

# 基于改进 VGGNet-16 网络新生儿胆红素水平测量方法

刘国玉, 王东颖, 侯桂军

(唐山市妇幼保健院, 河北 唐山 063000)

**摘要:** 针对目前新生儿胆红素含量升高造成的新生儿各种病症的弊端, 考虑到传统新生儿胆红素水平测试所带来的患者创伤与操作不便捷的问题, 在 VGGNet-16 神经网络的基础上, 提出了一种改进的 VGGNet-16 网络算法, 对已有的新生儿胆红素医疗记录数据进行分类研究, 不同的类别具备一定范围的胆红素水平值; 经过实际验证, 该算法在克服传统新生儿胆红素水平测定不便捷的同时, 保证了新生儿胆红素水平测试的高准确性, 平均准确率达 95.56%, 为新生儿胆红素水平测量方法提供了思路, 具有较高的推广应用价值。

**关键词:** 新生儿; 胆红素; VGGNet; 测量方法

## Neonatal Bilirubin Level Measurement Method Based on Improved VGGNet-16 Network

Liu Guoyu<sup>1</sup>, Wang Dongying, Hou Guijun

(Tangshan Maternal and Child Health Hospital, Tangshan 063000, China)

**Abstract:** In view of the shortcomings of various neonatal bilirubin levels caused by neonatal bilirubin, taking into account the problems of patient trauma and inconvenient operation caused by traditional neonatal bilirubin level test, an improved VGGNet-16 network algorithm was proposed to classify the existing neonatal bilirubin medical record data. Different categories have a certain range of bilirubin level values. After actual verification, the algorithm overcomes the inconvenience of traditional neonatal bilirubin level determination, and ensures the high accuracy of neonatal bilirubin level test. The average accuracy rate is 95.56%. It provides a way of thinking for the measurement method of neonatal bilirubin level. High promotion and application value.

**Keywords:** neonatal; bilirubin; VGGNet; measurement method

## 0 引言

在临床上, 新生儿黄疸是新生儿常见的症状之一, 主要病因是由于在新生儿体内胆红素 (Bilirubin) 代谢水平异常, 而造成新生儿血液中胆红素水平升高, 从而引起皮肤、粘膜及巩膜发黄的现象。一般情况下可以分为生理性和病理性两种<sup>[1-2]</sup>。正常情况下, 当新生儿血液中的胆红素浓度高于 2~3 mg/dL (34~51  $\mu$ mol/L) 时, 这些部分便会呈现肉眼可见的其他颜色。医学研究表明, 新生儿的发病率呈现逐年增加的趋势, 造成不可逆转的神经系统后遗症<sup>[3]</sup>, 给普通家庭带来了众多的灾难。国内外一系列研究表明, 较低浓度的胆红素足以影响到患者的神经系统、听力以及智力, 早产儿中特别常见。所以, 进行精确的检验, 辅以早且有效化的治疗, 能够在很大程度上防止高胆红素血症对患者的后遗症发生, 避免脑损伤的形成<sup>[4]</sup>, 如何高效且有效监测新生儿血液中的胆红素水平成为了首要的问题。

检验医学领域, 使用特定的仪器和经皮胆红素值测定两种主要的方法来测定新生儿血清总胆红素值, 前者测量方法虽然精确, 但是需要采集新生儿的血液, 造成有创伤口, 直接增加了新生儿的痛苦, 被父母所排斥; 后者测量方法虽然操作便捷, 一旦新生儿出院后, 无法实时在居家的环境下检测新生儿身体中较高的胆红素值, 无法做到及时有效的干预治疗<sup>[5-6]</sup>。因此需要选择一种合理的方法, 在不同的环境下均能实时便捷检测出新生儿血清中的胆红素水平, 在不造成新生儿创口的同时, 也要能保证检测的准确率。在哪吒保 APP 的启发下, 可以借助卷积神经网络 (CNN)<sup>[7]</sup>对新生儿图像特征进行识别, 与样本数据进行对比, 从而分析得出新生儿胆红素的含量, 实现便捷且高效地对新生儿胆红素水平的监控。

本文提出了一种改进的 VGGNet-16 网络<sup>[8]</sup>在新生儿胆红素水平准确性研究算法, 针对本院收集的新生儿胆红素特征照片以及医疗检测的记录, 将这些特征图片和胆红素含量指标进行分类, 然后再利用 VGGNet-16 网络的图像处理能力, 得到这些新生儿的特征, 借助全连接网络实现对特征图像的分类, 被分到某个类别的新生儿特征图像会具备此类特征图像的胆红素水平值范围, 给医疗人员提供一个有效的参考值。

