



## 基于多任务级联残差网络的枪支图像识别系统

周志飞<sup>1</sup>,吴金龙<sup>2</sup>,李轶映<sup>1</sup>,贾力榜<sup>3</sup>,沈玉杰<sup>1</sup>,张刚<sup>1</sup>,崔斌<sup>1</sup>

(1.公安部物证鉴定中心,北京 100038; 2.北京多维视通技术有限公司,北京 100070; 3.中国科学院自动化研究所,北京 100190)

**摘要:**针对枪支种属识别目前依赖检验人员经验、识别效率较低的问题,建立一种基于多任务级联深度残差网络的枪支图像自动识别模型。以ResNet18为基本构建单元,通过级联融合4个任务中的Softmax损失函数约束,实现对枪支图像从枪族到枪型的多维度聚类。在该模型的基础上,设计一套制式枪支图像智能检索系统,对拍摄上传的枪支图像种属信息进行自动识别。在自建的制式枪支图像数据集上进行实验,结果表明,与EfficientNet、NTS-net等模型相比,该模型的识别准确率更高,Rank-1、Rank-20识别准确率分别达到61.12%、95.28%,且其具有更好的鲁棒性。

**关键词:**枪支种属识别;深度学习;残差网络;细粒度图像识别;数据增广

开放科学(资源服务)标志码(OSID):



中文引用格式:周志飞,吴金龙,李轶映,等.基于多任务级联残差网络的枪支图像识别系统[J].计算机工程,2022,48(1):214-219.

英文引用格式:ZHOU Z F, WU J L, LI Y Y, et al. Firearm image recognition system based on multi-task cascaded residual network[J]. Computer Engineering, 2022, 48(1): 214-219.

## Firearm Image Recognition System Based on Multi-Task Cascaded Residual Network

ZHOU Zhifei<sup>1</sup>, WU Jinlong<sup>2</sup>, LI Yiyi<sup>1</sup>, JIA Libang<sup>3</sup>, SHEN Yujie<sup>1</sup>, ZHANG Gang<sup>1</sup>, CUI Bin<sup>1</sup>

(1. Institute of Forensic Science of China, Beijing 100038, China; 2. Beijing Visystem Co., Ltd., Beijing 100070, China;

3. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

**[Abstract]** The traditional firearm recognition methods rely heavily on expertise, and are limited in recognition accuracy. To address the problem, a model for automatic firearm image recognition is built using multi-task cascaded deep residual network. With ResNet18 as the basic build block, this model fuses the Softmax loss function constraints in the four cascaded tasks and realizes firearm image clustering, which is based on multiple dimensions ranging from the firearm family to the specific gun type. Based on the proposed model, a system for intelligent firearm image retrieval is designed, which can automatically recognize the type of the firearms in uploaded images. The experimental results on a self-made firearm image dataset show that the model displays a higher recognition accuracy in Rank-1 (61.12%) and Rank-20 (95.28%) than EfficientNet, NTS-net and other models. The proposed model also provides better robustness for gun image recognition in real scenes.

**[Key words]** firearm species recognition; deep learning; residual network; fine-grained image recognition; data augmentation  
DOI: 10.19678/j.issn.1000-3428.0059820

### 0 概述

枪支种属识别是涉枪案件物证检验的基础,对涉枪案件的侦查和检验具有重要意义<sup>[1-2]</sup>。枪支种属识别以外形特征作为主要判定因素,长期以来,检验人员进行识别时主要以工作经验为主,以查阅书籍、文献资料等为辅,导致查询效率较低。随着涉

枪案件现场物证勘查勘验对时效性的要求越来越高,枪弹痕迹检验领域迫切需要一种快速且准确的枪支识别技术。

境外一些国家和地区的法庭科学机构曾对以图像识别为基础的枪种识别展开研究,多数通过传统图像识别技术进行枪支识别,且大多利用外形特征提取某个特征点,然后利用特征进行匹配以获取检

