SCIENCE CHINA
Information Sciences

• REVIEW •

January 2020, Vol. 63 111101:1–111101:16
https://doi.org/10.1007/s11432-019-2708-1

Application of machine learning method in optical molecular imaging: a review

Yu AN[1], Hui MENG[1,2], Yuan GAO[1,2], Tong TONG[1,2], Chong ZHANG[1,2], Kun WANG[1] & Jie TIAN[1,3*]

[1]*CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing* 100190, *China;*
[2]*University of Chinese Academy of Sciences, Beijing* 100049, *China;*
[3]*Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing* 100191, *China*

Received 5 July 2019/Revised 17 September 2019/Accepted 22 October 2019/Published online 25 December 2019

Abstract Optical molecular imaging (OMI) is an imaging technology that uses an optical signal, such as near-infrared light, to detect biological tissue in organisms. Because of its specific and sensitive imaging performance, it is applied in both preclinical research and clinical surgery. However, it requires heavy data analysis and a complex mathematical model of tomographic imaging. In recent years, machine learning (ML)-based artificial intelligence has been used in different fields because of its ability to perform powerful data processing. Its analytical capability for processing complex and large data provides a feasible scheme for the requirement of OMI. In this paper, we review ML-based methods applied in different OMI modalities.

Keywords optical molecular imaging, machine learning, artificial intelligence

Citation An Y, Meng H, Gao Y, et al. Application of machine learning method in optical molecular imaging: a review. Sci China Inf Sci, 2020, 63(1): 111101, https://doi.org/10.1007/s11432-019-2708-1

1 Introduction

Optical imaging is one of the most powerful biomedical imaging technologies, which is widely used in both preclinical studies (e.g., cancer research and drug evaluation) and clinical practice (e.g., surgical navigation and pathological analysis). In the past 20 years, there has been a significant increase in the number of imaging modalities and their corresponding applications [1–4]. In particular, optical molecular imaging (OMI), which has produced a large number of breakthroughs, provides a specific and sensitive dynamic imaging technology for biomedical research [5–7].

OMI is a molecular imaging technology that uses an optical signal as the imaging medium. Based on the biochemical characteristics of tissue in organisms and related reactions, OMI acquires imaging signals by observing the distribution of photons on the surface of organisms or the interference of the laser with biochemical reactions. Thus, it performs well for detecting biomedical activities [5,8]. In 1977, Jöbsis [9] found that the hemoglobin and cytochrome in biological tissues absorb less near-infrared light at 650–900 nm, which means that there is a bio-optical window that allows photons to pass through the organism. Based on this research and subsequent theoretical discoveries, the near-infrared spectroscopy (NIS) technology is created and developed. With the advancements in computer science, molecular biology, and bio-photonics, new methods, such as the molecular probe technology, optical signal acquisition, and

* Corresponding author (email: tian@ieee.org)

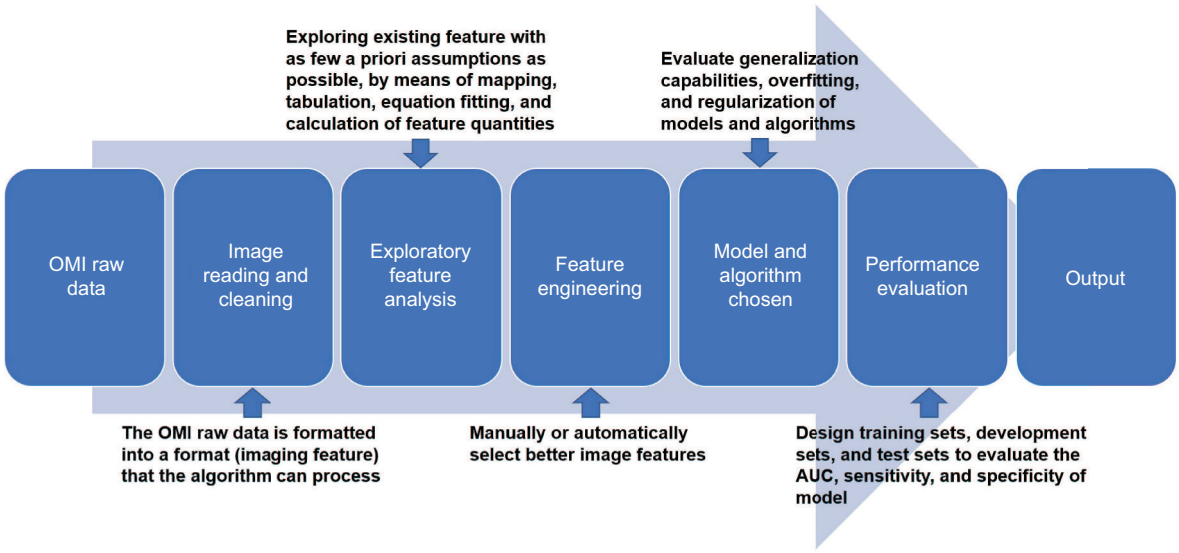


Figure 1 (Color online) Main flow chart of machine learning (ML) applied to optical molecular imaging (OMI).

amplification technology, have been gradually applied in NIS, significantly improving the near-infrared light acquisition capability of OMI. These improvements allow OMI to observe pathological information in organisms through optical signals.

However, there are still problems with OMI that need to be addressed, including limited imaging quality and image analysis accuracy. For example, the photon scattering and absorption of light propagation model lead to an ill-posed problem in optical scattering tomography [10, 11]. This problem limits the spatial resolution and source localization accuracy of the imaging result. Furthermore, in the field of optical image analysis, the diagnosis can be examiner-dependent, which leads to considerable inter-observer variability [12]. Thus, OMI requires a more effective method to improve the quality of images signal processing and analysis.

The rapid development of artificial intelligence (AI) technology could potentially address the problems encountered by OMI. AI is a science that endows artificial creations with cognitive, thinking and other behavioral patterns through human intervention. Through computer science, AI relies on bio-coding or physical coding and attempts to build bio-intelligence. Machine learning (ML) technology, which is an important part of AI, uses an artificial pattern recognition system or data-driven-based presentation learning method to achieve intelligent behavior [13]. In particular, the presentation learning method, which is based on implicit features learned from raw data, improves the complex data analysis performance of an AI system. Although this method has a powerful semantic representation capability, it relies on large amounts of data. However, with the development of computer storage technology, communication technology, and computer computing performance, it is now possible to collect and manage large-scale data that allows the bottleneck of the presentation learning method to be solved. Thus, presentation learning, especially the deep learning (DL) method, is now widely used. The DL-based method uses neural networks to extract features in different scales and spaces from natural data and then expresses the features in a cascading way. Because the DL performs well on high-dimensional data, it is being used in many domains, including image recognition [14–17], speech recognition [18–20], and natural language processing [21, 22]. It is also used in network information security [23, 24], industrial data processing [25], and drug molecular research [26]. The main flow chart of ML applied to OMI image processing is shown in Figure 1.

Owing to OMI's need for image data processing and analysis, ML has been applied to different optical modalities. It provides a new image processing tool for disease classification, lesion detection, segmentation, three-dimensional (3D) visualization, and tomographic image reconstruction. A summary of the different ML methods applied in OMI is listed in Table 1 [10, 12, 27–41]. In the following sections of

Table 1 Summary of the different ML methods applied in OMI

OMI modality[a)]	ML systems	Year	Model	AUC (%)[b)]	Sensitivity (%)	Specificity (%)
OCT	Guillaume et al. [27]	2016	Random Forest	NA	81.20	93.70
	Pratul et al. [28]	2014	Support Vector Machine	86.67	NA	NA
	Cecilia et al. [29]	2017	VGG-16	93.45	92.64	93.69
	Roy et al. [30]	2017	ReLayNet	99.00	NA	NA
	Jeffrey et al. [41]	2018	3D U-Net	99.21	NA	NA
	Abhijit et al. [31]	2016	Random Walks	97.86	NA	NA
	Zhao et al. [32]	2015	Bayesian Network	91	NA	NA
PAT	Johannes Schwab et al. [33]	2018	DALnet	93.30	NA	NA
	Andreas et al. [34]	2018	Model-based CNN	94.50	NA	NA
	Stephan et al. [35]	2019	NETT	89	NA	NA
OST	Chao H et al. [36]	2019	CNN and RNN	NA	NA	NA
	Yuan et al. [10]	2019	Multilayer Perceptron	NA	NA	NA
OIS	Andre et al. [37]	2011	Bag-of-Visual-Words	94.20	97.70	86.10
	Kamen et al. [38]	2016	Support Vector Machine	84.32	87.34	80.61
	Li et al. [12]	2018	CNN	99.49	NA	NA
	Daniele et al. [39]	2018	Context Specific Descriptor	NA	89.20	NA
	Chong et al. [40]	2019	Generative Adversarial Networks	91.64	NA	NA

a) OCT: optical coherent tomography; PAT: photoacoustic tomography; OST: optical scattering tomography; OIS: optical image-guided surgery.

b) Here, the index area under curve (AUC) includes the classification accuracy, dice efficiency, and structure similarity for different ML methods and their applications.

