

文章编号:1001-9383(2019)03-0028-09

# 基于深度学习的图像识别技术研究综述

张琦, 张荣梅, 陈彬

(河北经贸大学信息技术学院, 河北 石家庄 050061)

**摘要:**本文介绍了基于深度学习的图像识别算法,包括 R-CNN、SPP-Net、Fast R-CNN、Faster R-CNN、YOLO 以及 SDD 算法,讨论了深度学习在人脸识别、车牌识别和医学图像识别方面的应用,最后对深度学习图像识别技术的研究提出了问题与展望。

**关键词:**图像识别; CNN; R-CNN; SPP-Net; Fast R-CNN

**中图分类号:** TP181

**文献标识码:** A

**DOI:**10.16191/j.cnki.hbkx.2019.03.004

## Research review of image recognition technology based on deep learning

ZHANG Qi, ZHANG Rong-mei, CHEN Bin

(College of Information Technology, Hebei University of Economics and Business, Shijiazhuang Hebei 050061, China)

**Abstract:** This paper introduces several deep learning network models which was commonly used in image processing, such as R-CNN, SPP-Net, Fast R-CNN, Faster R-CNN, YOLO and SDD. Then the article discussed the application of deep learning in face recognition, license plate recognition and medical image recognition. Finally, the paper puts forward some problems and prospects for the research of deep learning image recognition technology.

**Keywords:** Image recognition; CNN; R-CNN; SPP-Net; Fast R-CNN

## 引言

随着计算机技术不断发展,图像数据量呈指数型增长,如何对海量的图像数据进行智能处理成为了学者们研究的热点。2012年, Hinton 教授以及他的团队第一次通过采用卷积神经网络实现了图像识别,推动了图像识别领域的发展<sup>[1]</sup>。现在,图像识别技术已经应用到交通、医疗、安防、刑侦等社会生活中的各个行业。

本文介绍了图像识别技术的发展史以及卷积神经网络模型;讨论了基于深度学习的图像识别算法,包括 R-CNN、SPP-Net、Fast R-CNN、Faster R-CNN、YOLO 以及 SDD 算法;对图像识别技术的几个典型应用进行了分析,并给出了图像处理技术存在的问题以及未来的发展方向。

收稿日期:2019-05-15

作者简介:张琦(1995-),女,河北保定人,硕士研究生,研究方向:深度学习、图像识别。

## 1 图像识别技术

### 1.1 图像识别过程

计算机的图像识别通常分为两大步骤:图像特征提取和图像分类预测。首先对输入图片进行预处理,处理为适合特征提取的形式,然后提取图像的特征,之后对特征图像进行分类预测,如图1所示。

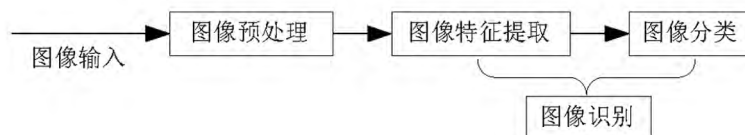


图1 图像识别过程

其中,图像预处理主要是为了消除干扰,增强目标图像信息,更好地进行图像特征提取。常见的预处理方法有:

- (1) 灰度化:将 RGB 图像改为灰度图像,主要有分量法、加权平均法等;
- (2) 几何变换:主要通过插值法对图像进行空间变换,减少图像误差信息;
- (3) 图像增强:增强目标图像信息,包括灰度变换法、直方图修正以及滤波等方法。

传统的图像识别方法需要人为提取图片特征,识别精度依赖于特征提取的准确度。它的图像特征提取方法包括 SIFT、HOG 以及 SURF 等,图像分类技术包括 KNN、SVM 以及朴素贝叶斯等分类算法。

随着计算机技术的发展,图片分辨率越来越高,传统的图像识别技术已经不适用于处理大数量高分辨率的数据图片。深度学习图像识别技术可以直接处理输入图像,避免了复杂的图像特征提取以及数据重建过程,因而得到了更为广泛的应用,成为当前图像识别领域的研究热点。

