

基于Transformer的新生儿肺部超声图像诊断

张磊^{1*}, 陈胜^{1#}, 姚莉萍²

¹上海理工大学光电信息与计算机工程学院, 上海

²上海市第一妇婴保健院超声科, 上海

收稿日期: 2022年9月30日; 录用日期: 2022年11月28日; 发布日期: 2022年12月8日

摘要

新生儿肺超声是近几年发展起来的新技术, 也是对新生儿肺部疾病进行观察与诊断的关键技术。本研究研讨基于深度学习的分类模型, 对新生儿肺部超声图像自动进行阴性与阳性的分类, 回顾性搜集了2020年在上海市第一妇婴保健院出生的新生儿的肺部超声图像, 基于Pytorch的Vision Transformer框架, 并结合卷积神经网络构建模型。在多次实验后, 测试集的准确率达到了95%, 敏感性为98%, 特异性、精确率均达到了93%, ROC曲线的AUC达到了99%。通过卷积与Transformer相结合构建的分类模型, 对新生儿肺部超声图像的分类效果较好, 有助于医生对患者的进一步诊断与治疗。

关键词

新生儿肺超声, 深度学习, Transformer, 图像分类

The Diagnosis of Neonatal Lung Disease Based on Ultrasound Images Using Transformer

Lei Zhang^{1*}, Sheng Chen^{1#}, Liping Yao²

¹School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

²Ultrasound Department, Shanghai First Maternity and Infant Hospital, Shanghai

Received: Sep. 30th, 2022; accepted: Nov. 28th, 2022; published: Dec. 8th, 2022

Abstract

Neonatal lung ultrasound is a new technique that has been developed in recent years, and it is also

*第一作者。

#通讯作者。

a key technology to observe and diagnose neonatal pulmonary diseases. This study examines a classification model based on deep learning to automatically classify negative and positive lung ultrasound images of newborns. We collected the ultrasonic images of the lungs of newborns born in the Shanghai First Maternity and Infant Hospital in 2020. Based on the Pytorch and Vision Transformer, and combined with convolution neural network, a model is built. After many experiments, the accuracy of the test set is 95%, the sensitivity is 98%, the specificity and accuracy are 93%, and the AUC of the ROC is 99%. The classification model constructed by the combination of convolution and Transformer has a good classification effect on neonatal lung ultrasound images, which is helpful for doctors to further diagnose and treat patients.

Keywords

Neonatal Lung Ultrasound, Deep Learning, Transformer, Image Classification

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着现代科学与医疗技术的不断发展,医学图像已成为临床诊断中不可或缺的材料之一。相关医护人员通过使用现代高性能成像医疗仪器,对患者相关部位进行扫描,来获取该部位的生理结构信息以及与诊断密切相关的病变信息[1]。当前,使用较多的成像技术有电子计算机断层扫描(Computed Tomography, CT)、核磁共振成像(Magnetic Resonance Imaging, MRI)、超声波成像(Ultrasound, US)以及 X 射线等。

近年来,由于超声检查是一种便捷、可靠、及时的诊断技术,相比于计算机断层扫描、核磁共振、X 射线,超声检查已成为健康人群或者亚健康人群的初步筛查的首选,已被广泛用于临床诊断。与此同时,相关的医务人员也可避免因工作带来的射线损伤。超声检查在临床的应用包括对复杂临床图像的诊断与治疗,例如血液动力学不稳[2]、急性呼吸衰竭[3]或心脏骤停[4]等。

床旁肺部超声(Lung Ultrasonography, LUS)是一种超声诊断肺部疾病技术。关于肺脏超声的首次描述可以追溯到上世纪 60 年代[5],关于肺超声的基本体征在 30 年前得到了系统的描述[6] [7] [8] [9]。但是由于肺泡内充满气体,超声波遇到气体时会发生全反射,因此超声技术在肺部检查时一直得不到广泛应用,几年来随着技术的发展,这一技术瓶颈逐渐被打破,过去需要依赖胸部 X 射线、CT 检查的多种疾病,现在均可借助超声检查做出明确诊断,且具有更高的特异性、准确性。且超声本身的无放射性,床旁肺部超声技术近十多年来得到飞速发展与广泛应用。

