Deep learning is the trendiest tool in a computational biologist’s toolbox. This exciting class of methods, based on artificial neural networks, quickly became popular due to its competitive performance in prediction problems. In pioneering early work, applying simple network architectures to abundant data already provided gains over traditional counterparts in functional genomics, image analysis, and medical diagnostics. Now, ideas for constructing and training networks and even off-the-shelf models have been adapted from the rapidly developing machine learning subfield to improve performance in a range of computational biology tasks. Here, we review some of these advances in the last 2 years.

**Introduction**

In 2017, it is impossible to avoid the buzz around deep learning. Deep neural networks appear to be a hammer that can crack any nut put in its way, and are thus applied in nearly all areas of research and industry. Originally inspired by models of brain function, neural networks comprise layers of interconnected compute units (neurons), each calculating a simple output function from weighted incoming information (Box 1 and references therein). Given a well-chosen number of neurons and their connectivity pattern, these networks have a seemingly magical ability to learn the features of input that discriminate between classes or capture structure in the data. All that is required is plenty of training examples for learning.

There are two main reasons why deep learning is appealing to computational biologists. First, this powerful class of models can, in principle, approximate nearly any input to output mapping if provided enough data [1]. For example, if the goal is to predict where a transcription factor binds, there is no need to restrict the expressivity of the model to only consider a single sequence motif. Second, deep neural networks can learn directly from raw input data, such as bases of DNA sequence or pixel intensities of a microscopy image. Contrary to the traditional machine learning approaches, this obviates the need for laborious feature crafting and extraction and, in principle, allows using the networks as off-the-shelf black box tools. As large-scale biological data are available from high-throughput assays, and methods for learning the thousands of network parameters have matured, the time is now ripe for taking advantage of these powerful models.

Here, we present the advances in applications of deep learning to computational biology problems in 2016 and in the first quarter of 2017. There are several reviews that broadly cover the content and history of deep learning [2,3], as well as the early applications in various domains of biology [4]. We do not attempt to replicate them here, but rather highlight interesting ideas, and recent notable studies that have applied deep neural networks on genomic, image, and medical data.

**Genomics**

The main focus of deep learning applications in computational biology has been functional genomics data. Three pioneering papers [5–7] generalized the traditional position weight matrix model to a convolutional neural network (Box 1, reviewed in ref. [4]), and demonstrated the utility for a range of readouts. All these studies used a multilayer network structure to combine base instances into sequence motifs, and motif instances into more complex signatures, followed by fully connected layers to learn the informative combinations of the signatures.
**New applications to functional genomics data**

After demonstrations that deep learning models can outperform traditional approaches in functional genomics, they were widely adopted. Similar convolutional architectures have been applied to predict DNA sequence conservation [8], identify promoters [9] and enhancers [10], detect genetic variants influencing DNA methylation [11], find translation initiation sites [12], map enhancer–promoter interactions [13], and predict transcription factor binding [14]. We present a list of recent studies in the Appendix to this article.

The applications of deep neural networks are not limited to genomic sequences. For example, CODA [15] applies a convolutional neural network to paired noisy and high-quality ChIP-seq datasets to learn a generalizable model that reduces the noise caused by low cell input, low sequencing depth, and low signal-to-noise ratio. Convolutional neural networks have also been used to predict genome-wide locations of transcription start sites from DNA sequence, RNA polymerase binding, nucleosome positioning and transcriptional data [16], as well as gene expression from histone modifications [17], 3D chromatin interactions from DNA sequence and chromatin accessibility [18], DNA methylation from single-cell bisulfite sequencing data [19], and protein binding to RNA from the primary, secondary, and tertiary structures [20] or other features [21].

Fully connected neural networks (Box 1) are often used for standard feature-based classification tasks. In genomics, they have been applied to predict the expression of all genes from a carefully selected subset of landmark genes [22], predict enhancers, [23] and to distinguish active enhancers and promoters from background sequences [24]. An early study also applied an architecture with three hidden layers and 60 neurons to estimate historical effective population size and selection for a genomic segment with reasonable results [25]. However, carefully chosen summary statistics were used as input, so there were limited gains from the traditional benefit of a network being able to figure out relevant features from raw data. While demonstrating good performance, these applications do not make use of the recent advances in neural network methodologies, and we do not describe them further.

**Variant calling from DNA sequencing**

With the development of high-throughput sequencing technology, models for the produced data and errors were created in parallel [26,27] and calibrated on huge datasets [28]. Perhaps surprisingly, deep neural networks provided with plenty of data can achieve high accuracies for variant calling without explicitly modeling sources of errors. A four-layer dense network considering only information at the candidate site can achieve reasonable performance [29,30]. Poplin and colleagues further converted the read pileup at a potential variable site into a 221 × 100-pixel RGB image, and then used Inception-v2 [31], a network architecture normally applied in image analysis tasks, to call mutation status [32]. Base identity, base quality, and strand information were encoded in the color channels, and no additional data were used. This approach won one of the categories of the Food and Drug Administration administered variant calling challenge; the authors ascribe its performance to the ability to model complex dependencies between reads that other methods do not account for.

The advantage of deep neural network models also seems to hold for other sequencing modalities. nanopore sequencing calls convert currents across a membrane with an embedded 5-mer containing pore into bases. One would, thus, expect that a hidden Markov model with four-base memory describes the data adequately, but a recurrent neural network (Box 1) with arbitrary length memory performs even better [33].

**Recent improvements to convolutional models**

Building on the successes mentioned above, the basic convolutional model has been improved for accuracy, learning rate, and interpretability by incorporating additional intuition from data and ideas from machine learning literature.

**Incorporating elements of recurrent neural networks**

Three convolutional layers could capture the effects of multiple nearby regulatory elements such as transcription factor binding sites [7]. DanQ [34] replaced the second and third convolutional layers with a recurrent neural network (Box 1), leading to a better performance. In principle, using a recurrent neural network allows extracting information from sequences of arbitrary length, thus better accounting for long-range dependencies in the data. While the DanQ model consisted of convolutional, pooling, recurrent, and dense layers, DeeperBind [35] omitted the pooling layers, thus allowing them to retain complete positional information in the intermediate layers. SPEID [13] further proposed an elegant way to modify the DanQ network by taking sequence pairs, rather than single-DNA sequences, as input, to predict enhancer–promoter interactions. In an
Box 1. Common neural network models

Neuron, activation function, and neural network

Synopsis: A neuron (left) is the basic compute unit of a neural network. Given the values $x_1 \ldots x_N$ of all $N$ inputs, it calculates its total input signal by weighting them with the learned weights $w_1 \ldots w_N$. The total input $w_1 x_1 + \cdots + w_N x_N$ is then passed to an activation function [e.g. rectified linear unit, pictured, $y = \max(0, w_1 x_1 + \cdots + w_N x_N)$] or sigmoid, $y = 1/(1 + \exp(-w_1 x_1 - \cdots - w_N x_N))$ that calculates the neuron output, propagated to be the input for the next layer of neurons. In a dense, multilayer network (right), the data are fed as input to the first layer, and the output is recorded from the final layer activations (green).

Useful for: general purpose function estimation. Fully connected neurons are often employed in final layer(s) to tune the network to the required task from features calculated in previous layers.

Classical analogy: hierarchical models, generalized linear models

In-depth review: ref. [2].

