袁浩成，丁德锐，汤文超. 基于图增强与双分支图神经网络的肺癌患者生存预测[J]. 智能计算机与应用，2026，16(8)：105-114. DOI：10.20169/j.issn.2095-2163.24122701

# 基于图增强与双分支图神经网络的肺癌患者生存预测

袁浩成，丁德锐，汤文超
（上海理工大学 光电信息与计算机工程学院，上海 200093）

**摘 要**：生存预测是利用高分辨率的全切片组织病理图像（Whole Slide Images，WSI）来预测患者的生存风险。然而，WSI 图像中的相似组织和层次结构给局部特征及全局特征的提取带来了巨大的挑战。基于图特征增强和图神经网络的相关研究，本文提出了一种改进的三阶段生存预测方法。首先，在图像预处理与特征提取阶段，提出了基于残差卷积与标准卷积并行融合的特征提取器 Dif-ResNet，提高了图像块特征的提取能力，为构建高质量的图结构提供基础；在图构建与图特征增强阶段，基于图节点间的条件概率分布，设计了一种新颖的特征增强方法 FEM，从而为每个图节点生成更多的关联特征，有效地强化了图结构以及优化了初始图节点特征表示；在图嵌入学习与生存预测阶段，设计了融合 GAT 分支与 GCN 分支的双分支图神经网络模型 DF-GNN，在保留上下文特征依赖性的同时，增强对 WSI 局部细节特征的捕获能力。实验结果证明本文提出的生存预测方法具有更高的预测精度，相比于当前先进生存预测方法，在 TCGA-LUAD 肺腺癌数据集上的一致性指数 C-Index 提高了 1.07%，证明了提出方法的有效性。

**关键词**：生存预测；图神经网络；图增强；卷积神经网络；全切片组织病理图像
**中图分类号**：TP391 **文献标志码**：A **文章编号**：2095-2163(2026)08-0105-10

# Survival prediction for lung cancer patients based on graph augmentation and dual-branch graph neural networks

YUAN Haocheng, DING Derui, TANG Wenchao
(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract**: Survival prediction leverages high-resolution Whole Slide Images (WSIs) to estimate patients′ survival risks. However, the similar tissue structures and hierarchical organization in WSIs pose significant challenges for extracting both local and global features. Based on research on graph feature augmentation and graph neural networks, this paper proposes an improved three-stage survival prediction method. First, in the image preprocessing and feature extraction stages, a feature extractor called Dif-ResNet is proposed, which combines residual convolution and standard convolution in a parallel fusion manner, enhancing the capability to extract patch-level features and laying a solid foundation for constructing high-quality graph structures. Second, in the graph construction and feature enhancement stage, a novel Feature Enhancement Method (FEM) is designed based on the conditional probability distribution between graph nodes. This method generates additional associated features for each graph node, effectively strengthening the graph structure and optimizing the initial feature representation of graph nodes. Finally, in the graph embedding learning and survival prediction stage, a dual-branch graph neural network model called DF-GNN, integrating GAT and GCN branches, is designed. This model preserves contextual feature dependencies while improving the ability to capture local detail features in WSIs. Experimental results demonstrate that the proposed survival prediction method achieves higher prediction accuracy. Compared to state-of-the-art methods, it improves the C-Index by 1.07% on the TCGA-LUAD lung adenocarcinoma dataset, confirming the effectiveness of the proposed approach.

**Key words**: survival prediction; graph neural networks; graph enhancement; convolutional neural networks; whole slide histopathological images

**基金项目**：国家自然科学基金(62373251)。

**作者简介**：袁浩成(1999—)，男，硕士研究生。主要研究方向：深度学习图像处理；汤文超(2002—)，男，硕士研究生。主要研究方向：医学图像分割。
**通信作者**：丁德锐(1981—)，男，博士，教授。主要研究方向：深度学习图像处理，智能算法研究。Email：deruiding2010@usst.edu.cn。

**收稿日期**：2024-12-27

## 0 引 言

肺癌是最常见的肺原发性恶性肿瘤。近几十年来,肺癌的发病率和病死率一直处于全球首位,其每年造成的死亡人数远超结直肠癌、乳腺癌和前列腺癌的总和[1]。及时发现并动态监测患者的病情进展是医生及学者们亟需解决的关键问题。近年来,随着计算机技术的快速迭代,计算机辅助诊断评估技术可有效降低误诊率,为肺癌临床诊断提供了高效的辅助手段。其中,生存预测模型是使用数据分析和机器学习技术建立的评估工具,能够为患者提供风险评估和预后分层支持[2]。在计算机技术与临床医学深度融合的背景下,数字化全切片组织病理图像(Whole Slide Image,WSI)已成为生存预测和计算机辅助诊断的重要数据基础[3]。组织病理学图像分析是肺癌诊断的"金标准",其反映了肿瘤环境的异质特征,传统上被用作诊断和预后的基础。WSI 可以在不同的放大倍数下进行观察和分析并加以保存,为临床与研究提供了多维度的分析视角。

基于深度学习方法强大的表示学习能力,模型能够自动从 WSI 中提取有效、关键特征,完成下游预测任务。然而 WSI 图像的分辨率极高,通常包含千兆像素,直接输入整张 WSI 会造成显存不足、算力受限等问题。为保留高分辨率图像的信息,如文献[4-5]的处理方式,一般将 WSI 图像分割成小的图像块(如 224×224 或 256×256 像素)作为输入。通过该方式提取到的特征侧重于局部信息,如癌症肿瘤细胞的形态变化或生长模式等,这类特征在生存预测、预后任务中是非常重要的生物学标志[6]。除此之外,基于不同的肿瘤微环境,免疫细胞也具有完全不同的作用,因此还需要考虑每一个局部病理特征之间的关联,同时提取能够准确反映周围环境和局部特征的有效上下文特征。

