

基于改进MobileNetXt的轻量级手势识别方法

陆洪奕¹, 宋瑾钰²

¹浙江理工大学启新学院, 浙江 杭州

²浙江理工大学计算机科学与技术学院, 浙江 杭州

收稿日期: 2024年4月25日; 录用日期: 2024年5月21日; 发布日期: 2024年5月29日

摘要

轻量级卷积神经网络在复杂背景下的手势图像识别任务中存在网络结构单一、特征提取能力差、识别精度低等问题, 本文针对上述问题提出一种基于改进MobileNetXt的轻量级手势图像识别方法。该方法在Sandglass沙漏块中引入CA注意力机制, 捕捉跨通道的方向和位置感知信息; 选用MetaAconC激活函数自适应地控制神经元的激活程度, 增强模型的特征学习能力和梯度传播效率。该方法在公开的NUS-II手势数据集上识别准确率达99.12%, 比原始的MobileNetXt网络提高了2.25%, 与GhostNet和EfficientNet经典轻量级卷积神经网络相比, 本模型在训练时间、模型参数量和识别准确率等网络指标上均有更优的性能表现。实验结果表明该方法有较强的信息表征和通道提取能力, 在面临多种复杂背景因素干扰的手势识别挑战时仍具有较高的识别准确率。

关键词

轻量级卷积神经网络, 手势识别, CA注意力机制, MetaAconC激活函数

Lightweight Gesture Recognition Method Based on Improved MobileNetXt

Hongyi Lu¹, Jinyu Song²

¹Qixin School, Zhejiang Sci-Tech University, Hangzhou Zhejiang

²School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: Apr. 25th, 2024; accepted: May 21st, 2024; published: May 29th, 2024

Abstract

Lightweight convolutional neural networks face problems such as a single network structure, poor feature extraction ability, and low recognition accuracy in gesture image recognition tasks in

complex backgrounds. This paper proposes a lightweight gesture image recognition method based on improved MobileNetXt to address these issues. This method introduces a CA attention mechanism into the Sandglass hourglass block to capture cross channel direction and position perception information; The MetaAconC activation function is selected to adaptively control the activation level of neurons, enhancing the model's feature learning ability and gradient propagation efficiency. This method has a recognition accuracy of 99.12% on the publicly available NUS-II gesture dataset, which is 2.25% higher than the original MobileNetXt network. Compared with GhostNet and EfficientNet classic lightweight convolutional neural networks, this model performs better in network metrics such as training time, model parameter count, and recognition accuracy. The experimental results show that this method has strong information representation and channel extraction capabilities, and still has high recognition accuracy in the face of gesture recognition challenges caused by various complex background factors.

Keywords

Lightweight Convolutional Neural Network, Gesture Recognition, CA attention Mechanism, MetaAconC Activation Function

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

手势是人类通过手掌和手指的灵活组合来传达特定信息的非语言表达方式,是聋哑人日常交流的重要方式。随着计算机视觉和人工智能技术的不断进步,手势识别方法作为人机交互领域的重要技术广泛应用于医疗保健、虚拟现实、自动驾驶等领域。早期的手势识别主要依赖于特制的物理设备,通过内置的传感器捕捉手指的弯曲程度和空间位置实现对手势的精确识别,而物理设备的佩戴不仅限制了用户的自然运动且成本较高,难以普及;近年来,基于计算机视觉的手势识别技术迅速发展,现有的手势识别任务解决方案主要可分为传统机器学习技术和深度学习神经网络两大类。

传统的手势识别方法以数据为驱动,通过自动化提取信息来构建模型。白煜等[1]采用设定阈值的 RGB 与 YCb Cr 混合颜色空间算法实现人体肤色的稳定分割,排除不同光照环境下亮度信息对彩色图像分割的影响;杜义浩等[2]提出一种基于遗传算法和弹性网络回归的 GA-BLS 方法,较 SVM、KNN、RF、LDA 4 种传统方法准确率高且耗时最短;王宇春等[3]设计了一种基于双权重粒子群算法优化支持向量机的分类模型 DWPSO-SVM,对 NIAPN12 类手势动作的识别准确率高达 93.274%。上述手势图像识别方法都在一定程度上提升了传统方法的性能,但其对于手势的多变性和复杂性处理能力有限,方法的健壮性和非线性表征能力难以满足现实场景的需求。

