

❖ 医学物理与工程学

Residual network intelligent bone age assessment model established base on China-05 attention and overlay texture

GUO Zisheng, WANG Jifang*, SHEN Xiaolong, SU Peng
(College of Mechanical & Electrical Engineering, Beijing Information
Science and Technology University, Beijing 100192, China)

[Abstract] **Objective** To establish a residual network model combining China-05 attention mechanism and multi-layer texture overlay, and to explore its value for assessing bone age. **Methods** Through combinational optimization to better characterize the global features of hand bone X-ray images with less information, a multi-layer overlay texture enhancement processing layer was introduced to reduce the interference of impurity information and release computing power. The China-05 spatial attention mechanism was designed, and the standard of human hand bone development in China was introduced to make the model intelligently focusing on ROI and automatically locating and learning the key information of images. The 50-layer deep residual network was built to integrate the overlay enhancement layer and the attention mechanism, and its value for assessing bone age was evaluated. **Results** The hybrid improved deep residual network model ZH05-DL-ResNet50 was successfully established, with the accuracy of 98.14% for assessing bone age, and the mean absolute error of 0.312 year. **Conclusion** A residual network intelligent bone age assessment model based on China-05 attention and overlay texture was successfully established, which was helpful to improving the accuracy of assessing bone age in Chinese population.

[Keywords] age determination by skeleton; residual network

DOI:10.13929/j.issn.1003-3289.2022.07.025

融合中华-05 注意力与叠加纹理构建残差 网络智能模型评估骨龄

郭子昇, 王吉芳*, 沈孝龙, 苏 鹏
(北京信息科技大学机电工程学院, 北京 100192)

[摘 要] **目的** 融合中华-05 注意力机制与多层纹理叠加建立残差网络模型, 观察其评估骨龄的价值。 **方法** 通过组合寻优引入多层叠加纹理增强处理层, 以更少信息量更好地表征手骨 X 线片全局特征, 减少杂质信息干扰并释放算力。设计中华-05 空间注意力机制, 引入我国人群手骨发育标准, 使模型智能化聚焦 ROI, 并自动定位、学习图像关键信息。建立 50 层深度残差网络, 集成融合叠加增强层与注意力机制, 观察其评估骨龄的价值。 **结果** 构建的混合式改进的深度残差网络模型 ZH05-DL-ResNet50 评估骨龄准确率达 98.14%, 平均绝对误差为 0.312 岁。 **结论** 成功建立了我国人群中中华-05 注意力与叠加纹理残差网络智能模型, 用于评估骨龄准确率高。

[关键词] 年龄测定, 骨骼; 残差网络

[中图分类号] R814; TP389.1 **[文献标识码]** A **[文章编号]** 1003-3289(2022)07-1070-07

[基金项目] 国家自然科学基金(52005045)、北京市自然科学基金-海淀原始创新联合基金项目(L192018)、北京高校高精尖学科建设项目(77D2111002)。

[第一作者] 郭子昇(1996—), 男, 河北邯郸人, 在读硕士。研究方向: 深度学习, 计算机视觉, 医学图像处理。E-mail: 1361803183@qq.com

[通信作者] 王吉芳, 北京信息科技大学机电工程学院, 100192。E-mail: wjffz@sina.com

[收稿日期] 2021-12-09 **[修回日期]** 2022-03-19

准确评估骨龄具有重要意义,目前我国所用主要方法为 G-P(Greulich-Pyle)图谱法^[1]和 TW(Tanner and Whitehouse)评分法^[2],均基于欧美青少年数据制定。2006 年提出的中华-05 评分法^[3]包括 TW3-C RUS(radius, ulna, and short bone)、TW3-C Carpal(腕骨)和 RUS-CHN 3 种方法,通过观察非惯用手 X 线片中的腕、掌及指骨等骨骼形态结构推断目标年龄,更适于我国人群,但仍存在耗时长、精度低及操作复杂等缺点。随着计算机技术的发展,骨龄评估逐渐向自动化及人工智能方向发展^[4],如基于深度学习(deep learning, DL)的卷积神经网络(convolutional neural networks, CNN)^[5]。本研究通过融合中华-05 注意力与多层纹理叠加建立残差网络智能模型,观察其评估骨龄的准确性。

1 网络模型设计与训练

1.1 数据集 采用北美放射学会(Radiological Society of North America, RSNA)公开数据集中的手骨 X 线片,受试者月龄范围为 1~228 个月,含 12 611 幅训练图像、1 425 幅验证图像和 200 幅测试图像;截取前 1 376 幅训练图像原始文件作为测试集,以其余图片为训练集;按原始文件骨龄编号,以 12 个月为 1 岁,将测试集、训练集图像的新编文件归为 0~18 岁,共计 19 类。

1.2 实验平台 CPU 为 Intel(R) Core (TM) i7-9750H, GPU 为 NVIDIA GeForce GTX 1650,内存 8G; Windows10 64 位系统。编程软件包括 Tensorflow2.1.0、Keras 框架、Python3.7 语言及 PyCharm2019.3.3 编辑器。

