

基于数据增强的小样本字符识别模型

乔雨轩, 方建安

东华大学信息科学与技术学院, 上海

收稿日期: 2022年4月13日; 录用日期: 2022年5月10日; 发布日期: 2022年5月18日

摘要

小样本学习是机器学习的重要研究方向, 过去的几十年里, 针对小样本学习已经展开了许多研究。针对一些银行支票原始数据积累不足、支票磁条码字符识别准确率低的问题, 提出了一种基于数据增强的小样本字符识别方法。通过原始少量样本设计了定制的数据增强方案, 有效解决了小样本冷启动的问题, 提高模型泛化能力, 实现磁条码的准确识别。经测试, 该方法的识别准确率达到95%以上。

关键词

机器学习, 小样本学习, 数据增强

Small Sample Character Recognition Model Based on Data Augmentation

Yuxuan Qiao, Jian'an Fang

College of Information Science and Technology, Donghua University, Shanghai

Received: Apr. 13th, 2022; accepted: May 10th, 2022; published: May 18th, 2022

Abstract

Small sample learning is an important research direction of machine learning. In the past decades, many researches have been carried out on small sample learning. Aiming at the problem of insufficient original data accumulation and low character recognition accuracy of magnetic bar code of check in some banks, a small sample character recognition method based on data enhancement is proposed. A customized data enhancement scheme is designed through a small number of original samples, which effectively solves the problem of cold start of small samples, improves the model

generalization ability, and realizes the accurate recognition of magnetic bar codes. Through testing, the recognition accuracy of this method is more than 95%.

Keywords

Machine Learning, Few-Shot Learning, Data Augmentation

Copyright © 2022 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

随着现代科学技术的发展, 当今社会已经进入信息化时代, 在金融行业领域, 每日传递的图形、文字等数字信息量可达到数以亿万计[1]。传统的人工处理方式面对这么大的数据信息量, 不仅效率低下, 而且不可控因素较多, 因此在金融行业内实现系统自动化处理金融票据的业务需求越来越高。实现系统自动化处理金融票据能够显著提高票据业务的处理效率, 并且有利于资金流动, 减少因人工处理而导致的金融风险, 对银行控制交易风险, 加快电子化业务进程意义重大[2]。

大数据时代的到来使基于深度学习[3]的各种方法如卷积神经网络, 在图像检测、人脸识别等计算机视觉相关领域中取得巨大成功, 通过深度学习模型来处理金融行业相关业务也成为了一个研究重点。但基于深度学习的方法很大程度上是依靠其自身足够深的网络来提取训练数据的深层特征, 是以一种数据为驱动的方法。而现实生活中, 某些应用场景只有少量原始数据, 不利于深度学习模型的训练。考虑到人类在识别某些未知事物时往往只需要学习少许的数据, 比如我们小时候看了几张猫的图片就能够准确识别出日常生活中所看到的猫, 我们也希望深度学习系统具备这样通过少量样本来学习的能力。小样本学习[4] (few-shot learning), 它要解决的问题就是如何通过少量样本快速学习。目前主流的小样本学习方法, 根据所采用方法的不同, 可以分为基于模型微调[5]、基于数据增强[6]和基于迁移学习[7]。

近年来, 越来越多的研究人员已经投入到小样本学习的研究中, 在图像相关任务中已经开发出很多性能优越的算法模型, 但也存在着一些待解决的问题[8]。经常使用的基于模型微调的方法, 虽然只需调整模型参数就可以实现, 但在样本量极少的情况下, 无论如何优化也无法学习到更健壮的数据特征, 极易过拟合。基于迁移学习的方法虽然是一个前沿方法, 但需要原始数据与迁移数据有较高的关联性才会有好的效果, 对于极少量的样本同样不适用。基于数据增强的方法主要通过数据扩充来增强数据特征, 能够解决深度学习算法数据不足的本质问题, 适合样本量很少的任务场景, 但对数据集扩充方案有较高要求, 很容易引入噪声等冗余信息, 使模型学到其他特征从而影响模型效果。

针对某些场景下金融票据原始数据积累不足, 常用的卷积神经网络等深度学习模型对含有特殊字符的磁条码识别准确率低的问题, 本文提出了一个基于数据增强的小样本字符识别模型的解决方案, 将模型整体分为定位模型和识别模型, 通过定制的数据增强方案扩充数据集, 提升了模型泛化能力, 有效实现了金融票据磁条码的准确识别。

