to CNNs in breast cancer identification, the goal being to improve the accuracy of the results by integrating different kinds of transcriptomics data.

3.3.3. Proteomics

Proteomics is the discipline which studies all the biological processes related to proteins. In medicine it plays an important role not only due to the importance of these molecules in biological processes, but also because it makes an important contribution to drug discovery [241,242]. In particular, very interesting areas of proteomics related to the *analysis of the structure of a protein* and *the study of the target–drug interaction*, where DL is able to provide a great accuracy in the results, as described in the introduction to this section.

In molecular structure analysis, the main DL models encountered in the literature are AEs and CNNs on account of their ability as feature extractors [10,225,243]. Moreover, in [243] Torrisi et al. present an alternative DL model for the structure analysis, which consists in a hybrid model obtained by the combination of a CNN and a LSTM. In particular, this model is also applied in other areas closely connected to the examination of the protein structure, such as the analysis of the functions of a protein [10], where a LSTM model is able to benefit from the information deriving from interactions among the other levels of the protein structure.

CNNs have a great relevance also in the target–drug interaction task, where these structures are used to exploit the interactions among the molecules on the basis of their spatial configuration, as in [241]. Moreover, an augmented model for the same purpose is illustrated in [242], where a CNN is combined with a pre-training GAN in order to overcome the lack of labelled data.

To summarize, in proteomics, CNNs are applied in problems where it is necessary to consider the spatial configuration of the molecules, AEs are used as feature extractors in the presence of missing values while RNNs are models commonly applied in relation to problems where the sequential dependency among the different molecules is required.

3.3.4. Metabolomics

This subsection describes the main tasks of the metabolism, where DL is making a great contribution to the construction of efficient statistical models to predict metabolic diseases [244] and to assist experts in the process of drug discovery [229,245].

In this field a main area of DL application can be highlighted, namely drug discovery, where the most widely used DL models are AEs and CNNs, which are applied not only to study biological processes but also to identify potential novel drugs. In [246], Wang et al. propose an AE architecture to determine the metabolites of a drug, a strategy that can be used not only to find possible new drugs but also to decrease the toxicity level of a drug. In [247] a CNN model is presented to identify the categories of biological reactions occurring in a cell. An alternative to AEs and CNN is presented in [248], where the authors develop a general DL model which analyzes molecular interactions in order to determine the toxicity levels of a drug.

This subsection has discussed the main application fields of DL in biomedicine, emphasizing the role of the different models according to the specific task analysed. In this context, CNNs are used to examine the spatial configuration of the molecules, RNNs are applied to exploit the sequential dependencies among the parts of a gene expression, and AEs are used to isolate the data features by reducing the level of noisiness of the data. The results of this subsection are summarized in Table 3, whose structure is the same of Table 1.

3.4. Electronic Health Records (EHR)

EHR is the standard digital version of the patients' health records, with the goal of optimizing the diagnostic process and of providing accurate medical results. For this reason, every EHR is characterized by the presence of several kinds of data, in particular unstructured data (medical notes, images, etc.), which pose several issues in terms of the performance of traditional algorithms. As shown in [256–258], DL can handle these data structures, obtaining promising results in their analysis, thanks to its capability in achieving high levels of abstraction in the learning process and in integrating different data types.

Starting from the previous considerations, in the following paragraphs the main DL applications in relation to EHR are discussed, which have been classified into three main categories: (i) *disease prediction*, which contains all the DL applications relative to the analysis of diseases such as cancer and heart failure; (ii) *risk analysis*, which groups together all the DL applications addressing general issues relative to the analysis of EHR data; and (iii) *treatment plans*, which regards all the applications of DL in personalized medicine.

An in-depth study of the literature reveals that the main DL models associated with these tasks prove to be LSTM networks and CNNs: the former are widely applied in EHR because, as RNN models, they are able to exploit the strong sequential dependence existing among EH records; the latter are mainly used due to their ability in extracting data from complex and heterogeneous databases.

With reference to the main tasks outlined, in disease prediction problem LSTM-based structures are the most popular method. Some remarkable examples can be found in [51,259], where the authors propose a LSTM to investigate and so prevent the possibility of heart failure, in [260], in which a LSTM provides accurate results by combining both unstructured and structured data in the study of acute kidney injury, in [261], where Cui et al. address the problem of dyslipidemia prediction and show that a LSTM provides the most accurate results by comparing its outcomes with the ones of a basic RNN, or in [262], where a multitask model, composed of a LSTM and GRU, is developed in order to estimate the mortality risk in hospitals using patients' historical information.

Other common models used in disease prediction are CNNs: in [263] the problem of cancer prediction is addressed by developing a CNN which integrates clinical images and patients' drug doses; in [264] Wang et al. propose a CNN whose goal is to alleviate the imbalance problem (see Section 5) in heart failure prediction; and in [265] the authors apply a CNN whose goal is to evaluate the impact of external databases on the model performance in sleep staging analysis. In relation to risk analysis, some strategies aimed at minimizing the probability of error in the study of EHR data are based on the use of particular GC-NNs, as in [266], where the network models the data on the basis of their semantic sphere. Finally, regarding the third group of tasks, some examples of interesting DL applications are provided in [267], where a LSTM has been used to construct a treatment plan for non-patients by using large and heterogeneous samples, and in [6], where a DL model, with a CNN as the feature extractor model, has been proposed in order to determine the optimal treatment plan by exploiting the correlations among patients' data.

What emerges from the above discussion is that the strategies to deal with EHR data are mainly based on LSTM and CNN structures, sometimes in combination with GANs when facing problems related to the lack of data. This result is not unexpected since the analysis of this type of highly heterogeneous information easily benefits from the ability of a LSTM to handle sequential data (as, for example, in text analysis), as well as from the feature extraction characteristics of CNN models. To summarize the discussion here presented, the main results reported in this part have been collected in Table 4, whose structure is the same of Table 1.

This section has described the main DL applications in medicine, which have been categorized, on the basis of the data type involved, into four groups: *Clinical Images*, *Biosignals*, *Biomedicine* and *EHR*. Successively, for each class, the DL applications have been grouped into several tasks and, finally, the main DL models and contributions for each task have been detected. The most commonly applied models in medicine are CNNs because they are optimal feature extractors and are able to preserve the spatial configuration of the data (for this reason, they are widely used in relation to clinical images and biomedicine). RNNs are adopted when the data are characterized by a high level of sequential dependence, as in relation to biosignals,

Table 3
For each medical area this table lists the medical tasks, the contributions obtained and the DL model most frequently used.

Area	Task	Contribution	Model
Genomics	DNA structure	Prediction of DNA missing values from dependences	RNN [230]
		Identification of DNA regions by exploiting the spatial configuration	GCNN [230]
		Simplification of DNA expressions by reducing noisiness	AE+RNN [232], AE [231]
	Disease prediction	High accuracy in the identification of sub-kinds of tumour; Personalized Treatments	Deep Triage [44]
		Improvement in cancer prediction in very sparse molecules	SAE [231]
		Parallel extraction of features from pure DNA expressions	CNN+RNN [234]
		Data integration	DAE [233], AE [232]
Transcriptomics	RNA structure	Reduction of data dimensionality and sparsity	AE [235,237]
		Exploitation of the spatial configuration of RNA molecules	CNN [235]
		Accurate classification of the RNA components	CNN [249], RNN [236]
	Disease prediction	Classification of tumour types	CNN [238]
		RNA variation analysis	CNN [239]
		Heterogeneous data integration	SAE [240]
	Drug discovery	New drug–target interaction identification	DL [250]
Proteomics	Protein structure	Molecular region identification	CNN [251], DeepGSH [252]
		Protein identification	CNN+LSTM [253], AE [254]
	Drug discovery	drug–target interactions	GAN [242]
		Scoring function construction	CNN [241]
Metabolomics	Diseases prediction	Improvement of prediction models	CNN [244]
	Drug discovery	Determination of optimal targets; decrease in drug toxicity	AE [246,255]
		Optimal molecular interactions	CNN [247]

Table 4
This table lists the medical tasks, the contributions provided and the DL models most frequently used in EHR.

Area	Task	Contribution	Model
EHR	Disease prediction	Integration of different medical data in cancer analysis	CNN [263]
		High accuracy in dyslipidemia prediction	LSTM [261]
		Improvement in the imbalance problem in heart failure	CNN [264]
		Impact of external databases in sleep staging evaluation	CNN [265]
		Mortality risk estimation by using patients' historical information	LSTM [262]
		Optimal heart failure prediction	LSTM [51,259]
		Integration of structured and unstructured data for the prediction of acute kidney injury subtypes	LSTM [260]
	Risk analysis	Organization of the data on the basis of semantic spheres	GCNN [266]
		Feature extraction by reducing the data dimensions	AE [268]
	Treatment plans	Construction of a treatment plans by using small population sets	LSTM [267]
		Estimation of a treatment plan by exploiting data correlations	CNN [6]

biomedicine and, in particular, EHR. AEs are mainly applied when the data are characterized by a high level of noisiness, as in EHR. GANs are frequently adopted when the problem of overfitting holds, as in relation to clinical images.

4. Data structure and hyperparameters optimization

DL methodologies in medicine, as well as in all in the other fields where these techniques can be applied, often require the analysis of some "algorithmic" problems that can arise from data or from the algorithm itself. For these reasons, in the following Section a brief discussion on two of the main issues involved in the usage of Neural network approaches are discussed: on one side the format of stored data usually affect the class of algorithms that can be used on a specific problem, so in Section 4.1 a concise description of different data structures categorized for typical challenges faced in medicine is provided; on the other side, in Section 4.2 some strategies for the hyperparameters optimization problem are discussed, since designing reliable procedures in finding sub-optimal hyperparameter configurations regulating the execution of a DL algorithm are often necessary to explore the possibilities of application of such algorithms to a real problem.

4.1. Data structures

ML, and in particular DL models, are based on the idea of designing algorithms capable of reconstructing the complex relations among the variables regulating a certain phenomenon, without any *a priori* assumption about its shape, extracting patterns and regularities of the process directly from data. In this context, the way information is stored becomes a relevant factor in terms of choosing the approach to address the problem, sometimes providing a preliminary selection of the family of models that can be used. Based on these considerations, this section provides a categorization of the data structures that can usually be found in relation to problems in medicine and can be addressed with DL algorithms.

It is worth emphasizing that data regarding the same field can be stored in different shapes according to different systems. Moreover, many techniques in changing the point of view on problems through data reshaping are present in the literature. Therefore, the categorization here provided should be understood as a guideline for the reader in terms of framing the problem that she/he is dealing with.

1-D structures

One-dimensional data structures are represented by the vectors $\mathbf{v} \in \mathbb{R}^n$ and often contain information about the temporal evolution of a dependent variable with respect to an independent one. Many *biosignals* can be found in this shape, such as ECG, single-channel EEG, phonocardiograms and spectroscopy signals, and they are usually addressed as reported in Section 3.2, mainly using CNNs, RNNs, or hybrid networks built from these two in combination. Moreover, metabolic signals can also be found in this shape, as in [269,270] where LSTM neural networks are used to predict patients' blood glucose levels in

order to detect hyperglycaemia and hypoglycaemia, or in [271] where, for the same purpose, a CNN-LSTM neural network is applied.

2-D structures

This type of data structure is pervasive in the field of DL, frequently applied on medical problems since all static images are represented in this way. Bi-dimensional data are stored in matrices $\mathbf{M} \in \mathbb{R}^{n\times m}$, where each cell contains a real value. In the case of *k-bit* grey-scale images the cells contain a value between 0 and $2^k - 1$ which represents the brightness of the pixels of the image (e.g. the images produced through X-ray, ultrasound, MRI and scintigraphy). Problems involving this kind of data are often addressed through CNN-based architectures, as reported in Section 3.1. Moreover, this data structure can also be found in relation to information about *multivariate biosignals* (e.g. multi-channel EEG) or quantitative physiological indicators, as in [272] where the authors evaluate the acceleration of biological ageing due to tobacco on the basis of blood biochemical values using a Dense Neural Network.

3-D structures

Tensors $\mathbf{T} \in \mathbb{R}^{n\times m\times l}$ can be used to represent both 3-D graphic reconstructions, as in the case of *computed axial tomography*, in which the x-ray pictures are taken from different angles and assembled through computer algorithms to create 3-dimensional views of tissues and organs, and to store evolutionary behaviours of a 2-D image, as in the case of *Doppler ultrasonography*, which employs the Doppler effect to generate imaging of the movement of human tissues and body fluids, and their relative velocity to the probe. Since this type of structure principally relates to image-types of data, the DL techniques most frequently used in this context are based on CNN architectures, as in [273] where the authors propose a multi-projection methodology for 3-D image segmentation through a network developed from the structure of a Convolutional Autoencoder network, or in [274] where the authors provide a novel Cervical Intra-epithelial Neoplasia (CIN) grade and cervical cancer classification method analysing sequential and multi-state cervigram images through a CNN-LSTM based architecture.

Mixed tabular datasets

Tables containing numerical and categorical features together with free text fields can be widely found in the medical field since information regarding the clinical history of patients and administrative data are usually stored in this kind of structure. Extracting knowledge from mixed tabular datasets often proves to be a challenging task due to the heterogeneity of the data contained in them. If the dataset only contains real-valued and categorical fields, an option is to transform the last ones in numerical values through techniques such as *one-hot encoding* and reframe the problem as one dealing with a 2-D real-valued matrix. If the dataset also contains free texts, usually mixed architectures with components derived from *Natural Language Processing* DL techniques are usually used, as in [275] where a word embedding technique is used to achieve the semantic description of the content of EHR then exploited through a CNN architecture or in [276] where the automatic extraction of data relating to morbid conditions from death certificates is achieved through an *Encoder–Decoder* architecture for *seq2seq* problems.

In this section we have provided a brief summary of the different data structures that can be found in relation to medical problems which are addressed with DL approaches. We want to emphasize that this categorical structure has to be understood as purely indicative. Due to the great flexibility offered by DL methodologies, problems are often reframed in order to exploit some particular aspect of the dynamics under examination, new features are created and added to the original problem or data are aggregated and reshaped to highlight patterns and regularities. Nevertheless, the categorization given above offers a general view on how information is usually stored in accordance with different problems, providing also some examples of the methodologies applied in the literature on such data. The results of this subsection are summarized in Table 5, which is structured as follows: (i) the first column, headed *Structure*, reports the data structures analysed in this subsection; (ii) the second column, headed *Data*, reports the kinds of medical data to which DL models are applied; (iii) the third column, headed *Example*, reports examples of medical data to which DL models are applied; (iv) the fourth column, headed *Model*, reports the DL models applied to the different kinds of data.

