doi:10.6041/j.issn.1000-1298.2025.02.018

基于轻量级 RepVIT 的农机具工况识别方法研究

安麒麟[1,2]　汪凤珠[1,2]　刘阳春[1,2]　邓　学[1,2]　周利明[1,2]　赵　博[1,2]　伟利国[1,2]
(1. 中国农业机械化科学研究院集团有限公司, 北京 100083; 2. 农业装备技术全国重点实验室, 北京 100083)

摘要: 为解决田间复杂环境下拖拉机机载农机具监测困难、模型参数量过大等问题,提出了一种基于轻量化 RepVIT 的农机具识别模型 TMAInet。利用自主开发的农机服务平台“农业机械化精准作业平台暨希望田野”收集了 6 种工作状态的农机具数据集,并通过 Copy-paste 等数据增强方法将训练集扩增至 6 627 幅。基于 RepVIT 网络模型框架,设计了一种卷积前馈模块(CFF)以提升不同尺度细粒度特征提取能力,引入了注意力机制 ECA 以优化模型参数结构并简化特征提取模块。通过 Pre-training + Fine-tuning(PF)迁移学习方法对模型进行了训练,并在 Jetson nano 边缘设备上进行了部署。实验结果表明,通过 PF 迁移学习方法,TMAInet 模型的识别准确率、F1 分数和召回率分别达到 99.13%、98.53% 和 98.78%,相较于原始的 RepVIT 模型分别提升 1.86、3.04、1.95 个百分点,在边缘设备端保持帧速率 73 f/s 的同时参数量降低至 7.3×10^6。TMAInet 能够在实际应用中准确、高效监测农机具常见类别,为无人化智慧农场的发展提供技术参考。

关键词: 拖拉机农机具; 工况识别; 无人化智慧农场; RepVIT; 迁移学习
中图分类号: TP391.41　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)02-0187-08　　**OSID**:

Lightweight RepVIT-based Working Condition Recognition Method for Agricultural Implements

AN Qilin[1,2]　WANG Fengzhu[1,2]　LIU Yangchun[1,2]　DENG Xue[1,2]　ZHOU Liming[1,2]　ZHAO Bo[1,2]　WEI Liguo[1,2]
(1. *Chinese Academy of Agricultural Mechanization Sciences Group Co., Ltd., Beijing* 100083, *China*
2. *State Key Laboratory of Agricultural Equipment Technology*, *Beijing* 100083, *China*)

Abstract: Aiming to address the problems of difficulty in monitoring tractor-mounted agricultural implements in complex field environments and the excessive amount of model parameters, a lightweight RepViT-based agricultural implements recognition model, tractor-mounted agricultural implements net (TMAInet), was proposed. Firstly, the self-developed agricultural machinery service platform 'Agricultural Mechanisation Precision Operation Platform' was used to collect the datasets of agricultural implements in six working states, and the training set was expanded to 6 627 frames by data enhancement methods such as copy-paste. Secondly, based on the RepVIT network model framework, a convolutional feed-forward module (CFF) was designed to improve the ability of fine-grained feature extraction at different scales, and an attention mechanism, ECA, was introduced to optimize the model parameter structure and simplify the feature extraction module. Finally, the model was trained by pre-training + fine-tuning (PF) migration learning method and deployed on Jetson nano edge devices. The experimental results showed that the recognition accuracy, F1 score and recall of the TMAInet model reached 99.13%, 98.53 and 98.78%, respectively, by the PF migration learning method. Compared with the original RepVIT model, the recognition accuracy, F1 score and recall were improved by 1.86 percentage points, 3.04 percentage points and 1.95 percentage points, respectively, and the number of parameters was reduced to 7.3×10^6 while maintaining 73 f/s at the edge device side. TMAInet was able to accurately and efficiently monitor the common categories of agricultural implements in practical applications, and it can provide a technical reference for the development of unmanned smart farms.

Key words: tractor-mounted agricultural implements; condition recognition; unmanned smart farms; RepVIT; transfer learning

收稿日期: 2024-10-10　修回日期: 2024-12-26
基金项目: 国家重点研发计划青年科学家项目(2022YFD2000300)
作者简介: 安麒麟(1998—),男,博士生,主要从事农业智能装备与信息化技术研究,E-mail: anqilin9@126.com
通信作者: 伟利国(1978—),男,研究员,博士生导师,主要从事农业智能装备与信息化技术研究,E-mail: weilg78@126.com

0 引言

农业机械化是建设现代农业的重要物质基础，也是实现农业现代化的重要标志与内容[1-2]。近年来，我国在政策补贴的支持下，农业机械化水平不断增长，截止 2023 年拖拉机保有量达到了 2 173.06 万台，配备农具 4 022.93 万部[3-4]。对于各类型农机具工况的识别是实现无人农场智能化管理的首要任务。因此，在实际应用环境下对农机具工况的识别尤为重要。

传统的农机具监测方法需要专用的设备，相关配套昂贵不便于实际的应用。文献[5]基于 ARM11 和 GNSS(全球导航卫星系统)开发了农机工况监测系统，主要对农机具工作面积和运行速度等信息进行监测。文献[6]基于 YOLO 模型开发了拖拉机安全监测系统，对农机具周围的障碍物进行检测和定位。文献[7]使用卷积神经网络对圆盘割草机进行监测，根据频谱振动数据检测出农机具故障状况。文献[8]基于优化极限学习遗传算法设计了一种农机田间地面识别模型，通过监测拖拉机上安装的加速度传感器以获取震动信号，从而通过算法对其地面信息进行分析和识别。文献[9]通过数据驱动的方式对犁地作业期间的工作状态进行全面分析，将实时捕获到的 CAN-Bus 数据与 GNSS 数据合并导入随机森林算法，以监测怠速、移动、转动和工作状态等不同犁地状态。

