武汉大学学报(理学版)
J. Wuhan Univ. (Nat. Sci. Ed.) 1~10

DOI: 10.14188/j.1671-8836.2026.0033

基于多任务学习模型的鱼群精准投喂方法

张 莹[1],杨绎好[2],陈 晶[1],蔡润基[1],周四维[1†]

1. 广东海洋大学 数学与计算机学院,广东 湛江 524088;

2. 湖北大学 数学与统计学学院,湖北 武汉 430000

收稿日期:2026-03-17 † 通信联系人 E-mail:Siweizhou@gdou.edu.cn

基金项目:广东省自然科学基金资助项目(2026A1515011448);广东省普通高校科研重点项目(2025ZDZX1007);广东省普通高校特色创新类项目(2025KTSCX047);广东省教育厅海洋牧场装备信息化与智能化创新团队项目(2023KCXTD016)

第一作者:张 莹,女,博士,副教授,现从事海洋大数据处理分析及智能模型算法研究。E-mail:zhangying@gdou.edu.cn

摘 要:为提高鱼群智能投喂精度并降低欠投喂、过量投喂风险,本文构建了一种融合摄食状态与残饵计数的多任务学习模型。该模型以共享特征网络提取鱼群摄食图像特征,利用分类分支识别摄食状态,利用回归分支预测残饵密度图,并通过对密度图求和估计残饵数量;采用基于任务不确定性的损失加权方法自适应平衡两项任务。在自建的750张鱼群摄食行为图像数据集上进行实验,结果表明,所提模型在测试集上的摄食状态准确率、精确率和召回率分别为94.67%、97.96%和94.12%,残饵计数平均绝对误差和均方根误差分别为14.85和22.19,整体优于单任务模型及常用对比模型。基于模型输出构建投喂策略映射规则,可为循环水养殖场景下的鱼群精准投喂决策提供方法参考。

关 键 词:鱼群精准投喂;多任务学习模型;摄食状态;残饵计数

中图分类号:S.951.2 **文献标志码**:A **文章编号**:1671-8836(XXXX)XX-0001-10

A Fish Precision Feeding Method Based on Multi-Task Learning Model

ZHANG Ying[1], YANG Yihao[2], CHEN Jing[1], CAI Runji[1], ZHOU Siwei[1†]

1. Faculty of Mathematics and Computer Science, Guangdong Ocean University, Zhanjiang 524088, Guangdong, China;

2. School of Mathematics and Statistics, Hubei University, Wuhan 430000, Hubei, China

Abstract: To improve the accuracy of intelligent fish feeding and reduce the risks of underfeeding and overfeeding, this study proposes a multi-task learning model that integrates feeding activity recognition and residual feed counting. The model employs a shared feature extraction network to learn visual representations from fish feeding images. A classification branch is used to identify feeding activity states, while a regression branch predicts a residual-feed density map, from which the number of uneaten feed pellets is estimated by summation. A task-uncertainty-based loss weighting strategy is further introduced to adaptively balance the two tasks. Experiments were conducted on a self-constructed dataset containing 750 images of fish-feeding behavior. The proposed model achieved an accuracy, precision, and recall of 94.67%, 97.96%, and 94.12%, respectively, for feeding activity recognition, while obtaining an MAE of 14.85 and an RMSE of 22.19 for residual feed counting. Overall, the proposed method outperformed both single-task models and commonly used baseline models. Furthermore, a feeding strategy mapping rule was developed based on the model outputs, providing a methodological reference for precise feeding decision-making in recirculating aquaculture systems.

Key words: precise fish feeding; multi-task learning model; feeding activity status; residual feed counting

引用格式:张莹,杨绎好,陈晶,等. 基于多任务学习模型的鱼群精准投喂方法[J]. 武汉大学学报(理学版),XXXX,XX(XX):1-10. DOI:10.14188/j.1671-8836.2026.0033.

ZHANG Ying, YANG Yihao, CHEN Jing, *et al*. A Fish Precision Feeding Method Based on Multi-Task Learning Model [J]. *J Wuhan Univ (Nat Sci Ed)*, XXXX, XX(XX):1-10. DOI:10.14188/j.1671-8836.2026.0033(Ch).

0 引 言

在水产养殖领域,饵料投喂的精准性直接影响养殖成本、鱼群生长及养殖水环境的稳定,探索鱼群的精准投喂方法一直是学术界与产业界关注的焦点[1]。传统的机械投喂方法是根据投喂人员的经验,定时定量投喂,容易造成欠投喂和过投喂等问题[2]。近年来,随着人工智能和计算机视觉技术的发展,基于图像的摄食行为识别与残饵计数为按需投喂提供了技术途径。

已有研究采用行为分类和目标分割等方法分析鱼群摄食行为。行为分类方法从图像中提取鱼群聚集、水面扰动等特征,并输出摄食状态或强度等级。Zhou等[3]首次应用深度学习模型LeNet-5将鱼群摄食状态划分为无、弱、中、强4个等级,分类准确率为90%;Zhang等[4]在MobileNetV3中引入多尺度信息融合,摄食行为分类准确率达到96.4%;Zhang等[5]提出MobileNetV2-SENet网络,用于识别鱼群摄食行为。目标分割方法则通过识别摄食过程中的鱼体目标描述群体行为。Yang等[6]针对鱼群遮挡、重叠和聚集问题,提出BlendMask-VoVNetV2模型,以改善工业化养殖场景下鱼群摄食行为的量化分析。这些研究从视觉行为特征评估鱼群摄食状态,为摄食状态识别提供了参考。

残饵计数可为判断投喂量与摄食需求是否匹配提供另一类信息。基于密度图回归的方法通过点标注生成密度图,并对预测密度图求和得到目标数量,已用于鱼饵、虾苗、虾卵和鱼苗等水产养殖对象的计数。Hou等[7]改进MCNN模型,并引入结构相似性损失,用于鱼饵颗粒计数。Peng等[8]提出多尺度注意力融合网络SFCNet,以降低虾苗计数误差。Zhang等[9]利用生成对抗网络扩充训练数据,并采用SECNet对红螯螯虾卵进行密度图回归计数。Li等[10]提出轻量化鱼苗计数网络MSENet,以支持便携式计数设备。上述研究为残饵颗粒的密度图回归计数提供了方法参考。