随着计算能力的提升和大规模数据集的出现,神经网络技术以深层次的网络结构自动学习数据的表征,展现出强大的特征提取与模式识别能力,特别是卷积神经网络(CNN),因其强大的抗干扰能力和可拓展性,在计算机视觉领域逐渐成为主流。王庆鑫[4]通过构建轻量级 SSD 网络结构并融入深度可分离卷积实现 CNN 模型的轻量化,在准确率和模型计算开销之间取得了良好的平衡;霍英等[5]将 YOLOv5 中的主干神经网络替换成轻量级神经网络 MobileNetV3,同时对手势骨骼的 21 个关键点进行特征提取,较原始的 YOLOv5 算法有 6%的提升;赵康等[6]以减少计算量和参数量为出发点,在(2+1)D 卷积结构的

基础上将其中的 3D 卷积替换为 3D 深度可分离卷积, 融合注意力机制弥补时空特征的表征不足, 在保证模型识别精度的前提下降低了计算成本; 王学慧等[7]设计了一种小波变换低频信息与 Xception 网络的静态手势识别方法, 大大减少了训练模型的计算参数并提高了运行速度。上述手势图像识别方法均利用轻量级模型的思想改进网络结构以减少模型的参数量, 但忽略了现有的轻量级网络模型因提取信息不够丰富而普遍存在的准确率不足的问题。

为了增强轻量级网络模型在手势识别任务中的性能, 本文基于 MobileNetXt 架构[8]提出一种新型的轻量级手势识别方法。该方法在 Sandglass 模块的基础上融合 CA 注意力机制, 引入精确的位置信息增强神经网络对输入数据的表征能力, 弥补原 MobileNetXt 网络使用小卷积核带来的通道间关系建模不足的问题; 同时采用 MetaAconC 激活函数, 在保留 ReLu6 函数在正值区域线性特性的基础上, 还在负值区域引入更为平滑的非线性变化, 避免 ReLu6 函数在负值区域置零导致的死区问题, 引入自学习参数加强模型的表征能力和泛化性能。本文通过对 MobileNetXt 网络结构进行优化, 在保证模型轻量级的基础上进一步提升手势图像识别的准确率。

2. 网络结构

2.1. MobileNetXt 网络结构

MobileNetXt 是一种轻量级卷积神经网络模型, 常用于移动设备和嵌入式设备完成有效的图像辨别与分类工作, 其主要的 SandglassBlock 模块即沙漏块结构如图 1 所示。该模块由 2 个 3×3 深度可分离卷积和 2 个 1×1 卷积共 4 个卷积模块组成, 3×3 深度可分离卷积用于空间信息变换, 2 个 1×1 卷积先后用于降维和升维, 使得模型在更高的维度上进行特征映射和空间转换, 缓解 MobileNetV2 网络中倒残差模块导致的梯度跨层传播能力降低、信息丢失和梯度混淆的问题。MobileNetXt 网络结构如表 1 所示, 模型具有 32 个卷积层, 每个卷积层后都有一个批量归一化层, 网络中使用的扩展率默认设置为 6。MobileNetXt 网络的轻量化设计使其网络结构相对简单, 捕捉图像的特征表征能力不足, 对于复杂场景下的物体识别能力有限。

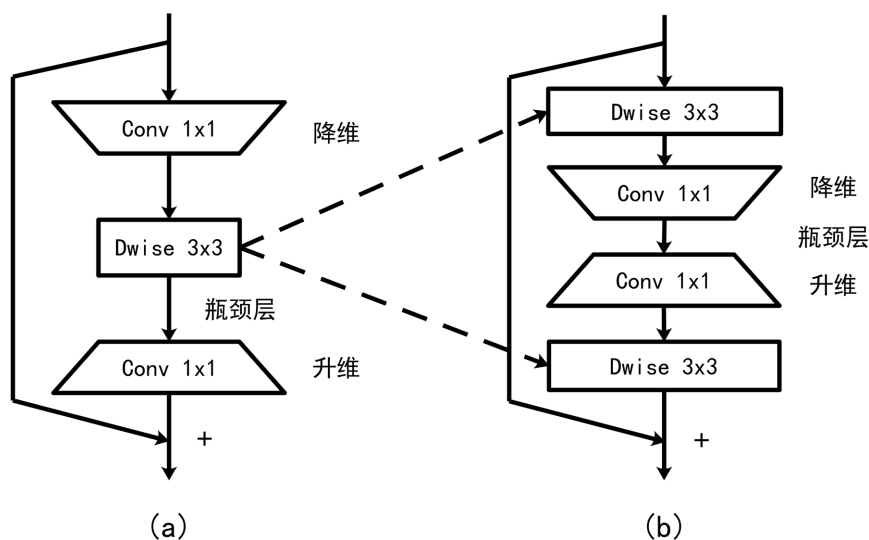


Figure 1. Different types of residual blocks. (a) The classic bottleneck structure of deep space convolution; (b) Sandglass module with bottleneck structure

