Computer Methods and Programs in Biomedicine xxx (xxxx) xxx

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

A transfer learning method with deep residual network for pediatric pneumonia diagnosis

Gaobo Liang, Lixin Zheng*

Fujian Provincial Academic Engineering Research Centre in Industrial Intellectual Techniques and Systems, College of Engineering, Huaqiao University, Quanzhou, China

ARTICLE INFO

Article history:
Received 30 April 2019
Revised 14 June 2019
Accepted 25 June 2019
Available online xxx

Keywords:
Pneumonia
Deep learning
Residual network
Image classification
Transfer Learning

ABSTRACT

Background and Objective: Computer aided diagnosis systems based on deep learning and medical imaging is increasingly becoming research hotspots. At the moment, the classical convolutional neural network generates classification results by hierarchically abstracting the original image. These abstract features are less sensitive to the position and orientation of the object, and this lack of spatial information limits the further improvement of image classification accuracy. Therefore, how to develop a suitable neural network framework and training strategy in practical clinical applications to avoid this problem is a topic that researchers need to continue to explore.

Methods: We propose a deep learning framework that combines residual thought and dilated convolution to diagnose and detect childhood pneumonia. Specifically, based on an understanding of the nature of the child pneumonia image classification task, the proposed method uses the residual structure to overcome the over-fitting and the degradation problems of the depth model, and utilizes dilated convolution to overcome the problem of loss of feature space information caused by the increment in depth of the model. Furthermore, in order to overcome the problem of difficulty in training model due to insufficient data and the negative impact of the introduction of structured noise on the performance of the model, we use the model parameters learned on large-scale datasets in the same field to initialize our model through transfer learning.

Results: Our proposed method has been evaluated for extracting texture features associated with pneumonia and for accurately identifying the performance of areas of the image that best indicate pneumonia. The experimental results of the test dataset show that the recall rate of the method on children pneumonia classification task is 96.7%, and the f1-score is 92.7%. Compared with the prior art methods, this approach can effectively solve the problem of low image resolution and partial occlusion of the inflammatory area in children chest X-ray images.

Conclusions: The novel framework focuses on the application of advanced classification that directly performs lesion characterization, and has high reliability in the classification task of children pneumonia.

1. Introduction

According to the World Health Organization (WHO), pneumonia is still the leading cause of death among children under five years of age in many infectious diseases. In 2016, more than 800,000 children died of pneumonia, most of who were no more than 2 years old, and the death toll was more than the sum of malaria, AIDS and measles [1,2]. It is noteworthy that more than 90% of newly diagnosed children with clinical pneumonia occur in developing countries where medical resources are scarce. Therefore, in order to save more children, accurate and timely diagnosis of pneumonia is imperative. Radiological diagnostic techniques for lung disease include chest X-ray, computed tomography (CT), and magnetic resonance imaging (MRI). As a painless and non-invasive method of examination, chest X-ray is one of the most common radiological tests used to screen and diagnose many lung diseases [3]. However, the current detection of pneumonia in chest X-rays is still largely dependent on the diagnostic level of the radiologist, and the reliability of the diagnostic results is still facing enormous challenges. This problem is particularly acute in areas where the incidence and mortality of children with pneumonia are much higher than the world average, such as the African and the South Asian subcontinent where medical resources are limited. On the other hand, the identification of chest X-ray images is primarily

* Corresponding Author: No. 269, Chenghua North Road, Fengze District, Quanzhou, Fujian Province, 362021, China.
E-mail address: zlx@hqu.edu.cn (L. Zheng).

https://doi.org/10.1016/j.cmpb.2019.06.023
0169-2607/

performed by highly trained and experienced physicians. Currently, in the field of medical image analysis, only a small number of systems can automatically identify the organs or tissues of the human body [4–6]. However, the implementation of these artificial intelligence (AI) systems still faces the challenge of reliability and interpretability [7]. The automatic classification and recognition of organs or lesions based on X-ray images has great application prospects in educational medicine, and the number of cases analyzed by students can be significantly increased using this method. In addition, it can promote the development of telemedicine training and save valuable human resources. Note that automatic classification can help mark and manage medical databases in different hospitals [8], so that the existing image data can be fully utilized and explored, and the development of AI systems in the field of medical imaging diagnosis is promoted. Therefore, this automated classification system can significantly alleviate problems caused by manual classification (e.g. subjectivity, economic cost, time, human resources). It can be seen that the development of auto-assisted diagnostic tools based on chest X-rays can improve the efficiency of radiologists, reduce medical costs, and help accelerate the diagnosis and treatment of pneumonia in children.

As an important branch of the field of AI, deep learning uses the raw data as the algorithm input, and the layered abstraction of the algorithm abstracts the original data layer by layer into the final feature representation required by its own task. Finally, with the learned features are mapped to the task target as the end, and the entire process does not require manual operations [9]. The development of deep learning is inseparable from the study of the cognitive principles of the brain. Deep learning is a branch of artificial neural network (ANN). Now, the ANN having a deep network structure is the earliest network model for deep learning [10]. Since the emergence of the concept of ANNs in 1943, researchers have begun to use mathematical models to theoretically model neurons in ANNs, and thus opening up the study of AI [11]. As the earliest and simplest ANN model, the perceptron network uses the Hebb learning rule or the least squares method to train the parameters [12]. Inspired by this unique structure, machine vision experts have proposed the earliest deep learning network model, which replaces the single layer structure of the perceptron by the depth structure of multiple hidden layers, and the multilayer perceptron model is one of the most representative [13]. Based on the above research, AI experts proposed convolutional neural network (CNN) which is one of the commonly used models for deep learning, and this model uses a back propagation algorithm to train [14]. However, these deep network models have long been difficult to train until the deep belief network and the subsequent efficient semi-supervised algorithm used to training network parameters bring the possibility of breaking the deadlock [15].

Encouraged by the amazing performance of deep learning algorithms on data analysis competition platforms such as the ImageNet competition and the Kaggle competition, researchers began to apply deep learning algorithms to the field of medical image analysis [16,17]. A Google research team successfully used machine learning to detect diabetic retinopathy, and the results showed that their algorithms performed as well as ophthalmologists [18]. Stanford's research team has proposed a new technology called CheXNet, whose deep learning algorithms have surpassed human physicians in identifying the accuracy of pneumonia and other diseases in chest radiographs [19]. Although recent advances in deep learning have shown that deep neural networks can successfully classify patient data and achieve better results than using non-data-driven methods. These convolutional networks for image classification will gradually reduce the resolution of the feature map, which will seriously damage the spatial structure of the scene, and the loss of spatial acuity will limit the further improvement of the image classification performance of the model [20]. Specially, under normal circumstances, the dose of chest X-ray radiation received by children is relatively low, so the shortcomings of low spatial resolution and overlapping of anatomical parts of X-ray chest radiographs will be more prominent. Due to the loss of spatial information, the typical CNN models are difficult to further improve performance on children chest X-ray dataset. For that reason, we propose a novel technique to effectively alleviate these problems in order to automatically classify children's chest X-ray images, identify and diagnose pneumonia.