人工智能技术在农机领域发展迅速，但由于农机具种类繁多、工况复杂，相关监测技术应用尚不成熟。文献[10]基于 TBOX 开发了一种农机作业监控图像采集系统，通过远程的摄像头对农机和机具进行监测，其主要通过人工在云端对深松深翻机具识别。文献[11]开发了一种农业机械识别神经网络模型，用以在耕整机上对机具进行识别，但其识别速度仅为 3.33 f/s，准确率也仅为 81%。文献[12]开发了一种农机图像识别网络，能够自动地区分不同环境下的农机，其网络结构相对简单，只有少数类型农机具能识别且准确率低于 90%。已有的模型参数量过大不便于边缘端设备的部署，因此本文重点研究模型的轻量化。

模型轻量化是视觉领域研究重点，其目标是在降低计算成本的同时保持优异性能[13]。近年，学者主要研究卷积神经网络 CNN 的轻量化，提出了残差结构、通道打乱、可分离卷积和结构参数化等策略，产生了 ReNet、MobileNets、ShuffleNet 等具有代表性的轻量化网络模型[14-16]。另一方面，Vision Transformers(VIT)模型成为高效视觉表征方法的另一种主要策略[17]。研究人员探索了对 VIT 模型结构的轻量化设计，提出了 EfficienFormer、FastViT、LcViT 和 ReVIT 等轻量化算法[18-20]。然而，现有的算法在农机具实际田间移动边缘设备端仍存在较大的延迟，且模型参数量不便于部署和应用。

针对模型参数量大、识别精度差、响应速度慢等问题，本文提出一种轻量化农机具识别模型 TMAInet(Tractor-mounted agricultural implements net)。基于云平台收集 6 类复杂环境下农机具数据集，包含深翻、深松、播种、喷药、收获和打捆。通过嵌入 ECA、设计卷积前馈模块 CFF 并简化 RepVIT 特征提取模块，建立轻量化农机具识别模型 TMAInet；结合迁移学习方法对模型进行优化和部署，将其部署于 Jetson nano 边缘设备端进行测试与应用，以期应用于实际田间环境时具备良好的可移植性和泛化性。

1 数据与方法

1.1 数据集建立

本研究基于“农业机械化精准作业平台暨希望田野”平台对数据集进行采集。该平台由中国农业机械化科学研究院集团有限公司开发，已在吉林、黑龙江、山东等多个省市自治区上线，并成为吉林省、青岛市等地的官方农业机械化服务平台。平台具备农机作业监测、社会化服务、作业补贴管理等功能，可对“耕”、“种”、“管”、“收”各环节的农机作业情况进行实时监测与管理。截至目前，平台累计生成日报 10 万份、作业面积约 1.7×10^6 hm^2，管理补贴资金逾 2.8 亿元。

数据集所使用的数据均来源于上述已接入平台服务的多省份区域，包括东北地区、华北与华东地区等地的田间作业现场。各区域在地形、土壤和气候条件上各不相同，具备典型的旱作与水田分布以及多元化的农作物轮作制度。数据集覆盖 2022 年 1 月至 2024 年 1 月，涵盖整地、播种、田间管理、秋季收获等不同季节性农事活动。并包含多种天气与光照条件下的影像数据，如晴天、阴雨、晨昏、强光与弱光等情境，从而提升模型泛化多样性，数据如图 1 所示。

数据由 APP 安装于拖拉机后端位置的摄像头定时采集，主要聚焦于玉-麦轮作体系中常用的深翻、深松、播种、喷药、收获和打捆 6 类典型农机具作业过程。原始数据共收集 3 157 幅(分辨率 320 像素×240 像素)(表 1)，以比例 7∶2∶1 随机划分为训练集 2 209 幅、测试集 631 幅和验证集 317 幅。此外，为了进一步提高模型在不同环境下的泛化能力，



图 1　机载农机具数据集

Fig. 1　TMAI datasets

表 1　农机具数据集样本数

Tab. 1　Details of TMAI dataset

类别号	种类	标签	原始数据
0	深松	shensong	473
1	深翻	shenfan	514
2	播种	bozhong	505
3	喷药	penyao	630
4	收获	shouhuo	610
5	打捆	dakun	425
总计			3 157

对训练集进行数据增强，通过翻转、高斯噪声、亮度锐度、Copy-paste 和 Mosaic 等方法，将训练集扩增至 6 627 幅。

1.2　方法

1.2.1　农机具识别模型框架

本研究参考 RepVIT 结构[21]设计了农机具识别模型 TMAInet。其中 RepVIT 通过结构重参数化整合了 ViTs 的关键设计理念，通过 early convolutions 取代了传统的 Stem 模块，分离了 Token Mixer 和 Channel Mixer，并增强了下采样层设计。此外，还简化了分类头，模型使用 3 ×3 大核卷积，同时交叉利用 SE 层。实现了在移动设备上更高效的视觉任务处理，显著提升了模型的性能和效率。图 2 为本研究所搭建的 TMAInet 模型结构，在宏观尺度上减少了冗余的特征提取模块的同时，嵌入了 RepVIT – ECA 和卷积前馈网络（Convolutional feedforward，CFF）模块。

模型结构的具体流程如下：当特征输入网络模型时，首先通过 3 ×3 的 early convolutions 增加通道数，以减少延迟并提高优化的稳定性。其次进入主干部分，该部分由主要由 RepViT – ECA 模块组成，每个 Block 由 3 ×3 深度卷积、1 ×1 卷积以及嵌入的 ECA 模块组成，通道数逐渐增大，从 48 依次提升至 96、192。在降采样模块中，使用 3 × 3DW 卷积和 1 ×1 卷积进行空间降采样，同时进一步增加通道数以准备进入下一阶段。最终，特征图通过全局平均池化层和线性层处理，生成农机具识别分类结果。在测试中发现，当模型迁移应用于农机具识别任务时，保持 3 个 Block 效果最佳，因此在主干中删除了冗余的特征提取模块。通过上述设计，本研究在农机具识别任务上保持高性能的同时，减少了模型延迟。

