We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

5,900
Open access books available

145,000
International authors and editors

180M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK
CITATION
INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Classification of Malaria-Infected Cells Using Deep Convolutional Neural Networks

W. David Pan, Yuhang Dong and Dongsheng Wu

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72426

Abstract

Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected mosquitoes. Automation of the diagnosis process will enable accurate diagnosis of the disease and hence holds the promise of delivering reliable healthcare to resource-scarce areas. Machine learning technologies have been used for automated diagnosis of malaria. We present some of our recent progresses on highly accurate classification of malaria-infected cells using deep convolutional neural networks. First, we describe image processing methods used for segmentation of red blood cells from wholeslide images. We then discuss the procedures of compiling a pathologists-curated image dataset for training deep neural network, as well as data augmentation methods used to significantly increase the size of the dataset, in light of the overfitting problem associated with training deep convolutional neural networks. We will then compare the classification accuracies obtained by deep convolutional neural networks through training, validating, and testing with various combinations of the datasets. These datasets include the original dataset and the significantly augmented datasets, which are obtained using direct interpolation, as well as indirect interpolation using automatically extracted features provided by stacked autoencoders. This chapter ends with a discussion of further research.

Keywords: deep learning, convolutional neural network, autoencoders, data augmentation, classification, wholeslide images, malaria

1. Introduction

Malaria is a widespread disease that has claimed millions of lives all over the world. According to the World Health Organization, approximately 438,000 deaths result from 214 million

infections in 2015 [1]. Endemic regions with widespread disease include Africa and South-East Asia. In these and other parts of the world where malaria mortality is significant, necessary resources such as reliable prevention, healthcare, and hygiene are far from adequate [1]. In most cases, the only available method of malaria diagnosis is manual examination of the microscopic slide [2]. In order to provide reliable diagnosis, extensive experience and training are required. Unfortunately, such specialized human resources are very often limited in rural areas where malaria has a marked predominance. Also, manual microscopy is subjective and suffers from a lack of standardization. This problem is further exacerbated by the large size of microscopic wholeslide images, which require a lengthy scanning.

## 1.1. The need for an automated malaria diagnosis process

The issues associated with manual diagnosis present the case for automation of the malaria diagnosis process. The automation of the diagnosis process will ensure accurate diagnosis of the disease and hence holds the promise of delivering reliable health-care to resource-scarce areas. Hence, rural areas suffering from lack of specialized infrastructure and trained manpower can benefit greatly from automated diagnosis. Automating the diagnosis of malaria involves adapting the methods, expertise, practices, and knowledge of conventional microscopy to a computerized system structure [3]. Early detection of malaria is essential for ensuring proper diagnosis and increasing chances of cure. In consideration of the severity and the number of fatalities claimed by this disease, it is rational to accept potential small implementation errors introduced by an automated system. An automated system consists of streamlined image processing techniques for initial filtering and segmentation and suite of pattern recognition and/or machine learning algorithms directed toward robustly recognizing infected cells in a light or wholeslide microscopic image [4]. Previous studies have shown that the degree of agreement between clinicians on the severity of the disease in a given patent's sample is very low. Hence, a computer-assisted system as a decision support system can be paramount to faster and reliable diagnosis. It can help provide a benchmark and standardized way of measuring the degree of infection of the disease [5].

1.2. Wholeslide images for computer-aided malaria infection classification

Among recent works on computer-aid diagnosis, two types of images have found prevalent use: light microscopic images and wholeslide images. The former has been in existence since a longer time frame compared to the latter, which has come into popular adoption recently. Because of recent advancements in computing power, improved cloud-based services and robust algorithms have enabled the widespread use of wholeslide images. For conventional light microscopy, the patient tissue image is acquired by means of incision and then examined under a light microscope. A diagnostic conclusion is arrived upon based interpretation of multiple slide samples [6]. This type of examination does not provide a good sensitivity and specificity for malaria diagnosis [7]. With an aim to standardize slide interpretations, wholeslide images were introduced. The wholeslide image is obtained by scanning an entire slide in one pass. The final image consists of several component images obtained by scanning the areas under the respective fields of sight of the microscope and stitched together. The most

widely used methods of scanning include tile scanning and line scanning. In tile scan, the component images are obtained in the form of 512 × 512 tiles. In the line scan method, the component images are generated in a strip-scan fashion.

