第一篇

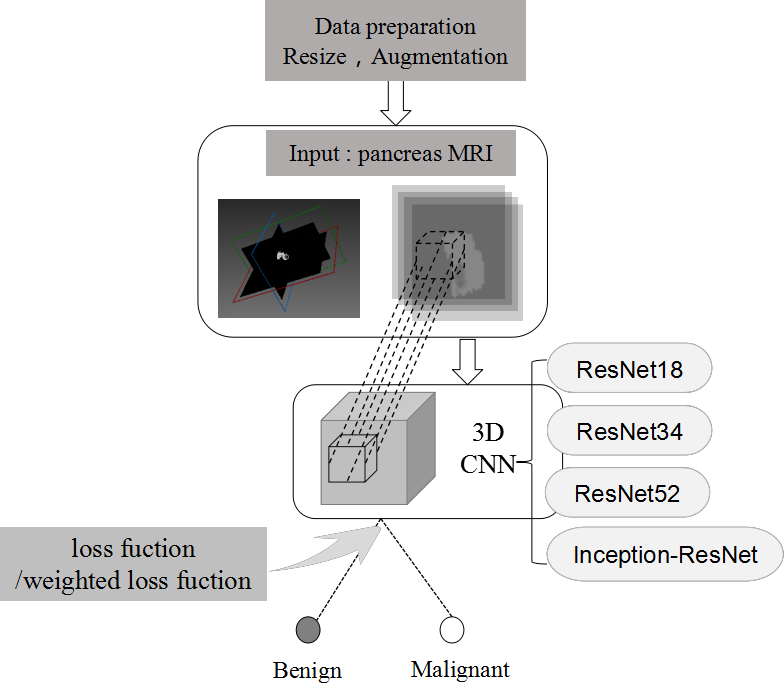
TITLE：Classification of Pancreatic Tumors based on MRI Images using 3D Convolutional Neural Networks

ABSTRACT:

Computer aided diagnosis of pancreatic cancer[胰腺癌的计算机辅助诊断] can help doctors improve diagnostic efficiency and accuracy, which does not depend on the doctor's subjective judgment and experience. Existing pancreatic tumors classification methods suffer from the problem of partial automation and ignore spatial and temporal characteristics[忽视肿瘤的时空特征]. In this paper, we used 3D versions of ResNet18, ResNet34, ResNet52 and Inception-ResNet[提出模型] for pancreatic magnetic resonance images (MRI) classification[胰腺磁共振图像(MRI)进行分类]. In order to alleviate the effect of class imbalance[减缓类别不均衡], we proposed a weighted loss function[加权损失函数]. Two sets of comparative experiments were performed to compare the effect of the proposed loss function. Experimental results show with weighted loss function, the performances of the four models are mostly improved except the precision of ResNet52. Moreover, the false negatives are reduced. Among them, ResNet18 performs best and achieves the accuracy of 91% .

METHOD:

We compare eight different approaches to pancreatic MRI classification with 3D CNNs[8种三维CNN实现胰腺MRI分类]. Taking into account the effect of the depth and width of the model, eight models are: ResNet18, ResNet34, ResNet52 and Inception-ResNet with binary crossentropy or the proposed weighted loss function[损失函数：二元交叉熵和提出的加权损失函数], respectively. The workflow of our method is as follows in Figure 1:



3D ResNet Network

Residual neural networks architecture[残差网络(微软研究院提出)]won the ImageNet contest in 2015 and demonstrated possibility to greatly improve the depth of the network while having fast convergence[面对深度网络可快速收敛]. There were already publications that showed their results for 3D brain MRI image classification. We use ResNet consisting of modules in Figure 4. As is shown in Figure 4(d), The output of the last ResNet module[ResNet最后一层输出] is sent to an average pooling layer to further reduce it to 1× 1× 1× 512[平均池化层], followed by a flatten layer with 512 hidden units[平坦层：512神经元] and a dense layer[全连接层] with an output for binary lassification[二分类输出] with sigmoid nonlinearity[sigmoid函数].

