Arch. Pharm. Res. (2019) 42:492–504
https://doi.org/10.1007/s12272-019-01162-9

Online ISSN 1976-3786
Print ISSN 0253-6269

Archives of Pharmacal Research
www.springer.com/12272

REVIEW

Applications of deep learning for the analysis of medical data

Hyun-Jong Jang[1] · **Kyung-Ok Cho**[2]

Received: 28 December 2018 / Accepted: 20 May 2019 / Published online: 28 May 2019

Abstract Over the past decade, deep learning has demonstrated superior performances in solving many problems in various fields of medicine compared with other machine learning methods. To understand how deep learning has surpassed traditional machine learning techniques, in this review, we briefly explore the basic learning algorithms underlying deep learning. In addition, the procedures for building deep learning-based classifiers for seizure electroencephalograms and gastric tissue slides are described as examples to demonstrate the simplicity and effectiveness of deep learning applications. Finally, we review the clinical applications of deep learning in radiology, pathology, and drug discovery, where deep learning has been actively adopted. Considering the great advantages of deep learning techniques, deep learning will be increasingly and widely utilized in a wide variety of different areas in medicine in the coming decades.

Keywords Artificial intelligence · Deep neural networks · Drug discovery · Medical image analysis

✉ Kyung-Ok Cho
kocho@catholic.ac.kr

[1] Department of Physiology, Department of Biomedicine & Health Sciences, Catholic Neuroscience Institute, College of Medicine, The Catholic University of Korea, Seoul 06591, South Korea

[2] Department of Pharmacology, Department of Biomedicine & Health Sciences, Catholic Neuroscience Institute, Institute of Aging and Metabolic Diseases, College of Medicine, The Catholic University of Korea, 222 Banpo-Daero, Seocho-Gu, Seoul 06591, South Korea

Introduction

Machine learning is a field of study which has amassed a collection of various methods that can automatically detect common or even subtle patterns hidden in datasets using a computer (Rajkomar et al. 2019). Deep learning is a subfield of machine learning, which is commonly characterized by algorithms consisting of multi-layered artificial neural networks, especially those containing more than four layers. Because deep learning can extract universal features in very complex datasets, it has exhibited improved performance in many tasks compared to other machine learning techniques (Lecun et al. 2015). These merits have allowed deep learning to become a new approach that can efficiently solve various problems in medicine (Mamoshina et al. 2016; Miotto et al. 2018; Yu et al. 2018). Recent advances in deep learning have demonstrated that diagnoses of diverse diseases based on the classification of radiologic images and histologic slides has almost surpassed human capabilities (Teare et al. 2017; Veta et al. 2015). Deep learning can also provide an outstanding accuracy on the detection of retinopathy from retinal fundus photographs (Gulshan et al. 2016). With the great promises of deep learning technology, the application of deep learning has been expanded to pharmaceutical research, including drug discovery (Chen et al. 2018). The active adoption of various traditional computational methods has enabled the pioneering application of deep learning in the field of drug discovery, where deep learning has already exceeded the performance of traditional computer-aided analyzers.

As big datasets are essential prerequisites for the deep learning technology, biomedical data quickly became an attractive candidate for the application of deep learning. However, given that proper data annotation and the

application of appropriate deep learning algorithms are necessary depending on the characteristics of medical datasets, it is critical that experts in both deep learning and medicine share fundamental knowledge of each field for the successful utilization of deep learning in medicine. Although there are several reviews on deep learning in biomedicine (Cao et al. 2018; Mamoshina et al. 2016; Ching et al. 2018), they are more appropriate for the experts in machine learning. Therefore, to provide general ideas of deep learning to the biomedical researchers and clinicians who have biologic backgrounds without professional knowledge of mathematics and computer sciences, in this review, we will first go over the basic concepts of deep learning algorithms, followed by two examples of how to build a deep learning-based classifier. In the second part, we will review recent applications of deep learning-based techniques in several fields of medicine. Finally, we will briefly discuss current limitations of deep learning and offer perspectives on the future of deep learning.

Basic principles of deep learning and neural networks

The basic concept of machine learning

Deep learning is a type of machine learning, so we first introduce the basic concepts of machine learning. Machine learning is one approach to data analysis that detects patterns in the data and then uses these patterns to predict future outcomes. When considering biomedical datasets, two main types of machine learning are widely utilized: the first is supervised learning and the second is unsupervised learning (Murphy 2012; Deo 2015). Supervised learning works by learning a mapping from inputs to outputs based on given labels for the set of input–output pairs. When the outputs are presented as categorical variables, this mapping task is known as a classification problem. For example, if a machine learning agent learns to distinguish normal and tumor tissues on pathologic slides based on the labels of either "normal" or "tumor," this is regarded as a classification problem in supervised learning. On the other hand, if the outputs are real-valued scalars, then this is known as a regression problem. Take for example a dataset regarding lipid metabolism-related gene expression profiles and blood cholesterol levels for thousands of patients. Here, a machine learning agent can learn to predict blood cholesterol levels for new patients based on these gene expression profiles. If the agent is trained using a dataset containing exact values of the cholesterol, this is an example of supervised learning solving a regression problem. Note that this example can instead be considered as a classification problem if the output format is represented as low, medium, or high levels of cholesterol instead of a continuous range.

The second approach to machine learning is unsupervised learning, which consists of learning without specific labels. One of the most common examples of unsupervised learning is known as clustering (Saxena et al. 2017). In more detail, consider a situation where there are thousands of nuclear images of various cells collected from histopathologic slides. These images can be automatically clustered into a fixed number of groups based on some measure of similarity identified by the machine learning agent. Because there is no specific label for this task, it is considered as unsupervised learning.

Taken together, biomedical datasets can be analyzed using two representative types of machine learning algorithms: supervised and unsupervised learning. Which approach is more appropriate is determined by the types of questions being asked and the various properties of the data being considered.

Optimization of model parameters

Among the many available machine learning algorithms, a general method of formalizing learning patterns in data is through function approximation. When the data can be understood using an unknown function f, the goal is to approximate f using the given data. This approximated function then acts as a model for the problem, and is thus known as a model-based approach to machine learning. After a model is constructed using a training dataset, it is validated using test datasets to verify the effectiveness of the model. For example, when we have some coordinates that appeared to be arranged in a line, we can try to explain it with a linear function, i.e., $y = ax$ (Fig. 1a). In this case, we should determine the proper value for a parameter 'a' to optimally explain the observed data. If the dataset (x, y) are consist of (0.5, 0.25), (1, 0.5), (3.5, 1.75), and (5.5, 2.75), the best one-parameter equation that can explain this dataset will be $y = 0.5x$ (Fig. 1a). Though analytical solutions exist for low-order polynomial fits, for illustrative purposes, computers can execute this fitting task by performing regression analysis in an iterative manner. If an initial value of 'a' is randomly assigned to 2, then the computer measures how wrong the current value of the parameter 'a' is in order to find a way to optimize it into 0.5 (Fig. 1b). Briefly, at each data-point, the error between the observed value (y) and the predicted one ($\hat{y}$) is measured and the overall error of a current model can be presented as the mean squared error (MSE). The MSE of the model can be computed as,

