A Survey on Deep Learning in Medical Image Analysis

Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A.W.M. van der Laak, Bram van Ginneken, Clara I. Sánchez

Diagnostic Image Analysis Group
Radboud University Medical Center
Nijmegen, The Netherlands

arXiv:1702.05747v2 [cs.CV] 4 Jun 2017

Abstract

Deep learning algorithms, in particular convolutional networks, have rapidly become a methodology of choice for analyzing medical images. This paper reviews the major deep learning concepts pertinent to medical image analysis and summarizes over 300 contributions to the field, most of which appeared in the last year. We survey the use of deep learning for image classification, object detection, segmentation, registration, and other tasks. Concise overviews are provided of studies per application area: neuro, retinal, pulmonary, digital pathology, breast, cardiac, abdominal, musculoskeletal. We end with a summary of the current state-of-the-art, a critical discussion of open challenges and directions for future research.

Keywords: deep learning, convolutional neural networks, medical imaging, survey

1. Introduction

As soon as it was possible to scan and load medical images into a computer, researchers have built systems for automated analysis. Initially, from the 1970s to the 1990s, medical image analysis was done with sequential application of low-level pixel processing (edge and line detector filters, region growing) and mathematical modeling (fitting lines, circles and ellipses) to construct compound rule-based systems that solved particular tasks. There is an analogy with expert systems with many if-then-else statements that were popular in artificial intelligence in the same period. These expert systems have been described as GOFAI (good old-fashioned artificial intelligence) (Haugeland, 1985) and were often brittle; similar to rule-based image processing systems.

At the end of the 1990s, supervised techniques, where training data is used to develop a system, were becoming increasingly popular in medical image analysis. Examples include active shape models (for segmentation), atlas methods (where the atlases that are fit to new data form the training data), and the concept of feature extraction and use of statistical classifiers (for computer-aided detection and diagnosis). This pattern recognition or machine learning approach is still very popular and forms the basis of many successful commercially available medical image analysis systems. Thus, we have seen a shift from systems that are completely designed by humans to systems that are trained by computers using example data from which feature vectors are extracted. Computer algorithms determine the optimal decision boundary in the high-dimensional feature space. A crucial step in the design of such systems is the extraction of discriminant features from the images. This process is still done by human researchers and, as such, one speaks of systems with *handcrafted* features.

A logical next step is to let computers learn the features that optimally represent the data for the problem at hand. This concept lies at the basis of many deep learning algorithms: models (networks) composed of many layers that transform input data (e.g. images) to outputs (e.g. disease present/absent) while learning increasingly higher level features. The most successful type of models for image analysis to date are convolutional neural networks (CNNs). CNNs contain many layers that transform their input with convolution filters of a small extent. Work on CNNs has been done since the late seventies (Fukushima, 1980) and they were already applied to medical image analysis in 1995 by Lo et al. (1995). They saw their first successful real-world application in LeNet (LeCun et al., 1998) for hand-written digit recog-

nition. Despite these initial successes, the use of CNNs did not gather momentum until various new techniques were developed for efficiently training deep networks, and advances were made in core computing systems. The watershed was the contribution of Krizhevsky et al. (2012) to the ImageNet challenge in December 2012. The proposed CNN, called AlexNet, won that competition by a large margin. In subsequent years, further progress has been made using related but deeper architectures (Russakovsky et al., 2014). In computer vision, deep convolutional networks have now become the technique of choice.

The medical image analysis community has taken notice of these pivotal developments. However, the transition from systems that use handcrafted features to systems that learn features from the data has been gradual. Before the breakthrough of AlexNet, many different techniques to learn features were popular. Bengio et al. (2013) provide a thorough review of these techniques. They include principal component analysis, clustering of image patches, dictionary approaches, and many more. Bengio et al. (2013) introduce CNNs that are trained end-to-end only at the end of their review in a section entitled *Global training of deep models*. In this survey, we focus particularly on such deep models, and do not include the more traditional feature learning approaches that have been applied to medical images. For a broader review on the application of deep learning in health informatics we refer to Ravi et al. (2017), where medical image analysis is briefly touched upon.

Applications of deep learning to medical image analysis first started to appear at workshops and conferences, and then in journals. The number of papers grew rapidly in 2015 and 2016. This is illustrated in Figure 1. The topic is now dominant at major conferences and a first special issue appeared of IEEE Transaction on Medical Imaging in May 2016 (Greenspan et al., 2016).

One dedicated review on application of deep learning to medical image analysis was published by Shen et al. (2017). Although they cover a substantial amount of work, we feel that important areas of the field were not represented. To give an example, no work on retinal image analysis was covered. The motivation for our review was to offer a comprehensive overview of (almost) all fields in medical imaging, both from an application and a methodology-drive perspective. This also includes overview tables of all publications which readers can use to quickly assess the field. Last, we leveraged our own experience with the application of deep learning methods to medical image analysis to provide readers with a dedicated discussion section covering the state-of-the-art, open challenges and overview of research directions and technologies that will become important in the future.

This survey includes over 300 papers, most of them recent, on a wide variety of applications of deep learning in medical image analysis. To identify relevant contributions PubMed was queried for papers containing (”convolutional” OR ”deep learning”) in title or abstract. ArXiv was searched for papers mentioning one of a set of terms related to medical imaging. Additionally, conference proceedings for MICCAI (including workshops), SPIE, ISBI and EMBC were searched based on titles of papers. We checked references in all selected papers and consulted colleagues. We excluded papers that did not report results on medical image data or only used standard feed-forward neural networks with handcrafted features. When overlapping work had been reported in multiple publications, only the publication(s) deemed most important were included. We expect the search terms used to cover most, if not all, of the work incorporating deep learning methods. The last update to the included papers was on February 1, 2017. The appendix describes the search process in more detail.

Summarizing, with this survey we aim to:

- show that deep learning techniques have permeated the entire field of medical image analysis;
- identify the challenges for successful application of deep learning to medical imaging tasks;
- highlight specific contributions which solve or circumvent these challenges.

The rest of this survey as structured as followed. In Section 2 we introduce the main deep learning techniques that have been used for medical image analysis and that are referred to throughout the survey. Section 3 describes the contributions of deep learning to canonical tasks in medical image analysis: classification, detection, segmentation, registration, retrieval, image generation and enhancement. Section 4 discusses obtained results and open challenges in different application areas: neuro, ophthalmic, pulmonary, digital pathology and cell imaging, breast, cardiac, abdominal, musculoskeletal, and remaining miscellaneous applications. We end with a summary, a critical discussion and an outlook for future research.

2. Overview of deep learning methods

The goal of this section is to provide a formal introduction and definition of the deep learning concepts,

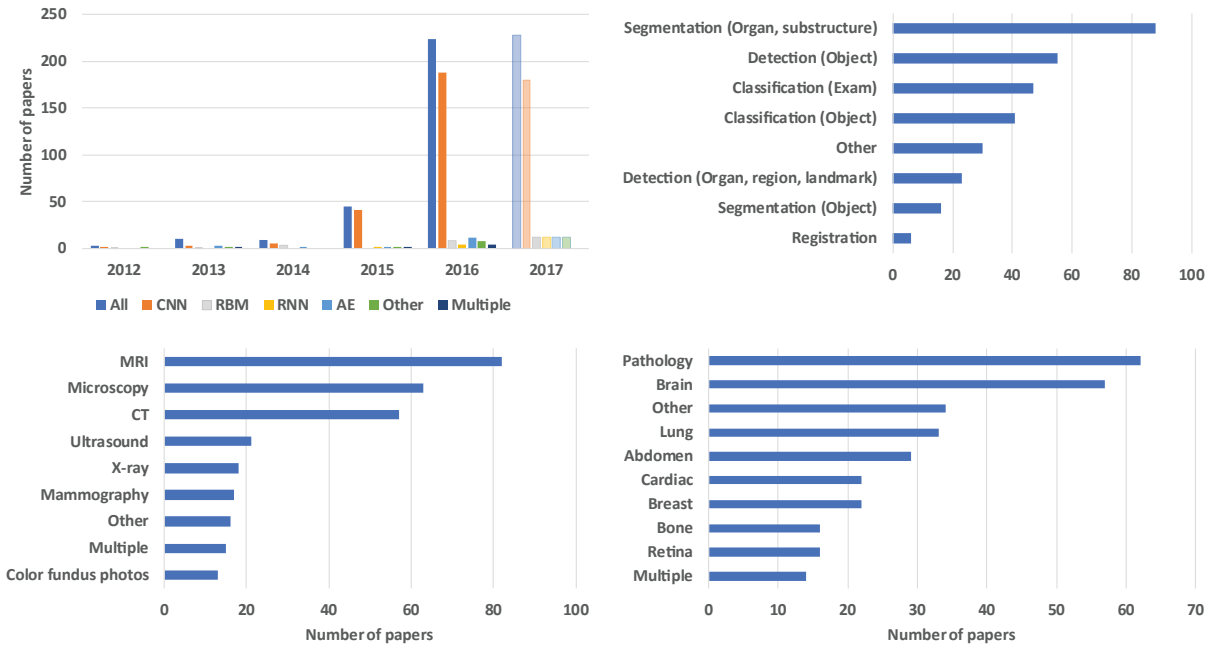

Figure 1: Breakdown of the papers included in this survey in year of publication, task addressed (Section 3), imaging modality, and application area (Section 4). The number of papers for 2017 has been extrapolated from the papers published in January.

techniques and architectures that we found in the medical image analysis papers surveyed in this work.

2.1. Learning algorithms

Machine learning methods are generally divided into *supervised* and *unsupervised learning* algorithms, although there are many nuances. In supervised learning, a model is presented with a dataset $\mathcal{D} = \{\mathbf{x}, y\}_{n=1}^{N}$ of input features $\mathbf{x}$ and label y pairs, where y typically represents an instance of a fixed set of classes. In the case of regression tasks y can also be a vector with continuous values. Supervised training typically amounts to finding model parameters Θ that best predict the data based on a loss function $L(y, \hat{y})$. Here $\hat{y}$ denotes the output of the model obtained by feeding a data point $\mathbf{x}$ to the function $f(\mathbf{x}; \Theta)$ that represents the model.

Unsupervised learning algorithms process data without labels and are trained to find patterns, such as latent subspaces. Examples of traditional unsupervised learning algorithms are principal component analysis and clustering methods. Unsupervised training can be performed under many different loss functions. One example is reconstruction loss $L(\mathbf{x}, \hat{\mathbf{x}})$ where the model has to learn to reconstruct its input, often through a lower-dimensional or noisy representation.

2.2. Neural Networks

Neural networks are a type of learning algorithm which forms the basis of most deep learning methods. A neural network comprises of neurons or units with some activation a and parameters $\Theta = \{\mathcal{W}, \mathcal{B}\}$, where $\mathcal{W}$ is a set of weights and $\mathcal{B}$ a set of biases. The activation represents a linear combination of the input $\mathbf{x}$ to the neuron and the parameters, followed by an element-wise non-linearity $\sigma(\cdot)$, referred to as a transfer function:

$$a = \sigma(\mathbf{w}^T\mathbf{x} + b). \tag{1}$$

Typical transfer functions for traditional neural networks are the sigmoid and hyperbolic tangent function. The multi-layered perceptrons (MLP), the most well-known of the traditional neural networks, have several layers of these transformations:

$$f(\mathbf{x}; \Theta) = \sigma(\mathbf{W}^T\sigma(\mathbf{W}^T \ldots \sigma(\mathbf{W}^T\mathbf{x} + b)) + b). \tag{2}$$

Here, $\mathbf{W}$ is a matrix comprising of columns $\mathbf{w}_k$, associated with activation k in the output. Layers in between the input and output are often referred to as 'hidden' layers. When a neural network contains multiple hidden layers it is typically considered a 'deep' neural network, hence the term 'deep learning'.

At the final layer of the network the activations are mapped to a distribution over classes $P(y|\mathbf{x}; \Theta)$ through

a *softmax* function:

$$P(y|\mathbf{x};\Theta) = \text{softmax}(\mathbf{x};\Theta) = \frac{e^{\mathbf{w}_i^T\mathbf{x}+b_i}}{\sum_{k=1}^{K} e^{\mathbf{w}_k^T\mathbf{x}+b_k}}, \quad (3)$$

where $\mathbf{w}_i$ indicates the weight vector leading to the output node associated with class i. A schematic representation of three-layer MLP is shown in Figure 2.

Maximum likelihood with stochastic gradient descent is currently the most popular method to fit parameters Θ to a dataset $\mathcal{D}$. In stochastic gradient descent a small subset of the data, a mini-batch, is used for each gradient update instead of the full data set. Optimizing maximum likelihood in practice amounts to minimizing the negative log-likelihood:

$$\arg\min_{\Theta} -\sum_{n=1}^{N} \log\left[P(y_n|\mathbf{x}_n;\Theta)\right]. \quad (4)$$

This results in the binary cross-entropy loss for two-class problems and the categorical cross-entropy for multi-class tasks. A downside of this approach is that it typically does not optimize the quantity we are interested in directly, such as area under the receiver-operating characteristic (ROC) curve or common evaluation measures for segmentation, such as the Dice coefficient.

For a long time, deep neural networks (DNN) were considered hard to train efficiently. They only gained popularity in 2006 (Bengio et al., 2007; Hinton and Salakhutdinov, 2006; Hinton et al., 2006) when it was shown that training DNNs layer-by-layer in an unsupervised manner (pre-training), followed by supervised fine-tuning of the stacked network, could result in good performance. Two popular architectures trained in such a way are stacked auto-encoders (SAEs) and deep belief networks (DBNs). However, these techniques are rather complex and require a significant amount of engineering to generate satisfactory results.

Currently, the most popular models are trained end-to-end in a supervised fashion, greatly simplifying the training process. The most popular architectures are convolutional neural networks (CNNs) and recurrent neural networks (RNNs). CNNs are currently most widely used in (medical) image analysis, although RNNs are gaining popularity. The following sections will give a brief overview of each of these methods, starting with the most popular ones, and discussing their differences and potential challenges when applied to medical problems.

2.3. Convolutional Neural Networks (CNNs)

There are two key differences between MLPs and CNNs. First, in CNNs weights in the network are shared in such a way that it the network performs convolution operations on images. This way, the model does not need to learn separate detectors for the same object occurring at different positions in an image, making the network equivariant with respect to translations of the input. It also drastically reduces the amount of parameters (i.e. the number of weights no longer depends on the size of the input image) that need to be learned. An example of a 1D CNN is shown in Figure 2.

At each layer, the input image is convolved with a set of K kernels $\mathcal{W} = \{\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_K\}$ and added biases $\mathcal{B} = \{b_1, \ldots, b_K\}$, each generating a new feature map $\mathbf{X}_k$. These features are subjected to an element-wise non-linear transform $\sigma(\cdot)$ and the same process is repeated for every convolutional layer l:

$$\mathbf{X}_k^l = \sigma(\mathbf{W}_k^{l-1} * \mathbf{X}^{l-1} + b_k^{l-1}). \quad (5)$$

The second key difference between CNNs and MLPs, is the typical incorporation of pooling layers in CNNs, where pixel values of neighborhoods are aggregated using a permutation invariant function, typically the max or mean operation. This induces a certain amount of translation invariance and again reduces the amount of parameters in the network. At the end of the convolutional stream of the network, fully-connected layers (i.e. regular neural network layers) are usually added, where weights are no longer shared. Similar to MLPs, a distribution over classes is generated by feeding the activations in the final layer through a softmax function and the network is trained using maximum likelihood.

2.4. Deep CNN Architectures

Given the prevalence of CNNs in medical image analysis, we elaborate on the most common architectures and architectural differences among the widely used models.

2.4.1. General classification architectures

LeNet (LeCun et al., 1998) and AlexNet (Krizhevsky et al., 2012), introduced over a decade later, were in essence very similar models. Both networks were relatively shallow, consisting of two and five convolutional layers, respectively, and employed kernels with large receptive fields in layers close to the input and smaller kernels closer to the output. AlexNet did incorporate rectified linear units instead of the hyperbolic tangent as activation function.

After 2012 the exploration of novel architectures took off, and in the last three years there is a preference for far deeper models. By stacking smaller kernels, instead of using a single layer of kernels with a large receptive field, a similar function can be represented with less parameters. These deeper architectures generally have a lower memory footprint during inference, which enable their deployment on mobile computing devices such as smartphones. Simonyan and Zisserman (2014) were the first to explore much deeper networks, and employed small, fixed size kernels in each layer. A 19-layer model often referred to as VGG19 or OxfordNet won the ImageNet challenge of 2014.

On top of the deeper networks, more complex building blocks have been introduced that improve the efficiency of the training procedure and again reduce the amount of parameters. Szegedy et al. (2014) introduced a 22-layer network named *GoogLeNet*, also referred to as Inception, which made use of so-called inception blocks (Lin et al., 2013), a module that replaces the mapping defined in Eq. (5) with a set of convolutions of different sizes. Similar to the stacking of small kernels, this allows a similar function to be represented with less parameters. The *ResNet* architecture (He et al., 2015) won the ImageNet challenge in 2015 and consisted of so-called ResNet-blocks. Rather than learning a function, the residual block only learns the residual and is thereby pre-conditioned towards learning mappings in each layer that are close to the identity function. This way, even deeper models can be trained effectively.

Since 2014, the performance on the ImageNet benchmark has saturated and it is difficult to assess whether the small increases in performance can really be attributed to 'better' and more sophisticated architectures. The advantage of the lower memory footprint these models provide is typically not as important for medical applications. Consequently, AlexNet or other simple models such as VGG are still popular for medical data, though recent landmark studies all use a version of GoogleNet called Inception v3 (Gulshan et al., 2016; Esteva et al., 2017; Liu et al., 2017). Whether this is due to a superior architecture or simply because the model is a default choice in popular software packages is again difficult to assess.

2.4.2. *Multi-stream architectures*

The default CNN architecture can easily accommodate multiple sources of information or representations of the input, in the form of channels presented to the input layer. This idea can be taken further and channels can be merged at any point in the network. Under the intuition that different tasks require different ways of fusion, *multi-stream* architectures are being explored. These models, also referred to as dual pathway architectures (Kamnitsas et al., 2017), have two main applications at the time of writing: (1) multi-scale image analysis and (2) 2.5D classification; both relevant for medical image processing tasks.

For the detection of abnormalities, context is often an important cue. The most straightforward way to increase context is to feed larger patches to the network, but this can significantly increase the amount of parameters and memory requirements of a network. Consequently, architectures have been investigated where context is added in a down-scaled representation in addition to high resolution local information. To the best of our knowledge, the multi-stream multi-scale architecture was first explored by Farabet et al. (2013), who used it for segmentation in natural images. Several medical applications have also successfully used this concept (Kamnitsas et al., 2017; Moeskops et al., 2016a; Song et al., 2015; Yang et al., 2016c).

As so much methodology is still developed on natural images, the challenge of applying deep learning techniques to the medical domain often lies in adapting existing architectures to, for instance, different input formats such as three-dimensional data. In early applications of CNNs to such volumetric data, full 3D convolutions and the resulting large amount of parameters were circumvented by dividing the Volume of Interest (VOI) into slices which are fed as different streams to a network. Prasoon et al. (2013) were the first to use this approach for knee cartilage segmentation. Similarly, the network can be fed with multiple angled patches from the 3D-space in a multi-stream fashion, which has been applied by various authors in the context of medical imaging (Roth et al., 2016b; Setio et al., 2016). These approaches are also referred to as 2.5D classification.

2.4.3. *Segmentation Architectures*

Segmentation is a common task in both natural and medical image analysis and to tackle this, CNNs can simply be used to classify each pixel in the image individually, by presenting it with patches extracted around the particular pixel. A drawback of this naive 'sliding-window' approach is that input patches from neighboring pixels have huge overlap and the same convolutions are computed many times. Fortunately, the convolution and dot product are both linear operators and thus inner products can be written as convolutions and vice versa. By rewriting the fully connected layers as convolutions, the CNN can take input images larger than it was trained on and produce a likelihood map, rather than an output for a single pixel. The resulting 'fully convolutional

network' (fCNN) can then be applied to an entire input image or volume in an efficient fashion.

However, because of pooling layers, this may result in output with a far lower resolution than the input. 'Shift-and-stitch' (Long et al., 2015) is one of several methods proposed to prevent this decrease in resolution. The fCNN is applied to shifted versions of the input image. By stitching the result together, one obtains a full resolution version of the final output, minus the pixels lost due to the 'valid' convolutions.

Ronneberger et al. (2015) took the idea of the fCNN one step further and proposed the U-net architecture, comprising a 'regular' fCNN followed by an upsampling part where 'up'-convolutions are used to increase the image size, coined contractive and expansive paths. Although this is not the first paper to introduce learned upsampling paths in convolutional neural networks (e.g. Long et al. (2015)), the authors combined it with so called skip-connections to directly connect opposing contracting and expanding convolutional layers. A similar approach was used by Çiçek et al. (2016) for 3D data. Milletari et al. (2016b) proposed an extension to the U-Net layout that incorporates ResNet-like residual blocks and a Dice loss layer, rather than the conventional cross-entropy, that directly minimizes this commonly used segmentation error measure.

2.5. Recurrent Neural Networks (RNNs)

Traditionally, RNNs were developed for discrete sequence analysis. They can be seen as a generalization of MLPs because both the input and output can be of varying length, making them suitable for tasks such as machine translation where a sentence of the source and target language are the input and output. In a classification setting, the model learns a distribution over classes $P(y|\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T; \Theta)$ given a sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T$, rather than a single input vector $\mathbf{x}$.

The plain RNN maintains a latent or hidden state $\mathbf{h}$ at time t that is the output of a non-linear mapping from its input $\mathbf{x}_t$ and the previous state $\mathbf{h}_{t-1}$:

$$\mathbf{h}_t = \sigma(\mathbf{W}\mathbf{x}_t + \mathbf{R}\mathbf{h}_{t-1} + \mathbf{b}), \tag{6}$$

where weight matrices $\mathbf{W}$ and $\mathbf{R}$ are shared over time. For classification, one or more fully connected layers are typically added followed by a softmax to map the sequence to a posterior over the classes.

$$P(y|\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T; \Theta) = \text{softmax}(\mathbf{h}_T; \mathbf{W}_{out}, \mathbf{b}_{out}). \tag{7}$$

Since the gradient needs to be backpropagated from the output through time, RNNs are inherently deep (in time) and consequently suffer from the same problems with training as regular deep neural networks (Bengio et al., 1994). To this end, several specialized memory units have been developed, the earliest and most popular being the Long Short Term Memory (LSTM) cell (Hochreiter and Schmidhuber, 1997). The Gated Recurrent Unit (Cho et al., 2014) is a recent simplification of the LSTM and is also commonly used.

Although initially proposed for one-dimensional input, RNNs are increasingly applied to images. In natural images 'pixelRNNs' are used as autoregressive models, generative models that can eventually produce new images similar to samples in the training set. For medical applications, they have been used for segmentation problems, with promising results (Stollenga et al., 2015) in the MRBrainS challenge.

2.6. Unsupervised models

2.6.1. Auto-encoders (AEs) and Stacked Auto-encoders (SAEs)

AEs are simple networks that are trained to reconstruct the input $\mathbf{x}$ on the output layer $\mathbf{x}'$ through one hidden layer $\mathbf{h}$. They are governed by a weight matrix $\mathbf{W}_{x,h}$ and bias $b_{x,h}$ from input to hidden state and $\mathbf{W}_{h,x'}$ with corresponding bias $b_{h,x'}$ from the hidden layer to the reconstruction. A non-linear function is used to compute the hidden activation:

$$\mathbf{h} = \sigma(\mathbf{W}_{x,h}\mathbf{x} + \mathbf{b}_{x,h}). \tag{8}$$

Additionally, the dimension of the hidden layer $|\mathbf{h}|$ is taken to be smaller than $|\mathbf{x}|$. This way, the data is projected onto a lower dimensional subspace representing a dominant latent structure in the input. Regularization or sparsity constraints can be employed to enhance the discovery process. If the hidden layer had the same size as the input and no further non-linearities were added, the model would simply learn the identity function.

The denoising auto-encoder (Vincent et al., 2010) is another solution to prevent the model from learning a trivial solution. Here the model is trained to reconstruct the input from a noise corrupted version (typically salt-and-pepper-noise). SAEs (or deep AEs) are formed by placing auto-encoder layers on top of each other. In medical applications surveyed in this work, auto-encoder layer were often trained individually ('greedily') after which the full network was fine-tuned using supervised training to make a prediction.