Convolutional Neural Networks
Synopsis: These networks harbor special convolutional neurons (‘filters’, different colors in A, F) that are applied one by one to different parts of the input (B–E for four example image parts) with the same weights. This allows the same pattern to be matched regardless of its position in the data (different image patches in example) and therefore reduces the number of parameters that need to be learned. Convolutional networks have one or more layers of convolutional neurons that are typically followed by deeper fully connected layers to produce the output (bottom).

Useful for: learning and detecting patterns. Convolutional neurons are usually added in lower-level layers to learn location-independent patterns and pattern combinations from data.

Classical analogy: position weight matrix (DNA sequence), Gabor filters (images)

In-depth review: ref. [4]

Recurrent Neural Networks

Synopsis: Recurrent neural networks typically take sequential data as input (bottom) and harbor connections between neurons that form a cycle. This way, a ‘memory’ can form as an activation state (darkness of neuron) and be retained over the input sequence thanks to its cyclical propagation.

Useful for: modeling distant dependencies in sequential data.

Classical analogy: Hidden Markov Models

In-depth review: ref. [36].

Autoencoders
interesting application, DeepCpG [19] combined a nucleotide-level convolutional neural network with a bidirectional recurrent neural network to predict binary DNA methylation states from single-cell bisulfite sequencing data. An important caveat to the general applicability of recurrent neural networks is that they can be difficult to train, even with the recent improvements in methodology [8,36].

Reverse complement parameter sharing
Shrikumar et al. [37] noted that convolutional networks for DNA learn separate representations for the forward and reverse complement sequences. This led to more complicated and less stable models that sometimes produced different predictions from the two strands of the same sequence. To overcome these limitations, they implemented new convolutional layers that explicitly share parameters between the forward and reverse complement strands. This improved model accuracy, increased learning rate, and led to a more interpretable internal motif representation.

Incorporating prior information
A key advantage of neural networks is that, given sufficient data, they learn relevant features directly. However, this also means that it is not straightforward to incorporate prior information into the models. For example, the binding preferences for many RNA- and DNA-binding proteins are already known and cataloged [38,39]. To take advantage of this information, the authors of OrbWeaver [40] fixed the first layer convolutional filters to 1320 known transcription factor motifs and found that on their small dataset of three cell types, this
configuration outperformed a classical network that tried to learn motifs from the data. Furthermore, the fixed motifs were easier to interpret with DeepLIFT [41]. Similarly, the authors of DanQ [34] increased the accuracy of the model by initializing 50% of the convolutional filters in the first layer with known transcription factor motifs, but allowing them to change during training.

**Biological image analysis**

As some of the most impressive feats of deep neural networks have been in image analysis tasks, the expectations are high for their utility in bioimage analyses. Microscopy images are processed with manufacturer’s software (e.g. PerkinElmer Acapella) or community-driven tools such as CellProfiler [42], EBImage [43], or Fiji [44] that have evolved to user demands over many years. What capabilities have neural networks recently added to this rich existing toolbox?

**Image segmentation**

Segmentation identifies regions of interest, such as cells or nuclei, within a microscopy image, a task equivalent to classifying each pixel as being inside or outside of the region. The early neural network applications trained a convolutional network on square image patches centered on labeled pixels [45] and performed well in open challenges [46]. Recently, Van Valen et al. adopted this approach in a high-content screening setting and used it to segment both mammalian and bacterial cells [47]. Perhaps most importantly, they identified the optimal input size to the neural network to be similar to the typical size of the region of interest.

An alternative to classifying the focal pixel within its surrounding region is to perform end-to-end image segmentation. U-net [48] achieved this with a fully convolutional design, where image patch features are calculated at a range of resolutions by convolution and pooling, and then combined across the resolutions to produce a prediction for each pixel. The architecture of the network, therefore, included links that feed the early layer outputs forward to deeper layers in order to retain the localization information. Segmentation approaches have since been extended to handle 3D images by applying U-net to 2D slices from the same volume [49], and by performing 3D convolutions [50].

Recent applications of deep neural networks to segment medical imaging data have been thoroughly reviewed elsewhere [51–53]; we cover some histopathology studies in the Appendix to this article.

**Cell and image phenotyping**

Segmenting regions of interest is the starting point of biological image analysis. One desired end product is a cell phenotype, which captures cell state either qualitatively or quantitatively [54]. Previous methods for obtaining phenotypes have ranged from low-level image processing transforms that can be applied to any image (Gabor or Zernicke filters, Haralick features, a range of signal processing tools, [55]), to bespoke crafting of features that precisely capture the desired image characteristic in a given dataset [56,57] and unsupervised clustering of full images [58]. An important intermediate approach is to learn informative features from a given dataset de novo, a task that deep neural networks excel at.

A recurring phenotyping problem is to identify the subcellular localization of a fluorescent protein. Pärnamaa and Parts used convolutional neural networks with a popular design (e.g. also applied for plant phenotyping, [59]) to solve this task with high accuracy for images of single yeast cells [60] obtained in a high-content screen [56]. They employed eight convolutional layers of $3 \times 3$ filters interspersed with pooling steps, which were followed by three fully connected layers that learn the feature combinations that discriminate organelles. The learned features were interpretable, capturing organelle characteristics, and robust, allowing us to predict previously unseen organelles after training on a few examples. The authors further combined cell-level predictions into a single, highly accurate, protein classification. A team from Toronto demonstrated on the same unsegmented data that are possible to identify a localization label within a region and an image-level label with convolutional neural networks in a single step [61]. This has the advantage that only image-level labels are used, precluding the need to perform cell segmentation first. The output of the model, thus, also provides a per-pixel localization probability that could further be processed to perform segmentation.

Much of the recent effort has been in obtaining qualitative descriptions of individual cells. Convolutional neural networks could accurately detect phototoxicity [62] and cell-cycle states [63] from images. An interesting architecture predicts lineage choice from brightfield timecourse imaging of differentiating primary hematopoietic progenitors by combining convolution for individual micrographs with recurrent connections between
timepoints [64]. Markedly, the lineage commitment can be predicted up to three generations before conventional molecular markers are observed.

Instead of a discrete label, a vector of quantitative features describing the cell or image can be useful in downstream applications. One approach to calculate this representation is to re-use a network trained on colossal datasets as a feature extractor. For example, cellular microscopy images can be phenotyped using the features obtained from such pre-trained networks [65]. Alternatively, autoencoders (Box 1) attempt to reconstruct the input by a neural network with a limited number of neurons in one of the layers, similar to an independent component analysis model. Neuron activations in the smallest layer can then be used as features for other machine learning methods; importantly, these are learned from data each time. This approach has been used to aid diagnoses for schizophrenia [66], brain tumors [67], lesions in the breast tissue [68,69], and atherosclerosis [70].

**Medical diagnostics**

The ultimate goal of much of biomedical research is to help diagnose, treat, and monitor patients. The popularity of deep learning has, thus, naturally led to public–private partnerships in diagnostics, with IBM’s Watson tackling cancer and Google’s DeepMind Health teaming up with the National Health Service in the U.K. While the models are being industrialized, many interesting advances in applications occurred over the last year.

**Self-diagnosis with deep learning**

Neural networks have become universally available through mobile applications and web services. Provided useful pre-trained models, this could allow everyone to self-diagnose on their phone and only refer to the hospital for the required treatments. As a first step toward this vision, the GoogLeNet convolutional neural network [71] was re-trained on ~130 000 images of skin lesions, each labeled with a malignancy indicator from a predefined taxonomy [72]. The classification performance on held-out data was on par with that of professionally trained dermatologists. Thus, this network could be capable of instantly analyzing and diagnosing birthmark images taken from regular smartphones, allowing us to detect skin cancer cases earlier and hence increase survival rates.