Specifically, X-ray images of children may be collected by different types of equipment and different operators in clinical practice. In addition, for some reason, it is difficult for some patients to obtain standard images in an ideal posture. These factors can cause the quality of the captured images to vary widely, posing challenges to the practical application of intelligent diagnostic tools. Therefore, in order to improve the generalization performance of the proposed model, we perform scale transformation, flipping and rotation operations on the image to expand the dataset. Then, we use the residual neural network [21] as the structural basis of the proposed model to increase the depth of the model, and introduce the dilated convolution [22] to solve the problem of spatial information loss caused by the increase of depth. Finally, in order to shorten the time of neural network training and accelerate convergence, we introduce a transfer learning method. The novelty of our approach is that we propose a more general automatic diagnostic algorithm that takes full advantage of the information contained in the input data. By using the augmented dataset to train the proposed model, we have achieved better classification performance.

The rest of this article is organized as follows. The related work is detailed in Section 2. The method proposed in this study is described in Section 3. In Section 4, we describe the results obtained. Based on the experimental results, we have carefully analyzed and discussed the proposed method in Section 5. Finally, the main conclusions of this work and the outlook for future work can be found in the last section.

2. Related work

The breakthrough progress of deep learning and large datasets has enabled algorithms to achieve a wide range of applications and stunning effects in radiology practice and research tasks, which include chest pathology [23], cardiovascular disease diagnosis [24], and cancer detection [25,26]. Among them, X-ray image analysis is an important application field of deep learning, which is a noteworthy fact. In order to explore the adaptability problem of the deep features of automatic learning compared to the visual features determined by domain experts in clinical practice. Some scholars have proposed and tested several deep learning methods on public X-ray scan datasets so as to automatically assess the bone age of patients from different races, age ranges and genders [27]. Their research has shown that deep learning does not require feature processing of input data compared to traditional supervised pattern recognition methods (such as SVM) , and model performance is not limited by the medical diagnostic experience owned by the designer, thus having an unparalleled advantage.

Besides human bone research, computer experts are also beginning to explore the feasibility of AI systems in medical image analysis. Investigators have proposed a method for screening for respiratory diseases [28]. Diagnosis can be achieved by designing a combination of two consecutive CNNs and training in a low-dose chest CT dataset. Specifically, the first neural network achieves efficient use of spatial information of learned features by using dilated convolution, and identifies and marks potential diseased region. The true diseased regions in the detected candidates are then screened out in the second network. Unlike other methods of performing tumor detection tasks that first perform detection

Please cite this article as: G. Liang and L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2019.06.023

on sensitive areas in the image, the method proposed in [28] can search for regions of interest directly in the entire image without using any pre-processing steps. Experimental results show that the neural network does not need to manually extract and input explicit spatial features. The spatial background in the three orthogonal 2D patches can be successfully identified by the CNN with dilated convolution characteristics.

With the dramatic increase in the size of the public chest X-ray dataset [3], scholars have begun to research into automated diagnostic algorithms for various chest diseases. To alleviate the contradiction between a limited number of highly skilled radiologists and the growing need for chest X-ray interpretation, researchers have developed a model called CheXNeXt based on the DenseNet 121 framework [29]. The framework can simultaneously detect the presence of 14 different lesions such as lung prostate cancer, pleural effusion, lung mass and nodules in the frontal lobe. Specifically, CheXNeXt was first trained and parameterized on open source datasets. Then, in order to verify the reliability of the intelligent system, the pathological examination results of CheXNeXt on the verification dataset were compared with the clinical diagnosis results of 9 radiologists. The study found that CheXNeXt has reached the diagnostic level of human physicians in 11 pathologies, and further proves the great potential of deep learning in the field of smart medical care in the future. In addition, deep learning algorithms for tuberculosis classification and pulmonary nodule detection have emerged in the field of chest disease diagnosis [23,30].

Inspired by the above research work, we proposed an automated method for the diagnosis of childhood pneumonia. The method uses the residual structure to deepen the model and realizes the utilization of the feature map spatial information by dilated convolution. The specific details will be elaborated in the methods section.

3. Methods

Before introducing the proposed deep learning algorithm, we first specifically describe the dataset that is used in this study. By interpreting the input data and the expected output before the actual method, it helps to better understand the method that is proposed in our paper.

3.1. Data

The dataset Kermany et al. [31] used to train and evaluate the proposed method was produced based on an X-ray scan database from pediatric patients from one to five years of age at the Guangzhou Women and Children's Medical Center [7]. The dataset contains a total of 5856 chest X-ray images collected and labeled from children. Of these, 5232 patients received AI system training (using 3883 pneumonia images, and 1349 normal images). After the training was completed, the performance of the proposed model was tested with 234 normal images and 390 pneumonia images from 624 patients. The specific X-ray scan image is shown in Fig. 1. Actually, the quality of captured images is different due to the scanning device, the operator's work habits, as well as problems caused by position and other inherent health issues. Given the distribution of the various samples in the dataset and the adaptability of the algorithm, data is added prior to formal training to maximize the number of training samples while balancing the distribution of samples in each category. In image classification tasks, data augment is often employed to increase the amount of training data and prevent over-fitting. To this end, we use the flip and rotate operations to extend the dataset to accommodate various non-standard images in clinical practice.