图 1. 不同类型的残差块。(a) 深度空间卷积的经典瓶颈结构; (b) 具有瓶颈结构的 Sandglass 模块

Table 1. MobileNetXt network architecture**表 1.** MobileNetXt 网络结构

扩展率	输出张量尺寸	步长	模块数量	输入张量尺寸	模块名称
-	$112 \times 112 \times 32$	2	1	$224 \times 224 \times 3$	Conv2d 3×3
2	$56 \times 56 \times 96$	2	1	$112 \times 112 \times 32$	Sandglass Block
6	$56 \times 56 \times 144$	1	1	$56 \times 56 \times 96$	Sandglass Block
6	$28 \times 28 \times 192$	2	3	$56 \times 56 \times 144$	Sandglass Block
6	$14 \times 14 \times 288$	2	3	$28 \times 28 \times 192$	Sandglass Block
6	$14 \times 14 \times 384$	1	4	$14 \times 14 \times 288$	Sandglass Block
6	$7 \times 7 \times 576$	2	4	$14 \times 14 \times 384$	Sandglass Block
6	$7 \times 7 \times 960$	1	2	$7 \times 7 \times 576$	Sandglass Block
6	$7 \times 7 \times 1280$	1	1	$7 \times 7 \times 960$	Sandglass Block
-	$1 \times 1 \times 1280$	-	1	$7 \times 7 \times 1280$	Avgpool 7×7
-	k	-	1	$1 \times 1 \times 1280$	Conv2d 1×1

2.2. CA-Sandglass 模块

注意力机制的合理使用对于提高 CNN 的网络性能具有显著作用。CA 模块[9]是注意力机制的一种,相较于常用的 SE 注意力机制[10],它不仅捕获跨通道的信息,还能捕获方向和位置感知的信息,灵活且轻量。CA 模块可分为 2 个部分,分别为坐标信息嵌入模块和坐标注意力生成模块。

CA 注意力机制通过精确的位置信息对通道关系和长程依赖进行编码,采用坐标信息嵌入模块将全局池化操作分解为一系列一维特征编码。对于输入 X ,使用尺寸 $(H, 1)$ 和 $(1, W)$ 的池化核在宽度和高度方向分别进行特征映射,得到高度 h 和宽度 w 的第 c 个通道的全局特征如下:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} X_c(h, i), z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} X_c(j, w) \quad (1)$$

级联坐标注意力生成模块,在得到的两个特征图 $z_c^h(h)$ 和 $z_c^w(w)$ 上进行 F1 变换,输出 $f = \delta(F_1([z_c^h(h), z_c^w(w)]))$ ($f \in R^{C/r \times (H+W)}$, r 表示下采样比例),即空间信息在水平方向和垂直方向的中间特征图。沿着空间维度将 f 切分为两个单独的张量 $f^h \in R^{C/r \times H}$ 和 $f^w \in R^{C/r \times W}$ 分离高度和宽度信息,利用卷积 F_h 和 F_w 将特征图 f^h 和 f^w 变换到和输入 X 相同的通道数,输出如下:

$$g^h = \sigma(F_h(f^h)), g^w = \sigma(F_w(f^w)) \quad (2)$$

最后将 g^h 和 g^w 作为注意力机制的权重进行拓展,得到最终的 CA 注意力机制模块如下:

$$y_c(i, j) = X_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (3)$$

改进网络中 CA-Sandglass 模块的模型结构如图 2 所示。在 SandglassBlock 模块最后的 3×3 深度可分离卷积后,输入尺寸为 $C_l \times H \times W$ 的特征张量。为了保留原始特征信息并引入新的特征变换,使用恒等映射将原始特征张量直接传递到后续层,筛选学习原始特征中的有效信息,使用空间变换在空间上对特征图重新排列,捕捉特征间的关联性和上下文信息,得到张量大小为 $C_o \times H \times W$ 。再基于残差连接分别

