Pattern Recognition 83 (2018) 134–149

Contents lists available at ScienceDirect

Pattern Recognition

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

Deep learning for image-based cancer detection and diagnosis – A survey

Zilong Hu[a], Jinshan Tang[a,b,c,*], Ziming Wang[b], Kai Zhang[a,c], Ling Zhang[a], Qingling Sun[d]

[a] School of Technology, Michigan Technological University, Houghton, MI 49931, United States
[b] Department of Electrical & Computer Engineering, Michigan Technological University, Houghton, MI 49931, United States
[c] College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan 430065, China
[d] Sun Technologies & Services, LLC, Clinton, MS, 39056 United States

ARTICLE INFO

Article history:
Received 27 September 2017
Revised 28 April 2018
Accepted 13 May 2018
Available online 23 May 2018

ABSTRACT

In this paper, we aim to provide a survey on the applications of deep learning for cancer detection and diagnosis and hope to provide an overview of the progress in this field. In the survey, we firstly provide an overview on deep learning and the popular architectures used for cancer detection and diagnosis. Especially we present four popular deep learning architectures, including convolutional neural networks, fully convolutional networks, auto-encoders, and deep belief networks in the survey. Secondly, we provide a survey on the studies exploiting deep learning for cancer detection and diagnosis. The surveys in this part are organized based on the types of cancers. Thirdly, we provide a summary and comments on the recent work on the applications of deep learning to cancer detection and diagnosis and propose some future research directions.

1. Introduction

Cancer is a major reason to cause death in the world [1] and a survey performed by American Cancer Society (ACS) shows that approximate 600,920 people are expected to die from cancers in USA in 2017 [2]. Thus, fighting against the cancers is a big challenge faced by both research scientists and clinic doctors [3].

Early detection plays a key role in cancer diagnosis and can improve long-term survival rates. Medical imaging is a very important technique for early cancer detection and diagnosis. As is well known, medical imaging has been widely employed for early cancer detection, monitoring, and follow-up after the treatments [4]. However, manual interpretation of enormous number of medical images can be tedious and time consuming and easily causes human bias and mistakes. Therefore, from early 1980s, computer-aided diagnosis (CAD) systems were introduced to assist doctors in interpreting medical images to improve their efficiency [5].

In CAD systems with medical imaging, machine learning techniques are widely employed for cancer detection and diagnosis. In order to adopt machine learning techniques, feature extraction is generally a key step. Different feature extraction methods have been investigated for different imaging modalities and different cancer types [6–22]. For example, in breast cancer detection, bilateral image subtraction, difference of Gaussian, and Laplacian of Gaussian filter have been adopted as feature extractor to detect mass regions in the mammograms [6–10]. However, the previous studies mainly focus on developing good feature descriptors combined with machine learning techniques for context learning from medical images. These methods based on feature extraction have a lot of weakness. The weakness limits the further improvement of performance of the CAD systems. In order to overcome the weakness and improve the performance of the CAD systems, the importance of representation learning has been emphasized instead of feature engineering in recent years [23,24]. Deep learning is one type of representation learning techniques that learns hierarchical feature representation from image data. One advantage of deep learning is that it can generate high level feature representation directly from the raw images. In addition, with the support of massive parallel architecture, Graphic Processing Unites (GPUs), deep learning techniques have gained enormous success in many fields in recent years, i.e. image recognition, object detection, and speech recognition. For example, recent studies show that CNNs [25] achieve promise performance in cancer detection and diagnosis.

This paper aims to introduce some popular deep learning techniques and provide an overview of the applications of deep learning for cancer detection and diagnosis.

* Corresponding author.
E-mail address: jinshant@mtu.edu (J. Tang).

https://doi.org/10.1016/j.patcog.2018.05.014
0031-3203/

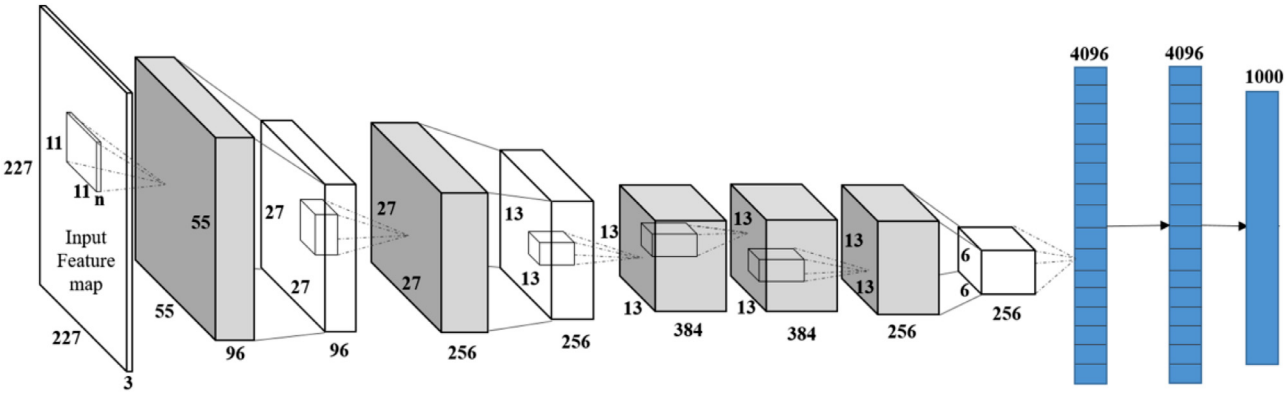

Fig. 1. Architecture of a single pipeline CNN.

2. Basic concepts of CNNs, FCNs, SSAE, and DBNs

Most of the successful image-based deep learning models were designed based on CNNs, FCNs, SSAE, and DBNs. This section introduces the basic concepts of the four architectures.

2.1. Convolutional neural networks (CCNs)

CNNs belong to feedforward neural networks where a signal flows through the network without forming cycles or loops, which can be expressed as [25]

$$F(x) = f_N(f_{N-1}(\dots(f_1(x)))) \tag{1}$$

where N denotes the number of hidden layers, and f_i represents the function in the corresponding layer i. In a typical CNN model, the main functional layers include convolutional layer, activation layer, pooling layer, fully connected layer, and predication layer.

In the convolutional layer, f is composed of multiple convolution kernels ($g^1 \dots g^{k-1}$, g^k). Each g^k represents a linear function in the kth kernel, which can be represented as follows[25]:

$$g^k(x,y) = \sum_{u=-m}^{m} \sum_{v=-n}^{n} \sum_{w=-d}^{w} W_k(u,v,w) I(x-u, y-v, z-w) \tag{2}$$

where (x, y, z) denotes the position of pixel in input I, W_k denotes the weight for the k-th kernel, m, n, and w denote the height, width, and depth of the filter. In the activation layer, f is a pixel-wise non-linear function, i.e. rectified linear unit (ReLU) [25–27]

$$f(x) = \max(0, x) \tag{3}$$

In the pooling layer, f is a layer-wise non-linear down-sampling function aiming at reducing progressively the size of the feature representation. A fully connected layer is considered to be a type of convolutional layer whose convolutional kernel has the size of 1×1. The predication layer, i.e. softmax, is often added to the last fully connected layer to compute the probabilities of I_i belonging to different classes.

An example of a single pipeline CNN model is shown in Fig. 1. It contains convolutional layers, ReLU activation layer, max-pooling layer, and fully connected layers. For each convolutional layer, it is followed by a ReLU activation layer. For the 1st, 2nd and 5th activation layer, they are followed by a max-pooling layer respectively. After the last max-pooling layer, three fully connected layers follows.

2.2. Fully convolutional networks (FCNs)

The major difference between FCNs and CNNs lies in that FCNs replace the fully connected layer with upsampling layer and deconvolutional layer [28]. Upsampling layer and deconvolutional layer can be considered as a backwards version of pooling layer and convolutional layer, and both of them are learnable. Instead of generating one probability score to each class to classify the whole image, FCNs manage to create a score map which has the exact size as the input image for each class, and classify each pixel of the image. In addition, FCNs combine the results of earlier layers through upsampling and deconvolutional results of the last layer(also called skip connection) to further improve the predication accuracy. Due to the ability of pixel-wise classification, FCNs are mainly adopted for image segmentation. The idea of upsampling, deconvolution, and skip connection also inspires the development of deep learning algorithms in many other applications [29–31]. Fig. 2 shows an example of a simple FCN model. The FCN model is based on the structure shown in Fig. 1. Different from the CNN model shown in Fig. 1, the last three fully connected layers are removed and changed to an upsampling layer and a deconvolutional layer in Fig. 2. The blue arrow indicates the skip connection where the output tensor of the second convolutional layer is concatenated to the output tensor of the last convolutional layer through upsampling. Here k denotes the number of possible classes the pixel belongs to.

2.3. Auto-Encoder (AEs)

Autoencoders (AEs) belong to another type of neural networks which are used for unsupervised learning [32]. The major goal of AEs is to learn a feature representation with lower dimensionality from the input data by itself [33]. A typical AE, as shown in Fig. 3, has a simple structure, which generally has three layers: an input layer, a hidden layer, and an output layer. The training of AEs often contains two stages: encoding stage and decoding stage. In the encoding stage, the input x is first encoded to a representation h by weight matrix $W_{x,h}$ and bias $b_{x,h}$:

$$h = \sigma\left(W_{x,h}\, x + b_{x,h}\right) \tag{4}$$

where σ is the activation function and sigmoid function is generally employed:

$$\sigma(x) = \frac{1}{1 + exp(-x)} \tag{5}$$

In the decoding stage, the representation h is decoded to reconstruct the output $\hat{x}$ through a new weight matrix $W_{h,\hat{x}}$ and a new bias $b_{h,\hat{x}}$:

$$\hat{x} = \sigma'\left(W_{h,\hat{x}}\, h + b_{h,\hat{x}}\right) \tag{6}$$

where σ' is the activation function. Note that $W_{h,\hat{x}}$ can be defined as the transpose of $W_{x,h}$, or as a new learnable parameter matrix depending on different applications. An AE is trained to minimize

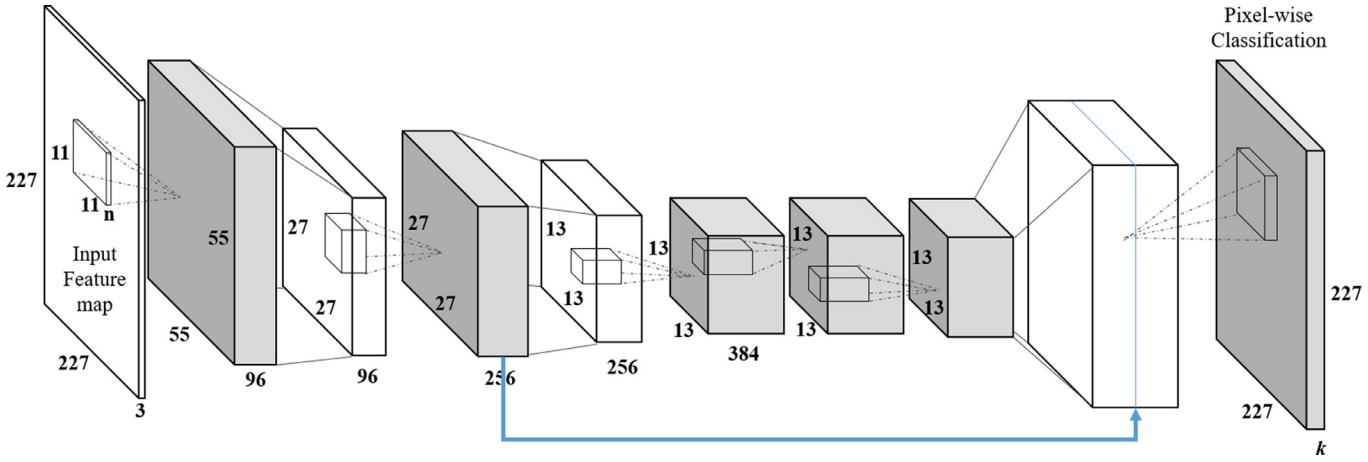

Fig. 2. Architecture of FCNs.

