

## 基于深度学习模型Mask R-CNN对M50轴承钢中碳化物的研究

孙瑞明<sup>1,2</sup> 李淑欣<sup>1,2</sup> 鲁思渊<sup>1,2</sup> 金永生<sup>3</sup> 肖华海<sup>3</sup>

(1. 宁波大学机械工程与力学学院, 宁波 315211)

(2. 浙江省零件轧制成形技术研究重点实验室, 宁波 315211)

(3. 宁波银球科技股份有限公司, 宁波 315207)

**摘要:** M50轴承钢中主要的碳化物类型为MC、M<sub>2</sub>C和M<sub>23</sub>C<sub>6</sub>。扫描电子显微镜(Scanning Electron Microscopy, SEM)下,3种碳化物的形状、尺寸和在材料中的分布存在明显的区别。有些碳化物的尺寸较大且分布不均匀。轴承承载过程中,这些碳化物会成为应力集中的区域,对轴承疲劳性能产生负面影响。为了高效地获得材料中的碳化物信息,提出一种改进的掩膜基于区域的卷积神经网络(Mask Region-based Convolutional Neural Network, Mask R-CNN)模型,可批量鉴别SEM图像中3种碳化物的种类,确定其尺寸大小及分布。网络模型输出的图像和数值结果显示,M50轴承钢中M<sub>2</sub>C型碳化物尺寸较大且分布不均匀,但总体尺寸最大的MC型碳化物和尺寸最小的M<sub>23</sub>C<sub>6</sub>型碳化物分布相对均匀。

**关键词:** 深度学习; Mask R-CNN; M50轴承钢; 碳化物

**中图分类号:** TH142; TP317 **DOI:** 10.16579/j.issn.1001.9669.2025.08.003

## 0 引言

M50轴承钢是一种高合金钢,具有优异的耐磨性、耐腐蚀性,广泛用于航空航天、船舶、汽车等领域<sup>[1]</sup>。该轴承钢主要成分为C、Cr、Mo、V、Ni等元素,其中V、Mo、Cr有利于碳化物的形成。碳化物和合金元素可提供二次硬化,使组织更加稳定<sup>[2-4]</sup>。碳化物按析出先后顺序分为一次碳化物和二次碳化物。一次碳化物在液态金属固化时形成,二次碳化物由固相转变析出。研究发现,M50轴承钢中存在较多尺寸偏大且分布不均匀的一次碳化物,其严重危害了其冲击韧性和疲劳寿命,不利于轴承的长期服役<sup>[5-7]</sup>。针对M50轴承钢中的碳化物展开了大量的研究<sup>[8][11][822][9][161755][10]105-110</sup>发现,M50轴承钢中的一次碳化物多为MC型和M<sub>2</sub>C型,二次碳化物主要为M<sub>23</sub>C<sub>6</sub>型<sup>[11]98-103</sup>。一次碳化物MC和M<sub>2</sub>C尺寸较大,二次碳化物M<sub>23</sub>C<sub>6</sub>尺寸相对较小。碳化物MC形状为类圆状或棒状,M<sub>2</sub>C形状为棒状或层片状<sup>[9][161755]</sup>,M<sub>23</sub>C<sub>6</sub>为颗粒状。MC的元素中富含V,M<sub>2</sub>C富含Mo,M<sub>23</sub>C<sub>6</sub>富含Cr。扫描电子显微镜(Scanning Electron Microscopy, SEM)图片中MC比M<sub>23</sub>C<sub>6</sub>衬度深<sup>[10]107</sup>。

近年来,在深度学习与材料研究的结合方面取得了一定的成果。DECOST等<sup>[12]</sup>使用视觉特征提取单元结合支持向量机(Support Vector Machine, SVM)算

法对黄铜、青铜、球墨铸铁、灰铸铁、亚共析钢、可锻铸铁、高温合金等7种材料的微观组织进行分类识别,准确率达80%。DURMAZ等<sup>[13]6272</sup>使用不同的Unet模型,对复杂金相组织中的贝氏体进行识别分割,准确率超过90%。YANG等<sup>[14]</sup>使用全连接神经网络(Fully Connected Neural Network, FCNN)、长短期记忆网络(Long Short-Term Memory, LSTM)、注意力机制等拟合了不同材料的疲劳寿命曲线,寿命数据在拟合寿命曲线的2倍分散带以内。李维刚等<sup>[15]</sup>利用不同的深度学习模型自动识别钢铁中的铁素体、马氏体、珠光体等,在测试集中的准确率达到100%。段献宝等<sup>[16]</sup>使用3种神经网络,完成了对钢材微观组织的高效、准确识别,分类精度达到99.6%。苏晨等<sup>[17]78-84</sup>将注意力机制与ResNet结合对钢材中残余奥氏体进行识别并评级。彭凡等<sup>[18]</sup>使用深度学习对螺栓连接失效载荷进行了非线性拟合。

综合来看,深度神经网络在材料学科的应用仍以图片识别、曲线拟合为主,对图片中内容的定量分析较少。本文设计了一种改进的掩膜基于区域的卷积神经网络(Mask Region-based Convolutional Neural Network, Mask R-CNN),用以分析M50轴承钢的碳化物,可以同时实现如下功能:①识别SEM图片中的碳化物并确定其类别;②对SEM图片中的碳化物进行分割

收稿日期: 2023-11-23 修回日期: 2024-01-16

基金项目: 国家自然科学基金项目(52075271); 宁波市重大科技任务攻关项目(2022Z050); 慈溪市行业共性技术攻关项目(CZ2022009)

作者简介: 孙瑞明,男,1993年生,河南鹤壁人,硕士研究生; 主要研究方向为机械结构强度; E-mail: 15658028610@163.com。

李淑欣(通信作者),女,1975年生,宁夏中卫人,博士,教授,博士研究生导师; 主要研究方向为机械结构强度; E-mail: lishuxin@nbu.edu.cn。

引用格式: 孙瑞明,李淑欣,鲁思渊,等. 基于深度学习模型Mask R-CNN对M50轴承钢中碳化物的研究[J]. 机械强度, 2025, 47(8): 19-27.

SUN Ruiming, LI Shuxin, LU Siyuan, et al. Research on carbides in M50 bearing steel based on Mask R-CNN deep learning model[J]. Journal of Mechanical Strength, 2025, 47(8): 19-27.