1.2.2　CFF 结构

本研究在模型的训练过程中对各个结构进行了优化和改进。原始的前馈网络 FFN 结构通常在视觉算法模型中用作通道混合器，主要由 2 个线性层和 1 个中间激活函数组成，以捕获特征信息，如图 3a 所示。该设计使得模型能够在通道维度上进行特征交换，其特征提取能力有限。本研究首先在主干激活函数前嵌入一个最小结构形式的 3 ×3 深度卷积，如图 3b 所示。其次，在主干中嵌入门控线性单元（GLU），该结构在相关模型中的效果均优于多层感知器（MLP），所嵌入的 GLU 结构由元素相乘的两个线性投影支路组成。将其转换为基于近邻特

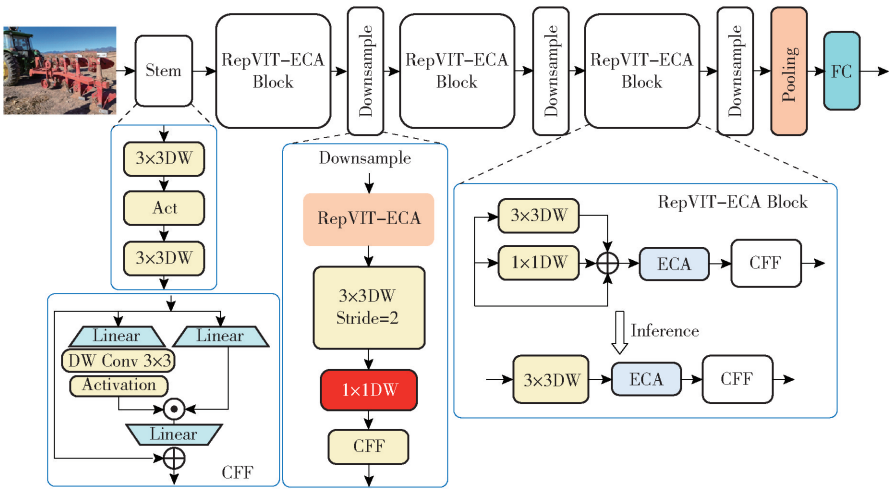


图 2　TMAInet 模型结构

Fig. 2　Model structure of tractor-mounted agricultural implements net（TMAInet）

征的门控通道特征提取网络,称为卷积前馈网络(Convolutional feedforward,CFF),嵌入在 RepVIT - ECA 模块中。

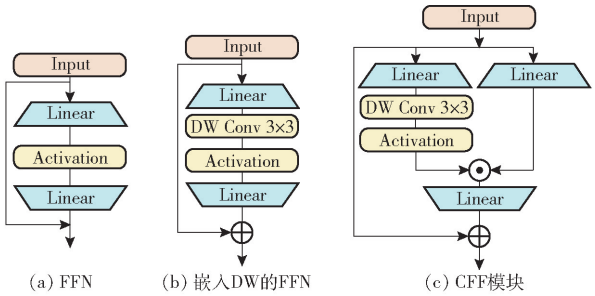


图 3 前馈网络结构对比

Fig. 3 Comparison of feedforward network structures

经过 CFF 结构处理的每个特征均由其最近的细粒度特征分配一个独特的门控信号,通过该信号缓解全局平均池化模块中因特征过度简化而导致的信息丢失问题。这种设计既能通过深度卷积为模型提供位置编码,又能使得模型保持与 MLP 和 GLU 相同的深度,保持良好的反向传播。CFF 模型复杂度为 $2RHWC^2 + 2RHWCk^2/3$(高度 H、宽度 W、通道数 C、扩展比 R、卷积核尺寸 $k \times k$),结构简单且稳健,使模型能够保持轻量化。

1.2.3 注意力机制

为提升模型识别准确性与泛化能力,引入轻量化注意力机制 ECA(Efficient channel attention)。该机制的工作原理是在卷积神经网络中对通道特征进行自适应再加权,从而通过局部跨通道交互来增强特征表征能力。

如图 4 所示,ECA 模块首先使用全局平均池化获取到全局的特征描述。然后使用 1D 卷积核对特征向量在通道维度上进行局部交互,以捕获相邻通道之间的关系。再将卷积输出通过 Sigmoid 函数进行归一化,最后再与原始特征向量进行逐通道相乘,得到加权后的特征向量。

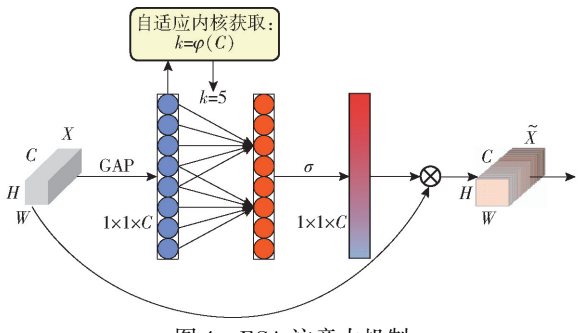


图 4 ECA 注意力机制

Fig. 4 Efficient channel attention (ECA)

ECA 模块具体流程如下:

(1)全局平均池化 GAP:当输入特征图时,ECA 模块首先通过 GAP 得到 $1 \times 1 \times C$ 的描述向量,GAP 将输入的特征图在空间维度上对其特征进行压缩,公式为

$$z_c = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} X_c(i,j) \tag{1}$$

其中 $X_c(i,j)$ 表示 C 个通道上位置(i,j)的特征值。

(2)局部跨通道交互:使用 1D 卷积在通道维度进行局部交互,1D 卷积核尺寸 k 基于通道数 C 动态获取,公式为

$$k = \varphi(C) = \left| \frac{\mathrm{lb}C}{\gamma} + b \right| \tag{2}$$

式中 γ——全局平均池化输出系数

b——自适应卷积核偏置

(3)自适应加权:将自适应卷积核嵌入一维卷积中,获取各通道自适应权重。进行卷积核尺寸为 k 的一维卷积操作,并通过 Sigmoid 激活函数得到各通道权重,公式为

$$S_c = \mathrm{Sig}(\mathrm{Conv1D}(z)) \tag{3}$$

式中 Sig——Sigmoid 激活函数

Conv1D——1D 卷积操作函数

z——式(1)中获取的全局平均池化结果

(4)特征重加权:将 S_c 与原始特征 X 逐通道相乘,为每个通道分配不同的权重,提升特征表示能力,公式为

$$\widetilde{X}_c = S_c X_c \tag{4}$$

1.3 训练策略

迁移学习是一种常用的训练策略,该方法在大模型数据集上进行预训练,再将主干特征模块迁移至特定数据集并进行微调,从而快速建立新模型。如图 5 所示,本研究使用 Stanford Cars 大型数据集作为网络预训练模型的源域,将模型的卷积参数和权重迁移至农机具识别任务中。Stanford Cars 大型数据集是由斯坦福大学人工智能实验室开发的细粒度车辆分类数据集,包含 196 种汽车类型数据,由 16 158 幅不同类型的特征数据组成。

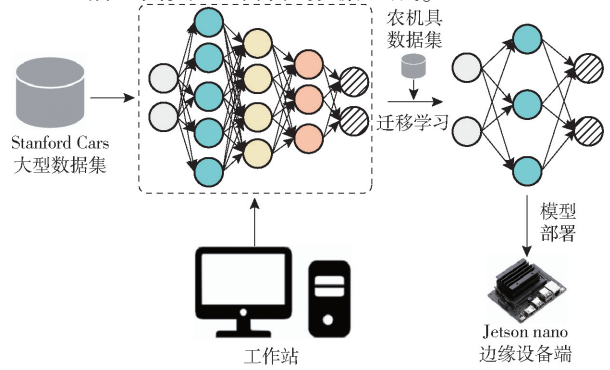


图 5 迁移学习训练流程

Fig. 5 Transfer learning training process

通过已训练得到的先验知识,可以在类似的识别任务中实现高效的模型建立。其次,在迁移训练过程中对模型的参数进行微调,并通过建立的农机具数据集进行新任务的训练。最终,将网络的输出转换为适合于农机具分类任务的 6 类输出,实现对农机具识别任务的准确预测。本研究以 Pre-training + Fine-tuning 融合的方式对模型进行迁移学习,通过 Pre-training 快速学习获取通用特征,丰富浅层特征提取能力。另一方面通过 Fine-tuning 优化对农机具小数据集的特征提取能力,增加模型准确性和泛化能力。

2　实验与验证

2.1　实验平台

实验的训练和测试分别在工作站和 Jetson nano 边缘设备端上进行。工作站配备了 RTX 3090 GPU、英特尔(R) Gold 6234 @3.30 GHz 处理器和 192 GB 内存。使用 Python 3.9 对算法进行编程,深度学习框架为 PyTorch 1.8.2 GPU 版本。将训练好的模型迁移至 Jetson nano 小型人工智能开发板,该设备属于边缘设备,体积较为小巧便于安装于各类设备上。Jetson nano 预装了 Ubuntu 18.04LTS 系统,搭载 128 核 Maxwell GPU 和 4 core ARM A57 @1.43 GHz 处理器。

2.2　训练方法对比实验

为了探索不同的迁移方式对模型精度的影响,从而选择出适用于 TMAInet 模型的最优迁移学习训练方法,选用 New learning(NL)、Frozen pretrained network(FPN)、Fine-tuning(FT)和 Pre-training + Fine-tuning(PF)共 4 种模型进行训练。除 NL 方法外,所有模型均在大数据集上进行了 300 轮初始训练,然后再通过农机具数据集进行 100 轮训练。

实验结果显示在表 2 中,NL 方法直接使用农机具数据集进行训练,虽然节省了时间,但其准确率和 F1 分数仅为 77.95% 和 77.33%,效果较差,无法实现对农机具的监测。FPN 方法通过冻结预训练前几层模块参数,从而仅对新任务中的最后几层进行训练,其准确率也仅达到 88.92%。

表 2　训练策略对比

Tab.2　Comparison of training methods　%

模型	准确率	F1 分数	召回率
NL	77.95	77.33	76.71
FPN	88.92	87.72	88.71
FT	93.83	93.49	93.16
PF	99.13	98.53	98.78

在相同条件下,本文所使用的 PF 方法效果最佳,首先通过相似的数据集 Standford Cars 对模型进行了 Frozen 处理,再通过 PF 方法迁移至农机具数据集对欠缺的网络模块进行 Fine-tuning 处理。Pre-training 提供了丰富的基础适应特性,Fine-tuning 使模型更好地适应农机具识别任务。使得 TMAInet 模型准确率、F1 分数和召回率高达 99.13%、98.53% 和 98.78%,优于其它迁移学习训练方法。

不同训练方法的混淆矩阵结果如图 6 所示,混淆矩阵通过数字的方式直观地展示出模型在不同类别上的识别能力。其中,对角线表示正确分类目标数量,数值越大模型识别性能越好。相反,非对角线表示错误的分类情况,反映了不同类别间的混淆程度,数值越大说明这些类别直接混淆越严重。从混淆矩阵中可以清晰地看出,PF 方法识别效果最好,特别是对于类别 5 的打捆机具。由于 PF 方法使用了预训练和微调融合的方式,增强了模型深层特征关联能力。