为充分学习 WSI 图像的上下文特征,近年来部分方法尝试利用图神经网络(GNN)挖掘其中所蕴含的图结构病理特征,在肿瘤分类和预后预测等任务中取得了先进的研究成果[7-8]。以往的工作通常利用卷积神经网络(CNN)直接提取图像块特征[9-11]。考虑到 WSI 图像特殊的组织病理学结构,这些网络可能无法充分挖掘更丰富的病理特征信息。将小分辨率图像块作为输入虽能提升效率,但图像块数量庞大,无法完全缓解硬件压力。为此,文献[9]提出了超级补丁方法,筛选并保留出最具代表性的图像块来构建图。但在生成超级补丁的过程中,原始图像块的排序与筛选会导致原本与某一节点相连的邻居节点数减少,可能丢失部分特征信息,进而影响生存预测结果。同时,超级补丁基于不同图像块特征向量间的相似性选取,这对特征提取器的性能提出了更高的要求。其次,部分工作聚焦于图结构特征的提取以获得高质量的初始特征表示。如文献[12]提出了局部增强方法,根据图节点与其邻居节点的分布,生成增强的图特征,丰富了初始节点表示。在图嵌入学习阶段,整体模型结构对捕获 WSI 局部与全局病理特征信息同样至关重要。文献[13]基于 GCN 聚合图像块特征信息,提升对局部细节特征的学习。文献[14]通过改进的图卷积操作使其中的消息传递机制能有效地在全局范围内进行层级特征聚合,有效捕获了全局特征信息。然而这些方法都只关注了局部或全局特征,缺乏对 WSI 全局与局部细节特征的互补。

为解决上述问题,本文提出基于图增强和双分支图神经网络的肺癌生存预测方法,主要的贡献如下:

(1) 设计了一种特征提取器 Dif-ResNet,采用并行的差分卷积与标准卷积构成的融合卷积,提升对 WSI 图像块高频信息和低频信息的提取能力,增强对 WSI 相似组织特征的捕获,增强了模型的表示能力和泛化能力;

(2)设计了基于条件分布与注意力融合的特征增强模块 FEM,在给定中心节点的前提下学习其相邻节点特征的条件分布,使得图节点(尤其是邻居节点数较少的节点)能够生成关联的特征向量,有效增强初始图节点特征表示;

(3)设计了基于 GAT 分支与 GCN 分支的双分支图神经网络 DF-GNN,充分捕捉 WSI 全局上下文特征并强化 WSI 局部细节特征的学习,实现信息互补,提升整体模型的学习能力。

(4)所提出的生存预测方法在公开的癌症数据集 TCGA-LUAD 上开展实验,相比一些先进的生存预测方法,在标准的评估指标 C-Index 和风险分组曲线上表现更好。

## 1 研究设计方案

基于图神经网络,本文提出的生存预测方法主要包括 3 个阶段:第一阶段(WSI 图像预处理与特征提取)构造了一种差分卷积与标准卷积的并行结构,并替代传统 ResNet 残差块中的 3×3 标准卷积,建立了新的特征提取器 Dif-ResNet,并行卷积结构可以从具有相似组织层次结构的 WSI 图像补丁中

提取高频与低频信息。第二阶段(图结构构建与图特征增强)对 Dif-ResNet 所提取到的特征进行计算、排序与筛选,生成具有代表性的超级补丁以缓解硬件压力;随后设计了一种新颖的特征增强模块 FEM,该模块在给定中心节点特征下学习其相邻节点特征的条件分布,为每个图节点生成更多的关联特征,强化了图特征表示。第三阶段(图嵌入学习与 WSI 生存预测),设计了双分支图神经网络模型 DF-GNN:其中 GAT 分支由三层 GAT 和残差连接构成,充分提取 WSI 图的上下文组织病理学特征;GCN 分支采用三层 GCN 聚合局部节点信息。最后,通过学习到的融合信息输出最终预测结果。整体框架如图 1 所示。

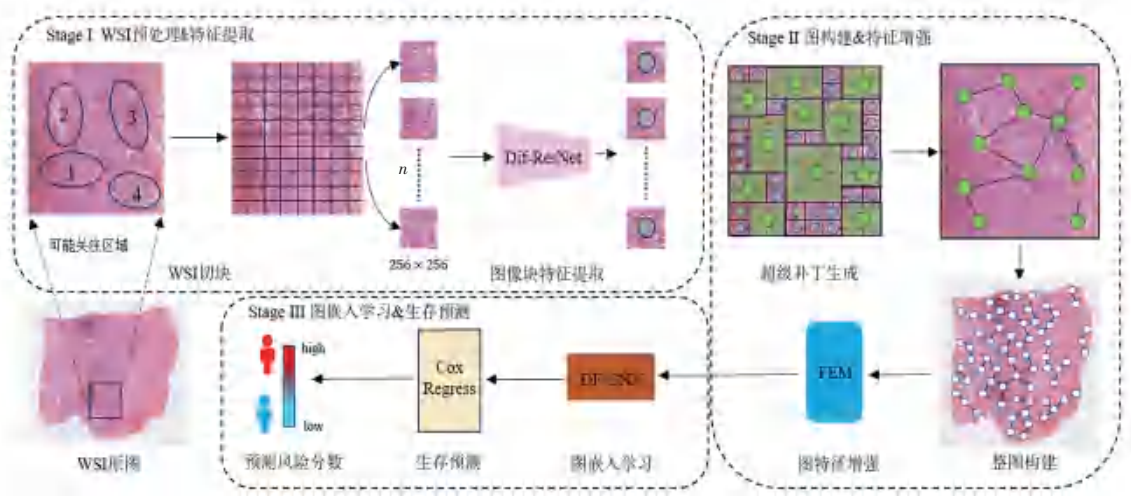



**图 1　生存预测方法整体流程**

**Fig. 1　The overall workflow of the survival prediction method**

### 1.1　基于差分卷积的特征提取器 Dif-ResNet

在 CNN 中,通常使用标准卷积(Vanilla Convolution, VC)对特征进行提取和学习。标准卷积一般强调低频信息(如颜色的变化),但对高频细节的表征能力较弱。与在视觉外观上具有复杂特征的自然图像或 CT 图像相比,全视野切片图像 WSI 往往以重复的结构和组织为主,呈现出一定的相似性,且致病区域的边缘、轮廓等高频信息可能与正常组织细胞结构存在相似,清晰的轮廓也往往难以勾勒。为了能在结构相似的场景下捕获更加丰富、清晰的纹理信息,受文献[15]启发,本文使用能够增强细节的融合卷积(FusionConv)替代 ResNet50[16] 中的 3×3 标准卷积层,从而设计了一个具有更优秀特征提取能力的特征提取器 Dif-ResNet。其依靠差分卷积与标准卷积的互补性,提高了低频和高频信息的提取能力,获取更加丰富的低级视觉特征,如纹理信息、边缘信息等,结构如图 2 所示。