4.2. Hyperparameter optimization in deep learning

The performance of ML approaches based on DL strongly depends on the choice of model hyperparameters with significant impact on the quality of the desired results and on the ability to correctly infer knowledge from novel information. In a huge number of applicative contexts, network features are tuned in a semi-automatic way. This makes published outcomes hard to reproduce and to generalize. In other words, it turns a strategy for the parameter selection of a designed model very close to being an art. This is particularly true in the medical research field where the results of a DL Neural Network need to be accurate and efficient. Discussing the general approaches to hyperparameter selection in-depth, the methodologies are grouped into two main classes: (i) hand-operated tuning; and (ii) automatic parameter selection algorithms.

Hand-operated hyperparameter tuning

The main target of manual hyperparameter optimization is to enhance the ability of a model: (i) to correctly represent the information; (ii) to help the learning algorithm to achieve the minimum of the cost function in the training step; and (iii) to regularize the overall learning process. A golden rule in a heuristic approach is, when it is possible, to follow the U-shaped curve for each hyperparameter. The high value of this function represents the underfitting regime, a low capacity of the model to learn with a high value of the training error. A small value of the U-shape gives low values of the training error but there is a gap between the generalization error and the training error. This is the overfitting problem. A good choice of a hyperparameter is, generally, a trade-off between the training and generalization error with relatively small gap. This last solution gives the optimal capacity of a model to learn. Obviously, the number of hyperparameters to manually select depends on the specific DL Network designed to solve a specific applicative problem and it may be extremely large. In the following paragraphs, we summarize the main parameters that influence the performance of a network starting from papers that have addressed the optimization task in the medicine applicative context. In particular, we consider: the Learning Rate (LR); the Hidden Layers (HL); the Convolutional Kernel Width (CKW); and the Dropout Rate (DR). The results of this subsection are summarized in Table 6, which is structured as follows: (i) the first column, headed *Value*, reports the main parameters treated in this subsection; (ii) the second column, headed *Description*, reports the description of every parameter; (iii) the third column, headed *Examples*, reports the main examples for every parameter treated in this subsection.

Automatic hyperparameter selection strategies

A desirable feature of a learning algorithm is the development of a procedure able to optimize automatically the hyperparameters of a model. This is an open challenge in ML research since the strategies, the targets, the input data types and the number of features vary considerably from one model to another. For these reasons, there is no unique best option for automatic hyperparameter selection in a specific medical application. A general technique to automatically optimize hyperparameters consists of considering the so-called evolutionary algorithms. Inspired by the Darwinian genetic selection, these

Table 5
Summary table of the different data structures, with some examples relating to the medical field and DL methodology applied.

Structure	Data	Example	Model
1-D $\mathbf{v} \in \mathbb{R}^n$	Time series, univariate data	ECG, single-channel EEG, phonocardiograms, spectroscopy signals, metabolic signals	LSTM [269,270], CNN-LSTM [271]
2-D $\mathbf{M} \in \mathbb{R}^{n \times m}$	Images, multivariate biosignals, physiological indicators	X-ray images, ultrasound images, MRI images, scintigraphy images, blood biochemical values	CNN-based architectures, MLP [272]
3-D $\mathbf{T} \in \mathbb{R}^{n \times m \times l}$	Tridimensional images, evolutionary images	Computed axial tomography, Doppler ultrasonography, sequential images	CNN [273], CNN-LSTM [274]
Mixed dataset	real-valued + categorical + free text data	EHR, Administrative Data	NLP + CNN [275], LSTM Encoder–Decoder [276]

Table 6
Remarks on manual hyperparameter selection.

Value	Description	Examples
LR	The LR defines how a network adjusts weights during training following suitable decisions. For LR values which are too small, the optimization task to minimize the loss function converges very slowly with tiny changes in the model weights. Conversely, with high values of LR the optimizer may overshoot the minimum and/or give convergence problems	In [277], a CNN has been designed to classify diseases in brain MRI images. The initial training of the model consisted in randomly setting the LR to a starting bound, with the model being tuned with a small number of epochs in order to reduce overfitting. Several restarting techniques were used to increase the accuracy in the training. Moreover, a data augmentation was adopted to retrain the model monitoring the LR by increasing the epochs.
HL	A HL strategy is to increase the HLs until the results improve. Unfortunately, the time and the memory cost of the implemented network increases dramatically. A trade-off between the efficiency of the computational resources and accuracy is needed.	A stacked denoising auto-encoder strategy is used in [278] to detect breast lesions in pulmonary nodules in CT images. Various nodules and lesions in the lungs are easily visible by using only a first hidden layer and by adding the second hidden layer. The network automatically encodes several hierarchical relations and hidden patterns.
CKW	CKW is used to increase the number of model parameters. It provokes a restricted output dimension, decreasing the capability of a model. CLW generally requires more memory to store the parameters and grow the runtime; a narrower output of the network gives a cheaper memory cost.	The effects of the choice of CKW and the methods to find a good trade-off between memory cost and number of parameters is reported in [279], where a CNN is adopted to deal with radiological tasks. This review describes techniques to address the overfitting problem and its potential benefits in diagnostic radiology.
DR	The DR value is generally used to regularize a model to reduce overfitting and improve the generalization error. DR is used to eliminate neurons from the network during training and it is often considered for network management.	The effect of DR on a multi-label classification problem from medical records is reported in [280] for several kinds of networks, as for example RNN and/or MLP.

procedures adopt scoring models to select the "best" parameter values of a model. This approach is likewise named a Genetic Algorithm (GA). In [281], a GA to implement this optimization task on a 3-layer CNN has been discussed. In the deep learning framework, two widely adopted procedures to determine hyperparameters are the grid search (GS) and the random search, presented in the second column of Table 7. The main difference between these two strategies is given by the rules adopted for varying the parameter values. An exhaustive dissertation about these algorithms is reported in the work of Bengio et al. in [282]. We will try to summarize the most common algorithms that are used in DL within the framework of medicine. In particular, we will consider: Grid Search (GS); Random Search (RC); and Bayesian Optimization (BO). The results of this subsection are summarized in Table 7, which is structured as follows: (i) the first column, headed *Value*, reports the main algorithms treated in this subsection; (ii) the second column, headed *Description*, reports the description of every algorithm; (iii) the third column, headed *Examples*, reports the main examples for every algorithm treated in this subsection.

5. Challenges and future research directions

DL is becoming the new paradigm in the analysis of medical data, as confirmed by the results discussed in Section 3. In addition, in recent years many other medical fields are beginning to benefit from the ability of these models to extract information from very different kinds of data. However, the complexity of DL models, the heterogeneity of medical data and the necessary interaction between machines and humans pose several issues, which must be taken into account in any assessment of future developments of DL in medicine. Starting from the previous considerations, this section overviews the main challenges and the future directions of DL in medicine. It is organized as follows: Section 5.1 addresses challenges in medicine, and Section 5.2 the main future directions of DL in medicine.

5.1. Challenges

DL has provided promising results in the processing of large amounts of noisy data, as in the case of biomedicine and biosignals, in extracting information from clinical images and in combining unstructured and structured data, as in EHR. However, some problems still arise from the use of DL in medicine, which prevent optimal results from being achieved. Starting from the previous considerations and from an in-depth study of the literature, in the following paragraphs we will investigate the main issues relative to DL, which have been classified into four parts: (i) *data quality*, which covers all the issues relative to the nature of medical data; (ii) *data distribution*, which covers all the issues relative to the quantity and the distribution of data in the input database; (iii) *model quality*, which covers the issues relative to the choice and validation of the model and to the lack of interpretability of the DL results; (iv) *explainable AI*, which covers all the issues relative to techniques to alleviate the black-box problem; and (v) *humans–machine interactions*, which covers the problems relative to the interactions among the figures involved in the diagnostic process. In detail, this subsection is structured as follows: in 5.1.1 data quality issues are discussed; in 5.1.2 data distribution issues; in 5.1.3 model quality issues; in 5.1.4 explainable AI; and, finally, in 5.1.5 humans–machines interactions issues.

5.1.1. Data quality

The strength of DL is its ability to extract information from large quantities of data, whose variability can impact on the levels of performance of DL. On the basis of this consideration, in this subsection we highlight the issues of every kind of medical data mentioned in Section 3.

Firstly, in the following paragraphs we focus on the two main issues relative to clinical images.

Table 7
Remarks on automatic hyperparameter selection.

Value	Description	Examples
GS	GS is generally applied for models described by a few parameters and is a brute force approach. One user assigns a small set of values for each parameter and tests all potential configuration within a range in a Cartesian grid. As the computational complexity grows exponentially with k hyperparameters that assume n values, the training and the evaluation of the cost function require $O(n^k)$ operations. In several cases it is also possible to deal with a great number of parameters with parallel computing strategies.	GPU implementation is reported in a multi-categorical network for classifying retinal images [283] GS ranges on a set of 10 categories were tested to achieve improvements in Random Forest and SVM algorithms. The application of the GS technique is described in the paper [284], where the detection and classification of cardiological arrhythmia levels from the ambulatory data of electrocardiograms are analysed.
RS	The RS strategy consists in selecting the values of the hyperparameters in a random way and it is typically favoured in the case of multiple features. RS is particularly useful when some hyperparameters influence the model learning capability more than others. This technique requires, for each parameter, the definition of a marginal probability distribution. However, it is not necessary to discretize the hyperparameter values. In this way it is possible to explore a large set of values without increasing the computational complexity of a mode	Uniform random exploration to study the impact of a combination of hyperparameters is reported in [285], with a classification of MRI data using deep learning. RS in this paper helps to minimize slice-wise classification errors and gives good results on a large MRI dataset. An open issue is to choose the probability distribution that fits with the hyperparameters. Several updating rules for an application of the RS strategy in the SVM classifiers for medical diagnosis are clearly described in [286].
BO	BO chooses the hyperparameters by using the function output in an iterative scheme in a global optimization strategy. BO is, generally, organized in two main subprocedures: first, hyperparameter values are generated randomly; in each iteration, a Bayesian probability model is computed to describe the conditional probability of the output values from the hyperparameters starting from past iterations.	A study on large scale cancer genomic profiling in [287] describes how BO is able to give good performances in terms of data classification, and, moreover an open source software is also given. In [288], BO is adopted to automatically optimize hyperparameters for CNN models on gastrointestinal datasets. This paper gives a powerful procedure to increase the performance in transfer learning in a large image dataset.

The first problem is that DL models require high storage resources to process clinical images, which impacts the space and time complexity of DL. This issue has been alleviated not only thanks to the introduction of very powerful hardware architectures, such as GPUs, but also by following new simpler approaches (for example, as discussed in [101], many $3D$-images are converted into $2D$-data).

The second problem consists in the lack of the papers dealing with colour analysis, which is a fundamental task especially in oncology in the classification of the kind and degree of a tumour [289].

Secondly, in the following paragraph the data quality problem in biosignals is described. These data provide the doctors with information only over a discrete set of instants, which cannot be generalized to a longer period because of the high degree of non-stationarity. This means that, in order to obtain accurate results, a patient should also be monitored after she/he is no longer in hospital. For this reason, in recent years Internet of Things (IoT) technologies are helping doctors in the monitoring of a patient's conditions when she/he stays at home. However, this poses a significant problem in terms of the reliability of the information obtained in this way, related to the possibility that a patient does not respect all the rules correctly.

Thirdly, the two main issues relative to biomedical data are investigated, where DL has provided good but not optimal results. The main problem is that the study of particular biological phenomena is not independent of manual intervention but still requires an expert's opinion. This aspect restricts the potential of DL, whose strength is to perform the process of the extraction of information from data automatically.

An additional problem is the high level of noisiness of the data, related to the presence of missing values and to their artificial origin (many biomedical data are produced in the laboratory). This problem has not only been alleviated but in fact solved completely by the use of DL, as discussed in 2.3.

Finally, in the following paragraph we will analyse the main issues relative to EHR. In addition to the presence of many unstructured data, as discussed in Section 2.4, the main problem is the absence of standardization protocols, those procedures in operation often relating not only to the specific rules of a particular hospital or department in the organization of a patient's data, but also to the doctor's personal practice in writing clinical notes. These issues, which complicate the merging of different kinds of medical data, can be addressed by creating public databases which share the information among departments and doctors. However, this poses several legal and ethical problems in relation to the type of data that can be shared and the possibility that the patient's privacy can be violated by hackers.

5.1.2. Data distribution

The large amount of data processed by DL poses not only problems in terms of data quality, described in 5.1.1, but also about data distribution, which are detailed in the following paragraphs.

The first problem is *overfitting*, which holds when the DL model underperforms because the number of model parameters is too high with respect to the available data. In medicine this phenomenon is very common when clinical images are involved because, differently from other image analysis tasks, the databases of the hospitals constitute the only data source and these contain an insufficient amount of data for a very complex model such as Dl. In particular, in some cases this lack of information is related to legal aspects, as in the USA, where the databases of the hospitals have an insufficient quantity of data because they do not contain information about people with no heath care. Several solutions have been proposed in order to overcome this issue, which are described in the following paragraph.

The first solution is the creation of public databases to share images among different hospitals but, as in the case of EHR, this poses several legal and ethical issues [290].

The second possible solution consists in an increase in the number of the input samples by generating new input instances or by integrating different kinds of medical data: in the first case, by using *data augmentation* techniques, such as generative models (DAE, GAN) [291] and *content-based image retrieval* (see Section 3.1), or *transfer learning*, in the second case by using a statistical approach called the *multimodal approach*.

The second problem relating to data distribution is the phenomenon of *class imbalance*, which holds when data categories, produced by a classification procedure, do not have the same dimension. Class imbalance introduces a loss of information in data analysis because the data contained in the smaller classes are interpreted by the model as noisy values. In medicine, two very common examples of class imbalance hold in the analysis of a tumour, where the number of sick people is generally higher than that of healthy people, and in drug discovery, where specific interactions between a target and a drug are not involved in the analysis process. In the literature stratification and data augmentation techniques have been proposed in order to alleviate the class imbalance problem [134,291].