1.3 深度残差网络模型设计

1.3.1 局部二值模式(local binary pattern, LBP)多层叠加纹理增强层 引入 LBP、圆形 LBP 及旋转不变 LBP 算子的多层叠加纹理增强处理层(图 1),计算公式依次为(1)~(3)。若计算值非整数,则以双线性插值计算插值点,见公式(4):

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c) \quad (1)$$

$$X_p = X_c + R \cos\left(\frac{2\pi p}{P}\right) \quad (2)$$

$$Y_p = Y_c + R \sin\left(\frac{2\pi p}{P}\right)$$

$$LBP_{P,R}^{rot} = \min \{ROR(LBP_{P,R}, i) \mid i = 0, \dots, P-1\} \quad (3)$$

$$f(X, Y) = [1 - X \ X] \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1 - Y \\ Y \end{bmatrix} \quad (4)$$

式(1)为 LBP 算子, (x_c, y_c) 为中心像素, i_c 为中心像素灰度值, i_p 为 P 点像素灰度值, s 为符号函数, p 为邻域 P 个采样点中第 p 个采样点。式(2)为圆形 LBP 算子,以可变半径的圆($P=16, R=2$)对近邻像素进行编码,其中 R 为采样半径, P 为第 P 个采样点, p 为采样数目。式(3)为旋转不变 LBP 算子,对圆形邻域进行二进制编码,再通过不断旋转及循环位移得到一系列初始定义值,取其最小值作为该邻域 LBP 值。

1.3.2 中华-05 空间注意力机制 基于中华-05 评分法设计空间注意力机制,包含尺、桡骨远端、第 1、3、5 掌骨、拇指近节指骨、中指和小指中节指骨、拇指、中指、小指远节指骨近端共 11 个易识别骨骼发育程度的区域。各骨骼连接处骨骺骨干宽度差距、钙化程度及融合程度所致相应手部 X 线片区域灰度值存在差异,将其作为空间注意力机制的 11 个 ROI 引入空间网络变压器^[6]与 11 个识别网络。

以中华-05 空间注意力机制根据图像特征自动搜索 11 个特定区域,资源分配倾向于关键区域,利用空间网络变压器对原始输入图像进行逐层空间区域学习,输出 11 个由更具辨识能力特征映射组成的子区域,再将其反馈于 11 个识别网络,后者在特征映射后的隐藏层加入压缩激活块^[7],并输入一幅手部 X 线片图像,随之输出一个 D 维向量;将所有网络输出连成一个 $N \times D$ 维向量组,分别设置 N 和 D 为 11 和 512,输入至全连接层中,由 softmax 层输出整图骨龄评分值(图 2)。