The problem, however, is that any one image with a malignant lesion could be marked as benign. A natural resolution to this issue is to further endow the convolutional neural network with an uncertainty estimate of its output [73]. This estimate is obtained by applying the model on the same image many times over, but with a different set of random neurons switched off each time (‘dropout’, [74]). The larger the changes in output in response to the randomization, the higher the model uncertainty, and importantly, the larger the observed prediction error. Images with large classification uncertainty could then be sent to human experts for further inspection, or simply re-photographed.

More than images can be captured using a phone. Chamberlain et al. [75] recorded 11 627 lung sounds from 284 patients using a mobile phone application and an electronic stethoscope, and trained an autoencoder (Box 1) to learn a useful representation of the data. Using the extracted features, and 890 labels obtained via a laborious process, two support vector machine classifiers were trained to accurately recognize wheezes and crackles, important clinical markers of pulmonary disease. As a stand-alone mobile application, these models could help doctors from around the world to recognize signs of the disease. In a similar vein, deep neural networks have been applied to diagnose Parkinson disease from voice recordings [76] and to classify infant cries into ‘hunger’, ‘sleep’, and ‘pain’ classes [77].

Other clinical assays that are relatively easy to perform independently could be analyzed automatically. For example, the heart rate and QT interval of 15 children with type 1 diabetes were monitored overnight and used to accurately predict low blood glucose with a deep neural network model [78]. Aging.ai, which uses an ensemble of deep neural networks on 41 standardized blood test measurements, has been trained to predict an individual’s chronological age [79].

**Using other medical data modalities**

Computer tomography (CT) is a precise, but costly and risky procedure, while magnetic resonance imaging (MRI) is safer, but noisier. Nie et al. [80] trained a model to generate CT scan images from MRI data. To do so, they employed a two-part model, where one convolutional neural network was trained to generate CT images from MRI information, while the other was trained to distinguish between true and generated ones. As a result, the MRI images could be converted to CT scans that qualitatively and quantitatively resembled the
true versions. This is the first application of generative adversarial networks (Box 1) [81], a recently popularized method, for medical data.

Electronic health records are a prime target for medical data models. In Doctor AI, past diagnoses, medication, and procedure codes were inputted to a recurrent neural network to predict diagnoses and medication categories for subsequent visits, beating several baselines [82]. Three layers of autoencoders were used to capture hierarchical dependencies in aggregated electronic health records of 700,000 patients from the Mount Sinai data warehouse [83]. This gave a quantitative latent description of patients which improved classification accuracy, and provided a compact data representation.

A range of other medical input signals has been usefully modeled with neural networks. Al Rahhal et al. [84] trained autoencoders to learn features from electrocardiogram signals and used them to detect various heart-related disorders. As a completely different input, a video recording of a patient’s face could be used to automatically estimate pain intensity with a recurrent convolutional neural network [85]. Just over the last year, there have been reports of applying convolutional neural networks in image-based diagnostics of age-related macular degeneration [86], diabetic retinopathy [87], breast cancer [88–90], brain tumors [91,92], cardiovascular disease [93], Alzheimer’s disease [94], and many more diseases (Appendix to this article).

**Discussion**

Deep learning has already permeated computational biology research. Yet its models remain opaque, as the inner workings of the deep networks are difficult to interpret. The layers of convolutional neural networks can be visualized in various ways to understand input features they capture, either by finding real inputs that maximize the neuron outputs, e.g. [60], generating synthetic inputs that maximize the neuron output [95], or mapping inputs that the neuron output is most sensitive to (saliency map, [96]; or alternative [97]). In this manner, neurons operating on sequences could be interpreted as detecting motifs and their combinations, or neurons in image analysis networks as pattern finders. All these descriptions are necessarily qualitative, so conclusive causal claims about network performance due to capturing a particular type of signal are to be taken with a grain of salt.

Computer performance in image recognition has reached human levels, owing to the volume of available high-quality training datasets [98]. The same scale of labeled biological data is usually not obtainable, so deep learning models trained on a single new experiment are bound to suffer from overfitting. However, one can use networks pre-trained on larger datasets in another domain to solve the problem in hand. This transfer learning can be used both as a means to extract features known to be informative in other applications and as a starting point for model fine-tuning. Repositories of pre-trained models are already emerging (e.g. Caffe Model Zoo) and first examples of transfer learning have been successful [72,99], so we expect many more projects to make use of this idea in the near future.

Will deep learning make all other models obsolete? Neural networks harbor hundreds of parameters to be learned from the data. Even if sufficient training data exist to make a model that can reliably estimate them, the issues with interpretability and generalization to data gathered in other laboratories under other conditions remain. While deep learning can produce exquisitely accurate predictors, the ultimate goal of research is understanding, which requires a mechanistic model of the world.

**Summary**

- Deep learning methods have penetrated computational biology research.
- Their applications have been fruitful across functional genomics, image analysis, and medical informatics.
- While trendy at the moment, they will eventually take a place in a list of possible tools to apply, and complement, not supplement, existing approaches.
## Appendix

### Short overview of computational biology deep learning papers published until the first quarter of 2017

#### Part 1 of 7

| Name       | Title                                                                 | Architecture | Input                              | Output                      | Highlight                                                                 | Category          |
|------------|-----------------------------------------------------------------------|--------------|------------------------------------|-----------------------------|---------------------------------------------------------------------------|-------------------|
| **FUNCTIONAL GENOMICS**                               |                                                         |              |                                    |                             |                                                                           |                   |
| DeepBind   | Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning [5] | CNN          | DNA sequence                       | TF binding                  | Arbitrary length sequences                                               | DNA binding       |
| DeeperBind | DeeperBind: enhancing prediction of sequence specificities of DNA binding proteins [35] | CNN-RNN      | DNA sequence                       | TF binding                  | Sequences of arbitrary length. Adds LSTM to DeepBind model.               | DNA binding       |
| DeepSEA    | Predicting effects of noncoding variants with deep learning-based sequence model [7] | CNN          | DNA sequence                       | TF binding                  | 3-layer CNN                                                             | DNA binding       |
| DanQ       | DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences [34] | CNN-RNN      | DNA sequence                       | TF binding                  | Adds LSTM layer to DeepSEA model                                         | DNA binding       |
| TFImpute   | Imputation for transcription factor binding predictions based on deep learning [14] | CNN          | DNA sequence; ChiP-seq             | TF binding                  | Impute TF binding in unmeasured cell types                                | DNA binding       |
| Basset     | Basset: learning the regulatory code of the accessible genome with deep convolutional neural networks [6] | CNN          | DNA sequence                       | Chromatin accessibility     | Uses DNAase-seq data from 164 cell types                                  | DNA binding       |
| OrbWeaver  | Impact of regulatory variation across human iPSCs and differentiated cells [40] | CNN          | DNA sequence                       | Chromatin accessibility     | Uses known TF motifs as fixed filters in the CNN                          | DNA binding       |
| CODA       | Denoising genome-wide histone ChiP-seq with convolutional neural networks [15] | CNN          | ChiP-seq                           | ChiP-seq                    | Denoise ChiP-seq data                                                    | DNA binding       |
| DeepEnhancer | DeepEnhancer: predicting enhancers by convolutional neural networks [10] | CNN          | DNA sequence                       | Enhancer prediction         | Convert convolutional filters to PWMs, compare to motif databases         | DNA binding       |
| TIDE       | TIDE: predicting translation initiation sites by deep learning [12]   | CNN-RNN      | RNA sequence                       | Translation initiation sites (QTI-seq) | DanQ model                                                               | RNA binding       |
| ROSE       | ROSE: a deep learning based framework for predicting ribosome stalling [101] | CNN          | RNA sequence                       | Ribosome stalling (ribosome profiling) | Parallel convolutions                                                    | RNA binding       |
| iDeep      | RNA-protein binding motifs mining with a new hybrid deep learning     | CNN-DBN      | RNA sequence; Known motifs         | RNA binding proteins (CLIP-seq) | Integrate multiple diverse data sources                                  | RNA binding       |
Short overview of computational biology deep learning papers published until the first quarter of 2017