2. 模型整体设计

下图 1 为本文的建模流程方案, 输入为任意的银行支票图片, 首先由定位模型对图片中的磁条码进行定位, 然后将定位出的磁条码由识别模型识别成文本显示出来。

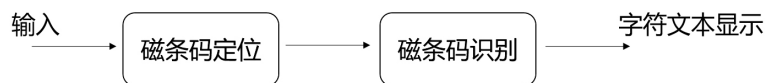


Figure 1. Modeling process design block diagram
图 1. 建模流程设计框图

下图 2 为该模型的整体方案, 具体内容将在下面各章节进行介绍。

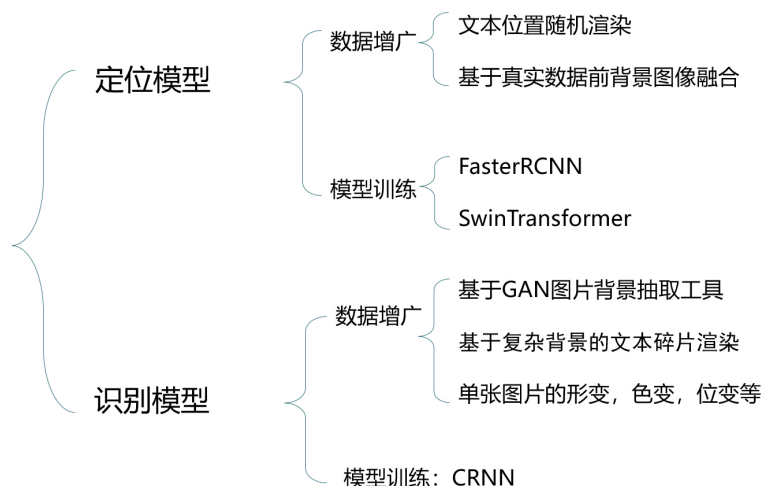


Figure 2. Overall plan
图 2. 整体方案

3. 定位模型设计

定位模型主要分为数据增强和模型训练两部分。

3.1. 定位模型数据增强方案

原始数据集只有 100 张, 因此在定位模型中分别使用两种数据增强方案文本位置随机渲染和基于真实数据前背景图像融合对数据集进行扩充。

3.1.1. 文本碎片随机渲染

文本碎片随机渲染主要通过 Text Renderer 工具实现, Text Renderer 是 GitHub 上的一个开源项目, 能够对文本内容随机组合, 结合模糊、倾斜、透视变换和加噪声、加背景等数据增强方法, 生成接近真实场景中的文字图片, 生成文本的数量、字体、大小和风格可控, 速度快, 能够用于生成用于训练深度学习模型的文本图像。

使用 Text Renderer 生成文本图像步骤分为数据准备和配置参数两步。

1) 数据准备

需要准备的有文本内容、字符集 TTF 文件、背景图。首先需要确定原始数据集样本中磁条码字符集的内容和格式, 下图 3 为原始数据集其中一个样本的磁条码字符集。

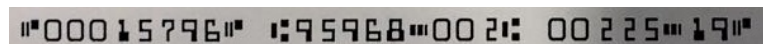


Figure 3. Magnetic barcode character set
图 3. 磁条码字符集

我国的磁条码是由 14 个字符组成, 分别为 0~9 和四个特殊符号。采用的字体格式为 E-13B, 根据字体格式可以下载对应的字符集 TTF 文件。最后还需要准备背景图像, 这里使用原始数据集中的背景样式。下图 4 为磁条码中的特殊符号。