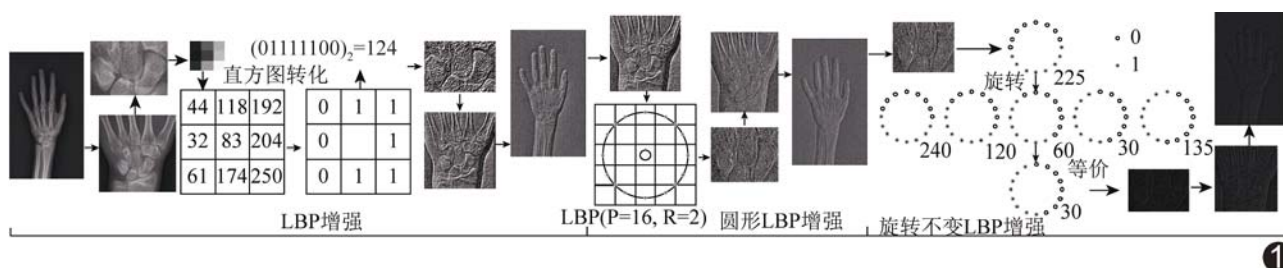


图 1 多层叠加纹理增强处理流程图

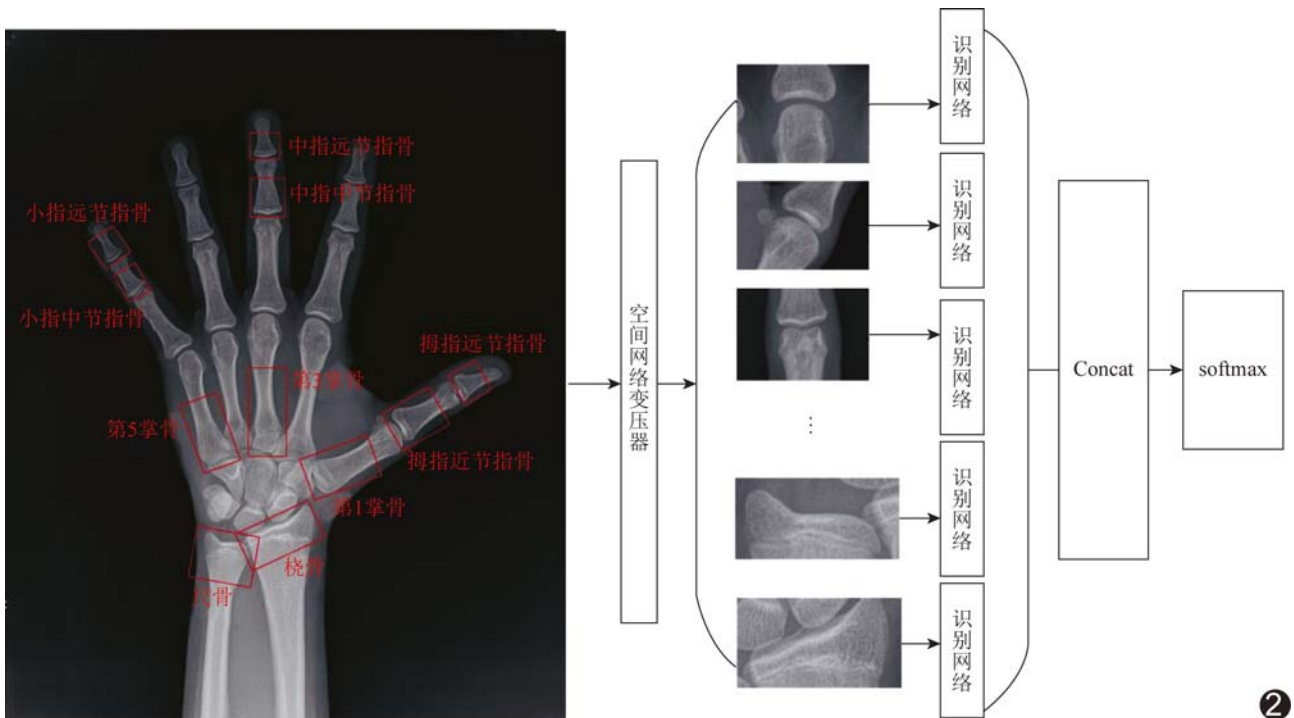


图2 中华-05 空间注意力机制设计图 (Concat:内部连接层;softmax:多分类的分类器)

1.3.3 集成化残差网络模型 参照文献[8]设计集成化残差网络模型,以残差块为基本单元,堆叠卷积层、批量处理归一化层和修正线性单元(rectified linear unit, ReLU);残差块计算公式如下:

$$x_{l+1} = f(x_l + F(x_l, W_l)) \quad (5)$$

式中, x_{l+1} 为第 $l+1$ 层残差单元输入, $F(x_l, W_l)$ 为残差函数, W_l 为第 l 层残差单元的权重, $f(\cdots)$ 为 ReLU。

构建 50 层深度残差网络,第 1 层设置卷积层,其后连接池化层,最后连接 16 个瓶颈残差模块(图 3A);分别于每个残差模块后连接 1 个池化层和 1 个激活层,最后连接全连接层,堆叠方式见图 3B。

1.3.4 ZH05-DL-ResNet50 构建集成融合多层叠加纹理增强层与中华-05 空间注意力机制的智能骨龄评估模型 ZH05-DL-ResNet50(图 4),于 50 层深度残差网络顶层,引入 LBP 多层叠加的纹理增强处理层,对原始数据集进行多层叠加纹理增强处理,并于底层加入测试集输入的 X 线片多层叠加纹理增强处理层,最后于池化层后引入中华-05 空间注意力机制。

1.4 模型训练与评价

1.4.1 数据增强 为优选 LBP 纹理增强层叠加顺序,验证引入改进处理层的有效性,对数据集依次进行 LBP 多层叠加纹理增强处理,对各增强处理图集进行遍历处理,并采用 CNN 依次设置 120 轮训练,以测试

集中平均绝对误差(mean absolute error, MAE)最低的组合顺序为最优,见公式(6)。

$$MAE = \frac{1}{N} \left(\sum_{i=1}^N |y_{true} - y_{pred}| \right) \quad (6)$$

式中, N 为样本数, y_{true} 为真实值, y_{pred} 为预测值。

1.4.2 训练策略 采用改进模型 ZH05-DL-ResNet50 对数据集进行多次训练,训练图片为三通道,设类别数为 19,每次喂入模型 16 幅图片,并调整输入图片宽、高皆为 224,设置缩放因子为 $1/255$,以增强模型泛化能力;设初始学习率为 0.000 1,配置 optimizers. Adam 优化器,结合 fit_generator 训练方式以节约内存;计算骨龄预测值和真实值的准确率(accuracy, AC)作为模型输出,并结合测试集所计算的 MAE 综合判断模型性能。

1.4.3 模型评价 对 ZH05-DL-ResNet50 模型与传统 CNN 训练时池化层、全连接层性能进行可视化比较。绘制偏置权重分布直方图,评价 ZH05-DL-ResNet50 模型与传统 CNN、50-layers ResNet、DL-ResNet50 及其他文献模型评估骨龄的性能。AC 为模型在训练集与测试集中正确分类样本占比,MAE 反映预测值与真实值的吻合程度,MAE 为 0 提示模型完美。权重表示神经元之间的连接强度,其值表示分类概率的大小。偏置值代表不断调整寻优使样本正确分类的值。