Finally, the last issue relative to data distribution is *data leakage*, which holds when some inputs are shared among training and validation sets [292]. The consequence of data leakage is that the DL model underperforms because no complete distinction between the two sets exists. In medicine data leakage is generally generated by the presence of duplicates, as in biomedical applications, a pre-training step, as in

clinical images, or a historical dependence among data, as in biosignals and EHR.

5.1.3. Model quality

The multi-layer structure provides a DL model with the ability to achieve high levels of abstraction from data, but also poses several issues in terms of the quality of the model, which are presented in the following paragraphs.

First, the great quantity of data involved in the process increases the *time and space complexity* of a DL model, which could discourage experts from the use of DL in medical applications. As shown in [293], this issue can be alleviated by applying specific optimization techniques, such as parameters sharing and data augmentation

Secondly, the choice of a DL model is not univocal because it depends on the formulation of the problem. The consequence is that the selection of the best model is not always definitive, even if the characteristics of the processed data can suggest a suitable model. An example is represented by choice of a DL model in the features extraction when clinical images are processed: even if CNNs are widespread models for this task, however, AEs are widely used in the case of very noisy images, and GANs have a great ability to extract information by generating new images from the available input data.

Thirdly, we deal with the problem of *model validation*, which is based on the *formulation of a significant clinical question* and on the *management of the database*. The first point requires an interaction and a comprehension between data scientists and doctors, which could be difficult because of many doctors' limited knowledge about DL. The second point requires a database containing only information relative to the particular problem to solve, but the identification of the relevant information is often difficult due to the heterogeneity of databases and the high degree of correlation among the data.

Finally, in the following paragraph the last problem relative to model quality is analysed, namely *lack of model interpretability*, which is the impossibility of explaining the steps of the prediction problem (for this reason, a DL model is often indicated with the term *black-box*). In literature many authors have proposed several approaches to alleviate this problem, named *Explainable AI*, which are discussed in 5.1.4.

5.1.4. Explainable AI

Explainable AI (XAI) is a set of ML procedures which can provide more information about the steps involving the prediction process [294]. This type of system is attracting the interest of many researchers because, as discussed in 5.1.3, the black-box problem hurts the quality of a DL model. In the following XAI approaches are presented whose goal is to alleviate the lack of interpretability of two DL models, CNNs and RNNs [295].

In the case of CNN's, the principal XAI methods mainly alleviate the black-box problem by following three approaches: (i) identifying the information contained in the input data which play the key role in the production of the output, (ii) analysing how the internal layers interpret the external world, without considering any specific input, and (iii) localizing the hidden layers where the activity level is high.

In the case of RNNs, the principal XAI methods mainly alleviate the black-box problem by using two approaches: (i) choosing a set of time instants randomly and analysing what the model has learned until every instant; (ii) evaluating the output of a DL model obtained by using a new input space, generated by perturbing the original values.

As discussed previously, XAI is gaining great popularity in the scientific community thanks to the results produced in the improvement of the prediction quality. However, because of the complexity of DL architectures and of processed data, several issues must be addressed to reach the optimality of the results. In particular, as observed in [295], the improvement of the interpretability of an XAI model could generate a decrease in its accuracy and performance, whose consequence is a degradation of the output quality.

5.1.5. Human–machine interactions

Classic diagnosis can be interpreted as an interaction between a doctor and a patient, where the doctor identifies the disease and suggests a treatment, and the patient decides if she/he accepts or rejects the medical opinion. This paradigm could be transformed by the development of DL: in such a case, the patient will give her/his personal data to a machine, which will produce as an output a diagnosis and best treatment. However, this scenario involves several issues, which reduce the possibility of a complete automatic data analysis process.

First, DL is not able to benefit completely from some kinds of medical data, such as omics data, due to the presence of missing values and the necessity for manual intervention in the analysis of any genomic data interactions and EHR, on account of the high degree of data heterogeneity.

Secondly, another issue consists in the contrast between the natures of a human and a machine. A machine provides a diagnosis of a disease considering only a rational component, while a patient takes a decision considering both a rational and irrational component [296].

A possible solution to alleviate these issues is to introduce a new medical paradigm, where the doctor is helped by the machine in providing a patient with a precise diagnosis. However, this solution is limited by the doctors' general ignorance of DL (for this reason, several online versions of DL models have been implemented).

The results of this subsection have been included in Table 8, which is structured as follows: (i) the first column, headed *Challenge*, reports the challenges corresponding to the different parts of this subsection; (ii) the second column, headed *Issue*, contains all the problems relative to each kind of challenge discussed in this subsection; (iii) the third column, headed *Description*, provides a brief description of the DL issues recorded in the previous columns.

5.2. Future directions

DL in medicine is not limited to the traditional applications which have been described in Section 3 but, thanks to its ability to outperform more traditional ML methods in the analysis of high-dimensional and complex data samples, the interest in DL is extending to other medical fields, as confirmed by an in-depth study of the literature.

Starting from the previous considerations, in the following paragraphs the new frontiers of DL in medicine are presented, which have been classified into the following groups: (i) *Drug Discovery*, where DL contributes to the identification of new targets and new biomarkers; (ii) *Personalized Medicine*, where DL contributes to the construction of specific treatment plans for each patient; (iii) *Cancer Staging*, where DL contributes to the classification of the degree of cancer from histopathological images; (iv) *Neurology*, where DL contributes to the prediction of particular diseases, such as Alzheimer's disease and epilepsy.

Such groups have been extracted from a thematic evolution map (see Fig. 5) we obtained by analysing a dataset of papers related to the application of DL in medicine.

In detail, this subsection is organized as follows: in 5.2.1 the main contributions of DL to Drug Discovery are presented; in 5.2.2 the main contributions of DL to Personalized Medicine; in 5.2.3 the main contributions of DL to Cancer Staging; and, finally, in 5.2.4 the main contributions of DL to Neurology.

5.2.1. Drug discovery

DL in drug discovery is not a new research topic [297], but in the recent years it has experienced a widespread diffusion thanks to the availability of large amounts of drug data [298–300]. This aspect has encouraged the major pharmaceutical companies to invest large amounts of money in developing new DL models in this field, thereby making drug discovery one of the most important and emerging sectors for DL.

Starting from the previous considerations, in the following paragraphs we will detail the main tasks of DL in drug discovery, namely *target identification* and *drug design*.

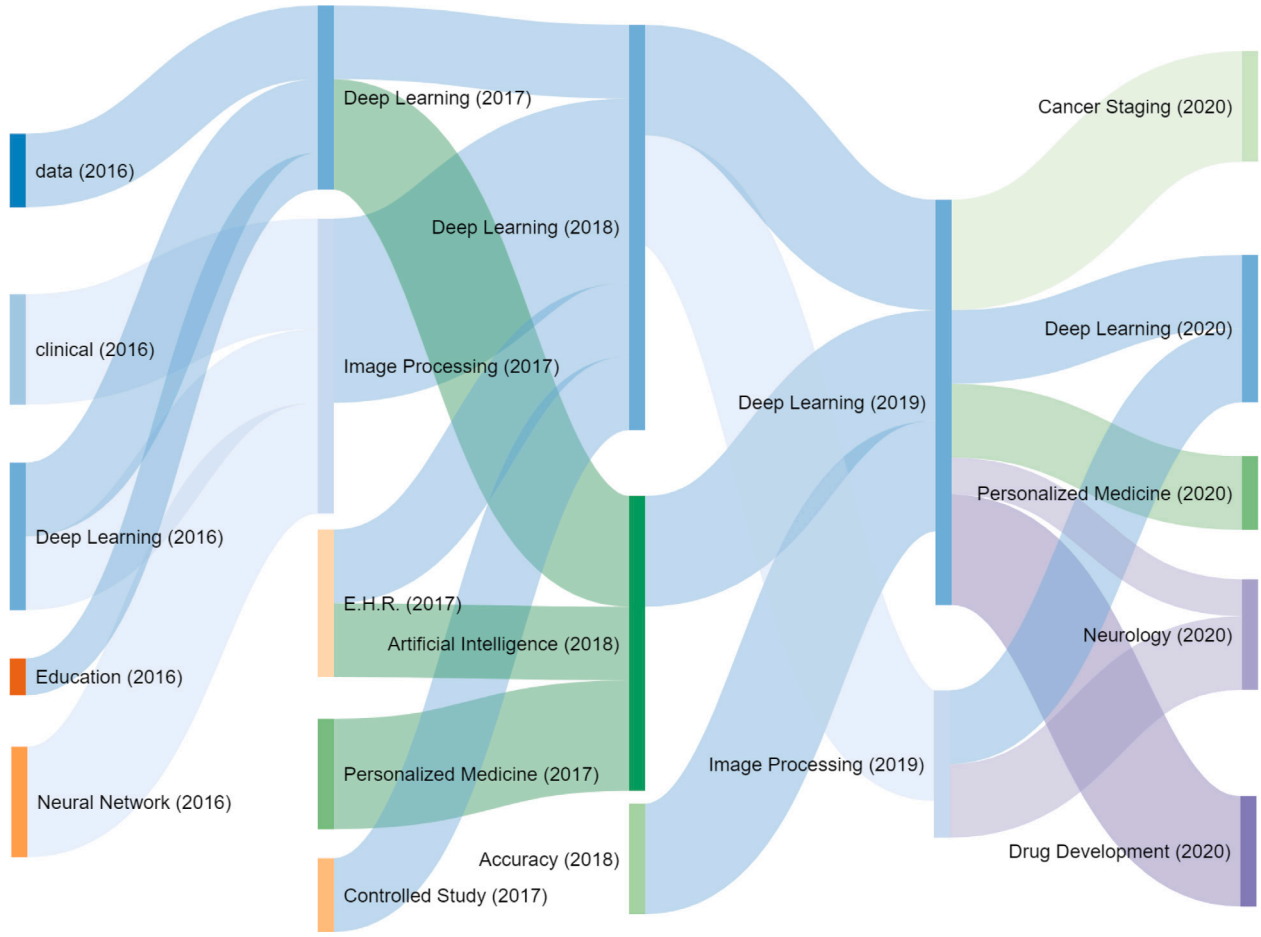

Fig. 5. A thematic evolution of Deep Learning in Medicine obtained through an analysis a dataset of papers related to the application of Deep Learning methodologies in Medicine. The map is based on co-word (keywords) network analysis and clustering.

Table 8
List of the kinds of DL challenge in medicine with details about the various related issues including a brief description.

Challenge	Issue	Description
Data quality	Colour analysis	Low number of papers about image colour problems
	Non-stationary	Information extracted from a biosignal is valid for a short time interval
	Feature extraction	Impossibility of performing the complete automatic features extraction step for biomedical data
	Unstructured data	Difficulty in gathering information from EHR unstructured data for the absence of a medical protocol standardization
Data distribution	Overfitting	Low availability of input data relating to the number of parameters
	Class imbalance	Classes with different dimensions
	Data leakage	Sharing of values between the training and validation sets
Model quality	Time and space complexity	Necessity of high storage and computational resources
	Model selection	Formulation of a significant clinical question; selection of an appropriate database
	Interpretability	Lack of information about the technological and logical principles of a process
Explainable AI	Set of techniques to alleviate the black-box issue	Identification of the most relevant input information, localization of the hidden layers with the highest activity, processing new perturbed data samples
Human–machine Interactions	Doctor vs. machine	Impossibility for the machine to provide a complete diagnosis without medical intervention
	Patient vs. machine	Irrationality of the patient's choices in opposition to the rationality of those made by the machine

First, in the following paragraphs the target identification task is analysed, which is the process to the interactions between a target and a disease. Recent scientific research has highlighted the role played by RNA and DNA sequences as an optimal kind of target and, consequently, the necessity to use sophisticated statistical models for the study of these data. As shown in 3.3.2, CNNs, RNNs and AEs have provided promising results in transcriptomics analysis thanks also to the integration of different kinds of omics data.

As a second step, we will consider DL in drug design, defined as the process of creating new drug candidates, starting from the analysis of drug–target interactions. The most common approaches in drug discovery are *de novo drug design* and *drug repositioning*, which are detailed in the following paragraphs.

De novo drug design consists in finding new drug candidates by reducing the use of information relative to existing drugs. This implies that this approach suffers from a problem of overfitting because of the low-dimensional data samples used in the drug discovery process. This issue is alleviated by applying HAs with generative models used at the pre-training step, as shown in Section 3.1.

In addition to de novo drug design, another approach is drug repositioning, which involves finding possible new applications for existing drugs. A crucial point in connection with drug repositioning is the study of the best orientation among the drug molecules, called *molecular docking*, which is performed by using particular mathematical functions, named *scoring functions*. The main drawback relative to molecular docking is the low number of available scoring functions, which results in the production of sub-optimal results in many specific cases. As shown in [301], DL outperforms all the existing statistical models in the creation of new scoring functions, which are able to capture the interactions existing among small molecules.

5.2.2. Personalized medicine

Personalized medicine deals with the creation of a specific treatment plan according to the patient's medical and biological characteristics [204,302]. On the basis of this definition, three main tasks of personalized medicine are highlighted: *disease classification*, *drug choice* and *dose estimation*. In the following paragraphs, the main contributions of DL to the previous tasks are presented.

In the disease classification task, the construction of an optimal treatment plan requires an accurate analysis of the specific characteristics of a disease, which can be obtained by using sophisticated technological tools. This aspect is very common in oncology, where the quality of treatment depends on the kind of tumour (benign or malign), on its grade and on the correct identification of the cancer sub-type. As shown in Section 3.1, CNNs are very common DL models for these tasks thanks to their great ability in extracting specific features from images characterized by a high degree of complexity.

Secondly, we will consider drug choice. This plays an important role in scientific research because it creates a bridge between personalized medicine and drug discovery and, consequently, a partnership among different kinds of scientists. As described in 5.2.1, DL is used as an efficient model for the identification of new drugs and in the analysis of target–disease interactions.

Finally, the dose estimation task is discussed. A central problem is personalized medicine is to determine the right drug to give a patient by using a set of available dose values. From a statistical point of view, this can be interpreted as the problem of estimating a probabilistic density from a prior distribution, given by the initial values. As discussed in Section 2, generative models, such as GANs and DAEs, have provided optimal results in sampling new input instances by processing large amounts of data.

