



基于STN与异构卷积滤波器的肝硬化识别

张欢¹, 赵希梅^{1,2}

(1. 青岛大学 计算机科学技术学院, 山东 青岛 266071; 2. 山东省数字医学与计算机辅助手术重点实验室, 山东 青岛 266000)

摘要: 卷积神经网络因缺乏空间不变性造成分类精度不高, 且由于复杂度过高导致分类效率较低。提出一种利用空间变换网络和异构卷积滤波器的SH_ImAlexNet网络, 应用于肝硬化样本识别。改进卷积神经网络AlexNet的结构和参数以满足肝硬化样本尺度要求, 引入空间变换网络层增强特征提取能力与空间不变性, 采用异构卷积滤波器替换部分卷积核降低复杂度并提升鲁棒性。实验结果表明, 该网络的分类效果较AlexNet、VGG等传统网络更优, 在小样本数据集和大样本数据集上的识别率分别达到98.28%和95.67%, 空间复杂度和时间复杂度更低且运行效率更高。

关键词: 空间变换网络; 异构卷积滤波器; AlexNet模型; 卷积神经网络; 肝硬化识别

开放科学(资源服务)标志码(OSID):



中文引用格式: 张欢, 赵希梅. 基于STN与异构卷积滤波器的肝硬化识别[J]. 计算机工程, 2021, 47(5): 301-307, 315.

英文引用格式: ZHANG Huan, ZHAO Ximei. Identification of liver cirrhosis based on STN and heterogeneous convolution filter[J]. Computer Engineering, 2021, 47(5): 301-307, 315.

Identification of Liver Cirrhosis Based on STN and Heterogeneous Convolution Filter

ZHANG Huan¹, ZHAO Ximei^{1,2}

(1. College of Computer Science and Technology, Qingdao University, Qingdao, Shandong 266071, China;

2. Shangdong Provincial Key Laboratory of Digital Medicine and Computer Aided Surgery, Qingdao, Shandong 266000, China)

[Abstract] The lack of space invariance of Convolutional Neural Network (CNN) results in low classification accuracy and low classification efficiency due to its high complexity. This paper proposes a SH_ImAlexNet network based on Spatial Transformer Network (STN) and Heterogeneous Convolution (HetConv) filter for the identification of liver cirrhosis samples. The structure and parameters of CNN AlexNet are optimized to fit into the size of liver cirrhosis samples, and the STN layer is introduced to enhance the feature extraction ability and spatial invariance. The HetConv filter is used to replace part of the convolution kernel to reduce the complexity and improve the robustness. Experimental results show that the classification performance of the proposed network is better than that of traditional networks such as AlexNet and VGG. The recognition rate of the network on small sample dataset and large sample dataset reaches 98.28% and 95.67% respectively. It provides a higher operation efficiency with reduced space complexity and time complexity.

[Key words] Spatial Transformer Network (STN); Heterogeneous Convolution (HetConv) filter; AlexNet model; Convolutional Neural Network (CNN); identification of liver cirrhosis

DOI: 10.19678/j.issn.1000-3428.0057341

0 概述

自人工神经网络^[1]诞生以来, 计算机辅助诊断 (Computer-Aided Diagnosis, CAD)^[2]技术逐渐受到关注。随着卷积神经网络 (Convolutional Neural Network, CNN)^[3]的发展, 该技术在医学影像特别是肝脏超声影像方面的应用日益广泛。肝脏超声影像可有效检测出肝脏组织的微小病变, 操作安全便捷。目前CAD辅助肝脏超声影像技术已成为检测肝脏病灶的重要手段,

肝病检测中病灶图像特征的提取方式, 也由基于机器学习^[4]的人工提取发展为基于深度学习^[5]的卷积自动提取。

目前, 国内外研究人员采用机器学习或深度学习方法进行肝脏病变研究。在机器学习方面, 文献[6]利用空间灰度独立矩阵、空间频率分解和分形结合两层BP神经网络, 对正常肝脏以及轻度、中度和重度脂肪肝4种肝脏超声影像的平均识别率达到95.33%。文献[7]提出一种改进SLBP特征与二维Gabor变换

基金项目: 国家自然科学基金(61303079)。

作者简介: 张欢(1995—), 女, 硕士研究生, 主研方向为医学图像处理; 赵希梅, 副教授、博士。

收稿日期: 2020-02-07 修回日期: 2020-03-31 E-mail: ZH15092028723@163.com

结合的方法,采用超限学习机 ELM 对肝硬化样本的识别率达到 95.4%。在深度学习方面,文献[8]提出基于深度视觉特征学习的肝脏病灶识别方法,并采用 CaffeNet 训练框架对原发性肝癌、肝硬化样本以及正常肝脏进行识别,其平均识别精度为 96.67%。