具体地,ResNet50 中的一个残差块由两个 1×1 卷积和一个 3×3 卷积构成,而融合卷积并行部署了 5 个卷积层(包括中心差分卷积 CDC、角度差分卷积 ADC、水平差分卷积 HDC、垂直差分卷积 VDC 和标准卷积 VC),并替换残差块中的标准 3×3 卷积。对于输入特征 $x$,将通过差分卷积 DCs 计算得到 4 组特征 $x_1^{'}, x_2^{'}, x_3^{'}$ 和 $x_4^{'}$ 与标准卷积计算得到的特征 $x_5^{'}$ 相加得到融合卷积最终的输出。具体表示为:

$$\begin{cases} x_1^{'} = \mathrm{ADConv}(x) \\ x_2^{'} = \mathrm{CDConv}(x) \\ x_3^{'} = \mathrm{VDConc}(x) \\ x_4^{'} = \mathrm{HDConv}(x) \\ x_5^{'} = \mathrm{VConv}(x) \end{cases} \tag{1}$$

在 ResNet50 的每一个残差块中进行上述操作后,将获得的特征图 $X_{\mathrm{Fconv}}$ 进行全局平均池化,再通过全连接层得到分类输出 $X_{\mathrm{out}}^{'}$,计算过程表示为:

$$X_{\mathrm{out}}^{'} = \mathrm{FC}(\mathrm{GAP}(X_{\mathrm{Fconv}})) \tag{2}$$

### 1.2　图结构构建 & 图特征增强

提取每个图像块的特征后,可得到与图像块数量相等的一维特征向量,其大小为 $F \in R^{1\times C}$。由于 WSI 图像的分辨率极大,其对于图形处理器 GPU 等计算机硬件有潜在的高要求,受文献[9]的启发,引入超级补丁方法进行图结构构建,结合图像块之间的空间距离和余弦相似度[17],删选最具有代表性的图像块作为图节点并生成边,这一操作可在适当压缩 WSI 图像的同时,不丢失其上下文特征,可有效降低实验的硬件与计算成本。

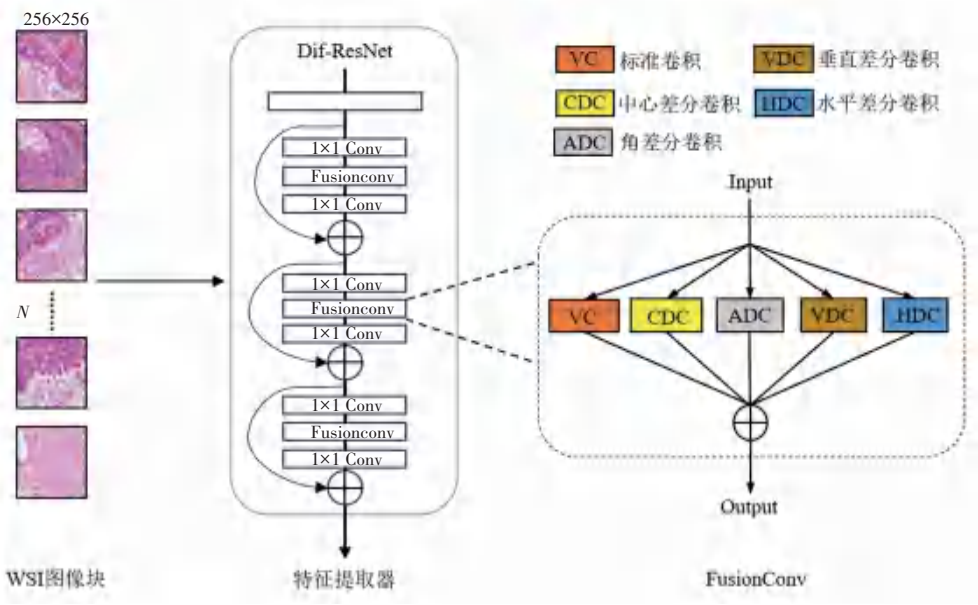



**图 2 基于差分卷积的特征提取器 Dif-ResNet 结构图**

**Fig. 2 Structural diagram of the Dif-ResNet feature extractor based on differential convolution**

然而,通过融合相似图像块构建图结构,会导致图节点的相邻节点特征信息被削弱,可能降低整体图结构质量。为解决这一问题,受文献[12]启发,在给定中心节点条件下,学习其相邻节点特征的条件分布,为图节点生成关联的特征向量,再通过注意力将原始图特征与生成图特征融合,使节点能够在局部域中生成更加丰富的特征,从而增强图初始节点特征的表达能力。其结构如图 3 所示。