### 1.2 图像识别技术发展历史

本文将分别从图像特征提取技术以及图像分类技术两方面介绍图像识别技术的发展历史。

#### 1.2.1 图像特征提取技术

1988 年,C. Harris 等人<sup>[2]</sup>提出了 Harris 角点检测算法,通过判断图像结构中的奇异值得到图像角点,提取图像特征。

1996 年,T. Ojala 等人<sup>[3]</sup>提出了 LBP(Local Binary Pattern,局部二值模式)算法,该算法通过比较局部像素值与中心像素值的大小生成 LBP 值,提取图像纹理信息。

1999 年,David Lowe<sup>[4-5]</sup>提出了 SIFT 算法(Scale-invariant feature transform,尺度不变特征转换),并在 2004 年进行了完善补充。该算法在不同尺度空间上选择关键特征点进行描述,生成图像特征。

2001 年,Viola 等人<sup>[6]</sup>提出了 Haar 特征提取方法,并结合 AdaBoost 分类算法实现了人脸检测。

2005 年,Dalal<sup>[7]</sup>提出了 HOG 算法(Histogram of Oriented Gradient,方向梯度直方图),通过边缘特征进行特征识别,与 SVM 分类器结合实现行人检测。

2008 年, Bay 等人<sup>[8]</sup>提出了 SURF(Speed Up Robust Features, 加速稳健特征), 作为对 SIFT 算法的改进, 极大地减小了程序的运行时间, 增加了算法的鲁棒性。

### 1.2.2 图像分类技术

1997 年, Osuna E<sup>[9]</sup>提出了基于 SVM 的人脸检测算法。通过测试 313 张高清人脸图片得到了 97.1% 的准确率。

1998 年, Lecun Y<sup>[10]</sup>提出了 Lenet-5 网络结构并实现了对 mnist 数据集的训练。但是由于当时计算机处理速度较慢, 无法训练大样本, 当时该模型并没有被广泛应用。

2007 年, 董立岩<sup>[11]</sup>等人实现了贝叶斯分类器对尿沉渣图像进行分类判断。该算法对 1500 张不同类别的样本进行分类识别, 准确率达到 94% 以上。余肖生<sup>[12]</sup>等人使用 KNN 模型实现了对医学图像的自动分类, 准确率达到 85%。

2008 年, 潘琛<sup>[13]</sup>等人使用决策树的方法对遥感图像进行分类, 并以徐州市卫星图像为数据, 平均准确率达到 95%。

2012 年, Alex Krizhevsky<sup>[1]</sup>提出了包含 7 层隐藏层的 AlexNet 网络模型, 逐层对图像特征进行提取, 在 ImageNet 数据集中失误率为 15.4%。

2014 年, Szegedy C 等人<sup>[14]</sup>设计了 GoogLeNet 网络模型, 提出了 inception 结构以及分支结构。该模型在 ImageNet 数据集中将错误率降低到了 6.7%。Simonyan K 等人<sup>[15]</sup>提出了 VGG-Net 模型, 该模型使用了 19 层卷积网络, 同时使用了小尺度的卷积核, 减少了模型参数的数量, 在 ImageNet 数据集中达到了 7.3% 的错误率。

2015 年, 何恺明等人<sup>[16]</sup>提出了 ResNet 网络模型, 在 ImageNet 数据集中实现了 3.6% 的错误率, 低于人类的平均识别水平。

## 2 图像识别的深度学习网络模型

常见的深度学习网络模型包括卷积神经网络、递归神经网络、以及生成对抗网络等。在图像识别技术中, 卷积神经网络(Convolutional Neural Network, CNN)的应用最广泛。

CNN 网络模型是一种用来处理具有网络结构数据的神经网络。它具有多层次结构, 包括卷积层、池化层以及全连接层, 如图 2 所示。首先将预处理后的图片输入到 CNN 网络后, 经过多个卷积池化操作, 提取图片的特征, 最后将图片特征送入全连接网络完成图像的分类识别。