5.2.3. Cancer staging

Cancer analysis is a central topic in medical research, whose main tasks are the classification and segmentation, as discussed in Section 2.1. However, thanks to the great technological developments of the last few years, new kinds of medical technique have been proposed in order to improve the accuracy of a diagnosis, such as *cancer staging*, which is the process of measuring the extension of a tumour. One of the main tools used in cancer staging is *histopathology*, which consists in the analysis of particular tissues via *computer-assisted diagnosis* (CAD) [117].

CNNs, and in particular AlexNet, ResNet and VGGNet, are becoming very popular in histopathological image analysis due to the high degree of complexity of these images and because, thanks to the use of GPUs, DL is able to reduce the time complexity with respect to classic ML algorithms. However, as in the case of clinical images, DL in histopathology suffers from a problem of overfitting, which can be alleviated by using classical data augmentation techniques.

5.2.4. Neurology

Neurology can be considered as one of the most interesting and promising medical sectors for DL thanks to its ability to process and integrate different kinds of data, such as clinical images and EEGs.

In addition, as shown in Section 3.3, DL provides promising results in the analysis of omics data, which are an important source of information for the prediction of several neurological diseases, such as Alzheimer's disease and cancer. In particular, DL is able to localize the presence of a tumour and to classify different kinds of cancer sub-types with a high degree of accuracy.

Moreover, DL is becoming very popular in several procedures whose goal is to prevent neurological diseases, namely *emotion recognition*, *mental workload*, *sleep stage scoring*, and *event related potential* [303]. In emotion recognition, DL transform face images and EEGs coming from people watching videos into values. In connection with mental workload, a DL classifies the level of the mind by using EEGs coming from people undertaking different mental tasks. In sleep stage scoring, DL identifies the different sleep stages by monitoring EEGs during the night in order to prevent possible diseases. In event related potential, particular letters or images stimulating a subject's brain are determined, which are analysed by DL in order to explore the working of the human brain.

6. Conclusions

DL is changing the cultural paradigm of medicine: its applications could become increasingly indispensable in terms of providing answers in contexts of high complexity and uncertainty and in order to allow doctors to have more time to take care of the medical needs of their patients. However, data are not values; any intervention based on data must be personalized, also taking into account the frequently contradictory nature of the knowledge provided in the literature. DL will be useful mainly insofar as it is complementary to the doctor, who will be able to delegate algorithmic operations to machines but will maintain control of the interpretation of complex phenomena and the consequent possible solutions [304]. DL systems must be considered as an instrument, like the microscope, phonendoscope and electrocardiogram, developed over time to compensate for the limitations in the perceptive capacity of doctors.

The best results are realized when DL works in support of health personnel, acting as a "second set of eyes", providing a means of cultural integration between humans and smart machines, thereby avoiding conflict, basically irrelevant, between cognitive, human and artificial methods, in short being more "intelligent". Physicians must, therefore, play the role of a guide, supervisor and monitor, exploiting the skills that make them superior to machines, in particular, abstraction, intuition, flexibility and empathy, the so-called soft skills [305]. In this way, they will be able to exercise a conservative and constructively critical approach, highlighting the enormous potential, often uncritically emphasized for commercial reasons, but also the limitations of DL (and, indeed, the possible threats, keeping in mind the science fiction dystopian nightmare of the "machines in power"!) [93]. This means, for example, detecting any lack of research studies on the effectiveness of DL in relation to critical clinical outcomes, such as reducing morbidity and mortality and improving the quality of life of patients. The objectives should also include assessing the level of satisfaction, both for doctors and patients, with the new relational context of the integration of the real and the digital world.

CRediT authorship contribution statement

Francesco Piccialli: Conceptualization, Methodology, Investigation, Writing, Visualization, Supervision. **Vittorio Di Somma:** Data curation, Writing, Investigation. **Fabio Giampaolo:** Writing, Resources, Review & editing. **Salvatore Cuomo:** Formal analysis, Writing - review & editing. **Giancarlo Fortino:** Writing - review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors dedicate this work to their friend and colleague Prof. Antonio Picariello, who passed away prematurely. The memory of his wonderful and kind soul will always remain in our hearts.

This work was supported by the *CUP-in-un-click* (CUP-in-One-Click) research, Italy project [Regione Campania - Bando RIS3 2018 - Fase 2]. The authors would like to thank the M.O.D.A.L. research laboratory (http://www.labdma.unina.it/index.php/modal/) for their efforts and support.

References

[1] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT press, 2016.
[2] L. Yue, D. Tian, W. Chen, X. Han, M. Yin, Deep learning for heterogeneous medical data analysis, World Wide Web (2020) 1–23, http://dx.doi.org/10.1007/s11280-019-00764-z.
[3] G. Fortino, D. Parisi, V. Pirrone, G. Di Fatta, Bodycloud: A saas approach for community body sensor networks, Future Gener. Comput. Syst. 35 (2014) 62–79.
[4] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, Nat. Med. 25 (1) (2019) 24–29.
[5] K. Gao, G. Mei, F. Piccialli, S. Cuomo, J. Tu, Z. Huo, Julia language in machine learning: Algorithms, applications, and open issues, Comp. Sci. Rev. 37 (2020) 100254, http://dx.doi.org/10.1016/j.cosrev.2020.100254, URL: http://www.sciencedirect.com/science/article/pii/S157401372030071X.
[6] D. Chen, S. Liu, P. Kingsbury, S. Sohn, C.B. Storlie, E.B. Habermann, J.M. Naessens, D.W. Larson, H. Liu, Deep learning and alternative learning strategies for retrospective real-world clinical data, NPJ Digit. Med. 2 (1) (2019) 1–5.
[7] A. Akay, H. Hess, Deep learning: current and emerging applications in medicine and technology, IEEE J. Biomed. Health Inf. 23 (3) (2019) 906–920.
[8] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[9] P. Meyer, V. Noblet, C. Mazzara, A. Lallement, Survey on deep learning for radiotherapy, Comput. Biol. Med. 98 (2018) 126–146.
[10] Z. Zhang, Y. Zhao, X. Liao, W. Shi, K. Li, Q. Zou, S. Peng, Deep learning in omics: a survey and guideline, Brief. Funct. Genom. 18 (1) (2019) 41–57.
[11] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Netw. 61 (2015) 85–117.
[12] M.A. Nielsen, Neural Networks and Deep Learning, Vol. 2018, Determination press, San Francisco, CA, USA, 2015.
[13] Y. LeCun, B.E. Boser, J.S. Denker, D. Henderson, R.E. Howard, W.E. Hubbard, L.D. Jackel, Handwritten digit recognition with a back-propagation network, in: Advances in Neural Information Processing Systems, 1990, pp. 396–404.
[14] G. Cheng, C. Yang, X. Yao, L. Guo, J. Han, When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs, IEEE Trans. Geosci. Remote Sens. 56 (5) (2018) 2811–2821.
[15] Z.-Q. Zhao, P. Zheng, S.-t. Xu, X. Wu, Object detection with deep learning: A review, IEEE Trans. Neural Netw. Learn. Syst. 30 (11) (2019) 3212–3232.
[16] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2017) 834–848.
[17] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: A comprehensive review, Neural Comput. 29 (9) (2017) 2352–2449.
[18] I. Ahmed, S. Din, G. Jeon, F. Piccialli, Exploring deep learning models for overhead view multiple object detection, IEEE Internet Things J. 7 (7) (2020) 5737–5744.
[19] Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, Q. Sun, Deep learning for image-based cancer detection and diagnosis- a survey, Pattern Recognit. 83 (2018) 134–149.
[20] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L. Fei-Fei, Large-scale video classification with convolutional neural networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 1725–1732, http://dx.doi.org/10.1109/CVPR.2014.223.
[21] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[22] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Vol. 9351, Springer, 2015, pp. 234–241.
[23] N. Akhtar, U. Ragavendran, Interpretation of intelligence in CNN-pooling processes: a methodological survey, Neural Comput. Appl. 32 (2019) 1–20.
[24] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
[25] N. Suda, V. Chandra, G. Dasika, A. Mohanty, Y. Ma, S. Vrudhula, J.-s. Seo, Y. Cao, Throughput-optimized OpenCL-based FPGA accelerator for large-scale convolutional neural networks, in: Proceedings of the 2016 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays, 2016, pp. 16–25.
[26] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 2015.
[27] D. Meng, L. Zhang, G. Cao, W. Cao, G. Zhang, B. Hu, Liver fibrosis classification based on transfer learning and FCNet for ultrasound images, Ieee Access 5 (2017) 5804–5810.
[28] Y. Yu, H. Lin, J. Meng, X. Wei, H. Guo, Z. Zhao, Deep transfer learning for modality classification of medical images, Information 8 (3) (2017) 91.
[29] Y. Tang, X. Wu, Scene text detection and segmentation based on cascaded convolution neural networks, IEEE Trans. Image Process. 26 (3) (2017) 1509–1520.
[30] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[31] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, X. Tang, Residual attention network for image classification, in: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Vol. 2017-January, 2017, pp. 6450–6458.
[32] S. Hershey, S. Chaudhuri, D. Ellis, J. Gemmeke, A. Jansen, R. Moore, M. Plakal, D. Platt, R. Saurous, B. Seybold, M. Slaney, R. Weiss, K. Wilson, CNN architectures for large-scale audio classification, in: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2017, pp. 131–135, http://dx.doi.org/10.1109/ICASSP.2017.7952132.
[33] H. Dong, G. Yang, F. Liu, Y. Mo, Y. Guo, Automatic brain tumor detection and segmentation using u-net based fully convolutional networks, in: Annual Conference on Medical Image Understanding and Analysis, Vol. 723, Springer, 2017, pp. 506–517.
[34] Ö. Çiçek, A. Abdulkadir, S.S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, in: LNCS, vol. 9901, Springer, 2016, pp. 424–432.
[35] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, in: LNCS, vol. 11045, Springer, 2018, pp. 3–11.
[36] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.
[37] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 580–587, http://dx.doi.org/10.1109/CVPR.2014.81.
[38] J. Dai, Y. Li, K. He, J. Sun, R-fcn: Object detection via region-based fully convolutional networks, in: Advances in Neural Information Processing Systems, 2016, pp. 379–387.
[39] R. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based convolutional networks for accurate object detection and segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 38 (1) (2015) 142–158.
[40] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 779–788.
[41] M.A. Al-antari, T.-S. Kim, Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital x-ray mammograms, Comput. Methods Programs Biomed. (2020) 105584.
[42] R. Cheng, N.A. Alexandridi, R.M. Smith, A. Shen, W. Gandler, E. McCreedy, M.J. McAuliffe, F.T. Sheehan, Fully automated patellofemoral MRI segmentation using holistically nested networks: Implications for evaluating patellofemoral osteoarthritis, pain, injury, pathology, and adolescent development, Magn. Reson. Med. 83 (1) (2020) 139–153.
[43] J. Zhou, O.G. Troyanskaya, Predicting effects of noncoding variants with deep learning–based sequence model, Nat. Methods 12 (10) (2015) 931–934.
[44] A. Beykikhoshk, T.P. Quinn, S.C. Lee, T. Tran, S. Venkatesh, Deeptriage: interpretable and individualised biomarker scores using attention mechanism for the classification of breast cancer sub-types, BMC Med. Genom. 13 (3) (2020) 1–10.
[45] L. Yao, C. Mao, Y. Luo, Graph convolutional networks for text classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 7370–7377.
[46] M. Zitnik, M. Agrawal, J. Leskovec, Modeling polypharmacy side effects with graph convolutional networks, Bioinformatics 34 (13) (2018) i457–i466.
[47] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536.

[48] T. Mikolov, M. Karafiát, L. Burget, J. Černockỳ, S. Khudanpur, Recurrent neural network based language model, in: Proceedings of the 11th Annual Conference of the International Speech Communication Association, INTERSPEECH 2010, 2010, pp. 1045–1048.

[49] X. Zhang, J. Zhao, Y. LeCun, Character-level convolutional networks for text classification, in: Advances in Neural Information Processing Systems, 2015, pp. 649–657.

[50] S. Lai, L. Xu, K. Liu, J. Zhao, Recurrent convolutional neural networks for text classification, in: Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.

[51] G. Maragatham, S. Devi, Lstm model for prediction of heart failure in big data, J. Med. Syst. 43 (5) (2019) 111.

[52] B.K. Reddy, D. Delen, Predicting hospital readmission for lupus patients: An RNN-LSTM-based deep-learning methodology, Comput. Biol. Med. 101 (2018) 199–209.

[53] A. Graves, A.-r. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2013, pp. 6645–6649.

[54] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[55] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.

[56] I. Sutskever, O. Vinyals, Q.V. Le, Sequence to sequence learning with neural networks, in: Advances in Neural Information Processing Systems, 2014, pp. 3104–3112.

[57] F. Piccialli, F. Giampaolo, G. Casolla, V.S.D. Cola, K. Li, A deep learning approach for path prediction in a location-based IoT system, Pervasive Mob. Comput. 66 (2020) 101210, http://dx.doi.org/10.1016/j.pmcj.2020.101210, URL: http://www.sciencedirect.com/science/article/pii/S1574119220300730.

[58] T. Pham, T. Tran, D. Phung, S. Venkatesh, Predicting healthcare trajectories from medical records: A deep learning approach, J. Biomed. Inform. 69 (2017) 218–229.

[59] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using RNN encoder-decoder for statistical machine translation, in: EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, Vol. 2018, 2014, pp. 1724–1734.

[60] T. Bansal, D. Belanger, A. McCallum, Ask the gru: Multi-task learning for deep text recommendations, in: Proceedings of the 10th ACM Conference on Recommender Systems, 2016, pp. 107–114.

[61] M. Schuster, K. Paliwal, Bidirectional recurrent neural networks, IEEE Trans. Signal Process. 45 (1997) 2673–2681, http://dx.doi.org/10.1109/78.650093.

[62] A. Graves, N. Jaitly, A.-r. Mohamed, Hybrid speech recognition with deep bidirectional LSTM, in: 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, IEEE, 2013, pp. 273–278.

[63] A. Graves, N. Jaitly, Towards end-to-end speech recognition with recurrent neural networks, in: International Conference on Machine Learning, 2014, pp. 1764–1772.

[64] D.H. Ackley, G.E. Hinton, T.J. Sejnowski, A learning algorithm for Boltzmann machines, Cogn. Sci. 9 (1) (1985) 147–169.