**基金项目:**中央级公益性科研院所基本科研业务费专项资金(2018JB020)。

**作者简介:**周志飞(1985—),男,助理研究员、硕士,主研方向为枪弹痕迹检验;吴金龙,工程师、硕士;李轶映,副研究员、硕士;贾力榜,工程师、硕士;沈玉杰,硕士;张刚,助理研究员、博士;崔斌,助理研究员、硕士。

**收稿日期:**2020-10-23 **修回日期:**2021-01-13 **E-mail:**zhouzhifei@cifs.gov.cn

索结果<sup>[3-5]</sup>。印度法庭科学实验室的MURTHY等<sup>[4]</sup>创建一个包含254张图像的9 mm手枪检索库,以枪支整体形状、握把、扳机等部位作为人工标定的特征点来进行检索和识别。中国台湾地区中央警校法庭科学部门的WEN等<sup>[5]</sup>对枪支图像进行外形轮廓特征提取和优化,计算枪支图像轮廓边缘到重心的距离分布,以前3个峰值点的位置作为对比特征,构建一个包含300支手枪库容量的检索库。目前,中国大陆较少有对以图像识别为基础的枪种识别展开研究,但有研究人员对枪支图像的特征提取与识别进行分析,大多利用传统图像匹配技术获取图像的特征值以进行枪支与非枪支识别,对于枪支集合下不同种类的枪支识别没有进一步研究。例如:冯艳平等<sup>[6]</sup>在太赫兹(THz)图像上进行阈值分割和边缘检测,计算枪的面积、长度、宽度以及枪管与枪托之间的夹角,以区分枪支与非枪支;和睿等<sup>[7]</sup>探索结合Canny特征和矩阵相似度理论的图像检索算法,并在手枪图像的检测和检索方面进行实验;张树江等<sup>[8]</sup>设计一套枪弹痕迹图像识别系统,通过利用基于小波直方图的图像检索方法在数据库中检索与目标图像最相似的样本图像,从而为枪弹检验人员提供参考。

上述研究主要采用传统的模式识别方法,通常包含特征提取和分类器设计2个步骤。其中,特征提取由人工设计特征,需要依靠设计者大量的先验知识,而对于枪支图像这类特殊的图像,先验知识难以挖掘;传统分类器大多结果较为简单,难以通过多个简单阈值完成复杂的图像识别任务。上述问题导致目前已有的枪支识别算法大多只能研究单一枪种,特征提取单一,数据库容量小,无法满足实际应用的需求。基于深度学习的方法<sup>[9]</sup>可以更好地处理图像分类任务,相比于传统机器学习,深度学习方法(如深度卷积神经网络<sup>[10]</sup>)采用更深、更多层的卷积结构,利用大规模的训练数据集以端对端的方式来训练网络中的权重参数,从而实现目标特征的自动提取和分类,克服人工提取设计特征的缺陷,且更容易学习到深层次的语义特征<sup>[10-11]</sup>。

一般认为,网络层数越深,模型表达能力越强,但是深层网络也带来了许多问题,如模型过拟合、梯度消失等。残差网络(ResNet)<sup>[12]</sup>通过引入捷径连接,学习信息的差值并增加梯度传播,从而有效避免网络层数增加导致的训练误差增大、模型退化问题。残差网络被广泛应用于目标检测、姿态估计、图像识别等任务<sup>[13]</sup>。例如:文献<sup>[14]</sup>提出一种改进的残差网络,结合交叉熵损失函数与中心损失函数完成人脸识别,其在测试数据集上取得了97.46%的准确率;文献<sup>[15]</sup>提出一种基于ResNet101的头部姿态估计方法,其利用残差网络结合回归损失预测欧拉

角以表示头部姿态;文献<sup>[16]</sup>采用152层残差网络架构,对花卉图像的特征进行自动提取,实现了一种基于迁移学习的花卉识别系统。

本文采用基于深度学习的方法,通过改进和优化深度残差网络模型的训练方式,使之适用于种类多、样本少的大规模枪支图像识别任务。

1 制式枪支数据集构建

枪支数据集是枪支图像检索的基础,枪支图像的质量影响枪支图像检索算法的精度。因此,为了保证枪支图像数据集的图像质量,本文制定枪支图像数据集的图像采集和图像增强规范。