由上述研究成果可知,与基于机器学习的人工提取方法相比,采用卷积自动提取特征的深度学习^[9]方法对肝脏病灶识别效果更优,因此,基于卷积神经网络的算法作为深度学习的重要算法^[10],逐渐成为肝脏影像处理领域的主流算法。然而,卷积神经网络也存在多种缺陷:卷积神经网络缺乏网络空间不变性,其对旋转、平移等操作输入特征的读取能力有限,造成网络分类效果较差且耗时较多;为提高网络分类精度并减少耗时,卷积神经网络不断加深,导致网络运行效率降低;传统卷积神经网络大部分采用同构内核执行卷积操作,其模型架构的堆叠易造成网络复杂度过高及分类效率较低。

为提高传统卷积神经网络的特征读取能力和分类精度、降低网络复杂度并提升运行效率,本文提出一种应用于肝硬化样本识别的 SH_ImAlexNet 网络。在改进 AlexNet 网络中加入空间变换网络(Spatial Transformer Network, STN)层增强空间不变性,在此基础上引入异构卷积滤波器降低网络复杂度,并将该网络与 AlexNet、VGG^[11]等传统深度学习网络的识别率及复杂度进行对比分析。

1 相关理论

1.1 空间变换网络

空间变换网络由 DEEPMIND 等人^[12]于 2015 年提出,其能将输入样本在空间进行对齐,以减少样本由于空间旋转、平移等几何变换对分类任务的影响,空间变换网络结构如图 1 所示。STN 以仿射变换矩阵为基础,允许神经网络学习输入样本或特征图执行空间变换的方式,以增强模型的几何不变性。

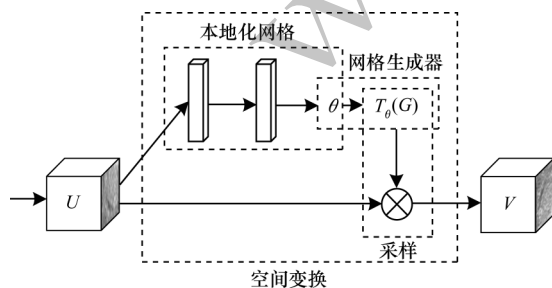


图 1 空间变换网络结构

Fig.1 Structure of spatial transformer network

空间变换模块是空间变换网络的核心,其主要包括本地化网络、参数采样网格和图像采样 3 部分。

1) 本地化网络

本地化网络的主要任务是确定输入所需变换的参数 θ 。将输入的特征图 $U \in \mathbb{R}^{(H \times W \times C)}$ 经过若干卷积或全连接操作后接入回归层,回归输出变换参数 $\theta = \int_{\text{loc}} (U), \theta \in \mathbb{R}^{2 \times 3}$ 。

2) 参数采样网格

参数采样网格主要利用本地化网络的输出参数 θ 对特征图进行仿射变换。根据输入和输出特征图的坐标位置得到仿射变换的特征关系 $T_\theta(G_i)$ 与结果 $V^t \in \mathbb{R}^{(H' \times W' \times C)}$ 。假设输入特征图 U 每个像素的坐标位置为 (x_i^s, y_i^s) , 经过空间转换后输出特征图每个像素的坐标位置为 (x_i^t, y_i^t) , 得到特征关系如下:

$$\begin{pmatrix} x_i^t \\ y_i^t \end{pmatrix} = T_\theta(G_i) = A_\theta \begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{21} & \theta_{22} & \theta_{23} \end{pmatrix} \begin{pmatrix} x_i^s \\ y_i^s \\ 1 \end{pmatrix} \quad (1)$$

其中, A_θ 为仿射变换矩阵。

3) 图像采样

图像采样是将经过本地化网络和参数采样网格得到的特征关系 $T_\theta(G_i)$ 与原始输入特征图 U 融合, 并经过以下对应关系生成最终结果 $V \in \mathbb{R}^{(H \times W \times C)}$:

$$\begin{cases} V_i^c = \sum_n^H \sum_m^W U_{nm}^c k(x_i^s - m; \phi_x) k(y_i^s - n; \phi_y) \\ \forall i \in [1, 2, \dots, H'W'] \\ \forall c \in [1, 2, \dots, C] \end{cases} \quad (2)$$

其中, $k(\cdot)$ 为采样内核, ϕ_x 和 ϕ_y 为输入参数, (n, m) 为特征图 U_{nm}^c 在通道 C 内的位置。

综上所述,特征图像 U 经过旋转、平移或拉伸等操作后,通过本地化网络得到仿射变换参数 θ , 在参数采样网格进行仿射变换,最终在图像采样处与原始特征图融合,得到具有空间不变性的新特征图像 V 。除了空间不变性外,空间变换网络还具有优化损失函数、计算速度快以及耗时少等特性。

1.2 异构卷积滤波器

随着卷积神经网络在计算机视觉和图像识别等领域的广泛应用,人们对卷积神经网络分类精度的要求越来越高。为提高识别率,采用同构滤波器^[13]的传统卷积神经网络深度不断增加,但导致网络复杂度逐渐增大。为了在提高网络分类精度的同时有效降低复杂度并提高网络运行效率,文献[13]提出一种采用异构内核进行卷积运算的深度学习模块,即异构卷积(Heterogeneous Convolution, HetConv)滤波器。