2.6.2. Restricted Boltzmann Machines (RBMs) and Deep Belief Networks (DBNs)

RBMs (Hinton, 2010) are a type of Markov Random Field (MRF), constituting an input layer or visible layer $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ and a hidden layer $\mathbf{h} = (h_1, h_2, \ldots, h_M)$ that carries the latent feature representation. The connections between the nodes are bi-directional, so given an input vector $\mathbf{x}$ one can obtain the latent feature representation $\mathbf{h}$ and also vice versa. As such, the RBM is a generative model, and we can sample from it and generate new data points. In analogy to physical systems, an energy function is defined for a particular state $(\mathbf{x}, \mathbf{h})$ of input and hidden units:

$$E(\mathbf{x}, \mathbf{h}) = \mathbf{h}^T \mathbf{W} \mathbf{x} - \mathbf{c}^T \mathbf{x} - \mathbf{b}^T \mathbf{h}, \tag{9}$$

with $\mathbf{c}$ and $\mathbf{b}$ bias terms. The probability of the 'state' of the system is defined by passing the energy to an exponential and normalizing:

$$p(\mathbf{x}, \mathbf{h}) = \frac{1}{Z} \exp\{-E(\mathbf{x}, \mathbf{h})\}. \tag{10}$$

Computing the partition function Z is generally intractable. However, conditional inference in the form of computing $\mathbf{h}$ conditioned on $\mathbf{v}$ or vice versa is tractable and results in a simple formula:

$$P(h_j|\mathbf{x}) = \frac{1}{1 + \exp\{-b_j - \mathbf{W}_j \mathbf{x}\}}. \tag{11}$$

Since the network is symmetric, a similar expression holds for $P(x_i|\mathbf{h})$.

DBNs (Bengio et al., 2007; Hinton et al., 2006) are essentially SAEs where the AE layers are replaced by RBMs. Training of the individual layers is, again, done in an unsupervised manner. Final fine-tuning is performed by adding a linear classifier to the top layer of the DBN and performing a supervised optimization.

2.6.3. Variational Auto-Encoders and Generative Adverserial Networks

Recently, two novel unsupervised architectures were introduced: the variational auto-encoder (VAE) (Kingma and Welling, 2013) and the generative adversarial network (GAN) (Goodfellow et al., 2014). There are no peer-reviewed papers applying these methods to medical images yet, but applications in natural images are promising. We will elaborate on their potential in the discussion.

2.7. Hardware and Software

One of the main contributors to steep rise of deep learning has been the widespread availability of GPU and GPU-computing libraries (CUDA, OpenCL). GPUs are highly parallel computing engines, which have an order of magnitude more execution threads than central processing units (CPUs). With current hardware, deep learning on GPUs is typically 10 to 30 times faster than on CPUs.

Next to hardware, the other driving force behind the popularity of deep learning methods is the wide availability of open source software packages. These libraries provide efficient GPU implementations of important operations in neural networks, such as convolutions; allowing the user to implement ideas at a high level rather than worrying about low-level efficient implementations. At the time of writing, the most popular packages were (in alphabetical order):

- **Caffe** (Jia et al., 2014). Provides C++ and Python interfaces, developed by graduate students at UC Berkeley.
- **Tensorflow** (Abadi et al., 2016). Provides C++ and Python and interfaces, developed by Google and is used by Google research.
- **Theano** (Bastien et al., 2012). Provides a Python interface, developed by MILA lab in Montreal.
- **Torch** (Collobert et al., 2011). Provides a Lua interface and is used by, among others, Facebook AI research.

There are third-party packages written on top of one or more of these frameworks, such as Lasagne (`https://github.com/Lasagne/Lasagne`) or Keras (`https://keras.io/`). It goes beyond the scope of this paper to discuss all these packages in detail.

3. Deep Learning Uses in Medical Imaging

3.1. Classification

3.1.1. Image/exam classification

Image or exam classification was one of the first areas in which deep learning made a major contribution to medical image analysis. In exam classification one typically has one or multiple images (an exam) as input with a single diagnostic variable as output (e.g., disease present or not). In such a setting, every diagnostic exam is a sample and dataset sizes are typically

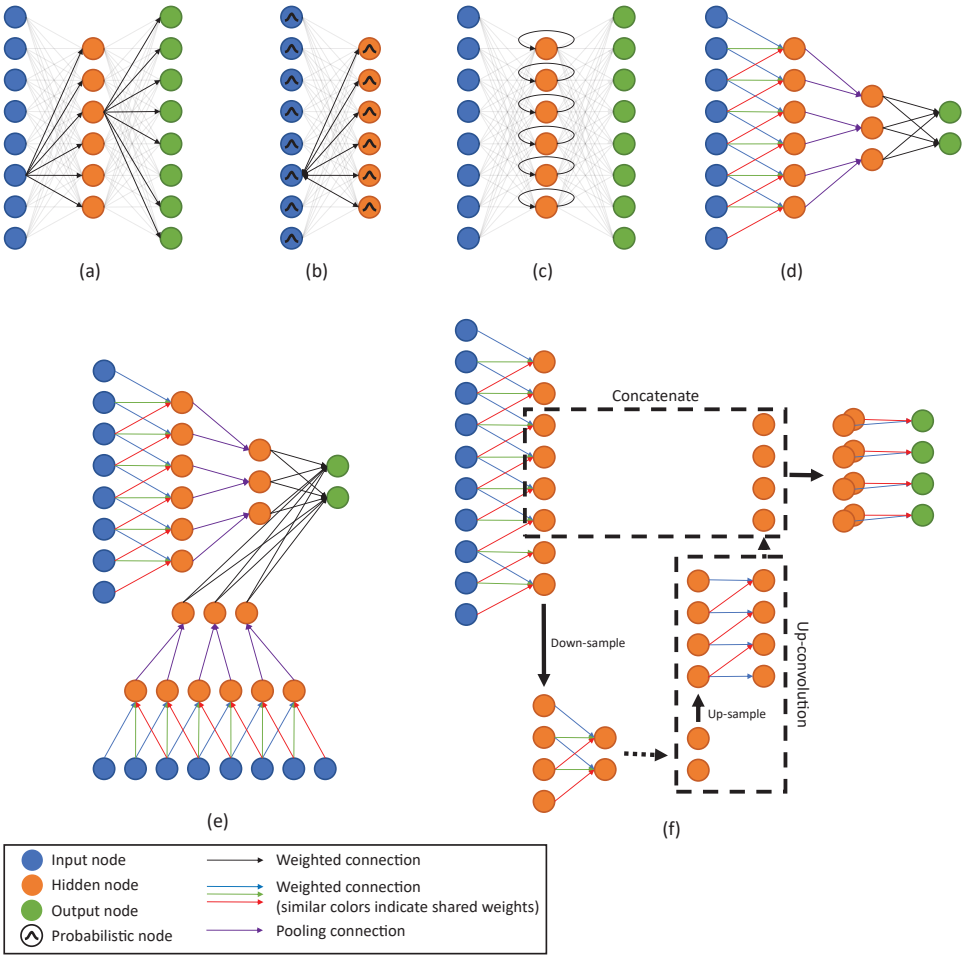

Figure 2: Node graphs of 1D representations of architectures commonly used in medical imaging. a) Auto-encoder, b) restricted Boltzmann machine, c) recurrent neural network, d) convolutional neural network, e) multi-stream convolutional neural network, f) U-net (with a single downsampling stage).

.

small compared to those in computer vision (e.g., hundreds/thousands vs. millions of samples). The popularity of transfer learning for such applications is therefore not surprising.

Transfer learning is essentially the use of pre-trained networks (typically on natural images) to try to work around the (perceived) requirement of large data sets for deep network training. Two transfer learning strategies were identified: (1) using a pre-trained network as a feature extractor and (2) fine-tuning a pre-trained network on medical data. The former strategy has the extra benefit of not requiring one to train a deep network at all, allowing the extracted features to be easily plugged in to existing image analysis pipelines. Both strategies are popular and have been widely applied. However, few authors perform a thorough investigation in which strategy gives the best result. The two papers that do, Antony et al. (2016) and Kim et al. (2016a), offer conflicting results. In the case of Antony et al. (2016), fine-tuning clearly outperformed feature extraction, achieving 57.6% accuracy in multi-class grade assessment of knee osteoarthritis versus 53.4%. Kim et al. (2016a), however, showed that using CNN as a feature extractor outperformed fine-tuning in cytopathology image classification accuracy (70.5% versus 69.1%). If any guidance can be given to which strategy might be most successful, we would refer the reader to two recent papers, published in high-ranking journals, which fine-tuned a pre-trained version of Google's Inception v3 architecture on medical data and achieved (near) human expert

performance (Esteva et al., 2017; Gulshan et al., 2016). As far as the authors are aware, such results have not yet been achieved by simply using pre-trained networks as feature extractors.

With respect to the type of deep networks that are commonly used in exam classification, a timeline similar to computer vision is apparent. The medical imaging community initially focused on unsupervised pre-training and network architectures like SAEs and RBMs. The first papers applying these techniques for exam classification appeared in 2013 and focused on neuroimaging. Brosch and Tam (2013), Plis et al. (2014), Suk and Shen (2013), and Suk et al. (2014) applied DBNs and SAEs to classify patients as having Alzheimer's disease based on brain Magnetic Resonance Imaging (MRI). Recently, a clear shift towards CNNs can be observed. Out of the 47 papers published on exam classification in 2015, 2016, and 2017, 36 are using CNNs, 5 are based on AEs and 6 on RBMs. The application areas of these methods are very diverse, ranging from brain MRI to retinal imaging and digital pathology to lung computed tomography (CT).

In the more recent papers using CNNs authors also often train their own network architectures from scratch instead of using pre-trained networks. Menegola et al. (2016) performed some experiments comparing training from scratch to fine-tuning of pre-trained networks and showed that fine-tuning worked better given a small data set of around a 1000 images of skin lesions. However, these experiments are too small scale to be able to draw any general conclusions from.

Three papers used an architecture leveraging the unique attributes of medical data: two use 3D convolutions (Hosseini-Asl et al., 2016; Payan and Montana, 2015) instead of 2D to classify patients as having Alzheimer; Kawahara et al. (2016b) applied a CNN-like architecture to a brain connectivity graph derived from MRI diffusion-tensor imaging (DTI). In order to do this, they developed several new layers which formed the basis of their network, so-called edge-to-edge, edge-to-node, and node-to-graph layers. They used their network to predict brain development and showed that they outperformed existing methods in assessing cognitive and motor scores.

Summarizing, in exam classification CNNs are the current standard techniques. Especially CNNs pre-trained on natural images have shown surprisingly strong results, challenging the accuracy of human experts in some tasks. Last, authors have shown that CNNs can be adapted to leverage intrinsic structure of medical images.

3.1.2. Object or lesion classification

Object classification usually focuses on the classification of a small (previously identified) part of the medical image into two or more classes (e.g. nodule classification in chest CT). For many of these tasks both local information on lesion appearance and global contextual information on lesion location are required for accurate classification. This combination is typically not possible in generic deep learning architectures. Several authors have used multi-stream architectures to resolve this in a multi-scale fashion (Section 2.4.2). Shen et al. (2015b) used three CNNs, each of which takes a nodule patch at a different scale as input. The resulting feature outputs of the three CNNs are then concatenated to form the final feature vector. A somewhat similar approach was followed by Kawahara and Hamarneh (2016) who used a multi-stream CNN to classify skin lesions, where each stream works on a different resolution of the image. Gao et al. (2015) proposed to use a combination of CNNs and RNNs for grading nuclear cataracts in slit-lamp images, where CNN filters were pre-trained. This combination allows the processing of all contextual information regardless of image size. Incorporating 3D information is also often a necessity for good performance in object classification tasks in medical imaging. As images in computer vision tend to be 2D natural images, networks developed in those scenarios do not directly leverage 3D information. Authors have used different approaches to integrate 3D in an effective manner with custom architectures. Setio et al. (2016) used a multi-stream CNN to classify points of interest in chest CT as a nodule or non-nodule. Up to nine differently oriented patches extracted from the candidate were used in separate streams and merged in the fully-connected layers to obtain the final classification output. In contrast, Nie et al. (2016c) exploited the 3D nature of MRI by training a 3D CNN to assess survival in patients suffering from high-grade gliomas.

Almost all recent papers prefer the use of end-to-end trained CNNs. In some cases other architectures and approaches are used, such as RBMs (van Tulder and de Bruijne, 2016; Zhang et al., 2016c), SAEs (Cheng et al., 2016a) and convolutional sparse auto-encoders (CSAE) (Kallenberg et al., 2016). The major difference between CSAE and a classic CNN is the usage of unsupervised pre-training with sparse auto-encoders.

An interesting approach, especially in cases where object annotation to generate training data is expensive, is the integration of multiple instance learning (MIL) and deep learning. Xu et al. (2014) investigated the use of a MIL-framework with both supervised and unsu-

pervised feature learning approaches as well as handcrafted features. The results demonstrated that the performance of the MIL-framework was superior to handcrafted features, which in turn closely approaches the performance of a fully supervised method. We expect such approaches to be popular in the future as well, as obtaining high-quality annotated medical data is challenging.

Overall, object classification sees less use of pretrained networks compared to exam classifications, mostly due to the need for incorporation of contextual or three-dimensional information. Several authors have found innovative solutions to add this information to deep networks with good results, and as such we expect deep learning to become even more prominent for this task in the near future.

3.2. Detection

3.2.1. Organ, region and landmark localization

Anatomical object localization (in space or time), such as organs or landmarks, has been an important preprocessing step in segmentation tasks or in the clinical workflow for therapy planning and intervention. Localization in medical imaging often requires parsing of 3D volumes. To solve 3D data parsing with deep learning algorithms, several approaches have been proposed that treat the 3D space as a composition of 2D orthogonal planes. Yang et al. (2015) identified landmarks on the distal femur surface by processing three independent sets of 2D MRI slices (one for each plane) with regular CNNs. The 3D position of the landmark was defined as the intersection of the three 2D slices with the highest classification output. de Vos et al. (2016b) went one step further and localized regions of interest (ROIs) around anatomical regions (heart, aortic arch, and descending aorta) by identifying a rectangular 3D bounding box after 2D parsing the 3D CT volume. Pretrained CNN architectures, as well as RBM, have been used for the same purpose (Cai et al., 2016b; Chen et al., 2015b; Kumar et al., 2016), overcoming the lack of data to learn better feature representations. All these studies cast the localization task as a classification task and as such generic deep learning architectures and learning processes can be leveraged.

Other authors try to modify the network learning process to directly predict locations. For example, Payer et al. (2016) proposed to directly regress landmark locations with CNNs. They used landmark maps, where each landmark is represented by a Gaussian, as ground truth input data and the network is directly trained to predict this landmark map. Another interesting approach was published by Ghesu et al. (2016a), in which reinforcement learning is applied to the identification of landmarks. The authors showed promising results in several tasks: 2D cardiac MRI and ultrasound (US) and 3D head/neck CT.

Due to its increased complexity, only a few methods addressed the direct localization of landmarks and regions in the 3D image space. Zheng et al. (2015) reduced this complexity by decomposing 3D convolution as three one-dimensional convolutions for carotid artery bifurcation detection in CT data. Ghesu et al. (2016b) proposed a sparse adaptive deep neural network powered by marginal space learning in order to deal with data complexity in the detection of the aortic valve in 3D transesophageal echocardiogram.

CNNs have also been used for the localization of scan planes or key frames in temporal data. Baumgartner et al. (2016) trained CNNs on video frame data to detect up to 12 standardized scan planes in mid-pregnancy fetal US. Furthermore, they used saliency maps to obtain a rough localization of the object of interest in the scan plan (e.g. brain, spine). RNNs, particularly LSTM-RNNs, have also been used to exploit the temporal information contained in medical videos, another type of high dimensional data. Chen et al. (2015a), for example, employed LSTM models to incorporate temporal information of consecutive sequence in US videos for fetal standard plane detection. Kong et al. (2016) combined an LSTM-RNN with a CNN to detect the end-diastole and end-systole frames in cine-MRI of the heart.

Concluding, localization through 2D image classification with CNNs seems to be the most popular strategy overall to identify organs, regions and landmarks, with good results. However, several recent papers expand on this concept by modifying the learning process such that accurate localization is directly emphasized, with promising results. We expect such strategies to be explored further as they show that deep learning techniques can be adapted to a wide range of localization tasks (e.g. multiple landmarks). RNNs have shown promise in localization in the temporal domain, and multi-dimensional RNNs could play a role in spatial localization as well.

3.2.2. Object or lesion detection

The detection of objects of interest or lesions in images is a key part of diagnosis and is one of the most labor-intensive for clinicians. Typically, the tasks consist of the localization and identification of small lesions in the full image space. There has been a long research tradition in computer-aided detection systems that are designed to automatically detect lesions, improving the

detection accuracy or decreasing the reading time of human experts. Interestingly, the first object detection system using CNNs was already proposed in 1995, using a CNN with four layers to detect nodules in x-ray images (Lo et al., 1995).

Most of the published deep learning object detection systems still uses CNNs to perform pixel (or voxel) classification, after which some form of post processing is applied to obtain object candidates. As the classification task performed at each pixel is essentially object classification, CNN architecture and methodology are very similar to those in section 3.1.2. The incorporation of contextual or 3D information is also handled using multi-stream CNNs (Section 2.4.2, for example by Barbu et al. (2016) and Roth et al. (2016b). Teramoto et al. (2016) used a multi-stream CNN to integrate CT and Positron Emission Tomography (PET) data. Dou et al. (2016c) used a 3D CNN to find micro-bleeds in brain MRI. Last, as the annotation burden to generate training data can be similarly significant compared to object classification, weakly-supervised deep learning has been explored by Hwang and Kim (2016), who adopted such a strategy for the detection of nodules in chest radiographs and lesions in mammography.

There are some aspects which are significantly different between object detection and object classification. One key point is that because every pixel is classified, typically the class balance is skewed severely towards the non-object class in a training setting. To add insult to injury, usually the majority of the non-object samples are easy to discriminate, preventing the deep learning method to focus on the challenging samples. van Grinsven et al. (2016) proposed a selective data sampling in which wrongly classified samples were fed back to the network more often to focus on challenging areas in retinal images. Last, as classifying each pixel in a sliding window fashion results in orders of magnitude of redundant calculation, fCNNs, as used in Wolterink et al. (2016), are important aspect of an object detection pipeline as well.

Challenges in meaningful application of deep learning algorithms in object detection are thus mostly similar to those in object classification. Only few papers directly address issues specific to object detection like class imbalance/hard-negative mining or efficient pixel/voxel-wise processing of images. We expect that more emphasis will be given to those areas in the near future, for example in the application of multi-stream networks in a fully convolutional fashion.

3.3. Segmentation

3.3.1. Organ and substructure segmentation

The segmentation of organs and other substructures in medical images allows quantitative analysis of clinical parameters related to volume and shape, as, for example, in cardiac or brain analysis. Furthermore, it is often an important first step in computer-aided detection pipelines. The task of segmentation is typically defined as identifying the set of voxels which make up either the contour or the interior of the object(s) of interest. Segmentation is the most common subject of papers applying deep learning to medical imaging (Figure 1), and as such has also seen the widest variety in methodology, including the development of unique CNN-based segmentation architectures and the wider application of RNNs.

The most well-known, in medical image analysis, of these novel CNN architectures is U-net, published by Ronneberger et al. (2015) (section 2.4.3). The two main architectural novelties in U-net are the combination of an equal amount of upsampling and downsampling layers. Although learned upsampling layers have been proposed before, U-net combines them with so-called skip connections between opposing convolution and deconvolution layers. This which concatenate features from the contracting and expanding paths. From a training perspective this means that entire images/scans can be processed by U-net in one forward pass, resulting in a segmentation map directly. This allows U-net to take into account the full context of the image, which can be an advantage in contrast to patch-based CNNs. Furthermore, in an extended paper by Çiçek et al. (2016), it is shown that a full 3D segmentation can be achieved by feeding U-net with a few 2D annotated slices from the same volume. Other authors have also built derivatives of the U-net architecture; Milletari et al. (2016b), for example, proposed a 3D-variant of U-net architecture, called V-net, performing 3D image segmentation using 3D convolutional layers with an objective function directly based on the Dice coefficient. Drozdzal et al. (2016) investigated the use of short ResNet-like skip connections in addition to the long skip-connections in a regular U-net.

RNNs have recently become more popular for segmentation tasks. For example, Xie et al. (2016b) used a spatial clockwork RNN to segment the perimysium in H&E-histopathology images. This network takes into account prior information from both the row and column predecessors of the current patch. To incorporate bidirectional information from both left/top and right/bottom neighbors, the RNN is applied four times

in different orientations and the end-result is concatenated and fed to a fully-connected layer. This produces the final output for a single patch. Stollenga et al. (2015) where the first to use a 3D LSTM-RNN with convolutional layers in six directions. Andermatt et al. (2016) used a 3D RNN with gated recurrent units to segment gray and white matter in a brain MRI data set. Chen et al. (2016d) combined bi-directional LSTM-RNNs with 2D U-net-like-architectures to segment structures in anisotropic 3D electron microscopy images. Last, Poudel et al. (2016) combined a 2D U-net architecture with a gated recurrent unit to perform 3D segmentation.

Although these specific segmentation architectures offered compelling advantages, many authors have also obtained excellent segmentation results with patch-trained neural networks. One of the earliest papers covering medical image segmentation with deep learning algorithms used such a strategy and was published by Ciresan et al. (2012). They applied pixel-wise segmentation of membranes in electron microscopy imagery in a sliding window fashion. Most recent papers now use fCNNs (subsection 2.4.3) in preference over sliding-window-based classification to reduce redundant computation.

fCNNs have also been extended to 3D and have been applied to multiple targets at once: Korez et al. (2016), used 3D fCNNs to generate vertebral body likelihood maps which drove deformable models for vertebral body segmentation in MR images, Zhou et al. (2016) segmented nineteen targets in the human torso, and Moeskops et al. (2016b) trained a single fCNN to segment brain MRI, the pectoral muscle in breast MRI, and the coronary arteries in cardiac CT angiography (CTA).

One challenge with voxel classification approaches is that they sometimes lead to spurious responses. To combat this, groups have tried to combine fCNNs with graphical models like MRFs (Shakeri et al., 2016; Song et al., 2015) and Conditional Random Fields (CRFs) (Alansary et al., 2016; Cai et al., 2016a; Christ et al., 2016; Dou et al., 2016c; Fu et al., 2016a; Gao et al., 2016c) to refine the segmentation output. In most of the cases, graphical models are applied on top of the likelihood map produced by CNNs or fCNNs and act as label regularizers.

Summarizing, segmentation in medical imaging has seen a huge influx of deep learning related methods. Custom architectures have been created to directly target the segmentation task. These have obtained promising results, rivaling and often improving over results obtained with fCNNs.

3.3.2. Lesion segmentation

Segmentation of lesions combines the challenges of object detection and organ and substructure segmentation in the application of deep learning algorithms. Global and local context are typically needed to perform accurate segmentation, such that multi-stream networks with different scales or non-uniformly sampled patches are used as in for example Kamnitsas et al. (2017) and Ghafoorian et al. (2016b). In lesion segmentation we have also seen the application of U-net and similar architectures to leverage both this global and local context. The architecture used by Wang et al. (2015), similar to the U-net, consists of the same downsampling and upsampling paths, but does not use skip connections. Another U-net-like architecture was used by Brosch et al. (2016) to segment white matter lesions in brain MRI. However, they used 3D convolutions and a single skip connection between the first convolutional and last deconvolutional layers.

One other challenge that lesion segmentation shares with object detection is class imbalance, as most voxels/pixels in an image are from the non-diseased class. Some papers combat this by adapting the loss function: Brosch et al. (2016) defined it to be a weighted combination of the sensitivity and the specificity, with a larger weight for the specificity to make it less sensitive to the data imbalance. Others balance the data set by performing data augmentation on positive samples (Kamnitsas et al., 2017; Litjens et al., 2016; Pereira et al., 2016).

Thus lesion segmentation sees a mixture of approaches used in object detection and organ segmentation. Developments in these two areas will most likely naturally propagate to lesion segmentation as the existing challenges are also mostly similar.

