



# A novel fused convolutional neural network for biomedical image classification

Shuchao Pang<sup>1,2</sup> · Anan Du<sup>3</sup> · Mehmet A. Orgun<sup>2</sup> · Zhezhou Yu<sup>1</sup> 

Received: 1 August 2017 / Accepted: 17 March 2018 / Published online: 12 July 2018  
© International Federation for Medical and Biological Engineering 2018

## Abstract

With the advent of biomedical imaging technology, the number of captured and stored biomedical images is rapidly increasing day by day in hospitals, imaging laboratories and biomedical institutions. Therefore, more robust biomedical image analysis technology is needed to meet the requirement of the diagnosis and classification of various kinds of diseases using biomedical images. However, the current biomedical image classification methods and general non-biomedical image classifiers cannot extract more compact biomedical image features or capture the tiny differences between similar images with different types of diseases from the same category. In this paper, we propose a novel fused convolutional neural network to develop a more accurate and highly efficient classifier for biomedical images, which combines shallow layer features and deep layer features from the proposed deep neural network architecture. In the analysis, it was observed that the shallow layers provided more detailed local features, which could distinguish different diseases in the same category, while the deep layers could convey more high-level semantic information used to classify the diseases among the various categories. A detailed comparison of our approach with traditional classification algorithms and popular deep classifiers across several public biomedical image datasets showed the superior performance of our proposed method for biomedical image classification. In addition, we also evaluated the performance of our method in modality classification of medical images using the ImageCLEFmed dataset.

**Keywords** Biomedical image classification · Convolutional neural networks · Deep learning · Deep feature · Shallow feature

## 1 Introduction

A very large number of biomedical images are generated in hospitals, imaging laboratories and biomedical institutions on a daily basis to assist physicians in the diagnosis of patients' health conditions. Types of biomedical images include those derived from ultrasound (US), magnetic resonance imaging (MRI), computed tomography (CT) and X-ray [1]. In the field of biomedical and medical analyses, images still play a key role in analysing the state of a disease and then in providing an

accurate diagnosis for the patients. However, it is a big challenge to manage and analyse substantial collections of large biomedical images produced every day; it is also extremely tedious and impractical to perform this analysis task solely by human annotation. Therefore, in order to better assist physicians, more robust biomedical image analysis technology is needed to assist medical practitioners to identify and distinguish the exact disease conditions in each patient's report.

### 1.1 Related work

Many biomedical image classification methods and techniques have been proposed to address this problem that primarily focus on using visual cues from the generated images [2–12]. The methods and techniques for the classification of biomedical images can be roughly divided into two kinds, namely, those based on a traditional model and those based on a deep model. The traditional model, as the name implies, combines traditional visual features and typical machine-learning classifiers. Here, such traditional features are named after hand-crafted image features, such as colour, shape,

✉ Zhezhou Yu  
yuzz@jlu.edu.cn

<sup>1</sup> Department of Computational Intelligence, College of Computer Science and Technology, Jilin University, Qianjin Street 2699, Changchun, Jilin Province, China

<sup>2</sup> Department of Computing, Macquarie University, Sydney, NSW 2109, Australia

<sup>3</sup> China Mobile (HangZhou) Information Technology Co., Ltd, Hangzhou, China

structure, texture, the widely used SIFT (scale-invariant feature transform), LBP (local binary pattern) and HOG (histogram of oriented gradient). The extraction of a feature is the first step towards representing each biomedical image, which is akin to an image dimensionality reduction process. If we want to identify and classify an image with its feature representation, it must be with the help of a classifier, which needs to be trained with a corresponding training dataset. Classifiers, which are widely used in biomedical image classification methods, include K-nearest neighbour, multiclass support vector machine (SVM), multiple kernel SVM and their corresponding variants.

In addition, as shown in previous works [2–5], the traditional classification algorithms include two independent parts, namely, a good feature extractor and a robust classifier. For example, Depeursinge et al. [2] proposed a new near-affine-invariant texture feature to extract lung tissue patterns and then added a one-versus-one support vector machine classifier with a Gaussian kernel to train the classification boundaries between any two categories. Song et al. [3] focused on designing a new classifier called the large margin local estimate model and combined multiple features that are rotation-invariant Gabor-local binary patterns (RGLBP), multi-coordinate histograms of oriented gradients (MCHOG) and intensity. Khachane et al. [4] used a texture feature descriptor and proposed a fuzzy rule-based system with 23 rules for classifying multimodal medical images. Ertuğrul et al. [5] presented a novel feature extraction approach, known as adaptive local binary pattern (aLBP), to represent surface electromyography (SEMG) signals, which can obtain a higher classification performance than other popular feature extraction approaches. However, when faced with large collections of imaging data, the traditional classification methods have difficulty in capturing more robust and discriminative biomedical features using only hand-crafted features and classical machine-learning classifiers.

With the further developments of deep neural networks, deep learning technology has found many practical applications [13], particularly in the computer vision area, such as in image classification and recognition [14], object detection [15] and image segmentation [16]. The main reason for the adoption of the deep learning technology is the availability of large annotated image datasets and high-performance computing power with GPUs in recent years. Several recent studies, as typical examples of deep models, have also introduced deep neural networks to biomedical image classification tasks [6–8]. Li et al. [6] proposed a convolutional neural network with a few layers to classify lung image patches, wherein there is only one added convolutional layer and three fully connected layers. Gao et al. [7] designed a specific deep neural network with three convolutional layers, three max-pooling layers and one fully connected layer for the classification of human epithelial-2 (HEp-2) cells with limited biomedical

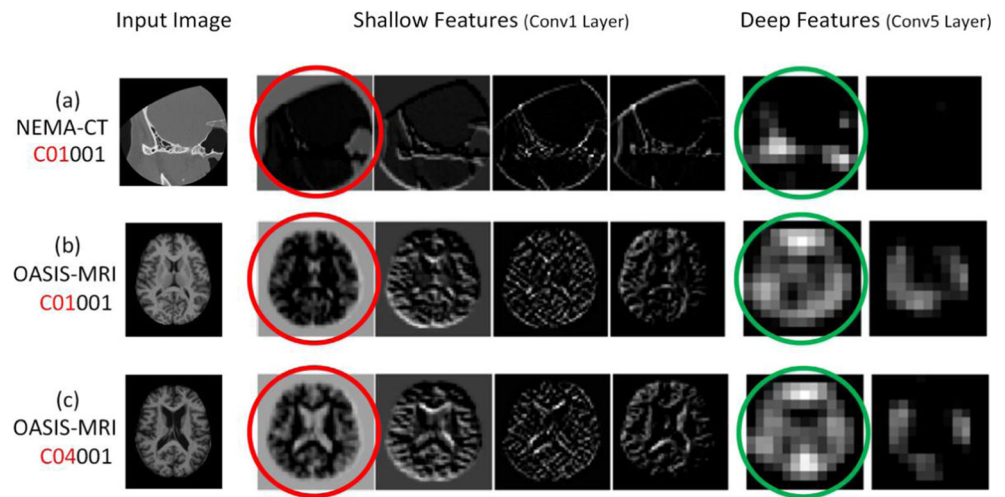
training data. Abbas et al. [8] utilized a semi-supervised multilayer deep learning algorithm to learn deep visual features, which can be used to classify and recognize five severity levels of diabetic retinopathy. In addition, transfer learning technique has also become popular in biomedical image processing by using different pre-trained deep neural networks. For instance, two very recent papers [9, 10] used the domain transfer method to utilize deep neural networks as a kind of a black box for extracting biomedical image features. Ahn et al. [9] proposed the use of a pre-trained deep convolutional network by Oxford's Visual Geometry Group (VGG) [17] as a deep transferred model to extract X-ray image features and then combined a sparse spatial pyramid model to classify X-ray images. Phan et al. [10] applied transfer learning to a pre-trained deep convolutional neural network model to extract general image features and further selected relevant features with the minimum redundancy maximum relevance algorithm to train three SVM classifiers for the classification of human epithelial-2 cell images. Furthermore, Pang et al. [11] exploited deep learning and transfer learning techniques to learn the deep features of biomedical images online and, in an end-to-end fashion, distinguish the categories of images in several public biomedical image databases.