从宽度和高度方向进行全局平均池化, 将得到的两个尺寸为 $C_o \times 1 \times 1$ 的特征图使用 Concat 连接拼接在一起实现特征融合, 使用共享的 1×1 卷积融合通道信息并降维。使用批量归一化缓解模型训练过程中的内部协变量偏移问题, 并融合 ReLu 激活函数增加模型的非线性。最后, 为了在保留原始特征空间结构信息的基础上调整通道数, 将特征图分离成宽度和高度方向的特征后分别融合 1×1 卷积和 Sigmoid 激活函数, 使用空间注意力机制扩展宽高特征使张量输出尺寸恢复为 $C_o \times H \times W$ 。

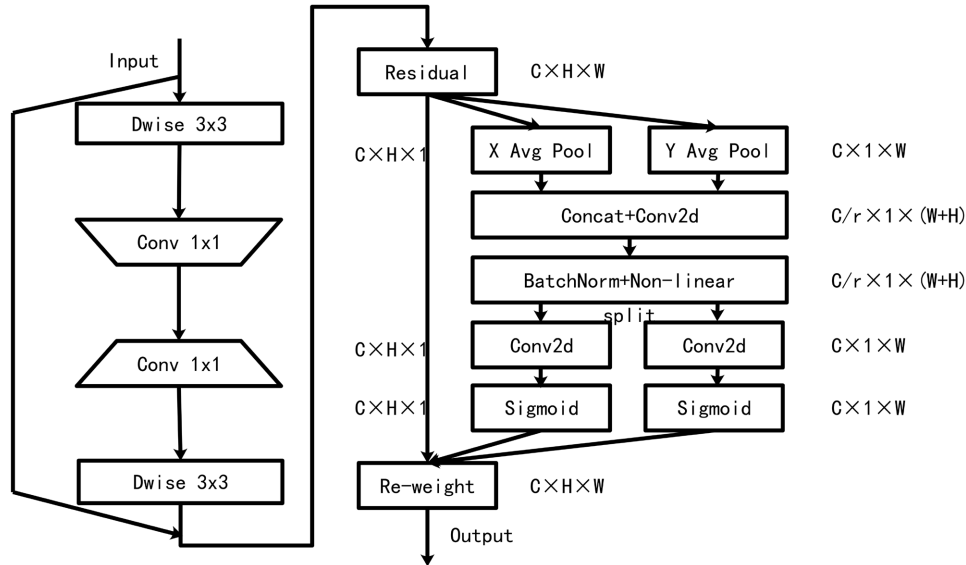


Figure 2. CA-Sandglass module structure
图 2. CA-Sandglass 模块结构

2.3. MetaAconC 激活函数

ReLU 激活函数[11]在很长一段时间都是最佳的神经网络激活函数, 主要是源于其非饱和、稀疏性等优秀的特性, 但是它也同样会产生神经元坏死的严重后果。Ningning Ma 等[12]从 Swish 激活函数和 ReLU 激活函数的公式出发, 挖掘其中的平滑近似原理, 并应用到 Maxout 系列激活函数上, 提出了一种新型激活函数即 Acon 激活函数, 网络结构如图 3 所示。

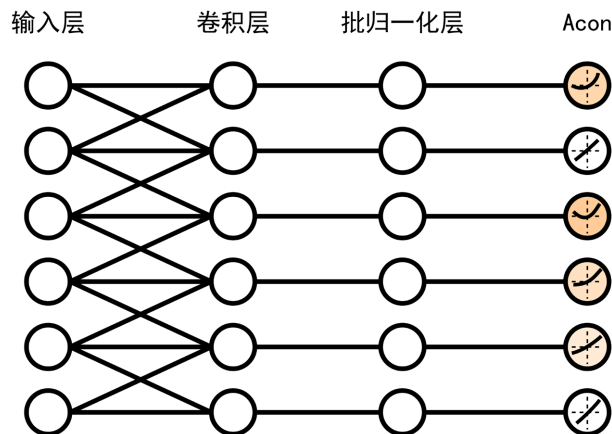


Figure 3. Acon module network structure
图 3. Acon 模块网络结构

Acon 函数之所以能轻易地在激活与非激活的状态间转化, 是因为其切换因子能适应控制函数为线性或非线性。因此, 相较于传统激活函数 ReLu, Acon 允许每个神经元的激活特性适应变化, 这种特定的激活行为有助于提高泛化和转移性能。在此基础上引入一个额外的可学习参数 $\beta_c = \sigma W_1 W_2 \sum_{h=1}^H \sum_{w=1}^W x_{c,h,w}$, 结合一个小型卷积网络和 Sigmoid 函数来动态控制激活函数的线性与非线性, 实现在训练过程中动态地调整函数形状以提高模型的性能。