3.4. Registration

Registration (i.e. spatial alignment) of medical images is a common image analysis task in which a coordinate transform is calculated from one medical image to another. Often this is performed in an iterative framework where a specific type of (non-)parametric transformation is assumed and a pre-determined metric (e.g. L2-norm) is optimized. Although segmentation and lesion detection are more popular topics for deep learning, researchers have found that deep networks can be beneficial in getting the best possible registration performance. Broadly speaking, two strategies are prevalent in current literature: (1) using deep-learning networks to estimate a similarity measure for two images to drive an iterative optimization strategy, and (2) to directly predict transformation parameters using deep regression networks.

Wu et al. (2013), Simonovsky et al. (2016), and Cheng et al. (2015) used the first strategy to try to optimize registration algorithms. Cheng et al. (2015) used two types of stacked auto-encoders to assess the local similarity between CT and MRI images of the head. Both auto-encoders take vectorized image patches of CT and MRI and reconstruct them through four layers. After the networks are pre-trained using unsupervised patch reconstruction they are fine-tuned using two prediction layers stacked on top of the third layer of the SAE. These prediction layers determine whether two patches are similar (class 1) or dissimilar (class 2). Simonovsky et al. (2016) used a similar strategy, albeit with CNNs, to estimate a similarity cost between two patches from differing modalities. However, they also presented a way to use the derivative of this metric to directly optimize the transformation parameters, which are decoupled from the network itself. Last, Wu et al. (2013) combined independent subspace analysis and convolutional layers to extract features from input patches in an unsupervised manner. The resultant feature vectors are used to drive the HAMMER registration algorithm instead of handcrafted features.

Miao et al. (2016) and Yang et al. (2016d) used deep learning algorithms to directly predict the registration transform parameters given input images. Miao et al. (2016) leveraged CNNs to perform 3D model to 2D x-ray registration to assess the pose and location of an implanted object during surgery. In total the transformation has 6 parameters, two translational, 1 scaling and 3 angular parameters. They parameterize the feature space in steps of 20 degrees for two angular parameters and train a separate CNN to predict the update to the transformation parameters given an digitally reconstructed x-ray of the 3D model and the actual interoperative x-ray. The CNNs are trained with artificial examples generated by manually adapting the transformation parameters for the input training data. They showed that their approach has significantly higher registration success rates than using traditional - purely intensity based - registration methods. Yang et al. (2016d) tackled the problem of prior/current registration in brain MRI using the OASIS data set. They used the large deformation diffeomorphic metric mapping (LDDMM) registration methodology as a basis. This method takes as input an initial momentum value for each pixel which is then evolved over time to obtain the final transformation. However, the calculation of the initial momentum map is often an expensive procure. The authors circumvent this by training a U-net like architecture to predict the x- and y-momentum map given the input images. They obtain visually similar results but with significantly improved execution time: 1500x speed-up for 2D and 66x speed-up for 3D.

In contrast to classification and segmentation, the research community seems not have yet settled on the best way to integrate deep learning techniques in registration methods. Not many papers have yet appeared on the subject and existing ones each have a distinctly different approach. Thus, giving recommendations on what method is most promising seems inappropriate. However, we expect to see many more contributions of deep learning to medical image registration in the near future.

3.5. Other tasks in medical imaging

3.5.1. Content-based image retrieval

Content-based image retrieval (CBIR) is a technique for knowledge discovery in massive databases and offers the possibility to identify similar case histories, understand rare disorders, and, ultimately, improve patient care. The major challenge in the development of CBIR methods is extracting effective feature representations from the pixel-level information and associating them with meaningful concepts. The ability of deep CNN models to learn rich features at multiple levels of abstraction has elicited interest from the CBIR community.

All current approaches use (pre-trained) CNNs to extract feature descriptors from medical images. Anavi et al. (2016) and Liu et al. (2016b) applied their methods to databases of X-ray images. Both used a five-layer CNN and extracted features from the fully-connected layers. Anavi et al. (2016) used the last layer and a pre-trained network. Their best results were obtained by feeding these features to a one-vs-all support vector machine (SVM) classifier to obtain the distance metric. They showed that incorporating gender information resulted in better performance than just CNN features. Liu et al. (2016b) used the penultimate fully-connected layer and a custom CNN trained to classify X-rays in 193 classes to obtain the descriptive feature vector. After descriptor binarization and data retrieval using Hamming separation values, the performance was inferior to the state of the art, which the authors attributed to small patch sizes of 96 pixels. The method proposed by Shah et al. (2016) combines CNN feature descriptors with hashing-forests. 1000 features were extracted for overlapping patches in prostate MRI volumes, after which a large feature matrix was constructed over all volumes. Hashing forests were then used to compress this into descriptors for each volume.

Content-based image retrieval as a whole has thus not seen many successful applications of deep learning

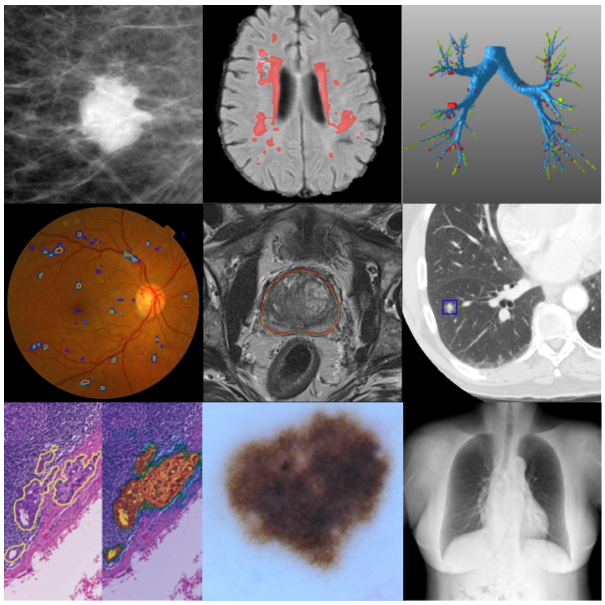

Figure 3: Collage of some medical imaging applications in which deep learning has achieved state-of-the-art results. From top-left to bottom-right: mammographic mass classification (Kooi et al., 2016), segmentation of lesions in the brain (top ranking in BRATS, ISLES and MRBrains challenges, image from Ghafoorian et al. (2016b), leak detection in airway tree segmentation (Charbonnier et al., 2017), diabetic retinopathy classification (Kaggle Diabetic Retinopathy challenge 2015, image from van Grinsven et al. (2016), prostate segmentation (top rank in PROMISE12 challenge), nodule classification (top ranking in LUNA16 challenge), breast cancer metastases detection in lymph nodes (top ranking and human expert performance in CAMELYON16), human expert performance in skin lesion classification (Esteva et al., 2017), and state-of-the-art bone suppression in x-rays, image from Yang et al. (2016c).

methods yet, but given the results in other areas it seems only a matter of time. An interesting avenue of research could be the direct training of deep networks for the retrieval task itself.

3.5.2. *Image Generation and Enhancement*

A variety of image generation and enhancement methods using deep architectures have been proposed, ranging from removing obstructing elements in images, normalizing images, improving image quality, data completion, and pattern discovery.

In image generation, 2D or 3D CNNs are used to convert one input image into another. Typically these architectures lack the pooling layers present in classification networks. These systems are then trained with a data set in which both the input and the desired output are present, defining the differences between the generated and desired output as the loss function. Examples are regular and bone-suppressed X-ray in Yang et al. (2016c), 3T and 7T brain MRI in Bahrami et al. (2016), PET from MRI in Li et al. (2014), and CT from MRI in Nie et al. (2016a). Li et al. (2014) even showed that one can use these generated images in computer-aided diagnosis systems for Alzheimer's disease when the original data is missing or not acquired.

With multi-stream CNNs super-resolution images can be generated from multiple low-resolution inputs (section 2.4.2). In Oktay et al. (2016), multi-stream networks reconstructed high-resolution cardiac MRI from one or more low-resolution input MRI volumes. Not only can this strategy be used to infer missing spatial information, but can also be leveraged in other domains; for example, inferring advanced MRI diffusion parameters from limited data (Golkov et al., 2016). Other image enhancement applications like intensity normalization and denoising have seen only limited application of deep learning algorithms. Janowczyk et al. (2016a) used SAEs to normalize H&E-stained histopathology images whereas Benou et al. (2016) used CNNs to perform denoising in DCE-MRI time-series.

Image generation has seen impressive results with very creative applications of deep networks in significantly differing tasks. One can only expect the number of tasks to increase further in the future.

3.5.3. *Combining Image Data With Reports*

The combination of text reports and medical image data has led to two avenues of research: (1) leveraging reports to improve image classification accuracy (Schlegl et al., 2015), and (2) generating text reports from images (Kisilev et al., 2016; Shin et al., 2015, 2016a; Wang et al., 2016e); the latter inspired by recent caption generation papers from natural images (Karpathy and Fei-Fei, 2015). To the best of our knowledge, the first step towards leveraging reports was taken by Schlegl et al. (2015), who argued that large amounts of annotated data may be difficult to acquire and proposed to add semantic descriptions from reports as labels. The system was trained on sets of images along with their textual descriptions and was taught to predict semantic class labels during test time. They showed that semantic information increases classification accuracy for a variety of pathologies in Optical Coherence Tomography (OCT) images.

Shin et al. (2015) and Wang et al. (2016e) mined semantic interactions between radiology reports and images from a large data set extracted from a PACS system. They employed latent Dirichlet allocation (LDA), a type of stochastic model that generates a distribution over a vocabulary of topics based on words in a document. In a later work, Shin et al. (2016a) proposed a sys-

Table 1: Overview of papers using deep learning techniques for brain image analysis. All works use MRI unless otherwise mentioned.

Reference	Method	Application; remarks
Disorder classification (AD, MCI, Schizophrenia)		
Brosch and Tam (2013)	DBN	AD/HC classification; Deep belief networks with convolutional RBMs for manifold learning
Plis et al. (2014)	DBN	Deep belief networks evaluated on brain network estimation, Schizophrenia and Huntington's disease classification
Suk and Shen (2013)	SAE	AD/MCI classification; Stacked auto encoders with supervised fine tuning
Suk et al. (2014)	RBM	AD/MCI/HC classification; Deep Boltzmann Machines on MRI and PET modalities
Payan and Montana (2015)	CNN	AD/MCI/HC classification; 3D CNN pre-trained with sparse auto-encoders
Suk et al. (2015)	SAE	AD/MCI/HC classification; SAE for latent feature extraction on a large set of hand-crafted features from MRI and PET
Hosseini-Asl et al. (2016)	CNN	AD/MCI/HC classification; 3D CNN pre-trained with a 3D convolutional auto-encoder on fMRI data
Kim et al. (2016b)	ANN	Schizophrenia/NH classification on fMRI; Neural network showing advantage of pre-training with SAEs, and L1 sparsification
Ortiz et al. (2016)	DBN	AD/MCI/HC classification; An ensemble of Deep belief networks, with their votes fused using an SVM classifier
Pinaya et al. (2016)	DBN	Schizophrenia/NH classification; DBN pre-training followed by supervised fine-tuning
Sarraf and Tofighi (2016)	CNN	AD/HC classification; Adapted Lenet-5 architecture on fMRI data
Suk et al. (2016)	SAE	MCI/HC classification of fMRI data; Stacked auto-encoders for feature extraction, HMM as a generative model on top
Suk and Shen (2016)	CNN	AD/MCI/HC classification; CNN on sparse representations created by regression models
Shi et al. (2017)	ANN	AD/MCI/HC classification; Multi-modal stacked deep polynomial networks with an SVM classifier on top using MRI and PET
Tissue/anatomy/lesion/tumor segmentation		
Guo et al. (2014)	SAE	Hippocampus segmentation; SAE for representation learning used for target/atlas patch similarity measurement
de Brebisson and Montana (2015)	CNN	Anatomical segmentation; fusing multi-scale 2D patches with a 3D patch using a CNN
Choi and Jin (2016)	CNN	Striatum segmentation; Two-stage (global/local) approximations with 3D CNNs
Stollenga et al. (2015)	RNN	Tissue segmentation; PyraMiD-LSTM, best brain segmentation results on MRBrainS13 (and competitive results on EM-ISBI12)
Zhang et al. (2015)	CNN	Tissue segmentation; multi-modal 2D CNN
Andermatt et al. (2016)	RNN	Tissue segmentation; two convolutional gated recurrent units in different directions for each dimension
Bao and Chung (2016)	CNN	Anatomical segmentation; Multi-scale late fusion CNN with random walker as a novel label consistency method
Birenbaum and Greenspan (2016)	CNN	Lesion segmentation; Multi-view (2.5D) CNN concatenating features from previous time step for a longitudinal analysis
Brosch et al. (2016)	CNN	Lesion segmentation; Convolutional encoder-decoder network with shortcut connections and convolutional RBM pretraining
Chen et al. (2016a)	CNN	Tissue segmentation; 3D res-net combining features from different layers
Ghafoorian et al. (2016b)	CNN	Lesion segmentation; CNN trained on non-uniformly sampled patch to integrate a larger context with a foviation effect
Ghafoorian et al. (2016a)	CNN	Lesion segmentation; multi-scale CNN with late fusion that integrates anatomical location information into network
Havaei et al. (2016b)	CNN	Tumor segmentation; CNN handling missing modalities with abstraction layer that transforms feature maps to their statistics
Havaei et al. (2016a)	CNN	Tumor segmentation; two-path way CNN with different receptive fields
Kamnitsas et al. (2017)	CNN	Tumor segmentation; 3D multi-scale fully convolutional network with CRF for label consistency
Kleesiek et al. (2016)	CNN	Brain extraction; 3D fully convolutional CNN on multi-modal input
Mansoor et al. (2016)	SAE	Visual pathway segmentation; Learning appearance features from SAE for steering the shape model for segmentation
Milletari et al. (2016a)	CNN	Anatomical segmentation on MRI and US; Hough-voting to acquire mapping from CNN features to full patch segmentations
Moeskops et al. (2016a)	CNN	Tissue segmentation; CNN trained on multiple patch sizes
Nie et al. (2016b)	CNN	Infant tissue segmentation; FCN with a late fusion method on different modalities
Pereira et al. (2016)	CNN	Tumor segmentation; CNN on multiple modality input
Shakeri et al. (2016)	CNN	Anatomical segmentation; FCN followed by Markov random fields
Zhao and Jia (2016)	CNN	Tumor segmentation; Multi-scale CNN with a late fusion architecture
Lesion/tumor detection and classification		
Pan et al. (2015)	CNN	Tumor grading; 2D tumor patch classification using a CNN
Dou et al. (2015)	ISA	Microbleed detection; 3D stacked Independent Subspace Analysis for candidate feature extraction, SVM classification
Dou et al. (2016c)	CNN	Microbleed detection; 3D FCN for candidate segmentation followed by a 3D CNN as false positive reduction
Ghafoorian et al. (2017)	CNN	Lacune detection; FCN for candidate segmentation then a multi-scale 3D CNN with anatomical features as false positive reduction
Survival/disease activity/development prediction		
Kawahara et al. (2016b)	CNN	Neurodevelopment prediction; CNN with specially-designed edge-to-edge, edge-to-node and node-to-graph conv. layers for brain nets
Nie et al. (2016c)	CNN	Survival prediction; features from a Multi-modal 3D CNN is fused with hand-crafted features to train an SVM
Yoo et al. (2016)	CNN	Disease activity prediction; Training a CNN on the Euclidean distance transform of the lesion masks as the input
van der Burgh et al. (2017)	CNN	Survival prediction; DBN on MRI and fusing it with clinical characteristics and structural connectivity data
Image construction/enhancement		
Li et al. (2014)	CNN	Image construction; 3D CNN for constructing PET from MR images
Bahrami et al. (2016)	CNN	Image construction; 3D CNN for constructing 7T-like images from 3T MRI
Benou et al. (2016)	SAE	Denoising DCE-MRI; using an ensemble of denoising SAE (pretrained with RBMs)
Golkov et al. (2016)	CNN	Image construction; Per-pixel neural network to predict complex diffusion parameters based on fewer measurements
Hoffmann et al. (2016)	ANN	Image construction; Deep neural nets with SRelu nonlinearity for thermal image construction
Nie et al. (2016a)	CNN	Image construction; 3D fully convolutional network for constructing CT from MR images
Sevetlidis et al. (2016)	ANN	Image construction; Encoder-decoder network for synthesizing one MR modality from another
Other		
Brosch et al. (2014)	DBN	Manifold Learning; DBN with conv. RBM layers for modeling the variability in brain morphology and lesion distribution in MS
Cheng et al. (2015)	ANN	Similarity measurement; neural network fusing the moving and reference image patches, pretrained with SAE
Huang et al. (2016)	RBM	fMRI blind source separation; RBM for both internal and functional interaction-induced latent sources detection
Simonovsky et al. (2016)	CNN	Similarity measurement; 3D CNN estimating similarity between reference and moving images stacked in the input
Wu et al. (2013)	ISA	Correspondence detection in deformable registration; stacked convolutional ISA for unsupervised feature learning
Yang et al. (2016d)	CNN	Image registration; Conv. encoder-decoder net. predicting momentum in x and y directions, given the moving and fixed image patches

tem to generate descriptions from chest X-rays. A CNN was employed to generate a representation of an image one label at a time, which was then used to train an RNN to generate sequence of MeSH keywords. Kisilev et al. (2016) used a completely different approach and predicted categorical BI-RADS descriptors for breast lesions. In their work they focused on three descriptors used in mammography: shape, margin, and density,

Table 2: Overview of papers using deep learning techniques for retinal image analysis. All works use CNNs.

Color fundus images: segmentation of anatomical structures and quality assessment	
Fu et al. (2016b)	Blood vessel segmentation; CNN combined with CRF to model long-range pixel interactions
Fu et al. (2016a)	Blood vessel segmentation; extending the approach by Fu et al. (2016b) by reformulating CRF as RNN
Mahapatra et al. (2016)	Image quality assessment; classification output using CNN-based features combined with the output using saliency maps
Maninis et al. (2016)	Segmentation of blood vessels and optic disk; VGG-19 network extended with specialized layers for each segmentation task
Wu et al. (2016)	Blood vessel segmentation; patch-based CNN followed by mapping PCA solution of last layer feature maps to full segmentation
Zilly et al. (2017)	Segmentation of the optic disk and the optic cup; simple CNN with filters sequentially learned using boosting
Color fundus images: detection of abnormalities and diseases	
Chen et al. (2015d)	Glaucoma detection; end-to-end CNN, the input is a patch centered at the optic disk
Abràmoff et al. (2016)	Diabetic retinopathy detection; end-to-end CNN, outperforms traditional method, evaluated on a public dataset
Burlina et al. (2016)	Age-related macular degeneration detection; uses overfeat pretrained network for feature extraction
van Grinsven et al. (2016)	Hemorrhage detection; CNN dynamically trained using selective data sampling to perform hard negative mining
Gulshan et al. (2016)	Diabetic retinopathy detection; Inception network, performance comparable to a panel of seven certified ophthalmologists
Prentasic and Loncaric (2016)	Hard exudate detection; end-to-end CNN combined with the outputs of traditional classifiers for detection of landmarks
Worrall et al. (2016)	Retinopathy of prematurity detection; fine-tuned ImageNet trained GoogLeNet, feature map visualization to highlight disease
Work in other imaging modalities	
Gao et al. (2015)	Cataract classification in slit lamp images; CNN followed by a set of recursive neural networks to extract higher order features
Schlegl et al. (2015)	Fluid segmentation in OCT; weakly supervised CNN improved with semantic descriptors from clinical reports
Prentasic et al. (2016)	Blood vessel segmentation in OCT angiography; simple CNN, segmentation of several capillary networks

where each have their own class label. The system was fed with the image data and region proposals and predicts the correct label for each descriptor (e.g. for shape either oval, round, or irregular).

Given the wealth of data that is available in PACS systems in terms of images and corresponding diagnostic reports, it seems like an ideal avenue for future deep learning research. One could expect that advances in captioning natural images will in time be applied to these data sets as well.

4. Anatomical application areas

This section presents an overview of deep learning contributions to the various application areas in medical imaging. We highlight some key contributions and discuss performance of systems on large data sets and on public challenge data sets. All these challenges are listed on http:\\www.grand-challenge.org.

4.1. Brain

DNNs have been extensively used for brain image analysis in several different application domains (Table 1). A large number of studies address classification of Alzheimer's disease and segmentation of brain tissue and anatomical structures (e.g. the hippocampus). Other important areas are detection and segmentation of lesions (e.g. tumors, white matter lesions, lacunes, micro-bleeds).

Apart from the methods that aim for a scan-level classification (e.g. Alzheimer diagnosis), most methods learn mappings from local patches to representations and subsequently from representations to labels. However, the local patches might lack the contextual information required for tasks where anatomical information is paramount (e.g. white matter lesion segmentation). To tackle this, Ghafoorian et al. (2016b) used non-uniformly sampled patches by gradually lowering sampling rate in patch sides to span a larger context. An alternative strategy used by many groups is multi-scale analysis and a fusion of representations in a fully-connected layer.

Even though brain images are 3D volumes in all surveyed studies, most methods work in 2D, analyzing the 3D volumes slice-by-slice. This is often motivated by either the reduced computational requirements or the thick slices relative to in-plane resolution in some data sets. More recent publications had also employed 3D networks.

DNNs have completely taken over many brain image analysis challenges. In the 2014 and 2015 brain tumor segmentation challenges (BRATS), the 2015 longitudinal multiple sclerosis lesion segmentation challenge, the 2015 ischemic stroke lesion segmentation challenge (ISLES), and the 2013 MR brain image segmentation challenge (MRBrains), the top ranking teams to date have all used CNNs. Almost all of the aforementioned methods are concentrating on brain MR images. We expect that other brain imaging modalities such as CT and US can also benefit from deep learning based analysis.

4.2. Eye

Ophthalmic imaging has developed rapidly over the past years, but only recently are deep learning algorithms being applied to eye image understanding. As summarized in Table 2, most works employ simple

Table 3: Overview of papers using deep learning techniques for chest x-ray image analysis.

Reference	Application	Remarks
Lo et al. (1995)	Nodule detection	Classifies candidates from small patches with two-layer CNN, each with 12 5×5 filters
Anavi et al. (2015)	Image retrieval	Combines classical features with those from pre-trained CNN for image retrieval using SVM
Bar et al. (2015)	Pathology detection	Features from a pre-trained CNN and low level features are used to detect various diseases
Anavi et al. (2016)	Image retrieval	Continuation of Anavi et al. (2015), adding age and gender as features
Bar et al. (2016)	Pathology detection	Continuation of Bar et al. (2015), more experiments and adding feature selection
Cicero et al. (2016)	Pathology detection	GoogLeNet CNN detects five common abnormalities, trained and validated on a large data set
Hwang et al. (2016)	Tuberculosis detection	Processes entire radiographs with a pre-trained fine-tuned network with 6 convolution layers
Kim and Hwang (2016)	Tuberculosis detection	MIL framework produces heat map of suspicious regions via deconvolution
Shin et al. (2016a)	Pathology detection	CNN detects 17 diseases, large data set (7k images), recurrent networks produce short captions
Rajkomar et al. (2017)	Frontal/lateral classification	Pre-trained CNN performs frontal/lateral classification task
Yang et al. (2016c)	Bone suppression	Cascade of CNNs at increasing resolution learns bone images from gradients of radiographs
Wang et al. (2016a)	Nodule classification	Combines classical features with CNN features from pre-trained ImageNet CNN

Table 4: Overview of papers using deep learning techniques for chest CT image analysis.