在实际投喂场景中,鱼群摄食状态与残饵数量共同反映鱼群摄食需求和饵料利用情况。仅依赖摄食状态,难以判断投喂是否过量;仅依赖残饵数量,也难以反映鱼群是否仍具有摄食意愿。因此,将摄食状态识别与残饵计数联合建模,有助于为投喂决策提供更全面的依据。同时,两项任务在视觉特征上存在相关性,均会受到鱼群聚集、水面波纹、颗粒遮挡和光照变化等因素影响,可据此探索共享特征的联合学习方法。

多任务学习通过共享表示同时优化相关任务,但任务间的损失尺度、训练进度和梯度方向可能存在差异。文献[11]利用任务不确定性自适应调整损失权重;文献[12]提出GradNorm,通过梯度归一化平衡任务的训练进度;文献[13]提出CAGrad,以缓解不同任务的梯度冲突;文献[14]提出Nash-MTL,将各任务梯度的组合建模为协商博弈,并利用纳什协商解确定共同的参数更新方向。这些方法分别从损失加权和梯度组合角度改善多任务优化,为本文的联合训练提供了参考。

摄食行为识别和残饵计数分别描述鱼群状态与饵料剩余情况,二者的联合利用有助于完善投喂决策依据。为此,本文面向循环水养殖场景,提出融合摄食状态识别与残饵计数的鱼群精准投喂方法,并基于模型输出构建投喂策略映射规则。本文的主要贡献如下:

1) 构建了同时包含摄食状态标签和残饵位置标注的鱼群多任务摄食行为数据集,为摄食状态识别与残饵计数联合训练提供数据基础。

2) 提出了融合摄食状态识别和残饵计数的多任务学习模型,将鱼群摄食行为状态和残饵数量信息纳入统一建模框架。

3) 引入基于方差不确定性的损失加权机制,自适应平衡分类损失和密度图回归损失,提升模型联合学习效果。基于摄食状态和残饵数量构建投喂策略映射规则,为循环水养殖场景下的鱼群精准投喂提供方法参考。

1 本文方法

1.1 数据集构建

现有相关研究分别关注鱼群摄食欲评估[15]、鱼群声信号分析[16]和密集鱼群检测计数[17]。为满足摄食状态识别与残饵计数的联合训练需求,本文构建鱼群多任务摄食行为数据集。数据集制作流程包含以下5个步骤:第1步,拍摄鱼群摄食行为视频;第2步,对视频进行采样和筛选,得到候选鱼群摄食图像;第3步,对图像进行人工标注;第4步,对标注后的图像进行预处理;第5步,划分训练集、验证集和测试集。

1.1.1 视频拍摄与采样

以草金鱼作为视频行为识别实验对象,因为其体色较为明显、群体上浮摄食反应较清晰,便于观察鱼群摄食行为、水面扰动及残饵分布。选用直径约2~3 mm的上浮型颗粒饵料,该粒径适合草金鱼

摄食,且颗粒在水面可见性较好,便于观察上浮摄食过程及残饵分布。实验在循环水养殖系统中进行,以维持相对稳定的水质条件;通过固定拍摄位置、光照条件和背景环境,减少环境波动对摄食行为观察及图像质量的影响。投喂频率为每日两次,固定于上午和下午相对稳定的时段,以维持相对稳定的摄食节律,降低采集时段差异带来的干扰。

视频拍摄持续两周,以获得多次重复投喂过程下的摄食行为样本,并尽量避免长周期养殖造成的鱼体生长和环境累积变化。

视频采集从投喂前20 s开始,并持续至鱼群明显摄食行为基本停止,以覆盖摄食前基线状态、摄食响应过程以及摄食后期残饵状态。由于连续视频中相邻帧存在较强冗余,本文每5 s抽取一帧作为候选图像,在减少重复样本和标注工作量的同时保留不同摄食阶段的代表性信息。随后,由水产养殖专家根据图像清晰度、鱼群摄食状态可辨性以及残饵分布可见性筛选图像,以提高后续标注数据的可辨性和可靠性。

1.1.2　标签标注、密度图生成和样本划分

数据集图像标签内容包含两部分:摄食状态标签和残饵位置标签。已有研究表明[15,17],鱼类摄食状态通常会反映在鱼群聚集程度、游动活跃性、摄食反应以及水面扰动等视觉特征上。为降低主观判断对摄食状态标注的影响,本文将摄食状态标签的判定过程量化为3个可观察指标:鱼群是否主动游向水面索食、鱼群游动是否积极以及水面是否出现明显摄食波纹。每个指标满足条件记为1分,否则记为0分,得到摄食状态评分S_{feed}。当$S_{\text{feed}} \geqslant 2$时,将该图像标注为活跃(Active);当$S_{\text{feed}} < 2$时,将该图像标注为不活跃(Not Active)。对视觉特征不明显、难以判定或标注人员意见不一致的样本,由两名标注人员独立复核;若复核结果仍存在分歧,则由水产养殖专家进行最终确认。残饵位置标签通过对图像中每个可见残余饵料颗粒进行中心点标注获得,并通过高斯密度函数将离散残饵位置点转化为连续密度图。图1展示了不同标签组合下的数据集可视化样本。表1列出了摄食状态标签的评分规则。