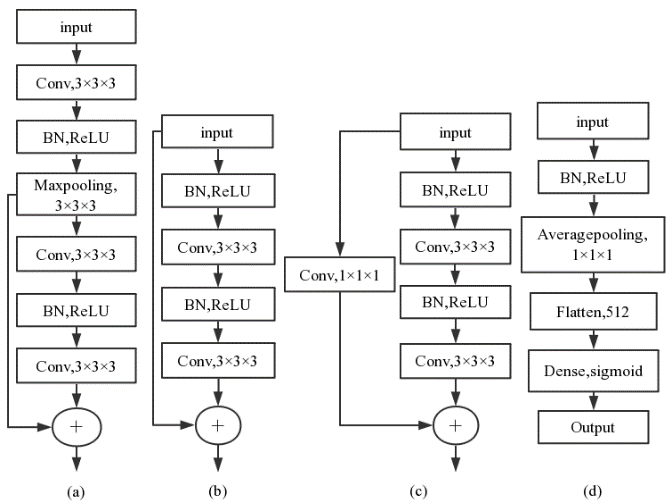
We use three ResNet architectures:

ResNet18 contains 1 module (a), 4 modules (b), 3 modules (c) and 1 module (d);

ResNet34 contained 1 module (a), 12 modules (b), 3 modules (c) and 1 module (d);

ResNet52 contains 1 module (a), 20 modules (b), 3 modules (c) and 1 module (d).

We train the final binary classification models using Adam with learning rate of 0.0001 and batchsize of 16 for 100 epochs.



3D Inception-ResNet:

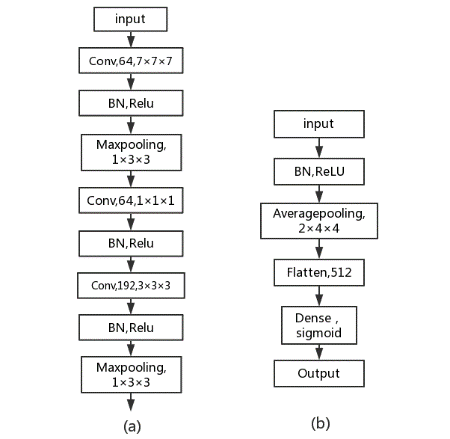
Inception-ResNet was proposed by Google, in which the top-5 error rate for the ImageNet classification challenge was reduced to 3%, with an ensemble of three residual and one Inception-v4[即融合残差网络和Inceptionv4].

We use:Inception-ResNet architecture contains 1 module (a), 9 modules (c) and 1 module (b) (see in Figure 5).

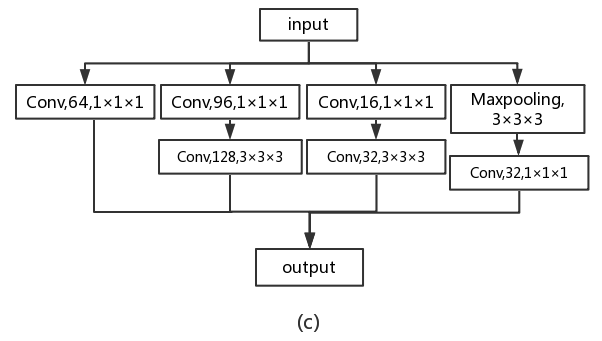
We insert:maxpooling layer after the second and seventh inception3d modules to reduce the dimension of the layer output.

As is shown in Figure 5(b),

The output of the last Inception-ResNet module is sent to an average pooling layer to further reduce it to 15× 1× 1×256, followed by a flatten layer with 3840 hidden units and a dense layer with an output for binary classification with sigmoid nonlinearity.



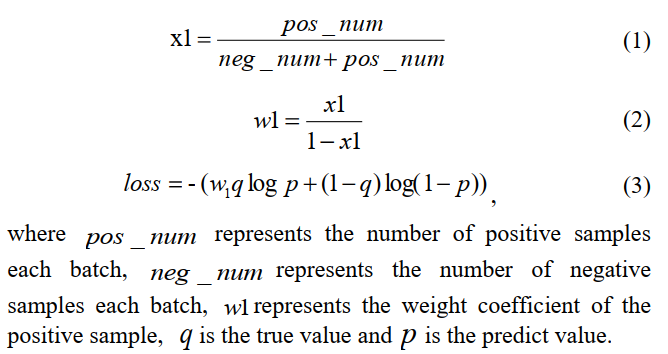
Inception-ResNet module (a) is the first input module. (b) is the output module.