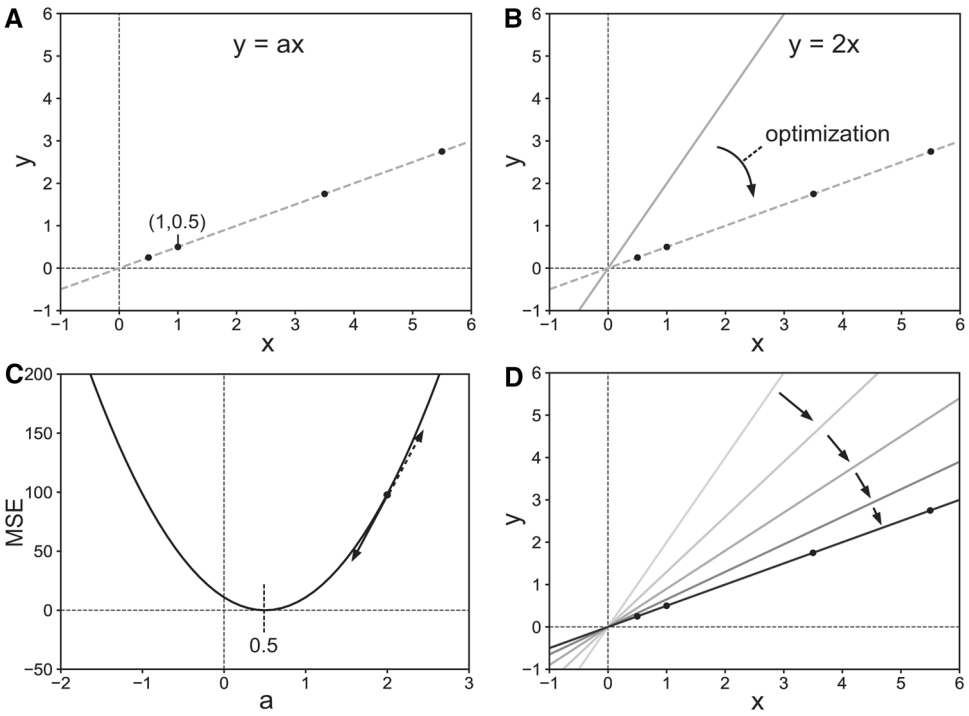

Fig. 1 Gradient descent optimization algorithm for parameter optimization. **a** As an example, the data (closed circle) was obtained from an unknown linear function y = ax. To explain the data, the parameter 'a' should be determined. **b** As one of the methods to determine the correct value of 'a', it can be iteratively optimized from a randomly assigned value. **c** The mean squared error (MSE) can measure the error of the current model. If the MSE is plotted against 'a', the gradient, ΔMSE/Δa (dotted arrow), is always opposite to the direction of the minimum MSE. **d** If the value of 'a' is adjusted along the opposite direction of the gradient, the optimal value of 'a' can be obtained

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n}(\mathrm{y_i} - \hat{\mathrm{y}}_\mathrm{i})^2$$

and thus,

$$\mathrm{MSE} = \frac{1}{\mathrm{n}}\sum_{i=1}^{n}(\mathrm{y_i} - \mathrm{ax_i})^2$$

When we plot calculated MSE values against all possible 'a' values, we can easily figure out that the gradient, ΔMSE/Δa, is always opposite to the direction of the minimum MSE (Fig. 1c). In this example, when 'a' is 2, the calculated gradient is 32.81, meaning that 'a' should be smaller than 2 because the gradient has a positive value. Thus, the correct value of 'a', resulting in the minimum MSE, can be obtained if the value of 'a' is adjusted following the direction opposite the gradient (Fig. 1d). When an error function, such as the MSE, is at its minimum value, the model is now considered to explain the data optimally. This procedure is called gradient descent optimization of parameters (Bishop 2006) and is the basis of the parameter optimization in deep learning. Even if the model is very complex and has many parameters, gradient descent can be applied to obtain optimal values for all parameters because the partial derivatives of each parameter in the error function can guide the direction of the minimum error for the parameter. The error-measuring function, like the MSE, is often called the error function, loss function, or objective function and should be designed to be differentiable for applying gradient descent method.

In summary, the parameters can be adjusted to build an optimal model for any observed phenomena by gradient descent, because the gradient of error function can guide the direction of the parameter modifications to minimize the error and thus maximize the validity of a model. One note of caution is that nonlinear optimization problems with a large parameter space typically have many local minima, and only one global minimum, and various techniques, from simulated annealing to genetic algorithms, have been applied to more rigorously minimize the error while using gradient descent methods.

The basic concept of deep neural networks

Deep learning is a type of machine learning that utilizes deep artificial neural networks. Ever since artificial neural networks were introduced more than 70 years ago (Mcculloch and Pitts 1943), the core concepts behind typical neural networks have not changed much, despite undergoing a few modifications (Fig. 2a). Basic artificial neural network structures use nodes and edges that correspond to neural cell bodies and neurites, respectively. Nodes in the input layer are connected to the other nodes in the next layer by edges (Fig. 2a). The edge has a weight, a parameter that reflects the strength of the connection between the two nodes. A typical neural network is composed of one input layer, one output layer, and variable numbers of hidden layers in between. If a neural network has multiple hidden layers, it is then called a deep neural network (Fig. 2b). Basically, the values in each node of the

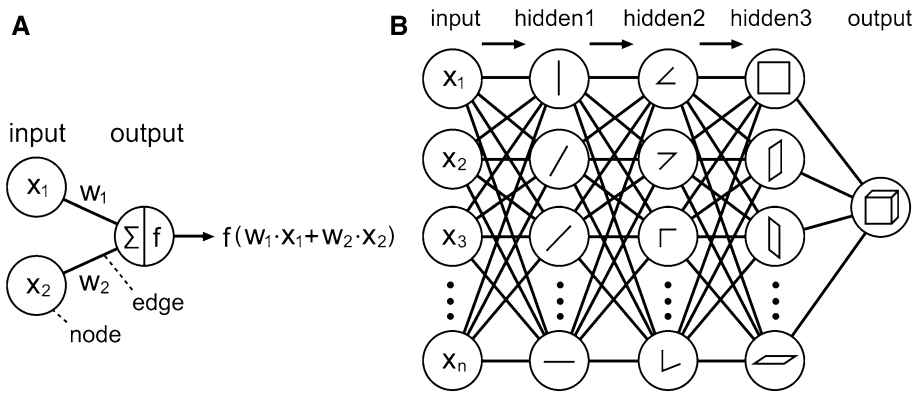

Fig. 2 Deep artificial neural network. **a** The basic component of an artificial neural network. Each input value is assigned to each node in the input layer. Nodes in the input layer are connected to output node by edges which have their own weights to determine how the inputs affect the output. The weighted sum of inputs is calculated for the output node, which then passes the sum through an activation function to produce an outcome. An activation function performs a nonlinear transformation and thus the neural network can learn to process nonlinear phenomena. **b** The basic structure of a deep neural network. Layers between the input and output layers are called hidden layers because their values are not directly visible. This figure schematically shows that nodes in layers learn a hierarchy of compositional features with the progression of layers. Diagrams in nodes represent extracted features, which become more complex by manipulating the weighted composition of features in previous layers. This figure is drawn for explaining purpose and is not a perfect reflection on how features are learned in real-world examples

input layer are multiplied by weights, which are summed to the nodes in the next layer. Thus, each node in the first hidden layer incorporates all the information in the input nodes with different weights, generating a variety of possible simplified representations to distinguish the differences in the dataset (Fig. 2b). Then, the information in the nodes of the first hidden layer is integrated into the nodes of the next hidden layer, i.e., the values in all the nodes of the first hidden layer are again multiplied by different weights to produce different values in each node of the second hidden layer. As this process is repeated in the multiple layers, more advanced criteria that can separate the differences in the dataset can be established, since the number of differential combinations of the nodes is increased. Training of deep neural networks aims to determine the weight values producing the minimum error function, which represents the optimal model to explain a dataset. During training, weights are optimized to transform the initial raw input variables into more useful features. Then, subsets of the initial features are selected to construct more abstract features in deeper layers, which are differential combinations of the original features (Fig. 2b). As these selected features contain the relevant information from the input data, the desired task can be performed by using this dimensionally reduced representation. This is how feature extraction and selection is carried out by deep neural networks.

