

基于深度学习的钢坯入炉前定位技术

季倬正¹, 邵允学¹, 吕 刚²

(1. 南京工业大学 计算机科学与技术学院, 江苏 南京 211816;

2. 上海策立工程技术有限公司, 上海 201900)

摘 要:随着深度学习技术的不断发展,深度学习与传统工业生产流程相融合成为近些年研究和应用的热点。在传统钢铁生产过程中,钢坯入炉位置是人工根据钢坯的横向位置是否正确来确认和控制的,存在控制不精确、生产效率低、人工成本高等不足之处。针对以上问题,文中通过机器视觉方法实现钢坯在炉前位置的实时精确定位,提出了一种基于钢坯横向位置分割的炉前定位方法,该方法以 U-Net 神经网络为基础结构,用 ResNet-34 网络的标准残差块对原 U-Net 网络进行特征提取与特征融合,这样既保留了 ResNet 的残差结构以加深网络层次,也使这两种网络有效的融合起来。经实验,横向分割网络的像素精度可以达到 99.7%,而标准的 ResNet 网络仅达到 92.2%,达到了现场应用精度要求。

关键词:炉前定位;图像分割;ResNet;U-Net;横向分割网络

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2022)0036-05

Positioning Technology of Billet before Furnace Based on Deep Learning

Ji Tan-zheng¹, Shao Yun-xue¹, Lyu Gang²

(1. School of Computer Science and Technology, Nanjing Tech University, Nanjing 211816, China;

2. Shanghai Celi Engineering Technology Co., Ltd., Shanghai 201900, China)

Abstract: With the continuous development of deep learning technology, the integration of deep learning and traditional industrial production process has become a research and application hotspot in recent years. In the traditional steel production process, the billet feeding position is manually confirmed and controlled according to whether the transverse position of the billet is correct. There are some shortcomings, such as inaccurate control, low production efficiency and high labor cost. To solve the above problems, we realize the real-time accurate positioning of billet in front of the furnace by machine vision method and propose a furnace positioning method based on billet transverse position segmentation. This method takes u-net neural network as the basic structure, and uses the standard residual block of resnet-34 network to extract and fuse the features of the original u-net network. It not only retains the residual structure of RESNET to deepen the network level, but also makes the two networks integrate effectively. Through experiments, the pixel accuracy of the horizontal segmentation network can reach 99.7%, while the standard RESNET network can only reach 92.2%, which meets the accuracy requirements of field application, and lays a foundation for the automation and intelligence of billet production process.

Key words: furnace front positioning; image segmentation; ResNet; U-Net; horizontal partition network

0 引 言

钢铁产业是中国工业发展中的支柱性产业,它是衡量一个国家国力的基本要素之一。目前,中国已经是世界钢铁大国,但距离钢铁强国仍有一定的差距,其中一部分原因在于中国钢铁生产过程中智能化程度偏弱,传统的钢铁产业工人面临着工作环境差、劳动强度大等问题,同时钢厂还会面临因为大量人工操作而导致操作失误或数据出错等问题。

该文主要研究内容是钢坯的入炉前定位问题,钢坯经轨道到达炉前,传统都是人工根据当前钢坯类型

确定入炉的位置,而人工操作经常会出现位置偏差大、钢坯来回挪动影响生产效率等不足,甚至出现连钢、错位等现象。

随着人工智能技术的飞速发展,它与工业的融合进程也在不断加深,该文提出了一种基于机器视觉的钢坯炉前定位方法。该方法是基于深度卷积神经网络对钢坯所在的横向位置进行分割,然后根据分割结果对钢坯的两端进行定位。实验结果和生产现场应用表明,该方法达到了现场应用的要求,提升生产效率和质量的同时,降低了人工成本。

收稿日期:2022-03-18

作者简介:季倬正(1999-),男,硕士研究生,研究方向为计算机视觉、模式识别与机器学习;通信作者:邵允学(1984-),男,博士,副教授,研究方向为计算机视觉、模式识别与机器学习。