图2为异构卷积滤波器与同构卷积滤波器的结构。由于两者的主要区别在内核,因此异构内核是异构卷积滤波器的核心。异构内核通常由 3×3 的分组卷积^[14]和 1×1 的逐点卷积^[15]构成。如果同构卷积滤波器为 $3 \times 3 \times C$ (C 为输入通道数),则异构卷积滤波器会利用比例系数 P 保留 C/P 的 3×3 卷积核尺寸,剩余的 $(C - C/P)$ 个卷积核尺寸则变为 1×1 ,以此替换同构卷积滤波器。

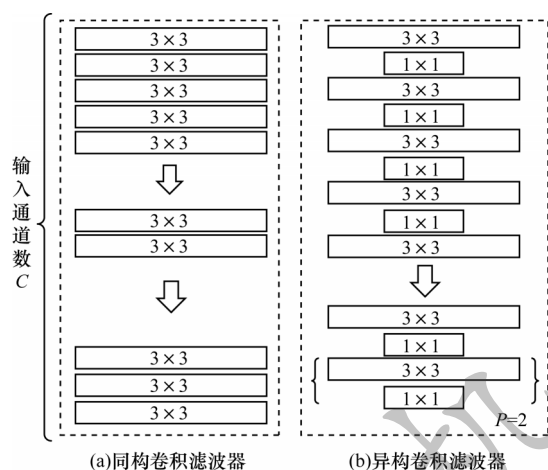


图2 异构卷积滤波器与同构卷积滤波器的结构

Fig.2 Structures of heterogeneous convolution filter and homogeneous convolution filter

如果一个 L 层的卷积核输出通道数为 D ,且每个通道均为 3×3 和 1×1 的异构内核,此时若比例系数 $P=4$,则异构卷积滤波器将会从该层的第一个滤波器的首位依次使用 3×3 和 1×1 的异构内核,如图3所示。

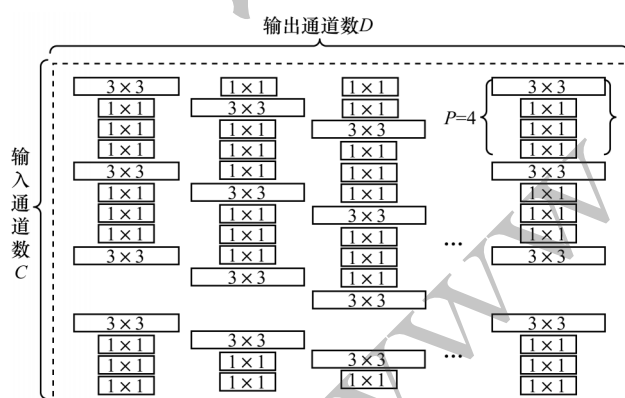


图3 L 层异构卷积滤波器结构

Fig.3 Structure of L -layer heterogeneous convolution filter

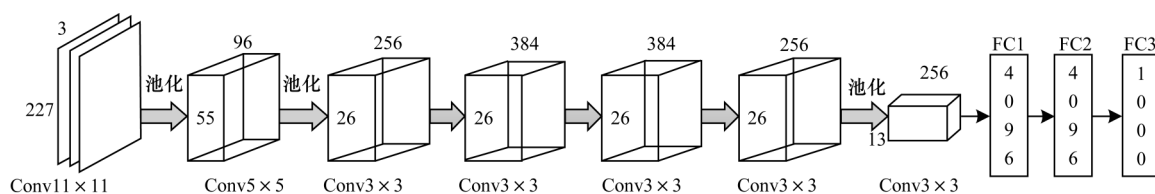


图4 AlexNet网络结构

Fig.4 Structure of AlexNet network

$K \times K$ 同构卷积滤波器的每秒浮点运算次数(FLOPS)的计算公式如下:

$$Fls = F_o \times F_o \times C \times D \times K \times K \quad (3)$$

其中, F_o 为卷积输出的特征图大小, C 为输入通道数, D 为输出通道数。

若将 L 层异构卷积滤波器的比例系数 P 视为一个整体,则 $K \times K$ (即 3×3)的内核只有 $(1/P)$ 个,含有 P 的 $K \times K$ 内核的FLOPS计算公式如下:

$$F_k = (F_o \times F_o \times C \times D \times K \times K) / P \quad (4)$$

实际上 $K \times K$ 内核个数为 (C/P) ,剩余 1×1 内核个数为 $(C - C/P)$,剩余 1×1 内核的FLOPS计算公式如下:

$$F_l = (F_o \times F_o \times D) \times (C - C/P) \quad (5)$$

L 层的异构卷积滤波器的FLOPS计算总量如下:

$$F_q = F_k + F_l \quad (6)$$

异构卷积滤波器与同构卷积滤波器相比,其减少计算量 F_h 为:

$$F_h = \frac{F_k + F_l}{Fls} = \frac{1}{P} + \frac{(1 - 1/P)}{K^2} \quad (7)$$

由式(7)可知,当 P 为1时,异构卷积滤波器也是同构卷积滤波器。

综上所述,异构卷积滤波器通过将一部分通道的同构卷积滤波器尺寸保留为 3×3 ,将另一部分通道的同构卷积滤波器尺寸减少为 1×1 ,从而确保覆盖前者的空间相关信息,并缩小后者的空间范围,同时得到相同或高于同构卷积滤波器的分类精度。此外,异构卷积滤波器直接插入卷积神经网络可降低网络复杂度。