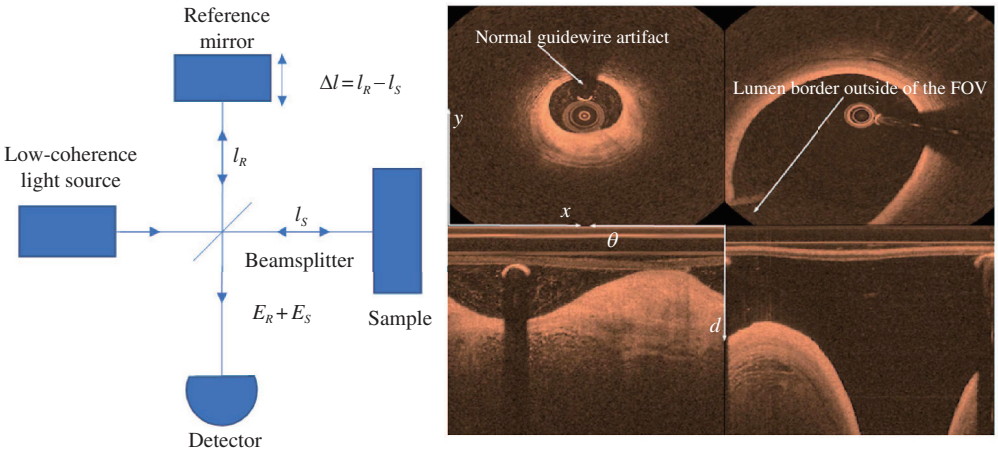


Figure 2 (Color online) Optical coherence tomography. (a) Imaging mechanism of OCT and (b) OCT images. These images are reproduced from [42, 43].

this paper, we introduce the research and practical application of the ML method in OMI. Section 2 reviews ML-based methods used in optical coherence tomography (OCT). Sections 3 and 4 present the ML technology used in the reconstruction of optical scattering tomography and photoacoustic tomography, respectively. Section 5 introduces the methods applied in optical image-guided surgery. Section 6 discusses issues and perspectives.

2 Optical coherence tomography

OCT is a technology that uses near-infrared light to obtain sub-surface images of opaque or translucent materials (Figure 2) [42, 43]. Based on low-time coherent optical interferometry or white-light coherence

technology, OCT provides a tomographic imaging modality that has a high spatial resolution, has no radiation, is non-invasive, has high imaging depth, and has fast imaging speed. With the development of imaging technology and the exploration of related biological applications, OCT has been used widely in the field of in vivo retinal imaging, coronary artery imaging, and gastrointestinal detection.

Recently, clinical OCT imaging data, such as age-related macular degeneration (AMD), coronary heart disease, and atherosclerosis, have increased rapidly, which requires the expertise and judgment of experienced physicians to interpret these data [44]. However, the progress of data interpretation is currently insufficient to meet the huge clinical requirements. The ML technology has been used to alleviate this problem.

OCT data have the following characteristics, which makes them suitable to be analyzed by the ML technology. Firstly, there is an increasing growth in the number of OCT regularly collected around the world, which means that a large amount of OCT data are available to feed the AI system. These data support the training of larger and more complex ML methods, such as the DL network, and allow the creation of more generalized and stronger AI machines. Secondly, the scanning field of OCT is almost fixed [45], which is different from real-world images. Thus, the 3D structure of OCT is consistent in different images, which reduces the complexity of the computer vision problem. Thirdly, structural details hidden in OCT images are difficult to be detected by conventional imaging techniques [29, 45]; these can be analyzed by the powerful feature extraction capability of the ML technology. The remainder of this section introduces different applications of ML technology in OCT images.

2.1 Disease classification

Traditional ML image analysis for OCT requires manual acquisition of convolutional maps to obtain key features that characterize a disease through edge detection and feature extraction. ML techniques applied to OCT disease classification based on manual acquisition of image features include principal component analysis, support vector machine (SVM), and random forest [27,28,46]. For example, Lemaître et al. [27] adopted a combination of random forest with local binary pattern features and different mapping strategies to classify diabetic macular edema versus normal subjects. Manually defined image features can effectively characterize certain characteristics of a disease, thereby improving the recognition of linear and nonlinear classifiers. However, the pathological significance of image features needs to be further explained to demonstrate the reliability of the classification results.

Recently, the DL method has been used in OCT image analysis, where a multilayer neural network is trained according to the intrinsic law and representation level of the training data. The information obtained in these learning processes is useful for the interpretation of data features. For example, Lee et al. [29] used VGG-16 [47] to automatically classify AMD data. They reported that their network, which was trained with 100000 OCT B-scan images, achieved an area under curve (AUC) of 0.97 in the validation.

2.2 Lesion region segmentation

In the application of lesion region segmentation, DL performs better than conventional methods. Several groups have used the DL network, such as U-Net, to segment intraretinal fluid cysts and subretinal fluid on OCT B-scans [30,48] (Figure 3(a)). Furthermore, DeepMind and the Moorfields Eye Hospital [41] proposed a novel AI framework for segmentation and classification of ocular disease (Figure 3(b)). They used the segmentation subnetwork to extract the morphological features of 15 different retinas and OCT acquisition artifacts. Based on these features, the AI system provides the referral triage decision (i.e., urgent, semi-urgent, routine, and observation) by a classification subnetwork. The AI framework provided an expert-level performance in 10 different kinds of OCT lesion detection (i.e., choroidal neovascularization (CNV), macular oedema without CNV, drusen, geographic atrophy, epiretinal membrane, vitreomacular traction, full-thickness macular hole, partial-thickness macular hole, central serous retinopathy, and 'normal').

