

基于轻量级网络的钢铁表面缺陷分类

史杨潇^{1*}, 章 军¹, 陈 鹏¹, 王 兵²

(1. 安徽大学 电气工程与自动化学院, 合肥 230601; 2. 安徽工业大学 电气信息学院, 安徽 马鞍山 243002)

(* 通信作者电子邮箱 1142752120@qq.com)

摘 要: 缺陷分类是钢铁表面缺陷检测的重要内容。在卷积神经网络(CNN)取得良好效果的同时, 网络日益增长的参数量耗费了大量计算成本, 为缺陷分类任务在个人计算机或低算力设备上的部署带来了巨大的挑战。针对上述问题, 提出了一种新颖的轻量级网络模型 Mix-Fusion。首先, 通过组卷积和通道洗牌两种操作, 在保持精度的同时有效降低计算成本; 其次, 利用一个狭窄的特征映射对组间信息进行融合编码, 并将生成的特征与原始网络结合, 从而有效解决了“稀疏连接”卷积阻碍组间信息交换的问题; 最后, 用一种新型的混合卷积(MixConv)替代了传统的深度卷积(DWConv), 以进一步提高模型的性能。在 NEU-CLS 数据集上的实验结果表明, Mix-Fusion 网络在缺陷分类任务中的浮点运算次数和分类准确率分别为 43.4 MFLOPs 和 98.61%。相较于 ShuffleNetV2 和 MobileNetV2 网络, Mix-Fusion 网络不仅降低了模型参数, 压缩了模型大小, 同时还得到了更好的分类精度。

关键词: 表面缺陷检测; 缺陷分类; 模型加速; 深度学习; 轻量级网络

中图分类号: TP391.4 **文献标志码:** A

Classification of steel surface defects based on lightweight network

SHI Yangxiao^{1*}, ZHANG Jun¹, CHEN Peng¹, WANG Bing²

(1. School of Electrical Engineering and Automation, Anhui University, Hefei Anhui 230601, China;

2. School of Electrical and Information, Anhui University of Technology, Ma'anshan Anhui 243002, China)

Abstract: Defect classification is an important part of steel surface defect detection. When the Convolutional Neural Network (CNN) has achieved good results, the increasing number of network parameters consumes a lot of computing cost, which brings great challenges to the deployment of defect classification tasks on personal computers or low computing power devices. Focusing on the above problem, a novel lightweight network model named Mix-Fusion was proposed. Firstly, two operations of group convolution and channel-shuffle were used to reduce the computational cost while maintaining the accuracy. Secondly, a narrow feature mapping was used to fuse and encode the information between the groups, and the generated features were combined with the original network, so as to effectively solve the problem that “sparse connection” convolution hindered the information exchange between the groups. Finally, a new type of Mixed depthwise Convolution (MixConv) was used to replace the traditional DepthWise Convolution (DWConv) to further improve the performance of the model. Experimental results on NEU-CLS dataset show that, the number of floating-point operations and classification accuracy of Mix-Fusion network in defect classification task is 43.4 Million Floating-point Operations Per second (MFLOPs) and 98.61% respectively. Compared to the networks of ShuffleNetV2 and MobileNetV2, the proposed Mix-Fusion network reduces the model parameters and compresses the model size effectively, as well as obtains the better classification accuracy.

Key words: surface defect detection; defect classification; model acceleration; deep learning; lightweight network

0 引言

钢铁表面缺陷分类是工业缺陷检测的关键环节。然而, 在传统工业中, 这个环节往往是手动执行的。为了取代手工操作, 人们希望机器能够利用计算机视觉技术自动检测钢铁表面缺陷^[1]。

由于钢铁表面缺陷图像受到光照和材质变化的影响, 并且钢铁表面类内缺陷在外观上存在较大差异, 类间缺陷又有相似的方面^[2], 利用计算机视觉技术进行缺陷分类, 仍然是一

个巨大的挑战。目前的图像分类方法主要为两类: 传统机器学习图像分类算法和基于卷积神经网络(Convolutional Neural Network, CNN)^[3]的深度学习。传统图像分类算法主要采用特征提取和分类器设计两大步骤来实现, 如 K 近邻(K -Nearest Neighbor, KNN)算法^[4]、支持向量机(Support Vector Machine, SVM)^[5]以及神经网络^[6]等。在实际缺陷分类应用中会面临各种复杂的情况, 使用传统的图像处理方法在准确率上很难达到要求。