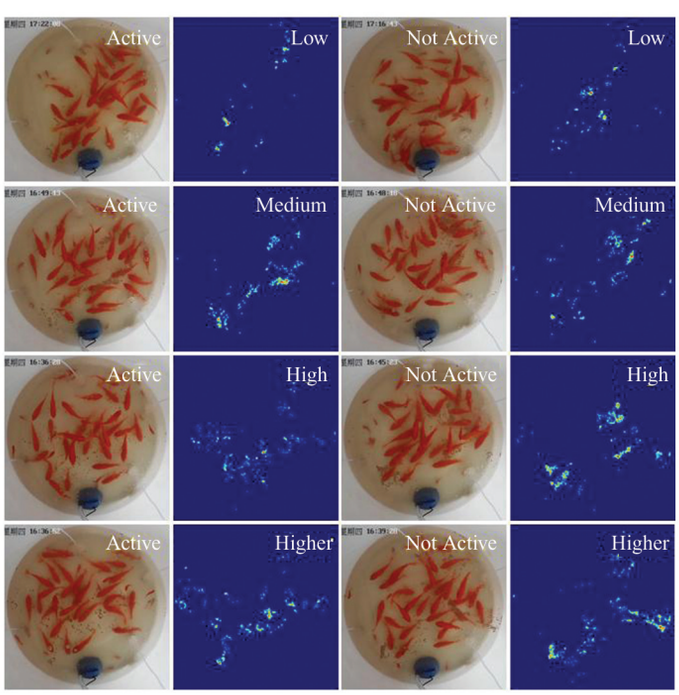


图1　不同标签组合的鱼群摄食图像及对应残饵密度图

Fig. 1　Examples of fish feeding images and density maps with different labels

表1　摄食状态标签的量化判定规则

Table 1　Quantitative criteria for feeding-intensity labels

指标	记1分条件	记0分条件
主动上浮索食	多数鱼体向水面或投喂区域聚集并主动接近饵料	鱼群无明显上浮或接近饵料行为
游动积极性	鱼群游动速度较快,方向变化频繁,摄食区域活动明显	鱼群游动缓慢或分散停留
水面波纹	水面出现由摄食或争食引起的连续明显波纹	水面波纹不明显或仅有偶发扰动

在完成图像的摄食活动标签和残饵中心点标注后,对原始图像进行预处理。原始图像分辨率为1 920×1 080,为减少背景区域、水体边缘及无关环境信息对模型训练的干扰,本文对图像中与鱼群摄食行为和残饵分布相关的有效区域进行裁剪,并统一缩放至512×512分辨率,以保证输入图像尺寸一致,提高后续模型训练的稳定性和计算效率。

最终,本文构建了包含750张图像的数据集。每张图像均同时标注摄食状态标签和残饵中心点位置,残饵密度等级依据残饵数量区间划分,摄食状态分为Active和Not Active两类。按照7∶1∶2的比例将数据集划分为训练集、验证集和测试集,如表2所示。

表2 数据集划分

Table 2 Dataset composition and partitioning

残饵密度等级	残饵数量区间	摄食状态	图像数量			残饵总数
			训练集	验证集	测试集	
低	$0 \leqslant N \leqslant 99$	Active	55	8	16	4 130
		Not Active	54	8	15	4 291
中	$100 \leqslant N \leqslant 199$	Active	84	12	24	19 163
		Not Active	55	8	15	12 248
高	$200 \leqslant N \leqslant 299$	Active	122	17	35	42 520
		Not Active	86	12	25	30 853
极高	$N \geqslant 300$	Active	40	6	12	20 666
		Not Active	29	4	8	14 090

1.2 多任务学习模型设计

针对鱼群投喂决策同时依赖摄食状态和残饵数量信息的问题,本文提出一种面向鱼群精准投喂的多任务学习模型。将摄食状态和残饵计数两个直接影响投喂决策的任务进行联合建模,并采用基于任务不确定性的损失加权方法平衡分类损失和密度图回归损失。

模型采用硬参数共享方式,由共享特征提取网络和任务分支组成。共享网络提取鱼群摄食图像的特征;任务分支中的分类分支识别摄食状态,回归分支预测残饵密度图。模型总体结构如图2所示。

设计前端网络的特征共享网络时,本研究选取ResNet34[18]作为待改进的网络基础版本。将ResNet34末端的全连接层和最大池化层移除,网络主体由一系列残差模块构成(BasicBlock_0和BasicBlock_1)。如图3所示,两种基础块都是由两个3×3卷积、两个批量归一化层(BatchNorm)和一个ReLU激活函数组成,区别在于BasicBlock_1中第二个卷积层的步幅为2,以达到对特征图的下采样。特征共享网络能够从输入图像中有效提取多层次的

