文章编号：1003-1421(2026)05-0184-11
引用格式：周通. 基于改进DB与SVTR模型的铁路集装箱箱号智能识别研究[J]. 铁道运输与经济，2026，48(5)：184-194.
ZHOU Tong. Research on Intelligent Recognition of Railway Container Numbers Based on Improved DB and SVTR Models[J]. Railway Transport and Economy，2026，48(5)：184-194.

基于改进DB与SVTR模型的铁路集装箱箱号智能识别研究

周 通

(中国铁道科学研究院集团有限公司 电子计算技术研究所，北京 100081)

摘 要：针对复杂场景下铁路集装箱箱号识别准确率低及常规深度学习模型计算复杂度较高等问题，结合现场业务规则提出铁路集装箱箱号智能识别的改进DB与SVTR模型。首先，在自定义数据集构建阶段，采用数据增强技术生成含破损、遮挡、模糊及夜间场景的多样化箱号数据集，以增强模型泛化能力。然后，在DB检测模型中FPN后加入HAFF模块，增强关键区域的特征表达能力，并结合跨尺度自注意力提升复杂背景下文本的检测能力。其次，使用标准余弦退火法和Warmup预热机制作为学习率更新策略改进SVTR识别模型，并通过箱号编码规则对识别结果进行校验修正。最后，通过自定义数据集对比实验表明，改进后模型的箱号检测精确率达到99.0%，箱号识别准确率达到95.9%。同时，改进SVTR模型的识别准确率和平均耗时均比基本SVTR模型、CRNN模型、Rosetta模型等表现更佳。

关键词：铁路集装箱箱号；深度学习；文本检测；文本识别；DB与SVTR模型
中图分类号：U294.3 **文献标识码**：A **DOI**：10.16668/j.cnki.issn.1003-1421.20250828001

Research on Intelligent Recognition of Railway Container Numbers Based on Improved DB and SVTR Models

ZHOU Tong

(Institute of Electronic Computing Technology, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China)

Abstract: Given the problems of low accuracy of railway container number recognition in complex scenarios and high computational complexity of conventional deep learning models, improved DB and SVTR models for intelligent recognition of railway container numbers based on on-site business rules were proposed. Firstly, in the custom dataset construction stage, data augmentation technologies were adopted to generate a diversified container number dataset containing damaged, occluded, blurred, and nighttime scenarios to enhance the generalization ability of the model. Then, the HAFF module was added after the FPN in the DB detection model to enhance the feature expression ability of key regions, and cross-scale self-attention was combined to improve the text detection ability in complex backgrounds. Secondly, the standard cosine annealing method and Warmup mechanism were used as learning rate update strategies to improve the SVTR recognition model, and the recognition results were verified and corrected through container number encoding rules.

收稿日期：2025-08-28 **修订日期**：2026-03-02
基金项目：中国国家铁路集团有限公司科技研究开发计划课题(P2024X003)

Finally, comparative experiments on custom datasets show that the container number detection precision of the improved model reaches 99.0%, and the container number recognition accuracy reaches 95.9%. Meanwhile, the improved SVTR model performs better than the basic SVTR model, CRNN model, and Rosetta model in terms of recognition accuracy and average time consumption.
Keywords: Railway Container Number; Deep Learning; Text Detection; Text Recognition; DB and SVTR Model

0　引言

铁路集装箱运输具有单位运量成本低、成本稳定性高、大运量适配、时效稳定性强、广域覆盖能力强、低碳环保等特点，发展铁路集装箱运输有助于减少交通运输碳排放，促进经济社会可持续发展。铁路集装箱运输信息化与智能化发展的前提是全面、准确掌握集装箱位置、状态等，关键是对全运输过程的集装箱箱号进行自动识别。

早期的集装箱箱号自动识别采用条形码技术，因存储量有限而被二维码技术替代，通过射频识别(Radio Frequency Identification，RFID)[1]电子标签传输数据实现自动识别，然而RFID系统存在标准不统一、成本较高和后台兼容性等问题，导致推广受限[2]。而通过使用计算机视觉技术实现集装箱箱号识别，由于其成本较低、部署更方便而得到广泛应用。早期通过形态学操作和连通区域分析等方法[3-6]获取箱号特征，再结合光学字符识别(Optical Character Recognition，OCR)技术将字符特征转换为箱号，但这种方法比较依赖图像质量与光照条件。

深度学习技术的快速发展在相当程度上提升了集装箱箱号自动识别的性能与效果。Jaderberg等[7-8]较早地将卷积神经网络用于自然场景文本检测与字符分类，为后续研究奠定理论基础。Ren等[9]提出的快速区域卷积神经网络(Faster Region-Based Convolutional Neural Networks，Faster R-CNN)框架，提高了检测效率。Wang等[10]通过多尺度特征融合策略，将模型的正确识别率提升至90.4%。Zhong等[11-12]通过引入区域候选网络(Region Proposal Network，RPN)来改进Faster R-CNN目标检测模型，进而提高复杂场景下算法的鲁棒性。崔循等[13-14]通过增加空间注意力机制来改进Faster R-CNN，既保证了检测精准度又兼顾检测速度。为满足检测实时性需求，Liu等[15]提出单次多框目标检测(Single Shot MultiBox Detector，SSD)算法。也有文献针对文本检测研究专门神经网络模型，如连接主义文本提议网络(Connectionist Text Proposal Network，CTPN)[16]，将文本行拆分为宽度固定的文本片段，再连接成完整文本行，能有效处理长文本；高效且准确的场景文本检测网络(Efficient and Accurate Scene Text Detector，EAST)[17]直接预测文本行的几何形状，无需候选区域生成，推理速度快；渐进式尺度扩展网络(Progressive Scale Expansion Network，PSENet)[18]采用分割的方法，结合渐进尺度扩展算法能精准分离相邻或重叠文本实例；可微分二值化(Differentiable Binarization，DB)文本检测网络[19]通过动态阈值图学习自适应二值化过程，提升对模糊、低对比度文本的检测精度。