| Name      | Title                                                                 | Architecture | Input                                      | Output                                           | Highlight                                                                 | Category       |
|-----------|----------------------------------------------------------------------|--------------|--------------------------------------------|------------------------------------------------|---------------------------------------------------------------------------|----------------|
| Deepnet-rbp | A deep learning framework for modeling structural features of RNA-binding protein targets [20] | DBN          | RNA sequence secondary structure, ternary structure | RNA binding proteins (CLIP-seq) | Uses k-mer counts instead of a CNN to capture RNA sequence features        | RNA binding    |
| SPEID     | Predicting enhancer-promoter interaction from genomic sequence with deep neural networks [13] | CNN-RNN      | DNA sequence                               | Promoter-enhancer interactions               | Inspired by DanQ                                                          | 3D interactions |
| Rambutan  | Nucleotide sequence and DNase sensitivity are predictive of 3D chromatin architecture [18] | CNN          | DNA sequence, DNAase-seq Genomic distance | HI-C interactions                              | Binarised input signal                                                   | 3D interactions |
| DeepChrome | A deep learning framework for modeling structural features of RNA-binding protein targets [20] | CNN          | Histone modification (ChIP-seq)             | Gene expression                                | Binary decision: expressed or not expressed                               | Transcription  |
| FIDDLE    | FIDDLE: An integrative deep learning framework for functional genomic data inference [16] | CNN          | DNA sequence RNA-seq, NET-seq, MNAse-seq, ChIP-seq | Transcription start sites (TSS-seq) | DNA sequences alone not sufficient for prediction, other data helps      | Transcription  |
| CNNProm   | Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks [9] | CNN          | DNA sequence                               | Promoter predictions                          | Predicts promoters from DNA sequence features                            | Transcription  |
| DeepCpG   | DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning [19] | CNN-GRU      | DNA sequence scRRBS-seq                     | DNA methylation state (binary)                | Predict DNA methylation state in single cells based on sequence content (CNN) and noisy measurement (GRU) | DNA methylation |
| CpGenie   | Predicting the impact of non-coding variants on DNA methylation [11] | CNN          | DNA sequence                               | DNA methylation state (binary)                | Predict genetic variants that regulate DNA methylation                   | DNA methylation |
| DNN-HMM   | De novo identification of replication-timing domains in the human genome by deep learning [102] | CNN          | Hidden markov model (HMM) combined with deep belief network (DBN) | Replicated DNA sequencing (Repli-seq) | Predict replication timing domains from Repli-seq data                   | Other          |
| DeepCons  | Understanding sequence conservation with deep learning [8] | CNN          | DNA sequence                               | Sequence conservation                         | Works on noncoding sequences only                                        | Other          |
| GMFR-CNN  | GMFR-CNN: an integration of gapped motif feature representation and deep learning approach for enhancer prediction [103] | CNN          | DNA sequence                               | TF binding                                    | Uses data from the DeepBind paper. Integrates gapped DNA motifs (as introduced by gkm-SVM) with a convolutional neural network | DNA binding    |

Continued
### Short overview of computational biology deep learning papers published until the first quarter of 2017

#### SEQUENCE DATA ANALYSIS

| Name          | Title                                                                 | Architecture | Input            | Output                                                                 | Highlight                                                                 | Category             |
|---------------|------------------------------------------------------------------------|--------------|------------------|----------------------------------------------------------------------|---------------------------------------------------------------------------|----------------------|
| DeepVariant   | Creating a universal SNP and small indel variant caller with deep neural networks [32] | CNN          | Image            | Assignment of low confidence variant call (Illumina sequencing)        | Turns sequence, base quality, and strand information into image          | Basecalling          |
| Goby          | Compression of structured high-throughput sequencing data [104]         | Dense Features | Basecall (Illumina sequencing) | Part of wider variant calling framework                              |                                                                           | Basecalling          |
| DeepNano      | DeepNano: Deep Recurrent Neural Networks for Base Calling in MinION Nanopore Reads [33] | RNN          | Current          | Base call (nanopore sequencing)                                       | Uses raw nanopore sequencing signal                                      | Basecalling          |
|               | Deep learning for population genetic inference [25]                    | Dense Features | Effective population size; selection coefficient | Estimate multiple population genetic parameters in one model         |                                                                           | Population genetics  |

#### MEDICAL DIAGNOSTICS

| Name          | Title                                                                 | Architecture | Input            | Output                                                                 | Highlight                                                                 | Category             |
|---------------|------------------------------------------------------------------------|--------------|------------------|----------------------------------------------------------------------|---------------------------------------------------------------------------|----------------------|
| Leveraging uncertainty information from deep neural networks for disease detection [73] | BCNN         | Image (retina)   | Disease probability | For each image estimates an uncertainty of the network, if this uncertainty is too high, discards image | Medical diagnostics     |                     |
| DRIU          | Deep retinal image understanding [105]                                | CNN          | Image (retina)   | Segmentation                                                          | Super-human performance, task customised layers                          | Retinal segmentation  |
| IDx-DR X2.1   | Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning [87] | CNN          | Image (retina)   | DR stages                                                              | Added DL component into the algorithm and reported its superior performance | DR detection         |
| Deep learning is effective for classifying normal versus age-related macular degeneration OCT images [86] | CNN (VGG16)  | Image (OCT)     | Normal versus Age-related macular degeneration | Visualised salience maps to confirm that areas of high interest for the network match pathology areas | Age-related macular degeneration classification                         |                     |
| Medical image synthesis with context-aware generative adversarial networks [80] | GAN          | Image (MR patch) | CT patch         | Predicts CT image from 3D MRI, could also be used for super-resolution, image denoising etc | Medical image synthesis                                                  |                     |
| DeepAD        | DeepAD: Alzheimer’s disease classification via deep convolutional neural networks using MRI and fMRI [94] | CNN          | Image (fMRI and MRI) | AD vs NC                                                              | 99.9% accuracy for LeNet architecture, fishy                             | Alzheimer’s disease classification                                  |                     |
| Brain tumor segmentation with deep neural networks [91] | CNN          | Image (MRI)     | Segmentation of the brain | Stacked CNNs, fast implementation                                      |                                                                           | Glioblastoma          |
| Name | Title | Architecture | Input | Output | Highlight | Category |
|------|-------|--------------|-------|--------|-----------|----------|
| Brain tumor segmentation using convolutional neural networks in MRI images | CNN | Image (MRI) | Segmentation of the brain | |
| A deep learning-based segmentation method for brain tumor in MR images | SDAE + DNN | Image (MRI) | Segmentation of the brain | |
| Classification of schizophrenia versus normal subjects using deep learning | SAE + SVM | Image (3D fMRI volume) | Disease probability | Works on directly on active voxel time series without conversion | Schizophrenia classification |
| Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker | 3D CNN | Image (minimally preprocessed raw T1-weighted MRI data) | Age | Almost no preprocessing, brain age was shown to be heritable | Age prediction |
| Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography | CNN | Image (mammography + DBT) | Disease probability | Network was first trained on mammography images, then first three conv. layers were fixed while other layers were initialised and trained again on DBT | Medical diagnostics + Transfer Learning |
| Large scale deep learning for computer aided detection of mammographic lesions | CNN + RF | Image (mammography patch) | Disease probability | Combines handcrafted features with learned by CNN to train RF | Mammography lesions classification |
| Breast mass classification from mammograms using deep convolutional neural networks | CNN | Image (mammography patch) | Disease probability | Transfer learning from pre-trained CNNs | Mammography lesions classification |
| Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring | CSAE | Image (mammogram) | Segmentation and classification of lesions | Developed a novel regulariser | Mammography segmentation and classification |
| A deep learning approach for the analysis of masses in mammograms with minimal user intervention | CNN + DBN | Image (mammogram) | Benign vs malignant class | End to end approach with minimal user intervention, some small tech innovation at each stage | Mammography segmentation and classification |
| Detecting cardiovascular disease from mammograms with deep learning | CNN | Image (mammogram patch) | BAC vs normal | Using mammograms for cardiovascular disease diagnosis | Breast arterial calcifications detection |
| Lung pattern classification for interstitial lung disease using a deep convolutional neural network | CNN | Image (CT patch) | 7 ILD classes | Maybe the first attempt to characterize lung tissue with deep CNN tailored for the problem | Medical diagnostics |