3.2. Proposed CNN

In order to determine the optimal architecture and configuration of the proposed model, it is first necessary to understand the nature of the pneumonia image classification task under consideration. The commonly used image features include target color, shape, texture, and spatial relationship. The main difference in pneumonia type is the lung texture feature; therefore, researchers generally choose local texture features to characterize the image. Although texture is an intuitive and simple visual concept, many papers do not currently have a general definition of the available textures, and the discussion of their definitions is far from over [32]. In addition, unlike image features such as grayscale and color, textures are represented by the grayscale distribution of the neighborhood of the pixel and its surrounding space. Here, we can define the texture as a constant repetition of the local space in the global region of the image. The convolution kernel of CNN is also referred to as a feature extractor. In the convolutional layer that is composed of filter stacks, the input of each neuron is connected to the local accepting domain of the previous layer, and the local features are extracted. Through the form of local connections, the framework has been successfully applied to many texture analysis tasks [33]. This encourages the use of CNN to identify effective texture features by modifying the convolutional

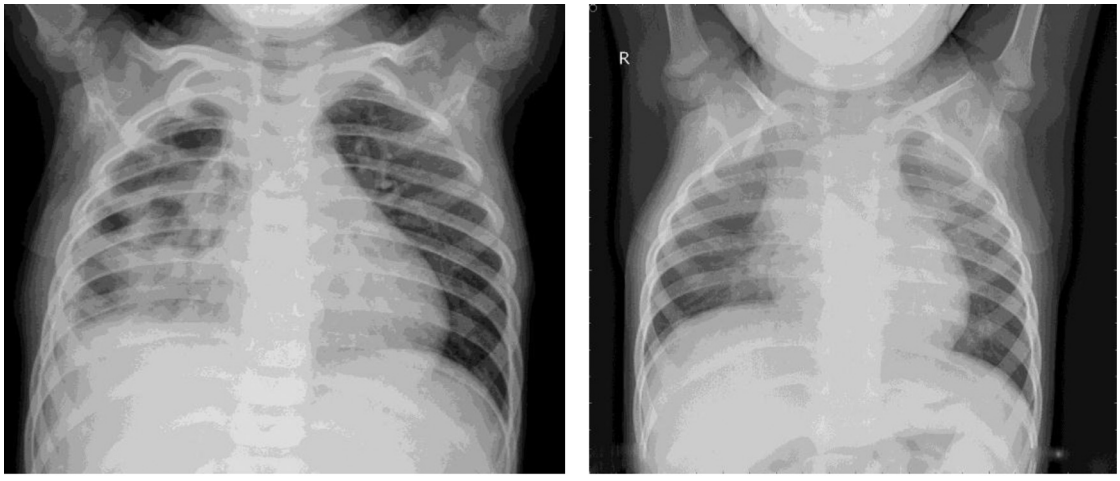


Fig. 1. Illustrative example of a child's chest X-rays used in this study whereby (a) is an image of pneumonia, and (b) is a normal chest radiograph depicting a clear lung with no abnormal turbid areas in the image.

Please cite this article as: G. Liang and L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2019.06.023

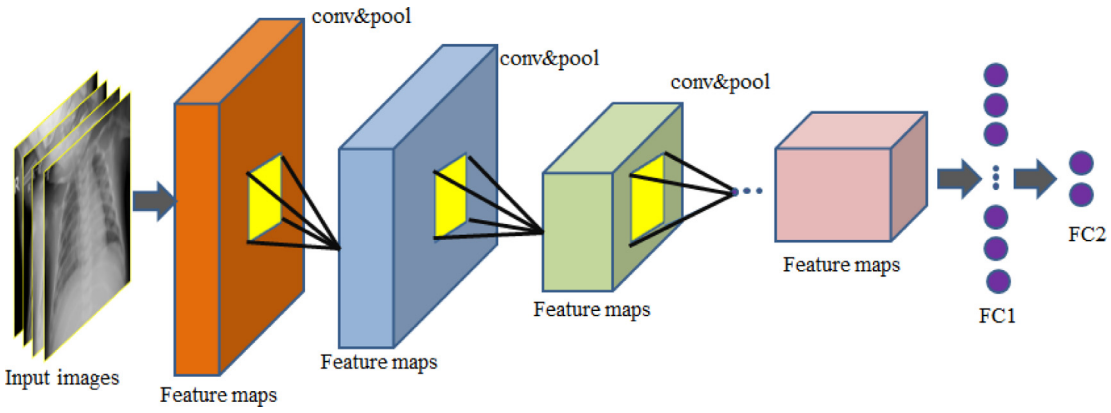


Fig. 2. Architecture of the proposed CNN with 49 convolutional layers and 2 dense layers for classification of children's lung patterns.

layer. Nevertheless, we should also note that the texture does not fully reflect the basic properties of the object, so it is not possible to use only texture features to obtain a higher level of image content. As the depth of the model increases, the resolution of the feature map gradually decreases, and the texture extracted by the convolution kernel may have a large deviation.

Architecture: After fully considering the above principles, we designed the network model shown in Fig. 2. The proposed method is based on a residual network. The residual is the difference between the actual observed value and the estimated value in mathematical statistics. By introducing the residual thought to remove the same part, thus highlighting minor changes, let our model focus on learning these tiny changes. Studies have shown that this method solves the problem that the fitting effect becomes worse as the number of layers of the neural network increases. The input of the network is adjusted to a 150×150 image block after preprocessing, and is convoluted by a series of 49 convolutional layers. Then, we select a filter bank of size 1×1 and 3×3 in the convolutional layer. Here, we set a convolution kernel smaller than 5×5 and 7×7 and choose the ReLU function for more nonlinear activation. In this way, we can ensure that each neuron is small enough relative to the total receptive field of the input to capture local texture features associated with the desired result. In this case, although the previous convolutional network performed well, the entire network could not achieve a very accurate result. Some of the information that is strongly related to the target is directly ignored by the previous convolutional network because of excessive dimensionality reduction and its small size. To this end, we introduce a dilated convolution in 49 convolutional layers to ensure that the resolution of the original network is constant and the loss of resolution of the image space is minimized. The principle of dilated convolution is shown in Fig. 3. From this, we can see that the dilated convolution supports increasing the receptive field of the convolution kernel without increasing the kernel parameters and avoiding excessive loss of resolution of the feature map. At the same time, the batch normalization layer is added after each convolution layer so that the input value of the nonlinear transformation function falls into the region sensitive to the input, thereby avoiding the problem of gradient disappearance caused by structural complexity and speeding up the training. This is followed by the global average pooling layer, which calculates the average of each feature map for the last convolutional layer output. The resulting feature is equal to the number of feature maps of the last layer and is fed to 2 fully connected layers. The first dense layer has 128 nodes, the second is the classification layer, and there are 2 nodes. In addition, to prevent over-fitting of the model, we add a dropout layer before the global mean pooling layer, with the parameter set to 0.5. It randomly stops half of the neurons from training each time the training is updated, and preventing hidden units from being dependent on specific inputs [34].