(c) is the inception\_3d module

A Weighted Loss Function

Considering that the data set used in the experiment is classimbalanced. We give the sample with a smaller sample size a little more weight to weaken this imbalance problem on the basis of using data augmentation. We use the cross-entropy function[交叉熵] of tensorflow and modify[调整Loss使其适应于不均衡数据] it to fit our model motivated by which introduced a class-balancing weight on per-pixel term basis for image segmentation and which optimized the cost of decision making. In our method, we just calculate the number of positive and negative categories every time the parameter is updated and iterated, then derive the weight coefficients for each category. The formula is as follows:



TITLE:3D Inception Convolutional Neural Networks For Automatic Lung Nodule Detection

ABSTRACT:

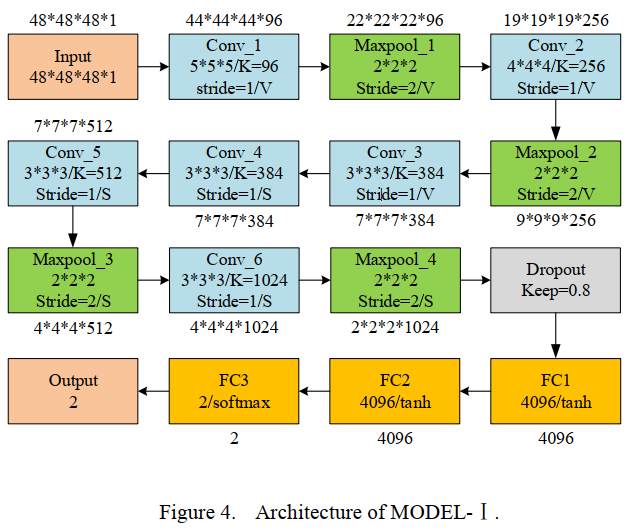
Lung cancer [肺癌]is the most common cause of cancer eath worldwide. Early detection of lung nodules in thoracic computed tomography (CT) scans[胸部计算机断层扫描(CT)] is currently the one of the most effective ways to predict and treat lung cancer. In practice, one CT image contains about 200 to 700 slices and this may cost radiologists a lot of time. In this paper, we propose two types of deep neural networks which are called MODEL-Ⅰ and MODEL-Ⅱ for automatically detecting lung nodule to help the radiologists with reading CT images. For the CT data pretreatment, we use filters to select the suspicious region for locating nodules in 2D images. we use downsampling and upsampling methods[上采样和下采样平衡数据集] to make dataset balanced. Inspired by Google 2D Inception module, we propose the inception block for 3D convolutional neural networks to accommodate the 3D nature of CT scans[提出3D卷积神经网络适应于3D-CT扫描图], which solve the gradient vanish problems and enhance the F1 score, experiment show that The MODEL-Ⅱ acquires the best F1 score of 0.979.

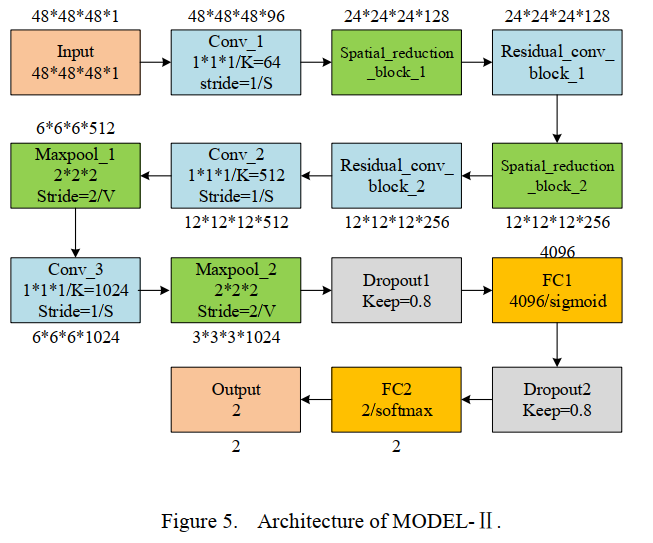
3D CNN Network Structure

We design two types of 3D CNN networks named MODEL-Ⅰ and MODEL-Ⅱ .