The file sizes of the wholeslide images are governed by the objective of the lens while scanning. Wholeslide images scanned 40× objective give rise to substantially large file size, for instance, approximately 2 GB. Magnification beyond the maximum level can result in pixelation [8]. Wholeslide images can be decomposed into a pyramid structure of different resolutions. The image at each magnification level is broken down into smaller constituent tiles and stored in respective folders. The image pyramid allows for real-time viewing of wholeslide images. The zoom levels are precalculated and stored in the metadata associated with the file [9]. Each tile can be viewed and analyzed individually. **Figure 1** shows an illustration [10]. The process of
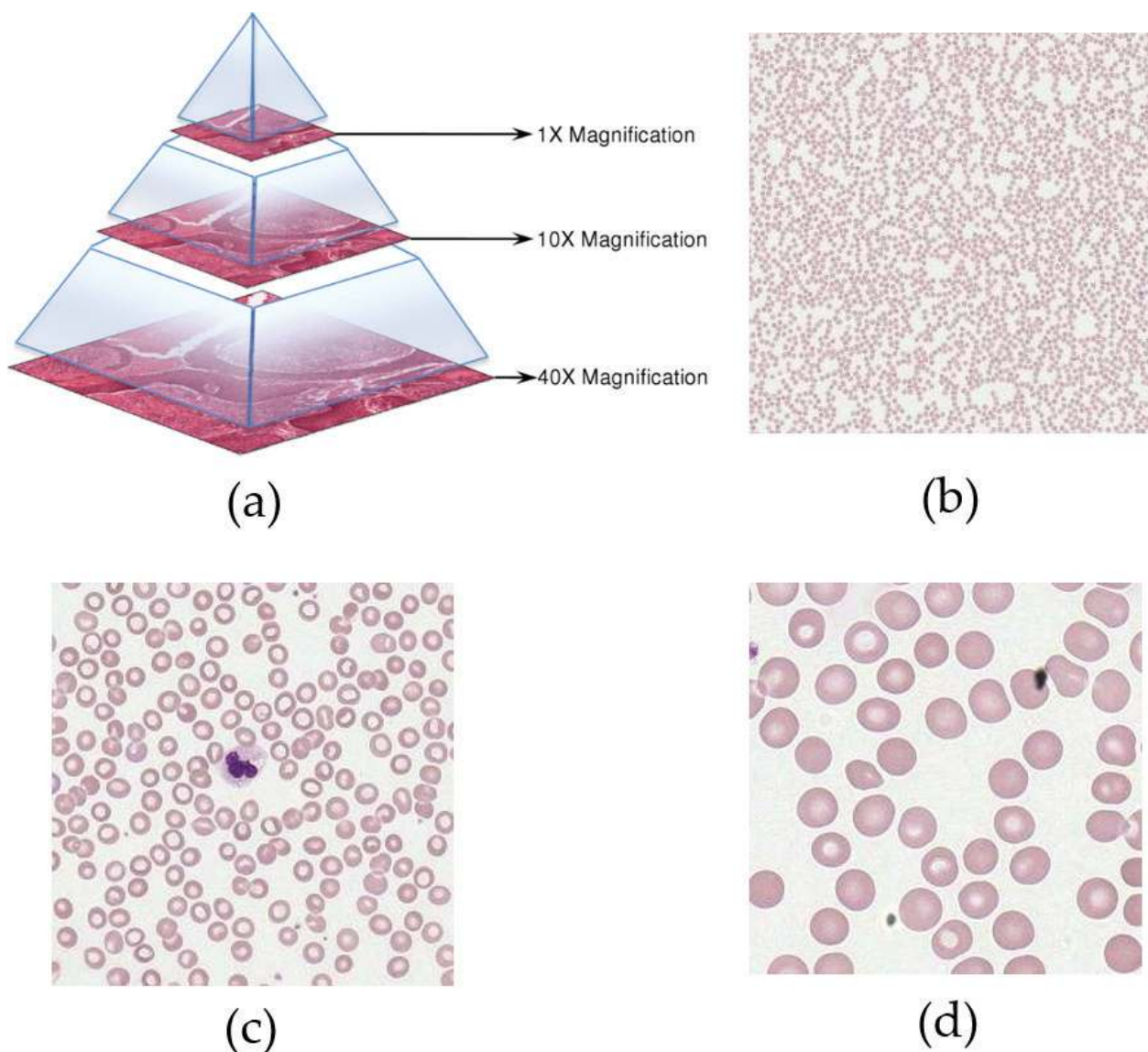
Figure 1. (a) Image pyramid of the DeepZoom structure. Example image tiles at various DeepZoom pyramid resolutions (b) Level 14, (c) Level 15, and (d) Level 16. Each tile image is 258 × 258, and magnification is 100×.

examining a wholeslide image is termed as "virtual microscopy", since analysis and examination can be performed through compatible software virtually on the computer. The DeepZoom structure is favorable in terms of storage and transmission. This arrangement of the original wholeslide image allows for smooth loading and panning using multiresolution images. Initially upon loading, a low-resolution version of the image is displayed. The higher resolution details get blended into the image as they become available. Thus, while viewing the image in DeepZoom, the user experiences a blurred image to sharp image transition. In terms of transmission, the DeepZoom structure is bandwidth efficient. Since, initially, a coarse low-resolution version is transmitted, the bandwidth overhead is reduced. At each level, each tile can be worked on individually [11].

1.3. Classification of malaria-infected red blood cells using deep learning

There has recently been an increasing amount of studies devoted to the application of computer vision and machine learning technologies to the automated diagnosis of malaria. Among the most recent related work [12–16], an automated analysis method was presented in [14] for detection and staging of red blood cells (RBCs) infected by the malaria parasite. In order to classify RBCs, three different types of machine learning algorithms were tested for prediction accuracy and speed as RBC classifiers. In [12], the authors built a low-cost automated digital microscope coupled with a set of computer vision and classification algorithms. Support vector machine (SVM) has been applied to detect malaria-infected cells using provided handcrafted features. In our prior work [17], we sought the best features from a set of 76 features organized into five categories extracted from the input data, in order to optimize SVM-based classification of wholeslide malarial smear images. We found that the binary SVM classifier yielded a superlative accuracy of 95.5% if the feature-selection is based on Kullback-Leibler distance. In contrast, deep learning has appeared as a genre of machine learning algorithms, which attempt to solve problems by learning abstraction in data following a stratified description paradigm based on non-linear transformation architectures. Recent advances in deep machine learning provide tools to automatically classify images and objects with (and occasionally exceeding) human-level accuracy. A key advantage of deep learning is its ability to perform semi-supervised or unsupervised feature extraction over massive datasets.

Deep learning has found exciting new applications in biomedicine [18], genomic medicine [19], bioinformatics [20], and medical imaging analysis [21–28]. However, there has been very sparse work on applying deep learning methods to computer-assisted malaria infection detection. In [16] were described point-of-care diagnostics using microscopes and smartphones, where deep convolutional neural network (CNN) was employed to identify image patches suspected to contain malaria-infected RBCs. The detection accuracy is similar to the results achieved with deep learning [15], where a CNN (with three convolutional layers and two fully connected layers) achieved a precision of 95.31% using images from dedicated microscope cameras [16]. Nevertheless, deep learning methods typically involve the calculation of tens of thousands of parameters, which in turn require large training datasets that may not be readily available. Thus, many commonly used machine learning methods such as support vector machine can outperform deep learning methods when experimental data is scarce. When the

datasets are not sufficiently large, one of the major challenges with training deep CNNs is to deal with the risk of overfitting. When training error is low but the test error is high, the model fails to learn a proper generalization of knowledge contained in data [18]. There are ways to regularize the deep network, such as randomized pruning of excessive connectivity, but overfitting is still a threat with small image datasets, especially with unbiased data.

In this chapter, we present some of our recent progresses on highly accurate classification of malaria-infected cells using deep convolutional neural networks. We will discuss the procedures of compiling a pathologists-curated image dataset for training deep neural network, as well as data augmentation methods used to significantly increase the size of the dataset, in light of the overfitting problem associated with training deep convolutional neural networks. In the next section, we describe image processing methods used for segmentation of red blood cells from wholeslide images.

2. Cell image pre-processing and compilation of dataset for deep learning

The images used in this work were wholeslide images provided in the PEIR-VM repository built by the University of Alabama in Birmingham. The original whole slide image data contain significant amount of redundant information. In order to achieve good classification accuracy, image segmentation and de-noising are needed to extract only blood cells and remove those redundant image pixels simultaneously. Several effective image processing techniques were used to accurately segment tiles into individual cells.