标记;③确定碳化物在图片中的占比;④统计图片中每个碳化物的尺寸,并输出碳化物的直径分布直方图。

### 1 对 Mask R-CNN 的改进

#### 1.1 Mask R-CNN 的结构

Mask R-CNN 模型是一种用于目标检测和实例分割的深度学习模型,是 R-CNN 系列模型的一种扩展,

由 HE 等<sup>[19]386-397</sup>于 2017 年提出。结合待识别碳化物图片的特点,本文在原 Mask R-CNN 模型的基础上进行了一系列的修改,具体细节见下文。本文还在最后程序中加入了数据处理模块,可对识别到的碳化物进行量化统计。整体网络主要由 Faster R-CNN Predictor 和 Mask Branch 这 2 个分支组成,网络结构如图 1 所示。

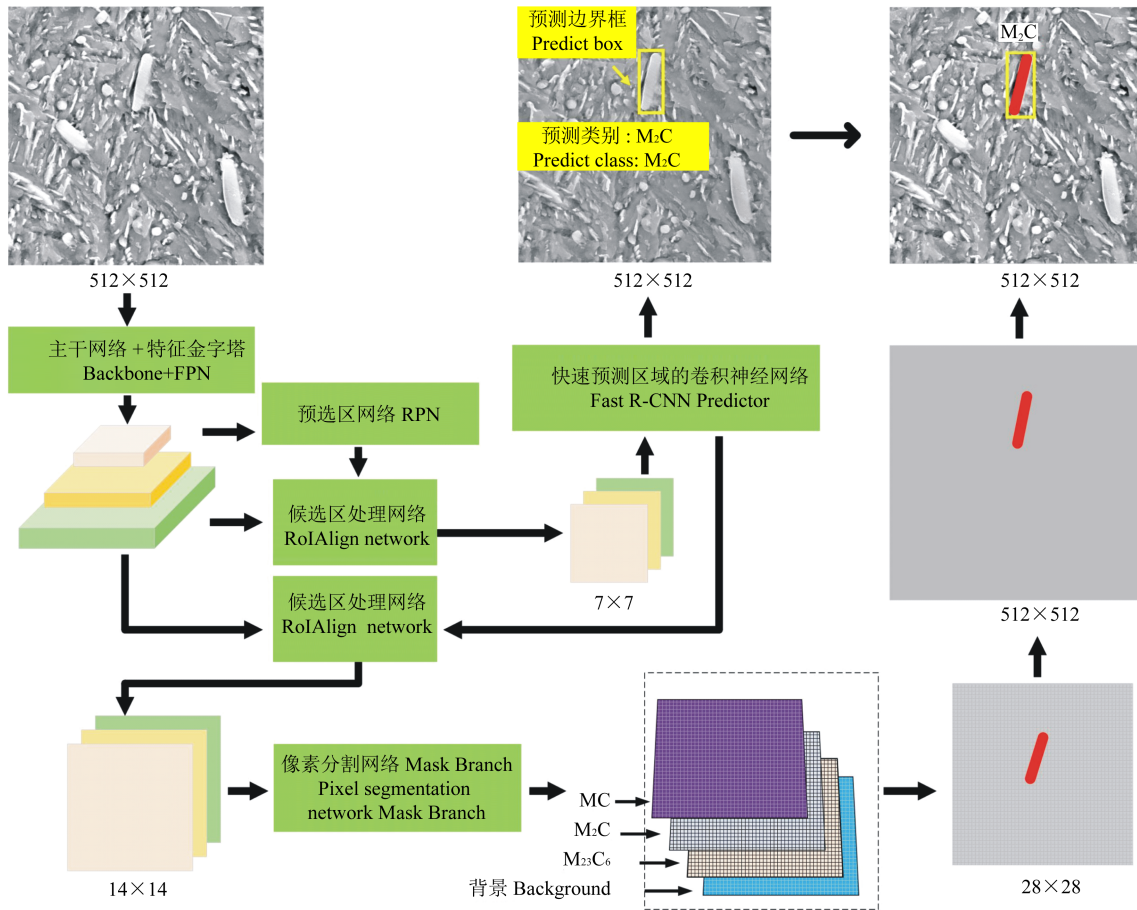


图1 Mask R-CNN 结构流程图

Fig. 1 Flow chart of the Mask R-CNN structure

结构中 Backbone+FPN 是特征提取模块。Backbone 使用的是残差网络 ResNet50。ResNet50 中使用了恒等映射结构。恒等映射单元如图 2 所示。对应式(1), 函数  $F(x_i, w_i)$  的输出值可为 0。当  $F(x_i, w_i)=0$  时, 就有  $x_{i+1}=x_i$ , 即输出和输入相同, 这样就实现了恒等映射。该结构保证了神经网络层数的增加, 不会使训练出的结果变差。特征金字塔网络 (Feature Pyramid Network, FPN)<sup>[20]</sup> 能够在不同的尺寸上提取特征。原始输入的 SEM 图片经 Backbone 部分处理后, 存在多次尺寸的变化, 使用特征金字塔结构可以较好地弥补图片在尺寸变化时部分特征信息丢失的缺点, 提高了模型的鲁棒性。

$$x_{i+1} = F(x_i, w_i) + x_i \quad (1)$$

区域建议网络 (Region Proposal Network, RPN) 的

作用是在图片的特定位置生成一系列预选边界框。

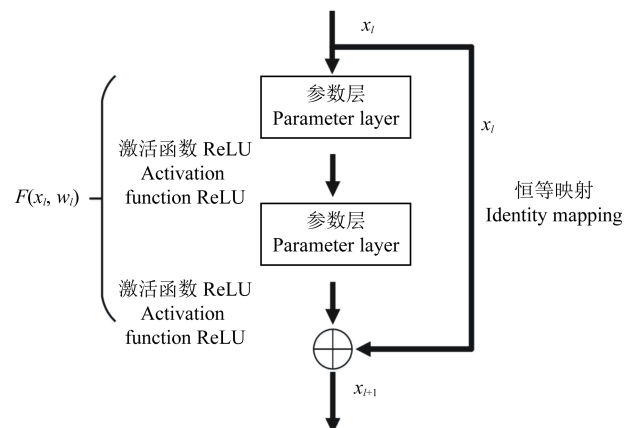


图2 恒等映射单元

Fig. 2 Identity mapping unit

RoIAlign 中 RoI 是 Region of Interest 的简写,即感兴趣区域。该模块的功能是将指定的边界框(来自 RPN 或者 Fast R-CNN Predictor)先映射到 Backbone+FPN 生成的特征图上,之后将对应区域变换为特定的大小,最后对变换后的区域进行最大池化处理。Fast R-CNN Predictor 模块使用的 RoIAlign(使用的边界框来自 RPN)将特征层处理后生成的图像大小是  $7 \times 7$ 。Mask Branch 模块使用的 RoIAlign(使用的边界框来自 Fast R-CNN Predictor)将特征层处理后生成的图像的大小是  $14 \times 14$ 。 $14 \times 14$  的图像可以在一定程度上保留更多和尺寸相关的特征信息,有利于对碳化物进行像素级的掩码分割。

Fast R-CNN Predictor 模块用于执行目标分类、确定边界框回归参数,即确定输入的 SEM 图片中哪些区域是碳化物,将非碳化物区域确定为背景,同时确定碳化物的边界框的大小及位置。