近年来, 基于深度学习的图像分类方法取得了很好的效

收稿日期: 2020-08-18; 修回日期: 2020-10-27; 录用日期: 2020-11-13。 基金项目: 国家自然科学基金资助项目(61872004)。

作者简介: 史杨潇(1997—), 男, 安徽合肥人, 硕士研究生, 主要研究方向: 深度学习、缺陷检测; 章军(1971—), 男, 安徽合肥人, 教授, 博士, 主要研究方向: 深度学习在视频分析中的应用、化学信息学; 陈鹏(1978—), 男, 安徽肥东人, 教授, 博士, 主要研究方向: 机器学习、生物信息学; 王兵(1976—), 男, 安徽怀宁人, 副教授, 博士, 主要研究方向: 模式识别、生物信息学。

果,例如 VGGNet (Visual Geometry Group Network)^[7]、ResNet (Residual Network)^[8]等。然而,最先进的 CNN 需要数十亿次浮点运算,这使得它们无法用于移动或嵌入式设备。例如,ResNet-101 的复杂度为 7.8×10^9 FLOPs (Floating-point Operations Per second),即使使用强大的 GPU 也无法实现实时检测。考虑到现代 CNN 的巨大计算成本,轻量级神经网络被提出部署在移动或嵌入式设备上。例如:MobileNetV1^[9]和 MobileNetV2^[10]采用深度可分离卷积来构建轻量级网络;ShuffleNet^[11]采用分组卷积和深度可分离卷积来构建轻量级网络;SqueezeNet^[12]利用核心模块 Fire 压缩模型参数,减小网络的深度,降低模型的大小;SENet (Squeeze-and-Excitation Network)^[13]提出的 SE 模块是一种轻量级注意力机制,通过学习通道重要性的方式自适应校准特征图,然而 SE 模块只关注了特征图通道方面的影响而忽略了空间维度的重要性。同时,目前已有研究者进行轻量级网络应用在缺陷和分类方面的研究,如:姚海明等^[14]提出了一种用于实时检测瓷片表面缺陷的 MagnetNet;张琪等^[15]提出了一种改进的用于对肝部病理组织进行分类的 MobileNet。

轻量级网络可以在有限的计算预算下获得相对较高的精度。然而现有轻量级网络倾向于使用“稀疏连接”卷积,例如深度卷积和群卷积,而不是标准的“完全连接”卷积。这种“稀疏连接”卷积在降低参数数量的同时,一定程度上会阻碍组间信息的交换,导致网络性能下降。而实用的钢铁缺陷分类算法需要部署在 CPU 甚至嵌入式系统上,因此需要采用一种可以避免组间信息丢失的具有较低计算复杂度,同时具有较高分类准确率的算法。

本文提出了一种新颖的 Mix-Fusion 网络模型,以 ShuffleNet 的通道洗牌单元和 MENet (Merging-Evolution Network) 的融合编码模块为核心,构建出具有三个分支的 MF (Mix-Fusion) 模块。该模块通过标准分支保留原有特征;通过降参分支降低计算成本的同时优化了模型精度;通过融合分支避免组间信息的丢失。同时,将该轻量级网络与混合卷积模块融合,提高了网络对于不同分辨率模式的捕获能力,获得了更好的模型精度及效率。在 NEU-CLS 数据集上进行实验验证,通过和其他方法的比较结果可以得出, Mix-Fusion 网络模型避免了组间信息的丢失,进一步降低了参数数量和计算量,显著提升了分类精度。

1 Mix-Fusion 网络的构建

由于深层神经网络计算量大、模型容量大,神经网络的压缩和加速问题已成为深度学习领域的研究热点。在嵌入式设备上运行高质量深层神经网络的需求不断增加,更是鼓励了对轻量级网络模型设计的研究。这些网络倾向于利用“稀疏连接”卷积,在减小计算成本的同时也会阻碍组间的信息交换。本文借鉴了 ShuffleNet 和 MENet^[16]的思想,并加入了 MixConv^[17]卷积模块,提出了一种 Mix-Fusion 网络。该网络在分类精度提高的同时,计算成本也有所下降。