1.3 AlexNet网络

2012年复杂卷积神经网络 AlexNet^[16]在 ImageNet 竞赛^[17]中夺冠,与 LeNet5^[18]网络相比, AlexNet 网络结构更深,图像识别效果更出色^[19],其结构如图4所示。 AlexNet 网络分别通过最大池化(MaxPool)、ReLU 激活函数、Dropout 函数以及数据增强处理,对浅层神经网络进行优化,避免训练时出现过拟合的现象,最终得到全连接(FC)层。此外, AlexNet 网络还具有传统卷积神经网络空间不变性的特点。由于肝硬化样本纹理信息复杂多样,且传统卷积神经网络在经过卷积(Conv)-池化训练后其特征会改变,因此尽管 AlexNet 网络能在自然图像识别上取得较好的识别效果,但仍会受样本图像和自身特性影响,出现网络识别率与运行效率降低的现象。

2 本文网络

2.1 改进的 AlexNet 网络

为提高图像识别率,本文对 AlexNet 网络结构和参数进行改进以适应肝硬化样本的尺寸,并将改进后的 AlexNet 网络记为 ImAlexNet 网络,其主要由 4 个最大池化层以及 3 个全连接层以及 7 个卷积层组成,卷积层中 6 个含有批标准化(Batch Normalization, BN)层。ImAlexNet 网络结构参数如表 1 所示(其中“—”表示该值不存在)。输入样本为 3×56 像素 $\times 56$ 像素(3 为通道数,下同),采用 64 个 3×3 的卷积核提取特征,然后采用 ReLU 函数作为激活函数进行处理,并经过最大池化层得到下一层卷积的输入为 64×14 像素 $\times 14$ 像素(64 为卷积核数,下同)。对于含有 BN 层的卷积层,在卷积提取完特征后,需要经过 BN 层的归一化处理,再采用 ReLU 激活函数进行训练。改进后的 AlexNet 网络采用尺寸为 3×3 的卷积核。

表 1 ImAlexNet 网络结构参数

网络层类型	卷积核大小	卷积核数量	输出维度
输入	—	—	(None, 3, 56, 56)
Conv1	3×3	64	(None, 64, 28, 28)
MaxPool	2×2	—	(None, 64, 14, 14)
Conv2/BN	3×3	192	(None, 192, 14, 14)
Conv3/BN	3×3	192	(None, 192, 14, 14)
MaxPool	2×2	—	(None, 192, 7, 7)
Conv4/BN	3×3	218	(None, 218, 7, 7)
Conv5/BN	3×3	218	(None, 218, 7, 7)
MaxPool	2×2	—	(None, 218, 3, 3)
Conv6/BN	3×3	256	(None, 256, 3, 3)
Conv7/BN	3×3	256	(None, 256, 3, 3)
MaxPool	2×2	—	(None, 256, 1, 1)
FC1~FC3	—	—	(None, 512)

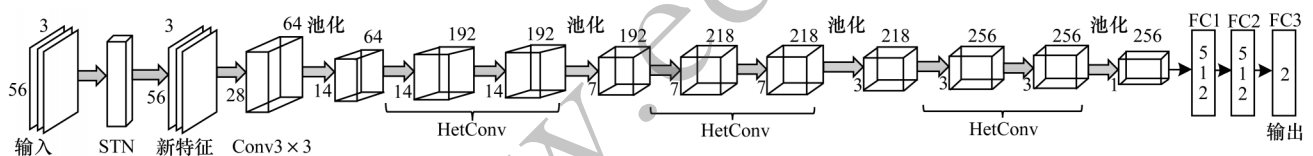


图 5 SH_ImAlexNet 网络结构

Fig.5 Structure of SH_ImAlexNet network

本文的 STN 层主要由本地化的 2 个卷积层(包括 8 个 7×7 卷积核和 10 个 5×5 卷积核)和 2 个全连接层(神经元数量分别为 32 和 6)组成。以改进的 ImAlexNet 网络为基础,在其输入层与第一个卷积层之间引入 STN,利用 STN 的仿射变换矩阵结构,对输入样本进行旋转生成并输出新的特征图像,从而增强卷积神经网络对样本输入特征的读取能力^[20],提高模型的空间不变性与分类精度。引入 STN 后,保留第一个卷积层和池化层的所有参数,将其余各层的卷积核替换为异构卷积滤波器,训练时每进行一次异构卷积就执行一次池化降维。根据 1.2 节中异构内核划分标准,设置比例系数 $P=2$,划分异构内核中

2.2 SH_ImAlexNet 网络

由于 AlexNet 网络对自然图像识别效果良好,因此本文以 AlexNet 网络为基础,在考虑了肝硬化纹理信息多样性的情况下,对 AlexNet 网络进行改进(见 2.1 节),所得 ImAlexNet 网络更适合肝硬化样本训练。此外,由于卷积神经网络在训练时不能保证网络的空间不变性,因此为提高网络分类精度,本文引入空间变换网络。该网络有较强鲁棒性,能动态地对样本执行空间变换,增强卷积神经网络平移、旋转及拉伸后的空间不变性,还能对样本的目标区域进行定位和优化^[20],以增强样本的分类效果。