Mask Branch 模块用于生成每个目标的精确掩码,即在目标边界框的内部确定每一个像素属于何种类别,是 MC、 $M_2C$ 、 $M_{23}C_6$  或者背景,以二值图像的形式表示出来。

最后,将 Mask Branch 生成的二值图像和 Fast R-CNN Predictor 生成的分类信息与边界框信息映射到原图上,得到具有碳化物类别信息、碳化物的边界框和碳化物像素分割信息的图像。

## 1.2 网络的预设

设定的神经网络的初始参数和数据处理方法直接影响训练的时间和最后的结果。在参考一些文献<sup>[13]6272[19]386-393[21]1025-1029</sup>和实践的基础上,对设定的参数、超参数、数据集的处理方法进行了如下思考。

### 1.2.1 迁移学习

迁移学习是深度学习中一个重要的方法。本文使用的 Mask R-CNN 结构较为复杂,若从 0 开始训练该网络,则需要较长的时间。迁移学习是将之前学习到的参数迁移到新的任务或领域中,从而改善新任务的学习效果,能在较短的时间内使网络收敛。迁移学习还可以减少数据需求<sup>[21]1028</sup>,提高模型的泛化能力,改善模型的表现等。通常情况下,迁移学习应用于两个相关但不同的任务,分别称为源任务和目标任务。

本文综合使用了特征提取和微调两种迁移学习方法,使用 Coco 数据集训练出的参数作为 Backbone+FPN 初始参数。Coco 数据集的图片数量达 328 000 张,有 82 个类别,经该数据集训练出来的参数有较好的特征提取能力。设定前 10 个 Epoch 为预训练阶段,该阶段使 Backbone+FPN 参数不变,仅在训练中更新 Fast R-CNN Predictor 和 Mask Branch 的参数。当大于 10 个 Epoch 后,解除对 Backbone+FPN 参数的冻结,整

体网络参数一起训练。

### 1.2.2 数据增强

数据增强是深度学习中一种常用的技术,通过对原始数据进行一系列的变换和处理,生成新的训练样本,从而扩充训练数据集的规模和多样性。数据增强技术对输入数据进行随机变换,提高了模型的泛化能力和鲁棒性。本文使用的 SEM 图片共有 240 张,与常见的数据集,如 ImageNet(100 万张)、MS Coco(328 000 张)、PASCAL VOC(11 540 张)对比,可知本文数据集的数量是相对较少的。因此使用数据增强是必要的。DURMAZ 等<sup>[13]6272</sup>使用数据增强的方法,用相对较少的图片完成了网络的训练,达到了预期的效果。

常见的数据增强技术有镜像翻转、随机裁剪、旋转和缩放、增加噪声、图像亮度和对比度调整等。本次研究的重点是图片中的碳化物种类和形状。在数据增强时,改变宽高比会明显地改变图片中碳化物的形状,改变颜色会影响原有的衬度,进而影响结果的准确性。因此,本文数据增强时仅采用水平翻转和垂直翻转,以保证碳化物的形状、衬度不变。

### 1.2.3 自适应学习率

在深度学习中,传统的优化算法,如梯度下降使用固定的学习率,但对于复杂的深度神经网络,固定学习率可能不够有效。自适应学习率是一种根据模型训练过程中的情况自动调整学习率的技术,可以更加高效地训练模型,加速模型的收敛,提高模型的性能。

常见的自适应学习率方法包括:动量法、学习率衰减、AdaGrad、RMSprop、AdaDelta、Adam 等。本次使用学习率衰减法,设定学习率随着训练次数的增加而减小。学习率越小,网络训练时参数变化越小,该做法保证了模型在训练到后期时的精度。本文将初始学习率设置为 0.002,学习率的下限设为 0.000 02。基于数据集图片的数量,本文选取的迭代次数为 300。

## 1.3 均值平均精度

混淆矩阵是深度学习中常用的评估模型性能的工具。混淆矩阵所包含的内容如表 1 所示。

表 1 混淆矩阵

Tab. 1 Confusion matrix

混淆矩阵 Confusion matrix		真实值 True value	
		阳性 Positive	阴性 Negative
预测值 Predicted value	阳性 Positive	真阳性 True positive	假阳性 False positive
	阴性 Negative	假阴性 False negative	真阴性 True negative

混淆矩阵是一个二维矩阵,用于比较分类模型的预测结果与真实标注之间的差异。通过混淆矩阵,可以计算出各种性能指标,如准确率(Accuracy)、精确率(Precision)、召回率(Recall)等。其中 Precision 和

Recall公式定义为

$$P_{\text{Precision}} = \frac{P_{\text{TP}}}{P_{\text{TP}} + P_{\text{FP}}} \quad (2)$$

$$R_{\text{Recall}} = \frac{P_{\text{TP}}}{P_{\text{TP}} + P_{\text{FN}}} \quad (3)$$

式中,  $P_{\text{TP}}$  为真阳性数量;  $P_{\text{FP}}$  为假阳性数量;  $P_{\text{FN}}$  为假阴性数量。由其定义可知  $P_{\text{Precision}}$  和  $R_{\text{Recall}}$  的值均小于 1。以  $P_{\text{Precision}}$  值为纵坐标、 $R_{\text{Recall}}$  值为横坐标作图, 便构成了评价神经网络的  $P_{\text{Precision}}-R_{\text{Recall}}$  曲线图。  $P_{\text{Precision}}-R_{\text{Recall}}$  曲线下的面积用  $P_{\text{AP}}$  (Average Precision, AP) 表示, 相关计算式为式 (4); 均值平均精度 (Mean Average Precision, MAP) 是所有类别的平均精度的均值, 相关计算式为式 (5)。

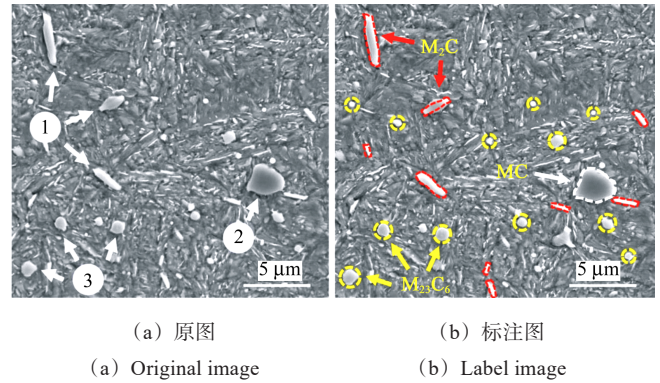
$$P_{\text{AP}} = \int_0^1 P_{\text{precision}}(R_{\text{Recall}}) dR_{\text{Recall}} \quad (4)$$

$$P_{\text{MAP}} = \frac{\sum P_{\text{AP}}}{n} \quad (5)$$

$P_{\text{Precision}}-R_{\text{Recall}}$  曲线下的面积越大, 说明网络输出结果与标注图越匹配, 即  $P_{\text{MAP}}$  的值越大说明网络的性能越好,  $P_{\text{MAP}}$  的上限为 1, 以  $P_{\text{MAP}}$  为指标衡量神经网络具有综合性; 因此, 本文选取  $P_{\text{MAP}}$  为网络的评价指标。