Figure 4. Four special characters in magnetic bar code
图 4. 磁条码中的 4 种特殊字符

2) 参数配置

参数主要控制生成图片的数量、生成文本的长度以及字体的大小, 还可以通过参数添加模糊、倾斜等数据增强方式。

文本图像具体生成过程如下图 5 所示。

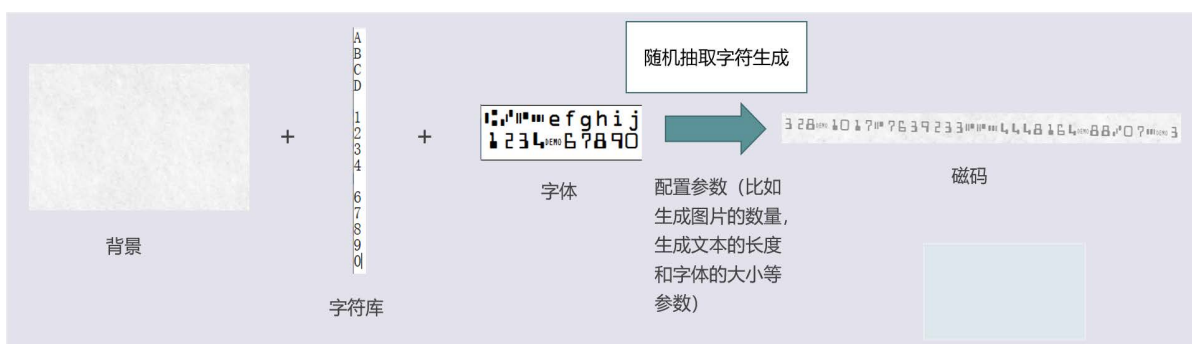


Figure 5. Text image generation process
图 5. 文本图像生成过程

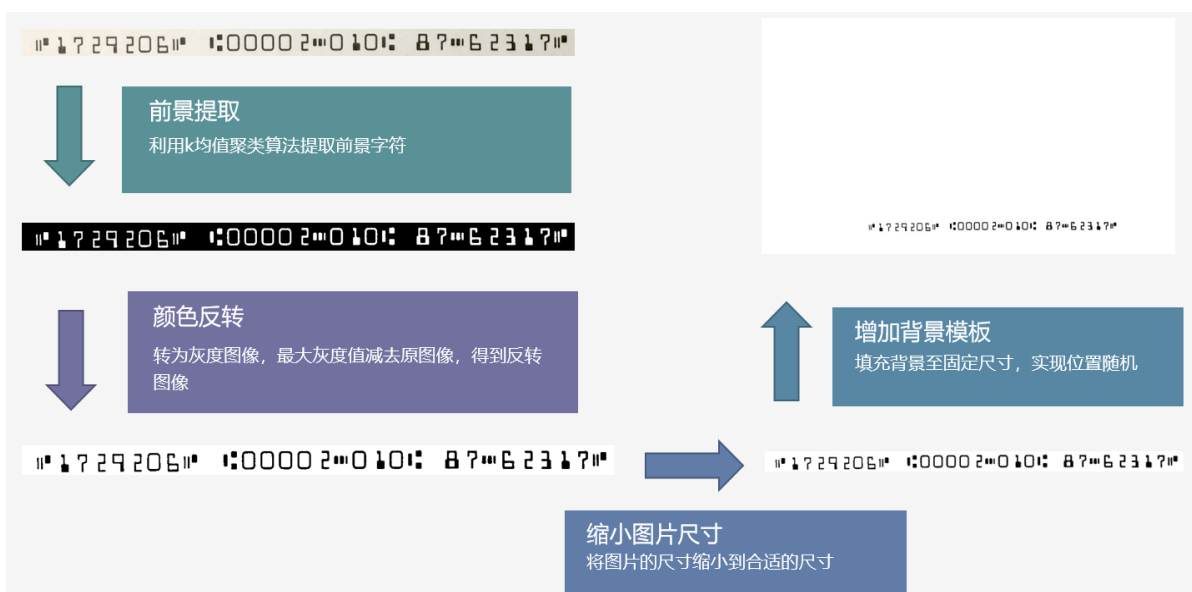


Figure 6. ImageFusion improvements
图 6. ImageFusion 改进方案

3.1.2. 基于真实数据前背景图像融合

基于真实数据前背景图像融合主要通过 ImageFusion 实现, ImageFusion 是一种图像融合算法, 在遥感探测、安全导航、医学图像分析等计算机视觉领域应用十分广泛。本文创建了基于图像融合方法的工具包, 使用真实的磁码和真实的背景进行组合生成数据集。但用此方法生成的部分图片融合的痕迹太过于明显, 会影响定位模型训练的效果, 因此本文对 ImageFusion 方法进行了改进, 改进方案由前景提取、颜色反转、缩小图片尺寸、增加背景模板四个步骤, 如图 6 所示。

首先利用 k 均值聚类[9]算法提取出磁条码的前景字符, 将磁条码图像从 RGB 空间转换到 HSV 空间下, 然后初始化三个聚类中心进行 K-means 聚类的迭代算法, 确定磁条码的位置, 提取出前景, 然后将其二值化转为灰度图像。用最大灰度值减去原图像, 得到反转图像, 再将图片的尺寸缩小到合适的尺寸, 填充背景至固定尺寸, 实现位置随机, 下图 7 为改进前后生成图片的对比, 可以看出改进后生成的图片能够符合数据集的要求, 模型也能很好的检测出磁条码的位置。



Figure 7. ImageFusion before and after improvement
图 7. ImageFusion 改进前后效果图

综合运用如上数据增强方案, 形成三套定位模型数据集, 如表 1 所示。

3.2. 定位模型训练

对于复杂的字符识别, 首先需要对字符位置进行定位检测, 本文主要使用两种检测网络对字符进行定位训练, 分别为 Swin Transformer 和 Faster R-CNN。