Continued...
Short overview of computational biology deep learning papers published until the first quarter of 2017

| Name | Title | Architecture | Input | Output | Highlight | Category |
|------|-------|--------------|-------|--------|-----------|----------|
| Multi-source transfer learning with convolutional neural networks for lung pattern analysis [109] | CNN | Image (CT patch) | 7 ILD classes | Transfer learning + ensemble | |
| Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning [110] | CNN | Image (CT) | ILD classes and Lung Node detection | Transfer learning, many architectures, IDL and LN detection | |
| Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans [69] | SDAE | Image (US and CT ROI) | Benign vs malignant class | Used the same SDAE for both breast lesions in US images and pulmonary nodules in CT scans, concatenated handcrafted features to original ROI pixels | CAD |
| Dermatologist-level classification of skin cancer with deep neural networks [72] | CNN | Image (Skin) | Disease classes | Could be potentially used on a server side to power self-diagnosis of skin cancer | Medical diagnostics |
| Early-stage atherosclerosis detection using deep learning over carotid ultrasound images [70] | AE | Image (US) | Segmentation and classification of arterial layers | Fully automatic US segmentation | Intima-media thickness measurement |
| Fusing deep learned and hand-crafted features of appearance, shape, and dynamics for automatic pain estimation [111] | CNN + LR | Image (Face) | Pain intensity | Combines handcrafted features with learned by CNN to train Linear regressor | Pain intensity estimation |
| Recurrent convolutional neural network regression for continuous pain intensity estimation in video [85] | RCNN | Video frames | Pain intensity | | Pain intensity estimation |
| Efficient diagnosis system for Parkinson’s disease using deep belief network [76] | DBN | Sound (Speech) | Parkinson vs normal | | Parkinson diagnosis |
| Application of semi-supervised deep learning to lung sound analysis [75] | DA + 2 SVM | Sound (Lung sounds) | Sound scores | Handling small data sets with DA + potential application | Pulmonary disease diagnosis |
| Application of deep learning for recognizing infant cries [77] | CNN | Sound (Infant cry) | Class scores | | Sound classification |
| Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes [78] | DBN | ECG | Hypoglycemic episode onset | Real-time episodes detection | Hypoglycemic episodes detection |

Continued
Short overview of computational biology deep learning papers published until the first quarter of 2017

| Name   | Title                                                                 | Architecture | Input                          | Output                          | Highlight                                                                 | Category                                      |
|--------|----------------------------------------------------------------------|--------------|--------------------------------|--------------------------------|---------------------------------------------------------------------------|-----------------------------------------------|
|        | Deep learning approach for active classification of electrocardiogram signals [84] | SDAE         | ECG                            | AAMI classes                   | Uses raw ECG                                                             | Classification of electrocardiogram signals |
| AgingAI| Deep biomarkers of human aging: application of deep neural networks to biomarker development [79] | 21 DNN       | Blood test measurements        | Age                            | Online tool which could be used to collect training data, 5 biomarkers for aging | Age prediction                                |

### BIOMEDICAL IMAGE ANALYSIS

#### Image segmentation

- **DeepCell**: Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments [47]
  - Architecture: CNN
  - Input: Microscopy images
  - Output: Cell segmentations
  - Highlight: Able to segment both mammalian and bacterial cells
  - Category: Segmentation

- **U-Net**: U-Net: convolutional networks for biomedical image segmentation [48]
  - Architecture: CNN
  - Input: Biomedical images
  - Output: Segmentations
  - Highlight: Won the ISBI 2015 EM segmentation challenge
  - Category: Segmentation

- **3D U-Net**: 3D U-Net: learning dense volumetric segmentation from sparse annotation [49]
  - Architecture: CNN
  - Input: Volumetric images
  - Output: 3D Segmentations
  - Highlight: Able to quickly volumetric images
  - Category: Segmentation

- **V-Net**: V-Net: Fully convolutional neural networks for volumetric medical image segmentation [50]
  - Architecture: CNN
  - Input: Volumetric images
  - Output: 3D Segmentations
  - Highlight: Performs 3D convolutions
  - Category: Segmentation

#### Cell and image phenotyping

- **DeepYeast**: Accurate classification of protein subcellular localization from high throughput microscopy images using deep learning [60]
  - Architecture: CNN
  - Input: Microscopy images
  - Output: Yeast protein localisation classification
  - Category: Automatic Phenotyping

- **Deep machine learning provides state-of-the-art performance in image-based plant phenotyping [59]**
  - Architecture: CNN
  - Input: Plant images
  - Output: Plant section phenotyping
  - Category: Automatic Phenotyping

- **Classifying and segmenting microscopy images with deep multiple instance learning [61]**
  - Architecture: CNN
  - Input: Microscopy images
  - Output: Yeast protein localisation classification
  - Highlight: Performs multi-instance localisation
  - Category: Automatic Phenotyping

- **DeadNet**: DeadNet: identifying phototoxicity from label-free microscopy images of cells using Deep ConvNets [62]
  - Architecture: CNN
  - Input: Microscopy images
  - Output: Phototoxicity identification
  - Category: Automatic Phenotyping

- **Deep learning for imaging flow cytometry: cell cycle analysis of Jurkat cells [63]**
  - Architecture: CNN
  - Input: Single cell microscopy images
  - Output: Cell-cycle prediction
  - Category: Automatic Phenotyping

- **Prospective identification of hematopoietic lineage choice by deep learning [64]**
  - Architecture: CNN
  - Input: Brightfield time course imaging
  - Output: Hematopoietic lineage choice
  - Highlight: Lineage choice can be detected up to three generations before conventional molecular markers are observable
  - Category: Automatic Phenotyping
Abbreviations
CT, computer tomography; MRI, magnetic resonance imaging.

Funding
W.J. was supported by a grant from the Wellcome Trust [109083/Z/15/Z]. D.F. was supported by Estonian Research Council grant IUT34-4, European Union through the Structural Fund [Project No. 2014-2020.4.01.16-0271, ELIXIR], and CoE of Estonian ICT research EXCITE. L.P. was supported by the Wellcome Trust and the Estonian Research Council [IUT34-4]. K.A. was supported by a grant from the Wellcome Trust [099754/Z/12/Z].

