

文章编号 1004-924X(2023)14-2123-12

## 结合改进 ConvNeXt 网络与知识蒸馏的天气识别

刘立波, 郝思宇, 邓 箴\*

(宁夏大学 信息工程学院, 宁夏 银川 750021)

**摘要:**为提升复杂交通场景下天气识别准确率的同时实现网络轻量化,提出了一种结合改进 ConvNeXt 网络与知识蒸馏的天气识别方法。首先,在 ConvNeXt 网络的每组 Block 特征提取块后加入 SimAm 注意力机制,构建 ConvNeXt\_F 网络,利用 SimAm 注意力机制对 Block 块提取的深层特征进行鉴权并校正权重,有效强化对天气判别性特征的捕获能力;其次,在网络训练过程中将 Equalized Focal Loss (EFL) 与 Mutual-Channel Loss (MCL) 采用平均占比的方式进行累加作为总损失函数,一方面利用 EFL 消除数据不均衡造成的影响,另一方面利用 MCL 减小同类天气下局部细节特征差异;最后,采用知识蒸馏技术将天气分类知识从 ConvNeXt\_F 网络迁移到轻量级 MobileNetV3 网络,虽然精度略微损失但网络参数量大幅减少。实验结果表明,与其他算法相比,所提方法在本文构建的宁夏高速公路场景下的天气数据集 weather-traffic 和公开的自然天气数据集 RSCM2017 上准确率分别达到 96.22%, 84.8%, FPS 分别达到 157.6 Hz, 137.6 Hz, FLOPs 和 Params 仅为 0.06 G 和 2.54 M, 识别精度、速度和网络的轻量化较原网络均有提高,能够更好地应用于储存和计算能力受限的实际场景中。

**关键词:**天气识别; ConvNeXt 网络; 注意力机制; 知识蒸馏

**中图分类号:** TP394.1; TH691.9 **文献标识码:** A **doi:** 10.37188/OPE.20233114.2123

## Weather recognition combining improved ConvNeXt models with knowledge distillation

LIU Libo, XI Siyu, DENG Zhen\*

(School of Information Engineering, Ningxia University, Yinchuan 750021, China)

\* Corresponding author, E-mail: dengzhen1025@163.com

**Abstract:** A weather recognition method combining an improved ConvNeXt network and knowledge distillation is proposed to improve the accuracy of weather recognition in complex traffic scenes while achieving network lightweighting. Firstly, the ConvNeXt\_F network was constructed, and the SimAm attention mechanism was added after each set of Block feature extraction of the ConvNeXt network to correct the weights of the extracted deep features and strengthen the ability to capture discriminative weather features. Secondly, during the network training, equalized focal loss (EFL) and mutual-channel loss (MCL) were aggregated as the total loss function by using the average occupancy ratio, eliminating the effect caused by data imbalance using EFL and reducing the difference of local detail features under similar weather using MCL. Finally, the knowledge distillation technique was used to migrate the weather classification knowledge from the ConvNeXt\_F network to the lightweight MobileNetV3 network, which has a

收稿日期: 2022-11-11; 修订日期: 2022-12-07.

基金项目: 宁夏回族自治区重点研发计划项目 (No. 2021BEG03024); 国家自然科学基金资助项目 (No. 62262053); 宁夏科技创新领军人才计划资助项目 (No. 2022GKLRX03)

marginal loss of accuracy but significant reduction in the number of network parameters. The experimental results showed that compared with other algorithms, the proposed method achieved 96.22% and 84.8% accuracy on the weather-traffic dataset of Ningxia expressway and publicly-available natural weather dataset RSCM2017, respectively; the FPSs were 157.6 Hz and 137.6 Hz and FLOPs and Params were 0.06 G and 2.54 M. Compared with the original network, the recognition accuracy, speed, and lightness of the network were improved, making it better applicable to practical scenarios with limited storage and computational power.

**Key words:** weather recognition; ConvNeXt network; attention mechanism; knowledge distillation

## 1 引言

恶劣天气是影响高速公路安全的重要因素之一。雨、雪、团雾等恶劣天气容易引发交通事故<sup>[1]</sup>,若能在交通路网的道路监控系统中,利用计算机视觉进行局部天气识别,实现天气识别的及时性和识别位置的具体性,是对高速公路监控系统下实时天气识别的有利支撑,也是对现有气象部门天气识别的有效补充,在提高效率和降低成本方面具有重要的研究价值和现实意义。

目前,基于计算机视觉的天气识别方法可分为两类:基于机器学习的天气识别方法和基于深度学习的天气识别方法。基于机器学习的天气识别方法,通过对图像进行定量分析,将图像中的每一个像元或区域归划为若干个类别中的一种。例如利用图像亮度、色调、位置、纹理和结构等特征,将不同信息域的特征组合成多维特征送入支持向量机、贝叶斯、K均值聚类机器学习方法中进行天气分类<sup>[2-8]</sup>。尽管基于机器学习的天气识别方法为天气分类任务提供了不少解决方案,但这些方法识别效果并不理想。

基于深度学习的天气识别方法,使图像识别技术避免了复杂的特征提取和数据重建过程,被研究人员广泛采用。陈思玮、刘奥强等<sup>[9-10]</sup>提出了基于本体网络的天气识别方法,陈思玮提出基于迁移 DenseNet 网络的天气分类方法、刘奥强提出基于迁移 ResNet 网络的天气分类方法,这两种方法解决了严重依靠特征提取、适用场景单一的问题。郭志强、左杰格等<sup>[11-12]</sup>将传统方法提取的浅层特征与 CNN 提取的深层特征融合用于天气图像分类。结合了天气图像浅层特征与深度特征,从不同角度更加详细地表达了图像信息,提高了天气识别的准确率。崔洪涛等<sup>[13]</sup>提出

了联合投票网络的交通场景天气分类方法,结合注意力与联合投票网络,提取丰富多样的特征信息,进一步提高了天气识别性能。这些方法虽然取得了较好的结果,但其模型性能的提升已经趋于饱和。ConvNeXt 网络突破了卷积神经网络所能达到的极限,且简单、高效,其先进的算法和网络结构能够为天气分类任务带来更好的效果。

由于 ConvNeXt 网络结构中,Block 采用单一的非线性激活层,使得模型对非线性特征拟合能力有所降低,在卷积过程中难以捕捉到复杂天气背景下的判别性特征。为了进一步提高网络的非线性拟合能力,使得提炼特征可以更聚焦主体目标,减少梯度消失风险,将 SimAM 注意力模块加入到 ConvNeXt 每组 Block 模块后,构建 ConvNeXt\_F 网络,以有效提高通道中判别性特征的权重。