此外，张添添[20]基于改进YOLO v3模型和自注意力机制，实现集装箱箱号定位和识别。李思佳等[21]设计开发采用智能相机配备红外自动补光功能的集装箱箱号识别模块，运用图像深度学习识别算法，通过调整相机曝光增益避免昼夜光感强度不同，提高昼夜工况下的箱号识别的准确率。钟彦鹏[22]通过深度学习算法对集装箱箱号进行定位、分割与识别。部分学者结合多种技术提高箱号识别准确率，涂昊等[23]通过减少原YOLO v3的网络完成箱号定位网络，并利用自注意力机制等方法完成箱号识别。李莉莉[24]基于YOLO v4的集装箱编号检测算法和字符识别的模板匹配算法，提高整个识别的速度和准确率。苗凌玮[25]通过将传统的最大稳定极值区域算法(Maximally Stable Extremal Regions，MSER)和EAST模型结合实现集装箱箱号检测，并采用基于字符分割的LeNet-5模块和卷积循环神经网络(Convolutional Recurrent Neural Network，CRNN)模块共同完成箱号识别。李文晖等[26]通过图像识别与读取RFID标签方式提高箱号识别准确率。肖健雄等[27]采用卷积神经网络(Convolutional Neural Networks，CNN)对箱号进行智能识别，并通过图像哈希技术与相似度检验算法相结合进行异常检测。雷宁昊

等[28]采用MobileNet V3作为检测骨干网络对集装箱号检测，再采用MobileNet V1 Enhance作为模型骨干网对箱号进行识别。

将深度学习技术应用于箱号识别已取得较好的效果，但仍面临以下挑战：一是现有方法对光照不均、箱体污损、反光或阴影干扰等复杂场景的适应性较弱，容易导致字符误检或漏检，识别效果与实际还有较大差距；二是主流深度学习模型计算复杂度高，多依赖高性能计算设备，难以满足现场室外作业管理的实时性和便携性的需求，无法支持现场作业过程自动化与智能化管理。针对上述问题，本研究提出采用改进DB与基于Swin Transformer的文字识别(Swin Transformer for Text Recognition，SVTR)模型相结合的铁路集装箱箱号智能识别方法。首先，根据公开数据集与现场采集图片，通过数据增强与标注方法，制作铁路集装箱检测与识别数据集。然后，为了提升复杂背景下提取箱号区域特征的能力，提出基于改进DB模型的集装箱箱号检测模型，通过在基础DB模型的特征金字塔网络(Feature Pyramid Network，FPN)结构中增加混合注意力特征融合(Hybrid Attention Feature Fusion，HAFF)模块，增强不同尺度特征的表达能力，提高检测模型在复杂背景下的鲁棒性。其次，为了提高识别的准确率，提出基于改进SVTR模型的集装箱箱号识别模型，通过动态调整学习率策略与正则化参数协同优化模型性能，更好地平衡训练速度与稳定性，提升模型整体稳定性。

1　铁路集装箱数据集自定义数据集制作

1.1　铁路集装箱图片构成

通用集装箱箱号与铁路专用集装箱箱号均满足ISO 6346标准要求，都是由11位英文字母和数字构成，依次为：“4位英文字母”“6位数字”“1位校验码”。为了提高模型的识别准确率和泛化能力，通过公开数据集和现场采集得到的集装箱箱号图片共同构成本研究的数据集，涵盖了集装箱箱号在不同位置的分布情况，集装箱箱号不同位置如图1所示。

1.2　检测数据集制作

（1）数据增强。为了模拟集装箱箱号破损、遮挡、模糊及黑夜条件下的情况，对箱号图片进行模糊、反光、亮度等增强处理，通过增加训练数据的规模、多样性和复杂性，提升模型的泛化能力和鲁棒性，集装箱箱号图片增强后部分效果如图2所示。

（2）数据标注。对于数据增强后的集装箱图片，采用PPOCRLabel工具对箱号区域进行四边形框标注，铁路集装箱箱号标注如图3所示。标注完成后进行结果导出，集装箱检测数据标签如图4所



a　集装箱箱体侧面箱号　　b　集装箱箱尾箱号
a　Container side number　　b　Container rear number

图1　集装箱箱号不同位置
Fig.1　Different positions of container numbers

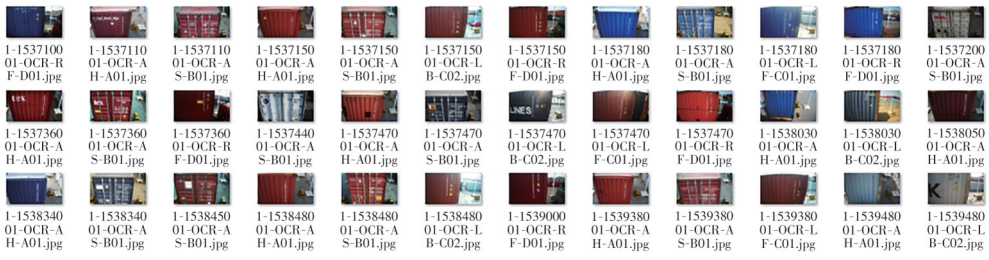


图2　集装箱箱号图片增强后部分效果
Fig.2　Partial effects of enhanced container number images

示，文本中的每行对应图片中一个箱号文本框的标注信息，由2个部分组成，前段为图像文件的名称，后段为经过json.dumps编码的标注内容，标注信息中的坐标值按照顺时针方向排列，依次表示文本框的左上角、右上角、右下角和左下角4个顶点的位置信息。



图3　铁路集装箱箱号标注

Fig.3　Railway container number annotation

1.3　识别数据集制作

（1）数据增强。为了模拟箱号破损、遮挡、模糊及黑夜条件下的情况，对箱号图片进行数据增强，通过增加训练数据的多样性和复杂性，提升模型的泛化能力和鲁棒性，集装箱箱号图片增强示例如图5所示。

（2）文本标注。铁路集装箱图片标注及识别数据标签如图6所示。如图6a所示，采用PPOCRLabel工具对箱号文本进行四点标注，将标注的四边形导出为用于箱号识别的图片数据集，共计得到5 453张箱号图片。标注完成后将结果导出为文本格式，如图6b所示，文本的每行对应图片中一个集装箱箱号，开头为图像文件名，之后为由5位英文字母和7位数字组成的箱号信息。