2.3　对比实验

为验证模型的有效性,选取 7 种较为流行的轻量级模型,在 TMAI 数据集上进行测试,包括 ReNet18、ShuffleNetV2、MobileNetV3、FastViT、EfficienFormer、RepVIT 和 TMAInet。通过参数量、耗时、召回率和准确率等指标进行评估,以全面对比各模型在 TMAI 识别任务中的表现。

训练时几种模型训练损失值、验证损失值、准确率和参数量曲线如图 7 所示。通过对比这些指标,能够深入分析不同模型在训练中的收敛速度、泛化能力以及模型复杂度之间的关系。图 7a、7b 为模型损失曲线,随着迭代次数的增加损失值和斜率都逐渐降低。当迭代 60 轮后损失值趋于稳定。TMAInet 都明显低于其他模型,能够迅速平滑下降且收敛速度快。这主要是因为本研究使用迁移学习训练策略并轻量化了模型结构。

在训练过程中精度方面,TMAInet 整体的稳定性较高,20 轮后模型准确率能够快速稳定在 80% 以上,这表明模型能够在较短的训练时间内达到较高的准确性,并且在训练后期能够维持其性能的稳定性。后续可高效迅速地将模型应用至其他农机具识别任务中。综上所述,TMAInet 模型在准确率、收敛速度和稳定性等方面均优于其他模型,是处理农机具识别任务时的最佳选择。

测试结果如表 3 所示,本文所设计的 TMAInet 模型在准确率和召回率 2 个方面均优于其他模型。虽然在参数量方面 TMAInet 较最轻量级的 MobileNetOne 方法多 2×10^5,但其综合性能均优于该模型。各种流行的轻量化模型均在参数量方面表