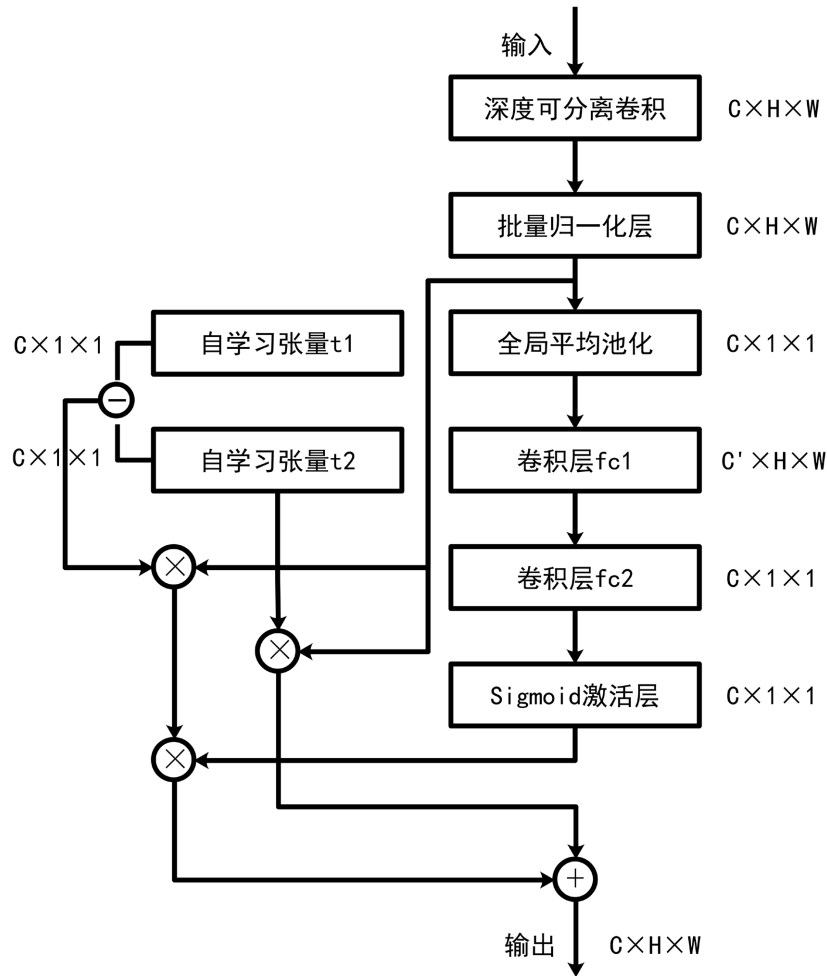


Figure 4. MetaAconC activation function fusion position and structure
图 4. MetaAconC 激活函数融合位置与结构

改进网络中 MetaAconC 激活函数的融合位置与结构如图 4 所示。融合位置位于深度可分离卷积的批量归一化层后。输入尺寸为 $C \times H \times W$ 的张量, 经过全局平均池化沿高度和宽度聚合信息生成一个尺寸为 $C \times 1 \times 1$ 的通道级别统计量, 将生成的统计量通过两个卷积层(fc1 和 fc2)进行特征变换, 进一步提取通道间的依赖关系并实现信息的非线性映射, 使用 sigmoid 激活函数生成每个通道的权重 beta, 反映不同通道在特征表示中的重要性。同时, 为了增加模型的灵活性和表达能力, 引入大小为 $C \times H \times W$ 的模块计算 dpx 作为输入特征图的线性变换, 其中变换依赖于两个尺寸为 $C \times 1 \times 1$ 的自学习参数 t1 和 t2。最终, 利用空间信息恢复特征图的尺寸为 $C \times H \times W$ 。每个通道的权重由 beta 和 dpx 共同决定, 使得网络能够根据输入数据的特性进行自适应调整, 大大提高模型的灵活性和表达能力。

3. 实验方法与结果分析

3.1. 数据集与实验环境

实验采用由新加坡国立大学(National University of Singapore)开源的 NUS-II 手势数据集[13]测试本文提出的方法。该数据集包含 2000 张 RGB 三通道图像, 涵盖字母 A~J 共 10 种不同手势类别, 每一类别均包含 200 张图像, 这些图像由 50 名年龄介于 22~56 岁、来自不同种族的男性和女性在多样化的复杂背景下拍摄完成, 部分图像示例如图 5 所示。在数据集预处理阶段, 以 7:2:1 的比例将数据集随机分割为训练集、测试集和验证集, 同时为了提升模型的泛化能力, 使用随机裁剪、旋转、缩放及平移等实时数据增强技术, 将训练样本的数量扩展至 8000 张。