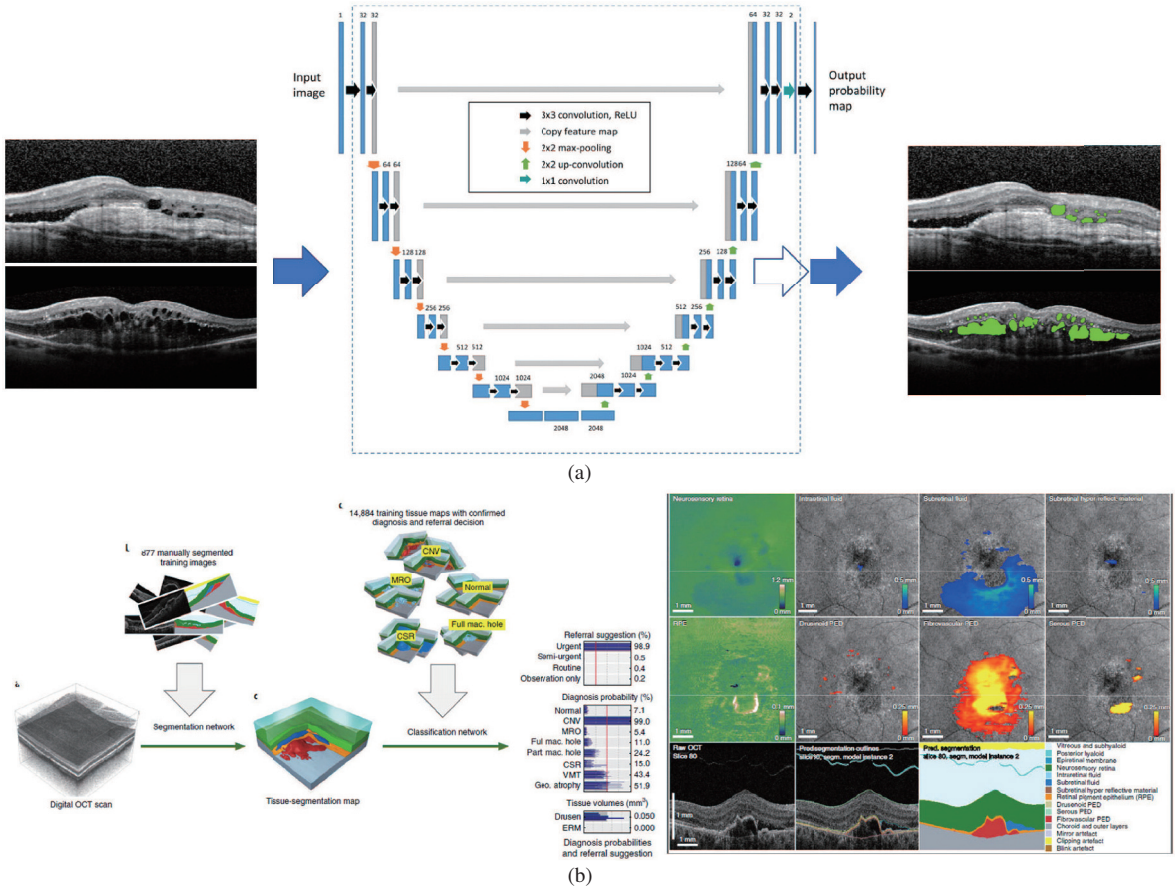


Figure 3 (Color online) Overview of ML-based algorithms for OCT segmentation. (a) U-Net-based intraretinal cystoid fluid segmentation reproduced from [48] and (b) artificial intelligence (AI) framework used in ocular disease and corresponding segmentation results reproduced from [41].

Furthermore, ML is also widely used in the detection of intravascular OCT (IVOCT) images such as automatic vascular cavity segmentation [31, 49], stent detection [32, 50], and plaque sediment detection [51]. The DL-based IVOCT analysis is also used to measure the thickness of the fibrous cap of an arterial wall, which is important to predict the potential follow-up myocardial infraction [52, 53].

Although these ML-based approaches have provided promising results in the field of OCT, there are still theoretical and technical issues that need to be addressed. For example, it is unclear whether the polar or Cartesian representation is the better choice for ML training [43]. In addition, artifacts in OCT images limit the performance of ML-based approaches [41, 43].

3 Photoacoustic tomography

Photoacoustic tomography (PAT) is an emerging and fast-growing biomedical imaging modality, which is used to reflect the optical absorption characteristics of biological tissue or molecule [54–57]. It uses a short-pulsed laser beam to excite the target tissue, and then detect the ultrasonic wave produced from the irradiated tissue to reconstruct the optical absorption in biological tissue (Figure 4) [58]. Based on its imaging mechanisms, PAT provides the optical properties of biological tissue with both optical contrast and acoustic.

3.1 Conventional reconstruction algorithms of PAT

Conventional reconstruction algorithms of PAT can be divided into three major types, i.e., the back-projection, time-reversal, and model-based algorithms. The back-projection algorithm, with its charac-

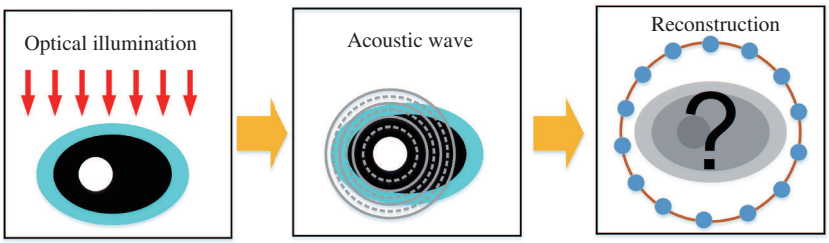


Figure 4 (Color online) Imaging mechanism of photoacoustic tomography (PAT) reproduced from [58].

teristics of easy implementation and fast imaging, is suitable for PAT reconstruction with high temporal resolution. Based on the analytical inversion formula of ultrasonic propagation, it can directly project the detected signals to the PAT image, which is applicable to different detection geometries [59–62]. However, it is prone to produce stripe artifacts and lose low-frequency information, which makes it not suitable for application to functional imaging or molecular imaging with precise quantification [63–65]. Meanwhile, the time-reversal algorithm builds a reconstruction model by simulating the back-propagation of ultrasonic waves in the time domain, which provides more accurate reconstruction results [66–68]. However, because of its complex mathematical model and numerical computation, it has an inefficient reconstruction performance in practical applications. On the other hand, the model-based algorithm depends on the photoacoustic propagation model in discretizing the imaging space. It creates PAT images by solving the optimization problem of the model calculated by different iterative optimization methods with different artificial regularization priors [63, 69–73]. Owing to the inherent noise interference of the photoacoustic signal acquisition, PAT reconstruction has not obtained ideal results with the existing regularization prior, especially when the detected signals are interfered by downsampling [73].

3.2 PAT reconstruction based on ML

To overcome the shortcomings of conventional methods, PAT reconstruction strategies based on ML have been proposed in recent years. They can be classified into three classes: using U-Net to improve the quality of coarse PAT images; learning the iterative process of conventional reconstruction based on a neural network; and constructing regularization terms by a neural network to guide the optimization.

The first class uses the U-Net network to improve the image quality of PAT images reconstructed by the filter back-projection algorithm [33, 58]. Based on the powerful denoising ability of the U-Net network, this class reduces artifacts and noise in images reconstructed using the conventional method (Figure 5(a)). Furthermore, residual training strategy has been used to improve the denoising process, and the reported experiment results demonstrate that the U-Net-based method performs better than the conventional total variation regularization method.

The second class was first proposed by Andreas Hauptmann's group [34, 74]. They used a neural network to simulate each loop of the iterative calculation. This method can be divided into two stages: reconstructing a coarse result by the conventional algorithm and fine-turning it by each layer of the neural network (Figure 5(b)). Compared with the first class of methods mentioned above, this method achieves better results for PAT reconstruction. Furthermore, the gradient information of photoacoustic is merged into the input signals of the neural network to improve the quality of PAT reconstruction.

The third class was first proposed by Stephan Antholzer's group [35]. They used a neural network as the regularization before the reconstruction. In this method, the result of each iterative loop is denoised by the network and produces the expectation of regularization in the next loop (Figure 5(c)). This method still relies on the conventional reconstruction strategy, and its performance is not as good as the U-net-based method for noisy data reconstruction.

Overall, the abovementioned ML-based reconstruction methods focus on the postprocessing of PAT images. The first two classes calculate the initial images based on conventional methods and then improve the image quality by neural networks. The third class merges ML into the conventional reconstruction strategy, which calculates the expected prior using a neural network to guide the calculation in each

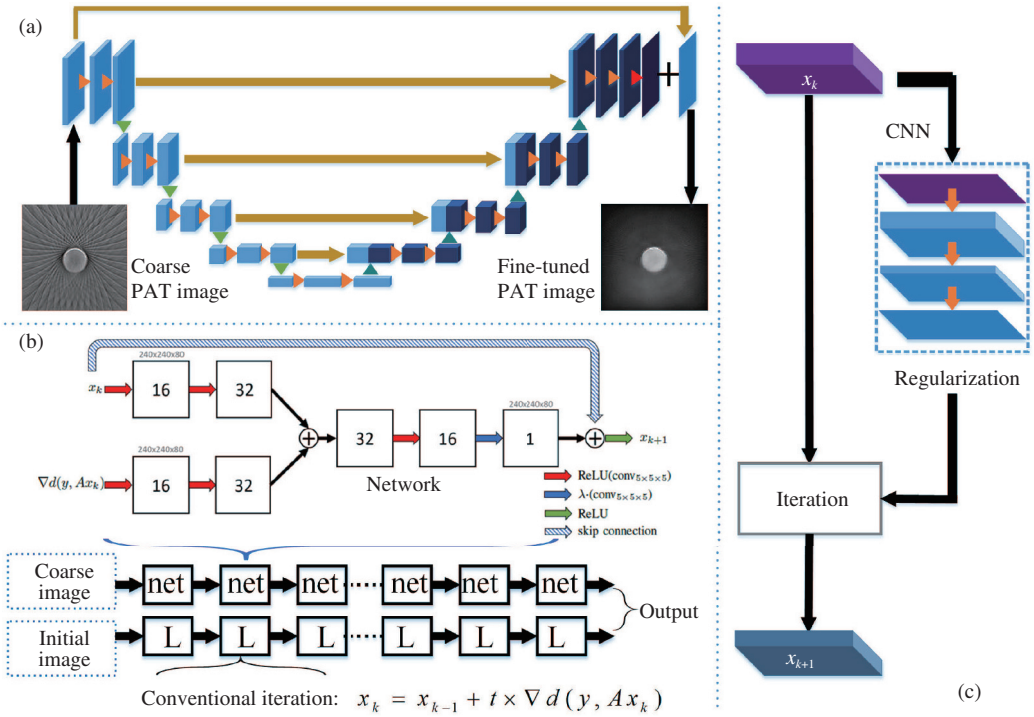


Figure 5 (Color online) Different methods of ML-based PAT reconstruction. (a) U-Net-based reconstruction method reproduced from [58]; (b) neural network-based method to simulate conventional iteration reproduced from [34]; and (c) convolutional neural network (CNN)-based regularization method reproduced from [35].

iteration as a regularization. All these methods perform reconstruction from the image space to image space and are very useful. However, the results are limited by the performance of the conventional reconstruction method, and no method can reconstruct PAT images from raw photoacoustic signals, i.e., PAT reconstruction from signal space to image space. Because the raw signal contains more information about the imaging object, the ML-based PAT method that focuses on the signal-to-image reconstruction obtains more details and will be one of the future research directions in this field.