1.1 枪支图像采集

为了保证枪支图像数据集的图像质量,本文参考《比例照相方法规则(GA/T 158-1996)》拍摄枪支物证照片的操作原则,制定图像采集规范如下:在拍摄前,解除枪支的附件或配件,如瞄准镜、消音器等,仅保留枪支本体,并且使枪支处于原始状态;在拍摄时,枪支主体与相机镜头角度垂直,枪支主体应位于图像的中心位置,枪管与水平面平行,枪支的外形特征清晰,光线均匀,背景为纯色;一般应拍摄左右两面的枪支照片,并且拍摄枪支标识等细目照片。

按照图像采集规范,本文通过实物拍摄、图册扫描、互联网查询下载等方式,采集到国内外54个国家和地区共7 364种枪支,38 156张枪支图像,其中,枪支整体图像数为21 128,枪支细目图像数为17 028,枪支种类包括手枪、步枪、猎枪、冲锋枪、机枪等,具体情况如表1所示。部分枪支图像样本示例如图1所示。

表1 枪支图像数据集信息

Table 1 Dataset information of firearms images

枪支类型	数量	所占比例/%
手枪	3 243	44.03
步枪	2 544	34.55
猎枪	1 146	15.56
冲锋枪	155	2.10
机枪	155	2.10
其他	121	1.64



图1 枪支图像数据集样本示例

Fig.1 Sample example of the firearms images dataset

## 1.2 枪支图像增强扩充

由于枪支图像样本较少,为了获取足够多的数据量以学习枪支型号的鲁棒信息,本文使用深度学习中常用的数据增广方法来对枪支图像进行增强<sup>[11]</sup>:首先,将原始图像统一调整到256像素×256像素;然后,对图像进行裁剪,截取主要以中心裁剪为主,取大小为224像素×224像素,每张截取的图像都进行左右翻转、上下翻转,调整图像亮度,亮度调整因子设置为0.2。

## 2 深度网络模型设计及检索系统

本文检索系统的开发目标是在自建的制式枪支图像数据集基础上,实现对涉案制式枪支图像中枪支种类信息的自动识别。一方面,由于制式枪支种类多,往往需要层数较深的卷积神经网络来实现分类;另一方面,由于每种枪支的图像样本数量非常有限,采用较深层的神经网络模型极易产生过拟合问题。相比于普通卷积神经网络,ResNet<sup>[13,17]</sup>增加了捷径连接,捷径连接可以学习信号的差值,从而有效避免网络层数增加导致的训练误差增大、模型退化问题。因此,本文采用残差网络作为枪支图像识别的基准模型。

### 2.1 深度残差网络基本结构

深度残差网络基本结构如图2所示,右侧分支称作残差分支,左侧部分称作本体分支, $x$ 为输入信号, $H(x)$ 是期望输出。通过捷径连接即残差分支,将输入 $x$ 跳变连接到输出层,此时残差网络相当于要学习目标值 $H(x)$ 和 $x$ 的差值,即残差 $F(x)=H(x)-x$ 。因此,通过不断地训练使残差结果逼近0,这样即使网络层数不断加深,准确率也不会下降,网络输出始终可以很好地表征网络输入的特征信息。残差的引入使得网络的参数量和计算复杂度并未提升,任何已有的网络模型都可以在原有框架的基础上引入残差模块。

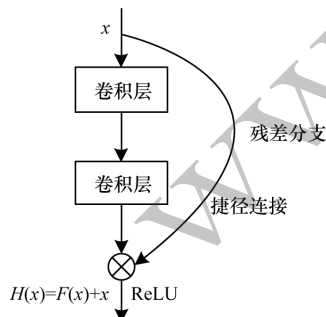


图2 深度残差网络基本结构

Fig.2 Basic structure of depth residual network

在深度残差网络中,为了保证网络的稳定性和准确性,每个残差单元都包含3个基本算子,分别为批归一化(Batch Normalization, BN)算子、激活算子、池化(Pooling)算子<sup>[18]</sup>。BN算子对激活函数的输