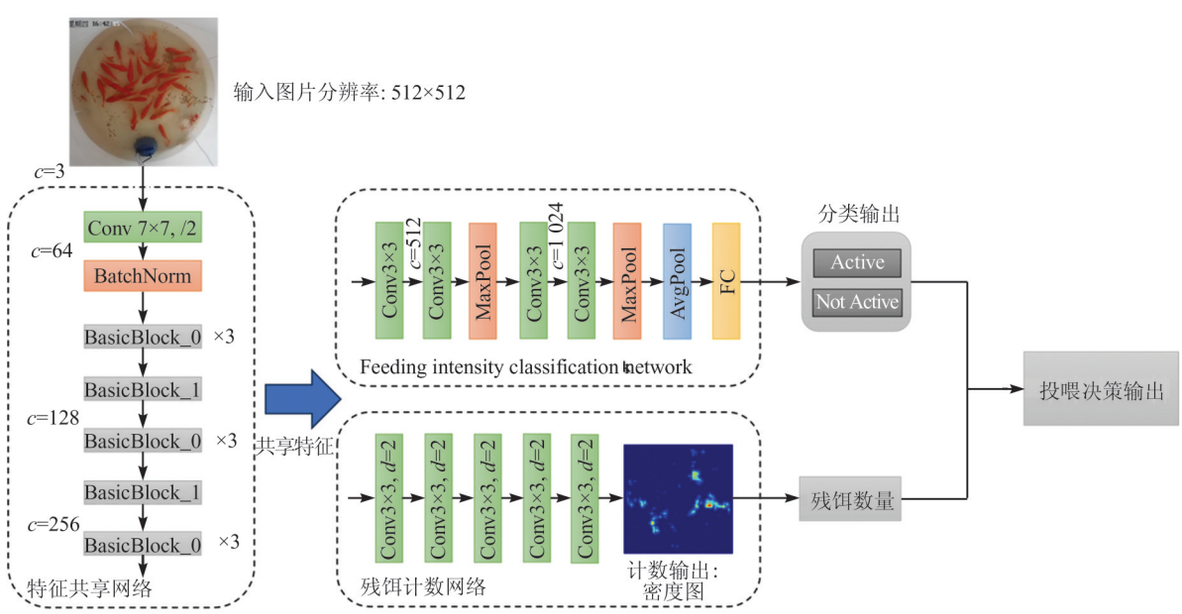


图2 多任务学习模型的总体结构

Fig. 2 The overall structure of the multi-task learning model

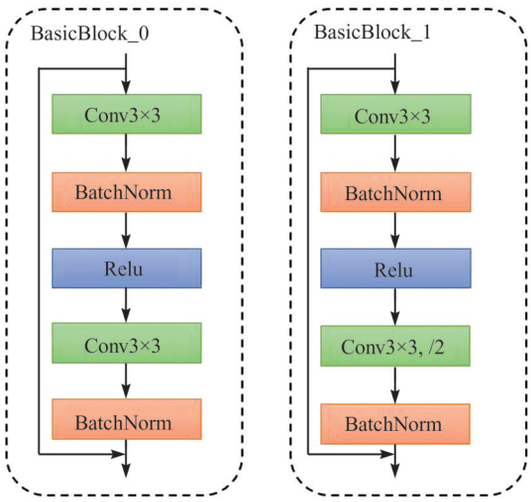


图 3　BasicBlock_0和BasicBlock_1的结构图

Fig. 3　The specific structure diagrams of BasicBlock_0 and BasicBlock_1

共享特征表示,并最终输出大小为原始输入图像1/8的特征图,供后续的任务特定网络使用。特征共享网络可在保留残饵颗粒空间位置信息的同时降低计算量,适合残饵小目标计数任务。

摄食状态网络主要用于处理摄食状态任务,如图2所示,该网络首先接收来自特征共享网络的特征图,并通过两次最大池化操作进一步进行特征浓缩和降维。然后,通过全局平均池化层将特征图整合为一维特征向量。最后经由一个全连接层输出每张图对应的预定义摄食状态类别("Active"或"Not Active")的概率。

在残饵计数网络中,连续引入5个3×3膨胀卷积层,并将膨胀率设置为2,如图2所示,连续3×3膨胀卷积在不增加参数量的前提下扩大感受野,有助于利用局部上下文区分残饵颗粒、鱼体边缘和水面反光噪声。对于单个3×3卷积核,当膨胀率d=2时,其有效感受野可扩展至5×5,但参数量仍与普通3×3卷积相同。因此,该结构能够在不显著增加参数量和计算复杂度的前提下扩大局部上下文感知范围。残饵颗粒在图像中常呈现密集分布、局部粘连或被鱼体部分遮挡等情况,仅依赖小范围纹理容易造成漏检或重复计数;膨胀卷积能够引入更大邻域内的空间上下文信息,使网络在生成残饵密度图时更好地区分残饵颗粒、鱼体边缘和水面噪声。计数分支最终输出残饵密度图D,通过对密度图像素值积分得到残饵数量估计值N_{hat},即$N_{hat}=\mathrm{sum}(D)$。

1.3　损失函数

在本文模型中,摄食状态属于二分类任务,残饵计数属于密度图回归任务。两类任务的损失函数形式和数值量级不同,若直接采用固定权重相加,可能导致某一任务在联合训练中占据主导,从而影响另一任务的收敛。因此,本文采用方差不确定性损失对分类损失和回归损失进行自适应加权。

摄食状态识别采用二元交叉熵损失,其表达式如下:

$$L_{cls}=-\frac{1}{N_T}\sum\left[\hat{y}^{cls}\log p(\hat{y}^{cls})+(1-\hat{y}^{cls})\log(1-p(\hat{y}^{cls}))\right] \tag{1}$$

其中,N为训练样本的数量,$\hat{y}^{cls}$为预测的分类类别,$p(\hat{y}^{cls})$为属于该类别的概率。而对于残饵计数,最终的回归损失可以表示为:

$$L_{cnt}=L_{mse}+\alpha L_{SSIM} \tag{2}$$

式中,L_{cnt}为残饵计数回归损失,L_{mse}为预测密度图与真实密度图之间的均方误差损失,L_{SSIM}[19]为结构相似性损失,α为结构相似性损失的权重系数,本文设置为0.01,均方误差损失L_{mse}和结构相似性损失L_{SSIM}的表达式分别为:

$$L_{mse}=\frac{1}{N_T}\sum\left(y^{cnt}-\hat{y}^{cnt}\right)^2 \tag{3}$$

$$L_{SSIM}=1-\frac{1}{N_T}\sum_x\frac{(2\mu_p\mu_g+c_1)(\sigma_{pg}+c_2)}{(\mu_p^2+\mu_g^2+c_1)(\sigma_p^2+\sigma_g^2+c_2)} \tag{4}$$

其中,y^{cnt}为密度图标签,$\hat{y}^{cnt}$为预测输出的密度图。μ_p和σ_p^2分别表示预测密度图的均值与方差,μ_g和σ_g^2分别表示密度图标签的均值和方差,σ_{pg}表示预测密度图和密度图标签的协方差,c_1和c_2为常数。

综合分类损失与残饵计数回归损失,多任务模型的联合损失L_{total}如下:

$$L_{total}=\frac{1}{2\sigma_1^2}L_{cls}+\frac{1}{2\sigma_2^2}L_{cnt}+\log(1+\sigma_1)+\log(1+\sigma_2) \tag{5}$$

其中,σ_1^2和σ_2^2分别表示分类任务和计数任务对应的任务不确定性,二者作为可学习参数与网络权重共同优化。当某一任务损失的不确定性较高时,其权重会相应降低;当任务不确定性较低时,其权重会相应提高,从而实现多任务训练过程中的自适应平衡。

1.4　评价指标

模型评价包括摄食状态分类和残饵计数两部分。分类任务采用准确率(A)、精确率(P)、召回率(R)和$F1$值($F1$)作为评价指标,计算公式如下:

$$\mathrm{Accuracy}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}}\times 100\% \tag{6}$$

$$P=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}}\times 100\% \tag{7}$$

$$R=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}\times 100\% \tag{8}$$

$$F1=\frac{2\times P\times R}{P+R} \tag{9}$$

其中,TP为真阳性样本数,TN为真阴性样本数,FP为假阳性样本数,FN为假阴性样本数。

残饵计数任务采用平均绝对误差(MAE)和均方根误差(RMSE)评价预测数量与实际数量之间的偏差,计算公式如下:

$$\mathrm{MAE}=\frac{1}{N_{\mathrm{E}}}\sum\left|y-\hat{y}\right| \tag{10}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{N_{\mathrm{E}}}\sum(y-\hat{y})^2} \tag{11}$$