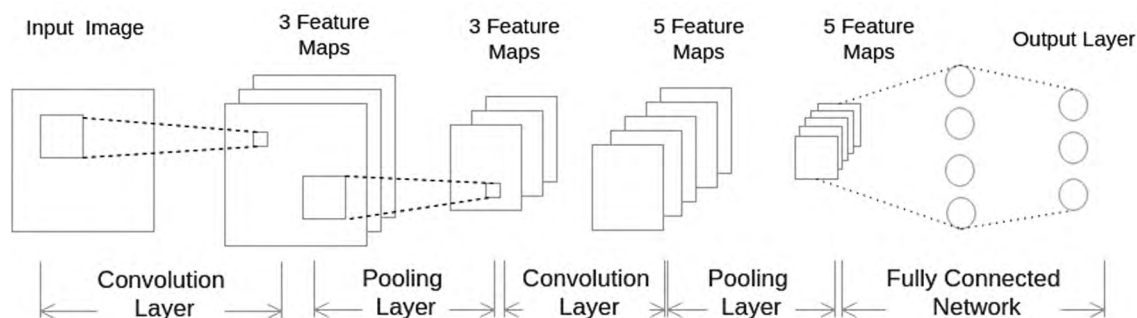


图 2 CNN 网络结构示意图

(1) 卷积层。卷积层的主要作用是生成图像的特征数据, 它的操作主要包括窗口滑动以

及局部关联两个方面。窗口滑动即通过卷积核在图像中滑动,与图像局部数据卷积,生成特征图;局部关联即每一个神经元只对周围局部感知,综合局部的特征信息得到全局特征。卷积操作后,需要使用 RELU 等激励函数对卷积结果进行非线性映射,保证网络模型的非线性。

(2) 池化层。池化层是对特征数据进行聚合统计,降低特征映射的维度,减少出现过拟合。池化的方法有最大池化和均值池化两种,根据检测目标的内容选择池化方法。最大池化的主要作用是对图片的纹理特征进行保留提取,而均值池化主要是对图片的背景特征进行提取。为了使学习到的数据特征更加全局化,数据会经过多层卷积池化操作,再输入到全连接层。

(3) 全连接层。全连接层会将池化后的多组数据特征组合成一组信号数据输出,进行图片类别识别。

### 3 基于深度学习的图像识别算法

#### 3.1 R-CNN 算法(Region Convolutional Neural Network)

一张图片中同时存在多个类别目标时,识别图片是就需要从背景中分离出感兴趣的目标,再确定该目标的类别和位置。传统的方法是使用滑窗法处理输入图片,以窗口的形式对图片遍历分割,最后将分割后的图片送入 CNN 网络模型识别分类。但是滑窗法是对图片进行随机遍历后选取目标区域,效率较低。

2014 年,Girshick 等人<sup>[17]</sup>提出 R-CNN 算法,该算法在 CNN 网络的基础上增加了选择性搜索操作来确定候选区域。算法首先对输入图片进行候选区域划分后,再通过 CNN 网络模型提取候选区域特征,进行分类识别。选择性搜索方法<sup>[14]</sup>是通过比较子区域之间的相似性并不断进行合并,得到候选区域,与滑窗法相比大大提高了搜索效率。

R-CNN 算法的结构包括:(1)候选区域:主要是从输入图片中提取可能出现物体的区域框,并对区域框归一化为固定大小,作为 CNN 网络模型的输入。(2)特征提取:将归一化后的候选区域输入到 CNN 网络模型,得到固定维度的特征输出,获取输入图片特征。(3)分类和回归:特征分类即通过图像特征进行分类,通常使用 SVM 分类器;边界回归即将目标区域精确化,可以使用线性回归方法。

2017 年,封晶<sup>[18]</sup>实现了利用 R-CNN 算法解决车辆检测问题,该算法将 Caltech1999 数据库内的图片作为训练集训练 CNN 网络模型,在 KITTI 数据集上的检测率达到了 98.62%。

