

基于网格重构学习的染色体分类模型

张林^{1,2} 易先鹏² 王广杰² 范心宇² 刘辉^{1,2} 王雪松^{1,2}

摘要 染色体的分类是核型分析的重要任务之一。因其柔软易弯曲，且类间差异小、类内差异大等特点，其精准分类仍然是一个具有挑战性的难题。对此，提出一种基于网格重构学习 (Grid reconstruction learning, GRiCoL) 的染色体分类模型。该模型首先将染色体图像网格化，提取局部分类特征；然后通过重构网络对全局特征进行二次提取；最后完成分类。相比于现有几种先进方法，GRiCoL 同时兼顾局部和全局特征提取更有效的分类特征，有效改善染色体弯曲导致的分类性能下降，参数规模合理。通过基于 G 带、荧光原位杂交 (Fluorescence in situ hybridization, FISH)、Q 带染色体公开数据集的实验表明：GRiCoL 能够更好地弱化染色体弯曲带来的影响，在不同数据集上的分类准确度均优于现有分类方法。

关键词 核型分析, 染色体分类, 特征重构, 网格化

引用格式 张林, 易先鹏, 王广杰, 范心宇, 刘辉, 王雪松. 基于网格重构学习的染色体分类模型. 自动化学报, 2024, 50(10): 2013–2021

DOI 10.16383/j.aas.c210303 **CSTR** 32138.14.j.aas.c210303

A Grid Reconstruction Learning Model for Chromosome Classification

ZHANG Lin^{1,2} YI Xian-Peng² WANG Guang-Jie² FAN Xin-Yu² LIU Hui^{1,2} WANG Xue-Song^{1,2}

Abstract Chromosome classification is one of the key tasks of karyotype analysis. However, due to chromosomes are flexible hence exhibit less difference between different types while significant difference within same type, accurate classification of chromosome remains a challenging issue. In this paper, a chromosome classification model based on grid reconstruction learning (GRiCoL) is proposed. To weaken the impact of the bendy state, chromosome images are first grid-enabled for feature extraction separately. Subsequently, global features are extracted for the second time by reconstruction network, which is followed by classification. Compared with the state-of-the-art methods, the proposed GRiCoL can get more efficient discriminable features based on both local and global features, therefore can overcome the adverse effects of bandy form of chromosome with reasonable parameter scale. Experiments on public G band, fluorescence in situ hybridization (FISH) as well as Q band chromosome datasets show that GRiCoL can extract discriminative features that weaken the bending of chromosomes more efficiently, meanwhile, higher performance was obtained as compared to current classification algorithms.

Key words Karyotype analysis, chromosome classification, feature reconstruction, gridding

Citation Zhang Lin, Yi Xian-Peng, Wang Guang-Jie, Fan Xin-Yu, Liu Hui, Wang Xue-Song. A grid reconstruction learning model for chromosome classification. *Acta Automatica Sinica*, 2024, 50(10): 2013–2021

人体健康体细胞内有 23 对染色体，包括 22 对常染色体和一对性染色体。核型分析通常对分裂中期的染色体进行扫描、拍摄、处理、分割后，获得染色体核型图用于染色体的分类及异常识别，为染色体变异相关疾病的诊断及未知基因型疾病的发现提

供有力参考。因此，核型分析已广泛应用于体外胚胎的细胞遗传学分析、产前诊断以及遗传病诊断。但目前其分割、分类、识别等各个步骤均严重依赖核型专家的人工分析，耗时耗力且依赖于经验。因此，不少学者开始研究相关的自动方法。

作为核型分析的核心环节，染色体分类一直是核型分析领域的研究重点。人类正常体细胞内不同类别的染色体形态相似，但细节纹理部分差异显著；且由于染色体的非刚性特点，各染色体长短臂可以呈现出不同的弯曲状态。针对核型图开展的染色体分类识别工作经历了传统方法^[1-4]和基于卷积神经网络 (Convolutional neural network, CNN)^[5]方法两大发展阶段。传统的染色体的自动分类方法一般由 3 个步骤构成^[6-8]，包括提取染色体中心轴、根据中心轴提取分类特征以及设计分类器^[9]。其中，染色

收稿日期 2021-04-09 录用日期 2021-12-02
Manuscript received April 9, 2021; accepted December 2, 2021
国家自然科学基金 (61971422, 31871337) 资助
Supported by National Natural Science Foundation of China (61971422, 31871337)
本文责任编辑 张道强
Recommended by Associate Editor ZHANG Dao-Qiang
1. 中国矿业大学地下空间智能控制教育部工程研究中心 徐州 221116 2. 中国矿业大学信息与控制工程学院 徐州 221116
1. Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou 221116 2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116

体中心轴的提取取决于染色体的弯曲程度,并决定后续的特征提取及处理.因此传统方法在处理弯曲染色体时,往往先作拉直处理,再提取中心轴.传统算法的各个环节具备较高的可解释性,但算法整体环节多、复杂度高,分类效果严重依赖中心轴的提取. CNN 的提出为图像处理等领域开辟了新的天地^[10-12],设计合理的 CNN 模型能够自主地挖掘数据中的有用特征,完成各类复杂任务,并陆续应用于染色体图像的处理.如 Sharma 等^[13]、Swati 等^[14]提出了基于 CNN 的提取染色体特征分类方法. Sharma 等^[13]针对不同染色体的弯曲点和弯曲程度不同的问题,先用传统方法^[15]将染色体拉直,然后再送入到卷积神经网络里进行分类,对比是否进行拉直处理的实验分别获得 68.5% 和 86.7% 的分类准确率. Swati 等^[14]同样认为染色体弯曲给分类带来极大的困难.因此采用提取中心轴和众包的方式拉直染色体,然后再送入到孪生网络中进行染色体分类,未拉直与拉直的准确率分别为 68.5% 与 85.5%. 这些方法未能端到端的完成任务,且准确率远远未达到要求.而 Zhang 等^[16]提出了一种基于利用改进的高分辨率网络模型预测弯曲染色体节点并分类的多任务方法.该方法获得了 98.1% 的准确率,但是需要预先标注弯曲节点,花费大量人工成本. Qin 等^[17]提出的方法不做染色体拉直的预处理,直接在高达 87 831 幅手工标注的 G 带图上使用 Variafocal-Net 模型,可获得 98.9% 的准确度.该模型采用两个残差模型串联,第一个残差模型训练分类的同时学习一个坐标,用于从原图切分出一个局部,这个局部再提取特征进行分类.而该数据集仅残差网络 (Residual neural network, ResNet)^[18]就得到 96.9% 的准确度.由于该模型没有针对染色体弯曲问题设计,在面对无法提供海量训练数据的实际问题时,较难达到预期的性能.为应对数据量少的情况, Wu 等^[19]提出应用生成对抗网络生成染色体图像,然后再送入到 CNN 中进行分类.由于生成的数据和真实数据仍有区别,效果提升有限,其在 5 474 幅图像上的准确率为 58.9%,扩充 250 倍后可提升至 63.5%. 染色体弯曲问题可通过大量的数据集来解决,但是大数据集标注代价高昂.因此,针对小规模染色体数据集,开发一种有效提升染色体识别性能的分类方法非常重要.