Swin Transformer 作为最新的 Transformer 模型, 在检测任务上展现了十分优秀的结果, 图 8 为原论文 Swin Transformer 的模型结构。Swin Transformer 不仅继承了 Transformer 关注全局自注意力的特点, 而且在此基础上对其改进, 采用了分层级设计与滑窗操作, 解决了 Transformer 全局自注意力计算量复杂的问题[10]。如图 8(a)所示, Swin Transformer 网络主要分为 4 个 stage, 输入图片首先会经过 patch partition 处理, 使像素矩阵变为 $H/4 \times W/4 \times 48$ 的三维矩阵, 然后在第一个 stage 部分, 经过 linear embedding 处理使 patch 的特征维度变成 C, 进入 Swin Transformer Block, 在 stage2 至 4 部分操作是相同的, 先经过 Patch Merging 降维, 节省计算量, 每次将分辨率降低到原来的 1/4, 最后送入 Swin Transformer Block。

图 8(b)为 Swin Transformer Block 结构, 由两个 Transformer 模块连续组成。输入特征进入第一个 Transformer 时, 首先经过 LN (Layer Normalization, 层归一化)处理, MSA (Multi-head Self-Attention, 多头自注意力机制)处理后, 然后通过残差结构后再次经过 LN 处理, MLP (Multi-layer perceptron, 多层感知机)处理, 最后再通过残差结构进入第二个 Transformer 模块, 经由第一个 Transformer 模块中同样的流程后将特征输出。

Table 1. Locate model datasets
表 1. 定位模型数据集

方案	数据量	效果
定位数据集 1 文本碎片随机渲染 (黑白背景)	10,000	
定位数据集 2 文本碎片随机渲染 (真实背景)	10,000	
定位数据集 3 图像融合	10,000	

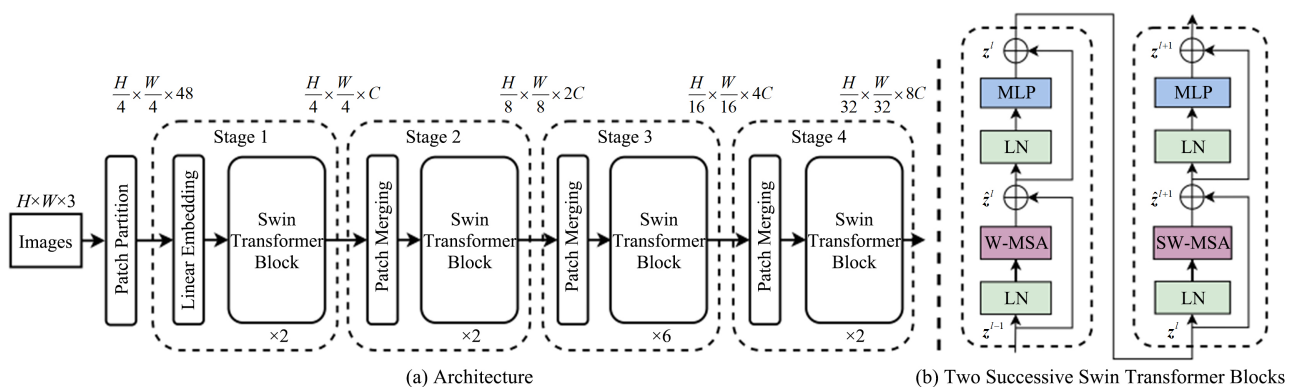


Figure 8. Swin Transformer model structure
图 8. Swin Transformer 模型结构

Faster RCNN 是目前主流的目标检测模型, 经过对 RCNN 和 Fast RCNN 的改进, 将 feature extraction, proposal, bounding box regression, classification 都整合在了一个网络中, 在精度优秀的前提下使速度得到了很大提升[11], 原论文模型结构如图 9 所示。Conv layers 包含了 13 个 conv 层, 13 个 relu 层和 4 个 pooling 层, 主要用于提取整张图片的特征; RPN 网络层的作用是生成候选定位检测框; RoI pooling 与 Classification 层主要用于对候选定位检测框进行分类, 并对候选框进行二次微调输出定位结果。

Faster RCNN 属于 two-stage 检测, 使用 Faster RCNN 训练时, 首先将图片输入到特征提取网络, 由特征提取网络提取出特征图将其送入 RPN 网络训练, 生成候选框, 并判断是否包含待检测目标并进行初步回归分析, 然后将其送入全连接层对候选框进行最终的精确分类并生成检测框。其中, 数据在经过 RPN 训练时可能会得到多个候选框, 通过调整 NMS (Non-Maximum Suppression, 非极大值抑制) 的参数, 可以将得到的不符合要求的候选框去除, 将最符合要求, 准度最高的候选框送入全连接层进一步训练, 以实现类别的最终敲定和检测框的精准定位, 这也正是双步检测算法精度普遍高于单步检测算法的原因[12]。