Figure 5. NUS-II gesture dataset example image

图 5. NUS-II 手势数据集示例图像

实验环境为 Ubuntu20.04, Python 3.8 版本, 使用 PyTorch 1.10.0 作为深度学习框架, 在 120GB 内存, RTX 4090 24GB 显卡的配置上进行训练。实验的初始学习率设置为 0.001, 选用 AdamW 优化器调整模型参数, 迭代周期 Epoch 为 50, 批处理大小 Batch Size 为 16。在训练之前先将原始的手势图像进行尺寸归一化处理, 转换为 $244 \times 244 \times 3$ 的三通道 RGB 图像, 再将 RGB 图像的每个通道像素值映射为浮点数进行标准化处理, 提高模型对不同光照环境下图像变化的适应性和泛化能力。使用回调函数在训练过程中动态调整学习率, 加速模型的收敛速度。将验证集的准确率作为回调函数的监测指标, 若连续 3 个迭代周期内验证集准确率没有显著提升, 则自动触发学习率调整机制将学习率减半。

3.2. 实验结果分析

3.2.1. 消融实验

为了评估改进后 MobileNetXt 网络模块的性能表现, 在改进的网络架构内, 对 CA-Sandglass 模块和 MetaAconC 激活函数进行消融对比实验, 验证各模块对网络性能的贡献和影响, 结果如表 2 所示。通过对比发现, 当所有模块融合时, 本文网络相较于原 MobileNetXt 网络计算参数量有所增长, 但识别准确率提升了 2.25%, 对于每一个单独模块的改进均使准确率有 1.06%~1.78% 的提升, 且参数量增长幅度较小。

Table 2. Results of ablation comparison experiment
表 2. 消融比较实验结果

CA-Sandglass 模块	MetaAconC 激活函数	参数量	准确率%
		1.94E+06	96.87
√		2.81E+06	97.93
	√	3.03E+06	98.65
√	√	3.89E+06	99.12

3.2.2. 激活函数对比实验

为了全面评估 MetaAconC 激活函数的性能优势, 本文使用 5 种不同的激活函数在改进后的 MobileNetXt 模型上进行实验比对。每种激活函数均进行 10 次独立的测试确保实验结果的可信度, 最终求得该模型下各激活函数的平均准确率如表 3 所示。

Table 3. The average accuracy of models under different activation functions
表 3. 不同激活函数下模型的平均准确率

激活函数	平均准确率/%
Sigmoid	90.52
ReLu6	96.77
Swish	98.61
SMU	98.91
MetaAconC	99.02

可以看出, 使用 MetaAconC 激活函数的模型平均准确率为 99.02%, 与 Sigmoid 函数和 ReLu6 函数相比分别有 8.5%和 2.25%的提升。Sigmoid 函数中梯度消失问题影响了模型的收敛速度和效果, 因函数的输出不以 0 为中心, 导致模型在训练过程中的权重更新不稳定; 相比之下, ReLu6 函数在一定程度上缓解了梯度消失问题, 在输入值大于 0 时保持了恒定的梯度, 但函数在输入值小于 0 时完全无激活, 容易导致部分神经元在训练过程中坏死, 不再对模型的输出产生影响; 而 MetaAconC 激活函数有效避免了梯度消失和神经元坏死的问题, 在模型训练过程中表现出较好的性能。同时, MetaAconC 激活函数的模型平均准确率与 Swish 函数和 SMU 函数相比分别有 0.41%和 0.11%的提升。Swish 函数是一种自门控激活函数, 结合了 Sigmoid 函数的门控特性和线性函数的优点, 且在不同输入区间内具有不同的激活强度, 使得函数在模型训练过程中能较好地适应数据的分布; SMU 函数是一种平滑的、可微的、有界的激活函数, 通过平滑的过渡来避免 ReLu 函数在 0 点处的不连续性; 尽管 Swish 和 SMU 函数都在一定程度上改进了传统激活函数的缺点, 但 MetaAconC 函数可以根据网络参数自适应地调整激活函数的形状, 进一步提升了模型的表达能力, 在深度学习任务中表现更稳定。