2　基于改进DB模型的铁路集装箱箱号检测

2.1　改进DB检测模型

基础DB检测模型是基于像素分割的文本检测网络，主要流程如下。

（1）主干(Backbone)网络通过多层卷积和下采样提取输入图像的多尺度特征。浅层特征保留高分辨率纹理信息，能够捕捉文本的笔画细节，中层特征在语义信息和位置信息之间取得平衡，用于识别字符部件，深层特征聚焦于文本区域的整体分布，提供全局语义信息。

（2）使用FPN对主干网络提取的多尺度特征进行融合，通过自顶向下的上采样与对应层级的特征相加，增强语义信息的传递，再通过跨层级的通道

```
1-154627001-OCR-RF-D01.jpg [{ " transcription " : " TCNU8171704 " ,  " points " : [900,  51],  [979,  48],  [985,  671],  [908,  673]]}]
1-132853001-OCR-LF-C01.jpg [{ " transcription " : " DVDU2800124 " ,  " points " : [973,  219],  [1 065,  220],  [1 004,  964],  [916,  963]]}]
1-133853001-OCR-AS-B01.jpg [{ " transcription " : " TEMU7683178 " ,  " points " : [1 045,  382],  [1 640,  381],  [1 630,  484],  [1 048,  483]]}]
1-125513001-OCR-AH-A01.jpg [( " transcription " : " SEGU5845118 " ,  " points " : [1 197,  192],  [1 293,  195],  [1 221,  884],  [1 137,  883]]}]
1-154535001-OCR-LB-C02.jpg [{ " transcription " : " TCKU3614214 " ,  " points " : [884,  222],  [976,  223],  [948,  966],  [863,  967]]}]
1-145021001-OCR-AS-B01.jpg [{ " transcription " : " WWWU2706054 " ,  " points " : [653,  491],  [1 193,  504],  [1 191,  595],  [656,  579]]}]
20191030-022430-01-4-R.jpg [{ " transcription " : " BMOU4624239 " ,  " points " : [903,  204],  [1 014,  204],  [995,  1 151],  [883,  1 153]]}]
1-150453001-OCR-AS-B01.jpg [{ " transcription " : " GESU3284522 " ,  " points " : [784,  471],  [1 272,  477],  [1 268,  574],  [783,  565]]}]
1-130933001-OCR-AS-B01.jpg [{ " transcription " : " EISU9475250 " ,  " points " : [1 022,  271],  [1 565,  266],  [1 556,  373],  [1 025,  374]]}]
1-154414001-OCR-LB-C02.jpg [{ " transcription " : " BMOU6448669 " ,  " points " : [[774,  0],  [879,  0],  [847,  862],  [749,  862]]}]
1-150911001-OCR-AS-B01.jpg [{ " transcription " : " TCKU4732187 " ,,  " points " : [979,  320],  [1 301,  320],  [1 301,  382],  [979,382]]}]
```

图4　集装箱检测数据标签

Fig.4　Container detection data label



a　原图

a　Original image

b　模糊处理

b　Blur processing

c　亮度处理

c　Brightness processing

d　遮挡处理

d　Occlusion processing

e　破损处理

e　Damage processing

图5　集装箱箱号图片增强示例

Fig.5　Container number image enhancement example



a　集装箱图片标注过程示意图
a　Schematic diagram of container number image recognition annotation

1-133750001-OCR-RF-D01_crop_0.jpg　YMLU3360534
1-130038001-OCR-AH-A01_crop_0.jpg　RWLU8276286
1-152819001-OCR-LF-C01_crop_0.jpg　TEMU0013847
1-155605001-OCR-AS-B01_crop_0.jpg　UETU2410747
1-154147001-OCR-AS-B01_crop_0.jpg　MRKU7981930
1-125858001-OCR-AH-A01_crop_0.jpg　CBHU9093747
1-133122001-OCR-AH-A01_crop_0.jpg　AXIU2933280
1-153848001-OCR-AS-B01_crop_0.jpg　TCNU5458922

b　集装箱图片识别数据标签示意图
b　Schematic diagram of container number image recognition data annotation results

图6　铁路集装箱图片标注及识别数据标签

Fig.6　Image annotation and data label of railway container number recognition

拼接聚合不同尺度的上下文特征，提升模型对复杂场景的适应能力。

（3）模型通过双分支检测头(Two-Branch Head)分别输出概率图(Probability Map)和阈值图(Threshold Map)2个特征图，利用可微分二值化算法将2个特征图融合为近似二值图。

基础DB文本检测模型中FPN缺乏对特征间深层关系的建模能力，无法有效关注关键区域并抑制背景干扰，同时跨尺度特征交互也比较弱。为了提升DB检测模型在复杂背景下提取箱号区域特征的能力，在基础FPN网络结构中增加HAFF模块，达到增强不同尺度特征的表达能力和提高模型在复杂场景下的鲁棒性。基于HAFF模块的改进DB模型网络结构如图7所示。

2.2　HAFF模块

HAFF模块改进FPN网络的处理流程如图8所示，具体流程如下。

（1）基础PFN特征融合。设FPN输出的多尺度特征列表(*features_list*)为$\{P_2, P_3, P_4, P_5\}$，$P_i \in \boldsymbol{R}^{H_i \times W_i \times C}$，$i=2,3,4,5$，则特征融合计算公式为

$$fuse_features = \mathrm{Conv}_{1\times1}(P_2 \oplus P_3 \oplus P_4 \oplus P_5) \tag{1}$$

式中：*fuse_features*为融合特征；$\mathrm{Conv}_{1\times1}$为1×1卷积；⊕为上采样至相同尺寸后的拼接操作。

（2）局部特征提取。对FPN输出融合后特征图*fuse_features*进行一次3×3卷积。

$$local_features = \mathrm{Conv}_{3\times3}(fuse_features) \in \boldsymbol{R}^{H\times W\times C} \tag{2}$$

式中：*local_features*为局部特征；$\mathrm{Conv}_{3\times3}$为3×3卷积操作；H，W，C分别为特征图的高度、宽度和通道数。