More recently, considering the lack of the availability of large collections of annotated biomedical images, alternative unsupervised deep learning method have also been proposed using only a few parameters of a deep model. For example, Shi et al. [12] improved an unsupervised principal component analysis network (PCANet) with random binary hashing and colour information to learn the deep feature of histopathological images and classify them with a matrix-form classifier on three datasets of small size (66, 100 and 66 images), which is another research direction of introducing deep learning into biomedical/medical image processing.

## 1.2 Observation and motivation

Although the deep models can capture more compact and hierarchical features with many hidden layers, a trained non-biomedical image classification model or standard deep architecture using deep neural networks cannot be directly used or trained for biomedical image classification tasks. The feature analysis, shown in Fig. 1 provides evidence for this observation. The trained deep model used is AlexNet [14], which claimed the championship in the ILSVRC-2012 competition with 1.2 million high-resolution images under 1000 different categories. The current popular deep models, such as AlexNet, aim to classify images or objects but in the presence of major differences between the different categories. However, on different biomedical images, the retrained AlexNet cannot distinguish similar images from different classes, such as images C01001 and C04001 that are from the OASIS-MRI dataset but belong to different health conditions. This is because the

**Fig. 1** Feature analysis of shallow and deep layers using the retrained AlexNet model on different biomedical images. Each red circle represents a shallow feature map extracted from the corresponding input image; each green circle represents a deep feature map capturing the semantic information of its input image. **a** shows an image from the NEMA-CT dataset. **b** and **c** show some images from the different categories of the same OASIS-MRI dataset



two deep feature maps with a green circle that AlexNet exploited and learned for the C01001 and C04001 images of the OASIS-MRI dataset are highly similar or even the same, as shown in Fig. 1. However, as we can see from Fig. 1, AlexNet can classify images from different datasets with obvious and different deep feature maps. Meanwhile, we can easily observe that the shallow features between the images from the OASIS-MRI dataset can be used to distinguish them, thus capturing more detailed local features than deep semantic features. Therefore, motivated by the analysis of the differences among various images, we propose a novel convolutional neural network that fuses shallow features and deep features for biomedical image classification tasks.

Moreover, if we would like to achieve an excellent performance on biomedical image analysis with deep learning, the main point at present is that a large number of labelled biomedical images should be available in the first place so that we can use them to train a deep neural network that has millions of parameters. However, there are two key problems that need to be addressed first. (1) Due to privacy considerations, it is extremely hard to obtain a sufficient number of annotated and labelled images that can be used to train such deep models in a given biomedical domain. (2) There may be slight differences between two given images from a biomedical or medical area, and this could mean that the two images may indicate two different types of diseases, as discussed before in Fig. 1b, c. Therefore, to address these two key problems for biomedical image classification, a fused convolutional neural network is proposed that learns pre-trained information from another domain and transfers it into the biomedical area using data augmentation, and subsequently constructs a deep model that adequately exploits shallow features and deep features simultaneously to distinguish biomedical images. Moreover, the shallow layers provide more detailed local features that can be used to distinguish different diseases from the same category, while the deep layers convey more high-level semantic

information, which can be used to classify the diseases among various categories.

### 1.3 Our contributions

The key contributions and highlights of our work are summarized below:

- 1) We propose a novel fused deep neural network by leveraging the low-level features from the shallow layers and the high-level features from the deep layers, which can accurately capture and discriminate the tiny differences between similar biomedical images from the same category that belong to different classes.
- 2) We further validate the importance of transfer learning for biomedical image classification tasks especially when there is a lack of annotated biomedical images for training deep neural networks.
- 3) We also observe that much deeper neural networks should not really be used to classify biomedical images. Much deeper neural networks tend to capture more abstract features and hence, tend to overlook the minor variances between similar images that may be the key to diagnosing different diseases.
- 4) Furthermore, this paper highlights the main difference between our fused convolutional neural network and the current regular deep convolutional neural networks in the aspect of designing a neural network architecture. Moreover, this paper explains why we do not directly use the current trained deep models to classify biomedical images due to the different image types and application purposes.
- 5) Finally, our novel biomedical image classification algorithm is evaluated using three public biomedical image datasets and the ImageCLEFmed dataset and it is compared with other traditional methods based on hand-

crafted features and current popular deep methods for biomedical image classification. The analysis of the classification accuracy shows that our proposed fused convolutional neural network approach for biomedical images has stable performance and notable improvement.

The rest of this paper is structured as follows. Section 2 presents our proposed deep architecture for biomedical image classification, including transfer learning technology, fused convolutional neural networks, parameter learning and data augmentation. Section 3 discusses several popular methods which are compared with our method, such as traditional approaches and deep models; it also introduces several public biomedical image datasets used as a benchmark in our experiments. Most importantly, the experimental results together with visualization analysis are presented in this section. Section 4 presents a comparison between the fused deep model and non-fused deep model and discusses misclassification cases and tiny differences between different classes. Finally, Section 5 concludes the paper with a brief summary of our contributions.

## 2 Methods

To overcome the two key problems of biomedical image analysis discussed above, transfer learning is used in our deep neural network architecture, wherein we transferred a model together with its millions of parameters, which are learned from a different generic image domain for the biomedical image domain. In this way, we offset the lack of training data for supervised learning from biomedical images. In addition, in order to better capture the tiny features and differences from the same category of images, we redesign the traditional deep convolutional neural network architecture for biomedical image classification, which achieves superior performance with accurate results. Finally, we introduce how to learn the parameters of the deep model and discussed why data augmentation is needed.

### 2.1 Transfer learning technology

Transfer learning aims to train a robust and discriminative model across different domains [18]. Simply put, transfer learning exploits the knowledge learned from a distinct domain to predict the probability distribution for a novel domain, which is, in our case, biomedical image classification. However, this kind of a method of transfer learning discussed in papers [9, 10], which use deep neural networks as a kind of a black box, only captures the general image features between different image domains. When there is no further learning from the knowledge transferred from a distinct domain, it is hard to achieve promising and stable performance in a novel domain.

Therefore, to fully exploit the deep transfer learning technology, we not only transfer the learned knowledge using deep convolutional neural networks from a different domain but also further tune the new knowledge for biomedical image classification for the novel domain. In this study, we use CaffeNet [19], which is pre-trained on a large ImageNet dataset that includes several millions of images with 1000 categories. Here, the natural question to ask was why we do not transfer much deeper models, such as VGG [17] or GoogLeNet [20], which have much better performance than AlexNet in the ImageNet image classification task. This is because biomedical images that are different from other general images have a great deal of tiny information and patterns inside as shown in Fig. 1. So, if we use the deeper layers to extract more semantic and abstract features, the model may ignore and lose the subtle differences between two similar images from the same category which may after all belong to two kinds of diseases.

Finally, deep transfer learning successfully alleviates the lack of training data to train a robust and deep neural network for the biomedical domain by transferring the learned knowledge from a different generic image domain. Moreover, on the basis of the domain transfer idea, we further propose our fused convolutional neural network for biomedical image classification in the next subsection and also discuss the differences between our deep model and other deep models and show how our deep model can effectively capture tiny biomedical image features.

### 2.2 Fused convolutional neural network

With the aim of designing a more robust and adaptive deep model for biomedical image classification, we propose a novel convolutional neural network architecture that fuses shallow features from the lower layers with the deep features from the higher layers and then is retrained with corresponding biomedical images before the classification task is performed. Different from the current CNN models, our proposed CNN model can adequately mine detailed and tiny local features to assist the biomedical image classification task and does not only rely on high-level semantic information to distinguish the different objects. Therefore, we redesign the traditional deep convolutional neural network architecture for biomedical image classification, which has been shown to achieve superior performance with accurate results in our experiments.

