**调研3篇涉及3D Inception的文献如下：**

（1）Classification of Pancreatic Tumors based on MRI Images using 3D Convolutional Neural Networks

（2）3D Inception Convolutional Neural Networks For Automatic Lung Nodule Detection

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

**总结内容如下：**

1. 3D Inception应用领域
2. 3D Inception相比于传统Inception或CNN的优势
3. 3D Inception常见的网络结构
4. 3D Inception具体实现（查看第3-(3)节中Inception模块图）

**1. 3D Inception应用领域**

（1）基于胰腺磁共振图像(MRI)胰腺癌的计算机辅助诊断

（2）基于胸部计算机断层扫描(CT)预测和治疗肺癌

（3）基于结构和功能磁共振成像和正电子发射层析成像用于早期发现阿尔茨海默病和轻度认知障碍

**2.3D Inception相对于传统Inception/CNN的优势原因**

（1）现有的胰腺肿瘤分类方法存在半自动缺陷以及会忽略肿瘤时空特征，尝试采用3D版本的ResNet18等4个网络模型优化。

（2）使用3D-CNN更能适应3D-CT扫描图，解决梯度弥散问题和提升F1度量指标。

（3）采用基于3D Inception的CNN网络结构更加充分利用计算机资源。

原文摘录：

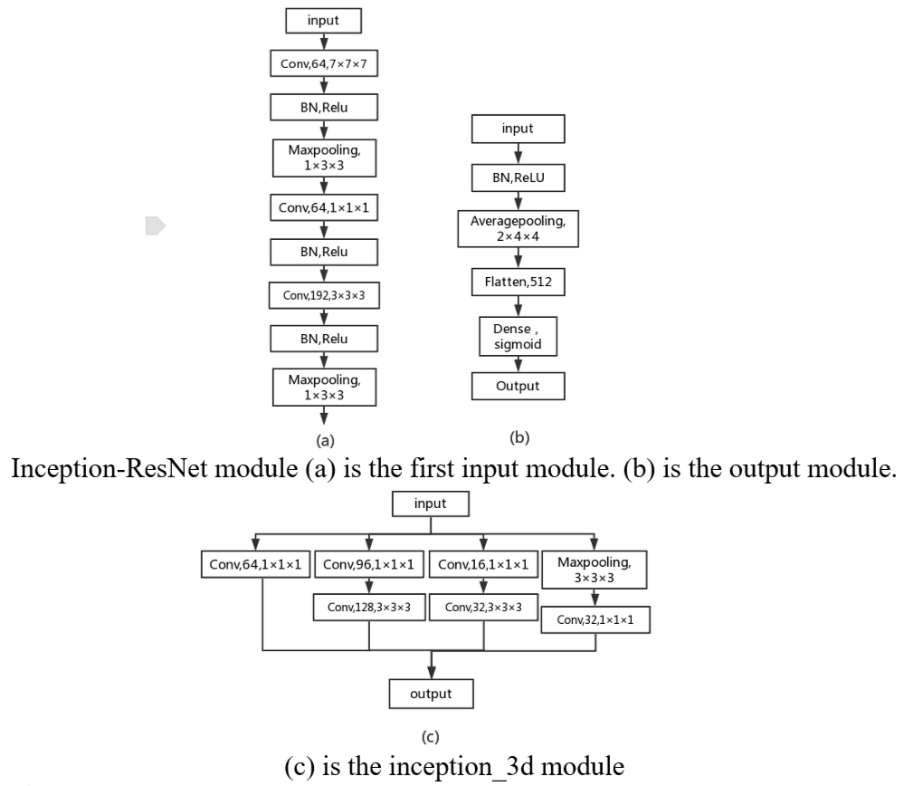
（1）原文：Existing pancreatic tumors classification methods suffer from the problem of partial automation and ignore spatial and temporal characteristics. In this paper, we used 3D versions of ResNet18, ResNet34, ResNet52 and Inception-ResNet for pancreatic magnetic resonance images (MRI) classification.

（2）原文：we propose the inception block for 3D convolutional neural networks to accommodate the 3D nature of CT scans, which solve the gradient vanish problems and enhance the F1 score.