近几年,机器学习技术得到迅速发展,其中的深度学习更是已成为热门研究领域,在图像处理、自然语言处理、机器翻译等方向已取得了很多成果。在医学方面,计算机辅助诊断也已经成为研究热点,其主要特点是提高人为诊断的工作效率与准确率。尽管基于机器学习的计算机辅助诊断技术实现了医学影像的分类,但这些医学影像需要人工提取特征再发送给分类器。而深度学习可以通过卷积神经网络自动提取特征,避免了人工提取特征这一过程,大幅度提升了工作效率。

本文实验先对医院采集的新生儿肺部超声图像进行预处理,再对 Vision Transformer 进行微调并将预处理的数据集输入至网络中,实现对新生儿肺部超声图像进行自动分类,分类结果为两组,为病情类(阳性)与健康类(阴性),可辅助医生诊断新生儿是否需要治疗。本文第二部分主要是对实验相关工作进行阐

述, 包括采样过程、数据集来源等。本文第三部分阐述模型结构; 第四部分描述实验具体内容与步骤; 第五部分为总结, 其中包括实验的结果, 不足点以及对未来的展望。

2. 关于数据集

本文实验选取了 2020 年在上海市第一妇婴保健院出生的 30 名新生儿, 日龄 0~3 天, 胎龄 27~40 周。本次实验通过上海市第一妇婴保健院认可, 且获得受试者监护人书面同意。

新生儿肺部疾病主要有: 新生儿呼吸窘迫综合征(neonatal respiratory distress syndrome, NRDS)、新生儿暂时性呼吸增快症(transient tachypnea of the newborn, TTN)、胎粪吸入综合征(meconium aspiration syndrome, MAS)。NRDS 是指由于各种原因引起肺泡表面活性物质(pulmonary surfactant, PS)原发或继发性缺乏, 出现肺不张, 导致新生儿出身不久后出现以进行性呼吸困难、青紫和呼吸衰竭为主要临床表现的严重肺部疾病, 该病情主要见于早产儿。TTN 又称新生儿湿肺(wet lung of the newborn), 主要高危因素有: 剖宫产分娩、出生体重大、母亲糖尿病等, 患儿主要表现为呼吸困难, 患儿出生时大作正常或有窒息史, 数小时(2~5 h)后出现呼吸困难, 轻者表现为呼吸增快, 重者表现为严重呼吸窘迫、吐沫、反应差。MAS 是因胎儿在宫内或产时吸入被胎粪污染的羊水而导致新生儿出生后出现以呼吸窘迫为主要临床表现的综合征, 主要表现为在出生数小时内出现呼吸困难, 且患儿皮肤、指甲、脐带等部位严重黄染。该病情常见于足月儿或过期产儿。总之, 当新生儿在出生几小时内出现呼吸困难时, 需要及时对其进行肺部超声检查。

医生使用的超声诊断仪是 GE LOGIC P5, 相应的探头为高频线阵探头, 频率为 10 MHz~12 MHz。新生儿处于安静状态下, 取仰卧位与侧卧位, 以腋前线、腋后线为基准, 将左右双肺各分为前、侧、后 3 个区域, 并将每个区域分为上下两部分, 总共 12 个区域。并根据从上而下, 先右后左, 从前到后的顺序依次一次扫查各个区域且储存规划范的图像。

对于新生儿肺部超声图像, 主要观察指标有胸膜线、A 线与 B 线[10]。胸膜线: 胸膜与肺表面的界面回声所形成的回声反射, 在正常情况下, 胸膜线光滑且厚度不超过 0.5 mm, 当胸膜线变粗糙、增厚或消失等, 为异常。A 线: 是胸膜下一系列间距相等且与胸膜线平行的线性高回声。B 线: 是起源于胸膜线的线性高回声, 且垂直于胸膜线并呈放射状向肺野深部发散。

正常情况下, 新生儿肺组织在超声下呈黑色低回声, 胸膜线呈现高回声, 光滑且清晰可见; A 线与胸膜线等间距平行排列, 呈现清晰、光滑且规则的线性高回声, 随着肺野深度增加回声强度逐渐降低; 同时没有 B 线或者仅有少数几条 B 线。如图 1 中(a)图所示。当新生儿有明显呼吸困难时, 显示的超声图像往往是胸膜线及 A 线模糊或消失, B 线清晰可见, 如图 1(b)~(d)所示, 其中, (b)图为患有 NRDS 的新生儿肺部超声图, (c)图为患有 TTN 的新生儿肺部超声图像, (d)图为患有 MAS 的新生儿肺部超声图像。下图 2 为超声下 NRDS 的恢复过程, 该新生儿在出生体重仅 1980 g, 因会呼吸困难出生 3 h 后入院, 此时肺部超声图像显示为图 A, 胸膜线异常且 A 线消失, 经过一定的治疗之后, 患儿情况有所好转, 肺部超声图像为图 C, 胸膜线恢复正常, 部分 A 线出现。