入进行归一化,相当于增加一个约束,限定网络的配置空间,从而有效解决输入数据发生偏移和增大的问题,避免梯度消失和梯度爆炸现象。对于输入的一批(Batch)样本 $x=[x_1, x_2, \dots, x_n]$ ,BN算子的操作如式(1)所示:

$$x'_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (1)$$

其中: $\mu_B = \bar{x}$ 为均值; $\sigma_B^2$ 为方差; $\epsilon$ 为一个极小值; $x'_i$ 为归一化后的激活值。结合枪支图像种类多、样本少的特点,本文采用最常用的线性修正(Rectified Linear Unit, ReLU)激活算子。

对于池化算子,针对枪支识别问题,本文在初始输入特征时采用最大池化(Max Pooling),可以更好地保留图像纹理特征;在最后一个卷积层的特征图上,采用全局平均池化(Average Pooling),这样更有利于保留整体数据信息,得到用于枪支检索的高级特征。

### 2.2 枪支图像识别网络模型设计

枪支图像数据集中枪支类别多(上千种类别),每种枪支图像的样本又很少(每种类别的枪支图片平均只有2~3张),导致难以实现稳定的网络训练,网络训练非常容易过拟合。因此,本文构建一种新的多任务级联残差网络模型,该网络模型共包含4个任务,或称为4个阶段,从而实现由粗到细的枪支图像分级识别和检索:第一阶段实现枪支图像的左右朝向分类;第二阶段从枪族的概念出发,将从同一个枪族发展变种而来的枪支类型合并为一类,称为枪族类( $ID_{family}$ );第三阶段依据每个大类中枪支图像的外观,将较为相似的枪支和明显不相似的枪支再次进行合并和分类,称为相似类( $ID_{similar}$ );第四阶段则细分到具体型号的枪支类型,称为原始类( $ID_{origin}$ )。网络模型中包含的这4个阶段均采用Softmax损失函数<sup>[19]</sup>输出其分值,如式(2)所示:

$$\text{Softmax}(f_y) = \frac{\exp(f_y)}{\sum_{i=1}^K \exp(f_i)} \quad (2)$$

最后,整个网络模型通过融合这4个阶段Softmax损失函数的输出,共同约束需要查找的枪支图像和库中的枪支图像的相似度,训练深度网络模型参数,从而获取高级检索特征。

此外,针对基于深度残差网络的枪支图像识别任务,本文尝试18层残差网络(ResNet18)、50层残差网络(ResNet50)和101层残差网络(ResNet101),对比实验结果发现,ResNet18网络可以满足应用要求,并且由于其网络层数较少,模型不容易出现过拟合问题。

改进后的多任务级联残差网络整体结构如图3所示,其中共包含4个阶段,每个阶段(Stage $N$ ,  $N=1, 2, 3, 4$ )中都包含一个2.1节所述的基础网络,该基础网络是一个包含BN、ReLU和Pooling的ResNet18。



在每个 Stage, 信息传递有 2 条支路, 右侧为残差分支, 左侧为本体分支。1×1 的卷积层加在残差分支, 可以在改变输出通道数的同时提高非线性特性, 进一步提升模型表现能力。