The main difficulty of designing a new deep neural network is to carefully consider the number of neurons and types of each layer, the size and number of convolutional kernels, and even how to match and connect any two layers. It is generally not easy to design a novel deep model for a specific domain and application. Our designed fused deep convolutional neural network architecture is shown in Fig. 2. For biomedical image classification, we use the Caffe



framework [19] and tools to construct our model and to train it with biomedical images, where this tool is exploited by the Berkeley Vision and Learning Center.

As shown in Fig. 2, the designed deep neural network architecture includes one trunk and two branches, which respectively represent one-layered deep feature from the fifth convolutional layer and two-layered shallow features from the first and the second convolutional layers. As the most significant highlight of this novel deep model, the fusion of these three layers is the key point to realize our idea for biomedical image classification. To address the fusion challenge, we analysed the size and type of each layer in detail and considered the size of convolutional kernels and average-pooling boxes. Finally, we respectively added a pooling layer with different average-pooling boxes after the first pooling layer and the second pooling layer and then made the output size of each shallow feature the same as the size of the fifth max-pooling layer. Before the fully connected layers, we need to add a concatenated layer to join these three feature outputs together. The above analysis is from the perspective of the characteristic and difference of our proposed deep model compared to other deep models that only have a main trunk to train for image classification tasks. As shown in Fig. 2, this overall deep convolutional neural network includes five convolutional layers (conv1, conv2, conv3, conv4 and conv5), three max-pooling layers (pool1, pool2 and pool5), two average-pooling layers (pool11 and pool21), two normalization layers (norm1 and norm2), one concatenated layer (concat) and three fully connected layers (fc6, fc7 and fc8) for biomedical image classification. For example, for the input image in Fig. 2, our proposed fused convolutional neural network gives an accurate classification prediction that is the fourth class of OASIS-MRI dataset (OASIS-MRI-4). In addition, for more detailed information about each kind of layer in the whole architecture shown in Fig. 2, we refer the reader to the introduction of the Caffe framework [19].

### 2.3 Parameter learning and data augmentation

Due to the numerous parameters of our model, we need to choose an efficient algorithm to tune and update them in order to accurately distinguish each biomedical image. Our proposed convolutional neural network exploits the back-propagation algorithm to update the network parameters  $\theta = \{W_b, b_i\}$  by minimizing the following loss function between the predicted results and the real labels of images:

$$L = -\frac{1}{|X|} \sum_{i=1}^{|X|} \ln(p(y^i|X^i)) \tag{1}$$

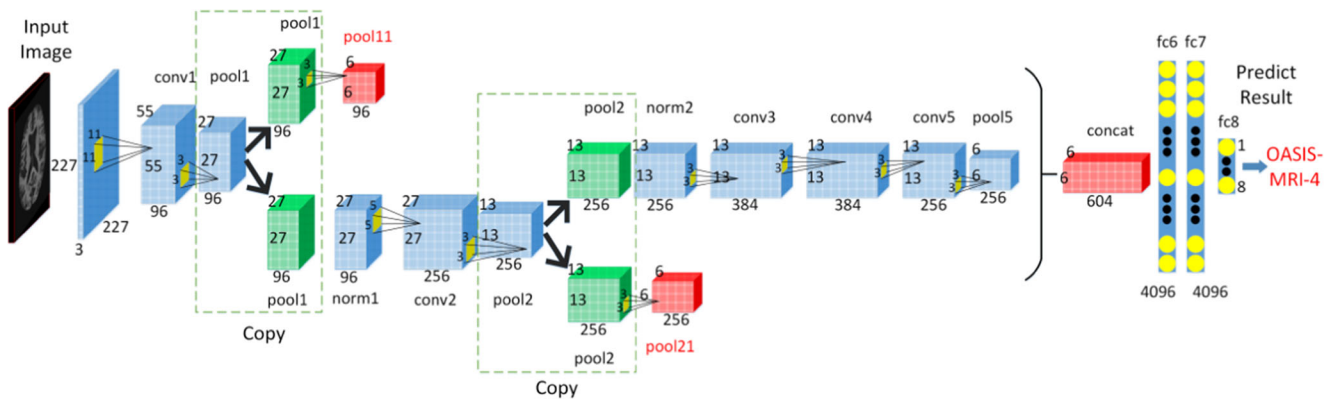
In this loss function  $L$ , we denote the number of training data by  $|X|$ , and  $X^i$  and  $y^i$  denote the  $i$ th training sample and its real label, respectively.

Here, the goal when we train a deep model is to make sure that the loss function is equal to zero when it is evaluated on the testing dataset, which means that all of the predicted results are the same as their real labels. To achieve this goal, the stochastic gradient descent method is used in our algorithm to compute and update the network parameters  $\theta$ . In this way, after each modification of parameters  $\theta$ , the loss function  $L$  reflects the corresponding change and then it is used to predict the result of all of the testing dataset again. In this way, a deep neural network will be trained after many iterations until the prediction result has 100% precision. Here, the updating process of the network parameters  $\theta$  is given in the following:

$$\theta(t + 1) = \theta(t) - \lambda \frac{\partial L}{\partial \theta} + \alpha \Delta \theta(t) - \beta \lambda \theta(t) \tag{2}$$

When the process is performed at iteration  $t + 1$ , the network parameters can be calculated based on the derivative at iteration  $t$ , joined together with the momentum weight and the weight decay to update it.  $\lambda$  is the learning rate that controls the learning speed. If  $\lambda$  is too large, the loss function could miss the optimal solution; if it is too small, the algorithm has to spend a great deal of time looking for an optimal solution, and it may even fall into the local optimum.  $\alpha$  is the momentum rate that speeds up the learning process and is helpful for obtaining the global optimal solution rather than the local optimum.  $\beta$  denotes the weight delay rate, which is used to slightly reduce the decayed weight parameters towards zero at each iteration and to improve the learning efficiency of the overall network parameters.

Although we successfully overcome the lack of biomedical training data by using the domain transferred learning method, the lack of sufficiently many labelled biomedical images needs to be addressed with other technology from the biomedical image perspective. In addition, the lack of sufficient training data will result in trained deep neural networks with an overfitting problem, which means that the trained model can only recognize the training data but does not know much about the testing data. To alleviate this overfitting phenomenon, we use two kinds of data augmentation methods in the process when training our fused deep convolutional neural network. First, each category should have the same number of images which is helpful to improve the classification accuracy of the deep CNN model. Here, we adopt a simple method of repeating the images from the same class. Another key method we used is that, during the training process, we resize the input image into a  $256 \times 256$ -pixel resolution but only a cropped image of a  $227 \times 227$ -pixel resolution is taken each time. As we can see from Fig. 2, the image that is input into our deep model is set to a  $227 \times 227$ -pixel resolution, and in this way, we can enlarge the number of the training data by different cropped samples to train our deep biomedical image classifier. In the future, we may also consider other data



**Fig. 2** The deep architecture proposed for fused biomedical image classification. It shows the whole architecture for biomedical image classification, where we can clearly see the feature-fusing process going from the shallow layers and the deep layers

augmentation methods; however, the simple data augmentation methods adopted in our work produces satisfactory results.

### 3 Results

This section first presents six compared methods, including traditional classification algorithms with perfect hand-crafted features and machine-learning classifiers, and deep classification models for biomedical images. In addition, several public biomedical image datasets were used for training and testing our fused deep convolutional neural network and comparing it with other approaches. Most importantly, we introduce our detailed implementation with model parameters, and then compare our fused convolutional neural network with other state-of-the-art methods for biomedical image classification. The evaluation criterion is chosen as the accuracy rate for classifying testing images. In addition, we also use t-SNE visualization analysis to show the classification performance by mapping the relationship from the high-dimensional space into the two-dimensional space.