Acknowledgements
We thank Oliver Stegle for the comments on the text.

Competing Interests
The Authors declare that there are no competing interests associated with the manuscript.

References
1 Hornik, K. (1991) Approximation capabilities of multilayer feedforward networks. Neural Networks 4, 251–257. https://doi.org/10.1016/0893-6080(91)90009-T
2 LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539
3 Schmidhuber, J. (2015) Deep learning in neural networks: an overview. Neural Networks 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003
4 Angermueller, C., Pärnamaa, T., Parts, L. and Stegle, O. (2016) Deep learning for computational biology. Mol. Syst. Biol. 12, 878. https://doi.org/10.15252/msb.20156651
5 Alipanahi, B., Delong, A., Weirauch, M.T. and Frey, B.J. (2015) Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nat. Biotechnol. 33, 831–838. https://doi.org/10.1038/nbt.3300
6 Kelley, D.R., Snoek, J. and Rinn, J.L. (2016) Basset: learning the regulatory code of the accessible genome with deep convolutional neural networks. Genome Res. 26, 990–999. https://doi.org/10.1101/gr.200535.115
7 Zhou, J. and Troyanskaya, O.G. (2015) Predicting effects of noncoding variants with deep learning-based sequence model. Nat. Methods 12, 931–934. https://doi.org/10.1038/nmeth.3547
8 Li, Y., Quang, D. and Xie, X. (2017) Understanding sequence conservation with deep learning. bioRxiv 103929 https://doi.org/10.1101/103929
9 Umarov, R.K. and Solovyev, V.V. (2017) Recognition of prokaryotic and eukaryotic promoters using convolutional deep learning neural networks. PLoS One 12, e0171410 https://doi.org/10.1371/journal.pone.0171410
10 Zhang, S., Hu, H., Jiang, T., Zhang, L. and Zeng, J. (2017) TIDE: predicting translation initiation sites by deep learning. bioRxiv 103374 https://doi.org/10.1101/103374
11 Singh, S., Yang, Y., Poczos, B. and Ma, J. (2016) Predicting enhancer-promoter interaction from genomic sequence with deep neural networks. bioRxiv 085241 https://doi.org/10.1101/085241
12 Qin, Q. and Feng, J. (2017) Imputation for transcription factor binding predictions based on deep learning. PLoS Comput. Biol. 13, e1005403 https://doi.org/10.1371/journal.pcbi.1005403
13 Koh, P.W., Piers, E. and Kundaje, A. (2017) Denoising genome-wide histone ChIP-seq with convolutional neural networks. bioRxiv https://doi.org/10.1101/052118
14 Eser, U. and Churchman, L.S. (2016) FIDDLE: an integrative deep learning framework for functional genomic data inference. bioRxiv 081380 https://doi.org/10.1101/081380
Singh, R., Lanchantin, J., Robins, G. and Qi, Y. (2016) Deepchrome: deep-learning for predicting gene expression from histone modifications. Bioinformatics 32, i639–i648. https://doi.org/10.1093/bioinformatics/btw427

Schreiber, J., Libbrecht, M., Bilmes, J. and Noble, W. (2017) Nucleotide sequence and DNaseI sensitivity are predictive of 3D chromatin architecture. bioRxiv 103614 https://doi.org/10.1101/103614

Angermueller, C., Lee, H.J., Reik, W. and Stegle, O. (2017) DeepCpG: accurate prediction of single-cell DNA methylation states using deep learning. Genome Biol. 18, 67. https://doi.org/10.1186/s13059-017-1189-z

Zhang, S., Zhou, J., Hu, H., Gong, H., Chen, L., Cheng, C. et al. (2016) A deep learning framework for modeling structural features of RNA-binding protein targets. Nucleic Acids Res. 44, e32. https://doi.org/10.1093/nar/gkw1025

Pan, X. and Shen, H.-B. (2017) RNA-protein binding motifs mining with a new hybrid deep learning based cross-domain knowledge integration approach. BMC bioinformatics 18, 126. https://doi.org/10.1186/s12859-017-1561-8

Chen, Y., Li, Y., Narayan, R., Subramanian, A. and Xie, X. (2016) Gene expression inference with deep learning. Bioinformatics 32, 1832–1839. https://doi.org/10.1093/bioinformatics/btv074

Liu, F., Li, H., Ren, C., Bo, X. and Shu, W. (2016) PELDA: predicting enhancers with a deep learning-based algorithmic framework. Sci. Rep. 6, 28517. https://doi.org/10.1038/srep28517

Li, Y., Shi, W. and Wasserman, W.W. (2016) Genome-wide prediction of cis-regulatory regions using supervised deep learning methods. bioRxiv 041616 https://doi.org/10.1101/041616

Sheehan, S. and Song, Y.S. (2016) Deep learning for population genetic inference. PLoS Comput. Biol. 12, e1004845 https://doi.org/10.1371/journal.pcbi.1004845

Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. Bioinformatics 27, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kermany, D., et al. (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303. https://doi.org/10.1101/gr.107524.110

1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M. et al. (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65. https://doi.org/10.1038/nature11632

Terracina, R. and Campagne, F. (2016) Training genotype callers with neural networks. bioRxiv 097469 https://doi.org/10.1101/097469

Terracina, R., Miserendino, L., Levine, S., Shklovskii, R., Hanson, M. and Campagne, F. (2016) Adaptive somatic mutations calls with deep learning and semi-simulated data. bioRxiv 079087 https://doi.org/10.1101/079087

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. (2016) Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA

Poplin, R., Newburger, D., Dijamco, J., Nguyen, N., Loy, D., Gross, S. et al. (2016) Creating a universal SNP and small indel variant caller with deep neural networks. bioRxiv 092890 https://doi.org/10.1101/092890

Boža, V., Brejová, B. and Vinař, T. (2017) Deepnanc: deep recurrent neural networks for base calling in MinION nanopore reads. PLoS ONE 12, e0178751 https://doi.org/10.1371/journal.pone.0178751

Quang, D. and Xie, X. (2016) DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. Nucleic Acids Res. 44, e107. https://doi.org/10.1093/nar/gkw226

Hassannazadeh, H.R. and Wang, M.D. (2016) DeepLearn: enhancing prediction of sequence specificities of DNA binding proteins. https://arxiv.org/abs/1611.05777

Lipton, Z.C., Berkowitz, J. and Elkan, C. (2015) A critical review of recurrent neural networks for sequence learning. https://arxiv.org/abs/1506.00019

Shrikumar, A., Greenside, P. and Kundaje, A. (2017) Reverse-complement parameter sharing improves deep learning models for genomics. bioRxiv 103663 https://doi.org/10.1101/103663

Matheier, A., Fornes, O., Arenillas, D.J., Chen, C.-Y., Denay, G., Lee, J. et al. (2016) JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 44, D101–D115. https://doi.org/10.1093/nar/gkw1176

Wiegrath, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Dräzen, P. et al. (2014) Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158, 1431–1443. https://doi.org/10.1016/j.cell.2014.08.009

Banovčič, N.E., Li, Y.L., Raj, A., Ward, M.C., Greenside, P., Calderon, D. et al. (2016) Impact of regulatory variation across human iPSCs and differentiated cells. bioRxiv 091660 https://doi.org/10.1101/091660

Shrikumar, A., Greenside, P., Shcherbina, A. and Kundaje, A. (2016) Not just a black box: learning important features through propagating activation differences. https://arxiv.org/abs/1605.01713

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O. et al. (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol. 7, R100. https://doi.org/10.1186/gb-2006-7-10-r100

Pau, G., Fuchs, F., Skyler, O., Boutros, M. and Huber, W. (2010) EBImage—an R package for image processing with applications to cellular phenotypes. Bioinformatics 26, 979–987. https://doi.org/10.1093/bioinformatics/btp046

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T. et al. (2012) Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Ning, F., Delhomme, D., LeCun, Y., Piano, F., Bottou, L. and Barbanera, P.E. (2005) Toward automatic phenotyping of developing embryos from videos. IEEE Trans. Image Process. 14, 1360–1371. https://doi.org/10.1109/TIP.2005.824740

Cresan, D., Giusti, A., Gambardella, L.M. and Schmidhuber, J. (2012) Deep neural networks segment neuronal membranes in electron microscopy images. In Advances in Neural Information Processing Systems 25 (Pereira, F., Burgos, C.J.C., Bottou, L. and Weinberger, K.Q., eds), pp. 2843–2851, Curran Associates, Inc.