1 相关工作

1.1 卷积神经网络

1998 年 LeCun Yann 推出了 LeNet^[1], 这一网络的问世确立了以后卷积神经网络(CNN)的基本结构, 成为了 CNN 的开山鼻祖。但 CNN 并没有从诞生就开始辉煌, 与此同时, 支持向量机^[2](SVM), 因其同样解决了线性不可分的问题, 且相较于神经网络, SVM 更高效, 没有梯度消失的风险且过拟合风险更小, 泛化能力更强。因此深度学习一时间进入了至暗时刻, 直到 2006 年 Hinton^[3] 提出了深度置信网络(DBN), 以及 2012 年赢得了当年的 ImageNet 冠军的 AlexNet^[4] 的推出, 深度学习以及卷积神经网络才重新掀起了热潮。

卷积神经网络是多层感知机的变种, 它有两个重要特点: 局部连接和共享权值。这两个特点使得卷积神经网络成为计算机视觉领域中最热门的网络^[12]。因为图像往往具有局部相关性, 即两个距离较近的像素之间通常具有相关性, 而距离较远的则通常无关, 因此引入卷积核使得每次计算都只在卷积核范围之内进行, 这样与原 BP 神经网络相比, 大大降低了参数规模, 这就是局部连接的好处。而权值共享则是在卷积时, 每一个卷积核将重复作用于整个感受野中, 卷积后的结果就构成了该幅图像的一个特征, 权重共享使得在对图像特征进行提取时不用再考虑局部特征的位置, 同样降低了网络的参数量。

1.2 ResNet

ResNet^[5] 是由何凯明团队于 2015 年提出, 并在当年的 ImageNet 竞赛中以远超过第二名的成绩取得冠军。按理论而言, 网络模型的层次越深, 其提取复杂特征的能力越强, 所表现出来的性能就会越好, 然而实际上, 当网络的深度增加到一定程度后再增加时, 网络反

而会出现退化问题。这主要是由网络的过拟合或梯度消失爆炸引起, 虽然有 ReLU 函数和 Batch Normalization^[6] 缓解了一定问题, 但其退化问题仍然十分严重。

ResNet 的创新点在于为网络模型提供了一条直连通道, 以便于将前一层的输入信息直接传入后层。对于一个堆积网络结构块, 当输入为 x 时, 预计其学习到的输出为 $H(x)$, 则希望其能学习到的残差为 $F(x) = H(x) - x$, 如图 1 所示。而之所以学习残差会比直接学习 $H(x)$ 简单是因为, 即便当残差为 0 时, 原网络也仅是做了一个恒等映射, 这样的跳跃连接也被称为 shortcut。

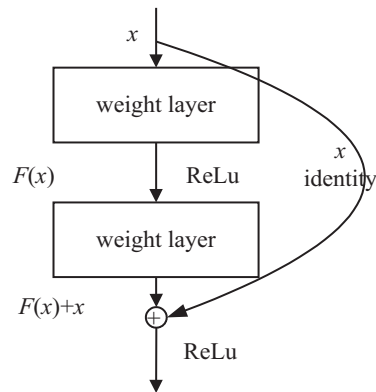


图 1 ResNet 残差块 ($F(x)$ 即为残差)

1.3 U-Net

U-Net^[7] 最早提出于 2015 年, 大量运用于医学影像的图像分割任务, 目前越来越多的语义分割和目标检测任务也关注到了这一 U 型结构。U-Net 继承了全卷积网络(FCN^[8])的思想, 并对其进行改进。U-Net 将一幅图先进行采样编码, 而后解码升采样, 输出其分割结果, 再根据与真实结果之间的比对, 反向传播更新训练网络。