Activations: The activation function plays an important role in the neural network model to learn and understand very complex and nonlinear tasks. As we know, the activation function introduces nonlinear factors to the neurons, allowing the neural network to arbitrarily approximate any nonlinear function, and the choice of the activation function can significantly affect the rate of convergence. Compared to the classic Sigmoid and Tanh alternatives, the use of the ReLU function has been shown to reduce computational costs and prevent gradient disappearance [35,36]. It is particularly noteworthy that ReLU causes the output of a subset of some neurons to be zero, which increases sparsity in the network and reduces the interdependence of parameters, thus alleviating the problem of over-fitting. In this study, we recommend using the ReLU function to activate each convolutional layer. For the fully connected layer portion of the network, standard ReLU activation is used for the first dense layer, and Sigmoid is used for the last layer to compress the 2-dimensional output into a classification probability distribution.

Transfer Learning: In traditional classification learning, the accuracy and reliability of the classification model can be guaranteed only when there are enough training data and test data from the same feature space and unified distribution, which means that the model will be retrained each time the data changes. However, in practical applications, we have found that the previously available labeled samples may be inconsistent with the distribution of new test samples for some reason. On the other hand, due to problems such as patient privacy, it is often difficult to obtain large-scale labeled medical image data. As a new method to solve different but related problems using existing knowledge, transfer learning has attracted extensive attention and research. Moreover, studies have shown that introducing random noise through transfer learning during training can improve model performance. However, the introduction of structured noise has the opposite effect, which destroys the association between the tag and the object, and degrades performance. In particular, this problem can be more serious when the noise is the same as the actual data source. The model will try to learn these completely different information that is not related to the task, which will eventually lead to confusion between the noise and the category [37]. Therefore, unlike the usual methods of pre-training on ImageNet datasets, we designed the network to be pre-trained on the recently released ChestX-ray14 dataset [3]. The dataset contains 112,120 frontal chest X-ray images labeled with up to 14 different chest diseases. The trained network weights are then saved and transferred to the children's pneumonia image classification task. We found in experiments that this transfer learning method can avoid the introduction of

Please cite this article as: G. Liang and L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2019.06.023

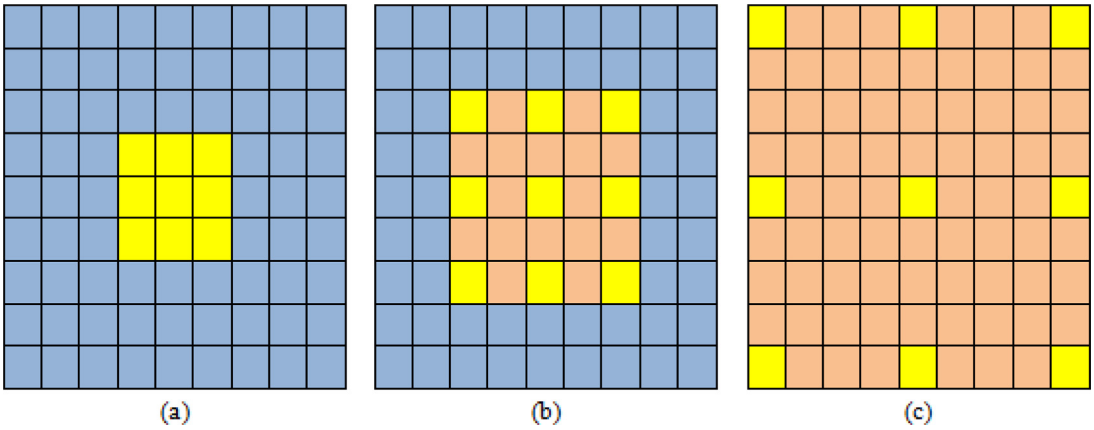


Fig. 3. Separate settings for three different dilated rates corresponding to: (a) 3×3 convolution kernel with dilated rate of 1, which is the same as a normal convolution operation, (b) 3×3 convolution kernel with dilated rate of 2, the receptive field is equivalent to a 5×5 normal convolution kernel, and (c) 3×3 convolution kernel with dilated rate of 4, the receptive field is equivalent to a 9×9 normal convolution kernel.