染色体核型图的分类、识别任务具有类间差距小、类内差距大的特点,因此其分类任务可借鉴自然图像的细粒度分类任务^[5, 20].目前细粒度分类任务多通过仅基于图像级标签实现的弱监督分类和带有边界框、零件关键点等先验信息的强监督分类两种思路实现^[21].在弱监督分类算法方面,基于显式

特征映射的池化框架利用核函数能够较好地捕捉 CNN 提取特征间的高阶信息,提升细粒度分类精度,但局部特征的提取方法仍有待改善^[22].基于两个 CNN 特征抽取器建立的双线性模型框架以平移不变的方式对局部成对特征进行提取,获得的特征更精细,但特征之间较难建立有效的关联,且模型参数量太大,不利于模型的实际应用^[23].另有一些破坏与重构学习模型和循环注意力卷积神经网络模型^[24-25],通过强化局部特征对分类任务的贡献,更好地习得具有判别性的特征,提高分类的准确性,但也未能充分建立局部特征间的有效关联.而在强监督分类方面, Wei 等^[26]建立全卷积网络结构,基于细粒度图像的局部标注定位有区别的部分,并生成目标及对应掩码,提取辨别性较高的特征.在此基础上,进一步搭建四路掩膜 CNN 模型,聚合提取出的特征,构建高精度分类器.但强监督分类需额外标注,代价太大.

综上所述,已有的染色体分类方法和细粒度分类思想在训练样本的数量、质量及模型对图像的局部特征提取能力、局部特征的综合利用程度、计算量和最终准确度方面都存在一定的局限.为此,本文针对染色体弯曲及类内差距大、类间差距小导致的分类难题,基于微分思想建立网格重构学习 (Grid reconstruction learning, GRiCoL) 模型提取细粒度局部特征,在不增加额外标注要求前提下有效提高分类性能.结果表明, GRiCoL 可有效提高染色体分类精度,在 G 带、荧光原位杂交 (Fluorescence in situ hybridization, FISH)、Q 带三个公共图像库上的分类准确率分别达到 0.995、0.973 和 0.972.

1 本文方法

本文提出如图 1 所示的 GRiCoL 模型实现染色体的自动分类.首先将染色体图像网格化,使得骨干网络获得相对平直的染色体图像块,从而解决染色体弯曲导致的类内差距大、类间差距小问题,同时,骨干网络能够学习到更多的局部细节特征.但网格化会破坏染色体结构的完整性,因此再设计网格重构模块,强化各个分块特征之间的联系并突出具有区分性的细节特征.通过上述染色体图像网格化及网格重构两个部分,有效提高染色体分类的准确率.

1.1 网格学习

染色体柔性易弯曲的特性给其分类带来较大挑战.针对此问题,本文基于微分近似的思想,将待分类染色体图像进行网格化处理,切分成多个相对平直的局部区域,以降低染色体弯曲对特征提取带来

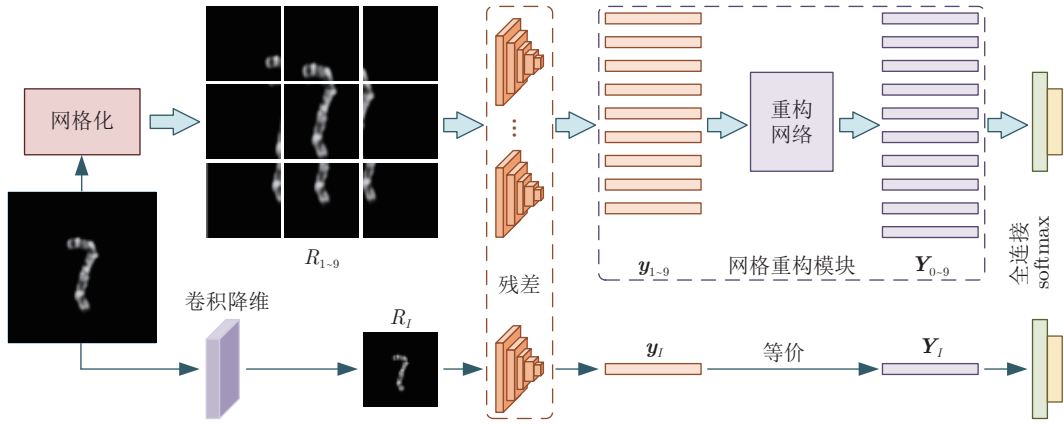


图 1 基于网格重构学习的染色体分类模型

Fig.1 Chromosome classification model based on grid reconstruction learning

的干扰.

以图 2 所示切分过程为例, 首先输入图像 I 划分为 $N \times N$ 个网格. 若 N 过小, 网格切分过粗会保留明显弯曲区域, 如图 2(a) 所示. 随着 N 的增加, 染色体弯曲的局部在网格内逐渐接近非弯曲状态, 与该染色体非弯曲状态下对应局部相比, 仅存在角度的差异, 如图 2(b) 所示. 但若 N 过大, 会产生过多的空白网格, 引入过多无效特征, 如图 2(c) 所示. 考虑染色体图像的一般分布, 本文设定 $N = 3$, 得到的 9 个网格分别用 R_k , $k \in [1, 9]$ 表示. 其中, 从左上角 R_1 开始按顺时针方向编号至 R_8 , 中间网格编号为 R_9 . 然而, 简单的网格切分会导致网格间关联信息的丢失. 因此, GRiCoL 在切割时在相邻网格邻接处设置交叠区域, 以保留各网格间的结构特征, 如图 2(d) 所示.