在图像分类任务中,数据集是不可或缺的因素,鉴于公开的天气数据集均来源于自然场景,本研究自行收集并构建了宁夏高速公路场景下的天气数据集 weather-traffic,但各类天气数据存在不均衡现象,影响网络分类精度。为解决该问题,提出了由 Equalized Focal Loss(EFL)与 Mutual-Channel Loss(MCL)按照平均占比方式累加组成的总损失函数。在训练过程中利用 EFL 动态调节样本损失,以解决数据集不均衡造成的精度不足;同时借助 MCL 进一步挖掘天气图像中细微的判别特征,使得网络的准确率和泛化能力得以提高。

改进后的 ConvNeXt\_F 网络对设备储存和计算能力要求高,难以部署在实际应用场景中。为了满足现实应用场景需求,本文将 ConvNeXt\_F 作为教师网络,利用知识蒸馏技术将天气分类知识迁移到轻量级网络 MobileNetV3,虽

精度微小损失但大幅减少了网络参数量,应用于实际生产中具备优越性。

## 2 本文方法

现有的天气识别方法难以捕捉到复杂交通背景下的判别性天气特征,且引入了较多计算量,网络结构复杂,在对设备储存和计算能力要求较高的实际生产中应用受限。本文利用知识蒸馏技术(Knowledge Distillation, KD),教师网络采用以 ConvNeXt 为基础进行改进后得到的 ConvNeXt\_F,将预训练好的 ConvNeXt\_F 所习得的暗知识蒸馏到网络结构更简单、参数量更少的 MobileNetV3 中,最终将蒸馏训练后的学生网络(MobileNetV3\_Z)用于交通天气识别。

其中, KD 是由 HINTON 等<sup>[14]</sup>提出的知识迁移技术,借助混合损失函数的反向传播将教师网络所习得的先验知识传授给学生网络,以轻微的精度损失换来大幅的网络参数减少,以达到压缩教师网络和增强学生网络性能的目的。知识蒸馏结构中学生网络的性能效果与教师网络高度相关,因此选择性能优秀的教师网络至关重要。ConvNeXt 网络是由 Zhuang Liu 等<sup>[15]</sup>从 ResNet 出发借鉴 Swin Transformer 所提出的,在多分类任务中精度和运算性能均优于 Swin Transformer,同时保持了标准 CNN 的简单性、高效性,是教师网络的良选。为了进一步强化网络的判别性特征提取能力,本文将对其改进后的 ConvNeXt\_F 作为知识蒸馏的教师网络。学生网络选择更适用存储空间和功耗受限的轻量级网络 MobileNetV3,经过 V1, V2 两代积累,其性能和速度表现更为优异。

### 2.1 ConvNeXt\_F 网络结构

交通天气识别存在两个难点:其一是在复杂交通场景下,天气特征受到光照、反射、季节等多种因素影响,导致网络难以捕捉到更具判别性的天气特征;其二是各类天气数据不均衡影响分类精度。针对以上两点问题,本文在原始 ConvNeXt 网络的基础上,提出了一种改进的 ConvNeXt\_F 网络。ConvNeXt 将 Swin Transformer 的设计优势与 CNN 的归纳偏置优势相结合,以 ResNet 为基础重新设计,其 Block 优化细节如图 1 所示。

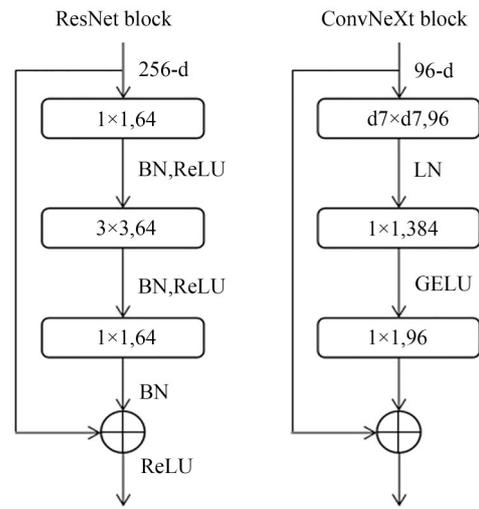


图 1 ResNet 与 ConvNeXt 的 Block 对比图

Fig. 1 Block comparison between ResNet and ConvNeXt

ConvNeXt 的 Block 将  $3 \times 3$  的卷积核替换为  $7 \times 7$  的大核卷积,卷积核位置上调至瓶颈部位,使用更少的 Batch Normalization 和 ReLU 激活函数,采用 Layer Normalization 替换 Batch Normalization,把 ReLU 替换为 GELU 激活函数。ConvNeXt 网络在图像分类领域中具有两个优势:一是特征提取能力较强,应用于下游分类任务泛化性高;二是网络简单高效,具备较好的可扩展性。

本文所提 ConvNeXt\_F 网络在原 ConvNeXt 网络卷积层、Block 层和线性层的基础上增加了注意力层,共 4 部分组成,结构如图 2 所示。利用注意力机制能够强化网络特征提取能力的特性,在每组 Block 后加入 SimAM 注意力机制,集成为新的特征提取块,负责提取天气判别特征。训练过程中利用 Equalized Focal Loss (EFL) 与 Mutual-Channel Loss (MCL) 按照平均占比的方式累加组成的损失函数,通过 EFL 动态调节样本贡献来消除数据不均衡造成的影响,同时利用 MCL 进一步强化网络对局部细节特征的判别能力。

具体流程是,先将天气图像转换为  $224 \times 224$  的特征图,进行 cutout 和均值化处理,利用 mixup 方法对数据增强,构建虚拟样本对,混合后输入网络。然后,使用  $4 \times 4$  的卷积核进行卷积操作,再进行线性归一化处理,通过四组重复堆叠的 Block 进行特征提取,提取的特征记为  $F_i, i \in \{1, 2, 3, 4\}$ ,每组 Block 提取的特征均经过 SimAM 注意力模块,对特征进行评估,评估过程用 *activa-*