## 2 碳化物类型鉴别和数据集制作

本文研究的 M50 轴承钢的化学成分 (质量分数, %) 为: 0.81~0.85C, 4.00~4.25Cr, 4.00~4.50Mo, 0.90~1.10V,  $\leq 0.15\text{Ni}$ , 0.15~0.35Mn,  $\leq 0.25\text{Si}$ ,  $\leq 0.25\text{Co}$ ,  $\leq 0.25\text{W}$ , 其余为 Fe。材料的热处理工艺影响碳化物的产生及分布, 本试验所用材料选取常规的热处理方法, 具体工艺是: 真空感应熔炼-自耗电极-真空自耗熔-自耗锭-高温均匀化-锻造-轧制棒材。所制成的坯料棒材的直径为 65 mm, 将  $\Phi 65$  mm 坯料经切割、车削、磨削等工艺制成  $\Phi 12$  mm 的试样轴, 将试样轴切片、打磨、抛光、腐蚀后在扫描电子显微镜 SU5000 下观察, 得到图 3(a) 所示的图片。图 3(a) 中的碳化物按外形可分为圆棒形、类圆形和颗粒形 3 种, 分别对应图 3(a) 中的 1、2、3。为进一步确定材料中碳化物的种类, 将材料经离子减薄制成可用透射电子显微镜 (Transmission Electron Microscope, TEM) JEM-F200 观察的样品。样品在 TEM 中经能量色散谱 (Energy Dispersive Spectroscopy, EDS) 分析得到图 4 所示结果。由图 4 可知, 圆棒形碳化物富含 Mo, 类圆形较大的碳化物富含 V, 相对较小的颗粒状碳化物富含 Cr。初步判断 3 种碳化物分别为  $\text{M}_2\text{C}$ 、MC 和  $\text{M}_{23}\text{C}_6$ 。对相同形状的碳化物做衍射斑分析, 结果如图 5 所示。标定结果显示, 3 种碳化物分别为  $\text{M}_2\text{C}$ 、MC 和  $\text{M}_{23}\text{C}_6$ 。上述结果也与文献 [9]<sup>161755</sup>、文献 [10]<sup>107-109</sup>、文献 [11]<sup>101-103</sup> 的描述相符。最终确定 SEM 图片的碳化物类型, 如图 3(b) 所示。



(a) 原图

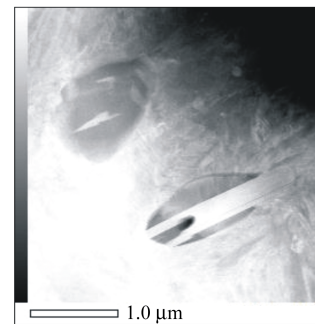
(b) 标注图

(a) Original image

(b) Label image

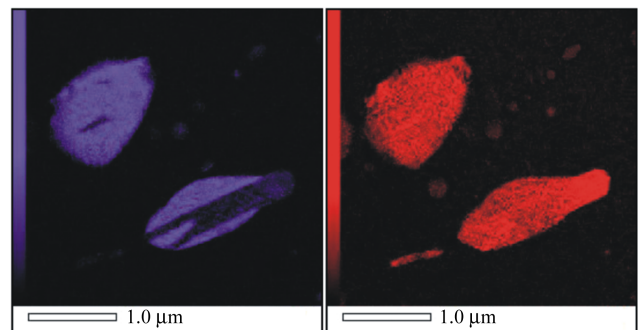
图 3 M50 轴承钢的 SEM 图像

Fig. 3 SEM images of M50 bearing steel



(a) 碳化物的 TEM 图像

(a) TEM image of carbides

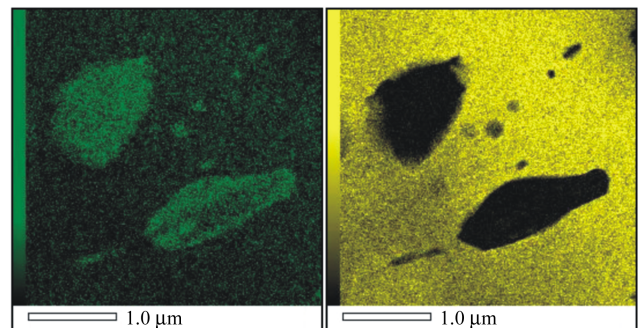


(b) 钒元素分布

(c) 钼元素分布

(b) Distribution of V element

(c) Distribution of Mo element



(d) 铬元素分布

(e) 铁元素分布

(d) Distribution of Cr element

(e) Distribution of Fe element

图 4 M50 轴承钢的 TEM-EDS 图像, 碳化物形貌及 V、Mo、Cr、Fe 元素分布

Fig. 4 TEM-EDS images of M50 bearing steel, carbides' morphology and distribution of V, Mo, Cr, Fe element

使用 SEM 观察并采集图像,随机在表面的不同位置获取多张图片,每张图像的分辨率是 1 280×960。受限于计算机的计算能力,输入神经网络的图片的尺寸不宜太大。在制作数据集时,去除原始图片中有文字的部分,将剩下的部分截取为两张,每张图片大小为 512×512。然后结合图 3(b)所示的结果,使用神经网络

标注工具 Labelme 对图片中的碳化物进行手动标注。将衬度较浅、面积较小的颗粒状碳化物标注为  $M_{23}C_6$ 。将衬度较深、面积相对较大的类圆形碳化物标注为 MC。将圆棒形的碳化物标注为  $M_2C$ 。标注前后的图片如图 6 所示。标注时忽略特别细小的碳化物。标注完成后使用转化程序将标注好的数据集转化为 json 格式,用于之后对模型的训练。