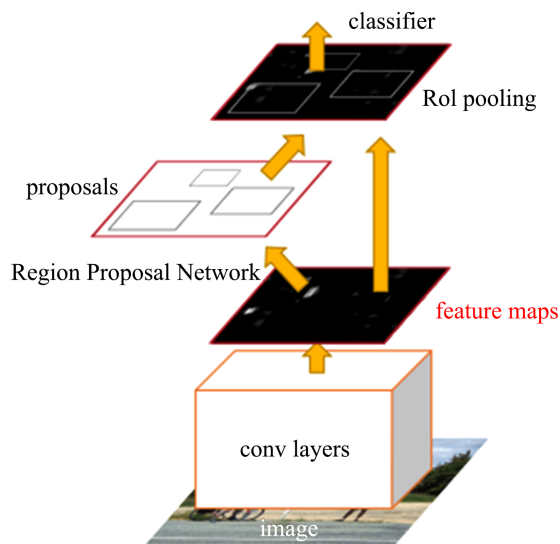


Figure 9. Faster RCNN model structure
图 9. Faster RCNN 模型结构

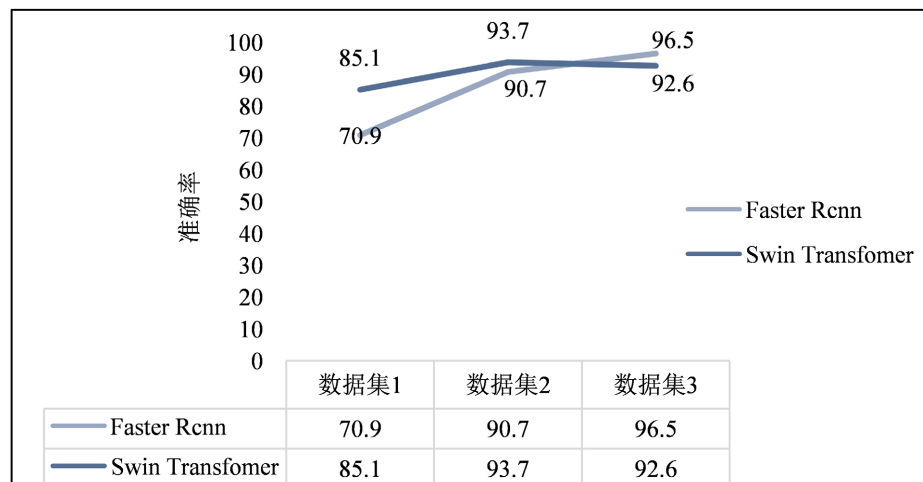


Figure 10. Comparison of training effect between Faster RCNN and Swin Transformer models
图 10. Faster RCNN 和 Swin Transformer 模型训练效果对比

本文使用 Faster RCNN 和 Swin Transformer 模型进行对比训练, 比较它们在测试集上的效果, 图 10 为两个模型在扩充的三个数据集上的训练效果, 最终选择效果较好的 Faster RCNN 作为检测模型, 在定位结果上经过处理后可以使模型在测试集上定位准确率达到 100%。

4. 识别模型设计

识别模型主要分为数据增强和模型训练两部分。

4.1. 识别模型数据增强方案

识别模型中使用如下三种数据增强方案: 基于 GAN 图片背景抽取工具, 基于复杂背景的文本碎片渲染和单张图片的形变, 色变, 位变等。

4.1.1. 基于 GAN 图片背景抽取工具

StyleText 是基于百度和华科合作研发的文本编辑算法[13], 是一种基于 GAN 图片背景抽取的工具, 通过该工具可以批量合成大量与目标场景类似的图像。StyleText 的主要框架主要包括: 1) 文本前景风格迁移模块; 2) 背景抽取模块; 3) 融合模块。文本前景风格迁移模块使用了一种骨架引导的学习机制, 能够将原图的文本风格转移到目标文本上去, 包括字体、颜色、位置以及规模大小, 即将原图文本在保留风格的情况下替换为目标文本。背景抽取模块用于去除原文本, 并以自下而上的特征融合方式填充适当的纹理。融合模块能够自动学习如何有效融合前景信息和背景纹理信息来生成编辑后的文本图像。