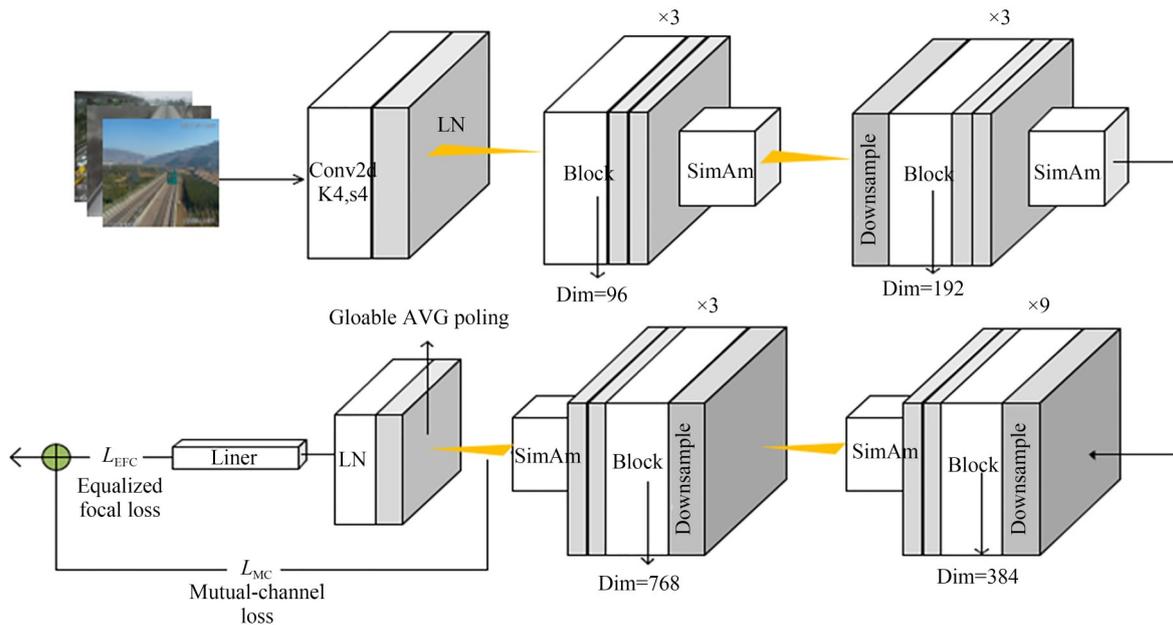


图2 融合 SimAM 注意力机制的 ConvNeXt\_F 网络结构图

Fig. 2 ConvNeXt\_F network structure diagram with SimAM attention mechanism

tion表示,再进行 *sigmoid* 激活后生成特征权值,与原特征相乘后输入下一组 Block,赋权过程记为  $F_i * \text{sigmoid}(\text{activation}(F_i))$ ,通过对  $F_1$  到  $F_4$  特征的不断赋权,可以提取到更具判别性的特征信息,最后经过线性层输出预测分类结果。训练过程中,预测分类结果与真实标签计算出 Equalized Focal Loss 损失  $L_{EFL}$ ,最后一组 Block 输出的特征  $F_4$  经过注意力模块评估后与真实标签计算出通道损失  $L_{MC}$ ,将  $L_{EFL}$  与  $L_{MC}$  以平均占比的方式累加后通过反向传播机制对网络参数进行更新。

## 2.2 SimAm 注意力机制的引入

基于卷积网络的天气识别方法在逐步降低分辨率时,容易丢失更具判别性的细节特征,使得天气分类准确性受到限制。针对这类问题,Li 等<sup>[16]</sup>基于残差思想提出了基于双线性注意机制 ConvNeXt 网络用于细粒度图像分类,有效保留了卷积过程中因维数的增加和减少所造成的信道间交互信息损失。Lingxiao 等<sup>[17]</sup>提出的 SimAM 注意力模块,嵌入到深度可分离卷积网络中,提炼的特征可以更好的聚焦主体目标。受此启发,为使网络在训练过程中关注更具判别性的天气特征,本文将骨干网络中的每组 Block 与 SimAM 注意力机制进行结合,以进一步强化特征提取,提高对细节信息的感知能力。

SimAM 注意力模块定义的能量函数,其最小化过程等效于度量每个通道中天气特征与其他特征的线性可分性,通过激活判别特征来抑制无效背景信息,进而寻找出最优天气特征,定义的能量函数如式(1)所示:

$$e^*_i = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda}, \quad (1)$$

其中: $t$ 是通道中的目标特征, $\hat{\sigma}^2$ 是通道中其他特征的方差, $\hat{\mu}$ 是通道中其他特征的均值, $\lambda$ 是添加的正则项。目标天气特征和周围非天气特征的区别越大,能量值  $e^*_i$  越小。因此可以用  $1/e^*_i$  来表示该特征的重要性。用缩放的操作来做特征提炼,整个模块的提炼过程如式(2)所示:

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{e^*_i}\right) \times X, \quad (2)$$

其中:*sigmoid* 是用来约束过大的值,不影响每个通道中特征值的相对大小。

## 2.3 损失函数的改进

在天气图像分类任务里,晴天图像相对于恶劣天气数量庞大且易于学习,虽然损失值小但在总损失中占比较大;恶劣天气由于数量较少且难以学习,容易被网络忽略而影响分类精度。

Lin 等<sup>[18]</sup>提出了 Focal Loss 损失函数,在交叉熵损失函数的基础上引入了调节因子,用来解决

数据不均衡问题,但是因为使用固定的因子值进行调整,无法处理长尾问题。公式如(3)所示:

$$FL(P_i) = -\alpha \cdot (1 - P_i)^\gamma \cdot \log(P_i), \quad (3)$$

其中: $\alpha$ 是类别权重, $\gamma$ 是难度权重指, $P_i$ 代表预测的概率。

Li等<sup>[19]</sup>提出的 Equalized Focal Loss,以 Focal Loss 为基础,在训练中利用动态调节因子方法,对样本的损失贡献进行平衡,解决了这一问题,公式如式(4)和式(5)所示:

$$EFL(P_i) = -\alpha \cdot (1 - P_i)^\gamma \cdot \log(P_i), \quad (4)$$

$$\begin{aligned} \gamma^j &= \gamma_b + \gamma_v^j \\ \gamma^j &= \gamma_b + s(2 - g^j) \end{aligned} \quad (5)$$

其中: $\gamma^j$ 表示第  $j$  类的 Focusing Factor,被解耦为 2 个组件, $\gamma_b$ 表示控制分类器基本行为的平衡数据场景中的 Focusing Factor, $\gamma_v^j$ 与类别相关,表示与第  $j$  类不平衡参数相关的变量参数, $g^j$ 表示第  $j$  类正样本与负样本的累积梯度比。

然而交叉熵损失采用了类间竞争机制,虽然可捕获到比较明显的类间差异,但只关心正确标签预测概率的准确性,忽略了其他非正确标签的差异,从而导致学习到的特征比较分散,难以关注到分类临界值模糊的判别性特征。天气识别实质上属于细粒度分类研究范畴,Chang等<sup>[20]</sup>提出的通道损失(Mutual-Channel Loss)由两个特定通道的组件组成:判别性组件  $L_{\text{dis}}$  和差异性组件  $L_{\text{div}}$ 。可以限制类与通道之间的特定关系,并保持类内和类间的可分离性,进而提高网络挖掘图像中细微判别特征的能力, $L_{\text{dis}}$ 和 $L_{\text{div}}$ 损失如式(6)~式(9)所示:

$$L_{\text{dis}}(F) = L_{\text{CE}}\left(y, \frac{\left[ e^{g(F_0)}, e^{g(F_1)}, \dots, e^{g(F_{c-1})} \right]^T}{\sum_{i=0}^{c-1} e^{g(F_i)}}\right), \quad (6)$$

$$g(F_i) = \frac{1}{WH} \sum_{k=1}^{WH} \max_{j=1,2,\dots,\xi} \left[ \underbrace{M_i \cdot F}_{\text{CWA}} \right], \quad (7)$$

$$L_{\text{div}}(F_i) = \frac{1}{c} \sum_{i=0}^{c-1} h(F_i), \quad (8)$$

$$h(F_i) = \sum_{k=1}^{WH} \max_{j=1,2,\dots,\xi} \left[ \underbrace{\frac{e^{F_{i,j,k}}}{\sum_{k'=1}^{WH} e^{F_{i,j,k'}}}}_{\text{Soft max}} \right], \quad (9)$$