此外,为确保提高分类精度和运行效率并降低网络复杂度,本文在引入空间变换网络的基础上融合异构卷积滤波器,以增强网络整体有效性。异构卷积滤波器采用比例系数 P 调整异构内核中 3×3 和 1×1 的卷积核数目,利用调整后的异构内核替换传统卷积神经网络的同构内核。异构内核的一部分通道在保留同构卷积滤波器尺度(3×3)的同时,也保留其空间的相关信息,另一部分通道则通过缩减同构卷积滤波器尺度来减少滤波器的空间范围,使得网络能在确保网络分类精度的情况下,提高网络运行效率并减少 FLOPS 计算量与网络参数量,从而提高网络整体有效性。本文融合 STN、HetConv 和 ImAlexNet 的优势,提出 SH_ImAlexNet 网络,其结构如图 5 所示。该网络主要由 1 个空间变换网络层、7 个卷积层(1 个 3×3 卷积层、6 个异构卷积层)以及 3 个全连接层组成。

各个 3×3 和 1×1 卷积核的数量。由样本采用 3 通道可知,最终得到的 3×3 和 1×1 卷积核数量一致,分别占通道总量的 $1/2$ 。本文提出的 SH_ImAlexNet 网络中各层卷积核的数量和 ImAlexNet 网络相同(见表 1),与原始 AlexNet 网络相比,其可有效降低网络复杂度,并增强网络的鲁棒性。

本文以 3×56 像素 $\times 56$ 像素的肝硬化超声影像样本为输入,通过空间变换网络生成 3×56 像素 $\times 56$ 像素的新特征,以新特征为输入对所保留的卷积层和池化层进行卷积-池化操作,获得维度为 $64 \times 14 \times 14$ 的输出特征。将该特征送入异构卷积神经网络,经过异构卷积、BN 层、ReLU 函数与最大

池化的多次交叠运算,最终得到全连接层输入维度为 $256 \times 1 \times 1$ 。

在网络训练中,先采用交叉熵函数 $H(p, q)$ 表示与期望值之间的差异,再利用Adam优化器优化网络结构,最后通过Softmax分类器对所训练样本的每类概率进行预测。假设 x_i 为训练样本 x 的第 i 个样本, $y \in \{1, 2, \dots, Y\}$ 为训练样本的每类概率(本文中 y 为2),则交叉熵函数表达式为:

$$H(p, q) = - \sum_{i=1}^m p(x_i) \ln(q(x_i)) \quad (8)$$

其中, $p(x_i)$ 为预测的概率值, $q(x_i)$ 为真实的概率值。

采用Adam优化器进行网络优化的相关计算公式如下:

$$M_t = \delta \times M_{t-1} + (1 - \delta) \times g_t \quad (9)$$

$$N_t = v \times N_{t-1} + (1 - v) \times g_t^2 \quad (10)$$

$$m_t = \frac{M_t}{1 - \delta^t} \quad (11)$$

$$n_t = \frac{N_t}{1 - v^t} \quad (12)$$

$$\Delta \theta_t = - \frac{m_t}{\sqrt{n_t} + c} \times \eta \quad (13)$$

其中, η 为学习率, c 为常数, m_t 为对梯度的一阶矩阵估计, n_t 为对梯度的二阶矩阵估计 N_t 的校正(用于理想图像的无偏差估计), $\Delta \theta_t$ 对学习率形成有明确范围的动态约束。

全连接层Softmax分类器的相关计算公式如下:

$$\text{Softmax}(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (14)$$

其中, j 为训练样本的种类个数。

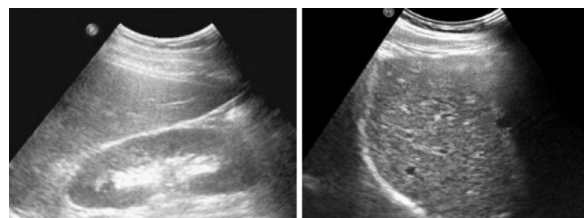
3 实验与结果分析

3.1 实验环境

本文实验基于pytorch1.2.0框架在Anaconda3环境的Spyder3.4平台进行,采用Win10 64位操作系统,Inter®Xeon™W-2133处理器,64 GB内存。通过R2018b版本的Matlab软件提取样本,并使用tensorboardX库、torchstat库和matplotlib库对实验结果进行可视化处理。

3.2 数据集及数据预处理

本文实验所用数据集是从青岛大学附属医院肝胆科选取的多幅正常肝脏和肝硬化样本的超声影像(部分肝脏超声影像示例见图6)。由于肝硬化超声影像的病变部位纹理受样本尺度影响较明显,为避免使用尺度较小样本(如16像素×16像素样本和28像素×28像素样本)造成纹理丢失,以及尺度较大样本(如128像素×128像素样本和224像素×224像素样本)包含过多胆囊等其他腹部组织的纹理信息导致肝硬化病灶识别准确率降低,本文采用matlab提取所选超声影像的感兴趣区域(Regions of Interest, ROI),得到1 200张初始样本,大小为56像素×56像素,提取后的部分肝脏感兴趣区域图像样本如图7所示。