One of the advantages of deep learning technology is that we can explain very complex phenomena by using a neural network model, provided that all of the weights of a deep neural network are successfully optimized. The weight optimization can be accomplished by the gradient descent described in the previous section, although the direct application of this method has a limitation in that the errors in the output layer are not directly linked to the parameters of earlier layers. Development of the back-propagation of error method (Rumelhart et al. 1986) nicely solved this issue, enabling complicated designs of deep neural networks for explaining complex data. Furthermore, continual developments of more robust optimization algorithms have allowed for faster learning of optimized parameters (Duchi et al. 2011; Kingma and Ba 2017), thus widening the applicability of deep learning.

The strength of deep learning comes from the multi-layered 'deep' architecture. As the data goes through multiple layers, more complex and abstract features in the input data can be extracted because each node in a layer can process the information collected from all the nodes in the previous layers (Fig. 2b). This type of learning is called representation learning, which can extract multiple high-level features in complex data. For this reason, deep learning could diminish the necessity of extensive feature engineering based on domain expertise, which has been a core task of traditional machine learning methods (Lecun et al. 2015). Despite the many highly significant advantages of deep learning technology, there has been another bottleneck impeding the widespread adoption of deep learning, the requirement of big datasets. Data size is a critical factor as adding more data points can result in a better model for representing the characteristics underlying a specific dataset (Fig. 3). Moreover, since the deep networks should learn all the features from a given dataset without any prior knowledge, it is essential that plenty of data be supplied to deep neural networks for successful learning. Recent accumulation of biomedical big data will

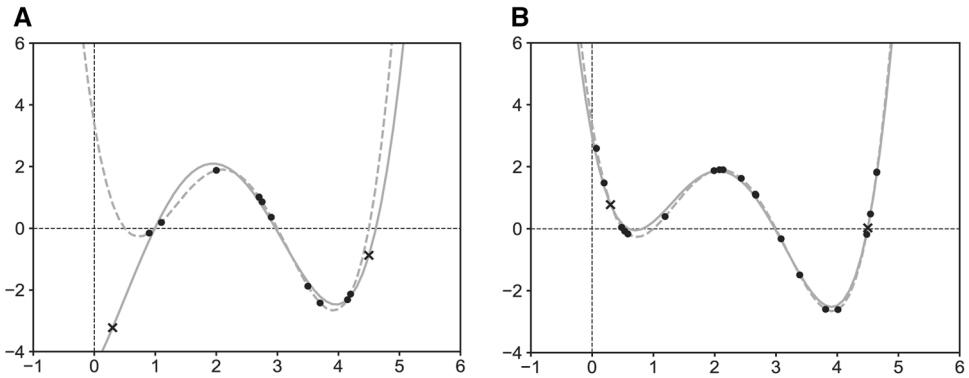

Fig. 3 Example demonstrating the relationship between the size of the data and the quality of the model. Data (filled circles) was obtained from a function represented by the dotted line. **a** A model (solid line) was made from 10 data points (filled circles) and did not correctly represent the original function. **b** A model was made from 20 data points that did faithfully represent the original function. Although this example is an oversimplified version, the necessity of a large quantity of data for deep learning is obvious, considering the complexity of real-world problems and the millions to billions of parameters that must be optimized in a deep neural network

ease and streamline the application of deep learning, providing great promise for the field of medicine.

There are many different types of deep neural networks (Goodfellow et al. 2016). The network architecture in Fig. 2b is called a fully connected feed-forward deep neural network and is a basic form of the deep neural network. Other network architectures can be formulated to enhance the learning by using specialized data structures. For example, a convolutional neural network (CNN) can be applied to data with a grid-structured topology, such as images (Shin et al. 2016b), whereas recurrent neural networks (RNNs) can be applied for sequential datasets such as language (Graves et al. 2013). As each type of neural network is specialized for specific learning, combining multiple existing frameworks or developing a new architecture will greatly improve the feasibility of neural networks for interpreting complicated phenomena.

Two examples of how to construct deep learning-based classifiers

In this section, we will describe how to build deep learning-based classifiers using the example of an EEG seizure auto-detecting algorithm (Fig. 4) (Jang and Cho 2019), and tumor classification algorithm for gastric cancer (Fig. 5). As a basic and standard neural network structure, we chose the fully connected deep neural network with two hidden layers to build a seizure autodetector. Then, we determined how to best supply EEG data as inputs to neural networks, as data preprocessing often greatly improves the efficiency of neural networks. Continuous EEG signals were split into 5-s segments and every segment was annotated as either a seizure or non-seizure segment, resulting in the production of 28,616 seizure and 76,464 non-seizure segments. Then, these EEG segment data was divided into two datasets for training and validation. To test the accuracy of a classifier, 10% of the entire seizure and non-seizure segments were randomly assigned for validation, i.e., data not used for training, in every trial. Since the EEG sensor was set to collect the data at 1000 Hz, the 5-s segments had 5000 data-points. Thus, we initially used all the 5000 raw data-points as inputs without any preprocessing and let the network learn to distinguish seizure or non-seizure data. Briefly, we labeled seizure segments with (1, 0) and non-seizure segments with (0, 1), and organized the output layer of the deep neural network to yield a pair of values between 0 and 1 (Fig. 4a). The parameters of the network were initially set to have random values. As a result, the accuracy of this network showing the correct answer was initially around 50% because the network randomly classified data into two categories. The error function adopted was categorical cross-entropy, which yields a minimum value when the label is identical with the output of the network. When the network classifies all of the input data correctly, this error function would necessarily become minimal. To accomplish that, gradient descent could automatically adjust the randomly initialized parameters to minimize the cross-entropy error function, making the deep neural network optimized for classifying seizure and non-seizure segments. This is the core process underlying deep learning. During the optimization processes, the nodes are automatically adjusted to extract proper features in the input data for distinguishing differences between groups (Fig. 2b). After learning, the accuracy for the validation dataset was around 95.5%, which was not acceptable for a state-of-the-art classifier. We speculated that each of the raw EEG inputs had a low signal to noise ratio in spite of 5000 data-points, failing to provide meaningful inputs to discriminate two different groups. To improve the learning accuracy, we decided to modify raw EEG traces by doing