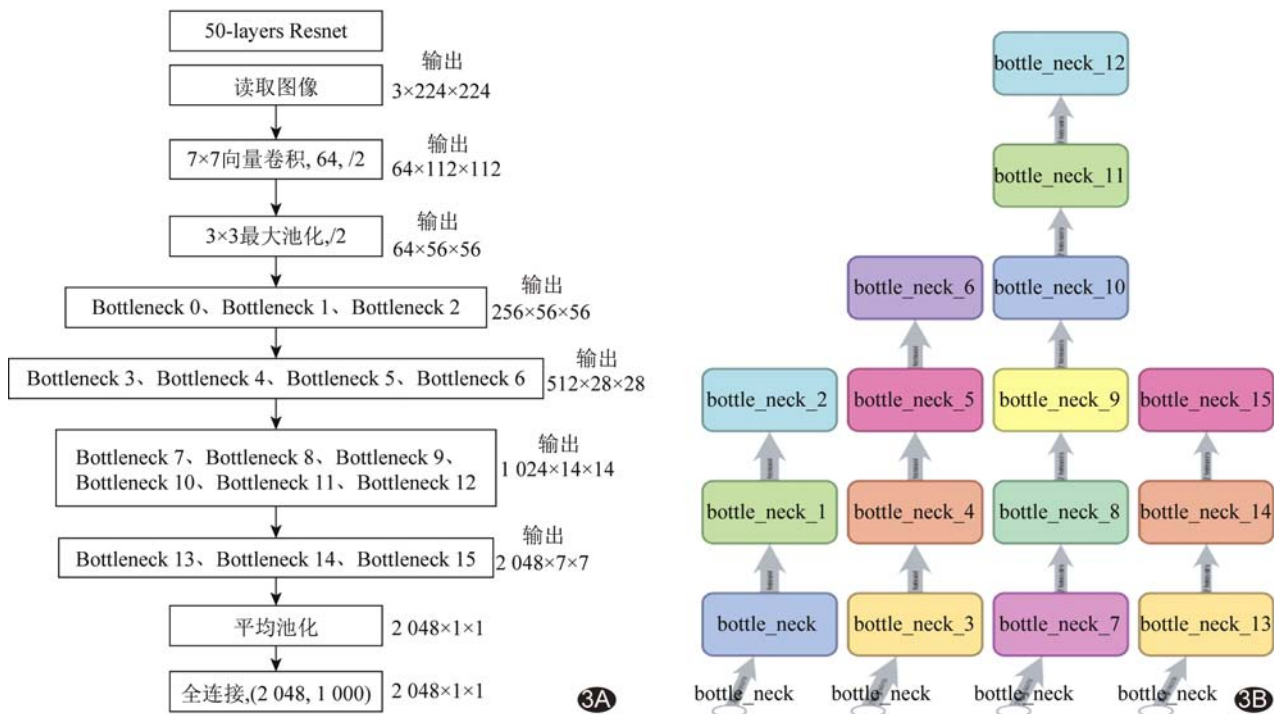


图 3 深度残差网络结构与堆叠方式 A. 50 层深度残差网络(50-layers ResNet)结构图; B. 瓶颈残差模块(bottleneck)堆叠方式可视化