1.1 通道洗牌网络单元

通道洗牌网络单元 (Channel-shuffle) 是 ShuffleNet 网络的核心,目的在于解决组卷积阻碍组间信息交换、导致性能下降的问题。如图 1 所示,将 3 组原始通道每组再次平均分为 3 组,用①~⑨表示,通道洗牌操作将 9 组通道打乱重置,使得第二卷积层中的每个组包含来自第一卷积层中每个组的通道,

在一定程度上实现了组间信息的交换。

然而,当每组通道数为 3 时,通道洗牌无法完全避免组间信息的丢失,第二卷积层中的每个组仅从第一卷积层的每个组接收一个通道,导致每个组中其他两个信道被忽略。因此,大部分组间信息无法利用。这个问题在更多的信道组中会更加严重。随着组数的增加,每组通道数增加,然而第二卷积层接收的通道数仍然保持为 1 个,同时每组忽略的通道数量也增加,造成组间信息丢失严重,网络性能大幅下降。

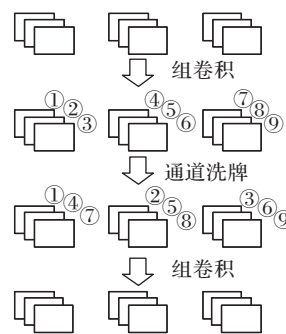


图1 通道洗牌网络单元示意图

Fig. 1 Schematic diagram of channel shuffle network unit

1.2 融合编码模块

为了解决组间信息丢失的问题,本文借鉴了 MENet 中合并和进化的思想。如图 2 所示,利用一个狭窄的特征映射对组间通道信息进行融合编码,并对其进行匹配变换后与原始网络相结合以获得更具区分性的特征。操作如下:

1) 通道融合。通道融合的目的是将所有通道特征聚合,并对组间信息编码,形成一个狭窄的特征映射。在组卷积生成的原始特征图 $F \in R^{C \times H \times W}$ 基础之上,网络对其进行融合编码变换 $T_F: R^{C \times H \times W} \rightarrow R^{C_m \times H \times W}$,达到对所有通道的特征进行聚合的目的。其中: C 为原始特征图的通道数; H 和 W 为原始特征图的宽和高; C_m 为融合特征图的通道数。由于 C 比较大,在不影响计算成本的前提下,很难对空间信息进行集成。因此本文首先利用 1×1 单点卷积完成融合编码变换,将同一空间位置上所有通道的特征聚合起来,同时降低通道数量并进行批处理规范化^[18]和 ReLU (Rectified Linear Unit) 激活。

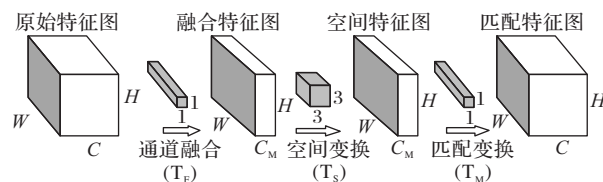


图2 融合编码模块示意图

Fig. 2 Schematic diagram of fusion coding module

2) 空间变换。由于计算成本的限制,通道融合操作未对空间信息进行集成,因此引入一个标准的 3×3 卷积核进行空间变换 $T_S: R^{C_m \times H \times W} \rightarrow R^{C_m \times H \times W}$ 。空间变换操作能够在不改变通道数的情况下提取更多的空间信息,之后进行批处理规范化和 ReLU 激活。

3) 匹配变换。为了将处理后的特征图与原始网络相结合以获得更具区分性的特征,网络对空间特征图进行匹配变换 $T_M: R^{C_m \times H \times W} \rightarrow R^{C \times H \times W}$,之后进行批处理规范化和 Sigmoid 激活,得到与原始特征图一样维度的匹配特征图。最终,将匹

配特征图作为神经元尺度因子,与原始网络以元素乘积的方式相结合,进一步提高特征在网络中的表达能力。在通道融合过程中,1×1单点卷积操作对于每一个通道的信息都进行了编码,因此在最终的匹配特征图中,变换后的通道都包含了来自原始特征图每一个通道的信息,这避免了卷积过程中组间信息的丢失。