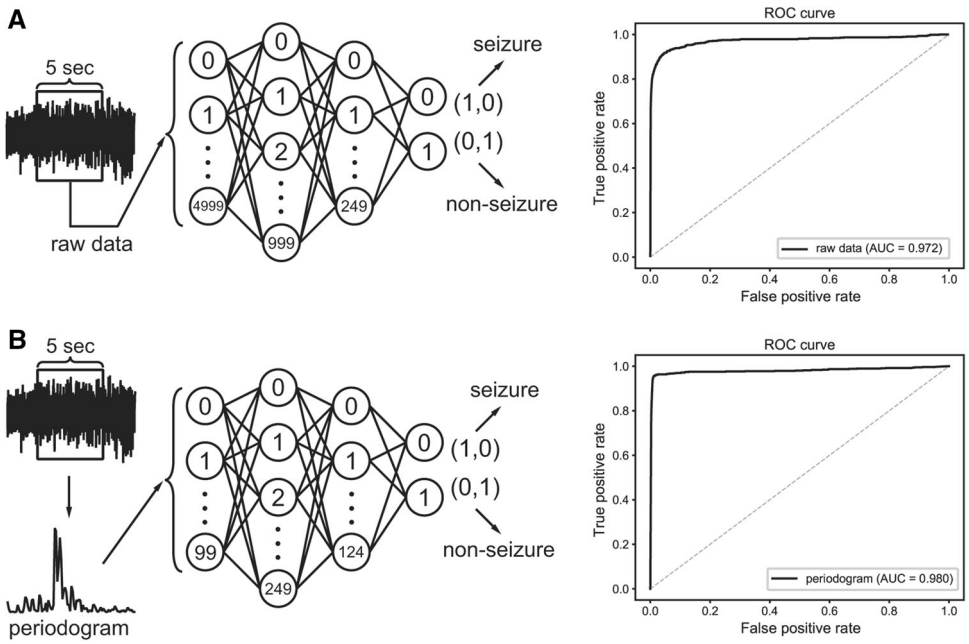

Fig. 4 Example of a deep neural network for classification of seizure electroencephalograms (modified from Jang and Cho 2019 with consent). **a** Left: The deep neural network used to classify raw input data. It receives 5000 raw data points for 5-s segment and processes the input with 2 hidden layers, containing 1000 and 250 nodes. Two output nodes yield the probability for seizure and non-seizure, respectively. The network was trained to yield the output (1, 0) for seizures and (0, 1) for a non-seizure EEG series. Right: Receiver operating characteristics (ROC) curve for the classifier of raw EEG inputs. Area under the curve (AUC) was 0.972. **b** Left: The deep neural network to classify the periodogram result from a 5-s segment. The input layer receives 100 values corresponding 0 to 99 Hz periodogram results. The hidden layers contain 250 and 125 nodes. Right: ROC curve for the classifier of periodogram result. AUC was 0.980

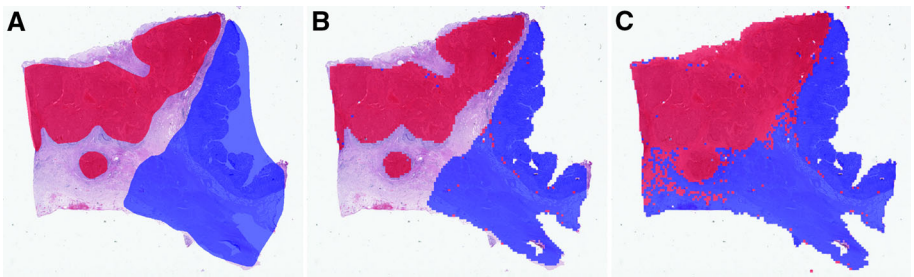

Fig. 5 Classification of a stomach tissue slide from the test set. **a** Annotation of normal (blue) and tumor (red) in the tissue section by pathologists. **b** A classification result between normal and tumor areas by trained CNN. **c** Full tissue classification results

spectral analysis, generating periodogram results that show a simple frequency-power distribution in each 5-s EEG segment (Fig. 4b). Since epileptic brain rhythms primarily have frequency bands less than 100 Hz, we obtained periodogram results between 0 and 99 Hz, leading to the reduction of the input number from 5000 to 100. When we fed the periodogram results into a new neural network with fewer nodes, the accuracy for the network was improved to around 99.5%. In summary, sensitivity and specificity for test set were 93.7% and 98.2% for raw input and 96.5% and 98.9% for periodogram, respectively. In this example, although we could achieve a better learning accuracy by choosing a different format of inputs to a deep neural network, there are many other adjustable factors in order to attain the best learning accuracy. For instance, when a CNN was used instead of fully connected neural networks, the accuracy increased to 99.9%, indicating that the endeavor to find the optimal factors is critical for building a successful deep neural network-based classifier.

Next, for constructing a tumor classifier for stomach tissue, we used TCGA-STAD dataset (https://portal.gdc.cancer.gov/projects/TCGA-STAD). TCGA-STAD dataset contains 442 digitally scanned whole slide images (WSI) of stomach tissues from patients with gastric cancer. We chose 379 scanned slides with a good quality where normal and the tumor area was annotated by pathologists (Fig. 5a). Because the size of WSI was very large, for example, a slide in Fig. 5 had 93,623 × 84,294 pixels in a 20× magnified view, it was unable to build a classifier based on a whole image. Thus, after the slide was divided into segments (i.e., 360 × 360 pixels), a classifier was built to distinguish between normal or tumor segments. We

randomly split 379 WSI into a training set with 342 images and a test set with 37 images. A CNN-based classifier was trained to distinguish normal and tumor segments collected from the 342 training WSI. We labeled tumor segments with (1, 0) and normal segments with (0, 1), and trained the CNN network with categorical cross-entropy error function. Thus, the training process was fundamentally almost the same as the training of seizure autodetector. After training, the classification accuracy of all the annotated segments collected from the test set was 95.4%. Since the classifier was built to distinguish small segments of WSI, we overlaid the classification results for segments in the annotated area for comparison (Fig. 5b). The overall results of the classifier were comparable to the annotation done by pathologists (Fig. 5a). When we overlaid the classification results on the whole slide, we could clearly see how tumor tissue was demarcated in the stomach tissue slides (Fig. 5c). In this example, even though we did not apply any feature engineering on slide images, the CNN network effectively learned to distinguish normal and tumor tissues.

Applications of deep learning in medicine

Application of deep learning for the analysis of radiologic images

As the picture archiving and communication system (PACS) has been widely used for more than 20 years (Huang 2011), deep learning can utilize an ample amount of digitized data. As a result, hundreds of different studies analyzing radiology images based on the deep learning technique have been recently reported (Litjens et al. 2017). Most of the studies utilized a CNN, an architecture of deep neural networks specialized for image data, as the CNN mimics the function of the visual cortex, which can extract local motifs such as lines and edges in order to construct visual features, and thus the CNN demonstrates improved performance on image-based input data. Application of deep learning on radiologic images can be largely divided into three tasks: classification, detection, and segmentation (Chartrand et al. 2017). Classification tasks can be defined as sorting out the radiologic images into discrete classes such as the presence or absence of the disease, or the judgment of malignant or benign tumors. For example, pulmonary tuberculosis can be successfully distinguished from normal chest X-ray findings (Lakhani and Sundaram 2017). Moreover, deep learning can classify whether breast lesions and pulmonary nodules are benign or malignant (Cheng et al. 2016), in addition to the excellent discrimination of prostate malignancies in the prostate magnetic resonance (MR) images (Wang et al. 2017b). Deep learning-assisted detection is the task to find focal lesions in the radiologic images with the information of where the lesions are located. For instance, deep learning algorithms can identify various pulmonary lesions including emphysema, fibrosis, or micronodules in lung images (Setio et al. 2016; Gao et al. 2018; Nam et al. 2018). In addition, enlarged lymph nodes or colonic polyps in computer tomography (CT) images (Roth et al. 2016), and cerebral microbleeding in MR images (Qi et al. 2016) can be detected by utilizing deep neural networks. Finally, segmentation identifies meaningful pixels or voxels composing an organ or a structure of interest. The final results of segmentation can be usually presented as the outline following the contour of the target structure. To apply segmentation to radiology, many studies tried to differentiate brain subregions (Lian et al. 2018; Moeskops et al. 2016; Akkus et al. 2017), or prostate substructures (Guo et al. 2016) in MR images. Efforts to segregate different body parts in transverse CT slices have also been directed at developing automatic segmentation of target organs (Zhennan et al. 2016). Recently, automatic identification of the lesions in multiple sclerosis has been reported by analyzing brain MR images (Brosch et al. 2016), demonstrating another great advancement in deep learning techniques to support the diagnosis of various diseases from radiologic images. However, considering the complex nature of various diseases and the different resolutions of diverse imaging tools, a full diagnosis system may be needed to utilize the multiple approaches combining classification, detection, and segmentation processes in deep learning to draw a precise conclusion (Kuan et al. 2017). We briefly summarized recent papers using deep learning to analyze radiologic images in Table 1. There are also more extensive reviews covering deep learning on radiologic images (Litjens et al. 2017; Bernal et al. 2019; Shen et al. 2017; Lee et al. 2017).