3. 模型结构

Transformer 在自然语言处理(Natural Language Processing, NLP)领域已取得了较大的成功。Vision Transformer (ViT)就是将 Transformer 在 NLP 领域的思想与架构运用至图像分类任务上。Alexey 等人[11]利用 ViT 网络先将一张图片分成多个小块的 patches, 再进行分类操作。本文没有将输入图像划分成若干个小块, 而是使通过使用卷积以及一些其他处理, 再输入至 Transformer 网络, 用于对新生儿肺部超声图像进行分类。模型结构如图 3 所示。

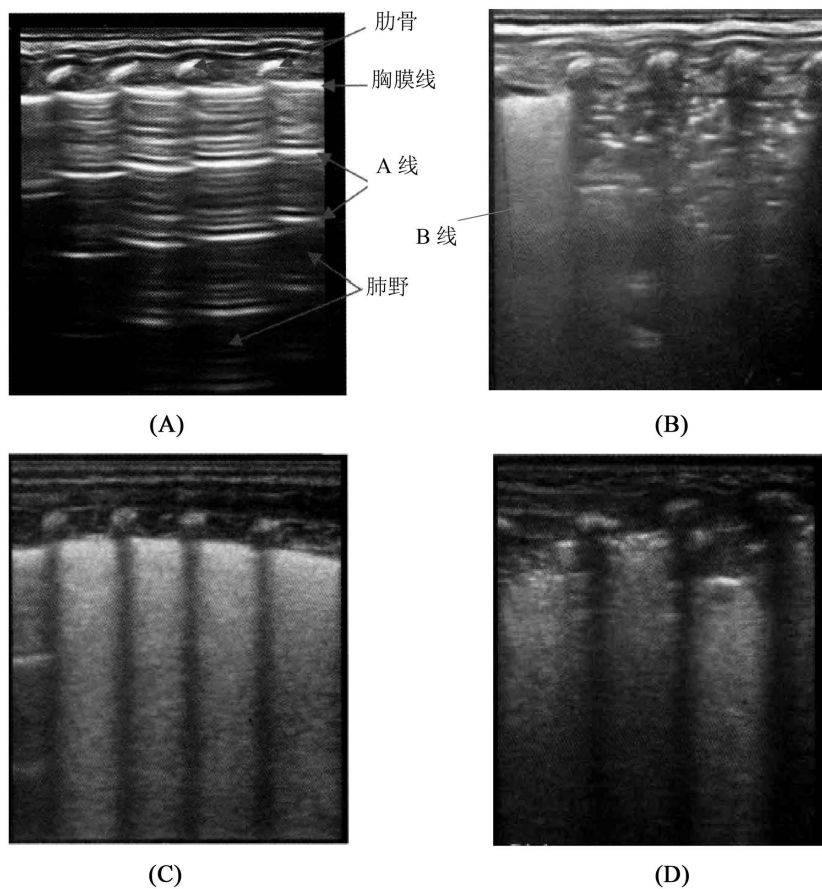


Figure 1. The images of neonatal lung ultrasound
图 1. 新生儿肺脏超声声像图表现

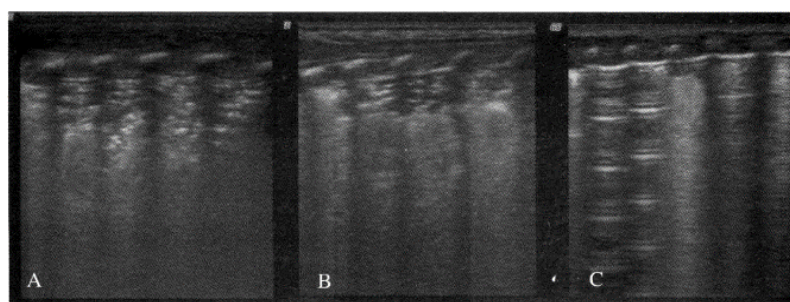


Figure 2. The recovery process of NRDS under ultrasound
图 2. 超声下 NRDS 恢复过程

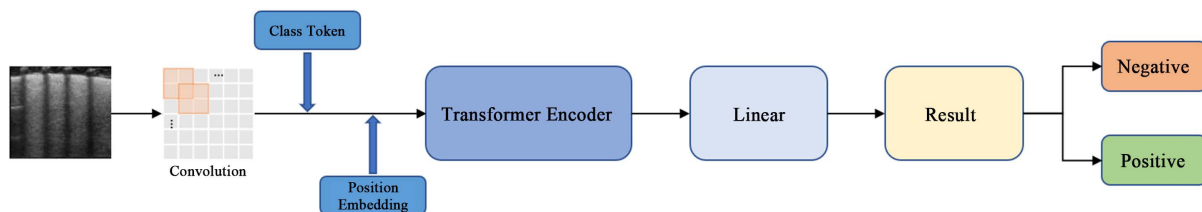


Figure 3. The structure of the model
图 3. 实验模型结构图

本实验模型输入的图像的数据格式为[224, 224, 3]的三维矩阵,但是 Transformer 模型要求输入的是二维矩阵[num_token, token_dim]。首先使用卷积核大小为 16×16 , 步距为 16, 卷积核个数为 768 的卷积将图像的数据格式从[224, 224, 3]变为[14, 14, 768], 接着进行展平处理, 把图像的宽与高两个维度展平, 得到的数据格式为[196, 768], 便能得到 Transformer 所需要的二维矩阵。同时, 在第一维度拼接一个长度为 768 的分类向量 class token, 其数据格式为[1, 768]。由于所有的 token 相互之间都在做信息交互, 因此分类向量 classtoken 可以从另外的 196 个 token 中学习到有用的信息, 最后只需根据 classtoken 的输出做出最后的分类判断便能得到最终分类结果, 拼接后得到的矩阵大小为[197, 768], 即得到 197 个 token, 每个 token 向量维度为 768。为了保持输入图像的空间位置信息, 需要添加位置编码向量, 此模型参考 Devlin 等人[12]的类标记, 使用标准的可学习 1D 位置编码, 其数据格式也为[197, 768], 其中分类向量 classtoken 的位置编码为 0。