MODEL-Ⅰ contains 6 3D CNN layers, 2 Maxpool layers and 3 Fully Connected layers FC1, FC2 and FC3. Inspired by the Inception structure, we design

MODEL-Ⅱ with residual convolutional block and spatial reduction block, both of the blocks use inception structure. The details of network architecture are shown in Fig.4. for MODEL- Ⅰ and Fig.5. for MODEL- Ⅱ . To prevent overfitting, both networks use Dropout layer. At the end of each convolutional layers we use ReLU as activation function.





In Fig.4, 5 and 6, the ‘K’ in Conv block indicates the number of convolutional kernels; the ‘S’ and ‘V’ in Conv and Maxpool block indicate the padding method of ‘SAME’ and ‘VALID’ respectively. Fig.6. shows the detail of the spatial reduction block and residual conv block in Fig.5.

TITLE:3D Inception-based CNN with sMRI and MD-DTI data fusion for Alzheimer’s Disease diagnostics

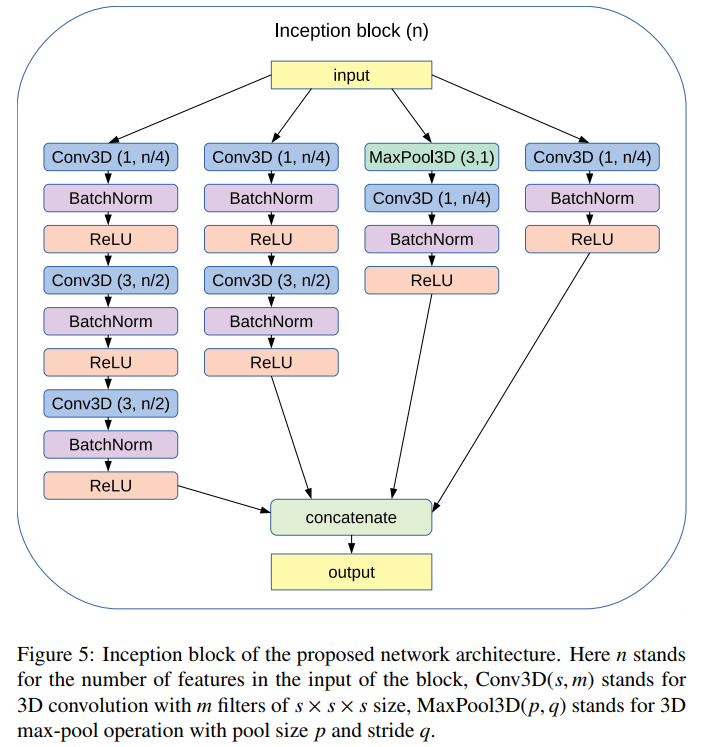
ABSTRACT:

In the last decade, computer-aided early diagnostics of Alzheimers Disease (AD)[阿尔茨海默病(AD)] and its prodromal form, Mild Cognitive Impairment (MCI)[轻度认知障碍], has been the subject of extensive research. Some recent studies have shown promising results in the AD and MCI determination using structural and functional Magnetic Resonance Imaging (sMRI, fMRI)[结构和功能磁共振成像], Positron Emission Tomography (PET)[正电子发射层析成像(PET)] and Diffusion Tensor Imaging (DTI)[扩散张量成像] modalities. Furthermore, fusion of imaging modalities in a supervised machine learning framework has shown promising direction of research. In this paper we first review major trends in automatic classification methods such as feature extraction based methods as well as deep learning approaches in medical image analysis applied to the field of Alzheimer’s Disease diagnostics. Then we propose our own design of a 3D Inception-based Convolutional Neural Network (CNN) for Alzheimer’s Disease diagnostics. The network is designed with an emphasis on the interior resource utilization and uses sMRI and DTI modalities fusion on hippocampal ROI. The comparison with the conventional AlexNet-based network[模型比较] using data from the Alzheimers Disease Neuroimaging Initiative (ADNI) dataset (<http://adni.loni.usc.edu>) demonstrates significantly better performance of the proposed 3D Inception-based CNN.