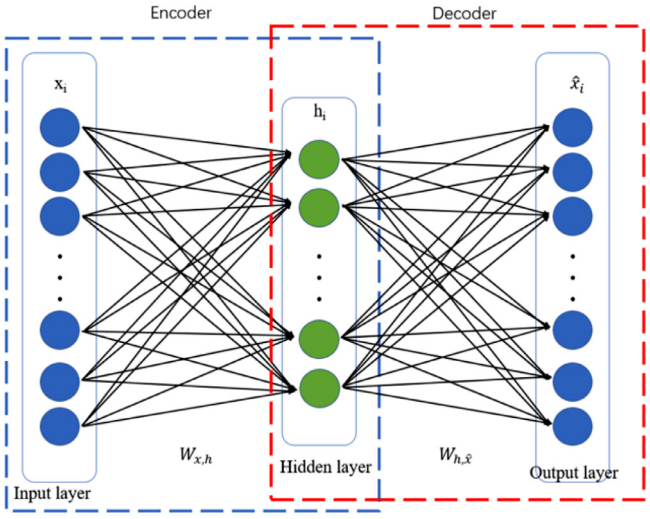

Fig. 3. Architecture of AEs.

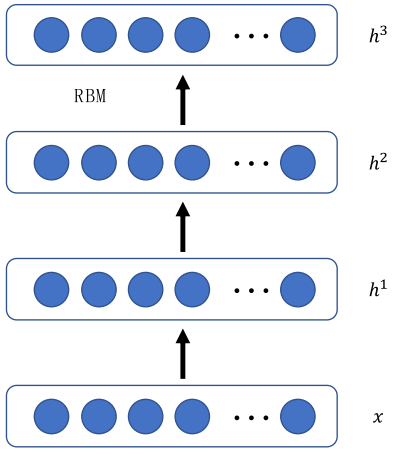

Fig. 4. Architecture of DBNs.

the following error:

$$\underset{W,b}{\operatorname{argmin}} \left|\left|x-\hat{x}\right|\right|_2^2 \tag{7}$$

Sparse autoencoders (SAEs) are a special type of AEs where sparsity is introduced into the hidden units by making the number of nodes in the hidden layer bigger than that in the input layer [33].

An SSAE is a stack of SAEs with only encoding part and they are often trained in a greedy fashion [33]: first train the hidden layer separately as a SAE, then the output of the current hidden layer is used as the input for training the successive layer. After stacking multiple SAEs together where the features extracted by a low-level SAE are fed to a high-level SAE to extract deeper features, SSAEs are able to learn deep feature representation from the data.

2.4. Deep belief network (DBNs)

Deep Belief Network (DBN) is a probabilistic generative model which is constructed by a stack of Restricted Boltzmann Machines (RBMs) instead of AEs [34]. An RBM has two layers: a visible layer and a hidden layer (see Fig. 4). In RBMs, an energy function is introduced which is expressed as:

$$E(v,h) = -a^T v - b^T h - v^T W h \tag{8}$$

where a and b are the bias vectors for the visible layer and the hidden layer respectively. An RBM is trained to maximize the product of the probability of visible vectors based on the following energy function:

$$\underset{W,a}{\operatorname{argmax}} P(v) = \frac{1}{Z}\sum_h \exp(-E(v,h)) \tag{10}$$

where Z is a partition function. Contrastive divergence algorithm, which combines Gibbs sampling and gradient descent, is used to optimize an RBM [35]. Due to the fact that the structures of RBMs and AEs share a certain degree of similarity, DBN is also trained in the greedy fashion.

In general, CNN has demonstrated its superior performance in image recognition problems. However, the input of CNN architecture is limited to relatively small images due to fully connected layers. It restricts its capability being directly used on large images. Instead, FCN does not have any fully connected layer and it can be applied to images of virtually any sizes. Compared with CNN and FCN, the major difference of DBN and SSAE is the training process. The training of DBN and SSAE includes two stages: pre-training stage and fine-tuning stage. In the pre-training stage, a network is trained in an unsupervised way and the learning for each individual layer at a time is done in a greedy fashion, which is also called greedy layer-wise training. In the fine-tuning stage, the pre-trained parameters are fine-tuned in a supervised fashion. The popularity of greedy layer-wise training in training deep networks has been reduced due to the introduction of GPU computing, ReLU, dropout, and batch normalization.

3. Cancer detection and diagnosis using deep learning techniques

This paper is intended to provide a comprehensive survey of recent studies on applying deep learning for cancer detection and diagnosis. For the survey, we organize the survey based on the type of cancers.

3.1. Breast cancer

In recent years, a bunch of papers have been published about the application of deep learning to breast cancer detection and diagnosis. In [36], Albayrak et al. developed a deep learning based feature extraction algorithm to detect mitosis in breast histopathological images. In the proposed algorithm, CNN model was used to extract features which were used to train a support vector machine (SVM) for the detection of mitosis. In [37], Spanhol et al. used AlexNet to construct an CNN model to classify benign or malignant tumors from the breast histopathological images [38]. In [39], Chen et al. proposed a deep cascade network for mitosis detection in breast histology slides. They first trained a FCN model to extract mitosis candidates from the whole histology slides and then fine-tuned a CaffeNet model [40] pre-trained on large scale images of ImageNet for the classification of mitosis. To further improve the robustness, three networks with different configurations of fully connected layers were trained to generate multiple scores/probabilities, and the scores were averaged to generate the final output. In [41], Albarqouni et al. explored deep CNNs for non-expert crowd annotations in a biomedical context. A multi-scale CNN architecture was developed. To combine a CNN with crowd annotations, an aggregation layer (AL) was introduced after the softmax layer to aggregate the prediction results with the annotation results from multiple participations. In [42], Xu et al. proposed a stacked sparse auto-encoder (SSAE) based algorithm to classify nuclei in breast cancer histopathology. During the training process, the SSAE was optimized using a greedy strategy, which one hidden layer was trained at a time and the previous layer's output was employed to train the next hidden layer. Besides the application of deep learning for breast cancer histopathology images, other studies were focused on exploiting deep learning models for the detection of breast cancer in mammographic images. In [43], Wichakam et al. proposed a system combining deep CNN and SVM for mass detection on digital mammograms. A CNN model was trained on mammographic patches and the output from the last fully connected layer was employed as the high-level feature representation of an image for training an SVM for classification. Due to the insufficient training images to train a deep CNN model, Suzuki et al. developed a transfer learning strategy to train CNN models for mass detection in mammographic images [44]. A fine-tuning strategy was applied on a pre-trained AlexNet [37] that was trained using the images from the ImageNet database [45]. In [46], Swiderski et al. presented a way to overcome the over-fitting of CNN models when the training data was limited. They used non-negative matrix factorization (NMF) and statistical self-similarity to enrich the training data. In [47], Ertosun et al. presented a deep learning based system which could determine whether a mammogram contain a mass or not and locate the masses after it was determined to contain masses. In [48], Kallenberg et al. proposed a stacked convolutional sparse auto-encoder (SCAE) to learn feature representation from mammographic images in multiple scales. In the model, a sparsity regularizer was incorporated into the model to enhance the robustness of the model. In [49], Dhungel et al. adopted a structured support vector machine to formulate a model that combined different types of potential functions, including the prior of the location, Gaussian mixture model, and a deep belief network for mass segmentation in mammograms. In another paper, Dhungel et al. proposed another algorithm for mass detection in mammograms [50]. In the proposed algorithm, a cascade of deep learning and random forest classifiers were used. In [51], Kim et al. proposed a 3-D multi-view deep CNN model to learn a latent bilateral feature representation of digital breast tomosynthesis (DBT) volume. The deep CNN model used the volume of interest (VOI) from the source volume and the VOI in the registered target volume as two separate inputs. Two individual CNNs were used to extract higher level features from two VOIs separately. Table 1 summarizes each paper reviewed above for breast cancer detection and diagnosis. In the table, we listed the specific applications, image modalities, deep learning architectures adopted, training methods, and test datasets for each paper.

3.2. Lung cancer

Deep learning has found applications in lung cancer detection and diagnosis and some research has been performed on different image modalities. Zhu et al. adopted deep convolutional neural networks (DCNNs) to predict patient's survival time directly from lung cancer pathological images [56]. Hua et al. used deep learning techniques for pulmonary nodule classification in 2D CT images [57]. They trained two end-to-end deep models, DBN and CNN, on raw lung images. Hussein et al. proposed an end-to-end trainable multi-view deep CNN model based on 3D CT images for nodule characterization [58]. A median intensity projection was used to generate three 2D patches corresponding to each dimension and patches were then concatenated to form a 3D tensor and fed to train the CNN model. Setio et al. proposed another multi-view deep CNN model based on 3D CT images for nodule classification [59]. In their model, multiple 2D patches were generated from the 3D image and each 2D patch was fed to an individual CNN model for feature extraction. Extracted features were fused together and fed to the classifiers eventually. Instead of generating 2D patches to train an CNN model, Dou et al. introduced a three-dimensional convolution neural network(3D CNN) that directly learned from 3D CT images [60]. To overcome the variant size of nodules, a multilevel model was introduced. Shen et al. proposed a multi-crop convolutional neural network (MC-CNN) to overcome the large size variation in nodule classification. Instead of training multiple CNNs parallel to generate multi-scale features, MC-CNNs introduced a multi-crop pooling operation to replace the traditional max pooling in CNN models to produce multi-scale features [61]. Paul et al. proposed using a pre-trained CNNs as a feature extractor to train classifiers [62] to detect cancer from lung CT images. In his algorithm, a CNN model that was pre-trained on a large-scale non-medical image dataset was adopted. Wang et al. [63]pointed out the possibility that a deep model could capture irrelevant information other than lung nodule and thus affected the classification results. Based on this assumption, they proposed to fuse the deep features extracted by a CNN model and other 26 hand-crafted features for the detection of lung nodules [63]. Instead of adopting a pre-trained CNN model to extract features, Hirayama et al. proposed to fine-tune the pre-trained CNN for ground glass opacity (GGO) candidate region selection [64]. Tajbakhsh et al. [65] investigated and compared the performance of massive-training artificial neural networks (MTANNs) and CNNs for lung nodule detection and distinction in CT images. Two training methods were used in the research: training with limited data and training with large datasets. His research showed that although the performance of CNN models was improved with large training dataset, MTANNs outperformed CNNs in the experiments. In addition to the CNN models, other deep learning models have also been adopted for lung cancer detection and diagnosis [66–69]. Table 2 summarizes each paper reviewed above for lung cancer detection and diagnosis.

Table 1
Summary of the papers for breast cancer detection and diagnosis.

Reference	Application	Imaging modality	Deep learning architecture	Training	Datasets
Albayrak et al [36].	Mitosis detection	Histopathology	CNN	End-to-end	MITOSATYPIA-14 [52]
Spanhol et al. [38]	Breast cancer classification	Histopathology	CNN	End-to-end	BreaKHis [53]
Chen et al. [39]	Mitosis detection	Histopathology	Hybrid (FCN + CNN)	Transfer learning	MITOSATYPIA-12,MITOSATYPIA-14 [52]
Albarqouni et al. [41]	Mitosis detection	Histopathology	CNN	End-to-end	MITOSATYPIA–13 [52]
Xu et al. [42]	Nuclei classification	Histopathology	SSAE	End-to-end	Unpublished dataset
Wichakam et al. [43]	Mass detection	Mammographic	CNN	End-to-end	INbreast [54]
Suzuki et al. [44]	Mass detection	Mammographic	CNN	Transfer learning	DDSM [55]
Swiderski et al. [46]	Lesion recognition	Mammographic	CNN	End-to-end	DDSM
Ertosun et al. [47]	Mass segmentation	Mammographic	CNN	End-to-end	DDSM
Kallenberg et al. [48]	Breast density segmentation & risk scoring	Mammographic	SSAE	End-to-end	Unpublished clinical dataset
Dhungel et al. [49]	Mass segmentation	Mammographic	DBN	End-to-end	DDSM [55] + INbreast
Dhungel et al. [50]	Mass detection	Mammographic	Hybrid (DBN + CNN)	End-to-end	DDSM + INbreast
Kim et al. [51]	Latent bilateral feature representation learning	Tomosynthesis	CNN	End-to-end	Unpublished clinical dataset

Table 2
Summary of the papers for lung cancer detection and diagnosis.