StyleText 的图片生成过程如下图 11 所示, 使用方便, 只要输入一个原始图片和一个文本, 就能将文本替换到原始图片中。

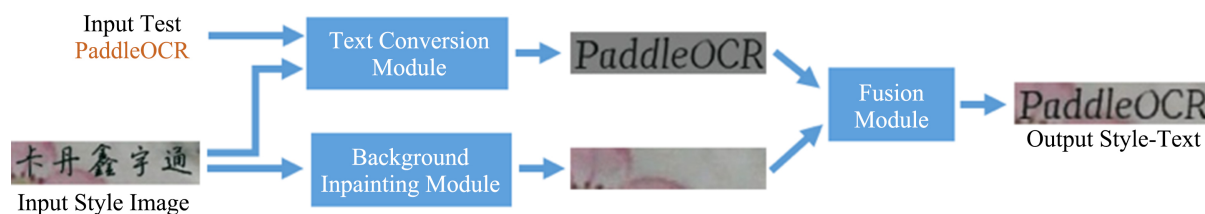


Figure 11. StyleText image generation process
图 11. StyleText 图片生成过程

StyleText 不支持特殊字符集, 直接使用会使生成的文本乱码。要想使用 StyleText, 就必须重新训练模型。分析 StyleText 原理代码后, 发现训练需要的数据复杂且耗时, 直接使用效果不好, 还需要和其他方法配合使用。

4.1.2. 基于复杂背景的文本碎片渲染

由于直接使用 StyleText 会出现乱码, 本文采用 PIL 库将特定字符直接贴在背景图片上, 这种方法简单直接, 只要准备好背景图片和文本库, 然后利用 StyleText 生成出的背景图片, 再加上纹理更复杂的数据集 VisualTexture, 就可以生成符合标准的磁条码数据集。

4.1.3. 单张图片的形变, 色变, 位变等

分别利用 paddleocr 算法随机图像增广和 ImgAug 工具进行定向增强, 总共为 6 种增强效果: 随机旋转、高斯模糊、中值模糊、均值模糊、锐化和椒盐噪点。

1) 随机旋转

随机旋转是一种比较常用的数据增强方法, 随机旋转是指以图片为中心, 围绕其旋转 1 到 359 度。

2) 高斯模糊

模糊, 就是对图像进行平滑化处理, 提取图像的重要信息。

高斯模糊就是使用高斯分布作为滤波函数, 由于图像是二维的, 因此需要二维的高斯分布函数, 公式如下:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

通过使用高斯模糊可以很好地消除图像中非椒盐类的噪声。

3) 中值模糊

中值模糊采用中值滤波的方式来去除特定的椒盐噪声, 是常见的图像去噪声与增强的方法之一, 是一种非线性平滑技术, 其原理是将每一像素点的灰度值设置为该点某邻域窗口内的所有像素点灰度值的中值。

4) 均值模糊

均值模糊使用均值滤波函数来实现, 公式如下:

$$h(x, y) = \begin{cases} \frac{1}{WH}, & |x| < \frac{W}{2}, |y| < \frac{H}{2} \\ 0, & \text{else} \end{cases} \quad (2)$$

均值滤波是典型的线性滤波算法, 通过均值滤波可以在图像上对目标像素套一个模板, 该模板以目标像素为中心的周围 8 个像素, 构成一个滤波模板, 即去掉目标像素本身, 再用模板中的全体像素的平均值来代替原来像素值。

5) 锐化

锐化与模糊是相反的操作, 锐化是通过增强高频分量来减少图像中的模糊, 增强图像细节边缘和轮廓, 增强灰度反差, 便于后期对目标的识别和处理。锐化处理在增强图像边缘的同时也增加了图像的噪声。本文使用高通滤波抑制低频分量, 从而达到图像锐化的目的, 公式如下:

$$g(\mu, \nu) = \sum_x \sum_y F(x, y) H(\mu - x + 1, \nu - y + 1) \quad (3)$$

6) 椒盐噪点

在图像中增加适量噪声可以增强学习能力, 椒盐噪声为一种随机出现的白点或者黑点, 可能是亮的区域有黑色像素或是在暗的区域有白色像素。

下图 12 为使用以上六种定向增强方式的效果。



Figure 12. Directional enhancement
图 12. 定向增强效果

综合使用如上基于 GAN 背景提取、基于复杂背景文本碎片渲染、单张图片的形变、色变、位变、训练算法随机增广等方案形成的四套识别模型数据集, 如下表 2 所示。

Table 2. Identify model datasets

表 2. 识别模型数据集

	方案	数据量
识别数据集 1	GAN 图片背景 + 文本碎片渲染 + 训练随机增强	9772
识别数据集 2	GAN 图片背景 + 文本碎片渲染 + 训练增强优化	9772
识别数据集 3	GAN 图片背景 + 文本碎片渲染 + 合成数据定向增广	9772 (包含合成数据定向增广 1000 张)
识别数据集 4	GAN 图片背景 + 文本碎片渲染 + 真实数据定向增广	9772 (包含合成数据定向增广 1000 张, 真实数据定向增广 324 张)