1.3 混合卷积模块

深度卷积(DepthWise Convolution, DWConv)在现代轻量级网络中越来越流行,常用的深度卷积将每个通道单独分为一组进行组卷积,从而极大降低了参数量和计算成本。然而传统做法都是简单地使用3×3卷积核^[19],忽视了卷积核的大小。本文借鉴 MixConv 多核结合的思想,使用不同大小组合的卷积核替代深度卷积,大卷积核能够在一定范围内提高模型精度,多卷积核则能提高模型在不同分辨率下的适应度。

如图3所示,不同于将单个内核应用于所有通道的深度卷积,混合卷积将通道平均划分为若干组,并对每个组应用不同大小的内核。网络既需要大卷积核来捕捉高分辨率模式,也需要小卷积核来捕捉低分辨率模式,以获得更好的模型精度和效率。相较于多分支网络集中改变神经网络的宏观结构以利用不同的卷积运算,如 Inception^[20]和 NASNet^[21],混合卷积在不改变网络结构的情况下,能够替换不同大小组合的卷积核以测试模型性能。

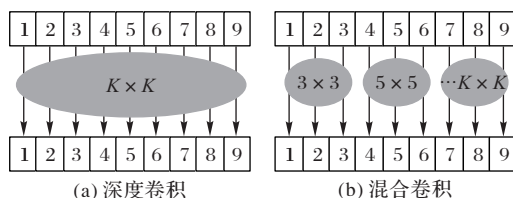


图3 深度卷积和混合卷积原理

Fig. 3 Principles of deep convolution and mixed convolution

1.4 Mix-Fusion 模块

基于上述单元,本文提出了 MF 模块。MF 模块由图4(a)从左至右三个分支组成:标准分支、降参分支和融合分支。标准分支是对原始特征图的直接映射。降参分支采用“稀疏连接”卷积,它由三层组成:第一层单点群卷积降低了计算成本,之后执行通道洗牌操作以减少组间信息丢失。第二层混合卷积添加在通道洗牌操作之后,利用混合卷积核在不同分辨率下的高适应度获得更加稳定的特征图,通过混合卷积中的大卷积核保留更多的特征信息;之后与融合分支处理后的特征图结合,成为连接降参分支和融合分支的桥梁。第三层单点群卷积是为了恢复通道维数以匹配标准分支。融合分支对网络进行融合编码,并在第三层单点群卷积前与降参分支以元素乘积的方式相结合,这种设计有助于降低第三层卷积过程中组间信息的丢失。

图4(b)为 MF 模块的下采样版本,做了以下修改:1)降参分支的混合卷积和融合分支中的3×3标准卷积的步长变为2。2)在标准分支中应用了步长为2的3×3平均池化,并用元素拼接(Concat)的方式代替了元素加法,将标准分支和降参分支结合起来。经过下采样 MF 模块后,特征图的空间维数被减半,而通道维数则增加了1倍。

1.5 Mix-Fusion 网络

基于 MF 模块,本文提出了一种新颖的网络结构 Mix-

Fusion,总体结构如表1所示。网络结构分为4个阶段:阶段一包括步长为2的一个3×3卷积层和max pooling层,这两层对输入图像执行4倍的下采样以降低计算成本;之后3个阶段都是由一个下采样 MF 模块和若干标准 MF 模块组成。模块类型后面的数字代表输出通道的数量。“×3”和“×7”分别表示 MF 模块重复3次或7次,“/2”表示步长为2,带有“/2”的 MF 模块执行下采样功能。表1中的“分类层”为采用核大小为7的全局平均池化层,通过全连接层输出6类缺陷的概率,生成预测分类。

