REVIEW

An overview of deep learning in medical imaging focusing on MRI

Alexander Selvikvåg Lundervold [a,b,*], **Arvid Lundervold** [a,c,d]

[a] Mohn Medical Imaging and Visualization Centre (MMIV), Haukeland University Hospital, Norway
[b] Department of Computing, Mathematics and Physics, Western Norway University of Applied Sciences, Norway
[c] Neuroinformatics and Image Analysis Laboratory, Department of Biomedicine, University of Bergen, Norway
[d] Department of Health and Functioning, Western Norway University of Applied Sciences, Norway

Abstract

What has happened in machine learning lately, and what does it mean for the future of medical image analysis? Machine learning has witnessed a tremendous amount of attention over the last few years. The current boom started around 2009 when so-called deep artificial neural networks began outperforming other established models on a number of important benchmarks. Deep neural networks are now the state-of-the-art machine learning models across a variety of areas, from image analysis to natural language processing, and widely deployed in academia and industry. These developments have a huge potential for medical imaging technology, medical data analysis, medical diagnostics and healthcare in general, slowly being realized. We provide a short overview of recent advances and some associated challenges in machine learning applied to medical image processing and image analysis. As this has become a very broad and fast expanding field we will not survey the entire landscape of applications, but put particular focus on deep learning in MRI.

Our aim is threefold: (i) give a brief introduction to deep learning with pointers to core references; (ii) indicate how deep learning has been applied to the entire MRI processing chain, from acquisition to image retrieval, from segmentation to disease prediction; (iii) provide a starting point for people interested in experimenting and perhaps contributing to the field of deep learning for medical imaging by pointing out good educational resources, state-of-the-art open-source code, and interesting sources of data and problems related medical imaging.

Keywords: Machine learning, Deep learning, Medical imaging, MRI

1 Introduction

Machine learning has seen some dramatic developments recently, leading to a lot of interest from industry, academia and popular culture. These are driven by breakthroughs in *artificial neural networks*, often termed *deep learning*, a set of techniques and algorithms that enable computers to discover complicated patterns in large data sets. Feeding the breakthroughs is the increased access to data ("big data"), user-friendly software frameworks, and an explosion of the available compute power, enabling the use of neural networks that are deeper than ever before. These models nowadays form the state-of-the-art approach to a wide variety of problems in computer vision, language modeling and robotics.

Deep learning rose to its prominent position in computer vision when neural networks started outperforming other methods on several high-profile image analysis benchmarks.

* Corresponding author: Alexander Selvikvåg Lundervold, Mohn Medical Imaging and Visualization Centre (MMIV), Haukeland University Hospital, Norway.

E-mail addresses: allu@hvl.no (A. Selvikvåg Lundervold),
mfyal@uib.no (A. Lundervold).

Most famously on the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) in 2012 [1] when a deep learning model (a *convolutional neural network*) halved the second best error rate on the image classification task. Enabling computers to recognize objects in natural images was until recently thought to be a very difficult task, but by now convolutional neural networks have surpassed even human performance on the ILSVRC, and reached a level where the ILSVRC classification task is essentially solved (i.e. with error rate close to the Bayes rate). Deep learning techniques have become the de facto standard for a wide variety of computer vision problems. They are, however, not limited to image processing and analysis but are outperforming other approaches in areas like natural language processing [2–4], speech recognition and synthesis [5,6],[1] and in the analysis of unstructured, tabular-type data using *entity embeddings* [7,8].[2]

The sudden progress and wide scope of deep learning, and the resulting surge of attention and multi-billion dollar investment, has led to a virtuous cycle of improvements and investments in the entire field of machine learning. It is now one of the hottest areas of study world-wide [14], and people with competence in machine learning are highly sought-after by both industry and academia.[3]

Healthcare providers generate and capture enormous amounts of data containing extremely valuable signals and information, at a pace far surpassing what "traditional" methods of analysis can process. Machine learning therefore quickly enters the picture, as it is one of the best ways to integrate, analyze and make predictions based on large, heterogeneous data sets (cf. health informatics [15]). Healthcare applications of deep learning range from one-dimensional biosignal analysis [16] and the prediction of medical events, e.g. seizures [17] and cardiac arrests [18], to computer-aided detection [19] and diagnosis [20] supporting clinical decision making and survival analysis [21], to drug discovery [22] and as an aid in therapy selection and pharmacogenomics [23], to increased operational efficiency [24], stratified care delivery [25], and analysis of electronic health records [26,27].

The use of machine learning in general and deep learning in particular within healthcare is still in its infancy, but there are several strong initiatives across academia, and multiple large companies are pursuing healthcare projects based on machine learning. Not only medical technology companies, but also for example Google Brain [28–30],[4] DeepMind [31],[5] Microsoft [32,33][6] and IBM [34].[7] There is also a plethora of small and medium-sized businesses in the field.[8]

2 Machine learning, artificial neural networks, deep learning

In machine learning one develops and studies methods that give computers the ability to solve problems by learning from experiences. The goal is to create mathematical models that can be *trained* to produce useful outputs when fed input data. Machine learning models are provided experiences in the form of *training data*, and are tuned to produce accurate predictions for the training data by an optimization algorithm. The main goal of the models are to be able to *generalize* their learned expertise, and deliver correct predictions for new, unseen data. A model's generalization ability is typically estimated during training using a separate data set, the validation set, and used as feedback for further tuning of the model. After several iterations of training and tuning, the final model is evaluated on a test set, used to simulate how the model will perform when faced with new, unseen data.

There are several kinds of machine learning, loosely categorized according to how the models utilize its input data during training. In *reinforcement learning* one constructs *agents* that learn from their environments through trial and error while optimizing some objective function. A famous recent application of reinforcement learning is AlphaGo and AlphaZero [35], the Go-playing machine learning systems developed by DeepMind. In *unsupervised learning* the computer is tasked with uncovering patterns in the data without our guidance. Clustering is a prime example. Most of today's machine learning systems belong to the class of *supervised learning*. Here, the computer is given a set of already labeled or annotated data, and asked to produce correct labels on new, previously unseen data sets based on the rules discovered in the labeled data set. From a set of input-output examples, the whole model is trained to perform specific data-processing tasks. Image annotation using human-labeled data, e.g. classifying skin lesions according to malignancy [36] or discovering

---

[1] Try it out here: https://deepmind.com/blog/wavenet-generative-model-raw-audio.

[2] As a perhaps unsurprising side-note, these modern deep learning methods have also entered the field of physics. Among other things, they are tasked with learning physics from raw data when no good mathematical models are available. For example in the analysis of gravitational waves where deep learning has been used for classification [9], anomaly detection [10] and denoising [11], using methods that are highly transferable across domains (think EEG and fMRI). They are also part of mathematical model and machine learning hybrids [12,13], formed to reduce computational costs by having the mathematical model train a machine learning model to perform its job, or to improve the fit with observations in settings where the mathematical model can't incorporate all details (think noise).

[3] See e.g. https://economicgraph.linkedin.com/research/LinkedIns-2017-US-Emerging-Jobs-Report for a study focused on the US job market.

[4] https://ai.google/research/teams/brain/healthcare-biosciences.

[5] https://deepmind.com/applied/deepmind-health/.

[6] https://www.microsoft.com/en-us/research/research-area/medical-health-genomics.

[7] https://www.research.ibm.com/healthcare-and-life-sciences.

[8] Aidoc, Arterys, Ayasdi, Babylon Healthcare Services, BenevolentAI, Enlitic, EnvoiAI, H2O, IDx, MaxQ AI, Mirada Medical, Viz.ai, Zebra Medical Vision, and many more.

cardiovascular risk factors from retinal fundus photographs [37], are two examples of the multitude of medical imaging related problems attacked using supervised learning.

Machine learning has a long history and is split into many sub-fields, of which deep learning is the one currently receiving the bulk of attention.

There are many excellent, openly available overviews and surveys of deep learning. For short general introductions to deep learning, see [38,39]. For an in-depth coverage, consult the freely available book [40].[9] For a broad overview of deep learning applied to medical imaging, see [41]. We will only mention some bare essentials of the field, hoping that these will serve as useful pointers to the areas that are currently the most influential in medical imaging.

2.1 Artificial neural networks

Artificial neural networks (ANNs) is one of the most famous machine learning models, introduced already in the 1950s, and actively studied since [40,Chapter 1.2].[10]

Roughly, a neural network consists of a number of connected computational units, called *neurons*, arranged in layers. There's an input layer where data enters the network, followed by one or more *hidden layers* transforming the data as it flows through, before ending at an output layer that produces the neural network's predictions. The network is trained to output useful predictions by identifying patterns in a set of labeled training data, fed through the network while the outputs are compared with the actual labels by an *objective function*. During training the network's parameters – the strength of each neuron – is tuned until the patterns identified by the network result in good predictions for the training data. Once the patterns are learned, the network can be used to make predictions on new, unseen data, i.e. generalize to new data.

It has long been known that ANNs are very flexible, able to model and solve complicated problems, but also that they are difficult and very computationally expensive to train.[11] This has lowered their practical utility and led people to, until recently, focus on other machine learning models. But by now,

artificial neural networks form one of the dominant methods in machine learning, and the most intensively studied. This change is thanks to the growth of *big data*, powerful processors for parallel computations (in particular, GPUs), some important tweaks to the algorithms used to construct and train the networks, and the development of easy-to-use software frameworks. The surge of interest in ANNs leads to an incredible pace of developments, which also drives other parts of machine learning with it.

The freely available books [40,49] are two of the many excellent sources to learn more about artificial neural networks. We'll only give a brief indication of how they are constructed and trained. The basic form of artificial neural networks,[12] the *feedforward neural networks*, are parametrized mathematical functions $y = f(\mathbf{x}; \theta)$ that maps an input $\mathbf{x}$ to an output $\mathbf{y}$ by feeding it through a number of nonlinear transformations: $f(\mathbf{x}) = (f_n \circ \cdots \circ f_1)(\mathbf{x})$. Here each component f_k, called a network *layer*, consists of a simple linear transformation of the previous component's output, followed by a nonlinear function: $f_k = \sigma_k(\theta_k^T f_{k-1})$. The nonlinear functions σ_k are typically sigmoid functions or ReLUs, as discussed below, and the θ_k are matrices of numbers, called the model's *weights*. During the training phase, the network is fed training data and tasked with making predictions at the output layer that match the known labels, each component of the network producing an expedient representation of its input. It has to learn how to best utilize the intermediate representations to form a complex hierarchical representation of the data, ending in correct predictions at the output layer. Training a neural network means changing its weights to optimize the outputs of the network. This is done using an optimization algorithm, called *gradient descent*, on a function measuring the correctness of the outputs, called a *cost function* or *loss function*. The basic ideas behind training neural networks are simple: as training data is fed through the network, compute the gradient of the loss function with respect to every weight using the chain rule, and reduce the loss by changing these weights using gradient descent. But one quickly meets huge computational challenges when faced with complicated networks with thousands or millions of parameters and an exponential number of paths between the nodes and the network output. The techniques designed to overcome these challenges gets quite complicated. See [40,Chapter 8,50,Chapters 3 and 4] for detailed descriptions of the techniques and practical issues involved in training neural networks.

Artificial neural networks are often depicted as a network of nodes, as in Fig. 1.[13]

---

[9] https://www.deeplearningbook.org/.

[10] The loose connection between artificial neural networks and neural networks in the brain is often mentioned, but quite over-blown considering the complexity of biological neural networks. However, there is some interesting recent work connecting neuroscience and artificial neural networks, indicating an increase in the cross-fertilization between the two fields [42–44].

[11] According to the famous *universal approximation theorem* for artificial neural networks [45–48], ANNs are mathematically able to approximate any continuous function defined on compact subspaces of $\mathbb{R}^n$, using finitely many neurons. There are some restrictions on the activation functions, but these can be relaxed (allowing for ReLUs for example) by restricting the function space. This is an existence theorem and successfully *training* a neural network to approximate a given function is another matter entirely. However, the theorem does suggest that neural networks are reasonable to study and develop further, at least as an engineering endeavour aimed at realizing their theoretical powers.

[12] These are basic when compared to for example *recurrent neural networks*, whose architectures are more involved.

[13] As we shall see, modern architectures are often significantly more complicated than captured by the illustration and equations above, with connections between non-consecutive layers, input fed in also at later layers, multiple outputs, and much more.
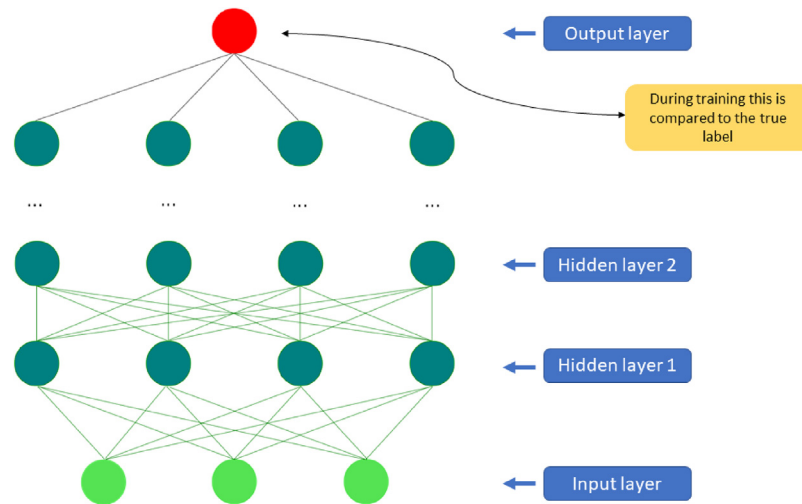
Figure 1. Artificial neural networks are built from simple linear functions followed by nonlinearities. One of the simplest class of neural network is the *multilayer perceptron*, or *feedforward neural network*, originating from the work of Rosenblatt in the 1950s [51]. It is based on simple computational units, called *neurons*, organized in *layers*. Writing i for the ith layer and j for the jth unit of that layer, the output of the jth unit at the ith layer is $z_j^{(i)} = \theta_j^{(i)T} x$. Here x consists of the outputs from the previous layer after they are fed through a simple nonlinear function called an *activation function*, typically a sigmoid function $\sigma(z) = 1/(1 + e^{-z})$ or a rectified linear unit $\text{ReLU}(z) = \max(0, z)$ or small variations thereof. Each layer therefore computes a weighted sum of the all the outputs from the neurons in the previous layers, followed by a nonlinearity. These are called the *layer activations*. Each layer activation is fed to the next layer in the network, which performs the same calculation, until you reach the *output layer*, where the network's predictions are produced. In the end, you obtain a hierarchical representation of the input data, where the earlier features tend to be very general, getting increasingly specific towards the output. By feeding the network training data, propagated through the layers, the network is *trained* to perform useful tasks. A training data point (or, typically, a small batch of training points) is fed to the network, the outputs and local derivatives at each node are recorded, and the difference between the output prediction and the true label is measured by an *objective function*, such as *mean absolute error* (L1), *mean squared error* (L2), *cross-entropy loss*, or *Dice loss*, depending on the application. The derivative of the objective function with respect to the output is calculated and used as a feedback signal. The discrepancy is propagated backwards through the network and all the weights are updated to reduce the error. This is achieved using *backward propagation* [52–54], which calculates the gradient of the objective function with respect to the weights in each node using the chain rule together with dynamic programming, and *gradient descent* [55], an optimization algorithm tasked with improving the weights.

2.2 Deep learning

Traditionally, machine learning models are trained to perform useful tasks based on manually designed features extracted from the raw data, or features learned by other simple machine learning models. In deep learning, the computers learn useful representations and features automatically, directly from the raw data, bypassing this manual and difficult step. By far the most common models in deep learning are various variants of artificial neural networks, but there are others. The main common characteristic of deep learning methods is their focus on *feature learning*: automatically learning representations of data. This is the primary difference between deep learning approaches and more "classical" machine learning. Discovering features and performing a task is merged into one problem, and therefore both improved during the same training process. See [38,40] for general overviews of the field.

In medical imaging the interest in deep learning is mostly triggered by *convolutional neural networks* (CNNs) [56],[14] a powerful way to learn useful representations of images and other structured data. Before it became possible to use CNNs efficiently, these features typically had to be engineered by hand, or created by less powerful machine learning models. Once it became possible to use features learned directly from the data, many of the handcrafted image features were typically left by the wayside as they turned out to be almost worthless compared to feature detectors found by CNNs.[15] There are some strong preferences embedded in CNNs based

---

[14] Interestingly, CNNs was applied in medical image analysis already in the early 90s, e.g. [57], but with limited success.
[15] However, combining hand-engineered features with CNN features is a very reasonable approach when low amounts of training data makes it difficult to learn good features automatically.

on how they are constructed, which helps us understand why they are so powerful. Let us therefore take a look at the building blocks of CNNs.