[65] A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis, Deep learning for computer vision: A brief review, Comput. Intell. Neurosci. 2018 (2018).

[66] H.-C. Shin, M.R. Orton, D.J. Collins, S.J. Doran, M.O. Leach, Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4d patient data, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2012) 1930–1943.

[67] M. Hayat, M. Bennamoun, S. An, Deep reconstruction models for image set classification, IEEE Trans. Pattern Anal. Mach. Intell. 37 (4) (2014) 713–727.

[68] Y. Bengio, I. Goodfellow, A. Courville, Deep Learning, Vol. 1, MIT press, 2017.

[69] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning Internal Representations by Error Propagation, Technical Report, California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

[70] L. Gondara, Medical image denoising using convolutional denoising autoencoders, in: 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), IEEE, 2016, pp. 241–246.

[71] T.M. Quan, T. Nguyen-Duc, W.-K. Jeong, Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss, IEEE Trans. Med. Imaging 37 (6) (2018) 1488–1497.

[72] N. Holzenberger, M. Du, J. Karadayi, R. Riad, E. Dupoux, Learning word embeddings: Unsupervised methods for fixed-size representations of variable-length speech segments, in: Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, Vol. 2018-September, 2018, pp. 2683–2687.

[73] M. Ribeiro, A.E. Lazzaretti, H.S. Lopes, A study of deep convolutional auto-encoders for anomaly detection in videos, Pattern Recognit. Lett. 105 (2018) 13–22.

[74] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.

[75] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, F.-Y. Wang, Generative adversarial networks: introduction and outlook, IEEE/CAA J. Autom. Sin. 4 (4) (2017) 588–598.

[76] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.

[77] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A.A. Bharath, Generative adversarial networks: An overview, IEEE Signal Process. Mag. 35 (1) (2018) 53–65.

[78] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved training of wasserstein GANs, Adv. Neural Inf. Process. Syst. 2017-December (2017) 5768–5778.

[79] G.E. Hinton, Learning multiple layers of representation, Trends Cogn. Sci. 11 (10) (2007) 428–434.

[80] G.B. Huang, H. Lee, E. Learned-Miller, Learning hierarchical representations for face verification with convolutional deep belief networks, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2012, pp. 2518–2525.

[81] R. Sarikaya, G.E. Hinton, A. Deoras, Application of deep belief networks for natural language understanding, IEEE/ACM Trans. Audio Speech Lang. Process. 22 (4) (2014) 778–784.

[82] M.M. Hassan, M.G.R. Alam, M.Z. Uddin, S. Huda, A. Almogren, G. Fortino, Human emotion recognition using deep belief network architecture, Inf. Fusion 51 (2019) 10–18, http://dx.doi.org/10.1016/j.inffus.2018.10.009, URL: http://www.sciencedirect.com/science/article/pii/S1566253518301301.

[83] C. Wang, F. Jiang, H. Yang, A hybrid framework for text modeling with convolutional RNN, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017, pp. 2061–2069.

[84] H. Li, P. Wang, C. Shen, Towards end-to-end text spotting with convolutional recurrent neural networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 5238–5246.

[85] C. Qin, J. Schlemper, J. Caballero, A.N. Price, J.V. Hajnal, D. Rueckert, Convolutional recurrent neural networks for dynamic MR image reconstruction, IEEE Trans. Med. Imaging 38 (1) (2018) 280–290.

[86] J. Cai, L. Lu, Z. Zhang, F. Xing, L. Yang, Q. Yin, Pancreas segmentation in MRI using graph-based decision fusion on convolutional neural networks, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 442–450.

[87] J. Yue-Hei Ng, M. Hausknecht, S. Vijayanarasimhan, O. Vinyals, R. Monga, G. Toderici, Beyond short snippets: Deep networks for video classification, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 4694–4702.

[88] N. Srivastava, E. Mansimov, R. Salakhudinov, Unsupervised learning of video representations using lstms, in: International Conference on Machine Learning, 2015, pp. 843–852.

[89] J. Geng, J. Fan, H. Wang, X. Ma, B. Li, F. Chen, High-resolution SAR image classification via deep convolutional autoencoders, IEEE Geosci. Remote Sens. Lett. 12 (11) (2015) 2351–2355.

[90] M. Zreik, N. Lessmann, R.W. van Hamersvelt, J.M. Wolterink, M. Voskuil, M.A. Viergever, T. Leiner, I. Išgum, Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis, Med. Image Anal. 44 (2018) 72–85.

[91] M. Caballo, D.R. Pangallo, R.M. Mann, I. Sechopoulos, Deep learning-based segmentation of breast masses in dedicated breast CT imaging: Radiomic feature stability between radiologists and artificial intelligence, Comput. Biol. Med. 118 (2020) 103629.

[92] Y. Lei, T. Wang, S. Tian, X. Dong, A.B. Jani, D. Schuster, W.J. Curran, P. Patel, T. Liu, X. Yang, Male pelvic multi-organ segmentation aided by CBCT-based synthetic MRI, Phys. Med. Biol. 65 (3) (2020) 035013.

[93] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities and challenges, Brief. Bioinform. 19 (6) (2018) 1236–1246.

[94] E.J. Topol, High-performance medicine: the convergence of human and artificial intelligence, Nat. Med. 25 (1) (2019) 44–56.

[95] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R.X. Gao, Deep learning and its applications to machine health monitoring, Mech. Syst. Signal Process. 115 (2019) 213–237.

[96] X. Du-Harpur, F. Watt, N. Luscombe, M. Lynch, What is AI? Applications of artificial intelligence to dermatology, Br. J. Dermatol. (2020) http://dx.doi.org/10.1111/bjd.18880.

[97] S. Kulkarni, N. Seneviratne, M.S. Baig, A.H.A. Khan, Artificial intelligence in medicine: where are we now? Acad. Radiol. 27 (1) (2020) 62–70.

[98] B.J. Erickson, P. Korfiatis, Z. Akkus, T.L. Kline, Machine learning for medical imaging, Radiographics 37 (2) (2017) 505–515.

[99] M.A. Mazurowski, M. Buda, A. Saha, M.R. Bashir, Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI, J. Magn. Reson. Imaging 49 (4) (2019) 939–954.

[100] I. Domingues, G. Pereira, P. Martins, H. Duarte, J. Santos, P.H. Abreu, Using deep learning techniques in medical imaging: a systematic review of applications on CT and PET, Artif. Intell. Rev. (2019) 1–68, http://dx.doi.org/10.1007/s10462-019-09788-3.

[101] A.S. Lundervold, A. Lundervold, An overview of deep learning in medical imaging focusing on MRI, Z. Med. Phys. 29 (2) (2019) 102–127.

[102] J. Liu, Y. Pan, M. Li, Z. Chen, L. Tang, C. Lu, J. Wang, Applications of deep learning to MRI images: A survey, Big Data Min. Anal. 1 (1) (2018) 1–18.

[103] W.L. Bi, A. Hosny, M.B. Schabath, M.L. Giger, N.J. Birkbak, A. Mehrtash, T. Allison, O. Arnaout, C. Abbosh, I.F. Dunn, et al., Artificial intelligence in cancer imaging: clinical challenges and applications, CA: Cancer J. Clin. 69 (2) (2019) 127–157.

[104] D.T. Hogarty, J.C. Su, K. Phan, M. Attia, M. Hossny, S. Nahavandi, P. Lenane, F.J. Moloney, A. Yazdabadi, Artificial intelligence in dermatology—Where we are and the way to the future: A review, Am. J. Clin. Dermatol. 21 (1) (2019) 1–7.

[105] K.-L. Hua, C.-H. Hsu, S.C. Hidayati, W.-H. Cheng, Y.-J. Chen, Computer-aided classification of lung nodules on computed tomography images via deep learning technique, OncoTargets Ther. 8 (2015) 2015–2022.

[106] R. Selvanambi, J. Natarajan, M. Karuppiah, S. Islam, M. Hassan, G. Fortino, Lung cancer prediction using higher-order recurrent neural network based on glowworm swarm optimization, Neural Comput. Appl. 32 (9) (2020) 4373–4386.

[107] F. Martino, S. Varricchio, D. Russo, F. Merolla, G. Ilardi, M. Mascolo, G. Dell'aversana, L. Califano, G. Toscano, G. De Pietro, M. Frucci, N. Brancati, F. Fraggetta, S. Staibano, A machine-learning approach for the assessment of the proliferative compartment of solid tumors on hematoxylin-eosin-stained sections, Cancers 12 (5) (2020).

[108] S. Chan, V. Reddy, B. Myers, Q. Thibodeaux, N. Brownstone, W. Liao, Machine learning in dermatology: Current applications, opportunities, and limitations, Dermatol. Ther. 10 (2020) 1–22.

[109] T.J. Kim, C.H. Kim, H.Y. Lee, M.J. Chung, S.H. Shin, K.J. Lee, K.S. Lee, Management of incidental pulmonary nodules: current strategies and future perspectives, Expert Rev. Respir. Med. 14 (2) (2020) 173–194.

[110] S.Z. Ramadan, Methods used in computer-aided diagnosis for breast cancer detection using mammograms: A review, J. Healthc. Eng. 2020 (2020).

[111] W. Lv, Y. Wang, C. Zhou, S. Huang, X. Fang, Q. Xu, Q. Zhang, C. Huang, X. Li, Z. Zhou, et al., Clinically applicable deep learning strategy for pulmonary nodule risk prediction: Insights into HONORS, medRxiv (2020).

[112] K. Kawauchi, S. Furuya, K. Hirata, C. Katoh, O. Manabe, K. Kobayashi, S. Watanabe, T. Shiga, A convolutional neural network-based system to classify patients using FDG PET/CT examinations, BMC Cancer 20 (1) (2020) 1–10.

[113] G. Savitha, P. Jidesh, A holistic deep learning approach for identification and classification of sub-solid lung nodules in computed tomographic scans, Comput. Electr. Eng. 84 (2020) 106626.

[114] A.-C. Woerl, M. Eckstein, J. Geiger, D.C. Wagner, T. Daher, P. Stenzel, A. Fernandez, A. Hartmann, M. Wand, W. Roth, et al., Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides, Eur. Urol. (2020) http://dx.doi.org/10.1016/j.eururo.2020.04.023.

[115] Z. Zhao, J. Zhao, K. Song, A. Hussain, Q. Du, Y. Dong, J. Liu, X. Yang, Joint DBN and fuzzy C-means unsupervised deep clustering for lung cancer patient stratification, Eng. Appl. Artif. Intell. 91 (2020) 103571.

[116] Z. Zhang, J. Xiao, S. Wu, F. Lv, J. Gong, L. Jiang, R. Yu, T. Luo, Deep convolutional radiomic features on diffusion tensor images for classification of glioma grades, J. Digit. Imaging (2020) 1–12, http://dx.doi.org/10.1007/s10278-020-00322-4.

[117] A. Das, M.S. Nair, S.D. Peter, Computer-aided histopathological image analysis techniques for automated nuclear atypia scoring of breast cancer: a review, J. Digit. Imaging (2020) 1–31, http://dx.doi.org/10.1007/s10278-019-00295-z.

[118] J. Fu, X. Zhong, N. Li, R. Van Dams, J. Lewis, K. Sung, A.C. Raldow, J. Jin, X.S. Qi, Deep learning-based radiomic features for improving neoadjuvant chemoradiation response prediction in locally advanced rectal cancer, Phys. Med. Biol. 65 (7) (2020) 075001.

[119] J. Zhou, Y. Zhang, K.-T. Chang, K.E. Lee, O. Wang, J. Li, Y. Lin, Z. Pan, P. Chang, D. Chow, et al., Diagnosis of benign and malignant breast lesions on DCE-MRI by using radiomics and deep learning with consideration of peritumor tissue, J. Magn. Reson. Imaging 51 (3) (2020) 798–809.

[120] W. Hua, T. Xiao, X. Jiang, Z. Liu, M. Wang, H. Zheng, S. Wang, Lymph-vascular space invasion prediction in cervical cancer: Exploring radiomics and deep learning multilevel features of tumor and peritumor tissue on multiparametric MRI, Biomed. Signal Process. Control 58 (2020) 101869.

[121] D. Dong, M.-J. Fang, L. Tang, X.-H. Shan, J.-B. Gao, F. Giganti, R.-P. Wang, X. Chen, X.-X. Wang, D. Palumbo, et al., Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multi-center study, Ann. Oncol. (2020) http://dx.doi.org/10.1016/j.annonc.2020.04.003.

[122] F. Xing, Y. Xie, H. Su, F. Liu, L. Yang, Deep learning in microscopy image analysis: A survey, IEEE Trans. Neural Netw. Learn. Syst. 29 (10) (2017) 4550–4568.

[123] B. Acs, M. Rantalainen, J. Hartman, Artificial intelligence as the next step towards precision pathology, J. Intern. Med. 288 (2020) 62–81.

[124] T. Nemoto, N. Futakami, M. Yagi, A. Kumabe, A. Takeda, E. Kunieda, N. Shigematsu, Efficacy evaluation of 2D, 3D U-Net semantic segmentation and atlas-based segmentation of normal lungs excluding the trachea and main bronchi, J. Radiat. Res. 61 (2) (2020) 257–264.

[125] N. Orlando, D.J. Gillies, I. Gyacskov, C. Romagnoli, D. D'Souza, A. Fenster, Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images, Med. Phys. 47 (6) (2020) 2413–2426.

[126] I.R.I. Haque, J. Neubert, Deep learning approaches to biomedical image segmentation, Inform. Med. Unlocked 18 (2020) 100297.

[127] T. Boers, Y. Hu, E. Gibson, D. Barratt, E. Bonmati, J. Krdzalic, F. van der Heijden, J. Hermans, H. Huisman, Interactive 3D U-net for the segmentation of the pancreas in computed tomography scans, Phys. Med. Biol. 65 (6) (2020) 065002.

[128] M. Buda, B. Wildman-Tobriner, K. Castor, J.K. Hoang, M.A. Mazurowski, Deep learning-based segmentation of nodules in thyroid ultrasound: Improving performance by utilizing markers present in the images, Ultrasound Med. Biol. 46 (2) (2020) 415–421.

[129] Z. Jiang, F.-F. Yin, Y. Ge, L. Ren, A multi-scale framework with unsupervised joint training of convolutional neural networks for pulmonary deformable image registration, Phys. Med. Biol. 65 (1) (2020) 015011.