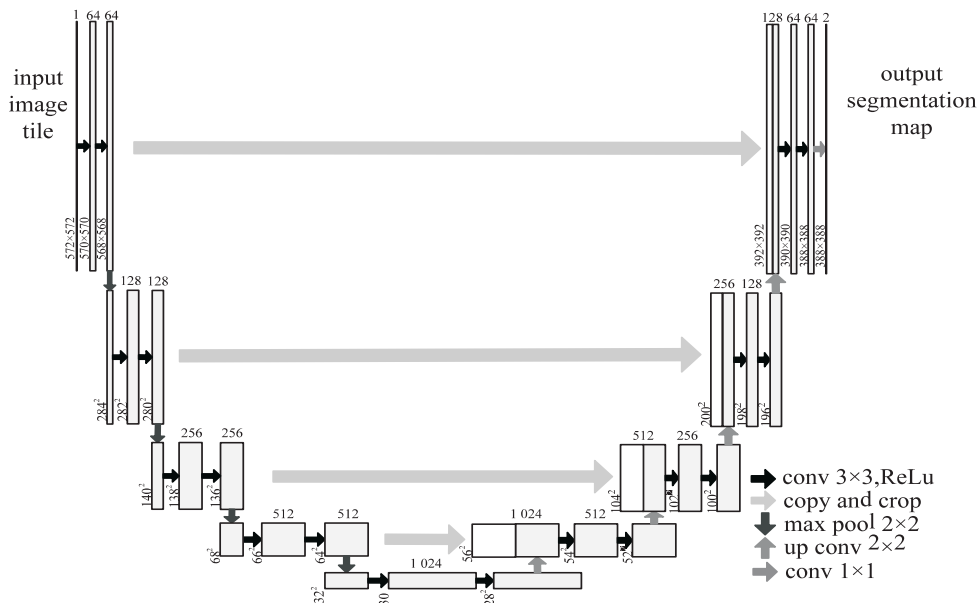


图 2 U-Net 结构

如图 2 所示,网络采用全卷积网络,可大致分为左右两部分。左边的部分为特征提取网络,即 Encoder,可通过卷积和最大池下采样来降低图像尺寸,用以提取一些浅显的特征。右边的部分为特征融合网络,即 Decoder,主要通过卷积和上采样以获得深层次的特征。其中卷积都是采用 valid 填充方式保证特征是在没有丢失上下文信息的情况下得到的。同时 U-Net

也采用了一种不同的融合方式:拼接 (Concat),它在 channel 维度上就将特征拼接在一起,令其在编码阶段获得的 feature map 与解码阶段获得的 feature map 结合起来,同时结合了深层次与浅层次的特征,细化了图像。最后与 1×1 的卷积核卷积做出分割。经典的 ResNet34 模型如图 3 所示。

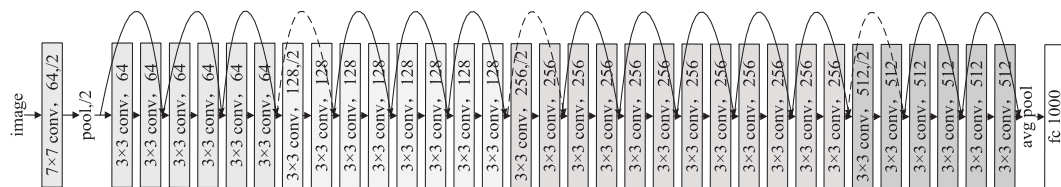


图 3 经典的 ResNet34 模型

2 钢坯定位方法

2.1 问题描述

本机器视觉系统包括一台工业相机、光源和视觉定位模型三部分,相机拍摄的图像如图 4 所示,根据现场应用需求,视觉定位模型的任务可以转化为判断钢坯两头在图像中的像素位置。这一任务可以采用两种方式实现,一种是按照图像分割的思路,由于任务不关心钢坯所在 y 轴坐标位置,所以设计的分割网络仅对钢坯沿 x 轴方向位置进行分割,即网络的输出是 x 轴上每个点对应的位置是否有钢坯,实现思路如图 5 所示,然后根据网络输出判断钢坯两端位置。另一种思路是按照目标检测的思路,直接输出钢坯两端的坐标位置,采用标准的 ResNet 网络实现钢坯端点检测任务,实现思路如图 6 所示。