2.3 Building blocks of convolutional neural networks

When applying neural networks to images one can in principle use the simple feedforward neural networks discussed above. However, having connections from all nodes of one layer to all nodes in the next is extremely inefficient. A careful pruning of the connections based on domain knowledge, i.e. the structure of images, leads to much better performance. A CNN is a particular kind of artificial neural network aimed at preserving spatial relationships in the data, with very few connections between the layers. The input to a CNN is arranged in a grid structure and then fed through layers that preserve these relationships, each layer operation operating on a small region of the previous layer (Fig. 2). CNNs are able to form highly efficient representation of the input data,[16] well-suited for image-oriented tasks. A CNN has multiple layers of *convolutions* and *activations*, often interspersed with *pooling* layers, and is trained using backpropagation and gradient descent as for standard artificial neural networks. See Section 2.1. In addition, CNNs typically have fully connected layers at the end, which compute the final outputs.[17]

(i) **Convolutional layers**: In the convolutional layers the activations from the previous layers are convolved with a set of small parameterized *filters*, frequently of size 3×3, collected in a tensor $W^{(j,i)}$, where j is the filter number and i is the layer number. By having each filter share the exact same weights across the whole input domain, i.e. translational equivariance at each layer, one achieves a drastic reduction in the number of weights that need to be learned. The motivation for this weight-sharing is that features appearing in one part of the image likely also appear in other parts. If you have a filter capable of detecting horizontal lines, say, then it can be used to detect them wherever they appear. Applying all the convolutional filters at all locations of the input to a convolutional layer produces a tensor of *feature maps*.

(ii) **Activation layer**: The feature maps from a convolutional layer are fed through nonlinear activation functions. This makes it possible for the entire neural network to approximate almost any nonlinear function [47,48].[18] The activation functions are generally the very simple rectified linear units, or ReLUs, defined as $\text{ReLU}(z) = \max(0, z)$, or variants like leaky ReLUs or parametric ReLUs.[19] See [59,60] for more information about these and other activation functions. Feeding the feature maps through an activation function produces new tensors, typically also called *feature maps*.

(iii) **Pooling**: Each feature map produced by feeding the data through one or more convolutional layer is then typically pooled in a *pooling layer*. Pooling operations take small grid regions as input and produce single numbers for each region. The number is usually computed by using the max function (*max-pooling*) or the average function (*average pooling*). Since a small shift of the input image results in small changes in the activation maps, the pooling layers gives the CNN some translational invariance.

A different way of getting the downsampling effect of pooling is to use convolutions with increased stride lengths. Removing the pooling layers simplifies the network architecture without necessarily sacrificing performance [61].

Other common elements in many modern CNNs include

(iv) **Dropout regularization**: A simple idea that gave a huge boost in the performance of CNNs. By averaging several models in an ensemble one tend to get better performance than when using single models. Dropout [62] is an averaging technique based on stochastic sampling of neural networks.[20] By randomly removing neurons during training one ends up using slightly different networks for each batch of training data, and the weights of the trained network are tuned based on optimization of multiple variations of the network.[21]

(v) **Batch normalization**: These layers are typically placed after activation layers, producing normalized activation maps by subtracting the mean and dividing by the standard deviation for each training batch. Including batch normalization layers forces the network to periodically change its activations to zero mean and unit standard deviation as the training batch hits these layers, which works as a regularizer for the network, speeds up training, and makes it less dependent on careful parameter initialization [66].

---

[16] It is interesting to compare this with the biological vision systems and their *receptive fields* of variable size (volumes in visual space) of neurons at different hierarchical levels.

[17] Lately, so-called *fully-convolution* CNNs have become popular, in which average pooling across the whole input after the final activation layer replaces the fully-connected layers, significantly reducing the total number of weights in the network.

[18] A neural network with only linear activations would only be able to perform linear approximation. Adding further layers wouldn't improve its expressiveness.

[19] Other options include exponential linear units (ELUs), and the now rarely used sigmoid or tanh activation functions.

[20] The idea of dropout is also used for other machine learning models, as in the DART technique for regression trees [63].

[21] In addition to increased model performance, dropout can also be used to produce robust uncertainty measures in neural networks. By leaving dropout turned on also during inference one effectively performs *variational inference* [58,64,65]. This relates standard deep neural networks to Bayesian neural networks, synthesized in the field of Bayesian deep learning.
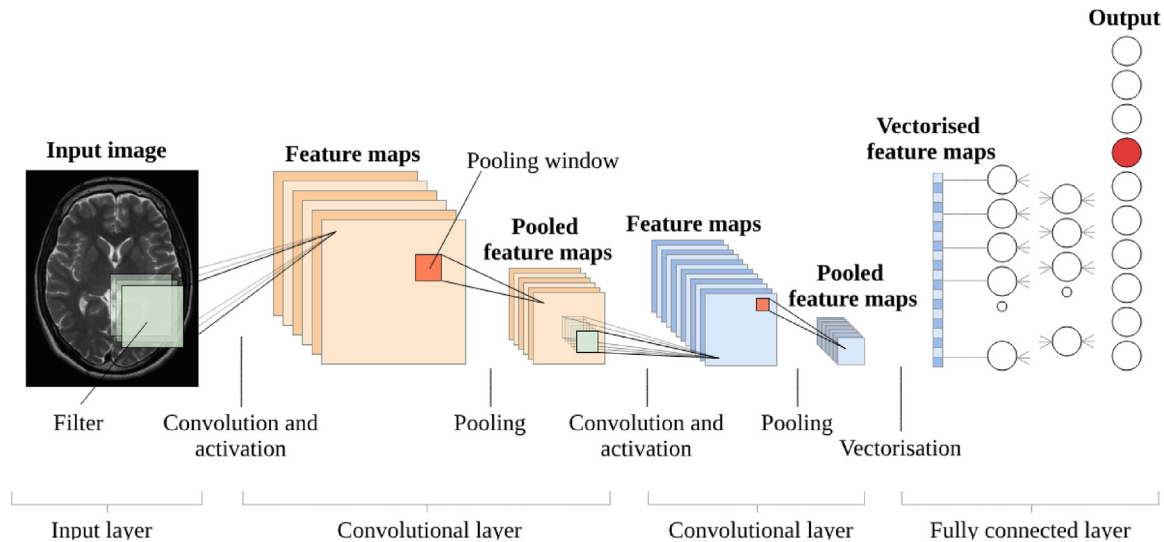
Figure 2. *Building blocks of a typical CNN*. A slight modification of a figure in [58], courtesy of the author.

In the design of new and improved CNN architectures, these components are combined in increasingly complicated and interconnected ways, or even replaced by other more convenient operations. When architecting a CNN for a particular task there are multiple factors to consider, including understanding the task to be solved and the requirements to be met, figuring out how to best feed the data to the network, and optimally utilizing one's budget for computation and memory consumption. In the early days of modern deep learning one tended to use very simple combinations of the building blocks, as in Lenet [56] and AlexNet [1]. Later network architectures are much more complex, each generation building on ideas and insights from previous architectures, resulting in updates to the state-of-the-art. Table 1 contains a short list of some famous CNN architectures, illustrating how the building blocks can be combined and how the field moves along.

These neural networks are typically implemented in one or more of a small number of software frameworks that dominates machine learning research, all built on top of NVIDIA's CUDA platform and the cuDNN library. Today's deep learning methods are almost exclusively implemented in either Tensor-Flow, a framework originating from Google Research, Keras, a deep learning library originally built by François Chollet and recently incorporated in TensorFlow, or Pytorch, a framework associated with Facebook Research. There are very few exceptions (YOLO built using the Darknet framework [85] is one of the rare ones). All the main frameworks are open source and under active development.

3 Deep learning, medical imaging and MRI

Deep learning methods are increasingly used to improve clinical practice, and the list of examples is long, growing daily. We will not attempt a comprehensive overview of deep learning in medical imaging, but merely sketch some of the landscape before going into a more systematic exposition of deep learning in MRI.

Convolutional neural networks can be used for efficiency improvement in radiology practices through protocol determination based on short-text classification [86]. They can also be used to reduce the gadolinium dose in contrast-enhanced brain MRI by an order of magnitude [87] without significant reduction in image quality. Deep learning is applied in radiotherapy [88], in PET-MRI attenuation correction [89,90], in *radiomics* [91,92] (see [93] for a review of radiomics related to radiooncology and medical physics), and for *theranostics* in neurosurgical imaging, combining confocal laser endomicroscopy with deep learning models for automatic detection of intraoperative CLE images on-the-fly [94].

Another important application area is advanced deformable image registration, enabling quantitative analysis across different physical imaging modalities and across time.[22] For example elastic registration between 3D MRI and transrectal ultrasound for guiding targeted prostate biopsy [95]; deformable registration for brain MRI where a "cue-aware deep regression network" learns from a given set of training images the displacement vector associated with a pair of reference-subject patches [96]; fast deformable image registration of brain MR image pairs by patch-wise prediction of the Large Deformation Diffeomorphic Metric Mapping model [97][23]; unsupervised convolutional neural network-based algorithm for deformable image registration of cone-beam CT to CT using a deep convolutional inverse graphics network [98]; deep learning-based 2D/3D

---

[22] E.g. test-retest examinations, or motion correction in dynamic imaging.
[23] Available at https://github.com/rkwitt/quicksilver.

Table 1
A far from exhaustive, non-chronological, list of CNN architectures and some high-level descriptions.

AlexNet	[1]	The network that launched the current deep learning boom by winning the 2012 ILSVRC competition by a huge margin. Notable features include the use of RELUs, dropout regularization, splitting the computations on multiple GPUs, and using data augmentation during training. ZFNet [67], a relatively minor modification of AlexNet, won the 2013 ILSVRC competition.
VGG	[68]	Popularized the idea of using smaller filter kernels and therefore deeper networks (up to 19 layers for VGG19, compared to 7 for AlexNet and ZFNet), and training the deeper networks using pre-training on shallower versions.
GoogLeNet	[69]	Promoted the idea of stacking the layers in CNNs more creatively, as *networks in networks*, building on the idea of [70]. Inside a relatively standard architecture (called the *stem*), GoogLeNet contains multiple *inception modules*, in which multiple different filter sizes are applied to the input and their results concatenated. This multi-scale processing allows the module to extract features at different levels of detail simultaneously. GoogLeNet also popularized the idea of not using fully-connected layers at the end, but rather global average pooling, significantly reducing the number of model parameters. It won the 2014 ILSVRC competition.
ResNet	[71]	Introduced *skip connections*, which makes it possible to train much deeper networks. A 152 layer deep ResNet won the 2015 ILSVRC competition, and the authors also successfully trained a version with *1001* layers. Having skip connections in addition to the standard pathway gives the network the option to simply copy the activations from layer to layer (more precisely, from ResNet block to ResNet block), preserving information as data goes through the layers. Some features are best constructed in shallow networks, while others require more depth. The skip connections facilitate both at the same time, increasing the network's flexibility when fed input data. As the skip connections make the network learn residuals, ResNets perform a kind of boosting.
Highway nets	[72]	Another way to increase depth based on *gating units*, an idea from Long Short Term Memory (LSTM) recurrent networks, enabling optimization of the skip connections in the network. The gates can be trained to find useful combinations of the identity function (as in ResNets) and the standard nonlinearity through which to feed its input.
DenseNet	[73]	Builds on the ideas of ResNet, but instead of adding the activations produced by one layer to later layers, they are simply concatenated together. The original inputs in addition to the activations from previous layers are therefore kept at each layer (again, more precisely, between blocks of layers), preserving some kind of global state. This encourages feature reuse and lowers the number of parameters for a given depth. DenseNets are therefore particularly well-suited for smaller data sets (outperforming others on e.g. Cifar-10 and Cifar-100).
ResNext	[74]	Builds on ResNet and GoogLeNet by using inception modules between skip connections.
SENets	[75]	Squeeze-and-Excitation Networks, which won the ILSVRC 2017 competition, builds on ResNext but adds trainable parameters that the network can use to weigh each feature map, where earlier networks simply added them up. These SE-blocks allows the network to model the channel and spatial information separately, increasing the model capacity. SE-blocks can easily be added to any CNN model, with negligible increase in computational costs.
NASNet	[76]	A CNN architecture designed by a neural network, beating all the previous human-designed networks at the ILSVRC competition. It was created using AutoML,[a] Google Brain's reinforcement learning approach to architecture design [77]. A controller network (a recurrent neural network) proposes architectures aimed to perform at a specific level for a particular task, and by trial and error learns to propose better and better models. NASNet was based on Cifar-10, and has relatively modest computational demands, but still outperformed the previous state-of-the-art on ILSVRC data.
YOLO	[78]	Introduced a new, simplified way to do simultaneous object detection and classification in images. It uses a single CNN operating directly on the image and outputting bounding boxes and class probabilities. It incorporates several elements from the above networks, including inception modules and pretraining a smaller version of the network. It's fast enough to enable real-time processing.[b] YOLO makes it easy to trade accuracy for speed by reducing the model size. YOLOv3-tiny was able to process images at over 200 frames per second on a standard benchmark data set, while still producing reasonable predictions.
GANs	[79]	A generative adversarial network consists of two neural networks pitted against each other. The *generative network* G is tasked with creating samples that the *discriminative network D* is supposed to classify as coming from the generative network or the training data. The networks are trained simultaneously, where G aims to maximize the probability that D makes a mistake while D aims for high classification accuracy.
Siamese nets	[80]	An old idea (e.g. [81]) that's recently been shown to enable *one-shot learning*, i.e. learning from a single example. A Siamese network consists of two identical neural networks, both the architecture and the weights, attached at the end. They are trained together to *differentiate* pairs of inputs. Once trained, the features of the networks can be used to perform one-shot learning without retraining.
U-net	[82]	A very popular and successful network for segmentation in 2D images. When fed an input image, it is first downsampled through a "traditional" CNN, before being upsampled using transpose convolutions until it reaches its original size. In addition, based on the ideas of ResNet, there are skip connections that concatenates features from the downsampling to the upsampling paths. It is a fully-convolutional network, using the ideas first introduced in [83].
V-net	[84]	A three-dimensional version of U-net with volumetric convolutions and skip connections as in ResNet.

[a] https://cloud.google.com/automl.

[b] You can watch YOLO in action here https://youtu.be/VOC3huqHrss.

registration frame-work for registration of preoperative 3D data and intraoperative 2D X-ray images in image-guided therapy [99]; real-time prostate segmentation during targeted prostate biopsy, utilizing temporal information in the series of ultrasound images [100].

This is just a tiny sliver of the many applications of deep learning to central problems in medical imaging. There are several thorough reviews and overviews of the field to consult for more information, across modalities and organs, and with different points of view and level of technical details. For example the comprehensive review [101],[24] covering both medicine and biology and spanning from imaging applications in healthcare to protein-protein interaction and uncertainty quantification; key concepts of deep learning for clinical radiologists [102–111], including radiomics and imaging genomics (radiogenomics) [112], and toolkits and libraries for deep learning [113]; deep learning in neuroimaging and neuroradiology [114]; brain segmentation [115]; stroke imaging [116,117]; neuropsychiatric disorders [118]; breast cancer [119,120]; chest imaging [121]; imaging in oncology [122–124]; medical ultrasound [125,126]; and more technical surveys of deep learning in medical image analysis [41,127–129]. Finally, for those who like to be hands-on, there are many instructive introductory deep learning tutorials available online. For example [130], with accompanying code available at https://github.com/paras42/Hello_World_Deep_Learning, where you'll be guided through the construction of a system that can differentiate a chest X-ray from an abdominal X-ray using the `Keras/TensorFlow` framework through a Jupyter Notebook. Other nice tutorials are http://bit.ly/adltktutorial, based on the Deep Learning Toolkit (DLTK) [131], and https://github.com/usuyama/pydata-medical-image, based on the Microsoft Cognitive Toolkit (CNTK).

Let's now turn to the field of MRI, in which deep learning has seen applications at each step of entire workflows. From acquisition to image retrieval, from segmentation to disease prediction. We divide this into two parts: (**i**) the signal processing chain close to the physics of MRI, including image restoration and multimodal image registration (Fig. 3), and (**ii**) the use of deep learning in MR image segmentation, disease detection, disease prediction and systems based on images and text data (reports), addressing a few selected organs such as the brain, the kidney, the prostate and the spine (Fig. 4).

3.1 From image acquisition to image registration

Deep learning in MRI has typically been focused on segmentation and classification of reconstructed magnitude images. Its penetration into the lower levels of MRI

measurement techniques is more recent, but already impressive. From MR image acquisition and signal processing in MR fingerprinting, to denoising and super-resolution, and into image synthesis.