[130] S.M. Galib, H.K. Lee, C.L. Guy, M.J. Riblett, G.D. Hugo, A fast and scalable method for quality assurance of deformable image registration on lung CT scans using convolutional neural networks, Med. Phys. 47 (1) (2020) 99–109.

[131] A. Babier, R. Mahmood, A.L. McNiven, A. Diamant, T.C. Chan, Knowledge-based automated planning with three-dimensional generative adversarial networks, Med. Phys. 47 (2) (2020) 297–306.

[132] S. Kazemifar, A.M. Barragán Montero, K. Souris, S.T. Rivas, R. Timmerman, Y.K. Park, S. Jiang, X. Geets, E. Sterpin, A. Owrangi, Dosimetric evaluation of synthetic CT generated with GANs for MRI-only proton therapy treatment planning of brain tumors, J. Appl. Clin. Med. Phys. 21 (5) (2020) 76–86.

[133] Y. Lee, K. Veerubhotla, M.H. Jeong, C.H. Lee, Deep learning in personalization of cardiovascular stents, J. Cardiovasc. Pharmacol. Ther. 25 (2) (2020) 110–120.

[134] D.D. Miller, Machine intelligence in cardiovascular medicine, Cardiol. Rev. 28 (2) (2020) 53–64.

[135] P. Bizopoulos, D. Koutsouris, Deep learning in cardiology, IEEE Rev. Biomed. Eng. 12 (2018) 168–193.

[136] C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, T. Kitai, Artificial intelligence in precision cardiovascular medicine, J. Am. Coll. Cardiol. 69 (21) (2017) 2657–2664.

[137] K.W. Johnson, J.T. Soto, B.S. Glicksberg, K. Shameer, R. Miotto, M. Ali, E. Ashley, J.T. Dudley, Artificial intelligence in cardiology, J. Am. Coll. Cardiol. 71 (23) (2018) 2668–2679.

[138] K. Naseer Qureshi, S. Din, G. Jeon, F. Piccialli, An accurate and dynamic predictive model for a smart m-health system using machine learning, Inform. Sci. 538 (2020) 486–502, http://dx.doi.org/10.1016/j.ins.2020.06.025.

[139] K.K. Wong, G. Fortino, D. Abbott, Deep learning-based cardiovascular image diagnosis: A promising challenge, Future Gener. Comput. Syst. 110 (2020) 802–811.

[140] H.-H. Chen, C.-M. Liu, S.-L. Chang, P.Y.-C. Chang, W.-S. Chen, Y.-M. Pan, S.-T. Fang, S.-Q. Zhan, C.-M. Chuang, Y.-J. Lin, et al., Automated extraction of left atrial volumes from two-dimensional computer tomography images using a deep learning technique, Int. J. Cardiol. (2020) http://dx.doi.org/10.1016/j.ijcard.2020.03.075.

[141] A. Ghorbani, D. Ouyang, A. Abid, B. He, J.H. Chen, R.A. Harrington, D.H. Liang, E.A. Ashley, J.Y. Zou, Deep learning interpretation of echocardiograms, NPJ Digit. Med. 3 (1) (2020) 1–10.

[142] L. Ma, R. Shuai, X. Ran, W. Liu, C. Ye, Combining DC-GAN with resnet for blood cell image classification, Med. Biol. Eng. Comput. 58 (6) (2020) 1251–1264.

[143] S. Chibuta, A.C. Acar, Real-time malaria parasite screening in thick blood smears for low-resource setting, J. Digit. Imaging 33 (3) (2020) 763–775.

[144] S. Hahn, M. Perry, C.S. Morris, S. Wshah, D.J. Bertges, Machine deep learning accurately detects endoleak after endovascular abdominal aortic aneurysm repair, JVS: Vasc. Sci. 1 (2020) 5–12.

[145] D. Xie, Y. Li, H. Yang, L. Bai, T. Wang, F. Zhou, L. Zhang, Z. Wang, Denoising arterial spin labeling perfusion MRI with deep machine learning, Magn. Reson. Imaging 68 (2020) 95–105.

[146] D.C. Benz, G. Benetos, G. Rampidis, E. Von Felten, A. Bakula, A. Sustar, K. Kudura, M. Messerli, T.A. Fuchs, C. Gebhard, et al., Validation of deep-learning image reconstruction for coronary computed tomography angiography: Impact on noise, image quality and diagnostic accuracy, J. Cardiovasc. Comput. Tomogr. (2020) http://dx.doi.org/10.1016/j.jcct.2020.01.002.

[147] T. Vu, M. Li, H. Humayun, Y. Zhou, J. Yao, A generative adversarial network for artifact removal in photoacoustic computed tomography with a linear-array transducer, Exp. Biol. Med. 245 (7) (2020) 597–605.

[148] M. Antico, F. Sasazawa, M. Dunnhofer, S. Camps, A. Jaiprakash, A. Pandey, R. Crawford, G. Carneiro, D. Fontanarosa, Deep learning-based femoral cartilage automatic segmentation in ultrasound imaging for guidance in robotic knee arthroscopy, Ultrasound Med. Biol. 46 (2) (2020) 422–435.

[149] A.S. Chaudhari, K.J. Stevens, J.P. Wood, A.K. Chakraborty, E.K. Gibbons, Z. Fang, A.D. Desai, J.H. Lee, G.E. Gold, B.A. Hargreaves, Utility of deep learning super-resolution in the context of osteoarthritis MRI biomarkers, J. Magn. Reson. Imaging 51 (3) (2020) 768–779.
[150] M. Byra, M. Wu, X. Zhang, H. Jang, Y.-J. Ma, E.Y. Chang, S. Shah, J. Du, Knee menisci segmentation and relaxometry of 3D ultrashort echo time cones MR imaging using attention U-Net with transfer learning, Magn. Reson. Med. 83 (3) (2020) 1109–1122.
[151] S. Gaj, M. Yang, K. Nakamura, X. Li, Automated cartilage and meniscus segmentation of knee mri with conditional generative adversarial networks, Magn. Reson. Med. 84 (1) (2020) 437–449.
[152] A. Jodeiri, R.A. Zoroofi, Y. Hiasa, M. Takao, N. Sugano, S. Yoshinobu, Y. Otake, Fully automatic estimation of pelvic sagittal inclination from anterior-posterior radiography image using deep learning framework, Comput. Methods Programs Biomed. 184 (2020) 105282.
[153] J. Guo, J. Zhu, H. Du, B. Qiu, A bone age assessment system for real-world x-ray images based on convolutional neural networks, Comput. Electr. Eng. 81 (2020) 106529.
[154] J. Xu, D. Xu, Q. Wei, Y. Zhou, Automatic classification of male and female skeletal muscles using ultrasound imaging, Biomed. Signal Process. Control 57 (2020) 101731.
[155] J. Huang, H. Shen, J. Wu, X. Hu, Z. Zhu, X. Lv, Y. Liu, Y. Wang, Spine explorer: a deep learning based fully automated program for efficient and reliable quantifications of the vertebrae and discs on sagittal lumbar spine mr images, Spine J. 20 (4) (2019) 590–599.
[156] E. Lin, C.-H. Lin, H.-Y. Lane, Precision psychiatry applications with pharmacogenomics: Artificial intelligence and machine learning approaches, Int. J. Mol. Sci. 21 (3) (2020) 969.
[157] Y. Li, H. Zhang, C. Bermudez, Y. Chen, B.A. Landman, Y. Vorobeychik, Anatomical context protects deep learning from adversarial perturbations in medical imaging, Neurocomputing 379 (2020) 370–378.
[158] Z. Yang, Z. Liu, The risk prediction of alzheimer's disease based on the deep learning model of brain 18f-fdg positron emission tomography, Saudi J. Biol. Sci. 27 (2) (2020) 659–665.
[159] R.A. Brown, D. Fetco, R. Fratila, G. Fadda, S. Jiang, N.M. Alkhawajah, E.A. Yeh, B. Banwell, A. Bar-Or, D.L. Arnold, et al., Deep learning segmentation of orbital fat to calibrate conventional mri for longitudinal studies, NeuroImage 208 (2020) 116442.
[160] K.T. Oh, S. Lee, H. Lee, M. Yun, S.K. Yoo, Semantic segmentation of white matter in fdg-pet using generative adversarial network, J. Digit. Imaging (2020) 1–10.
[161] Y. Han, J. Yoo, H.H. Kim, H.J. Shin, K. Sung, J.C. Ye, Deep learning with domain adaptation for accelerated projection-reconstruction mr, Magn. Reson. Med. 80 (3) (2018) 1189–1205.
[162] Q. Liu, Q. Yang, H. Cheng, S. Wang, M. Zhang, D. Liang, Highly undersampled magnetic resonance imaging reconstruction using autoencoding priors, Magn. Reson. Med. 83 (1) (2020) 322–336.
[163] S.U.H. Dar, M. Özbey, A.B. Çatlı, T. Çukur, A transfer-learning approach for accelerated mri using deep neural networks, Magn. Reson. Med. 84 (2) (2020) 663–685.
[164] G. Briganti, O. Le Moine, Artificial intelligence in medicine: Today and tomorrow, Front. Med. 7 (2020) 27.
[165] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, IEEE Trans. Med. Imaging 35 (5) (2016) 1207–1216.
[166] R.H. Abiyev, M.K.S. Ma'aitah, Deep convolutional neural networks for chest diseases detection, J. Healthc. Eng. 2018 (2018).
[167] L.Y. Tang, H.O. Coxson, S. Lam, J. Leipsic, R.C. Tam, D.D. Sin, Towards large-scale case-finding: training and validation of residual networks for detection of chronic obstructive pulmonary disease using low-dose ct, Lancet Digit. Health (2020).
[168] S. Bharati, P. Podder, M.H. Mondal, Hybrid deep learning for detecting lung diseases from x-ray images, Inform. Med. Unlocked (2020) 100391.
[169] J. Ying, J. Dutta, N. Guo, C. Hu, D. Zhou, A. Sitek, Q. Li, Classification of exacerbation frequency in the copdgene cohort using deep learning with deep belief networks, IEEE J. Biomed. Health Inform. (2016).
[170] J. Park, J. Yun, N. Kim, B. Park, Y. Cho, H.J. Park, M. Song, M. Lee, J.B. Seo, Fully automated lung lobe segmentation in volumetric chest ct with 3d u-net: validation with intra-and extra-datasets, J. Digit. Imaging 33 (1) (2020) 221–230.
[171] C. Le Berre, W.J. Sandborn, S. Aridhi, M.-D. Devignes, L. Fournier, M. Smaïl-Tabbone, S. Danese, L. Peyrin-Biroulet, Application of artificial intelligence to gastroenterology and hepatology, Gastroenterology 158 (1) (2020) 76–94.
[172] X. Wang, H. Qian, E.J. Ciaccio, S.K. Lewis, G. Bhagat, P.H. Green, S. Xu, L. Huang, R. Gao, Y. Liu, Celiac disease diagnosis from videocapsule endoscopy images with residual learning and deep feature extraction, Comput. Methods Programs Biomed. 187 (2020) 105236.
[173] W. Cao, X. An, L. Cong, C. Lyu, Q. Zhou, R. Guo, Application of deep learning in quantitative analysis of 2-dimensional ultrasound imaging of nonalcoholic fatty liver disease, J. Ultrasound Med. 39 (1) (2020) 51–59.
[174] S. Park, L. Chu, E. Fishman, A. Yuille, B. Vogelstein, K. Kinzler, K. Horton, R. Hruban, E. Zinreich, D.F. Fouladi, et al., Annotated normal ct data of the abdomen for deep learning: Challenges and strategies for implementation, Diagn. Interv. Imaging 101 (1) (2020) 35–44.
[175] A.J. de Groof, M.R. Struyvenberg, J. van der Putten, F. van der Sommen, K.N. Fockens, W.L. Curvers, S. Zinger, R.E. Pouw, E. Coron, F. Baldaque-Silva, et al., Deep-learning system detects neoplasia in patients with barrett's esophagus with higher accuracy than endoscopists in a multistep training and validation study with benchmarking, Gastroenterology 158 (4) (2020) 915–929.
[176] X. Yi, S. Adams, P. Babyn, A. Elnajmi, Automatic catheter and tube detection in pediatric x-ray images using a scale-recurrent network and synthetic data, J. Digit. Imaging 33 (1) (2020) 181–190.
[177] E. Long, H. Lin, Z. Liu, X. Wu, L. Wang, J. Jiang, Y. An, Z. Lin, X. Li, J. Chen, et al., An artificial intelligence platform for the multihospital collaborative management of congenital cataracts, Nat. Biomed. Eng. 1 (2) (2017) 1–8.
[178] M.M. Islam, H.-C. Yang, T.N. Poly, W.-S. Jian, Y.-C.J. Li, Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis, Comput. Methods Programs Biomed. 191 (2020) 105320.
[179] Y.-T. Hsieh, L.-M. Chuang, Y.-D. Jiang, T.-J. Chang, C.-M. Yang, C.-H. Yang, L.-W. Chan, T.-Y. Kao, T.-C. Chen, H.-C. Lin, et al., Application of deep learning image assessment software verisee for diabetic retinopathy screening, J. Formos. Med. Assoc. (2020).
[180] V. Sathananthavathi, G. Indumathi, A. Ranjani, Parallel architecture of fully convolved neural network for retinal vessel segmentation, J. Digit. Imaging 33 (1) (2020) 168–180.
[181] K. Zhang, X. Li, L. He, C. Guo, Y. Yang, Z. Dong, H. Yang, Y. Zhu, C. Chen, X. Zhou, et al., A human-in-the-loop deep learning paradigm for synergic visual evaluation in children, Neural Netw. 122 (2020) 163–173.
[182] S. Estrada, R. Lu, S. Conjeti, X. Orozco-Ruiz, J. Panos-Willuhn, M.M. Breteler, M. Reuter, Fatsegnet: A fully automated deep learning pipeline for adipose tissue segmentation on abdominal dixon mri, Magn. Reson. Med. 83 (4) (2020) 1471–1483.
[183] A.D. Missert, L. Yu, S. Leng, J.G. Fletcher, C.H. McCollough, Synthesizing images from multiple kernels using a deep convolutional neural network, Med. Phys. 47 (2) (2020) 422–430.
[184] E. Kang, J. Min, J.C. Ye, A deep convolutional neural network using directional wavelets for low-dose x-ray ct reconstruction, Med. Phys. 44 (10) (2017) e360–e375.
[185] P.M. Cheng, H.S. Malhi, Transfer learning with convolutional neural networks for classification of abdominal ultrasound images, J. Digit. Imaging 30 (2) (2017) 234–243.
[186] F.G. Zöllner, A. Šerifović-Trbalić, G. Kabelitz, M. Kociński, A. Materka, P. Rogelj, Image registration in dynamic renal mri—current status and prospects, Magn. Reson. Mater. Phys. Biol. Med. 33 (1) (2020) 33–48.
[187] D. Arefan, A.A. Mohamed, W.A. Berg, M.L. Zuley, J.H. Sumkin, S. Wu, Deep learning modeling using normal mammograms for predicting breast cancer risk, Med. Phys. 47 (1) (2020) 110–118.
[188] T. Apiparakoon, N. Rakratchatakul, M. Chantadisai, U. Vutrapongwatana, K. Kingpetch, S. Sirisalipoch, Y. Rakvongthai, T. Chaiwatanarat, E. Chuangsuwanich, Malignet: Semisupervised learning for bone lesion instance segmentation using bone scintigraphy, IEEE Access 8 (2020) 27047–27066.
[189] G. Yang, J. Chen, Z. Gao, S. Li, H. Ni, E. Angelini, T. Wong, R. Mohiaddin, E. Nyktari, R. Wage, et al., Simultaneous left atrium anatomy and scar segmentations via deep learning in multiview information with attention, Future Gener. Comput. Syst. 107 (2020) 215–228.
[190] D. Dreizin, Y. Zhou, Y. Zhang, N. Tirada, A.L. Yuille, Performance of a deep learning algorithm for automated segmentation and quantification of traumatic pelvic hematomas on ct, J. Digit. Imaging 33 (1) (2020) 243–251.
[191] F. Rehman, S.I.A. Shah, M.N. Riaz, S.O. Gilani, R. Faiza, A region-based deep level set formulation for vertebral bone segmentation of osteoporotic fractures, J. Digit. Imaging 33 (1) (2020) 191–203.
[192] T.-H. Kim, H. Oh, K. Kim, Y. Lee, Investigating single image super-resolution algorithm with deep learning using convolutional neural network for chest digital tomosynthesis, Optik 203 (2020) 164070.
[193] S. Saha, A. Pagnozzi, P. Bourgeat, J.M. George, D. Bradford, P.B. Colditz, R.N. Boyd, S.E. Rose, J. Fripp, K. Pannek, Predicting motor outcome in preterm infants from very early brain diffusion mri using a deep learning convolutional neural network (cnn) model, NeuroImage (2020) 116807.
[194] N. Ganapathy, R. Swaminathan, T.M. Deserno, Deep learning on 1-d biosignals: a taxonomy-based survey, Yearb. Med. Inform. 27 (01) (2018) 098–109.
[195] B. Rim, N.-J. Sung, S. Min, M. Hong, Deep learning in physiological signal data: A survey, Sensors 20 (4) (2020) 969.
[196] Z. Lv, L. Qiao, Q. Wang, F. Piccialli, Advanced machine-learning methods for brain-computer interfacing, IEEE/ACM Trans. Comput. Biol. Bioinform. (2020) 1.
[197] Y. Zhang, R. Gravina, H. Lu, M. Villari, G. Fortino, Pea: Parallel electrocardiogram-based authentication for smart healthcare systems, J. Netw. Comput. Appl. 117 (2018) 10–16.