Reference	Application; remarks
Segmentation	
Charbonnier et al. (2017)	Airway segmentation where multi-view CNN classifies candidate branches as true airways or leaks
Nodule detection and analysis	
Ciompi et al. (2015)	Used a standard feature extractor and a pre-trained CNN to classify detected lesions as benign peri-fissural nodules
van Ginneken et al. (2015)	Detects nodules with pre-trained CNN features from orthogonal patches around candidate, classified with SVM
Shen et al. (2015b)	Three CNNs at different scales estimate nodule malignancy scores of radiologists (LIDC-IDRI data set)
Chen et al. (2016e)	Combines features from CNN, SDAE and classical features to characterize nodules from LIDC-IDRI data set
Ciompi et al. (2016)	Multi-stream CNN to classify nodules into subtypes: solid, part-solid, non-solid, calcified, spiculated, perifissural
Dou et al. (2016b)	Uses 3D CNN around nodule candidates; ranks #1 in LUNA16 nodule detection challenge
Li et al. (2016a)	Detects nodules with 2D CNN that processes small patches around a nodule
Setio et al. (2016)	Detects nodules with end-to-end trained multi-stream CNN with 9 patches per candidate
Shen et al. (2016)	3D CNN classifies volume centered on nodule as benign/malignant, results are combined to patient level prediction
Sun et al. (2016b)	Same dataset as Shen et al. (2015b), compares CNN, DBN, SDAE and classical computer-aided diagnosis schemes
Teramoto et al. (2016)	Combines features extracted from 2 orthogonal CT patches and a PET patch
Interstitial lung disease	
Anthimopoulos et al. (2016)	Classification of 2D patches into interstitial lung texture classes using a standard CNN
Christodoulidis et al. (2017)	2D interstitial pattern classification with CNNs pre-trained with a variety of texture data sets
Gao et al. (2016c)	Propagates manually drawn segmentations using CNN and CRF for more accurate interstitial lung disease reference
Gao et al. (2016a)	AlexNet applied to large parts of 2D CT slices to detect presence of interstitial patterns
Gao et al. (2016b)	Uses regression to predict area covered in 2D slice with a particular interstitial pattern
Tarando et al. (2016)	Combines existing computer-aided diagnosis system and CNN to classify lung texture patterns.
van Tulder and de Bruijne (2016)	Classification of lung texture and airways using an optimal set of filters derived from DBNs and RBMs
Other applications	
Tajbakhsh et al. (2015a)	Multi-stream CNN to detect pulmonary embolism from candidates obtained from a tobogganing algorithm
Carneiro et al. (2016)	Predicts 5-year mortality from thick slice CT scans and segmentation masks
de Vos et al. (2016a)	Identifies the slice of interest and determine the distance between CT slices

CNNs for the analysis of color fundus imaging (CFI). A wide variety of applications are addressed: segmentation of anatomical structures, segmentation and detection of retinal abnormalities, diagnosis of eye diseases, and image quality assessment.

In 2015, Kaggle organized a diabetic retinopathy detection competition: Over 35,000 color fundus images were provided to train algorithms to predict the severity of disease in 53,000 test images. The majority of the 661 teams that entered the competition applied deep learning and four teams achieved performance above that of humans, all using end-to-end CNNs. Recently Gulshan et al. (2016) performed a thorough analysis of the performance of a Google Inception v3 network for diabetic retinopathy detection, showing performance comparable to a panel of seven certified ophthalmologists.

4.3. Chest

In thoracic image analysis of both radiography and computed tomography, the detection, characterization, and classification of nodules is the most commonly addressed application. Many works add features derived from deep networks to existing feature sets or compare

Table 5: Overview of papers using deep learning for digital pathology images. The staining and imaging modality abbreviations used in the table are as follows: H&E: hematoxylin and eosin staining, TIL: Tumor-infiltrating lymphocytes, BCC: Basal cell carcinoma, IHC: immunohistochemistry, RM: Romanowsky, EM: Electron microscopy, PC: Phase contrast, FL: Fluorescent, IFL: Immunofluorescent, TPM: Two-photon microscopy, CM: Confocal microscopy, Pap: Papanicolaou.

Reference	Topic	Staining\Modality	Method
Nucleus detection, segmentation, and classification			
Cireşan et al. (2013)	Mitosis detection	H&E	CNN-based pixel classifier
Cruz-Roa et al. (2013)	Detection of basal cell carcinoma	H&E	Convolutional auto-encoder neural network
Malon and Cosatto (2013)	Mitosis detection	H&E	Combines shapebased features with CNN
Wang et al. (2014)	Mitosis detection	H&E	Cascaded ensemble of CNN and handcrafted features
Ferrari et al. (2015)	Bacterial colony counting	Culture plate	CNN-based patch classifier
Ronneberger et al. (2015)	Cell segmentation	EM	U-Net with deformation augmentation
Shkolyar et al. (2015)	Mitosis detection	Live-imaging	CNN-based patch classifier
Song et al. (2015)	Segmentation of cytoplasm and nuclei	H&E	Multi-scale CNN and graph-partitioning-based method
Xie et al. (2015a)	Nucleus detection	Ki-67	CNN model that learns the voting offset vectors and voting confidence
Xie et al. (2015b)	Nucleus detection	H&E, Ki-67	CNN-based structured regression model for cell detection
Akram et al. (2016)	Cell segmentation	FL, PC, H&E	fCNN for cell bounding box proposal and CNN for segmentation
Albarqouni et al. (2016)	Mitosis detection	H&E	Incorporated 'crowd sourcing' layer into the CNN framework
Bauer et al. (2016)	Nucleus classification	IHC	CNN-based patch classifier
Chen et al. (2016b)	Mitosis detection	H&E	Deep regression network (DRN)
Gao et al. (2016e)	Nucleus classification	IFL	Classification of Hep2-cells with CNN
Han et al. (2016)	Nucleus classification	IFL	Classification of Hep2-cells with CNN
Janowczyk et al. (2016b)	Nucleus segmentation	H&E	Resolution adaptive deep hierarchical learning scheme
Kashif et al. (2016)	Nucleus detection	H&E	Combination of CNN and hand-crafted features
Mao and Yin (2016)	Mitosis detection	PC	Hierarchical CNNs for patch sequence classification
Mishra et al. (2016)	Classification of mitochondria	EM	CNN-based patch classifier
Phan et al. (2016)	Nucleus classification	FL	Classification of Hep2-cells using transfer learning (pre-trained CNN)
Romo-Bucheli et al. (2016)	Tubule nuclei detection	H&E	CNN-based classification of pre-selected candidate nuclei
Sirinukunwattana et al. (2016)	Nucleus detection and classification	H&E	CNN with spatially constrained regression
Song et al. (2017)	Cell segmentation	H&E	Multi-scale CNN
Turkki et al. (2016)	TIL detection	H&E	CNN-based classification of superpixels
Veta et al. (2016)	Nuclear area measurement	H&E	A CNN directly measures nucleus area without requiring segmentation
Wang et al. (2016d)	Subtype cell detection	H&E	Combination of two CNNs for joint cell detection and classification
Xie et al. (2016a)	Nucleus detection and cell counting	FL and H&E	Microscopy cell counting with fully convolutional regression networks
Xing et al. (2016)	Nucleus segmentation	H&E, IHC	CNN and selection-based sparse shape model
Xu et al. (2016b)	Nucleus detection	H&E	Stacked sparse auto-encoders (SSAE)
Xu and Huang (2016)	Nucleus detection	Various	General deep learning framework to detect cells in whole-slide images
Yang et al. (2016b)	Glial cell segmentation	TPM	fCNN with an iterative k-terminal cut algorithm
Yao et al. (2016)	Nucleus classification	H&E	Classifies cellular tissue into tumor, lymphocyte, and stromal
Zhao et al. (2016)	Classification of leukocytes	RM	CNN-based patch classifier
Large organ segmentation			
Ciresan et al. (2012)	Segmentation of neuronal membranes	EM	Ensemble of several CNNs with different architectures
Kainz et al. (2015)	Segmentation of colon glands	H&E	Used two CNNs to segment glands and their separating structures
Apou et al. (2016)	Detection of lobular structures in breast	IHC	Combined the outputs of a CNN and a texture classification system
BenTaieb and Hamarneh (2016)	Segmentation of colon glands	H&E	fCNN with a loss accounting for smoothness and object interactions
BenTaieb et al. (2016)	Segmentation of colon glands	H&E	A multi-loss fCNN to perform both segmentation and classification
Chen et al. (2016d)	Neuronal membrane and fungus segmentation	EM	Combination of bi-directional LSTM-RNNs and kU-Nets
Chen et al. (2017)	Segmentation of colon glands	H&E	Deep contour-aware CNN
Çiçek et al. (2016)	Segmentation of xenopus kidney	CM	3D U-Net
Drozdzal et al. (2016)	Segmentation of neuronal structures	EM	fCNN with skip connections
Li et al. (2016b)	Segmentation of colon glands	H&E	Compares CNN with an SVM using hand-crafted features
Teikari et al. (2016)	Volumetric vascular segmentation	FL	Hybrid 2D-3D CNN architecture
Wang et al. (2016c)	Segmentation of messy and muscle regions	H&E	Conditional random field jointly trained with an fCNN
Xie et al. (2016b)	Perimysium segmentation	H&E	2D spatial clockwork RNN
Xu et al. (2016d)	Segmentation of colon glands	H&E	Used three CNNs to predict gland and contour pixels
Xu et al. (2016a)	Segmenting epithelium & stroma	H&E, IHC	CNNs applied to over-segmented image regions (superpixels)
Detection and classification of disease			
Cruz-Roa et al. (2014)	Detection of invasive ductal carcinoma	H&E	CNN-based patch classifier
Xu et al. (2014)	Patch-level classification of colon cancer	H&E	Multiple instance learning framework with CNN features
Bychkov et al. (2016)	Outcome prediction of colorectal cancer	H&E	Extracted CNN features from epithelial tissue for prediction
Chang et al. (2017)	Multiple cancer tissue classification	Various	Transfer learning using multi-Scale convolutional sparse coding
Günhan Ertosun and Rubin (2015)	Grading glioma	H&E	Ensemble of CNNs
Källén et al. (2016)	Predicting Gleason score	H&E	OverFeat pre-trained network as feature extractor
Kim et al. (2016a)	Thyroid cytopathology classification	H&E, RM & Pap	Fine-tuning pre-trained AlexNet
Litjens et al. (2016)	Detection of prostate and breast cancer	H&E	fCNN-based pixel classifier
Quinn et al. (2016)	Malaria, tuberculosis and parasites detection	Light microscopy	CNN-based patch classifier
Rezaeilouyeh et al. (2016)	Gleason grading and breast cancer detection	H&E	The system incorporates shearlet features inside a CNN
Schaumberg et al. (2016)	SPOP mutation prediction of prostate cancer	H&E	Ensemble of ResNets
Wang et al. (2016b)	Metastases detection in lymph node	H&E	Ensemble of CNNs with hard negative mining
Other pathology applications			
Janowczyk et al. (2016a)	Stain normalization	H&E	Used SAE for classifying tissue and subsequent histogram matching
Janowczyk and Madabhushi (2016)	Deep learning tutorial	Various	Covers different detecting, segmentation, and classification tasks
Sethi et al. (2016)	Comparison of normalization algorithms	H&E	Presents effectiveness of stain normalization for application of CNNs

CNNs with classical machine learning approaches using handcrafted features. In chest X-ray, several groups detect multiple diseases with a single system. In CT the detection of textural patterns indicative of interstitial lung diseases is also a popular research topic.

Chest radiography is the most common radiological exam; several works use a large set of images with text reports to train systems that combine CNNs for image analysis and RNNs for text analysis. This is a branch of research we expect to see more of in the near future.

In a recent challenge for nodule detection in CT, LUNA16, CNN architectures were used by all top performing systems. This is in contrast with a previous lung nodule detection challenge, ANODE09, where handcrafted features were used to classify nodule candidates. The best systems in LUNA16 still rely on nodule candidates computed by rule-based image processing, but systems that use deep networks for candidate detection also performed very well (e.g. U-net). Estimating the probability that an individual has lung cancer from a CT scan is an important topic: It is the objective of the Kaggle Data Science Bowl 2017, with $1 million in prizes and more than one thousand participating teams.

4.4. Digital pathology and microscopy

The growing availability of large scale gigapixel whole-slide images (WSI) of tissue specimen has made digital pathology and microscopy a very popular application area for deep learning techniques. The developed techniques applied to this domain focus on three broad challenges: (1) Detecting, segmenting, or classifying nuclei, (2) segmentation of large organs, and (3) detecting and classifying the disease of interest at the lesion- or WSI-level. Table 5 presents an overview for each of these categories.

Deep learning techniques have also been applied for normalization of histopathology images. Color normalization is an important research area in histopathology image analysis. In Janowczyk et al. (2016a), a method for stain normalization of hematoxylin and eosin (H&E) stained histopathology images was presented based on deep sparse auto-encoders. Recently, the importance of color normalization was demonstrated by Sethi et al. (2016) for CNN based tissue classification in H&E stained images.

The introduction of grand challenges in digital pathology has fostered the development of computerized digital pathology techniques. The challenges that evaluated existing and new approaches for analysis of digital pathology images are: EM segmentation challenge 2012 for the 2D segmentation of neuronal processes, mitosis detection challenges in ICPR 2012 and AMIDA 2013, GLAS for gland segmentation and, CAMELYON16 and TUPAC for processing breast cancer tissue samples.

In both ICPR 2012 and the AMIDA13 challenges on mitosis detection the IDSIA team outperformed other algorithms with a CNN based approach (Cireşan et al., 2013). The same team had the highest performing system in EM 2012 (Ciresan et al., 2012) for 2D segmentation of neuronal processes. In their approach, the task of segmenting membranes of neurons was performed by mild smoothing and thresholding of the output of a CNN, which computes pixel probabilities.

GLAS addressed the problem of gland instance segmentation in colorectal cancer tissue samples. Xu et al. (2016d) achieved the highest rank using three CNN models. The first CNN classifies pixels as gland versus non-gland. From each feature map of the first CNN, edge information is extracted using the holistically nested edge technique, which uses side convolutions to produce an edge map. Finally, a third CNN merges gland and edge maps to produce the final segmentation.

CAMELYON16 was the first challenge to provide participants with WSIs. Contrary to other medical imaging applications, the availability of large amount of annotated data in this challenge allowed for training very deep models such as 22-layer GoogLeNet (Szegedy et al., 2014), 16-layer VGG-Net (Simonyan and Zisserman, 2014), and 101-layer ResNet (He et al., 2015). The top-five performing systems used one of these architectures. The best performing solution in the Camelyon16 challenge was presented in Wang et al. (2016b). This method is based on an ensemble of two GoogLeNet architectures, one trained with and one without hard-negative mining to tackle the challenge. The latest submission of this team using the WSI standardization algorithm by Ehteshami Bejnordi et al. (2016) achieved an AUC of 0.9935, for task 2, which outperformed the AUC of a pathologist (AUC = 0.966) who independently scored the complete test set.

The recently held TUPAC challenge addressed detection of mitosis in breast cancer tissue, and prediction of tumor grading at the WSI level. The top performing system by Paeng et al. (2016) achieved the highest performance in all tasks. The method has three main components: (1) Finding high cell density regions, (2) using a CNN to detect mitoses in the regions of interest, (3) converting the results of mitosis detection to a feature vector for each WSI and using an SVM classifier to compute the tumor proliferation and molecular data scores.

Table 6: Overview of papers using deep learning techniques for breast image analysis. MG = mammography; TS = tomosynthesis; US = ultrasound; ADN = Adaptive Deconvolution Network.

Reference	Modality	Method	Application; remarks
Sahiner et al. (1996)	MG	CNN	First application of a CNN to mammography
Jamieson et al. (2012)	MG, US	ADN	Four layer ADN, an early form of CNN for mass classification
Fonseca et al. (2015)	MG	CNN	Pre-trained network extracted features classified with SVM for breast density estimation
Akselrod-Ballin et al. (2016)	MG	CNN	Use a modified region proposal CNN (R-CNN) for the localization and classification of masses
Arevalo et al. (2016)	MG	CNN	Lesion classification, combination with hand-crafted features gave the best performance
Dalmis et al. (2017)	MRI	CNN	Breast and fibroglandular tissue segmentation
Dubrovina et al. (2016)	MG	CNN	Tissue classification using regular CNNs
Dhungel et al. (2016)	MG	CNN	Combination of different CNNs combined with hand-crafted features
Fotin et al. (2016)	TS	CNN	Improved state-of-the art for mass detection in tomosynthesis
Hwang and Kim (2016)	MG	CNN	Weakly supervised CNN for localization of masses
Huynh et al. (2016)	MG	CNN	Pre-trained CNN on natural image patches applied to mass classification
Kallenberg et al. (2016)	MG	SAE	Unsupervised CNN feature learning with SAE for breast density classification
Kisilev et al. (2016)	MG	CNN	R-CNN combined with multi-class loss trained on semantic descriptions of potential masses
Kooi et al. (2016)	MG	CNN	Improved the state-of-the art for mass detection and show human performance on a patch level
Qiu et al. (2016)	MG	CNN	CNN for direct classification of future risk of developing cancer based on negative mammograms
Samala et al. (2016a)	TS	CNN	Microcalcification detection
Samala et al. (2016b)	TS	CNN	Pre-trained CNN on mammographic masses transfered to tomosynthesis
Sun et al. (2016a)	MG	CNN	Semi-supervised CNN for classification of masses
Zhang et al. (2016c)	US	RBM	Classification benign vs. malignant with shear wave elastography
Kooi et al. (2017)	MG	CNN	Pre-trained CNN on mass/normal patches to discriminate malignant masses from (benign) cysts
Wang et al. (2017)	MG	CNN	Detection of cardiovascular disease based on vessel calcification

Table 7: Overview of papers using deep learning techniques for cardiac image analysis.

Reference	Modality	Method	Application; remarks
Emad et al. (2015)	MRI	CNN	Left ventricle slice detection; simple CNN indicates if structure is present
Avendi et al. (2016)	MRI	CNN	Left ventricle segmentation; AE used to initialize filters because training data set was small
Kong et al. (2016)	MRI	RNN	Identification of end-diastole and end-systole frames from cardiac sequences
Oktay et al. (2016)	MRI	CNN	Super-resolution; U-net/ResNet hybrid, compares favorably with standard superresolution methods
Poudel et al. (2016)	MRI	RNN	Left ventricle segmentation; RNN processes stack of slices, evaluated on several public datasets
Rupprecht et al. (2016)	MRI	CNN	Cardiac structure segmentation; patch-based CNNs integrated in active contour framework
Tran (2016)	MRI	CNN	Left and right ventricle segmentation; 2D fCNN architecture, evaluated on several public data sets
Yang et al. (2016a)	MRI	CNN	Left ventricle segmentation; CNN combined with multi-atlas segmentation
Zhang et al. (2016b)	MRI	CNN	Identifying presence of apex and base slices in cardiac exam for quality assessment
Ngo et al. (2017)	MRI	DBN	Left ventricle segmentation; DBN is used to initialize a level set framework
Carneiro et al. (2012)	US	DBN	Left ventricle segmentation; DBN embedded in system using landmarks and non-rigid registration
Carneiro and Nascimento (2013)	US	DBN	Left ventricle tracking; extension of Carneiro et al. (2012) for tracking
Chen et al. (2016c)	US	CNN	Structure segmentation in 5 different 2D views; uses transfer learning
Ghesu et al. (2016b)	US	CNN	3D aortic valve detection and segmentation; uses shallow and deeper sparse networks
Nascimento and Carneiro (2016)	US	DBN	Left ventricle segmentation; DBN applied to patches steers multi-atlas segmentation process
Moradi et al. (2016a)	US	CNN	Automatic generation of text descriptions for Doppler US images of cardiac valves using doc2vec
Gülsün et al. (2016)	CT	CNN	Coronary centerline extraction; CNN classifies paths as correct or leakages
Lessmann et al. (2016)	CT	CNN	Coronary calcium detection in low dose ungated CT using multi-stream CNN (3 views)
Moradi et al. (2016b)	CT	CNN	Labeling of 2D slices from cardiac CT exams; comparison with handcrafted features
de Vos et al. (2016b)	CT	CNN	Detect bounding boxes by slice classification and combining 3 orthogonal 2D CNNs
Wolterink et al. (2016)	CT	CNN	Coronary calcium detection in gated CTA; compares 3D CNN with multi-stream 2D CNNs
Zreik et al. (2016)	CT	CNN	Left ventricle segmentation; multi-stream CNN (3 views) voxel classification

4.5. Breast

One of the earliest DNN applications from Sahiner et al. (1996) was on breast imaging. Recently, interest has returned which resulted in significant advances over the state of the art, achieving the performance of human readers on ROIs (Kooi et al., 2016). Since most breast imaging techniques are two dimensional, methods successful in natural images can easily be transferred. With one exception, the only task addressed is the detection of breast cancer; this consisted of three subtasks: (1) detection and classification of mass-like lesions, (2) detection and classification of micro-calcifications, and (3) breast cancer risk scoring of images. Mammography is by far the most common modality and has consequently enjoyed the most attention. Work on tomosynthesis, US, and shear wave elastography is still scarce, and we have only one paper that analyzed breast MRI with deep learning; these other modalities will likely receive more

attention in the next few years. Table 6 summarizes the literature and main messages.

Since many countries have screening initiatives for breast cancer, there should be massive amounts of data available, especially for mammography, and therefore enough opportunities for deep models to flourish. Unfortunately, large public digital databases are unavailable and consequently older scanned screen-film data sets are still in use. Challenges such as the recently launched DREAM challenge have not yet had the desired success.

As a result, many papers used small data sets resulting in mixed performance. Several projects have addressed this issue by exploring semi-supervised learning (Sun et al., 2016a), weakly supervised learning (Hwang and Kim, 2016), and transfer learning (Kooi et al., 2017; Samala et al., 2016b)). Another method combines deep models with handcrafted features (Dhungel et al., 2016), which have been shown to be complementary still, even for very big data sets (Kooi et al., 2016). State of the art techniques for mass-like lesion detection and classification tend to follow a two-stage pipeline with a candidate detector; this design reduces the image to a set of potentially malignant lesions, which are fed to a deep CNN (Fotin et al., 2016; Kooi et al., 2016). Alternatives use a region proposal network (R-CNN) that bypasses the cascaded approach (Akselrod-Ballin et al., 2016; Kisilev et al., 2016).

When large data sets are available, good results can be obtained. At the SPIE Medical Imaging conference of 2016, a researcher from a leading company in the mammography CAD field told a packed conference room how a few weeks of experiments with a standard architecture (AlexNet) - trained on the company's proprietary database - yielded a performance that was superior to what years of engineering handcrafted feature systems had achieved (Fotin et al., 2016).

4.6. Cardiac

Deep learning has been applied to many aspects of cardiac image analysis; the literature is summarized in Table 7. MRI is the most researched modality and left ventricle segmentation the most common task, but the number of applications is highly diverse: segmentation, tracking, slice classification, image quality assessment, automated calcium scoring and coronary centerline tracking, and super-resolution.

Most papers used simple 2D CNNs and analyzed the 3D and often 4D data slice by slice; the exception is Wolterink et al. (2016) where 3D CNNs were used. DBNs are used in four papers, but these all originated from the same author group. The DBNs are only used for feature extraction and are integrated in compound segmentation frameworks. Two papers are exceptional because they combined CNNs with RNNs: Poudel et al. (2016) introduced a recurrent connection within the U-net architecture to segment the left ventricle slice by slice and learn what information to remember from the previous slices when segmenting the next one. Kong et al. (2016) used an architecture with a standard 2D CNN and an LSTM to perform temporal regression to identify specific frames and a cardiac sequence. Many papers use publicly available data. The largest challenge in this field was the 2015 Kaggle Data Science Bowl where the goal was to automatically measure end-systolic and end-diastolic volumes in cardiac MRI. 192 teams competed for $200,000 in prize money and the top ranking teams all used deep learning, in particular fCNN or U-net segmentation schemes.

4.7. Abdomen

Most papers on the abdomen aimed to localize and segment organs, mainly the liver, kidneys, bladder, and pancreas (Table 8). Two papers address liver tumor segmentation. The main modality is MRI for prostate analysis and CT for all other organs. The colon is the only area where various applications were addressed, but always in a straightforward manner: A CNN was used as a feature extractor and these features were used for classification.

It is interesting to note that in two segmentation challenges - SLIVER07 for liver and PROMISE12 for prostate - more traditional image analysis methods were dominant up until 2016. In PROMISE12, the current second and third in rank among the automatic methods used active appearance models. The algorithm from IMorphics was ranked first for almost five years (now ranked second). However, a 3D fCNN similar to U-net (Yu et al., 2017) has recently taken the top position. This paper has an interesting approach where a sum-operation was used instead of the concatenation operation used in U-net, making it a hybrid between a ResNet and U-net architecture. Also in SLIVER07 - a 10-year-old liver segmentation challenge - CNNs have started to appear in 2016 at the top of the leaderboard, replacing previously dominant methods focused on shape and appearance modeling.