1.2 网格重构及特征提取

与自然图像不同, 染色体图像类间 (特别是局部区域) 差异小的鲜明特点导致任何局部网格均无法独立涵盖原染色体的主体语义信息. 因此, GRiCoL 在提取独立网格特征的同时, 也需具备感知染色体局部和整体间隶属关系的能力. 切分时特意保留的相邻网格间的重叠冗余信息一定程度上保留了全局结构. 但若特征的提取偏向重叠区域, 则无法获得全局最优解, 影响后续分类的准确率. 为此, GRiCoL 特别设计了网格重构模块, 补偿网格化导致的全局特征丢失.

GRiCoL 首先采用共享权重 Θ 的残差网络 $y = F(\cdot)$, 通过两条支路分别提取特征图后再经全局池化获得网格局部特征 $y_{1 \sim 9}$ 及全局特征 y_I .

$$y_k = F(\Theta, R_k), \quad k \in \{1 \sim 9, I\} \quad (1)$$

R_I 为降维变换后的全局图像; y_k 为残差网络提取

的特征向量, 其通道数 C 由残差网络输出层卷积核数量决定.

随后, GRiCoL 通过图 3 所示的注意力机制重构网络^[27]. 每个网格对最终分类贡献的权重由式 (2) 训练获得, 并构造特征向量 v_k .

$$v_k = \frac{1}{1 + \exp(\mathbf{w}_k \times \mathbf{y}_k + \mathbf{b}_k)} \otimes \mathbf{y}_k \quad (2)$$

其中, \mathbf{w}_k 和 \mathbf{b}_k 为网格 k 的模型权重和偏置参数; “ \otimes ”表示哈达玛乘法.

局部特征的重组通过两个分支实现. 一方面, 通过式 (3) 融合相邻网格筛选后的特征加强关联, 并进一步进行特征筛选.

$$\mathbf{Y}_k = \mathbf{w}'_k(\mathbf{y}_k + \mathbf{v}_{k-1}) + \mathbf{b}'_k \quad (3)$$

此处, \mathbf{v}_{k-1} 是 \mathbf{y}_{k-1} 的特征筛选结果, \mathbf{w}'_k , \mathbf{b}'_k 为全连接层的参数.

另一方面, $N \times N$ 个 C 通道特征向量 \mathbf{v}_k 按照切分顺序, 重组为 $N \times N \times C$ 的特征图, 再通过最大池化得到全局特征, 即

$$\mathbf{Y}_0 = \max(\mathbf{v}_k^m), \quad m \in [1, C] \quad (4)$$

1.3 损失函数设计

待分类的图像中, 染色体多处于图像的中部, 而各网格覆盖区域不同, 包含染色体的信息量会有所区别, 因此, 对分类的贡献度也就有差异. 所以, 在训练过程中, GRiCoL 采用式 (5) 设置上支路各特征的损失权重 w''_k .

$$w''_k = \begin{cases} k, & k \in \{1 \sim 9\} \\ 10, & k = 0 \end{cases} \quad (5)$$

其中, \mathbf{Y}_0 作为全局特征, 权重最大. 此外, 由于 \mathbf{v}_k 总是包含 \mathbf{v}_{k-1} 的信息, 因此特征 $\mathbf{Y}_{1 \sim 9}$ 按顺时针方向, 其损失权重依次增大. 而下支路学习整幅图像, 训练初始赋予较大权重. 因此模型采用比例因子 β