[198] S.W. Baalman, F.E. Schroevers, A.J. Oakley, T.F. Brouwer, W. van der Stuijt, H. Bleijendaal, L.A. Ramos, R.R. Lopes, H.A. Marquering, R.E. Knops, et al., A morphology based deep learning model for atrial fibrillation detection using single cycle electrocardiographic samples, Int. J. Cardiol. (2020).

[199] S. Parvaneh, J. Rubin, A. Rahman, B. Conroy, S. Babaeizadeh, Analyzing single-lead short ecg recordings using dense convolutional neural networks and feature-based post-processing to detect atrial fibrillation, Physiol. Meas. 39 (8) (2018) 084003.

[200] K.-C. Chang, P.-H. Hsieh, M.-Y. Wu, Y.-C. Wang, J.-Y. Chen, F.-J. Tsai, E.S. Shih, M.-J. Hwang, T.-C. Huang, Usefulness of machine-learning-based detection and classification of cardiac arrhythmias with 12-lead electrocardiograms, Can. J. Cardiol. (2020).

[201] P.A. Noseworthy, Z.I. Attia, L.C. Brewer, S.N. Hayes, X. Yao, S. Kapa, P.A. Friedman, F. Lopez-Jimenez, Assessing and mitigating bias in medical artificial intelligence: the effects of race and ethnicity on a deep learning model for ecg analysis, Circ.: Arrhythm. Electrophysiol. 13 (3) (2020) e007988.

[202] Z. Xiong, M.P. Nash, E. Cheng, V.V. Fedorov, M.K. Stiles, J. Zhao, Ecg signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network, Physiol. Meas. 39 (9) (2018) 094006.

[203] E. Urtnasan, J.-U. Park, K.-J. Lee, Multiclass classification of obstructive sleep apnea/hypopnea based on a convolutional neural network from a single-lead electrocardiogram, Physiol. Meas. 39 (6) (2018) 065003.

[204] M. Porumb, S. Stranges, A. Pescapè, L. Pecchia, Precision medicine and artificial intelligence: A pilot study on deep learning for hypoglycemic events detection based on ecg, Sci. Rep. 10 (1) (2020) 1–16.

[205] F. Movahedi, J.L. Coyle, E. Sejdić, Deep belief networks for electroencephalography: A review of recent contributions and future outlooks, IEEE J. Biomed. Health Inform. 22 (3) (2017) 642–652.

[206] O. Tsinalis, P.M. Matthews, Y. Guo, Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders, Ann. Biomed. Eng. 44 (5) (2016) 1587–1597.

[207] M.-P. Hosseini, T.X. Tran, D. Pompili, K. Elisevich, H. Soltanian-Zadeh, Multimodal data analysis of epileptic eeg and rs-fmri via deep learning and edge computing, Artif. Intell. Med. 104 (2020) 101813.

[208] H. Li, X. Wang, C. Liu, Q. Zeng, Y. Zheng, X. Chu, L. Yao, J. Wang, Y. Jiao, C. Karmakar, A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection, Comput. Biol. Med. 120 (2020) 103733.

[209] V. Maknickas, A. Maknickas, Recognition of normal–abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients, Physiol. Meas. 38 (8) (2017) 1671.

[210] P.T. Krishnan, P. Balasubramanian, S. Umapathy, Automated heart sound classification system from unsegmented phonocardiogram (pcg) using deep neural network, Phys. Eng. Sci. Med. (2020) 1–11.

[211] M. Radha, K. De Groot, N. Rajani, C.C. Wong, N. Kobold, V. Vos, P. Fonseca, N. Mastellos, P.A. Wark, N. Velthoven, et al., Estimating blood pressure trends and the nocturnal dip from photoplethysmography, Physiol. Meas. 40 (2) (2019) 025006.

[212] M.M. Islam, S.M. Maniur, Design and implementation of a wearable system for non-invasive glucose level monitoring, in: 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), IEEE, 2019, pp. 29–32.

[213] J. Chen, H. Huang, W. Hao, J. Xu, A machine learning method correlating pulse pressure wave data with pregnancy, Int. J. Numer. Methods Biomed. Eng. 36 (1) (2020) e3272.

[214] N. Duan, L.-Z. Liu, X.-J. Yu, Q. Li, S.-C. Yeh, Classification of multichannel surface-electromyography signals based on convolutional neural networks, J. Ind. Inf. Integr. 15 (2019) 201–206.

[215] E.A. Chung, M.E. Benalcázar, Real-time hand gesture recognition model using deep learning techniques and emg signals, in: 2019 27th European Signal Processing Conference (EUSIPCO), IEEE, 2019, pp. 1–5.

[216] H.H. Lee, H. Kim, Intact metabolite spectrum mining by deep learning in proton magnetic resonance spectroscopy of the brain, Magn. Reson. Med. 82 (1) (2019) 33–48.

[217] H. Teng, M.D. Cao, M.B. Hall, T. Duarte, S. Wang, L.J. Coin, Chiron: translating nanopore raw signal directly into nucleotide sequence using deep learning, GigaScience 7 (5) (2018) giy037.

[218] Q. Li, C. Liu, Q. Li, S.P. Shashikumar, S. Nemati, Z. Shen, G.D. Clifford, Ventricular ectopic beat detection using a wavelet transform and a convolutional neural network, Physiol. Meas. 40 (5) (2019) 055002.

[219] S. Hong, Y. Zhou, M. Wu, J. Shang, Q. Wang, H. Li, J. Xie, Combining deep neural networks and engineered features for cardiac arrhythmia detection from ecg recordings, Physiol. Meas. 40 (5) (2019) 054009.

[220] R. Kamaleswaran, R. Mahajan, O. Akbilgic, A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length, Physiol. Meas. 39 (3) (2018) 035006.

[221] W. Zhong, L. Liao, X. Guo, G. Wang, Fetal electrocardiography extraction with residual convolutional encoder–decoder networks, Australas. Phys. Eng. Sci. Med. 42 (4) (2019) 1081–1089.

[222] J. Teo, C.L. Hou, J. Mountstephens, Preference classification using electroencephalography (eeg) and deep learning, J. Telecommun. Electron. Comput. Eng. (JTEC) 10 (1–11) (2018) 87–91.

[223] S.M. Usman, S. Khalid, M.H. Aslam, Epileptic seizures prediction using deep learning techniques, IEEE Access 8 (2020) 39998–40007.

[224] P. Mamoshina, A. Vieira, E. Putin, A. Zhavoronkov, Applications of deep learning in biomedicine, Mol. Pharm. 13 (5) (2016) 1445–1454.

[225] C. Cao, F. Liu, H. Tan, D. Song, W. Shu, W. Li, Y. Zhou, X. Bo, Z. Xie, Deep learning and its applications in biomedicine, Genom. Proteom. Bioinform. 16 (1) (2018) 17–32.

[226] W. Zhu, L. Xie, J. Han, X. Guo, The application of deep learning in cancer prognosis prediction, Cancers 12 (3) (2020) 603.

[227] R.S. Sealfon, L.H. Mariani, M. Kretzler, O.G. Troyanskaya, Machine learning, the kidney, and genotype-phenotype analysis, Kidney Int. 97 (6) (2020) 1141–1149.

[228] R. Dias, A. Torkamani, Artificial intelligence in clinical and genomic diagnostics, Genome Med. 11 (1) (2019) 1–12.

[229] L. Alberghina, G. Piccialli, From computational genomics to systems metabolomics for precision cancer medicine and drug discovery., Pharmacol. Res. 151 (2020) 104479.

[230] G. Eraslan, Ž. Avsec, J. Gagneur, F.J. Theis, Deep learning: new computational modelling techniques for genomics, Nature Rev. Genet. 20 (7) (2019) 389–403.

[231] A. Alzubaidi, J. Tepper, et al., A novel deep mining model for effective knowledge discovery from omics data, Artif. Intell. Med. 104 (2020) 101821.

[232] K. Chaudhary, O.B. Poirion, L. Lu, L.X. Garmire, Deep learning–based multi-omics integration robustly predicts survival in liver cancer, Clin. Cancer Res. 24 (6) (2018) 1248–1259.

[233] H. Failmezger, S. Muralidhar, A. Rullan, C.E. de Andrea, E. Sahai, Y. Yuan, Topological tumor graphs: a graph-based spatial model to infer stromal recruitment for immunosuppression in melanoma histology, Cancer Res. 80 (5) (2020) 1199–1209.

[234] Y. Guo, D. Zhou, R. Nie, X. Ruan, W. Li, Deepanf: A deep attentive neural framework with distributed representation for chromatin accessibility prediction, Neurocomputing 379 (2020) 305–318.

[235] M. Jost, D.A. Santos, R.A. Saunders, M.A. Horlbeck, J.S. Hawkins, S.M. Scaria, T.M. Norman, J.A. Hussmann, C.R. Liem, C.A. Gross, et al., Titrating gene expression using libraries of systematically attenuated crispr guide rnas, Nature Biotechnol. 38 (3) (2020) 355–364.

[236] Y. Deng, F. Bao, Q. Dai, L.F. Wu, S.J. Altschuler, Scalable analysis of cell-type composition from single-cell transcriptomics using deep recurrent learning, Nat. Methods 16 (4) (2019) 311–314.

[237] S. Zhang, X. Li, Q. Lin, J. Lin, K.-C. Wong, Uncovering the key dimensions of high-throughput biomolecular data using deep learning, Nucleic Acids Res. 48 (10) (2020) e56.

[238] N.E.M. Khalifa, M.H.N. Taha, D.E. Ali, A. Slowik, A.E. Hassanien, Artificial intelligence technique for gene expression by tumor rna-seq data: A novel optimized deep learning approach, IEEE Access 8 (2020) 22874–22883.

[239] J. Zhou, C.Y. Park, C.L. Theesfeld, A.K. Wong, Y. Yuan, C. Scheckel, J.J. Fak, J. Funk, K. Yao, Y. Tajima, et al., Whole-genome deep-learning analysis identifies contribution of noncoding mutations to autism risk, Nat. Genet. 51 (6) (2019) 973–980.

[240] Y. Guo, X. Shang, Z. Li, Identification of cancer subtypes by integrating multiple types of transcriptomics data with deep learning in breast cancer, Neurocomputing 324 (2019) 20–30.

[241] C. Shen, J. Ding, Z. Wang, D. Cao, X. Ding, T. Hou, From machine learning to deep learning: Advances in scoring functions for protein–ligand docking, Wiley Interdiscip. Rev.: Comput. Mol. Sci. 10 (1) (2020) e1429.

[242] L. Zhao, J. Wang, L. Pang, Y. Liu, J. Zhang, Gansdta: Predicting drug-target binding affinity using gans, Front. Genet. 10 (2019).

[243] M. Torrisi, G. Pollastri, Q. Le, Deep learning methods in protein structure prediction, Comput. Struct. Biotechnol. J. (2020) http://dx.doi.org/10.1016/j.csbj.2019.12.011.