收稿日期: 2019-11-18; 修回日期: 2019-12-30。

基金项目: 河北省卫生厅科研基金项目(20181331)。

作者简介: 刘国玉(1982-), 女, 河北唐山人, 硕士, 主治医师, 主要从事新生儿胆红素检测与诊断方向的研究。

### 1 算法构建

借助卷积神经网络对图像的特征提取，识别出新生儿胆红素照片的不同特征，然后使用全连接网络对这些不同的特征进行分类研究，从而给出测试的样本数据对应的类别胆红素水平的属性值。

#### 1.1 卷积神经网络

卷积神经网络的主要核心是卷积核，卷积核是由特定大小组成的特征加权小区域，可以有效对图像的特征进行提取，卷积核的几个重要的特征包括卷积核的大小，卷积核的通道数以及卷积核的运算法则，包括是否进行边缘填充和卷积核的步长。单纯从理论的角度看待卷积核的通道数，其直接作用于上层图像的特征矩阵，能得到相应通道数的特征图，一般情况下，提取得到的特征图越多，特征空间的维度会越大，表明学习到的图像特征越多，得到的最终结果更加准确。但是另一方面，会增加卷积模型的参数个数，导致模型更加复杂，训练所需的计算量也更加庞大，有时甚至导致过拟合的现象，因此并不是卷积核的数量越多越好，而是要根据图像的特征来确定卷积核的数量。

对于使用卷积核对图像特征的提取，需要设定卷积核的尺寸，假定卷积核的规格为  $w \times w$ ，设置步长为  $k$  个像素，处理的图像尺寸为  $(n_h \times n_w)$ ，不采用边缘填充的方法，将会得到一个大小为  $\frac{n_h - w + k}{k} \times \frac{n_w - w + k}{k}$  的特征图，如果卷积核的数量为  $m$  个，将会得到  $m$  个这样的特征图，卷积核运算的过程如图 1 所示。理论情况下，所采用的的卷积核尺寸越小，特征提取的质量也会越高，但是在实际的运用过程中，对于卷积核的选取还得依靠分析图像的特征决定。

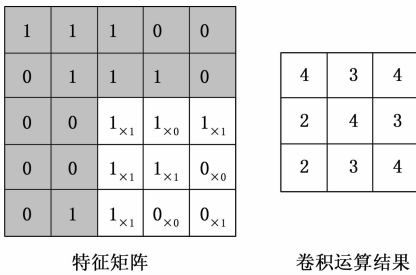


图 1 图像卷积示意图

从图 1 中可以看出，对于卷积的运算是将对应的数值与卷积核各位对应相乘并求取总和。经过卷积运算后的数值特征矩阵不能直接用于下一阶段的处理，还需要进行池化<sup>[9]</sup>操作。所谓的池化操作，就是对经过卷积核运算后的特征图小邻域内进行下采样得到新的特征，这些新的特征在减少模型参数的情况下（特征维度降低，参数也会降低），也增强了相关的特征，使得最终的特征能准确表征图像的某些不变性，比对旋转、伸缩和平移等特征，池化的过程本质上是一个降维的过程。常用的池化操作分为均值采样（mean-pooling）、最大采样（max-pooling），一

般情况下均值采样可以减小邻域大小受限造成的估计值方差增大造成的特征提取不准确的弊端，更能保留图像的背景信息；而最大采样可以减小卷积层参数误差造成估计均值的偏移，使得特征更能保留图像的纹理特征，由于本文在特征纹理上没有太多的要求，所以在池化的操作上采用均值采样（mean-pooling）操作。

#### 1.2 激活函数

为了使得神经网络模型具备非线性映射的特点，需要在神经网络模型中使用激活函数，常用的激活函数有 Sigmoid 函数、Tanh 函数、ReLU 函数等<sup>[10]</sup>，其中 Sigmoid 和 Tanh 函数在 BP 神经网络中运用广泛，而 ReLU 函数被广泛应用于深度学习中。假设一个神经单元的激活函数为  $h^{(i)}$ ，其中  $i$  表示隐含层单元的个数， $w^{(i)}$  表示隐含单元的权值，那么 ReLU 函数的表达式如公式（1）所示。

$$h^{(i)} = \max((w^{(i)})^T x, 0) = \begin{cases} (w^{(i)})^T x & (w^{(i)})^T x > 0 \\ 0 & \text{else} \end{cases} \quad (1)$$