R-CNN 网络存在一定的局限性,它需要在进行区域选择后将选到的区域归一化为统一的尺寸送入 CNN 网络中,导致丢失图片信息。

#### 3.2 SPP-Net 算法(Spatial Pyramid Pooling Network)

针对 R-CNN 中的问题,2014 年 He K 等人<sup>[19]</sup>提出了 SPP-Net 算法,减少了图片归一化的过程,解决了图片信息丢失以及存储的问题。

SPP-Net 算法是在 CNN 网络模型的基础上添加了 SPP 层,放在第一个全连接层前面。SPP 层可以改变卷积后的特征图像尺寸,避免了图片归一化操作。SPP-Net 算法可以将整个图片送入卷积层提取特征图像,再通过候选区域映射特征图,从而得到候选区域的卷积特征并输入到全连接网络。

2018 年,侯明伟<sup>[20]</sup>实现了用 SPP-Net 算法解决图像分类识别问题,并利用 mnist 数据集

以及 Cifar10 数据集训练预测,准确率分别为 99.54%和 84.31%。

但是,SPP-Net 算法依然存在一些问题,整个算法的训练过程是隔离的,需要存储大量的中间结果,占用磁盘空间大。

### 3.3 Fast R-CNN 算法(Fast Region Convolutional Neural Network)

2015 年,Girshick R<sup>[21]</sup>提出了 Fast R-CNN 算法,该算法引用了感兴趣区域(Region of Interest, RoI)以及多任务损失函数方法,并用 Softmax 和 SmoothLoss 代替 SVM 分类与线性回归,实现分类与回归统一,减少了磁盘空间占用,图 3 为 Fast R-CNN 算法的基础网络结构。

2018 年,潘广贞<sup>[22]</sup>等人利用 Fast R-CNN 算法实现了检测车辆的移动阴影,该算法采用 VGG16 模型对提取图像特征,使用 Pascal VOC2012 数据集训练数据。在不同光照条件下,该算法识别率均高于 94%。

### 3.4 Faster R-CNN 算法(Faster Region Convolutional Neural Network)

Fast R-CNN 算法需要对图像提取候选区域,耗费时间。为了解决这一问题,2015 年 Ren S<sup>[23]</sup>提出了 Faster R-CNN 算法,该算法采用区域建议网络(Region Proposal Network, RPN)直接对卷积后的特征图提取候选区域。Faster R-CNN 算法中,输入图片首先通过卷积池化层生成特征图像,特征图像再经过 RPN 网络生成候选区域,送入 RoI 层生成固定大小的特征向量,最后经过分类回归对图片进行识别分类。

2019 年,张汇<sup>[24]</sup>等人实现了利用 Faster RCNN 模型检测行人目标,该算法采用 K-means 算法以及 RPN 网络提取图像中行人的候选区域,并用 Softmax 进行判别分类。该算法利用 INRIA 数据集进行训练,准确率达到 92.7%。

### 3.5 YOLO 算法(You Only Look Once)

虽然 Faster R-CNN 算法识别准确率高,但是需要利用候选区域的特征进行分类识别,无法满足实时性。2015 年,Redmon J<sup>[25]</sup>提出了 YOLO 算法,该算法可以一次性地识别图片内多个物品的类别和位置,实现了端到端的图像识别。YOLO 算法首先对输入图片进行网格划分,并计算每个网格内存在目标物体的置信度以及分类概率,同时通过阈值去除没有目标物体的网格。YOLO 算法运行速度快,但是识别的准确率低。

2019 年,于秀萍<sup>[26]</sup>等人提出了改进的 YOLO 多目标分类算法。该算法在 YOLO 的基础上增加了一层卷积层,用来提取更全面的图片特征,同时使用 voc2007 数据集进行训练,平均识别率为 51%。