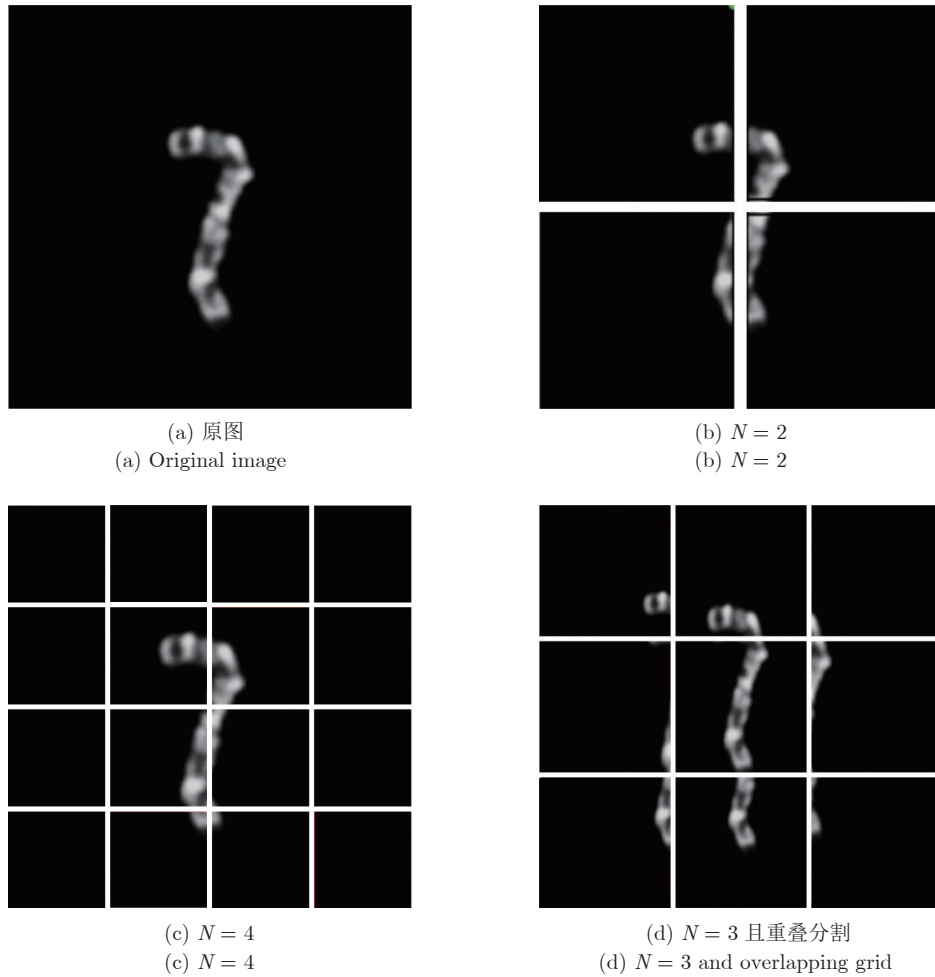


图 2 染色体图像网格化效果
Fig.2 Gridding effect of chromosome image

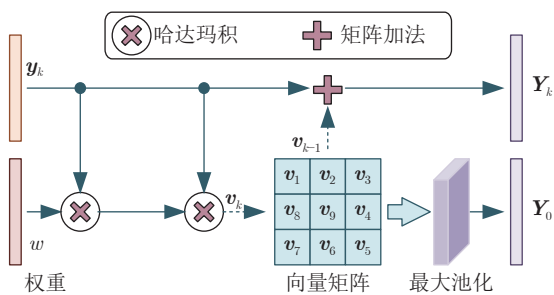


图 3 重构网络模型
Fig.3 Reconstruction network model

调节上下支路的权重比例. 训练初始下支路损失 L_I 权重大, 当训练轮次 e 达到 λ 后, 上下支路逐渐上升至相当权重. 总的损失函数如式 (6) 所示.

$$L_s = \sum_{k=0}^9 \min(1, \exp(e - \lambda) + w_k'' \times \beta) L_k + L_I \quad (6)$$

其中, L_k, L_I 均可通过式 (7) 计算交叉熵得到, 即

$$L = \frac{1}{B} \sum_{b=1}^B t_b \log(p_b) \quad (7)$$

其中, B 表示网格 k 或者全局 I 中像素的个数, t_b 表示第 b 个像素所属的真实类别, p_b 则表示模型对该像素的判别类别.

1.4 GRiCoL 的实现步骤

模型训练样本由原始输入图像 I 及其对应染色体编号 t 构成, 记为 $\langle I, t \rangle, t \in \{1, \dots, 22, X, Y\}$. GRiCoL 模型即通过函数 $\mathbf{p} = \Phi(I, \Omega)$ 将输入图像 I 映射为其属于各类染色体的概率向量 \mathbf{p} , 其中, Ω (包含 Θ) 表示分类模型中所有的可学习参数.

步骤 1. 将完整图像 I 切分成 9 个相同尺寸的网格 $\{R_k\}, k \in [1, 9]$. 同时, 通过卷积将完整图像 I 下采样至网格同尺寸, 表示为 R_I .

步骤 2. 将 $\{R_I, R_k\}$ 分别送入残差卷积单元 (各单元共享权重), 得到全局特征 \mathbf{y}_I 及网格特征向量 \mathbf{y}_k .

步骤 3. 特征向量 \mathbf{y}_k 经过重构网络, 得到重构

后的特征 $Y_k, k \in [0, 9]$.

步骤 4. 利用线性分类器对重构特征 Y_k 以及完整图像特征 $Y_I = y_I$ 进行分类, 得到 11 个分类概率向量 p .

步骤 5. 基于分类概率向量 p 完成分类并计算交叉熵损失.

步骤 6. 采用自适应动量估计优化器优化网络参数.

步骤 7. 推理时, 待分类图像 I 无需切分, 直接通过模型下支路完成染色体分类识别.

2 实验与分析

为了测试 GRiCoL 的性能及对不同显带技术的适应能力, 本文采用 3 组公开数据集对模型进行评估. 训练及测试的硬件环境为 Nvidia 2080Ti GPU, 软件环境为 Centos7 及 PyTorch 1.3.1.

2.1 实验数据库

本文分别在 G 带^[28]、FISH 图像^[29] 和 Q 带^[30] 三个染色体公共库上开展实验及分析.

考虑到不同数据库样本及残差网络模块等特点, GRiCoL 模型采用填充、线性插值等方法, 将所有图像的高、宽及通道数调整为 $448 \times 456 \times 1$, 作为 GRiCoL 模型的输入. 其中, G 带图像库包含来自 65 个细胞 (33 名男性、32 名女性) 通过 Giemsa 染料染色后的 2986 幅染色体图像, 原始尺寸为 224×224 像素, 直接通过双线性插值的方式将图像调整至 448×456 像素, 该类图像沿染色体中心轴呈现出较清晰的深浅相间带纹, 如图 4(a) 所示. FISH 图像库包含来自 91 个细胞 (30 名男性、61 名女性) 通过荧光染色成像的 4184 条染色体图像, 原始图像大小均为 100×200 像素, 首先通过补 0 扩增至 224×224 像素, 再采用双线性插值调整为 448×456 像素, 该组图像无带纹标记, 如图 4(b) 所示. Q 带图像库则包含来自 119 个细胞 (45 名男性、74 名女

性) 的通过喹吖因荧光染料染色的 5474 幅大小不一的图像, 图像首先填充至 224×224 像素, 再通过双线性插值调整为 448×456 像素, 该组图像的染色体则呈现亮暗不同的带纹, 如图 4(c) 所示.

2.2 评估方法

本文从分类准确率和特征提取能力两方面对 GRiCoL 模型进行性能评价. 模型的分类准确率 A_r 定义为式 (8)^[31].

$$A_r = \frac{T_d}{N_d} \quad (8)$$

其中, T_d 表示预测正确的正样本数量, N_d 表示测试集的样本数量.

特征有效性的评估采用了导向反向传播^[32] 以及 t-SNE (t-distributed stochastic neighbor embedding)^[33] 数据降维两种方式. 其中, 导向反向传播利用输出相对于输入的梯度信息, 限制小于 0 的梯度信息回传. 在进行特征显示时, 将原图和梯度信息相结合, 以此显示出梯度大于 0 所对应位置的特征区域, 从而凸显出有区分度的特征. t-SNE 对模型提取的 1024 维特征向量非线性降维, 将其映射到 2 维空间表示, 观察各类样本的分布.