2.1. Image pre-processing tasks

Most image tiles may easily be visualized as having no malaria-infected cells, so preselection of noninfected tiles can be used to significantly reduce overall processing runtime. Given the contrast between the darkly purple/blue-stained nuclei of malaria and the light pink color of normal cells, pixel color information is used for preliminary selection of "infected" tiles. In order to estimate the color of infected cells, we conducted statistical analysis on the collected cell pixels. The maximal and minimal RGB values of infected cells were selected as two thresholds for "suspect" tiles. Considering the risk of excluding infected cells, we expanded the selected RGB value range to include more tiles. In this work, 24,648 of the original 85,094 tiles (29%) were marked as suspect and require further analysis.

For the suspected tile, thresholding is performed on the binarized image using Otsu's method. An example is shown in **Figure 2**. We can see that noise not only exists in the image background but also inside RBCs. A series of morphological steps were applied to fill the isolated dots and holes to finally obtain the individual cell samples.

In our work, only RBCs will provide features in the wholeslide image to the following classification. Therefore, we only keep RBCs and remove everything else using a combination of morphological operations. After all RBCs are processed, we then obtain all clean RBC samples for further classification. **Figure 3** shows some normal and infected RBC samples.
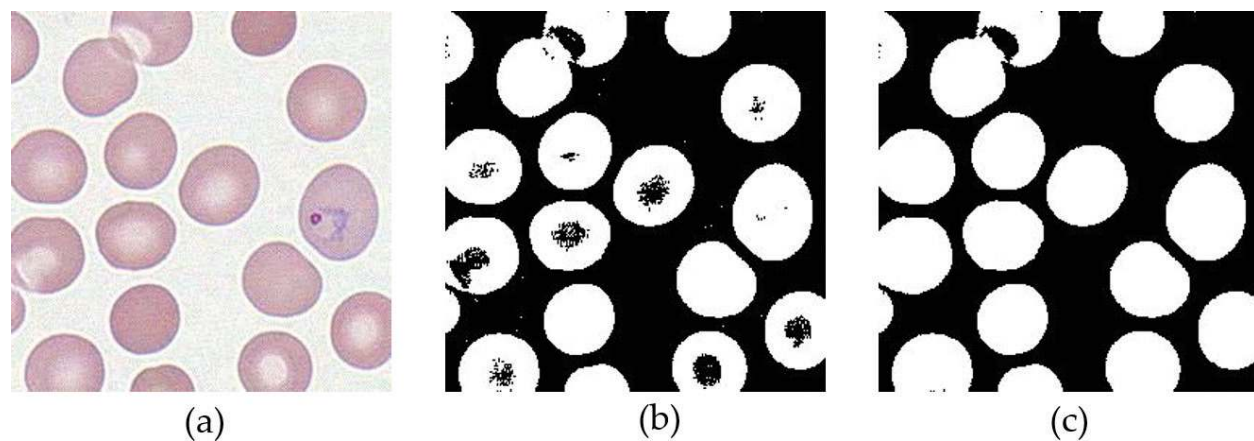
Figure 2. Steps of image pre-processing. (a) An image tile of interest; (b) Otsu thresholded image; and (c) morphologically filled image.
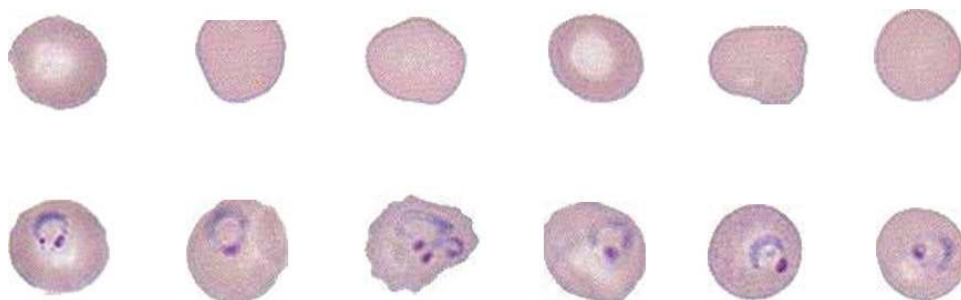
Figure 3. Some example segmented red blood cell images. (Upper row) normal cells and (lower row) infected cells.

2.2. Construction of an image dataset

There is no sufficiently large, high-quality image dataset of pathologically annotated cell images available to fully train multiple-layer neural networks. The only reasonably large, publicly available dataset in [16] we are aware of contains only 2703 images. However, these images were taken from thick blood smears, showing blurry patches rather than extractable RBCs found in high-resolution wholeslide images scanned from thin blood smears. Therefore, we worked with a team of pathologists to construct a dataset. After the data preprocessing, we randomly selected a large number of cell images and provided them to pathologists at the University of Alabama at Birmingham. The entire whole slide image dataset have been divided into four segments evenly. Each of four pathologists is assigned with two segments so that each cell image will be viewed and labeled by at least two experienced pathologists. One cell image can only be considered as infected and included in our final dataset if all the reviewers mark it positively whereas it will be excluded otherwise. The same selection rule also applies to the normal cells in our dataset.

3. Convolution neural network

Convolutional neural network is an artificial neural network inspired by the animal visual system [29]. Convolutional layer, pooling layer, and fully connection layer are the three main types of layers used to construct the CNN architecture. Compared to traditional neural networks, CNNs can extract features without losing much spatial correlations of the input. Each layer consists of neurons that have learnable weights and biases. The optimal model is achieved after feeding data into the network and minimizing the loss function at the top layer. Several different architectures of CNN have been proposed. In this work, we used LeNet-5. LeNet-5 [30] was first used in handwritten digit recognition and achieved an impressive error rate as low as 0.8%. **Figure 4** shows the architecture of the LeNet-5 convolutional neural network used for classification of the red blood cell images.