4 Optical scattering tomography

Optical scattering tomography (OST) is a multimodal imaging technique that combines two-dimensional (2D) near-infrared light imaging and high-resolution anatomical tomography (Figure 5) [75]. The anatomical information provided by X-ray computed tomography (X-CT) or magnetic resonance imaging (MRI) is used as the 3D imaging space, and the surface photon distribution is obtained by planar optical scattering imaging such as fluorescence molecular imaging (FMI) [76] and bioluminescence imaging (BLI) [77]. These tomographic optical imaging methods, such as fluorescence molecular tomography (FMT) and bioluminescence tomography (BLT), use a mathematic model to describe the photon propagation and reconstruct the 3D distribution of light sources. These technologies provide more tomographic information than conventional 2D optical scattering imaging (Figure 6) [78].

4.1 Conventional reconstruction algorithms of OST

The conventional reconstruction method is based on a model of photon propagation such as the radiation transfer equation (RTE) [79–81]. The propagation model is simplified into a low-order approximation and transformed into a linear matrix by the variational method and finite element analysis. This matrix is used to describe the linear relation between the surface light information of the imaging object and sources in biological tissue, and its inverse problem is used to reconstruct the source distribution. However,

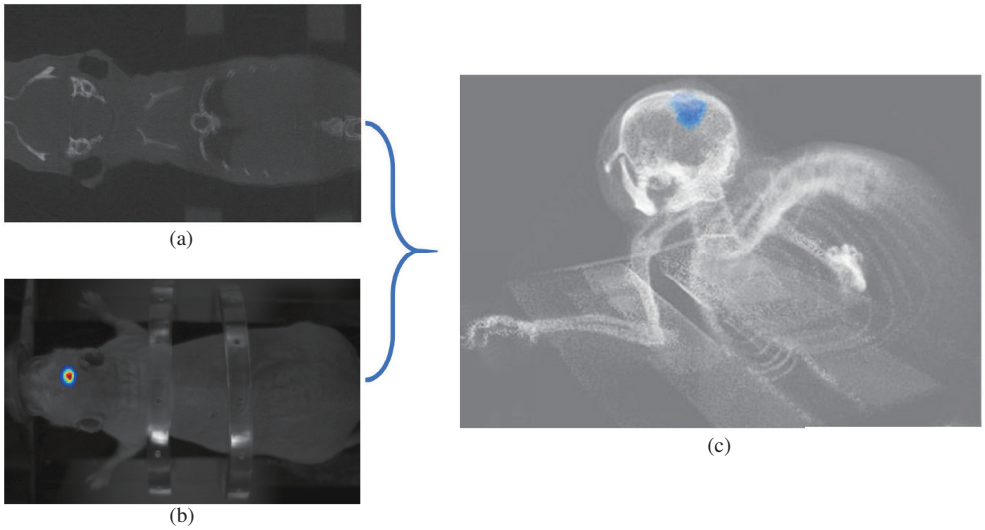


Figure 6 (Color online) Optical scattering tomography. (a) Anatomical information of X-CT; (b) planar optical imaging; and (c) images of OST. These images are reproduced from [78].

the scattering and absorption effects of photon propagation complicate the reconstruction and limit the imaging quality of OST. Many methods have been proposed to overcome this limitation. One of them uses a sparse prior for reconstruction, which assumes that light sources are sparse compared to the entire reconstruction region. Based on this prior, sparse regularization terms (e.g., L0, L1, Lp $(0 < p < 2)$) [82,83] are employed to improve the image quality. In addition, the greedy strategies-based matching pursuit algorithm [84,85], the total variation regularization penalized method [86], and the L2, 1-norm-based optimization [11] method have been proposed to improve the accuracy of source localization. Moreover, another effective method is to use the prior information of a tumor region segmented from other structural imaging modalities (e.g., CT and MRI) to guide the reconstruction, which is referred to as the guided method [87]. Furthermore, methods such as the hard-prior regularization [88,89], soft-prior regularization [87], hierarchal [90], and kernel methods [91] use a tumor region segmented manually or automatically by an algorithm as the prior region to improve the structural resolution of reconstruction results.

However, reconstruction methods based on the model of photon propagation have many limitations. Firstly, the low-order approximation equation is simplified in the RTE, which cannot accurately describe the process of photon propagation. This limits the accuracy of OST reconstruction. Secondly, owing to the ill-posed problem of reconstruction, the 2D imaging of light distribution on the object's surface cannot completely reflect the location of light sources in a biological tissue, which limits the performance of conventional methods based on the residual error between the reconstructed and detected surface light.

4.2 OST reconstruction based on ML

To overcome the limitations of conventional reconstruction methods, Yuan et al. [10] proposed a data-driven-based strategy to reconstruct bioluminescence tomography, which uses a multilayer perceptron (MLP) (Figure 7(a)). They collected training data by simulating a Monte Carlo process and trained the network with the loss between the reconstructed and real sources. As an extension, Huang et al. [36] proposed an FMT reconstruction framework that contains a convolutional neural network (CNN) and a recurrent neural network (RNN). The CNN subnet uses the VGG16 network to extract the features of 2D fluorescence images, and the gated recurrent unit (GRU) [92] is employed to combine these multi-angle fluorescence features. Finally, these features are used as input to the MLP to reconstruct the FMT images (Figure 7(b)). This method can achieve FMT reconstruction based on 2D fluorescence images, which avoids errors caused by mesh registration in the conventional method.

Although these ML-based approaches address the ill-posed problem of OST reconstruction, there are

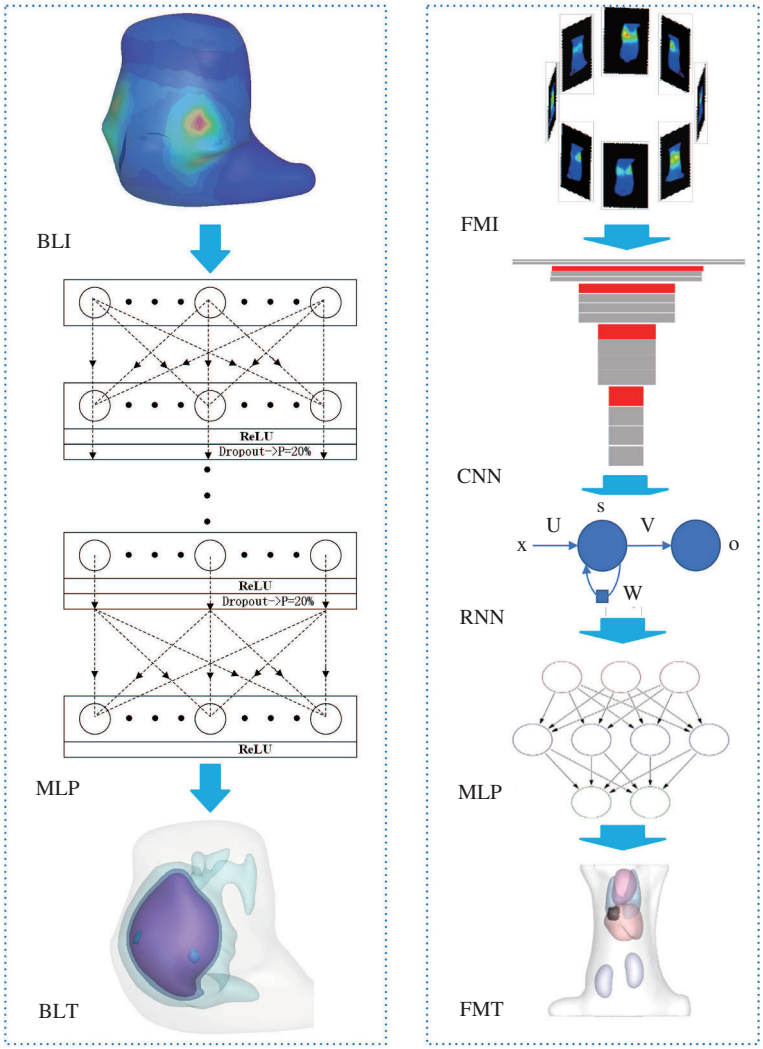
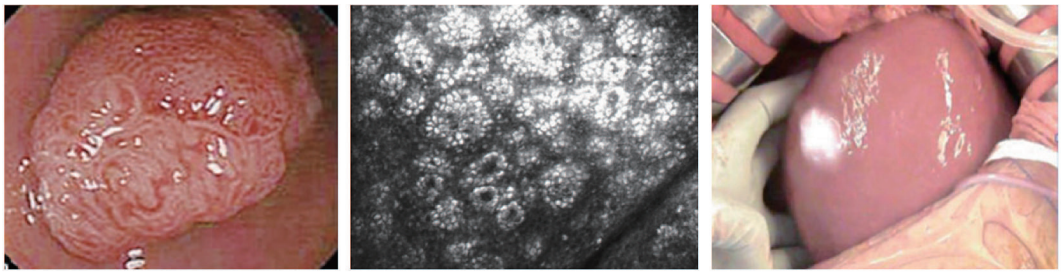