2.3 实验及结果分析

考虑染色体图像大多为背景内容简单的灰度图, 最多包含 24 种类别, 与自然图像分类任务相比计算规模较小, GRiCoL 对骨干网络 ResNet50 进行精剪, 减少 50% 残差模块卷积核的通道数, 获得 1024 维的特征向量输出. 在模型训练过程中, 初始学习率设置为 0.001, 每次输入图像 28 幅, 权重比例 β 为 0.1, λ 为 150. 综合考虑模型性能的提升及计算负荷, GRiCoL 在对染色体图像分类推理时, 可选择只采用下支路. 与上下支路同时使用的模型相比, 分类准确率会稍有下降, 但会大大降低推理负荷.

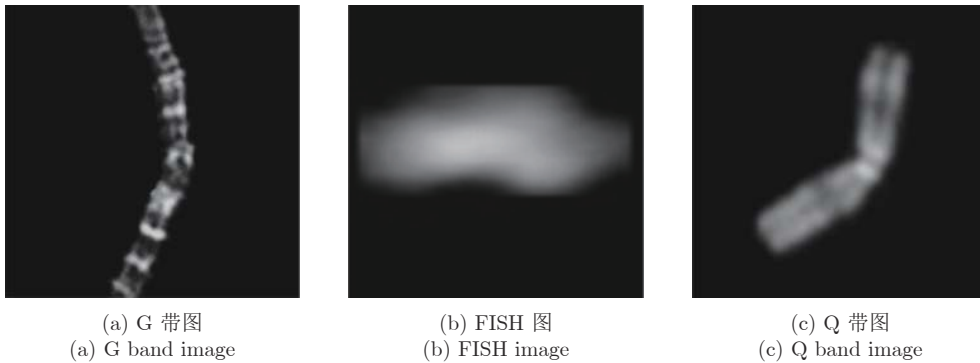


图 4 染色体图像

Fig.4 Chromosome images

本文首先通过实验比较图像切分过程中有无交叠对最终分类性能的影响, 并通过 t 检验的 p 值对比两种模型的 100 次运行结果的差异性, 结果如表 1 所示.

表 1 交叠网格设计的分类性能对比

Table 1 Classification performance comparison between grid with and without overlapping

模型	G 带	FISH	Q 带
无交叠 GRiCoL	98.1%	96.2%	95.3%
GRiCoL	99.5%	97.2%	97.3%
p 值	2.66e-22	0.52	1.71e-8

可以看出, 在 G 带和 Q 带两组数据库上, 含有交叠分块的 GRiCoL 方法与没有交叠的方法的分类性能表现出显著差异; 而 FISH 图数据库上, 两者没有明显区别. 进一步深入分析可知, G 带图和 Q 带图均呈现明显的条带特征, 因此, 通过含有交叠分块的 GRiCoL 方法强化了更多的分类相关特征信息, 因而有助于提高其分类的性能, 而 FISH 图像基于荧光技术, 局部的分类相关特征较少. 也就是说, 本文所提 GRiCoL 方法更适用于细节特征丰富的染色体图像的分类.

因此, 本文后续均采用交叠切分设计, 在此基础上, 进一步分析网格切分参数 N 对性能的影响, 结果如表 2 所示.

表 2 不同 N 数量下分类性能的对比Table 2 Classification performance comparison between grids with different N

N	G 带 (%)	FISH (%)	Q 带 (%)	Gflops	参数量 (M)
2	98.5	96.1	95.8	11.5	22.1
3	99.5	97.2	97.3	26.0	27.5
4	99.2	97.8	97.6	46.3	35.0

可以看出, 当 $N = 3$ 或 $N = 4$ 时, GRiCoL 的分类性能明显优于 $N = 2$, 但是 $N = 4$ 时, 由于网格数目的增加, 导致模型训练和推理的计算量与参数量增加. 因此, GRiCoL 模型对图像实行 3×3 交叠切分, 且每个网格均为 224×224 , 行方向的切割点设定在 112、224, 列方向的切割点设定在 120、232. 此时, 第 1 行网格的下沿和第 3 行的上沿相邻, 在避免染色体过度弯曲导致过多空白网格的同时强化了各网格间的关联.

本文以 ResNet50 作为骨干网络的算法为基线^[34-35]. 具体地, 文献 [34] 对 ResNet50 中第 4 层的 3 个 block 输出的特征图水平切分. 第 1 个 block 分别切分为 8 份和 4 份; 第 2 个 block 切分成 2 份; 第 3 个不切分, 全局平均池化后再利用卷积对通道降维, 最后将 15 个特征向量通道拼接分类; 文献 [35]

则将图像分成 4 个象限及中心, 外加完整图像分别通过 ResNet50 骨干网络提取特征, 再按照卷积通道拼接并分类. 此外, 为比较 GRiCoL 与现有同类任务的性能及重构模块的作用, 本文在 3 组图像数据上分别实现了 CIR-Net (Classification inception ResNet)^[28]、ResNet50 等 8 种方法, 并进行 5 折交叉验证实验. 其中, 网格学习模型 (GRid Learning, GRiL) 去除上支路中重构网络, 仅保留 GRiCoL 模型的完整下支路. 8 种方法在 3 个公共图像库上的分类结果如表 3 所示.

表 3 不同模型分类性能对比 (%)

Table 3 Classification performance comparison between different models (%)

模型	G 带	FISH	Q 带
基线 ^[34]	93.4	92.0	87.8
基线 ^[35]	95.3	93.4	91.7
CIRNet	96.0	83.3	86.5
ResNet50	86.4	92.6	95.3
文献 [29]	94.7	93.7	87.7
文献 [30]	—	—	94.0
GRiL	98.3	95.8	95.9
GRiCoL	99.5	97.2	97.3

可以看出, 基于网格划分的方法 GRiL、GRiCoL 通过网格局部和图像整体的细粒度特征提取, 有效克服了弯曲造成的分类难题, 提高了分类性能. 比较 GRiL 和 GRiCoL 的分类结果可以看出, 重构模块有效补偿了网格化导致的全局特征丢失, 进一步提高了各染色体图像的分类性能. 基线方法切分图像时未考虑染色体弯曲的情况及各切分块局部特征间的联系, 导致错分概率较大.