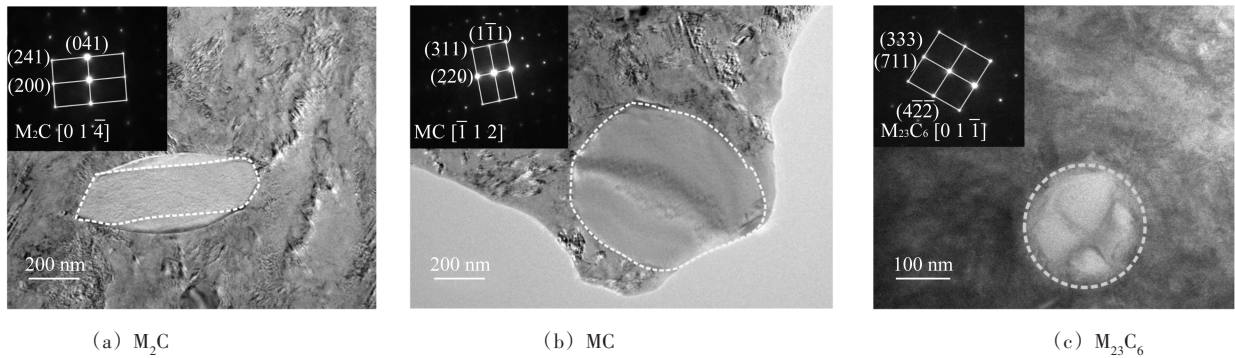
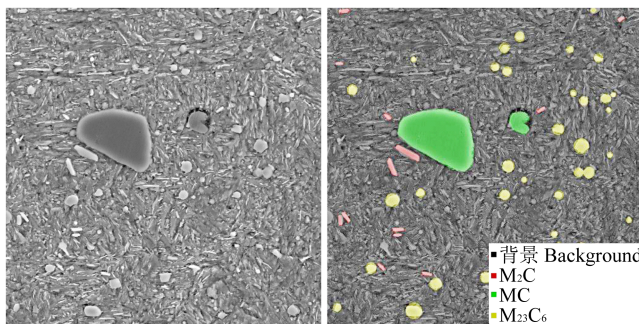


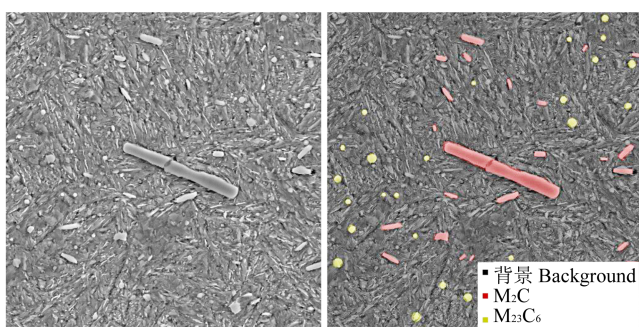
图 5 M50 轴承钢中  $M_2C$ 、MC、 $M_{23}C_6$  的 TEM 图像及衍射花样

Fig. 5 TEM images of  $M_2C$ , MC and  $M_{23}C_6$  and their diffraction patterns in M50 bearing steel



(a) 原始图片  
(a) Original image

(b) 图 6(a) 的标注图  
(b) Label image of Fig. 6(a)



(c) 原始图片  
(c) Original image

(d) 图 6(c) 的标注图  
(d) Label image of Fig. 6(c)

图 6 M50 轴承钢的 SEM 原图及标注图

Fig. 6 Original SEM images of M50 bearing steel and their label images

### 3 神经网络对数据集中碳化物的识别和量化统计

本文使用的图片分为训练集、验证集和测试集,分别为 160 张、40 张和 40 张。训练集和验证集都是带

有标注的图片,用于前期对网络模型的训练。测试集用于最后对模型的检验。为了提高训练速度,本次使用图形处理器(Graphics Processing Unit, GPU)进行并行训练,优化器选用了 Adam<sup>[22]</sup>,使用的框架是 PyTorch。经过 8 h 完成 300 个 Epoch 的训练。训练后网络达到了使用要求,结果见下文。

#### 3.1 对 SEM 图片中碳化物的识别

将测试集的 40 张图片输入训练好的网络,取置信度为 50%,得到所有的图片处理结果,其中的 3 张结果如图 7 所示。图 7 中用不同的颜色标识出不同的碳化物,碳化物的种类在边界框的左上角标出。由图 7 可看出,网络能较为准确地识别出图中的碳化物并正确地确定其种类。由图 7 还可直观地看出 3 种碳化物在图片内的分布。图 7(c)展示了在材料的局部区域多个  $M_2C$  型碳化物集中出现的现象,该结果说明 M50 轴承钢材料内存在  $M_2C$  分布不均的情况。

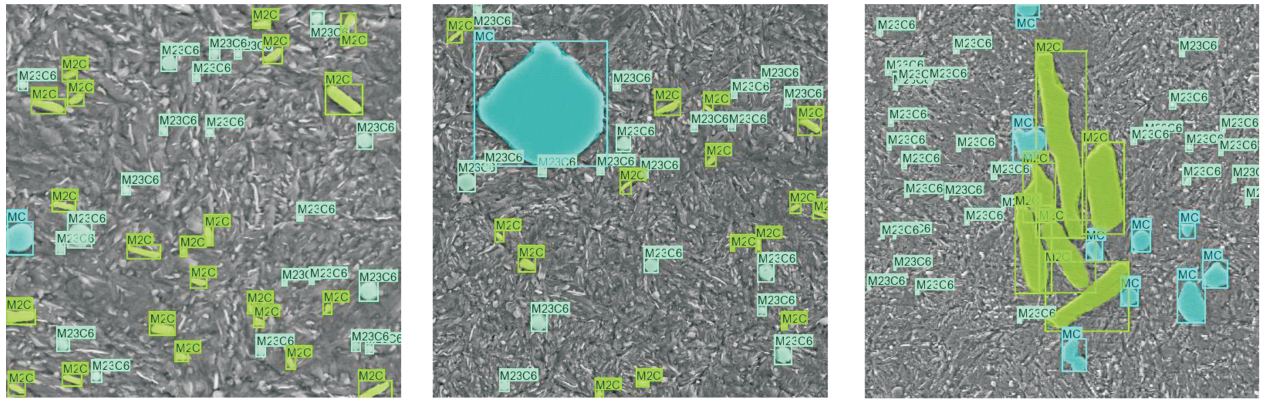
#### 3.2 图片中碳化物的占比和尺寸

材料中碳化物的多少也是影响材料性能的一个关键因素。本文的模型以像素面积占比计算 40 张测试集图片中碳化物的面积占比,所得的结果为 5%。相关表达式为

$$P(C) = \frac{\sum N_c}{\sum N_p} \times 100\% \quad (6)$$

式中,  $N_c$  为一张图片内属于碳化物的像素点的总数;  $N_p$  为一张图片中像素点的总数。

碳化物的尺寸直接影响了材料的性能。明确碳化物的尺寸对分析材料的品质、判断材料的性能具有重要意义。本文的网络使用两种尺寸表示碳化物的大小,分别为当量直径和费雷特直径。