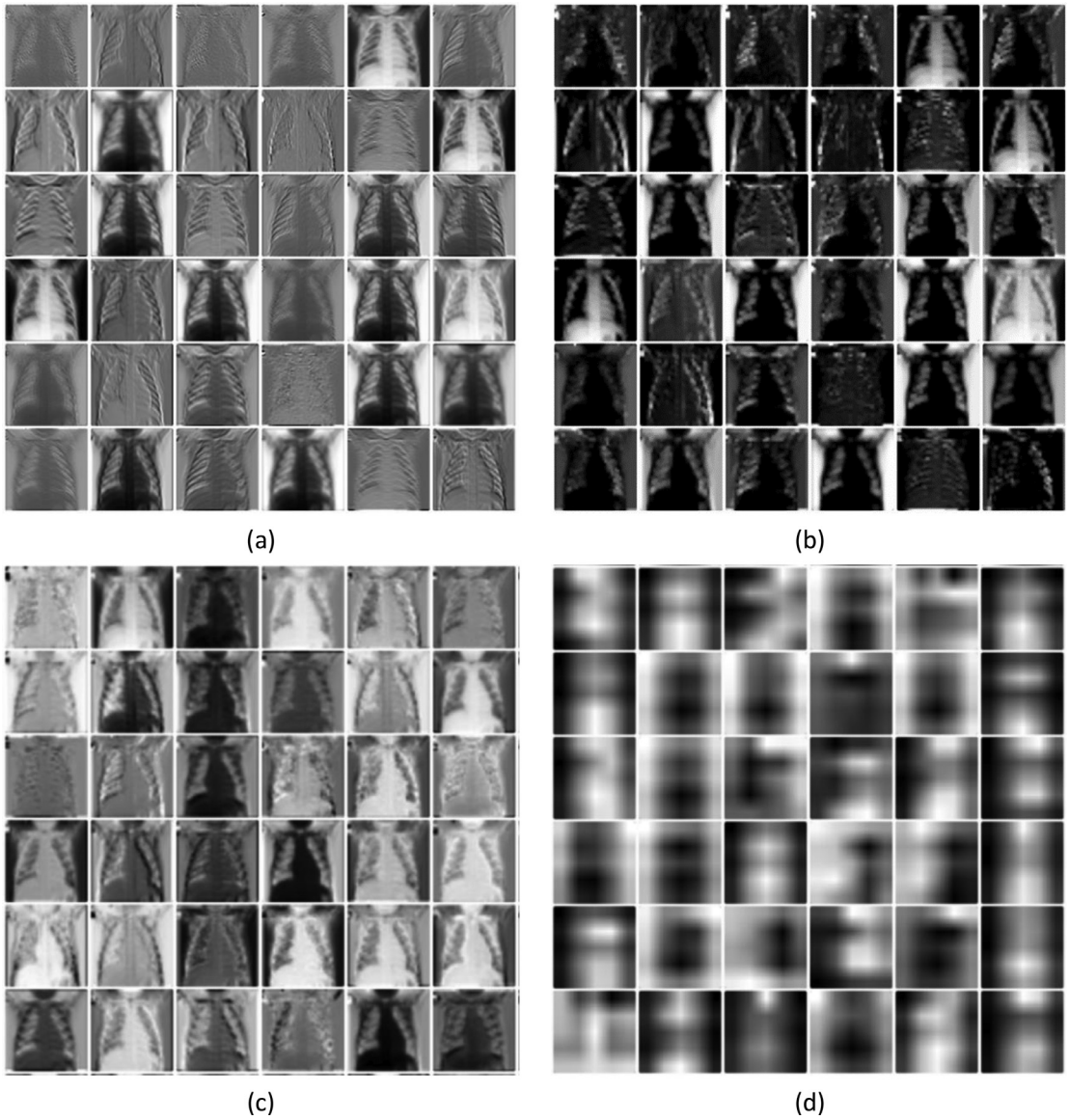


Fig. 4. Partial visualization results of the feature map in the best model of the proposed method: (a) the convolution layer, (b) the max-pooling layer, (c) the added layer of the first bottleneck design module, (d) the last adding layer.

Please cite this article as: G. Liang and L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2019.06.023

structured noise and is effective in accelerating model convergence and improving classification performance.

Training Method: In the training of neural networks, the most important is to define the loss function and choose the algorithm for backpropagation optimization. For this study, we used the Adam optimizer [38] to minimize the cross entropy loss function. The Adam optimization algorithm is an extension of the stochastic gradient descent algorithm, which iteratively updates the neural network weights based on the training data. Unlike traditional stochastic gradient descent algorithms, Adam designs independent adaptive learning rates for different parameters by calculating the first-order moment estimates and second-order moment estimates of the gradient. The value of the cross entropy loss function represents the distance between the actual output and the expected output, and the magnitude of the value directly reflects the proximity of their probability distribution. For the parameter settings of the Adam optimizer, we have retained the parameters as default values after relevant experimental tests. In the initialization phase of the model, we transferred the pre-trained weight parameters on the ChestX-ray14 dataset to the current neural network model for performing the child pneumonia classification task. In addition, we perform weight update with small batch data, and the batch size is set to 64. When the accuracy of the model on the verification set no longer increases, the learning rate is reduced by 10 times, and if the model performance improvement is not seen in the 6 epoch, the learning rate is also reduced. The training is over when the network is continuously training 100 epochs.

4. Experiment and results

In this section we describe the selected evaluation strategy and some details about the implementation of the method, and introduce the experimental results of this study.

4.1. Experimental setup

Evaluation: The evaluation of the Child chest x-ray image classification approaches is based on a train-validation-test scheme. We train our model on the training dataset, fine tune the hyperparameters through the performance of the network on the validation set, and then use the test set to evaluate the overall performance of the proposed method. We use three metrics based on the confusion matrix to evaluate the response of the model in the two-category task: precision, recall, f1-score, and calculate the accuracy of the overall classification. In addition, we also plotted the ROC curve and calculated the corresponding AUC value. It should be noted that in order to further evaluate the classification effect of the proposed model in this work, we compared the five most commonly used CNN architectures for medical image classification - VGG16 [39], DenseNet121 [40], Xception [41], and InceptionV3 [42].

Implementation: The deep learning algorithms that are run in this experiment, including the proposed method, are implemented using the Keras framework, and the methods that do not involve convolutional networks are coded in Python. All experiments were performed on a Linux operating system on a workstation with GPU NVIDIA GeForce GTX 1080Ti and 32 GB RAM.

4.2. Test results

Here, we visualize the proposed model to improve our understanding of its performance. Fig. 4 shows the visualization of the partial filter output feature maps of the first convolutional layer of the best model. We will notice that the convolutional layer has edge detection capabilities, and their different properties capture the basic edge modes. In addition to the visualization of the convolutional layer, Fig. 4 also shows a partial output feature map of the max-pooling layer in the visualization network. Next, we show the visualization results of the partial output feature maps of the added layer of the first bottleneck design module and the last bottleneck design module respectively.

Table 1
Comparison of proposed methods with other cnn models.

Model	Accuracy	Precision	Recall	F1-score	AUC
VGG16	0.742	0.723	0.951	0.822	0.840
DenseNet121	0.819	0.792	0.964	0.869	0.769
InceptionV3	0.853	0.916	0.841	0.877	0.655
Xception	0.878	0.857	0.967	0.908	0.930
Proposed Method	**0.905**	**0.891**	**0.967**	**0.927**	**0.953**

Table 1 provides a comparison of our approach to some typical CNNs, including VGG16, DenseNet121, Xception, and InceptionV3. To the best of our knowledge, the VGG16 and DenseNet121 are designed for 224×224 color image classifications, while Xception and InceptionV3 are designed for 299×299 color image classification. It should be emphasized again that our proposed deep learning algorithm needs to adjust the size of the original input image to a size of 150×150. At the same time, in order to overcome the problem of insufficient data, we fine-tuned the deep network that has been trained on ImageNet, and then tested these models with pre-training weights with test sets. According to the test results shown in Table 1, we can find that in the dataset used in this study, the proposed method achieves the highest classification accuracy and f1-score. Furthermore, we also show the confusion matrix of the method proposed in Fig. 5, and in order to highlight the advantages of the algorithm, the confusion matrix of the other CNNs for comparison is given in Fig. 6.