（3）空间注意力计算。空间注意力(Spatial Attention，SA)计算是对*fuse_features*进行均值池化后通过一系列卷积和Sigmoid激活函数获得权重，从而增强关键区域的特征。

$$avg_pool = \mathrm{MeanPool}(f_{conv}) \in \boldsymbol{R}^{H\times W\times 1} \tag{3}$$

$$conv_sa = \mathrm{Conv}_{3\times3}(avg_pool) \tag{4}$$

$$M_s = \mathrm{Sigmoid}(conv_sa) \in \boldsymbol{R}^{H\times W\times 1} \tag{5}$$

$$sa_fetures = local_fetures \otimes M_s \tag{6}$$

式中：*avg_pool*为平均池化；MeanPool为通道维度的均值池化操作；*conv_sa*为平均池化卷积；M_s为空间

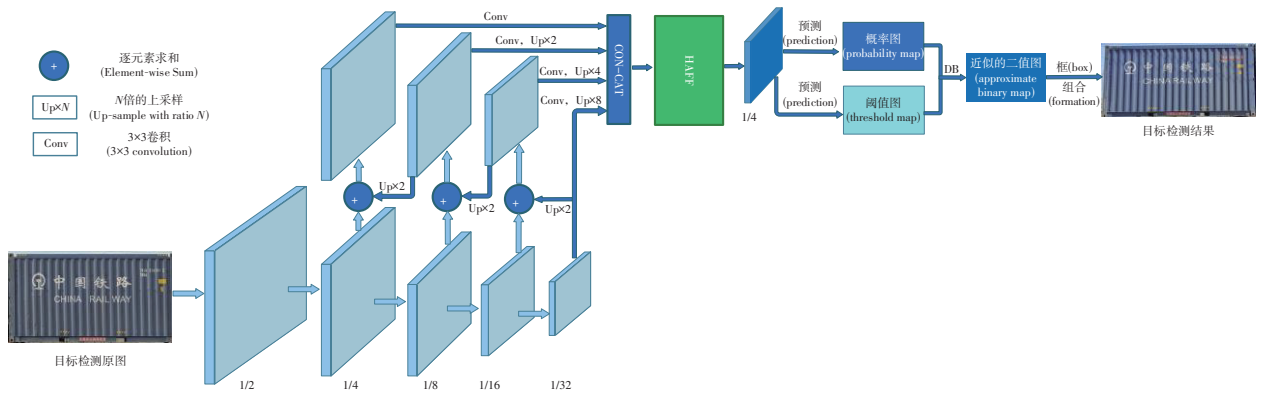


图7　基于HAFF模块的改进DB模型网络结构

Fig.7　Improved DB model network structure based on HAFF module

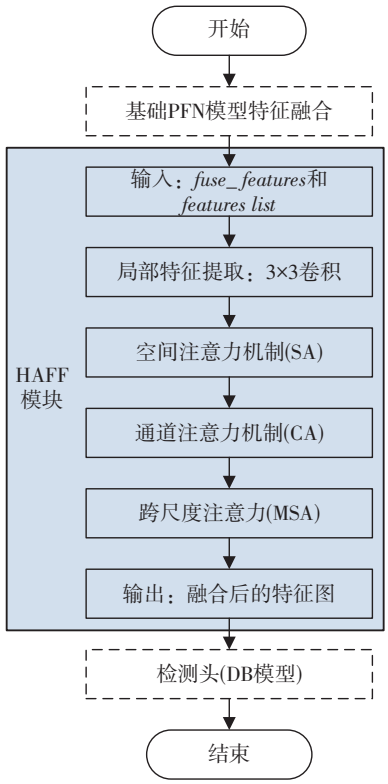


图8　HAFF模块改进FPN网络的处理流程

Fig.8　Processing flow of improved FPN network by HAFF module

注意力权重，用于突出关键区域特征；Sigmoid(·)为Logistic函数，取值范围为(0，1)；*sa_fetures*为空间增强特征；⊗为逐元素相乘。

（4）通道注意力计算。通道注意力(Channel Attention，CA)通过1×1卷积，计算通道注意力权重，将特征列表*features_list*中的每个特征图与对应的注意力权重相乘，并将结果拼接。对于特征列表*features_list*中的第$i(i=2,3,4,5)$个特征图P_i，计算公式为

$$avg_pool_{c,\,i}=\mathrm{GlobalAvgPool}(P_i)\in\mathbf{R}^{1\times1\times C_i} \tag{7}$$

$$M_{c,\,i}=\mathrm{Sigmoid}(\mathrm{Conv}_{1\times1}(avg_pool_{c,\,i}))\in\mathbf{R}^{1\times1\times C_i} \tag{8}$$

$$f_{i_ca}=P_i\otimes M_{c,\,i} \tag{9}$$

$$ca_fetures=\mathrm{Concat}(f_{2_ca},\ f_{3_ca},\ f_{4_ca},\ f_{5_ca}) \tag{10}$$

式中：$avg_pool_{c,\,i}$为第i个特征图的平均池化；GlobalAvgPool为全局平均池化；$M_{c,\,i}$为第i个特征图的通道注意力权重；f_{i_ca}为第i个特征图的通道加权；*ca_fetures*为通道增强特征；Concat为通道维度拼接操作。

（5）跨尺度自注意力计算。跨尺度自注意力模块(Multi-Scale Attention，MSA)，建模不同尺度特征间的依赖关系，融合了多个尺度的特征，并进行自注意力计算，提升对局部和全局特征的建模能力，增强特征之间的相互关系。跨尺度自注意力计算步骤如下。①尺度分割，将*sa_fetures*和*ca_fetures*分割为多个尺度特征。②生成查询：Q；键：K；值：V，如式(11)所示。③跨尺度注意力权重计算，如式(12)所示。④加权聚合，如式(13)所示。

$$\begin{cases}Q=\mathrm{Conv}_{1\times1}(sa_fetures)\\K=\mathrm{Conv}_{1\times1}(ca_fetures)\\V=\mathrm{Conv}_{1\times1}(ca_fetures)\end{cases} \tag{11}$$

$$\alpha=\mathrm{Softmax}\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}}\right) \tag{12}$$

$$msa_fetures=\alpha V \tag{13}$$

式中：Softmax(·)为归一化指数函数；d_k为Key向量维度，用于缩放点积结果；*msa_fetures*为跨尺度增强特征。

（6）最终特征融合。将空间注意力特征与跨尺度自注意力特征拼接，形成最终融合特征。

$$final_fetures=\mathrm{Concat}(sa_fetures,\ msa_fetures) \tag{14}$$

式中：*final_fetures*为最终增强特征。

3　基于改进SVTR模型的铁路集装箱箱号识别

为了提高铁路集装箱箱号识别的准确率，通过动态调整学习率策略与正则化参数，协同改进SVTR模型，并结合箱号编码规则对结果进行检验与修正。

3.1　SVTR模型改进

SVTR文本识别算法仅依靠视觉特征就能够完成文字识别任务，与主流的模型相比，速度更快，参数量更低。SVTR模型的结构主要由4个核心部分组成：块编码(Patch Embedding)模块、混合块(Mixing Blocks)、下采样模块和线性预测模块[29]。SVRT模型网络结构如图9所示。