本文尝试通过导向反向传播对模型 GRiCoL、GRiL 和 ResNet50 提取的特征进行进一步解析. 图 5(a) 为对上述 3 种模型提取的 G 带图分类特征. 可以看出, ResNet50 所提取的特征与原图简单重复, 且携带有明显的噪声. 相对而言, GRiCoL 所提取的特征则更为丰富、清晰. 图 5(b) 和图 5(c) 分别为 3 种模型提取的 FISH 图和 Q 带图分类特征. 可明显看出, ResNet50 所提取的特征过分强调了图像的边缘, 但这些边缘并非有效的染色体分类特征, 而是掺杂了图像原始处理时产生的噪声. 因此, 本文方法提取的特征更加关注有区分度的细节, 是更有区分度的特征.

为进一步比较 3 种模型提取特征的分类性能, 本文基于 t -SNE 方法对各图像库中同一染色体图像的分类特征降维表示, 如图 6 所示. 可以看出, ResNet50 提取的 G 带图特征分布分散, 类间明显混合, GRiL 与 GRiCoL 提取的特征类内明显集中,

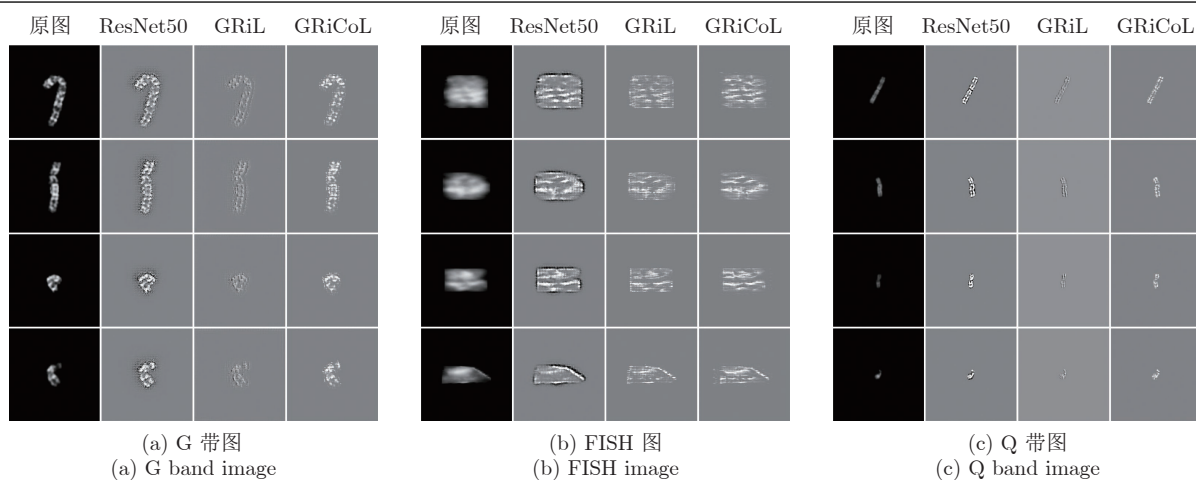


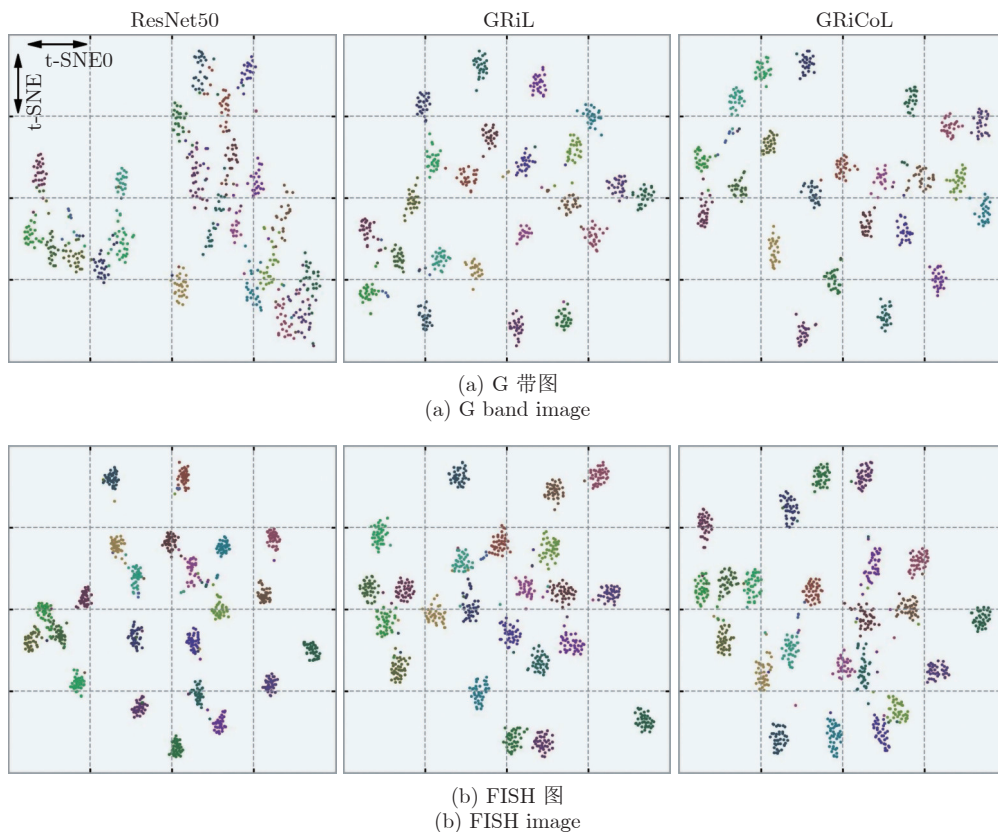
图 5 骨干网络第 50 层特征的导向反向传播可视化

Fig. 5 Visualization of guided back propagation of the 50th layer features of the backbone network

且 GRiCoL 的类间间距明显大于 GRiL. 这与表 3 中 G 带图分类 ResNet50 准确率最低、GRiCoL 最高相吻合. 而对 FISH 图像和 Q 带图像, ResNet50 所提取的特征类内间距也很小, 但类间混杂较为严重, GRiL 与 GRiCoL 尽管类内间距没有 ResNet50 提取的特征表现得那么紧凑, 但类间区别明显, 这也是 GRiCoL 模型在 FISH 图像库能够取得最佳分类性能的原因.