### 3.6 SSD 算法(Single Shot multibox Detector)

2015 年,Liu W<sup>[27]</sup>提出了 SSD 算法,该算法是一种直接预测目标类别的多目标检测算法。它首先通过 CNN 网络模型提取特征图,再对特征图进行回归分类。同时,SSD 算法增加了多

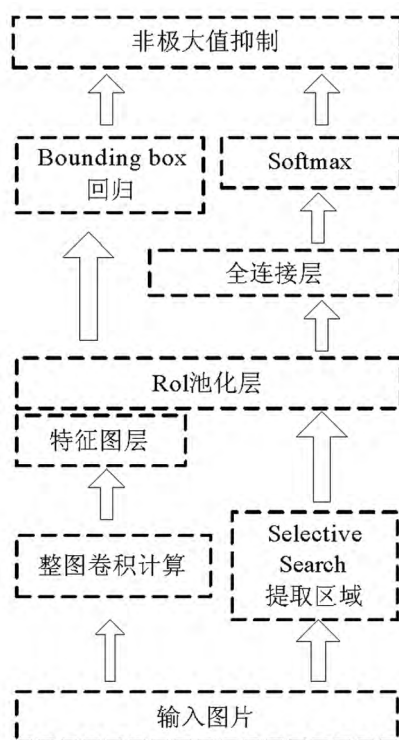


图 3 Fast R-CNN 算法模型

尺度特征图功能,能够在不同层次的特征图上回归不同尺寸的候选框,检测不同大小的目标,提高了识别准确度。

姚红革<sup>[28]</sup>等人通过 SSD 算法实现了刑侦图像的目标检测,该算法在 SSD 算法的基础上添加了卷积层,实现对大目标图像以及小目标图像分别进行特征提取,提高了对小目标物体的识别效率。该算法对于刑侦图像的识别精度为 94.8%,并且识别速率高。

表 1 图像识别算法对比

算法名称	主要特点	存在问题	识别率
R-CNN <sup>[17]</sup>	Selective Search+CNN+SVM	需要对候选区域进行归一化	66%(mAP)
SPP-Net <sup>[19]</sup>	Selective Search+CNN+SPP+SVM	在训练过程中大量的中间结果需要存储	63.1%(mAP)
Fast R-CNN <sup>[21]</sup>	Selective Search+CNN+ROI	对候选区域提取时耗费大量时间	70%(mAP)
Faster R-CNN <sup>[23]</sup>	RPN+CNN+ROI	运行速度慢,无法实现实时性	73.2%(mAP)
YOLO <sup>[25]</sup>	基于深度学习回归方法	运行速度快,但目标检测精确度低	57.9%(mAP)
SSD <sup>[27]</sup>	YOLO+Proposal+多尺度	——	72.1%(mAP)

4 深度学习在图像识别中的应用

目前,深度学习技术在图像识别领域有了长足的发展,各种应用层出不穷,如人脸识别、车牌识别、医学影像识别等不同行业。

4.1 人脸识别

人脸识别技术是根据脸部特征信息确定人类身份的一种图像识别技术,目前已经应用到各个领域。基于深度学习的人脸识别技术是人脸识别中的研究热点,为人脸识别领域的发展提供了动力。

2014 年,Taigman Y 等人<sup>[29]</sup>构建了 DeepFace 网络模型,该模型由 8 层网络结构构成,包含 5 层卷积层、1 层池化层和 2 层全连接层。该模型训练了 4 百万张人脸图像,并在 LFW 数据集上预测,准确率为 97.35%。LFW 数据集是目前常用的人脸图像数据集,包含 13233 幅图像,5749 个人。