Patch Embedding模块通过线性变换层将特征映射到更高维的空间，以增强特征表达能力。Mixing Blocks模块包含局部混合块(Local Mixing Block)和全局混合块(Global Mixing Block)，能够在保留局部细节的同时捕捉全局上下文信息，提升模型对复杂文

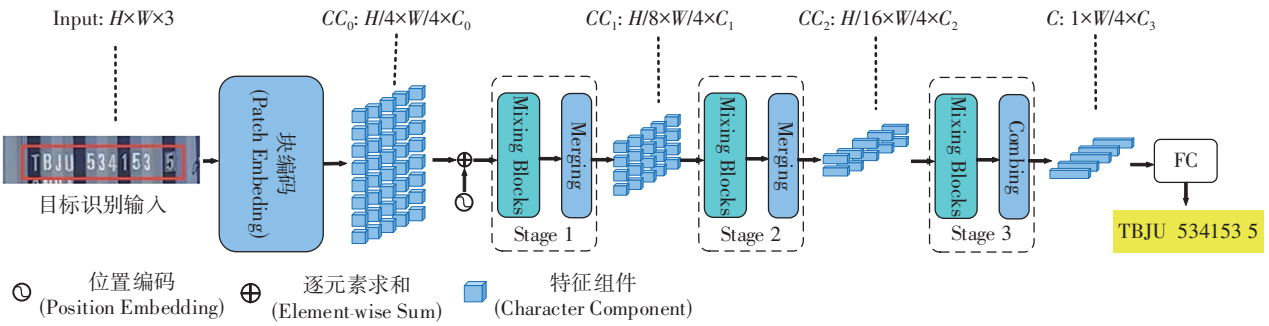


图9　SVRT模型网络结构

Fig.9　SVRT model network structure

本场景的识别能力。下采样模块通过卷积和池化操作逐步降低特征图的分辨率。在Stage 1和Stage 2前2个阶段的合并(Merging)环节使用卷积将注意力图谱的高度减半；在Stage 3的组合(Combing)环节采用池化操作将特征图的高度压缩为1，进一步压缩特征维度，并生成序列特征。

为了进一步优化训练效率和提升模型性能，使用标准余弦退火法结合预热机制(Warmup)作为学习率调整策略，学习率变化曲线如图10所示。与固定步长衰减法和多步长衰减法相比，该方法能够实现学习率的平滑过渡，避免因突变而导致损失函数波动；相比单一的指数衰减法和余弦退火法，在训练初始阶段通过线性增长的Warmup策略逐步提升学习率至峰值0.000 5，使模型参数在初始阶段以小幅度进行更新，缓解模型参数因随机初始化导致的早期震荡问题；随后通过余弦退火函数周期性地平滑降低学习率，使模型在训练后期能更精细地逼近最优解，同时避免因固定学习率过大引发的震荡或过小导致的收敛停滞。

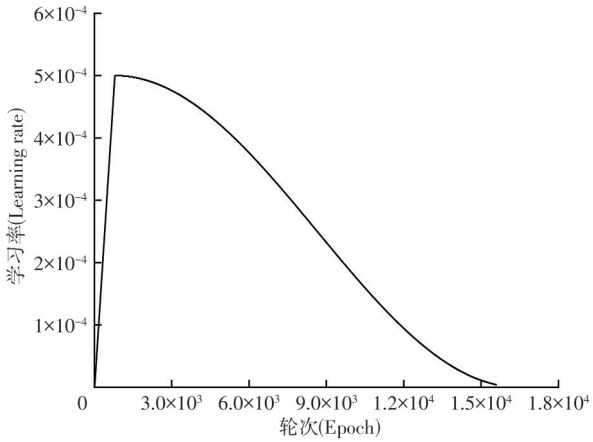


图10　学习率变化曲线

Fig.10　Learning rate variation curve

为增强模型泛化能力，同步优化L2正则化系数，初始值为0.000 03，根据训练集与验证集表现的动态差异，过拟合时增大，欠拟合时降低，以能够更好地平衡训练速度与稳定性，提升整体稳定性。

3.2　箱号检验与修正

为了进一步提高识别的准确率，结合ISO 6346标准中集装箱箱号编码规则对识别结果进行检验与修正：构建一个相似字符对照表，如字母“O”与数字“0”、字母“I”与数字“1”等易混淆字符的对应关系；在校验识别结果时，若发现前4位字符中出现数字，系统会基于相似字符表将其替换为对应的大写字母；同样，若后7位中出现字母，系统也会将其替换为相应的数字。完成校正后，系统会根据集装箱箱号的校验码规则对修正后的箱号进行进一步验证。

4　实验验证

4.1　数据预处理

为了确保图像数据及其对应的标注信息能够有效支持神经网络的学习，在模型训练前对图像数据进行图像解码、标签解码、基础数据增广、获取阈值图标签、获取概率图标签、归一化、通道变换等预处理操作。通过预处理操作不仅保证了数据的一致性，还提升模型的性能、收敛速度和泛化能力，使其能够更好地适应实际应用场景的需求。