图 4 相机采集样本示例

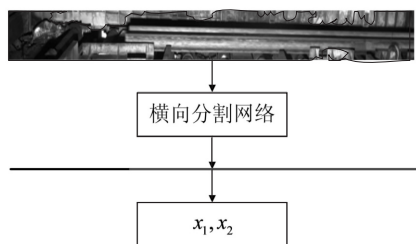


图 5 基于分割网络的钢坯端点检测示意图

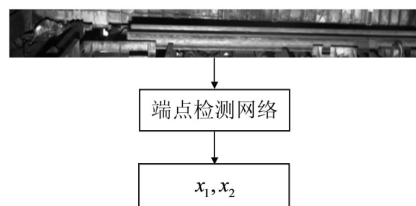


图 6 基于检测网络的钢坯端点检测示意图

2.2 基于 U-Net 的钢坯横向分割网络

分割网络结构融合了 ResNet 和 U-Net 的特点,采

用了 U-Net 的上采样编码与下采样解码部分以及 ResNet 的残差块部分。网络从数据集中获取已经过预处理的图片 ($H \times W \times 1$, 其中 $H \times W$ 代表图片的高宽, 1 代表通道数) 后, 将进入 U-Net 编码器的六个阶段, 特征图信息依次变成 $H \times W \times 4$, $\frac{H}{2} \times \frac{W}{2} \times 8$, $\frac{H}{4} \times \frac{W}{4} \times 16$, $\frac{H}{8} \times \frac{W}{8} \times 32$, $\frac{H}{16} \times \frac{W}{16} \times 64$, $\frac{H}{32} \times \frac{W}{32} \times 64$ 。在下采样完成后, 将最后一层输出 $\frac{H}{32} \times \frac{W}{32} \times 64$ 作为解码器第一阶段的输入, 经过第一阶段后, 输出重新变为 $\frac{H}{16} \times \frac{W}{16} \times 64$, 然后再将解码器这一阶段的输出和编码器有与之相同大小的输出拼接起来作为下一阶段解码器的输入, 与此往复直到输出重新变为 $H \times W \times 8$, 经过两层卷积变为 $H \times W \times 1$, 最后通过自适应均值池化后得到最终输出 $W \times 1$ 。然后再与 label 向量进行比对更新训练网络。与 U-Net 不同的是, 该网络并不是采取原有简单的卷积和最大池下采样来降低图片尺寸, 而是采用了 ResNet 残差块 (如图 7), 输入在残差块内先进行一次卷积和 BatchNorm, 再用 Relu 函数对其进行激活, 再将激活后的输出继续卷积和 BatchNorm, 得到残差 $F(x)$, 将得到的 $F(x)$ 加上本次原输入的 shortcut, 最后对整体进行一次 Relu 激活, 这其中的下采样过程由卷积中的 stride 参数控制完成。

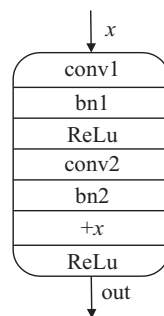


图 7 残差块内部结构

2.3 基于 ResNet 钢坯端点检测网络

标准的 ResNet 网络采用了单纯的下采样,网络的输入仍是经同样预处理后的图片(图片大小 $H \times W \times 1$),其中 $H=300$, $W=1\ 600$,经过 5 次残差块卷积和下采样后(其中残差块结构与图 7 一致),特征图信息分别变为 $H \times W \times 4$, $\frac{H}{2} \times \frac{W}{2} \times 8$, $\frac{H}{4} \times \frac{W}{4} \times 16$, $\frac{H}{8} \times$

$\frac{W}{8} \times 32$, $\frac{H}{16} \times \frac{W}{16} \times 64$,在最后一层输出后,特征图信息变为 $19 \times 100 \times 64$,然后将最后一层多维的特征图输出平铺为一维,再经过一层 Linear 线性变换,最终输出一个 $1 \times 2 \times 1$ 的向量,也可以看成两个点。再将这两个点与 label 标定比对更新网络参数。