The number of positive and negative samples of the validation set used to test model performance is not balanced. Therefore, to more fully assess the performance of the proposed method, we plot the receiver operating characteristic (ROC) curve of our method for the diagnosis of childhood pneumonia in Fig. 7, and

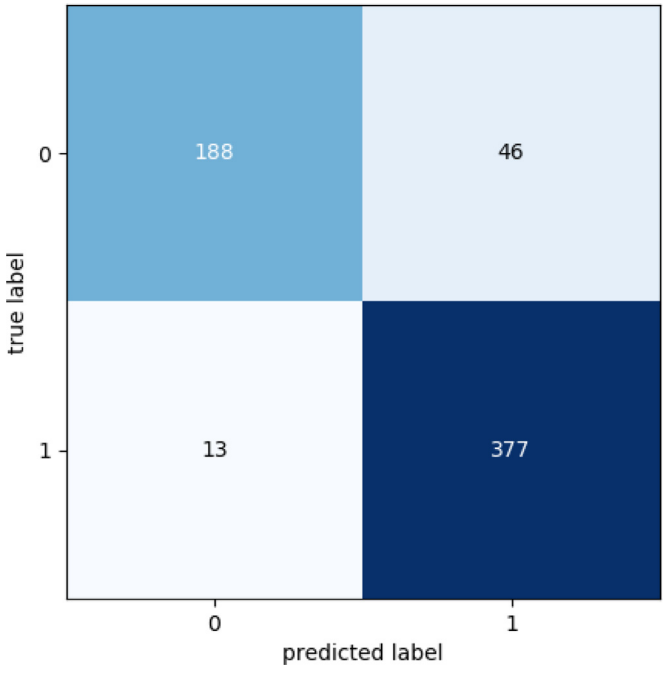


Fig. 5. Confusion matrix of the best model obtained through training and parameter adjustment.

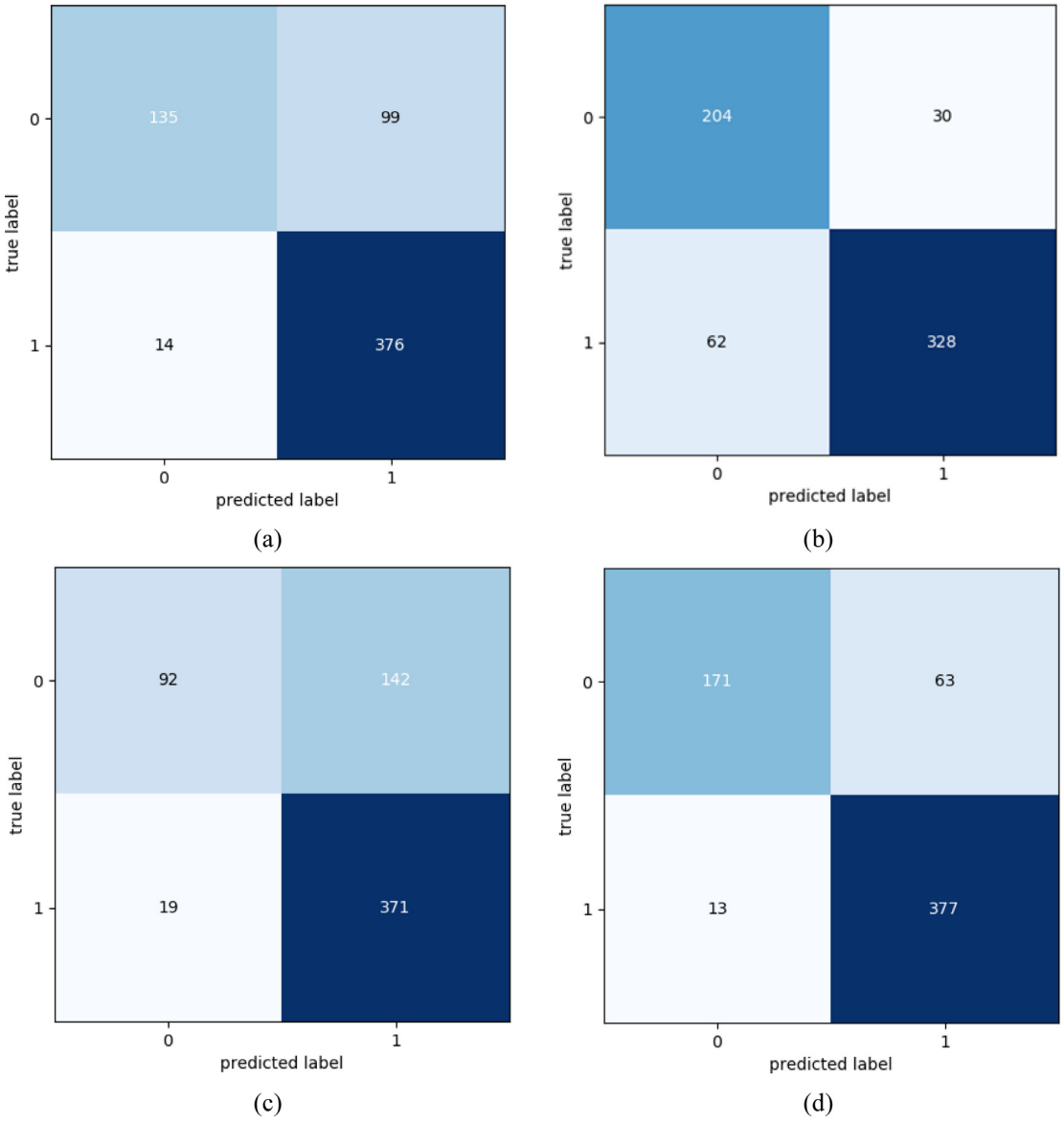


Fig. 6. Confusion matrices based on: (a) DenseNet121, (b) InceptionV3, (c) VGG16, and (d) Xception.

calculate the area under the curve (AUC). Furthermore, we also evaluated the detection performance of the proposed model for pneumonia lesions by a heat-map as shown in Fig. 8.

Based on the above experimental results, we find that the designed model can learn the texture features related to pneumonia, and accurately locate the image area that best indicates pneumonia. Compared with other CNN models, our method has the best classification performance in the field of X-ray image based children's pneumonia diagnosis, having and AUC of 95.3%.

5. Discussion

In this section, we will systematically analyze the proposed method compared with other CNN based on the experimental results in the previous section, and discuss the advantages and possible problems of the method in the field of childhood pneumonia diagnosis.