4.2　检测模型训练结果与分析

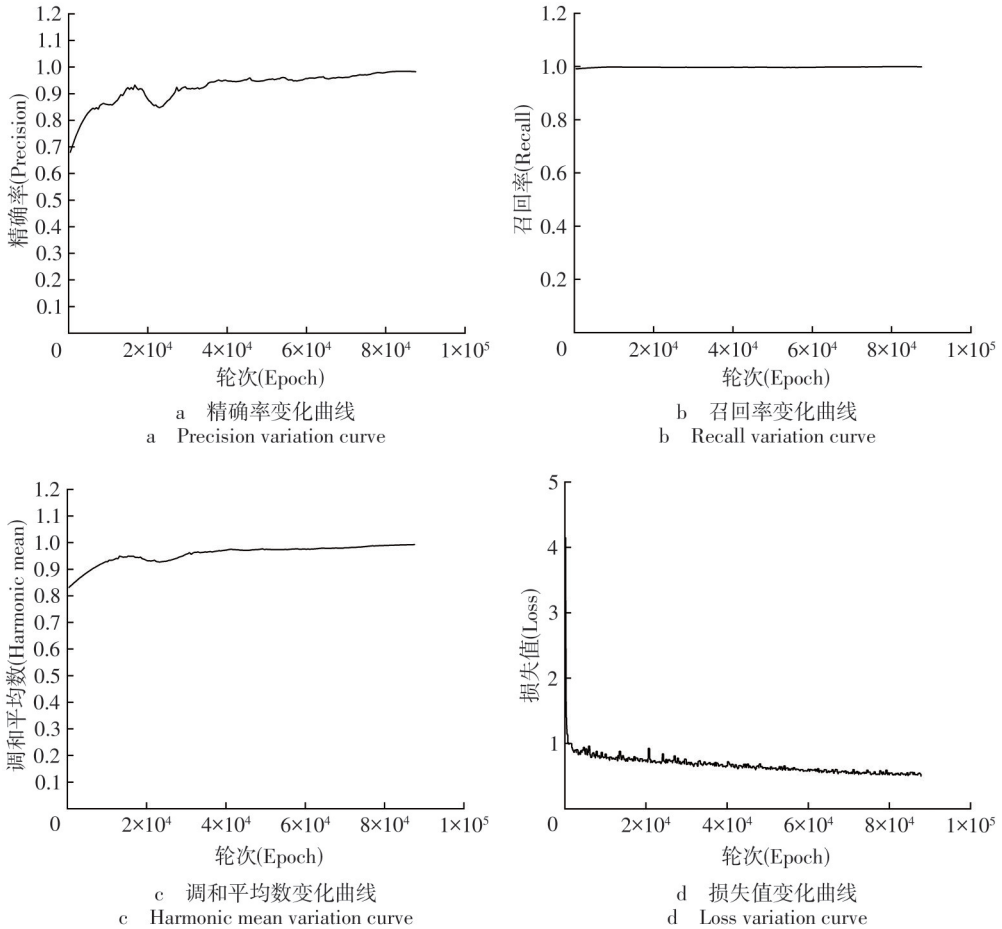
4.2.1　检测模型训练及评估

铁路集装箱箱号训练集包含5 038张图片，验证集包含415张图片。模型训练采用MobileNet V3作为骨干网络，批次大小(Batch Size)设置为16，工

作进程数(Num Works)设置为8，模型共迭代训练200个轮次(Epoch)，其中每个Epoch代表使用小批量梯度下降法对所有训练数据进行一次完整的训练过程。箱号检测模型训练过程中各参数变化曲线如图11所示。

其中箱号检测模型在训练过程中最好的精准度达到了99.0%，具有很好的检测效果，在训练模型时增大后处理阈值(unclip_ratio)，不同后处理阈值下箱号检测效果如图12所示。

从图中可以看出即使在背光条件下，改进后的DB检测模型对于箱号也具有良好的检测效果，unclip_ratio为扩张系数，值越大检测出的文本框就越大，增大unclip_ratio后可以看到检测框变宽，避免检测框边界处与箱号字符重合的情况，避免后续识别过程中造成字符缺失的情况，提高后续识别的准确率。



a　精确率变化曲线
a　Precision variation curve

b　召回率变化曲线
b　Recall variation curve

c　调和平均数变化曲线
c　Harmonic mean variation curve

d　损失值变化曲线
d　Loss variation curve

图11　箱号检测模型训练过程中各参数变化曲线

Fig.11　Parameter variation curves during container number detection model training



a　unclip_ratio为1.5的检测框效果图
a　Detection box result with unclip_ratio=1.5

b　unclip_ratio为1.9的检测框效果图
b　Detection box result with unclip_ratio=1.9

图12　不同后处理阈值下箱号检测效果

Fig.12　Container number detection effects under different post-processing thresholds

4.2.2　检测模型性能与分析

为评估集装箱箱号检测模型性能，选取精确率(Precision)、召回率(Recall)和调和平均数(Hmean)3个指标，分别对改进后的DB模型、基础DB模型、EAST模型、PSE模型及文献[25]的模型进行对比。集装箱箱号检测结果对比如表1所示。从表1可以看出本研究的检测模型在Precison上比DB模型提升2.2%，相较于其他模型有较高的提升，在Recall上比DB模型提升1.3%。对比实验结果表明，针对基础DB检测模型的改进是有效的。

表1　集装箱箱号检测结果对比　%

Tab.1　Comparison of container number detection results

检测模型	Precison	Recall	Hmean
EAST	40.4	71.9	51.7
PSE	98.3	1.0	99.1
基础DB	96.8	98.7	97.8
文献[25]的模型	89.4	93.2	91.3
改进后的DB模型	99.0	100.0	99.5