2015 年,Zhou E 等人<sup>[30]</sup>构建了 Face++ 网络模型,该模型是包含 10 个卷积层的 CNN 网络。同时,Zhou E 等人构建了包含 5 百万张的人脸图片的数据集进行训练,并在 LFW 数据集上测试,准确率为 99.5%。Sun Y 等人<sup>[31]</sup>构建了 DeepID 模型,分别对 VGG-Net 结构和 GoogLeNet 结构进行了改进,两种结构在 LFW 数据集中的准确率都高于 97.45%。Schroff F 等人<sup>[32]</sup>提出了 FaceNet 模型,该模型以 CNN 网络为基础,将人脸图像映射到欧几里得空间特征向量中,根据空间距离判断人脸图像的相似性。同时该模型训练了包含 800 万人的脸数据,在 LFW 数据集上的准确率为 99.63%。Liu J 等人<sup>[33]</sup>提出了百度模型,该模型使用多个 GPU 对切分后的面部图像同时进行识别,每个网络包含 9 个卷积层。该模型在 LFW 数据集中的准确度为 99.77%。

目前,许多学者都在研究人脸识别问题。2019 年,胡亚洲<sup>[34]</sup>等人将改进的 Inception-Res-Net-V1 网络结构应用到人脸识别问题中,该模型使用 CASIA 数据集进行训练,并在 LFW 数

据集中的准确率达到 99.22%。彭先霖<sup>[35]</sup>等人提出一种多任务学习的人脸表情识别方法,该模型在 CNN 网络模型的基础上使用双层树分类器代替 Softmax 分类器,实现了人脸表情标签与人脸标签同步分类,提高了人脸表情的识别率。同时模型利用 CK+人脸表情数据集进行训练,准确率为 96.62%。

#### 4.2 车牌识别

相比于传统的识别方法,利用深度神经网络进行车牌识别能够实现自动提取车牌图片特征,准确率高。目前,深度学习技术是车牌识别研究中使用最广泛的方法。

2018 年 6 月,陈利<sup>[36]</sup>实现了基于 Faster R-CNN 模型的车牌识别算法。该算法利用 Faster R-CNN 对车牌图像进行定位,再利用垂直投影法和模板匹配法实现车牌的字符分割以及字符识别。该算法将 16217 张车牌图片作为数据集进行训练,准确率为 95.3%。

2018 年 9 月,曹正凤<sup>[37]</sup>等人在 CNN 网络模型的基础上,构建了一个卷积层、RPN 网络、ROI 池化、全连接以及 RNN 网络的车牌识别模型,能够实现直接对车牌图像处理,省去了字符分割步骤。该模型利用构建的 3 万张车辆图片作为数据集进行训练,准确率达到 95.4%。

2018 年 10 月,李晓燃<sup>[38]</sup>提出了倾斜车牌矫正识别方法,利用 CNN 网络处理车牌识别图像,同时用霍夫变换方法对倾斜的车牌图像进行矫正识别。该模型训练了 150 张模糊倾斜车牌图片进行训练,正确率达到了 98.5%。

#### 4.3 医学图像识别

深度学习技术可以应用于医学影像领域很多方面,如 CAD(Computer Aided Diagnosis, 计算机辅助诊断)、图像分割以及图像匹配与融合等。

2014 年,Roth<sup>[39]</sup>等人采用 CNN 网络识别临床淋巴结病变,通过随机旋转采样提高了 CNN 的分类性能。该模型将 388 个纵隔淋巴结和 595 个腹腔淋巴结进行训练,预测准确率达到 83%。

2015 年,Gao<sup>[40]</sup>将 CNN 网络与 SVM 分类器结合,诊断核性白内障病变的程度。该模型训练了 5378 幅人群图像,准确率为 99.1%。

2016 年,Dou<sup>[41]</sup>提出 3D CNN 模型对脑微出血图像进行自动检测,该模型构建了飞利浦医疗系统中 SWI-CMB 数据集进行训练,准确率为 93.16%。

2017 年,斯坦福团队<sup>[42]</sup>通过 CNN 网络模型实现了对皮肤癌的识别分类,该模型整合了 129450 张临床图像作为数据集,诊断准确率超过了人工诊断准确率。

2018 年,张泽中<sup>[43]</sup>等人针对胃癌图像分类的问题,将 AlexNet 模型与 GoogLeNet 模型结合对胃癌图像进行分类。该模型通过训练大赛中的数据集,对胃癌图像识别分类准确率为 99.4%。