5.1. Comparison with the prior art

According to the results shown in Table 1 and Fig. 7, the classification accuracy of the method is 0.905, the f1 score is 0.927, and the AUC value is 0.953. Compared with VGG16, DenseNet121, Xception, and InceptionV3, our method has the best classification performance in this dataset. Based on the results presented in Figs. 5 and 6, and in conjunction with the analysis of the nature of the pneumonia image classification task, we provide the reasons as to why the prior art exhibits poor performance. Firstly, the four typical CNNs that are used for comparison are deep networks, and the spatial resolution of the feature map of their last convolutional layer output have greatly diminished, which will limit the classification performance of the model. Secondly, these network filter size designs are not suitable for this task, resulting in the neuron's total receptive field relative to the input being too large, and some key features are therefore directly ignored. In addition, other algorithm choices will also affect the performance of the model,

Please cite this article as: G. Liang and L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2019.06.023

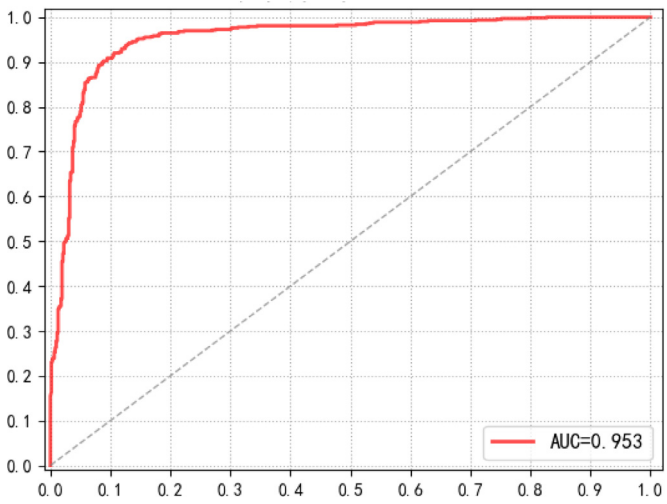


Fig. 7. ROC curve used to evaluate the performance of the proposed method that is drawn from two variables: false positive rate (horizontal axis) and true positive rate (vertical axis).

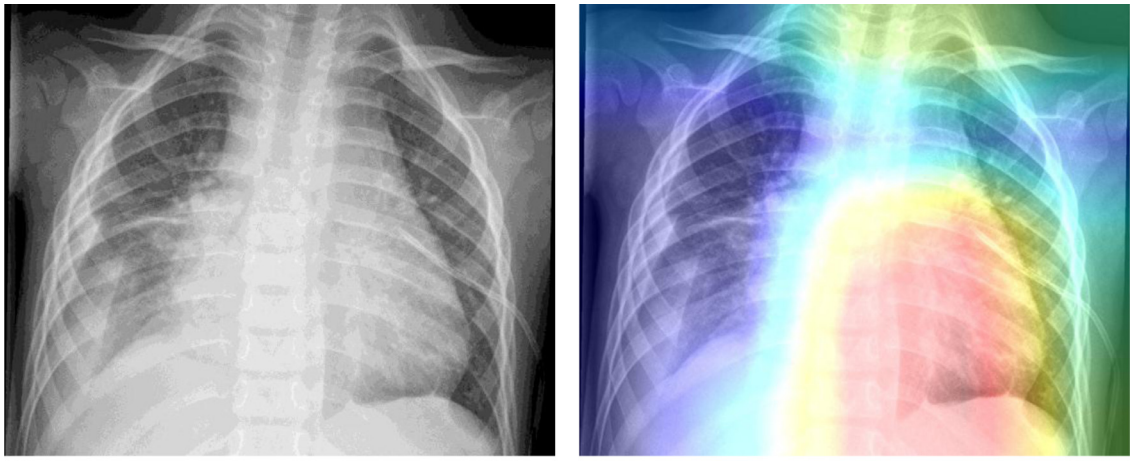
such as data augment mode, selection of activation function and loss function, learning rate and other hyper-parameters settings. In summary, our deep learning model has the following contributions compared to these existing methods: We propose an end-to-end neural network model for the classification of pediatric pneumonia images in combination with the residual thought and dilated convolution. The model effectively solves the problem of low resolution, partial occlusion or overlap in the inflammatory region of children chest X-ray images. In addition, we use the transfer learning method to initialize the model through weight parameters learned on large data sets in the same research fields. This method accelerates the convergence of the model while effectively avoiding the negative impact of the introduction of structured noise on the performance of our model and further improving the performance.

5.2. Performance analysis of the proposed method

We provide a further analysis of the performance of the proposed method. Although only partial visualization results are shown in the experimental results due to the large number of feature maps, it is still noted that the convolutional layer of the network can learn the edge mode of the image. Moreover, these patterns are continuously abstracted as they pass through successive convolutional layers until the microstructure that characterizes the lung texture is described in the last layer. The heat-map of Fig. 8 further demonstrates that our method can learn features that are highly correlated with the class and accurately detect areas of inflammation in the lungs. In practice, based on the understanding of the original input data, we use the residual network as the structural basis, combined with the dilated convolution design to propose an automated diagnosis algorithm for children with pneumonia. This method can maintain a certain resolution of the picture space, and achieves classification performance beyond the prior art through migration learning and hyper-parameter adjustment. However, despite the good performance of the proposed method, it still has problems that cannot be ignored, and should be considered when interpreting the results. Our model has a deeper convolutional layer, and this relatively large network has a problem of low prediction efficiency in practical applications. According to related research, slower predictions increase the actual operating time by at least 7 times [33]. This suggests that for true clinical practice, we need to further optimize the network structure and reduce the scale of the parameters. In addition, the dataset evaluated in this study comes from a single organization, so it is necessary to continue to study the generalization performance of the method based on data from different institutions.

6. Conclusion

In this paper, we present an automated diagnostic tool that classifies children's chest X-ray images into normal and pneumonia. In order to learn the effective texture characteristics of lung tissue, we designed novel network architecture with residual structures. The network consists of 49 convolutional layers and the ReLU activation, followed by only one global average pooling layer and two dense layers. Training is performed by introducing dilated convolutions and using the Adam optimizer to minimize the cross entropy loss function. The proposed method is able to avoid the loss of feature space information caused thereby while maintaining the depth of the model. In addition, we use transfer learning to accelerate neural network training and overcome the problem of insufficient data. On the X-ray scan database from the Women and Children's Medical Center, our method demonstrated classification performance superior to the prior art. In future research, we will continue to optimize our approach and plan to apply it to medical



(a) (b)

Fig. 8. Chest X-ray image is input into the proposed network to output a heat-map based on: (a) the chest X-ray scan of the child with pneumonia, and (b) the visualization output.

Please cite this article as: G. Liang and L. Zheng, A transfer learning method with deep residual network for pediatric pneumonia diagnosis, Computer Methods and Programs in Biomedicine, https://doi.org/10.1016/j.cmpb.2019.06.023

image analysis based on CT and MRI. At the same time, the method can also be considered to be extended to economically underdeveloped areas to improve the level of local medical image diagnosis and to promote the development of telemedicine training.