Reference	Application	Modality	Deep learning architecture	Training	Dataset
Zhu et al. [56]	Survival analysis	Histopathology	CNN	End-to-end	Unpublished dataset
Hua et al. [57]	Nodule classification	Computed tomography slices	DBN & CNN	End-to-end	LIDC-IDRI [67]
Hussein et al. [58]	Nodule characterization	Volumetric computed tomography	CNN	End-to-end	LIDC-IDRI
Setio et al. [59]	pulmonary nodules detection	Volumetric computed tomography	CNN	End-to-end	LIDC-IDRI, NODE09 [68], DLCST [69]
Dou et al. [60]	pulmonary nodules detection	Volumetric computed tomography	CNN	End-to-end	LIDC-IDRI
Shen et al. [61]	lungnodulemalignancy suspiciousness classification	Volumetric computed tomography	CNN	End-to-end	LIDC-IDRI
Paul et al. [62]	Survival prediction	Computed tomography slices	CNN	Transfer learning	Unpublished dataset
Wang et al. [63]	Lung nodule classification	Computed tomography slices	CNN	Transfer learning	JSRT [70]
Hirayama et [64]	Extraction of ground glass opacity (GGO) candidate region	Computed tomography slices	CNN	Transfer learning	LIDC
Tajbakhsh et al. [65]	Lung nodule detection and classification	Computed tomography slices	MTANN & CNN	End-to-end	Unpublished dataset
Kim et al. [66]	pulmonary nodules classification	Computed tomography slices	SSAE	End-to-end	Unpublished dataset

3.3. Skin cancer

Skin cancer is a typical common cancers and some efforts based on deep learning have been done to develop algorithms to help diagnose the disease in recent years. In [70], Pomponiu et al. proposed an algorithm for skin cancer classification by applying a pre-trained CNNs, AlexNet, to generate high-level feature representation of skin samples. The features which were extracted from the last three fully connected layers were used to train a k nearest neighbor (NN) classifier [71] for skin cancer classification. In [72], Esteva et al. studied pre-trained CNNs for skin cancer classification and a big dataset (129,450 clinical images) was used in their study. In [73], Mahbod et al. studied skin lesion classification using pre-trained CNN. In their algorithm, a pre-trained AlexNet and VGG-16 [74] architecture were adopted to extract deep features from dermoscopic images for skin lesion classification [73]. In addition to use the pre-trained CNNs, some papers also developed their own CNNs for their systems. In [75], Massod et al. proposed a semi-supervised, self-advised learning model for melanoma detection in dermoscopic images. In the proposed system, a deep belief network and two self-advised support vector machines (SA-SVMs) with radial basis function (RBF) kernel and polynomial kernel respectively were trained on three different datasets. The three datasets were generated from both labeled data and unlabeled data. Fine-tuning strategy with an exponential loss function was adopted in the training stage to maximize the separation of the labeled data. In [76], Majtner et al. proposed a classification system combining hand-crafted features and deep features for the recognition of melanoma. Two SVM classifiers, one is trained on rotated speeded-up robust features (RSurf) and local binary patterns (LBPs) extracted from grayscale images, and the other is trained on deep features extracted using CNN model trained on raw color images, are adopted to generate probability scores, and the final result was determined based on the higher scores. In [67], Sabbaghi et al. proposed a deep neural network that learned high-level image representation, and mapped images into bag-of-features (BoF) space to enhance the classification accuracy [77]. In [78], Demyanov et al. proposed using deep CNNs to detect two types of patterns (typical network and regular globules) in dermoscopic skin images. In the proposed method, the CNN was trained by the standard stochastic gradient descent algorithm. In [79], Yu et al. proposed using deep residual networks for the recognition of melanoma from dermoscopy images. The proposed system contained two networks. One was a fully convolutional residual network (FCRN), which was

Table 3
Summary of the papers for skin cancer detection.

Reference	Application	Modality	Deep learning architecture	Training	Dataset
Pomponiu et al. [71]	Skin mole lesion classificaiton	dermoscopy	CNN	Transfer learning	DermIS [82]; DermQuest [83]
Esteva et al. [72]	Dermotologist-level skin cancer classification	dermoscopy	CNN	Transfer learning	Open-access online dataset, Unpublished clinical dataset
Mahbod et al.'s [73]	Skin lesion classification	dermoscopy	CNN	Transfer learning	ISIC [84]
Massod et al.'s [75]	Skin lesion classification	dermoscopy	DBN	End-to-end	Unpublished dataset
Majtner et al. [76]	Skin lesion classification	dermoscopy	CNN	End-to-end	ISIC
Sabbaghi et al. [77]	Melanomas classification	dermoscopy	SSAE	End-to-end	Unpublished dataset
Demyanov et al. [78]	Dermoscopy patterns classification	dermoscopy	CNN	End-to-end	ISIC
Yu et al. [79]	Melanoma recognition	dermoscopy	Hybrid (FCN + CNN)	End-to-end	ISIC
Nasr-Esfahani et al. [80]	Melanoma detection	clinical photography	CNN	End-to-end	MED-NODE [85]
Sabouri et al. [81]	Lesion border detection	clinical photography	CNN	End-to-end	DermIS, DermQuest, Online dataset [86–88]

an alternative of FCN by replacing traditional convolutional layers with residual blocks. It was used to generate a score map to segment skin lesion from dermoscopy image. Next, the region of interest, containing skin lesion, was cropped and resized, and passed to a deep residual network for classification. Besides the research proposed above, in [80] Nasr-Esfahani et al. proposed a melanoma lesion detection system that feeded preprocessed clinical images to CNN models and in [81] Sabouri et al. proposed a CNNs based border detection system for skin lesions recognition. Table 3 summarizes each paper reviewed above for skin cancer detection and diagnosis.

3.4. Prostate cancer

Prostate cancer has a highly rate of diagnosed among male and is the third leading cause of death in men [89]. Successful segmentation of the prostate is key and challenging in prostate cancer radiotherapy. Deep learning technology has been used for prostate segmentation. In [89], Guo et al. unified deep feature learning with sparse patch matching for prostate segmentation. An SSAE was introduced to learn the latent feature representation from prostate MR images, and then it was further improved by fine-tuning the SSAE model in a supervised way. In [90], Yan et al. proposed to use SAE as a classifier to detect prostate cancer regions in MR images. In the proposed method, an energy minimization procedure was introduced to refine the recognition map based on the relationship among neighbor pixels. Besides the work in [72], Yan et al. also proposed another prostate segmentation model based on the features extracted using an CNN [91]. In his method, AlexNet was adopted to extract deep features from a set of pre-selected prostate proposals for boundary refinement in a finer scale. In [92], Tian et al. proposed to use deep fully convolutional networks (FCNs) for prostate segmentation in MR images. Due to insufficient training data, a pre-trained FCN model [28] was adopted and the model was fine-tuned using the medical image dataset. Yu et al. proposed a volumetric convolutional neural network with mixed residual connections for the segmentation of prostate from 3D MR images [93]. In their model, a FCN model with residual blocks was extended to enable volume-to-volume prediction. In addition, auxiliary lost was introduced during the training stage to accelerate the convergence speed. Milletari et al. trained a CNN from end to end on MRI volumes for the segmentation of prostate [94]. In order to solve the strong imbalance between interest and background voxels in 3D images, a new objective function based on Dice coefficient was introduced for the optimization of parameters in the training stage. Cheng et al. applied Holistically-Nested Networks (HNNs), an end-to-end edge detection system inspired by FCNs, for prostate segmentation [95]. CT is another image modality being used for prostate detection. Maa et al. proposed to use a patch-based deep CNN model to segment prostate from the preselected regions of interest (ROI), followed by a multi-atlas label fusion to generate the final segmentation results [96]. In addition to MR images and CT images, deep learning techniques have also been adopted for detecting prostate cancers from pathological images. Kwak et al. proposed to use a deep CNN model for lumen segmentation and generated maps for the classification of prostate cancers [97]. Gummeson et al. proposed using deep CNN model to segment microscopic images automatically. The algorithm can segment the Haemotoxylin and Eosin(H&E) stained tissue into four categories: benign, gleason grade 3, gleason grade 4, and gleason grade 5 [98]. Källén et al. proposed to use a patch-based CNN model for gleason grading on pathological images [99]. In their research, a pre-trained CNNs model developed by OverFeat in [100] was adopted to extract features from image patches to train a classifier. The classification of the whole image was determined based on all patches through majority voting. Table 4 summarizes each paper reviewed above for prostate cancer detection and diagnosis.

3.5. Brain cancer

Brain tumor is a solid mess that grows uncontrolled, and it may occur anywhere in the brain. Some research has been done on the applications of deep learning for brain cancer detection and diagnosis. Gao et al. studies deep learning for the classification of tumor cells in CT brain images [102]. In the algorithm, two convolutional neural networks, a 2D CNN and a 3D CNN were trained separately based on 2D slice images and 3D images and then the final prediction result was generated by fusing the outputs from the 2D and 3D CNNs. The proposed method was better than some previous algorithms with 2D and 3D scale-invariant feature transform (SIFT) and KAZE feature [103]. Pereira et al. presented a CNN based method for automatic MR image segmentation [104]. In the proposed algorithm, intensity normalization and augmentation were investigated for brain tumor segmentation. Havaei et al. presented a brain tumor segmentation algorithm using CNN models for MR images [105]. The proposed CNNs exploited both local features and global contextual feature and a final layer which was a convolutional implementation of fully connected layer was introduced in the architecture to significantly speed up the process.

Table 4
Summary of papers for prostate cancer detection.

Reference	Application	Modality	Deep learning architecture	Training	Dataset
Guo et al. [89]	Deformable prostate segmentation	Magnetic resonance image	SSAE	End-to-end	Unpublished dataset
Yan et al. [90]	Prostate recognition	Magnetic resonance image	SSAE	End-to-end	PROMISE12 [101]
Yan et al. [91]	Prostate segmentation	Magnetic resonance image	CNN	End-to-end	PROMISE12
Tian et al. [92]	Prostate segmentation	Magnetic resonance image	FCN	Transfer learning	Unpublished dataset
Yu et al. [93]	Prostate segmentation	3D Magnetic resonance volume	FCN	End-to-end	PROMISE12
Milletari et al. [94]	Prostate segmentation	3D Magnetic resonance volume	CNN	End-to-end	PROMISE12
Cheng et al. [95]	Prostate segmentation	Magnetic resonance image	HNN	End-to-end	Unpublished dataset
Maa et al. [96]	Prostate segmentation	Computed tomography slices	CNN	End-to-end	Unpublished dataset
Kwak et al. [97]	Lumen-based prostate cancer detection	Histopathology	CNN	End-to-end	Unpublished dataset
Gummeson et al. [98]	Gleason grading	Histopathology	CNN	End-to-end	Unpublished dataset
Källén et al. [99]	Gleason grading	Histopathology	CNN	Transfer learning	Unpublished dataset

Table 5
Summary of the papers for brain cancer detection.