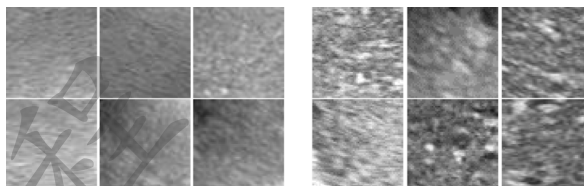


(a)正常肝脏

(b)肝硬化样本

图6 肝脏超声影像示例图

Fig.6 Example diagrams of liver ultrasound images



(a)正常肝脏

(b)肝硬化样本

图7 部分肝脏感兴趣区域图像样本

Fig.7 Partial image samples of region of interest of liver

为确保实验的有效性与真实性,对现有的1 200张样本进行归一化处理,以增强图像的对亮度。同时,对部分样本分别以 5° 、 10° 、 15° 和 20° 进行逆时针旋转以增强数据,得到的实验样本共2 480张。本文将样本分为S1和S2两组进行组合实验。S1组共880张,将其中320张正常肝脏样本和320张肝硬化样本作为训练集,120张正常肝脏样本和120张肝硬化样本作为测试集。S2组共1 600张,将其中300张正常肝脏样本和300张肝硬化样本作为训练集,将500张正常肝脏样本和500张肝硬化样本作为测试集。

3.3 实验参数和评价指标

本文采用AlexNet和VGG11两种网络与本文网络进行对比分析。为使对比网络符合肝脏样本尺度(3×56像素×56像素)的需要,在不改变原始网络结构的基础上优化这两种网络的参数,优化后其步长为1,卷积核大小均为3×3,卷积核个数分别为64、192、218和256,全连接层的神经元个数均为512。本文采用S1和S2两组样本进行组合实验。实验参数设置为:迭代次数为100,每次迭代输入样本的批尺寸为16,学习率初始值设置为0.001,训练时每迭代7次调整1次学习率,此时gamma取0.1。

将改进的AlexNet和VGG11与STN、HetConv相融合,分别得到STN_AlexNet、STN_VGG、SH_AlexNet和SH_VGG。将改进的AlexNet、VGG11和上述4种网络,与本文的ImAlexNet、SH_ImAlexNet以及融合STN的STN_ImAlexNet,从网络识别率和复杂度两方面进行对比分析。

网络识别率定义如下:

$$\text{Accuracy} = \frac{n_{\text{correct}}}{n_{\text{total}}} \times 100\% \quad (15)$$

其中, n_{correct} 为正确分类的样本数量, n_{total} 为样本总数量。

网络复杂度分为时间复杂度和空间复杂度。时间复杂度指模型的运算次数,主要体现实际运行时

CPU的运算能力。

单个卷积层的时间复杂度为:

$$\text{Time} \sim O(M^2 \cdot K^2 \cdot C_{\text{in}} \cdot C_{\text{out}}) \quad (16)$$

其中, M 为每个卷积核输出的特征图大小, K 为卷积核大小, C_{in} 为输入通道数, C_{out} 为输出通道数。

卷积神经网络整体的时间复杂度为:

$$\text{Time} \sim O\left(\sum_{l=1}^D M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l\right) \quad (17)$$

其中, l 表示网络深度为 D 的第 l 层, C_{l-1} 为上一层的卷积核个数(即输出通道数), C_l 为第 l 个卷积层的卷积核个数。

空间复杂度主要包括总参数量和各层输出的特征图,其表达式如下:

$$\text{Space} \sim O\left(\sum_{l=1}^D K_l^2 \cdot C_{l-1} \cdot C_l \cdot \sum_{l=1}^D M^2 \cdot C_l\right) \quad (18)$$

由式(18)可知,空间复杂度的总参数量与卷积核大小 K 、通道数 C 以及层数 D 有关,而各层输出的特征图仅与输入数据的大小 M 以及通道 C 有关。

由上述分析结果可知,时间复杂度和空间复杂度分别决定网络的训练耗时和参数量。如果网络复杂度过高,则会造成网络训练耗时和参数量过多,使训练网络所需数据量增加,并在样本较少时导致网络出现过拟合现象。

3.4 结果分析

由于 ImAlexNet 是构成 SH_ImAlexNet 的基础网络,其复杂度和识别效果对 SH_ImAlexNet 有直接影响,因此先对 ImAlexNet 与原始网络(即优化后的 AlexNet 和 VGG)进行分析,实验结果如表 2 所示。可以看出,在样本一致且不考虑训练耗时情况下,ImAlexNet 的识别率略高于原始网络。若考虑训练耗时,当样本一致时,ImAlexNet 的训练耗时最多为 AlexNet 的 2 倍、VGG 的 1.3 倍,但当样本不一致时,在 S1 组和 S2 组样本中 ImAlexNet 的识别率仅较 AlexNet 分别提高 3.75 个百分点和 5.00 个百分点。