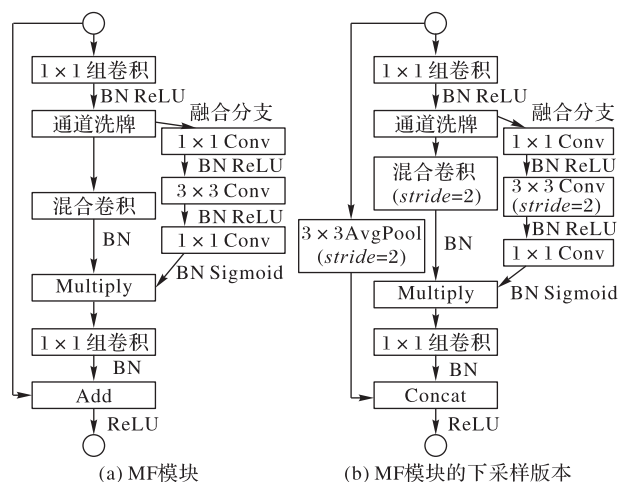


图4 Mix-Fusion 模块示意图

Fig. 4 Schematic diagram of Mix-Fusion module

表1 Mix-Fusion 网络结构

Tab. 1 Network structure of Mix-Fusion

阶段	Mix-Fusion	混合卷积核	输出(维度)
	输入图像		224×224
阶段一	3×3 conv, 24, /2 3×3 max pool, /2	[3, 5, 7]	112×112 56×56
阶段二	MF 模块, 108, /2 MF 模块, 108, ×3	[3, 5, 7]	28×28
阶段三	MF 模块, 216, /2 MF 模块, 216, ×7	[3, 5, 7]	14×14
阶段四	MF 模块, 432, /2 MF 模块, 432, ×3	[3, 5, 7]	7×7
分类层	global average pool, 6-d fc, softmax		1×1

输出通道的数量在同一阶段中保持不变,在下一阶段中增加1倍。此外,降参分支中的通道数被设置为同一 MF 模块中输出信道的1/4以节省计算成本。最后,本文将降参分支中组卷积的组数设置为3,进而增加降参分支的连接稀疏度。

2 实验设计及结果分析

2.1 实验数据集

本文采用东北大学收集的热轧带钢表面缺陷数据集 NEU-CLS^[22]。该数据集收集了六种典型热轧带钢表面缺陷,包括裂纹(Cr)、夹杂(In)、裂斑(Pa)、麻点(PS)、轧屑(RS)和划痕(SC)。每一类缺陷有300个样本,每幅图像的原始分辨率为200像素×200像素。该数据集主要面临两个挑战:1)类内缺陷在外观上存在较大差异;2)类间缺陷有相似的方面,由于缺陷图像受到光照和材料变化的影响,类间缺陷图像的灰度也会发生变化。部分缺陷图像示例样本如图5所示。

2.2 基准实验

本文在NEU-CLS数据集上提取了一种常用的传统纹理特征并进行基准测试,即灰度共生矩阵(Gray-Level Co-occurrence Matrix, GLCM)^[23]。GLCM描述了具有某种空间位置关系的两个像素灰度的联合分布,它可以反映像素的分布特征以及图片的纹理特征。本文选取对比度、差异性、同质性、熵、相关性、能量六种灰度共生矩阵统计量的组合作为需要提取的特征,并选取支持向量机(SVM)作为分类器,svm. SVC作为SVM类,linear作为核函数,惩罚因子C设置为0.5,最终分类结果达到了90.81%。

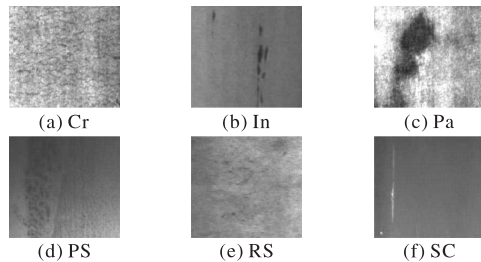


图5 示例缺陷图像及对应标签
Fig. 5 Sample defect images and corresponding labels

这种由特征提取算法加分类器设计的传统方法是目前工业界主流的一种表面缺陷分类方法,将此SVM+GLCM作为基准实验并与Mix-Fusion网络比较可以保证实验的可靠性。

2.3 环境与训练

本文的算法是在PyTorch框架上进行的,实验环境配置为:Inter Core i7-6700 CPU@ 3.40 GHz 处理器,16 GB 内存,NVIDIA GeForce GTX 1080 显卡,操作系统为Windows 10。

实验从数据集中随机选取1440张图片作为训练集,剩余360张图片作为测试集。网络在训练阶段采用Adam优化算法基于训练数据迭代地更新神经网络的权重,每一个批次(batch)包含32张图片,权值的初始学习率0.02,每隔10个epoch学习率衰减一次,衰减系数设置为0.9。在训练前将数据图像边缘调整为256像素,之后中心裁剪为224×224像素进行实验。