#### 3.1 Traditional methods, deep methods and several public datasets

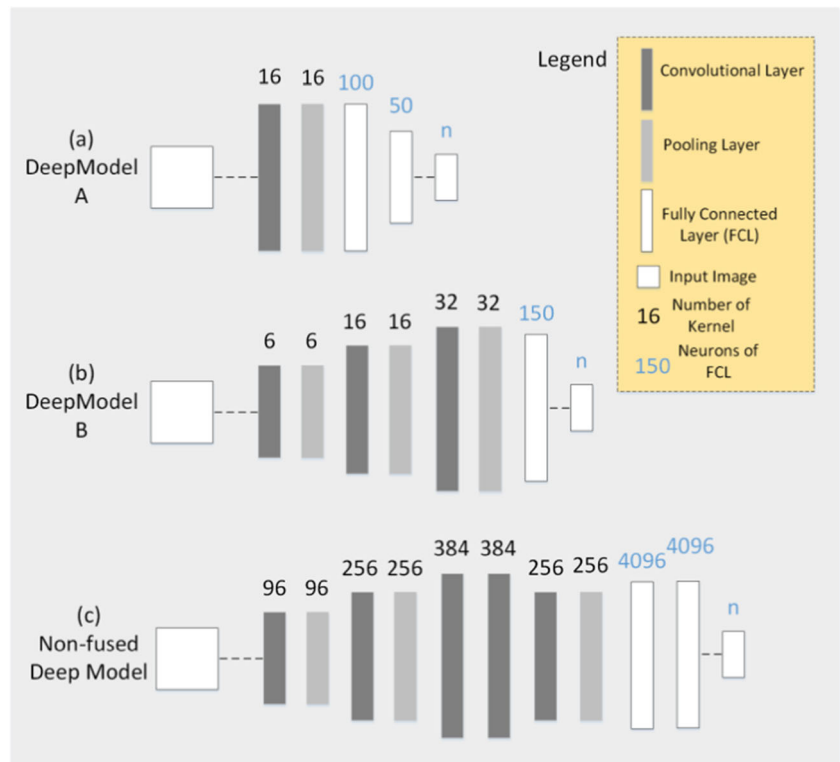
The key points of traditional methods for image classification is the design of a good feature descriptor and training a robust classifier. The state-of-the-art features widely used in biomedical image classification [2–4] include colour, LBP and HOG, which have an excellent discriminative power to represent different images and are the basis of training classifiers. In this paper, we choose three kinds of traditional algorithms to evaluate the classification performance for biomedical images. The first one uses the colour histogram to extract the biomedical image features and then trains a multiclass support vector machine with the same training data. The second method uses a local binary pattern descriptor, which is a very popular

texture feature in biomedical image analysis [21, 22] and then combines it with a *k*-nearest neighbour classifier trained with the same training data to label each testing image. The last traditional model exploits the best hand-designed feature, histograms of oriented gradients, into the biomedical image classification task and then trains multiple SVM classifiers with one-versus-one strategy.

The deep methods used in the computer vision area have achieved superior performance to address numerous visual problems, including image classification [14–16]. The advantage of deep models is that multiple hidden layers from deep neural networks can represent many different levels of features, such as the visual cortex of a human being. Therefore, the deep features from different layers contain richer and more compact features to recognize and distinguish an image than a single hand-designed feature. Based on successful applications in image classification, several newest studies have exploited deep learning in biomedical image classification tasks. To compare them with our proposed method, we coded their algorithms following the designed models in detail from their papers. DeepModel A [6] and DeepModel B [7] are good attempts to use deep learning methods to address the biomedical image classification problem. As mentioned in the introduction part and shown in Fig. 3, those deep models only design their deep architectures with different combinations among convolutional layers, pooling layers and fully connected layers. In effect, there is no major difference in essence between these models and other popular and mature CNN models, such as AlexNet and VGG.

To evaluate the performance of our proposed method, we have tested all the biomedical image classification methods mentioned above on three publicly available black-and-white biomedical datasets, that is, NEMA-CT, TCIA-CT and OASIS-MRI. Moreover, to avoid the randomness of the results, we have performed a 5-fold cross-validation on these three biomedical datasets. By grouping all the images from each dataset into five groups, we chose any four groups as the training set to implement five fully independent experiments

**Fig. 3** The deep architectures compared in this paper for biomedical image classification. **a** shows the deep model named DeepModel A from [6], which only includes one convolutional layer and three fully connected layers. **b** shows the deep CNN model for biomedical image classification from [7] that consists of three convolutional layers and two fully connected layers. **c** shows the popular deep CNN architecture used in many computer vision tasks, where we also select this model [14] as a non-fused deep model to compare with our proposed fused deep CNN model



and use the last group to test the classifier. For the grouping method, we adopt that sequences of five images are placed into a group, so there are a total of five groups.

The NEMA-CT dataset is from the National Electrical Manufacturers Association [23], which involves different body parts with various pictures that are in the DICOM format. This widely used dataset includes 499 images with a  $512 \times 512$ -pixel resolution. Based on the visual cues of these images, these biomedical images are divided into eight categories (104, 46, 29, 71, 108, 39, 33 and 69 images of each class). For each cross-validation, as an example, we could obtain 696 images as our training set using the simple data augmentation technology mentioned before and 87 images as our testing set to train and test each biomedical image classifier.

The TCIA-CT dataset is built by the National Cancer Institute and Washington University to support the National Institute of Health’s call and secondary research [24]. Following the same settings with other papers [21], 604 colon images in the dataset are also divided into 8 categories with 44, 43, 54, 88, 116, 96, 76 and 87 images respectively. Finally, we construct the training dataset of 744 images and the testing dataset of 117 images as one of the 5-fold cross-validations.

The OASIS-MRI dataset supplied by the Open Access Series of Imaging Studies (OASIS) contains magnetic resonance imaging (MRI) biomedical images that can be used only for study and research [25]. This dataset contains 416 images from subjects 18 to 96 years old; however, it is very

difficult to distinguish the images even if we observe them very carefully. Finally, as was done in previous works, based on the shape of the ventricular of images, they are divided into 4 classes, which have 141, 123, 86 and 66 images respectively. For this dataset, we select 452 images as the training dataset and 82 images as the testing dataset with the simple data augmentation methods, as an example, to perform a cross-validation.

Furthermore, in order to test the robustness of our fused convolutional neural network, the ImageCLEFmed dataset is used to evaluate the performance on modality classification of medical images as a subtask of ImageCLEF2015. This subfigure classification task was proposed first in ImageCLEF2015 [26, 27]. This colourful classification dataset includes not only regular diagnostic images (e.g. magnetic resonance imaging, ultrasound, computerized tomography, X-ray) but also generic biomedical illustration images (e.g. tables and forms, screenshots, flowcharts, gene sequences, statistical figures-graphs-charts, hand-drawn sketches), which significantly increases the difficulty of biomedical/medical image classification. Therefore, our aim is to classify each image from the testing set into the 30 modalities. To fairly compare the classification accuracy of our proposed algorithm with the competition results from ImageCLEF2015, we only use the 4532-training set to train each classifier and then evaluate it on the 2244 testing set of ImageCLEFmed dataset. For the number of each class from the training set, please refer to Table 1.