Reference	Application	Modality	Deep learning architecture	Training	Dataset
Gao et al. [102]	Early detection of Alzheimer's disease	Volume computed tomography	CNN	End-to-end	Unpublished clinical dataset
Pereira et al. [104]	Brain tumor segmentation	Magnetic resonance image	CNN	End-to-end	BRATS [113,114]
Havaei et al. [105]	Brain tumor segmentation	Magnetic resonance image	CNN	End-to-end	BRATS
Zhao et al. [106]	Prostate segmentation	Magnetic resonance image	FCN	End-to-end	BRATS
Kamnitsas et al. [107]	Prostate segmentation	Magnetic resonance image	CNN	End-to-end	BRATS
Zhao et al. [108]	Brain tumor segmentation	Magnetic resonance image	CNN	End-to-end	BRATS
Paredes et al. [109]	Brain tumor segmentation	Magnetic resonance image	CNN	End-to-end	TCIA [115]
Liu et al. [110]	Feature representation learning of brain tumor	Magnetic resonance image	CNN	Transfer learning	Unpublished dataset
Ahmed et al. [112]	Brain tumor classification	Magnetic resonance image	CNN	Transfer learning	Unpublished dataset

Zhao et al. proposed a novel brain tumor segmentation method that integrated FCNs and conditional random field (CRFs) [106]. In the proposed method, a FCN was first trained using image patches and then a CRF was trained. After that, the whole framework was fine-tuned directly using image slices. Kamnitsas et al. proposed a deep 3-D CNN for brain lesion segmentation [107]. The proposed CNN used a new dense training scheme which joined the processing of adjacent image patches into one pass. After the 3-D image was segmented using deep CNN, a 3D fully connected Conditional Random Field was employed to remove the false positives. Zhao et al. presented a CNN based algorithm for 3D voxel classification [108]. In the algorithm, multi-modality information from T1, T1C, T2, and fluid-attenuated inversion recovery (FLAIR) images were combined to train the proposed CNN. 3D data was transferred to 2D slices and the patches with different scales were extracted from the 2D image and were fed to different CNNs to learn respectively. In [109], Paredes et al. explored the potential impact of different training image datasets on the performance of deep CNNs. The datasets were captured by two different institutions using different imaging equipment or different contrast protocols or different patient populations. In their experiments, MRI data of glioblastoma (GBM) patients were used for training and testing. The procedure was as follows [94]: half of the time the CNN was tested on data from the same institution that was used for training and half of the time it was tested on another institution. In the procedure, the size of the training set and testing set was assumed to be constant. Their experiments found that the accuracy for the test on the same institution was higher than the accuracy for the test on another institution. Besides brain tumor segmentation, there is some work on survival prediction using deep CNNs [110–112]. Table 5 summarizes each paper reviewed above for brain cancer detection and diagnosis.

3.6. Colonial cancer

Colorectal cancer (CRC) is the third most common cancer in both men and women. In order to help medical institutions diagnosing the colorectal cancer, researchers are trying to use deep learning methods for colonic polyp detection. Ribeiro et al. explored CNNs for colonic polyp classification [116]. They investigated different configurations of CNNs, including different filter sizes for filtering and different strides for overlapping patches. Navarro et al. also studied using CNN models for colonic polyp classification. They compared three different algorithms based on CNN models and a baseline algorithm based on hand-crafted features [117]. Zhang et al proposed using a pre-trained CNN, CaffeNet [40], trained on non-medical source domain as feature extractor and used support vector machine for the classification of colorectal polyps [118]. Yuana et al. proposed a CNN based algorithm to detect polyp automatically in colonoscopy videos [119]. In the proposed algorithm, the frames from a real-time colonoscopy video database were first pre-processed using edge detection and morphology operations. Then each connected component (edge contour) was extracted as one candidate patch. After that, a CNN with AlexNet architecture was adopted to classify each candidate into with-polyp or non-polyp class. Yu et al proposed an offline and online 3-D deep learning integration framework based on FCNs for polyp detection [120]. The offline 3D FCN aimed at learning spatio-temporal features from the colonoscopy videos while the online 3D-FCN aimed at removing false positives generated from

Table 6
Summary of the papers for colonial cancer detection and diagnosis.

Reference	Application	Modality	Deep learning architecture	Training	Dataset
Ribeiro et al. [116]	Colonic polyp classification	Colonoscopy	CNN	End-to-end	Unpublished dataset
Navarro et al. [117]	Polyp classification	Colonoscopy	CNN	Transfer learning	Unpublished dataset
Zhang et al. [118]	Colorectal polyps detection and classification	Colonoscopy	CNN	Transfer learning	Unpublished dataset
Yuana et al. [119]	Polyp detection in colonoscopy videos	Colonoscopy	CNN	Transfer learning	Unpublished dataset
Yu et al. [120]	Polyp detection in colonoscopy videos	Colonoscopy	FCN	End-to-end	ASU-Mayo Clinic Colonoscopy Video (c) Database [126]
Chowdhurya et al. [121]	Grade classification in colon cancer	Histopathology	CNN	Transfer learning	Unpublished dataset
Sirinukunwattana et al. [122]	Detection and classification of nuclei	Histopathology	CNN	End-to-end	CRCHistoPhenotypes [122]
Kashif et al. [123]	Detection of tumor cells	Histopathology	CNN	End-to-end	CRCHistoPhenotypes [122]
Xu et al. [124]	Feature representation learning with minimum manual annotation	Histopathology	CNN	End-to-end	Unpublished dataset
Haj-Hassan et al. [125]	Classification of tissue cells related to the progression of Colorectal cancer	Histopathology	CNN	End-to-end	Unpublished dataset

the offline 3D FCN. Chowdhurya et al. proposed using a pre-trained CNN model, AlexNet trained on the ImageNet database, as the feature extractor [121] for grade classification of colon cancer. The deep features obtained by the CNN were combined with the handcrafted features obtained by different feature extractors. SVM was used as the classifier in their research. Sirinukunwattana et al. proposed a spatially constrained convolutional neural network (SC-CNN) for nucleus detection in routine colon cancer histology images [122]. In the algorithm, SC-CNN was first used to compute the probability of a pixel being the center of a nucleus. Then a neighboring ensemble predictor (NEP) coupled with the CNN was developed to find the class label of the detected cell nuclei. Kashif et al. developed a new CNN model which was called SC-CNN(spatially constrained CNN) for the detection of tumor cells in histology images [123]. In the SC-CNN model, two layers were added to the standard CNN model which aimed at imposing spatial constrains to make it not only capable of extracting color characteristics but capable of extracting texture information [123]. In [124], Xu et al. investigated deep learning for automatic feature extraction. They studied two CNN based learning frameworks(One was called full supervised feature learning while the other was called unsupervised feature learning) to extract deep features from pathological image patches for classification [124]. In [125], Haj-Hassan et al. proposed using the segmentation results by an active contour model to train a CNN model for the classification of the CRC tissues from multispectral biopsy images [125]. Table 6 summarizes each paper reviewed above for colonial cancer detection and diagnosis.

3.7. Deep learning for other types of cancers

Deep learning has also found applications in other types of cancers, such as cervical cancer, bladder cancer, liver cancer, and so on. The cytology based screening method, such as Pap tests, is a common method for the detection of cervical cancer in its early stage [127]. However, cytology screening process highly depends on experts' domain knowledge and is time consuming and labor intensive [128]. Thus, computer aided methods have been investigated. Deep learning is one of these computer aided methods being investigated. Song et al. proposed a multiscale convolution neural network (MSCN) and graph-partitioning-based algorithm for the segmentation of cervical cytoplasm and nuclei in Pap smear images [129]. The proposed MSCN was used for rough segmentation and graph partitioning method was used for refining the segmentation results. In addition to [105], Song et al. also proposed a new approach to refine the boundary of overlapped cells by introducing an overlapping constraint level set [130]. Xu et al. proposed to adopt a deep CNN model for cervical dysplasia diagnosis [131]. A pre-trained CNN model was fine-tuned on a small cervigram dataset for feature learning. The extracted feature vector was then combined with non-image multimodal clinical information and fed to train a classifier.

Bladder cancer has the highest recurrence rate among all the cancers. Diagnosis and bladder cancer could be developed through deep learning methods. Cha et al. proposed a deep CNN based algorithm for bladder segmentation to aid the detection of bladder cancer in CT urography (CTU) [132]. In their algorithm, a CNN model was trained to compute the probability map of the patterns inside and outside the bladder in a CTU image. Then the probability map was used to guide level set segmentation. In another paper, Cha et al. explored a CAD system with deep CNN model for bladder cancer treatment response assessment with CT imaging. Deep CNNs with different network sizes and transfer learning [133] were investigated. Gordon et al. proposed an algorithm using CNNs with level sets to segment bladder wall from the internal bladder region and structures outside the bladder [134]. In the algorithm, a CNN model that contained two convolutional layers, followed by a pooling layer, two locally connected convolutional layers, and a fully connected layer, was used to distinguish whether the area was inside the bladder wall or not.

Liver cancer is a primary cancer, which means the cancer starts in the liver rather than migrating to the liver from another organ or section of the body. Gibson et al proposed a CNN based system to achieve automatic segmentation of liver from laparoscopic videos which could help diagnosing liver cancer [135] and Li et al presented a deep CNN based automatic segmentation procedure for liver tumor from CT images [136].

Accurate segmentation of glands is an important task for morphological statistics of gland tumors. Current CAD systems for gland analysis are facing several challenges, such as variation of glandular morphology, difficulty of separating individual objects, and degeneration of glandular structures in malignant cells [137]. Deep learning has been investigated for this task. For example, BenTaieb et al. [138] proposed a multi-objective learning method for gland segmentation and classification and Chen et al. proposed to integrate multi-level contextual features generated by CNN models to accurately segment glands from histology images [139].

The number of circulating tumor cells (CTC) in blood plays an important role in early diagnosis of cancer when tumors are invisible. Mao et al. presented a deep CNN based system for automatic imaged-based CTC detection. In the proposed system, a training

Table 7
Summary of the papers for other types of cancer detection and diagnosis.

Reference	Application	Modality	Deep learning architecture	Training	Dataset
Song et al. [129]	Segmentation of cervical cytoplasm and nuclei	Histopathology	CNN	End-to-end	Unpublished clinical dataset
Song et al. [130]	Segmentation of overlapping cervical cells	Histopathology	CNN	End-to-end	ISBI challenge 2015,Unpublished clinical dataset
Xu et al. [131]	Cervical dysplasia diagnosis	Digital cervicography	CNN	End-to-end	Unpublished dataset
Cha et al. [132]	Urinary bladder segmentation	Computed tomography slices	CNN	End-to-end	Unpublished dataset
Cha et al. [133]	Bladder cancer treatment response assessment	Computed tomography slices	CNN	Transfer learning	Unpublished dataset
Gordon et al. [134]	Segmentation of inner and outer bladder wall	Computed tomography slices	CNN	End-to-end	Unpublished dataset
Gibson et al. [135]	Liver segmentation on laparoscopic videos	Laparoscopy	CNN	End-to-end	Unpublished dataset
Li et al. [136]	Segmentation of liver tumor	Computed tomography slices	CNN	End-to-end	Unpublished dataset
BenTaieb et al. [138]	Joint classification and segmentation of colon adenocarcinoma glands	Histopathology	CNN	End-to-end	Warwick-QU [137,142]
Chen et al. [139]	Gland segmentation	Histopathology	CNN	End-to-end	Warwick-QU
Mao et al. [140]	Circulating tumor cell detection	Histopathology	CNN	End-to-end	Unpublished dataset
Xing et al. [141]	Nucleus segmentation	Histopathology	CNN	End-to-end	Unpublished dataset

method to define the classification boundary between positive and negative samples was developed [140]. Xing et al. proposed a CNN based algorithm for robust and automatic nucleus segmentation with shape preservation [141]. In the paper, CNN models were applied to obtain the likelihood map of the initial nuclei regions. Table 7 summarizes each paper reviewed above for other types of cancer detection and diagnosis.

Besides the papers mentioned above for specific cancer detection and diagnosis, there are also some review papers summarizing the image-based methods for cancer detection or diagnosis. Bernal et al. presented a complete validation study comparing eight different polyp detection methods that were proposed on endoscopic vision challenge in MICCAI 2015 [143]. Among the eight methods, one method was based on hand-crafted features; four methods were based on representation learning and all of them used CNN; three methods were based on hybrid methods. However, none of CNN based methods reviewed in the paper was published. Menze et al. [114] compared twenty different tumor segmentation algorithms on BRATS benchmark, and indicated that the highest dice scores of whole tumor segmentation can be achieved by 80% with available algorithms. However, none of the twenty reviewed algorithms used deep learning models.