3.2.3. 改进网络与其它网络比较

为了验证本文网络的可行性, 在 NUS-II 手势数据集上与原始 MobileNetXt、ShuffleNet、GhostNet、EfficientNet 和 SqueezeNet 5 个经典轻量级卷积神经网络进行比较, 综合考虑网络的训练时间、模型大小、模型参数量和准确率等指标对模型性能进行评判, 实验结果如表 4 所示。

Table 4. Comparison of network model performance
表 4. 网络模型性能比较

网络模型	训练时间/s	模型大小/MB	模型参数量	准确率/%
MobileNetXt	3160.07	7.76	1.94E+06	96.87
ShuffleNet	3556.46	9.78	2.56E+06	94.81
GhostNet	3204.61	15.29	4.01E+06	97.12
EfficientNet	3668.47	15.37	4.03E+06	98.97
SqueezeNet	3519.23	2.83	7.41E+05	95.63
本文网络	3116.99	14.83	3.89E+06	99.12

可以看出, 本文网络的训练时间最短, 在训练过程中具有更高的计算效率, 能更快地收敛到理想的性能水平。SqueezeNet 使用“Squeeze”层压缩特征图的通道数, 使用“Expand”层增加通道数量产生表征特征, 多次使用 1×1 卷积核降低模型的参数数量, 极大程度上缩小了模型的大小, 但其网络深度较浅, 在训练过程中容易受到梯度消失或爆炸的影响, 在处理复杂任务、学习深层特征方面表现较差。ShuffleNet 通过引入通道混洗操作来增强模型的信息流动能力, 当使用宽度因子加宽网络时, 计算分支容易产生通道冗余, 相较之下本文网络具有更好的表征能力, 能够捕捉精确的位置信息和通道关系编码到神经网络中。尽管 SqueezeNet 和 ShuffleNet 在模型参数量上占有一定优势, 但网络的识别准确率仅有 95.63% 和 94.81%, 均远低于本文网络, 在一些对分类精度要求较高的任务中难以满足预期标准。GhostNet 通过简单的线性变换生成冗余的特征图, 以较小的计算代价获得更丰富的特征信息, EfficientNet 通过复合缩放方法和 Compound Scaling 技术, 在网络的深度、宽度和分辨率三个维度上进行联合优化, 确保网络在扩大规模时能够保持各维度之间的协调和平衡, 实现有限计算资源下的性能提升, 但从训练时间、模型大小、计算参数量和准确率等指标来看, 上述网络的模型性能均逊于本文网络。