Application of deep learning for the analysis of pathology slides

Computer-aided diagnostic procedures in pathology have attracted considerable attention for decades (Mendelsohn et al. 1965; Bloom and Weinstein 1985). In a simple preparation, such as pap smear screening, traditional machine learning could show an acceptable diagnostic accuracy (Wilbur et al. 1998), suggesting the high potential of artificial intelligence in the field of pathology. Recently, as whole slide imaging has become a routine process in developed countries (Farahani et al. 2015), the accumulation of digital images has enabled deep learning-based approaches for detecting pathologic findings. For example, deep learning showed superior performance for mitosis detection in breast cancer histopathology images (Veta

Table 1 Brief overview of papers using deep learning to analyze radiologic images

Organ	Modality	Task	References
Abdomen	CT	Detection (lymph node metastasis)	Roth et al. (2016)
Abdomen	CT	Detection (polyp)	Nappi et al. (2016)
Abdomen	CT	Segmentation (bladder cancer)	Cha et al. (2016)
Abdomen	CT	Segmentation (liver)	Lu et al. (2017)
Abdomen	MRI	Classification (prostate cancer)	Wang et al. (2017b)
Abdomen	MRI	Segmentation (kidney)	Haghighi et al. (2018)
Abdomen	MRI	Segmentation (pancreas)	Cai et al. (2016)
Abdomen	MRI	Segmentation (prostate)	Milletari et al. (2016)
Brain	fMRI	Classification (Alzheimer's disease)	Hosseini-Asl et al. (2018)
Brain	MRI	Detection (microbleed detection)	Qi et al. (2016)
Brain	MRI	Segmentation (anatomical structures)	Moeskops et al. (2016)
Brain	MRI	Segmentation (brain tumor)	Akkus et al. (2017)
Brain	MRI	Segmentation (brain tumors and ischemic stroke)	Kamnitsas et al. (2017)
Brain	MRI	Segmentation (multiple sclerosis lesion)	Brosch et al. (2016)
Brain	MRI	Segmentation (perivascular space)	Lian et al. (2018)
Brain	MRI	Segmentation (striatum segmentation)	Choi and Jin (2016)
Breast	MG	Classification (breast mass)	Sun et al. (2017)
Breast	MG	Detection (breast mass)	Kooi et al. (2017)
Heart	CT	Detection (coronary calcification)	Wolterink et al. (2016)
Heart	MRI	Segmentation (cardiac chambers)	Bai et al. (2018)
Lung	CT	Classification (interstitial lung disease)	Christodoulidis et al. (2017)
Lung	CT	Classification (lung nodules)	Ciompi et al. (2017)
Lung	CT	Detection (emphysema, fibrosis, and etc.)	Gao et al. (2018)
Lung	CT	Detection (lung cancer)	Kuan et al. (2017)
Lung	CT	Detection (lung nodules)	Setio et al. (2016)
Lung	CT	Segmentation (airway)	Charbonnier et al. (2017)
Lung	X-ray	Classification (lung nodules)	Wang et al. (2017a)
Lung	X-ray	Classification (tuberculosis)	Lakhani and Sundaram (2017)
Lung	X-ray	Detection (lung nodules)	Nam et al. (2018)
Lung	X-ray	Detection (various pathologic condition including emphysema, granuloma, and etc.)	Shin et al. (2016a)

CT computed tomography, *MG* mammography, *MRI* magnetic resonance image, *fMRI* functional MRI

et al. 2015). The detection of cancer metastasis can be also facilitated by deep learning (Wang et al. 2016). Moreover, deep learning can assess and score the expression of diagnostic markers such as HER2 in tissue slides (Vandenberghe et al. 2017), in addition to the prediction of genetic mutations (Coudray et al. 2018), providing additional information to support medical decisions. In addition to the diagnostic applications, the functional status of major organs including the kidney, liver, and lung can be estimated from slide images. A recent study nicely demonstrated that renal function could be predicted from kidney biopsy images using a CNN (Ledbetter et al. 2017). These various examples can illustrate how widely deep learning can be applied. We briefly summarized exemplary articles applying deep learning for analyzing pathologic slides in Table 2. For more specialized information, we recommend to read more extensive reviews of deep learning in pathology (Litjens et al. 2017; Komura and Ishikawa 2018; Janowczyk and Madabhushi 2016; Chang et al. 2019). Since deep learning can already perform versatile tasks including the cancer diagnosis and subtype distinction to the level of human physicians (Coudray et al. 2018; Esteva et al. 2017) and even well-trained pathologists often fail to achieve diagnostic concordance (Elmore et al. 2015; Brimo et al. 2010), deep learning can become a major supportive tool to the pathologists for increasing the accuracy and efficiency of histopathological diagnosis in a near future (Litjens et al. 2016).

Table 2 Brief overview of papers using deep learning to analyze pathology slides

Tissue or cell type	Modality	Task	References
Breast tissue	H&E	Detection (mitosis)	Veta et al. (2015)
Breast tissue	H&E	Detection (invasive ductal carcinoma)	Cruz-Roa et al. (2014)
Breast tissue	IHC	Detection and segmentation (lobular structures)	Apou et al. (2016)
Colorectal tissue	H&E	Classification (colorectal polyps)	Korbar et al. (2017)
Colorectal tissue	H&E	Classification and detection (nucleus)	Sirinukunwattana et al. (2016)
Colorectal tissue	H&E	Prediction (outcome in colorectal cancer)	Bychkov et al. (2018)
Colorectal tissue	H&E	Segmentation (colon glands)	Kainz et al. (2017)
Kidney tissue	H&E	Prediction (kidney function)	Ledbetter et al. (2017)
Lung tissue	H&E	Classification (cancer type) Prediction (genetic mutation)	Coudray et al. (2018)
Neuronal tissue	EM	Segmentation (cell)	Ronneberger et al. (2015)
Prostate tissue	H&E	Detection (prostate cancer)	Litjens et al. (2016)
Renal and prostate tissue	IHC	Classification (nucleus)	Bauer et al. (2016)
Sentinel lymph node	H&E	Detection (metastatic breast cancer)	Wang et al. (2016)
Thyroid tissue	H&E	Classification (thyroid pathology)	Kim et al. (2016a)

EM electron microscopy, *H&E* hematoxylin and eosin stain, *IHC* immunohistochemistry