其中,y表示实际的残饵数量,$\hat{y}$为预测的残饵数量,N_{E}为评估的图像数量。

2 实验结果与分析

2.1 实验设置

本文对多任务模型、对比模型及消融实验模型采用相同的训练配置,以提高实验结果的可比性。表3列出了多任务学习模型的训练参数,训练过程中每50轮保存一次模型权重,并保存最优结果对应的模型权重。

表3 模型训练参数配置

Table 3 Training configuration

参数	值
学习率	1.0×10^{-6}
批量大小	8
训练轮次	300
优化器	随机梯度下降(SGD)
随机数种子	2 468
权重衰减率	5.0×10^{-4}
动量	0.95

2.2 多任务模型性能评估

2.2.1 模型训练过程

模型训练的主要步骤是,1)将待训练图像输入特征共享网络,提取图像数据中的底层语义特征;2)将特征共享网络输出的特征图作为输入分别输入摄食状态网络和残饵计数网络中;3)摄食状态网络输出属于每个类别的概率,残饵计数网络输出每张图像关于残饵分布的密度图;4)对密度图中所有像素值求和即可获得每张图的残饵数量。

多任务网络模型的训练损失曲线如图4所示。图4(a)(b)和(c)分别对应摄食状态分类损失L_{cls}、残饵计数回归损失L_{cnt}和(5)式定义的联合损失L_{total}。随着训练推进,三类损失均呈下降趋势;约100轮后,回归损失和联合损失的下降速度减缓,分类损失仍伴随一定波动继续下降。该结果反映了模型在当前训练设置下的损失变化过程。

2.2.2 摄食状态性能比较

为评估本研究所提的多任务模型在摄食状态任务上的性能,将其与AlexNet[20]、ShuffleNetV2[21]、ResNet34[18]以及MobileNetV3[22]等主流分类模型进行对比。最终训练的分类模型在测试集上进行性能比较,比较结果如表4所示,从表4中可以看出,本文模型在准确率、精确率、召回率和$F1$值等指标上均取得较优结果;FPS(Frames Per Second)为39.06,折算单帧推理时间约为25.6 ms,能够满足视频帧级摄食状态监测和投喂策略更新的基本需求。

表4 本文模型与其他模型在测试集上的分类性能比较

Table 4 Comparison of classification performance on the test set between multi-task models and other models

模型	A/%	P/%	R/%	$F1$	FPS
AlexNet[20]	81.33	68.42	92.85	0.788	258.07
ShuffleNetV2[21]	72.00	73.08	84.44	0.784	108.41
ResNet34[18]	86.67	76.47	92.86	0.839	107.41
MobileNetV3[22]	90.67	83.87	92.86	0.881	148.38
Multi-Task(ours)	94.67	97.96	94.12	0.960	39.06

图5展示了不同分类模型在训练过程中于验证集上的准确率变化。本文多任务模型的准确率在训练初期快速提高,随后在大部分训练阶段保持较高水平,且波动相对较小。相比之下,AlexNet、ResNet34和MobileNetV3的验证集准确率整体较低,其中MobileNetV3在部分训练阶段出现较明显的波动;ShuffleNetV2的准确率则长期维持在较低水平。上述结果表明,在本文数据集和统一训练配置下,多任务模型能够获得较高且相对稳定的验证集分类表现。

2.2.3 残饵计数性能比较

为评估残饵计数性能,本文将多任务模型与MCNN[23]、CSRNet[24]及基于VGG16[25]的计数模型进行比较。图6展示了4个残饵密度等级下的原图、密度图标签及各模型的预测结果,用于观察预测残饵数量与空间分布的差异。本文多任务模型在测试集上的MAE为14.85,RMSE为22.19。共享特征对计数任务的作用通过第2.3节的单任务与多任务对比实验进行分析。

图7比较了不同模型在验证集和测试集上的逐幅图像残饵计数误差,误差越接近零,表示预测数量与

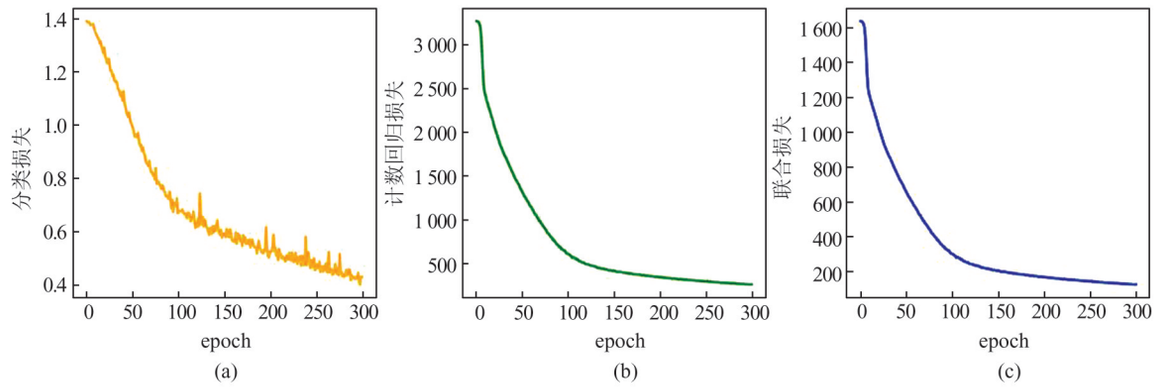


图 4　分类损失、计数回归损失及联合损失的训练曲线

Fig. 4　Training curves for the classification, counting, and total losses

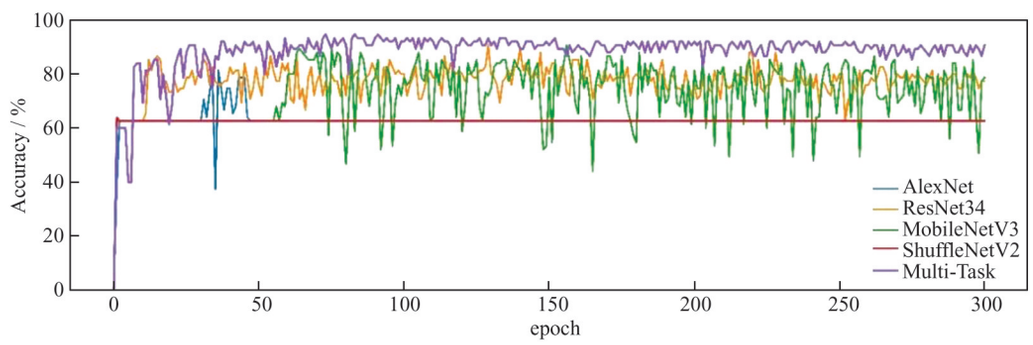


图 5　不同分类模型训练过程中在验证集上准确率的收敛情况

Fig. 5　The convergence of accuracy rates on the validation set during the training process of different classification models