ReLU 函数的表示如图 2 所示。

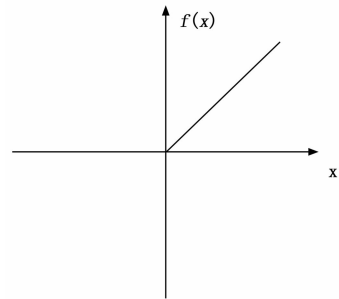


图 2 ReLU 函数的图像

经过大量的深度学习研究表明，ReLU 函数具有非饱和、线性的特征，能够对单侧的特征值进行抑制，对另一侧的边界兴奋，分散了神经单元的激活性，相比 Sigmoid 和 Tanh 函数<sup>[10]</sup>在卷积神经网络中性能优异。

#### 1.3 全连接神经网络

在对 CNN 网络提取的特征图像进行分类的时候，需要对已经提取得到的特征值进行全连接网络的分类处理，全连接网络的结构图如图 3 所示，其输入是已经得到的图像特征值，输出是处理分类的信息。

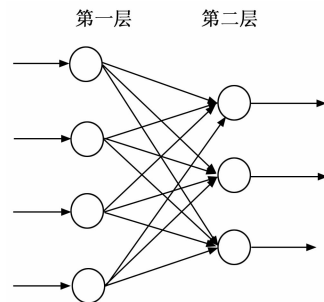


图 3 全连接网络层

图 3 中的隐藏层包含了 2 层的全连接神经网络模型，本文构造的 BP 神经网络由 1 个输入层，3 个隐藏层和 1 个输

出层组成。对于第二层神经网络中的每一个神经节点来说, 对于第一层输入的数据, 经过线性变换, 线性变换的公式为:

$$h_l^2 = \sum_{j=1}^9 x_{ij} * w_{ij} + b_l \quad (2)$$

其中:  $h_l^2$  表示第二层的第  $l$  个节点参数的运算,  $x_{ij}$  表示第  $i$  个训练数据的第  $j$  个特征值,  $w_{ij}$  表示第一层神经网络的第  $j$  个节点与第二层第  $l$  个节点的连接参数,  $b_l$  表示第二层第  $l$  个节点的偏置值。

若隐藏层不采用激活函数, 构造的全连接神经网络模型将会是一个线性模型, 只是简单具备数据的线性拟合能力, 无法具备逼近真实的分类函数能力, 本文选用的激活函数为 Sigmoid 激活函数<sup>[10]</sup>, 公式如 (3) 所示。

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

公式 (2) 可以将线性转化的实数转化到 0~1 之间的输出, 具体来说也就是将越大的负数转化到越接近 0, 越大的正数转化到越接近 1。

#### 1.4 改进的 VGGNet-16 网络算法

使用的 CNN 神经网络模型中流行的 VGGNet-16 作为网络的主体, 因为其识别的精度较高, 适合用于图像特征的识别, 在此之上进行改进, 将最大化池化原则更换为平均池化原则, 主要是为了更多发现新生儿照片中的背景颜色特征, 降低图像纹理特征的干扰。如图 4 所示, 是本文选用的网络模型。

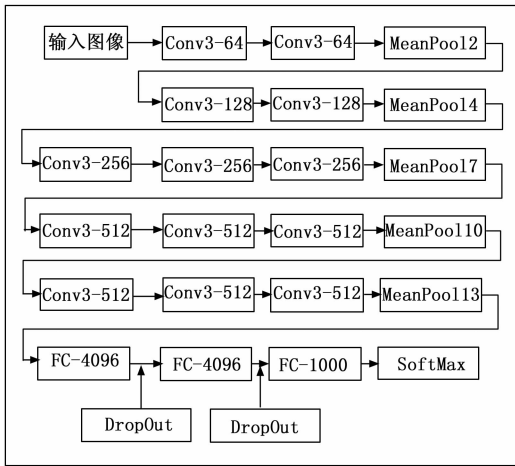


图4 算法采用模型

如图 4 所示, “Conv3-64” 表示采用的是尺寸为  $3 \times 3$  的卷积神经网络, 深度为 64, “MeanPool2” 表示采用的是平均池化的策略, 而且池化层的尺寸大小为  $2 \times 2$ , 这样可以有效降低特征维度的长宽尺寸; 进一步分析图 4 可以得到, 采用这样的神经网络模型, 深度越来越深, 而特征维度的长和宽在减少, 有利于提取新生儿胆红素图像的更多特征。