4. Summary and discussion

4.1. Summary

Among the thirteen surveyed papers on breast cancer, five studies were focused on cancer diagnosis based on digital pathological images, seven studies were focused on early cancer detection based on mammograms, and one study was focused on cancer detection with Tomosynthesis [51]. In the cancer detection with pathological images, four of five studies used open source databases, including BreaKHis [53] and MITOSATYPIA [52], and only one paper used private dataset collected from hospital. For the papers on cancer detection based on mammograms, the open source databases used were INbreast [54] and DDSM [55].

Among eleven surveyed papers on lung cancer, only one study was based on digital pathological images while the others were based on CT images. Except [57] (using DBN and CNN models) and [66] (using SSAEs), all other papers used CNN models. For papers based on CT images, several open source databases, including LIDC-IDRI [67], ANODE09 [68], DLCST [69], and JSRT [70], were used for evaluation. Besides open source databases, several papers [62,65,66] used unpublished clinical datasets to evaluate their methods.

Among ten surveyed papers on skin cancer detection, eight papers were based on dermoscopic images [71–73, 75–79] and two papers were based on clinical photography [80,81]. Most of the studies based on dermoscopic images used CNN models. For clinical photography based studies, all of them used CNN models and most of them used online open resource databases [82–88].

Among the eleven surveyed papers on prostate cancer, seven papers were based on MR images [89–95]. The studies in [90,91] used the open source database from PROMISE12 challenge [101] while the study in [96] used an unpublished dataset. The rest of three papers were based on digital pathological images [97–99], and all of them used CNN models. None of them used open source databases for the evaluation of the algorithms.

Among the nine surveyed papers for the application of brain tumor, only one paper was based on CT images [102] and all the others were based on MR images. Three papers adopted CNN models for classification tasks and five papers adopted CNN models for segmentation tasks. The open source database BRATS [113,114] was used in [104–108] while the open source database TCIA [115] was used in [109] for the evaluation.

Among the ten surveyed papers colonial cancer, five papers were based on colonoscopy. Four of them adopted CNN models and only one study used FCN models. The only open source dataset used for evaluation is ASU-Mayo Clinic Colonoscopy Video (c) Database [126] used by [120]. Besides the studies based on colonoscopy, the other papers were based on digital pathological images and all of them adopted CNN models. An open source dataset CRCHistoPhenotypes [122] was used in [122] and [123] to evaluate the proposed systems.

Among the three surveyed papers on cervical cancer detection, two papers were based on digital pathological images [129,130]. Both of the two papers exploited CNN models for segmentation tasks. The other paper was based on cervigram images, and it applied CNN models for the classification task [131].

All three surveyed papers on bladder cancer detection were based on CT images. All the tested databases in the three papers were unpublished datasets collected from clinic organization. Among the two surveyed papers on liver cancer detection, one paper was based on laparoscopy images [135], and the other one was based on CT images [136]. Both papers exploited CNN models for segmentation tasks. The databases used for the evaluation of the proposed methods in those papers were also collected from clinic organizations. All the four survey papers for others were

Table 8
Summary of open source datasets used by studies in this review.

Name	Images information	Usage	Ground truth	Quantity	Accessibility	Link
MITOS-ATYPIA [52]	Breast cancer biopsy slides, stained with standard hematoxylin and eosin (H&E) dyes	Nuclear Atypia Scoring (× 20 frames); Mitotic Count (x40 frames)	Six criteria, each with 3 scores, to evaluate the nuclear atypia	408 (20 × frames); 1632 (40 × frames)	No registration required	https://mitos-atypia-14.grand-challenge.org/dataset/
BreaKHis [53]	Microscopic images of breast tumor tissue collected from 82 patients	Benign and malignant breast tumor recognition	Tumor class and tumor type	1995 (40 ×); 2081 (100 ×); 2013 (200 ×); 1820 (400 ×)	Registration required	https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/
INbreast [54]	Mammograms	Detection and diagnosis of mammary lesions	Six BI-BADS categories; Accurate contours made by specialists	90 cases from women with both breasts (4 images per case); 25 cases from mastectomy patients (2 images per case)	Registration required	http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database
DDSM [55]	Digitized film-screen mammograms	Detection and diagnosis of mammary lesions	Severity of the finding: normal, benign without callback, benign, and malignant; Marked suspicious regions included	2620 cases (4 images per case)	No registration required	http://marathon.csee.usf.edu/Mammography/Database.html
LIDC-IDRI [67]	Diagnostic and lung cancer screening thoracic computed tomography (CT) scans	Development, training, and evaluation of computer-assisted diagnostic (CAD) methods for lung cancer detection and diagnosis	Marked lesions belong to one of three categories: 'nodule ≥ 3mm', 'nodule < 3mm', and 'non-nodule ≥ 3mm'	Number of patients: 1010; Number of studies: 1308; Number of series: 1018 CT Number of images: 224,527	No registration required	https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI#940027f1a8a845d0a61a1b5b5083567e
ANODE09 [68]	Thoracic CT scans	Automatic detection of pulmonary nodules in thoracic CT scans	Example scans (5) contains information about found lung nodules (≥ 4 mm) as well as irrelevant findings	55 scans in total, including 5 example scans (with ground truth) and 50 testing scans	Registration required	https://anode09.grand-challenge.org/details/
DLCST [69]	CT screening of smokers between the age of 50 and 70 years	Evaluate if annual CT screening can reduce lung cancer mortality; evaluate psychological effects of screening; evaluate possible effects on smoking behavior		4101 participates	Restricted access	https://clinicaltrials.gov/ct2/show/NCT00496977
JSRT [70]	Chest X-ray images	Lung nodule detection	Coordinates of nodule; Diagnosis results (malignant or benign)	154 nodule images and 93 non-nodule images	Registration required	http://db.jsrt.or.jp/eng.php
DermIS-BioGPS [82]	Collection of dermoscopic image datasets	Skin lesion detection			No registration required	http://biogps.org/dataset/tag/dermis/
DermQuest [83]	Dermatology image library	Skin lesion detection	Diagnosis result		No registration required	https://www.dermquest.com/image-library/
ISIC [84]	Dermatology images	Diagnoses of skin lesions (melanoma, nevus, and seborrheic keratosis)	Skin lesions labels: melanoma, nevus, and seborrheic keratosis	2000 public training images (374 melanoma, 254 seborrheic keratosis, and 1372 benign nevi); 150 validation images; 600 testing images	No registration required	https://challenge.kitware.com/#phase/5840f53ccad3a51cc66c8dab
MED-NODE [85]	Dermatology images	Development and testing of MED-NODE system	Skin lesions labels: melanoma and nevus	170 images (70 melanoma, and 100 nevi cases)	No registration required	http://www.cs.rug.nl/~imaging/databases/melanoma_naevi/

(*continued on next page*)

Table 8 (*continued*)

Name	Images information	Usage	Ground truth	Quantity	Accessibility	Link
PROMISE12 [101]	Transversal T2-weighted MR images of prostate	Compare interactive and (semi)-automatic segmentation algorithms for MRI of the prostate	Reference segmentation result for each case	50 public training cases; 20 unseen testing cases	Registration required	https://promise12.grand-challenge.org/home/
BraTS [113,114]	Four MRI sequences: T1, T1c, T2, and FLAIR	Segmentation of intrinsically heterogeneous brain tumors; Prediction of patient overall survival	Manual segmentation results	65 MR scans, including 30 public training data	Registration required	http://www.med.upenn.edu/sbia/brats2017/data.html
TCIA collections [115]	Cancer-related imaging	Support research, development, and educational initiatives utilizing advanced medical imaging of cancer	Cancer type and/or anatomical site		No registration required	https://wiki.cancerimagingarchive.net/display/Public/RIDER+NEURO+MRI
ASU-Mayo Clinic Colonoscopy Video (c) Database [126]	Colonoscopy videos	Automatic Polyp Detection in Colonoscopy Videos.	Segmentation of polyp	20 short colonoscopy videos for training; 18 videos for testing	Registration required	https://polyp.grand-challenge.org/site/Polyp/AsuMayo/
CRCHistoPhenotypes [122]	H&E stained histology images of colorectal adenocarcinomas	Detection and Classification of Nuclei in Routine Colon Cancer Histology Images	Marked nuclei at/around the center, associated with label class	100 images, including 29,756 marked nuclei; 22,444 nuclei have associated class label	No registration required	https://warwick.ac.uk/fac/sci/dcs/research/tia/data/crchistolabelednucleihe/
Warwick-QU [137,142]	Images of Hematoxylin and Eosin (H&E) stained slides	Segmentation of glands	Segmentation results	165 images, including 85 training data and 80 testing data	No registration required	https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/about/

based on digital pathological images. Two studies used open source database Warwick-QU [137,142] to evaluate their methods in the papers.

Note that 37 of 76 surveyed papers used open source databases, the rest of papers tested and evaluated their methods on the datasets collected from medical organizations, such as medical universities, hospitals, and cancer research centers. The lack of large training data in the open source datasets was probably the main reason why those papers chose to use clinical data. Another possible reason is that the information, except images data, provided by open source datasets was limited for some specific applications. Table 8 presented a summary of open source datasets used in the studies reviewed in this paper. Note that the information of some datasets listed in the table is based on the newly updated versions, which may be different from the version being used in the studies reviewed in this paper. For example, the description of dataset in BraTS challenge is based on BraTS 2017, which is significantly different from the data provided during the previous BraTS challenges.

Among all 76 studies surveyed in this paper, 63 studies adopted CNN models, 6 studies adopted FCN models, 6 studies adopted SSAE models, 4 studies adopted DBN models, and 3 studies proposed hybrid model based on multiple types of deep learning models. Each study and the corresponding deep learning methods being used in the paper are listed in Table 9. Comparison results show that CNN has been widely studied and adopted for different types of cancer detection tasks.

Table 9
Summary of deep learning models and studies used in the paper.

Deep learning models	References	Total number
CNN	[36,38,39,41,43,44,46,47,50,51,56–65,71–73,76,78–81,91,94,96–99,102,104,105,107–110,112,116–119,121–125,129–136,138–141]	63
FCN	[39,79,92,93,106,120]	6
SSAE	[42,48,66,77,89,90]	6
DBN	[49,50,57,75]	4

4.2. Discussion and future directions

From the surveyed papers, we found that one big challenge of training deep learning models for medical image analysis was the lack of large training datasets. Although the popularization of picture archiving and communication system (PACS) in hospitals has helped gather millions of medical images, most of them include confidential information of patients and they are stored in hospitals. In order to make those datasets available for research uses, more efforts are needed on those data, such as de-identification and data transportation. Many surveyed papers used different datasets collected from hospitals or cancer research organizations to test and evaluate deep learning models. The main drawback is that it is difficult to compare the performance of deep learning models among different studies. Open source medical image datasets have been provided for public research on different types of cancers in recent years. Among all the surveyed papers, the work of Esteva et al. built up the largest dataset by collecting images from different resources and the dataset included 127,463 training and validation images and 1,942 biopsy-labelled test images. However, it is worth noting that, for some types of cancer research, the number of case studies (patients) in the dataset is too small [72]. In addition, some of open source datasets only contained raw image data, extra efforts from expert domain are required to generate ground truth for the purpose of the model training as well as evaluation. Therefore, it is desirable to build up larger and more systematic open source datasets for different applications.