**Table 1** The number of images in the training set of each class in the ImageCLEFmed dataset. The name of each class is the abbreviation of each subfigure modality. For the full name of each class, please refer to the overview paper of [26]

Class	Number	Class	Number	Class	Number
D3DR	89	DRUS	26	GFLO	14
DMEL	148	DRXR	31	GGEL	228
DMFL	594	DSEC	10	GGEN	6
DMLI	430	DSEE	4	GHDR	84
DMTR	223	DSEM	4	GMAT	7
DRAN	10	DVDM	17	GNCP	54
DRCO	6	DVEN	12	GPLI	1
DRCT	55	DVOR	38	GSCR	20
DRMR	96	GCHE	53	GSYS	25
DRPE	10	GFIG	2190	GTAB	47

### 3.2 Model parameters

We train the traditional methods, DeepModel A [6] and DeepModel B [7], following the parameter settings from the relevant papers but by using the same training data and testing data we used for our proposed algorithm. However, in order to fairly compare our fused deep model and non-fused deep model in the next section, we have used the same model parameters. For each dataset, the training mini-batch size is the same, with 32 images each time. The learning rate is  $1e-6$  at the beginning and it is then divided by 10 with a step learning rate policy. In addition, the momentum and weight decay are set to 0.9 and 0.0002, respectively. Moreover, in order to focus on the convergent behaviour of each deep model, we control the iteration times that makes each deep model attain its best status on each dataset. Empirically, the pooling layers are used with max pooling in the main trunk and average pooling in the two branches, and the activation function in our proposed deep model is the rectified linear unit (ReLU), which is a kind of a non-saturating non-linear function that can prevent the output gradient from dropping close to zero. In the next set of experiments, we have fixed the values of all the parameters discussed above in each experiment and investigated the benefits of our fused deep model for biomedical image classification. Note that all the experiments were performed on a computer with Intel(R) Core(TM) i7-4710HQ CPU @ 2.50 GHz, 16.0 GB RAM and 64-bit Windows 8.1 Operating System.

### 3.3 Classification results

To comprehensively evaluate our fused convolutional neural network for biomedical image classification and to compare it with other popular classification methods, we have performed a sufficient number of experiments on three public biomedical

image datasets with 5-fold cross-validation. The classification accuracy rates for all the methods are presented in Table 2. As can be seen from the results, our proposed novel deep model has achieved the best classification performance on each dataset. Our work confirms that deep models have an advantage in classifying biomedical images over the traditional methods, combining feature extractors and typical classifiers, respectively. In fact, the traditional method based on the HOG feature and the SVM classifier also showed its powerful ability among all the traditional models. In particular, the traditional methods usually lack stable performance on different datasets. When the biomedical images are hard to distinguish, such as those in the TCIA-CT and OASIS-MRI datasets, the first two methods cannot give a satisfying answer and produce a large fluctuation in classification accuracy rates. In addition, only our proposed method has achieved 100% classification results on the TCIA-CT dataset. Moreover, our novel deep architecture has yielded better results than the two other compared deep models. In other words, a popular deep architecture designed with different layers one by one is not much useful to capture the shallow and tiny features of images from the same category, and it is not suitable for classifying biomedical images, particularly those disease images with slight differences but belonging to the same disease. For example, the colon images in Fig. 8 are difficult to classify because the differences from the CT images are very small. In a nutshell, our proposed fused convolutional neural network is more accurate and stable than other traditional methods and popular deep models for biomedical image classification.

In addition, we also compared our proposed algorithm with other classifiers from the ImageCLEFmed dataset from ImageCLEF2015 regarding the modality classification of medical images. As seen in Table 1, it is very hard to train a good classifier with the training set, due to the serious imbalance in the number of samples in each training class. When we look into the dataset closely, we observe that the smallest class (GPLI) has one training sample and the largest one (GFIG) 2190. As a result, based on the competition of ImageCLEF2015, we can obtain the subfigure classification results [26, 28–30]. Furthermore, we have also evaluated the performance of our fused CNN model and other traditional and deep classifiers, and the results can be seen in Table 3. As we can conclude, the classification accuracy (70.24%) of our model has outperformed the best competition result (67.60%) from ImageCLEF2015 and was also better than the other compared classifiers. Thus, for the modality classification task of medical images, our algorithm also has a good classification accuracy and strong competitiveness.

### 3.4 Visualization analysis

To evaluate our proposed algorithm for biomedical image classification tasks, we have further studied classification



**Table 2** Accuracy (%) comparison of classifiers across three public biomedical image databases, where the accuracy rate is denoted with the mean value and standard deviation (mean ± SD) on a 5-fold cross-validation

Traditional and deep methods	Biomedical image datasets		
	NEMA-CT	TCIA-CT	OASIS-MRI
ColorwithSVM	50.09 ± 38.58	47.22 ± 9.42	76.22 ± 4.15
LBPwithKNN	99.61 ± 0.54	71.38 ± 4.09	42.72 ± 6.01
HOGwithSVM	100.00 ± 0	98.68 ± 0.72	81.01 ± 6.65
DeepModel_A	100.00 ± 0	98.33 ± 0.08	87.72 ± 0.10
DeepModel_B	100.00 ± 0	99.47 ± 0.03	89.23 ± 0.08
The proposed model	100.00 ± 0	100.00 ± 0	93.44 ± 0.04

performance from the perspective of image visualization. Here, we have used a t-distributed stochastic neighbour embedding (t-SNE) map to show the classification ability of our fused convolutional neural network. The t-SNE algorithm [31] is very suitable for the visualization of a high-dimensional feature space by the dimensionality reduction method, which exploits the Barnes-Hut approximation strategy [32] to map a high-dimensional space into a two-dimensional space. Furthermore, the t-SNE method can transfer the approximate relation between images from the high-dimensional feature space into the two-dimensional plane. In this way, it becomes easy to observe whether similar images are distributed into the same area or not, which in turn reveals whether classification of biomedical images by our method is good or not. Therefore, we have first introduced the t-SNE evaluation algorithm into the biomedical image classification task for visual analysis.

Figure 4 shows the performance of biomedical image classification based on our proposed deep fused convolutional neural network. In particular, the t-SNE visualization map is a kind of image exploration using the manifold learning idea, which transfers the spatial relationship from a high-dimensional feature representation into a low-dimensional feature representation. However, we observe that the resulting effect is very convincing and proves the classification ability of our model of biomedical images. For the NEMA-CT and OASIS-MRI datasets, we have observed that the clustered

images of each category are very clear as a whole, which shows that our fused deep model is effective and accurate in distinguishing different kinds of biomedical images. For the TCIA-CT database, the resulting shape was similar to a manifold structure, which also accurately classifies different images because this overall dataset was generated in a sequential manner based on time. In other words, several images are divided into two neighbouring categories. Moreover, the local enlargement on a part of the whole map shown in Fig. 4, with a five-pointed star, also has confirmed that the images that are similar on the high-dimensional feature space are also close on the two-dimensional plane. The detailed visualization maps are available in our supplementary documents.

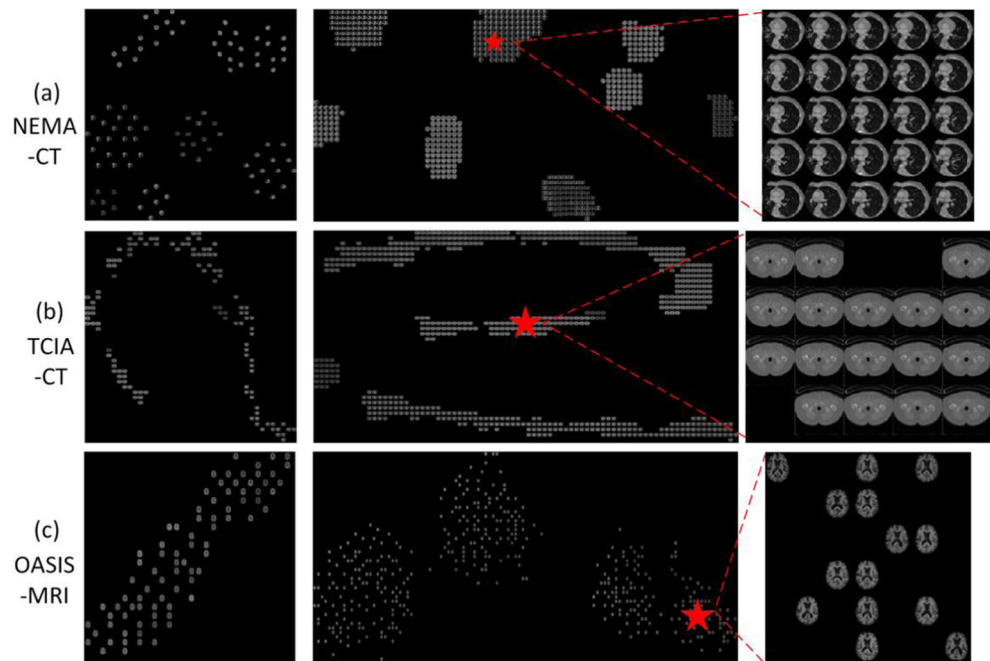
### 4 Discussion

To further evaluate the performance of our proposed fused convolutional neural networks for biomedical image classification, we have analysed the differences between our fused deep model and non-fused deep model on all the biomedical image datasets in detail. In addition, in order to identify and elucidate the reason why the classification accuracy of our proposed algorithm was not 100% on the OASIS-MRI dataset, the misclassification cases are shown in this section with a detailed description. Again, the OASIS-MRI dataset is the most difficult dataset for classification as we have