[244] M. Cuperlovic-Culf, Machine learning methods for analysis of metabolic data and metabolic pathway modeling, Metabolites 8 (1) (2018) 4.

[245] A. Dasgupta, N. Chowdhury, R.K. De, Metabolic pathway engineering: Perspectives and applications, Comput. Methods Programs Biomed. 192 (2020) 105436.

[246] D. Wang, W. Liu, Z. Shen, L. Jiang, J. Wang, S. Li, H. Li, Deep learning based drug metabolites prediction, Front. Pharmacol. 10 (2020).

[247] M. Baranwal, A. Magner, P. Elvati, J. Saldinger, A. Violi, A.O. Hero, A deep learning architecture for metabolic pathway prediction, Bioinformatics 36 (8) (2020) 2547–2553.

[248] N.L. Dang, M.K. Matlock, T.B. Hughes, S.J. Swamidass, The metabolic rainbow: Deep learning phase i metabolism in five colors, J. Chem. Inf. Model. 60 (3) (2020) 1146–1164.

[249] H. Tayara, M. Tahir, K.T. Chong, Identification of prokaryotic promoters and their strength by integrating heterogeneous features, Genomics 112 (2) (2020) 1396–1403.

[250] L. Xie, S. He, X. Song, X. Bo, Z. Zhang, Deep learning-based transcriptome data classification for drug-target interaction prediction, BMC Genom. 19 (7) (2018) 667.

[251] N. Thapa, S. Hiroto, K. Roy, R.H. Newman, K. Dukka, et al., Rf-malosite and dl-malosite: Methods based on random forest and deep learning to identify malonylation sites, Comput. Struct. Biotechnol. J. 18 (2020) 852.
[252] S. Li, K. Yu, D. Wang, Q. Zhang, L. Zhao, Z.-X. Liu, H. Cheng, Deep learning based prediction of species-specific protein s-glutathionylation sites, Biochim. Biophys. Acta (BBA)-Proteins Proteom. 1868 (7) (2020) 140422.
[253] Y. Yang, X. Liu, C. Shen, Y. Lin, P. Yang, L. Qiao, In silico spectral libraries by deep learning facilitate data-independent acquisition proteomics, Nat. Commun. 11 (1) (2020) 1–11.
[254] S. Gessulat, T. Schmidt, D.P. Zolg, P. Samaras, K. Schnatbaum, J. Zerweck, T. Knaute, J. Rechenberger, B. Delanghe, A. Huhmer, et al., Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning, Nat. Methods 16 (6) (2019) 509–518.
[255] G. Zampieri, S. Vijayakumar, E. Yaneske, C. Angione, Machine and deep learning meet genome-scale metabolic modeling, PLoS Comput. Biol. 15 (7) (2019) e1007084.
[256] B. Shickel, P.J. Tighe, A. Bihorac, P. Rashidi, Deep ehr: a survey of recent advances in deep learning techniques for electronic health record (ehr) analysis, IEEE J. Biomed. Health Inform. 22 (5) (2017) 1589–1604.
[257] J.R.A. Solares, F.E.D. Raimondi, Y. Zhu, F. Rahimian, D. Canoy, J. Tran, A.C.P. Gomes, A.H. Payberah, M. Zottoli, M. Nazarzadeh, et al., Deep learning for electronic health records: A comparative review of multiple deep neural architectures, J. Biomed. Inform. 101 (2020) 103337.
[258] W.-C. Lin, J.S. Chen, M.F. Chiang, M.R. Hribar, Applications of artificial intelligence to electronic health record data in ophthalmology, Transl. Vis. Sci. Technol. 9 (2) (2020) 13.
[259] D.-H. Jang, J. Kim, Y.H. Jo, J.H. Lee, J.E. Hwang, S.M. Park, D.K. Lee, I. Park, D. Kim, H. Chang, Developing neural network models for early detection of cardiac arrest in emergency department, Amer. J. Emerg. Med. 38 (1) (2020) 43–49.
[260] Z. Xu, J. Chou, X.S. Zhang, Y. Luo, T. Isakova, P. Adekkanattu, J.S. Ancker, G. Jiang, R.C. Kiefer, J.A. Pacheco, et al., Identifying sub-phenotypes of acute kidney injury using structured and unstructured electronic health record data with memory networks, J. Biomed. Inform. 102 (2020) 103361.
[261] S. Cui, C. Li, Z. Chen, J. Wang, J. Yuan, Research on risk prediction of dyslipidemia in steel workers based on recurrent neural network and lstm neural network, IEEE Access 8 (2020) 34153–34161.
[262] R. Yu, Y. Zheng, R. Zhang, Y. Jiang, C.C. Poon, Using a multi-task recurrent neural network with attention mechanisms to predict hospital mortality of patients, IEEE J. Biomed. Health Inf. 24 (2) (2019) 486–492.
[263] M.L. Welch, C. McIntosh, A. McNiven, S.H. Huang, B.-B. Zhang, L. Wee, A. Traverso, B. O'Sullivan, F. Hoebers, A. Dekker, et al., User-controlled pipelines for feature integration and head and neck radiation therapy outcome predictions, Phys. Med. 70 (2020) 145–152.
[264] Z. Wang, Y. Zhu, D. Li, Y. Yin, J. Zhang, Feature rearrangement based deep learning system for predicting heart failure mortality, Comput. Methods Programs Biomed. 191 (2020) 105383.
[265] D. Alvarez-Estevez, I. Fernández-Varela, Dealing with the database variability problem in learning from medical data: An ensemble-based approach using convolutional neural networks and a case of study applied to automatic sleep scoring, Comput. Biol. Med. (2020) 103697.
[266] D. Lee, X. Jiang, H. Yu, Harmonized representation learning on dynamic EHR graphs, J. Biomed. Inform. (2020) 103426.
[267] K. Rough, A.M. Dai, K. Zhang, Y. Xue, L.M. Vardoulakis, C. Cui, A.J. Butte, M.D. Howell, A. Rajkomar, Predicting inpatient medication orders from electronic health record data, Clin. Pharmacol. Therapeutics 108 (1) (2020) 145–154.
[268] L. Wang, L. Tong, D. Davis, T. Arnold, T. Esposito, The application of unsupervised deep learning in predictive models using electronic health records, BMC Med. Res. Methodol. 20 (1) (2020) 1–9.
[269] T.E. Idriss, A. Idri, I. Abnane, Z. Bakkoury, Predicting blood glucose using an lstm neural network, in: 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), 2019, pp. 35–41.
[270] W. Song, W. Cai, J. Li, F. Jiang, S. He, Predicting blood glucose levels with emd and lstm based cgm data, in: 2019 6th International Conference on Systems and Informatics (ICSAI), 2019, pp. 1443–1448.
[271] K. Li, J. Daniels, C. Liu, P. Herrero, P. Georgiou, Convolutional recurrent neural networks for glucose prediction, IEEE J. Biomed. Health Inf. 24 (2) (2020) 603–613.
[272] P. Mamoshina, K. Kochetov, F. Cortese, A. Kovalchuk, A. Aliper, E. Putin, M. Scheibye-Knudsen, C.R. Cantor, N.M. Skjodt, O. Kovalchuk, A. Zhavoronkov, Blood biochemistry analysis to detect smoking status and quantify accelerated aging in smokers, Sci. Rep. 9 (1) (2019) 142.
[273] R. Indraswari, T. Kurita, A.Z. Arifin, N. Suciati, E.R. Astuti, Multi-projection deep learning network for segmentation of 3d medical images, Pattern Recognit. Lett. 125 (2019) 791–797.
[274] Z. Yue, S. Ding, W. Zhao, H. Wang, J. Ma, Y. Zhang, Y. Zhang, Automatic cin grades prediction of sequential cervigram image using lstm with multistate cnn features, IEEE J. Biomed. Health Inf. 24 (3) (2020) 844–854.
[275] X. Li, H. Wang, H. He, J. Du, J. Chen, J. Wu, Intelligent diagnosis with chinese electronic medical records based on convolutional neural networks, BMC Bioinformatics 20 (1) (2019).
[276] A. Atutxa, A.D. de Ilarraza, K. Gojenola, M. Oronoz, O.P. de Viñasprе, Interpretable deep learning to map diagnostic texts to icd-10 codes, Int. J. Med. Inform. 129 (2019) 49–59.
[277] M. Talo, U.B. Baloglu, Ö. Yıldırım, U.R. Acharya, Application of deep transfer learning for automated brain abnormality classification using mr images, Cogn. Syst. Res. 54 (2019) 176–188.
[278] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, C.-S. Huang, D. Shen, C.-M. Chen, Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in ct scans, Sci. Rep. 6 (1) (2016) 1–13.
[279] R. Yamashita, M. Nishio, R.K.G. Do, K. Togashi, Convolutional neural networks: an overview and application in radiology, Insights Imaging 9 (4) (2018) 611–629.
[280] P. Nigam, Applying Deep Learning to ICD-9 Multi-Label Classification from Medical Records, Technical Report, Technical report, Stanford University, 2016.
[281] S.R. Young, D.C. Rose, T.P. Karnowski, S.-H. Lim, R.M. Patton, Optimizing deep learning hyper-parameters through an evolutionary algorithm, in: Proceedings of the Workshop on Machine Learning in High-Performance Computing Environments, 2015, pp. 1–5.
[282] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, J. Mach. Learn. Res. 13 (1) (2012) 281–305.
[283] J.Y. Choi, T.K. Yoo, J.G. Seo, J. Kwak, T.T. Um, T.H. Rim, Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database, PLoS One 12 (11) (2017) e0187336.
[284] A.Y. Hannun, P. Rajpurkar, M. Haghpanahi, G.H. Tison, C. Bourn, M.P. Turakhia, A.Y. Ng, Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network, Nat. Med. 25 (1) (2019) 65.
[285] H. Bertrand, R. Ardon, M. Perrot, I. Bloch, Hyperparameter optimization of deep neural networks: combining hperband with Bayesian model selection, in: CAp, Grenoble, France, 2017, URL: https://hal.telecom-paris.fr/hal-02412262.
[286] A. Rojas-Domínguez, L.C. Padierna, J.M.C. Valadez, H.J. Puga-Soberanes, H.J. Fraire, Optimal hyper-parameter tuning of svm classifiers with application to medical diagnosis, IEEE Access 6 (2017) 7164–7176.
[287] S. Yousefi, F. Amrollahi, M. Amgad, C. Dong, J.E. Lewis, C. Song, D.A. Gutman, S.H. Halani, J.E.V. Vega, D.J. Brat, et al., Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models, Sci. Rep. 7 (1) (2017) 1–11.
[288] R.J. Borgli, H.K. Stensland, M.A. Riegler, P. Halvorsen, Automatic hyperparameter optimization for transfer learning on medical image datasets using Bayesian optimization, in: 2019 13th International Symposium on Medical Information and Communication Technology (ISMICT), IEEE, 2019, pp. 1–6.
[289] M. Celebi, N. Codella, A. Halpern, Dermoscopy image analysis: overview and future directions, IEEE J. Biomed. Health Inform. 23 (2) (2019) 474–478.
[290] P. Mamoshina, L. Ojomoko, Y. Yanovich, A. Ostrovski, A. Botezatu, P. Prikhodko, E. Izumchenko, A. Aliper, K. Romantsov, A. Zhebrak, et al., Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare, Oncotarget 9 (5) (2018) 5665.
[291] C. Shorten, T.M. Khoshgoftaar, A survey on image data augmentation for deep learning, J. Big Data 6 (1) (2019) 60.
[292] T. Ching, D.S. Himmelstein, B.K. Beaulieu-Jones, A.A. Kalinin, B.T. Do, G.P. Way, E. Ferrero, P.-M. Agapow, M. Zietz, M.M. Hoffman, et al., Opportunities and obstacles for deep learning in biology and medicine, J. R. Soc. Interface 15 (141) (2018) 20170387.
[293] F. Murat, O. Yildirim, M. Talo, U.B. Baloglu, Y. Demir, U.R. Acharya, Application of deep learning techniques for heartbeats detection using ecg signals-analysis and review, Comput. Biol. Med. (2020) 103726.
[294] D. Gunning, Explainable artificial intelligence (xai), Defense Advanced Research Projects Agency (DARPA), nd Web, Vol. 2, 2017, p. 2.
[295] A.B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai, Inf. Fusion 58 (2020) 82–115.
[296] S. Dash, B.R. Acharya, M. Mittal, A. Abraham, A. Kelemen, Deep Learning Techniques for Biomedical and Health Informatics, Springer, 2020.
[297] E. Gawehn, J.A. Hiss, G. Schneider, Deep learning in drug discovery, Mol. Inform. 35 (1) (2016) 3–14.
[298] A. Zhavoronkov, Artificial Intelligence for Drug Discovery, Biomarker Development, and Generation of Novel Chemistry, ACS Publications, 2018.
[299] B. Fan, W. Fan, C. Smith, et al., Adverse drug event detection and extraction from open data: A deep learning approach, Inf. Process. Manage. 57 (1) (2020) 102131.
[300] H. Chen, O. Engkvist, Y. Wang, M. Olivecrona, T. Blaschke, The rise of deep learning in drug discovery, Drug Discov. Today 23 (6) (2018) 1241–1250.

[301] L. Zhou, Z. Li, J. Yang, G. Tian, F. Liu, H. Wen, L. Peng, M. Chen, J. Xiang, L. Peng, Revealing drug-target interactions with computational models and algorithms, Molecules 24 (9) (2019) 1714.
[302] N. Harder, R. Schönmeyer, K. Nekolla, A. Meier, N. Brieu, C. Vanegas, G. Madonna, M. Capone, G. Botti, P.A. Ascierto, et al., Automatic discovery of image-based signatures for ipilimumab response prediction in malignant melanoma, Sci. Rep. 9 (1) (2019) 1–19.
[303] A. Craik, Y. He, J.L. Contreras-Vidal, Deep learning for electroencephalogram (eeg) classification tasks: a review, J. Neural Eng. 16 (3) (2019) 031001.
[304] J. Goldhahn, V. Rampton, G.A. Spinas, Could artificial intelligence make doctors obsolete? Bmj 363 (2018) k4563.
[305] A.S. Ahuja, The impact of artificial intelligence in medicine on the future role of the physician, PeerJ 7 (2019) e7702.