4.8. Musculoskeletal

Musculoskeletal images have also been analyzed by deep learning algorithms for segmentation and identification of bone, joint, and associated soft tissue abnormalities in diverse imaging modalities. The works are summarized in Table 9.

Table 8: Overview of papers using deep learning for abdominal image analysis.

Reference	Topic	Modality	Method	Remarks
Multiple				
Hu et al. (2016a)	Segmentation	CT	CNN	3D CNN with time-implicit level sets for segmentation of liver, spleen and kidneys
Segmentation tasks in liver imaging				
Li et al. (2015)	Lesion	CT	CNN	2D 17×17 patch-based classification, Ben-Cohen et al. (2016) repeats this approach
Vivanti et al. (2015)	Lesion	CT	CNN	2D CNN for liver tumor segmentation in follow-up CT taking baseline CT as input
Ben-Cohen et al. (2016)	Liver	CT	CNN	2D CNN similar to U-net, but without cross-connections; good results on SLIVER07
Christ et al. (2016)	Liver & tumor	CT	CNN	U-net, cascaded fCNN and dense 3D CRF
Dou et al. (2016a)	Liver	CT	CNN	3D CNN with conditional random field; good results on SLIVER07
Hoogi et al. (2016)	Lesion	CT/MRI	CNN	2D CNN obtained probabilities are used to drive active contour model
Hu et al. (2016b)	Liver	CT	CNN	3D CNN with surface evolution of a shape prior; good results on SLIVER07
Lu et al. (2017)	Liver	CT	CNN	3D CNN, competitive results on SLIVER07
Kidneys				
Lu et al. (2016)	Localization	CT	CNN	Combines local patch and slice based CNN
Ravishankar et al. (2016b)	Localization	US	CNN	Combines CNN with classical features to detect regions around kidneys
Thong et al. (2016)	Segmentation	CT	CNN	2D CCN with 43×43 patches, tested on 20 scans
Pancreas segmentation in CT				
Farag et al. (2015)	Segmentation	CT	CNN	Approach with elements similar to Roth et al. (2015b)
Roth et al. (2015b)	Segmentation	CT	CNN	Orthogonal patches from superpixel regions are fed into CNNs in three different ways
Cai et al. (2016a)	Segmentation	CT	CNN	2 CNNs detect inside and boundary of organ, initializes conditional random field
Roth et al. (2016a)	Segmentation	CT	CNN	2 CNNs detect inside and boundary of pancreas, combined with random forests
Colon				
Tajbakhsh et al. (2015b)	Polyp detection	Colonoscopy	CNN	CNN computes additional features, improving existing scheme
Liu et al. (2016a)	Colitis detection	CT	CNN	Pre-trained ImageNet CNN generates features for linear SVM
Nappi et al. (2016)	Polyp detection	CT	CNN	Substantial reduction of false positives using pre-trained and fine-tuned CNN
Tachibana et al. (2016)	Electronic cleansing	CT	CNN	Voxel classification in dual energy CT, material other than soft tissue is removed
Zhang et al. (2017)	Polyp detection	Colonoscopy	CNN	Pre-trained ImageNet CNN for feature extraction, two SVMs for cascaded classification
Prostate segmentation in MRI				
Liao et al. (2013)	Application of stacked independent subspace analysis networks			
Cheng et al. (2016b)	CNN produces energy map for 2D slice based active appearance segmentation			
Guo et al. (2016)	Stacked sparse auto-encoders extract features from patches, input to atlas matching and a deformable model			
Milletari et al. (2016b)	3D U-net based CNN architecture with objective function that directly optimizes Dice coefficient, ranks #5 in PROMISE12			
Yu et al. (2017)	3D fully convolutional network, hybrid between a ResNet and U-net architecture, ranks #1 on PROMISE12			
Prostate				
Azizi et al. (2016))	Lesion classification	US	DBN	DBN learns features from temporal US to classify prostate lesions benign/malignant
Shah et al. (2016)	CBIR	MRI	CNN	Features from pre-trained CNN combined with features from hashing forest
Zhu et al. (2017)	Lesion classification	MRI	SAE	Learns features from multiple modalities, hierarchical random forest for classification
Bladder				
Cha et al. (2016)	Segmentation	CT	CNN	CNN patch classification used as initialization for level set

Table 9: Overview of papers using deep learning for musculoskeletal image analysis.

Reference	Modality	Application; remarks
Prasoon et al. (2013)	MRI	Knee cartilage segmentation using multi-stream CNNs
Chen et al. (2015c)	CT	Vertebrae localization; joint learning of vertebrae appearance and dependency on neighbors using CNN
Roth et al. (2015c)	CT	Sclerotic metastases detection; random 2D views are analyzed by CNN and aggregated
Shen et al. (2015a)	CT	Vertebrae localization and segmentation; CNN for segmenting vertebrae and for center detection
Suzani et al. (2015)	MRI	Vertebrae localization, identification and segmentation of vertebrae; CNN used for initial localization
Yang et al. (2015)	MRI	Anatomical landmark detection; uses CNN for slice classification for presence of landmark
Antony et al. (2016)	X-ray	Osteoarthritis grading; pre-trained ImageNet CNN fine-tuned on knee X-rays
Cai et al. (2016b)	CT, MRI	Vertebrae localization; RBM determines position, orientation and label of vertebrae
Golan et al. (2016)	US	Hip dysplasia detection; CNN with adversarial component detects structures and performs measurements
Korez et al. (2016)	MRI	Vertebral bodies segmentation; voxel probabilities obtained with a 3D CNN are input to deformable model
Jamaludin et al. (2016)	MRI	Automatic spine scoring; VGG-19 CNN analyzes vertebral discs and finds lesion hotspots
Miao et al. (2016)	X-ray	Total Knee Arthroplasty kinematics by real-time 2D/3D registration using CNN
Roth et al. (2016c)	CT	Posterior-element fractures detection; CNN for 2.5D patch-based analysis
Štern et al. (2016)	MRI	Hand age estimation; 2D regression CNN analyzes 13 bones
Forsberg et al. (2017)	MRI	Vertebrae detection and labeling; outputs of two CNNs are input to graphical model
Spampinato et al. (2017)	X-ray	Skeletal bone age assessment; comparison among several deep learning approaches for the task at hand

A surprising number of complete applications with promising results are available; one that stands out is Jamaludin et al. (2016) who trained their system with 12K discs and claimed near-human performances across four different radiological scoring tasks.

4.9. Other

This final section lists papers that address multiple applications (Table 10) and a variety of other applications (Table 11).

It is remarkable that one single architecture or approach based on deep learning can be applied without modifications to different tasks; this illustrates the versatility of deep learning and its general applicability. In some works, pre-trained architectures are used, sometimes trained with images from a completely different domain. Several authors analyze the effect of fine-tuning a network by training it with a small data set of images from the intended application domain. Combining features extracted by a CNN with 'traditional' features is also commonly seen.

From Table 11, the large number of papers that address obstetric applications stand out. Most papers address the groundwork, such as selecting an appropriate frame from an US stream. More work on automated measurements with deep learning in these US sequences is likely to follow.

The second area where CNNs are rapidly improving the state of the art is dermoscopic image analysis. For a long time, diagnosing skin cancer from photographs was considered very difficult and out of reach for computers. Many studies focused only on images obtained with specialized cameras, and recent systems based on deep networks produced promising results. A recent work by Esteva et al. (2017) demonstrated excellent results with training a recent standard architecture (Google's Inception v3) on a data set of both dermoscopic and standard photographic images. This data set was two orders of magnitude larger than what was used in literature before. In a thorough evaluation, the proposed system performed on par with 30 board certified dermatologists.

5. Discussion

Overview

From the 308 papers reviewed in this survey, it is evident that deep learning has pervaded every aspect of medical image analysis. This has happened extremely quickly: the vast majority of contributions, 242 papers, were published in 2016 or the first month of 2017. A large diversity of deep architectures are covered. The earliest studies used pre-trained CNNs as feature extractors. The fact that these pre-trained networks could simply be downloaded and directly applied to any medical image facilitated their use. Moreover, in this approach already existing systems based on handcrafted features could simply be extended. In the last two years, however, we have seen that end-to-end trained CNNs have become the preferred approach for medical imaging interpretation (see Figure 1). Such CNNs are often integrated into existing image analysis pipelines and replace traditional handcrafted machine learning methods. This is the approach followed by the largest group of papers in this survey and we can confidently state that this is the current standard practice.

Key aspects of successful deep learning methods

After reviewing so many papers one would expect to be able to distill the perfect deep learning method and architecture for each individual task and application area. Although convolutional neural networks (and derivatives) are now clearly the top performers in most medical image analysis competitions, one striking conclusion we can draw is that the exact architecture is not the most important determinant in getting a good solution. We have seen, for example in challenges like the Kaggle Diabetic Retinopathy Challenge, that many researchers use the exact same architectures, the same type of networks, but have widely varying results. A key aspect that is often overlooked is that expert knowledge about the task to be solved can provide advantages that go beyond adding more layers to a CNN. Groups and researchers that obtain good performance when applying deep learning algorithms often differentiate themselves in aspects outside of the deep network, like novel data preprocessing or augmentation techniques. An example is that the best performing method in the CAMELYON16-challenge improved significantly (AUC from 0.92 to 0.99) by adding a stain normalization pre-processing step to improve generalization without changing the CNN. Other papers focus on data augmentation strategies to make networks more robust, and they report that these strategies are essential to obtain good performance. An example is the elastic deformations that were applied in the original U-Net paper (Ronneberger et al., 2015).

Augmentation and pre-processing are, of course, not the only key contributors to good solutions. Several researchers have shown that designing architectures incorporating unique task-specific properties can obtain better results than straightforward CNNs. Two examples which we encountered several times are multi-view

Table 10: Overview of papers using a single deep learning approach for different tasks. DQN = Deep Q-Network

Reference	Task	Modality	Method	Remarks
Shin et al. (2013)	Heart, kidney, liver segmentation	MRI	SAE	SAE to learn temporal/spatial features on 2D + time DCE-MRI
Roth et al. (2015a)	2D slice classification	CT	CNN	Automatically classifying slices in 5 anatomical regions
Shin et al. (2015)	2D key image labeling	CT, MRI	CNN	Text and 2D image analysis on a diverse set of 780 thousand images
Cheng et al. (2016a)	Various detection tasks	US, CT	AE, CNN	Detection of breast lesions in US and pulmonary nodules in CT
Ghesu et al. (2016a)	Landmark detection	US, CT, MRI	CNN, DQN	Reinforcement learning with CNN features, cardiac MR/US, head&neck CT
Liu et al. (2016b)	Image retrieval	X-ray	CNN	Combines CNN feature with Radon transform, evaluated on IRMA database
Merkow et al. (2016)	Vascular network segmentation	CT, MRI	CNN	Framework to find various vascular networks
Moeskops et al. (2016b)	Various segmentation tasks	MRI, CT	CNN	Single architecture to segment 6 brain tissues, pectoral muscle & coronaries
Roth et al. (2016b)	Various detection tasks	CT	CNN	Multi-stream CNN to detect sclerotic lesions, lymph nodes and polyps
Shin et al. (2016b)	Abnormality detection	CT	CNN	Compares architectures for detecting interstitial disease and lymph nodes
Tajbakhsh et al. (2016)	Abnormality detection	CT, US	CNN	Compares pre-trained with fully trained networks for three detection tasks
Wang et al. (2016e)	2D key image labeling	CT, MRI	CNN	Text concept clustering, related to Shin et al. (2015)
Yan et al. (2016)	2D slice classification	CT	CNN	Automatically classifying CT slices in 12 anatomical regions
Zhou et al. (2016)	Thorax-abdomen segmentation	CT	CNN	21 structures are segmented with 3 orthogonal 2D fCNNs and majority voting

Table 11: Overview of papers using deep learning for various image analysis tasks.

Reference	Task	Modality	Method	Remarks
Fetal imaging				
Chen et al. (2015b)	Frame labeling	US	CNN	Locates abdominal plane from fetal ultrasound videos
Chen et al. (2015a)	Frame labeling	US	RNN	Same task as Chen et al. (2015b), now using RNNs
Baumgartner et al. (2016)	Frame labeling	US	CNN	Labeling 12 standard frames in 1003 mid pregnancy fetal US videos
Gao et al. (2016d)	Frame labeling	US	CNN	4 class frame classification using transfer learning with pre-trained networks
Kumar et al. (2016)	Frame labeling	US	CNN	12 standard anatomical planes, CNN extracts features for support vector machine
Rajchl et al. (2016b)	Segmentation with non expert labels	MRI	CNN	Crowd-sourcing annotation efforts to segment brain structures
Rajchl et al. (2016a)	Segmentation given bounding box	MRI	CNN	CNN and CRF for segmentation of structures
Ravishankar et al. (2016a)	Quantification	US	CNN	Hybrid system using CNN and texture features to find abdominal circumference
Yu et al. (2016b)	Left ventricle segmentation	US	CNN	Frame-by-frame segmentation by dynamically fine-tuning CNN to the latest frame
Dermatology				
Codella et al. (2015)	Melanoma detection in dermoscopic images		CNN	Features from pre-trained CNN combined with other features
Demyanov et al. (2016)	Pattern identification in dermoscopic images		CNN	Comparison to simpler networks and simple machine learning
Kawahara et al. (2016a)	5 and 10-class classification photographic images		CNN	Pre-trained CNN for feature extraction at two image resolutions
Kawahara and Hamarneh (2016)	10-class classification photographic images		CNN	Extending Kawahara et al. (2016a) now training multi-resolution CNN end-to-end
Yu et al. (2016a)	Melanoma detection in dermoscopic images		CNN	Deep residual networks for lesion segmentation and classification, winner ISIC16
Menegola et al. (2016)	Classification of dermoscopic images		CNN	Various pre-training and fine-tuning strategies are compared
Esteva et al. (2017)	Classification of photographic and dermoscopic images		CNN	Inception CNN trained on 129k images; compares favorably to 29 dermatologists
Lymph nodes				
Roth et al. (2014)	Lymph node detection	CT	CNN	Introduces multi-stream framework of 2D CNNs with orthogonal patches
Barbu et al. (2016)	Lymph node detection	CT	CNN	Compares effect of different loss functions
Nogues et al. (2016)	Lymph node detection	CT	CNN	2 fCNNs, for inside and for contour of lymph nodes, are combined in a CRF
Other				
Wang et al. (2015)	Wound segmentation	photographs	CNN	Additional detection of infection risk and healing progress
Ypsilantis et al. (2015)	Chemotherapy response prediction	PET	CNN	CNN outperforms classical radiomics features in patients with esophageal cancer
Zheng et al. (2015)	Carotid artery bifurcation detection	CT	CNN	Two stage detection process, CNNs combined with Haar features
Alansary et al. (2016)	Placenta segmentation	MRI	CNN	3D multi-stream CNN with extension for motion correction
Fritscher et al. (2016)	Head&Neck tumor segmentation	CT	CNN	3 orthogonal patches in 2D CNNs, combined with other features
Jaumard-Hakoun et al. (2016)	Tongue contour extraction	US	RBM	Analysis of tongue motion during speech, combines auto-encoders with RBMs
Payer et al. (2016)	Hand landmark detection	X-ray	CNN	Various architectures are compared
Quinn et al. (2016)	Disease detection	microscopy	CNN	Smartphone mounted on microscope detects malaria, tuberculosis & parasite eggs
Smistad and Løvstakken (2016)	Vessel detection and segmentation	US	CNN	Femoral and carotid vessels analyzed with standard fCNN
Twinanda et al. (2017)	Task recognition in laparoscopy	Videos	CNN	Fine-tuned AlexNet applied to video frames
Xu et al. (2016c)	Cervical dysplasia	cervigrams	CNN	Fine-tuned pre-trained network with added non-imaging features
Xue et al. (2016)	Esophageal microvessel classification	Microscopy	CNN	Simple CNN used for feature extraction
Zhang et al. (2016a)	Image reconstruction	CT	CNN	Reconstructing from limited angle measurements, reducing reconstruction artefacts
Lekadir et al. (2017)	Carotid plaque classification	US	CNN	Simple CNN for characterization of carotid plaque composition in ultrasound
Ma et al. (2017)	Thyroid nodule detection	US	CNN	CNN and standard features combines for 2D US analysis

and multi-scale networks. Other, often underestimated, parts of network design are the network input size and receptive field (i.e. the area in input space that contributes to a single output unit). Input sizes should be selected considering for example the required resolution and context to solve a problem. One might increase the size of the patch to obtain more context, but without changing the receptive field of the network this might not be beneficial. As a standard sanity check researchers could perform the same task themselves via visual assessment of the network input. If they, or domain experts, cannot achieve good performance, the chance that you need to modify your network input or architecture is high.

The last aspect we want to touch on is model hyperparameter optimization (e.g. learning rate, dropout rate), which can help squeeze out extra performance from a network. We believe this is of secondary im-

portance with respect to performance to the previously discussed topics and training data quality. Disappointingly, no clear recipe can be given to obtain the best set of hyper-parameters as it is a highly empirical exercise. Most researchers fall back to an intuition-based random search (Bergstra and Bengio, 2012), which often seems to work well enough. Some basic tips have been covered before by Bengio (2012). Researchers have also looked at Bayesian methods for hyper-parameter optimization (Snoek et al., 2012), but this has not been applied in medical image analysis as far as we are aware of.

Unique challenges in medical image analysis

It is clear that applying deep learning algorithms to medical image analysis presents several unique challenges. The lack of large training data sets is often mentioned as an obstacle. However, this notion is only partially correct. The use of PACS systems in radiology has been routine in most western hospitals for at least a decade and these are filled with millions of images. There are few other domains where this magnitude of imaging data, acquired for specific purposes, are digitally available in well-structured archives. PACS-like systems are not as broadly used for other specialties in medicine, like ophthalmology and pathology, but this is changing as imaging becomes more prevalent across disciplines. We are also seeing that increasingly large public data sets are made available: Esteva et al. (2017) used 18 public data sets and more than 10^5 training images; in the Kaggle diabetic retinopathy competition a similar number of retinal images were released; and several chest x-ray studies used more than 10^4 images.

The main challenge is thus not the availability of image data itself, but the acquisition of relevant annotations/labeling for these images. Traditionally PACS systems store free-text reports by radiologists describing their findings. Turning these reports into accurate annotations or structured labels in an automated manner requires sophisticated text-mining methods, which is an important field of study in itself where deep learning is also widely used nowadays. With the introduction of structured reporting into several areas of medicine, extracting labels to data is expected to become easier in the future. For example, there are already papers appearing which directly leverage BI-RADS categorizations by radiologist to train deep networks (Kisilev et al., 2016) or semantic descriptions in analyzing optical coherence tomography images (Schlegl et al., 2015). We expect the amount of research in optimally leveraging free-text and structured reports for network training to increase in the near future.

Given the complexity of leveraging free-text reports from PACS or similar systems to train algorithms, generally researchers request domain experts (e.g. radiologist, pathologists) to make task-specific annotations for the image data. Labeling a sufficiently large dataset can take a significant amount of time, and this is problematic. For example, to train deep learning systems for segmentation in radiology often 3D, slice-by-slice annotations need to be made and this is very time consuming. Thus, learning efficiently from limited data is an important area of research in medical image analysis. A recent paper focused on training a deep learning segmentation system for 3D segmentation using only sparse 2D segmentations (Çiçek et al., 2016). Multiple-instance or active learning approaches might also offer benefit in some cases, and have recently been pursued in the context of deep learning (Yan et al., 2016). One can also consider leveraging non-expert labels via crowd-sourcing (Rajchl et al., 2016b). Other potential solutions can be found within the medical field itself; in histopathology one can sometimes use specific immunohistochemical stains to highlight regions of interest, reducing the need for expert experience (Turkki et al., 2016).

Even when data is annotated by domain expert, label noise can be a significant limiting factor in developing algorithms, whereas in computer vision the noise in the labeling of images is typically relatively low. To give an example, a widely used dataset for evaluating image analysis algorithms to detect nodules in lung CT is the LIDC-IDRI dataset (Armato et al., 2011). In this dataset pulmonary nodules were annotated by four radiologists independently. Subsequently the readers reviewed each others annotations but no consensus was forced. It turned out that the number of nodules they did not unanimously agreed on to be a nodule, was three times larger than the number they did fully agree on. Training a deep learning system on such data requires careful consideration of how to deal with noise and uncertainty in the reference standard. One could think of solutions like incorporating labeling uncertainty directly in the loss function, but this is still an open challenge.

In medical imaging often classification or segmentation is presented as a binary task: normal versus abnormal, object versus background. However, this is often a gross simplification as both classes can be highly heterogeneous. For example, the normal category often consists of completely normal tissue but also several categories of benign findings, which can be rare, and may occasionally include a wide variety of imaging artifacts. This often leads to systems that are ex-

tremely good at excluding the most common normal subclasses, but fail miserably on several rare ones. A straightforward solution would be to turn the deep learning system in a multi-class system by providing it with detailed annotations of all possible subclasses. Obviously this again compounds the issue of limited availability of expert time for annotating and is therefore often simply not feasible. Some researchers have specifically looked into tackling this imbalance by incorporating intelligence in the training process itself, by applying selective sampling (van Grinsven et al., 2016) or hard negative mining (Wang et al., 2016b). However, such strategies typically fail when there is substantial noise in the reference standard. Additional methods for dealing with within-class heterogeneity would be highly welcome.

Another data-related challenge is class imbalance. In medical imaging, images for the abnormal class might be challenging to find, depending on the task at hand. As an example, the implementation of breast cancer screening programs has resulted in vast databases of mammograms that have been established at many locations world-wide. However, the majority of these images are normal and do not contain any suspicious lesions. When a mammogram does contain a suspicious lesion this is often not cancerous, and even most cancerous lesions will not lead to the death of a patient. Designing deep learning systems that are adept at handling this class imbalance is another important area of research. A typical strategy we encountered in current literature is the application of specific data augmentation algorithms to just the underrepresented class, for example scaling and rotation transforms to generate new lesions. Pereira et al. (2016) performed a thorough evaluation of data augmentation strategies for brain lesion segmentation to combat class imbalance.

In medical image analysis useful information is not just contained within the images themselves. Physicians often leverage a wealth of data on patient history, age, demographics and others to arrive at better decisions. Some authors have already investigated combining this information into deep learning networks in a straightforward manner (Kooi et al., 2017). However, as these authors note, the improvements that were obtained were not as large as expected. One of the challenges is to balance the number of imaging features in the deep learning network (typically thousands) with the number of clinical features (typically only a handful) to prevent the clinical features from being drowned out. Physicians often also need to use anatomical information to come to an accurate diagnosis. However, many deep learning systems in medical imaging are still based on patch classification, where the anatomical location of the patch is often unknown to network. One solution would be to feed the entire image to the deep network and use a different type of evaluation to drive learning, as was done by, for example, Milletari et al. (2016b), who designed a loss function based on the Dice coefficient. This also takes advantage of the fact that medical images are often acquired using a relatively static protocol, where the anatomy is always roughly in the same position and at the same scale. However, as mentioned above, if the receptive field of the network is small feeding in the entire image offers no benefit. Furthermore, feeding full images to the network is not always feasible due to, for example, memory constraints. In some cases this might be solved in the near future due to advances in GPU technology, but in others, for example digital pathology with its gigapixel-sized images, other strategies have to be invented.

Outlook

Although most of the challenges mentioned above have not been adequately tackled yet, several high-profile successes of deep learning in medical imaging have been reported, such as the work by Esteva et al. (2017) and Gulshan et al. (2016) in the fields of dermatology and ophthalmology. Both papers show that it is possible to outperform medical experts in certain tasks using deep learning for image classification. However, we feel it is important to put these papers into context relative to medical image analysis in general, as most tasks can by no means be considered 'solved'. One aspect to consider is that both Esteva et al. (2017) and Gulshan et al. (2016) focus on small 2D color image classification, which is relatively similar to the tasks that have been tackled in computer vision (e.g. ImageNet). This allows them to take advantage of well-explored network architectures like ResNet and VGG-Net which have shown to have excellent results in these tasks. However, there is no guarantee that these architectures are optimal in for example regressions/detection tasks. It also allowed the authors to use networks that were pre-trained on a very well-labeled dataset of millions of natural images, which helps combat the lack of similarly large, labeled medical datasets. In contrast, in most medical imaging tasks 3D gray-scale or multi-channel images are used for which pre-trained networks or architectures dont exist. In addition this data typically has very specific challenges, like anisotropic voxel sizes, small registration errors between varying channels (e.g. in multi-parametric MRI) or varying intensity ranges. Although many tasks in medical image analysis can be postulated as a classification problem, this might not always be the

optimal strategy as it typically requires some form of post-processing with non-deep learning methods (e.g. counting, segmentation or regression tasks). An interesting example is the paper by Sirinukunwattana et al. (2016), which details a method directly predicting the center locations of nuclei and shows that this outperforms classification-based center localization. Nonetheless, the papers by Esteva et al. (2017) and Gulshan et al. (2016) do show what ideally is possible with deep learning methods that are well-engineered for specific medical image analysis tasks.