在 Transformer Encoder 中, 堆叠 Encoder Block 12 次。主要由 Layer Norm、Multi-Head Attention、MLP Block 组成。其结构图如图 4 所示。Layer Norm 用于对输入的数据归一化, 可提升训练的稳定性。Multi-Head Attention (多头注意力机制) [13]是基于自注意力机制(Self-Attention), Self-Attention 是注意力机制中的一特例。Self-Attention 公式如下式(1), 式(2)所示:

$$\begin{cases} Q = XW^q \\ K = XW^k \\ V = XW^v \end{cases} \quad (1)$$

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

式(1)中, 输入矩阵 X 通过矩阵乘法线性映射到注意力空间矩阵 Q (查询矩阵), K (键矩阵), V (值矩阵), W^q 、 W^k 、 W^v 为可学习的超参数矩阵。式(2)中 d_k 表示键的维度, $\frac{1}{\sqrt{d_k}}$ 表示缩放因子, 用来防止 Q 与 K^T 的乘积过大。最终得到基于查询和键的相似度与值得加权和。

Multi-Head Attention 是将 Q 、 K 、 V 映射到不相同子空间上, 每个子空间独立进行 Self-Attention 计算, 最后将子空间的各个输出拼接在一起, 最终得到多头注意力机制的结果。公式如下式(3)、式(4)所示。

$$\text{head}_i = \text{Attention}(QW_i^q, KW_i^k, VW_i^v) \quad i = 1, 2, \dots, 12 \quad (3)$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_{12})W^o \quad (4)$$

式中 W_i^q 、 W_i^k 、 VW_i^v 是可学习的超参数矩阵, W^o 是为了将多头注意力进一步融合, 保证输入输出 mlti-head attention 的向量长度保持不变, 用一个输入输出大小全为 768 的全连接层实现。

MLP Block 由全连接层、GELU 激活函数、Dropout 组成。其结构如图 5 所示, 其中输入至第一个全连接层后, 二维矩阵中 token_dim 的值变成原来的 4 倍, 即矩阵[197, 3072], 经过第二个全连接层之后, 数据格式还原为原来的[197, 768]。其公式为式(5)、式(6)。

$$x'_i = \text{MSA}(\text{LN}(x_{i-1})) + x_{i-1} \quad i = 1, 2, \dots, 12 \quad (5)$$

$$x_i = \text{MLP}(\text{LN}(x'_i)) + x'_i \quad i = 1, 2, \dots, 12 \quad (6)$$

式中 MSA 表示多头注意力操作, LN 表示 Layer Norm, MLP 表示 MLPBlock 层对应的操作, x'_i 是多头注意力的结果, x_i 表示每次 Transformer Block 的输出。

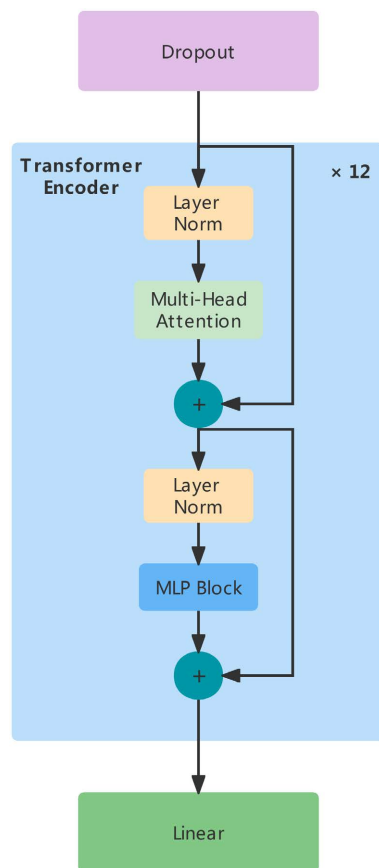


Figure 4. The structure of the Transformer Encoder
图 4. Transformer Encoder 结构图

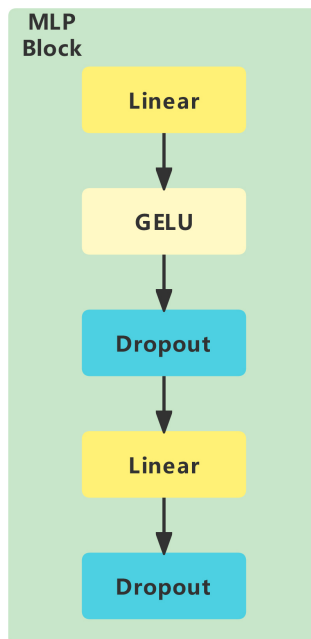


Figure 5. The structure of MLP Block
图 5. MLP Block 结构图

对于 Transformer Encoder 输出的数据, 取位置编码维度中索引为 0 的分类向量 classtoken, 输入至全连接层, 实现对新生儿肺部超声图像的分类, 最终得到诊断结果。公式如下式(7)所示。

$$y = LN(x_i^0) \quad (7)$$

式中 x_i^0 即为分类向量 class token 所对应的输出。

4. 实验步骤

4.1. 数据预处理

在超声诊断仪采集到的图像存在诸多与模型训练无关的干扰信息, 因此需要在不丢失图像关键信息的基础上对原始图像进行手动降维。原始图像大小为 816×614 像素, 保留原始图像中的信息, 使用 Matlab 中的 imcrop 函数, 将图像的剪切成 500×500 , 再通过 imresize 函数, 将图像等比例缩小至 224×224 像素。原始图像如图 6 所示, 预处理后图像如图 7 所示。

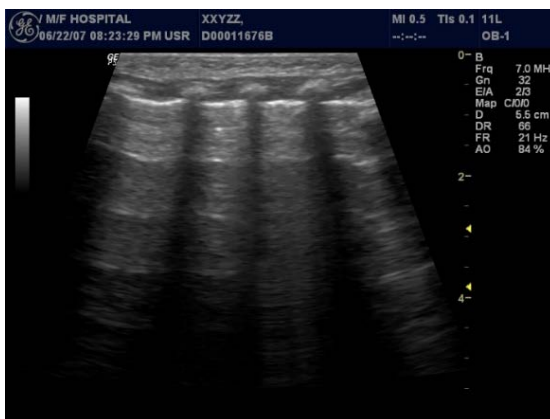


Figure 6. The original image captured by ultrasonography
图 6. 超声诊断仪采集的原始图像

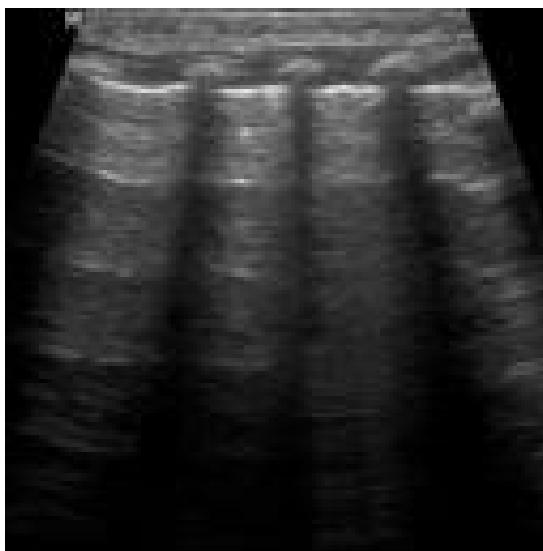


Figure 7. The image after pre-processing
图 7. 数据预处理后的图像

4.2. 模型训练

本次实验数据集总共 600 张新生儿肺部超声图, 包括 300 张正常的超声图像和 300 张有病的超声图像。在 300 张正常超声图像中, 270 张图用作训练与验证, 30 张图用作测试集; 300 张有病的超声图像中, 270 张图用作训练与验证, 剩余 30 张图用作测试集。同时, 在 540 张用于训练的图像中, 按 8:2 的比例划分训练集与验证集。在完成数据集的预处理之后, 基于 Pytorch 环境, 使用的显卡是 NVIDIA GTX 3090。学习率设置为 0.01, 每次输入模型 100 张数据集、即 BATCH_SIZE 设为 100, STEP 设为 600, 使用 ViT 网络在 ImageNet-21K 上预训练的权重, 对模型进行训练。本文共进行了 5 次实验, 5 次实验独立进行, 每次实验前重新随机划分训练集与测试集。

4.3. 评价指标

为了验证模型对新生儿肺部超声图像的分类效果, 采取四种不相同的指标对模型进行评估[14] [15] [16]: 正确率(Accuracy, Acc)、精确度(Precision, Pre)、敏感性(Sensitivity, Sen)、特异性(Specificity, Spe)。Acc 公式如下式子(8)所示, 是指正确识别数据集的数目在数据集总数的占比; Pre 公式如下式子(9)所示, 是指正确识别阳性数据集数目在所有被认定为阳性数据集的占比。Sen 公式如下式子(10)所示, 是指在所有阳性数据集中被正确判别出来的比率; Spe 公式如下式子(11)所示, 是指在所有阴性数据集中被正确识别出来的比率。

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$Pre = \frac{TP}{TP + FP} \quad (9)$$

$$Sen = \frac{TP}{TP + FN} \quad (10)$$

$$Spe = \frac{TN}{TN + FP} \quad (11)$$

其中, TP (True Positive)是指将阳性样本正确分类的个数; TN (True Negative)是指将阴性样本正确分类的个数; FP (False Positive)是指将阴性样本错误分类的个数, FN (False Negative)是指将阳性样本错误分类的个数。其混淆矩阵图如图 8 所示。

		Truth	
		0	1
Prediction	0	TN	FN
	1	FP	TP

Figure 8. The confusion matrix of indicators

图 8. 评估指标的混淆矩阵

另外,受试者工作特征曲线(Receiver Operating Characteristic Curve, ROC)反映着敏感性与特异性连续变量的综合指标,在一定程度上可以显示出分类器的分类效果; AUC 表示 ROC 曲线下的面积,可以更加直观地反映出分类效果。

4.4. 实验结果

在这 5 次的模型训练中,训练集的平均准确率 acc 为 0.935,平均损失函数 loss 为 0.210,验证集的平均准确率为 0.893,平均 loss 为 0.279。对于测试集,每组有 60 张图片,其结果如下表 1 所示。Acc、Pre、Sen、Spe 四类模型评估指标如下表 2 所示。同时,在表 3 中,我们还对本次实验提出的模型与目前一些主流分类模型做比较,有残差网络 ResNet,基于逐点分组卷积的轻量级神经网络 ShuffleNet 以及基于深度可分离卷积的轻量级神经网络 MobileNet-V2,来验证本次模型的实验效果。

Table 1. The results of five experiments

表 1. 5 次实验测试集测试结果

实验次数 \ 指标	TP	TN	FP	FN
1	30	25	5	0
2	29	29	1	1
3	29	30	1	1
4	30	28	2	0
5	29	28	2	1

Table 2. The four different indexes of five experiments

表 2. 5 次实验评估指标

实验次数 \ 评估指标	Acc	Pre	Sen	Spe
1	91.67%	85.71%	100%	83.33%
2	96.67%	96.67%	96.67%	96.67%
3	98.33%	100%	96.67%	100%
4	96.67%	93.75%	100%	93.33%
5	95%	93.55%	96.67%	93.33%
Average	95.68%	93.94%	98.00%	93.32%

Table 3. Comparison with other classification models

表 3. 与其他模型对比

分类模型	Acc	Pre	Sen	Spe
ResNet	94.00%	92.65%	90.51%	83.33%
ShuffleNet	91.50%	96.63%	85.76%	97.07%
MobileNet-V2	93.66%	98.54%	88.47%	98.69%
Our approach	98.33%	100%	96.67%	100%

从表 2 中可以看出, 经过 5 次实验, 平均准确率 Acc 已达到了 95.68%, 敏感性 Sen、精确度 Pre、特异性 Spc 均达 90% 以上。表 3 中, 选取了其他几个模型效果最好时的评估指标值, 可以看出, 效果都不如本文的模型。

根据实验结果, 绘制出 ROC 曲线, 如图 9 所示。从图中可以看出, 平均 AUC 达到了 99.53%, 说明该模型对新生儿肺部超声图像的分类效果较好。

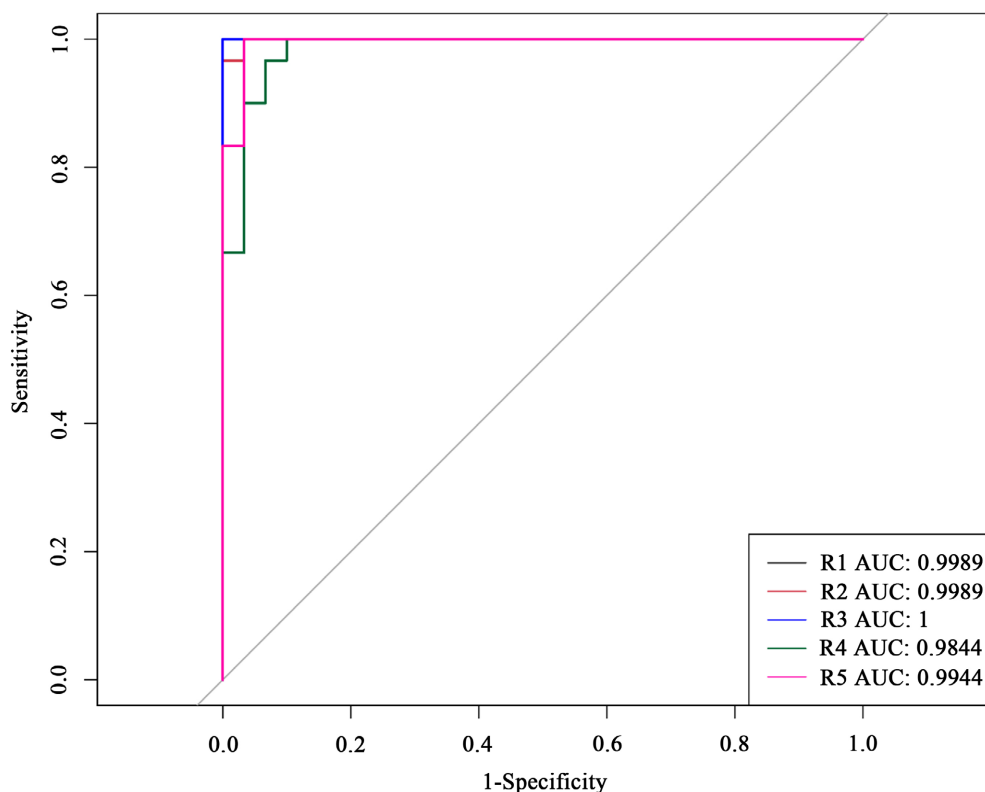


Figure 9. ROC curve of five experiments
图 9. 五次实验的 ROC 曲线

5. 总结

实验结果表明, 相对于传统的诊断方法, 本文提出的一种卷积与 ViT 相结合的模型可以实现对新生儿肺部超声图像的特征提取及分类, 并且分类效果较好。本文分类模型能够辅助医生诊断新生儿是否需要进一步的诊断, 提高主治医师的工作效率, 减轻工作负担, 同时能够减轻新生儿父母的心理负担。但还存在着一些不足之处, 本次实验数据集的样本较少, 实验的结果容易出现过拟合。同时, 需要对模型结构进一步优化与完善。实验结果如何与医生的临床诊断相结合, 也需要不断地探索与研究。

基金项目

国家自然科学基金 81101116。

参考文献

- [1] 邢晓芬, 徐向民, 黄晓泓, 黄建敬. 基于内容的医学图像分类研究[J]. 科学技术与工程, 2007, 7(1): 85-90.
- [2] Volpicelli, G., Lamorte, A., Tullio, M., *et al.* (2013) Point-of-Care Multiorgan Ultrasonography for the Evaluation of

- Undifferentiated Hypotension in the Emergency Department. *Intensive Care Medicine*, **39**, 1290-1298. <https://doi.org/10.1007/s00134-013-2919-7>
- [3] Lichtenstein, D.A. and Meziere, G.A. (2008) Relevance of Lung Ultrasound in the Diagnosis of Acute Respiratory Failure: The BLUE Protocol. *CHEST*, **134**, 117-125. <https://doi.org/10.1378/chest.07-2800>
- [4] Lien, W.-C., Hsu, S.-H., Chong, K.-M., *et al.* (2018) US-CAB Protocol for Ultrasonographic Evaluation during Cardiopulmonary Resuscitation: Validation and Potential Impact. *Resuscitation*, **127**, 125-131. <https://doi.org/10.1016/j.resuscitation.2018.01.051>
- [5] Ross, A.M., Genton, E. and Holmes, J.H. (1968) Ultrasonic Examination of the Lung. *The Journal of Laboratory and Clinical Medicine*, **72**, 556-564.
- [6] Lichtenstein, D. and Meziere, G. (1998) A Lung Ultrasound Sign Allowing Bedside Distinction between Pulmonary Edema and COPD: The Comet-Tail Artifact. *Intensive Care Medicine*, **24**, 1331-1334. <https://doi.org/10.1007/s001340050771>
- [7] Lichtenstein, D., Meziere, G., Biderman, P. and Gepner, A. (1999) The Comet-Tail Artifact: An Ultrasound Sign Ruling out Pneumothorax. *Intensive Care Medicine*, **25**, 383-388. <https://doi.org/10.1007/s001340050862>
- [8] Lichtenstein, D., Meziere, G., Biderman, P., Gepner, A. and Barré, O. (1997) The Comet-Tail Artifact. An Ultrasound Sign of Alveolar-Interstitial Syndrome. *American Journal of Respiratory and Critical Care Medicine*, **156**, 1640-1646. <https://doi.org/10.1164/ajrcem.156.5.96-07096>
- [9] Lichtenstein, D.A. and Menu, Y. (1995) A Bedside Ultrasound Sign Ruling out Pneumothorax in the Critically III: Lung Sliding. *CHEST*, **108**, 1345-1348. <https://doi.org/10.1378/chest.108.5.1345>
- [10] 刘敬, 曹海英, 程秀永. 新生儿肺脏疾病超声诊断学[M]. 郑州: 河南科学技术出版社出版, 2013.
- [11] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. ArXiv: 2010.11929v2.
- [12] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2018) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. ArXiv: 1810.04805v1.
- [13] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. *31st Annual Conference on Neural Information Processing Systems (NIPS)*, Long Beach, 4-9 December 2017.
- [14] Anwar, S.M., Majid, M., Qayyum, A., *et al.* (2018) Medical Image Analysis Using Convolutional Neural Networks: A Review. *Journal of Medical Systems*, **42**, Article No. 226. <https://doi.org/10.1007/s10916-018-1088-1>
- [15] Hamidinekoo, A., Denton, E., Rampun, A., Honnor, K. and Zwigelaar, R. (2018) Deep Learning in Mammography and Breast Histology, an Overview and Future Trends. *Medical Image Analysis*, **47**, 45-67. <https://doi.org/10.1016/j.media.2018.03.006>
- [16] Litjens, G., Kooi, T., Bejnordi, B.E., *et al.* (2017) A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis*, **42**, 60-88.