**Table 3** Subfigure classification accuracy (%) on the ImageCLEFmed dataset with different classifiers for modality classification of medical images

ImageCLEF2015			
Methods	Accuracy (%)	Methods	Accuracy (%)
ColorwithSVM	5.08	BCSGwithMixed [28]	67.60
LBPwithKNN	44.70	BCSGwithTextual [28]	60.91
HOGwithSVM	53.97		
DeepModel_A [6]	50.47	BCSGwithVisual [28]	60.91
		CMTECH [29]	52.98
DeepModel_B [7]	49.35		
		BMET [30]	45.63
The proposed model	70.24		

**Fig. 4** Visualization analysis of our proposed method on three public biomedical image datasets. Left (a–c): the classification results of the testing data. Middle (a–c): the space relationship of all data as a whole. Right (a–c): a detailed local enlargement of the corresponding part annotated with a five-point star in the middle image



explained in Section 3. Moreover, we have also analysed the tiny differences between similar images belonging to different categories.

#### 4.1 Fused deep model vs non-fused deep model

In this experiment, we have attempted a detailed comparison for the differences between our proposed fused deep method and non-fused deep method. In Section 2, we have introduced our proposed convolutional neural network by fusing the shallow features from the lower layers and the deep features from the higher layers. To further demonstrate the superiority of our fused biomedical classification model, we have performed an experiment to compare the fused deep model and the non-fused deep model on all the biomedical image datasets. For the non-fused deep model, we have only needed to reduce and move the fused layers added in our deep framework, which are the pool11 layer, pool21 layer and concat layer. Then, the non-fused deep model can be tuned into one stream deep convolutional neural network. Other parameter settings and experimental data are the same between them and then the non-fused deep model was trained iteratively. Finally, we have obtained the classification accuracy curves with the increase of iteration counts, as shown in Fig. 5.

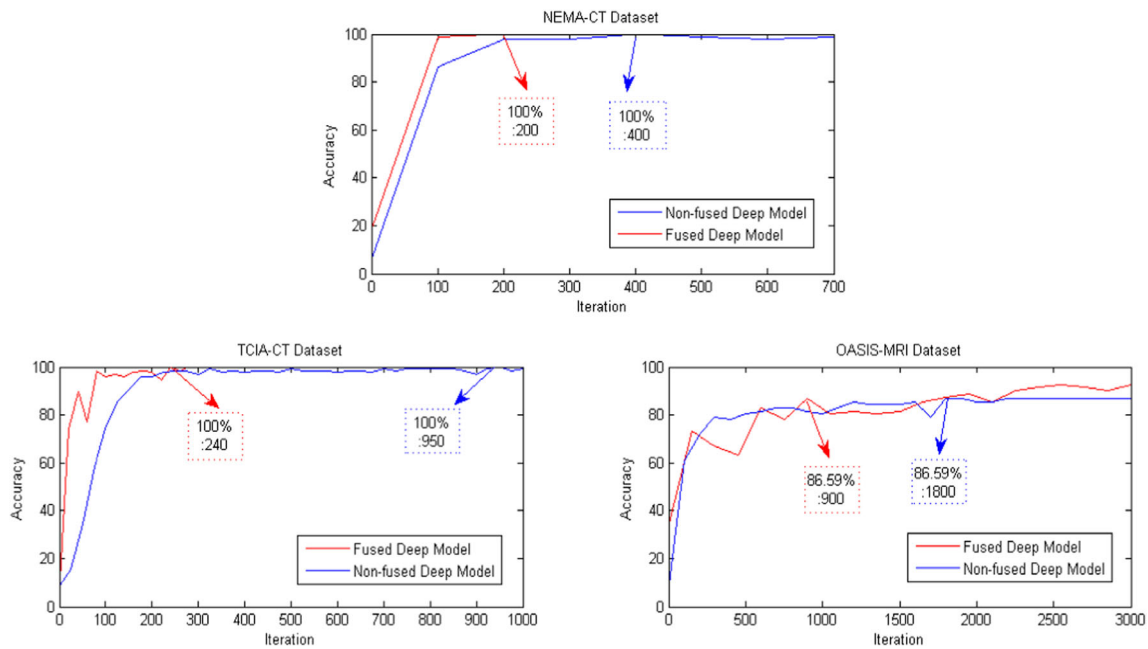
For the NEMA-CT dataset, we have evaluated the classification performance between the fused deep model and the non-fused deep model to show that the convergence speed of the former is much quicker than the latter. The fused deep model only needed 200 iterations to achieve a 100% classification accuracy, whereas the non-fused deep model needed up to 400 iterations to achieve a 100% classification accuracy. In

addition, we also applied these two deep models to the TCIA-CT dataset with successively produced biomedical images to evaluate their performance. The experimental results validated that our fused deep model can achieve a quick convergence status within 100 iterations, and it only needs 240 iterations to achieve a 100% classification accuracy, in contrast with the non-fused deep model that required 950 iterations. For the same reason, these two deep models are also evaluated with the OASIS-MRI dataset for biomedical image classification. We emphasize that the accuracy of 86.59% is first obtained at iteration 900 from our fused deep model; however, the non-fused deep model only achieves that accuracy when it reaches iteration 1800. Finally, our fused deep model converges to 92.68% for classification accuracy on the OASIS-MRI dataset, but the 86.59% accuracy is the highest classification result for the non-fused deep model.

Several interesting observations were made from the above experimental results: fusing the shallow features from lower layers yielded better accuracy than only using the higher layers; the convergence speed combining the fused deep model was much quicker than that of the non-fused one. Therefore, our fused deep algorithm for biomedical image classification was helpful in improving the classification accuracy and in accelerating the convergence speed compared to the normal and non-fused deep convolutional neural networks.

#### 4.2 Investigation of misclassification cases

To provide further analysis of the OASIS-MRI dataset, we recorded all the prediction results for each tested brain image and included the probability of each class that our algorithm



**Fig. 5** Performance comparison of our proposed fused deep model and non-fused deep model for biomedical image classification, for the NEMA-CT dataset, the TCIA-CT dataset and the OASIS-MRI dataset

predicted, and then we counted the confusion matrices on all the tested images from the OASIS-MRI dataset based on our proposed biomedical image classification method. Figure 6 shows the confusion matrix for all four of the considered classes. For instance, there were four images from class C2 that were misclassified as class C1 in this experiment. The relative confusion and high misclassification cases between classes C1 and C2 could be justified by the fact that the similarity between them is hard to discover. Figure 7 presents the difficult cases of these brain images that were misclassified, together with the corresponding prediction output of our fused deep model.