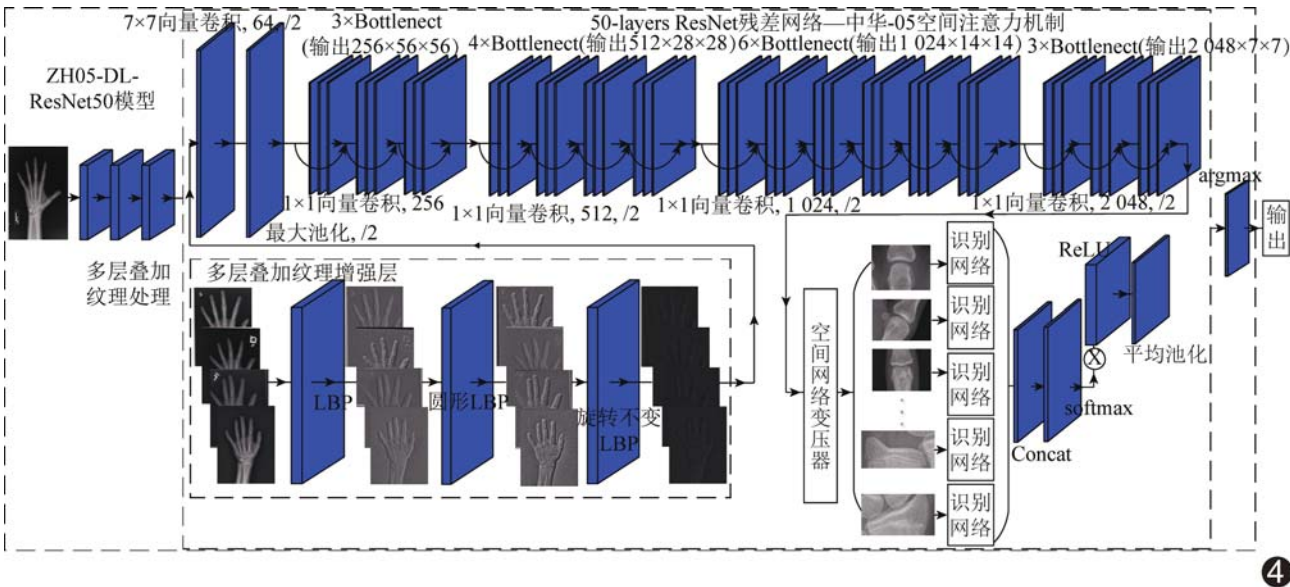


图 4 ZH05-DL-ResNet50 智能评估骨龄模型结构图

2 结果

2.1 数据增强处理层对比结果 旋转不变 LBP-LBP-圆形 LBP 组合测试集的 MAE 最低,故以此组合顺序设置多层叠加纹理增强层,见图 5。

2.2 模型可视化分析 50-layers ResNet 的 AC 较 CNN 有所提升,且 AC 曲线收敛明显加快;DL-ResNet50 的 AC 较 50-layers ResNet 进一步提升;ZH05-DL-ResNet50 曲线收敛最快,AC 最高(98.14%),

预测测试集骨龄的 MAE 计算值为 0.312 岁,预测效果极佳。见表 1 及图 6。

表 1 网络模型训练结果

网络模型	AC(%)	损失值
CNN	93.20	0.165
50-layers ResNet	96.37	0.091
DL-ResNet50	97.31	0.049
ZH05-DL-ResNet50	98.14	0.035