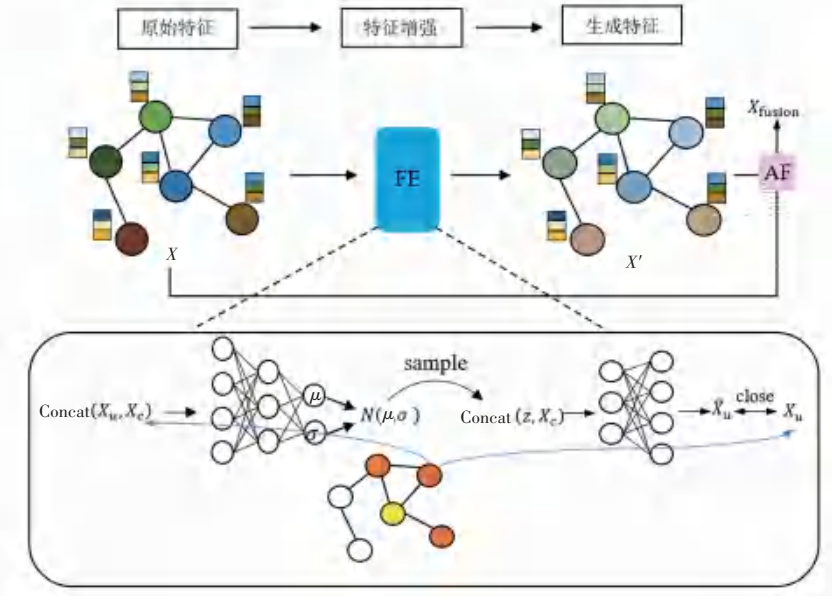



**图 3 图特征增强模块(FEM)结构**

**Fig. 3 Structural diagram of the graph feature enhancement module(FEM)**

具体地,在学习过程中利用条件变分编码器学习给定中心节点 $c$ 的相连邻居 $u(u \in N_c)$ 的节点特征的条件分布。由于 $X_u$ 的分布依赖中心节点特征 $X_c$,因此将 $X_c$ 作为条件变量。依据文献[18]的方法,隐变量 $z$ 是由先验分布 $p_\theta(z|X_c)$ 生成的,表示节点 $c$ 与其邻居节点表示的分布规律。$X_u$ 则由以 $z$ 为条件的生成分布 $p_\theta(X|X_c,z)$ 生成,由此可得:$X_c:z \sim p_\theta(z|X_c)$,$X_u \sim p_\theta(X|X_c,z)$。分别设 $\phi$ 为变分参数,$\theta$ 为生成参数。变分下界(Evidence Lower Bound, ELBO)可表示为:

$$\mathcal{L}(X_u,X_c,\theta,\phi) = -\mathrm{KL}(q_\phi(z|X_u,X_c) \| p_\theta(z|X_c)) + \frac{1}{L}\sum_{l=1}^{L}\log p_\theta(X_u|X_c,z^{(l)}) \quad (3)$$

其中，$z^{(l)}=g_\phi(X_c,X_u,\epsilon^{(l)})$，$\epsilon^{(l)}\sim N(0,I)$。需要注意的是，本文为所有节点共享训练一个增强特征生成器。在训练阶段，目标是使用相邻节点对 $(X_c,X_u,u\in N_c)$ 作为输入来最大化 ELBO，即公式(3)。在生成阶段，使用节点特征 $X_c$ 作为条件并采样了一个隐变量 $z\sim N(0,I)$ 作为解码器的输入。最终可以生成与节点 $c$ 相关联的增强节点特征向量 $\boldsymbol{X}'\in R^{NxC}$。

得到的增强特征图 $\boldsymbol{X}'$ 提供了潜在的补充信息，为了充分整合特征的互补优势，考虑利用注意力融合 AF 对原始特征图和增强特征图进行融合。融合过程表示为：

$$X_{\text{fusion}}=\sigma(W_1X+W_2\boldsymbol{X}'+\tau) \tag{4}$$

其中，$W_1$、$W_2$、$\tau$ 表示可学习参数；$X$ 表示原始图特征；$\boldsymbol{X}'$ 表示增强图特征；$\sigma$ 是激活函数。通过注意力机制可以自适应地分配原始图特征与增强图特征的权重，动态挖掘两类特征在生存预测中的重要性，保留有效关键信息、过滤冗余或无关特征的干扰，以此构建更具判别力的图结构表示。

### 1.3　双分支图神经网络 DF-GNN

在图嵌入学习阶段，先通过 FEM 模块生成增强图特征，利用特征融合策略将原始图特征与增强特征相融合，再用三层多层感知机(MLP)[19]对输入特征进行处理，将特征向量降维至 200 维，并输入 DF-GNN 模型，如图 4 所示。对于 WSI 图的上下文特征学习，考虑使用 GNN，其在处理具有非欧几里得结构的图数据时具有显著的优势和性能。为充分捕捉上下文特征并最大化 GNN 的学习能力，本文采用了图注意力网络(Graph Attention Network，GAT)[20]。该网络可以学习每条边的注意力分数，在聚合领域特征的过程中区分节点的权重。具体而言，该模型由三层 GAT 组成，每层使用两个注意力头，更广的邻域范围捕捉上下文环境特征。此外，为防止过拟合，每一层都使用 LayerNorm 进行归一化，并采用 PReLU 激活函数作为激活层，同时引入残差连接重新利用早期的输入特征，提高特征利用率和模型表达能力，以缓解梯度消失和梯度爆炸等问题。

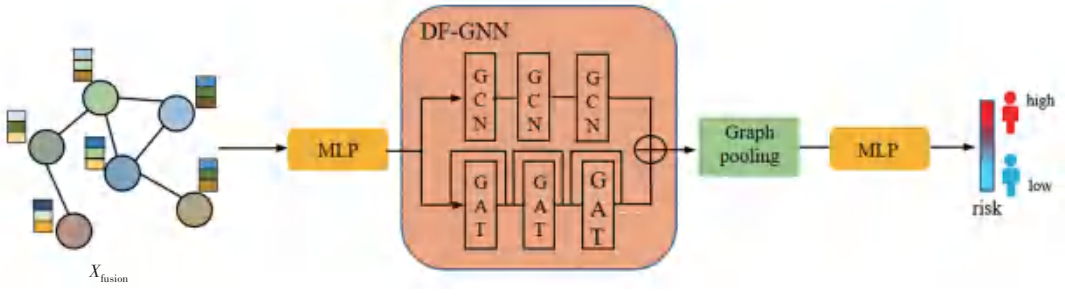



**图 4　图嵌入学习及生存预测结构**

**Fig. 4　Structural diagram of graph embedding learning and survival prediction**

需要注意的是，GAT 层的输入 $h$ 是从 MLP 的最后一层提取的组织病理学特征 $h_i$，其中 $N$ 是节点的数量，$F_n$ 是特征数，可表示为：

$$h=\{h_1,h_2,\cdots,h_N\},h_i\in R^{F_n} \tag{5}$$

此外，本文通过可学习的参数将距离 $d$ 和角度 $\theta$ 嵌入到超级补丁的边特征中，可表示为：

$$d=\{d_{11},d_{12},\cdots,d_{NN-1},d_{NN}\},d_{ij}\in R^{F_e} \tag{6}$$

$$\theta_{ij}=\{\theta_{11},\theta_{12},\cdots,\theta_{NN-1},\theta_{NN}\},\theta_{ij}\in R^{F_e} \tag{7}$$