Figure 7 (Color online) Structure of the networks used in OST reconstruction and the corresponding reconstruction results. (a) MLP-based BLT reconstruction network reproduced from [10] and (b) CNN-RNN-based FMT reconstruction framework reproduced from [36].

still many limitations in practical applications. One major shortcoming is that the training data are only generated by simulations and do not use data from in vivo experiments. Because the gold standard of in vivo experiments is hard to obtain, it is important to build a more efficient data collection scheme. In addition, the size of 3D input data is too large to be processed by a neural network, and the tetrahedron-based data downsampling mesh does not describe the orthogonal spatial relationships in physical space. Therefore, a novel data structure or a 2D mathematical model to describe photon propagation needs to be developed.

5 Optical image-guided surgery

Optical image-guided surgery is an important navigation technology in clinical surgery. It provides powerful support for operators to observe the pathological tissue clearly and protect important organs from iatrogenic injury. Intraoperative optical imaging technology is an interdisciplinary field of imaging technology and clinical medicine (Figure 8) [93–95]. White-light endoscopy imaging is one of the major optical imaging technologies for minimally invasive surgery (MIS), which provides important image information for surgery. However, the image it produces only provides information about the tissue under white-light,



(a) (b) (c)

Figure 8 (Color online) Different imaging modalities used in surgery. (a) Narrowband cavity mirror imaging; (b) laser confocal microscopy imaging; and (c) near-infrared fluorescence imaging. These images are reproduced from [93–95].

which is full of red or dark red blood. Hence, it is very difficult to clearly observe the lesion tissue. Thus, doctors rely on their knowledge of anatomy and subjective experience to judge the structure of tumors and blood vessels via touch and vision. There is still a lack of an objective surgery navigation method to achieve real-time intraoperative identification of the structure of complex tissues [96]. To improve the clinical effect of real-time localization of pathological tissues, vascular nerves, and delineation of boundaries, a variety of OMI techniques with better specific imaging performance have been proposed and used in both preclinical research and clinical applications; this includes narrowband cavity mirror imaging [93, 97], laser confocal microscopy imaging [94], and near-infrared fluorescence imaging [95, 98–102]. Furthermore, ML-based methods are also applied to improve the quality of imaging and the efficiency of diagnostic information extraction.

5.1 Image analysis

ML is used to assist doctors in intraoperative optical imaging analysis, such as lesion type classification and grading, to avoid the problem of having an examiner-dependent diagnosis. In probe-based confocal laser endomicroscopy (pCLE), André et al. [37] proposed a bag-of-visual-words (WoB)-based image and video retrieval framework to classify neoplastic and benign colorectal polyps in pCLE images (Figure 9(a)). They then extended this study using high-level pathological interpretation of pCLE images as a prior to improve the accuracy of classification [103]. Kamen et al. [38] extracted the features from pCLE imaging of the brain by using encoding schemes to classify meningitis and pytoblastoma. They used an SVM as the prediction model and achieved an accuracy of more than 83%. Recently, Li et al. [12] proposed a CNN-RNN-based video-classification framework to classify intraoperative brain tumor. This framework extracts the feature of each frame from the pCLE video by the CNN part and fuses all the features by the RNN part. They found that their method performs better than conventional methods (accuracy = 99.49%) (Figure 9(c)).

5.2 Optical image enhancement and registration

ML is also used in optical image enhancement and image registration to improve optical image quality. In the field of image enhancement, Raví et al. [39] proposed a novel synthetic data generation approach to construct the ground-truth data and used these data to evaluate the performance of different exemplar-based deep neural networks in obtaining optical images with super-resolution. Recently, Zhang et al. [40] developed a postprocessing method for fluorescence image enhancement, which employs a generative adversarial network to improve the image resolution. To overcome the drawback of fake texture generation in traditional neural networks, they proposed a total gradient loss for network training and applied a fine-tuning training procedure to further improve the network architecture. For image registration, Mountney et al. [104] proposed a context-specific-based feature descriptor and used a decision tree to present the feature point and match these points based on their likelihood (Figure 10). They used this method to describe the 3D space of endoscopic images in MIS and reported that the registration results are robust to drift, occlusion, and changes in orientation and scale.