采用 VGGNet-16 网络应用于图像分类任务时, 需要在后面的全连接网络层连接一个 Softmax 分类器<sup>[11]</sup>, 用于生成对图像分类标签的预测, 结合本文的实际应用场景,

本文的模型中将会分为 6 类。在 Softmax 回归中, 解决的是多分类问题 (相对于 logistic 回归解决的二分类问题), 标签值  $y$  可以取  $k$  个不同的值, 每个标签是一个独立的类别。因此, 对于不同的训练集  $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$ , 最终分类的结果都会得到  $y^{(i)} \in \{1, 2, \dots, k\}$  的不同概率, 用  $p(y = j | x)$  表示样本相对不同分类标签的而不同概率, 输出是一个  $k$  维的具有分类信息的向量 (概率和为 1), 使用  $h_\theta(x)$  表示这  $k$  维分类特征, 其表达式如公式 (4) 所示。

$$h_\theta(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \dots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \dots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \quad (4)$$

公式 (4) 中使用  $\theta$  符号表示模型的全部参数,  $\theta$  是由  $k \times (n+1)$  的矩阵组成的, 该矩阵是通过  $\theta_1, \theta_2, \dots, \theta_k$  参数进行堆叠得到的, 如公式 (5) 所示。

$$\theta = \begin{bmatrix} \theta_1^T \\ \theta_2^T \\ \dots \\ \theta_k^T \end{bmatrix} \quad (5)$$

由上式子可得样本  $x^{(i)}$  属于  $j$  的概率公式如 (6) 所示。

$$p(y = y^{(i)} | x^{(i)}; \theta) = \frac{e^{\theta_{y^{(i)}}^T x^{(i)}}}{\sum_{i=1}^k e^{\theta_i^T x^{(i)}}} \quad (6)$$

当每个样本所属类别的条件概率  $p(y = y^{(i)} | x^{(i)}; \theta)$  都最大时, Softmax 的分类识度是最高的, 此时等价于最大化如公式 (7) 所示的似然函数。

$$L(\theta | x) = \prod_{i=1}^m p(y = y^{(i)} | x^{(i)}; \theta) \quad (7)$$

为了降低计算量和防止溢出, 对似然函数取对数, 并适当变形得到公式 (8)。

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^k e^{\theta_i^T x^{(i)}}} \right] \quad (8)$$

其中:  $1\{\cdot\}$  称为指示性函数, 表达的意思是, 当  $\{true\}$  时, 其值为 1, 否则为 0。于是最大化似然函数  $L(\theta | x)$  可以转化为最小化代价函数  $J(\theta)$  的问题, 这也是深度学习中常使用的转化方法。对于求解最小化  $J(\theta)$  的方法, 常用的就是梯度下降法, 给出使得  $J(\theta)$  满足极小值时参数  $\theta$  的值, 但是理论上极小值很难求解。代价函数  $J(\theta)$  的梯度如公式 (9) 所示。

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} (1\{y^{(i)} = j\} - p(y^{(i)} = j | x^{(i)}; \theta))] \quad (9)$$

在实际应用中, 本文的模型为了防止过拟合问题, 采用了正则化的思想, 在代价函数后加正则化项  $\frac{\lambda}{2} \sum_{i=1}^m \sum_{j=0}^n \theta_{ij}^2$  (L2 正则化), 这样代价函数变为如公式 (10) 所示,

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^m \sum_{j=0}^n \theta_{ij}^2 \quad (10)$$

在公式 (10) 中之所以加上第二项，是因为它会对偏大的权重值进行一定程度的惩罚，通常又称为权值衰减项，通过对超参数  $\lambda$  值的调控，可以降低权重值的数量级，防止训练模型过拟合。

使用反向传播算法进行梯度的运算，借助梯度信息对模型参数进行更新，主要的方法有随机梯度下降法<sup>[12]</sup> (stochastic gradient decent, SGD)，自适应矩估计法<sup>[13]</sup> (adaptive moment estimation, Adam)，本文使用的是 SGD 算法，通常情况下训练数据集会比较大，如何一次性装载所有训练样本进行训练，往往会出现内存溢出问题，因此随机抽取数据集的一个小样本数据集<sup>[14]</sup> (mini-batch，数量为  $N \ll |D|$  进行训练是很有必要的，具体的代价函数如公式 (11) 所示。

$$J(\theta) = \frac{1}{|N|} \sum_i^{N|} f_{\theta}(x^{(i)}) + \lambda r(\theta) \quad (11)$$