其中，$F_e$ 表示可学习参数的数量。对于 WSI 图像而言，每个组织病理学特征的空间位置至关重要。因其反映了不同组织细胞间的层次结构和连通性，从而在计算注意力分数时，可基于这两类特征在节点间传递位置信息。GAT 的注意力分数计算方法可表示为：

$$\alpha_{ij}=\frac{\exp(\text{LeakyReLU}(a_s^{\mathrm{T}}W_sh_i+a_r^{\mathrm{T}}W_rh_j+a_d^{\mathrm{T}}W_dd_{ij}+a_a^{\mathrm{T}}W_a\theta_{ij}))}{\sum_{k\in N_i}\exp(\text{LeakyReLU}(a_s^{\mathrm{T}}W_sh_i+a_r^{\mathrm{T}}W_rh_k+a_d^{\mathrm{T}}W_dd_{ik}+a_a^{\mathrm{T}}W_a\theta_{ik}))} \tag{8}$$

其中，$W_s$、$W_r$、$W_d$、$W_a$ 分别是层中源节点、接收节点、距离和角度特征的初始线性变换；$a_s$、$a_r$、$a_d$、$a_a$ 分别表示源节点、接收节点、距离和角度特征的为构建注意力打分的单层前馈网络参数；$N_i$ 表示节点 $i$ 的邻居节点集群。利用得到的归一化注意力系数对邻居节点特征加权聚合，并使用 PReLU 函数增强非线性表达能力，通过引入残差连接利用早期的输入特征，增强全局上下文表示能力，得到的每一层 GAT 输出表示为：

$$\begin{cases} h_i^{(n)} = \sum_{j \in N_i} (a_{ij} \boldsymbol{W} h_j) \\ h_{\text{out}}^{(n)} = \text{PReLU}(h_i^{(n)}) + h_i^{(n-1)} \end{cases} \tag{9}$$

其中，$\boldsymbol{W}$ 是可学习参数矩阵；$h_i$ 是节点 $i$ 聚合后的特征；$h_{\text{out}}$ 是融合前一层特征后的输出；最后将每层输出与原始特征拼接得到最终输出特征，可表示为：

$$X_{\text{out1}} = \text{cat}(h_{\text{out}}^{(0)}, h_{\text{out}}^{(1)}, h_{\text{out}}^{(2)}, h_{\text{out}}^{(3)}) \tag{10}$$

此外，为加强对 WSI 图局部细节特征的表示能力，本文引入基于 GCN[21] 的辅助分支。GCN 通过图卷积操作聚合邻居节点特征，能够有效捕获局部范围内的组织病理特征的信息。具体而言，采用深度为 3 的 GCN 分支，由于其聚合函数引入了拉普拉斯矩阵，各节点之间特征的信息传递较快，导致 2-3 层后 GCN 节点间特征的捕获能力衰减[22]，因此设置三层 GCN 充分提取局部细节病理特征。其节点聚合过程可表示为：

$$\begin{cases} H_i^{(0)} = h_i \\ H_{\text{out}}^{(l)} = \text{PReLU}(\widetilde{\boldsymbol{D}}^{-\frac{1}{2}} \widetilde{\boldsymbol{A}} \widetilde{\boldsymbol{D}}^{-\frac{1}{2}} H_{\text{out}}^{(l-1)} \boldsymbol{W}^{(l)}), l \geqslant 1 \end{cases} \tag{11}$$

其中，$h_i$ 是初始输入图特征，$H_{\text{out}}^{(l)}$ 表示第 $L$ 层 GCN 输出特征；$\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{I}$ 表示加入自连接的邻接矩阵；$\widetilde{\boldsymbol{D}}$ 表示度矩阵；$\boldsymbol{W}^{(l)}$ 表示第 $l$ 层的可学习权重矩阵；PReLU(·) 是非线性激活函数。最后同样将每层输出与原始特征拼接，可表示为：

$$X_{\text{out2}} = \text{Cat}(H_{\text{out}}^{(0)}, H_{\text{out}}^{(1)}, H_{\text{out}}^{(2)}, H_{\text{out}}^{(3)}) \tag{12}$$

GAT 分支通过将角度、距离融入节点的边特征和跳跃式信息融合结构，对组织病理特征进行全局上下文感知学习，GCN 分支则补充局部细节组织病理特征。通过融合双分支特征信息，使整体网络在捕获局部和全局特征的同时，兼顾计算效率与鲁棒性。最终将两个支路的特征信息相加，再通过全局平均池化和 MLP 后处理，得到融合输出特征，其表示为：

$$X_{\text{final}} = \text{MLP}(\text{GAP}(X_{\text{out1}} + X_{\text{out2}})) \tag{13}$$

### 1.4 损失函数

生存预测任务中常常存在删失问题。删失是指部分患者在研究结束时，未能观测到目标事件（如死亡或疾病复发）的确切发生时间，这类数据称为删失数据。删失数据的存在会导致生存分析结果的偏倚，若未妥善处理，将影响模型的预测准确性和稳健性。为解决该问题，本文采用负对数似然函数[23]，其表达式为：

$$L_{\text{NLL}} = \sum_{i=1}^{M} \delta_i (\log \sum_{j \in \{j: r_j^g \leqslant r_i^g\}} \exp(r_j^p) + (-r_i^p)) \tag{14}$$

其中，$r_i^p$ 和 $r_i^g$ 分别表示预测结果与真实值；M 表示可比较匹配的患者对数量；$\delta_i$ 表示删失事件指示符，若事件发生则值为 1，否则为 0。该函数的意义在于通过拟合生存时间的概率分布，能够生成更精细的个体化生存预测。

## 2 实验及结果分析

### 2.1 数据集

本文在 TCGA-LUAD 公开数据集[24] 上进行训练与评估，在特征提取阶段，Dif-ResNet 的预训练数据集为结肠癌和肺癌组织病理学图像 LC25000[25]。

LC25000 数据集包含 25 000 张来自肺癌和结肠癌患者的高分辨率组织病理切片图像，图像大小均为 768×768，其主要划分为 5 类，分别为结肠腺癌、良性结肠组织、肺腺癌、肺鳞状细胞癌、良性肺组织。