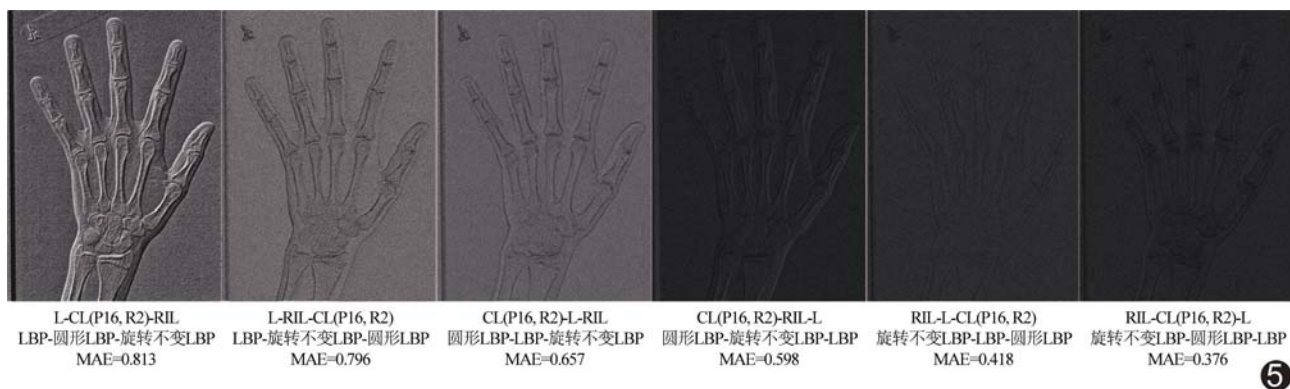


图 5 多层叠加纹理增强处理结果图

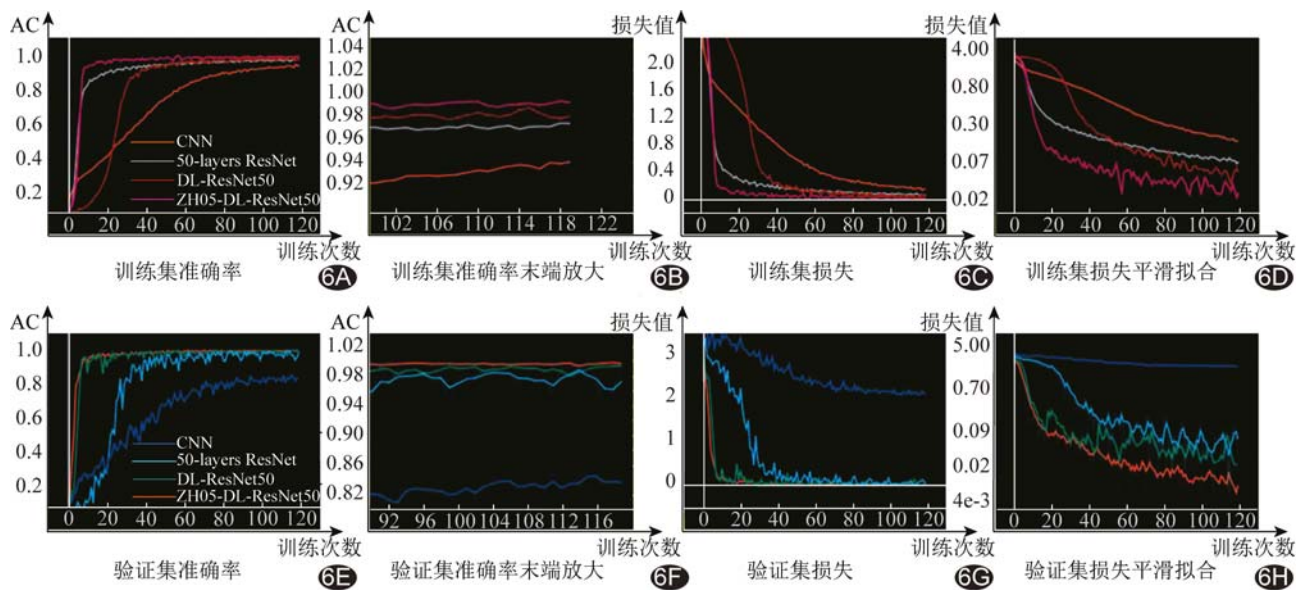


图 6 网络模型训练结果可视化分析图 A. 训练集 AC; B. 训练集 AC 末端放大; C. 训练集损失值; D. 训练集损失值平滑拟合; E. 验证集 AC; F. 验证集 AC 末端放大; G. 验证集损失值; H. 验证集损失值平滑拟合

2.3 模型评估与精度 ZH05-DL-ResNet50 批量数据均值方差期望走势波动较 CNN 明显减小,全连接层卷积核权重值皆在 0 附近颜色最深,即在 0 附近区域权重值取值频次最高,而 ZH05-DL-ResNet50 波动范围明显更小,见图 7。传统 CNN 训练过程卷积层的权重曲线缺乏结构性,规律性差;而 ZH05-DL-ResNet50 的卷积层权重曲线规律、平滑,具有结构性,见图 8。

与其他骨龄评估模型相比, ZH05-DL-ResNet50 改进模型的 MAE 更低而 AC 更高,见表 2。

3 讨论

基于 DL 的 CNN 用于检测图像和分类任务表现出色。IGLOVIKOV 等^[20]引入 CNN,提出针对整幅手部 X 线片的 DL 骨龄评估模型,因无针对性 ROI,整体评估准确性一般;SPAMPINATO 等^[12]提出 DL 骨

表 2 ZH05-DL-ResNet50 与文献中的其他模型评估骨龄结果比较

网络模型	AC(%)	MAE(岁)
CaffeNet ^[9]	—	1.575
ST-ResNet ^[10]	—	0.980
BoneXpert ^[11]	—	0.800
BoNet ^[12]	—	0.800
Inception-Bone ^[13]	—	0.730
Xception ^[14]	—	0.633
U-Net-ID-MF-HE ^[15]	—	0.453
Inception-ResNetV2 ^[16]	—	0.377
SGD ^[17]	68.303	0.423
YOLOv3-SPP-ResNet(RReLU) ^[18]	86.265	0.560
RIL-MobileNetV3 Large ^[19]	94.204	0.350
ZH05-DL-ResNet50	98.137	0.312

龄评估模型 BoNet,以 TW 法细化提取 ROI,并以涵盖 0 ~18 岁多种族人群手部 X 线片的公共数据集进行验证,

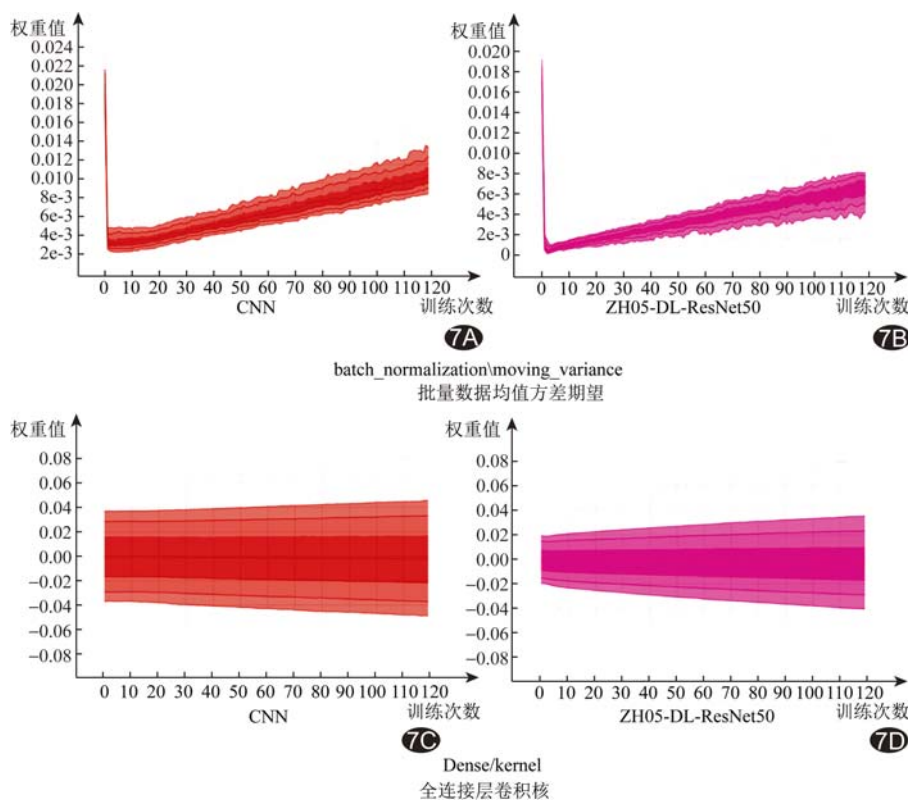


图7 训练数据分布图 A、B, CNN(A)及 ZH05-DL-ResNet50(B)池化层处理批量数据时的均值方差期望权重值; C、D, CNN(C)及 ZH05-DL-ResNet50(D)全连接层卷积核训练时的权重值