4.2. 识别模型训练

在定位模型检测到磁条码字符位置后, 就需要识别模型对其进行文本识别。本文选择 CRNN 作为识别模型训练网络, 全称为 Convolutional Recurrent Neural Network, 网络结构如下图 13 所示, 它继承了卷积神经网络 CNN 和循环神经网络 RNN 的优点, 能够获取不同尺寸的输入图像并对其长度进行预测[14]。它直接在中文字符级别上运行。此外, 由于 CRNN 放弃了传统神经网络中使用的完全连接的层, 因此它导致了更加紧凑和有效的模型, 所有这些特性使 CRNN 成为基于图像的序列识别的绝佳方法[15]。在具体的应用阶段, CNN 用于完成文本图像的特征提取, RNN 依靠 Bi-LSTM 融合特征向量, 再提取文字序列特征, 最后由转录层 CTC 输出序列的概率分布, 借此来预测最优的文本序列值[16]。

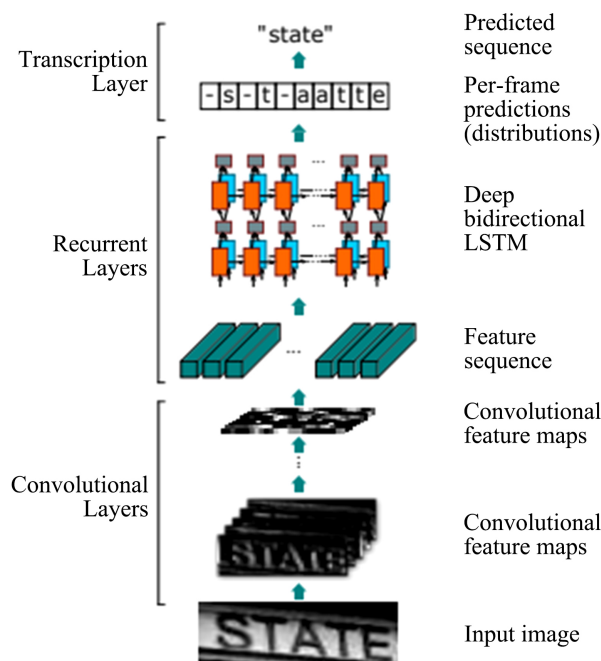


Figure 13. CRNN network structure

图 13. CRNN 网络结构

5. 实验结果与分析

实验采用 CRNN 对原始数据集和不同方案增强后的数据集分别训练网络进行识别准确率对比, 下表 3 为 CRNN 在不同数据集的识别结果。

Table 3. Identify model datasets

表 3. CRNN 在不同数据集的识别结果

方案		CRNN 识别准确率(%)
原始数据集		38.81
识别数据集 1	GAN 图片背景 + 文本碎片渲染 + 训练随机增强	89.90
识别数据集 2	GAN 图片背景 + 文本碎片渲染 + 训练增强优化	87.51
识别数据集 3	GAN 图片背景 + 文本碎片渲染 + 合成数据定向增广	79.83
识别数据集 4	GAN 图片背景 + 文本碎片渲染 + 真实数据定向增广	96.37

由表 3 可知, 由于原始数据只有 100 张, 数量太少, 导致模型训练效果很差, 只有 38.31% 的识别准确率。经过数据增强方法扩充得到的四个识别数据集相较原始数据集识别准确率都得到有效提升, 分别提升了 51.09%、48.7%、41.02%、57.56%, 充分证明本文方法对提升模型识别准确率的有效性。为进一步验证本文数据增强方案的有效性, 本文采用 CRNN、FasterRCNN + DCNN 和 SVM 对不同方案增强后的数据集和原始数据集进行识别效果对比。FasterRCNN + DCNN 为文献[17]所提方法, 使用 FasterRCNN 与 DCNN 相结合的双网络模型进行票据中字符的定位识别。SVM 为文献[18]所提, 使用传统的 SVM 对小样本字符集进行识别。由于 SVM 提取特征较慢, 识别过程较长, 不适用于大样本数据集, 因此本次对比实验在不同方案增强后的数据集中分别随机抽取 1000 张作为训练集, 1000 张作为测试集, 下表 4 为实验结果。