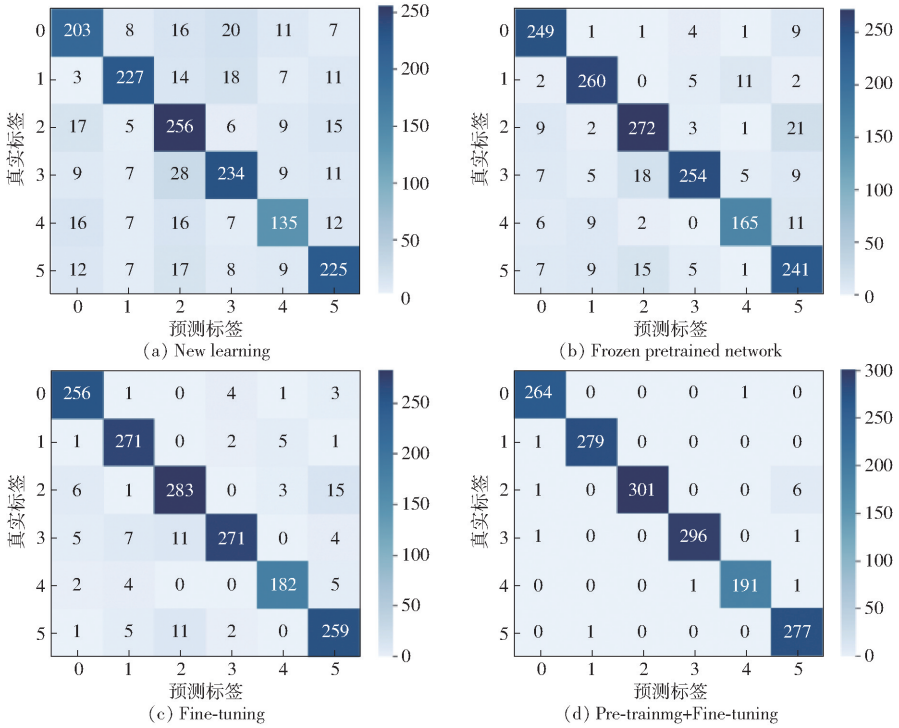


图 6 不同训练方法混淆矩阵

Fig. 6 Training method confusion matrixes

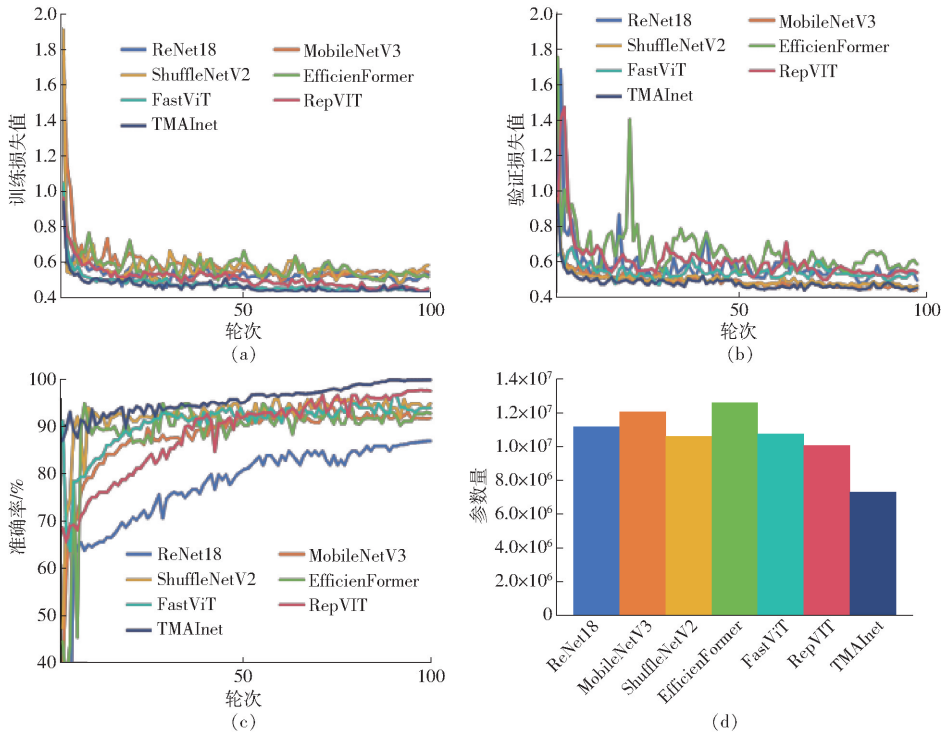


图 7 不同模型训练过程性能

Fig. 7 Visualization of training process

现出优异的效果,尤其是 MobileNetOne 专注于压缩和优化模型,其设计更加强调计算资源的利用和能效,其参数量仅为 7.1×10^6,但其准确率仅为 96.36%,且应用于边缘设备时延时较高。此外,通过实验表明使用 CNN 结构综合性能较优,Hybrid 和 Attention 结构由于 Transformer 结构消耗资源较多,尤其是选用 DW 深度卷积,在保持准确率的同时平衡了模型速度和参数量。

在延时方面 TMAInet 仅 16 ms,这是由于 TMAInet 模型在迁移方法中通过 PF 方法减少了宏观尺度上冗余的特征提取模块。其他算法模型延时均较高,尤其是 ReNet18 和 ShuffleNetV2 模型帧速

表 3　不同算法对比实验结果

Tab. 3　Comparison experiments of different algorithms

模型	类型	参数量	准确率/%	召回率/%	帧速率/(f·s^{-1})
ReNet18	Conv	1.12×10^7	86.76	85.72	15
ShuffleNetV2	Conv	1.07×10^7	94.37	94.27	36
MobileNetV3	Conv	1.21×10^6	95.46	94.21	49
MobileNetOne	Conv	7.10×10^6	96.36	95.75	55
LeViT	Hybrid	1.09×10^7	91.74	91.33	41
FastViT	Hybrid	1.08×10^7	96.35	96.09	49
EfficienFormer	Attention	1.26×10^7	93.64	92.31	47
ReVIT	Conv	1.01×10^7	97.27	96.83	65
TMAInet	Conv	7.30×10^6	99.13	98.78	73

率仅 15 f/s 和 36 f/s，且其准确率仅为 86.76% 和 94.37%。此外，通过 APP 随机选取不同光照、地域、季节等复杂田间环境农机具数据进行全面的测试与应用，以验证其算法的实时性能。

如图 8 所示，本研究直接将输出置信度可视化至图中，多次测试发现原始 RepVIT 能够识别出农机具的作业类型，但其置信度较低。例如，在深松作业场景中，RepVIT 模型置信度仅为 0.37，这表明模型对其判断的信心不足。而通过 TMAInet 模型显著提升了置信度，达到了 0.87。这种提升不仅反映了模型对特定作业类型的更高辨识能力，还大大增强了模型在实际应用中的可靠性。

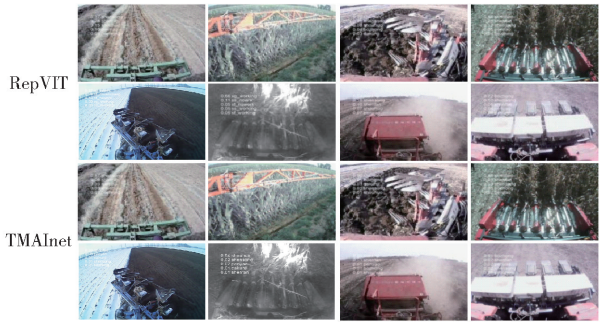


图 8　测试结果对比

Fig. 8　Comparison of test results

在复杂且具有挑战性的农机作业环境中，地域、环境、季节等因素会严重影响模型的识别能力。特别是尘土飞扬、杂物遮挡严重的情况下，打捆作业原有 RepVIT 模型置信度仅为 0.53，这可能导致误识别或不稳定的输出。而在相同环境条件下，本研究所提出的模型表现出较优异的鲁棒性，置信度提升至 0.95。在夜间和冬季雪天等不同场景进行工作时，TMAInet 模型置信度均高于 0.9，证明模型在恶劣条件下高效的适应能力和监测能力。

2.4　消融实验

为验证各改进模块对模型性能的提升效果，探究不同模块对识别准确性的影响，本研究分别对数据增强 DA、ECA 注意力机制和 CFF 模块进行了消融测试。表 4 实验结果表明，在训练前使用数据增强可以有效提升模型，使准确率、F1 分数和召回率分别提升 0.41、1.82、0.64 个百分点。通过数据增强的方法可有效扩增整体数据集，加强模型的泛化能力和鲁棒性。此外，嵌入轻量化注意力机制 ECA，可使模型参数量降低至 7.84×10^6。

表 4　消融实验对比

Tab. 4　Comparison of ablation experiments

模型	准确率/%	F1 分数/%	召回率/%	参数量
RepVIT	97.27	95.49	96.83	1.01×10^7
+ DA	97.68	97.31	97.47	1.01×10^7
+ ECA	98.69	97.90	98.12	7.84×10^6
+ CFF(TMAInet)	99.13	98.53	98.78	7.30×10^6

原始的模型使用 SE 通道注意力机制，其通过压缩和激励的方式来增强神经网络的表达能力。ECA 模型除去了 SE 机制中的冗余全连接层。通过一维卷积对通道权重进行局部交互，在不显著增加参数的情况下实现通道权重学习，并使模型准确率和 F1 分数达到 98.69% 和 97.90%。提升通道注意力权重可以有效地提升 TMAInet 模型的性能，当加入空间注意力机制时模型较原始模型参数量增长且模型性能提升有限。最终，通过嵌入 CFF 模块使模型整体准确率、F1 分数和召回率分别提升 1.86、3.04、1.95 个百分点，参数量降低至 7.3×10^6。本研究所设计的 CFF 模块通过 GLU 结构和 DW 卷积来提升整体结构的特征提取能力，加强模型对细粒度特征的权重，使模型能够更好地识别农机具类型。