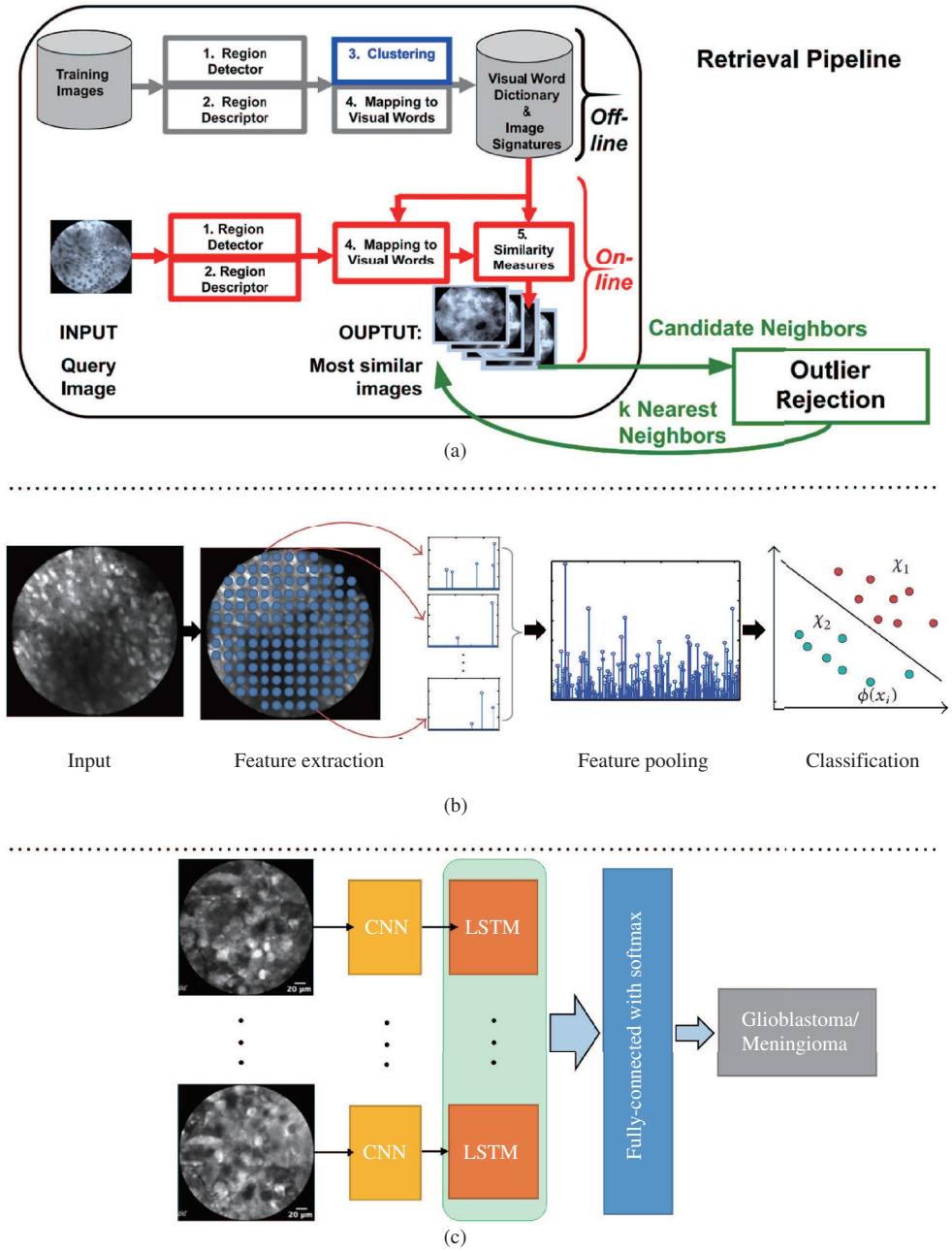


Figure 9 (Color online) ML-based methods in the application of minimally invasive surgery (MIS) imaging classification. (a) WoB-based method in [37]; (b) SVM-based classification method in [38]; and (c) CNN-RNN framework in [12].

6 Issues and perspectives

Although many researches have tried to apply the ML technology in OMI and achieved promising breakthroughs, the application of ML-based methods in clinical surgery still requires more theoretical research and clinical experiments. Because the structured data of most clinical applications are insufficient to support the complexity of data-driven-based ML technology, further research is needed for collecting more data and designing novel ML methods based on small-scale data learning. Furthermore, there is still no ideal method that explains the mechanism of a neural network. The features extracted by exciting methods are difficult to illustrate the relevant theories of medicine or optical imaging. This is an urgent problem of neural networks that needs to be overcome.

However, ML-based AI still has a very broad space of research and applications. With the expansion of

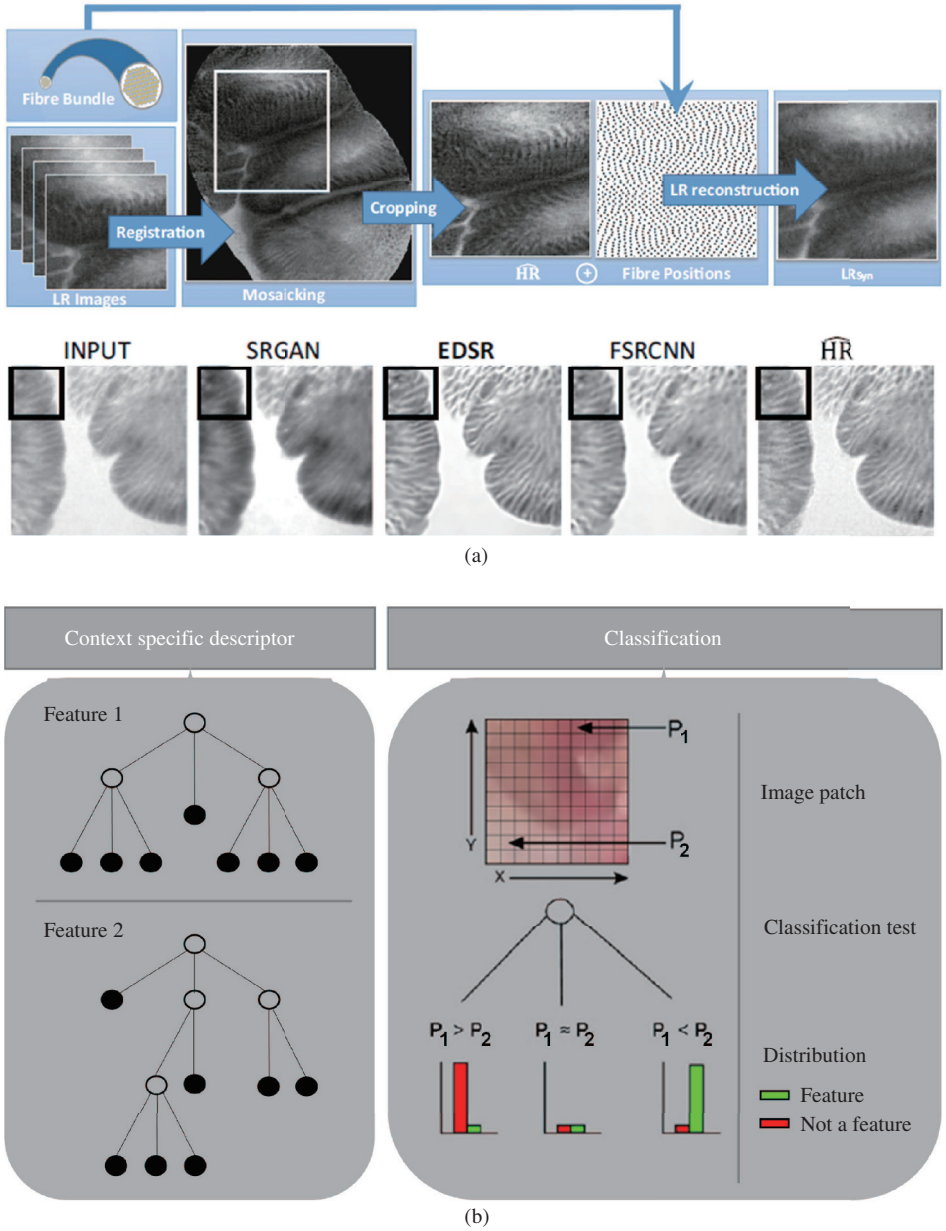


Figure 10 (Color online) Overview of methods applied in MIS imaging enhancement. (a) Ground-truth building strategy and super-resolution results of different DL-based methods and (b) registration method based on a decision tree. These images are reproduced from [39, 104].

standardized data, the generalization and robustness of ML methods, such as DL, are expected to improve. Furthermore, attention mechanism and other strategies [105–107] have been proposed to visualize the prediction basis of neural networks, which provide a feasible scheme for network interpretation. With the continuous developments in related research, the application of ML in OMI remains a promising technique.

Acknowledgements This work was supported by Ministry of Science and Technology of China (Grant Nos. 2018YFC0910602, 2017YFA0205200, 2017YFA0700401, 2016YFA0100902, 2016YFC0103702), National Natural Science Foundation of China (Grant Nos. 61901472, 61671449, 81227901, 81527805), the Strategic Priority Research Program of Chinese Academy of Sciences (Grant Nos. XDB32030200, XDB01030200), Chinese Academy of Sciences (Grant Nos. GJJSTD20170004, YJKYYQ20180048, KFJ-STS-ZDTP-059, QYZDJ-SSW-JSC005), Beijing Municipal Science & Technology Commission (Grant Nos. Z161100002616022, Z171100000117023), and General Financial Grant from the China Postdoctoral Science

Foundation (Grant No. 2017M620952). The authors would like to acknowledge the instrumental and technical support of Multi-modal biomedical imaging experimental platform, Institute of Automation, Chinese Academy of Sciences.