Referring to the correct sample images from Fig. 7a, the misclassified image C01140 could also be recognized as class

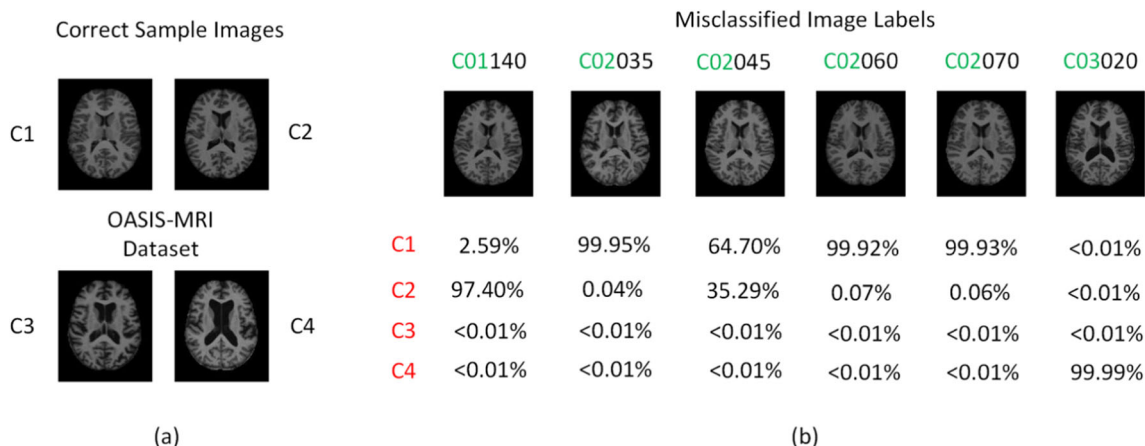
C2 because it is very similar to C1, and it is really difficult to classify the images from C1 and C2. Meanwhile, images C02035, C02045, C02060 and C02070 also look as though they belong to class C1. In the same way, we can also recognize image C03020 as class C4 based on the large shape of the ventricle in the image. As we can conclude, some especially similar biomedical images are very hard for a deep classifier to distinguish, whereas it is a normal phenomenon for any classifier without enough training in a biomedical image set. Obviously, this point highlights the importance of collecting large volumes of biomedical images annotated by physicians and experts.

### 4.3 Tiny differences between the classes

In this subsection, the aim was to demonstrate the powerful ability of our fused deep CNN algorithm to capture the tiny differences between similar images from different classes in the TCIA-CT dataset. As part of the motivation of this work, we focused on how to design a biomedical image classifier to capture much smaller, but key, differences between different categories. In other words, the general deep architecture designed with different layers one by one is not very useful to capture the shallow and tiny features of images from the same category, and it is not suitable for classifying biomedical images, such as those disease images with slight differences but belonging to the same disease. For example, compared with DeepModel B [7], which is a typical deep model used in various tasks, our fused deep CNN model was more robust in the detection of the key differences between similar images

		Predicted Label			
		C1	C2	C3	C4
Categories	X \ Y				
	C1	27	1	0	0
	C2	4	20	0	0
	C3	0	0	16	1
C4	0	0	0	13	

**Fig. 6** The confusion matrices based on the prediction results of all the testing images from the OASIS-MRI dataset. The entry in the X-th row and Y-th column corresponds to the number of images from class CX that were classified as class CY



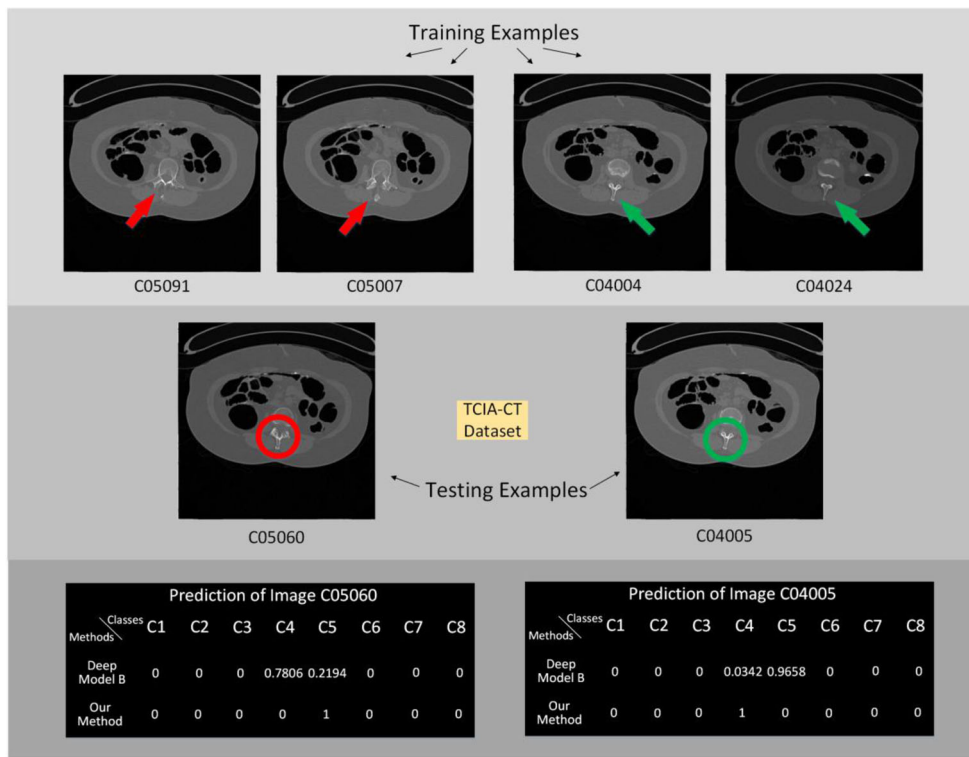
**Fig. 7** Correctly classified images and misclassified images of the OASIS-MRI dataset. **a** The correct sample images from each category. **b** All the misclassified images with their possibility outputs of our proposed fused CNN below each image

from different categories in the TCIA-CT dataset, as shown in Fig. 8.

As we can see from the performance results, there were several significant observations to be made: for these two testing images, DeepModel B could not correctly discriminate their class; moreover, our fused deep model, by combining the low-level features and the high-level features, exhibited better performance and correct classification results. Furthermore, as it is analysed from the several training images, the tiny differences between classes C4 and C5 are the green and red arrow positions with similar backgrounds between them. Therefore, we also observed from all the training images from these two

categories that testing image C04005 should belong to class C4 and image C05060 should belong to class C5. Due to the tiny, but critical, differences between the biomedical images, a generally designed deep CNN model, such as DeepModel B, could not detect or classify them using only high-level features, but our deep CNN model could accomplish this by fusing the low-level features and the high-level features from the shallow layers and the deep layers, respectively. Therefore, our fused algorithm was better than other general deep models in biomedical image classification with a higher accuracy because the fusion strategy helps us to effectively take the shallow features and the deep features into account together.

**Fig. 8** The performance of capturing tiny differences from two classes of images in the TCIA-CT dataset using DeepModel B and our proposed deep CNN model. Here, we list the misclassified images using DeepModel B to classify colon images. Meanwhile, we also give the prediction and classification results for our proposed method





## 5 Conclusion

In this paper, we have proposed a novel convolutional neural network architecture for biomedical image classification, which fuses shallow feature layers and deep feature layers together. In this way, our trained fused deep model not only accurately distinguishes images from different categories but also captures more detailed but tiny differences between the images from the same category, which is essential for identifying biomedical image characteristics. In contrast, the existing methods and general non-biomedical image deep models usually ignore the local features when only using higher semantic layers to classify different objects. Moreover, we have proposed the use of a domain transferred learning strategy to alleviate the lack of supervised biomedical images. Finally, through sufficient experiments and visualization analysis over three public datasets and through extensive comparisons with the traditional models and the two deep models, our fused convolutional neural network was shown to have superior performance in biomedical image classification. In addition, our proposed model has also shown a stronger competitiveness than other classifiers in modality classification of medical images by evaluating them using the ImageCLEFmed dataset.

In our future work, we will further employ our novel neural network architecture to handle other biomedical image classification problems. We also plan to extend our innovative approach by fusing shallow feature layers and deep feature layers in other deep models to tackle many more image processing problems in other domains of interest.

**Funding information** This work started while Shuchao Pang was a visiting scholar at Macquarie University from Mar. 2016 to Mar. 2017. This work has also been supported by the project of Science and Technology Development Plan of Jilin Province, China (Grant 20150204007GX) and Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant 20120061110045).