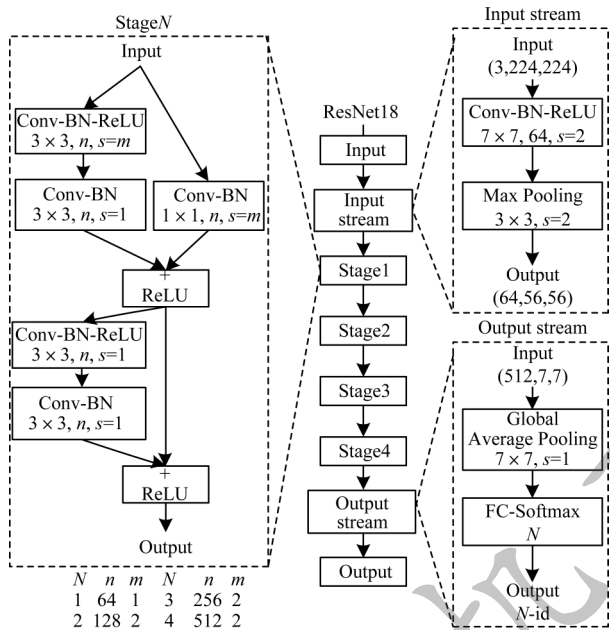


图3 枪支识别系统网络架构

Fig.3 Firearms identification system network architecture

整个枪支识别系统模型的步骤如下：

**步骤1** 创建一定数量的枪支图像样本, 采用代码自动化处理及人工处理相结合的方式将枪支图像粗分为整枪、枪支部件和标记错误图像, 并剔除无法识别的图像。

**步骤2** 对图像样本进行剪裁及翻转, 再将样本按人工经验提取出部分数据, 按所属类别进行分类, 拆分成训练集和测试集。

**步骤3** 将处理好的图像输入进行 7×7 卷积、BN、ReLU 激活及最大池化操作。

**步骤4** 瓶颈块结构如图 3 中 StageN 模块所示, 右侧为残差分支, 左侧为本体分支。对于本体分支, 首先对输入数据进行 3×3 卷积、BN 及第一次 ReLU 非线性激活, 接着将得到的特征送入下一层网络, 再进行一次卷积和 BN 操作; 对于残差分支, 对输入数据进行 1×1 卷积、BN 然后作为输出, 与左侧分支特征图进行合并, 进行第二次 ReLU 激活; 随后将第二次 ReLU 非线性激活的输出结果作为输入, 送入本体分支进行 3×3 卷积、BN 及第三次 ReLU 非线性激活, 进一步进行 3×3 卷积、BN 处理得到相应的特征图。本体分支特征图与残差分支特征图相合并, 进行第四次 ReLU 激活, 整个过程构成网络的一个阶段 (Stage)。

**步骤5** 枪支识别网络模型由依次相连的 4 个 Stage 组成, 每一段网络结构类似。对第四阶段 (Stage4) 输出的枪支二维特征图, 在池化层进行全局平均池化, 然后接入全连接层, 并通过 Softmax 层

进行归一化处理, 采用 4 个 Softmax 损失函数联合作用, 约束枪支的左右朝向、枪族属性 (ID<sub>Family</sub>)、相似枪支类 (ID<sub>Similar</sub>) 和精确的枪支型号 (ID<sub>origin</sub>)。

**步骤6** 将 ImageNet<sup>[20]</sup> 上预训练模型的权重加载到改进 ResNet 上对应的层, 设置学习率、Batch 大小、权重衰减正则项等参数。对训练数据进行迭代训练, 保存多次训练中最高识别精度的模型权重, 并在测试集上加载训练好的模型进行测试, 从而验证模型精度。

2.3 枪支图像自动识别与检索系统

根据枪支图像识别算法的应用需求, 本文设计一种枪支图像检索系统, 该系统基于 B/S 架构, 部署于公安网, 采用 Java 开发语言。系统以枪支图像数据集和枪支基础信息数据库数据作为支撑, 以枪支图像检索算法为核心功能, 通过上传枪支物证图像, 为用户提供枪支图像自动识别和信息查询等服务。枪支图像检索系统架构如图 4 所示。