References

1 Conway J R W, Carragher N O, Timpson P. Developments in preclinical cancer imaging: innovating the discovery of therapeutics. Nat Rev Cancer, 2014, 14: 314–328
2 Maldiney T, Bessiére A, Seguin J, et al. The in vivo activation of persistent nanophosphors for optical imaging of vascularization, tumours and grafted cells. Nat Mater, 2014, 13: 418–426
3 Ellenbroek S I J, van Rheenen J. Imaging hallmarks of cancer in living mice. Nat Rev Cancer, 2014, 14: 406–418
4 Weissleder R, Pittet M J. Imaging in the era of molecular oncology. Nature, 2008, 452: 580–589
5 Massoud T F, Gambhir S S. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. Genes Dev, 2003, 17: 545–580
6 Fan-Minogue H, Cao Z W, Paulmurugan R, et al. Noninvasive molecular imaging of c-Myc activation in living mice. Proc Natl Acad Sci USA, 2010, 107: 15892–15897
7 Nguyen Q T, Tsien R Y. Fluorescence-guided surgery with live molecular navigation–a new cutting edge. Nat Rev Cancer, 2013, 13: 653–662
8 Weissleder R. Molecular imaging in cancer. Science, 2006, 312: 1168–1171
9 Jobsis F F. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. Science, 1977, 198: 1264–1267
10 Gao Y, Wang K, An Y, et al. Nonmodel-based bioluminescence tomography using a machine-learning reconstruction strategy. Optica, 2018, 5: 1451–1454
11 Jiang S X, Liu J, Zhang G L, et al. Reconstruction of fluorescence molecular tomography via a fused LASSO method based on group sparsity prior. IEEE Trans Biomed Eng, 2019, 66: 1361–1371
12 Li Y C, Charalampaki P, Liu Y, et al. Context aware decision support in neurosurgical oncology based on an efficient classification of endomicroscopic data. Int J Comput Assist Radiol Surg, 2018, 13: 1187–1199
13 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521: 436–444
14 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. Commun ACM, 2017, 60: 84–90
15 Farabet C, Couprie C, Najman L, et al. Learning hierarchical features for scene labeling. IEEE Trans Pattern Anal Mach Intell, 2013, 35: 1915–1929
16 Tompson J J, Jain A, LeCun Y, et al. Joint training of a convolutional network and a graphical model for human pose estimation. In: Proceedings of Advances in Neural Information Processing Systems 27. 2014
17 Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions. In: Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, 2015. 1–9
18 Mikolov T, Deoras A, Povey D, et al. Strategies for training large scale neural network language models. In: Proceedings of 2011 IEEE Workshop on Automatic Speech Recognition & Understanding, Waikoloa, 2011. 196–201
19 Hinton G, Deng L, Yu D, et al. Deep neural networks for acoustic modeling in speech recognition. IEEE Signal Process Magaz, 2012, 29: 82–97
20 Sainath T N, Kingsbury B, Saon G, et al. Deep convolutional neural networks for large-scale speech tasks. Neural Netw, 2015, 64: 39–48
21 Bengio Y, Ducharme R, Vincent P. A neural probabilistic language model. J Mach Learn Res, 2003, 3: 1137–1155
22 Sutskever I, Vinyals O, Le Q V. Sequence to sequence learning with neural networks. 2014. ArXiv: 1409.3215
23 Quan W Z, Wang K, Yan D M, et al. Distinguishing between natural and computer-generated images using convolutional neural networks. IEEE Trans Inform Forensic Secur, 2018, 13: 2772–2787
24 Bayar B, Stamm M C. Constrained convolutional neural networks: a new approach towards general purpose image manipulation detection. IEEE Trans Inform Forensic Secur, 2018, 13: 2691–2706
25 Yang Y, Zhang W S, He Z W, et al. Locator slope calculation via deep representations based on monocular vision. Neural Comput Applic, 2019, 31: 2781–2794
26 Ma J S, Sheridan R P, Liaw A, et al. Deep neural nets as a method for quantitative structure-activity relationships. J Chem Inf Model, 2015, 55: 263–274
27 Lemaître G, Rastgoo M, Massich J, et al. Classification of SD-OCT volumes using local binary patterns: experimental validation for DME DETECtion. J Ophthalmology, 2016, 2016: 1–14
28 Srinivasan P P, Kim L A, Mettu P S, et al. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. Biomed Opt Express, 2014, 5: 3568–3577
29 Lee C S, Baughman D M, Lee A Y. Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration. Ophthalmology Retina, 2017, 1: 322–327
30 Roy A G, Conjeti S, Karri S P K, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. Biomed Opt Express, 2017, 8: 3627
31 Roy A G, Conjeti S, Carlier S G, et al. Lumen segmentation in intravascular optical coherence tomography using backscattering tracked and initialized random walks. IEEE J Biomed Health Inform, 2016, 20: 606–614
32 Wang Z, Jenkins M W, Linderman G C, et al. 3-D stent detection in intravascular OCT using a Bayesian network and graph search. IEEE Trans Med Imag, 2015, 34: 1549–1561
33 Schwab J, Antholzer S, Nuster R, et al. Real-time photoacoustic projection imaging using deep learning. 2018.

ArXiv: 1801.06693
34 Hauptmann A, Lucka F, Betcke M, et al. Model-based learning for accelerated, limited-view 3-D photoacoustic tomography. IEEE Trans Med Imag, 2018, 37: 1382–1393
35 Antholzer S, Schwab J, Bauer-Marschallinger J, et al. Nett regularization for compressed sensing photoacoustic tomography. In: Proceedings of SPIE, 2019. 10878
36 Huang C, Meng H, Gao Y, et al. Fast and robust reconstruction method for fluorescence molecular tomography based on deep neural network. In: Proceedings of SPIE, 2019. 10881
37 André B, Vercauteren T, Buchner A M, et al. A smart atlas for endomicroscopy using automated video retrieval. Med Image Anal, 2011, 15: 460–476
38 Kamen A, Sun S H, Wan S H, et al. Automatic tissue differentiation based on confocal endomicroscopic images for intraoperative guidance in neurosurgery. Biomed Res Int, 2016, 2016: 1–8
39 Raví D, Szczotka A B, Shakir D I, et al. Effective deep learning training for single-image super-resolution in endomicroscopy exploiting video-registration-based reconstruction. Int J Comput Assist Radiol Surg, 2018, 13: 917–924
40 Zhang C, Wang K, An Y, et al. Improved generative adversarial networks using the total gradient loss for the resolution enhancement of fluorescence images. Biomed Opt Express, 2019, 10: 4742–4756
41 de Fauw J, Ledsam J R, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. Nat Med, 2018, 24: 1342–1350
42 Wang L V, Wu H I, Masters B R. Biomedical optics, principles and imaging. J Biomed Opt, 2008, 13: 049902
43 Gessert N, Lutz M, Heyder M, et al. Automatic plaque detection in IVOCT pullbacks using convolutional neural networks. IEEE Trans Med Imag, 2019, 38: 426–434
44 Foot B, MacEwen C. Surveillance of sight loss due to delay in ophthalmic treatment or review: frequency, cause and outcome. Eye, 2017, 31: 771–775
45 Ting D S W, Pasquale L R, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. British J Ophthalmol, 2019, 103: 167–175
46 Liu Y Y, Chen M, Ishikawa H, et al. Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding. Med Image Anal, 2011, 15: 748–759
47 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014. ArXiv: 1409.1556
48 Venhuizen F G, van Ginneken B, Liefers B, et al. Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography. Biomed Opt Express, 2018, 9: 1545
49 Tsantis S, Kagadis G C, Katsanos K, et al. Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography. Med Phys, 2012, 39: 503–513
50 Lu H, Gargesha M, Wang Z, et al. Automatic stent detection in intravascular OCT images using bagged decision trees. Biomed Opt Express, 2012, 3: 2809–2824
51 Yabushita H, Bouma B E, Houser S L, et al. Characterization of human atherosclerosis by optical coherence tomography. Circulation, 2002, 106: 1640–1645
52 Wang Z, Chamie D, Bezerra H G, et al. Volumetric quantification of fibrous caps using intravascular optical coherence tomography. Biomed Opt Express, 2012, 3: 1413–1426
53 Zahnd G, Karanasos A, van Soest G, et al. Quantification of fibrous cap thickness in intracoronary optical coherence tomography with a contour segmentation method based on dynamic programming. Int J Comput Assist Radiol Surg, 2015, 10: 1383–1394
54 Wang L V. Multiscale photoacoustic microscopy and computed tomography. Nat Photon, 2009, 3: 503–509
55 Kruger R A, Liu P Y, Fang Y R, et al. Photoacoustic ultrasound (PAUS)–reconstruction tomography. Med Phys, 1995, 22: 1605–1609
56 Karabutov A A, Podymova N B, Letokhov V S. Time-resolved laser optoacoustic tomography of inhomogeneous media. Appl Phys B-Lasers Opt, 1996, 63: 545–563
57 Ntziachristos V, Razansky D. Molecular imaging by means of multispectral optoacoustic tomography (MSOT). Chem Rev, 2010, 110: 2783–2794
58 Antholzer S, Haltmeier M, Schwab J. Deep learning for photoacoustic tomography from sparse data. Inverse Problems Sci Eng, 2019, 27: 987–1005
59 Xu M H, Wang L V. Universal back-projection algorithm for photoacoustic computed tomography. Phys Rev E, 2005, 71: 016706
60 Burgholzer P, Bauer-Marschallinger J, Grün H, et al. Temporal back-projection algorithms for photoacoustic tomography with integrating line detectors. Inverse Problems, 2007, 23: S65–S80
61 Zeng L, Xing D, Gu H M, et al. High antinoise photoacoustic tomography based on a modified filtered backprojection algorithm with combination wavelet. Med Phys, 2007, 34: 556–563
62 Hoelen C G A, de Mul F F M. Image reconstruction for photoacoustic scanning of tissue structures. Appl Opt, 2000, 39: 5872–5883
63 Rosenthal A, Razansky D, Ntziachristos V. Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography. IEEE Trans Med Imag, 2010, 29: 1275–1285
64 Paltauf G, Viator J A, Prahl S A, et al. Iterative reconstruction algorithm for optoacoustic imaging. J Acoust Soc Am, 2002, 112: 1536–1544
65 Jetzfellner T, Rosenthal A, Englmeier K H, et al. Interpolated model-matrix optoacoustic tomography of the mouse brain. Appl Phys Lett, 2011, 98: 163701