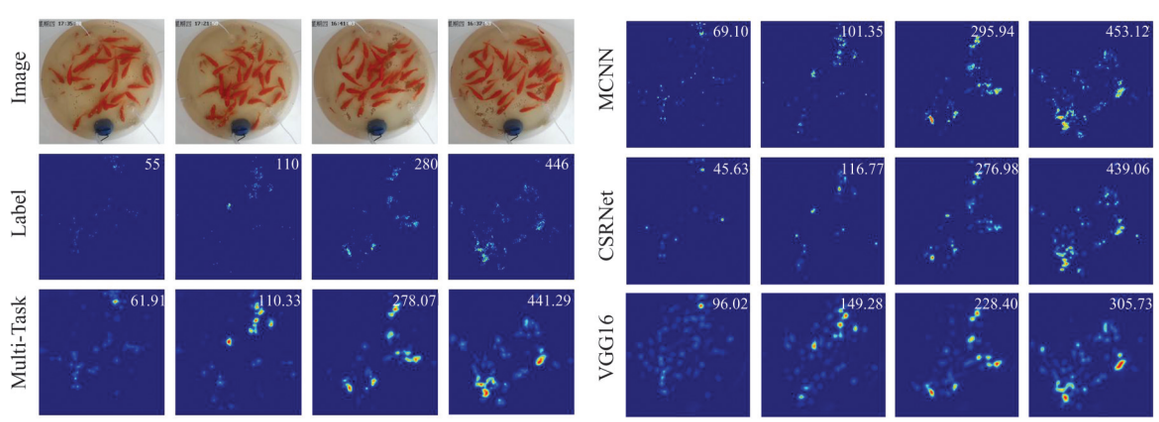


图 6　4个残饵密度等级的原图、密度图标签及不同模型的预测密度图

Fig. 6　Original images, ground-truth density maps, and predictions of different models at four residual feed density levels

真实数量越接近。由图可见,VGG16在多个样本上出现较大的正向或负向偏差,误差波动较为明显。相比之下,本文多任务模型的误差总体更集中于零误差线附近,部分大幅偏差得到减小。与MCNN和CSRNet相比,本文模型在部分样本上具有较小的计数偏差,但并非在所有图像上均取得最小误差。结合测试集上MAE为14.85、RMSE为22.19的结果,可以看出本文模型在当前数据集上具有较好的残饵计数表现。

2.3　消融实验分析

为验证所采用的基于方差不确定性的损失加权策略的有效性,以及本文多任务网络相较于多任务模型中单分支网络上的性能,本研究进行了消融实验。表5展示了使用平均权重、通过手动调参得到的最优估计权重以及使用方差不确定性方法损失函数的测试结果。平均权重方法实现简单,但无法反映分类任务和回归任务之间的损失量级差异;手动设定权重能

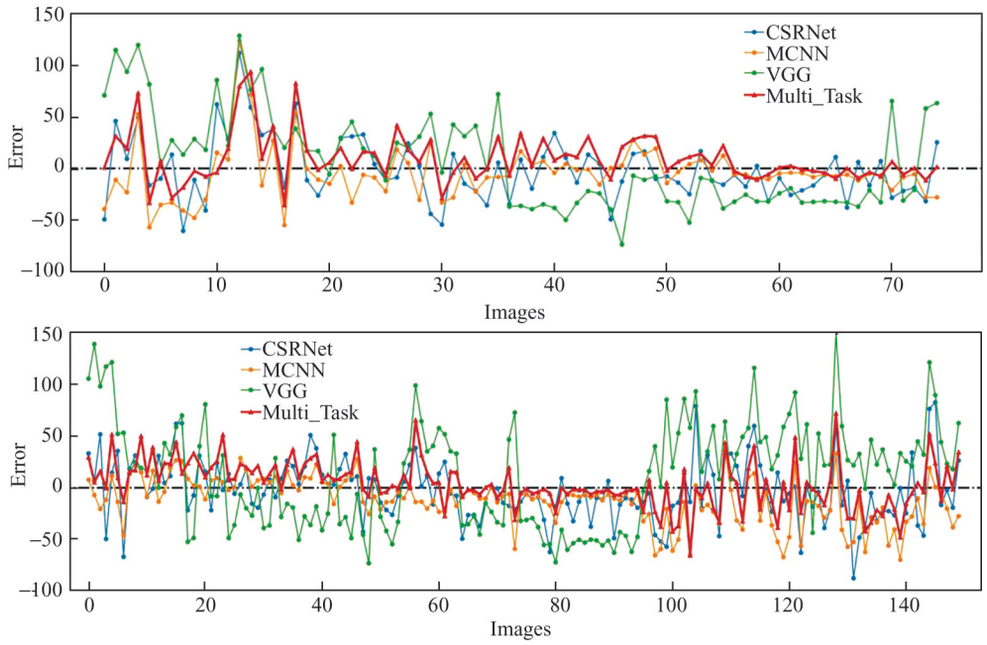


图 7 不同模型在验证集和测试集上的残饵计数误差比较

Fig. 7 Residual feed counting errors of different models on the validation and test sets

够在一定程度上改善联合训练效果,但需要依赖人工调参,且所得权重未必适用于新的数据分布。相比之下,方差不确定性损失将任务权重作为可学习参数进行优化,能够根据训练过程中不同任务的不确定性自适应调整损失权重。由表5可见,在测试集上,该策略在分类性能和残饵计数误差之间取得了较好的平衡。

表 5 使用不同损失策略的多任务模型评价指标性能比较

Table 5 Performance comparison of evaluation metrics for multi-task models using different loss strategies