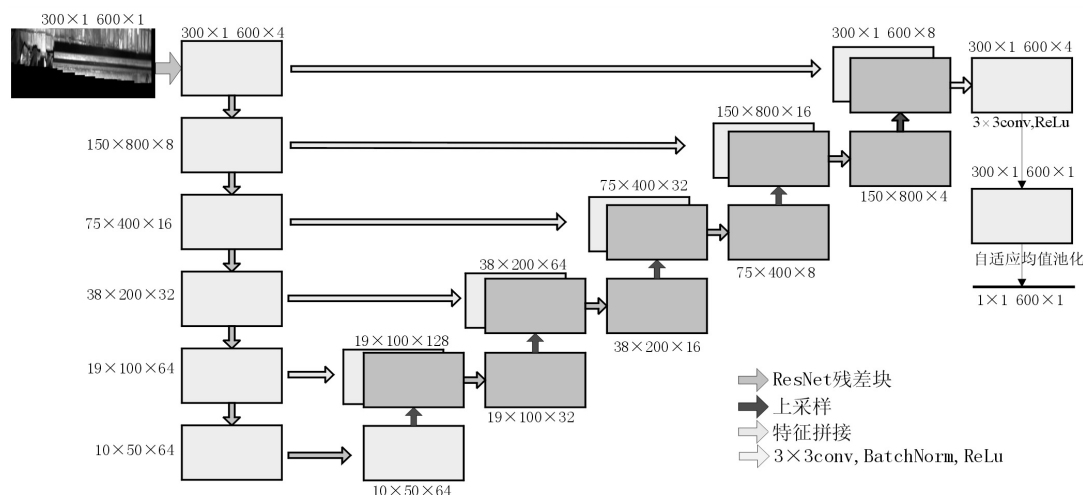


图 8 U-Net 钢坯纵向分割网络示意图

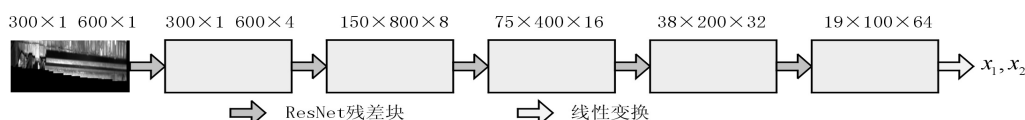


图 9 ResNet 钢坯端点检测网络结构示意图

3 实验结果和分析

3.1 数据集构建

本次数据集照片采集的是某钢厂锅炉前的钢坯的图片,为相机设置好参数,将相机固定在可以看到推钢全过程的地方,使得相机能 24 小时不间断拍摄照片,每隔 1 s 采集一张图片,总共采集了一周时间,然后对这些样本进行筛选,总共筛选出了 34 994 张有效样本,然后利用自己开发了标定软件对图片进行标定。由于本文要完成的任务是对钢坯的端点的定位,所以只需要对钢坯的首尾两端点进行标定即可,如图 10 所示,每张图片标定完成后会生成一个 txt 文本,txt 文本内记录着钢坯在图片中的相对位置信息。标定完成后,我们将整个样本划分成训练集与测试集,其中训练集 31 799 张样本,测试集 3 195 张样本,图片分辨率均为 596×300 。



图 10 样本标定示例

3.2 数据集预处理

由于轨道靠近相机一侧的挡板经常发生变动为了

减少挡板对网络模型的影响,制造了一张掩码图,用来将挡板发生变动的区域去掉,从而消除额外的影响因素,如图 11 所示。



图 11 预处理结束的图片

3.3 训练网络细节

采用相同的设置来进行两种网络的训练,使用的损失函数为均方损失函数 (MSELoss),优化器为 Adam 优化器,设置的初始学习率为 0.000 1,并使用 StepLR 有序调整学习率,设置的 step-size 为 30,即为每迭代 30 次,学习率降低 10 倍,网络的 batch-size 为 32,共计 100 次迭代,使用的 GPU 为 2080Ti,各参数如表 1 所示。

表 1 参数表

Name	Parameters and settings
GPU	RTX 2080Ti
Number of iterations	100
Loss	MSELoss
Optimizer	Adam
Learning rate	0.000 1

4 评价指标

该文采用分割任务中常用的 PA 评价指标来评价提出的两种网络的性能,对于基于检测的网络的输出,首先根据其输出的坐标结果将其转换为和分割网络输出一致的向量,然后再对它们进行评价。