2.5　田间应用测试

本研究将 TMAInet 模型优化部署于嵌入式设备 Jetson nano 端，在北京市平谷区(40.196 902° N，117.218 513°E)田间进行测试与应用。对模型进行 ONNX 导出和 TensorRT 加速，再将优化后的轻量化模型部署至 Jetson nano 端。其次，通过摄像头在拖拉机后端实时采集作业视频流，并对农机具状态进行快速识别分类。经由 5G 通信模块将识别结果附带时间、地理位置等信息上传至云端服务器，并在“农业机械化精准作业平台暨希望田野”平台上实时可视化。

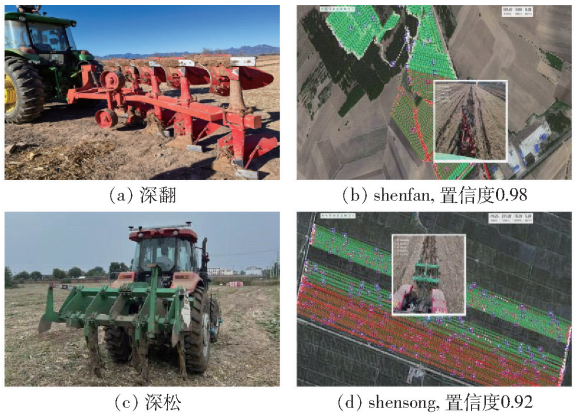
此外还测试了 TMAInet 模型在工作站和 Jetson nano 端的综合性能，具体信息如表 5 所示。由于边缘端性能有限，准度率和置信度均有所下降，延迟为 13.7 ms，其综合性能仍优于其他模型。其次，在实际田间对深翻与深松两种工况进行测试，在连续工作工

程中模型对工况识别置信度均高于0.9,帧速率最高可达到73 f/s,如图9所示。综上所述,所提出的模型展现了优越的适应性,能够应对不同环境中的挑战,确保了农机具识别任务的准确性和稳定性。TMAInet不仅提高了农机具操作的智能化水平,还为智能农业的进一步发展提供了坚实的技术支持。

表5 部署测试对比

Tab.5 Deployment test comparison

应用平台	帧速率/(f·s^{-1})	准确率/%	置信度	延迟/ms
工作站	159	99.84	0.98	6.3
Jetson nano	73	99.07	0.93	13.7



(a) 深翻 (b) shenfan, 置信度0.98

(c) 深松 (d) shensong, 置信度0.92

图9 平台应用界面

Fig.9 App application interface

3 结论

(1)与New learning(NL)、Frozen pretrained network(FPN)和Fine-tuning(FT)相比,本研究所引入的Pre-training+Fine-tuning(PF)融合的迁移学习方法收敛速度快、泛化能力强,准确率高达99.13%,整体模型训练性能最优。

(2)对比ReNet18、EfficienFormer及ReVIT等7种模型的农机具识别任务性能,所提出的TMAInet模型综合性能最佳,在Jetson nano端帧速率最高可达到73 f/s。在田间环境实验中,平均耗时仅13.7 ms。

(3)通过消融实验验证各模块对模型性能的提升效果。通过数据增强、嵌入轻量化注意力机制ECA以及卷积前馈模块CFF,使TMAInet模型准确率、F1分数以及召回率分别提升1.86、3.04、1.95个百分点。模型参数量降低至7.3×10^6。

参 考 文 献

[1] 沈跃,张亚飞,刘慧,等.农业装备自动控制技术研究综述[J].农业机械学报,2023,54(8):1-18.
SHEN Yue, ZHANG Yafei, LIU Hui, et al. Research review of agricultural equipment automatic control technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(8): 1-18. (in Chinese)

[2] SHAHAB H, IQBAL M, SOHAIB A, et al. IoT-based agriculture management techniques for sustainable farming: a comprehensive review[J]. Computers and Electronics in Agriculture, 2024, 220: 108851.

[3] 赵博,张巍朋,苑严伟,等.农业装备运维与作业服务管理信息化技术研究进展[J].农业机械学报,2023,54(12):1-26.
ZHAO Bo, ZHANG Weipeng, YUAN Yanwei, et al. Research progress in information technology for agricultural equipment maintenance and operation service management[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(12): 1-26. (in Chinese)

[4] 刘成良,林洪振,李彦明,等.农业装备智能控制技术研究现状与发展趋势分析[J].农业机械学报,2020,51(1):1-18.
LIU Chengliang, LIN Hongzhen, LI Yanming, et al. Analysis on status and development trend of intelligent control technology for agricultural equipment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(1): 1-18. (in Chinese)

[5] XIANG M, WEI S, ZHANG M, et al. Real-time monitoring system of agricultural machinery operation information based on ARM11 and GNSS[J]. IFAC-PapersOnLine, 2016, 49(16): 121-126.

[6] CHEN S, NOGUCHI N. Remote safety system for a robot tractor using a monocular camera and a YOLO-based method[J]. Computers and Electronics in Agriculture, 2023, 215: 108409.

[7] JAUMANN M, OLCAY E, OKSANEN T. Condition monitoring using convolutional neural network in agricultural machinery-use case: disc mower[J]. IFAC-PapersOnLine, 2022, 55(32): 235-240.