其中: $F$ 表示特征, $i$ 为给定的目标类别, $\xi$ 用于表示每一个类的 channel 数, $g(\cdot)$ 定义如公式(7)所

示, $h(\cdot)$ 定义如式(9)所示, $GAP$ 、 $CCMP$ 和 $CWA$ 分别是全局平均池化,跨通道最大池化和通道维度的注意力机制。总损失  $L_{\text{MC}}$ 如式(10)所示:

$$L_{\text{MC}}(F) = L_{\text{dis}}(F) - \lambda \times L_{\text{div}}(F). \quad (10)$$

为了在平衡各类天气数据的同时进一步挖掘同类天气间的细微差异,本文将公式(4)和公式(10)的损失函数进行组合来开展训练,因为(4)计算出的损失与(10)计算出的损失相近,且互不影响,所以采用了平均占比的方式对两个损失进行累加,总目标损失为:

$$L_f = L_{\text{EFL}} + L_{\text{MC}}. \quad (11)$$

具体来说,在训练过程中利用 Equalized Focal Loss 约束预测值与真实值的差距,并将这种差距的模式定义为梯度的模,通过反向传播更新参数来改善数据不平衡,并将最后一组 Block 输出的深层特征与真实标签计算 Mutual-Channel Loss,探索他们的互补性,制约相似特征分布,强化网络的特征提取能力。

## 2.4 结合知识蒸馏的网络迁移

为了实现天气识别网络的轻量化,满足实际生产需求,本文利用知识蒸馏技术将 ConvNeXt\_F 习得的暗知识传递到轻量级 MobelnetV3 网络用于天气识别,方法流程如图 3 所示。首先,对初始天气图进行预处理,然后分别输入训练完备的教师网络(ConvNeXt\_F)和轻量级学生网络(MobileNetV3),生成天气预测的概率分布  $P_{ti}$ 和 $P_{si}$ ,并通过蒸馏温度  $T$ 软化概率分布,公式如(12)所示。

$$P_i' = \frac{\exp(P_i/T)}{\sum_{i=1}^n \exp(P_i/T)}. \quad (12)$$

将  $P_{ti}$ 和 $P_{si}$ 分别带入公式(12)中的  $P_i$ ,可得到软化后的标签  $P_{ti}'$ , $P_{si}'$ ,计算两者之间的交叉熵损失构成第一个目标函数  $L_{\text{soft}}$ ,公式如(13)所示:

$$L_{\text{soft}} = - \sum_{i=1}^K P_{ti}' \log P_{si}', \quad (13)$$

其次,计算学生预测  $P_{si}$ 和天气图像正确标签  $target$ 之间的交叉熵损失构成第二个目标函数,公式如(14)所示:

$$L_{\text{hard}} = - \sum_{i=1}^K target_i \log P_{si}. \quad (14)$$

最后,通过  $\alpha$ 调节  $L_{\text{hard}}$ 与 $L_{\text{soft}}$ 的占比组成总损失函数  $Loss$ ,公式如(15)所示:

$$Loss = \alpha L_{soft} + (1 - \alpha) L_{hard}, \quad (15)$$

其中,  $\alpha$  为权重参数, 取值范围  $(0, 1)$ , 由于知识蒸馏混合损失中, student 网络拟合  $L_{hard}$  为任务, 占比较大, 所以  $\alpha$  取值较小。利用总损失进行梯度反向传播, 更新学生网络参数, 实现天气分类知识从教师网络到学生网络的传递。

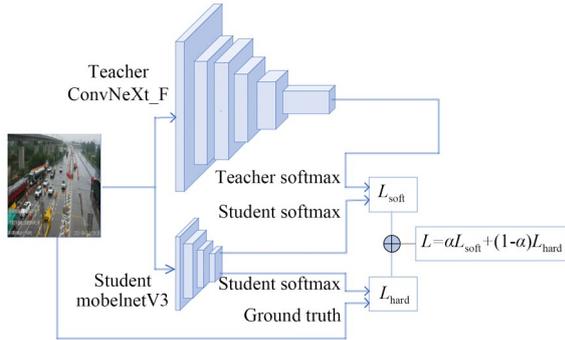


图 3 知识蒸馏整体流程图

Fig. 3 Overall flow chart of knowledge distillation

### 3 数据集及实验参数设置

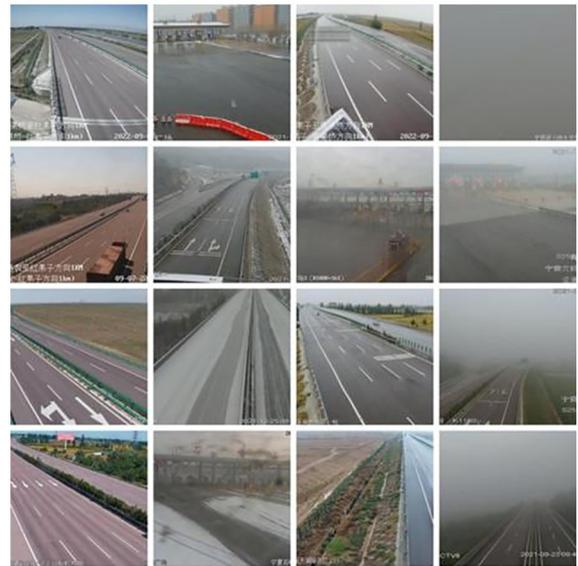
为进行公平比较, 在自建数据集 weather-traffic 和公开数据集 RSCM2017 上分别进行实验。weather-traffic 数据集的原始数据来源于宁夏交通厅提供的宁夏省内高速路摄像头视频, 收集包含晴、雪、雨、雾天气下的宁夏高速公路视频数据共计 200G。利用 python 脚本对所获取视频按帧截取处理, 初步获得 4 万张图像。由于视频质量、按帧截取获得的图像之间相似度极高等问题, 会严重干扰天气的识别准确率, 故需要进一步对获得的图像进行人工筛选工作, 首先删除因网络丢包导致花屏的图像, 然后删除无车辆时段背景过于相似的图片, 最后筛掉不带天气特征(雾气、雨滴、雪花)或者天气特征不明显的图像, 如雾天图像的近景区域, 雾气很淡, 近似于晴天图像。最终筛选出包含晴、雪、雨、雾四类天气的高速公路场景天气数据集, 共计 8 000 张, 分类情况如表 1 所示, 图 4 给出了各类天气在不同路段场景下的图片展示。RSCM2017 数据集由 Lin 等<sup>[21]</sup>构建, 该数据集为自然场景下的天气数据集, 包含 6 个天气类别(晴天、多云、下雨、下雪、薄雾和雷雨)总计 6.5 万张图像。