3.1.1 Data acquisition and image reconstruction

Research on CNN and RNN-based image reconstruction methods is rapidly increasing, pioneered by Yang et al. [132] at NIPS 2016 and Wang et al. [133] at ISBI 2016. Recent applications addresses e.g. convolutional recurrent neural networks for dynamic MR image reconstruction [134], reconstructing good quality cardiac MR images from highly undersampled complex-valued k-space data by learning spatio-temporal dependencies, outperforming 3D CNN approaches and compressed sensing-based dynamic MRI reconstruction algorithms in computational complexity, reconstruction accuracy and speed for different undersampling rates. Schlemper et al. [135] created a *deep cascade* of concatenated CNNs for dynamic MR image reconstruction, making use of data augmentation, both rigid and elastic deformations, to increase the variation of the examples seen by the network and reduce overfitting.[25] Using *variational networks* for single-shot fast spin-echo MRI with variable density sampling, Chen et al. [136] enabled real-time (200 ms per section) image reconstruction, outperforming conventional parallel imaging and compressed sensing reconstruction. In [137], the authors explored the potential for *transfer learning* (pretrained models) and assessed the generalization of learned image reconstruction regarding image contrast, SNR, sampling pattern and image content, using a variational network and true measurement k-space data from patient knee MRI recordings and synthetic k-space data generated from images in the Berkeley Segmentation Data Set and Benchmarks. Employing least-squares generative adversarial networks (GANs) that learns texture details and suppresses high-frequency noise, [138] created a novel compressed sensing framework that can produce diagnostic quality reconstructions "on the fly" (30 ms).[26] A unified framework for image reconstruction [139], called *automated transform by manifold approximation* (AUTOMAP) consisting of a feedforward deep neural network with fully connected layers followed by a sparse convolutional autoencoder, formulate image reconstruction generically as a data-driven supervised learning task that generates a mapping between the sensor and the image domain based on an appropriate collection of training data (e.g. MRI examinations collected from the Human Connectome Project, transformed to the k-space sensor domain).

---

[24] A continuous collaborative manuscript (https://greenelab.github.io/deep-review) with >500 references.

[25] Code available at https://github.com/js3611/Deep-MRI-Reconstruction.

[26] In their GAN setting, a generator network is used to map undersampled data to a realistic-looking image with high measurement fidelity, while a discriminator network is trained jointly to score the quality of the reconstructed image.
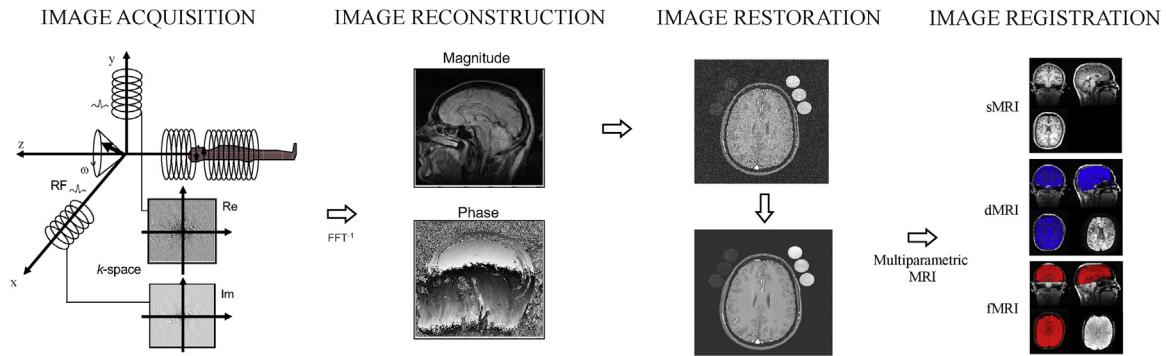
Figure 3. Deep learning in the MR signal processing chain, from image acquisition (in complex-valued *k*-space) and image reconstruction, to image restoration (e.g. denoising) and image registration. The rightmost column illustrates coregistration of multimodal brain MRI. sMRI = structural 3D T1-weighted MRI, dMRI = diffusion weighted MRI (stack of slices in blue superimposed on sMRI), fMRI = functional BOLD MRI (in red).
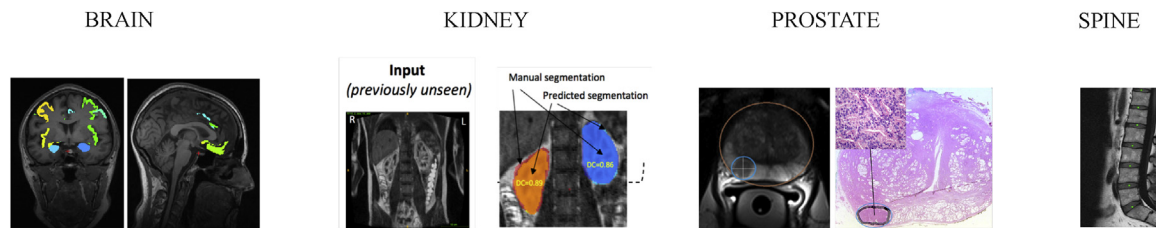

Figure 4. Deep learning for MR image analysis in selected organs, partly from ongoing work at MMIV.

There are also other approaches and reports on deep learning in MR image reconstruction, e.g. [140–143], a fundamental field rapidly progressing.

3.1.2 Quantitative parameters – QSM and MR fingerprinting

Another area that is developing within deep learning for MRI is the estimation of quantitative tissue parameters from recorded complex-valued data. For example within *quantitative susceptibility mapping*, and in the exciting field of *magnetic resonance fingerprinting*.

Quantitative susceptibility mapping (QSM) is a growing field of research in MRI, aiming to noninvasively estimate the magnetic susceptibility of biological tissue [144,145]. The technique is based on solving the difficult, ill-posed inverse problem of determining the magnetic susceptibility from local magnetic fields. Recently Yoon et al. [146] constructed a three-dimensional CNN, named QSMnet and based on the U-Net architecture, able to generate high quality susceptibility source maps from single orientation data. The authors generated training data by using the gold-standard for QSM: the so-called COSMOS method [147]. The data was based on 60 scans from 12 healthy volunteers. The resulting model both simplified and improved the state-of-the-art for QSM. Rasmussen and coworkers [148] took a different approach. They also used a U-Net-based convolutional neural network to perform field-to-source inversion, called *DeepQSM*, but it

was trained on synthetically generated data containing simple geometric shapes such as cubes, rectangles and spheres. After training their model on synthetic data it was able to generalize to real-world clinical brain MRI data, computing susceptibility maps within seconds end-to-end. The authors conclude that their method, combined with fast imaging sequences, could make QSM feasible in standard clinical practice.

Magnetic resonance fingerprinting (MRF) was introduced a little more than five years ago [149], and has been called "a promising new approach to obtain standardized imaging biomarkers from MRI" by the European Society of Radiology [150]. It uses a pseudo-randomized acquisition that causes the signals from different tissues to have a unique signal evolution ("fingerprint") that is a function of the multiple material properties being investigated. Mapping the signals back to known tissue parameters (T1, T2 and proton density) is then a rather difficult inverse problem. MRF is closely related to the idea of compressed sensing [151] in MRI [152] in that MRF undersamples data in *k*-space producing aliasing artifacts in the reconstructed images that can be suppressed by compressed sensing.[27] It can be regarded as a quantitative multiparametric MRI analysis, and with recent acquisition schemes using a single-shot spiral trajectory with undersampling, whole-brain

---

[27] See [153–157] for recent perspectives and developments connecting deep learning-based reconstruction methods to the more general research field of inverse problems.

coverage of T_1, T_2 and proton density maps can be acquired at $1.2 \times 1.2 \times 3\,mm^3$ voxel resolution in less than 5 min [158].

The processing of MRF after acquisition usually involves using various pattern recognition algorithms that try to match the fingerprints to a predefined dictionary of predicted signal evolutions,[28] created using the Bloch equations [149,163].

Recently, deep learning methodology has been applied to MR fingerprinting. Cohen et al. [164] reformulated the MRF reconstruction problem as learning an optimal function that maps the recorded signal magnitudes to the corresponding tissue parameter values, trained on a sparse set of dictionary entries. To achieve this they fed voxel-wise MRI data acquired with an MRF sequence (MRF-EPI, 25 frames in ∼3 s; or MRF-FISP, 600 frames in ∼7.5 s) to a four-layer neural network consisting of two hidden layers with 300×300 fully connected nodes and two nodes in the output layer, considering only T_1 and T_2 parametric maps. The network, called MRF Deep RecOnstruction NEtwork (DRONE), was trained by an adaptive moment estimation stochastic gradient descent algorithm with a mean squared error loss function. Their dictionary consisted of ∼70,000 entries (product of discretized T_1 and T_2 values) and training the network to convergence with this dictionary (∼10 MB for MRF-EPI and ∼300 MB for MRF-FISP) required 10 to 70 min using an NVIDIA K80 GPU with 2 GB memory. They found their reconstruction time (10 to 70 ms per slice) to be 300 to 5000 times faster than conventional dictionary-matching techniques, using both well-characterized calibrated ISMRM/NIST phantoms and in vivo human brains.

A similar deep learning approach to predict quantitative parameter values (T_1 and T_2) from MRF time series was taken by Hoppe et al. [165]. In their experiments they used 2D MRF-FISP data with variable TR (12–15 ms), flip angles (5–74°) and 3000 repetitions, recorded on a MAGNETOM 3T Skyra. A high resolution dictionary was simulated to generate a large collection of training and testing data, using tissues T_1 and T_2 relaxation time ranges as present in normal brain at 3T (e.g. [166]) resulting in ∼1.2×10^5 time series. In contrast to [164], their deep neural network architecture was inspired from the domain of speech recognition due to the similarity of the two tasks. The architecture with the smallest average error for validation data was a standard convolutional neural network consisting of an input layer of 3000 nodes (number of samples in the recorded time series), four hidden layers, and an output layers with two nodes (T_1 and T_2). Matching one time series was about 100 times faster than the conventional [149] matching method and with very small mean absolute deviations from ground truth values.

In the same context, Fang et al. [167] used a deep learning method to extract tissue properties from highly undersampled 2D MRF-FISP data in brain imaging, where 2300 time points were acquired from each measurement and each time point consisted of data from one spiral readout only. The real and imaginary parts of the complex signal were separated into two channels. They used MRF signal from a patch of 32×32 pixels to incorporate correlated information between neighboring pixels. In their work they designed a standard three-layer CNN with T_1 and T_2 as output.

Virtue et al. [168] investigated a different approach to MRF. By generating 100,000 synthetic MRI signals using a Bloch equation simulator they were able to train feedforward deep neural networks to map new MRI signals to the tissue parameters directly, producing approximate solutions to the inverse mapping problem of MRF. In their work they designed a new complex activation function, the complex cardioid, that was used to construct a complex-valued feedforward neural network. This three-layer network outperformed both the standard MRF techniques based on dictionary matching, and also the analogous real neural network operating on the real and imaginary components separately. This suggested that complex-valued networks are better suited at uncovering information in complex data.[29]

3.1.3 Image restoration (denoising, artifact detection)

Estimation of noise and image denoising in MRI has been an important field of research for many years [171,172], employing a plethora of methods. For example Bayesian Markov random field models [173], rough set theory [174], higher-order singular value decomposition [175], wavelets [176], independent component analysis [177], or higher order PDEs [178].

Recently, deep learning approaches have been introduced to denoising. In their work on learning implicit brain MRI manifolds using deep neural networks, Bermudez et al. [179] implemented an autoencoder with skip connections for image denoising, testing their approach with adding various levels of Gaussian noise to more than 500 T1-weighted brain MR images from healthy controls in the Baltimore Longitudinal Study of Aging. Their autoencoder network outperformed the current FSL SUSAN denoising software according to peak signal-to-noise ratios. Benou et al. [180] addressed spatio-temporal denoising of dynamic contrast-enhanced MRI of the brain with bolus injection of contrast agent (CA), proposing a novel approach using ensembles of deep neural networks for noise reduction. Each DNN was trained on a different range of SNRs and types of CA concentration time curves (denoted "pathology experts", "healthy experts", "vessel experts") to

---

[28] A dictionary of time series for every possible combination of *parameters* like (discretized) T_1 and T_2 relaxation times, spin-density (M_0), B_0, off-resonance (Δf), and also voxel-wise cerebral blood volume (CBV), mean vessel radius (R), blood oxygen saturation (SO_2) and T_2^* [159–161], and more, e.g. MFR-ASL [162].

---

[29] Complex-valued deep learning is also getting some attention in a broader community of researchers, and has been shown to lead to improved models. See e.g. [169,170] and the references therein.

generate a reconstruction hypothesis from noisy input by using a classification DNN to select the most likely hypothesis and provide a "clean output" curve. Training data was generated synthetically using a three-parameter Tofts pharmacokinetic (PK) model and noise realizations. To improve this model, accounting for spatial dependencies of PK pharmacokinetics, they used concatenated noisy time curves from first-order neighbourhood pixels in their expert DNNs and ensemble hypothesis DNN, collecting neighboring reconstructions before a boosting procedure produced the final clean output for the pixel of interest. They tested their trained ensemble model on 33 patients from two different DCE-MRI databases with either stroke or recurrent glioblastoma (RIDER NEURO[30]), acquired at different sites, with different imaging protocols, and with different scanner vendors and field strengths. The qualitative and quantitative (MSE) denoising results were better than spatiotemporal Beltrami, moving average, the dynamic Non Local Means method [181], and stacked denoising autoencoders [182]. The run-time comparisons were also in favor of the proposed sDNN. In this context of DCE-MRI, it's tempting to speculate whether deep neural network approaches could be used for *direct estimation* of tracer-kinetic parameter maps from highly undersampled (k, t)-space data in dynamic recordings [183,184], a powerful way to by-pass 4D DCE-MRI reconstruction altogether and map sensor data directly to spatially resolved pharmacokinetic parameters, e.g. K^{trans}, v_p, v_e in the extended Tofts model or parameters in other classic models [185]. A related approach in the domain of diffusion MRI, by-passing the model-fitting steps and computing voxelwise scalar tissue properties (e.g. radial kurtosis, fiber orientation dispersion index) directly from the subsampled DWIs was taken by Golkov et al. [186] in their proposed "q-space deep learning" family of methods.

Deep learning methods has also been applied to MR *artifact detection*, e.g. poor quality spectra in MRSI [187]; detection and removal of ghosting artifacts in MR spectroscopy [188]; and automated reference-free detection of patient motion artifacts in MRI [189].

3.1.4 Image super-resolution

Image super-resolution, reconstructing a higher-resolution image or image sequence from the observed low-resolution image [190], is an exciting application of deep learning methods.[31]

Super-resolution for MRI have been around for almost 10 years [191,192] and can be used to improve the trade-off between resolution, SNR, and acquisition time [193], generate 7T-like MR images on 3T MRI scanners [194], or obtain super-resolution T_1 maps from a set of low resolution T_1

weighted images [195]. Recently deep learning approaches has been introduced, e.g. generating super-resolution single (no reference information) and multi-contrast (applying a high-resolution image of another modality as reference) brain MR images using CNNs [196]; constructing superresolution brain MRI by a CNN stacked by multi-scale fusion units [197]; and super-resolution musculoskeletal MRI ("DeepResolve") [198]. In DeepResolve thin (0.7 mm) slices in knee images (DESS) from 124 patients included in the Osteoarthritis Initiative were used for training and 17 patients for testing, with a 10 s inference time per 3D ($344 \times 344 \times 160$) volume. The resulting images were evaluated both quantitatively (MSE, PSNR, and the perceptual window-based structural similarity SSIM[32] index) and qualitatively by expert radiologists.

3.1.5 Image synthesis

Image synthesis in MRI have traditionally been seen as a method to derive new parametric images or new tissue contrast from a collection of MR acquisition performed at the same imaging session, i.e. "an intensity transformation applied to a given set of input images to generate a new image with a specific tissue contrast" [199]. Another avenue of MRI synthesis is related to quantitative imaging and the development and use of physical phantoms, imaging calibration/standard test objects with specific material properties. This is done in order to assess the performance of an MRI scanner or to assess imaging biomarkers reliably with application-specific phantoms such as a structural brain imaging phantom, DCE-MRI perfusion phantom, diffusion phantom, flow phantom, breast phantom or a proton-density fat fraction phantom [200]. The *in silico* modeling of MR images with certain underlying properties, e.g. [201,202], or model-based generation of large databases of (cardiac) images from real healthy cases [203] is also part of this endeavour. In this context, deep learning approaches have accelerated research and the amount of costly training data.

The last couple of years have seen impressive results for photo-realistic image synthesis using deep learning techniques, especially generative adversarial networks (GANs, introduced by Goodfellow et al. in 2014 [79]), e.g. [204–206]. These can also be used for biological image synthesis [207,208] and text-to-image synthesis [209–211].[33] Recently, a group of researchers from NVIDIA, MGH & BWH Center for Clinical Data Science in Boston, and the Mayo Clinic in Rochester [212] designed a clever approach to generate synthetic abnormal MRI images with brain tumors by training a GAN based on pix2pix[34] using two publicly available data sets of brain MRI (ADNI and the BRATS'15 Challenge, and later also the Ischemic Stroke Lesion Segmentation ISLES'2018

Challenge). This approach is highly interesting as medical imaging datasets are often imbalanced, with few pathological findings, limiting the training of deep learning models. Such generative models for image synthesis serve as a form of data augmentation, and also as an anonymization tool. The authors achieved comparable tumor segmentation results when trained on the synthetic data rather than on real patient data. A related approach to brain tumor segmentation using coarse-to-fine GANs was taken by Mok and Chung [213]. Guibas et al. [214] used a two-stage pipeline for generating synthetic medical images from a pair of GANs, addressing retinal fundus images, and provided an online repository (SynthMed) for synthetic medical images. Kitchen and Seah [215] used GANs to synthetize realistic prostate lesions in T_2, ADC, K^{trans} resembling the SPIE-AAPM-NCI ProstateX Challenge 2016[35] training data.