3 结束语

染色体由于其柔性易弯曲的特点, 目前其自动分类存在较大困难. 为解决该问题, 本文提出 GRiCoL 模型, 利用微分思想对染色体图像进行合理切分, 弱化弯曲的影响, 通过残差网络、重构网络等模块提取特征后, 完成分类模型的构建. 本文分别基于 G 带图、FISH 图和 Q 带图三个染色体公共图像库



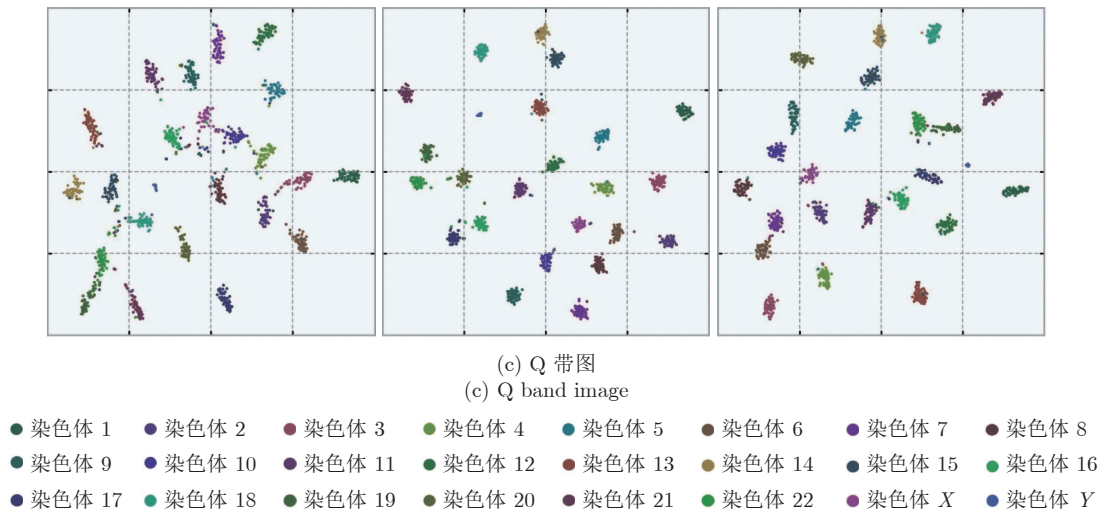


图 6 特征的 t-SNE 降维表示

Fig. 6 Representation of features dimensionality reduced by t-SNE

对 GRiCoL 模型进行训练和评估, 对比实验结果显示, 本文提出的 GRiCoL 能够有效提高染色体图像的分类准确率, 且表现出更强的泛化能力及对不同显带成像技术的适应性. 此外, GRiCoL 模型可以在较小的数据集上完成有效的训练, 所提取的特征也呈现出较好的解释性. 因此, 本文所提出的 GRiCoL 网络模型能够较好地解决弯曲染色体的分类问题, 并为染色体的端到端自动核型分析提供新的思路.

References

- Madian N, Jayanthi K B. Analysis of human chromosome classification using centromere position. *Measurement*, 2014, **47**: 287–295
- Abid F, Hamami L. A survey of neural network based automated systems for human chromosome classification. *Artificial Intelligence Review*, 2018, **49**(1): 41–56
- Poletti E, Grisan E, Ruggeri A. A modular framework for the automatic classification of chromosomes in Q-band images. *Computer Methods and Programs in Biomedicine*, 2012, **105**(2): 120–130
- Khan S, DSouza A, Sanches J, Ventura R. Geometric correction of deformed chromosomes for automatic Karyotyping. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. San Diego, USA: IEEE, 2012. 4438–4441
- Sharif Razavian A, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: An astounding baseline for recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Columbus, USA: IEEE, 2014. 512–519
- Lerner B, Guterman H, Dinstein I, Romem Y. Medial axis transform-based features and a neural network for human chromosome classification. *Pattern Recognition*, 1995, **28**(11): 1673–1683
- Ming D L, Tian J W. Automatic pattern extraction and classification for chromosome images. *Journal of Infrared, Millimeter, and Terahertz Waves*, 2010, **31**(7): 866–877
- Markou C M C, Delopoulos A. Automatic chromosome classification using support vector machines. In: Proceedings of the IEEE Information Technology, Networking, Electronic and Automation Control Conference. 2012. 1–24
- Gao D Y, Madden M, Chambers D, Lyons G. Bayesian ANN classifier for ECG arrhythmia diagnostic system: A comparison study. In: Proceedings of the IEEE International Joint Conference on Neural Networks. Montreal, Canada: IEEE, 2005. 2383–2388
- Tian Juan-Xiu, Liu Guo-Cai, Gu Shan-Shan, Ju Zhong-Jian, Liu Jin-Guang, Gu Dong-Dong. Deep learning in medical image analysis and its challenges. *Acta Automatica Sinica*, 2018, **44**(3): 401–424
(田娟秀, 刘国才, 谷珊珊, 鞠中建, 刘劲光, 顾冬冬. 医学图像分析深度学习研究方法研究与挑战. *自动化学报*, 2018, **44**(3): 401–424)
- Song Yan, Wang Yong. Multi-stage attention-based capsule networks for image classification. *Acta Automatica Sinica*, 2024, **50**(9): 1804–1817
(宋燕, 王勇. 多阶段注意力胶囊网络的图像分类. *自动化学报*, 2024, **50**(9): 1804–1817)
- Feng Jian-Zhou, Ma Xiang-Cong. Fine-grained entity type classification based on transfer learning. *Acta Automatica Sinica*, 2020, **46**(8): 1759–1766
(冯建周, 马祥聪. 基于迁移学习的细粒度实体分类方法的研究. *自动化学报*, 2020, **46**(8): 1759–1766)
- Sharma M, Saha O, Sriraman A, Hebbalaguppe R, Vig L, Karande S. Crowdsourcing for chromosome segmentation and deep classification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu, USA: IEEE, 2017. 786–793
- Swati, Gupta G, Yadav M, Sharma M, Vig L. Siamese networks for chromosome classification. In: Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW). Venice, Italy: IEEE, 2017. 72–81
- Javan-Roshkhar M, Setarehdan S K. A new approach to automatic classification of the curved chromosomes. In: Proceedings of the 5th International Symposium on Image and Signal Processing and Analysis. Istanbul, Turkey: IEEE, 2007. 19–24
- Zhang J P, Hu W J, Li S Y, Wen Y F, Bao Y, Huang H F, et al. Chromosome classification and straightening based on an interleaved and multi-task network. *IEEE Journal of Biomedical and Health Informatics*, 2021, **25**(8): 3240–3251
- Qin Y L, Wen J, Zheng H, Huang X L, Yang J, Song N, et al. Varifocal-net: A chromosome classification approach using deep convolutional networks. *IEEE Transactions on Medical Imaging*, 2019, **38**(11): 2569–2581
- He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770–778
- Wu Y R, Yue Y S, Tan X, Wang W, Lu T. End-to-end chromosome karyotyping with data augmentation using GAN. In: Proceedings of the 25th IEEE International Conference on Image Processing (ICIP). Athens, Greece: IEEE, 2018. 2456–2460
- Yang Z, Luo T G, Wang D, Hu Z Q, Gao J, Wang L W. Learning to navigate for fine-grained classification. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 438–454
- Luo Jian-Hao, Wu Jian-Xin. A survey on fine-grained image cat-