Given the current limitation of the training dataset, some surveyed papers proposed using data augmentation approaches, such as special filtering, adding noise, rotation, cropping to increase the size of training data. Transfer learning is another way to avoid overfitting given a small training dataset to train deep learning model, and they were used in 17 studies among the surveyed papers. Among those studies, 10 studies directly used pre-trained deep learning model as the feature extractor to extract high-level features to train a classifier [62,63,71,73,99,110,117,118,121,133], and 7 studies fine-tuned the pre-trained models by changing the last layer and trained models on new medical datasets [39,44,64,72,92,112,119]. In addition, complex CNN models were also used for transfer learning, i.e. Esteva et al. [72] used a pre-trained inception v3 in transfer learning, which included 98 convolutional layers and 14 pooling layers. It is expected to see more studies applying more complex deep learning models and transfer learning for cancer detection in the future studies. The study of Cha et al. [133] investigated and compared two CNN models as feature extractor, one was pre-trained on natural scene images and the other was pre-trained on bladder images, and the experimental results indicated that the feature extractors trained on bladder images achieved better performance. It indicates the promise potential of end-to-end trainable deep learning model in the future with large training data.

We note that there were few studies investigating the effect of intrinsic characteristics (contrast, signal to noise ratio, etc) of medical image on the performance of deep learning models. It was mentioned in [144] that most deep neural network architectures performed poorly on blur and noisy nature scene images, even if they were trained on low quality images. Thus how to improve the performance of deep learning based cancer detection and diagnosis when the images have low contrast and signal to noise ratio is an important research direction. In addition, it was mentioned in [109] that the performance of brain tumor segmentation using deep learning model suffered moderate decrease when the model was trained with multi-institutional data. Therefore, how to improve the system so as to maintain its performance on multi-institutional data is worthwhile paying more attention.

In the current transfer learning approach, the pre-trained models were mostly trained on non-medical images and no research was done to investigate the potential for transferring knowledge from one medical image modality to a another medical image modality. Future work can be done in this direction for transfer learning.

Another problem about the medical image dataset is that the ratio of positive and negative in the dataset is often heavily imbalanced. Training models directly on imbalanced data may bias the prediction towards the more common classes. We found that most studies ignored this problem in the training stage. Only the work of Pereira et al. proposed to use all samples from the underrepresented classes and down-sample other classes so that the same number of samples were used to train the model [104].

Another big challenge using CNN models for cancer detection was the size variation of target objects within the images. To overcome this problem, several studies proposed to train the same CNN models using different scales of image data, and fused the outputs of multiple models to gain final result [41,48,50]. The study of Shen et al. [61] introduced a multi-crop pooling operation to replace the traditional pooling layer to directly capture multi-scale features from the image. We note that there is a lack of studies to compare the performance and efficiency of different methods. It is desirable to develop approaches robust to the size variation of target objects.

Several studies investigated the potential of combining multi-modality information for cancer detection. The study of Xu et al. combined extracted deep features and non-image multimodal clinical information to train classifier [131]. Zhao et al. trained CNN models using four different types of MR images and combined their results for tumor segmentation [108]. We expect to see more studies applying multi-modality information for cancer detection in the future research.

The studies surveyed in this review used different datasets as well as different imaging modalities, thus it is difficult to conduct a comparison on the performance of all the methods (specificity and sensitivity) with clinical standard practice for cancer diagnosis. However, the ground truth of most datasets are provided through expert consensus or pathology report information, therefore it is reasonable to rely on the evaluation results to demonstrate the potential of deep learning algorithms for cancer detection and diagnosis.

In summary, deep learning has shown a significant improvement compared with many other machine learning methods in different applications. The success of deep learning in natural scene image classification and segmentation stimulates the research of adopting it in image-based cancer detection and diagnosis. One major advantage of deep learning is that it reduce the need of feature engineering, which is one of the most complicated and time-consuming parts in machine learning practice, especially in processing redundant image data. In addition, it is relatively easy to adapt or modify existing deep learning architectures on new applications. However, it is worth noting that there are also some disadvantages in adopting deep learning in real practice: (1) deep learning models often require a large amount of training data to achieve superior performance than other methods. (2) Training process is extremely computational expensive and it is quite time consuming to train a deep and complex model even with the support of most powerful GPU hardware. (3) The body of trained deep learning model is like a black box, we still lack the perfect methodology to fully comprehend its deep structure.

5. Conclusion

In this paper, we surveyed most recent studies on the subject of applying deep learning techniques in image based cancer detection and diagnosis. These applications are organized in nine categories depending upon specific types of cancers, including breast cancer, lung cancer, skin cancer, prostate cancer, brain cancer, colonial cancer, cervical cancer, bladder cancer, and liver cancer. In addition, one extra category is introduced in the paper which contains studies related to applying deep learning in general image-based cancer detection. Four popular image-based deep learning models, including convolutional neural networks, fully convolutional networks, auto-encoders, and deep belief networks are highlighted in the survey. The uniqueness of past studies and some potential topics for future study are discussed.

References

[1] L.A. Torre, F. Bray, R.L. Siegel, J. Ferlay, J. Lortet–Tieulent, A. Jemal, Global cancer statistics, 2012, CA, Cancer J. Clin. 65 (2015) 87–108.
[2] Cancer facts & figures 2017, American Cancer Society, 2017, (2017).
[3] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2016, CA, Cancer J. Clin. 66 (2016) 7–30.
[4] L. Fass, Imaging and cancer: a review, Mol. Oncol. 2 (2008) 115–152.
[5] K. Doi, Computer-aided diagnosis in medical imaging: historical review, current status and future potential, Comput. Med. Imaging Graph. 31 (2007) 198–211.
[6] F.F. Yin, M.L. Giger, C.J. Vyborny, R.A. Schmidt, Computerized detection of masses in digital mammograms: automated alignment of breast images and its effect on bilateral-subtraction technique, Med. Phys. 21 (1994) 445–452.
[7] M. Beller, R. Stotzka, T. Müller, H. Gemmeke, An example-based system to support the segmentation of stellate lesions, Bildverarb. Med. 2005 (2005) 475–479.
[8] G.M. te Brake, N. Karssemeijer, J.H. Hendriks, An automatic method to discriminate malignant masses from normal tissue in digital mammograms1, Phys. Med. Biol. 45 (2000) 2843.

[9] N.H. Eltonsy, G.D. Tourassi, A.S. Elmaghraby, A concentric morphology model for the detection of masses in mammography, IEEE Trans. Med. Imag. 26 (2007) 880–889.

[10] J. Wei, B. Sahiner, L.M. Hadjiiski, H.P. Chan, N. Petrick, M.A. Helvie, M.A. Roubidoux, J. Ge, C. Zhou, Computer-aided detection of breast masses on full field digital mammograms, Med. Phys. 32 (2005) 2827–2838.

[11] S.H. Hawkins, J.N. Korecki, Y. Balagurunathan, Y. Gu, V. Kumar, S. Basu, L.O. Hall, D.B. Goldgof, R.A. Gatenby, R.J. Gillies, Predicting outcomes of nonsmall cell lung cancer using CT image features, IEEE Access 2 (2014) 1418–1426.

[12] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, C. Parmar, P. Grossmann, S. Cavalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat. Commun. 5 (2014) 1–8 4006.

[13] Y. Balagurunathan, Y. Gu, H. Wang, V. Kumar, O. Grove, S. Hawkins, J. Kim, D.B. Goldgof, L.O. Hall, R.A. Gatenby, Reproducibility and prognosis of quantitative features extracted from CT images, Transl. Oncol. 7 (2014) 72–87.

[14] F. Han, H. Wang, G. Zhang, H. Han, B. Song, L. Li, W. Moore, H. Lu, H. Zhao, Z. Liang, Texture feature analysis for computer-aided diagnosis on pulmonary nodules, J. Digit. Imaging 28 (2015) 99–115.

[15] C. Barata, J.S. Marques, M.E. Celebi, Improving dermoscopy image analysis using color constancy, Proceedings of 2014 IEEE International Conference on Image Processing (ICIP) 2014, pp. 3527–3531.

[16] C. Barata, J.S. Marques, J. Rozeira, A system for the detection of pigment network in dermoscopy images using directional filters, IEEE Trans. Biomed. Eng. 59 (2012) 2744–2754.

[17] C. Barata, M. Ruela, T. Mendonça, J.S. Marques, A bag-of-features approach for the classification of melanomas in dermoscopy images: the role of color and texture descriptors,, in: Computer Vision Techniques For the Diagnosis of Skin Cancer, Springer, 2014, pp. 49–69.

[18] M. Sadeghi, T.K. Lee, D. McLean, H. Lui, M.S. Atkins, Detection and analysis of irregular streaks in dermoscopic images of skin lesions, IEEE Trans. Med. Imaging 32 (2013) 849–861.

[19] D. Zikic, B. Glocker, E. Konukoglu, A. Criminisi, C. Demiralp, J. Shotton, O.M. Thomas, T. Das, R. Jena, S.J. Price, Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel MR, in: Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2012, pp. 369–376.

[20] R. Meier, S. Bauer, J. Slotboom, R. Wiest, M. Reyes, A hybrid model for multimodal brain tumor segmentation, Multimod. Brain Tumor Segm. (2013) 31.

[21] A. Pinto, S. Pereira, H. Correia, J. Oliveira, D.M. Rasteiro, C.A. Silva, Brain tumour segmentation based on extremely randomized forest with high-level features, Proceedings of the 37th Annual International Conference on IEEE Engineering in Medicine and Biology Society (EMBC, 2015), pp. 3037–3040.

[22] N.J. Tustison, K. Shrinidhi, M. Wintermark, C.R. Durst, B.M. Kandel, J.C. Gee, M.C. Grossman, B.B. Avants, Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTsR, Neuroinformatics 13 (2015) 209–225.

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

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

[25] B. Xu, N. Wang, T. Chen, M. Li, Empirical evaluation of rectified activations in convolutional network, arXiv preprint arXiv:1505.00853, (2015).

[26] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 807–814.

[27] D. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980, (2014).

[28] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.

[29] H. Noh, S. Hong, B. Han, Learning deconvolution network for semantic segmentation, Proceedings of the IEEE International Conference on Computer Vision2015, 1520–1528.

[30] C. Dong, C.C. Loy, K. He, X. Tang, Learning a deep convolutional network for image super-resolution, in: Proceedings of European Conference on Computer Vision, Springer, 2014, pp. 184–199.

[31] V. Jain, S. Seung, Natural image denoising with convolutional networks, Adv. Neural Inf. Process. Syst. 21 (2009) 769–776.

[32] C.M. Bishop, Neural Networks For Pattern Recognition, Oxford University Press, 1995.

[33] A. Ng, Sparse autoencoder, CS294A Lect. Notes 72 (2011) 1–19.

[34] G.E. Hinton, Deep belief networks, Scholarpedia 4 (2009) 5947.

[35] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (2006) 1527–1554.

[36] A. Albayrak, G. Bilgin, Mitosis detection using convolutional neural network based features, in: Proceedings of IEEE Seventeenth International Symposium on Computational Intelligence and Informatics (CINTI), 2016, pp. 000335–000340.

[37] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. (2012) 1097–1105.

[38] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using convolutional neural networks, in: Proceedings of 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pp. 2560–2567.

[39] H. Chen, Q. Dou, X. Wang, J. Qin, P.-A. Heng, Mitosis detection in breast cancer histology images via deep cascaded networks, AAAI (2016) 1160–1166.

[40] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, T. Darrell, Caffe: convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM international conference on Multimedia, ACM, 2014, pp. 675–678.

[41] S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, N. Navab, Aggnet: deep learning from crowds for mitosis detection in breast cancer histology images, IEEE Trans. Med. Imaging 35 (2016) 1313–1321.

[42] J. Xu, L. Xiang, R. Hang, J. Wu, Stacked Sparse Autoencoder (SSAE) based framework for nuclei patch classification on breast cancer histopathology, in: Proceedings of IEEE Eleventh International Symposium on Biomedical Imaging (ISBI), 2014, pp. 999–1002.

[43] I. Wichakam, P. Vateekul, Combining deep convolutional networks and SVMs for mass detection on digital mammograms, in: Proceedings of Eighth International Conference on Knowledge and Smart Technology (KST), 2016, pp. 239–244.