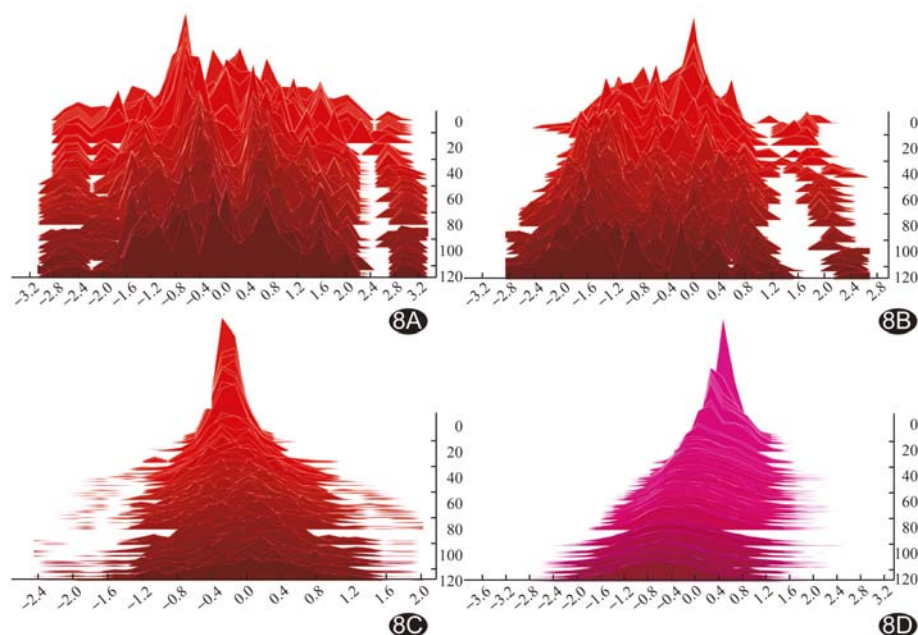


图8 偏置权重分布直方图 A. CNN; B. 50-layers ResNet; C. DL-ResNet50; D. ZH05-DL-ResNet50

法根据尺、桡骨不同发育阶段的不同特征评估骨龄。以上研究均有所缺陷,如基于整幅手部图像或某特征区域进行学习训练可致评估精度较低,依据欧洲标准的 TW 或 G-P 法训练 ROI 可能影响模型泛化能力、导致用于我国人群存在不足,训练及预测评估模型时过度关注数据集的传统数据增强手段而忽略医学图像的单通道、小对比度、复杂纹理及细粒度图像识别等特征,可致模型训练和预测效果不佳。

本研究引入 LBP 算子的多层叠加纹理增强处理层,使图像灰度不变,减少无关信息对训练及预测精度的干扰;依照中华-05 评分法设计新的空间注意力机制,针对我国人群手骨特征进行训练。相比 G-P 图谱法及 TW 评分法,中华-05 评分法所测同年龄段骨龄与生活年龄差值的平均数和中位数更接近 0 值水平线,大部分箱体相对较小,且介于 $-1 \sim 1$ 之间^[23]。通道注意力、轻量级卷积块注意力、挤压和激发注意力及空间注意力^[24]等机制用于处理自然语言、分类图像及分割语义等任务表现出色,相比原始卷积网络可减小处理高维输入数据计算负担,通过结构化选取输入子集,降低数据维度,使网络更专注于训练 ROI 和其针对性特征,且各 ROI 空间注意力权重更高,使模型训练更加聚焦;以空间网络变压器针对原图进行逐层学习,并截取分辨率适宜的子图,等同于放大手部 X 线片中的关键部位,以清晰展现局部信息,进而提取表达能力更强的特

征,为进一步处理提供基础。

其 MAE 为 0.8 岁;边增亚^[21]提出基于腕骨的骨龄评估方法,以 CNN 对腕骨 ROI 进行训练并评估骨龄;WANG 等^[22]提出基于 CNN 的骨龄分类系统,以 TW

针对性特征改进残差网络模型现已用于智能诊疗领域中的众多医学成像计算机视觉任务^[6]。残差网络

易于优化,跳跃式连接的残差块可将输入信息直接传至输出层,缓解网络深度增加带来的梯度消失及爆炸问题。为融合注意力机制与叠加纹理层,测试混合改进模型性能,本研究引入残差网络,并在相同参数下可视化分析数据分布、权重及偏置等重要指标,观测训练过程中的模型性能。