## References

- Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ (2015) Big data for health. *IEEE JBHI* 19(4):1193–1208
- Depeursinge A, Van de Ville D, Platon A, Geissbuhler A, Poletti PA, Muller H (2012) Near-affine-invariant texture learning for lung tissue analysis using isotropic wavelet frames. *IEEE Trans Inf Technol Biomed* 16(4):665–675
- Song Y, Cai W, Huang H, Zhou Y, Feng DD, Wang Y, Fulham MJ, Chen M (2015) Large margin local estimate with applications to medical image classification. *IEEE Trans Med Imaging* 34(6):1362–1377
- Khachane MY, Ramteke RJ (2016) Modality based medical image classification. In *Emerging research in computing, information, communication and applications* 597–606
- Ertugrul ÖF, Kaya Y, Tekin R (2016) A novel approach for SEMG signal classification with adaptive local binary patterns. *Med Biol Eng Comput* 54(7):1137–1146
- Li Q, Cai W, Wang X, Zhou Y, Feng DD, Chen M (2014) Medical image classification with convolutional neural network. In *Control Automation Robotics & Vision (ICARCV), 13th IEEE International Conference on* 844–848
- Gao Z, Zhang J, Zhou L, Wang L (2014) HEP-2 cell image classification with convolutional neural networks. In *Pattern recognition techniques for indirect immunofluorescence images (I3A), 2014 1st IEEE Workshop on* 24–28
- Abbas Q, Fondon I, Sarmiento A, Jiménez S, Alemany P (2017) Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features. *Med Biol Eng Comput*:1–16
- Ahn E, Kumar A, Kim J, Li C, Feng D, Fulham M (2016) X-ray image classification using domain transferred convolutional neural networks and local sparse spatial pyramid. *IEEE 13th ISBI* 855–858
- Phan HTH, Kumar A, Kim J, Feng D (2016) Transfer learning of a convolutional neural network for HEP-2 cell image classification. *IEEE 13th ISBI* 1208–1211
- Pang S, Yu Z, Orgun MA (2017) A novel end-to-end classifier using domain transferred deep convolutional neural networks for biomedical images. *Comput Methods Prog Biomed* 140:283–293
- Shi J, Wu J, Li Y, Zhang Q, Ying S (2017) Histopathological image classification with color pattern random binary hashing-based PCANet and matrix-form classifier. *IEEE J Biomed Health Inform* 21(5):1327–1337
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444
- Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* 1097–1105
- Ouyang W, Wang X, Zhang C, Yang X (2016) Factors in finetuning deep model for object detection with long-tail distribution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 864–873
- Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 3431–3440
- Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*
- Long M, Wang J, Jordan MI (2016) Deep transfer learning with joint adaptation networks. *arXiv preprint arXiv:1605.06636*
- Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia* 675–678
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 1–9
- Murala S, Wu QM (2013) Peak valley edge patterns: a new descriptor for biomedical image indexing and retrieval. *Proc. IEEE CVPR Workshops* 444–449
- Dubey SR, Singh S, Singh R (2015) Local bit-plane decoded pattern: a novel feature descriptor for biomedical image retrieval. *IEEE J Biomed Health Inform* 20(4):1139–1147
- NEMA–CT image database (2012) Available: <ftp://medical.nema.org/medical/Dicom/Multiframe/>
- Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L (2013) The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *J Digit Imaging* 26(6):1045–1057
- Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL (2007) Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged,

- nondemented, and demented older adults. *J Cogn Neurosci* 19(9): 1498–1507
26. de Herrera AGS, Müller H, Bromuri S (2015) Overview of the ImageCLEF 2015 Medical Classification Task. In *CLEF (Working Notes)*
  27. Kalpathy-Cramer J, de Herrera AGS, Demner-Fushman D, Antani S, Bedrick S, Müller H (2015) Evaluating performance of biomedical image retrieval systems—an overview of the medical image retrieval task at ImageCLEF 2004–2013. *Comput Med Imaging Graph* 39:55–61
  28. Pelka O, Friederich CM (2015) FHDO Biomedical Computer Science Group at Medical Classification Task of ImageCLEF 2015. In *CLEF (Working Notes)*
  29. Cirujeda P, Binefa X (2015) Medical image classification via 2D color feature based covariance descriptors. In *CLEF (Working Notes)*
  30. Lyndon D, Kumar A, Kim J, Leong PHW, Feng D (2015) Convolutional neural networks for subfigure classification. In *CLEF (Working Notes)*
  31. Van DM, Laurens (2014) Accelerating t-SNE using tree-based algorithms. *J Mach Learn Res* 15(1):3221–3245
  32. Rippel O, Paluri M, Dollar P, Bourdev L (2015) Metric learning with adaptive density discrimination. *arXiv preprint arXiv:1511.05939*



**Shuchao Pang** is currently a PhD student at Jilin University, Changchun, China. In June 2011, he received the BSc degree in computer science and technology from Jilin University of Finance and Economics, Changchun, China. In September 2011, he was enrolled at the College of Computer Science and Technology as a Master of Philosophy student, majoring in Computer Application Technology. Due to his excellent academic record and out-

standing scientific research capability, he has participated in a MPhil-PhD program through individual application, college selection and school approval. He completed his MPhil course in August 2013, and began his PhD's study in September 2013. Between Oct. 2013 and Feb. 2014, he visited Saint-Petersburg National Research University of Information Technologies, Mechanics and Optics (ITMO University), Saint-Petersburg, Russia, as a Joint Training Research Student. From Jul. 2015 to Jan. 2016, he was a Visiting Research Doctoral Student in the Machine Learning Group, Artificial Intelligence Center, University of Oviedo, Gijon, Asturias, Spain. During Mar. 2016 to Mar. 2017, he was a Visiting Scholar in the Department of Computing, Macquarie University, Sydney, NSW, Australia. His current research interests include Biomedical Image Processing, Object Tracking, Face Recognition, Deep Learning and Preference Learning.



Deep Learning, especially OCR, Face Recognition and Object tracking.

**Anan Du** received the MSc degree in Computer Application Technology from the College of Computer Science and Technology, Jilin University, Changchun, China in 2015 and BSc degree in Mathematics and Applied Mathematics from Shandong University of Finance and Economics in 2012. She is currently an engineer at China Mobile (HangZhou) Information Technology Co., Ltd., Hangzhou, China. Her current research interests include Computer Vision and



and senior program committee memberships of numerous national and international conferences. Recently, he was the program co-chair of The 14th Pacific-Rim International Conference on Artificial Intelligence (PRICAI 2010) and conference co-chair of The 7th and 8th International Conferences on Security of Information and Networks (SIN 2014 and SIN 2015). He is a Senior Member of IEEE (SMIEEE). His current research interests include knowledge discovery, multi-agent systems and trusted and secure systems.

**Mehmet A. Orgun** received the BSc and MSc degrees in computer science and engineering from Hacettepe University, Ankara, Turkey, in 1982 and 1985 respectively; and the PhD degree in computer science from the University of Victoria, Canada, in 1991. He is currently a professor at the Department of Computing, Macquarie University, Sydney. His professional service includes editorial and review board memberships of several leading journals and program committee



**Zhezhou Yu** is currently a Professor and Doctoral Tutor at the Department of Computational Intelligence, College of Computer Science and Technology, Jilin University, Changchun, China. He received the BSc degree in System Structure from the Department of Computer Science and Technology in Jilin University, Changchun, China in 1982; and in 1988, he received the MSc degree from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese

Academy of Sciences; and the PhD degree from College of Computer Science and Technology, Jilin University, in 2007. He has mainly researched in Computational Intelligence Theory and Embedded System Technology, and he has participated in three items of National Natural Science Fund Project and presided over more than six items of National Project and Provincial and Ministerial Level Scientific Research Project. At present, he is interested in Human Behavior Recognition, Face Recognition and other relevant Computer Vision technology research. He is currently a member of the editorial board of Journal of Jilin University (Science Edition) and Jilin Province Organizing Committee and the expert group of National College Students Electronic Design Competition. He has published more than 50 papers in SCI/EI journals and holds 4 copyrights and 3 patents.