### 5 问题与展望

尽管深度学习已经成为了图像识别方面的主要应用技术,但是它的发展并不完善,还存在很多问题需要研究者去解决。

(1) 理论模型不完善。基于深度学习的图像识别技术在国内发展时间相对较短,国内对于模型的构建还不完善,需要更多的理论研究。在之后的研究中,可以通过各种模型的组合,

取长补短来构建合适的网络模型。

(2) 模型参数优化问题。基于深度神经网络的图像识别技术需要训练大量参数提取图片特征,这样就占用了大量的运行时间与计算机的存储内存。研究者应该改进模型结构,在保证图片识别的准确率的前提下提高模型的时间复杂度,优化模型参数。

(3) 训练数据优化问题。深度学习网络模型依赖大量的训练集进行特征提取,训练数据集不平衡甚至缺失,会极大的限制深度学习技术的应用。在以后的研究方向中应该考虑如何解决训练数据问题。

(4) 非监督学习的完善。对于有监督学习算法,需要训练数据进行大量的人工数据标注,浪费精力。之后的研究应该加强非监督学习算法的构建,解决数据标注问题。

(5) 对视频数据进行分类预测。目前,大多数算法都是基于静态图片数据进行处理分析的,对视频数据的处理较少。因此,加强对于视频数据识别检测的理论研究是一个重要的方向。

深度学习技术最大的缺点就是对于训练数据的依赖性,当训练数据不完整时,网络模型是否还能保持识别准确率是一个重要的研究方向。为了进一步完善图像识别技术,我们应该关注深度神经网络模型的构造,优化训练数据,减少模型训练参数,提高模型准确率。

#### 参考文献:

- [1] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[C]// NIPS. Curran Associates Inc. 2012.
- [2] Chris Harris, Mike Stephens, A Combined Corner and Edge Detector, 4th Alvey Vision Conference, 1988, pp147—151
- [3] Ojala T, Pietikinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions[J]. Pattern Recognition: The Journal of the Pattern Recognition Society, 1996, 29(1): 51—59.
- [4] David G. Lowe Object Recognition from Local Scale-Invariant Features. 1999
- [5] David G. Lowe Distinctive Image Features from Scale-Invariant Keypoints. January 5, 2004.
- [6] Viola P, Jones M. Rapid Object Detection using a Boosted Cascade of Simple Features[C]//null. IEEE Computer Society, 2001.
- [7] Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005, pp. 886—893.
- [8] Bay H, Ess A, Tuytelaars T, et al. Speeded-Up Robust Features (SURF)[J]. Computer Vision and Image Understanding, 2008, 110(3): 346—359.
- [9] Osuna E. Training support vector machines; an application to face detection[J]. Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1997, 1997.
- [10] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278—2324.
- [11] 董立岩, 苑森森, 刘光远, 贾书洪. 基于贝叶斯分类器的图像分类[J]. 吉林大学学报(理学版), 2007(02): 249—253.
- [12] 余肖生, 周宁, 张芳芳. 基于 KNN 的图像自动分类模型研究[J]. 中国图书馆学报, 2007(01): 74—76.
- [13] 潘琛, 杜培军, 张海荣. 决策树分类法及其在遥感图像处理中的应用[J]. 测绘科学, 2008(01): 208—211+253.
- [14] Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions[J]. 2014.
- [15] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.
- [16] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015.
- [17] Girshick, Ross, Donahue, Jeff, Darrell, Trevor, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]. //2014 IEEE Conference on Computer Vision and Pattern Recognition; 2014 27th IEEE Conference on