Looking at current trends in the machine learning community with respect to deep learning, we identify a key area which can be highly relevant for medical imaging and is receiving (renewed) interest: unsupervised learning. The renaissance of neural networks started around 2006 with the popularization of greedy layer-wise pre-training of neural networks in an unsupervised manner. This was quickly superseded by fully supervised methods which became the standard after the success of AlexNet during the ImageNet competition of 2012, and most papers in this survey follow a supervised approach. However, interest in unsupervised training strategies has remained and recently has regained traction.

Unsupervised methods are attractive as they allow (initial) network training with the wealth of unlabeled data available in the world. Another reason to assume that unsupervised methods will still have a significant role to play is the analogue to human learning, which seems to be much more data efficient and also happens to some extent in an unsupervised manner; we can learn to recognize objects and structures without knowing the specific label. We only need very limited supervision to categorize these recognized objects into classes. Two novel unsupervised strategies which we expect to have an impact in medical imaging are variational auto-encoders (VAEs), introduced by Kingma and Welling (2013) and generative adversarial networks (GANs), introduced by Goodfellow et al. (2014). The former merges variational Bayesian graphical models with neural networks as encoders/decoders. The latter uses two competing convolutional neural networks where one is generating artificial data samples and the other is discriminating artificial from real samples. Both have stochastic components and are generative networks. Most importantly, they can be trained end-to-end and learn representative features in a completely unsupervised manner. As we discussed in previous paragraphs, obtaining large amounts of unlabeled medical data is generally much easier than labeled data and unsupervised methods like VAEs and GANs could optimally leverage this wealth of information.

Finally, deep learning methods have often been described as 'black boxes'. Especially in medicine, where accountability is important and can have serious legal consequences, it is often not enough to have a good prediction system. This system also has to be able to articulate itself in a certain way. Several strategies have been developed to understand what intermediate layers of convolutional networks are responding to, for example deconvolution networks (Zeiler and Fergus, 2014), guided back-propagation (Springenberg et al., 2014) or deep Taylor composition (Montavon et al., 2017). Other researchers have tied prediction to textual representations of the image (i.e. captioning) (Karpathy and Fei-Fei, 2015), which is another useful avenue to understand what a network is perceiving. Last, some groups have tried to combine Bayesian statistics with deep networks to obtain true network uncertainty estimates Kendall and Gal (2017). This would allow physicians to assess when the network is giving unreliable predictions. Leveraging these techniques in the application of deep learning methods to medical image analysis could accelerate acceptance of deep learning applications among clinicians, and among patients. We also foresee deep learning approaches will be used for related tasks in medical imaging, mostly unexplored, such as image reconstruction (Wang, 2016). Deep learning will thus not only have a great impact in medical image analysis, but in medical imaging as a whole.

Acknowledgments

The authors would like to thank members of the Diagnostic Image Analysis Group for discussions and suggestions. This research was funded by grants KUN 2012-5577, KUN 2014-7032, and KUN 2015-7970 of the Dutch Cancer Society.

Appendix A: Literature selection

PubMed was searched for papers containing "convolutional" OR "deep learning" in any field. We specifically did not include the term neural network here as this would result in an enormous amount of 'false positive' papers covering brain research. This search initially gave over 700 hits. ArXiv was searched for papers mentioning one of a set of terms related to medical imaging. The exact search string was: 'abs:((medical OR mri OR "magnetic resonance" OR CT OR "computed tomography" OR ultrasound OR pathology OR xray OR x-ray OR radiograph OR mammography OR

fundus OR OCT) AND (”deep learning” OR convolutional OR cnn OR ”neural network”))’. Conference proceedings for MICCAI (including workshops), SPIE, ISBI and EMBC were searched based on titles of papers. Again we looked for mentions of ’deep learning’ or ’convolutional’ or ’neural network’. We went over all these papers and excluded the ones that did not discuss medical imaging (e.g. applications to genetics, chemistry), only used handcrafted features in combination with neural networks, or only referenced deep learning as future work. When in doubt whether a paper should be included we read the abstract and when the exact methodology was still unclear we read the paper itself. We checked references in all selected papers iteratively and consulted colleagues to identify any papers which were missed by our initial search. When largely overlapping work had been reported in multiple publications, only the publication deemed most important was included. A typical example here was arXiv preprints that were subsequently published or conference contributions which were expanded and published in journals.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv:1603.04467.

Abràmoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C., Niemeijer, M., 2016. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci 57 (13), 5200–5206.

Akram, S. U., Kannala, J., Eklund, L., Heikkilä, J., 2016. Cell segmentation proposal network for microscopy image analysis. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 21–29.

Akselrod-Ballin, A., Karlinsky, L., Alpert, S., Hasoul, S., Ben-Ari, R., Barkan, E., 2016. A region based convolutional network for tumor detection and classification in breast mammography. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 197–205.

Alansary, A., Kamnitsas, K., Davidson, A., Khlebnikov, R., Rajchl, M., Malamateniou, C., Rutherford, M., Hajnal, J. V., Glocker, B., Rueckert, D., Kainz, B., 2016. Fast fully automatic segmentation of the human placenta from motion corrupted MRI. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 589–597.

Albarqouni, S., Baur, C., Achilles, F., Belagiannis, V., Demirci, S., Navab, N., 2016. AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images. IEEE Trans Med Imaging 35, 1313–1321.

Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2015. A comparative study for chest radiograph image retrieval using binary texture and deep learning classification. In: Conf Proc IEEE Eng Med Biol Soc. pp. 2940–2943.

Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2016. Visualizing and enhancing a deep learning framework using patients age and gender for chest X-ray image retrieval. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 978510.

Andermatt, S., Pezold, S., Cattin, P., 2016. Multi-dimensional gated recurrent units for the segmentation of biomedical 3D-data. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 142–151.

Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., Mougiakakou, S., 2016. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans Med Imaging 35 (5), 1207–1216.

Antony, J., McGuinness, K., Connor, N. E. O., Moran, K., 2016. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. arXiv:1609.02469.

Apou, G., Schaadt, N. S., Naegel, B., Forestier, G., Schönmeyer, R., Feuerhake, F., Wemmert, C., Grote, A., 2016. Detection of lobular structures in normal breast tissue. Comput Biol Med 74, 91–102.

Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L., Guevara Lopez, M. A., 2016. Representation learning for mammography mass lesion classification with convolutional neural networks. Comput Methods Programs Biomed 127, 248–257.

Armato, S. G., McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., Zhao, B., Aberle, D. R., Henschke, C. I., Hoffman, E. A., Kazerooni, E. A., MacMahon, H., Beek, E. J. R. V., Yankelevitz, D., Biancardi, A. M., Bland, P. H., Brown, M. S., Engelmann, R. M., Laderach, G. E., Max, D., Pais, R. C., Qing, D. P. Y., Roberts, R. Y., Smith, A. R., Starkey, A., Batrah, P., Caligiuri, P., Farooqi, A., Gladish, G. W., Jude, C. M., Munden, R. F., Petkovska, I., Quint, L. E., Schwartz, L. H., Sundaram, B., Dodd, L. E., Fenimore, C., Gur, D., Petrick, N., Freymann, J., Kirby, J., Hughes, B., Casteele, A. V., Gupte, S., Sallamm, M., Heath, M. D., Kuhn, M. H., Dharaiya, E., Burns, R., Fryd, D. S., Salganicoff, M., Anand, V., Shreter, U., Vastagh, S., Croft, B. Y., 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Med Phys 38, 915–931.

Avendi, M., Kheradvar, A., Jafarkhani, H., 2016. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. Med Image Anal 30, 108–119.

Azizi, S., Imani, F., Ghavidel, S., Tahmasebi, A., Kwak, J. T., Xu, S., Turkbey, B., Choyke, P., Pinto, P., Wood, B., Mousavi, P., Abolmaesumi, P., 2016. Detection of prostate cancer using temporal sequences of ultrasound data: a large clinical feasibility study. Int J Comput Assist Radiol Surg 11 (6), 947–956.

Bahrami, K., Shi, F., Rekik, I., Shen, D., 2016. Convolutional neural network for reconstruction of 7T-like images from 3T MRI using appearance and anatomical features. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 39–47.

Bao, S., Chung, A. C., 2016. Multi-scale structured CNN with label consistency for brain MR image segmentation. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–5.

Bar, Y., Diamant, I., Wolf, L., Greenspan, H., 2015. Deep learning with non-medical training used for chest pathology identification. In: Medical Imaging. Vol. 9414 of Proceedings of the SPIE. p. 94140V.

Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., Greenspan, H., 2016. Chest pathology identification using deep feature selection with non-medical training. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–5.

Barbu, A., Lu, L., Roth, H., Seff, A., Summers, R. M., 2016. An analysis of robust cost functions for CNN in computer-aided diagnosis.

Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 2016, 1–6.

Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I., Bergeron, A., Bouchard, N., Warde-Farley, D., Bengio, Y., 2012. Theano: new features and speed improvements. In: Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop.

Bauer, S., Carion, N., Schäffler, P., Fuchs, T., Wild, P., Buhmann, J. M., 2016. Multi-organ cancer classification and survival analysis. arXiv:1606.00897.

Baumgartner, C. F., Kamnitsas, K., Matthew, J., Smith, S., Kainz, B., Rueckert, D., 2016. Real-time standard scan plane detection and localisation in fetal ultrasound using fully convolutional neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 203–211.

Ben-Cohen, A., Diamant, I., Klang, E., Amitai, M., Greenspan, H., 2016. Dlmia. In: International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis. Vol. 10008 of Lect Notes Comput Sci. pp. 77–85.

Bengio, Y., 2012. Practical recommendations for gradient-based training of deep architectures. In: Neural Networks: Tricks of the Trade. Springer Berlin Heidelberg, pp. 437–478.

Bengio, Y., Courville, A., Vincent, P., 2013. Representation learning: A review and new perspectives. IEEE Trans Pattern Anal Mach Intell 35 (8), 1798–1828.

Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., 2007. Greedy layer-wise training of deep networks. In: Advances in Neural Information Processing Systems. pp. 153–160.

Bengio, Y., Simard, P., Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw 5, 157–166.

Benou, A., Veksler, R., Friedman, A., Raviv, T. R., 2016. De-noising of contrast-enhanced mri sequences by an ensemble of expert deep neural networks. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 95–110.

BenTaieb, A., Hamarneh, G., 2016. Topology aware fully convolutional networks for histology gland segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 460–468.

BenTaieb, A., Kawahara, J., Hamarneh, G., 2016. Multi-loss convolutional networks for gland analysis in microscopy. In: IEEE Int Symp Biomedical Imaging. pp. 642–645.

Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. J Mach Learn Res 13 (1), 281–305.

Birenbaum, A., Greenspan, H., 2016. Longitudinal multiple sclerosis lesion segmentation using multi-view convolutional neural networks. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 58–67.

Brosch, T., Tam, R., 2013. Manifold learning of brain MRIs by deep learning. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 633–640.

Brosch, T., Tang, L. Y., Yoo, Y., Li, D. K., Traboulsee, A., Tam, R., 2016. Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to Multiple Sclerosis lesion segmentation. IEEE Trans Med Imaging 35 (5), 1229–1239.

Brosch, T., Yoo, Y., Li, D. K. B., Traboulsee, A., Tam, R., 2014. Modeling the variability in brain morphology and lesion distribution in multiple sclerosis by deep learning. In: Med Image Comput Comput Assist Interv. Vol. 8674 of Lect Notes Comput Sci. pp. 462–469.

Burlina, P., Freund, D. E., Joshi, N., Wolfson, Y., Bressler, N. M., 2016. Detection of age-related macular degeneration via deep learning. In: IEEE Int Symp Biomedical Imaging. pp. 184–188.

Bychkov, D., Turkki, R., Haglund, C., Linder, N., Lundin, J., 2016. Deep learning for tissue microarray image-based outcome prediction in patients with colorectal cancer. In: Medical Imaging. Vol. 9791 of Proceedings of the SPIE. p. 979115.

Cai, J., Lu, L., Zhang, Z., Xing, F., Yang, L., Yin, Q., 2016a. Pancreas segmentation in mri using graph-based decision fusion on convolutional neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 442–450.

Cai, Y., Landis, M., Laidley, D. T., Kornecki, A., Lum, A., Li, S., 2016b. Multi-modal vertebrae recognition using transformed deep convolution network. Comput Med Imaging Graph 51, 11–19.

Carneiro, G., Nascimento, J. C., 2013. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. IEEE Trans Pattern Anal Mach Intell 35, 2592–2607.

Carneiro, G., Nascimento, J. C., Freitas, A., 2012. The segmentation of the left ventricle of the heart from ultrasound data using deep learning architectures and derivative-based search methods. IEEE Trans Image Process, 968–982.

Carneiro, G., Oakden-Rayner, L., Bradley, A. P., Nascimento, J., Palmer, L., 2016. Automated 5-year mortality prediction using deep learning and radiomics features from chest computed tomography. arXiv:1607.00267.

Cha, K. H., Hadjiiski, L. M., Samala, R. K., Chan, H.-P., Cohan, R. H., Caoili, E. M., Paramagul, C., Alva, A., Weizer, A. Z., Dec. 2016. Bladder cancer segmentation in CT for treatment response assessment: Application of deep-learning convolution neural network-a pilot study. Tomography 2, 421–429.

Chang, H., Han, J., Zhong, C., Snijders, A., Mao, J.-H., Jan. 2017. Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications. IEEE transactions on pattern analysis and machine intelligence.

Charbonnier, J., van Rikxoort, E., Setio, A., Schaefer-Prokop, C., van Ginneken, B., Ciompi, F., 2017. Improving airway segmentation in computed tomography using leak detection with convolutional networks. Med Image Anal 36, 52–60.

Chen, H., Dou, Q., Ni, D., Cheng, J.-Z., Qin, J., Li, S., Heng, P.-A., 2015a. Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9349 of Lect Notes Comput Sci. Cham, pp. 507–514.

Chen, H., Dou, Q., Yu, L., Heng, P.-A., 2016a. VoxResNet: Deep voxelwise residual networks for volumetric brain segmentation. arXiv:1608.05895.

Chen, H., Ni, D., Qin, J., Li, S., Yang, X., Wang, T., Heng, P. A., 2015b. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. IEEE J Biomed Health Inform 19 (5), 1627–1636.

Chen, H., Qi, X., Yu, L., Heng, P.-A., 2017. DCAN: Deep contour-aware networks for accurate gland segmentation. Med Image Anal 36, 135–146.

Chen, H., Shen, C., Qin, J., Ni, D., Shi, L., Cheng, J. C. Y., Heng, P.-A., 2015c. Automatic localization and identification of vertebrae in spine CT via a joint learning model with deep neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9349 of Lect Notes Comput Sci. pp. 515–522.

Chen, H., Wang, X., Heng, P. A., 2016b. Automated mitosis detection with deep regression networks. In: IEEE Int Symp Biomedical Imaging. pp. 1204–1207.

Chen, H., Zheng, Y., Park, J.-H., Heng, P.-A., Zhou, S. K., 2016c. Iterative multi-domain regularized deep learning for anatomical structure detection and segmentation from ultrasound images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 487–495.

Chen, J., Yang, L., Zhang, Y., Alber, M., Chen, D. Z., 2016d. Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation. In: Advances in Neural Information Processing Systems. pp. 3036–3044.

Chen, S., Qin, J., Ji, X., Lei, B., Wang, T., Ni, D., Cheng, J.-Z., 2016e. Automatic scoring of multiple semantic attributes with multi-task feature leverage: A study on pulmonary nodules in CT images. IEEE Trans Med Imaging, in press.

Chen, X., Xu, Y., Wong, D. W. K., Wong, T. Y., Liu, J., 2015d. Glaucoma detection based on deep convolutional neural network. In: Conf Proc IEEE Eng Med Biol Soc. pp. 715–718.

Cheng, J.-Z., Ni, D., Chou, Y.-H., Qin, J., Tiu, C.-M., Chang, Y.-C., Huang, C.-S., Shen, D., Chen, C.-M., 2016a. Computer-Aided Diagnosis with deep learning architecture: Applications to breast lesions in US images and pulmonary nodules in CT scans. Nat Sci Rep 6, 24454.

Cheng, R., Roth, H. R., Lu, L., Wang, S., Turkbey, B., Gandler, W., McCreedy, E. S., Agarwal, H. K., Choyke, P., Summers, R. M., McAuliffe, M. J., 2016b. Active appearance model and deep learning for more accurate prostate segmentation on MRI. In: Medical Imaging. Vol. 9784 of Proceedings of the SPIE. p. 97842I.

Cheng, X., Zhang, L., Zheng, Y., 2015. Deep similarity learning for multimodal medical images. Computer Methods in Biomechanics and Biomedical Engineering, 1–5.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning phrase representations using rnn encoder-decoder for statistical machine translation. arXiv:1406.1078.

Choi, H., Jin, K. H., 2016. Fast and robust segmentation of the striatum using deep convolutional neural networks. Journal of Neuroscience Methods 274, 146–153.

Christ, P. F., Elshaer, M. E. A., Ettlinger, F., Tatavarty, S., Bickel, M., Bilic, P., Rempfler, M., Armbruster, M., Hofmann, F., D'Anastasi, M., et al., 2016. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 415–423.

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

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., Ronneberger, O., 2016. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. Springer, pp. 424–432.

Cicero, M., Bilbily, A., Colak, E., Dowdell, T., Gray, B., Perampaladas, K., Barfett, J., 2016. Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs. Invest Radiol, in press.

Ciompi, F., Chung, K., van Riel, S. J., Setio, A. A. A., Gerke, P. K., Jacobs, C., Scholten, E. T., Schaefer-Prokop, C. M., Wille, M. M. W., Marchiano, A., Pastorino, U., Prokop, M., van Ginneken, B., 2016. Towards automatic pulmonary nodule management in lung cancer screening with deep learning. arXiv:1610.09157.

Ciompi, F., de Hoop, B., van Riel, S. J., Chung, K., Scholten, E. T., Oudkerk, M., de Jong, P. A., Prokop, M., van Ginneken, B., 2015. Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. Med Image Anal 26, 195–202.

Cireşan, D. C., Giusti, A., Gambardella, L. M., Schmidhuber, J., 2013. Mitosis detection in breast cancer histology images with deep neural networks. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 411–418.

Ciresan, D., Giusti, A., Gambardella, L. M., Schmidhuber, J., 2012. Deep neural networks segment neuronal membranes in electron microscopy images. In: Advances in Neural Information Processing Systems. pp. 2843–2851.

Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., Smith, J. R., 2015. Deep learning, sparse coding, and svm for melanoma recognition in dermoscopy images. In: International Workshop on Machine Learning in Medical Imaging. pp. 118–126.

Collobert, R., Kavukcuoglu, K., Farabet, C., 2011. Torch7: A matlab-like environment for machine learning. In: Advances in Neural Information Processing Systems.

Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A., 2014. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Medical Imaging. Vol. 9041 of Proceedings of the SPIE. p. 904103.

Cruz-Roa, A. A., Ovalle, J. E. A., Madabhushi, A., Osorio, F. A. G., 2013. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 403–410.

Dalmis, M., Litjens, G., Holland, K., Setio, A., Mann, R., Karssemeijer, N., Gubern-Mérida, A., Feb. 2017. Using deep learning to segment breast and fibroglandular tissue in mri volumes. Medical physics 44, 533–546.

de Brebisson, A., Montana, G., 2015. Deep neural networks for anatomical brain segmentation. In: Comput Vis Pattern Recognit. pp. 20–28.

de Vos, B. D., Viergever, M. A., de Jong, P. A., Išgum, I., 2016a. Automatic slice identification in 3D medical images with a ConvNet regressor. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 161–169.

de Vos, B. D., Wolterink, J. M., de Jong, P. A., Viergever, M. A., Išgum, I., 2016b. 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In: Medical Imaging. Vol. 9784 of Proceedings of the SPIE. p. 97841Y.

Demyanov, S., Chakravorty, R., Abedini, M., Halpern, A., Garnavi, R., 2016. Classification of dermoscopy patterns using deep convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 364–368.

Dhungel, N., Carneiro, G., Bradley, A. P., 2016. The automated learning of deep features for breast mass classification from mammograms. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. Springer, pp. 106–114.

Dou, Q., Chen, H., Jin, Y., Yu, L., Qin, J., Heng, P.-A., 2016a. 3D deeply supervised network for automatic liver segmentation from CT volumes. arXiv:1607.00582.

Dou, Q., Chen, H., Yu, L., Qin, J., Heng, P. A., 2016b. Multi-level contextual 3D CNNs for false positive reduction in pulmonary nodule detection, in press.

Dou, Q., Chen, H., Yu, L., Shi, L., Wang, D., Mok, V. C., Heng, P. A., 2015. Automatic cerebral microbleeds detection from MR images via independent subspace analysis based hierarchical features. Conf Proc IEEE Eng Med Biol Soc, 7933–7936.

Dou, Q., Chen, H., Yu, L., Zhao, L., Qin, J., Wang, D., Mok, V. C., Shi, L., Heng, P.-A., 2016c. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 35, 1182–1195.

Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., Pal, C., 2016. The importance of skip connections in biomedical image segmentation. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 179–187.

Dubrovina, A., Kisilev, P., Ginsburg, B., Hashoul, S., Kimmel, R., 2016. Computational mammography using deep neural networks. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–5.

Ehteshami Bejnordi, B., Litjens, G., Timofeeva, N., Otte-Holler, I., Homeyer, A., Karssemeijer, N., van der Laak, J., Sep 2016. Stain

specific standardization of whole-slide histopathological images. IEEE Trans Med Imaging 35 (2), 404–415.
URL http://dx.doi.org/10.1109/TMI.2015.2476509

Emad, O., Yassine, I. A., Fahmy, A. S., 2015. Automatic localization of the left ventricle in cardiac MRI images using deep learning. In: Conf Proc IEEE Eng Med Biol Soc. pp. 683–686.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118.

Farabet, C., Couprie, C., Najman, L., LeCun, Y., 2013. Learning hierarchical features for scene labeling. IEEE Trans Pattern Anal Mach Intell 35 (8), 1915–1929.

Farag, A., Lu, L., Roth, H. R., Liu, J., Turkbey, E., Summers, R. M., 2015. A bottom-up approach for pancreas segmentation using cascaded superpixels and (deep) image patch labeling. arXiv:1505.06236.

Ferrari, A., Lombardi, S., Signoroni, A., 2015. Bacterial colony counting by convolutional neural networks. Conf Proc IEEE Eng Med Biol Soc, 7458–7461.

Fonseca, P., Mendoza, J., Wainer, J., Ferrer, J., Pinto, J., Guerrero, J.and Castaneda, B., 2015. Automatic breast density classification using a convolutional neural network architecture search procedure. In: Medical Imaging. Vol. 9413 of Proceedings of the SPIE. p. 941428.

Forsberg, D., Sjöblom, E., Sunshine, J. L., 2017. Detection and labeling of vertebrae in MR images using deep learning with clinical annotations as training data. J Digit Imaging, in press.

Fotin, S. V., Yin, Y., Haldankar, H., Hoffmeister, J. W., Periaswamy, S., 2016. Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 97850X.

Fritscher, K., Raudaschl, P., Zaffino, P., Spadea, M. F., Sharp, G. C., Schubert, R., 2016. Deep neural networks for fast segmentation of 3D medical images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 158–165.

Fu, H., Xu, Y., Lin, S., Kee Wong, D. W., Liu, J., 2016a. Deepvessel: Retinal vessel segmentation via?deep learning and conditional random?field. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 132–139.