2.4 结果分析

为测试网络模型在钢铁表面缺陷分类任务中的综合性能,本文引入了四种评价指标:计算力,即每秒所执行的浮点运算次数(Floating-point Operations Per second, FLOPs)、参数量、精度以及平均运行时间,并使用PyTorch-OpCounter工具测试出网络的FLOPs以及参数量。其中,FLOPs表示浮点运算数,用来衡量模型的复杂度,复杂度越低,模型越轻便,一般轻量级网络的FLOPs可以降到 150×10^6 以下。精度反映了模型在缺陷分类任务中的准确率,运行平均时间为单张图像连续运行12次,去掉一个最大值和一个最小值后取10次运行的平均结果,衡量了模型的运行速度。表2展示了在NEU-CLS数据集上Mix-Fusion和一些最先进的网络结构以及基准实验关于四种评价指标的比较情况。表2中的MFLOPs(Million FLOPs)用来衡量FLOPs,指每秒浮点运算次数为 10^6 。

为验证提出的不同模块对于网络性能的影响,本文设计了一个只使用传统深度卷积的Mix-Fusion(Base)网络。基础网络结构与图4保持一致,仅仅将模块示意图中的混合卷积替换为卷积核尺度一致的3×3深度卷积,结果表明该网络测试效果明显优于基准实验,分类精度达96.67%。之后将Mix-

Fusion(Base)网络与三种经典的流行网络(GooGleNet^[24]、ResNet-50以及AlexNet^[25])作比较。其中AlexNet及ResNet-50分别取得了95.00%和95.56%的分类精度,而GooGleNet取得了稍好的精度96.38%。相比较之下,Mix-Fusion(Base)相较AlexNet精度提高了1.67个百分点的同时FLOPs数量减少为原来的1/17.2,相较ResNet-50精度提高了1.11个百分点的同时FLOPs数量减少为原来的1/99.7,相较GooGleNet精度提高了0.29个百分点的同时FLOPs数量减少为原来的1/36.5。上述结果充分说明了组卷积和通道洗牌操作具有降低网络参数量的作用,且融合编码操作打通了组间信息交流,具有降低性能损失的有效性。

表2 不同网络综合性能对比

Tab. 2 Comparison of comprehensive performance of different networks

模型	计算力/ MFLOPs	参数量/ 10^6	精度/%	运行平均 时间/s
SVM+GLCM	—	—	90.81	—
Mix-Fusion(Base)	41.2	0.22	96.67	0.93
AlexNet	710.6	57.02	95.00	1.17
GooGleNet	1 503.9	5.61	96.38	1.54
ResNet-50	4 109.5	23.52	95.56	2.33
SE-ResNet-50	4 117.5	23.04	97.52	2.59
SE-ResNet-101	7 847.7	49.29	98.07	2.94
Mix-Fusion(3-5)	35.0	0.11	97.22	0.96
Mix-Fusion(3-5-7)	43.4	0.25	98.61	1.05
Mix-Fusion(3-5-7-9)	52.7	0.41	97.50	1.22

之后,本文实验了不同大小混合卷积核对于Mix-Fusion网络性能的影响。如表2所示,网络名称的数字后缀代表了混合卷积核的组合大小,如Mix-Fusion(3-5)代表将通道平均划分为两组,分别应用 $\{(3 \times 3), (5 \times 5)\}$ 的混合卷积核。实验结果显示,随着混合卷积模块的加入,模型精度有所提高,其中Mix-Fusion(3-5-7)取得了最好的精度98.61%,表明混合卷积可以有效降低网络对大卷积核的敏感度,提高网络稳定性,优化模型性能。

为做出全面的比较实验,本文引入带有通道注意力机制的SENet进行比较。如表2所示,尽管SE-ResNet-50和SE-ResNet-101的分类精度高出Mix-Fusion(3-5)和Mix-Fusion(3-5-7-9),但是Mix-Fusion(3-5-7)的分类精度依然略微胜出,同时Mix-Fusion网络的运行速度要远快于SE-ResNet网络,SE-ResNet网络的参数量和复杂度远大于Mix-Fusion网络。