实验在 unubtu18.04 系统、NVIDIA Ge-

表 1 数据集内容描述

Tab. 1 Datasets content description

种类	描述	数量
Sunny	晴天	3 180
Snow	雪天	1 540
Rainy	雨天	1 580
Haze	雾天	1 700



(a) Sunny (b) Snow (c) Rainy (d) Haze

图 4 四类天气示例图片

Fig. 4 Example pictures of four types of weather

Force RTX3090, python 3.8, pytorch1.11, cuda11.4 环境下进行。同时利用深度学习框架 Pytorch 搭建多个网络。为保证实验结果公平性和变量唯一性, 设置相同网络参数: 输入图像尺寸为  $224 \times 224$ , 初始学习率设置为 0.000 1, 所有网络统一使用 Adam 优化器, 训练批次大小(batch size)为 64, 共进行 50 次 Epoch。两个数据集的训练集和测试集均按 7:3 进行划分。

图 4 中, 从左到右四列图片的场景依次为晴、雪、雨、雾。

## 4 实验结果及分析

### 4.1 评价指标

为了验证网络性能, 本文利用准确率(Accuracy)、FPS、网络浮点运算次数(FLOPs)和网络参数量(Params)这 4 个指标来评估性能。

准确率 (Accuracy) 主要由 TP (True Positive), FP (False Positive), FN (False Negative) 和 TN (True Negative) 4 个基础指标计算,其混淆矩阵如表 2 所示,计算如式 (16) 所示:

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP}. \quad (16)$$

表 2 混淆矩阵表

Tab. 2 Confusion matrix table

真是情况	预测结果(正类)	预测结果(负类)
正类	TP(真正例)	FN(假反例)
负类	FP(假正例)	TN(真反例)

FPS (Frames Per Second) 表示深度学习网络每秒能够识别出的图片数量,计算如式 (17) 所示:

$$FPS = \frac{frameCount}{elapsedTime}. \quad (17)$$

FLOPs (Floating Point Operations) 浮点运算次数,用来衡量网络计算复杂度,计算如式 (18) 所示:

$$FLOPs = 2 * H * W * (C_{in} * K * K) * C_{out}. \quad (18)$$

Params 网络参数量,用来衡量网络的大小,单个卷积核的参数量计算如式 (19) 所示:

$$Params = (K * K * C_{in}) * C_{out} + C_{out}. \quad (19)$$

公式 (17) 和公式 (18) 中,  $C_{in}$  是指卷积层输入 tensor 的通道数,  $C_{out}$  是指卷积层输出 tensor 的通道数,  $K$  是指卷积核大小。

## 4.2 对比试验

为证明所提方法应用于天气图像分类的有效性,本文选取 8 种方法用于对比试验,其中包括 DenseNet<sup>[9]</sup>, ResNet<sup>[10]</sup>, AlexNet<sup>[11]</sup>, VGG16-TL<sup>[12]</sup>, JVNet<sup>[13]</sup> 5 种天气识别网络, MobileNetV3<sup>[23]</sup> 1 种经典网络,以及 2021 年 Ze Liu 等<sup>[22]</sup> 提出的 Swin Transformer 和 2022 年 Zhuang Liu 等<sup>[15]</sup> 提出的 ConvNeXt。其中 DenseNet<sup>[9]</sup>, ResNet<sup>[10]</sup> 均是基于本体网络的天气识别方法, DenseNet 密集连接的结构能够强化特征传播有效缓解梯度消失问题, ResNet 借助残差网络结构以解决深度网络的退化问题; AlexNet<sup>[11]</sup>, VGG16-TL<sup>[12]</sup> 将传统方法提取的浅层特征与 CNN 提取的深层特征融合用于天气图像分类; JVNet<sup>[13]</sup> 将注意力机制与联合投票网络相结合,

提取丰富多样的特征信息; Swin transformer 将当前最火的 Transformer 应用与图像分类, ConvNeXt 基于 Swin Transformer 策略重新设计,以保留 CNN 的简单与高效。本文选用 Accuracy、FPS、网络浮点运算次数 (FLOPs) 和网络参数量 (Params) 作为量化指标在 weather-traffic 数据集上进行实验,结果如表 3 所示。

表 3 Weather-traffic 数据集不同方法的天气分类结果

Tab. 3 Weather classification results of different algorithms for weather-traffic dataset

Model	Accuracy/%	FPS/Hz	FLOPs/G	Params/M
DenseNet <sup>[9]</sup>	93.44	33.5	2.88	7.9
ResNet50 <sup>[10]</sup>	93.5	95.6	4.11	25.56
AlexNet <sup>[11]</sup>	93.5	95.6	0.72	61.1
VGG16-TL <sup>[12]</sup>	92.81	33.6	15.65	138.36
JVNet <sup>[13]</sup>	94.81	11.9	10.11	75.92
Swin transformer <sup>[22]</sup>	96.56	45.1	15.14	87.77
ConvNeXt <sup>[15]</sup>	96.83	108.9	4.46	28.57
<b>ConvNeXt_F</b>	<b>98.79</b>	<b>111.8</b>	<b>4.46</b>	<b>28.39</b>
MobileNetV3 <sup>[23]</sup>	83.58	132.7	0.06	2.54
<b>MobileNetV3_Z (ours)</b>	<b>96.22</b>	<b>157.6</b>	<b>0.06</b>	<b>2.54</b>