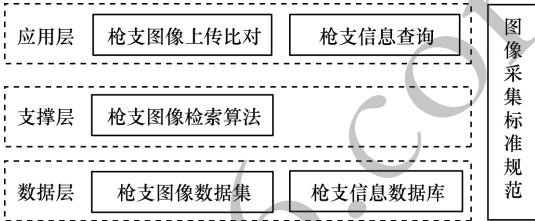


图4 枪支图像检索系统架构

Fig.4 Firearm images retrieval system architecture

枪支图像检索系统包括图像上传、图像比对、信息查询 3 个功能模块。用户按照要求将图像上传到系统后, 枪支图像检索算法自动提取所上传枪支图像的特征, 并与系统中枪支图像数据集的图像特征进行比对, 按照相似度由高到低依次排列识别到的枪支图像。用户可以选择枪支图像查阅枪支详情并进一步确认查询结果, 如图 5 所示。



图5 对比结果页面

Fig.5 Comparison results pages

3 实验结果与分析

3.1 模型训练和实验设计

本文所提制式枪支图像识别系统基于 PyTorch 深度学习框架构建, 采用 Java 语言搭建实现前端业务系统。系统测试环境为 Ubuntu16.04 操作系统, Intel Core i7 处理器, GTX1080Ti 显卡, 显存 11 GB。模型训练采用

Adam 优化算法,批次大小 (Batch size) 设为 32,初始学习率设为  $10^{-4}$ ,在经过 30 个轮次的训练后,学习率逐步衰减到  $10^{-6}$ 。评估指标主要包括:

1)前  $n$  位命中概率 (Rank- $n$ ),即将待检索的图像输入识别网络提取特征后,计算其与数据库中各类枪种图像特征的欧式距离,以该距离作为相似度从而对检索结果进行排序,排序结果中最靠前的  $n$  张图像中有正确结果的概率。Rank- $n$  具体计算如式 (3) 所示:

$$R(n)=\frac{\sum_{i=1}^N A(n_i)}{N}$$
$$A(n_i)=\begin{cases} 1, & \text{前 } n \text{ 个结果中有第 } i \text{ 张图像所属类别} \\ 0, & \text{前 } n \text{ 个结果中没有第 } i \text{ 张图像所属类别} \end{cases}$$

(3)

其中:  $N$  为待检索的图像总数量。

2)模型推理速度,即每秒可以处理的图像帧数 (Frame Per Second, FPS)。

测试样本从枪支测试数据集中选取,检索数据集为本文 2.1 节所述的自建枪支数据集,共包含 7 364 种类型的枪支。在枪种的挑选上,从实战应用的角度出发,用于测试的枪支图像涵盖了国内常见制式枪支,如 51 式 7.62 mm 手枪、54 式 7.62 mm 手枪、56 式 7.62 mm 冲锋枪、64 式 7.62 mm 手枪、92 式 9 mm 手枪、56 式半自动手枪等,同时包括国外知名枪支,如格洛克 17 手枪、伯莱塔 92F 手枪、西格绍尔 P229 手枪、鲁格 P85 手枪等,共计 874 种枪支类型,1 018 张枪支图像。

3.2 结果分析

将 3.1 节所选取的 1 018 张测试枪支图像逐一输入本文所搭建的制式枪支图像检索系统,提取测试图像的深度特征,并与数据库中每类枪支图像的深度特征进行比较,计算其欧式距离,按照欧式距离的相似度对库中检索的每类枪支图像排序输出,记录排序结果,获取前  $n$  位命中的检索结果。同时,除了与经典残差网络 (ResNet) 模型<sup>[12]</sup> 进行比较,还和近期主流分类网络模型 EfficientNet<sup>[21]</sup>、专门用于细粒度图像分类的网络模型 NTS-net<sup>[22]</sup> 进行比较。表 2 所示为采用本文模型、残差网络 ResNet18、ResNet50、EfficientNet 和 NTS-net 对相同样本进行训练后的测试结果,其中,检索库枪支种类数为 7 364,检索库枪支图像数量为 21 128,测试库图像数量为 1 018。