表3展示了Mix-Fusion网络对于数据的依赖性分析,训练数据占比代表新的训练集在原始训练集中所占百分比。当训练数据占比为50%和25%时网络精度略微下降,当训练数据占比为10%及以下时网络性能会极大下降,并出现轻微的过拟合现象,表明本文网络对数据量有一定的依赖性。

表3 不同训练数据占比的网络精度对比

Tab. 3 Comparison of network accuracy with different training data ratios

训练数据占比/%	精度/%
100	98.61
50	97.72
25	94.78
10	85.56

为分析网络的收敛性及稳定性,图6展示了Mix-Fusion、ShuffleNetV2以及MobileNetV2三种网络的训练损失曲线和验

证准确率曲线。

如图 6(a) 所示, Mix-Fusion 网络收敛最快, loss 稳定在 0.004; ShuffleNetV2 收敛速度次之, loss 稳定在 0.005; MobileNetV2 收敛最慢, loss 稳定在 0.009。图 6(b) 中, Mix-Fusion 的验证准确率明显高于另外两个网络, 同时准确率曲线趋势显示 Mix-Fusion 和 ShuffleNetV2 的稳定性较高, MobileNetV2 次之。

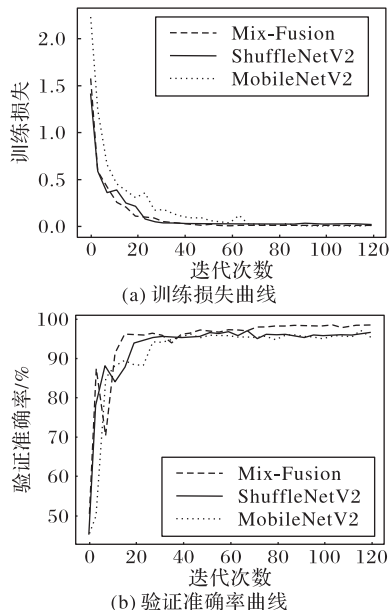


图 6 不同模型性能比较

Fig. 6 Performance comparison of different models

为了验证模型的有效性, 本文进一步比较了 Mix-Fusion 与三种先进的轻量级网络 ShuffleNet、ShuffleNetV2 以及 MobileNetV2 的综合性能, 如表 4 所示。由表 4 可以看出, Mix-Fusion 的性能优于 ShuffleNet。考虑到在 Mix-Fusion 中降参分支的通道比在 ShuffleNet 中要少, 将这种改进归因于所提出的融合编码操作。尽管 ShuffleNet 拥有更多的通道, 但它依然遭受着组间信息丢失的困扰, 而 Mix-Fusion 有效地利用了组间信息。因此, Mix-Fusion 比 ShuffleNet 产生了更具有区分性的特征, 克服了性能下降的缺陷。同时与 ShuffleNetV2、MobileNetV2 相比, Mix-Fusion 的分类精度在最小计算成本的基础上取得了分别高于前两者 1.36 个百分点以及 1.67 个百分点的好成绩。

表 4 不同轻量级网络综合性能对比

Tab. 4 Comparison of comprehensive performance of different lightweight networks

模型	计算力/MFLOPs	参数量/ 10^6	精度/%
ShuffleNet	140.7	0.91	96.32
ShuffleNetV2	146.0	2.26	97.25
MobileNetV2	300.2	3.40	96.94
Mix-Fusion(3-5-7)	43.4	0.25	98.61

综上所述, Mix-Fusion 无论是在与传统纹理特征提取方法 (GLCM), 还是其他几种经典网络或是先进的轻量级网络的对比中都能够以更小的计算成本代价下取得更高的分类精度和更快的运行速度。一方面是因为 MF 模块的融合分支有效地解决了“稀疏连接”卷积阻碍组间信息交换的问题, 降参

分支极大降低了网络的计算量; 另一方面是因为混合卷积相较于传统的深度卷积, 降低了网络对大卷积核的敏感性, 提高了网络的稳定性。

3 结语

针对卷积神经网络参数量、计算成本日益增长以及现有轻量级网络难以完全避免组间信息丢失的问题, 本文借鉴了 ShuffleNet 和 MENet 的思想, 提出了一种新颖的轻量级网络 Mix-Fusion。实验结果表明, 该网络避免了组间信息丢失, 提高了分类精度, 降低了计算成本和参数量, 相较其他网络综合性能有明显提升, 为钢铁表面缺陷分类任务在移动端的部署提供了有力支持。我们接下来的工作是进一步验证模型的泛化性能, 并开展优化算法、适配设备等方面的工作以满足产品化的需求。