(a) 输出结果图1  
(a) Output image 1

(b) 输出结果图2  
(b) Output image 2

(c) 输出结果图3  
(c) Output image 3

图7 3张测试结果图

Fig. 7 Three images of testing results

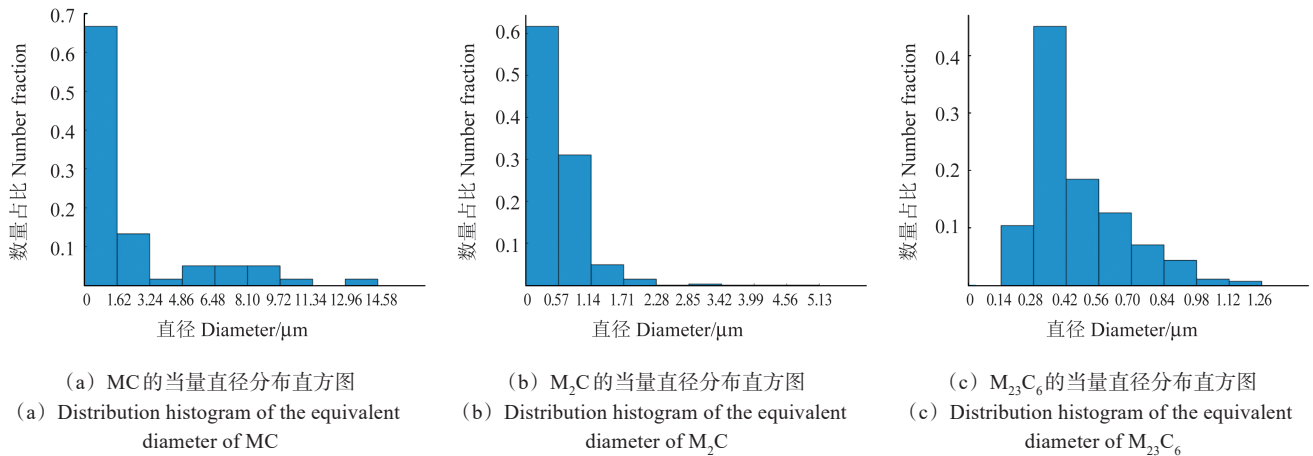
### 3.2.1 当量直径

SEM图片中单个碳化物的当量直径计算式为

$$D_{eq} = 2\sqrt{N/\pi} / I \quad (7)$$

式中,  $N$  为1个碳化物的像素面积;  $I$  为图片比例尺。

网络能输出每个碳化物的当量直径。为了直观展示,本次将所有碳化物的直径以直方图的形式输出。使用测试集图片得到的当量直径分布直方图如图8所示。



(a) MC的当量直径分布直方图

(a) Distribution histogram of the equivalent diameter of MC

(b)  $M_2C$ 的当量直径分布直方图

(b) Distribution histogram of the equivalent diameter of  $M_2C$

(c)  $M_{23}C_6$ 的当量直径分布直方图

(c) Distribution histogram of the equivalent diameter of  $M_{23}C_6$

图8 MC,  $M_2C$ 和 $M_{23}C_6$ 的当量直径分布直方图

Fig. 8 Distribution histogram of equivalent diameters of MC,  $M_2C$  and  $M_{23}C_6$

### 3.2.2 费雷特直径

在图7中,每个碳化物的边界框的4个顶点都有相应的坐标,通过坐标可以计算边界框的长和宽。本文选取边界框中较长边作为对应碳化物的费雷特直径。基于此,可求得测试集40张图片中的3种碳化物的费雷特直径。用得到的3种碳化物的费雷特直径生成费雷特直径分布直方图,如图9所示。

## 4 结果与讨论

### 4.1 网络的性能指标

改进的Mask R-CNN共进行了300轮的迭代,学习率和损失值分别反映了网络的训练精度和测试结果的准确度。图10所示为学习率随迭代次数变化的曲线,图11所示为损失值<sup>[23]</sup>随迭代次数变化的曲线。

前10个Epoch是预训练阶段,该阶段Backbone+FPN模块的参数保持不变,仅更新网络其他部分的参数。图10中,在第200个Epoch时学习率从0.002衰减到0.0002,在第270个Epoch时学习率衰减为最小值0.00002,之后保持不变。由图11可看出,在前50轮迭代时损失值快速下降。在第200个Epoch后损失值逐渐稳定,最后数值约为0.2。该损失值为分类损失、边界框损失、图片分割损失的平均值。损失值为0.2,该数值较小,满足使用要求。

图12所示为网络的 $P_{MAP}$ 值随着迭代次数的变化曲线。网络的MAP值在迭代了220个Epoch后趋于稳定,最后数值约为0.9,即平均精度达90%。该精确度达到了同类网络的较高水平<sup>[13]6272[17]83</sup>,满足实验室测试的精度要求。

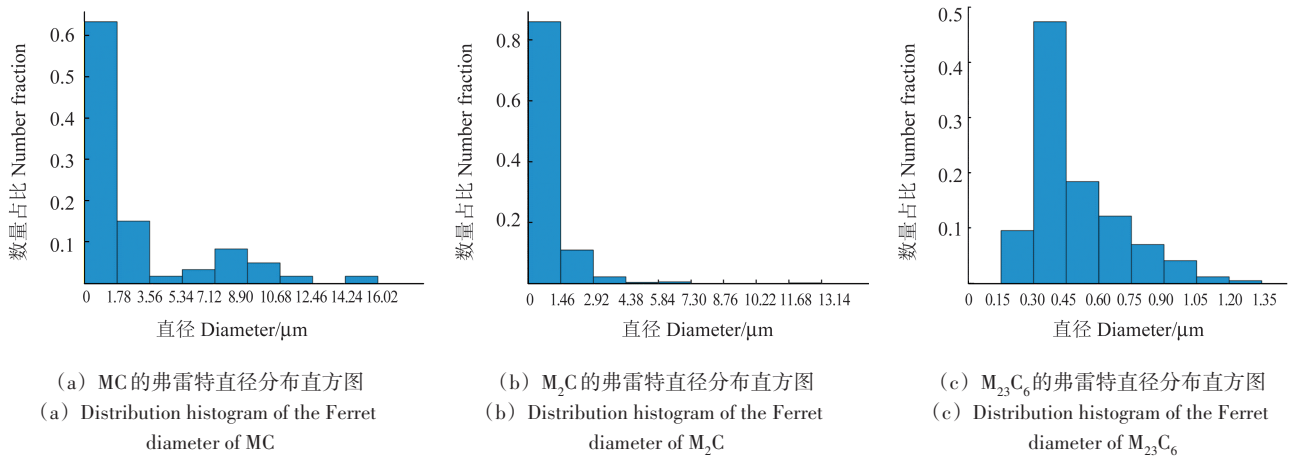
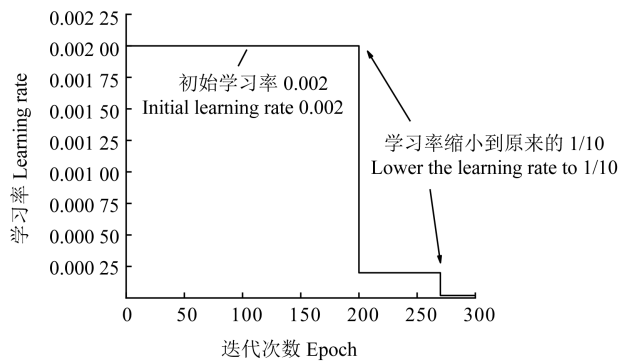
图9 MC、 $M_2C$ 和 $M_{23}C_6$ 的弗雷特直径分布直方图Fig. 9 Distribution histogram of Ferret diameters of MC,  $M_2C$  and  $M_{23}C_6$ 