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}}$$

PA 指标常用于评估图像分割网络,代表正确分类像素数量占像素总数量的比例。其中 p_{ii} 表示将第 i 类分成第 i 类的像素数量(正确分类的像素数量), p_{ij} 表示将第 i 类分成第 j 类的像素数量(所有像素数量)。

采用相同的预处理和训练策略,表 2 是这两组实验的训练测试损失变化情况,以及在 PA 指标下两者的表现。

表 2 两种网络的分割结果对比

Network	PA/%	Training Time/Day
Ours	99.7	2.5
ResNet	92.2	2.5

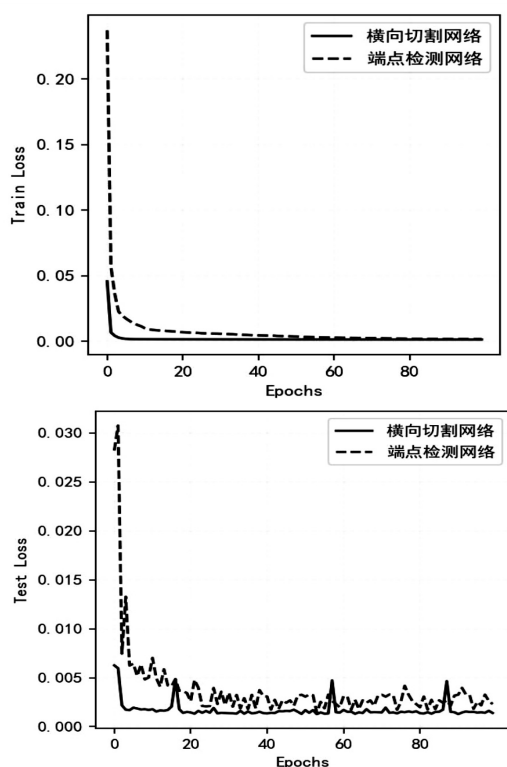


图 12 损失函数值(横坐标代表迭代次数,纵坐标代表训练损失函数值)

图 12 上图是横向分割网络与端点检测网络的训练集损失变化,下图是横向分割网络与端点检测网络的测试集损失变化。从图 12 中的结果可以看出,基于 U-Net 的分割网络收敛速度更快,从 PA 评价指标结果可以看出,基于 U-Net 的分割网络的钢坯定位方法

取得了更高的检测精度,现场应用情况表明,该方法达到了实际应用的要求。

5 结束语

提出了一个融合了 ResNet 残差块与 U-Net 上下采样结构的图像分割网络,并将它用于在钢坯的炉前定位技术上,加快了钢坯生产的自动化进程。指标结果证明它在本实验中拥有比 ResNet 检测网络取得了更好的性能表现,成功的将 PA 精度从 92% 提升到了 99% 以上,达到了实际应用要求。

参考文献:

- [1] 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.
- [2] LUTS J, OJEDA F, PLAS R V, et al. A tutorial on support vector machine-based methods for classification problems in chemometrics[J]. Analytica Chimica Acta, 2010, 665(2): 129-145.
- [3] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communication of ACM, 2012, 60(6): 1-9.
- [5] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE conference on computer vision and pattern recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.
- [6] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd international conference on machine learning. Lille: [s. n.], 2015.
- [7] ZHOU Z W, SIDDIQUEE M M R, TAJBAKSHI N, et al. UNet++: a nested u-net architecture for medical image segmentation[C]//4th deep learning in medical image analysis (DLMIA) workshop. Granada: [s. n.], 2018.
- [8] LONG J, SELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [9] KOO G, YUN J P, SANG J L, et al. End-to-end billet identification number recognition system[J]. ISIJ International, 2019, 59(1): 98-103.
- [10] HSU C Y, KANG L W, LIN C Y, et al. Vision-based detection of steel billet surface defects via fusion of multiple image features[J]. Frontiers in Artificial Intelligence and Applications, 2015, 274(20): 1239-1247.
- [11] ZHAO P, ZHAO X Y, ZHAO C. Image denoising based on

(下转第 57 页)