Fu, H., Xu, Y., Wong, D. W. K., Liu, J., 2016b. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: IEEE Int Symp Biomedical Imaging. pp. 698–701.

Fukushima, K., 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36 (4), 193–202.

Gao, M., Bagci, U., Lu, L., Wu, A., Buty, M., Shin, H.-C., Roth, H., Papadakis, G. Z., Depeursinge, A., Summers, R. M., Xu, Z., Mollura, D. J., 2016a. Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–6.

Gao, M., Xu, Z., Lu, L., Harrison, A. P., Summers, R. M., Mollura, D. J., 2016b. Multi-label deep regression and unordered pooling for holistic interstitial lung disease pattern detection. In: Machine Learning in Medical Imaging. Vol. 10019 of Lect Notes Comput Sci. pp. 147–155.

Gao, M., Xu, Z., Lu, L., Nogues, I., Summers, R., Mollura, D., 2016c. Segmentation label propagation using deep convolutional neural networks and dense conditional random field. In: IEEE Int Symp Biomedical Imaging. pp. 1265–1268.

Gao, X., Lin, S., Wong, T. Y., 2015. Automatic feature learning to grade nuclear cataracts based on deep learning. IEEE Trans Biomed Eng 62 (11), 2693–2701.

Gao, Y., Maraci, M. A., Noble, J. A., 2016d. Describing ultrasound video content using deep convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 787–790.

Gao, Z., Wang, L., Zhou, L., Zhang, J., 2016e. Hep-2 cell image classification with deep convolutional neural networks. Journal of Biomedical and Health Informatics.

Ghafoorian, M., Karssemeijer, N., Heskes, T., Bergkamp, M., Wissink, J., Obels, J., Keizer, K., de Leeuw, F.-E., van Ginneken, B., Marchiori, E., Platel, B., 2017. Deep multi-scale location-aware 3d convolutional neural networks for automated detection of lacunes of presumed vascular origin. NeuroImage: Clinical, in press.

Ghafoorian, M., Karssemeijer, N., Heskes, T., van Uden, I., Sanchez, C., Litjens, G., de Leeuw, F.-E., van Ginneken, B., Marchiori, E., Platel, B., 2016a. Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities. arXiv:1610.04834.

Ghafoorian, M., Karssemeijer, N., Heskes, T., van Uden, I. W. M., de Leeuw, F.-E., Marchiori, E., van Ginneken, B., Platel, B., 2016b. Non-uniform patch sampling with deep convolutional neural networks for white matter hyperintensity segmentation. In: IEEE Int Symp Biomedical Imaging. pp. 1414–1417.

Ghesu, F. C., Georgescu, B., Mansi, T., Neumann, D., Hornegger, J., Comaniciu, D., 2016a. An artificial agent for anatomical landmark detection in medical images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci.

Ghesu, F. C., Krubasik, E., Georgescu, B., Singh, V., Zheng, Y., Hornegger, J., Comaniciu, D., 2016b. Marginal space deep learning: Efficient architecture for volumetric image parsing. IEEE Trans Med Imaging 35, 1217–1228.

Golan, D., Donner, Y., Mansi, C., Jaremko, J., Ramachandran, M., 2016. Fully automating Graf's method for DDH diagnosis using deep convolutional neural networks. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 130–141.

Golkov, V., Dosovitskiy, A., Sperl, J., Menzel, M., Czisch, M., Samann, P., Brox, T., Cremers, D., 2016. q-Space deep learning: Twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans Med Imaging 35, 1344 – 1351.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative adversarial nets. arXiv:1406.2661.

Greenspan, H., Summers, R. M., van Ginneken, B., 2016. Deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Trans Med Imaging 35 (5), 1153–1159.

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

Gülsün, M. A., Funka-Lea, G., Sharma, P., Rapaka, S., Zheng, Y., 2016. Coronary centerline extraction via optimal flow paths and CNN path pruning. In: Med Image Comput Comput Assist Interv. Vol. 9902 of Lect Notes Comput Sci. Springer, pp. 317–325.

Günhan Ertosun, M., Rubin, D. L., 2015. Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. In: AMIA Annual Symposium. pp. 1899–1908.

Guo, Y., Gao, Y., Shen, D., 2016. Deformable MR prostate segmentation via deep feature learning and sparse patch matching. IEEE Trans Med Imaging 35 (4), 1077–1089.

Guo, Y., Wu, G., Commander, L. A., Szary, S., Jewells, V., Lin, W., Shen, D., 2014. Segmenting hippocampus from infant brains by sparse patch matching with deep-learned features. In: Med Image Comput Comput Assist Interv. Vol. 8674 of Lect Notes Comput

Sci. pp. 308–315.
Han, X.-H., Lei, J., Chen, Y.-W., 2016. HEp-2 cell classification using K-support spatial pooling in deep CNNs. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 3–11.
Haugeland, J., 1985. Artificial intelligence: the very idea. The MIT Press, Cambridge, Mass.
Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., Larochelle, H., 2016a. Brain tumor segmentation with Deep Neural Networks. Med Image Anal 35, 18–31.
Havaei, M., Guizard, N., Chapados, N., Bengio, Y., 2016b. HeMIS: Hetero-modal image segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 469–477.
He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep residual learning for image recognition. arXiv:1512.03385.
Hinton, G., 2010. A practical guide to training restricted Boltzmann machines. Momentum 9 (1), 926.
Hinton, G. E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comput 18, 1527–1554.
Hinton, G. E., Salakhutdinov, R. R., 2006. Reducing the dimensionality of data with neural networks. Science 313, 504–507.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Computation 9 (8), 1735–1780.
Hoffmann, N., Koch, E., Steiner, G., Petersohn, U., Kirsch, M., 2016. Learning thermal process representations for intraoperative analysis of cortical perfusion during ischemic strokes. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 152–160.
Hoogi, A., Subramaniam, A., Veerapaneni, R., Rubin, D., 2016. Adaptive estimation of active contour parameters using convolutional neural networks and texture analysis. IEEE Trans Med Imaging.
Hosseini-Asl, E., Gimel'farb, G., El-Baz, A., 2016. Alzheimer's disease diagnostics by a deeply supervised adaptable 3D convolutional network. arXiv:1607.00556.
Hu, P., Wu, F., Peng, J., Bao, Y., Chen, F., Kong, D., Nov. 2016a. Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets. Int J Comput Assist Radiol Surg.
Hu, P., Wu, F., Peng, J., Liang, P., Kong, D., Dec. 2016b. Automatic 3D liver segmentation based on deep learning and globally optimized surface evolution. Phys Med Biol 61, 8676–8698.
Huang, H., Hu, X., Han, J., Lv, J., Liu, N., Guo, L., Liu, T., 2016. Latent source mining in FMRI data via deep neural network. In: IEEE Int Symp Biomedical Imaging. pp. 638–641.
Huynh, B. Q., Li, H., Giger, M. L., Jul 2016. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J Med Imaging 3, 034501.
Hwang, S., Kim, H., 2016. Self-transfer learning for fully weakly supervised object localization. arXiv:1602.01625.
Hwang, S., Kim, H.-E., Jeong, J., Kim, H.-J., 2016. A novel approach for tuberculosis screening based on deep convolutional neural networks. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. pp. 97852W–1.
Jamaludin, A., Kadir, T., Zisserman, A., 2016. SpineNet: Automatically pinpointing classification evidence in spinal MRIs. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 166–175.
Jamieson, A. R., Drukker, K., Giger, M. L., 2012. Breast image feature learning with adaptive deconvolutional networks. In: Medical Imaging. Vol. 8315 of Proceedings of the SPIE. p. 831506.
Janowczyk, A., Basavanhally, A., Madabhushi, A., 2016a. Stain normalization using sparse autoencoders (StaNoSA): Application to digital pathology. Comput Med Imaging Graph, in press.
Janowczyk, A., Doyle, S., Gilmore, H., Madabhushi, A., 2016b. A resolution adaptive deep hierarchical (RADHicaL) learning scheme applied to nuclear segmentation of digital pathology images. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–7.
Janowczyk, A., Madabhushi, A., 2016. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. Journal of pathology informatics 7, 29.
Jaumard-Hakoun, A., Xu, K., Roussel-Ragot, P., Dreyfus, G., Denby, B., 2016. Tongue contour extraction from ultrasound images based on deep neural network. arXiv:1605.05912.
Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: Convolutional architecture for fast feature embedding. In: Proceedings of the 22nd ACM International Conference on Multimedia. pp. 675–678.
Kainz, P., Pfeiffer, M., Urschler, M., 2015. Semantic segmentation of colon glands with deep convolutional neural networks and total variation segmentation. arXiv:1511.06919.
Källén, H., Molin, J., Heyden, A., Lundstr, C., Aström, K., 2016. Towards grading gleason score using generically trained deep convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 1163–1167.
Kallenberg, M., Petersen, K., Nielsen, M., Ng, A., Diao, P., Igel, C., Vachon, C., Holland, K., Karssemeijer, N., Lillholm, M., 2016. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. IEEE Trans Med Imaging 35, 1322–1331.
Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., Glocker, B., 2017. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Med Image Anal 36, 61–78.
Karpathy, A., Fei-Fei, L., June 2015. Deep visual-semantic alignments for generating image descriptions. In: Comput Vis Pattern Recognit. ArXiv:1412.2306.
Kashif, M. N., Raza, S. E. A., Sirinukunwattana, K., Arif, M., Rajpoot, N., 2016. Handcrafted features with convolutional neural networks for detection of tumor cells in histology images. In: IEEE Int Symp Biomedical Imaging. pp. 1029–1032.
Kawahara, J., BenTaieb, A., Hamarneh, G., 2016a. Deep features to classify skin lesions. In: IEEE Int Symp Biomedical Imaging. pp. 1397–1400.
Kawahara, J., Brown, C. J., Miller, S. P., Booth, B. G., Chau, V., Grunau, R. E., Zwicker, J. G., Hamarneh, G., 2016b. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. NeuroImage.
Kawahara, J., Hamarneh, G., 2016. Multi-resolution-tract CNN with hybrid pretrained and skin-lesion trained layers. In: Machine Learning in Medical Imaging. Vol. 10019 of Lect Notes Comput Sci. pp. 164–171.
Kendall, A., Gal, Y., 2017. What uncertainties do we need in bayesian deep learning for computer vision? arXiv:1703.04977.
Kim, E., Cortre-Real, M., Baloch, Z., 2016a. A deep semantic mobile application for thyroid cytopathology. In: Medical Imaging. Vol. 9789 of Proceedings of the SPIE. p. 97890A.
Kim, H., Hwang, S., 2016. Scale-invariant feature learning using deconvolutional neural networks for weakly-supervised semantic segmentation. arXiv:1602.04984.
Kim, J., Calhoun, V. D., Shim, E., Lee, J.-H., 2016b. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. NeuroImage 124, 127–146.
Kingma, D. P., Welling, M., 2013. Auto-encoding variational bayes. arXiv:1312.6114.
Kisilev, P., Sason, E., Barkan, E., Hashoul, S., 2016. Medical image description using multi-task-loss CNN. In: International Workshop

on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis. Springer, pp. 121–129.

Kleesiek, J., Urban, G., Hubert, A., Schwarz, D., Maier-Hein, K., Bendszus, M., Biller, A., 2016. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. NeuroImage 129, 460–469.

Kong, B., Zhan, Y., Shin, M., Denny, T., Zhang, S., 2016. Recognizing end-diastole and end-systole frames via deep temporal regression network. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 264–272.

Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I., Mann, R., den Heeten, A., Karssemeijer, N., 2016. Large scale deep learning for computer aided detection of mammographic lesions. Med Image Anal 35, 303–312.

Kooi, T., van Ginneken, B., Karssemeijer, N., den Heeten, A., 2017. Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network. Medical Physics.

Korez, R., Likar, B., Pernuš, F., Vrtovec, T., 2016. Model-based segmentation of vertebral bodies from MR images with 3D CNNs. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. Springer, pp. 433–441.

Krizhevsky, A., Sutskever, I., Hinton, G., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105.

Kumar, A., Sridar, P., Quinton, A., Kumar, R. K., Feng, D., Nanan, R., Kim, J., 2016. Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 791–794.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278–2324.

Lekadir, K., Galimzianova, A., Betriu, A., Del Mar Vila, M., Igual, L., Rubin, D. L., Fernandez, E., Radeva, P., Napel, S., Jan. 2017. A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. IEEE J Biomed Health Inform 21, 48–55.

Lessmann, N., Isgum, I., Setio, A. A., de Vos, B. D., Ciompi, F., de Jong, P. A., Oudkerk, M., Mali, W. P. T. M., Viergever, M. A., van Ginneken, B., 2016. Deep convolutional neural networks for automatic coronary calcium scoring in a screening study with low-dose chest CT. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. pp. 978511–1 – 978511–6.

Li, R., Zhang, W., Suk, H.-I., Wang, L., Li, J., Shen, D., Ji, S., 2014. Deep learning based imaging data completion for improved brain disease diagnosis. In: Med Image Comput Comput Assist Interv. Vol. 8675 of Lect Notes Comput Sci. pp. 305–312.

Li, W., Cao, P., Zhao, D., Wang, J., 2016a. Pulmonary nodule classification with deep convolutional neural networks on computed tomography images. Computational and Mathematical Methods in Medicine, 6215085.

Li, W., Jia, F., Hu, Q., 2015. Automatic segmentation of liver tumor in CT images with deep convolutional neural networks. Journal of Computer and Communications 3 (11), 146–151.

Li, W., Manivannan, S., Akbar, S., Zhang, J., Trucco, E., McKenna, S. J., 2016b. Gland segmentation in colon histology images using hand-crafted features and convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 1405–1408.

Liao, S., Gao, Y., Oto, A., Shen, D., 2013. Representation learning: A unified deep learning framework for automatic prostate mr segmentation. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 254–261.

Lin, M., Chen, Q., Yan, S., 2013. Network in network. arXiv:1312.4400.

Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., Hulsbergen-van de Kaa, C., Bult, P., van Ginneken, B., van der Laak, J., 2016. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Nat Sci Rep 6, 26286.

Liu, J., Wang, D., Wei, Z., Lu, L., Kim, L., Turkbey, E., Summers, R. M., 2016a. Colitis detection on computed tomography using regional convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 863–866.

Liu, X., Tizhoosh, H. R., Kofman, J., 2016b. Generating binary tags for fast medical image retrieval based on convolutional nets and Radon transform. In: International Joint Conference on Neural Networks. ArXiv:1604.04676.

Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., Venugopalan, S., Timofeev, A., Nelson, P. Q., Corrado, G. S., Hipp, J. D., Peng, L., Stumpe, M. C., 2017. Detecting cancer metastases on gigapixel pathology images. arXiv:1703.02442.

Lo, S.-C., Lou, S.-L., Lin, J.-S., Freedman, M. T., Chien, M. V., Mun, S. K., 1995. Artificial convolution neural network techniques and applications for lung nodule detection. IEEE Trans Med Imaging 14, 711–718.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. arXiv:1411.4038.

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

Lu, X., Xu, D., Liu, D., 2016. Robust 3d organ localization with dual learning architectures and fusion. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 12–20.

Ma, J., Wu, F., Zhu, J., Xu, D., Kong, D., Jan 2017. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. Ultrasonics 73, 221–230.

Mahapatra, D., Roy, P. K., Sedai, S., Garnavi, R., 2016. Retinal image quality classification using saliency maps and CNNs. In: Machine Learning in Medical Imaging. Vol. 10019 of Lect Notes Comput Sci. pp. 172–179.

Malon, C. D., Cosatto, E., 2013. Classification of mitotic figures with convolutional neural networks and seeded blob features. Journal of pathology informatics.

Maninis, K.-K., Pont-Tuset, J., Arbeláez, P., Gool, L., 2016. Deep retinal image understanding. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 140–148.

Mansoor, A., Cerrolaza, J., Idrees, R., Biggs, E., Alsharid, M., Avery, R., Linguraru, M. G., 2016. Deep learning guided partitioned shape model for anterior visual pathway segmentation. IEEE Trans Med Imaging 35 (8), 1856–1865.

Mao, Y., Yin, Z., 2016. A hierarchical convolutional neural network for mitosis detection in phase-contrast microscopy images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 685–692.

Menegola, A., Fornaciali, M., Pires, R., Avila, S., Valle, E., 2016. Towards automated melanoma screening: Exploring transfer learning schemes. arXiv:1609.01228.

Merkow, J., Kriegman, D., Marsden, A., Tu, Z., 2016. Dense volume-to-volume vascular boundary detection. arXiv:1605.08401.

Miao, S., Wang, Z. J., Liao, R., 2016. A CNN regression approach for real-time 2D/3D registration. IEEE Trans Med Imaging 35 (5), 1352–1363.

Milletari, F., Ahmadi, S.-A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., Levin, J., Dietrich, O., Ertl-Wagner, B., Bötzel, K., Navab, N., 2016a. Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. arXiv:1601.07014.

Milletari, F., Navab, N., Ahmadi, S.-A., 2016b. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. arXiv:1606.04797.

Mishra, M., Schmitt, S., Wang, L., Strasser, M. K., Marr, C., Navab,

N., Zischka, H., Peng, T., 2016. Structure-based assessment of cancerous mitochondria using deep networks. In: IEEE Int Symp Biomedical Imaging. pp. 545–548.

Moeskops, P., Viergever, M. A., Mendrik, A. M., de Vries, L. S., Benders, M. J. N. L., Isgum, I., 2016a. Automatic segmentation of MR brain images with a convolutional neural network. IEEE Trans Med Imaging 35 (5), 1252–1262.

Moeskops, P., Wolterink, J. M., Velden, B. H. M., Gilhuijs, K. G. A., Leiner, T., Viergever, M. A., Isgum, I., 2016b. Deep learning for multi-task medical image segmentation in multiple modalities. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 478–486.

Montavon, G., Lapuschkin, S., Binder, A., Samek, W., Müller, K.-R., 2017. Explaining nonlinear classification decisions with deep taylor decomposition. Pattern Recognition 65, 211–222.

Moradi, M., Guo, Y., Gur, Y., Negahdar, M., Syeda-Mahmood, T., 2016a. A cross-modality neural network transform for semi-automatic medical image annotation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 300–307.

Moradi, M., Gur, Y., Wang, H., Prasanna, P., Syeda-Mahmood, T., 2016b. A hybrid learning approach for semantic labeling of cardiac CT slices and recognition of body position. In: IEEE Int Symp Biomedical Imaging.

Nappi, J. J., Hironaka, T., Regge, D., Yoshida, H., 2016. Deep transfer learning of virtual endoluminal views for the detection of polyps in CT colonography. In: Medical Imaging. Proceedings of the SPIE. p. 97852B.

Nascimento, J. C., Carneiro, G., 2016. Multi-atlas segmentation using manifold learning with deep belief networks. In: IEEE Int Symp Biomedical Imaging. pp. 867–871.

Ngo, T. A., Lu, Z., Carneiro, G., 2017. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Med Image Anal 35, 159–171.

Nie, D., Cao, X., Gao, Y., Wang, L., Shen, D., 2016a. Estimating CT image from MRI data using 3D fully convolutional networks. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 170–178.

Nie, D., Wang, L., Gao, Y., Shen, D., 2016b. Fully convolutional networks for multi-modality isointense infant brain image segmentation. In: IEEE Int Symp Biomedical Imaging. pp. 1342–1345.

Nie, D., Zhang, H., Adeli, E., Liu, L., Shen, D., 2016c. 3D deep learning for multi-modal imaging-guided survival time prediction of brain tumor patients. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 212–220.

Nogues, I., Lu, L., Wang, X., Roth, H., Bertasius, G., Lay, N., Shi, J., Tsehay, Y., Summers, R. M., 2016. Automatic lymph node cluster segmentation using holistically-nested neural networks and structured optimization in CT images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 388–397.

Oktay, O., Bai, W., Lee, M., Guerrero, R., Kamnitsas, K., Caballero, J., Marvao, A., Cook, S., O'Regan, D., Rueckert, D., 2016. Multi-input cardiac image super-resolution using convolutional neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9902 of Lect Notes Comput Sci. pp. 246–254.

Ortiz, A., Munilla, J., Górriz, J. M., Ramírez, J., 2016. Ensembles of deep learning architectures for the early diagnosis of the Alzheimer's disease. International Journal of Neural Systems 26, 1650025.

Paeng, K., Hwang, S., Park, S., Kim, M., Kim, S., 2016. A unified framework for tumor proliferation score prediction in breast histopathology. arXiv:1612.07180.

Pan, Y., Huang, W., Lin, Z., Zhu, W., Zhou, J., Wong, J., Ding, Z., 2015. Brain tumor grading based on neural networks and convolutional neural networks. In: Conf Proc IEEE Eng Med Biol Soc. pp. 699–702.

Payan, A., Montana, G., 2015. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks. arXiv:1502.02506.

Payer, C., Stern, D., Bischof, H., Urschler, M., 2016. Regressing heatmaps for multiple landmark localization using CNNs. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 230–238.

Pereira, S., Pinto, A., Alves, V., Silva, C. A., 2016. Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Trans Med Imaging.

Phan, H. T. H., Kumar, A., Kim, J., Feng, D., 2016. Transfer learning of a convolutional neural network for HEp-2 cell image classification. In: IEEE Int Symp Biomedical Imaging. pp. 1208–1211.

Pinaya, W. H. L., Gadelha, A., Doyle, O. M., Noto, C., Zugman, A., Cordeiro, Q., Jackowski, A. P., Bressan, R. A., Sato, J. R., Dec. 2016. Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. Nat Sci Rep 6, 38897.

Plis, S. M., Hjelm, D. R., Salakhutdinov, R., Allen, E. A., Bockholt, H. J., Long, J. D., Johnson, H. J., Paulsen, J. S., Turner, J. A., Calhoun, V. D., 2014. Deep learning for neuroimaging: a validation study. Frontiers in Neuroscience.

Poudel, R. P. K., Lamata, P., Montana, G., 2016. Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation. arXiv:1608.03974.

Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., Nielsen, M., 2013. Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 246–253.

Prentasic, P., Heisler, M., Mammo, Z., Lee, S., Merkur, A., Navajas, E., Beg, M. F., Sarunic, M., Loncaric, S., 2016. Segmentation of the foveal microvasculature using deep learning networks. Journal of Biomedical Optics 21, 75008.

Prentasic, P., Loncaric, S., 2016. Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion. Comput Methods Programs Biomed 137, 281–292.

Qiu, Y., Wang, Y., Yan, S., Tan, M., Cheng, S., Liu, H., Zheng, B., 2016. An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 978521.

Quinn, J. A., Nakasi, R., Mugagga, P. K. B., Byanyima, P., Lubega, W., Andama, A., 2016. Deep convolutional neural networks for microscopy-based point of care diagnostics. arXiv:1608.02989.

Rajchl, M., Lee, M. C., Oktay, O., Kamnitsas, K., Passerat-Palmbach, J., Bai, W., Kainz, B., Rueckert, D., 2016a. DeepCut: Object segmentation from bounding box annotations using convolutional neural networks. IEEE Trans Med Imaging, in press.

Rajchl, M., Lee, M. C., Schrans, F., Davidson, A., Passerat-Palmbach, J., Tarroni, G., Alansary, A., Oktay, O., Kainz, B., Rueckert, D., 2016b. Learning under distributed weak supervision. arXiv:1606.01100.

Rajkomar, A., Lingam, S., Taylor, A. G., Blum, M., Mongan, J., 2017. High-throughput classification of radiographs using deep convolutional neural networks. J Digit Imaging 30, 95–101.

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

Ravishankar, H., Prabhu, S. M., Vaidya, V., Singhal, N., 2016a. Hybrid approach for automatic segmentation of fetal abdomen from ultrasound images using deep learning. In: IEEE Int Symp Biomedical Imaging. pp. 779–782.

Ravishankar, H., Sudhakar, P., Venkataramani, R., Thiruvenkadam,

S., Annangi, P., Babu, N., Vaidya, V., 2016b. Understanding the mechanisms of deep transfer learning for medical images. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 188–196.

Rezaeilouyeh, H., Mollahosseini, A., Mahoor, M. H., 2016. Microscopic medical image classification framework via deep learning and shearlet transform. Journal of Medical Imaging 3 (4), 044501.