One of the major challenges of the research is that the current image dataset is still too small, which could lead to overfitting when used for training deep convolutional neural network. To this end, we consider data augmentation. More similar images can be added to the dataset by applying to the existing images operations such as rotation, translation, flip, zoom, and color perturbations. Other methods include data augmentation in the spatial domain by learning the statistical models of data transformation [31], as well as data augmentation through interpolation and extrapolation in the feature domain ([32, 33]). In the following, we present our work in augmenting the image dataset of the red blood cells and discuss the impact of the data augmentation on the image classification accuracies using deep convolutional neural network.
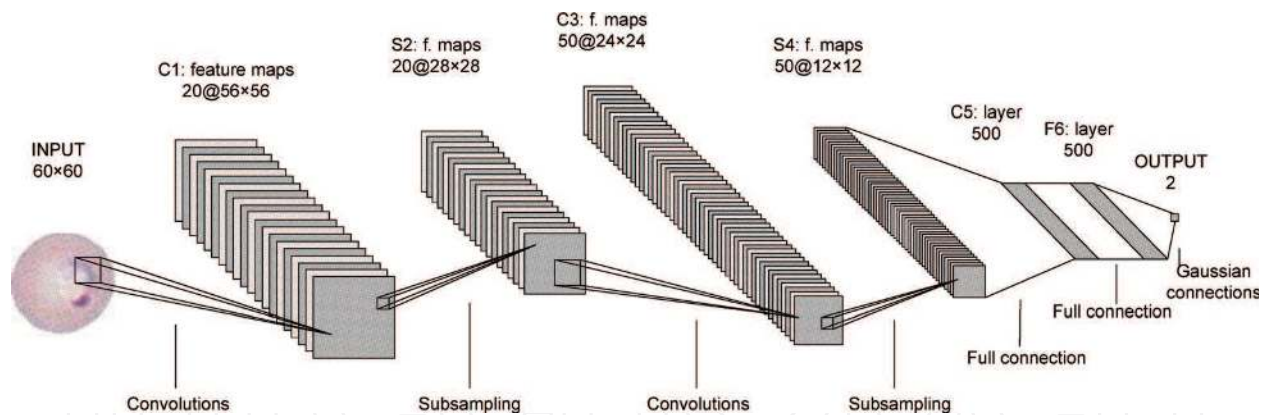
Figure 4. LeNet-5 convolutional neural network architecture. There are two convolution layers (C1 and C3), two subsampling (pooling) layers (S2 and S4), and two fully connected layers.

4. Image data augmentation

The set of infected red blood cell images has 800 images, each with size of $50 \times 50 \times 3$ (for red, green, and blue channels). Only the red channel pixel values were used. Since we want to

evaluate the quality of the augmented data set, we used only half of the infected cell images (400 images) for data augmentation, with the remaining 400 images untouched. The same configuration applies to the set of normal red blood cell images, which contains 4000 images. Only half (2000 images) of the dataset were used for augmentation.

We first describe the algorithms for data augmentation by using image interpolation in the spatial domain (Section 4.1), and in the feature domain (Section 4.2), respectively. As a comparison, we then present some example read blood cell images to show the effect of image interpolation in the spatial and feature domains at the end of Section 4.2.

4.1. Image interpolation in the spatial domain

For any two images A and B in the dataset, we can generate a new image by finding a weighted average C. Specifically, the pixel at location (i,j) in C can be obtained by

$$C(i,j) = \min[A(i,j), B(i,j)] + k \times \{\max[A(i,j), B(i,j)] - \min[A(i,j), B(i,j)]\}, \qquad (1)$$

where k is a weight ranging between 0 and 1. It can be seen that $C(i,j) = \min[A(i,j), B(i,j)]$ for $k = 0$; $C(i,j) = \max[A(i,j), B(i,j)]$ for $k = 1$. By varying the k values, for example, from 0 to 1 with a step size of 0.1, we can create 11 different images for any two input images. Assume the number of images in the dataset to be augmented is N, we can generate $\frac{11N(N-1)}{2}$ images, which can lead to a much enlarged dataset.

4.2. Image interpolation in the feature domain

To obtain the features of the red blood cell images, we used Hinton's autoencoder [34], which in essence is artificial neural network that performs unsupervised learning on the input data [35]. In the encoding phase, low-dimensional representations of the input data are learned through training the neural network. These learned representations are extracted features of the input image. The features can then be used to reconstruct the original data (decoding). The training algorithm will seek to optimize the neural network by minimizing the reconstruction loss as a cost function on sufficiently large amount of data. Moreover, a deep neural network can be constructed by concatenating multiple autoencoders. This would allow for a hierarchical representation of the data through a multilayer architecture. In [34], the Restricted Boltzmann Machine (RBM) was used as an autoencoder, which serves as a building block of a deep autoencoder network. Each RBM was pretrained and unrolled. Then, back propagation was carried out to fine-tune the entire stacked autoencoder based on cross entropy as the cost function.

In our implementation, the numbers of neurons in each of the four layers are 2500-1500-500-30. The maximum number of epochs for training the autoencoder was set to 1000, and back propagation was set to 500 iterations. **Figure 5** shows the architecture of the stacked autoencoders and autodecoders, where data interpolation is performed on the 30-point feature vectors. **Figure 6** shows some examples of reconstructed images.
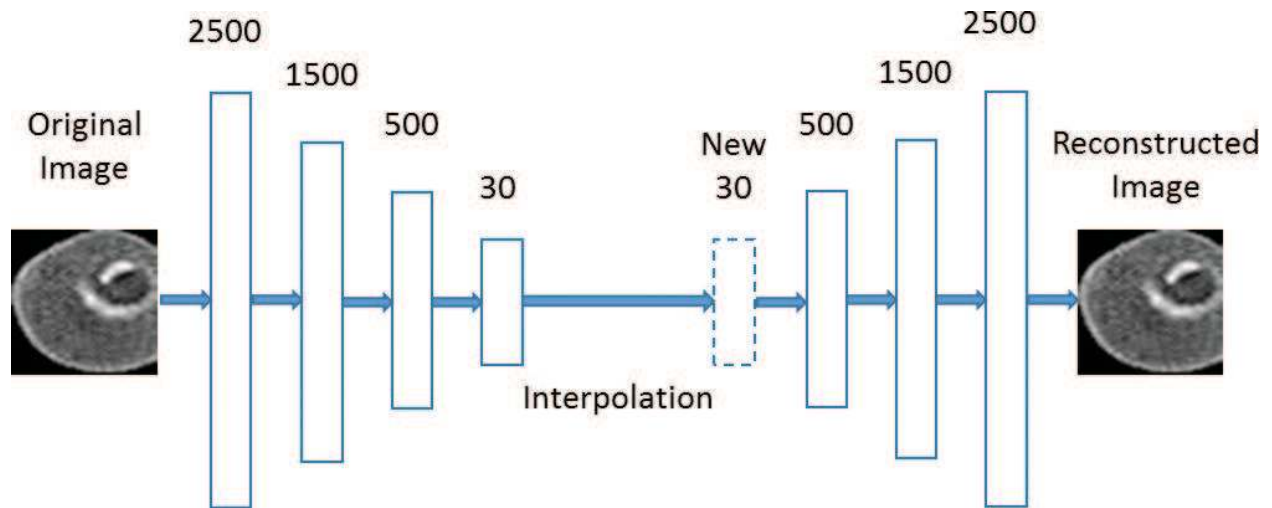
Figure 5. A four-layer stacked autoencoders and autodecoders used to extract the 30-point features of the input image in an unsupervised learning manner. After network training, interpolation of the input images is performed using their 30-point features. The resulting feature vector was then used to reconstruct the red blood cell image by the autodecoders.