66 Treeby B E, Cox B T. k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields. J Biomed Opt, 2010, 15: 021314
67 Xu Y, Wang L V. Time reversal and its application to tomography with diffracting sources. Phys Rev Lett, 2004, 92: 033902
68 Hristova Y, Kuchment P, Nguyen L. Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media. Inverse Problems, 2008, 24: 055006
69 Dean-Ben X L, Ntziachristos V, Razansky D. Acceleration of optoacoustic model-based reconstruction using angular image discretization. IEEE Trans Med Imag, 2012, 31: 1154–1162
70 Dean-Ben X L, Buehler A, Ntziachristos V, et al. Accurate model-based reconstruction algorithm for three-dimensional optoacoustic tomography. IEEE Trans Med Imag, 2012, 31: 1922–1928
71 Huang C, Wang K, Nie L M, et al. Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media. IEEE Trans Med Imag, 2013, 32: 1097–1110
72 Arridge S R, Betcke M M, Cox B T, et al. On the adjoint operator in photoacoustic tomography. Inverse Problems, 2016, 32: 115012
73 Arridge S R, Beard P, Betcke M, et al. Accelerated high-resolution photoacoustic tomography via compressed sensing. Phys Med Biol, 2016, 61: 8908–8940
74 Hauptmann A, Cox B, Lucka F, et al. Approximate k-space models and deep learning for fast photoacoustic reconstruction. In: Machine Learning for Medical Image Reconstruction. Berlin: Springer, 2018. 103–111
75 Ntziachristos V, Ripoll J, Wang L V, et al. Looking and listening to light: the evolution of whole-body photonic imaging. Nat Biotechnol, 2005, 23: 313–320
76 Ntziachristos V, Bremer C, Weissleder R. Fluorescence imaging with near-infrared light: New technological advances that enable in vivo molecular imaging. Eur Radiol, 2003, 13: 195–208
77 Wang G, Li Y, Jiang M. Uniqueness theorems in bioluminescence tomography. Med Phys, 2004, 31: 2289–2299
78 Gao Y, Wang K, Jiang S X, et al. Bioluminescence tomography based on gaussian weighted laplace prior regularization for in vivo morphological imaging of glioma. IEEE Trans Med Imag, 2017, 36: 2343–2354
79 Qin C H, Zhu S P, Feng J C, et al. Comparison of permissible source region and multispectral data using efficient bioluminescence tomography method. J Biophoton, 2011, 4: 824–839
80 Arridge S R, Schweiger M, Hiraoka M, et al. A finite element approach for modeling photon transport in tissue. Med Phys, 1993, 20: 299–309
81 Arridge S R. Optical tomography in medical imaging. Inverse Problems, 1999, 15: R41–R93
82 Lu Y J, Zhang X Q, Douraghy A, et al. Source reconstruction for spectrally-resolved bioluminescence tomography with sparse a priori information. Opt Express, 2009, 17: 8062–8080
83 Liu K, Tian J, Qin C H, et al. Tomographic bioluminescence imaging reconstruction via a dynamically sparse regularized global method in mouse models. J Biomed Opt, 2011, 16: 046016
84 Chehade M, Srivastava A K, Bulte J W M. Co-registration of bioluminescence tomography, computed tomography, and magnetic resonance imaging for multimodal in vivo stem cell tracking. Tomography, 2016, 2: 158–165
85 Zhang X Q, Lu Y J, Chan T. A novel sparsity reconstruction method from poisson data for 3D bioluminescence tomography. J Sci Comput, 2012, 50: 519–535
86 Dutta J, Ahn S, Li C Q, et al. Joint l1 and total variation regularization for fluorescence molecular tomography. Phys Med Biol, 2015, 57: 1459–1476
87 Davis S C, Samkoe K S, O'Hara J A, et al. Comparing implementations of magnetic-resonance-guided fluorescence molecular tomography for diagnostic classification of brain tumors. J Biomed Opt, 2010, 15: 051602
88 Davis S C, Samkoe K S, Tichauer K M, et al. Dynamic dual-tracer MRI-guided fluorescence tomography to quantify receptor density in vivo. Proc Natl Acad Sci USA, 2013, 110: 9025–9030
89 Holt R W, Demers J L H, Sexton K J, et al. Tomography of epidermal growth factor receptor binding to fluorescent Affibody in vivo studied with magnetic resonance guided fluorescence recovery in varying orthotopic glioma sizes. J Biomed Opt, 2015, 20: 026001
90 Schulz R B, Ale A, Sarantopoulos A, et al. Hybrid system for simultaneous fluorescence and x-ray computed tomography. IEEE Trans Med Imag, 2010, 29: 465–473
91 Baikejiang R, Zhao Y, Fite B Z, et al. Anatomical image-guided fluorescence molecular tomography reconstruction using kernel method. J Biomed Opt, 2017, 22: 055001
92 Cho K, van Merrienboer B, Gulcehre C, et al. Learning phrase representations using rnn encoder-decoder for statistical machine translation. 2014. ArXiv: 1406.1078
93 Machida H, Sano Y, Hamamoto Y, et al. Narrow-band imaging in the diagnosis of colorectal mucosal lesions: a pilot study. Endoscopy, 2004, 36: 1094–1098
94 Gerger A, Koller S, Weger W, et al. Sensitivity and specificity of confocal laser-scanning microscopy for in vivo diagnosis of malignant skin tumors. Cancer, 2006, 107: 193–200
95 Gotoh K, Kobayashi S, Marubashi S, et al. Intraoperative detection of hepatocellular carcinoma using indocyanine green fluorescence imaging. In: ICG Fluorescence Imaging and Navigation Surgery. Tokyo: Springer, 2016. 325–334
96 Glatz J, Garcia-Allende P B, Becker V, et al. Near-infrared fluorescence cholangiopancreatoscopy: initial clinical feasibility results. Gastrointest Endosc, 2014, 79: 664–668
97 Adler A, Pohl H, Papanikolaou I S, et al. A prospective randomised study on narrow-band imaging versus conventional colonoscopy for adenoma detection: does narrow-band imaging induce a learning effect? Gut, 2007, 57: 59–64
98 Vahrmeijer A L, Hutteman M, van der Vorst J R, et al. Image-guided cancer surgery using near-infrared fluorescence.

Nat Rev Clin Oncol, 2013, 10: 507–518
99 Schaafsma B E, Mieog J S D, Hutteman M, et al. The clinical use of indocyanine green as a near-infrared fluorescent contrast agent for image-guided oncologic surgery. J Surg Oncol, 2011, 104: 323–332
100 Kitai T, Inomoto T, Miwa M, et al. Fluorescence navigation with indocyanine green for detecting sentinel lymph nodes in breast cancer. Breast Cancer, 2005, 12: 211–215
101 Tummers Q R J G, Verbeek F P R, Schaafsma B E, et al. Real-time intraoperative detection of breast cancer using near-infrared fluorescence imaging and methylene blue. Eur J Surgical Oncology, 2014, 40: 850–858
102 Keereweer S, van Driel P B A A, Snoeks T J A, et al. Optical image-guided cancer surgery: challenges and limitations. Clin Cancer Res, 2013, 19: 3745–3754
103 Andre B, Vercauteren T, Buchner A M, et al. Learning semantic and visual similarity for endomicroscopy video retrieval. IEEE Trans Med Imag, 2012, 31: 1276–1288
104 Mountney P, Yang G Z. Context specific descriptors for tracking deforming tissue. Med Image Anal, 2012, 16: 550–561
105 Hu J, Shen L, Sun G, et al. Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018. 7132–7141
106 Xu T, Zhang P C, Huang Q Y, et al. Attngan: fine-grained text to image generation with attentional generative adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018. 1316–1324
107 Woo S, Park J, Lee J Y, et al. Cbam: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), 2018. 3–19