本研究成功建立融合中华-05 空间注意力机制与 LBP 多层纹理叠加的混合改进式残差网络模型,50-layers ResNet 的 AC 较传统网络模型 CNN 有所提升,表明引入残差网络可有效解决梯度消失及爆炸问题,协助提高训练性能;DL-ResNet50 的 AC 进一步提升,提示 LBP 多层纹理叠加增强层对骨龄值划分具有明显效果,可将更细致的手部骨骼纹理特征作为网络输入端;ZH05-DL-ResNet50 的 AC 最高,表明中华-05 空间注意力机制使网络更专注于训练 11 个易识别骨骼发育程度的 ROI 及其针对性特征,ROI 内骨髓及骨干宽度、钙化情况、融合程度等特有图像特征对分类训练残差网络更具针对性,使算力分配更倾向于手骨特征的重要区域,可释放网络整体计算负担,进一步降低无关信息干扰,并协助提高训练 AC。

综上,本研究成功建立的中华-05 注意力与叠加纹理的残差网络智能模型可提高评估我国人群骨龄的准确率。但本研究仅对集成模型的性能进行了技术改进,未对数据集样本分布、性别差异及网络初始学习速率加以处理,尚有一定提升空间,有待后续进一步完善。

[参考文献]

- [1] GREULICH W, PYLE S. Radiographic atlas of skeletal development of the hands and wrists[M]. Palo Alto, Stanford University Press and Oxford University Press, 1999:31-33.
- [2] HÄGG U, TARANGER J. Skeletal stages of the hand and wrist as indicators of the pubertal growth spurt[J]. Acta Odontol Scand, 1980,38(3):187-200.
- [3] 张绍岩,刘丽娟,吴真列,等.中国人手腕骨发育标准—中华 05 I. TW₃-CRUS, TW₃-C 腕骨和 RUS-CHN 方法[J]. 中国运动医学杂志, 2006,25(5):509-516.
- [4] 江光前,李红.传统方法、新兴方法及人工智能方法评估骨龄的研究进展[J]. 中国介入影像与治疗学, 2019,16(6):376-379.
- [5] 康敏,王齐艳,蒲杨梅,等.人工智能骨龄评系统评估儿童腕骨骨龄[J]. 中国医学影像技术, 2019,35(12):1804-1807.
- [6] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[EB/OL]. (2015-12-07) [2021-12-09]. <https://dl.acm.org/doi/10.5555/2969442.2969465>.
- [7] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal Mach Intell, 2020, 42(8):2011-2023.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[EB/OL]. (2016-12-12) [2021-12-09]. <https://ieeexplore.ieee.org/document/7780459>.
- [9] LEE J H, KIM K G. Applying deep learning in medical images: The case of bone age estimation[J]. Healthc Inform Res, 2018, 24(1):86-92.
- [10] 周文祥.基于 X-ray 图像的骨龄自动评估[D].成都:电子科技大学, 2018:62-66.
- [11] THODBERG H H, KREIBORG S, JUUL A, et al. The BoneXpert method for automated determination of skeletal maturity[J]. IEEE Trans Med Imaging, 2009,28:52-66.
- [12] SPAMPINATO C, PALAZZO S, GIORDANO D, et al. Deep learning for automated skeletal bone age assessment in X-ray images[J]. Med Image Anal, 2017,36:41-51.
- [13] 曾冬冬.基于 TW2 法和卷积神经网络的骨龄辅助评测技术[D].成都:电子科技大学, 2019:51-55.
- [14] 傅贤君,汪婵婵.基于深度学习的骨龄评估方法研究[J]. 电脑知识与技术, 2021,17(12):183-185.
- [15] 苏叶,李婧,徐寅林,等.手骨 X 光片骨龄预测中图像预处理的研究[J]. 南京师范大学学报(工程技术版), 2021,21(2):54-59.
- [16] 王嘉庆,梅礼晔,张俊华,等.基于深度学习的手骨 X 射线图像骨龄评估[J]. 计算机工程, 2021,47(1):291-297.
- [17] 王景樟.基于深度学习的骨龄自动评测系统的研究与实现[D].北京:北京邮电大学, 2018:33-36.
- [18] 管骏.基于卷积神经网络的骨龄辅助评测技术[J]. 科学技术创新, 2021,4:53-55.
- [19] 郭子昇,王吉芳,苏鹏.基于深度学习的智能骨龄分类器[J]. 计算机系统应用, 2022,31(6):339-346.
- [20] IGLOVNIKOV V I, RAKHLIN A, KALININ A A, et al. Paediatric bone age assessment using deep convolutional neural networks[EB/OL]. (2018-09-20) [2021-12-09]. https://link.springer.com/chapter/10.1007/978-3-030-00889-5_34.
- [21] 边增亚.基于深度卷积神经网络的自动骨龄识别方法研究[D].北京:北京交通大学, 2019:3-5.
- [22] WANG S, SHEN Y, ZENG D, et al. Bone age assessment using convolutional neural networks [EB/OL]. (2018-06-18) [2021-12-09]. <https://ieeexplore.ieee.org/document/8396189/citations# citations>.
- [23] 潘其乐,张洪,周慧康,等. Greulich-Pyle 图谱法、CHN 法和中华 05 法评估儿童青少年骨龄的比较[J]. 中国组织工程研究, 2021, 25(5):662-667.
- [24] CHEN L, ZHANG H, XIAO J, et al. SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning [EB/OL]. (2017-11-19) [2021-12-09]. <https://ieeexplore.ieee.org/document/8100150>.