(a)

(b)
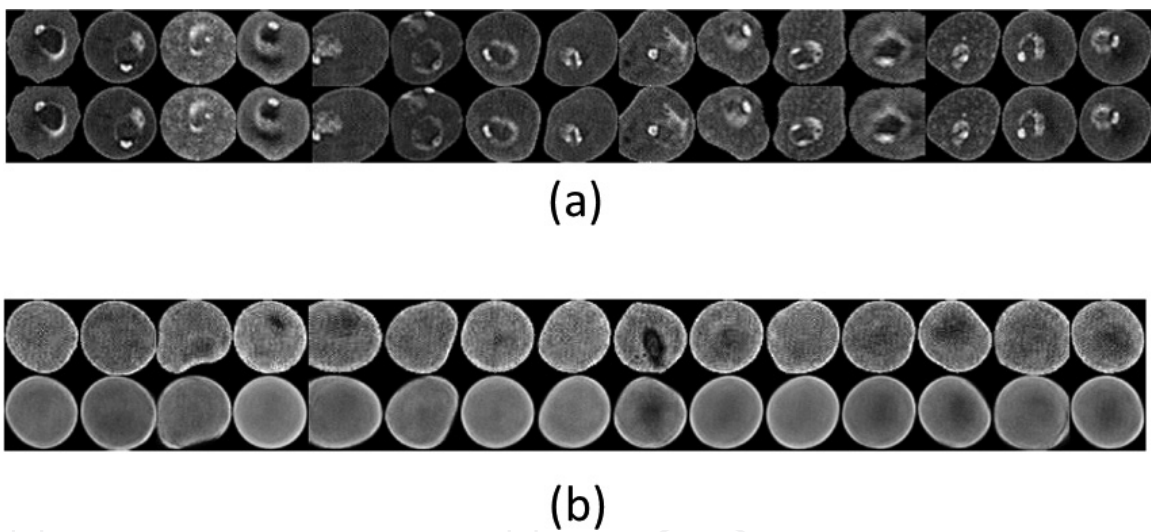
Figure 6. Original images (top row) and the reconstructed images (bottom row) using the stacked autodecoders. (a) Malaria-infected cells and (b) normal cells.

For any two images A and B in the dataset, two 30-point feature vectors F_A and F_B can be obtained by the stacked autoencoders that have been trained. Similar to the image interpolation in the spatial domain, we can generate a new 30-point vector F_C by finding a weighted average. Specifically, the pixel at location (i, j) in C can be obtained by

$$F_C = \min[F_A, F_B] + k \times \{\max[F_A, F_B] - \min[F_A, F_B]\}, \tag{2}$$

where k is a weight varied between 0 and 1 with a step size of 0.1. The newly generated feature vectors are then fed into the trained autodecoders to reconstruct the image $\overline{C}$ in the spatial domain.

4.3. Results of image interpolation

As a visual comparison of the effect of image interpolation in the spatial and feature domains, **Figure 7** shows the result of interpolating from two example red blood cell images.

4.4. Image classification using the original and augmented datasets

The original image dataset is split into two halves, as shown in **Table 1**. The first half (Dataset 1) was used for data augmentation. Using data augmentation methods discussed above, images of 400 infected cells were increased to 4000 cells, and images of 1000 normal cells were increased to 10,000 cells. Consequently, we created two datasets, one as the result of using spatial domain interpolation, the other as the result of using feature domain (via stacked autoencoders) interpolation, as shown in **Table 2**. Note that the samples for validation were randomly selected.

(a)

(b)
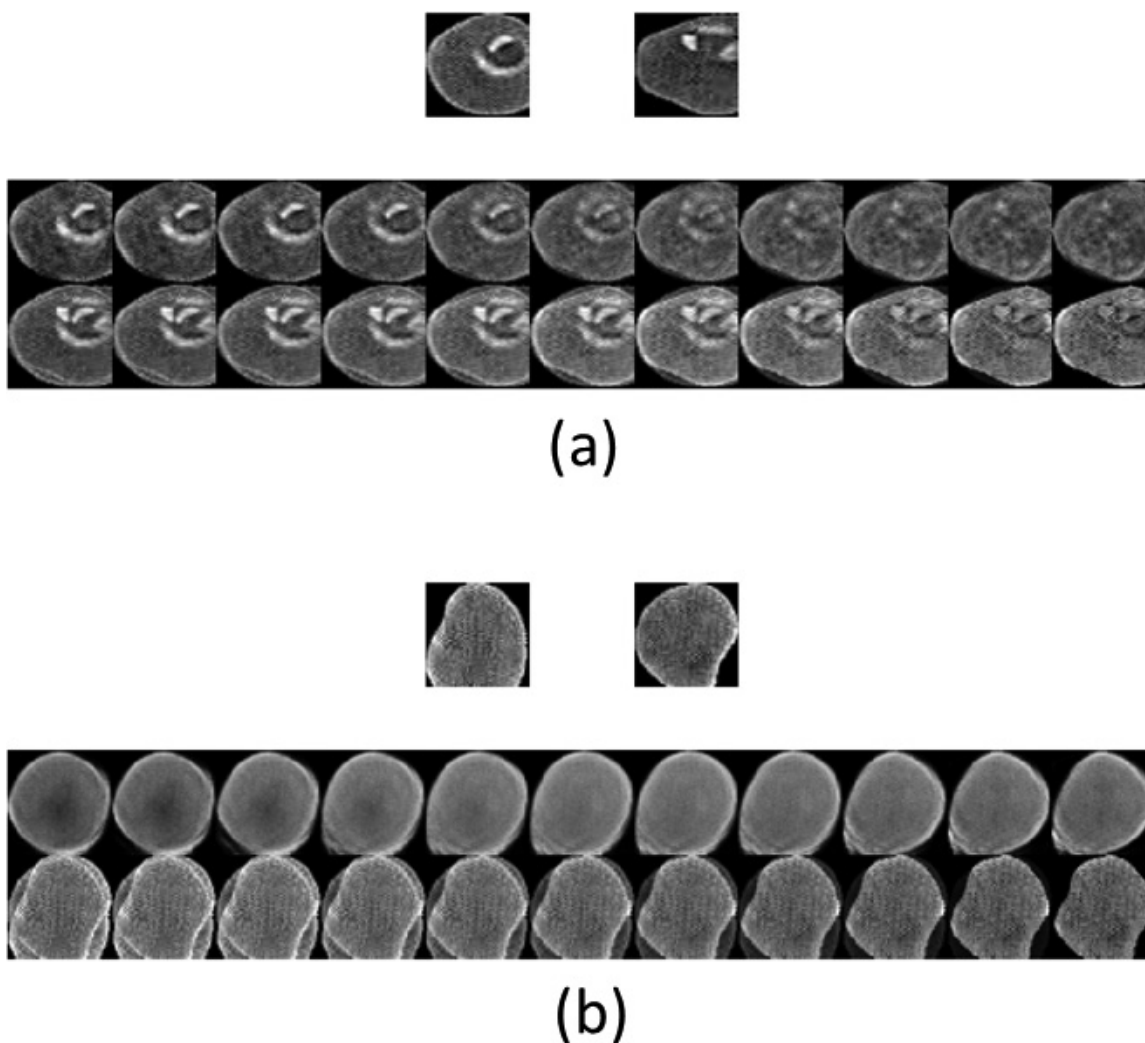
Figure 7. Result of interpolation using two example images. (a) Top row: two images of malaria-infected red blood cells used to generate a new image using interpolation. Middle row: 11 images obtained by interpolation using Eq. (1), where k is a weight varied between 0 and 1 with a step size of 0.1. Bottom row: 11 images obtained by interpolation in the feature domain using Eq. (2), where k is a weight varied between 0 and 1 with a step size of 0.1. (b) Similar to (a), images obtained by interpolation in the spatial (middle row) and the feature domain (bottom row).