[8] 王宏伟,温昌凯,刘孟楠,等.拖拉机作业工况参数检测系统研究[J].农业机械学报,2023,54(增刊2):409-416.
WANG Hongwei, WEN Changkai, LIU Mengnan, et al. Tractor operating condition parameter testing system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(Supp.2): 409-416. (in Chinese)

[9] BETTUCCI F, SOZZI M, BENETTI M, et al. A data-driven approach to agricultural machinery working states analysis during ploughing operations[J]. Smart Agricultural Technology, 2024, 8: 100511.

[10] 赵建国. 基于TBOX图像采集的农机作业监控系统设计[J]. 南方农机, 2024, 55(12): 55-58.
ZHAO Jianguo. Design of agricultural machinery operation monitoring system based on TBOX image acquisition[J]. Southern Agricultural Machinery, 2024, 55(12): 55-58. (in Chinese)

[11] KUN Y, HUI L, PEI W, et al. Convolutional neural network-based automatic image recognition for agricultural machinery[J]. International Journal of Agricultural and Biological Engineering, 2018, 11(4): 200-206.

(下转第205页)

[21] 刘月峰，刘博，暴祥，等. 基于超轻量化孪生网络的自然场景奶牛单目标跟踪方法[J]. 农业机械学报，2023，54(10)：282－293.
LIU Yuefeng，LIU Bo，BAO Xiang，et al. Single target tracking method for dairy cows in natural scenes[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(10)：282－293.（in Chinese）
[22] SUN D，ZHANG L，WANG J，et al. Efficient and accurate detection of herd pigs based on Ghost－YOLO v7－SIoU[J]. Neural Computing & Applications，2024，36(5)：2339－2352.
[23] GEVORGYAN Z. SloU loss：more powerful learning for bounding box regressiony[J]. arXiv Preprint，arXiv：2205.12740，2022.
[24] JU J Y，CHEN G Q，WANG J F，et al. Design and experiment of an adaptive cruise weeding robot for paddy fields based on improved YOLOv5[J]. Computers and Electronics in Agriculture，2024，219：108824.
[25] 李民赞，任新建，杨玮，等. 基于树莓派的农田表土层土壤容重检测系统研究[J]. 农业机械学报，2021，52(增刊)：329－335，376.
LI Minzan，REN Xinjian，YANG Wei，et al. Soil bulk density detection system of farmland topsoil based on Raspberry Pi[J]. Transactions of the Chinese Society for Agricultural Machinery，2021，52(Supp.)：329－335，376.（in Chinese）
[26] 朱铭敏，张国平，谭建军，等. 基于 YOLOv5s 的轻量级茶叶嫩芽终端检测模型[J]. 浙江农业学报，2024，36(6)：1413－1424.
ZHU Mingmin，ZHANG Guoping，TAN Jianjun，et al. A lightweight tea shoot terminal detection model based on YOLOv5s[J]. Acta Agriculturae Zhejiangensis，2024，36(6)：1413－1424.（in Chinese）
[27] 李海龙. 大田对靶施药除草机器人系统研究[D]. 哈尔滨：东北农业大学，2023.
LI Hailong. Research on field targeted herbicide spraying robot system[D]. Harbin：Northeast Agricultural University，2023.（in Chinese）
[28] 周福君，王文明，李小利，等. 凸轮摇杆式摆动型玉米株间除草装置设计与试验[J]. 农业机械学报，2018，49(1)：77－85.
ZHOU Fujun，WANG Wenming，LI Xiaoli，et al. Design and experiment of cam rocker swing intra-row weeding device for maize[J]. Transactions of the Chinese Society for Agricultural Machinery，2018，49(1)：77－85.（in Chinese）
[29] 权龙哲，王建森，奚德君，等. 靶向灭草机器人药液喷洒空气动力学模型建立与验证[J]. 农业工程学报，2017，33(15)：72－80.
QUAN Longzhe，WANG Jiansen，XI Dejun，et al. Aerodynamics modeling and validation on liquid medicine spraying of target weeding robot[J]. Transactions of the CSAE，2017，33(15)：72－80.（in Chinese）

(上接第 194 页)

[12] ZHANG Z，LIU H，MENG Z，et al. Deep learning-based automatic recognition network of agricultural machinery images[J]. Computers and Electronics in Agriculture，2019，166：104978.
[13] 宋怀波，尚钰莹，何东健. 果实目标深度学习识别技术研究进展[J]. 农业机械学报，2023，54(1)：1－19.
SONG Huaibo，SHANG Yuying，HE Dongjian. Review on deep learning technology for fruit target recognition[J]. Transactions of the Chinese Society for Agricultural Machinery，2023，54(1)：1－19.（in Chinese）
[14] WU Caicong，LI Dong，ZHANG Xiaoqiang，et al. China's agricultural machinery operation big data system[J]. Computers and Electronics in Agriculture，2023，205：107594.
[15] ZHANG X，ZHOU X，LIN M，et al. ShuffleNet：an extremely efficient convolutional neural network for mobile devices[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition，2018：6848－6856.
[16] DOSOVITSKIY A. An image is worth 16x16 words：transformers for image recognition at scale[J]. arXiv preprint arXiv：2010.11929，2020.
[17] HE K，ZHANG X，REN S，et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)，2016：770－778.
[18] VASU P K A，GABRIEL J，ZHU J，et al. FastViT：a fast hybrid vision transformer using structural reparameterization[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision，2023：5785－5795.
[19] GRAHAM B，EL-NOUBY A，TOUVRON H，et al. LeViT：a vision transformer in ConvNet's clothing for faster inference[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV)，2021：12239－12249.
[20] VASU P K A，GABRIEL J，ZHU J，et al. MobileOne：an improved one millisecond mobile backbone[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)，2023：7907－7917.
[21] WANG A，CHEN H，LIN Z，et al. RepViT：revisiting mobile CNN from ViT perspective[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)，2024：15909－15920.