由表 3 可知,从网络性能来看,使用本文方法训练的网络均取得了较好的分类精度,证明所选用的训练策略具有优秀的鲁棒性。在现有的 5 种天气识别方法 DenseNet<sup>[9]</sup>, ResNet<sup>[10]</sup>, AlexNet<sup>[11]</sup>, VGG16-TL<sup>[12]</sup>, JVNet<sup>[13]</sup> 中, JVNet 准确率最高,达到了 94.81%, 但 FPS 最低仅有 11.9 Hz, ConvNeXt 因其设计策略与 Swin\_transformer 相关,因此分类准确率相近,但其利用深度可分离卷积在很大程度上减小了网络的大小,比 Swin\_transformer 更简单高效,本文网络 ConvNeXt\_F 准确率最高,达到 98.78%, 通过蒸馏训练后的 MobileNetV3\_Z 准确率达到 96.22%, 相比 ConvNeXt\_F 减少了 2.57%, 但相比原网络 MobileNetV3 提高了 12.64%, 其 FPS 为 157.6 Hz, 达到最高。从网络规格看, MobileNetV3\_Z 相比 ConvNeXt\_F 网络,在微小精度损失的前提下大幅度减小了网络的 FLOPs 和 Params, 表现出更优的综合性能。

为了验证本文方法的泛化性,对比实验同时在公开数据集 RSCM2017 下进行验证,训练策略不变,由于 FLOPs 和 Params 与表 3 一致,下表中将不再赘述。实验结果如表 4 所示。在公开数据集下,因为数据量大且分类目标多导致各个方法准确率相较于 weather-traffic 数据集较低,但从性能表现来看,本文方法仍优于其他方法。教师网络 ConvNeXt\_F 准确率依然最高,达到了 90.47%,蒸馏训练后的 MobileNetV3\_Z 准确率为 84.8%,相比教师网络 ConvNeXt\_F 降低了 5.67%,相比原网络 MobileNetV3 准确率提高了 11.53%,FPS 为 137.6 Hz,达到最高。FLOPs 和 Params 与表 3 一致,远低于其他网络,其综合性表现最佳。

表 4 RSCM2017 数据集不同方法的天气分类结果

Tab.4 Weather classification results of different algorithms for RSCM2017 dataset

Model	Accuracy/%	FPS/Hz
DenseNet <sup>[9]</sup>	76.87	33.3
ResNet50 <sup>[10]</sup>	82.63	86.5
AlexNet <sup>[11]</sup>	80.82	44.8
VGG16-TL <sup>[12]</sup>	80.3	31.6
JVNet <sup>[13]</sup>	85.15	81.6
Swin transformer <sup>[22]</sup>	87.31	25.6
ConvNeXt <sup>[15]</sup>	88.49	99.1
ConvNeXt_F	90.47	96.5
MobileNetV3 <sup>[23]</sup>	73.27	115.4
MobileNetV3_Z(ours)	84.80	137.6

综上所述,从表 3 和表 4 的实验结果可以发现本文改进后的教师网络 ConvNeXt\_F 在 weather-traffic 和 RSCM2017 两个数据集上的分类精度对比其他方法表现最优。蒸馏后的学生网络 MobileNetV3\_Z 尽可能的保留了教师网络的精度,同时大幅减小了网络的 FLOPs 和 Params,适用于储存资源有限和硬件水平低的生产场景中。

### 4.3 消融试验

注意力机制可以看作是类人脑模拟视觉信号处理的机制,是计算机视觉领域常用的特征强化方法。其中,SE 通道注意力可以有效捕获通道间的依赖关系,但只考虑通道信息,没有考虑位置信息。CBAM 由通道注意力机制和空间

注意力机制组成,有助于网络提取更深层次的特征信息,但只能在通道或者空间维度中的一个维度对特征进行精炼,在空间和通道同时变化的空间缺乏灵活性。SimAm 注意力机制无需向原始网络添加参数,利用能量函数可直接评估特征的重要性。为了进一步客观验证本文所用 SimAm 注意力机制的优越性,分别在 ConvNeXt 网络中加入 SENet, CBMA 和 SimAm 注意力机制进行对比实验。结果如表 5 所示。在 weather-traffic 和 RSCM2017 两个数据集上,引入 SENet 注意力机制后的准确率分别为 97.14%, 89.18%, 引入 CBMA 注意力机制后准确率分别为 97.36% 和 88.28%,本文引入的 SimAm 注意力机制的网络 ConvNeXt\_F 分类效果最佳,分别为 98.29% 和 90.29%。

表 5 不同注意力模块对比实验结果表

Tab.5 Comparison experiment results of different attention modules

Model	Accuracy(%)	
	weather-traffic	RSCM2017
+SENet	97.14	89.18
+CBMA	97.36	88.28
+SimAm	98.29	90.29

为了更直观地展示其作用,将不同注意力模块提取的天气特征生成热力图,结果如图 5 所示。

通过可视化结果的分析,可以得到与客观结果相符合的结论。晴天和雨天图像中,SimAm 注意力机制能够使网络更关注公路区域的特征信息,雪天图像中 SimAm 注意力机制和其他两个注意力机制关注的特征区域相近,雾天图像中,SimAm 注意力机制和 CBMA 注意力机制关注天气特征区域相近,部分焦距与公路区域,优于 SENet 注意力机制,综合来看,SimAm 注意力机制的性能优于 CBMA 和 SENet 注意力机制,能够更关注于公路区域天气特征,从而抑制背景冗余信息。

为了证明教师网络每个改进步骤的有效性,本文分别在两个数据集上对改进的注意力和损失函数分别进行验证,探索其对网络分类效果的作用。教师网络改进部分的消融实验如表 6 所