图10 Mask R-CNN的学习率

Fig. 10 Learning rate of the Mask R-CNN

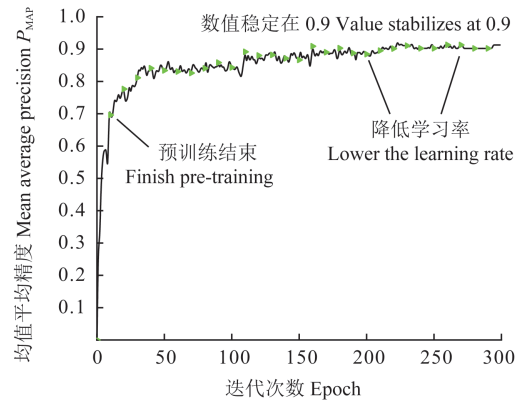
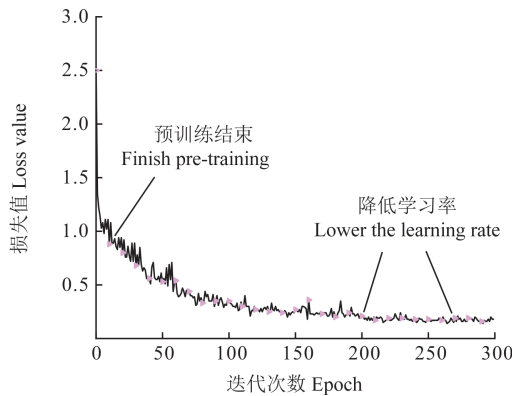
图12 Mask R-CNN的 $P_{MAP}$ Fig. 12  $P_{MAP}$  of the Mask R-CNN

图11 Mask R-CNN的损失值

Fig. 11 Loss value of the Mask R-CNN

## 4.2 对量化结果的分析

图8、图9分别展示了3种碳化物的当量直径和弗雷特直径。从每一个直方图上看,在数量占比上,3种碳化物中,都是小直径的碳化物占比大,大直径的碳化物占比小。该结果能直观地看出材料中每种碳化物的细化情况,可帮助研究者判断材料的品质。将三种碳化物的两种直径进行对比,汇总结果如表2所示。对比同一种碳化物的2种直径可知,MC和

$M_{23}C_6$ 的2种直径差异不大, $M_2C$ 的2种直径差异明显。一个碳化物的当量直径和弗雷特直径的差异越大,其宽高比越大,宽高比越大的碳化物越容易断裂<sup>[8]111822</sup>。由此可知, $M_2C$ 更容易断裂,断裂的碳化物易成为材料裂纹的萌生点。碳化物发生断裂的频率与碳化物的尺寸、强度、形状相关<sup>[8]111822</sup>。考虑到3种碳化物外形的不规则性,相较于当量直径,弗雷特直径能更准确地从数字上表达碳化物的大小。因此,在从尺寸和形状上判断碳化物的断裂难易性时,弗雷特直径可作为判断依据。在同一类直径3种碳化物之间相互对比可知,MC和 $M_2C$ 的尺寸比 $M_{23}C_6$ 大,且3种碳化物的直径最大值分别是16、13、1.3  $\mu\text{m}$ 。碳化物的尺寸越大,越容易引起局部应力集中,进而导致材料产生裂纹<sup>[24]</sup>。因此,较大的一次碳化物MC和 $M_2C$ 更容易引起材料内部应力集中,而 $M_{23}C_6$ 的影响较小。综上所述,无论是同一种碳化物2类直径的对比还是同一类直径几种碳化物之间的对比,都能进一步获得材料中微观组织的信息,该信息可为材料的优化提供指导。

表2 3种碳化物的直径分布

Tab. 2 Distribution of three kinds of carbides' diameters

碳化物 Carbide	当量直径 Equivalent diameter/ $\mu\text{m}$			费雷特直径 Ferret diameter/ $\mu\text{m}$		
	分布范围	最大值	平均值	分布范围	最大值	平均值
	Distribution range	Maximum value	Average value	Distribution range	Maximum value	Average value
MC	0~14.58	14	1.17	0~16.02	16	1.47
M <sub>2</sub> C	0~5.13	5	0.59	0~13.14	13	0.96
M <sub>23</sub> C <sub>6</sub>	0~1.26	1.2	0.45	0~1.35	1.3	0.45

## 5 结论

本文改进的Mask R-CNN使用训练集160张图片、验证集40张图片,经300轮迭代完成了模型的训练。用训练好的网络对测试集的40张SEM图片进行处理,得到了碳化物分割结果图和直径分布直方图。对输出的图像和数值结果进行分析,得到以下主要结论:

1) M50轴承钢的SEM图片中的碳化物MC、M<sub>2</sub>C、M<sub>23</sub>C<sub>6</sub>能被网络识别并标记,平均准确率为90%。

2) 网络输出的碳化物的图片面积占比是5%,可以粗略推断本次研究的材料中碳化物的体积占比为5%。

3) 直径分布直方图显示,M50轴承钢中一次碳化物MC和M<sub>2</sub>C的尺寸明显比M<sub>23</sub>C<sub>6</sub>的大。MC、M<sub>2</sub>C、M<sub>23</sub>C<sub>6</sub>的费雷特直径均值分别为1.47、0.96、0.45  $\mu\text{m}$ 。

本文验证了深度神经网络作为一种工具,既可以用来识别材料的微观组织,帮助研究者认识学习,也可以定量分析特定的微观组织,帮助研究者判断材料的品质。在分析或处理较多的数据时,深度神经网络明显具有准确、高效的优势。