We conducted various simulations based on the configuration shown in **Table 3**. For example, we used the augmented images in Dataset 3 to train the LeNet-5 convolutional neural network, and tested the original images in Dataset 2 using the trained network in order to classify the images into two categories: either infected or normal cells. Inversely, we trained the LeNet-5 using the original dataset and tested using the augmented datasets, in order to see how the trained classifier would perform on the augmented datasets.

Simulation results are shown in **Figures 8** and **9**.

It can be seen in **Figure 8** that training and validation using the augmented dataset provides fairly high accuracy (above 90%) when testing using the original dataset, implying the augmented data agree reasonably well statistically with the original data. Besides, feature domain interpolation seems to offer higher accuracy than spatial domain interpolation. Furthermore, the classification accuracies vary more significantly with the interpolation (mixing) coefficient k for spatial domain interpolation than for feature domain interpolation. For both interpolation

Original dataset	# of infected cells	# of normal cells	# of infected cells for training (T) and validation (V)	# of normal cells for training (T) and validation (V)
Dataset 1	400	1000	N/A	N/A
Dataset 2	400	1000	(T:320, V: 80)	(T: 800, V:200)

The first half (Dataset 1) was used for data augmentation. The second half (Dataset 2) was used for training and testing.

Table 1. The original image dataset is split into two halves.

| Augmented dataset	# of infected cells	# of normal cells	# of infected cells for training (T) and validation (V)	# of normal cells for training (T) and validation (V)
Dataset 3	4000	10,000	(T:3200, V:800)	(T:8000, V:2000)
Dataset 4	4000	10,000	(T:3200, V:800)	(T:8000, V:2000)

Dataset 3 was obtained using the spatial domain interpolation and Dataset 4 was obtained using the feature domain (via stacked autoencoders) interpolation.

Table 2. The augmented image dataset from dataset 1 in **Table 1**.

Training (infected, normal)	Validation (infected, normal)	Testing (infected, normal)
Dataset 3 (3200, 8000)	Dataset 3 (800, 2000)	Dataset 2 (400, 1000)
Dataset 4 (3200, 8000)	Dataset 3 (800, 2000)	Dataset 2 (400, 1000)
Dataset 2 (320, 800)	Dataset 2 (80, 200)	Dataset 3 (4000, 10,000)
Dataset 2 (320, 800)	Dataset 2 (80, 200)	Dataset 4 (4000, 10,000)

