

# 一种基于端点顺序预测的手写体笔画恢复方法

张 瑞<sup>1</sup> 湛永松<sup>2</sup> 杨明浩<sup>3</sup>

(桂林电子科技大学广西信息科学实验中心 广西 桂林 541004)<sup>1</sup>

(桂林电子科技大学广西可信软件重点实验室 广西 桂林 541004)<sup>2</sup>

(中国科学院自动化研究所模式识别国家重点实验室 北京 100190)<sup>3</sup>

**摘 要** 针对汉字手写体的笔画动态序列恢复问题,文中提出了一种基于端点顺序预测的手写体笔画顺序恢复模型。首先对经过数字化处理后的手写体图像进行细化、笔画片段分割、图像坐标提取和规整等预处理,然后利用预处理后的图像和对应的书写坐标序列生成网络训练的样本,样本由静态手写体图像和包含字体书写顺序的热力图标签组成,该模型采用一种端到端的卷积神经网络结构,最后使用训练好的网络模型对静态手写体图像进行预测,从而得到字体原先的书写顺序。实验结果表明,该方法能够有效地对 5 笔以内的手写字体进行书写顺序的恢复,具有较高的准确率和处理速度。

**关键词** 手写字体,时序信息,深度学习,笔画恢复,卷积神经网络

中图分类号 TP183 文献标识码 A

## Handwritten Drawing Order Recovery Method Based on Endpoint Sequential Prediction

ZHANG Rui<sup>1</sup> ZHAN Yong-song<sup>2</sup> YANG Ming-hao<sup>3</sup>

(Guangxi Experiment Center of Information Science, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China)<sup>1</sup>

(Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China)<sup>2</sup>

(The National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)<sup>3</sup>

**Abstract** To address the problem of dynamic sequential recovery for Chinese handwritten, a handwritten drawing order recovery model based on deep learning method was designed. First, the handwritten image is preprocessed by coordinate regularization, refinement, and interruption of intersections, then the preprocessed image and the corresponding written coordinate sequence are used to generate the sample of the network. The sample consists of a static handwritten image and a heat map label containing the font writing order. The model uses an end-to-end convolutional neural network. Finally, the trained network model is used to predict the static handwritten image to get the original writing order of the font. The experimental results show that the method can effectively recovery the drawing order of handwritten fonts that less than five strokes.

**Keywords** Handwriting, Time series information, Deep learning, Order recovery, Convolutional neural networks

## 1 引言

手写文字的识别是手写字体信息化管理的重要基础,同时也是模式识别中一个重要的领域和关键应用。根据识别对象数据采集方式的不同,手写文字的识别可以分为离线识别和在线识别<sup>[1]</sup>。离线识别所处理的对象是通过扫描设备以及图像采集设备生成的二维静态图像,在线识别所处理的对象是通过手写输入设备记录的动态坐标序列,同时在线识别还有可能用到如书写速度、倾角、书写压力等其他辅助信息。相比离线识别,在线手写文字识别具有更高的处理速度和精度,但在现实应用场景中,由于离线识别所需要的数据获取较为方便,其应用更为广泛。可以对静态的二维图像数据进行动态信息恢复,从而获得识别文字的书写笔画顺序,将离线手写数据转化为在线数据是提高手写文字识别的一种有效方

法<sup>[2-3]</sup>。除此之外,书写动态信息恢复技术在智能书写机器人、智能机械臂等方面也有广泛的应用<sup>[4-5]</sup>。

目前,大多数的动态恢复技术都是基于图模型的方法<sup>[6-11]</sup>。该类方法通常先将待恢复的图像数据进行平滑、细化等预处理,然后对预处理后的图像采用图模型的方式进行建模,将笔画顺序的恢复问题转换为图搜索的问题,即根据某种规则如最小能量代价准则<sup>[9]</sup>,在图上找到符合条件的若干条路径,使得找到的该路径满足某个目标函数的最大化或最小化问题<sup>[12]</sup>。基于图模型的方法的原理简单,便于实现,但在复杂笔画情况下存在搜索时间较长、计算代价大等问题。也有一部分方法采用基于马尔可夫链的概率图模型<sup>[13]</sup>、基于 RNN 的序列生成模型<sup>[14-15]</sup>,基于概率模型和 RNN 序列生成的方法主要是针对英文或拉丁文的手写字体和签字,这些方法在某些特定的英文手写数据集上取得了不错的效果,对于

本文受广西自然科学基金项目(2017GXNSFAA198226),广西重点研发计划(AB17195027, AC16380124, AB18126053),桂林电子科技大学研究生创新教育项目(2018YJCX43)资助。

张 瑞(1991—),男,硕士生,主要研究方向为图像处理、机器学习;湛永松(1979—),博士,教授,CCF 会员,主要研究方向为多媒体技术, E-mail:yszhan@guet.edu.cn(通信作者)。

复杂手写汉字恢复效果并不理想。对于手写汉字来说,由于汉字字符种类较多、手写字体的不规范和书写风格差异化等原因,造成了汉字手写体笔画恢复的难度较高,目前还没有一种通用的方法能有效地恢复手写体汉字。