## 参考文献(References)

- [1] BHADOSHIA H K D H. Steels for bearings[J]. Progress in Materials Science, 2012, 57(2): 268-435.
- [2] KINKUS T J, OLSON G B. Microanalytical evaluation of a prototype stainless bearing steel[J]. Surface Science, 1992, 266(1/2/3): 391-396.
- [3] BARROW A T W, RIVERA-DÍAZ-DEL-CASTILLO P E J. Nano-precipitation in bearing steels[J]. Acta Materialia, 2011, 59(19): 7155-7167.
- [4] KANG J H, RIVERA-DÍAZ-DEL-CASTILLO P E J. Carbide dissolution in bearing steels[J]. Computational Materials Science, 2013, 67: 364-372.
- [5] GUETARD G, TODA-CARABALLO I, RIVERA-DÍAZ-DEL-CASTILLO P E J. Damage evolution around primary carbides under rolling contact fatigue in VIM-VAR M50[J]. International Journal of Fatigue, 2016, 91: 59-67.
- [6] GUAN J, WANG L Q, ZHANG Z Q, et al. Fatigue crack nucleation and propagation at clustered metallic carbides in M50 bearing steel [J]. Tribology International, 2018, 119: 165-174.
- [7] JELITA RYDEL J, TODA-CARABALLO I, GUETARD G, et al. Understanding the factors controlling rolling contact fatigue damage in VIM-VAR M50 steel [J]. International Journal of Fatigue, 2018, 108: 68-78.
- [8] DU N Y, LIU H H, CAO Y F, et al. In situ investigation of the fracture of primary carbides and its mechanism in M50 steel [J]. Materials Characterization, 2022, 186: 111822.
- [9] LIU W F, GUO Y F, CAO Y F, et al. Transformation behavior of primary MC and M<sub>2</sub>C carbides in Cr<sub>4</sub>Mo<sub>4</sub>V steel [J]. Journal of Alloys and Compounds, 2021, 889: 161755.
- [10] 侯志远, 刘威峰, 徐斌, 等. 热加工过程中M50轴承钢一次碳化物的形成与演化[J]. 热加工工艺, 2024, 53(1): 105-110. HOU Zhiyuan, LIU Weifeng, XU Bin, et al. Formation and evolution of primary carbides in M50 bearing steel during hot working [J]. Hot Working Technology, 2024, 53(1): 105-110. (In Chinese)
- [11] 吴玉成, 姜宏伟, 胡园, 等. 多向锻造对M50钢一次碳化物破碎机制的影响[J]. 中国冶金, 2020, 30(9): 98-103. WU Yucheng, JIANG Hongwei, HU Yuan, et al. Effect of multi-direction forging on carbide evolution of M50 steel [J]. China Metallurgy, 2020, 30(9): 98-103. (In Chinese)
- [12] DECOST B L, HOLM E A. A computer vision approach for automated analysis and classification of microstructural image data [J]. Computational Materials Science, 2015, 110: 126-133.
- [13] DURMAZ A R, MÜLLER M, LEI B, et al. A deep learning approach for complex microstructure inference [J]. Nature Communications, 2021, 12(1): 6272.
- [14] YANG J Y, KANG G Z, KAN Q H. A novel deep learning approach of multiaxial fatigue life-prediction with a self-attention mechanism characterizing the effects of loading history and varying temperature [J]. International Journal of Fatigue, 2022, 162: 106851.
- [15] 李维刚, 谌竟成, 范丽霞, 等. 基于卷积神经网络的钢铁材料微观组织自动辨识[J]. 钢铁研究学报, 2020, 32(1): 33-43. LI Weigang, CHEN Jingcheng, FAN Lixia, et al. Automatic identification of microstructure of iron and steel material based on convolutional neural network [J]. Journal of Iron and Steel Research, 2020, 32(1): 33-43. (In Chinese)
- [16] 段献宝, 何惠珍, 李平平, 等. 基于深度卷积神经网络的钢材微观组织分类识别[J]. 铁道车辆, 2022, 60(1): 43-47. DUAN Xianbao, HE Huizhen, LI Pingping, et al. Classification and

- identification of steel microstructure based on deep convolution neural network[J]. *Rolling Stock*, 2022, 60(1): 43-47. (In Chinese)
- [17] 苏晨,任志俊,范彪,等. 基于注意力机制与 ResNet 的残余奥氏体评级研究[J]. *轻工机械*, 2023, 41(2): 78-84.
- SU Chen, REN Zhijun, FAN Biao, et al. Research on retained austenite rating based on attention mechanism and ResNet [J]. *Light Industry Machinery*, 2023, 41(2): 78-84. (In Chinese)
- [18] 彭凡,邹司农,任毅如. 基于深度学习的复合材料螺栓连接失效预测[J]. *机械强度*, 2023, 45(2): 447-453.
- PENG Fan, ZOU Sinong, REN Yiru. Failure prediction of bolted connection of composite materials based on deep learning [J]. *Journal of Mechanical Strength*, 2023, 45(2): 447-453. (In Chinese)
- [19] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 386-397.
- [20] WANG C Y, ZHONG C M. Adaptive feature pyramid networks for object detection [J]. *IEEE Access*, 2021, 9: 107024-107032.
- [21] 吴定海,任国全,王怀光,等. 基于卷积神经网络的机械故障诊断方法综述[J]. *机械强度*, 2020, 42(5): 1024-1032.
- WU Dinghai, REN Guoquan, WANG Huaiguang, et al. The review of mechanical fault diagnosis methods based on convolutional neural network [J]. *Journal of Mechanical Strength*, 2020, 42(5): 1024-1032. (In Chinese)
- [22] KINGMA D P, BA J L. Adam: a method for stochastic optimization [C]// 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. Ithaca: ArXiv, 2015: 1-15.
- [23] HE G Q, HUO Y C, HE M Y, et al. A novel orthogonality loss for deep hierarchical multi-task learning [J]. *IEEE Access*, 2020, 8: 67735-67744.
- [24] 高楚寒,葛思楠,李万明,等. 高速钢碳化物偏析的研究现状[J]. *中国冶金*, 2019, 29(5): 1-5.
- GAO Chuhan, GE Sinan, LI Wanming, et al. Research status of carbide segregation in high speed steel [J]. *China Metallurgy*, 2019, 29(5): 1-5. (In Chinese)

## Research on carbides in M50 bearing steel based on Mask R-CNN deep learning model

SUN Ruiming<sup>1,2</sup> LI Shuxin<sup>1,2</sup> LU Siyuan<sup>1,2</sup> JIN Yongsheng<sup>3</sup> XIAO Huahai<sup>3</sup>

(1. Faculty of Mechanical Engineering and Mechanics, Ningbo University, Ningbo 315211, China)

(2. Part Rolling Key Laboratory of Zhejiang Province, Ningbo 315211, China)

(3. Ningbo Silver-ball Technology Co., Ltd., Ningbo 315207, China)

**Abstract:** The main types of carbides in M50 bearing steel are MC,  $M_2C$  and  $M_{23}C_6$ . Under the scanning electron microscopy (SEM), they exhibit significant differences in the shape, size, and distribution. Some carbides have larger sizes and uneven distribution. They become areas of stress concentration under loading, which has a negative impact on the bearing fatigue performance. So an improved mask region-based convolutional neural network (Mask R-CNN) model was proposed which can batch identify the types of three kinds of carbides in SEM pictures, the diameters of carbides were measured, and the distribution of carbides was showed. The output images and histogram results show that the size of  $M_2C$  carbide in M50 bearing steel is large and unevenly distributed, but the distribution of MC carbide with the largest size and  $M_{23}C_6$  with the smallest size is reasonably uniform.

**Key words:** Deep learning; Mask R-CNN; M50 bearing steel; Carbide

Corresponding author: LI Shuxin, E-mail: lishuxin@nbu.edu.cn

Fund: Natural Science Foundation of China (52075271); Major Science and Technology Task Tackling Project of Ningbo (2022Z050); Cixi Industry-Specific Technology Research and Development Project (CZ2022009)

Received: 2023-11-23      Revised: 2024-01-16