Conflict of interests

The authors declare that there is no conflict of interests regarding the publication of this article.

Acknowledgments

This work was supported in part by the Quanzhou Municipal Bureau of Science and Technology under grant 2017G036, in part by the Xiamen Municipal Bureau of Science and Technology under grant 3502Z20173045. The author would like to thank the Innovation Center of Engineering College of Huaqiao University and the Robotics Research Center for providing high performance computing platforms and servers to run our code.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2019.06.023.

References

[1] I. Rudan, C. Boschi-Pinto, Z. Biloglav, K. Mulholland, H. Campbell, Epidemiology and etiology of childhood pneumonia, Bull. World Health Organ. 86 (2008) 408–416.
[2] R.A. Adegbola, Childhood pneumonia as a global health priority and the strategic interest of the Bill & Melinda gates foundation, Clin. Infect. Dis. 54 (2012) S89–S92.
[3] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, R.M. Summers, Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2097–2106.
[4] M. Haloi, K. R. Rajalakshmi, and P. Walia, Towards Radiologist-Level Accurate Deep Learning System for Pulmonary Screening, 2018, arXiv preprint arXiv:1807.03120.
[5] A. González-Briones, G. Villarrubia, J.F. De Paz, J.M. Corchado, A multi-agent system for the classification of gender and age from images, Comput. Vis. Image Underst. 172 (2018) 98–106.
[6] M.A. Al-masni, M.A. Al-antari, J.-M. Park, G. Gi, T.-Y. Kim, P. Rivera, et al., Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system, Comput Methods Programs Biomed 157 (2018) 85–94.
[7] D.S. Kermany, M. Goldbaum, W. Cai, C.C. Valentim, H. Liang, S.L. Baxter, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172 (2018) 1122–1131 e9.
[8] I. Hernández Aguado, M. Porta Serra, M. Miralles, F. García Benavides, F. Bolúmar, La cuantificación de la variabilidad en las observaciones clínicas, Med. Clin. (Barc.) 95 (1990) 424–429.
[9] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436.
[10] J.E. Dayhoff, J.M. DeLeo, Artificial neural networks: opening the black box, Cancer Interdiscip. Int. J. Am. Cancer Soc. 91 (2001) 1615–1635.
[11] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (1943) 115–133.
[12] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, Psychol. Rev. 65 (1958) 386.
[13] S. Singhal, L. Wu, Training multilayer perceptrons with the extended Kalman algorithm, in: Advances in Neural Information Processing Systems, 1989, pp. 133–140.
[14] Y. LeCun, Y. Bengio, Convolutional networks for images, speech, and time series, Handb. Brain Theory Neural Netw. 3361 (1995) 1995.
[15] G.E. Hinton, S. Osindero, Y.-W. Teh, A fast learning algorithm for deep belief nets, Neural Comput. 18 (2006) 1527–1554.
[16] G. Liang, H. Hong, W. Xie, L. Zheng, Combining convolutional neural network with recursive neural network for blood cell image classification, IEEE Access 6 (2018) 36188–36197.
[17] S. Pang, Z. Yu, M.A. Orgun, A novel end-to-end classifier using domain transferred deep convolutional neural networks for biomedical images, Comput. Methods Programs Biomed. 140 (2017) 283–293.
[18] V. Gulshan, L. Peng, M. Coram, M.C. Stumpe, D. Wu, A. Narayanaswamy, et al., Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs, JAMA 316 (2016) 2402–2410.
[19] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, et al., Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning, 2017, arXiv preprint arXiv:1711.05225.
[20] F. Yu, V. Koltun, T. Funkhouser, Dilated residual networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 472–480.
[21] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[22] F. Yu and V. Koltun, Multi-scale context aggregation by dilated convolutions, 2015, arXiv preprint arXiv:1511.07122.
[23] P. Lakhani, B. Sundaram, Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks, Radiology 284 (2017) 574–582.
[24] J. Wang, H. Ding, F.A. Bidgoli, B. Zhou, C. Iribarren, S. Molloi, et al., Detecting cardiovascular disease from mammograms with deep learning, IEEE Trans. Med. Imaging 36 (2017) 1172–1181.
[25] A.S. Becker, M. Marcon, S. Ghafoor, M.C. Wurnig, T. Frauenfelder, A. Boss, Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer, Invest. Radiol. 52 (2017) 434–440.
[26] Y. Xiao, J. Wu, Z. Lin, X. Zhao, A deep learning-based multi-model ensemble method for cancer prediction, Comput. Methods Programs Biomed. 153 (2018) 1–9.
[27] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, R. Leonardi, Deep learning for automated skeletal bone age assessment in X-ray images, Med Image Anal 36 (2017) 41–51.
[28] N. Lessmann, B. van Ginneken, M. Zreik, P.A. de Jong, B.D. de Vos, M.A. Viergever, et al., Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions, IEEE Trans Med Imaging 37 (2018) 615–625.
[29] P. Rajpurkar, J. Irvin, R.L. Ball, K. Zhu, B. Yang, H. Mehta, et al., Deep learning for chest radiograph diagnosis: a retrospective comparison of the chexnext algorithm to practicing radiologists, PLoS Med. 15 (2018) e1002686.
[30] P. Huang, S. Park, R. Yan, J. Lee, L.C. Chu, C.T. Lin, et al., Added value of computer-aided CT image features for early lung cancer diagnosis with small pulmonary nodules: a matched case-control study, Radiology 286 (2017) 286–295.
[31] D. Kermany, K. Zhang, M. Goldbaum, Labeled optical coherence tomography (OCT) and chest X-Ray images for classification, Mendeley Data v2 (2018).
[32] M. Tuceryan, A.K. Jain, The handbook of pattern recognition and computer Vision, chapter 2.1, Texture Anal. World Sci. Co. (1998) 207–248.
[33] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung diseases using a deep convolutional neural network, IEEE Trans. Med. Imaging 35 (2016) 1207–1216.
[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958.
[35] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[36] 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.
[37] D. Rolnick, A. Veit, S. Belongie, and N. Shavit, Deep learning is robust to massive label noise, 2017, arXiv preprint arXiv:1705.10694.
[38] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, 2014, arXiv preprint arXiv:1412.6980.
[39] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, arXiv preprint arXiv:1409.1556.
[40] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[41] F. Chollet, Xception: deep learning with depthwise separable convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1251–1258.
[42] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.