针对汉字手写字体的笔画恢复问题,本文提出了一种基于 CNN 卷积神经网络<sup>[16-18]</sup>的手写体笔画恢复方法。该方法首先对手写体图像进行细化、笔画片段分割等预处理,从而得到没有交叉点的单像素骨架图,然后利用卷积神经网络高效的特征提取能力,建立字体的笔画特征和书写顺序之间的映射关系,最后利用预先训练好的模型对 5 笔以内的手写汉字进行恢复,获得了比较好的恢复效果,实验结果证实了本文方法的有效性。

## 2 手写体笔画恢复模型

### 2.1 手写体图像的预处理

手写体图像的预处理包括骨架图像的提取、笔画片段分割、图像相应坐标序列的删除 3 个部分。首先从在线手写字体数据库中获取手写字体的图像数据和该字体书写时对应的坐标序列,图像数据为固定大小的二值图像,坐标序列如式(1)所示,坐标序列对应着手写字体的书写过程,坐标统一规整到图像大小的范围。

$$[x_1, y_1], [x_2, y_2], \dots, [x_n, y_n] \quad (1)$$

从在线手写体数据库中获取的手写字体是书写者在数据采集设备上书写的原始字体经数字化处理后的结果,如图 1(a)所示,字体笔画的平均宽度大于单个像素,为了方便地把图像数据转化为带有时序信息的动态序列,可以对字体笔画的宽度进行调整,使笔画的宽度在每一笔中都相同。对手写字体原图进行细化处理可以把原图像中的笔画变成单像素的宽度,常见的手写体骨架提取算法有形态细化法、中轴变换法、索引算法等。本文使用索引算法对手写字体进行细化,算法针对阈值化的二值图像,其主要思想是按一定的规则遍历二值图像的边缘,根据边缘点周围邻接的 8 邻域的像素分布决定是否移除该点,二值图像中点的取值为 0 或 1,因此在点的 8 邻域内可能有 256 种不同的分布情况,索引算法将这 256 种不同的分布即 256 个模板映射到一个大小为 256 的索引表中,在索引表中有两种不同的取值 0 或 1,取值为 1 表示可以删除,0 表示不能删除。图 1(b)给出了使用骨架提取算法对手写字体图像的原图进行细化处理后的结果。细化后的手写体图像中,每个像素在 8 邻域内最多只有 3 个相邻的像素,这样在细化图像中每一笔都是一段有着明确的起始端点和结尾端点的直线或曲线线段,在对该手写字体进行笔画恢复时只要确定了每个笔画端点的顺序就能由笔画的起始点和末尾点来确定该直线的具体位置和书写的顺序。在对字体的图像数据进行细化处理的同时,也对字体对应的坐标序列进行处理,使得细化后的图像和字体的坐标训练相对应。

经过细化算法处理后的手写字体变成单像素宽度,从而能够将字体从图像数据变成代表书写过程的坐标序列,序列中坐标的位置代表了书写过程中笔画的先后顺序,但细化后的手写体图像笔画间存在着交叉点,由于这些交叉点的存在,从字体的书写坐标序列中我们无法确定单独的笔画片段,因此无法确定每一个笔画的起始端点和末尾端点。为了利用笔画片段的顺序恢复整个字体的书写顺序,需要提取出该字体的笔画。笔画间一个重要的特征就是字体书写过程中产生的交叉点,因此可以先找到字体书写过程中产生的所有交叉点,

利用这些交叉点来进行笔画片段的划分,为了便于笔画片段首尾端点的提取,我们将字体从交叉点处打断,同时将交叉点从该字体的笔画序列中删除,最后得到没有交叉的笔画片段图和相应的坐标序列,经过笔画片段分割处理后的结果如图 1(c)所示。得到该字体的书写笔画片段图像和相应的书写坐标序列后,就可以提取每个笔画片段的首尾端点和具体的位置,利用训练样本中笔画片段的先后顺序,可以通过深度学习学习方法学习到相似字形或字体笔画的先后顺序,从而达到笔画恢复的目的。

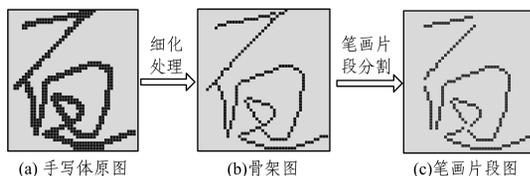


图 1 手写体图像的预处理过程

### 2.2 手写体笔画恢复处理流程

本文提出的基于深度学习的手写字体笔画恢复流程如图 2 所示。该方法的处理流程主要分为两个部分,笔画恢复网络的训练和手写体汉字的笔画恢复。在笔画恢复网络的训练和手写体汉字的笔画恢复。在笔画恢复网络的训练部分对手写字体数据库中的数据进行预处理和样本的生成,然后使用带有笔画片段和相应书写顺序信息的样本对笔画恢复网络进行训练,从而得到一个参数最优化的笔画预测网络。在手写体汉字的笔画恢复部分,固定训练好的网络权重,对要恢复的手写字体图像进行和第一部分相同的预处理过程,然后利用训练好的网络对手写字体图像进行预测,从而得到相应的预测结果,根据预测结果得到该手写字体的笔画顺序。