Other applications are unsupervised synthesis of T1-weighted brain MRI using a GAN [179]; image synthesis with context-aware GANs [216]; synthesis of patient-specific transmission image for PET attenuation correction in PET/MR imaging of the brain using a CNN [217]; pseudo-CT synthesis for pelvis PET/MR attenuation correction using a Dixon-VIBE Deep Learning (DIVIDE) network [218]; image synthesis with GANs for tissue recognition [219]; synthetic data augmentation using a GAN for improved liver lesion classification [220]; and deep MR to CT synthesis using unpaired data [221].

3.1.6 Image registration

Image registration[36] is an increasingly important field within MR image processing and analysis as more complementary and multiparametric tissue information are collected in space and time within shorter acquisition times, at higher spatial (and temporal) resolutions, often longitudinally, and across patient groups, larger cohorts, or atlases. Traditionally one has divided the tasks of image registration into dichotomies: intra vs. inter-modality, intra vs. inter-subject, rigid vs. deformable, geometry-based vs. intensity-based, and prospective vs. retrospective image registration. Mathematically, registration is a challenging mix of geometry (spatial transformations), analysis (similarity measures), optimization strategies, and numerical schemes. In prospective motion correction, real-time MR physics is also an important part of the picture [223,224]. A wide range of methodological approaches have been developed and tested for various organs

and applications[37] [228–237], including "previous generation" artificial neural networks [238].

Recently, deep learning methods have been applied to image registration in order to improve accuracy and speed (e.g. Section 3.4 in [41]). For example: deformable image registration [97,239]; model-to-image registration [240,241]; MRI-based attenuation correction for PET [242,243]; PET/MRI dose calculation [244]; unsupervised end-to-end learning for deformable registration of 2D CT/MR images [245]; an unsupervised learning model for deformable, pairwise 3D medical image registration by Balakrishnan et al. [246][38]; and a deep learning framework for unsupervised affine and deformable image registration [247].

3.2 From image segmentation to diagnosis and prediction

We leave the lower-level applications of deep learning in MRI to consider higher-level (down-stream) applications such as fast and accurate image segmentation, disease prediction in selected organs (brain, kidney, prostate, and spine) and content-based image retrieval, typically applied to reconstructed magnitude images. We have chosen to focus our overview on deep learning applications close to the MR physics and will be brief in the present section, even if the following applications are very interesting and clinically important.

3.2.1 Image segmentation

Image segmentation, the holy grail of quantitative image analysis, is the process of partitioning an image into multiple regions that share similar attributes, enabling localization and quantification.[39] It has an almost 50 years long history, and has become the biggest target for deep learning approaches in medical imaging. The multispectral tissue classification report by Vannier et al. in 1985 [248], using statistical pattern recognition techniques (and satellite image processing software from NASA), represented one of the most seminal works leading up to today's machine learning in medical imaging segmentation. In this early era, we also had the opportunity to contribute with supervised and unsupervised machine learning approaches for MR image segmentation and tissue classification [249–252]. An impressive range of segmentation methods and approaches have been reported (especially for brain segmentation) and reviewed, e.g. [253–261]. MR image segmentation using deep learning approaches, typically CNNs, are now penetrating the whole field of applications.

---

[36] Image registration can be defined as "the determination of a one-to-one mapping between the coordinates in one space and those in another, such that points in the two spaces that correspond to the same anatomical point are mapped to each other" (C.R Maurer [222], 1993).

[37] And different hardware e.g. GPUs [225–227] as image registration is often computationally time consuming.
[38] With code available at https://github.com/voxelmorph/voxelmorph.
[39] Segmentation is also crucial for functional imaging, enabling tissue physiology quantification with preservation of anatomical specificity.

For example acute ischemic lesion segmentation in DWI [262]; brain tumor segmentation [263]; segmentation of the striatum [264]; segmentation of organs-at-risks in head and neck CT images [265]; and fully automated segmentation of polycystic kidneys [266]; deformable segmentation of the prostate [267]; and spine segmentation with 3D multiscale CNNs [268]. See [41,101] for more comprehensive lists.

3.2.2 Diagnosis and prediction

A presumably complete list of papers up to 2017 using deep learning techniques for brain image analysis is provided as Table 1 in Litjens at al. [41]. In the Table 2 we add some more recent work on organ-specific deep learning using MRI, restricting ourselves to brain, kidney, prostate and spine.

3.3 Content-based image retrieval

The objective of content-based image retrieval (CBIR) in radiology is to provide medical cases similar to a given image in order to assist radiologists in the decision-making process. It typically involves large case databases, clever image representations and lesion annotations, and algorithms that are able to quickly and reliably match and retrieve the most similar images and their annotations in the case database. CBIR has been an active area of research in medical imaging for many years, addressing a wide range of applications, imaging modalities, organs, and methodological approaches, e.g. [297–303], and at a larger scale outside the medical field using deep learning techniques, e.g. at Microsoft, Apple, Facebook, and Google (reverse image search[40]), and others. See e.g. [304–308] and the code repositories https://github.com/topics/image-retrieval. One of the first application of deep learning for CBIR in the medical domain came in 2015 when Sklan et al. [309] trained a CNN to perform CBIR with more than one million random MR and CT images, with disappointing results (true positive rate of 20%) on their independent test set of 2100 labeled images. Medical CBIR is now, however, dominated by deep learning algorithms [310–312]. As an example, by retrieving medical cases similar to a given image, Pizarro et al. [278] developed a CNN for automatically inferring the contrast of MRI scans based on the image intensity of multiple slices. Recently, deep learning methods have also been used for automated generation of radiology reports, typically incorporating long-short-term-memory (LSTM) network models to generate the

textual paragraphs [313–316], and also to identify findings in radiology reports [317–319].

4 Open science and reproducible research in machine learning for medical imaging

Machine learning is moving at a breakneck speed, too fast for the standard peer-review process to keep up. Many of the most celebrated and impactful papers in machine learning over the past few years are only available as preprints, or published in conference proceedings long after their results are well-known and incorporated in the research of others. Bypassing peer-review has some downsides, of course, but these are somewhat mitigated by researchers' willingness to share code and data.[41]

Most of the main new ideas and methods are posted to the arXiv preprint server,[42] and the accompanying code shared on the GitHub platform.[43] The data sets used are often openly available through various repositories. This, in addition to the many excellent online educational resources,[44] makes it easy to get started in the field. Select a problem you find interesting based on openly available data, a method described in a preprint, and an implementation uploaded to GitHub. This forms a good starting point for an interesting machine learning project.

Another interesting aspect about modern machine learning and data science is the prevalence of *competitions*, with the annual ImageNet ILSVRC competition as the main driver of progress in deep learning for computer vision since 2012. Each competition typically draws large number of participants, and the top results often push the state-of-the art to a new level. In addition to inspiring new ideas, competitions also provide natural entry points to modern machine learning. It is interesting to note how deep learning-based models are completely dominating the leaderboards of essentially all image-based competitions. Other machine learning models, or non-machine learning-based techniques, have largely been outclassed.

What's true about the openness of machine learning in general is increasingly true also for the sub-field of machine learning for medical image analysis. We've listed a few examples of openly available implementations, data sets and challenges in Tables 3–5.

---

[40] See "search by image" https://images.google.com/https://developers.google.com/custom-search, and also https://tineye.com, indexing more than 30 billion images.

[41] In the spirit of sharing and open science, we've created a GitHub repository to accompany our article, available at https://github.com/MMIV-ML/DLMI2018.

[42] http://arxiv.org.

[43] https://github.com.

[44] For example http://www.fast.ai, https://www.deeplearning.ai, http://cs231n.stanford.edu, https://developers.google.com/machine-learning/crash-course.

Table 2

A short list of deep learning applications per *organ*, *task*, *reference* and *description*.

Brain

Brain extraction	[269]	A 3D CNN for skull stripping
Functional connectomes	[270]	Transfer learning approach to enhance deep neural network classification of brain functional connectomes
	[271]	Multisite diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI
Structural connectomes	[272]	A convolutional neural network-based approach (https://github.com/MIC-DKFZ/TractSeg) that directly segments tracts in the field of fiber orientation distribution function (fODF) peaks without using tractography, image registration or parcellation. Tested on 105 subjects from the Human Connectome Project
Brain age	[273]	Chronological age prediction from raw brain T1-MRI data, also testing the heritability of brain-predicted age using a sample of 62 monozygotic and dizygotic twins
Alzheimer's disease	[274]	Landmark-based deep multi-instance learning evaluated on 1526 subjects from three public datasets (ADNI-1, ADNI-2, MIRIAD)
	[275]	Identify different stages of AD
	[276]	Multimodal and multiscale deep neural networks for the early diagnosis of AD using structural MR and FDG-PET images
Vascular lesions	[277]	Evaluation of a deep learning approach for the segmentation of brain tissues and white matter hyperintensities of presumed vascular origin in MRI
Identification of MRI contrast	[278]	Using deep learning algorithms to automatically identify the brain MRI contrast, with implications for managing large databases
Meningioma	[279]	Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI
Glioma	[280]	Glioblastoma segmentation using heterogeneous MRI data from clinical routine
	[281]	Deep learning for segmentation of brain tumors and impact of cross-institutional training and testing
	[282]	Automatic semantic segmentation of brain gliomas from MRI using a deep cascaded neural network
	[283]	AdaptAhead optimization algorithm for learning deep CNN applied to MRI segmentation of glioblastomas (BRATS)
Multiple sclerosis	[284]	Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls

Kidney

Abdominal organs	[285]	CNNs to improve abdominal organ segmentation, including left kidney, right kidney, liver, spleen, and stomach in T_2-weighted MR images
Cyst segmentation	[266]	An artificial multi-observer deep neural network for fully automated segmentation of polycystic kidneys
Renal transplant	[286]	A deep-learning-based classifier with stacked non-negative constrained autoencoders to distinguish between rejected and non-rejected renal transplants in DWI recordings

Prostate

Cancer (PCa)	[287]	Proposed a method for end-to-end prostate segmentation by integrating holistically (image-to-image) nested edge detection with fully convolutional networks. Their nested networks automatically learn a hierarchical representation that can improve prostate boundary detection. Obtained very good results (Dice coefficient, 5-fold cross validation) on MRI scans from 250 patients
	[288]	Computer-aided diagnosis with a CNN, deciding 'cancer' 'no cancer' trained on data from 301 patients with a prostate-specific antigen level of <20 ng/mL who underwent MRI and extended systematic prostate biopsy with or without MRI-targeted biopsy
	[289]	Automatic approach based on deep CNN, inspired from VGG, to classify PCa and noncancerous tissues with multiparametric MRI using data from the PROSTATEx database
	[290]	Deep CNN and a non-deep learning using feature detection (the scale-invariant feature transform and the bag-of-words model, a representative method for image recognition and analysis) were used to distinguish pathologically confirmed PCa patients from prostate benign conditions patients with prostatitis or prostate benign hyperplasia in a collection of 172 patients with more than 2500 morphologic 2D T_2-w MR images
	[291]	Designed a system which can concurrently identify the presence of PCa in an image and localize lesions based on deep CNN features (co-trained CNNs consisting of two parallel convolutional networks for ADC and T_2-w images respectively) and a single-stage SVM classifier for automated detection of PCa in multiparametric MRI. Evaluated on a dataset of 160 patients
	[292]	Designed and tested multimodel CNNs, using clinical data from 364 patients with a total of 463 PCa lesions and 450 identified noncancerous image patches. Carefully investigated three critical factors which could greatly affect the performance of their multimodal CNNs but had not been carefully studied previously: (1) Given limited training data, how can these be augmented in sufficient numbers and variety for fine-tuning deep CNN networks for PCa diagnosis? (2) How can multimodal mp-MRI information be effectively combined in CNNs? (3) What is the impact of different CNN architectures on the accuracy of PCa diagnosis?

Table 2 (*Continued*)

Spine

Vertebrae labeling	[293]	Designed a CNN for detection and labeling of vertebrae in MR images with clinical annotations as training data
Intervertebral disc localization	[268]	3D multi-scale fully connected CNNs with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images
Disc-level labeling, spinal stenosis grading	[294]	CNN model denoted DeepSPINE, having a U-Net architecture combined with a spine-curve fitting method for automated lumbar vertebral segmentation, disc-level designation, and spinal stenosis grading with a natural language processing scheme
Lumbal neural forminal stenosis (LNFS)	[295]	Addressed the challenge of automated pathogenesis-based diagnosis, simultaneously localizing and grading multiple spinal structures (neural foramina, vertebrae, intervertebral discs) for diagnosing LNFS and discover pathogenic factors. Proposed a deep multiscale multitask learning network integrating a multiscale multi-output learning and a multitask regression learning into a fully convolutional network where (i) a DMML-Net merges semantic representations to reinforce the salience of numerous target organs (ii) a DMML-Net extends multiscale convolutional layers as multiple output layers to boost the scale-invariance for various organs, and (iii) a DMML-Net joins the multitask regression module and the multitask loss module to combine the mutual benefit between tasks
Spondylitis vs tuberculosis	[296]	CNN model for differentiating between tuberculous and pyogenic spondylitis in MR images. Compared their CNN performance with that of three skilled radiologists using spine MRIs from 80 patients
Metastasis	[290]	A multi-resolution approach for spinal metastasis detection using deep Siamese neural networks comprising three identical subnetworks for multi-resolution analysis and detection. Detection performance was evaluated on a set of 26 cases using a free-response receiver operating characteristic analysis (observer is free to mark and rate as many suspicious regions as are considered clinically reportable)

Table 3

A short list of openly available code for ML in medical imaging.

Summary	Reference	Implementation
NiftyNet. An open source convolutional neural networks platform for medical image analysis and image-guided therapy	[320,321]	http://niftynet.io
DLTK. State of the art reference implementations for deep learning on medical images	[131]	https://github.com/DLTK/DLTK
DeepMedic	[322]	https://github.com/Kamnitsask/deepmedic
U-Net: Convolutional Networks for Biomedical Image Segmentation	[323]	https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net
V-net	[84]	https://github.com/faustomilletari/VNet
SegNet: A Deep Convolutional EncoderDecoder Architecture for Robust Semantic Pixel-Wise Labelling	[324]	https://mi.eng.cam.ac.uk/projects/segnet
Brain lesion synthesis using GANs	[212]	https://github.com/khcs/brain-synthesis-lesion-segmentation
GANCS: Compressed Sensing MRI based on Deep Generative Adversarial Network	[325]	https://github.com/gongenhao/GANCS
Deep MRI Reconstruction	[135]	https://github.com/js3611/Deep-MRI-Reconstruction
Graph Convolutional Networks for brain analysis in populations, combining imaging and non-imaging data	[326]	https://github.com/parisots/population-gcn

5 Challenges, limitations and future perspectives

It is clear that deep neural networks are very useful when one is tasked with producing accurate decisions based on complicated data sets. But they come with some significant challenges and limitations that you either have to accept or try to overcome. Some are general: from technical challenges related to the lack of mathematical and theoretical underpinnings of many central deep learning models and techniques, and the resulting difficulty in deciding exactly what it is that makes one model better than another, to societal challenges related to maximization and spread of the technological benefits [327,328] and the problems related to the tremendous

amounts of hype and excitement.[45] Others are more domain-specific.

In deep learning for standard computer vision tasks, like object recognition and localization, powerful models and a set of best practices have been developed over the last few years. The pace of development is still incredibly high, but certain things seem to be settled, at least momentarily. Using

---

[45] Lipton: *Machine Learning: The Opportunity and the Opportunists* https://www.technologyreview.com/video/612109, Jordan: *Artificial Intelligence – The Revolution Hasn't Happened Yet* https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7.

Table 4

A short list of medical imaging data sets and repositories.