4.3　识别模型训练结果与分析

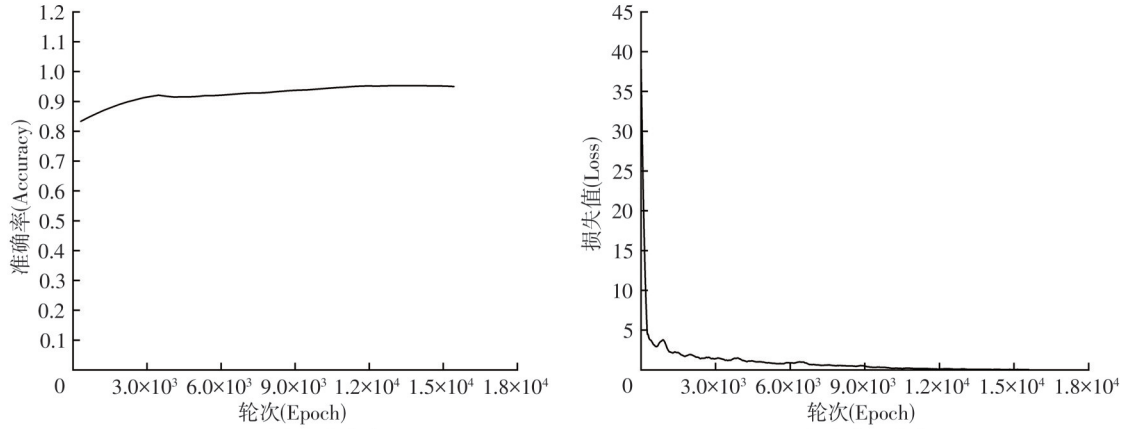
4.3.1　识别模型训练及评估

Batch Size设置为32，Num Works设置为8，模型共迭代训练100个Epoch。箱号识别模型训练过程中各参数变化曲线如图13所示。

箱号文本识别模型在训练过程中最好的准确率达到了95.9%，部分箱号文本识别结果如14所示。图中左侧列举了模糊、遮挡、破损和夜晚等复杂背景下箱号文本图像，右侧包括识别出的文本结果及对应置信度，由结果可以看出优化后的SVTR文本识别具有很好的识别效果。

4.3.2　识别模型性能与分析

为评估集装箱箱号识别模块性能，在自定义数据集上，分别对改进后的SVTR模型、基础SVTR模型、CRNN模型、Rosetta模型及文献[25]的模型，从准确率(Acc)与平均耗时上进行对比，集装箱箱号识别结果对比如表2所示。由表2可以看出，改进后的SVTR模型箱号识别准确率为95.9%，在这几个模型中处于最高水平，每张图片平均识别时间5.55 ms，既满足精度的同时也兼顾了速度。对比实

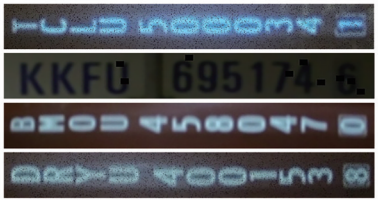


a　准确率变化曲线　　b　损失值变化曲线

a　Accuracy variation curve　　b　Loss variation curve

图13　箱号识别模型训练过程中各参数变化曲线

Fig.13　Parameter variation curves during container number recognition model training



TCLU5000341　0.9999164938926697

KKFU6951746　0.9991464018821716

BMOU4580470　0.9680913686752319

DRYU4001538　0.9999982714653015

a　原箱号图片　　b　箱号文本识别结果与置信度

a　Original container number images　　b　Container number recognition results and confidence levels

图14　部分箱号文本识别结果

Fig.14　Partial container number recognition results

验结果表明，对SVTR识别模型的改进是有效的。

表2　集装箱箱号识别结果对比

Tab.2　Comparison of container number recognition results

识别模型	Acc/%	平均耗时/ms
Rosetta	93.2	8.56
CRNN	94.0	5.31
基础SVTR	95.2	5.41
文献[25]的模型	94.2	—
改进后的SVRT模型	95.9	5.55

5　结论

针对铁路集装箱箱号智能识别技术在复杂场景下识别准确度低、计算复杂度高及难以在资源受限的手持机设备上部署的问题，提出一种基于改进DB与SVTR的集装箱箱号识别方法，主要结论如下。

（1）采集铁路集装箱图片，利用数据增强技术，生成箱号破损、遮挡、模糊及黑夜条件下的照片，以增强模型的泛化能力，并制作数据集，为模型训练与验证提供基础。

（2）对于箱号检测模块，通过在DB模型中FPN网络后加入HAFF混合注意力特征融合模块，增强复杂背景下文本的检测能力，并增大检测框尺寸以避免检测框边界与字符重合的情况，提高后续识别的准确率。

（3）对于箱号识别模块，通过使用标准余弦退火法和Warmup预热机制作为学习率的更新策略，结合调整正则化参数，协同优化SVTR文本识别模型，并结合集装箱箱号的编码规则，对箱号识别结果进行校验与修正，提高识别结果的可靠性。

（4）经过实验验证，改进后模型的箱号检测精确率达到99.0%，箱号识别准确率达到95.9%，每张图片平均识别时间5.55 ms，满足现场对集装箱箱号识别的准确性与时效性需求。

参考文献：

[1] 吕石磊，梁尹聪，李　震，等. 基于超高频RFID双天线双标签对照的果园单轨运输机定位[J]. 农业工程学报，2018，34(4)：71-79.
LYU Shilei，LIANG Yincong，LI Zhen，et al. Orchard Monorail Conveyer Location Based on Ultra High Frequency RFID Dual Antennas and Dual Tags[J]. Transactions of the Chinese Society of Agricultural Engineering，2018，34(4)：71-79.

[2] 周受钦. RFID技术在集装箱行业的应用以及几个问题的思考[J]. 射频世界，2006，1(3)：54-59.

[3] 王　娜，王俊平，朱俊辉. 基于矩阵偏序关系的形态学算子[J]. 电子学报，2023，51(1)：213-221.
WANG Na，WANG Junping，ZHU Junhui. Morphological Operators Based on Matrix Partial Ordering Relation[J]. Acta Electronica Sinica，2023，51(1)：213-221.

[4] 江苏蓬，向　伟，刘云鹏，等. 采用多特征共生矩阵的模板匹配[J]. 光学精密工程，2021，29(6)：1459-1467.
JIANG Supeng，XIANG Wei，LIU Yunpeng，et al. Template Matching with Multi-Feature Co-Occurrence Matrix[J]. Optics and Precision Engineering，2021，29(6)：1459-1467.

[5] 曹林根，宓　超. 集装箱箱号字符识别算法研究[J]. 计算机工程与应用，2021，57(15)：178-185.
CAO Lingen，MI Chao. Research on Recognition Algorithms for Container Code Character[J]. Computer Engineering and Applications，2021，57(15)：178-185.

[6] 牛　晗，伍希志. 基于大津算法连通域的松果多目标识别定位[J]. 江苏农业科学，2021，49(15)：193-198.
NIU Han，WU Xizhi. Pinecone Multi-Target Identification and Location Based on Connected Domain of Otsu Algorithm[J]. Jiangsu Agricultural Sciences，2021，49(15)：193-198.