Table 4. Recognition results of different methods on recognition data sets

表 4. 不同方法在识别数据集上的识别结果

	CRNN (%)	FasterRCNN + DCNN (%)	SVM (%)
原始数据集	38.81	17.87	41.91
识别数据集 1	81.17	75.90	67.71
识别数据集 2	80.73	71.51	63.12
识别数据集 3	74.61	67.83	59.23
识别数据集 4	91.10	81.37	76.38

可以看出, FasterRCNN + DCNN 和传统 SVM 方法在经过数据增强后的数据集上的识别准确率相比原始数据集都得到大大提升, 充分说明了本文数据增强方案的有效性。并且相比 FasterRCNN + DCNN、传统 SVM 方法, 本文的方法在效果最好的识别数据集 4 上识别准确率提高了 9.73% 和 14.72%, 进一步验证了本文方法的有效性。

6. 总结

本文通过综合尝试 6 个数据增强方案(2 大定位数据增强, 4 大识别数据增强), 将数据集 20 张扩展到近 10,000 张, 通过数据增强技术扩充训练数据集, 有效解决了因数据不足造成模型精度不高的问题, 实现了识别准确率从 38.81% 到 96.37% 的大幅提升。此外, 本文所采用的数据增强方案如图像色彩恢复、文本碎片渲染、图像前后背景融合等适用于在大多数通用的视觉任务中, 为提升建模效率和精度提供保障。

参考文献

- [1] 徐伟伟. 基于光学字符识别技术的支票识别系统设计[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2015.
- [2] 陈楠. 支票磁码自动识别技术研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2014.
- [3] 孙志军, 薛磊, 许阳明, 王正. 深度学习研究综述[J]. 计算机应用研究, 2012, 29(8): 2806-2810.
- [4] Jankowski (2011) *Meta-Learning in Computational Intelligence*. Springer, Berlin. <https://doi.org/10.1007/978-3-642-20980-2>
- [5] 李锦明, 曲毅, 裴禹豪, 宸泽江. 预训练卷积神经网络模型微调的行人重识别[J]. 计算机工程与应用, 2018, 54(20): 219-222+229.
- [6] 高友文, 周本君, 胡晓飞. 基于数据增强的卷积神经网络图像识别研究[J]. 计算机技术与发展, 2018, 28(8): 62-65.
- [7] 刘鑫鹏, 栾悉道, 谢毓湘, 黄明哲. 迁移学习研究和算法综述[J]. 长沙大学学报, 2018, 32(5): 28-31+36.
- [8] 赵凯琳, 靳小龙, 王元卓. 小样本学习研究综述[J]. 软件学报, 2021, 32(2): 349-369.
- [9] 刘靖明, 韩丽川, 侯立文. 基于粒子群的 K 均值聚类算法[J]. 系统工程理论与实践, 2005(6): 54-58.
- [10] Liu, Z., Lin, Y.T., Cao, Y., Hu, H., Wei, Y.X., Zhang, Z., Lin, S. and Guo, B.N. (2018) SWIN Transformer: Hierarchical Vision Transformer Using Shifted Windows.
- [11] 刘成, 李正辉, 高基豪. 基于深度学习的银行卡号识别研究与应用[J]. 湖南邮电职业技术学院学报, 2020, 19(4): 35-38.
- [12] 王璐, 王雷欧, 王东辉. 基于 Faster-rcnn 的水下目标检测算法研究[J]. 网络新媒体技术, 2021, 10(5): 43-51+58.
- [13] Wu, L., Zhang, C.Q., Liu, J.M., Han, J.Y., Liu, J.T., Ding, E. and Bai, X. (2019) Editing Text in the Wild. *Proceedings of the 27th ACM International Conference on Multimedia*, Nice, 21-25 October 2019, 1500-1508. <https://doi.org/10.1145/3343031.3350929>
- [14] Shi, B., Xiang, B. and Cong, Y. (2016) An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **39**, 2298-2304. <https://doi.org/10.1109/TPAMI.2016.2646371>
- [15] 黄子涵, 黄豪, 林丹盈, 曹洋瑞, 张泽填, 肖振球. 基于 CRNN 的自然场景多语言文本检测方法[J]. 电脑知识与技术, 2022, 18(1): 98-100.
- [16] 华春梦, 臧艳辉, 马伙财. 一种基于 CRNN 的车牌识别算法研究与应用[J]. 现代信息科技, 2021, 5(20): 78-81+86.
- [17] 郑祖兵, 盛冠群, 谢凯, 唐新功, 文畅, 李长晟. 双网络模型下的智能医疗票据识别方法[J]. 计算机工程与应用, 2020, 56(12): 141-148.
- [18] 刘明, 王命延. 支持向量机在小字符集识别中的应用[J]. 计算机与现代化, 2006(12): 7-9.