参考文献 (References)

- [1] HE Y, SONG K, MENG Q, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(4): 1493-1504.
- [2] DONG H, SONG K, HE Y, et al. PGA-Net: pyramid feature fusion and global context attention network for automated surface defect detection [J]. IEEE Transactions on Industrial Informatics, 2020, 16(12): 7448-7458.
- [3] RAZAVIAN A S, AZIZPOUR H, SULLIVAN J, et al. CNN features off-the-shelf: an astounding baseline for recognition [C]// Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2014: 512-519.
- [4] AMATO G, FALCHI F, GENNARO C. Geometric consistency checks for k NN based image classification relying on local features [C]// Proceedings of the 2011 4th International Conference on Similarity Search and Applications. New York: ACM, 2011: 81-88.
- [5] LIN Y, LV F, ZHU S, et al. Large-scale image classification: fast feature extraction and SVM training [C]// Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2011: 1689-1696.
- [6] WANG J, YANG Y, MAO J, et al. CNN-RNN: a unified framework for multi-label image classification [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 2285-2294.
- [7] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2020-07-20]. <http://arxiv.org/pdf/1409.1556.pdf>.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.
- [9] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. [2020-06-08]. <https://arxiv.org/pdf/1704.04861.pdf>.
- [10] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 4510-4520.
- [11] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely

- efficient convolutional neural network for mobile devices [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 6848-6856.
- [12] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size [EB/OL]. [2020-07-07]. <https://arxiv.org/pdf/1602.07360.pdf>.
- [13] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [14] 姚明海, 陈志浩. 基于深度主动学习的磁片表面缺陷检测[J]. 计算机测量与控制, 2018, 26(9): 29-33. (YAO M H, CHEN Z H. Deep active learning in detection of surface defects on magnetic sheet [J]. Computer Measurement and Control, 2018, 26(9): 29-33.)
- [15] 张琪, 王国栋, 赵希梅, 等. 基于轻量级卷积神经网络的肝部病理组织切片分类[J]. 青岛大学学报(自然科学版), 2018, 31(4): 76-82. (ZHANG Q, WANG G D, ZHAO X M, et al. Classification of liver pathological tissue sections based on the lightweight convolutional neural network [J]. Journal of Qingdao University (Natural Science Edition), 2018, 31(4): 76-82.)
- [16] QIN Z, ZHANG Z, ZHANG S, et al. Merging and evolution: improving convolutional neural networks for mobile applications [EB/OL]. [2020-07-12]. <https://arxiv.org/pdf/1803.09127.pdf>.
- [17] TAN M, LE Q V. MixConv: mixed depthwise convolutional kernels [EB/OL]. [2020-08-02]. <https://arxiv.org/pdf/1907.09595.pdf>.
- [18] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]// Proceedings of the 2015 32nd International Conference on Machine Learning. New York: ACM, 2015: 448-456.
- [19] CHOLLET F. Xception: deep learning with depthwise separable convolutions [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 1800-1807.
- [20] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 2818-2826.
- [21] ZOPH B, VASUDEVAN V, SHLENS J, et al. Learning transferable architectures for scalable image recognition [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 8697-8710.
- [22] SONG K, YAN Y. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects [J]. Applied Surface Science, 2013, 285(Pt B): 858-864.
- [23] REN R, HUNG T, TAN K C, et al. A generic deep-learning-based approach for automated surface inspection [J]. IEEE Transactions on Cybernetics, 2018, 48(3): 929-940.
- [24] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 1-9.
- [25] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]// Proceedings of the 2012 25th International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2012: 1097-1105.

This work is partially supported by the National Natural Science Foundation of China (61872004).

SHI Yangxiao, born in 1997, M. S. candidate. His research interests include deep learning, defect detection.

ZHANG Jun, born in 1971, Ph. D., professor. His research interests include application of deep learning in video analysis, chemoinformatics.

CHEN Peng, born in 1978, Ph. D., professor. His research interests include machine learning, bioinformatics.

WANG Bing, born in 1976, Ph. D., associate professor. His research interests include pattern recognition, bioinformatics.