Application of deep learning for the drug discovery

Structure-based virtual screening is a computational approach to discover new drug candidates with the basic knowledge of the 3D structure of the target molecule and has been widely studied as a cost-effective alternative to labor-intensive experimental methods (Danishuddin and Khan 2015). Many computer algorithms have been developed to estimate target structures and docking mechanisms against a target, which can be further improved by deep learning such as AlphaFold, a DeepMind's software to predict protein 3D structures (Anderson 2003; Senior et al. 2018). In contrast, ligand-based drug design relies on the knowledge of ligand molecules that bind to the biological target of interest (Acharya et al. 2011). The quantitative structure–activity relationship (QSAR) or quantitative structure–property relationship (QSPR) is one of the most widely used tools to identify lead compounds from many candidates when the receptor protein structure is not known (Svetnik et al. 2003; Tropsha 2010). QSAR/QSPR in ligand-based virtual screening is based on the assumption that for a group of similar compounds, the structural molecular characteristics (descriptors) such as the geometric, steric, and electronic properties would be quantitatively correlated with biological activities (Acharya et al. 2011). However, traditional machine learning approaches could test only a few molecular descriptors at once because these methods could not process a very large quantity of multi-dimensional data. Consequently, the selection of suitable molecular descriptors became crucial for the reliability of the model. Thus, it could be a time-consuming and rigorous task to find out critical combinations of molecular descriptors for a faithful model (Zhang et al. 2017). On top of that, recent advancements of high-throughput screening techniques in drug discovery accelerated the generation of big data, requiring high-dimensional data processing. As a result, deep learning has drawn attention, since it can automatically extract critical features from high-dimensional heterogeneous data, without manually selecting molecular descriptors (Chen et al. 2018). Moreover, deep learning-based virtual screening can process a massive volume of structural data in an efficient manner for discriminating the desired characteristics in molecules (Lusci et al. 2013). Thus, the capability of deep learning that recognizes representative features directly from structures without predefined sets of descriptors proved it had distinct advantages over other machine learning methods. Not only can deep learning assess the effectiveness of candidate molecules, but deep learning can also estimate drug-target interactions (Wen et al. 2017), which is useful for pharmacodynamics modeling. In addition, possible adverse drug events can be predicted by building a deep neural network-based identifier for drug-drug interactions (Ryu et al. 2018), suggesting that the application of deep learning can be extended to the field of pharmacokinetic profiling. Taken together, drug discovery will be accelerated by utilizing deep learning techniques, especially for in silico drug design and computer-aided ADMET (absorption, distribution, metabolism, excretion, and toxicity) profiling. The advantage of deep learning in big data processing, coupled with rapid accumulation of chemical data (Papadatos et al. 2015; Gilson et al. 2016;

Kim et al. 2016b) and other medical data, will synergistically lead to the development of powerful virtual screening systems that can open up new potentials in the field of drug discovery.

Concluding remarks

We briefly reviewed the learning mechanisms of deep neural networks and the application of deep learning techniques in several research fields where it shows great potential. Although we selected radiology, pathology and drug discovery as representative examples, deep learning has been applied for wider variety of problems in biomedical and health informatics domains including bioinformatics, medical imaging, pervasive sensing, medical informatics, and public health issues (Yao et al. 2018; Ravi et al. 2017). In bioinformatics, deep learning can be applied for the analysis of DNA sequence, gene expression profile, and protein structures to help diagnosis of diseases, prognostic prediction, and drug design. Huge amounts of data from various sensors including EEG, electrocardiogram, and many implantable and wearable devices can also be effectively analyzed by deep learning technique. Application of deep learning on other medical big data such as electronic health records, lab test results, and even lifelong data can reveal unexpected important factors to influence disease status hidden in the big data. Because deep learning spans such a wide range of applications, we anticipate that practice of medicine would be greatly affected by deep learning in a near future. For example, future pathologists may not be required to review the entire pathologic slides daily generated from patients. Since deep learning-based classifiers can yield a probability map of the tissue indicating cancer (Fig. 6), only ambiguous slides (Fig. 6b) will be necessary to be evaluated by pathologists.

Although deep learning has come into the limelight, there are a few obstacles that deep learning should overcome to be more widely utilized in medical practice. First, for maximizing the potential of deep learning, the creation of the large databases of well-annotated medical data is crucial. Although the digitization of medical data has been strongly encouraged in hospitals, current forms of digitized data are not standardized to train the deep neural network. Because deep learning requires very specific and precise labeling of the data for successful training, medical experts have to reexamine and annotate huge amounts of medical data to make them suitable for deep learning. As it is a time consuming and costly task, this has been one of the major hurdles for the application of deep learning in wider fields of medicine. Cooperative efforts to build large datasets of labeled medical data, such as The Cancer Imaging Archive (http://www.cancerimagingarchive.net) and Genomic Data Commons Data Portal (https://portal.gdc.cancer.gov), might help to overcome this difficulty. Second, it would be ideal that a computer-aided diagnostic system clearly justify its outcomes as physicians take the final responsibility for their medical judgment when they make a decision with the aid of deep learning techniques. However, deep learning is currently a black-box-like system where the decision process cannot be clearly explained. Because of the complexity of incorporating millions to billions of parameters, it is difficult to understand how deep learning analyzes data to draw a decision in detail. This opaque nature is another hurdle that deep learning should ideally overcome to hasten the practical utilization of neural networks in clinical settings.

The era of deep learning-assisted medical practice is now dawning. Currently, deep learning has been applied to solve a few specialized tasks in medicine. However, if the aforementioned obstacles can be overcome in the near future and more physicians can appreciate the benefits of

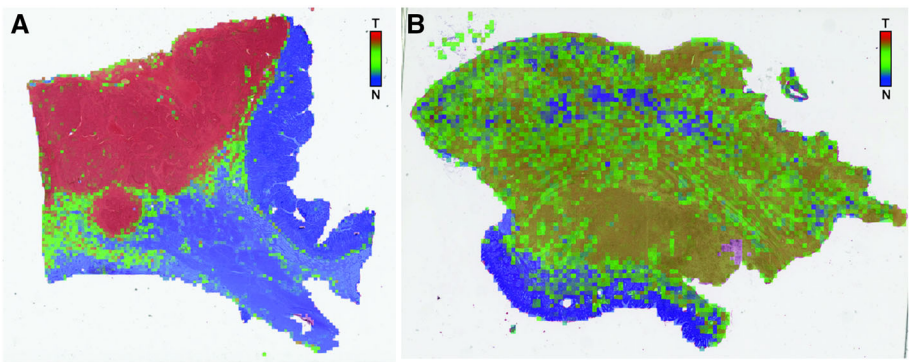

Fig. 6 Tumor probability map of two stomach tissue slides. Normal segments were overlaid with blue color. Tumor segments were overlaid with red color. The segments with 50% of tumor probability were overlaid with green color. **a** Slide with mass of high tumor probability region. **b** Slide with plenty of ambiguous segments. Insets demonstrate color distribution between normal (N) and tumor (T)

deep learning, deep learning techniques will greatly change how medicine is practiced, considering the precise diagnostic tools and new drugs developed by deep learning. Physicians of the next generation will likely make a decision based on deep learning-guided supporting systems in a variety of tasks in medicine and give prescriptions of new drugs discovered by deep learning-assisted virtual screening.

Acknowledgements This work was supported by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health & Welfare (HI15C2854).

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

Acharya C, Coop A, Polli JE, Mackerell AD Jr (2011) Recent advances in ligand-based drug design: relevance and utility of the conformationally sampled pharmacophore approach. Curr Comput Aided Drug Des 7:10–22

Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ (2017) Deep learning for brain MRI segmentation: state of the art and future directions. J Digit Imaging 30:449–459

Anderson AC (2003) The process of structure-based drug design. Chem Biol 10:787–797

Apou G, Schaadt NS, Naegel B, Forestier G, Schonmeyer R, Feuerhake F, Wemmert C, Grote A (2016) Detection of lobular structures in normal breast tissue. Comput Biol Med 74:91–102

Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N, Lukaschuk E, Sanghvi MM, Zemrak F, Fung K, Paiva JM, Carapella V, Kim YJ, Suzuki H, Kainz B, Matthews PM, Petersen SE, Piechnik SK, Neubauer S, Glocker B, Rueckert D (2018) Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. J Cardiovasc Magn Reson 20:65

Bauer S, Carion N, Schaffler P, Fuchs T, Wild P, Buhmann JM (2016) Multi-organ cancer classification and survival analysis. https://arxiv.org/abs/1606.00897. Accessed 2 Jun 2016

Bernal J, Kushibar K, Asfaw DS, Valverde S, Oliver A, Marti R, Llado X (2019) Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. Artif Intell Med 95:64–81