- Computer Vision and Pattern Recognition (CVPR 2014), 23–28 June 2014, Columbus, Ohio, 2014; 580–587.
- [18] 封磊. 基于区域卷积神经网络的车辆检测方法[J]. 科技广场, 2017(03): 10–14.
- [19] He K, Zhang X, Ren S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904–16.
- [20] 侯明伟. 基于空间金字塔池化的卷积神经网络图像分类算法[D]. 武汉大学, 2018.
- [21] Girshick R. Fast R-CNN[J]. Computer Science, 2015.
- [22] 潘广贞, 孙艳青, 王凤. 基于 Fast RCNN 模型的车辆阴影去除[J]. 计算机工程与设计, 2018, 39(03): 819–823.
- [23] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. 2015.
- [24] 张汇, 杜煜, 宁淑荣, 张永华, 杨硕, 杜晨. 基于 Faster RCNN 的行人检测方法[J]. 传感器与微系统, 2019, 38(02): 147–149+153.
- [25] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[J]. 2015.
- [26] 于秀萍, 吕淑平, 陈志韬. 基于 YOLO 算法的多类目标识别[J]. 实验室研究与探索, 2019, 38(03): 34–36+76.
- [27] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[J]. 2015.
- [28] 姚红革, 白小军, 杨浩琪. 基于 SSD 的多特征刑侦场景识别[J]. 西安邮电大学学报, 2018, 23(04): 28–33.
- [29] Taigman Y, Yang M, Ranzato M, et al. DeepFace: Closing the Gap to Human-Level Performance in Face Verification [C]// Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2014.
- [30] Zhou E, Cao Z, Yin Q. Naive-Deep Face Recognition: Touching the Limit of LFW Benchmark or Not? [J]. Computer Science, 2015.
- [31] Sun Y, Liang D, Wang X, et al. DeepID3: Face Recognition with Very Deep Neural Networks[J]. Computer Science, 2015.
- [32] Schroff F, Kalenichenko D, Philbin J. [IEEE 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)-Boston, MA, USA (2015. 6. 7–2015. 6. 12)] 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)-FaceNet: A unified embedding for face recognition and clustering[J]. 2015; 815–823.
- [33] Liu J, Deng Y, Bai T, et al. Targeting Ultimate Accuracy: Face Recognition via Deep Embedding[J]. 2015.
- [34] 胡亚洲, 周亚丽, 张奇志. 基于深度学习的人脸识别算法研究[J/OL]. 计算机应用研究: 1–7 [2019-05-14]. 陈利. 基于深度学习的车牌识别系统设计[J]. 计算机技术与发展, 2018, 28(06): 85–89.
- [35] 彭先霖, 张海曦, 胡琦瑶. 基于多任务深度卷积神经网络的人脸/面瘫表情识别方法[J]. 西北大学学报(自然科学版), 2019, 49(02): 187–192.
- [36] 陈利. 基于深度学习的车牌识别系统设计[J]. 计算机技术与发展, 2018, 28(06): 85–89.
- [37] 曹正凤, 李少丁, 王栋梁, 徐建波, 刘晗, 卓晴. 基于深度学习的端到端车牌检测识别系统[J]. 中国交通信息化, 2018(09): 88–91.
- [38] 李晓燃. 基于深度学习的倾斜车牌矫正识别[J]. 软件, 2018, 39(10): 215–219.
- [39] Roth H. R. et al. (2014) A New 2.5D Representation for Lymph Node Detection Using Random Sets of Deep Convolutional Neural Network Observations. In: Golland P., Hata N., Barillot C., Hornegger J., Howe R. (eds) Medical Image Computing and Computer-Assisted Intervention-MICCAI 2014. MICCAI 2014. Lecture Notes in Computer Science, vol 8673. Springer, Cham
- [40] Gao Xinting, Lin Stephen, Wong Tien Yin. Automatic Feature Learning to Grade Nuclear Cataracts Based on Deep Learning[J]. IEEE transactions on bio-medical engineering, 2015, Vol. 62 (11), pp. 2693–701
- [41] Qi Dou, Hao Chen, et al. Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks[J]. IEEE Trans Med Imaging. 2016 May; 35(5): 1182–1195.
- [42] Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115–118.
- [43] 张泽中, 高敬阳, 吕纲, 赵地. 基于深度学习的胃癌病理图像分类方法[J]. 计算机科学, 2018, 45(S2): 263–268.