Table 3. Several combinations of datasets used in training, validation and testing of the convolutional neural network.
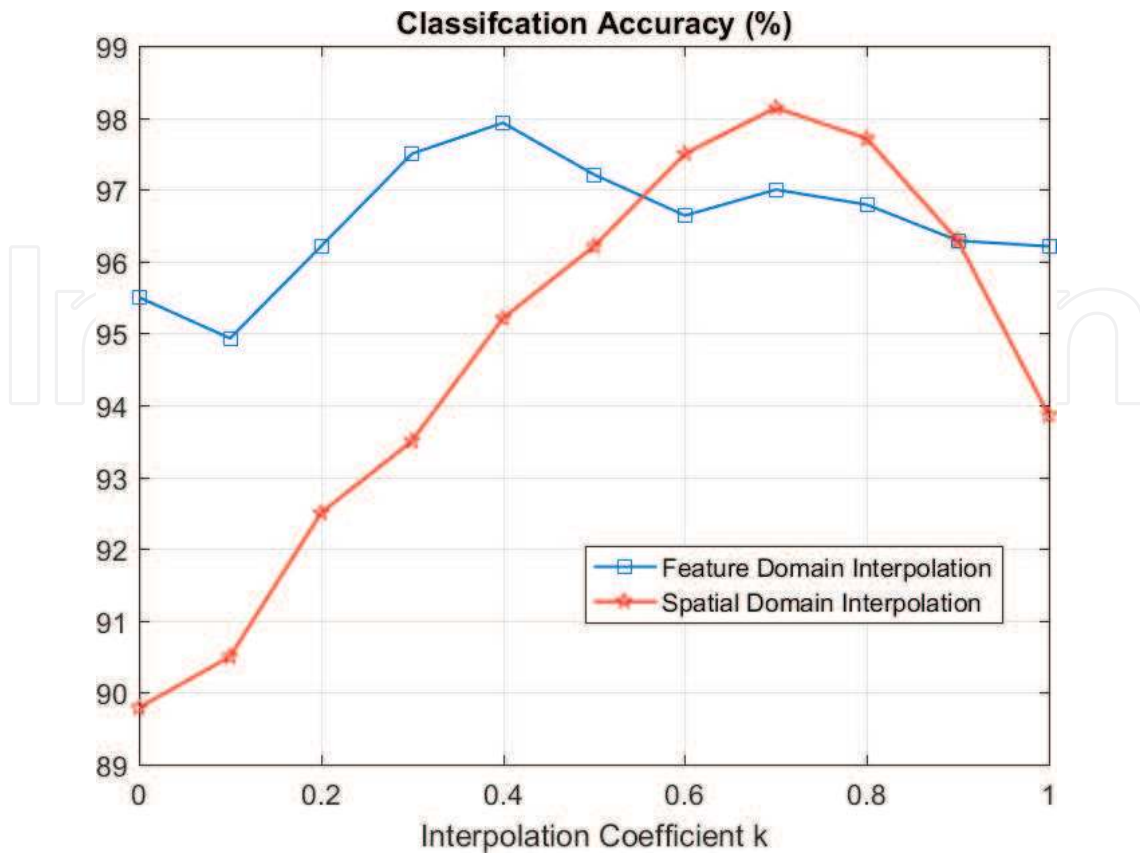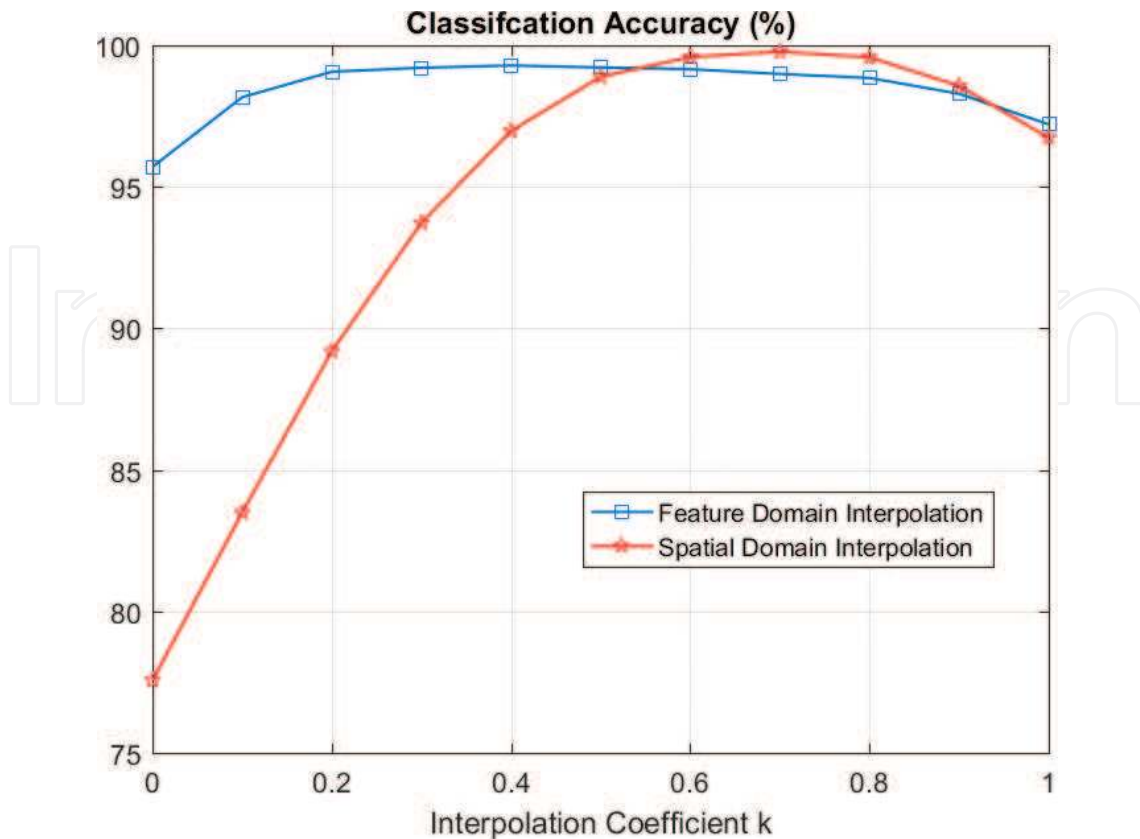
Figure 8. Classification accuracies as a function of the interpolation coefficient k in Eqs. (1) and (2), when we trained the LeNet-5 using the augmented datasets and tested the classifier using the original dataset. The augmented datasets are dataset 3 with spatial domain interpolation and dataset 4 with feature domain interpolation, respectively.

methods, there exists an optimal coefficient such that the classification accuracy reaches its maximum.

It can be seen in **Figure 9** that by using the classifier trained on the original dataset, we might get very low (below 80%) classification accuracy on the input images in the augmented dataset (obtained using interpolation in the spatial domain). This highlights the importance of using data augmentation in order to attain a more balanced estimation of the generalization ability of the classifier. This generalization ability seems to depend heavily on the varying new image samples that were interpolated using a different interpolation (mixing) coefficient from the original dataset (e.g., when $k = 0.7$, the accuracy can reach about 99%). **Figure 9** also shows that Dataset 4 (feature domain interpolation) seems to be a less challenging dataset than Dataset 3 (spatial domain interpolation), in that all accuracies are above 95%, possibly suggesting that mixing images using their features extracted by the stacked autoencoder would generate less diverse images than directly mixing images in the spatial domain.

Figure 9. Classification accuracies as a function of the interpolation coefficient k in Eqs. (1) and (2), when we trained the LeNet-5 using the original dataset (dataset 2) and tested the classifier using the augmented datasets, where dataset 3 is with spatial domain interpolation and dataset 4 with feature domain interpolation.