随机梯度下降法每次输入一个微型集 (mini-batch) 对网络进行训练，由于每次的微型集都是随机选取的，所以每次迭代的代价函数会不同，当前 batch 的梯度对网络参数的更新影响较大，为了减少这种影响，本文选用引入动量系数对传统的随机梯度下降法进行改进，更新的时候在一定程度上保留之前更新的方向，同时利用当前 batch 的梯度微调最终的更新方向，这样在一定程度上增加了模型的稳定性，促进模型学习地更快，具备一定的消除局部最优解的能力。加动量的随机梯度下降算法迭代公式如公式 (12) ~ (13) 所示。

$$V_{t+1} = \mu V_t - \eta \nabla J(\theta_t) \quad (12)$$

$$\theta_{t+1} = \theta_t + V_{t+1} \quad (13)$$

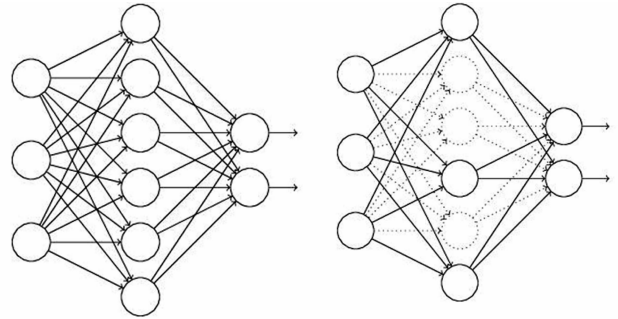
其中： $V_t$  是上一轮的权值更新量， $\mu$  为动量系数，表示要在多大程度上保留原来的更新方向，这个值在  $0 \sim 1$  之间， $\eta$  为学习率。值得注意的是，本文还使用了 Dropout 策略<sup>[2]</sup> 进行优化，Dropout 通过随机修改神经网络的结构来实现的，属于神经网络训练过程中的优化方法，核心思想是在训练过程中随机隐藏部分神经单元，使之不参与运算，保留它们的权重值，暂时不进行更新操作，其他神经单元进行正常的更新。在实际的训练过程中，Dropout 优化策略采用  $(1-p)$  的概率将隐含的神经单元输出值设置为 0，在反向传播更新权值阶段，不再更新与该节点相连的权值，Dropout 优化策略图如图 5 所示。

本文提出的改进的 VGGNet-16 网络新生儿胆红素水平准确性研究算法选用的 Dropout 的比例值是 0.5。

## 2 实际应用

### 2.1 数据准备

为了将已经采样得到的新生儿胆红素检测照片（主要是心脏部位的皮肤照片）与通过仪器设备测定的胆红素含



(a) 无Dropout的神经网络结构 (b) 设置Dropout的神经网络结构

图 5 Dropout 优化策略示意图

量进行分类，从已有的本院数据库中导出了 1 300 份有关的数据条例，借助文献 [15] 的部分理论结果，将数据分为 6 大类，每大类的具体信息如表 1 所示。

表 1 训练数据的分类表

类别	胆红素水平/( $\mu\text{mol/L}$ )	样本数量
类别 1	$\leq 51$	213
类别 2	$70 \pm 20$	195
类别 3	$110 \pm 20$	354
类别 4	$150 \pm 20$	268
类别 5	$190 \pm 20$	198
类别 6	$> 210$	72

将本院已有的新生儿医疗记录信息供 1 300 条数据按照表 1 的分类方式进行分类，然后按照训练集：测试集的 7:3 的比例按照交叉验证<sup>[17]</sup>的方法对没个类别进行数据的随机选取与训练，每个类别中剩下的 30% 比例的数据做为测试集进行精确分类的测试集，随机进行 10 次试验，取 10 次验证数据集的平均值作为验证的准确率，这样做的好处是可以充分利用小数量的训练数据集，克服模型泛化能力较差的弊端，取 10 次中验证数据最好的网络作为新生儿胆红素准确性研究的算法模型。

### 2.2 试验结果

经过一系列的试验进行验证，得到的 10 次的交叉验证的实验数据集如表 2 所示。

表 2 实验结果的验证精确度

次数	验证集数量	正确分类数量	分类准确率
1	390	371	95.13%
2	390	376	96.41%
3	390	381	97.69%
4	390	372	95.38%
5	390	374	95.90%
6	390	357	91.54%
7	390	362	92.82%
8	390	382	97.44%
9	390	378	96.92%
10	390	374	95.90%
平均准确率	——	——	95.56%