表 2 5 种模型的测试结果

模型	前 $n$ 位命中概率				%
	第 1 位	前 10 位	前 20 位	20 位以外	
ResNet18	41.68	79.18	86.67	13.33	
ResNet50	45.14	81.06	86.55	13.45	
NTS-net	60.40	91.29	94.32	5.68	
EfficientNet	59.01	89.64	92.24	7.76	
本文模型	61.78	91.85	95.19	4.81	

从表 2 可以看出:采用 ResNet50 虽然在训练集中能够获得比 ResNet18 更好的准确率,但是在测试集中前 20 位命中概率比 ResNet18 略微下降,说明对于细粒度小样本分类问题,采用 ResNet50 会造成过学习;标准的 EfficientNet 模型结果优于 ResNet;专门用于细粒度图像分类的 NTS-net 模型准确率优于 EfficientNet,并且非常接近本文模型。图 6、图 7 所示为 2 把枪支 (64 式 7.62 mm 手枪、格洛克 17 手枪) 图像检索识别结果,从中可以看出,本文基于多任务级联残差网络模型可以准确识别出枪支的型号,并且将与之相似的枪支型号的图像均按照相似度进行排序展示。



图 6 64 式手枪检索结果

Fig.6 Retrieval results of 64 pistol



图 7 格洛克 17 手枪检索结果

Fig.7 Retrieval results of Glock 17 pistol

表 3 所示为不同模型的参数量大小以及推理平均帧率,从表 3 可以看出:当批次大小为 32 时,本文模型的帧率虽然略微低于 ResNet50,但基本也可满足实际需求,并且在本文所构建的枪支库中也能够实现秒级检索返回;EfficientNet 虽然有较小的模型参数量,但是其推理速度并无太大优势,这主要是由于采用 EfficientNet 生成的特征向量维数较大。

表 3 模型参数量及推理速度比较

模型	模型参数量/ $10^6$	平均帧率/FPS
ResNet18	51.0	46.62
ResNet50	109.8	16.35
NTS-net	192.8	3.60
EfficientNet	24.0	12.35
本文模型	128.0	15.77

#### 4 结束语

传统的枪支识别技术依赖检验人员的经验, 查询速度慢且精度低。本文建立一种基于多任务级联深度残差网络的枪支图像识别模型, 该模型以 ResNet18 为基本构建单元, 通过对枪支图像从像素级特征到语义级特征的分级划分, 并融合每个任务的损失函数, 以共同训练网络参数, 实现针对制式枪支图像的自动识别。在此基础上, 采用 Java 语言构建一套枪支图像自动识别系统。实验结果表明, 该模型对真实场景枪支图像的 Rank-20 识别准确率超过 95%, 并且具有较好的鲁棒性。下一步将对深度网络模型的轻量化进行研究, 降低识别模型的参数量并提高模型的推理速度。此外, 探索并实现针对非制式枪支图像的识别和检索系统也是今后的研究方向。