表 2 ImAlexNet 与原始网络的识别率对比

Table 2 Comparison of recognition rates between ImAlexNet and original networks

网络	S1 识别率 /%	S1 训练耗时 /min	S2 识别率 /%	S2 训练耗时 /min
AlexNet	87.50	1.41	88.67	2.37
VGG	87.19	2.00	85.83	2.59
ImAlexNet	91.25	2.80	93.67	3.10

由表 2 的分析结果可知,ImAlexNet 的识别效果并不理想,因此,为增强图像的特征表示、优化分类效果以及减少训练耗时,将 STN 与表 2 中的网络分别进行融合,实验结果如表 3 所示。可以看出,在样本一致时,

STN_ImAlexNet 的识别率略高于 STN_AlexNet 和 STN_VGG,其时间损耗比 STN_AlexNet 更少,与 STN_VGG 的耗时差距较小。与表 2 中 AlexNet 和 ImAlexNet 相比,STN_AlexNet 的识别率显著提升,STN_ImAlexNet 的识别率也小幅提升。与表 2 中 VGG 相比,STN_VGG 的耗时更多且识别率更低,整体识别效果较差。综上可知,STN 虽然可以提高网络分类效果,但对部分网络存在识别效果不稳定的现象,易导致识别率过低。

表 3 融合 STN 后不同网络的识别率对比

Table 3 Comparison of recognition rates of different networks after fusion of STN

网络	S1 识别率 /%	S1 训练耗时 /min	S2 识别率 /%	S2 训练耗时 /min
STN_AlexNet	91.72	2.80	91.17	3.90
STN_VGG	84.53	2.31	84.00	3.31
STN_ImAlexNet	92.81	2.33	93.83	3.38

由表 2 和表 3 的分析结果可知,ImAlexNet 与 STN_ImAlexNet 的识别率不高,网络整体识别效果较差。为优化网络结构,提高其识别率与运行效率,本文提出融合 ImAlexNet、STN 和异构卷积滤波器(即 SH_ImAlexNet),并将其与融合 STN 和 HetConv 的 AlexNet、VGG(即 SH_AlexNet 和 SH_VGG)进行对比分析,实验结果如表 4 所示。可以看出,SH_ImAlexNet 对 S1 和 S2 两组样本进行训练后的识别率均显著高于 SH_AlexNet 和 SH_VGG。训练 S1 组样本时 SH_ImAlexNet 的耗时最低,而在训练 S2 组样本时,SH_ImAlexNet 的耗时虽少于 SH_AlexNet,但仍略高于 SH_VGG。

表 4 融合 STN 和 HetConv 后不同网络的识别率对比

Table 4 Comparison of recognition rates of different networks after fusion of STN and HetConv

网络	S1 识别率 /%	S1 训练耗时 /min	S2 识别率 /%	S2 训练耗时 /min
SH_AlexNet	92.03	2.50	88.83	3.80
SH_VGG	92.34	2.45	89.67	3.40
SH_ImAlexNet	98.28	2.40	95.67	3.49

由上述分析可知,本文提出的 SH_ImAlexNet 具有较高的识别率,其训练 S1 组样本的耗时较 ImAlexNet 更少,与 STN_ImAlexNet 的耗时差距较小。由网络识别率和训练耗时分析结果可知,SH_ImAlexNet 具有一定的鲁棒性。

为验证 SH_ImAlexNet 的有效性,本文将其与其他 6 种网络在复杂度上进行对比,实验结果如表 5 所示。可以看出,SH_ImAlexNet 的总参数量更少,从而可知其空间复杂度更低。SH_ImAlexNet 的时间

复杂度仅高于SH_AlexNet,然而SH_AlexNet的总参数量过多,造成网络空间复杂度较高,且其对S1组和S2组样本训练所得识别率和耗时均不如SH_ImAlexNet。综合考虑网络识别效果、空间复杂度和时间复杂度可知,SH_ImAlexNet识别率较高,空间复杂度较低,可有效避免网络出现过拟合现象,且时间复杂度也较低。由此可见,SH_ImAlexNet具有一定的有效性和鲁棒性。

表5 不同网络的复杂度对比

Table 5 Complexity comparison of different networks

网络	总参数量	时间复杂度/ 10^6
AlexNet	3 061 372	100.70
VGG	3 422 970	216.99
ImAlexNet	2 739 938	138.31
STN_AlexNet	71 294 114	2 400.00
STN_VGG	3 458 394	265.90
STN_ImAlexNet	2 775 368	142.21
SH_AlexNet	2 446 242	64.21
SH_VGG	2 280 680	166.52
SH_ImAlexNet	1 864 464	89.24

4 结束语

本文在改进 AlexNet 网络的基础上,提出一种应用于肝硬化样本识别的 SH_ImAlexNet 网络。在改进 AlexNet 网络中增加空间变换网络层提高特征提取能力,同时引入异构卷积滤波器减少网络参数量并提升运行效率。实验结果表明,该网络具有一定的有效性和鲁棒性,分类效果较 AlexNet、VGG 等传统网络更优。然而其时间复杂度略高且两组样本的识别率变化不稳定,后续将结合剪枝算法和轻量化网络进行研究,进一步提高网络有效性和分类效果。