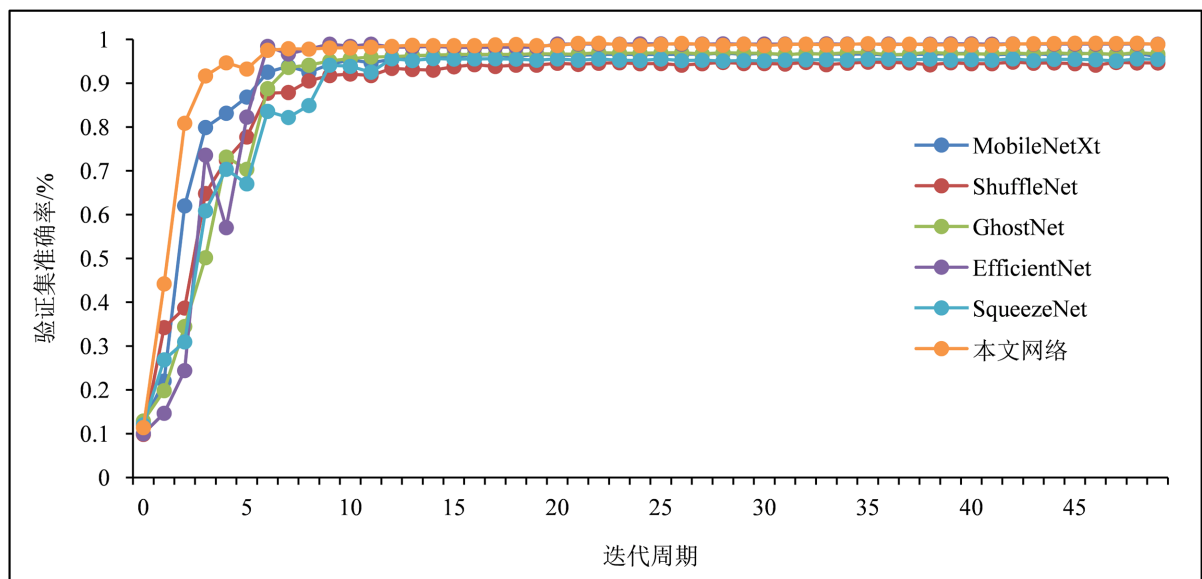


Figure 6. Network model accuracy curve
图 6. 网络模型准确率曲线

各网络模型在训练过程中的准确率变化曲线如图 6 所示。显而易见, 虽然本文网络的计算参数量较

原 MobileNetXt、ShuffleNet、SqueezeNet 网络模型大, 但模型训练收敛速度最快, 约在第 8 个 Epochs 后验证集准确率曲线趋于稳定。EfficientNet、SqueezeNet 和 GhostNet 网络在收敛前数据波动较大, 但所有模型均在 13 个 Epochs 后趋于稳定。

4. 结论

本文提出了一种基于改进 MobileNetXt 网络的手势识别方法, 通过 CA 注意力机制改进 Sandglass 模块, 在保留小卷积核减少模型参数的基础上强化了特征通道间的信息交互; 使用 MetaAconC 激活函数加强模型的特征学习能力和梯度传播效率。本文方法在 NUS-II 手势数据集上识别率达 99.12%, 训练时间为 3117 s, 模型大小为 14.83 MB, 模型参数量为 $3.89\text{E}+06$; 消融实验证明了 CA-Sandglass 模块和 MetaAconC 激活函数对模型的识别准确率均有提升; 激活函数对比实验证明了 MetaAconC 函数通过自学习参数更能提升模型的表征能力, 在深度学习任务中表现更稳定; 通过与其它轻量级网络进行比较, 证明本文方法在保持模型轻量化的基础上准确率更高, 在训练过程中收敛速度最快, 具有较好的鲁棒性和泛化能力。本文网络是基于 MobileNetXt 的轻量级小型网络, 后续将在保证识别准确率的同时通过模型剪枝和知识蒸馏等方法进一步压缩模型参数量, 提升模型的可拓展性和可部署性, 并尝试将其应用于移动端解决实际问题。

基金项目

浙江省一流课程建设项目。

参考文献

- [1] 白煜, 李香萍, 张雨菡. 基于机器视觉的静态手势识别实验设计[J]. 实验技术与管理, 2023, 40(6): 187-191+198.
- [2] 杜义浩, 曹添福, 范强, 王孝冉. 基于 GA-BLS 方法的手势识别研究[J]. 计量学报, 2024, 45(1): 121-127.
- [3] 王宇春, 王敏, 袁东华, 赵化启. 基于 DWPSO-SVM 的 sEMG 手势动作识别[J]. 智能计算机与应用, 2023, 13(12): 158-164.
- [4] 王庆鑫. 基于轻量级 CNN 的手势识别算法研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨理工大学, 2023.
- [5] 霍英, 张嘉明, 陈台兴, 林鸿森, 彭季雨. 基于改进 YOLOv5+Mobile Net 的动态手势识别方法[J]. 软件, 2023, 44(11): 47-52.
- [6] 赵康, 黎向锋, 李高扬, 左敦稳. 轻量级(2+1)D 卷积结构的动态手势识别研究[J]. 微电子学与计算机, 2022, 39(9): 46-54.
- [7] 王学慧, 赵跃鹏, 王嘉炜, 李振, 田秋红. 小波变换低频信息与 Xception 网络的静态手势识别[J]. 软件导刊, 2021, 20(8): 12-19.
- [8] Zhou, D., Hou, Q., Chen, Y., *et al.* (2020) Rethinking Bottleneck Structure for Efficient Mobile Network Design. *Computer Vision—ECCV 2020: 16th European Conference*, Glasgow, 23-28 August 2020, 680-697. https://doi.org/10.1007/978-3-030-58580-8_40
- [9] Hou, Q., Zhou, D., Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 13713-13722. <https://doi.org/10.1109/CVPR46437.2021.01350>
- [10] Hu, J., Shen, L., Sun, G. (2018) Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [11] Agarap, A.F. (2018) Deep Learning Using Rectified Linear Units (Relu). arXiv preprint arXiv:1803.08375.
- [12] Ma, N., Zhang, X., Liu, M., *et al.* (2021) Activate or Not: Learning Customized Activation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 8032-8042. <https://doi.org/10.1109/CVPR46437.2021.00794>
- [13] Pisharady, P.K., Vadakkepat, P. and Poh, L.A. (2017).The NUS Hand Posture Datasets II. <https://doi.org/10.25540/AWJS-GMJB>