Romo-Bucheli, D., Janowczyk, A., Gilmore, H., Romero, E., Madabhushi, A., Sep 2016. Automated tubule nuclei quantification and correlation with Oncotype DX risk categories in ER+ breast cancer whole slide images. Nat Sci Rep 6, 32706.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9351 of Lect Notes Comput Sci. pp. 234–241.

Roth, H. R., Lee, C. T., Shin, H.-C., Seff, A., Kim, L., Yao, J., Lu, L., Summers, R. M., 2015a. Anatomy-specific classification of medical images using deep convolutional nets. In: IEEE Int Symp Biomedical Imaging. pp. 101–104.

Roth, H. R., Lu, L., Farag, A., Shin, H.-C., Liu, J., Turkbey, E. B., Summers, R. M., 2015b. DeepOrgan: Multi-level deep convolutional networks for automated pancreas segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9349 of Lect Notes Comput Sci. pp. 556–564.

Roth, H. R., Lu, L., Farag, A., Sohn, A., Summers, R. M., 2016a. Spatial aggregation of holistically-nested networks for automated pancreas segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 451–459.

Roth, H. R., Lu, L., Liu, J., Yao, J., Seff, A., Cherry, K., Kim, L., Summers, R. M., 2016b. Improving computer-aided detection using convolutional neural networks and random view aggregation. IEEE Trans Med Imaging 35 (5), 1170–1181.

Roth, H. R., Lu, L., Seff, A., Cherry, K. M., Hoffman, J., Wang, S., Liu, J., Turkbey, E., Summers, R. M., 2014. A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. In: Med Image Comput Comput Assist Interv. Vol. 8673 of Lect Notes Comput Sci. pp. 520–527.

Roth, H. R., Wang, Y., Yao, J., Lu, L., Burns, J. E., Summers, R. M., 2016c. Deep convolutional networks for automated detection of posterior-element fractures on spine CT. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 97850P.

Roth, H. R., Yao, J., Lu, L., Stieger, J., Burns, J. E., Summers, R. M., 2015c. Detection of sclerotic spine metastases via random aggregation of deep convolutional?neural network classifications. In: Recent Advances in Computational Methods and Clinical Applications for Spine Imaging. Vol. 20 of Lecture Notes in Computational Vision and Biomechanics. pp. 3–12.

Rupprecht, C., Huaroc, E., Baust, M., Navab, N., 2016. Deep active contours. arXiv:1607.05074.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., Fei-Fei, L., 2014. ImageNet large scale visual recognition challenge. Int J Comput Vis 115 (3), 1–42.

Sahiner, B., Chan, H.-P., Petrick, N., Wei, D., Helvie, M. A., Adler, D. D., Goodsitt, M. M., 1996. Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. IEEE Trans Med Imaging 15, 598–610.

Samala, R. K., Chan, H.-P., Hadjiiski, L., Cha, K., Helvie, M. A., 2016a. Deep-learning convolution neural network for computer-aided detection of microcalcifications in digital breast tomosynthesis. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 97850Y.

Samala, R. K., Chan, H.-P., Hadjiiski, L., Helvie, M. A., Wei, J., Cha, K., 2016b. Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography. Medical Physics 43 (12), 6654–6666.

Sarraf, S., Tofighi, G., 2016. Classification of Alzheimer's disease using fMRI data and deep learning convolutional neural networks. arXiv:1603.08631.

Schaumberg, A. J., Rubin, M. A., Fuchs, T. J., 2016. H&E-stained whole slide deep learning predicts SPOP mutation state in prostate cancer. bioRxiv:064279.

Schlegl, T., Waldstein, S. M., Vogl, W.-D., Schmidt-Erfurth, U., Langs, G., 2015. Predicting semantic descriptions from medical images with convolutional neural networks. In: Inf Process Med Imaging. Vol. 9123 of Lect Notes Comput Sci. pp. 437–448.

Sethi, A., Sha, L., Vahadane, A. R., Deaton, R. J., Kumar, N., Macias, V., Gann, P. H., 2016. Empirical comparison of color normalization methods for epithelial-stromal classification in H and E images. J Pathol Inform 7, 17.

Setio, A. A. A., Ciompi, F., Litjens, G., Gerke, P., Jacobs, C., van Riel, S., Wille, M. W., Naqibullah, M., Sanchez, C., van Ginneken, B., 2016. Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Trans Med Imaging 35 (5), 1160–1169.

Sevetlidis, V., Giuffrida, M. V., Tsaftaris, S. A., Jan. 2016. Whole image synthesis using a deep encoder-decoder network. In: Simulation and Synthesis in Medical Imaging. Vol. 9968 of Lect Notes Comput Sci. pp. 127–137.

Shah, A., Conjeti, S., Navab, N., Katouzian, A., 2016. Deeply learnt hashing forests for content based image retrieval in prostate MR images. In: Medical Imaging. Vol. 9784 of Proceedings of the SPIE. p. 978414.

Shakeri, M., Tsogkas, S., Ferrante, E., Lippe, S., Kadoury, S., Paragios, N., Kokkinos, I., 2016. Sub-cortical brain structure segmentation using F-CNNs. In: IEEE Int Symp Biomedical Imaging. pp. 269–272.

Shen, D., Wu, G., Suk, H.-I., Mar. 2017. Deep learning in medical image analysis. Annu Rev Biomed Eng.

Shen, W., Yang, F., Mu, W., Yang, C., Yang, X., Tian, J., 2015a. Automatic localization of vertebrae based on convolutional neural networks. In: Medical Imaging. Vol. 9413 of Proceedings of the SPIE. p. 94132E.

Shen, W., Zhou, M., Yang, F., Dong, D., Yang, C., Zang, Y., Tian, J., 2016. Learning from experts: Developing transferable deep features for patient-level lung cancer prediction. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 124–131.

Shen, W., Zhou, M., Yang, F., Yang, C., Tian, J., 2015b. Multi-scale convolutional neural networks for lung nodule classification. In: Inf Process Med Imaging. Vol. 9123 of Lect Notes Comput Sci. pp. 588–599.

Shi, J., Zheng, X., Li, Y., Zhang, Q., Ying, S., Jan. 2017. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease. IEEE J Biomed Health Inform, in press.

Shin, H.-C., Lu, L., Kim, L., Seff, A., Yao, J., Summers, R. M., 2015. Interleaved text/image deep mining on a very large-scale radiology database. In: Comput Vis Pattern Recognit. pp. 1090–1099.

Shin, H.-C., Orton, M. R., Collins, D. J., Doran, S. J., Leach, M. O., 2013. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. IEEE Trans Pattern Anal Mach Intell 35, 1930–1943.

Shin, H.-C., Roberts, K., Lu, L., Demner-Fushman, D., Yao, J., Summers, R. M., 2016a. Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. arXiv:1603.08486.

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., Summers, R. M., 2016b. Deep convolu-

tional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 35 (5), 1285–1298.

Shkolyar, A., Gefen, A., Benayahu, D., Greenspan, H., 2015. Automatic detection of cell divisions (mitosis) in live-imaging microscopy images using convolutional neural networks. In: Conf Proc IEEE Eng Med Biol Soc. pp. 743–746.

Simonovsky, M., Gutiérrez-Becker, B., Mateus, D., Navab, N., Komodakis, N., 2016. A deep metric for multimodal registration. In: Med Image Comput Comput Assist Interv. Vol. 9902 of Lect Notes Comput Sci. pp. 10–18.

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

Sirinukunwattana, K., Raza, S. E. A., Tsang, Y.-W., Snead, D. R., Cree, I. A., Rajpoot, N. M., 2016. Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. IEEE Trans Med Imaging 35 (5), 1196–1206.

Smistad, E., Løvstakken, L., 2016. Vessel detection in ultrasound images using deep convolutional neural networks. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 30–38.

Snoek, J., Larochelle, H., Adams, R. P., 2012. Practical bayesian optimization of machine learning algorithms. In: Advances in Neural Information Processing Systems. pp. 2951–2959.

Song, Y., Tan, E.-L., Jiang, X., Cheng, J.-Z., Ni, D., Chen, S., Lei, B., Wang, T., Sep 2017. Accurate cervical cell segmentation from overlapping clumps in pap smear images. IEEE Trans Med Imaging 36, 288–300.

Song, Y., Zhang, L., Chen, S., Ni, D., Lei, B., Wang, T., 2015. Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning. IEEE Trans Biomed Eng 62 (10), 2421–2433.

Spampinato, C., Palazzo, S., Giordano, D., Aldinucci, M., Leonardi, R., Feb. 2017. Deep learning for automated skeletal bone age assessment in X-ray images. Med Image Anal 36, 41–51.

Springenberg, J. T., Dosovitskiy, A., Brox, T., Riedmiller, M., 2014. Striving for simplicity: The all convolutional net. arXiv preprint arXiv:1412.6806.

Štern, D., Payer, C., Lepetit, V., Urschler, M., 2016. Automated age estimation from hand MRI volumes using deep learning. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 194–202.

Stollenga, M. F., Byeon, W., Liwicki, M., Schmidhuber, J., 2015. Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation. In: Advances in Neural Information Processing Systems. pp. 2998–3006.

Suk, H.-I., Lee, S.-W., Shen, D., 2014. Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. NeuroImage 101, 569–582.

Suk, H.-I., Lee, S.-W., Shen, D., 2015. Latent feature representation with stacked auto-encoder for AD/MCI diagnosis. Brain Struct Funct 220, 841–859.

Suk, H.-I., Shen, D., 2013. Deep learning-based feature representation for AD/MCI classification. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 583–590.

Suk, H.-I., Shen, D., 2016. Deep ensemble sparse regression network for Alzheimer's disease diagnosis. In: Med Image Comput Comput Assist Interv. Vol. 10019 of Lect Notes Comput Sci. pp. 113–121.

Suk, H.-I., Wee, C.-Y., Lee, S.-W., Shen, D., 2016. State-space model with deep learning for functional dynamics estimation in resting-state fMRI. NeuroImage 129, 292–307.

Sun, W., Tseng, T.-L. B., Zhang, J., Qian, W., 2016a. Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. Comput Med Imaging Graph.

Sun, W., Zheng, B., Qian, W., 2016b. Computer aided lung cancer diagnosis with deep learning algorithms. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 97850Z.

Suzani, A., Rasoulian, A., Seitel, A., Fels, S., Rohling, R., Abolmaesumi, P., 2015. Deep learning for automatic localization, identification, and segmentation of vertebral bodies in volumetric mr images. In: Medical Imaging. Vol. 9415 of Proceedings of the SPIE. p. 941514.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2014. Going deeper with convolutions. arXiv:1409.4842.

Tachibana, R., Näppi, J. J., Hironaka, T., Kim, S. H., Yoshida, H., 2016. Deep learning for electronic cleansing in dual-energy ct colonography. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. p. 97851M.

Tajbakhsh, N., Gotway, M. B., Liang, J., 2015a. Computer-aided pulmonary embolism detection using a novel vessel-aligned multiplanar image representation and convolutional neural networks. In: Med Image Comput Comput Assist Interv. Vol. 9350 of Lect Notes Comput Sci. pp. 62–69.

Tajbakhsh, N., Gurudu, S. R., Liang, J., 2015b. A comprehensive computer-aided polyp detection system for colonoscopy videos. In: Inf Process Med Imaging. Vol. 9123 of Lect Notes Comput Sci. pp. 327–338.

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B., Liang, J., 2016. Convolutional neural networks for medical image analysis: Fine tuning or full training? IEEE Trans Med Imaging 35 (5), 1299–1312.

Tarando, S. R., Fetita, C., Faccinetto, A., Yves, P., 2016. Increasing CAD system efficacy for lung texture analysis using a convolutional network. In: Medical Imaging. Vol. 9785 of Proceedings of the SPIE. pp. 97850Q–97850Q.

Teikari, P., Santos, M., Poon, C., Hynynen, K., 2016. Deep learning convolutional networks for multiphoton microscopy vasculature segmentation. arXiv:1606.02382.

Teramoto, A., Fujita, H., Yamamuro, O., Tamaki, T., 2016. Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. Med Phys 43, 2821–2827.

Thong, W., Kadoury, S., Piché, N., Pal, C. J., 2016. Convolutional networks for kidney segmentation in contrast-enhanced CT scans. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–6.

Tran, P. V., 2016. A fully convolutional neural network for cardiac segmentation in short-axis MRI. arXiv:1604.00494.

Turkki, R., Linder, N., Kovanen, P. E., Pellinen, T., Lundin, J., 2016. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. Journal of pathology informatics 7, 38.

Twinanda, A. P., Shehata, S., Mutter, D., Marescaux, J., de Mathelin, M., Padoy, N., 2017. Endonet: A deep architecture for recognition tasks on laparoscopic videos. IEEE Trans Med Imaging 36, 86–97.

van der Burgh, H. K., Schmidt, R., Westeneng, H.-J., de Reus, M. A., van den Berg, L. H., van den Heuvel, M. P., 2017. Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis. NeuroImage. Clinical 13, 361–369.

van Ginneken, B., Setio, A. A., Jacobs, C., Ciompi, F., 2015. Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. In: IEEE Int Symp Biomedical Imaging. pp. 286–289.

van Grinsven, M. J. J. P., van Ginneken, B., Hoyng, C. B., Theelen, T., Sánchez, C. I., 2016. Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images. IEEE Trans Med Imaging 35 (5), 1273–1284.

van Tulder, G., de Bruijne, M., 2016. Combining generative and discriminative representation learning for lung CT analysis with

convolutional Restricted Boltzmann Machines. IEEE Trans Med Imaging 35 (5), 1262–1272.

Veta, M., van Diest, P. J., Pluim, J. P. W., 2016. Cutting out the middleman: measuring nuclear area in histopathology slides without segmentation. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 632–639.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., Manzagol, P.-A., 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. J Mach Learn Res 11, 3371–3408.

Vivanti, R., Ephrat, A., Joskowicz, L., Karaaslan, O., Lev-Cohain, N., Sosna, J., 2015. Automatic liver tumor segmentation in follow-up ct studies using convolutional neural networks. In: Proc. Patch-Based Methods in Medical Image Processing Workshop, MICCAI.–2015. pp. 54–61.

Wang, C., Elazab, A., Wu, J., Hu, Q., Nov. 2016a. Lung nodule classification using deep feature fusion in chest radiography. Comput Med Imaging Graph.

Wang, C., Yan, X., Smith, M., Kochhar, K., Rubin, M., Warren, S. M., Wrobel, J., Lee, H., 2015. A unified framework for automatic wound segmentation and analysis with deep convolutional neural networks. In: Conf Proc IEEE Eng Med Biol Soc. pp. 2415–2418.

Wang, D., Khosla, A., Gargeya, R., Irshad, H., Beck, A. H., 2016b. Deep learning for identifying metastatic breast cancer. arXiv:1606.05718.

Wang, G., 2016. A perspective on deep imaging. IEEE Access 4, 8914–8924.

Wang, H., Cruz-Roa, A., Basavanhally, A., Gilmore, H., Shih, N., Feldman, M., Tomaszewski, J., Gonzalez, F., Madabhushi, A., 2014. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features. J Med Imaging 1, 034003.

Wang, J., Ding, H., Azamian, F., Zhou, B., Iribarren, C., Molloi, S., Baldi, P., 2017. Detecting cardiovascular disease from mammograms with deep learning. IEEE Trans Med Imaging.

Wang, J., MacKenzie, J. D., Ramachandran, R., Chen, D. Z., 2016c. A deep learning approach for semantic segmentation in histology tissue images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. Springer, pp. 176–184.

Wang, S., Yao, J., Xu, Z., Huang, J., 2016d. Subtype cell detection with an accelerated deep convolution neural network. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 640–648.

Wang, X., Lu, L., Shin, H.-c., Kim, L., Nogues, I., Yao, J., Summers, R., 2016e. Unsupervised category discovery via looped deep pseudo-task optimization using a large scale radiology image database. arXiv:1603.07965.

Wolterink, J. M., Leiner, T., de Vos, B. D., van Hamersvelt, R. W., Viergever, M. A., Isgum, I., 2016. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal 34, 123–136.

Worrall, D. E., Wilson, C. M., Brostow, G. J., 2016. Automated retinopathy of prematurity case detection with convolutional neural networks. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 68–76.

Wu, A., Xu, Z., Gao, M., Buty, M., Mollura, D. J., 2016. Deep vessel tracking: A generalized probabilistic approach via deep learning. In: IEEE Int Symp Biomedical Imaging. pp. 1363–1367.

Wu, G., Kim, M., Wang, Q., Gao, Y., Liao, S., Shen, D., 2013. Unsupervised deep feature learning for deformable registration of MR brain images. In: Med Image Comput Comput Assist Interv. Vol. 8150 of Lect Notes Comput Sci. pp. 649–656.

Xie, W., Noble, J. A., Zisserman, A., 2016a. Microscopy cell counting and detection with fully convolutional regression networks. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 1–10.

Xie, Y., Kong, X., Xing, F., Liu, F., Su, H., Yang, L., 2015a. Deep voting: A robust approach toward nucleus localization in microscopy images. In: Med Image Comput Comput Assist Interv. Vol. 9351 of Lect Notes Comput Sci. pp. 374–382.

Xie, Y., Xing, F., Kong, X., Su, H., Yang, L., 2015b. Beyond classification: Structured regression for robust cell detection using convolutional neural network. In: Med Image Comput Comput Assist Interv. Vol. 9351 of Lect Notes Comput Sci. pp. 358–365.

Xie, Y., Zhang, Z., Sapkota, M., Yang, L., 2016b. Spatial clockwork recurrent neural network for muscle perimysium segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Vol. 9901 of Lect Notes Comput Sci. Springer, pp. 185–193.

Xing, F., Xie, Y., Yang, L., 2016. An automatic learning-based framework for robust nucleus segmentation. IEEE Trans Med Imaging 35 (2), 550–566.

Xu, J., Luo, X., Wang, G., Gilmore, H., Madabhushi, A., 2016a. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing 191, 214–223.

Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., Madabhushi, A., 2016b. Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. IEEE Trans Med Imaging 35, 119–130.

Xu, T., Zhang, H., Huang, X., Zhang, S., Metaxas, D. N., 2016c. Multimodal deep learning for cervical dysplasia diagnosis. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 115–123.

Xu, Y., Li, Y., Liu, M., Wang, Y., Lai, M., Chang, E. I.-C., 2016d. Gland instance segmentation by deep multichannel side supervision. arXiv:1607.03222.

Xu, Y., Mo, T., Feng, Q., Zhong, P., Lai, M., Chang, E. I. C., 2014. Deep learning of feature representation with multiple instance learning for medical image analysis. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1626–1630.

Xu, Z., Huang, J., 2016. Detecting 10,000 Cells in one second. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 676–684.

Xue, D.-X., Zhang, R., Feng, H., Wang, Y.-L., 2016. CNN-SVM for microvascular morphological type recognition with data augmentation. J Med Biol Eng 36, 755–764.

Yan, Z., Zhan, Y., Peng, Z., Liao, S., Shinagawa, Y., Zhang, S., Metaxas, D. N., Zhou, X. S., 2016. Multi-instance deep learning: Discover discriminative local anatomies for bodypart recognition. IEEE Trans Med Imaging 35 (5), 1332–1343.

Yang, D., Zhang, S., Yan, Z., Tan, C., Li, K., Metaxas, D., 2015. Automated anatomical landmark detection on distal femur surface using convolutional neural network. In: IEEE Int Symp Biomedical Imaging. pp. 17–21.

Yang, H., Sun, J., Li, H., Wang, L., Xu, Z., 2016a. Deep fusion net for multi-atlas segmentation: Application to cardiac mr images. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 521–528.

Yang, L., Zhang, Y., Guldner, I. H., Zhang, S., Chen, D. Z., 2016b. 3d segmentation of glial cells using fully convolutional networks and k-terminal cut. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. Springer, pp. 658–666.

Yang, W., Chen, Y., Liu, Y., Zhong, L., Qin, G., Lu, Z., Feng, Q., Chen, W., 2016c. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain. Med Image Anal 35, 421–433.

Yang, X., Kwitt, R., Niethammer, M., 2016d. Fast predictive image registration. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci.

pp. 48–57.

Yao, J., Wang, S., Zhu, X., Huang, J., 2016. Imaging biomarker discovery for lung cancer survival prediction. In: Med Image Comput Comput Assist Interv. Vol. 9901 of Lect Notes Comput Sci. pp. 649–657.

Yoo, Y., Tang, L. W., Brosch, T., Li, D. K. B., Metz, L., Traboulsee, A., Tam, R., 2016. Deep learning of brain lesion patterns for predicting future disease activity in patients with early symptoms of multiple sclerosis. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 86–94.

Ypsilantis, P.-P., Siddique, M., Sohn, H.-M., Davies, A., Cook, G., Goh, V., Montana, G., 2015. Predicting response to neoadjuvant chemotherapy with pet imaging using convolutional neural networks. PLoS ONE 10 (9), 1–18.

Yu, L., Chen, H., Dou, Q., Qin, J., Heng, P. A., 2016a. Automated melanoma recognition in dermoscopy images via very deep residual networks. IEEE Trans Med Imaging, in press.

Yu, L., Guo, Y., Wang, Y., Yu, J., Chen, P., Nov. 2016b. Segmentation of fetal left ventricle in echocardiographic sequences based on dynamic convolutional neural networks. IEEE Trans Biomed Eng, in press.

Yu, L., Yang, X., Chen, H., Qin, J., Heng, P. A., 2017. Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MR images. In: Thirty-First AAAI Conference on Artificial Intelligence.

Zeiler, M. D., Fergus, R., 2014. Visualizing and understanding convolutional networks. In: European Conference on Computer Vision. pp. 818–833.

Zhang, H., Li, L., Qiao, K., Wang, L., Yan, B., Li, L., Hu, G., 2016a. Image prediction for limited-angle tomography via deep learning with convolutional neural network. arXiv:1607.08707.

Zhang, L., Gooya, A., Dong, B. H. R., Petersen, S. E., Medrano-Gracia, K. P., Frangi, A. F., 2016b. Automated quality assessment of cardiac MR images using convolutional neural networks. In: SASHIMI. Vol. 9968 of Lect Notes Comput Sci. pp. 138–145.

Zhang, Q., Xiao, Y., Dai, W., Suo, J., Wang, C., Shi, J., Zheng, H., 2016c. Deep learning based classification of breast tumors with shear-wave elastography. Ultrasonics 72, 150–157.

Zhang, R., Zheng, Y., Mak, T. W. C., Yu, R., Wong, S. H., Lau, J. Y. W., Poon, C. C. Y., Jan. 2017. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. IEEE J Biomed Health Inform 21, 41–47.

Zhang, W., Li, R., Deng, H., Wang, L., Lin, W., Ji, S., Shen, D., 2015. Deep convolutional neural networks for multi-modality isointense infant brain image segmentation. NeuroImage 108, 214–224.

Zhao, J., Zhang, M., Zhou, Z., Chu, J., Cao, F., Nov. 2016. Automatic detection and classification of leukocytes using convolutional neural networks. Medical & Biological Engineering & Computing.

Zhao, L., Jia, K., 2016. Multiscale CNNs for brain tumor segmentation and diagnosis. Computational and Mathematical Methods in Medicine 2016, 8356294.

Zheng, Y., Liu, D., Georgescu, B., Nguyen, H., Comaniciu, D., 2015. 3D deep learning for efficient and robust landmark detection in volumetric data. In: Med Image Comput Comput Assist Interv. Vol. 9349 of Lect Notes Comput Sci. pp. 565–572.

Zhou, X., Ito, T., Takayama, R., Wang, S., Hara, T., Fujita, H., 2016. Three-dimensional CT image segmentation by combining 2D fully convolutional network with 3D majority voting. In: DLMIA. Vol. 10008 of Lect Notes Comput Sci. pp. 111–120.

Zhu, Y., Wang, L., Liu, M., Qian, C., Yousuf, A., Oto, A., Shen, D., Jan. 2017. MRI based prostate cancer detection with high-level representation and hierarchical classification. Med Phys, in press.

Zilly, J., Buhmann, J. M., Mahapatra, D., 2017. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. Comput Med Imaging Graph 55, 28–41.

Zreik, M., Leiner, T., de Vos, B., van Hamersvelt, R., Viergever, M., Isgum, I., 2016. Automatic segmentation of the left ventricle in cardiac CT angiography using convolutional neural networks. In: IEEE Int Symp Biomedical Imaging. pp. 40–43.