Name	Summary	Link
OpenNeuro	An open platform for sharing neuroimaging data under the public domain license. Contains brain images from 168 studies (4,718 participants) with various imaging modalities and acquisition protocols.	https://openneuro.org[a]
UK Biobank	Health data from half a million participants. Contains MRI images from 15,000 participants, aiming to reach 100,000.	http://www.ukbiobank.ac.uk/
TCIA	The cancer imaging archive hosts a large archive of medical images of cancer accessible for public download. Currently contains images from 14,355 patients across 77 collections.	http://www.cancerimagingarchive.net
ABIDE	The autism brain imaging data exchange. Contains 1114 datasets from 521 individuals with Autism Spectrum Disorder and 593 controls.	http://fcon_1000.projects.nitrc.org/indi/abide
ADNI	The Alzheimer's disease neuroimaging initiative. Contains image data from almost 2000 participants (controls, early MCI, MCI, late MCI, AD)	http://adni.loni.usc.edu/

[a] Data can be downloaded from the AWS S3 Bucket https://registry.opendata.aws/openneuro.

Table 5

A short list of medical imaging competitions.

Name	Summary	Link
Grand-Challenges	Grand challenges in biomedical image analysis. Hosts and lists a large number of competitions	https://grand-challenge.org/
RSNA Pneumonia Detection Challenge	Automatically locate lung opacities on chest radiographs	https://www.kaggle.com/c/rsna-pneumonia-detection-challenge
HVSMR 2016	Segment the blood pool and myocardium from a 3D cardio-vascular magnetic resonance image	http://segchd.csail.mit.edu/
ISLES 2018	Ischemic Stroke Lesion Segmentation 2018. The goal is to segment stroke lesions based on acute CT perfusion data.	http://www.isles-challenge.org/
BraTS 2018	Multimodal Brain Tumor Segmentation. The goal is to segment brain tumors in multimodal MRI scans.	http://www.med.upenn.edu/sbia/brats2018.html
CAMELYON17	The goal is to develop algorithms for automated detection and classification of breast cancer metastases in whole-slide images of histological lymph node sections.	https://camelyon17.grand-challenge.org/Home
ISIC 2018	Skin Lesion Analysis Towards Melanoma Detection	https://challenge2018.isic-archive.com/
Kaggle's 2018 Data Science Bowl	Spot Nuclei. Speed Cures.	https://www.kaggle.com/c/data-science-bowl-2018
Kaggle's 2017 Data Science Bowl	Turning Machine Intelligence Against Lung Cancer	https://www.kaggle.com/c/data-science-bowl-2017
Kaggle's 2016 Data Science Bowl	Transforming How We Diagnose Heart Disease	https://www.kaggle.com/c/second-annual-data-science-bowl
MURA	Determine whether a bone X-ray is normal or abnormal	https://stanfordmlgroup.github.io/competitions/mura/

the basic building blocks described above, placed according to the ideas behind, say, ResNet and SENet, will easily result in close to state-of-the-art performance on two-dimensional object detection, image classification and segmentation tasks.

However, the story for deep learning in medical imaging is not quite as settled. One issue is that medical images are often three-dimensional, and three-dimensional convolutional neural networks are as well-developed as their 2D counterparts. One quickly meet challenges associated to

memory and compute consumption when using CNNs with higher-dimensional image data, challenges that researchers are trying various approaches to deal with (treating 3D as stacks of 2Ds, patch- or segment-based training and inference, downscaling, etc.). It is clear that the ideas behind state-of-the-art two-dimensional CNNs can be lifted to three dimensions, but also that adding a third spatial dimension results in additional constraints. Other important challenges are related to data, trust, interpretability, workflow integration, and regulations, as discussed below.

5.1 Data

This is a crucially important obstacle for deep neural networks, especially in medical data analysis. When deploying deep neural networks, or any other machine learning model, one is instantly faced with challenges related to data access, privacy issues, data protection, and more.

As privacy and data protection is often a requirement when dealing with medical data, new techniques for training models without exposing the underlying training data to the user of the model are necessary. It is not enough to merely restrict access to the training set used to construct the model, as it is easy to use the model itself to discover details about the training set [329]. Even hiding the model and only exposing a prediction interface would still leave it open to attack, for example in the form of model-inversion [330] and membership attacks [331]. Most current work on deep learning for medical data analysis use either open, anonymized data sets (as those in Table 4), or locally obtained anonymized research data, making these issues less relevant. However, the general deep learning community are focusing a lot of attention on the issue of privacy, and new techniques and frameworks for *federated learning* [332][46] and *differential privacy* [333–335] are rapidly improving. There are a few examples of these ideas entering the medical machine learning community, as in [336] where the distribution of deep learning models among several medical institutions was investigated, but then without considering the above privacy issues. As machine learning systems in medicine grows to larger scales, perhaps even including computations and learning on the "edge", federated learning and differential privacy will likely become the focus of much research in our community.

If you are able to surmount these obstacles, you will be confronted with deep neural networks' insatiable appetite for training data. These are very inefficient models, requiring large number of training samples before they can produce anything remotely useful, and labeled training data is typically both expensive and difficult to produce. In addition, the training data has to be representative of the data the network will meet in the future. If the training samples are from a data distribution that is very different from the one met in the real world, then the network's generalization performance will be lower than expected. See [337] for a recent exploration of this issue. Considering the large difference between the high-quality images one typically work with when doing research and the messiness of the real, clinical world, this can be a major obstacle when putting deep learning systems into production.

Luckily there are ways to alleviate these problems somewhat. A widely used technique is *transfer learning*, also called fine-tuning or pre-training: first you train a network to perform a task where there is an abundance of data, and then you copy weights from this network to a network designed for the task at hand. For two-dimensional images one will almost always use a network that has been pre-trained on the ImageNet data set. The basic features in the earlier layers of the neural network found from this data set typically retain their usefulness in any other image-related task (or are at least form a better starting point than random initialization of the weights, which is the alternative). Starting from weights tuned on a larger training data set can also make the network more robust. Focusing the weight updates during training on later layers requires less data than having to do significant updates throughout the entire network. One can also do interorgan transfer learning in 3D, an idea we have used for kidney segmentation, where pre-training a network to do brain segmentation decreased the number of annotated kidneys needed to achieve good segmentation performance [338]. The idea of pre-training networks is not restricted to images. Pre-training entire models has recently been demonstrated to greatly impact the performance of natural language processing systems [2–4].

Another widely used technique is *augmenting* the training data set by applying various transformations that preserves the labels, as in rotations, scalings and intensity shifts of images, or more advanced data augmentation techniques like anatomically sound deformations, or other data set specific operations (for example in our work on kidney segmentation from DCE-MRI, where we used image registration to propagate labels through a time course of images [339]). *Data synthesis*, as in [212], is another interesting approach.

In short, as expert annotators are expensive, or simply not available, spending large computational resources to expand your labeled training data set, e.g. indirectly through transfer learning or directly through data augmentation, is typically worthwhile. But whatever you do, the way current deep neural networks are constructed and trained results in significant data size requirements. There are new ways of constructing more data-efficient deep neural networks on the horizon, for example by encoding more domain-specific elements in the neural network structure as in the *capsule* systems of [340,341], which adds *viewpoint invariance*. It is also possible to add *attention mechanisms* to neural networks [342,343], enabling them to focus their resources on the most informative components of each layer input.

However, the networks that are most frequently used, and with the best raw performance, remain the data-hungry standard deep neural networks.

5.2 Interpretability, trust and safety

As deep neural networks relies on complicated interconnected hierarchical representations of the training data to produce its predictions, interpreting these predictions becomes very difficult. This is the "black box" problem of deep neural networks [344]. They are capable of producing extremely accurate predictions, but how can you trust predictions based on features you cannot understand?

---

[46] See for example https://ai.googleblog.com/2017/04/federated-learning-collaborative.html.

Considerable effort goes into developing new ways to deal with this problem, including DARPA launching a whole program "*Explainable AI* "[47] dedicated to this issue, and lots of research going into enhancing interpretability [345,346], and finding new ways to measure sensitivity and visualize features [67,347–350].

Another way to increase their trustworthiness is to make them produce robust uncertainty estimates in addition to predictions. The field of *Bayesian Deep Learning* aims to combine deep learning and Bayesian approaches to uncertainty. The ideas date back to the early 90s [351–353], but the field has recently seen renewed interest from the machine learning community at large, as new ways of computing uncertainty estimates from state of the art deep learning models have been developed [58,64,354]. In addition to producing valuable measures that function as uncertainty measures [65,355,356], these techniques can also lessen deep neural networks susceptibility to *adversarial attacks* [354,357].

5.3 Workflow integration, regulations

Another stumbling block for successful incorporation of deep learning methods is *workflow integration*. It is possible to end up developing clever machine learning system for clinical use that turn out to be practically useless for actual clinicians. Attempting to augment already established procedures necessitates knowledge of the entire workflow. Involving the end-user in the process of creating and evaluating systems can make this a little less of an issue, and can also increase the end users' trust in the systems,[48] as you can establish a feedback loop during the development process. But still, even if there is interest on the "ground floor" and one is able to get prototype systems into the hands of clinicians, there are many higher-ups to convince and regulatory, ethical and legal hurdles to overcome.

5.4 Perspectives and future expectations

Deep learning in medical data analysis is here to stay. Even though there are many challenges associated to the introduction of deep learning in clinical settings, the methods produce results that are too valuable to discard. This is illustrated by the tremendous amounts of high-impact publications in top-journals dealing with deep learning in medical imaging (for example [16,20,29,31,39,89,136,139,161,272,284], all published in 2018). As machine learning researchers and practitioners gain more experience, it will become easier to classify problems according to what solution approach is the most reasonable: (i) best approached using deep learning techniques end-to-end, (ii) best tackled by a combination of deep learning with other techniques, or (iii) no deep learning component at all.

Beyond the application of machine learning in medical imaging, we believe that the attention in the medical community can also be leveraged to strengthen the general computational mindset among medical researchers and practitioners, mainstreaming the field of *computational medicine*.[49] Once there are enough high-impact software-systems based on mathematics, computer science, physics and engineering entering the daily workflow in the clinic, the acceptance for other such systems will likely grow. The access to bio-sensors and (edge) computing on *wearable devices* for monitoring disease or lifestyle, plus an ecosystem of machine learning and other computational medicine-based technologies, will then likely facilitate the transition to a *new medical paradigm* that is **p**redictive, **p**reventive, **p**ersonalized, and **p**articipatory – P4 medicine [359].[50]

Acknowledgements

References

[1] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Advances in neural information processing systems 25. Curran Associates, Inc.; 2012. p. 1097–105.

[2] Peters M, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, et al. Deep contextualized word representations. In: Proceedings of the 2018 conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (long papers), vol. 1. 2018. p. 2227–37.

[3] Howard J, Ruder S. Universal language model fine-tuning for text classification. In: Proceedings of the 56th annual meeting of the Association for Computational Linguistics (volume 1: long papers). 2018. p. 328–39.

[4] Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language understanding by generative pre-training; 2018.

[5] Xiong W, Wu L, Alleva F, Droppo J, Huang X, Stolcke A. The Microsoft 2017 conversational speech recognition system. In: Proc. speech and signal processing (ICASSP) 2018 IEEE int. conf. acoustics. 2018. p. 5934–8.

[6] van den Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, et al. WaveNet: a generative model for raw audio; 2016, arXiv:1609.03499v2.

[7] Guo C, Berkhahn F. Entity embeddings of categorical variables; 2016, arXiv:1604.06737.

---

[8] De Brébisson A, Simon É, Auvolat A, Vincent P, Bengio Y. Artificial neural networks applied to taxi destination prediction; 2015, arXiv preprint arXiv:1508.00021.

[9] George D, Huerta E. Deep learning for real-time gravitational wave detection and parameter estimation: results with advanced LIGO data. Phys Lett B 2018;778:64–70.

[10] George D, Shen H, Huerta E. Classification and unsupervised clustering of LIGO data with deep transfer learning. Phys Rev D 2018;97:101501.

[11] Shen H, George D, Huerta E, Zhao Z. Denoising gravitational waves using deep learning with recurrent denoising autoencoders, arXiv:1711.09919 (2017).

[12] Raissi M, Karniadakis GE. Hidden physics models: machine learning of nonlinear partial differential equations. J Comput Phys 2018;357:125–41.

[13] Karpatne A, Atluri G, Faghmous JH, Steinbach M, Banerjee A, Ganguly A, et al. Theory-guided data science: a new paradigm for scientific discovery from data. IEEE Trans Knowl Data Eng 2017;29:2318–31.

[14] Gartner. Top strategic technology trends for 2018; 2018.

[15] Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B, et al. Deep learning for health informatics. IEEE J Biomed Health Inform 2017;21:4–21.

[16] Ganapathy N, Swaminathan R, Deserno TM. Deep learning on 1-D biosignals: a taxonomy-based survey. Yearbook Med Inform 2018;27:98–109.

[17] Kuhlmann L, Lehnertz K, Richardson MP, Schelter B, Zaveri HP. Seizure prediction – ready for a new era. Nat Rev Neurol 2018.

[18] Kwon J-M, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. J Am Heart Assoc 2018;7.

[19] Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 2016;35:1285–98.

[20] Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 2018;172:1122–31, e9.

[21] Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol 2018;18:24.

[22] Jimènez J, Škalič M, Martínez-Rosell G, De Fabritiis G. KDEEP: protein-ligand absolute binding affinity prediction via 3D-Convolutional Neural Networks. J Chem Inf Model 2018;58:287–96.

[23] Kalinin AA, Higgins GA, Reamaroon N, Soroushmehr S, Allyn-Feuer A, Dinov ID, et al. Deep learning in pharmacogenomics: from gene regulation to patient stratification. Pharmacogenomics 2018;19:629–50.

[24] Jiang S, Chin K-S, Tsui KL. A universal deep learning approach for modeling the flow of patients under different severities. Comput Methods Programs Biomed 2018;154:191–203.

[25] Vranas KC, Jopling JK, Sweeney TE, Ramsey MC, Milstein AS, Slatore CG, et al. Identifying distinct subgroups of ICU patients: a machine learning approach. Crit Care Med 2017;45:1607–15.

[26] Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. NPJ Dig Med 2018;1:18.

[27] Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. IEEE J Biomed Health Inform 2017.

[28] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 2016;316:2402–10.

[29] Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell M, Corrado G, et al. Predicting cardiovascular risk factors in retinal fundus photographs using deep learning. Nat Biomed Eng 2018.

[30] Poplin R, Chang P-C, Alexander D, Schwartz S, Colthurst T, Ku A, et al. A universal SNP and small-indel variant caller using deep neural networks. Nat Biotechnol 2018.

[31] De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med 2018;24:1342.

[32] Qin Y, Kamnitsas K, Ancha S, Nanavati J, Cottrell G, Criminisi A, et al. Autofocus layer for semantic segmentation, arXiv:1805.08403 (2018).

[33] Kamnitsas K, Baumgartner C, Ledig C, Newcombe V, Simpson J, Kane A, et al. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In: International conference on information processing in medical imaging. 2017. p. 597–609.

[34] Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. J Am Med Inform Assoc 2018.

[35] Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, et al. Mastering the game of Go without human knowledge. Nature 2017;550:354.

[36] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8.

[37] Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng 2018;2:158.

[38] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436.

[39] Hinton G. Deep learning a technology with the potential to transform health care; 2018. p. 1–2.

[40] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016. www.deeplearningbook.org.

[41] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.

[42] Marblestone AH, Wayne G, Kording KP. Toward an integration of deep learning and neuroscience. Front Comput Neurosci 2016;10:94.

[43] Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. Neuron 2017;95:245–58.

[44] Banino A, Barry C, Uria B, Blundell C, Lillicrap T, Mirowski P, et al. Vector-based navigation using grid-like representations in artificial agents. Nature 2018;557:429.

[45] Cybenko G. Approximation by superpositions of a sigmoidal function. Math Control Signals Syst 1989;2:303–14.

[46] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw 1989;2:359–66.

[47] Leshno M, Lin VY, Pinkus A, Schocken S. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. Neural Netw 1993;6:861–7.

[48] Sonoda S, Murata N. Neural network with unbounded activation functions is universal approximator. Appl Comput Harm Anal 2017;43:233–68.

[49] Nielsen MA. Neural networks and deep learning. Determination Press; 2015.

[50] Aggarwal CC. Neural networks and deep learning. Springer; 2018.

[51] Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev 1958;65:386.

[52] Linnainmaa S. The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors [Master's thesis]. Univ. Helsinki; 1970. p. 6–7 [in Finnish].

[53] Werbos P. Beyond regression: new tools for prediction and analysis in the behavioral sciences [Ph.D. dissertation]. Harvard University; 1974.

[54] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533.

[55] Cauchy A. Méthode générale pour la résolution des systemes déquations simultanées. C R Sci Paris 1847;25:536–8.

[56] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86:2278–324.

[57] Lo SC, Freedman MT, Lin JS, Mun SK. Automatic lung nodule detection using profile matching and back-propagation neural network techniques. J Digit Imaging 1993;6:48–54.

[58] Murray S. An exploratory analysis of multi-class uncertainty approximation in Bayesian convolution neural networks [Master's thesis]. University of Bergen; 2018.

[59] Clevert D-A, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (ELUS), arXiv:1511.07289 (2015).

[60] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. 2015. p. 1026–34.

[61] Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M. Striving for simplicity: the all convolutional net, arXiv:1412.6806 (2014).