Proposed network architecture:

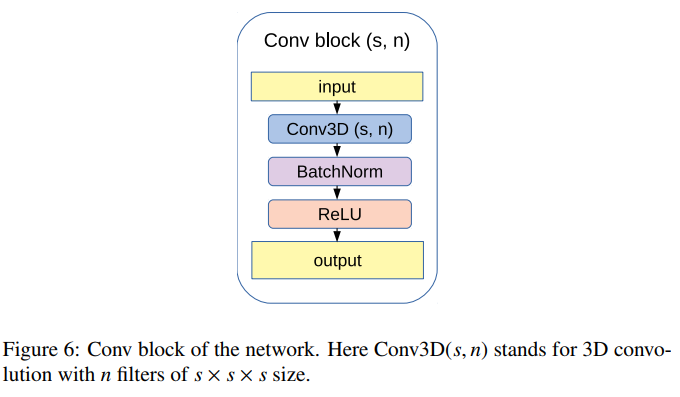
In this work we propose a new 3D Inception-based convolutional neural network architecture, based on the idea of improved utilization of the computer resources inside the network[提升计算机资源利用率], first mentioned in for 2D case.

The main building block of the network is an Inception block (Fig. 5). To eliminate the need of choosing the specific layer type at each level of the network Inception block uses 4 different bands of layers simultaneously[同时选用4个不同层带]. Besides that, a number of 1 × 1 × 1 convolution filters are used to significantly reduce the number of network parameters[使用1×1×1可降低模型参数] by decreasing the dimension of the feature space. In particular, the first band of the block performs a two successive 3 × 3 × 3 convolutions (equivalent to 5 × 5 × 5 filter)[第1波], the second band performs one 3 × 3 × 3 convolution[第2波], third band performs a max-pool operation[第3波], fourth band performs 1×1×1 convolution[第4波]. Besides that, first three bands use 1×1×1 convolution at the beginning[前三波起初用1×1×1卷积]. Each convolution layer is followed with batch normalization layer[每层卷积层后添加批处理规范化层] and a ReLU[激活函数]. The number of features in each convolution depends on the input and is shown in Fig.5. Thus, the output of the Inception block increases the feature dimension of data in 1.5 times compared to its input[Inception块输出的数据维度是输入的1.5倍]. All these tricks substantially reduce the number of parameters inside the network, while at the same time batch normalization layers accelerate network training.[以上优化可以极大地减少模型参数，批处理规范化层可以加快网络训练]



A preliminary 3 × 3 × 3 Conv block with the 4 sequent combinations of Inception block with 3D max-pooling layer form a pipeline of the proposed network architecture[第一层（含4波）构成网络模型的管道]. This pipeline transforms the source spatial data to the feature space[管道目的：将空间数据转为特征空间]. The last modification to reduce the number of network parameters compared to the conventional AlexNet-like networks[对比网络] is to place a 3D average-pooling layer at the end of the pipeline instead of the fully-connected layer[与AlexNet-like网络相比，在末层添加3D平均池化层可减少网络参数]. For each ROI in the brain scan and for each modality we use a separate described above pipeline[对于大脑扫描ROI及模式独立采用管道]. Finally, all pipelines are concatenated and with the following dropout, fully-connected and softmax layers produce the classification result[管道拼接及Dropout,全连接层,softmax层生成分类结果] (Fig. 7). Thus the described network is a siamese network[孪生网络：共享网络权重]which performs the late fusion of the data from input ROIs[对输入的ROI的数据融合].

The usage of batch normalization as mentioned earlier allows us to speed up the network training process and according to eliminate the necessity of using the pretraining techniques[使用批处理规范化可以加快网络训练速度和节省预训练手段] (e.g. autoencoders). Batch normalization partially plays a role of regularization[BN含正则化作用] as it allows each layer of a network to be trained less dependent on other layers[使每层更少依赖于其它层].



Proposed network architecture:

