

IMRT QA 中基于后融合卷积神经网络的MLC误差分类预测

闫冰¹, 余泉¹, 王帅¹, 吴爱林², 张红雁¹, 吴爱东¹

1. 中国科学技术大学附属第一医院(安徽省立医院)放疗科, 安徽 合肥 230001; 2. 中国科学技术大学附属第一医院西区(安徽省肿瘤医院)放疗科, 安徽 合肥 230001

【摘要】目的:评估基于剂量差图与Gamma分布图的多模态多通路卷积神经网络用于分类调强放射治疗(IMRT)质量保证(QA)中多叶准直器(MLC)误差的可行性及优势。**方法:**首先通过修改选取的68例IMRT放疗计划原始无误差照射野的MLC叶片位置用于模拟4种误差类型:平移误差、外扩误差、内收误差、随机误差,并将原始无误差计划及4种引入MLC误差计划重新导入TPS,计算PTW 729模体中的剂量分布;其次从测量和计算的剂量分布中创建剂量差图和两种通过率标准下的Gamma图作为数据集建立并训练多模态多通路卷积神经网络,其中330个剂量误差图和660个Gamma图用于测试集,其余数据集按照五折交叉验证划分为训练集与验证集。根据测试集的预测结果,计算其总体分类准确度、宏平均F1值以及归一化混淆矩阵用于评估模型性能。**结果:**后融合卷积神经网络具有最高的总体分类准确度(0.855)和宏平均F1值(0.853),根据归一化混淆矩阵,MLC内收误差、外扩误差、无误差、随机误差、平移误差的平均分类准确度分别为0.98、1.00、0.66、0.63、1.00。**结论:**基于多模态误差图像的后融合卷积神经网络,其在总体分类准确度、宏平均F1值以及每种误差类型的特定分类准确度等方面均显示了该方法的可行性及准确性。

【关键词】深度学习;误差分类;调强放射治疗;质量保证

【中图分类号】R318;R811.1

【文献标志码】A

【文章编号】1005-202X(2023)08-0925-08

MLC error detection using a late-fusion convolutional neural network in quality assurance for IMRT

YAN Bing¹, YU Xiao¹, WANG Shuai¹, WU Ailin², ZHANG Hongyan¹, WU Aidong¹

1. Department of Radiation Oncology, the First Affiliated Hospital of University of Science and Technology of China (Anhui Provincial Hospital), Hefei 230001, China; 2. Department of Radiation Oncology, West Branch of the First Affiliated Hospital of University of Science and Technology of China (Anhui Cancer Hospital), Hefei 230001, China

Abstract: Objective To evaluate the feasibility and advantages of multi-channel multi-path DenseNet (MCMP-DenseNet) for detecting multi-leaf collimator (MLC) errors in quality assurance (QA) for intensity-modulated radiotherapy (IMRT) from dose difference maps and Gamma maps. **Methods** The MLC positions of 98 error-free IMRT plans were modified to simulate translation, extension, shift, and random errors. The plans with and without errors were re-imported into TPS for calculating the dose distributions in the PTW 729 phantom. The dose difference maps and the Gamma maps with two Gamma criteria which were created from the measured and calculated dose distributions were used for dataset establishment and MCMP-DenseNet training. Among them, 330 dose difference maps and 660 Gamma maps was adopted for the test set, and the remaining were divided into training and validation sets according to 5-fold cross-validation. Based on the prediction results of the test set, the overall classification accuracy, Macro-F1, and normalized confusion matrix were calculated for evaluating the model performance. **Results** MCMP-DenseNet had the highest overall classification accuracy (0.855) and Macro-F1 (0.853). The normalized confusion matrix revealed that the average classification accuracies of the MLC shift error, expansion error, error-free, random error, and translation error were 0.98, 1.00, 0.66, 0.63, and 1.00, respectively. **Conclusion** The study demonstrates the feasibility and accuracy of MCMP-Densenet in terms of the overall classification accuracy, Macro-F1, and specific classification accuracy.

Keywords: deep learning; error classification; intensity-modulated radiotherapy; quality assurance

【收稿日期】2023-02-19

【基金项目】国家自然科学基金(11805198)

【作者简介】闫冰,高级工程师,研究方向:肿瘤放射物理及放射治疗技术,E-mail: cpreeo@126.com

【通信作者】吴爱东,高级工程师,研究方向:肿瘤放射物理及放射治疗技术,E-mail: flkaidongwu@163.com

前言

调强放射治疗(Intensity Modulated Radiation Therapy, IMRT)是肿瘤放射治疗最重要的治疗方式之一,其具有较强的剂量调制能力,可以给予肿瘤组织较高照射剂量的同时最大程度降低周围危及器官的受照剂量^[1]。由于IMRT具有较多的控制点用于调控剂量分布,因此在患者接受IMRT治疗之前通常需针对该计划行剂量方面的质量保证(Quality Assurance, QA),从而保证患者受照剂量与计划系统计算的剂量分布在允许的误差范围之内^[2]。针对相应IMRT计划进行剂量QA时,通常采用Gamma指数中的剂量偏差和距离满意度法分析计算剂量与测量剂量之间的差别,在3%/2 mm和剂量阈值10%条件下,Gamma通过率大于90%被认为是QA允许的偏差量^[3]。由于Gamma通过率是单一数值,在临床使用中通常是比较该值是否达到容差限值或干预限值,但QA不通过时,很难通过Gamma数值分析误差产生的来源^[4]。基于此,国内外有许多研究将机器学习方法应用于患者计划QA中,用于判别多叶准直器(Multi-Leaf Controller, MLC)误差类型^[5-7]。但这些研究大多是基于单一的数据类型(Gamma图或者剂量差图)^[8-11],因此本研究旨在构建一个多模态多通路卷积神经网络(Convolutional Neural Networks, CNN),可以将Gamma图和剂量差图等多种数据类型应用于机器学习中,实现对MLC误差类型的分类,同时对比评估其分类性能以验证模型的可行性及优势。

1 资料与方法

1.1 计划选取与测量

选取中国科学技术大学附属第一医院68例IMRT计划,涵盖头颈部、胸部、腹盆腔肿瘤。所有计划均使用Monaco放疗计划系统(瑞典医科达公司)以DMC固定照射野方式完成计划设计,计划结果符合临床要求。所有IMRT计划均使用配置为80叶MLC等中心宽度为1 cm的直线加速器(瑞典医科达)完成照射。68例IMRT计划中共计包含366个照射野,每例计划所含照射野的数量为3~13个。将所有IMRT计划移植至PTW 729二维矩阵(德国PTW公司),为了减少探测器的角度影响,所有照射野角度归一至0°,68例IMRT计划使用PTW 729探测器在0°对每个单野进行测量,共计获得366个照射野的剂量分布图。

1.2 MLC误差模拟

为了尽可能模拟MLC的位置误差,本研究模拟4种不同类型MLC位置误差计划,包括:平移误差、

外扩误差、内收误差、随机误差。平移误差:对于照射野的每一控制点,所有MLC沿同一方向移动2 mm,与原始无误差计划相比,该种误差相当于所有控制点的照射野大小不变,照射野平移2 mm。外扩误差:对于照射野的每一控制点,所有MLC沿远离等中心点方向移动2 mm,与原始无误差计划相比,该种误差计划相当于照射野内所有控制点的射野尺寸扩大4 mm。内收误差:对于照射野的每一控制点,所有MLC沿靠近等中心点方向移动2 mm,与原始无误差计划相比,该种误差计划相当于照射野内所有控制点的射野尺寸内收4 mm。随机误差:对于照射野的每一控制点,所有MLC按照高斯分布($\delta=1$ mm)误差随机移动。

为了模拟这些误差,将原始无误差计划的RT-plan数据从TPS中导出。RT-plan数据中所有控制点的MLC位置通过MATLAB(美国Math Works)自编程序被修改为指定的位移量,然后将修改后的RT-plan数据重新导入至TPS中。共计1464个含有MLC位置误差的照射野被重新移植至PTW 729二维矩阵,并提取等中心层面的剂量分布,如图1所示。

1.3 数据处理

通过QA测量与误差计划模拟,共获得366个测量剂量分布与1830个计算剂量分布(原始无误差、平移误差、外扩误差、内收误差、随机误差计划各366个)。为了获得剂量差异图和Gamma图用于深度学习的输入,上述测量与计算的剂量分布通过以下方式处理,其中一个典型的照射野处理后的剂量差图与Gamma图如图2所示。剂量差图:首先将各种MLC类型的计算剂量分布与测量剂量分布的分辨率调整一致;随后将计算剂量分布和与之对应的测量剂量分布按照测量剂量分布中的最大剂量进行归一,最终将归一化的计算剂量分布与测量剂量分布相减获得剂量差图。为了方便输入至深度学习网络,以-0.2和0.2为阈值,将上述处理后的剂量差图重新归一化至-1~1。Gamma图:将各种MLC类型的计算剂量分布与测量分布进行比较,其中剂量阈值10%,全局归一,Gamma标准分别取3%/2 mm和2%/1 mm。同样为了方便输入至深度学习网络,以Gamma值-1.5和1.5为阈值,将Gamma图归一化至-1~1。Gamma计算在Python环境中使用Pygamma完成,具体代码可在GitHub查阅(github.com/christopherpoole/pygamma)^[12]。

1.4 网络结构

多模态图像或数据一般采用前融合或后融合两种方式输入至CNN。前融合方式指多模态数据在CNN前期层进行特征融合^[13],而后融合方式与之相

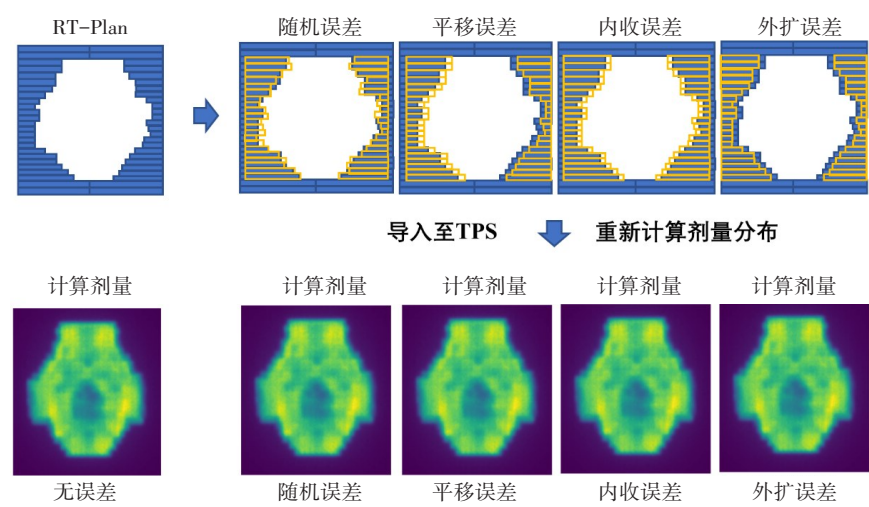


图1 误差计划模拟流程(4种MLC位置误差)
Figure 1 Simulation of error plans (4 types of MLC errors)

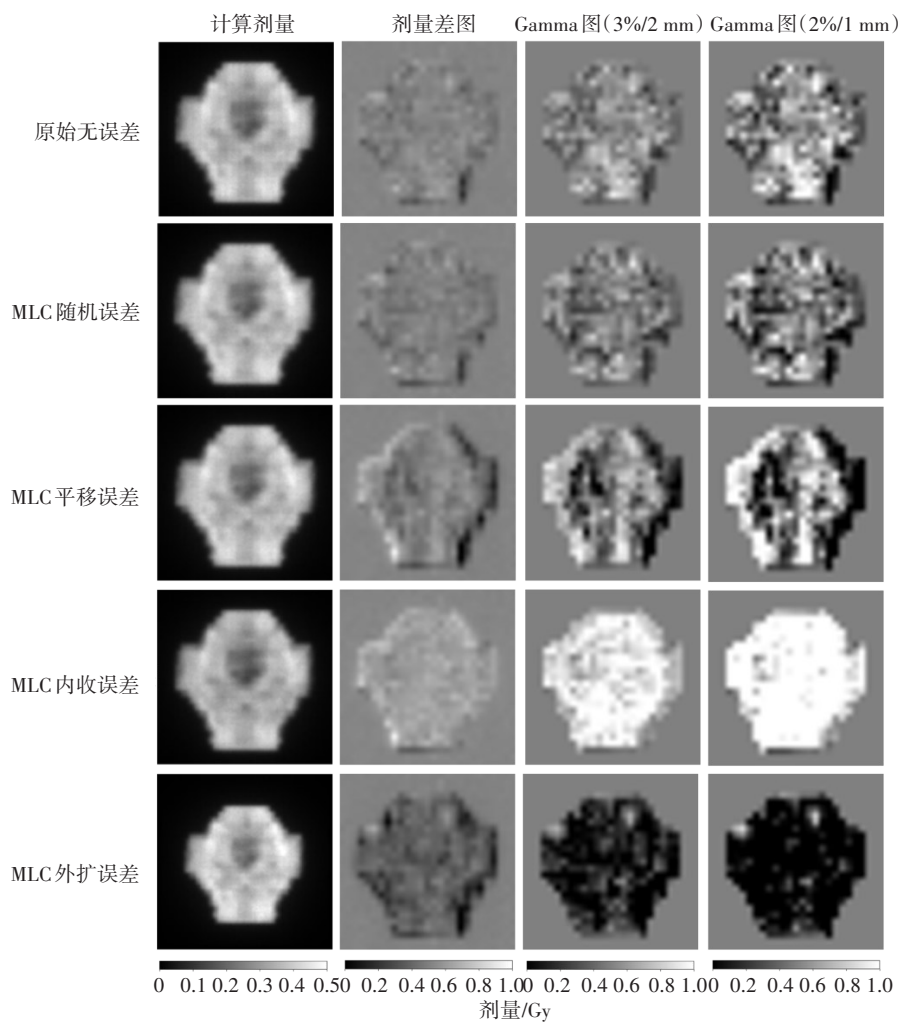


图2 原始无误差计划与4种误差计划在数据处理后的剂量差图和Gamma图
Figure 2 Gamma maps and dose difference maps of MLC error and free-error plans

反,其特征融合在CNN后期层^[14-15]。为了探讨剂量差图、Gamma图以何种融合方式更有利于MLC误差类型的分类,本研究分别建立并训练了前融合CNN

和后融合CNN。
前融合CNN网络由一个特征提取器和一个全连接层两部分组成。特征提取器采用了Huang等^[16]提出的

密集卷积网络 (Densely Connected Convolutional Networks, DenseNet),包括1个卷积层、4个密集模块和3个传输层。与原始网络有所区别的是由于用于输入的剂量差图和Gamma图像尺寸较小,因此首个卷积层的卷积核大小由原始的7×7调整为3×3用于降低感受野,同时取消了原网络中随后的平均池化操作。4个密集模块依次包含了6、12、24、16个稠密层,每个密集层均包含了两个卷积核大小为1×1和3×3的卷积层。传输层使用卷积核大小为1×1的卷积层用于降低channel数量,每个传输层后均使用平均池化操作用于降低图像尺寸。所有卷积层后均采用批量归一化和修正线性单元操作。最后使用全连接层输出5种MLC误差类型的预测概率,其中预测值最大的一类被用作最终的预测结果。由于该网络使用了3种模态的数据,但仅使用

了一个特征提取通路,将该网络记为MCSP-DenseNet (Multi-channel Single-path DenseNet)。后融合CNN网络由3个特征提取器、1个融合模块、1个全连接层组成。3个特征提取器与前融合CNN网络中的特征提取器结构大体一致,不同的是将最后一次平均池化操作移至融合模块内,3个特征提取器各自独立地从剂量误差图以及两种通过率标准下的Gamma图中提取特征信息。融合模块包含拼接操作、卷积层和平均池化,3种独立的特征在拼接操作后,通过卷积核大小为1×1的卷积层融合特征。融合后的特征输入至全连接层,预测最终分类结果。由于该网络使用了3种模态的数据,且使用了3个特征提取通路,将该网络记为MCMP-DenseNet (Multi-channel Multi-path DenseNet),前融合和后融合CNN网络结构具体见图3。

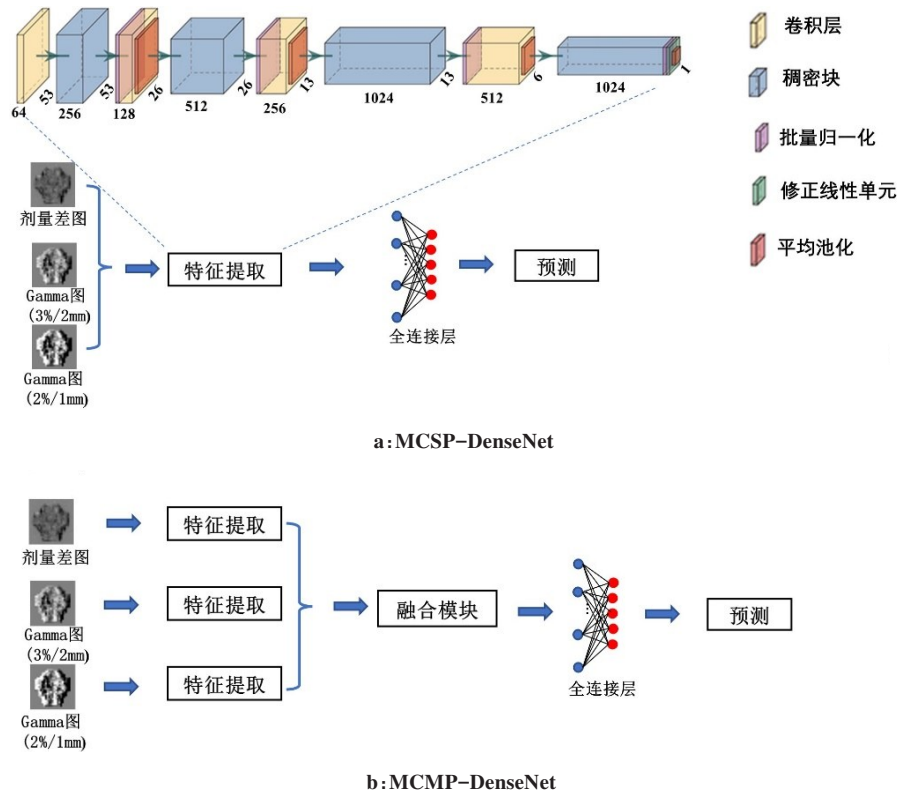


图3 前融合与后融合CNN网络结构图
Figure 3 Architectures of the early fusion and late fusion CNN models

为了比较MLC误差类型分类中多模态图像与单一模态图像的优劣,本研究亦建立了应用单模态图像的CNN用于对比研究。单模态CNN在网络结构上与MCSP-DenseNet大致一致,不同的是网络输入仅使用一种模态的图像。这些网络单独以剂量差图、Gamma图(3%/2 mm)、Gamma图(2%/1 mm)作为输入,被依次命名为SCSP-DenseNet1(单通道单路径DenseNet)、SCSP-DenseNet2和SCSP-DenseNet3。

1.5 网络训练

本研究随机选择66个照射野,将其所对应的共计330个剂量差图和660个Gamma图用于测试集,剩余的300个照射野共计1 500个剂量差图和3 000个Gamma图按照五折交叉验证划分训练集与验证集。每次交叉验证所获得的网络模型均使用测试集验证其性能表现,5次测试集验证的平均值被认为是网络性能表现的最终结果。所有网络均在Ubuntu18.04.4

LTS 系统和 Tesla V100 GPU 上使用 PyTorch 实现。网络训练使用 Adam 优化器^[17], epoch 为 200, 学习率为 0.000 1, batch 大小为 42, 采用随机水平和垂直翻转用于数据增强。在训练过程中, 根据验证集的精度保存最优网络参数用于测试集。

1.6 评价指标

本研究采用常用的多分类评价指标衡量模型的性能, 包括准确度 (Accuracy)、宏平均 F1 值 (Macro-F1) 及归一化混淆矩阵^[18]。其中准确度定义为测试集中分类正确的个数与总数的比值; 宏平均 F1 值表示在 MLC 分类中, 每种类别 F1 值 (F1-Score_{*i*}) 的算数平均值, 其中每种类别 F1 值定义为该类别精度 (Precision) 和召回率 (Recall) 的调和平均, 宏平均 F1 值和每种类别的 F1 值越大表示模型的分类性能越好; 归一化混淆矩阵元素 (*i, j*) 代表类别 *i* 被分类为类别 *j* 的个数与测试集中类别 *i* 总数的比值, 即每种类别的特定分类准确度 (Specific Classification Accuracy, SCA)。上述参数的计算公式如下:

$$\text{Recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i}$$

(1)

$$\text{Precision}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i}$$

(2)

$$\text{F1 - Score}_i = 2 \times \frac{\text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}$$

(3)

$$\text{Macro - F1} = \sum_{i=1}^5 \frac{\text{F1 - Score}_i}{5}$$

(4)

其中, Precision_{*i*}、Recall_{*i*}、F1 - Score_{*i*} 分别表示第 *i* 类的精度、召回率、F1 值; TP_{*i*}、FN_{*i*}、FP_{*i*} 分别为第 *i* 类的真阳数 (实际为 *i* 类被预测为 *i* 类)、假阴数 (实际为 *i* 类被预测为非 *i* 类)、假阳数 (实际为非 *i* 类被预测为 *i* 类)。

2 结果

3 种 SCSP-DenseNet、MCSP-DenseNet 和 MCMP-DenseNet 的五折交叉验证总体分类准确度如表 1 所示, 其中 MCMP-DenseNet 的分类准确度在 5 种网络中最高, 达到 0.855。

表 1 五折交叉 3 种 SCSP-DenseNet、MCSP-DenseNet 以及 MCMP-DenseNet 的分类准确度
Table 1 Classification accuracies of MCSP-DenseNet, MCMP-DenseNet and 3 types of SCSP-DenseNet in 5-fold cross-validation

交叉组合	SCSP-DenseNet1	SCSP-DenseNet2	SCSP-DenseNet3	MCSP-DenseNet	MCMP-DenseNet
组合 1	0.845	0.821	0.803	0.845	0.855
组合 2	0.794	0.855	0.839	0.827	0.845
组合 3	0.845	0.839	0.839	0.839	0.867
组合 4	0.803	0.842	0.812	0.830	0.852
组合 5	0.848	0.845	0.796	0.836	0.855
平均值	0.827±0.026	0.840±0.012	0.818±0.020	0.835±0.007	0.855±0.008

3 种 SCSP-DenseNet、MCSP-DenseNet 以及 MCMP-DenseNet 的宏平均 F1 值及每种类别的 F1 值如表 2 所示, MCMP-DenseNet 具有最高的宏平均 F1 值 (0.853), 同时

在 MLC 无误差及随机误差中具有最高的分类 F1 值, 分别为 0.653 和 0.639。SCSP-DenseNet2 在 3 种基于单模态图像的网络中具有最高的宏平均 F1 值。

表 2 3 种 SCSP-DenseNet、MCSP-DenseNet 以及 MCMP-DenseNet 在 5 种 MLC 位置分类中每种类别的 F1 值与宏平均 F1 值
Table 2 F1-score and Macro-F1 of MCSP-DenseNet, MCMP-DenseNet and 3 types of SCSP-DenseNet for different MLC errors

参数类型	SCSP-DenseNet1 (剂量差图)	SCSP-DenseNet2 (Gamma 3%/2 mm)	SCSP-DenseNet3 (Gamma 2%/1 mm)	MCSP-DenseNet	MCMP-DenseNet
内收误差	0.983	0.992	0.991	0.986	0.991
外扩误差	1.000	0.997	0.982	0.991	0.988
无误差	0.627	0.595	0.562	0.592	0.653
随机误差	0.541	0.602	0.568	0.602	0.639
平移误差	0.966	0.997	0.949	0.991	0.994
宏平均 F1 值	0.823	0.837	0.810	0.832	0.853

5种网络模型的五折平均归一化混淆矩阵及五折累计混淆矩阵如图4和表3所示。对于所有的模型,在预测MLC外扩、内收和平移误差上的表现相差不大(SCA:0.96~1.00),且均明显优于预测MLC无误差及随机误差(SCA:0.50~0.69)。对于MLC无误差分类预测

方面,SCSP-DenseNet1具有最高的SCA(0.69),MCMP-DenseNet次之(0.66),SCSP-DenseNet3表现最差(0.55)。而对于MLC随机误差分类预测方面,MCMP-DenseNet具有最高的SCA(0.63),MCSP-DenseNet次之(0.62),SCSP-DenseNet1的性能最差(0.50)。

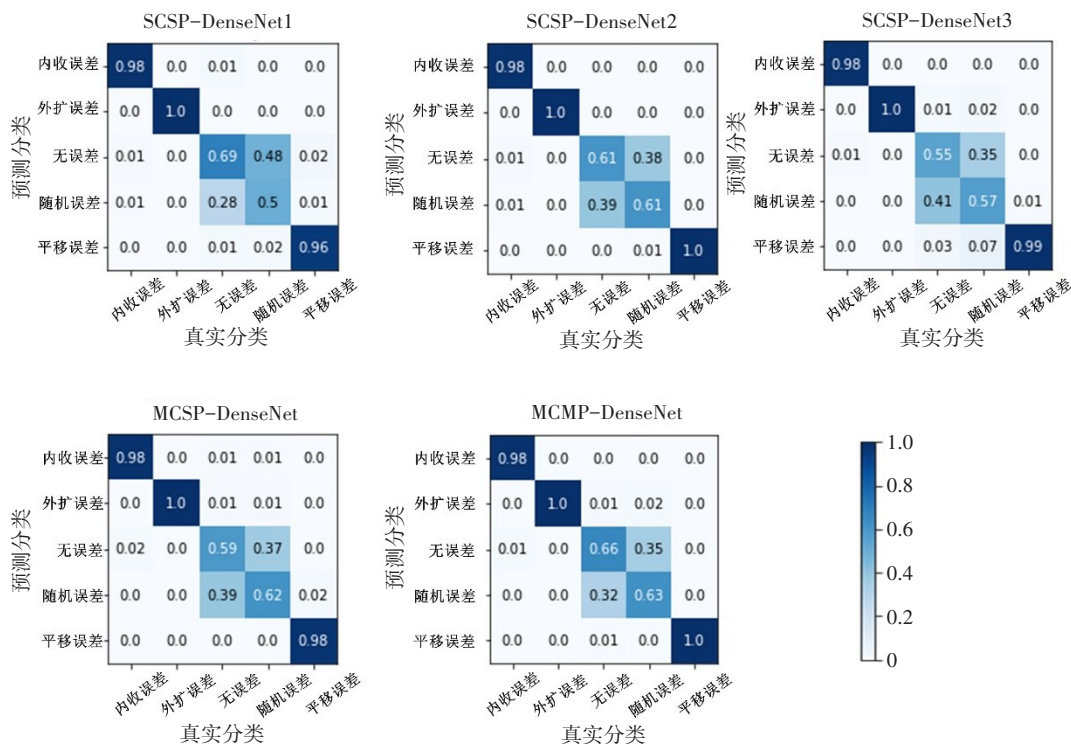


图4 3种SCSP-DenseNet、MCSP-DenseNet以及MCMP-DenseNet的归一化混淆矩阵

Figure 4 Normalized confusion matrixes of MCSP-DenseNet, MCMP-DenseNet, and 3 types of SCSP-DenseNet

3 讨论

本研究构建了一个多模态多通路CNN,通过对Gamma图和剂量差图等多种数据类型学习,实现对MLC误差类型的分类预测,同时对比评估这种方法在IMRT患者QA中的可行性及优越性。本研究结果显示,MCMP-DenseNet与SCSP-DenseNet相比具有更高的总体分类准确度、宏平均F1值,且MCMP-DenseNet与使用Gamma图的两种SCSP-DenseNet相比,MCMP-DenseNet均具有最高的SCA。尽管基于剂量差图的SCSP-DenseNet1在MLC无误差上的SCA最高(0.69),但在随机误差上的SCA仅有0.50,远低于MCSP-DenseNet(0.63)。因此,总体而言MCMP-DenseNet具有最高的分类性能表现。这也表明在MLC误差分类预测中,MCMP-DenseNet联合3种不同图像,较单一类型图像更具有优势。

与前融合网络MCSP-DenseNet相比,两者在随机误差上的SCA差异不大,但在无误差上的SCA相差7%,同时MCMP-DenseNet模型在总体分类准确

度、宏平均F1值也均具有明显优势,这得益于剂量差图与两种通过率标准下的Gamma分布图在网络中各自独立提取特征,避免3种不同类型图像在网络的低级阶段相互融合,从而全连接网络可完整接收3种不同类型分布图的特征信息。MCSP-DenseNet虽然也使用了3种类型的图像,但其处理方式是在网络的输入阶段相互拼接,类似于RGB图像的处理方式,不同图像的独立特征可能会在网络的高水平上消失^[19]。

3种SCSP-DenseNet之间相比,基于3%/2 mm Gamma分布图的网络具有最佳的总体分类准确度,达到0.84。Kimura等^[20]研究发现CNN网络使用剂量差图的表现整体优于使用Gamma图,而本研究显示使用Gamma图的网络表现优于剂量差图,分析原因是由于Kimura等^[20]研究的Gamma值是使用Delta4软件计算,其Gamma值均为正值,与剂量差图相比缺失了冷热点信息,而本研究使用Pygamma,可根据冷热点信息在Gamma值中加入符号,Gamma图并未出现信息丢失。因此,采用剂量差图的SCSP-DenseNet

表3 3种SCSP-DenseNet、MCSP-DenseNet及MCMP-DenseNet五折交叉验证共计330个照射野的混淆矩阵
Table 3 Confusion matrixes of MCSP-DenseNet, MCMP-DenseNet and 3 types of SCSP-DenseNet for 330 beams in 5-fold cross-validation

方法	金标\预测	内收误差	外扩误差	无误差	随机误差	平移误差
SCSP-DenseNet1	内收误差	325	0	3	2	0
	外扩误差	0	330	0	0	0
	无误差	4	0	228	94	4
	随机误差	1	0	159	165	5
	平移误差	0	0	8	4	318
SCSP-DenseNet2	内收误差	325	0	3	2	0
	外扩误差	0	330	0	0	0
	无误差	0	1	200	129	0
	随机误差	0	1	125	202	2
	平移误差	0	0	0	0	330
SCSP-DenseNet3	内收误差	325	0	3	1	1
	外扩误差	0	330	0	0	0
	无误差	1	4	181	134	10
	随机误差	0	5	114	187	24
	平移误差	0	1	0	2	327
MCSP-DenseNet	内收误差	325	0	5	0	0
	外扩误差	0	330	0	0	0
	无误差	2	3	196	129	0
	随机误差	2	3	121	204	0
	平移误差	0	0	1	5	324
MCMP-DenseNet	内收误差	325	0	4	1	0
	外扩误差	0	330	0	0	0
	无误差	1	3	219	105	2
	随机误差	0	5	117	208	0
	平移误差	0	0	0	1	329

在分类准确性方面并未表现出明显优势。

机器学习方法应用于患者放疗计划QA分析,国内外许多学者展开相关研究,其中基于影像组学的机器学习方法在IMRT的QA中被证明是检测MLC位置错误的一种有优势的方法。Kimura等^[20]使用基于Delta4所得到的剂量误差图作为卷积神经网络的输入,对MLC的位置错误取得较好的分类结果,五折验证的平均模型精度为0.855,与本研究的结果相同。与Kimura等^[20]研究不同的是本研究使用PTW 729二维探测器的测量结果作为卷积神经网络的输入,进一步扩大了深度学习在MLC误差检测中的应用范围,同时在其3种MLC误差类型的基础上增加至5种。

由于电子射野影像设备(Electronic Portal Imaging Device, EPID)采集的影像具有较高的空间分辨率^[21-23],较多学者采用EPID影像作为机器学习的输入,用于MLC位置错误的分类。Sakai等^[10]利用EPID测量所得到的Gamma图进行定量分析,结果与传统的Gamma分析相比,基于影像组学的机器学习模型在检测单一类型的MLC建模错误和MLC位置错误方面显示出更高的灵敏度和特异性。Ma等^[24]通过对Gamma图的结构相似性的子指数图进行特征提取,结合机器学习方法,结果表明该方法在IMRT QA中取得较高的错误分类准确度,并且与传统的Gamma阈值方法相比,在检测错误方面表现出卓越的敏感性。但上述研究均是基于EPID为测量设备展开的,而本研究所采用的探测器在空间分辨率上较低,因此很难将本研究结果与其相互比较,且本研究

旨在分析将Gamma图和剂量差图相结合应用于卷积神经网络,以证明多种类型参数卷积神经网络较单一类型更具有优势。

虽然本研究提出的卷积神经网络与Gamma分析法相比具有一定的优势,但仍存在一些问题需要解决:一是误差类型较为简单,仅模拟了MLC的5种误差类型,其它误差因素并未考虑,如机架角度误差等^[25];二是模型输入并未考虑到IMRT计划本身的复杂度与Gamma通过率之间的关系。因此,研究的下一步将会探讨更多的误差类型,并将放疗计划本身的复杂度考虑到新模型中,与本研究的模型结果进行对比。

综上所述,本研究提出并实现一种后融合的多模态多通道卷积神经网络,与单模态及前融合卷积神经网络相比,在总体分类准确度、敏感性等方面均显示了方法的可行性及准确性。

【参考文献】

- [1] Brahme A. Optimization of stationary and moving beam radiation therapy techniques[J]. *Radiother Oncol*, 1988, 12(2): 129-140.
- [2] Portelance L, Chao KS, Grigsby PW, et al. Intensity modulated radiation therapy (IMRT) reduces small bowel, rectum, and bladder doses in patients with cervical cancer receiving pelvic and para-aortic irradiation[J]. *Int J Radiat Oncol Biol Phys*, 2001, 51(1): 261-266.
- [3] Miften M, Olch A, Mihlids D, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM Task Group No. 218[J]. *Med Phys*, 2016, 43(6): 3852-3853.
- [4] Nelms BE, Chan MF, Jarry G, et al. Evaluating IMRT and VMAT dose accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels[J]. *Med Phys*, 2013, 40(11): 111722.
- [5] Interian Y, Rideout V, Kearney VP, et al. Deep nets vs expert designed features in medical physics: an IMRT QA case study[J]. *Med Phys*, 2018, 45(6): 2672-2680.
- [6] Dai GY, Zhang XB, Liu WJ, et al. Analysis of EPID transmission fluence maps using machine learning models and CNN for identifying position errors in the treatment of GO patients[J]. *Front Oncol*, 2021, 11: 721591.
- [7] Nyflot MJ, Thammasorn P, Wootton LS, et al. Deep learning for patient-specific quality assurance: identifying errors in radiotherapy delivery by radiomic analysis of gamma images with convolutional neural networks[J]. *Med Phys*, 2019, 46(2): 456-464.
- [8] 段欢欢, 李书舟, 曹瑛, 等. 基于深度学习方法预测IMRT计划射野的 γ 通过率[J]. *中国医学物理学杂志*, 2021, 38(6): 677-681.
Duan HH, Li SZ, Cao Y, et al. Predicting gamma passing rates for intensity-modulated radiotherapy fields based on deep learning method[J]. *Chinese Journal of Medical Physics*, 2021, 38(6): 677-681.
- [9] Lizar JC, Yaly CC, Colello Bruno A, et al. Patient-specific IMRT QA verification using machine learning and gamma radiomics[J]. *Phys Medica*, 2021, 82: 100-108.
- [10] Sakai M, Nakano H, Kawahara D, et al. Detecting MLC modeling errors using radiomics-based machine learning in patient-specific QA with an EPID for intensity-modulated radiation therapy[J]. *Med Phys*, 2021, 48(3): 991-1002.
- [11] Wootton LS, Nyflot MJ, Chaovaitwongse WA, et al. Error detection in intensity-modulated radiation therapy quality assurance using radiomic analysis of gamma distributions[J]. *Int J Radiat Oncol Biol Phys*, 2018, 102(1): 219-228.
- [12] Mohammadi M, Rostampour N, Rutten TP. Modification of the gamma function for the recognition of over- and under-dose regions in three dimensions[J]. *J Med Phys*, 2012, 37(4): 200-206.
- [13] 张瑞萍, 刘应龙, 张文静, 等. 基于人工智能的多模态影像辅助海马体自动勾画研究[J]. *中国医学物理学杂志*, 2022, 39(3): 390-396.
Zhang RP, Liu YL, Zhang WJ, et al. Auto-segmentation of the hippocampus in multimodal image using artificial intelligence[J]. *Chinese Journal of Medical Physics*, 2022, 39(3): 390-396.
- [14] 谷珊珊, 吴青南, 刘国才, 等. 基于深度学习融合模型的算法在宫颈癌术后患者直肠自动分割中的应用研究[J]. *医疗卫生装备*, 2020, 41(7): 30-35.
Gu SS, Wu QN, Liu GC, et al. Application of deep learning fusion model algorithm in rectal auto segmentation of postoperative patients with cervical cancer[J]. *Chinese Medical Equipment Journal*, 2020, 41(7): 30-35.
- [15] 康丽霞, 马建芬, 张朝霞. 基于多特征后期融合的声学场景分类[J]. *计算机工程与设计*, 2023, 44(1): 141-147.
Kang LX, Ma JF, Zhang CX. Acoustic scene classification based on multi feature post-fusion[J]. *Computer Engineering and Design*, 2023, 44(1): 141-147.
- [16] Huang G, Liu Z, Pleiss G, et al. Convolutional networks with dense connectivity[J]. *IEEE Trans Pattern Anal*, 2022, 44(12): 8704-8716.
- [17] Zhang Z. Improved adam optimizer for deep neural networks[C]// 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS). ACM, 2018.
- [18] 薛哲, 赵宗耀, 陈家旭, 等. 以六种疾病为例研究基于统计注意力的神经网络模型在证名诊断中的应用[J]. *北京中医药大学学报*, 2021, 44(4): 358-365.
Xue Z, Zhao ZY, Chen JX, et al. Applied research of SANN model in the diagnosis of traditional Chinese patterns with six diseases data as examples[J]. *Journal of Beijing University of Traditional Chinese Medicine*, 2021, 44(4): 358-365.
- [19] Tie X, Lam SK, Zhang Y, et al. Pseudo-CT generation from multi-parametric MRI using a novel multi-channel multi-path conditional generative adversarial network for nasopharyngeal carcinoma patients[J]. *Med Phys*, 2020, 47(4): 1750-1762.
- [20] Kimura Y, Kadoya N, Tomori S, et al. Error detection using a convolutional neural network with dose difference maps in patient-specific quality assurance for volumetric modulated arc therapy[J]. *Phys Medica*, 2020, 73: 57-64.
- [21] Ritter TA, Schultz B, Barnes M, et al. Automated EPID-based measurement of MLC leaf offset as a quality control tool[J]. *Biomed Phys Eng Express*, 2018, 4(2): 027008.
- [22] Lam D, Zhang X, Li H, et al. Predicting gamma passing rates for portal dosimetry-based IMRT QA using machine learning[J]. *Med Phys*, 2019, 46(10): 4666-4675.
- [23] Liang B, Liu B, Zhou F, et al. Comparisons of volumetric modulated arc therapy (VMAT) quality assurance (QA) systems: sensitivity analysis to machine errors[J]. *Radiat Oncol*, 2016, 11(1): 146.
- [24] Ma C, Wang R, Zhou S, et al. The structural similarity index for IMRT quality assurance: radiomics-based error classification[J]. *Med Phys*, 2021, 48(1): 80-93.
- [25] Klein EE, Hanley J, Bayouth J, et al. Task Group 142 report: quality assurance of medical accelerators[J]. *Med Phys*, 2009, 36(9 Part 1): 4197-4212.

(编辑:陈丽霞)