Bishop C (2006) Pattern recognition and machine learning. Springer, Berlin

Bloom KJ, Weinstein RS (1985) Expert systems: robot physicians of the future? Hum Pathol 16:1082–1084

Brimo F, Schultz L, Epstein JI (2010) The value of mandatory second opinion pathology review of prostate needle biopsy interpretation before radical prostatectomy. J Urol 184:126–130

Brosch T, Tang LY, Youngjin Y, Li DK, Traboulsee A, Tam R (2016) Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation. IEEE Trans Med Imaging 35:1229–1239

Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, Walliander M, Lundin M, Haglund C, Lundin J (2018) Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep 8:3395

Cai J, Lu L, Zhang Z, Xing F, Yang L, Yin Q (2016) Pancreas segmentation in MRI using graph-based decision fusion on convolutional neural networks. Med Image Comput Comput Assist Interv 9901:442–450

Cao C, Liu F, Tan H, Song D, Shu W, Li W, Zhou Y, Bo X, Xie Z (2018) Deep learning and its applications in biomedicine. Genomics Proteom Bioinf 16:17–32

Cha KH, Hadjiiski LM, Samala RK, Chan HP, Cohan RH, Caoili EM, Paramagul C, Alva A, Weizer AZ (2016) Bladder cancer segmentation in CT for treatment response assessment: application of deep-learning convolution neural network—a pilot study. Tomography 2:421–429

Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, Kwak TY (2019) Artificial intelligence in pathology. J Pathol Transl Med 53:1–12

Charbonnier JP, Rikxoort EMV, AaA Setio, Schaefer-Prokop CM, Ginneken BV, Ciompi F (2017) Improving airway segmentation in computed tomography using leak detection with convolutional networks. Med Image Anal 36:52–60

Chartrand G, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, Pal CJ, Kadoury S, Tang A (2017) Deep Learning: a primer for radiologists. Radiographics 37:2113–2131

Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T (2018) The rise of deep learning in drug discovery. Drug Discov Today 23:1241–1250

Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC, Huang CS, Shen D, Chen CM (2016) Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci Rep 6:24454

Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, Hoffman MM, Xie W, Rosen GL, Lengerich BJ, Israeli J, Lanchantin J, Woloszynek S, Carpenter AE, Shrikumar A, Xu J, Cofer EM, Lavender CA, Turaga SC, Alexandari AM, Lu Z, Harris DJ, Decaprio D, Qi Y, Kundaje A, Peng Y, Wiley LK, Segler MHS, Boca SM, Swamidass SJ, Huang A, Gitter A, Greene CS (2018) Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 15:20170387

Choi H, Jin KH (2016) Fast and robust segmentation of the striatum using deep convolutional neural networks. J Neurosci Methods 274:146–153

Christodoulidis S, Anthimopoulos M, Ebner L, Christe A, Mougiakakou S (2017) Multi-source transfer learning with convolutional neural networks for lung pattern analysis. IEEE J Biomed Health Inform 21:76–84

Ciompi F, Chung K, Van Riel SJ, AaA Setio, Gerke PK, Jacobs C, Scholten ET, Schaefer-Prokop C, Wille MMW, Marchiano A, Pastorino U, Prokop M, Van Ginneken B (2017) Towards automatic pulmonary nodule management in lung cancer screening with deep learning. Sci Rep 7:46479

Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, Moreira AL, Razavian N, Tsirigos A (2018) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nat Med 24:1559–1567

Cruz-Roa A, Basavanhally A, Gonz´Alez F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A (2014) Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. SPIE Mecial Imaging, San Diego, California, USA, pp. 904103

Danishuddin M, Khan AU (2015) Structure based virtual screening to discover putative drug candidates: necessary considerations and successful case studies. Methods 71:135–145

Deo RC (2015) Machine Learning in Medicine. Circulation 132:1920–1930

Duchi J, Hazan E, Singer Y (2011) Adaptive subgradient methods for online learning and stochastic optimization. J Mach Learn Res 12:2121–2159

Elmore JG, Longton GM, Carney PA, Geller BM, Onega T, Tosteson AN, Nelson HD, Pepe MS, Allison KH, Schnitt SJ, O'malley FP, Weaver DL (2015) Diagnostic concordance among pathologists interpreting breast biopsy specimens. JAMA 313:1122–1132

Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542:115–118

Farahani N, Parwani AV, Pantanowitz L (2015) Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. Pathol Lab Med Int 7:23–33

Gao M, Bagci U, Lu L, Wu A, Buty M, Shin HC, Roth H, Papadakis GZ, Depeursinge A, Summers RM, Xu Z, Mollura DJ (2018) Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. Comput Methods Biomech Biomed Eng Imaging Vis 6:1–6

Gilson MK, Liu T, Baitaluk M, Nicola G, Hwang L, Chong J (2016) BindingDB in 2015: a public database for medicinal chemistry, computational chemistry and systems pharmacology. Nucleic Acids Res 44:D1045–1053

Goodfellow I, Bengio Y, Courville (2016) A deep learning. The MIT Press, Cambridge

Graves A, Mohamed AR, Hinton G (2013) Speech recognition with deep recurrent neural networks. https://arxiv.org/abs/1303.5778. Accessed 22 Mar 2013

Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316:2402–2410

Guo Y, Gao Y, Shen D (2016) Deformable MR prostate segmentation via deep feature learning and sparse patch matching. IEEE Trans Med Imaging 35:1077–1089

Haghighi M, Warfield SK, Kurugol S (2018) Automatic renal segmentation in Dce-Mri using convolutional neural networks. Proc IEEE Int Symp Biomed Imaging 2018:1534–1537

Hosseini-Asl E, Ghazal M, Mahmoud A, Aslantas A, Shalaby AM, Casanova MF, Barnes GN, Gimel'farb G, Keynton R, El-Baz A (2018) Alzheimer's disease diagnostics by a 3D deeply supervised adaptable convolutional network. Front Biosci (Landmark Ed) 23:584–596

Huang HK (2011) Short history of PACS. Part I: USA. Eur J Radiol 78:163–176

Jang HJ, Cho KO (2019) Dual deep neural network-based classifiers to detect experimental seizures. Korean J Physiol Pharmacol 23:131–139

Janowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform 7:29

Kainz P, Pfeiffer M, Urschler M (2017) Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. PeerJ 5:e3874

Kamnitsas K, Ledig C, Newcombe VFJ, Simpson JP, Kane AD, Menon DK, Rueckert D, Glocker B (2017) Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal 36:61–78

Kim E, Corte-Real M, Baloch Z (2016a) A deep semantic mobile application for thyroid cytopathology. In: Medical imaging. SPIE Mecial Imaging, San Diego, California, USA, pp. 97890A

Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH (2016b) PubChem substance and compound databases. Nucleic Acids Res 44:D1202–1213

Kingma DP, Ba J (2017) Adam: a method for stochastic optimization. https://arxiv.org/abs/1412.6980. Accessed 30 Jan 2017

Komura D, Ishikawa S (2018) Machine learning methods for histopathological image analysis. Comput Struct Biotechnol J 16:34–42

Kooi T, Litjens G, Van Ginneken B, Gubern-Merida A, Sanchez CI, Mann R, Den Heeten A, Karssemeijer N (2017) Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal 35:303–312

Korbar B, Olofson AM, Miraflor AP, Nicka CM, Suriawinata MA, Torresani L, Suriawinata AA, Hassanpour S (2017) Deep learning for classification of colorectal polyps on whole-slide images. J Pathol Inform 8:30

Kuan K, Ravaut M, Manek G, Chen H, Lin J, Nazir B, Chen C, Howe TC, Zeng Z, Chandrasekhar V (2017) Deep learning for lung cancer detection: tackling the Kaggle Data science bowl 2017 challenge. https://arxiv.org/abs/1705.09435. Accessed 26 May 2017

Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284:574–582

Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444

Ledbetter D, Ho L, Lemley KV (2017) Prediction of kidney function from biopsy images using convolutional neural networks. https://arxiv.org/abs/1702.01816. Accessed 6 Feb 2017

Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, Kim N (2017) Deep learning in medical imaging: general overview. Korean J Radiol 18:570–584

Lian C, Zhang J, Liu M, Zong X, Hung SC, Lin W, Shen D (2018) Multi-channel multi-scale fully convolutional network for 3D perivascular spaces segmentation in 7T MR images. Med Image Anal 46:106–117

Litjens G, Sanchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-Van De Kaa C, Bult P, Van Ginneken B, Van Der Laak J (2016) Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep 6:26286

Litjens G, Kooi T, Bejnordi BE, AaA Setio, Ciompi F, Ghafoorian M, Van Der Laak J, Van Ginneken B, Sanchez CI (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88

Lu F, Wu F, Hu P, Peng Z, Kong D (2017) Automatic 3D liver location and segmentation via convolutional neural network and graph cut. Int J Comput Assist Radiol Surg 12:171–182

Lusci A, Pollastri G, Baldi P (2013) Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. J Chem Inf Model 53:1563–1575

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

Mcculloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biol 5:115–133

Mendelsohn ML, Kolman WA, Perry B, Prewitt JM (1965) Morphological analysis of cells and chromosomes by digital computer. Methods Inf Med 4:163–167

Milletari F, Navab N, Ahmadi S-A (2016) V-Net: fully convolutional neural networks for volumetric medical image segmentation. https://arxiv.org/abs/1606.04797. Accessed 15 Jun 2016

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

Moeskops P, Viergever MA, Mendrik AM, De Vries LS, Benders MJ, Isgum I (2016) Automatic segmentation of MR brain images with a convolutional neural network. IEEE Trans Med Imaging 35:1252–1261

Murphy KP (2012) Machine learning: a probabilistic perspective. The MIT Press, Cambridge

Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, Vu TH, Sohn JH, Hwang S, Goo JM, Park CM (2018) Development and validation of deep learning-based automatic detection algorithm for malignant pulmonary nodules on chest radiographs. Radiology 290:180–237

Nappi JJ, Hironaka T, Regge D, Yoshida H (2016) Deep transfer learning of virtual endoluminal views for the detection of polyps in CT colonography. SPIE Medical Imaging, San Diego, California, USA, pp. 97852B

Papadatos G, Gaulton A, Hersey A, Overington JP (2015) Activity, assay and target data curation and quality in the ChEMBL database. J Comput Aided Mol Des 29:885–896

Qi D, Hao C, Lequan Y, Lei Z, Jing Q, Defeng W, Mok VC, Lin S, Pheng-Ann H (2016) Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 35:1182–1195

Rajkomar A, Dean J, Kohane I (2019) Machine learning in medicine. N Engl J Med 380:1347–1358

Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, Yang GZ (2017) Deep learning for health informatics. IEEE J Biomed Health Inform 21:4–21

Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. https://arxiv.org/abs/1505.04597. Accessed 18 May 2015

Roth HR, Lu L, Liu J, Yao J, Seff A, Cherry K, Kim L, Summers RM (2016) Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans Med Imaging 35:1170–1181

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323:533–536

Ryu JY, Kim HU, Lee SY (2018) Deep learning improves prediction of drug-drug and drug-food interactions. Proc Natl Acad Sci USA 115:E4304–E4311

Saxena A, Prasad M, Gupta A, Bharill N, Patel OP, Tiwari A, Er MJ, Ding WP, Lin CT (2017) A review of clustering techniques and developments. Neurocomputing 267:664–681

Senior A, Jumper J, Hassabis D (2018) AlphaFold: using AI for scientific discovery. https://deepmind.com/blog/alphafold/. 12.2

Setio AA, Ciompi F, Litjens G, Gerke P, Jacobs C, Van Riel SJ, Wille MM, Naqibullah M, Sanchez CI, Van Ginneken B (2016) Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 35:1160–1169

Shen D, Wu G, Suk HI (2017) Deep learning in medical image analysis. Annu Rev Biomed Eng 19:221–248

Shin H-C, Roberts K, Lu L, Demner-Fushman D, Yao J, Summers RM (2016a) Learning to read chest X-rays: recurrent neural cascade model for automated image annotation. https://arxiv.org/abs/1603.08486. Accessed 28 May 2016

Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016b) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35:1285–1298

Sirinukunwattana K, Ahmed Raza SE, Yee-Wah T, Snead DR, Cree IA, Rajpoot NM (2016) Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 35:1196–1206

Sun W, Tseng TB, Zhang J, Qian W (2017) Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Comput Med Imaging Graph 57:4–9

Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP (2003) Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci 43:1947–1958

Teare P, Fishman M, Benzaquen O, Toledano E, Elnekave E (2017) Malignancy detection on mammography using dual deep convolutional neural networks and genetically discovered false color input enhancement. J Digit Imaging 30:499–505

Tropsha A (2010) Best practices for QSAR model development, validation, and exploitation. Mol Inform 29:476–488

Vandenberghe ME, Scott ML, Scorer PW, Soderberg M, Balcerzak D, Barker C (2017) Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. Sci Rep 7:45938

Veta M, Van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A, Gonzalez F, Larsen AB, Vestergaard JS, Dahl AB, Ciresan DC, Schmidhuber J, Giusti A, Gambardella LM, Tek FB, Walter T, Wang CW, Kondo S, Matuszewski BJ, Precioso F, Snell V, Kittler J, De Campos TE, Khan AM, Rajpoot NM, Arkoumani E, Lacle MM, Viergever MA, Pluim JP (2015) Assessment of algorithms for mitosis detection in breast cancer histopathology images. Med Image Anal 20:237–248

Wang D, Khosla A, Gargeya R, Irshad H, Beck AH (2016) Deep learning for identifying metastatic breast cancer. https://arxiv.org/abs/1606.05718. Accessed 18 Jun 2016

Wang C, Elazab A, Wu J, Hu Q (2017a) Lung nodule classification using deep feature fusion in chest radiography. Comput Med Imaging Graph 57:10–18

Wang X, Yang W, Weinreb J, Han J, Li Q, Kong X, Yan Y, Ke Z, Luo B, Liu T, Wang L (2017b) Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. Sci Rep 7:15415

Wen M, Zhang Z, Niu S, Sha H, Yang R, Yun Y, Lu H (2017) Deep-learning-based drug-target interaction prediction. J Proteome Res 16:1401–1409

Wilbur DC, Prey MU, Miller WM, Pawlick GF, Colgan TJ (1998) The AutoPap system for primary screening in cervical cytology. Comparing the results of a prospective, intended-use study with routine manual practice. Acta Cytol 42:214–220

Wolterink JM, Leiner T, De Vos BD, Van Hamersvelt RW, Viergever MA, Isgum I (2016) Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal 34:123–136

Yao Z-J, Bi J, Chen Y-X (2018) Applying deep learning to individual and community health monitoring data: a survey. Int J Aut Comput 15:643–655

Yu KH, Beam AL, Kohane IS (2018) Artificial intelligence in healthcare. Nat Biomed Eng 2:719–731

Zhang L, Tan J, Han D, Zhu H (2017) From machine learning to deep learning: progress in machine intelligence for rational drug discovery. Drug Discov Today 22:1680–1685

Zhennan Y, Yiqiang Z, Zhigang P, Shu L, Shinagawa Y, Shaoting Z, Metaxas DN, Xiang Sean Z (2016) Multi-instance deep learning: discover discriminative local anatomies for bodypart recognition. IEEE Trans Med Imaging 35:1332–1343