Van Valen, D.A., Kudt, T., Lane, K.M., Macklin, D.N., Ouach, N.T., DeFelice, M.M. et al. (2016) Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. PLoS Comput. Biol. 12, e1005177 https://doi.org/10.1371/journal.pcbi.1005177

Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention — MICCAI 2015 (Navab, N., Hornegger, J., Wells, W.M. and Frangi, A.F., eds), pp. 234–241, Springer, CRC Press. https://www.crcpress.com/Phenomics/Hancock/p/book/9781466590953

Emerging Topics in Life Sciences (2017) 1 133–150
https://doi.org/10.1042/ETLS20160025

© 2017 The Author(s). This is an open access article published by Portland Press Limited on behalf of the Biochemical Society and the Royal Society of Biology and distributed under the Creative Commons Attribution License 4.0 (CC BY).
Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T. and Ronneberger, O. (2016) 3D U-Net: learning dense volumetric segmentation from sparse annotation. In Medical Image Computing and Computer-Assisted Intervention — MICCAI 2016, pp. 424–432, Springer, Cham, Switzerland

Milletari, F., Navab, N. and Ahmadi, S.A. (2016) V-Net: fully convolutional neural networks for volumetric medical image segmentation. 2016 Fourth International Conference on 3D Vision (3DV), Stanford University, California, USA, pp. 565–571

Greenspan, H., van Ginneken, B. and Summers, R.M. (2016) Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique. IEEE Trans. Med. Imaging 35, 1153–1159 https://doi.org/10.1109/TMI.2016.2553401

Kevin Zhou, S., Greenspan, H. and Shen, D. (2017) Deep Learning for Medical Image Analysis. Academic Press

Liljens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M. et al. (2017) A survey on deep learning in medical image analysis. Med. Image Anal. 42, 60–88 https://doi.org/10.1016/j.media.2017.07.005

Heinicke, J. K. (2014) Systematic cell phenotyping. In Phenomics (John M. Hancock, ed.), pp. 86–110

Orlov, N., Shamir, L., Macura, T., Johnston, J., Eckley, D.M. and Goldberg, I.G. (2008) WND-CHARM: multi-purpose image classification using compound image transforms. Pattern Recognit. Lett. 29, 1684–1693 https://doi.org/10.1016/j.patrec.2008.04.013

Chong, Y.T., K{"o}t, J.L.Y., Friesen, H., Duffy, S.K., Cox, M.J., Moses A. et al. (2015) Yeast proteome dynamics from single cell imaging and automated analysis. Cell 161, 1413–1424 https://doi.org/10.1016/j.cell.2015.04.051

Handfield, L.-F., Strome, B., Chong, Y.T. and Moses, A.M. (2015) Local statistics allow quantification of cell-to-cell variability from high-throughput microscope images. Bioinformatics 31, 940–947 https://doi.org/10.1093/bioinformatics/btu759

Lu, A.X. and Moses, A.M. (2016) An unsupervised kNN method to systematically detect changes in protein localization in high-throughput microscopy images. Proc IEEE 110, e1058712 https://doi.org/10.1109/journal.pone.01058712

Parnamai, T. and Parts, L. (2017) Accurate classification of protein subcellular localization from high throughput microscopy images using deep learning. G3 7, 1385–1392 https://doi.org/10.1534/g3.116.033654

Kraus, O.Z., Ba, J.L. and Frey, B.J. (2016) Classifying and segmenting microscopy images with deep multiple instance learning. Bioinformatics 32, i52–i59 https://doi.org/10.1093/bioinformatics/btw252

Richmond, D., Jost, A.P.-T., Lambert, T., Waters, J. and Elliott, H. (2017) DeadNet: identifying phototoxicity from label-free microscopy images of cells using Deep ConvNets. https://arxiv.org/abs/1701.06109

Eulenberg, P., Koehler, N., Blasi, T., Filby, A., Carpenter, A.E., Rees, P. et al. (2016) Deep learning for imaging flow cytometry: cell cycle analysis of Jurkat cells. bioRxiv 081364 https://doi.org/10.1101/081364

Buggenthin, F., Buettner, F., Hoppe, P.S., Endele, M., Kroiss, M., Strasser, M. et al. (2017) Prospective identification of hematopoietic lineage choice by deep learning. Nat. Methods 14, 403–406 https://doi.org/10.1038/nmeth.4182

Pawloski, N., Caicedo, J.C., Singh, S., Carpenter, A.E. and Storkey, A. (2016) Automating morphological profiling with generic deep convolutional networks. bioRxiv 085118 https://doi.org/10.1101/085118

Patel, P., Aggarwal, P. and Gupta, A. (2016) Classification of schizophrenia versus normal subjects using deep learning. Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, New York, NY, U.S.A. ACM, pp. 28:1–28:6

Xiao, Z., Huang, R., Ding, Y., Lan, T., Dong, R., Qin, Z. et al. (2016) A deep learning-based segmentation method for brain tumor in MR images. 2016 IEEE 6th International Conference on Computational Advances in Bio and Medical Sciences (ICAABS), Atlanta, USA

Kallenberg, M., Petersen, K., Nielsen, M., Ng, A.Y., Diao, P., Igel, C. et al. (2016) Unsupervised deep learning applied to breast density segmentation using cohort image transforms. IEEE Trans. Med. Imaging 35, 1322–1331 https://doi.org/10.1109/TMI.2016.25322122

Cheng, J.-Z., Ni, D., Zhou, Y.-H., Qin, J., Tiu, C.-M., Chang, Y.-C. et al. (2016) Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. Sci. Rep. 6, 24454 https://doi.org/10.1038/srep24454

Menchón-Lara, R.-M., Sancho-Gómez, J. L. and Bueno-Crespo, A. (2016) Early-stage atherosclerosis detection using deep learning over carotid ultrasound images. Appl. Soft Comput. 49, 616–628 https://doi.org/10.1016/j.asoc.2016.08.055

Nakagawa, T., Sagedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. et al. (2016) Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, pp. 1–9

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M. et al. (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118 https://doi.org/10.1038/nature21056

Leibig, C., Aiken, V., Berens, P. and Wahl, S. (2016) Leveraging uncertainty information from deep neural networks for disease detection. bioRxiv 084210 https://doi.org/10.1101/084210

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014) Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958

Chamberlain, D., Koduvel, R., Ganelin, D., Miglani, V. and Fletcher, R.R. (2016) Application of semi-supervised deep learning to lung sound analysis. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA