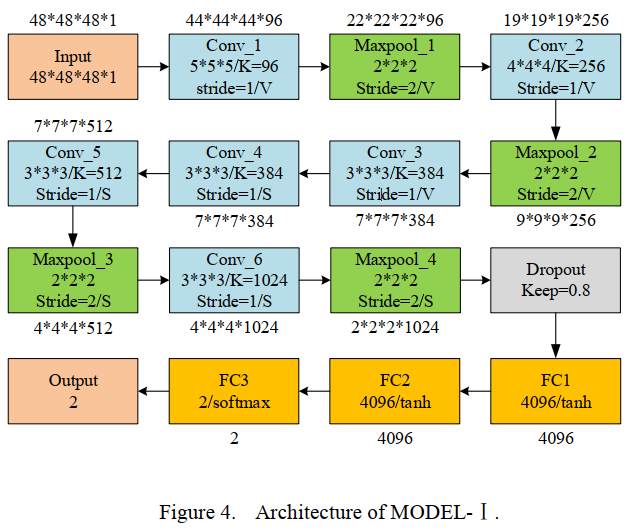
（3）原文：In this work we propose a new 3D Inception-based convolutional neural network architecture, based on the idea of improved utilization of the computer resources inside the network, first mentioned in for 2D case.

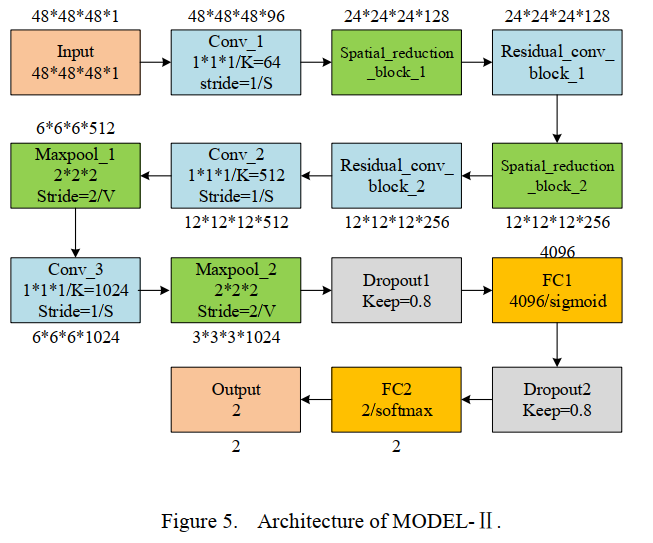
**3.3D Inception常见的网络结构**

（1）3D Inception-ResNet：包含1个a模块，9个c模块，1个b模型，并在第2和第7的Inception 3d模块后加入最大池化层（目的减少输出层维度），在最后一个Inception-ResNet模块后添加平均池化层（减少维度至15× 1× 1×256），连接含3840个神经隐藏单元的平坦层，再接入带有sigmoid函数的全连接层实现2分类预测。