[62] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15:1929–58.

[63] Rashmi K, Gilad-Bachrach R. Dart: dropouts meet multiple additive regression trees. In: International conference on artificial intelligence and statistics. 2015. p. 489–97.

[64] Gal Y. Uncertainty in deep learning [Ph.D. thesis]. University of Cambridge; 2016.

[65] Wickstrøm K, Kampffmeyer M, Jenssen R. Uncertainty modeling and interpretability in convolutional neural networks for polyp segmentation. In: 2018 IEEE 28th international workshop on machine learning for signal processing (MLSP), IEEE. 2018. p. 1–6.

[66] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning. 2015. p. 448–56.

[67] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: European conference on computer vision, Springer. 2014. p. 818–33.

[68] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556 (2014).

[69] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. p. 1–9.

[70] Lin M, Chen Q, Yan S. Network in network, arXiv:1312.4400 (2013).

[71] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 770–8.

[72] Srivastava RK, Greff K, Schmidhuber J. Training very deep networks. In: Advances in neural information processing systems. 2015. p. 2377–85.

[73] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: CVPR, vol. 1. 2016. p. 3.

[74] Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR), IEEE. 2016. p. 5987–95.

[75] Hu J, Shen L, Sun G. Squeeze-and-excitation networks, arXiv:1709.01507 (2017).

[76] Zoph B, Vasudevan V, Shlens J, Le QV. Learning transferable architectures for scalable image recognition, arXiv:1707.07012 2 (2017).

[77] Bello I, Zoph B, Vasudevan V, Le QV. Neural optimizer search with reinforcement learning. In: Precup D, Teh YW, editors. Proceedings of the 34th international conference on machine learning. Proceedings of machine learning research, vol. 70. Sydney, Australia: PMLR, International Convention Centre; 2017. p. 459–68.

[78] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 779–88.

[79] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, editors. Advances in neural information processing systems 27. Curran Associates, Inc.; 2014. p. 2672–80.

[80] Koch G, Zemel R, Salakhutdinov R. Siamese neural networks for one-shot image recognition. In: ICML deep learning workshop, vol. 2. 2015.

[81] Bromley J, Guyon I, LeCun Y, Säckinger E, Shah R. Signature verification using a "Siamese" time delay neural network. In: Advances in neural information processing systems. 1993. p. 737–44.

[82] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. 2015. p. 234–41.

[83] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 3431–40.

[84] Milletari F, Navab N, Ahmadi S-A. V-net: fully convolutional neural networks for volumetric medical image segmentation. In: 2016 fourth international conference on 3D Vision (3DV), IEEE. 2016. p. 565–71.

[85] Redmon J. Darknet: open source neural networks in C. http://pjreddie.com/darknet/, 2013–2016.

[86] Lee YH. Efficiency improvement in a busy radiology practice: determination of musculoskeletal magnetic resonance imaging protocol using deep learning convolutional neural networks. J Digit Imaging 2018.

[87] Gong E, Pauly JM, Wintermark M, Zaharchuk G. Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI. J Magn Reson Imaging 2018;48:330–40.

[88] Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. Comput Biol Med 2018;98:126–46.

[89] Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. Radiology 2018;286:676–84.

[90] Mehranian A, Arabi H, Zaidi H. Vision 20/20: magnetic resonance imaging-guided attenuation correction in PET/MRI: challenges, solutions, and opportunities. Med Phys 2016;43:1130–55.

[91] Lao J, Chen Y, Li Z-C, Li Q, Zhang J, Liu J, et al. A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. Sci Rep 2017;7:10353.

[92] Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley AP, Palmer LJ. Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. Sci Rep 2017;7:1648.

[93] Peeken JC, Bernhofer M, Wiestler B, Goldberg T, Cremers D, Rost B, et al. Radiomics in radiooncology – challenging the medical physicist. Phys Med 2018;48:27–36.

[94] Izadyyazdanabadi M, Belykh E, Mooney MA, Eschbacher JM, Nakaji P, Yang Y, et al. Prospects for theranostics in neurosurgical imaging: empowering confocal laser endomicroscopy diagnostics via deep learning. Front Oncol 2018;8:240.

[95] Haskins G, Kruecker J, Kruger U, Xu S, Pinto PA, Wood BJ, Yan P. Learning deep similarity metric for 3D MR-TRUS registration, arXiv:1806.04548v1 (2018).

[96] Cao X, Yang J, Zhang J, Wang Q, Yap P-T, Shen D. Deformable image registration using a cue-aware deep regression network. IEEE Trans Bio-med Eng 2018;65:1900–11.

[97] Yang X, Kwitt R, Styner M, Niethammer M. Quicksilver: Fast predictive image registration – a deep learning approach. Neuroimage 2017;158:378–96.

[98] Kearney VP, Haaf S, Sudhyadhom A, Valdes G, Solberg D. An unsupervised convolutional neural network-based algorithm for deformable image registration. Phys Med Biol 2018.

[99] Zheng J, Miao S, Jane Wang Z, Liao R. Pairwise domain adaptation module for CNN-based 2-D/3-D registration. J Med Imaging (Bellingham, Wash.) 2018;5:021204.

[100] Anas EMA, Mousavi P, Abolmaesumi P. A deep learning approach for real time prostate segmentation in freehand ultrasound guided biopsy. Med Image Anal 2018;48:107–16.

[101] Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin A, Do BT, Way GP, et al. Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 2018;15.

[102] Lee J-G, Jun S, Cho Y-W, Lee H, Kim GB, Seo JB, et al. Deep learning in medical imaging: general overview. Korean J Radiol 2017;18:570–84.

[103] Rueckert D, Glocker B, Kainz B. Learning clinically useful information from images: past, present and future. Med Image Anal 2016;33:13–8.

[104] Chartrand G, Cheng PM, Vorontsov E, Drozdzal M, Turcotte S, Pal CJ, et al. Deep learning: a primer for radiologists. Radiographics 2017;37:2113–31.

[105] Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics 2017;37:505–15.

[106] Mazurowski MA, Buda M, Saha A, Bashir MR. Deep learning in radiology: an overview of the concepts and a survey of the state of the art, arXiv:1802.08717v1 (2018).

[107] McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, et al. Deep learning in radiology. Acad Radiol 2018.

[108] Savadjiev P, Chong J, Dohan A, Vakalopoulou M, Reinhold C, Paragios N, et al. Demystification of AI-driven medical image interpretation: past, present and future. Eur Radiol 2018.

[109] Thrall JH, Li X, Li Q, Cruz C, Do S, Dreyer K, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. J Am Coll Radiol: JACR 2018;15:504–8.

[110] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018.

[111] Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Deep learning with convolutional neural network in radiology. Jpn J Radiol 2018;36:257–72.

[112] Giger ML. Machine learning in medical imaging. J Am Coll Radiol: JACR 2018;15:512–20.

[113] Erickson BJ, Korfiatis P, Akkus Z, Kline T, Philbrick K. Toolkits and libraries for deep learning. J Digit Imaging 2017;30:400–5.

[114] Zaharchuk G, Gong E, Wintermark M, Rubin D, Langlotz CP. Deep learning in neuroradiology. AJNR Am J Neuroradiol 2018.

[115] Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ. Deep learning for brain MRI segmentation: state of the art and future directions. J Digit Imaging 2017;30:449–59.

[116] Lee E-J, Kim Y-H, Kim N, Kang D-W. Deep into the brain: artificial intelligence in stroke imaging. J Stroke 2017;19:277–85.

[117] Feng R, Badgeley M, Mocco J, Oermann EK. Deep learning guided stroke management: a review of clinical applications. J Neurointervent Surg 2018;10:358–62.

[118] Vieira S, Pinaya WHL, Mechelli A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. Neurosci Biobehav Rev 2017;74:58–75.

[119] Burt JR, Torosdagli N, Khosravan N, RaviPrakash H, Mortazi A, Tissavirasingham F, et al. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. Br J Radiol 2018;91:20170545.

[120] Samala RK, Chan H-P, Hadjiiski LM, Helvie MA, Cha KH, Richter CD. Multi-task transfer learning deep convolutional neural network: application to computer-aided diagnosis of breast cancer on mammograms. Phys Med Biol 2017;62:8894–908.

[121] van Ginneken B. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. Radiol Phys Technol 2017;10:23–32.

[122] Morin O, Vallires M, Jochems A, Woodruff HC, Valdes G, Braunstein SE, et al. A deep look into the future of quantitative imaging in oncology: a statement of working principles and proposal for change. Int J Radiat Oncol Biol Phys 2018.

[123] Parmar C, Barry JD, Hosny A, Quackenbush J, Aerts JWL. Data analysis strategies in medical imaging. Clin Cancer Res 2018;24:3492–9.

[124] Xue Y, Chen S, Qin J, Liu Y, Huang B, Chen H. Application of deep learning in automated analysis of molecular images in cancer: a survey. Contrast Media Mol Imaging 2017;2017:9512370.

[125] Brattain LJ, Telfer BA, Dhyani M, Grajo JR, Samir AE. Machine learning for medical ultrasound: status, methods, and future opportunities. Abdom Radiol 2018;43:786–99.

[126] Huang Q, Zhang F, Li X. Machine learning in ultrasound computer-aided diagnostic systems: a survey. BioMed Res Int 2018;2018:5137904.

[127] Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. Annu Rev Biomed Eng 2017;19:221–48.

[128] Suzuki K. Overview of deep learning in medical imaging. Radiol Phys Technol 2017;10:257–73.

[129] Cao C, Liu F, Tan H, Song D, Shu W, Li W, et al. Deep learning and its applications in biomedicine. Genomics Proteomics Bioinform 2018;16:17–32.

[130] Lakhani P, Gray DL, Pett CR, Nagy P, Shih G. Hello world deep learning in medical imaging. J Digit Imaging 2018.

[131] Pawlowski N, Ktena SI, Lee MC, Kainz B, Rueckert D, Glocker B, et al. DLTK: state of the art reference implementations for deep learning on medical images, arXiv:1711.06853 (2017).

[132] Yang Y, Sun J, Li H, Xu Z. Deep ADMM-Net for compressive sensing MRI. In: Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R, editors. Advances in neural information processing systems 29. Curran Associates, Inc.; 2016. p. 10–8.

[133] Wang S, Su Z, Ying L, Peng X, Zhu S, Liang F, et al. Accelerating magnetic resonance imaging via deep learning. In: 2016 IEEE 13th international symposium on biomedical imaging (ISBI), IEEE. 2016. p. 514–7.

[134] Qin C, Hajnal JV, Rueckert D, Schlemper J, Caballero J, Price AN. Convolutional recurrent neural networks for dynamic MR image reconstruction. IEEE Trans Med Imaging 2018.

[135] Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. IEEE Trans Med Imaging 2018;37:491–503.

[136] Chen F, Taviani V, Malkiel I, Cheng JY, Tamir JI, Shaikh J, et al. Variable-density single-shot fast Spin-Echo MRI with deep learning reconstruction by using variational networks. Radiology 2018:180445.

[137] Knoll F, Hammernik K, Kobler E, Pock T, Recht MP, Sodickson DK. Assessment of the generalization of learned image reconstruction and the potential for transfer learning. Magn Reson Med 2018.

[138] Mardani M, Gong E, Cheng JY, Vasanawala SS, Zaharchuk G, Xing L, et al. Deep generative adversarial neural networks for compressive sensing (GANCS) MRI. IEEE Trans Med Imaging 2018.

[139] Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature 2018;555:487–92.

[140] Eo T, Jun Y, Kim T, Jang J, Lee H-J, Hwang D. KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images. Magn Reson Med 2018;80:2188–201.

[141] Han Y, Yoo J, Kim HH, Shin HJ, Sung K, Ye JC. Deep learning with domain adaptation for accelerated projection-reconstruction MR. Magn Reson Med 2018;80:1189–205.

[142] Shi J, Liu Q, Wang C, Zhang Q, Ying S, Xu H. Super-resolution reconstruction of MR image with a novel residual learning network algorithm. Phys Med Biol 2018;63:085011.

[143] Yang G, Yu S, Dong H, Slabaugh G, Dragotti PL, Ye X, et al. DAGAN: deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. IEEE Trans Med Imaging 2018;37:1310–21.

[144] Deistung A, Schäfer A, Schweser F, Biedermann U, Turner R, Reichenbach JR. Toward in vivo histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and r2-imaging at ultra-high magnetic field strength. Neuroimage 2013;65:299–314.

[145] Deistung A, Schweser F, Reichenbach JR. Overview of quantitative susceptibility mapping. NMR Biomed 2017;30:e3569.

[146] Yoon J, Gong E, Chatnuntawech I, Bilgic B, Lee J, Jung W, et al. Quantitative susceptibility mapping using deep neural network: QSMnet. Neuroimage 2018;179:199–206.

[147] Liu T, Spincemaille P, De Rochefort L, Kressler B, Wang Y. Calculation of susceptibility through multiple orientation sampling (COSMOS): a method for conditioning the inverse problem from measured magnetic field map to susceptibility source image in MRI. Magn Reson Med 2009;61:196–204.

[148] Rasmussen KGB, Kristensen MJ, Blendal RG, Ostergaard LR, Plocharski M, O'Brien K, et al. DeepQSM-using deep learning to solve the dipole inversion for MRI susceptibility mapping. Biorxiv 2018:278036.

[149] Ma D, Gulani V, Seiberlich N, Liu K, Sunshine JL, Duerk JL, et al. Magnetic resonance fingerprinting. Nature 2013;495:187–92.

[150] E. S. of Radiology (ESR). Magnetic resonance fingerprinting – a promising new approach to obtain standardized imaging biomarkers from MRI. Insights Imaging 2015;6:163–5.

[151] Donoho DL. Compressed sensing. IEEE Trans Inf Theory 2006;52:1289–306.

[152] Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. Magn Reson Med 2007;58:1182–95.

[153] McCann MT, Jin KH, Unser M. Convolutional neural networks for inverse problems in imaging: a review. IEEE Signal Process Mag 2017;34:85–95.

[154] Shah V, Hegde C. Solving linear inverse problems using GAN priors: an algorithm with provable guarantees, arXiv:1802.08406 (2018).

[155] Lucas A, Iliadis M, Molina R, Katsaggelos AK. Using deep neural networks for inverse problems in imaging: beyond analytical methods. IEEE Signal Process Mag 2018;35:20–36.

[156] Aggarwal HK, Mani MP, Jacob M. Modl: model based deep learning architecture for inverse problems. IEEE Trans Med Imaging 2018.

[157] Li H, Schwab J, Antholzer S, Haltmeier M. Nett: solving inverse problems with deep neural networks, arXiv:1803.00092 (2018).

[158] Ma D, Jiang Y, Chen Y, McGivney D, Mehta B, Gulani V, et al. Fast 3D magnetic resonance fingerprinting for a whole-brain coverage. Magn Reson Med 2018;79:2190–7.

[159] Christen T, Pannetier NA, Ni WW, Qiu D, Moseley ME, Schuff N, et al. MR vascular fingerprinting: a new approach to compute cerebral blood volume, mean vessel radius, and oxygenation maps in the human brain. Neuroimage 2014;89:262–70.

[160] Lemasson B, Pannetier N, Coquery N, Boisserand LSB, Collomb N, Schuff N, et al. MR vascular fingerprinting in stroke and brain tumors models. Sci Rep 2016;6:37071.

[161] Rieger B, Akakaya M, Pariente JC, Llufriu S, Martinez-Heras E, Weingrtner S, et al. Time efficient whole-brain coverage with MR fingerprinting using slice-interleaved echo-planar imaging. Sci Rep 2018;8:6667.

[162] Wright KL, Jiang Y, Ma D, Noll DC, Griswold MA, Gulani V, et al. Estimation of perfusion properties with MR fingerprinting arterial spin labeling. Magn Reson Imaging 2018;50:68–77.

[163] Panda A, Mehta BB, Coppo S, Jiang Y, Ma D, Seiberlich N, et al. Magnetic resonance fingerprinting-an overview. Curr Opin Biomed Eng 2017;3:56–66.

[164] Cohen O, Zhu B, Rosen MS. MR fingerprinting deep reconstruction network (DRONE). Magn Reson Med 2018;80:885–94.

[165] Hoppe E, Krzdrfer G, Wrfl T, Wetzl J, Lugauer F, Pfeuffer J, et al. Deep learning for magnetic resonance fingerprinting: a new approach for predicting quantitative parameter values from time series. Stud Health Technol Inform 2017;243:202–6.

[166] Bojorquez JZ, Bricq S, Acquitter C, Brunotte F, Walker PM, Lalande A. What are normal relaxation times of tissues at 3T? Magn Reson Imaging 2017;35:69–80.

[167] Fang Z, Chen Y, Lin W, Shen D. Quantification of relaxation times in MR fingerprinting using deep learning. In: Proceedings of the International Society for Magnetic Resonance in Medicine. Scientific meeting and exhibition 25. 2017.