Al-Fattawi, A.H., Jadari, M.H. and Ling, S.H. (2016) Efficient diagnosis system for Parkinson’s disease using deep belief network. 2016 IEEE Congress on Evolutionary Computation (CEC), Vancouver, Canada

Chang, C.-Y. and Li, J.-J. (2016) Application of deep learning for recognizing infant cries. 2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TM), Nantou, Taiwan

San, P.P., Ding, S.H. and Nguyen, H.T. (2016) Deep learning framework for detection of hypoglycemic episodes in children with type 1 diabetes. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA

Putin, E., Mamoshina, P., Aliper, A., Korzinik, M., Mokshalev, A., Kolosov, A. et al. (2016) Deep biomarkers of macular degeneration: application of deep neural networks to biomarker development. Aging 8, 1021–1033 https://doi.org/10.18632/aging.100968

Nie, D., Trullo, R., Pettitane, C., Ruan, S. and Shen, D. (2016) Medical image synthesis with context-aware generative adversarial networks. https://arxiv.org/abs/1612.05362

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. et al. (2014) Generative adversarial networks. https://arxiv.org/abs/1406.2661
82 Choi, E., Bahadori, M.T., Schuetz, A., Stewart, W.F. and Sun, J. (2016) Doctor AI: predicting clinical events via recurrent neural networks. Proceedings of the 1st Machine Learning for Healthcare Conference, Northeastern University, Boston, MA, USA, pp. 301–318
83 Miotto, R., Li, L., Kidd, B.A. and Dudley, J.T. (2016) Deep learning approach for active classification of electrocardiogram signals. Inf. Sci. 345, 340–354 https://doi.org/10.1016/j.ins.2016.01.082
84 Zhou, J., Hong, X., Su, F. and Zhao, G. (2016) Recurrent convolutional neural network regression for continuous pain intensity estimation in video. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, USA
85 Lee, C.S., Baughman, D.M. and Lee, A.Y. (2017) Deep learning is effective for classifying normal versus age-related macular degeneration OCT images. Ophthalmol. Retina 1, 322–327 https://doi.org/10.1016/j.oret.2016.12.009
86 Abramoff, M.D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Fol, J.C. et al. (2016) Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest. Ophthalmol. Vis. Sci. 57, 5200–5206 https://doi.org/10.1167/iovs.16-19964
87 Dhungel, N., Carneiro, G. and Bradley, A.P. (2017) A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Med. Image Anal. 37, 114–128 https://doi.org/10.1016/j.media.2017.01.009
88 Levy, D. and Jain, A. Breast mass classification from mammograms using deep convolutional neural networks. https://arxiv.org/abs/1612.00542
89 Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C.I., Mann, R. et al. (2017) Large scale deep learning for computer aided detection of mammographic lesions. Med. Image Anal. 35, 303–312 https://doi.org/10.1016/j.media.2016.07.007
90 Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I. et al. (2016) DeepAD: Alzheimer’s disease classification via deep convolutional neural networks using MRI and fMRI. bioRxiv https://doi.org/10.1101/067108
91 Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y. et al. (2017) Brain tumor segmentation with deep neural networks. Med. Image Anal. 35, 18–31 https://doi.org/10.1016/j.media.2016.05.004
92 Pereira, S., Pinto, A., Aves, V. and Silva, C.A. (2016) Brain tumor segmentation using convolutional neural networks in MR images. IEEE Trans. Med. Imaging 35, 1240–1251 https://doi.org/10.1109/TMI.2016.2538465
93 Wang, J., Ding, H., Bidgoli, F.A., Zhou, B., Inbarren, C., Moloi, S. et al. (2017) Detecting cardiovascular disease from mammograms with deep learning. IEEE Trans. Med. Imaging 36, 1172–1181 https://doi.org/10.1109/TMI.2017.2655486
94 Sarraf, S., DeSouza, B.D., Anderson, J. and Tootgah, G. (2017) DeepAO: Alzheimer’s disease classification via deep convolutional neural networks using MRI and fMRI. bioRxiv https://doi.org/10.1101/070441
95 Mordvintsev, A., Olah, C. and Tyka, M. (2015) Deepdream—a code example for visualizing Neural Networks. https://research.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html
96 Simonyan, K., Vedaldi, A. and Zisserman, A. (2013) Deep inside convolutional networks: visualising image classification models and saliency maps. https://arxiv.org/abs/1312.6034
97 Shriram, A., Greeneipe, P. and Kundaje, A. (2017) Learning important features through propagating activation differences. https://arxiv.org/abs/1704.02685
98 He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, pp. 770–778
99 Kraus, O.Z., Grys, B.T., Ba, J., Chong, Y., Frey, B.J., Boone, C. et al. (2017) Automated analysis of high-content microscopy data with deep learning. Mol. Syst. Biol. 13, 924 https://doi.org/10.15252/msb.20177551
100 Bengio, Y. (2009) Learning Deep Architectures for AI. Now Publishers Inc.
101 Zhang, S., Hu, H., Zhou, J., He, X., Jiang, T. and Zeng, J. (2016) ROSE: a deep learning based framework for predicting ribosome stalling. bioRxiv https://doi.org/10.1101/067108
102 Liu, F., Ren, C., Li, H., Zhou, P., Bo, X. and Shu, W. (2016) De novo identification of replication-timing domains in the human genome by deep learning. Bioinformatics 32, 641–649 https://doi.org/10.1093/bioinformatics/btv643
103 Wang, Y.S., Lee, N.K. and Omar N. (2016). GMFR-CNN.
104 Campagne, F., Doff, K.C., Chamber, N., Robinson, J.T. and Mesirov, J.P. (2013) Compression of structured high-throughput sequencing data. PLoS One 8, e79871 https://doi.org/10.1371/journal.pone.0079871
105 Maninis, K.K., Pont-Tuset, J., Arbeláez, P. and Van Gool, L. (2016) Deep retinal image understanding. In Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science, (Ourselin S., Joskowicz L., Sabuncu M., Ulna G., Wells W., eds) vol 9901. Springer. https://doi.org/10.1007/978-3-319-46723-8_17
106 Cole, J.H., Poulet, R.P.K., Tsagkraoulis, D., Caan, M.W.A., Stevens, C., Spector, T.D. and Montana, G. (2017) Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. NeuroImage 163, 115–124
107 Samala, R.K., Chan, H.-P., Hadjiiski, L., Helvie, M.A., Wei, J. and Cha, K. (2016) Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography. Med. Phys. 43, 6654 https://doi.org/10.1118/1.4967345
108 Anthimosoupolos, M., Christoudoulis, S., Enner, L., Christie, A. and Mougiakakou, S. (2016) Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Trans. Med. Imaging 35, 1207–1216 https://doi.org/10.1109/TMI.2016.2535865
109 Christoudoulis, S., Anthimosoupolos, M., Enner, L., Christie, A. and Mougiakakou, S. (2017) Multi-source transfer learning with convolutional neural networks for lung pattern analysis. IEEE J. Biomed. Health Informatics 21, 76–84 https://doi.org/10.1109/JBIH.2016.2636929
110 Shin, H.-C., Roth, H.R., Gao, M., Lu, L., Xu, Z., Nogues, I. et al. (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans. Med. Imaging 35, 1285–1298 https://doi.org/10.1109/TMI.2016.2528162
111 Egede, J., Valstar, M. and Martinez, B. (2017) Fusing deep learned and hand-crafted features of appearance, shape, and dynamics for automatic pain estimation. 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017) https://doi.org/10.1109/fg.2017.87