TCGA-LUAD 是由美国国家癌症研究所和美国国家人类基因组研究所于 2006 年联合启动的 TCGA 项目中的癌症类型，包含与肺腺癌相关的基因组属性和临床信息。该数据集记录了 443 位肺腺癌患者的 H&E 染色组织图像，其中记录的死亡患者数为 158 例，共有 541 张 WSI 图像。实验中将其随机划分为训练集（70%）、验证集（10%）、测试集（20%），并采用 5 次交叉验证进行评估。

### 2.2 实验环境

#### 2.2.1 实施细节

图像预处理：在分割出感兴趣区域（简称 ROI）后，使用 Otsu 阈值法[26] 过滤掉 WSI 图像中的无关背景区域，实验过程中 Otsu 掩码参数设置为 0.65。然后将 40 倍放大倍数下的 WSI 切成大小为 256×256 的图像块集群。

实验硬件环境为搭载 NVIDIA GeForce RTX 4060 Ti（16 GB）显卡，64 位 GPU、Intel Core i5-13600KF3.50 GHz 处理器的设备；软件环境采用 Pytorch 深度学习框架搭建。训练过程中，选择 Adam 优化器更新参数，实验在批量大小为 8 的训练数据上运行 200 个 epoch，初始学习率为 0.000 1，权重衰减率为 0.000 05，此外在构建超级补丁时相关阈值为 0.65。当验证集的一致性指数在一段时间后不再增加时终止训练。

#### 2.2.2 评估指标

生存预测的评估指标采用一致性指数（C-Index）[27]，用于衡量模型对事件发生时间排序的预测能力，评估模型预测的风险值与实际生存时间之

间的一致性。具体而言，若预测生存率高的患者的生存时间长于生存概率低的患者，则认为该组预测结果与实际结果一致，其公式为：

$$C_{\text{ind}} = \frac{1}{M}\sum_{i:\sigma_i=1}\sum_{j:t_i<t_j}\xi(\hat{y}_i(x_i) > \hat{y}_j(x_j)) \quad (15)$$

其中，$M$ 表示所有可能的患者比较对；$\xi(\cdot)$ 为指示函数；$C_{\text{ind}}$ 的范围为 (0,1)，该指标的值越大，预测的一致性越强。

## 2.3　结果与分析

### 2.3.1　C-Index 结果分析

为验证本文生存预测整体框架的有效性，本文将其与多种先进的模型方法进行对比，对比方法包括：Attention MIL[28]、BDOCOX[4]、PatchGCN[14]、SeTranSurv[29]、WSISA[10]、DeepGraphSurv[13]、TEA-Graph[9]和 HGSurvNet[11]。为保证实验对比的公平性，所有实验均采用相同的训练设置。

基于 TCGA-LUAD 数据集的一致性指数 C-Index，结果见表 1，其中，Standard deviation 代表标准差，加粗表示最优结果。

**表 1　不同方法在 TCGA-LUAD 数据集上的结果**

**Table 1　Results of different methods on the TCGA-LUAD dataset**

| 网络模型 | C-Index | Standard deviation |
|---|---|---|
| WSISA(2017) | 0.584 | 0.033 |
| DeepGraphSurv(2018) | 0.624 | 0.040 |
| Attention MIL(2018) | 0.620 | 0.048 |
| BDOCOX(2020) | 0.613 | 0.035 |
| SeTranSurv(2021) | 0.635 | 0.032 |
| PatchGCN(2022) | 0.667 | 0.044 |
| HGSurvNet(2023) | 0.699 | 0.038 |
| TEA-Graph(2023) | 0.704 | 0.040 |
| 本文 | 0.715 | 0.039 |

从表中数据对比可知，基于 GNN 的生存预测方法普遍优于基于 CNN 的方法和基于多实例学习的方法。表明通过构建图结构可以充分捕捉 WSI 图像块的全局依赖性，并通过节点和边的多层信息传播逐步聚合领域节点特征，提高对肿瘤微环境等病理特征的表征能力，同时 GNN 在处理高维、异质性强的 WSI 图像时表现出更好的鲁棒性。在同样基于 GNN 的研究方法中，本文提出的生存预测方法框架在 TCGA-LUAD 数据集上取得了最好的效果，相比于近期表现优异的 TEA-Graph 提高了 1.1%。

### 2.3.2　预后风险分组

除评估分析 C-Index 指标外，预后风险分组曲线也是生存预测任务中常用的评估可视化工具，本节进一步分析该曲线并验证所提出方法的有效性。

预后风险分组是生存预测分析中常用的标准。若患者的预测风险大于临界值，则将患者分为高风险组(短期幸存者，存活时间更短)；否则，将其分类为低风险组(长期幸存者，存活时间更长)。本文将提出的生存预测方法与部分现有方法在 TCGA-LUAD 数据集上进行比较。$P$ 值由对数秩检验计算，即风险组之间的生存差异。结果如图 5 所示。

从图 5 中可以看出，在 TCGA-LUAD 数据集上，本文提出的方法表现出色（$P$ 值为 $2.41\times10^{-5}$），能够明显区分出高风险患者和低风险患者之间的差异性。其表现与 TEA-Graph 相当，能显著区分高生存率和低生存率风险（$P$ 值为 $2.06\times10^{-5}$）。其它方法，如 WSISA（$P$ 值为 0.004）、Attention MIL（$P$ 值为 0.002）、PatchGCN（$P$ 值为 $2\times10^{-4}$）、HGSurvNet（$P$ 值为 $3\times10^{-4}$）相较之下，对患者的高低风险分组差异较小，不能显著区分患者的肺癌风险情况。

## 2.4　消融实验

为验证本文提出的不同阶段的改进方法对整体生存预测框架模型的影响，本节在 TCGA-LUAD 数据集上进行了消融实验。本次实验从 3 个方面展开：不同特征提取器对预后预测的影响、FEM 模块及融合策略对预后预测的影响、图嵌入学习网络结构对预后预测的影响。

文献[30]的研究发现特征提取器对图基础结构具有影响。表 2 展示了 VGG[31]、ResNet50、EfficientNet-b4[32]和 Dif-ResNet 作为特征提取器对预测结果的影响。结果表明 Dif-ResNet 相比于 VGG、ResNet50 及 EfficientNet-b4 在 C-Index 指标上分别提高 0.056、0.009 和 0.014，取得了最优的生存预测结果(加粗表示)。