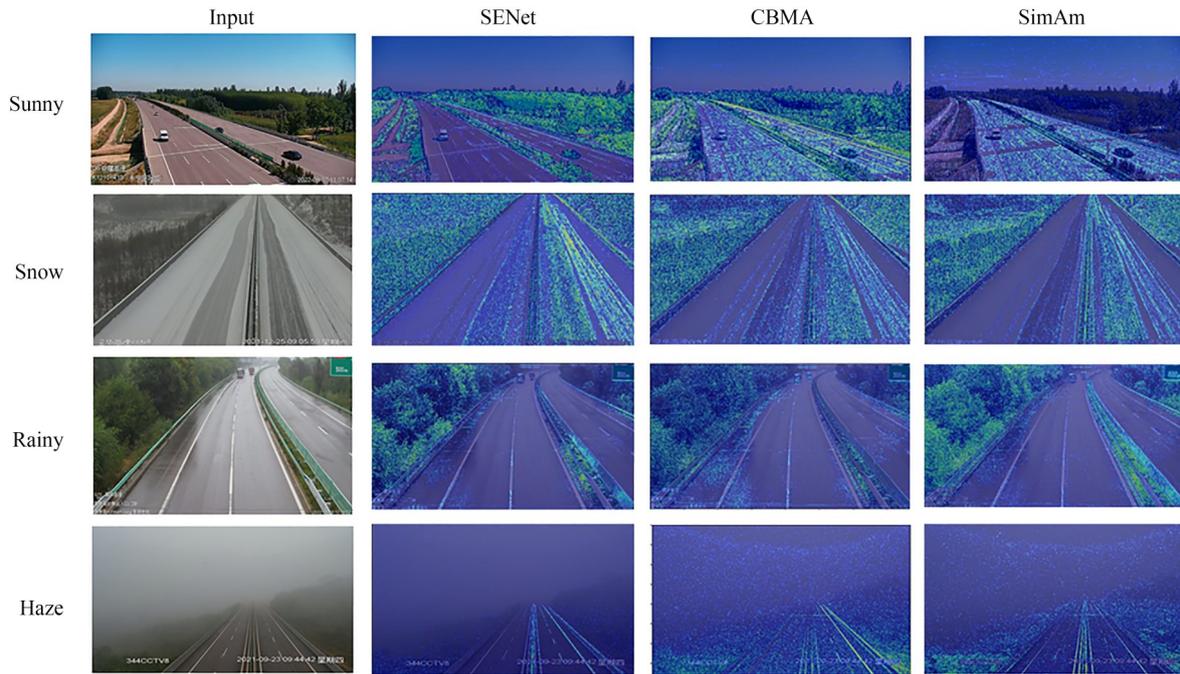


图 5 不同注意力机制热力图对比

Fig. 5 Comparison of heat maps of different attention mechanisms

示。在 weather-traffic 和 RSCM2017 两个数据集上,以 ConvNeXt 为基准网络,利用组合损失函数  $L_f$  作为目标函数进行训练,网络准确率分别达到 98.06% 和 89.29%,相比基准网络分别提高 1.12%,0.8%,引入 SimAm 注意力机制让网络在特征提取阶段充分捕获关键数据信息,抑制冗余信息干扰,网络准确率分别为 98.29% 和 90.29%,相比基准网络分别提高 1.35%,1.8%,通过实验结果可以证明本文改进方法有效,均能提升网络分类准确率。

表 6 教师网络消融实验结果

Tab. 6 Results of teacher model ablation experiment

Model	Accuracy/%	
	weather-traffic	RSCM2017
ConvNeXt	96.94	88.49
$L_f$	98.06	89.29
SimAm	98.29	90.29
$L_f + \text{SimAm}$	98.79	90.47

#### 4.4 知识蒸馏试验

蒸馏训练中,借助教师网络所习得的天气概率分布,通过蒸馏温度  $T$  对其软化后,利用组合

损失进行知识迁移,为探索蒸馏温度  $T$  对蒸馏学生网络性能的影响,本文依次设置  $T$  的取值为  $\{3, 5, 7, 9, 11, 15\}$  进行验证。表 7 为学生网络在不同蒸馏温度下的分类精度。

表 7 不同蒸馏温度实验结果

Tab. 7 Table of results of experiments at different distillation temperatures

蒸馏温度 $T$	Accuracy(%)	
	weather-traffic	RSCM2017
3	95.19	83.47
5	96.14	84.29
7	96.22	84.80
9	96.22	84.38
11	96.03	84.33
15	96.00	83.80

由表 7 可知,学生网络 MobileNetV3 分类精度随着蒸馏温度  $T$  的增大先有所提升后下降,这是由于随着  $T$  对目标软化力度增加,分类结果的泛化性逐渐增强,当到达一定临界值之后,对准确率的影晌会增大,从而降低分类精度,通过实验可知蒸馏温度  $T$  取值 7 时效果最佳。知识蒸馏借助混合损失函数的反向传播将教师网络所习

得的先验知识传授给学生网络,混合损失通过 $\alpha$ 调节 $L_{\text{hard}}$ 与 $L_{\text{soft}}$ 的占比组成,为了确定 $\alpha$ 值的大小范围对结果的影响,设置蒸馏温度为7的基础上,探索权重参数 $\alpha$ 对蒸馏学生网络性能的影响,因为student拟合 $L_{\text{hard}}$ 为主线任务,所以 $1-\alpha$ 占比较大, $\alpha$ 占比较小,因此本文依次设置 $\alpha$ 为 $\{0.2, 0.3, 0.4\}$ 。

表 8 不同混合损失权重参数实验结果

Tab.8 Experimental results of different mixed loss weight parameters

权重参数 $\alpha$	Accuracy(%)	
	weather-traffic	RSCM2017
0.2	95.83	84.07
0.3	96.22	84.80
0.4	95.68	84.31

由表 8 可知,在 weather-traffic 和 RSCM2017 两个数据集上,当 $\alpha$ 为 0.3, $1-\alpha$ 为 0.7 时,通过混合损失训练后学生网络 MobileNetV3 的拟合效果最佳。

## 5 结 论

为提高天气识别精度的同时实现网络的轻量化,本文提出结合改进 ConvNeXt 网络与知识

蒸馏的天气识别方法。首先构建了高速公路场景下的天气数据集 weather-traffic;其次在 ConvNeXt 网络的每组 Block 特征提取块后加入 SimAm 注意力机制,构建 ConvNeXt\_F 网络,提高网络的判别性特征提取能力,并在训练过程中采用平均占比方式累加 Equalized Focal Loss 和 Mutual-Channel Loss 作为总损失函数,消除数据不平衡影响的同时减小同类天气间的特征差异,进一步提高天气分类精度;最后利用知识蒸馏技术,将 ConvNeXt\_F 网络的天气分类知识迁移到轻量级网络 MobileNetV3\_Z,用于天气识别,提高了识别精度的同时实现网络的轻量化,能够满足实际生产需求。实验结果表明,在本文数据集 weather-traffic 和公开数据集 RSCM2017 上,本文方法 ConvNeXt\_F 准确率分别为 98.78%, 90.47%,蒸馏训练后的 MobileNetV3\_Z 网络准确率分别为 96.22%, 84.8%,相比 MobileNetV3 分别提高了 12.64%, 11.53%,FPS 分别为 157.6 Hz 和 137.6 Hz,网络的 FLOPs 和 Params 仅为 0.06 G 和 2.54 M,其综合性能优于现有方法。

但本文方法在应用中也存在问题,通常天气情况并非单一出现,如雨雾天、雨雪天等两种情况并存的恶劣天气,单标签识别就会存在局限性,未来将进一步探索视频下多标签天气识别,以期更好的应用于实践中。

## 参考文献:

- [1] 路欣,田辉,严玉彬,等. 低温雨雪冰冻对交通的影响与防范[C]. 第 27 届中国气象学会年会重大天气气候事件与应急气象服务分会场论文集, 2010: 344-347.
- LU X, TIAN H, YAN Y B, *et al.* Impact of Low Temperature, Rain, Snow and Freezing on Traffic and Prevention[C]. *Proceedings of the 27th Annual Meeting of the Chinese Meteorological Society on Major Weather and Climate Events and Emergency Meteorological Services*, 2010: 344-347. (in Chinese)
- [2] 张红艳,李茵茵,王伟. 改进 K 近邻和支持向量机相融合的天气识别[J]. 计算机工程与应用, 2014, 50(14): 148-151, 167.
- ZHANG H Y, LI Y Y, WAN W. Weather identifi-