egorization using deep convolutional features. *Acta Automatica Sinica*, 2017, **43**(8): 1306–1318
(罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. 自动化学报, 2017, **43**(8): 1306–1318)

- 22 Cui Y, Zhou F, Wang J, Liu X, Lin Y Q, Belongie S. Kernel pooling for convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 3049–3058
- 23 Lin T Y, RoyChowdhury A, Maji S. Bilinear CNN models for fine-grained visual recognition. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1449–1457
- 24 Chen Y, Bai Y L, Zhang W, Mei T. Destruction and construction learning for fine-grained image recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 5152–5161
- 25 Fu J L, Zheng H L, Mei T. Look closer to see better: Recurrent attention convolutional neural network for fine-grained image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 4476–4484
- 26 Wei X S, Xie C W, Wu J X, Shen C H. Mask-CNN: Localizing parts and selecting descriptors for fine-grained bird species categorization. *Pattern Recognition*, 2018, **76**: 704–714
- 27 Fukui H, Hirakawa T, Yamashita T, Fujiyoshi H. Attention branch network: Learning of attention mechanism for visual explanation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 10697–10706
- 28 Lin C C, Zhao G S, Yang Z R, Yin A H, Wang X M, Guo L, et al. CIR-Net: Automatic classification of human chromosome based on inception-ResNet architecture. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2022, **19**(3): 1285–1293
- 29 Hu X, Yi W L, Jiang L, Wu S J, Zhang Y, Du J Q, et al. Classification of metaphase chromosomes using deep convolutional neural network. *Journal of Computational Biology*, 2019, **26**(5): 473–484
- 30 Wang X W, Zheng B, Li S B, Mulvihill J J, Wood M C, Liu H. Automated classification of metaphase chromosomes: Optimization of an adaptive computerized scheme. *Journal of Biomedical Informatics*, 2009, **42**(1): 22–31
- 31 Liu Jia-Min, Su Yuan-Qi, Wei Ping, Liu Yue-Hu. Video-EEG based collaborative emotion recognition using LSTM and information-attention. *Acta Automatica Sinica*, 2020, **46**(10): 2137–2147
(刘嘉敏, 苏远歧, 魏平, 刘跃虎. 基于长短记忆与信息注意的视频-脑电交互协同情感识别. 自动化学报, 2020, **46**(10): 2137–2147)
- 32 Springenberg J T, Dosovitskiy A, Brox T, Riedmiller M A. Striving for simplicity: The all convolutional net. In: Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: 2015.
- 33 Van der Maaten L, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, **9**: 2579–2605
- 34 Zhang Y Z, Liu S W, Qi L, Coleman S, Kerr D, Shi W D. Multi-level and multi-scale horizontal pooling network for person re-identification. *Multimedia Tools and Applications*, 2020, **79**(39): 28603–28619
- 35 Wang Ya-Mei, Wang Zhen-You. Deep convolutional neural network model based on multi-region feature. *Advances in Applied Mathematics*, 2019, **8**(11): 1753–1765
(王雅涓, 王振友. 基于多区域特征的深度卷积神经网络模型. 应用数学进展, 2019, **8**(11): 1753–1765)



张林 中国矿业大学信息与控制工程学院教授. 主要研究方向为生物信息学, 医学图像处理, 机器学习.
E-mail: lin.zhang@cumt.edu.cn
(**ZHANG Lin** Professor at the School of Information and Control Engineering, China University of

Mining and Technology. Her research interest covers bioinformatics, medical image processing, and machine learning.)



易先鹏 中国矿业大学信息与控制工程学院硕士研究生. 主要研究方向为医学图像处理.

E-mail: xianpeng.yi@cumt.edu.cn

(**YI Xian-Peng** Master student at the School of Information and Control Engineering, China University of Mining and Technology. His main research interest is medical image processing.)



王广杰 中国矿业大学信息与控制工程学院硕士研究生. 主要研究方向为医学图像处理.

E-mail: guangjie.wang@cumt.edu.cn

(**WANG Guang-Jie** Master student at the School of Information and Control Engineering, China University of Mining and Technology. His main research interest is medical image processing.)



范心宇 中国矿业大学信息与控制工程学院博士研究生. 主要研究方向为图像处理.

E-mail: xinyu.fan@cumt.edu.cn

(**FAN Xin-Yu** Ph.D. candidate at the School of Information and Control Engineering, China University of Mining and Technology. Her main research interest is image processing.)



刘辉 中国矿业大学信息与控制工程学院副教授. 主要研究方向为生物信息学, 医学图像处理, 机器学习.

E-mail: hui.liu@cumt.edu.cn

(**LIU Hui** Associate professor at the School of Information and Control Engineering, China University of Mining and Technology. His research interest covers bioinformatics, medical image processing, and machine learning.)



王雪松 中国矿业大学信息与控制工程学院教授. 主要研究方向为人工智能, 机器学习. 本文通信作者.

E-mail: wangxuesongcumt@163.com

(**WANG Xue-Song** Professor at the School of Information and Control Engineering, China University of Mining and Technology. Her research interest covers artificial intelligence and machine learning. Corresponding author of this paper.)