**表 2　不同特征提取器在 TCGA-LUAD 数据集上预测的结果**

**Table 2　Comparison of prediction results from different feature extractors on the TCGA-LUAD**

| 特征提取器 | C-Index | Standard deviation |
|---|---|---|
| VGG | 0.659 | 0.042 |
| ResNet50 | 0.706 | 0.032 |
| EfficientNet-b4 | 0.701 | 0.040 |
| **Dif-ResNet** | **0.715** | 0.039 |

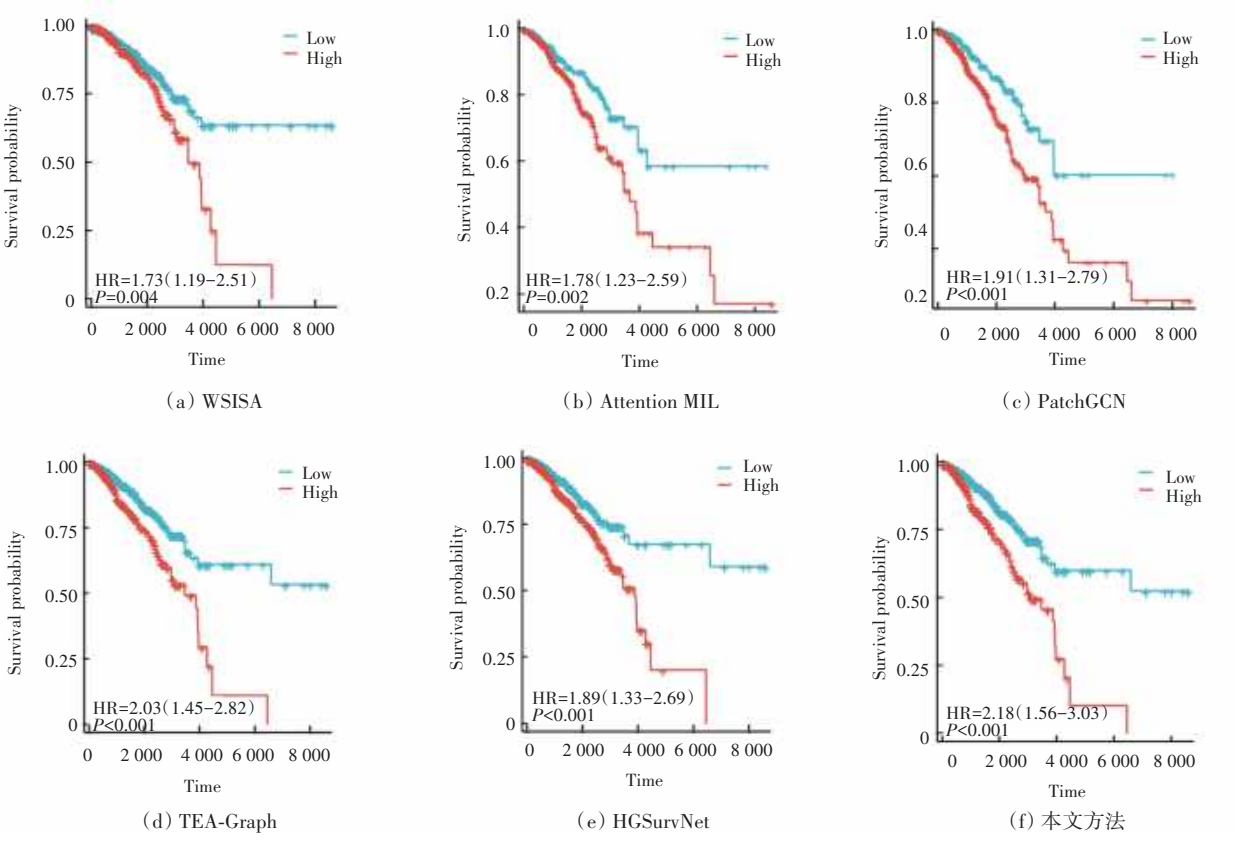



**图 5 不同方法预后风险分组曲线图**

**Fig. 5 Prognostic risk grouping curves of different methods**

表 3 展示了 FEM 模块及不同融合策略下对预后预测结果的影响。直接对提取的图像块特征进行图构建,得到的预测结果明显低于采用 FEM 模块的结果,这是因为在构建过程中,会引入噪声或丢失重要的结构信息,导致图连通性不足或质量较低。引入 FEM 模块并通过相加策略将原始图特征与生成图特征融合,预测结果提升了 0.059。在此基础上进一步引入融合模块 AF 后,预测结果提升了 0.062,表明基于注意力机制的融合策略关注了重要特征信息,抑制了无关特征干扰,有效提升图结构特征质量及最终预测性能。

**表 3 FEM 模块及不同融合策略预测结果**

**Table 3 Prediction results of the FEM module and different fusion strategies**

| FEM | Add | AF | C-Index | Standard deviation |
|---|---|---|---|---|
| | | | 0.653 | 0.049 |
| ✓ | ✓ | | 0.712 | 0.033 |
| ✓ | | ✓ | **0.715** | 0.039 |

表 4 展示了不同图神经网络模型对预测结果的影响。在 C-Index 指标下,相较于仅使用简单节点信息聚合的模型,引入残差连接的 GAT、GCN、并结合双分支结构,预测性能分别提升了 0.048 和 0.040,其中,拼接浅层特征的 GCN 结构可以捕捉多尺度信息,强化节点的局部特征表示;融入边、角度信息并引入跳跃式结构,可以提升模型对肿瘤环境异质性的学习能力,增强全局特征表示。最终采用双分支结构取得了最佳的预测结果,表明局部特征信息与全局特征信息的互补作用,提升了生存预测准确性。

**表 4 不同图神经网络结构预测结果**

**Table 4 Prediction results of different graph neural network structures**

| GCN | GAT | Residual | DF | C-Index | Standard deviation |
|---|---|---|---|---|---|
| √ | | | | 0.667 | 0.049 |
| | √ | | | 0.675 | 0.046 |
| | | √ | | 0.699 | 0.044 |
| √ | √ | √ | √ | **0.715** | 0.039 |