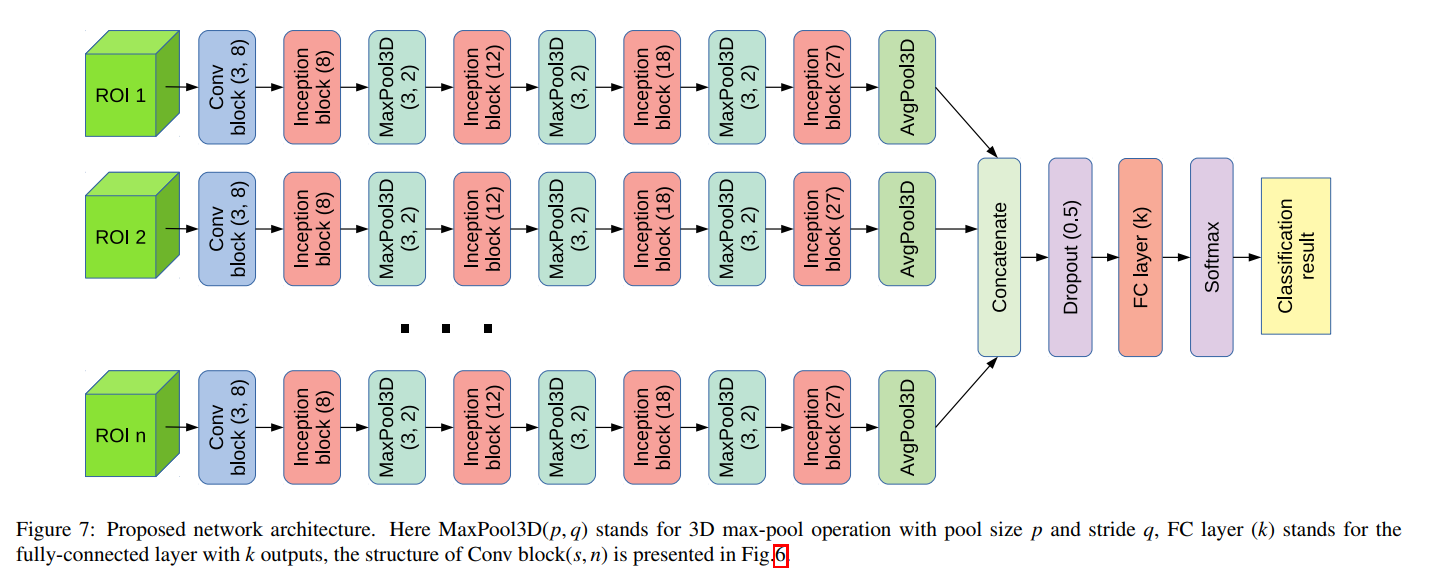


（2）3D CNN Network Structure：两种3D-CNN(MODEL-I & MODEL-II)，其中MODEL-I包含6层3D CNN 层，2层最大池化和3层全连接层(FC1,FC2,FC3)；MODEL-II结合了Inception的思想，包含残差卷积块和空间缩减块且均采用Inception结构，为避免过拟合，两个网络模型均使用Dropout层，且卷积层后拼接ReLu激活函数。

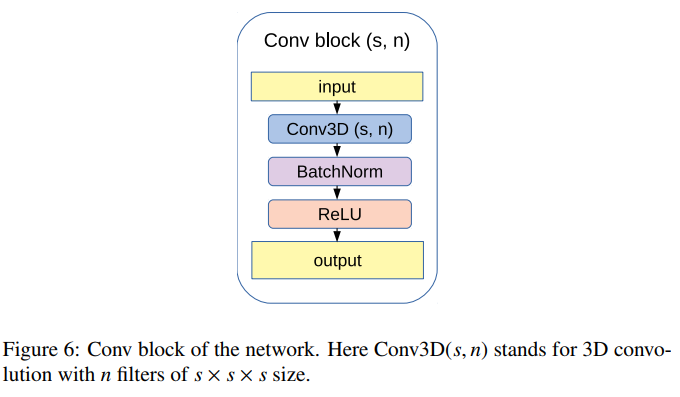




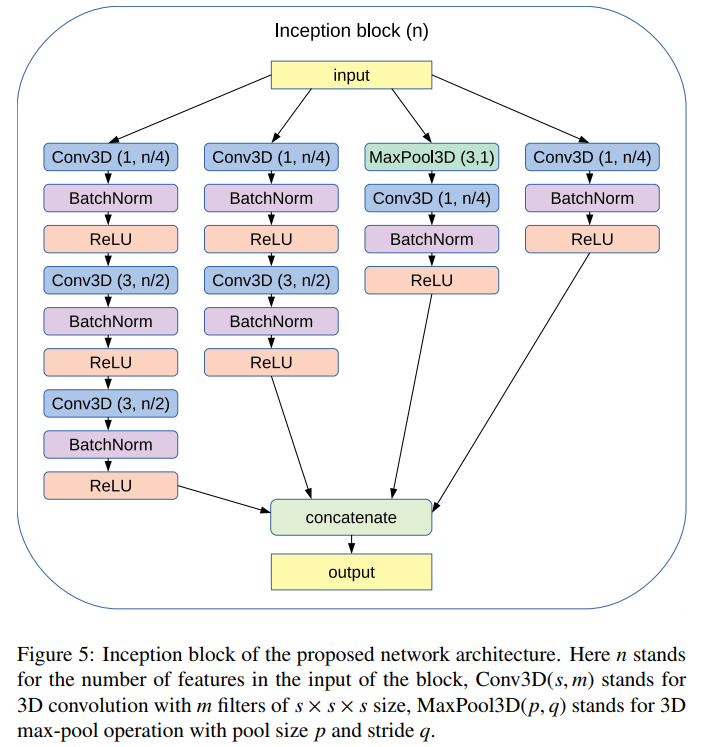
（3）3D Inception-based convolutional neural network：是一个孪生网络（网络间共享权重），其中卷积模块和Inception模块具体可查看后2图。（文中具体介绍了Inception模块的细节）



其中Conv块为：其中n表示输入特征数量



其中Inception块：



原文摘录：

（1）原文：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.

（2）原文：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.

（3）原文：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. Besides that, a number of 1 × 1 × 1 convolution filters are used to significantly reduce the number of network parameters 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), the second band performs one 3 × 3 × 3 convolution, third band performs a max-pool operation, fourth band performs 1×1×1 convolution. Besides that, first three bands use 1×1×1 convolution at the beginning. Each convolution layer is followed with batch normalization layer [48] and a ReLU. 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. All these tricks substantially reduce the number of parameters inside the network, while at the same time batch normalization layers accelerate network training.