[44] S. Suzuki, X. Zhang, N. Homma, K. Ichiji, N. Sugita, Y. Kawasumi, T. Ishibashi, M. Yoshizawa, Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis, in: Proceedings of the 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 2016, pp. 1382–1386.

[45] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition CVPR 2009, 2009, pp. 248–255.

[46] B. Swiderski, J. Kurek, S. Osowski, M. Kruk, W. Barhoumi, Deep learning and non-negative matrix factorization in recognition of mammograms, in: Proceedings of Eighth International Conference on Graphic and Image Processing, International Society for Optics and Photonics, 2017 102250B-102250B-102257.

[47] M.G. Ertosun, D.L. Rubin, Probabilistic visual search for masses within mammography images using deep learning, in: Proceedings of 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2015, pp. 1310–1315.

[48] M. Kallenberg, K. Petersen, M. Nielsen, A.Y. Ng, P. Diao, C. Igel, C.M. Vachon, K. Holland, R.R. Winkel, N. Karssemeijer, Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring, IEEE Trans. Med. Imaging 35 (2016) 1322–1331.

[49] N. Dhungel, G. Carneiro, A.P. Bradley, Deep structured learning for mass segmentation from mammograms, in: Proceedings of IEEE International Conference on Image Processing (ICIP), 2015, pp. 2950–2954.

[50] N. Dhungel, G. Carneiro, A.P. Bradley, Automated mass detection in mammograms using cascaded deep learning and random forests, in: Proceedings of 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 2015, pp. 1–8.

[51] D.H. Kim, S.T. Kim, Y.M. Ro, Latent feature representation with 3-D multi-view deep convolutional neural network for bilateral analysis in digital breast tomosynthesis, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 927–931.

[52] M. Veta, P.J. Van Diest, S.M. Willems, H. Wang, A. Madabhushi, A. Cruz-Roa, F. Gonzalez, A.B. Larsen, J.S. Vestergaard, A.B. Dahl, Assessment of algorithms for mitosis detection in breast cancer histopathology images, Med. Image Anal. 20 (2015) 237–248.

[53] F.A. Spanhol, L.S. Oliveira, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification, IEEE Trans. Biomed. Eng. 63 (2016) 1455–1462.

[54] I.C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M.J. Cardoso, J.S. Cardoso, Inbreast: toward a full-field digital mammographic database, Acad. Radiol. 19 (2012) 236–248.

[55] M. Heath, K. Bowyer, D. Kopans, R. Moore, W.P. Kegelmeyer, The digital database for screening mammography, in: Proceedings of the 5th international workshop on digital mammography, Medical Physics Publishing, 2000, pp. 212–218.

[56] X. Zhu, J. Yao, J. Huang, Deep convolutional neural network for survival analysis with pathological images, in: Proceedings of 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2016, pp. 544–547.

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

[58] S. Hussein, R. Gillies, K. Cao, Q. Song, U. Bagci, TumorNet: Lung Nodule Characterization Using Multi-View Convolutional Neural Network with Gaussian Process, in: Proceedings of IEEE 14th International Symposium on Biomedical Imaging (ISBI), 2017, pp. 1007–1010.

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

[60] Q. Dou, H. Chen, L. Yu, J. Qin, P.-A. Heng, Multilevel contextual 3-D CNNs for false positive reduction in pulmonary nodule detection, IEEE Trans. Biomed. Eng. 64 (2017) 1558–1567.

[61] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, J. Tian, Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification, Pattern Recognit. 61 (2017) 663–673.

[62] R. Paul, S.H. Hawkins, L.O. Hall, D.B. Goldgof, R.J. Gillies, Combining deep neural network and traditional image features to improve survival prediction accuracy for lung cancer patients from diagnostic CT, in: Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2016, pp. 002570–002575.

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

[64] K. Hirayama, J.K. Tan, H. Kim, Extraction of GGO candidate regions from the LIDC database using deep learning, in: Proceedings of Sixteenth International Conference on Control, Automation and Systems (ICCAS), 2016, pp. 724–727.

[65] N. Tajbakhsh, K. Suzuki, Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: MTANNs vs. CNNs, Pattern Recognit. 63 (2017) 476–486.

[66] B.-C. Kim, Y.S. Sung, H.-I. Suk, Deep feature learning for pulmonary nodule classification in a lung CT, in: Proceedings of Fourth International Winter Conference on Brain-Computer Interface (BCI), 2016, pp. 1–3.

[67] S.G. Armato, G. McLennan, L. Bidaut, M.F. McNitt-Gray, C.R. Meyer, A.P. Reeves, B. Zhao, D.R. Aberle, C.I. Henschke, E.A. Hoffman, 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 (2011) 915–931.

[68] B. Van Ginneken, S.G. Armato, B. de Hoop, S. van Amelsvoort-van de Vorst, T. Duindam, M. Niemeijer, K. Murphy, A. Schilham, A. Retico, M.E. Fantacci, Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: the ANODE09 study, Med. image Anal. 14 (2010) 707–722.

[69] J.H. Pedersen, H. Ashraf, A. Dirksen, K. Bach, H. Hansen, P. Toennesen, H. Thorsen, J. Brodersen, B.G. Skov, M. Døssing, The Danish randomized lung cancer CT screening trial—overall design and results of the prevalence round, J.Thorac. Oncol. 4 (2009) 608–614.

[70] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K.-i Komatsu, M. Matsui, H. Fujita, Y. Kodera, K. Doi, Development of a digital image database for chest radiographs with and without a lung nodule: receiver operating characteristic analysis of radiologists' detection of pulmonary nodules, Am. J. Roentgenol. 174 (2000) 71–74.

[71] V. Pomponiu, H. Nejati, N.-M. Cheung, Deepmole: deep neural networks for skin mole lesion classification, in: Proceedings of IEEE International Conference on Image Processing (ICIP), 2016, pp. 2623–2627.

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

[73] A. Mahbod, R. Ecker, I. Ellinger, Skin lesion classification using hybrid deep neural networks, arXiv preprint arXiv:1702.08434, (2017).

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

[75] A. Masood, A. Al-Jumaily, K. Anam, Self-supervised learning model for skin cancer diagnosis, in: Proceedings of Seventh International IEEE/EMBS Conference on Neural Engineering (NER), 2015, pp. 1012–1015.

[76] T. Majtner, S. Yildirim-Yayilgan, J.Y. Hardeberg, Combining deep learning and hand-crafted features for skin lesion classification, in: Proceedings of the Sixth International Conference on Image Processing Theory Tools and Applications (IPTA), 2016, pp. 1–6.

[77] S. Sabbaghi, M. Aldeen, R. Garnavi, A deep bag-of-features model for the classification of melanomas in dermoscopy images, in: Proceedings of IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 2016, pp. 1369–1372.

[78] S. Demyanov, R. Chakravorty, M. Abedini, A. Halpern, R. Garnavi, Classification of dermoscopy patterns using deep convolutional neural networks, in: Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI), 2016, pp. 364–368.

[79] L. Yu, H. Chen, Q. Dou, J. Qin, P.-A. Heng, Automated melanoma recognition in dermoscopy images via very deep residual networks, IEEE Trans. Med. Imaging 36 (2017) 994–1004.

[80] E. Nasr-Esfahani, S. Samavi, N. Karimi, S.M.R. Soroushmehr, M.H. Jafari, K. Ward, K. Najarian, Melanoma detection by analysis of clinical images using convolutional neural network, in: Proceedings of IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 2016, pp. 1373–1376.

[81] P. Sabouri, H. GholamHosseini, Lesion border detection using deep learning, in: Proceedings of 2016 IEEE Congress on Evolutionary Computation (CEC), 2016, pp. 1416–1421.

[82] T.K. Dey, Curve and Surface reconstruction: Algorithms With Mathematical Analysis, Cambridge University Press, 2006.

[83] DermQuest, Online medical resource, http://www.dermquest.com.

[84] D. Gutman, N.C. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, A. Halpern, Skin lesion analysis toward melanoma detection: a challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC), arXiv preprint arXiv:1605.01397, (2016).

[85] I. Giotis, N. Molders, S. Land, M. Biehl, M.F. Jonkman, N. Petkov, MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images, Expert Syst. Appl. 42 (2015) 6578–6585.

[86] An atlas of clinical dermatology, 2014. http://www.danderm.dk/atlas/.

[87] Dermnetnz, Online medical resources, 2014. http://www.dermnetnz.org.

[88] Interactive dermatology atlas, 2014. http://www.dermatlas.net/atlas/index.cfm.

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

[90] K. Yan, C. Li, X. Wang, Y. Yuan, A. Li, J. Kim, B. Li, D. Feng, Comprehensive autoencoder for prostate recognition on MR images, in: Proceedings of IEEE Thirteenth International Symposium on Biomedical Imaging (ISBI), 2016, pp. 1190–1194.

[91] K. Yan, C. Li, X. Wang, A. Li, Y. Yuan, D. Feng, M. Khadra, J. Kim, Automatic prostate segmentation on MR images with deep network and graph model, in: Proceedings of 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), 2016, pp. 635–638.

[92] Z. Tian, L. Liu, B. Fei, Deep convolutional neural network for prostate MR segmentation, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 101351L-101351L-101356.

[93] L. Yu, X. Yang, H. Chen, J. Qin, P.-A. Heng, Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images, AAAI, 2017, pp. 66–72.

[94] F. Milletari, N. Navab, S.-A. Ahmadi, V-net: fully convolutional neural networks for volumetric medical image segmentation, in: Proceedings of Fourth International Conference on 3D Vision (3DV), 2016, pp. 565–571.

[95] R. Cheng, H.R. Roth, N. Lay, L. Lu, B.I. Turkbey, W. Gandler, E.S. McCreedy, P. Choyke, R.M. Summers, M.J. McAuliffe, Automatic MR prostate segmentation by deep learning with holistically-nested networks, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 101332H-101332H-101336.

[96] L. Maa, R. Guoa, G. Zhanga, F. Tadea, D.M. Schustera, P. Niehc, V. Masterc, B. Fei, Automatic segmentation of the prostate on CT images using deep learning and multi-atlas fusion, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 101332O-101332O-101339.

[97] J.T. Kwak, S.M. Hewittb, Lumen-based detection of prostate cancer via convolutional neural networks, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 1014008-1014008-1014006.

[98] A. Gummeson, I. Arvidsson, M. Ohlsson, N.C. Overgaard, A. Krzyzanowska, A. Heyden, A. Bjartell, K. Aström, Automatic Gleason grading of H&E stained microscopic prostate images using deep convolutional neural networks, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 101400S-101400S-101407.

[99] H. Källén, J. Molin, A. Heyden, C. Lundström, K. Åström, Towards grading gleason score using generically trained deep convolutional neural networks, in: Proceedings of Thirteenth International Symposium on Biomedical Imaging (ISBI), 2016, pp. 1163–1167.

[100] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun, Overfeat: integrated recognition, localization and detection using convolutional networks, arXiv preprint arXiv:1312.6229, (2013).

[101] G. Litjens, R. Toth, W. van de Ven, C. Hoeks, S. Kerkstra, B. van Ginneken, G. Vincent, G. Guillard, N. Birbeck, J. Zhang, Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge, Med. Image Anal. 18 (2014) 359–373.

[102] X.W. Gao, R. Hui, Z. Tian, Classification of CT brain images based on deep learning networks, Comput. Methods Programs Biomed. 138 (2017) 49–56.

[103] P.F. Alcantarilla, A. Bartoli, A.J. Davison, KAZE features, in: Proceedings of European Conference on Computer Vision, Springer, 2012, pp. 214–227.

[104] S. Pereira, A. Pinto, V. Alves, C.A. Silva, Brain tumor segmentation using convolutional neural networks in MRI images, IEEE Trans. Med. Imaging 35 (2016) 1240–1251.

[105] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks, Med. Image Anal. 35 (2017) 18–31.

[106] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, Y. Fan, A deep learning model integrating FCNNs and CRFs for brain tumor segmentation, Med. Image Anal. 43 (2018) 98–111.