[7] JADERBERG M，VEDALDI A，ZISSERMAN A. Deep Features for Text Spotting[EB/OL]. [2025-01-01]. https://link.springer.com/chapter/10.1007/978-3-319-10593-2_34.

[8] JADERBERG M，SIMONYAN K，VEDALDI A，et al. Reading Text in the Wild with Convolutional Neural Networks[J]. International Journal of Computer Vision，2016，116(1)：1-20.

[9] REN S Q，HE K M，GIRSHICK R，et al. Faster R-CNN：Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence，2017，39(6)：1137-1149.

[10] WANG B Y，XU J Q，LI J B，et al. Scene Text Recognition Algorithm Based on Faster RCNN[EB/OL]. [2025-01-01]. https://ieeexplore.ieee.org/document/8298720.

[11] ZHONG Z Y，JIN L W，HUANG S P. DeepText：A New Approach for Text Proposal Generation and Text Detection in Natural Images[EB/OL]. [2025-01-01]. https://ieeexplore.ieee.org/document/7952348.

[12] 陈心怡，王雪娇，胡　石. 改进的Faster R-CNN文本检测算法[J]. 安庆师范大学学报(自然科学版)，2021，27(3)：

67-71.
CHEN Xinyi, WANG Xuejiao, HU Shi. Improved Faster R-CNN Text Detection Algorithm[J]. Journal of Anqing Normal University (Natural Science Edition), 2021, 27(3): 67-71.

[13] 崔　循，景文博，于洪洋，等. 基于改进Faster R-CNN的集装箱箱号定位算法[J]. 长春理工大学学报(自然科学版)，2020，43(1)：45-50.
CUI Xun, JING Wenbo, YU Hongyang, et al. Container Code Number Location Algorithm Based on Improved Faster R-CNN[J]. Journal of Changchun University of Science and Technology, 2020, 43(1): 45-50.

[14] WANG Z P, WANG Y L. FRCA: High-Efficiency Container Number Detection and Recognition Algorithm with Enhanced Attention[EB/OL]. [2025-01-01]. https://www.spiedigitallibrary.org/conference-proceedings-of-spie/11373/1137306/FRCA--High-efficiency-container-number-detection-and-recognition-algorithm/10.1117/12.2557197.short.

[15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[EB/OL]. [2025-01-01]. https://arxiv.org/abs/1512.02325.

[16] TIAN Z, HUANG W L, HE T, et al. Detecting Text in Natural Image with Connectionist Text Proposal Network[EB/OL]. [2025-01-01]. https://arxiv.org/abs/1609.03605.

[17] ZHOU X Y, YAO C, WEN H, et al. EAST: An Efficient and Accurate Scene Text Detector[EB/OL]. [2025-01-01]. https://arxiv.org/abs/1704.03155.

[18] WANG W H, XIE E Z, LI X, et al. Shape Robust Text Detection with Progressive Scale Expansion Network[EB/OL]. [2025-01-01]. https://arxiv.org/abs/1903.12473.

[19] LIAO M H, WAN Z Y, YAO C, et al. Real-Time Scene Text Detection with Differentiable Binarization[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 11474-11481.

[20] 张添添. 复杂字符干扰下铁路集装箱箱号检测与识别方法研究[D]. 抚州：东华理工大学，2023.

[21] 李思佳，孔德顺，姜会增. 基于正面吊计量检测的集装箱装卸作业信息化研究与应用[J]. 铁道运输与经济，2023，45(3)：103-110.
LI Sijia, KONG Deshun, JIANG Huizeng. Research and Application of Container Handling Informatization Based on Measurement and Detection of Reach Stacker[J]. Railway Transport and Economy, 2023, 45(3): 103-110.

[22] 钟彦鹏. 集装箱作业流程智能化控制系统设计与研究[J]. 铁道货运，2023，41(11)：54-60.
ZHONG Yanpeng. Design and Research of Intelligent Control System for Container Operation Process[J]. Railway Freight Transport, 2023, 41(11): 54-60.

[23] 涂　昊，刁云峰，程文明，等. 基于深度学习的实时集装箱箱号识别算法[J]. 起重运输机械，2021(13)：58-63.
TU Hao, DIAO Yunfeng, CHENG Wenming, et al. Real-Time Container Number Recognition Algorithm Based on Deep Learning[J]. Hoisting and Conveying Machinery, 2021(13): 58-63.

[24] 李莉莉. 基于深度学习的集装箱编号识别技术研究与实现[D]. 上海：东华大学，2021.

[25] 苗凌玮. 基于计算机视觉的集装箱箱号识别技术研究[D]. 成都：西南交通大学，2022.

[26] 李文晖，杨文韬，张义川. 铁路智慧箱管体系构建及应用研究[J]. 铁道运输与经济，2021，43(8)：43-50.
LI Wenhui, YANG Wentao, ZHANG Yichuan. Construction and Application of Smart Rail Container Management System[J]. Railway Transport and Economy, 2021, 43(8): 43-50.

[27] 肖健雄，刘　冰，程广远，等. 基于人工智能的中欧班列装箱照片智能识别技术应用研究[J]. 铁路物流，2025，43(5)：48-52.
XIAO Jianxiong, LIU Bing, CHENG Guangyuan, et al. Application of Packing Photos-Based Intelligent Recognition Technology for China Railway Express Based on Artificial Intelligence[J]. Railway Logistics, 2025, 43(5): 48-52.

[28] 雷宁昊，刘　辉，裴正楷，等. 基于北斗的铁路集装箱定位识别系统设计与实现[J]. 铁路计算机应用，2025，34(12)：50-54.
LEI Ninghao, LIU Hui, PEI Zhengkai, et al. Railway Container Positioning and Number Identification System Based on BeiDou Satellite Navigation System[J]. Railway Computer Application, 2025, 34(12): 50-54.

[29] DU Y K, CHEN Z N, JIA C Y, et al. SVTR: Scene Text Recognition with a Single Visual Model[EB/OL]. [2025-01-01]. https://arxiv.org/abs/2205.00159.

(责任编辑　李丹丹)