## 3 结束语

本文提出的生存预测框架 EG-GraphSurv,有效解决了 WSI 图像造成的低级特征冗余、图结构质量较差及难以学习关键图特征等问题。通过 DifResNet 对癌症分类数据集预训练并保留权重,以提

取 WSI 图像特征，解决初级特征质量偏低问题，为后续图结构构建和学习提供具有判别力的特征信息；FEM 模块利用条件分布为中心节点生成更多特征信息，并通过注意力融合得到增强图结构，优化了图初始节点表示；双分支图神经网络通过引入位置信息和残差连接的 GAT 分支和浅层 GCN 分支提升对图局部特征和全局特征的表征能力，进一步优化了生存预测效果。将所提出的方法与生存预测领域常见的方法模型进行对比，结果表明本章方法具备更优的性能，同时，通过消融实验也验证了生存预测任务不同阶段中设计的有效性。

## 参考文献


[1] Siegel R, Giaquinto A, Jemal A. Cancer statistics 2024[J]. CA: A Cancer Journal for Clinicians, 2024,74(1): 12-49.

[2] Pantanowitz L, Evans A, Pfeifer J, et al. Review of the current state of whole slide imaging in pathology[J]. Journal of Pathology Informatics, 2011,2(1):36. DOI:10.4103/2153-3539.83746.

[3] Richard C, Cheng C, Yi L, et al. Scaling vision transformers to gigapixel images via hierarchical self-supervised learning[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2022: 16123-16134.

[4] Shao W, Wang T, Huang Z, et al. Weakly supervised deep ordinal cox model for survival prediction from whole-slide pathological images[J]. IEEE Transactions on Medical Imaging, 2021, 40(12):3739-3747.

[5] Humphrey P A, Moch H, Cubilla A, et al. The 2016 who classification of tumours of the urinary system andmale genital organs-part A: renal, penile, and testicular tumours[J]. European Urology, 2016, 70(1): 93-105.

[6] Lu W, Graham S, Bilal M, et al. Capturing cellular topology in multi-gigapixel pathology images[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 1049-1058.

[7] Zhao Y, Yang F, Fang Y, et al. Predicting lymph node metastasis using histopathological images based on multiple instance learning with deep graph convolution[C]//Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 4837-4846.

[8] Ozen Y, Aksoy S, Ksemehmetolu K, et al. Self-supervised learning with graph neural networks for region of interest retrieval in histopathology[C]//Proceedings of the 25th International Conference on Pattern Recognition. Piscataway, NJ: IEEE, 2021: 6329-6334.

[9] Lee Y, Park J, Oh S, et al. Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning[J]. Nature Biomedical Engineering, 2022: 1-15.

[10] Zhu X, Yao J, Zhu F, et al. WSISA: Making survival prediction from whole slide histopathological images[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 6855-6863.

[11] Di D, Zou C, Feng Y, et al. Generating hypergraph-based high-order representationsof whole-slide histopathological images for survival prediction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(5): 5800-5815.

[12] Liu S, Dong H, Li L, et al. Local augmentation for graph neural networks[C]//Proceedings of the International Conference on Machine Learning. IMLS, 2021: 14054-14072.

[13] Li R, Yao J, Zhu X, et al. Graph CNN for survival analysis on whole slide pathological images[C]//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2018: 174-182.

[14] Chen R, Lu M, Shaban M, et al. Whole slide imagesare 2D point clouds: Context-aware survival predictionusing patch based graph convolutional networks[C]//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham:Springer, 2021: 339-349.

[15] Chen Z, He Z, Lu Z. DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention[J]. IEEE Transactions on Image Processing, 2023(33): 1002-1015.

[16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 770-778.

[17] 吴桂玲. 基于欧氏距离和余弦相似度特征选择的入侵检测模型[J]. 中小企业管理与科技, 2010(4): 231-232.

[18] Sohn K, Lee H, Yan X. Learning structured output representation using deep conditional generative models[C]//Proceedings of the International Conference on Neural Information Processing Systems. NeurIPS, 2015: 28.

[19] Thimm G, Fiesler E. High order and multilayer perceptron initialization[J]. IEEE Transactions on Neural Networks, 1997, 8(2): 349-359.

[20] Velickovic P, Cucurull G, Casanova A, et al. Graph attention network[J]. arXiv preprint arXiv,1710.10903, 2018.

[21] Jiang B, Zhang Z, Lin D, et al. Semi-supervised learning with graph learning-convolutional networks[C]//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019: 11305-11312.

[22] Li G, Muller M, Thabet A, et al. DeepGCNs: Can GCNs go as deep as CNNs? [C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2019: 9266-9275.

[23] Zhao L, Hou R, Zhao W Y, et al. Self-supervised learning guided transformer for survival prediction of lung cancer using pathological images[C]//2023 IEEE 20th International Symposium on Biomedical Imaging. Piscataway, NJ: IEEE, 2023: 1-5.

[24] Kandoth C, Mclellan M D, Vandin F, et al. Mutational landscape and significance across 12 major cancer types[J]. Nature, 2013, 502(7471): 333-339.

[25] Andrew A, Bui M, Thomas L, et al. Lung and colon cancer histopathological image dataset (LC25000)[J]. arXiv preprint arXiv,1912.12142, 2019.

[26] Otsu N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man and Cybernetic, 1979, 9(1): 62-66.

[27] Hai J, Xuan Z, Yang R, et al. R2rnet: Low-light image enhancement via real-low to real-normal network[J]. Visual Communication and Image Representation, 2023, 90:103712.

[28] Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning[C]//Proceedings of the International Conference on Machine Learning. IMLS, 2018: 2127-2136.

[29] Huang Z, Chai H, Wang R, et al. Integration of patch features through self-supervised learning and transformer for survival analysis on whole slide images[C]//Medical Image Computing and Computer Assisted Intervention, 2021: 561-570.

[30] 李杜蕾,杨波,杨博雄,等. 图卷积神经网络基础结构影响研究[J]. 小型微型计算机系统, 2021, 42(4): 891-896

[31] Ding X, Zhang X, Ma N, et al. Repvgg: Making vgg-style convnets great again [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2021: 13728-13737.

[32] Tan M, Le Q. EfficientNet: Rethinking models scaling for convolutional neural networks [C]//Proceedings of the 36th International Conference on Machine Learning. IMLS, 2019: 6105-6114.