#### 参考文献

- [1] 李岳松, 罗亚平. 运用 Evofinder 枪弹自动识别系统对建档手枪弹头数字化建档的研究[J]. 刑事技术, 2018, 43(5): 363-368.  
LI Y S, LUO Y P. Digital filing into the registered-bullet imaging database by ballistic identification system Evofinder[J]. Forensic Science and Technology, 2018, 43(5): 363-368. (in Chinese)
- [2] 周志飞, 李铁舫, 鲍立垠, 等. 非制式枪射击弹头痕迹自动识别系统及性能测试[J]. 刑事技术, 2016, 41(2): 99-102.  
ZHOU Z F, LI Y Y, BAO L Y, et al. Introduction to the automatic identification system of non-standard firearms' bullet marks and its performance[J]. Forensic Science and Technology, 2016, 41(2): 99-102. (in Chinese)
- [3] GERULES G, BHATIA S K, JACKSON D E. A survey of image processing techniques and statistics for ballistic specimens in forensic science[J]. Science & Justice, 2013, 53(2): 236-250.
- [4] MURTHY S S, MAZUMDAR C, RAO M C, et al. Intensity-based image matching for regular firearms[J]. International Journal of Imaging Systems and Technology, 2002, 12(2): 68-72.
- [5] WEN C Y, YAO J Y. Pistol image retrieval by shape representation[J]. Forensic Science International, 2006, 155(1): 35-50.
- [6] 冯艳平, 王徽. 基于阈值分割和边缘检测的枪支 THz 图像识别[J]. 红外, 2011, 32(12): 23-26.  
FENG Y P, WANG H. Feature recognition of pistol in THz image based on threshold segmentation and edge detection[J]. Infrared, 2011, 32(12): 23-26. (in Chinese)
- [7] 和睿, 孙永奇. Canny 特征在基于内容的图像检索中的应用[J]. 云南大学学报(自然科学版), 2011, 33(6): 651-657.  
HE R, SUN Y Q. A content-based image retrieval method based on Canny feature[J]. Journal of Yunan University (Natural Science Edition), 2011, 33(6): 651-657. (in Chinese)
- [8] 张树江, 颜景龙, 邢慧. 一种基于图像检索的枪弹识别系统[J]. 兵工学报, 2008(4): 459-463.  
ZHANG S J, YAN J L, XING H. A cartridge identification system based on image retrieval[J]. Acta Armamentarii, 2008(4): 459-463. (in Chinese)
- [9] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [10] SZEGEDY C, WEI L, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2015: 1-9.
- [11] SHORTEN C, KHOSHGOFTAAAR T M. A survey on image data augmentation for deep learning[J]. Journal of Big Data, 2019, 6(1): 10-15.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 156-163.
- [13] 郭玥秀, 杨伟, 刘琦, 等. 残差网络研究综述[J]. 计算机应用研究, 2020, 37(5): 1292-1297.  
GUO Y X, YANG W, LIU Q, et al. Survey of residual network[J]. Application Research of Computers, 2020, 37(5): 1292-1297. (in Chinese)
- [14] 张枫, 田联房, 杜启亮. 基于残差网络与中心损失的人脸识别[J]. 计算机工程与设计, 2019, 40(6): 1689-1695.  
ZHANG F, TIAN L F, DU Q L. Face recognition based on ResNet and center loss[J]. Computer Engineering and Design, 2019, 40(6): 1689-1695. (in Chinese)
- [15] 齐永锋, 马中玉. 基于深度残差网络的多损失头部姿态估计[J]. 计算机工程, 2020, 46(12): 247-253.  
QI Y F, MA Z Y. Multi-loss head posture estimation based on deep residual network[J]. Computer Engineering, 2020, 46(12): 247-253. (in Chinese)
- [16] 关胤. 基于残差网络迁移学习的花卉识别系统[J]. 计算机工程与应用, 2019, 55(1): 174-179.  
GUAN Y. Flower species recognition system based on residual network transfer learning[J]. Computer Engineering and Applications, 2019, 55(1): 174-179. (in Chinese)
- [17] 倪志文, 马小虎, 孙霄, 等. 结合显式和隐式特征交互的深度融合模型[J]. 计算机工程, 2020, 46(3): 87-92, 98.  
NI Z W, MA X H, SUN X, et al. Deep fusion model combining explicit and implicit feature interactions[J]. Computer Engineering, 2020, 46(3): 87-92, 98. (in Chinese)
- [18] SERGEY I, CHRISTIAN S. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. [2020-09-05]. <https://arxiv.org/pdf/1502.03167v2.pdf>.
- [19] FENG W, CHEN J, LIU W Y, et al. Additive margin Softmax for face verification[J]. IEEE Signal Processing Letters, 2018, 25(7): 926-930.
- [20] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [21] TAN M, LE Q. EfficientNet: rethinking model scaling for convolutional neural networks [C]//Proceedings of the 36th International Conference on Machine Learning. Washington D. C., USA: IEEE Press, 2019: 6105-6114.
- [22] YANG Z, LUO T, WANG D, et al. Learning to navigate for fine-grained classification [C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2018: 438-454.

编辑 吴云芳