[168] Virtue P, Yu SX, Lustig M. Better than real: complex-valued neural nets for MRI fingerprinting. In: Proc. IEEE int. conf. image processing (ICIP). 2017. p. 3953–7.

[169] Tygert M, Bruna J, Chintala S, LeCun Y, Piantino S, Szlam A. A mathematical motivation for complex-valued convolutional networks. Neural Comput 2016;28:815–25.

[170] Trabelsi C, Bilaniuk O, Zhang Y, Serdyuk D, Subramanian S, Santos JF, et al. Deep complex networks, arXiv:1705.09792 (2017).

[171] Sijbers J, den Dekker AJ, Van Audekerke J, Verhoye M, Van Dyck D. Estimation of the noise in magnitude MR images. Magn Reson Imaging 1998;16:87–90.

[172] McVeigh ER, Henkelman RM, Bronskill MJ. Noise and filtration in magnetic resonance imaging. Med Phys 1985;12:586–91.

[173] Baselice F, Ferraioli G, Pascazio V, Sorriso A. Bayesian MRI denoising in complex domain. Magn Reson Imaging 2017;38:112–22.

[174] Phophalia A, Mitra SK. 3d MR image denoising using rough set and kernel PCA method. Magn Reson Imaging 2017;36:135–45.

[175] Zhang X, Xu Z, Jia N, Yang W, Feng Q, Chen W, et al. Denoising of 3D magnetic resonance images by using higher-order singular value decomposition. Med Image Anal 2015;19:75–86.

[176] Van De Ville D, Seghier ML, Lazeyras F, Blu T, Unser M. Wspm: wavelet-based statistical parametric mapping. Neuroimage 2007;37:1205–17.

[177] Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 2014;90:449–68.

[178] Lysaker M, Lundervold A, Tai X-C. Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. IEEE Trans Image Process 2003;12:1579–90.

[179] Bermudez C, Plassard AJ, Davis TL, Newton AT, Resnick SM, Landman BA. Learning implicit brain MRI manifolds with deep learning. Proc SPIE 2018;10574.

[180] Benou A, Veksler R, Friedman A, Riklin Raviv T. Ensemble of expert deep neural networks for spatiotemporal denoising of contrast-enhanced MRI sequences. Med Image Anal 2017;42:145–59.

[181] Gal Y, Mehnert AJH, Bradley AP, McMahon K, Kennedy D, Crozier S. Denoising of dynamic contrast-enhanced MR images using dynamic non-local means. IEEE Trans Med Imaging 2010;29:302–10.

[182] Vincent P, Larochelle H, Lajoie I, Bengio Y, Manzagol P-A. Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res (JMLR) 2010;11:3371–408.

[183] Dikaios N, Arridge S, Hamy V, Punwani S, Atkinson D. Direct parametric reconstruction from undersampled (k,t)-space data in dynamic contrast enhanced MRI. Med Image Anal 2014;18:989–1001.

[184] Guo Y, Lingala SG, Zhu Y, Lebel RM, Nayak KS. Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI. Magn Reson Med 2017;78:1566–78.

[185] Sourbron SP, Buckley DL. Classic models for dynamic contrast-enhanced MRI. NMR Biomed 2013;26:1004–27.

[186] Golkov V, Dosovitskiy A, Sperl JI, Menzel MI, Czisch M, Samann P, et al. q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans Med Imaging 2016;35:1344–51.

[187] Gurbani SS, Schreibmann E, Maudsley AA, Cordova JS, Soher BJ, Poptani H, et al. A convolutional neural network to filter artifacts in spectroscopic MRI. Magn Reson Med 2018;80:1765–75.

[188] Kyathanahally SP, Dring A, Kreis R. Deep learning approaches for detection and removal of ghosting artifacts in MR spectroscopy. Magn Reson Med 2018;80:851–63.

[189] Küstner T, Liebgott A, Mauch L, Martirosian P, Bamberg F, Nikolaou K, et al. Automated reference-free detection of motion artifacts in magnetic resonance images. MAGMA 2018;31:243–56.

[190] Yue L, Shen H, Li J, Yuan Q, Zhang H, Zhang L. Image super-resolution: the techniques, applications, and future. Signal Process 2016;128:389–408.

[191] Shilling RZ, Robbie TQ, Bailloeul T, Mewes K, Mersereau RM, Brummer ME. A superresolution framework for 3-D high-resolution and high-contrast imaging using 2-D multislice MRI. IEEE Trans Med Imaging 2009;28:633–44.

[192] Ropele S, Ebner F, Fazekas F, Reishofer G. Super-resolution MRI using microscopic spatial modulation of magnetization. Magn Reson Med 2010;64:1671–5.

[193] Plenge E, Poot DHJ, Bernsen M, Kotek G, Houston G, Wielopolski P, et al. Super-resolution methods in MRI: can they improve the trade-off between resolution, signal-to-noise ratio, and acquisition time? Magn Reson Med 2012;68:1983–93.

[194] Bahrami K, Shi F, Rekik I, Gao Y, Shen D. 7T-guided super-resolution of 3T MRI. Med Phys 2017;44:1661–77.

[195] Van Steenkiste G, Poot DHJ, Jeurissen B, den Dekker AJ, Vanhevel F, Parizel PM, et al. Super-resolution T_1 estimation: quantitative high resolution T_1 mapping from a set of low resolution T_1-weighted images with different slice orientations. Magn Reson Med 2017;77:1818–30.

[196] Zeng K, Zheng H, Cai C, Yang Y, Zhang K, Chen Z. Simultaneous single-and multi-contrast super-resolution for brain MRI images based on a convolutional neural network. Comput Biol Med 2018;99:133–41.

[197] Liu C, Wu X, Yu X, Tang Y, Zhang J, Zhou J. Fusing multi-scale information in convolution network for MR image super-resolution reconstruction. Biomed Eng Online 2018;17:114.

[198] Chaudhari AS, Fang Z, Kogan F, Wood J, Stevens KJ, Gibbons EK, et al. Super-resolution musculoskeletal MRI using deep learning. Magn Reson Med 2018;80:2139–54.

[199] Jog A, Carass A, Roy S, Pham DL, Prince JL. Random forest regression for magnetic resonance image synthesis. Med Image Anal 2017;35:475–88.

[200] Keenan KE, Ainslie M, Barker AJ, Boss MA, Cecil KM, Charles C, et al. Quantitative magnetic resonance imaging phantoms: a review and the need for a system phantom. Magn Reson Med 2018;79:48–61.

[201] Jurczuk K, Kretowski M, Eliat P-A, Saint-Jalmes H, Bezy-Wendling J. In silico modeling of magnetic resonance flow imaging in complex vascular networks. IEEE Trans Med Imaging 2014;33:2191–209.

[202] Zhou Y, Giffard-Roisin S, De Craene M, Camarasu-Pop S, D'Hooge J, Alessandrini M, et al. A framework for the generation of realistic synthetic cardiac ultrasound and magnetic resonance imaging sequences from the same virtual patients. IEEE Trans Med Imaging 2018;37:741–54.

[203] Duchateau N, Sermesant M, Delingette H, Ayache N. Model-based generation of large databases of cardiac images: synthesis of pathological cine MR sequences from real healthy cases. IEEE Trans Med Imaging 2018;37:755–66.

[204] Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B, Bharath AA. Generative adversarial networks: an overview. IEEE Signal Process Mag 2018;35:53–65.

[205] Hong Y, Hwang U, Yoo J, Yoon S. How generative adversarial networks and their variants work: an overview of GAN, arXiv:1711.05914v7 (2017).

[206] Huang H, Yu PS, Wang C. An introduction to image synthesis with generative adversarial nets, arXiv:1803.04469v1 (2018).

[207] Osokin A, Chessel A, Salas REC, Vaggi F. GANs for biological image synthesis. In: Proc. IEEE int. conf. computer vision (ICCV). 2017. p. 2252–61.

[208] Antipov G, Baccouche M, Dugelay J. Face aging with conditional generative adversarial networks. In: Proc. IEEE int. conf. image processing (ICIP). 2017. p. 2089–93.

[209] Bodnar C. Text to image synthesis using generative adversarial networks, arXiv:1805.00676v1 (2018).

[210] Dong H, Yu S, Wu C, Guo Y. Semantic image synthesis via adversarial learning, arXiv:1707.06873v1 (2017).

[211] Reed S, Akata Z, Yan X, Logeswaran L, Schiele B, Lee H. Generative adversarial text to image synthesis, arXiv:1605.05396v2 (2016).

[212] Shin H-C, Tenenholtz NA, Rogers JK, Schwarz CG, Senjem ML, Gunter JL, et al. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In: International workshop on simulation and synthesis in medical imaging. 2018. p. 1–11.

[213] Mok TCW, Chung ACS. Learning data augmentation for brain tumor segmentation with coarse-to-fine generative adversarial networks, arXiv:1805.11291 (2018).

[214] Guibas JT, Virdi TS, Li PS. Synthetic medical images from dual generative adversarial networks, arXiv:1709.01872 (2017).

[215] Kitchen A, Seah J. Deep generative adversarial neural networks for realistic prostate lesion MRI synthesis, arXiv:1708.00129 (2017).

[216] Nie D, Trullo R, Lian J, Petitjean C, Ruan S, Wang Q, et al. Medical image synthesis with context-aware generative adversarial networks, Medical image computing and computer-assisted intervention: MICCAI. In: International conference on medical image computing and computer-assisted intervention 10435. 2017. p. 417–25.

[217] Spuhler KD, Gardus J, Gao Y, DeLorenzo C, Parsey C, Huang. Synthesis of patient-specific transmission image for PET attenuation correction for PET/MR imaging of the brain using a convolutional neural network. J Nucl Med 2018.

[218] Torrado-Carvajal A, Vera-Olmos J, Izquierdo-Garcia D, Catalano OA, Morales MA, Margolin J, et al. Dixon-VIBE deep learning (DIVIDE) pseudo-CT synthesis for pelvis PET/MR attenuation correction. J Nucl Med 2018.

[219] Zhang Q, Wang H, Lu H, Won D, Yoon SW. Medical image synthesis with generative adversarial networks for tissue recognition. In: Proc. IEEE int. conf. healthcare informatics (ICHI). 2018. p. 199–207.

[220] Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H. Synthetic data augmentation using GAN for improved liver lesion classification. In: Proc. IEEE 15th int. symp. biomedical imaging (ISBI 2018). 2018. p. 289–93.

[221] Wolterink JM, Dinkla AM, Savenije MHF, Seevinck PR, van den Berg CAT, Isgum I. Deep MR to CT synthesis using unpaired data, arXiv:1708.01155v1 (2017).

[222] Fitzpatrick JM, Maurer Jr CR. A review of medical image registration; 1993.

[223] Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: a review. Magn Reson Med 2013;69:621–36.

[224] Zaitsev M, Akin B, LeVan P, Knowles BR. Prospective motion correction in functional MRI. Neuroimage 2017;154:33–42.

[225] Fluck O, Vetter C, Wein W, Kamen A, Preim B, Westermann R. A survey of medical image registration on graphics hardware. Comput Methods Programs Biomed 2011;104:e45–57.

[226] Shi L, Liu W, Zhang H, Xie Y, Wang D. A survey of GPU-based medical image computing techniques. Quant Imaging Med Surg 2012;2:188–206.

[227] Eklund A, Dufort P, Forsberg D, LaConte SM. Medical image processing on the GPU – past, present and future. Med Image Anal 2013;17:1073–94.

[228] Maintz JB, Viergever MA. A survey of medical image registration. Med Image Anal 1998;2:1–36.

[229] Glocker B, Sotiras A, Komodakis N, Paragios N. Deformable medical image registration: setting the state of the art with discrete methods. Annu Rev Biomed Eng 2011;13:219–44.

[230] Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: a survey. IEEE Trans Med Imaging 2013;32:1153–90.

[231] Oliveira FPM, Tavares JMRS. Medical image registration: a review. Comput Methods Biomech Biomed Eng 2014;17:73–93.

[232] Saha PK, Strand R, Borgefors G. Digital topology and geometry in medical imaging: a survey. IEEE Trans Med Imaging 2015;34:1940–64.

[233] Viergever MA, Maintz JBA, Klein S, Murphy K, Staring M, Pluim JPW. A survey of medical image registration – under review. Med Image Anal 2016;33:140–4.

[234] Song G, Han J, Zhao Y, Wang Z, Du H. A review on medical image registration as an optimization problem. Curr Med Imaging Rev 2017;13:274–83.

[235] Ferrante E, Paragios N. Slice-to-volume medical image registration: a survey. Med Image Anal 2017;39:101–23.

[236] Keszei AP, Berkels B, Deserno TM. Survey of non-rigid registration tools in medicine. J Digit Imaging 2017;30:102–16.

[237] Nag S. Image registration techniques: a survey, arXiv:1712.07540v1 (2017).

[238] Jiang J, Trundle P, Ren J. Medical image analysis with artificial neural networks. Comput Med Imaging Graph 2010;34:617–31.

[239] Wu G, Kim M, Wang Q, Munsell BC, Shen D. Scalable high-performance image registration framework by unsupervised deep feature representations learning. IEEE Trans Biomed Eng 2016;63:1505–16.

[240] Salehi SSM, Khan S, Erdogmus D, Gholipour A. Real-time deep pose estimation with geodesic loss for image-to-template rigid registration. IEEE Trans Med Imaging 2018.

[241] Toth D, Miao S, Kurzendorfer T, Rinaldi CA, Liao R, Mansi T, et al. 3D/2D model-to-image registration by imitation learning for cardiac procedures. Int J Comput Assist Radiol Surg 2018.

[242] Han X. MR-based synthetic CT generation using a deep convolutional neural network method. Med Phys 2017;44:1408–19.

[243] Liu M, Cheng D, Wang K, Wang Y, Initiative ADN. Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis. Neuroinformatics 2018;16:295–308.

[244] Xiang L, Qiao Y, Nie D, An L, Wang Q, Shen D. Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI. Neurocomputing 2017;267:406–16.

[245] Shan S, Yan W, Guo X, Chang EI-C, Fan Y, Xu Y. Unsupervised end-to-end learning for deformable medical image registration, arXiv:1711.08608v2 (2017).

[246] Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. An unsupervised learning model for deformable medical image registration, arXiv:1802.02604v3 (2018).

[247] de Vos BD, Berendsen FF, Viergever MA, Sokooti H, Staring M, Isgum I. A deep learning framework for unsupervised affine and deformable image registration, arXiv:1809.06130v1 (2018).

[248] Vannier MW, Butterfield RL, Jordan D, Murphy WA, Levitt RG, Gado M. Multispectral analysis of magnetic resonance images. Radiology 1985;154:221–4.

[249] Lundervold A, Moen K, Taxt T. Automatic recognition of normal and pathological tissue types in MR images. In: Proc. of the NOBIM conference. 1988.

[250] Taxt T, Lundervold A, Fuglaas B, Lien H, Abeler V. Multispectral analysis of uterine corpus tumors in magnetic resonance imaging. Magn Reson Med 1992;23:55–76.

[251] Taxt T, Lundervold A. Multispectral analysis of the brain using magnetic resonance imaging. IEEE Trans Med Imaging 1994;13:470–81.

[252] Lundervold A, Storvik G. Segmentation of brain parenchyma and cerebrospinal fluid in multispectral magnetic resonance images. IEEE Trans Med Imaging 1995;14:339–49.

[253] Cabezas M, Oliver A, Llad X, Freixenet J, Cuadra MB. A review of atlas-based segmentation for magnetic resonance brain images. Comput Methods Programs Biomed 2011;104:e158–77.

[254] Garca-Lorenzo D, Francis S, Narayanan S, Arnold DL, Collins DL. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. Med Image Anal 2013;17:1–18.

[255] Smistad E, Falch TL, Bozorgi M, Elster AC, Lindseth F. Medical image segmentation on GPUs – a comprehensive review. Med Image Anal 2015;20:1–18.

[256] Bernal J, Kushibar K, Asfaw DS, Valverde S, Oliver A, Mart R, et al. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review, arXiv:1712.03747v3 (2017).

[257] Dora L, Agrawal S, Panda R, Abraham A. State-of-the-art methods for brain tissue segmentation: a review. IEEE Rev Biomed Eng 2017;10:235–49.

[258] Torres HR, Queiros S, Morais P, Oliveira B, Fonseca JC, Vilaa JL. Kidney segmentation in ultrasound, magnetic resonance and computed tomography images: a systematic review. Comput Methods Programs Biomed 2018;157:49–67.

[259] Bernal J, Kushibar K, Asfaw DS, Valverde S, Oliver A, Mart R, et al. Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review. Artif Intell Med 2018.

[260] Moccia S, De Momi E, El Hadji S, Mattos LS. Blood vessel segmentation algorithms-review of methods, datasets and evaluation metrics. Comput Methods Programs Biomed 2018;158:71–91.

[261] Makropoulos A, Counsell SJ, Rueckert D. A review on automatic fetal and neonatal brain MRI segmentation. Neuroimage 2018;170:231–48.