5. Conclusions and further research

Malaria is a widespread disease that has claimed millions of lives all over the world. Automation of the diagnosis process will provide accurate diagnosis of the disease, which will benefit health-care to resource-scarce areas. We showed that the deep convolutional network based on LeNet-5 was capable of achieving very high classification accuracies for automated malaria diagnosis, by automatically learn the features from the input image data. We briefly described the workflow of classification of the red blood cell images, and discussed in details the data augmentation methods we proposed to deal with the issue with training deep convolutional neural networks with a small dataset. We then compared the classification accuracies associated with training, validating, and testing with various combinations of the original dataset and the significantly augmented datasets, which were obtained using direct interpolation in the spatial domain, as well as indirect interpolation using automatically extracted features provided by stacked autoencoders. This comparative study indicated that data augmentation in the feature domain seemed to be more robust in terms of preserving

the high classification accuracies. We plan to expand the existing dataset by including more pathologist-curated cell images and further evaluate the effectiveness of the proposed data augmentation methods.

Author details

W. David Pan[1]*, Yuhang Dong[2] and Dongsheng Wu[2]

*Address all correspondence to: pand@uah.edu

1 Department of Electrical and Computer Engineering, University of Alabama in Huntsville, Huntsville, AL, USA

2 Department of Mathematical Sciences, University of Alabama in Huntsville, Huntsville, AL, USA

References

[1] Chan M. World Malaria Report. Geneva: World Health Organization; 2015

[2] Kettlehut MM et al. External quality assessment schemes raise standards: Evidence from the UKNEQAS parasitology subschemes. Journal of Clinical Pathology. 2003 Dec;**56**:927-932

[3] Tek FB. Computerised diagnosis of malaria [PhD Thesis]. University of Westminster; 2007

[4] Tek FB, Dempster AG, Kale I. Computer vision for microscopy diagnosis of malaria. Malaria Journal. 2009 Dec;**8**(1):1-14

[5] Mitiku K, Mengistu G, Gelaw B. The reliability of blood film examination for malaria at the peripheral health unit. Ethiopian Journal of Health Development. 2003;**17**(3):197-204

[6] Center for Devices and Radiological Health. Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices. Center for Devices and Radiological Health. Silver Spring, MD, USA: Food and Drug Administration; 2016

[7] World Health Organization. Microscopy. Available: http://www.who.int/malaria/areas/diagnosis/microscopy/en/

[8] Farahani N, Parwani AV, Pantanowitz L. Whole slide imaging in pathology: Advantages, limitations, and emerging perspectives. Pathology and Laboratory Medicine International. 2015 June;**2015**(7):23-33

[9] Cornish TC, Swapp RE, Kaplan KJ. Whole-slide imaging: Routine pathologic diagnosis. Advances In Anatomic Pathology. 2012 May;**19**(3):152-159

[10] Consortium for Open Medical Image Computing. ISBI Challenge on Cancer Metastasis Detection in Lymph Node. Available: http://grand-challenge.org/site/camelyon16/data/

[11] Microsoft Corporation. Deep Zoom. Available: https://msdn.microsoft.com/en-us/library/cc645050(VS.95).aspx

[12] Delahunt CB, Mehanian C, Hu L, McGuire SK, Champlin CR, Horning MP, et al. Automated microscopy and machine learning for expert-level malaria field diagnosis. In: Global Humanitarian Technology Conference (GHTC). 2015 IEEE; 2015. pp. 393-399

[13] Muralidharan V, Dong Y, Pan WD. Comparison of feature selection techniques for machine learning based automatic malarial cell recognition in wholeslide images. In: Proc. of IEEE International Conference on Biomedical and Health Informatics. Las Vegas, NV; 2016

[14] Park HS, Rinehart MT, Walzer KA, Chi JTA, Wax A. Automated detection of P. falciparum using machine learning algorithms with quantitative phase images of unstained cells. PLoS One. 2016;**11**(9):e0163045

[15] Sanchez CS. Deep Learning for Identifying Malaria Parasites in Images. MSc Thesis. Edinburgh, UK: University of Edinburgh; 2015

[16] Quinn JA, Nakasi R, Mugagga PKB, Byanyima P, Lubega W, Andama A. Deep convolutional neural networks for microscopy-based point of care diagnostics. CoRR. 2016; abs/1608.02989. Available from: http://arxiv.org/abs/1608.02989

[17] Muralidharan V, Dong Y, Pan WD. A comparison of feature selection methods for machine learning based automatic malarial cell recognition in wholeslide images. In: 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); 2016. pp. 216-219

[18] Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of deep learning in biomedicine. Molecular Pharmaceutics. 2016;**13**(5):1445-1454. PMID: 27007977. Available from:. DOI: 10.1021/acs.molpharmaceut.5b00982

[19] Leung MKK, Delong A, Alipanahi B, Frey BJ. Machine learning in genomic medicine: A review of computational problems and data sets. Proceedings of the IEEE. 2016 Jan;**104**(1): 176-197

[20] Min S, Lee B, Yoon S. Deep learning in bioinformatics. Briefings in Bioinformatics. 2017 Sep 1;**18**(5):851-869

[21] Greenspan H, van Ginneken B, Summers RM. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging. 2016 May;**35**(5):1153–1159

[22] Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging. 2016 May;**35**(5):1285-1298

[23] Lai M. Deep learning for medical image segmentation. CoRR. 2015;abs/1505.02000. Available from: http://arxiv.org/abs/1505.02000

[24] Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging. 2016 May;**35**(5):1299-1312

[25] Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE Transactions on Medical Imaging. 2016 May;**35**(5):1207-1216

[26] Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N. EndoNet: A deep architecture for recognition tasks on laparoscopic videos. IEEE Transactions on Medical Imaging. 2016;PP(99):1-1

[27] Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S, Navab N. AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images. IEEE Transactions on Medical Imaging. 2016 May;**35**(5):1313-1321

[28] Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. Arxiv. 2016

[29] Hubel DH, Wiesel TN. Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology. 1968;**195**(1):215-243

[30] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proceedings of the IEEE. 1998;**86**(11):2278-2324

[31] Hauberg S, Freifeld O, Larsen ABL, Fisher J, Hansen L. Dreaming more data: Class-dependent distributions over diffeomorphisms for learned data augmentation. In: Proceedings of the 19th International Conference on Artificial Intelligence and Statistics (AISTATS), Cadiz, Spain; 2016. pp. 342-350

[32] Wong SC, Gatt A, Stamatescu V, McDonnell MD. Understanding data augmentation for classification: When to warp? In: 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE; 2016. pp. 1-6

[33] DeVries T, Taylor GW. Dataset Augmentation in Feature Space. arXiv preprint arXiv: 170205538; 2017

[34] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science. 2006 Jul 28;**313**(5786):504-507

[35] Goodfellow I, Bengio Y, Courville A. Deep Learning. Cambridge, Massachusetts, USA: MIT Press; 2016