[107] K. Kamnitsas, C. Ledig, V.F. Newcombe, J.P. Simpson, A.D. Kane, D.K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Med. Image Anal. 36 (2017) 61–78.

[108] L. Zhao, K. Jia, Deep feature learning with discrimination mechanism for brain tumor segmentation and diagnosis, in: Proceedings of International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2015, pp. 306–309.

[109] D. Paredes, A. Saha, M.A. Mazurowski, Deep learning for segmentation of brain tumors: can we train with images from different institutions? SPIE Medical Imaging (2017) 101341P-101341P-101346.

[110] R. Liu, L.O. Hall, D.B. Goldgof, M. Zhou, R.A. Gatenby, K.B. Ahmed, Exploring deep features from brain tumor magnetic resonance images via transfer learning, in: Proceedings of International Joint Conference on Neural Networks (IJCNN), 2016, pp. 235–242.

[111] K. Chatfield, K. Simonyan, A. Vedaldi, A. Zisserman, Return of the devil in the details: Delving deep into convolutional nets, arXiv preprint arXiv:1405.3531, (2014).

[112] K.B. Ahmed, L.O. Hall, D.B. Goldgof, R. Liub, R.A. Gatenby, Fine-tuning convolutional deep features for MRI based brain tumor classification, in: SPIE Proceedings Vol. 10134: Medical Imaging 2017: Computer-Aided Diagnosis, 2017, pp. 101342E–1101341.

[113] M. Kistler, S. Bonaretti, M. Pfahrer, R. Niklaus, P. Büchler, The virtual skeleton database: an open access repository for biomedical research and collaboration, J. Med. Internet Res. 15 (11) (2013) e245.

[114] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, Y. Burren, N. Porz, J. Slotboom, R. Wiest, The multimodal brain tumor image segmentation benchmark (BRATS), IEEE Trans. Med. Imaging 34 (2015) 1993–2024.
[115] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository, J. Digit. Imaging 26 (2013) 1045–1057.
[116] E. Ribeiro, A. Uhl, M. Häfner, Colonic polyp classification with convolutional neural networks, in: Proceedings of IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS), IEEE, 2016, pp. 253–258.
[117] F. Navarro, Y. Saint-Hill-Febles, J. Renner, P. Klare, S. von Delius, N. Navab, D. Mateus, Computer assisted optical biopsy for colorectal polyps, SPIE Medical Imaging, International Society for Optics and Photonics, 2017.
[118] R. Zhang, Y. Zheng, T.W.C. Mak, R. Yu, S.H. Wong, J.Y. Lau, C.C. Poon, Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain, IEEE J. Biomed. Health Inform. 21 (2017) 41–47.
[119] Z. Yuana, M. IzadyYazdanabadia, D. Mokkapatib, R. Panvalkarb, J.Y. Shinc, N. Tajbakhshc, S. Gurudud, J. Liangc, Automatic polyp detection in colonoscopy videos, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 101332K-101332K-101310.
[120] L. Yu, H. Chen, Q. Dou, J. Qin, P.A. Heng, Integrating online and offline three--dimensional deep learning for automated polyp detection in colonoscopy videos, IEEE J. Biomed. Health Inform. 21 (2017) 65–75.
[121] A. Chowdhury, C.J. Sevinsky, A. Santamaria-Pang, B. Yener, A computational study on convolutional feature combination strategies for grade classification in colon cancer using fluorescence microscopy data, in: Proceedings of SPIE, 10140, 2017(2017) id. 101400Q 5140.
[122] K. Sirinukunwattana, S.E.A. Raza, Y.-W. Tsang, D.R. Snead, I.A. Cree, N.M. Rajpoot, Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images, IEEE Trans. Med. Imaging 35 (2016) 1196–1206.
[123] M.N. Kashif, S.E.A. Raza, K. Sirinukunwattana, M. Arif, N. Rajpoot, Handcrafted features with convolutional neural networks for detection of tumor cells in histology images, in: Proceedings of IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016, pp. 1029–1032.
[124] Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, I. Eric, C. Chang, Deep learning of feature representation with multiple instance learning for medical image analysis, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014, pp. 1626–1630.
[125] H. Haj-Hassan, A. Chaddad, Y. Harkouss, C. Desrosiers, M. Toews, C. Tanougast, Classifications of multispectral colorectal cancer tissues using convolution neural network, J. Pathol. Inform. (2017) 8.
[126] N. Tajbakhsh, S.R. Gurudu, J. Liang, Automated polyp detection in colonoscopy videos using shape and context information, IEEE Trans. Med. Imaging 35 (2016) 630–644.
[127] D. Saslow, D. Solomon, H.W. Lawson, M. Killackey, S.L. Kulasingam, J. Cain, F.A. Garcia, A.T. Moriarty, A.G. Waxman, D.C. Wilbur, American Cancer Society, CA American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer, Cancer J. Clin. 62 (2012) 147–172.
[128] G.G. Birdsong, Automated screening of cervical cytology specimens, Hum. Pathol. 27 (1996) 468–481.
[129] Y. Song, L. Zhang, S. Chen, D. Ni, B. Lei, T. Wang, Accurate segmentation of cervical cytoplasm and nuclei based on multiscale convolutional network and graph partitioning, IEEE Trans. Biomed. Eng. 62 (2015) 2421–2433.
[130] Y. Song, J.-Z. Cheng, D. Ni, S. Chen, B. Lei, T. Wang, Segmenting overlapping cervical cell in pap smear images, in: Proceedings of IEEE Thirteenth International Symposium on Biomedical Imaging (ISBI), 2016, pp. 1159–1162.
[131] T. Xu, H. Zhang, X. Huang, S. Zhang, D.N. Metaxas, Multimodal deep learning for cervical dysplasia diagnosis, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 115–123.
[132] K.H. Cha, L. Hadjiiski, R.K. Samala, H.P. Chan, E.M. Caoili, R.H. Cohan, Urinary bladder segmentation in CT urography using deep-learning convolutional neural network and level sets, Med. Phys. 43 (2016) 1882–1896.
[133] K.H. Cha, L.M. Hadjiiski, H.-P. Chan, R.K. Samala, R.H. Cohan, E.M. Caoili, C. Paramagul, A. Alva, A.Z. Weizer, Bladder cancer treatment response assessment using deep learning in CT with transfer learning, SPIE Medical Imaging, International Society for Optics and Photonics„ 2017.
[134] M. Gordon, L. Hadjiiski, K. Cha, H.-P. Chan, R. Samala, R.H. Cohan, E.M. Caoili, Segmentation of inner and outer bladder wall using deep-learning convolutional neural networks in CT urography, SPIE Medical Imaging, International Society for Optics and Photonics, 2017.
[135] E. Gibson, M.R. Robu, S. Thompson, P.E. Edwards, C. Schneider, K. Gurusamy, B. Davidson, D.J. Hawkes, D.C. Barratt, M.J. Clarkson, Deep residual networks for automatic segmentation of laparoscopic videos of the liver, SPIE Medical Imaging, International Society for Optics and Photonics, 2017 101351M-101351M-101356.
[136] W. Li, F. Jia, Q. Hu, Automatic segmentation of liver tumor in CT images with deep convolutional neural networks, J. Comput. Commun. 3 (2015) 146.
[137] K. Sirinukunwattana, D.R. Snead, N.M. Rajpoot, A stochastic polygons model for glandular structures in colon histology images, IEEE Trans. Med. Imaging 34 (2015) 2366–2378.
[138] A. BenTaieb, J. Kawahara, G. Hamarneh, Multi-loss convolutional networks for gland analysis in microscopy, in: Proceedings of the IEEE Thirteenth International Symposium on Biomedical Imaging (ISBI), 2016, pp. 642–645.
[139] H. Chen, X. Qi, L. Yu, P.-A. Heng, DCAN: deep contour-aware networks for accurate gland segmentation, in: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2016, pp. 2487–2496.
[140] Y. Mao, Z. Yin, J. Schober, A deep convolutional neural network trained on representative samples for circulating tumor cell detection, in: Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV), 2016, pp. 1–6.
[141] F. Xing, Y. Xie, L. Yang, An automatic learning-based framework for robust nucleus segmentation, IEEE Trans. Med. Imaging 35 (2016) 550–566.
[142] K. Sirinukunwattana, J.P. Pluim, H. Chen, X. Qi, P.-A. Heng, Y.B. Guo, L.Y. Wang, B.J. Matuszewski, E. Bruni, U. Sanchez, Gland segmentation in colon histology images: the glas challenge contest, Med. Image Anal. 35 (2017) 489–502.
[143] J. Bernal, N. Tajkbaksh, F.J. Sánchez, B.J. Matuszewski, H. Chen, L. Yu, Q. Angermann, O. Romain, B. Rustad, I. Balasingham, Comparative validation of polyp detection methods in video colonoscopy: Results from the miccai 2015 endoscopic vision challenge, IEEE Trans. Med. Imaging 36 (2017) 1231–1249.
[144] S. Dodge, L. Karam, Understanding how image quality affects deep neural networks, in: Proceedings of Eighth International Conference on Quality of Multimedia Experience (QoMEX), 2016, pp. 1–6.

Dr. Zilong Hu got his Ph.D. in 2018 in Computational Science & Engineering at Michigan Tech University, Houghton, MI, USA. He received his B.S. degree in automation from Tianjin University, Tianjin, China in 2011, and his M.S. degree in medical informatics from Michigan Tech University in 2014. His research is focused on medical image processing, pattern recognition and classification. He is particularly interested in machine learning/deep learning on pattern recognition. His other major research interest is the implementation of GPU technique on digital image processing.

Dr. Jinshan Tang is currently a professor at Michigan Technological University. He received his Ph.D. in 1998 from Beijing University of Posts and Telecommunications, and got post-doctoral training in Harvard Medical School and National Institute of Health. His research interests include biomedical image processing, biomedical imaging, and computer aided cancer detection. He has obtained more than two million dollars grants in the past years as a PI or Co-PI. He has published more than 100 refereed journal and conference papers. He has published two edited books on medical image analysis. One is Computer Aided Cancer Detection: Recent Advance and the other is Electronic Imaging Applications in Mobile Healthcare. He is a leading guest editor of several journals on medical image processing and computer aided cancer detection. He is a senior member of IEEE and Co-chair of the Technical Committee on Information Assurance and Intelligent Multimedia-Mobile Communications, IEEE SMC society. His research has been supported by USDA, DoD, NIH, Air force, DoT, and DHS.

Ziming Wang is currently a master student in Electronic & Computer Engineering in Michigan Technological University, Houghton, Michigan, United States. He received his B.S degree in automation and communication engineering from Jilin University, Jilin, China in 2010. His research interests include image processing and deep learning.

Dr. Kai Zhang is a professor of School of Computer Science and Technology at Wuhan University of Science and Technology. He received his PhD degree from Huazhong University of Science and Technology in 2003. He got post-doctoral training in the School of Electronics Engineering and Computer Science at Peking University from 2008 to 2010. His major research interests include artificial intelligence, pattern recognition and multiobjective objective optimization.

Ling Zhang is currently a second-year graduate student major in Data Science at Michigan Technological University. He received his B.S degrees in 2016 from the 2+2 program between Wuhan Institute of Technology and Indiana State University. He got B.S degree in Electrical Engineering and Automation from Wuhan Institute of Technology, Wuhan province, China. Besides, he acquired B.S degree in Computer Engineering with minor in Electrical Engineering from Indiana State University. He is doing research work under his advisor Dr. Tang. His research interests include data mining and machine learning.

Qingling Sun is currently the chief software engineer and the manager of Sun Technologies & Services, LLC. She received her master degree from University of Virginia. She received her Ph.D. study in University of Southern Mississippi. She provided sub-contract service to DoD sponsored project and provided consulting service to USDA sponsored project. Her research interests include: medical informatics, image database, data mining, comprehensive web based systems, etc. Till now, she has published about 10 papers.