参考文献

- [1] YU Daoheng. Artificial neural network[J]. World Science and Technology Research and Development, 1996, 18(2): 47-49, 54. (in Chinese)
余道衡. 人工神经网络[J]. 世界科技研究与发展, 1996, 18(2): 47-49, 54.
- [2] CHEN Tao, TU Shaoxiong, WANG Haolu, et al. Computer-aided diagnosis of gallbladder polyps based on high resolution ultrasonography[J]. Computer Methods and Programs in Biomedicine, 2020, 185(3): 105-118.
- [3] ZHANG Shun, GONG Yihong, WANG Jinjun. The development of deep convolution neural network and its applications on computer vision[J]. Chinese Journal of Computers, 2019, 42(3): 453-482. (in Chinese)
张顺, 龚怡宏, 王进军. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报, 2019, 42(3): 453-482.
- [4] YANG Jianfeng, QIAO Peirui, LI Yongmei, et al. A review of machine-learning classification and algorithms[J]. Statistics and Decision, 2019, 35(6): 36-40. (in Chinese)
杨剑锋, 乔佩蕊, 李永梅, 等. 机器学习分类问题及算法研究综述[J]. 统计与决策, 2019, 35(6): 36-40.
- [5] SU Fu, LÜ Qin, LUO Renze. Review of image classification based on deep learning[J]. Telecommunications Science, 2019, 35(11): 58-74. (in Chinese)
苏赋, 吕沁, 罗仁泽. 基于深度学习的图像分类研究综述[J]. 电信科学, 2019, 35(11): 58-74.
- [6] CHEN Fei. The application research of neural network in ultrasound fatty image recognition[J]. Control and Automation, 2007, 23(12): 302-303, 278. (in Chinese)
陈菲. 神经网络在超声脂肪肝图像识别中的应用研究[J]. 微计算机信息, 2007, 23(12): 302-303, 278.
- [7] LEI Yiming, ZHAO Ximei, WANG Guodong, et al. Cirrhosis recognition based on improved LBP algorithm and extreme learning machine[J]. Computer Science, 2017, 44(10): 45-50. (in Chinese)
雷一鸣, 赵希梅, 王国栋, 等. 基于一种改进的 LBP 算法和超限学习机的肝硬化识别[J]. 计算机科学, 2017, 44(10): 45-50.
- [8] ZHAI Shengqing. OU Wenbo, YANG Yusi, et al. Hepatic lesion recognition based on deep visual feature learning[EB/OL]. [2020-12-22]. <https://www.mdpi.com/2071-1050/13/3/1224/htm>.
- [9] SHI En, LI Qian, GU Daquan, et al. Convolutional neural network model based on local feature[J]. Computer Engineering, 2018, 44(2): 282-286. (in Chinese)
施恩, 李骞, 顾大权, 等. 基于局部特征的卷积神经网络模型[J]. 计算机工程, 2018, 44(2): 282-286.
- [10] ZHOU Linyong, XIE Xiaoyao, LIU Zhijie, et al. Research on pooling method of convolution neural network[J]. Computer Engineering, 2019, 45(4): 211-216. (in Chinese)
周林勇, 谢晓尧, 刘志杰, 等. 卷积神经网络池化方法研究[J]. 计算机工程, 2019, 45(4): 211-216.
- [11] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2020-12-22]. <https://arxiv.org/abs/1409.1556>.
- [12] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[EB/OL]. [2020-12-22]. <https://arxiv.org/abs/1506.02025>.
- [13] SINGH P, VERMA V K, RAI P, et al. HetConv: heterogeneous kernel-based convolutions for deep CNNs[C]// Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2019: 4830-4839.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[EB/OL]. [2020-12-22]. <https://www.mdpi.com/2071-1050/13/3/1224/htm>.
- [15] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]// Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 1-9.

(下转第315页)

(上接第 307 页)

- [16] GU Shenshen, DING Lu, YANG Yue, et al. A new deep learning method based on AlexNet model and SSD model for tennis ball recognition[C]//Proceedings of 2017 IEEE International Workshop on Computational Intelligence and Applications. Washington D. C. , USA; IEEE Press, 2017; 32-39.
- [17] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [18] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[C]//Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [19] ZHAO Limeng, ZHU Bing, BAI Tao, et al. Human behavior recognition based on image recognition technology[J]. Industrial Control Computer, 2021, 34(2): 107-108, 111. (in Chinese)
赵丽梦, 朱冰, 白涛, 等. 基于图像识别技术的人体行为识别[J]. 工业控制计算机, 2021, 34(2): 107-108, 111.
- [20] CHENG Ran, SHI Jianfang. Research of convolutional neural network gesture recognition algorithm[J]. Electronic Design Engineering, 2021, 28(2): 179-184. (in Chinese)
程冉, 史健芳. 基于卷积神经网络的手势识别算法研究[J]. 电子设计工程, 2021, 28(2): 179-184.

编辑 宋 圆