[262] Chen L, Bentley P, Rueckert D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. NeuroImage Clin 2017;15:633–43.

[263] Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, et al. Brain tumor segmentation with deep neural networks. Med Image Anal 2017;35:18–31.

[264] Choi H, Jin KH. Fast and robust segmentation of the striatum using deep convolutional neural networks. J Neurosci Methods 2016;274:146–53.

[265] Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Med Phys 2017;44:547–57.

[266] Kline TL, Korfiatis P, Edwards ME, Blais JD, Czerwiec FS, Harris PC, et al. Performance of an artificial multi-observer deep neural network for fully automated segmentation of polycystic kidneys. J Digit Imaging 2017;30:442–8.

[267] Guo Y, Gao Y, Shen D. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. IEEE Trans Med Imaging 2016;35:1077–89.

[268] Li X, Dou Q, Chen H, Fu C-W, Qi X, Belav DL, et al. 3D multi-scale FCN with random modality voxel dropout learning for intervertebral disc localization and segmentation from multi-modality MR images. Med Image Anal 2018;45:41–54.

[269] Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M, et al. Deep MRI brain extraction: a 3D convolutional neural network for skull stripping. Neuroimage 2016;129:460–9.

[270] Li H, Parikh NA, He L. A novel transfer learning approach to enhance deep neural network classification of brain functional connectomes. Front Neurosci 2018;12:491.

[271] Zeng L-L, Wang H, Hu P, Yang B, Pu W, Shen H, et al. Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI. EBioMedicine 2018;30:74–85.

[272] Wasserthal J, Neher P, Maier-Hein KH. Tract-Seg-fast and accurate white matter tract segmentation. Neuroimage 2018;183:239–53.

[273] Cole JH, Poudel RPK, Tsagkrasoulis D, Caan MWA, Steves C, Spector TD, et al. Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. Neuroimage 2017;163:115–24.

[274] Liu M, Zhang J, Adeli E, Shen D. Landmark-based deep multi-instance learning for brain disease diagnosis. Med Image Anal 2018;43:157–68.

[275] Islam J, Zhang Y. Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. Brain Inform 2018;5:2.

[276] Lu D, Popuri K, Ding GW, Balachandar R, Beg MF. Multimodal and multiscale deep neural networks for the early diagnosis of Alzheimer's disease using structural MR and FDG-PET images. Sci Rep 2018;8:5697.

[277] Moeskops P, de Bresser J, Kuijf HJ, Mendrik AM, Biessels GJ, Pluim JPW, et al. Evaluation of a deep learning approach for the segmentation of brain tissues and white matter hyperintensities of presumed vascular origin in MRI. NeuroImage Clin 2018;17:251–62.

[278] Pizarro R, Assemlal H-E, De Nigris D, Elliott C, Antel S, Arnold D, et al. Using deep learning algorithms to automatically identify

the brain MRI contrast: implications for managing large databases. Neuroinformatics 2018.

[279] Laukamp KR, Thiele F, Shakirin G, Zopfs D, Faymonville A, Timmer M, et al. Fully automated detection and segmentation of meningiomas using deep learning on routine multiparametric MRI. Eur Radiol 2018.

[280] Perkuhn M, Stavrinou P, Thiele F, Shakirin G, Mohan M, Garmpis D, et al. Clinical evaluation of a multiparametric deep learning model for glioblastoma segmentation using heterogeneous magnetic resonance imaging data from clinical routine. Invest Radiol 2018.

[281] AlBadawy EA, Saha A, Mazurowski MA. Deep learning for segmentation of brain tumors: impact of cross-institutional training and testing. Med Phys 2018;45:1150–8.

[282] Cui S, Mao L, Jiang J, Liu C, Xiong S. Automatic semantic segmentation of brain gliomas from MRI images using a deep cascaded neural network. J Healthc Eng 2018;2018:4940593.

[283] Hoseini F, Shahbahrami A, Bayat P. Adaptahead optimization algorithm for learning deep CNN applied to MRI segmentation. J Digit Imaging 2018.

[284] Yoo Y, Tang LYW, Brosch T, Li DKB, Kolind S, Vavasour I, et al. Deep learning of joint myelin and T1w MRI features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. NeuroImage Clin 2018;17:169–78.

[285] Bobo MF, Bao S, Huo Y, Yao Y, Virostko J, Plassard AJ, et al. Fully convolutional neural networks improve abdominal organ segmentation. Proc SPIE 2018;10574.

[286] Shehata M, Khalifa F, Soliman A, Ghazal M, Taher M, Abou El-Ghar A, et al. Computer-aided diagnostic system for early detection of acute renal transplant rejection using diffusion-weighted MRI. IEEE Trans Bio-med Eng 2018.

[287] Cheng R, Roth HR, Lay N, Lu L, Turkbey B, Gandler W, et al. Automatic magnetic resonance prostate segmentation by deep learning with holistically nested networks. J Med Imaging 2017;4:041302.

[288] Ishioka J, Matsuoka Y, Uehara S, Yasuda Y, Kijima T, Yoshida S, et al. Computer-aided diagnosis of prostate cancer on magnetic resonance imaging using a convolutional neural network algorithm. BJU Int 2018.

[289] Song Y, Zhang Y-D, Yan X, Liu H, Zhou M, Hu B, et al. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. J Magn Reson Imaging: JMRI 2018.

[290] Wang X, Yang W, Weinreb J, Han J, Li Q, Kong X, et al. Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning. Sci Rep 2017;7:15415.

[291] Yang X, Liu C, Wang Z, Yang J, Min HL, Wang L, et al. Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI. Med Image Anal 2017;42:212–27.

[292] Le MH, Chen J, Wang L, Wang Z, Liu W, Cheng K-TT, et al. Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks. Phys Med Biol 2017;62:6497–514.

[293] Forsberg D, Sjöblom E, Sunshine JL. Detection and labeling of vertebrae in MR images using deep learning with clinical annotations as training data. J Digit Imaging 2017;30:406–12.

[294] Lu J-T, Pedemonte S, Bizzo B, Doyle S, Andriole KP, Michalski MH, et al. DeepSPINE: automated lumbar vertebral segmentation, disc-level designation, and spinal stenosis grading using deep learning, arXiv:1807.10215v1 (2018).

[295] Han Z, Wei B, Leung S, Nachum IB, Laidley D, Li S. Automated pathogenesis-based diagnosis of lumbar neural foraminal stenosis via deep multiscale multitask learning. Neuroinformatics 2018;16:325–37.

[296] Kim KH, Do W-J, Park S-H. Improving resolution of MR images with an adversarial network incorporating images with different contrast. Med Phys 2018;45:3120–31.

[297] Pilevar AH. CBMIR: content-based image retrieval algorithm for medical image databases. J Med Signals Sens 2011;1:12–8.

[298] Kumar A, Kim J, Cai W, Fulham M, Feng D. Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. J Digit Imaging 2013;26:1025–39.

[299] Faria AV, Oishi K, Yoshida S, Hillis A, Miller I, Mori S. Content-based image retrieval for brain MRI: an image-searching engine and population-based analysis to utilize past clinical data for future diagnosis. NeuroImage Clin 2015;7:367–76.

[300] Kumar A, Nette F, Klein K, Fulham M, Kim J. A visual analytics approach using the exploration of multidimensional feature spaces for content-based medical image retrieval. IEEE J Biomed Health Inform 2015;19:1734–46.

[301] Bedo MVN, Pereira Dos Santos D, Ponciano-Silva M, de Azevedo-Marques PM, Ferreira de Carvalho APdL, Traina C. Endowing a content-based medical image retrieval system with perceptual similarity using ensemble strategy. J Digit Imaging 2016; 29:22–37.

[302] Muramatsu C. Overview on subjective similarity of images for content-based medical image retrieval. Radiol Phys Technol 2018.

[303] Spanier AB, Caplan N, Sosna J, Acar B, Joskowicz. A fully automatic end-to-end method for content-based image retrieval of CT scans with similar liver lesion annotations. Int J Comput Assist Radiol Surg 2018;13:165–74.

[304] Gordo A, Almazan J, Revaud J, Larlus D. End-to-end learning of deep visual representations for image retrieval; 2017.

[305] Liu P, Guo J, Wu C, Cai D. Fusion of deep learning and compressed domain features for content-based image retrieval. IEEE Trans Image Process 2017;26:5706–17.

[306] Han J, Zhang D, Cheng G, Liu N, Xu D. Advanced deep-learning techniques for salient and category-specific object detection: a survey. IEEE Signal Process Mag 2018;35:84–100.

[307] Piplani T, Bamman D. Deepseek: content based image search & retrieval, arXiv:1801.03406v2 (2018).

[308] Yang J, Liang J, Shen H, Wang K, Rosin PL, Yang. Dynamic match kernel with deep convolutional features for image retrieval. IEEE Trans Image Process 2018;27:5288–302.

[309] Sklan JES, Plassard AJ, Fabbri D, Landman BA. Toward content based image retrieval with deep convolutional neural networks. Proc SPIE 2015;9417.

[310] Bressan RS, Alves DHA, Valerio LM, Bugatti PH, Saito PTM. DOC-ToR: the role of deep features in content-based mammographic image retrieval. In: Proc. IEEE 31st int. symp. computer-based medical systems (CBMS). 2018. p. 158–63.

[311] Qayyum A, Anwar SM, Awais M, Majid M. Medical image retrieval using deep convolutional neural network, arXiv:1703.08472v1 (2017).

[312] Chung Y-A, Weng W-H. Learning deep representations of medical images using Siamese CNNs with application to content-based image retrieval, arXiv:1711.08490v2 (2017).

[313] Jing B, Xie P, Xing E. On the automatic generation of medical imaging reports, arXiv:1711.08195v3 (2017).

[314] Li CY, Liang X, Hu Z, Xing EP. Hybrid retrieval-generation reinforced agent for medical image report generation, arXiv:1805.08298v1 (2018).

[315] Moradi M, Madani A, Gur Y, Guo Y, Syeda-Mahmood T. Bimodal network architectures for automatic generation of image annotation from text, arXiv:1809.01610v1 (2018).

[316] Zhang Y, Ding DY, Qian T, Manning CD, Langlotz CP. Learning to summarize radiology findings, arXiv:1809.04698v1.

[317] Pons E, Braun LMM, Hunink MGM, Kors JA. Natural language processing in radiology: a systematic review. Radiology 2016;279:329–43.

[318] Zech J, Pain M, Titano J, Badgeley M, Schefflein J, Su A, et al. Natural language-based machine learning models for the annotation of clinical radiology reports. Radiology 2018;287:570–80.

[319] Goff DJ, Loehfelm TW. Automated radiology report summarization using an open-source natural language processing pipeline. J Digit Imaging 2018;31:185–92.

[320] Gibson E, Li W, Sudre C, Fidon L, Shakir DI, Wang G, et al. NiftyNet: a deep-learning platform for medical imaging. Comput Methods Programs Biomed 2018;158:113–22.

[321] Li W, Wang G, Fidon L, Ourselin S, Cardoso MJ, Vercauteren T. On the compactness, efficiency, and representation of 3d convolutional networks: brain parcellation as a pretext task. In: International conference on information processing in medical imaging (IPMI). 2017.

[322] Kamnitsas K, Ledig C, Newcombe VF, Simpson JP, Kane AD, Menon DK, et al. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal 2017; 36:61–78.

[323] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention (MICCAI). LNCS, vol. 9351. Springer; 2015. p. 234–41, arXiv:1505.04597 [cs.CV].

[324] Badrinarayanan V, Kendall A, Cipolla R. Seg-Net: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017.

[325] Mardani M, Gong E, Cheng JY, Vasanawala S, Zaharchuk G Alley M, et al. Deep generative adversarial networks for compressed sensing automates MRI. arXiv:1706.00051 (2017).

[326] Parisot S, Ktena SI, Ferrante E, Lee M, Moreno RG, Glocker B, et al. Spectral graph convolutions for population-based disease prediction. In: International conference on medical image computing and computer-assisted intervention. 2017. p. 177–85.

[327] Marcus G. Deep learning: a critical appraisal, arXiv:1801.00631 (2018).

[328] Lipton ZC, Steinhardt J. Troubling trends in machine learning scholarship; 2018.

[329] Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning requires rethinking generalization, arXiv:1611.03530 (2016).

[330] Fredrikson M, Jha S, Ristenpart T. Model inversion attacks that exploit confidence information and basic countermeasures. In: Proceedings of the 22nd ACM SIGSAC conference on computer and communications security, ACM. 2015. p. 1322–33.

[331] Shokri R, Stronati M, Song C, Shmatikov V. Membership inference attacks against machine learning models. In: 2017 IEEE symposium on security and privacy (SP), IEEE. 2017. p. 3–18.

[332] McMahan B, Moore E, Ramage D, Hampson S, Arcas BA. Communication-efficient learning of deep networks from decentralized data. In: Singh A, Zhu J, editors. Proceedings of the 20th international conference on artificial intelligence and statistics, Proceedings of machine learning research, vol. 54. Fort Lauderdale, FL, USA: PMLR; 2017. p. 1273–82.

[333] Papernot N, Abadi M, Erlingsson U, Goodfellow I, Talwar K. Semi-supervised knowledge transfer for deep learning from private training data, arXiv:1610.05755 (2016).

[334] Papernot N, Song S, Mironov I, Raghunathan A, Talwar K, Erlingsson Ú. Scalable private learning with PATE, arXiv:1802.08908 (2018).

[335] McMahan HB, Ramage D, Talwar K, Zhang L. Learning differentially private recurrent language models. In: International conference on learning representations. 2017.

[336] Chang K, Balachandar N, Lam C, Yi D, Brown J, Beers A, et al. Distributed deep learning networks among institutions for medical imaging. J Am Med Inform Assoc: JAMIA 2018;25:945–54.

[337] Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: a cross-sectional study. PLoS Med 2018;15:e1002683.

[338] Lundervold A, Lundervold A, Rørvik J. Fast semi-supervised segmentation of the kidneys in DCE-MRI using convolutional neural networks and transfer learning; 2017.

[339] Lundervold A, Sprawka K, Lundervold A. Fast estimation of kidney volumes and time courses in DCE-MRI using convolutional neural networks; 2018.

[340] Hinton GE, Krizhevsky A, Wang SD. Transforming auto-encoders. In: International conference on artificial neural networks. 2011. p. 44–51.

[341] Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In: Advances in neural information processing systems. 2017. p. 3856–66.

[342] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention. In: Advances in neural information processing systems. 2014. p. 2204–12.

[343] Xu K, Ba J, Kiros R, Cho K, Courville A, Salakhudinov R, et al. Show, attend and tell: neural image caption generation with visual attention. In: International conference on machine learning. 2015. p. 2048–57.

[344] Castelvecchi D. Can we open the black box of AI? Nat News 2016;538:20.

[345] Olah C, Satyanarayan A, Johnson I, Carter S, Schubert L, Ye K, et al. The building blocks of interpretability. Distill 2018;3:e10.

[346] Montavon G, Samek W, Müller K-R. Methods for interpreting and understanding deep neural networks. Digit Signal Process 2017.

[347] Yosinski J, Clune J, Nguyen A, Fuchs T, Lipson H. Understanding neural networks through deep visualization. In: Deep learning workshop, 31st international conference on machine learning. 2015.

[348] Olah C, Mordvintsev A, Schubert L. Feature visualization. Distill 2017;2:e7.

[349] Hohman FM, Kahng M, Pienta R, Chau DH. Visual analytics in deep learning: an interrogative survey for the next frontiers. IEEE Trans Vis Comput Graph 2018.

[350] Bach S, Binder A, Montavon G, Klauschen F, Müller K-R, Samek W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. PLOS ONE 2015;10.

[351] Neal RM. Bayesian learning for neural networks [Ph.D. thesis]. University of Toronto; 1995.

[352] MacKay DJ. A practical Bayesian framework for backpropagation networks. Neural Comput 1992;4:448–72.

[353] Dayan P, Hinton GE, Neal RM, Zemel RS. The Helmholtz machine. Neural Comput 1995;7:889–904.

[354] Li Y, Gal Y. Dropout inference in Bayesian neural networks with alpha-divergences. In: International conference on machine learning. 2017. p. 2052–61.

[355] Leibig C, Allken V, Ayhan MS, Berens P, Wahl S. Leveraging uncertainty information from deep neural networks for disease detection. Sci Rep 2017;7:17816.

[356] Kendall A, Badrinarayanan V, Cipolla R. Bayesian segnet: model uncertainty in deep convolutional encoder-decoder architectures for scene understanding, arXiv:1511.02680 (2015).

[357] Feinman R, Curtin RR, Shintre S, Gardner AB. Detecting adversarial samples from artifacts, arXiv:1703.00410 (2017).

[358] Sharp P, Hockfield S. Convergence: the future of health. Science 2017;355:589.

[359] Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. New Biotechnol 2012;29:613–24.