损失策略	任务权重		A/%	P/%	R%	$F1$	MAE	RMSE
	分类	回归						
平均权重	0.5	0.5	90.67	96.81	89.22	0.929	20.66	25.04
最优估计权重	0.99	0.01	93.33	93.40	97.06	0.952	19.94	26.31
方差不确定性	—	—	94.67	97.96	94.12	0.960	14.85	22.19

注:“—”表示不存在。

为进一步证明多任务联合训练对于提升各独立任务性能的积极作用,将本文多任务模型与其两个任务分支(摄食状态和残饵计数)分别进行单任务训练时的性能进行对比。如表6所示,引入多任务学习后,摄食状态的准确率从单任务的86.00%提升至94.67%(提升了8.67%),残饵计数的MAE也从单任务的18.03降低至14.85(降低了3.18),这说明在本文数据集上,摄食状态与残饵计数之间存在一定互补关系,多任务联合训练能够利用共享特征提升两个子任务的综合表现。

表 6 多任务与不同分支单任务训练模型在测试集上的评价指标性能比较

Table 6 Comparison of evaluation metric performance of multi-task models and single-task models with different branches on the test set

方法	任务类型	A/%	P%	R%	$F1$	MAE	RMSE
单任务	摄食状态	86.00	87.85	92.16	0.899	—	—
	残饵计数	—	—	—	—	18.03	22.84
多任务	—	94.67	97.96	94.12	0.960	14.85	22.19

注:“—”表示不存在。

2.4 鱼群智能投喂策略

基于多任务学习模型的输出结果,本文进一步建立鱼群投喂策略映射规则。模型分类分支输出摄食状态概率 $p_{\mathrm{cls}}=[P(\mathrm{Active}),\ P(\mathrm{Not\ Active})]$,当 $P(\mathrm{Active})\geqslant 0.5$ 时判定为摄食 Active,否则判定为 Not Active。该阈值是二分类任务中常用的默认概

率判定阈值,用于将模型输出概率转换为不同的摄食状态状态。本文根据数据集中残饵数量标注的分布范围、不同残饵积累状态下的图像可辨识差异以及投喂过程中残饵积累对水质和继续投喂风险的影响,将N_{hat}划分为低($0\leqslant N_{hat}\leqslant 99$)、中($100\leqslant N_{hat}\leqslant 199$)、高($200\leqslant N_{hat}\leqslant 299$)和极高($N_{hat}\geqslant 300$)4个残饵密度等级。最后,根据表7确定投喂策略。本研究将投喂策略细化为4种主要类型:

1) 持续投喂:当模型判定摄食状态为Active,且残饵密度等级为低时,说明鱼群摄食需求仍较强且水体中可利用残饵较少,可继续小剂量、短间隔投喂。

2) 投喂停滞:当模型判定为Active且残饵密度为中或高时,鱼群仍可能存在摄食需求,但水体中已有一定残饵;当模型判定为Not Active且残饵密度为低时,残饵较少,但鱼群摄食响应不足。按照表7的映射规则,上述两类情况均短时暂停投喂并继续观察。

3) 投喂中止:当模型判定为Active但残饵密度达到极高,或判定为Not Active且残饵密度为中时,说明继续投喂可能增加残饵积累,应中止当前投喂过程并观察鱼群后续摄食状态。

4) 投喂结束:当模型判定为Not Active,且残饵密度为高或极高时,说明鱼群摄食响应较弱且水体中残饵较多,本轮投喂可判定为结束,并应停止投喂、检查残饵积累和水质状态。

表 7　鱼群在不同摄食状态以及不同残饵密度等级下的投喂策略

Table 7　The feeding strategies of fish schools under different feeding active and different levels of residual food density

P(Active)	摄食状态	N_{hat}数量区间	残饵密度	投喂策略
P(Active)≥0.5	Active	$0\leqslant N_{hat}\leqslant 99$	低	持续投喂
		$100\leqslant N_{hat}\leqslant 199$	中	投喂停滞
		$200\leqslant N_{hat}\leqslant 299$	高	投喂停滞
		$N_{hat}\geqslant 300$	极高	投喂中止
P(Active)<0.5	Not Active	$0\leqslant N_{hat}\leqslant 99$	低	投喂停滞
		$100\leqslant N_{hat}\leqslant 199$	中	投喂中止
		$200\leqslant N_{hat}\leqslant 299$	高	投喂结束
		$N_{hat}\geqslant 300$	极高	投喂结束

3　结　语

本文提出了一种基于多任务学习模型的鱼群精准投喂方法,将共享特征提取网络、摄食状态分支和残饵计数分支整合到统一框架中,并采用基于任务不确定性的损失加权方法平衡分类任务与密度图回归任务。实验结果表明,在测试集上,本文多任务模型的摄食状态准确率达到94.67%,残饵计数MAE为14.85,整体表现优于所选单任务模型和计数对比模型。消融实验结果进一步表明,在本文数据集上,摄食状态与残饵计数之间具有一定互补关系,联合训练有助于提升模型综合表现。基于模型输出,本文进一步建立了摄食状态、残饵密度等级与投喂策略之间的映射规则,为鱼群精准投喂提供了决策参考。

需要指出的是,本文数据集规模仍然有限,且样本主要来自特定循环水养殖系统和草金鱼投喂场景,模型在不同鱼种、不同养殖容器、不同光照和水体扰动条件下的泛化能力仍需进一步验证。此外,本文尚未系统比较GradNorm、CAGrad、Nash-MTL等多任务优化方法,也未开展跨场景交叉验证和统计显著性检验。后续研究将进一步扩充多场景数据集,结合水质、声学或时序行为信息,验证模型在复杂养殖环境下的稳定性和实际投喂控制效果。同时,需评估光照波动、水体扰动等环境因素对识别与计数的影响,以及图像采集、预处理、模型推理、策略判断、通信和执行机构响应共同构成的闭环控制时延;结合验证结果研究动态阈值调整机制,并在循环水养殖系统中开展实际投喂试验。

参考文献:

[1] 朱明, 张镇府, 黄凰, 等. 鱼类养殖智能投喂方法研究进展[J]. 农业工程学报, 2022, **38**(7): 38-47. DOI: 10.11975/j.issn.1002-6819.2022.07.005.
ZHU M, ZHANG Z F, HUANG H, *et al*. Research progress on intelligent feeding methods in fish farming[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, **38**(7): 38-47. DOI: 10.11975/j. issn. 1002-6819.2022.07.005 (Ch)

[2] 刘思,俞国燕. 工厂化养殖自动投喂饵料系统研究进展[J]. 渔业现代化, 2017, **44**(2): 1-5.
LIU S, YU G. Progress of research on automatic feeding system infactory aquaculture[J]. *Fishery Modernization*, 2017, **44**(2): 1-5 (Ch).