- cation based on improved K nearest neighbor and support vector machine[J]. *Computer Engineering and Applications*, 2014, 50(14): 148-151, 167. (in Chinese)
- [3] 孟凡军,李天伟,徐冠雷,等. 基于 K 均值聚类算法的雾天识别方法研究[J]. 现代电子技术, 2015, 38(22): 80-83.
- MENG F J, LI T W, XU G L, *et al.* Research on method of foggy weather recognition based on K-means clustering algorithm[J]. *Modern Electronics Technique*, 2015, 38(22): 80-83. (in Chinese)
- [4] 于浩,王孝通,徐冠雷. 基于贝叶斯分类的雾天和雨天两类天气图像自动识别[J]. 舰船电子工程, 2016, 36(9): 73-75, 91.
- YU H, WANG X T, XU G L. Automatic recognition of two different weathers of foggy and rainy based on Bayesian classification[J]. *Ship Electronic*

- Engineering*, 2016, 36(9):73-75, 91. (in Chinese)
- [5] ZHANG Z, MA H, FU H, *et al.* Scene-free multi-class weather classification on single images [J]. *Neurocomputing*, 2016, 207: 365-373.
- [6] 史静, 朱虹, 韩勇. 户外天气状况分类识别[J]. 计算机系统应用, 2018, 27(4):259-263.
- SHI J, ZHU H, HAN Y. Outdoor weather classification [J]. *Computer Systems & Applications*, 2018, 27(4):259-263. (in Chinese)
- [7] 李鹏程, 吕昌峰, 于向茹, 等. 基于分区域特征提取的单张图像天气识别[J]. 济南大学学报(自然科学版), 2020, 34(4): 321-327.
- LIP C, LYU C F, YU X R, *et al.* Weather classification of single image based on sub-region feature extraction[J]. *Journal of University of Jinan (Science and Technology)*, 2020, 34(4): 321-327. (in Chinese)
- [8] 李冰村, 唐晓文, 何建新, 等. 基于机器学习的冰雹天气识别研究[J]. 气象科学, 2022, 42(5): 581-590.
- LI B C, TANG X W, HE J X, *et al.* Hail identification based on machine learning methods[J]. *Journal of the Meteorological Sciences*, 2022, 42(5): 581-590. (in Chinese)
- [9] 陈思玮, 贾克斌, 王聪聪, 等. 深度学习在多天气分类算法中的研究与应用[J]. 高技术通讯, 2020, 30(10):1010-1017.
- CHEN S W, JIA K B, WANG C C, *et al.* Research and application of deep learning in multi-weather classification algorithms[J]. *Chinese High Technology Letters*, 2020, 30(10): 1010-1017. (in Chinese)
- [10] 刘奥强, 张旭. 基于 ResNet50 的恶劣天气识别研究[J]. 智能计算机与应用, 2022, 12(4): 168-171.
- LIU A Q, ZHANG X. Research on severe weather recognition based on ResNet50 [J]. *Intelligent Computer and Applications*, 2022, 12(4): 168-171. (in Chinese)
- [11] 郭志强, 胡永武, 刘鹏, 等. 基于特征融合的室外天气图像分类[J]. 计算机应用, 2020, 40(4): 1023-1029.
- GUO ZH Q, HU Y W, LIU P, *et al.* Outdoor weather image classification based on feature fusion [J]. *Journal of Computer Applications*, 2020, 40(4):1023-1029. (in Chinese)
- [12] 左杰格, 柳晓鸣, 蔡兵. 基于图像分块与特征融合的户外图像天气识别[J]. 计算机科学, 2022, 49(3):197-203.
- ZUO J G, LIU X M, CAI B. Outdoor image weather recognition based on image blocks and feature fusion[J]. *Computer Science*, 2022, 49(3): 197-203. (in Chinese)
- [13] 崔洪涛, 曹科, 张虎, 等. 基于联合投票网络的交通场景天气分类方法[J]. 西南交通大学学报, 2021, 56(3):579-586.
- CUI H T, CAO K, ZHANG H, *et al.* Weather classification in traffic scene based on joint voting network[J]. *Journal of Southwest Jiaotong University*, 2021, 56(3):579-586. (in Chinese)
- [14] HINTON G, VINYALS O, DEAN J. Distilling the Knowledge in a Neural Network [EB/OL]. 2015: *arXiv*: 1503.02531. <https://arxiv.org/abs/1503.02531>
- [15] LIU Z, MAO H Z, WU C Y, *et al.* A ConvNet for the 2020s [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 11966-11976.
- [16] LI J, WANG C, HUANG B, *et al.* ConvNeXt-backbone HoVerNet for nuclei segmentation and classification [EB/OL]. 2022: *arXiv*: 2202.13560. <https://arxiv.org/abs/2202.13560>
- [17] YANG L, ZHANG R Y, Li L, *et al.* SimAM: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks [C]. *International conference on machine learning*. PMLR, 2021: 11863-11874.
- [18] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.
- [19] LI B, YAO Y Q, TAN J R, *et al.* Equalized Focal Loss for Dense Long-Tailed Object Detection [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 6980-6989.
- [20] CHANG D L, DING Y F, XIE J Y, *et al.* The devil is in the channels: mutual-channel loss for fine-grained image classification [J]. *IEEE Transactions on Image Processing*, 2020, 29: 4683-4695.
- [21] LIN D, LU C W, HUANG H, *et al.* RSCM: re-

- gion selection and concurrency model for multi-class weather recognition [J]. *IEEE Transactions on Image Processing*, 2017, 26(9): 4154-4167.
- [22] LIU Z, LIN Y T, CAO Y, *et al.* Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows [C]. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 10-17, 2021, *Montreal, QC, Canada*. IEEE, 2022: 9992-10002.
- [23] HOWARD A, SANDLER M, CHEN B, *et al.* Searching for MobileNetV3 [C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27 - November 2, 2019, *Seoul, Korea (South)*. IEEE, 2020: 1314-1324.

#### 作者简介:



刘立波(1974—),女,宁夏平罗人,博士,教授,博士生导师,2012年于北京理工大学博士后出站,主要从事智能信息处理、计算机视觉方面的研究。E-mail: liulib@163.com



郝思宇(1992—),男,宁夏中卫人,硕士研究生,主要研究方向为基于深度学习的天气识别。E-mail: xisiyu852@qq.com

#### 通讯作者:



邓 箴(1984—),女,河南三门峡人,博士,讲师,2006年于西安工业大学获得学士学位,2009年于北京科技大学获得硕士学位,2022年于宁夏大学获得博士学位,主要从事图像处理、机器视觉方面的研究。E-mail: dengzhen1025@163.com