[3] ZHOU C, XU D M, CHEN L, *et al*. Evaluation of fish feeding intensity in aquaculture using a convolutional neural network and machine vision[J]. *Aquaculture*, 2019, **507**: 457-465. DOI: 10.1016/j.aquaculture.2019.04.056.

[4] ZHANG Y Q, XU C, DU R X, *et al*. MSIF-MobileNetV3: An improved MobileNetV3 based on multi-scale information fusion for fish feeding behavior analysis[J].

Aquacultural Engineering, 2023, **102**: 102338. DOI:10.1016/j.aquaeng.2023.102338.

[5] ZHANG L, WANG J P, LI B B, *et al*. A MobileNetV2-SENet-based method for identifying fish school feeding behavior[J]. *Aquacultural Engineering*, 2022, **99**: 102288. DOI:10.1016/j.aquaeng.2022.102288.

[6] YANG L, CHEN Y Y, SHEN T, *et al*. A BlendMask-VoVNetV2 method for quantifying fish school feeding behavior in industrial aquaculture[J]. *Computers and Electronics in Agriculture*, 2023, **211**: 108005.

[7] HOU S Y, LIU J C, WANG Y Q, *et al*. Research on fish bait particles counting model based on improved MCNN[J]. *Computers and Electronics in Agriculture*, 2022, **196**: 106858. DOI:10.1016/j.compag.2022.106858.

[8] PENG X H, ZHOU T Y, ZHANG Y, *et al*. Automatic shrimp fry counting method using multi-scale attention fusion [J]. *Sensors*, 2024, **24**(9): 2916. DOI:10.3390/s24092916.

[9] ZHANG J J, YANG G W, SUN L H, *et al*. Shrimp egg counting with fully convolutional regression network and generative adversarial network [J]. *Aquacultural Engineering*, 2021, **94**: 102175. DOI:10.1016/j.aquaeng.2021.102175.

[10] LI W R, ZHU Q, ZHANG H Y, *et al*. A lightweight network for portable fry counting devices[J]. *Applied Soft Computing*, 2023, **136**: 110140. DOI:10.1016/j.asoc.2023.110140.

[11] CIPOLLA R, GAL Y, KENDALL A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 7482-7491. DOI:10.1109/CVPR.2018.00781.

[12] CHEN Z, BADRINARAYANAN V, LEE C Y, *et al*. GradNorm: Gradient normalization for adaptive loss balancing in deep multitask networks[C/OL].[2026-09-02]. https://proceedings.mlr.press/v80/chen18a/chen18a.pdf.

[14] NAVON A, SHAMSIAN A, ACHITUVE I, *et al*. Multi-task learning as a bargaining game[PP/OL].[2026-09-02].https://doi.org/10.48550/arXiv.2202.01017.

[13] LYU H, SHA N Y, QIN S Y, *et al*. Advances in neural information processing systems [C/OL]. [2026-09-02]. https://par.nsf.gov/servlets/purl/10195511.

[15] ZHAO H X ,CUI H W ,QU K M, *et al*.A fish appetite assessment method based on improved ByteTrack and spatiotemporal graph convolutional network [J]. *Biosystems Engineering*,2024,**240**:46-55.DOI:10.1016/j.biosystemseng.2024.02.011.

[16] 刘欢, 张培珍, 沈晨, 等. 大型网箱养殖鱼群声信号短时频谱特性[J]. 广东海洋大学学报, 2023, **43**(3): 17-25. DOI:10.3969/j.issn.1673-9159.2023.03.003.
LIU H,ZHANG P Z, SHEN C, *et al*. Short-time characteristics of acoustic signals of fish farming in large cages [J]. *Journal of Guangdong Ocean University*, 2023, **43**(3): 17-25. DOI:10.3969/j.issn.1673-9159.2023.03.003 (Ch).

[17] 李尹佳, 胡泽元, 涂万, 等. 基于改进YOLOv7的密集鱼群计数检测[J]. 广东海洋大学学报, 2024, **44**(2): 115-123. DOI:10.3969/j.issn.1673-9159.2024.02.015.
LI Y J, HU Z Y, TU W, *et al*. Dense fish population counting detection based on improved YOLOv7[J]. *Journal of Guangdong Ocean University*, 2024, **44**(2): 115-123. DOI:10.3969/j.issn.1673-9159.2024.02.015 (Ch).

[18] HE K M, ZHANG X Y, REN S Q, *et al*. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE Press, 2016: 770-778. DOI:10.1109/CVPR.2016.90.

[19] WANG Z, BOVIK A C, SHEIKH H R, *et al*. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600-612. DOI:10.1109/TIP.2003.819861.

[20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, **60**(6): 84-90. DOI:10.1145/3065386.

[21] MA N N, ZHANG X Y, ZHENG H T, *et al*. ShuffleNet V2: practical guidelines for efficient CNN architecture design [C]//Computer Vision—ECCV 2018. Cham: Springer, 2018: 122-138. DOI:10.1007/978-3-030-01264-9_8.

[22] HOWARD A, SANDLER M, CHU G, *et al*. Searching for MobileNetV3 [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). New York: IEEE Press, 2019: 1314-1324. DOI:10.1109/iccv.2019.00140.

[23] ZHANG Y Y, ZHOU D S, CHEN S Q, *et al*. Single-image crowd counting *via* multi-column convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE Press, 2016: 589-597. DOI:10.1109/CVPR.2016.70.

[24] LI Y H, ZHANG X F, CHEN D M. CSRNet: Dilated convolutional neural networks for understanding the highly congested scenes [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2018: 1091-1100. DOI:10.1109/CVPR.2018.00120.

[25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[PP/OL].[2026-09-02].https://doi.org/10.48550/arXiv.1409.1556.

□