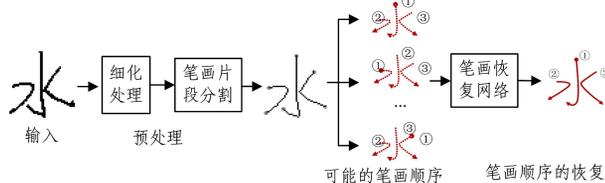


图 2 手写字体笔画恢复流程

笔画恢复网络的训练又分为图像数据的预处理、网络训练样本和相对应的热力图标签的生成、网络参数的训练和优化等。笔画恢复网络使用的图像数据大小为 4848 的二值图像以及该字体书写时的坐标序列。图像中,像素值为 1 的部分为手写字体,像素值为 0 的部分为字体的背景。在图像数据预处理阶段使用图像细化算法提取二值图像的骨架,由于字体的骨架保留了字体的字形和结构信息,因此使用手写字体图像的骨架作为笔画顺序恢复的对象,获得图像的骨架后进一步进行笔画片段分割处理,使得手写字体变成不连续的单像素笔画片段,笔画恢复网络正是对这些笔画片段的书写顺序进行恢复。图像数据所对应的标签是和图像相同大小的二维数组,数组中的值由该字体书写时的坐标序列按书写时间计算得出,最终得到手写体图像相对应的热力图标签。样本生成后使用常用的深度网络优化算法来完成网络权重的优化和更新。

在手写体汉字笔画恢复阶段,对任意给定的手写体图像进行细化和笔画片段分割处理,此时将处理好的图像输入网络中,得到该图像对应的预测值,预测值为原图像大小的数值矩阵,手写体图像中非字体部分在数值矩阵中的数值为 0,字体部分为 0~1 之间的数值,数值的大小代表笔画书写的先后顺序,利用预测值的大小可以得到笔画片段的书写顺序,从而

完成手写字体笔画顺序恢复。

### 2.3 手写体笔画恢复网络

手写体笔画恢复网络采用一种端到端的卷积神经网络模型结构,如图3所示。该网络的输入输出大小相同,网络的输入为经过预处理后的手写体图像,图像是4848的二值图像,其中字体像素部分为1,非字体部分为0。网络的输出为4848大小的数值矩阵,矩阵中有数值的部分对应输入图像中的笔画片段,数值的大小表示该字体像素在书写时的先后顺序。网络的训练过程如式(2)所示:

$$\arg \min_{w,b} \frac{1}{2|D|} \sum \|F(x_i) - y_i\|_2^2 \quad (2)$$

其中, $w,b$ 为网络优化的权重参数, $D$ 为整个训练的数据集, $F$ 为网络模型, $x_i$ 为输入的第*i*个手写体图像, $y_i$ 为输入图像对应的标签,网络使用数据集*D*上所有样本的最小均方差的方法来优化整个模型的参数。

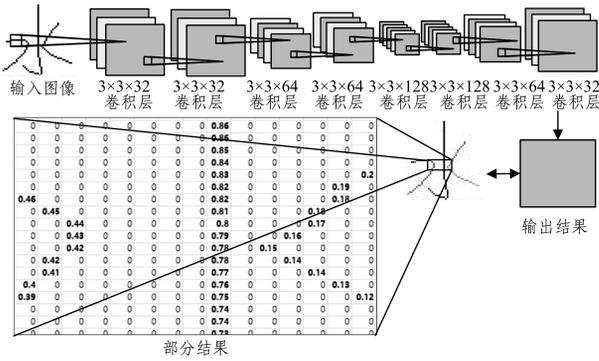


图3 手写体笔画恢复网络结构图

整个网络模型由一系列的卷积层、修正线性激活层(ReLU)和上采样层组成。网络前一部分采用多级卷积层作为图像特征的提取,共有6层级联的卷积层。其中,第一个和第二个卷积层的卷积核个数均为32,大小为 $3 \times 3$ ,卷积步幅为1,第三层卷积核个数为64,大小为 $3 \times 3$ ,卷积步幅为2,第四层卷积核个数为64,大小为 $3 \times 3$ ,卷积步幅为2,第五层卷积核个数为128,大小为 $3 \times 3$ ,卷积步幅为2,第六层卷积核个数为128,大小为 $3 \times 3$ ,卷积步幅为1。为了增加网络的非线性因素,在经过每个卷积层后使用非线性激活层,其中激活层使用修正线性激活函数(ReLU)。在经过第一部分的卷积层提取特征后,特征图的大小会变得与原图像不一致,为了保证输出和输入图像的大小一致,网络中使用上采样来改变特征图的尺寸,因此在第六层卷积层后面加入大小为 $2 \times 2$ 的上采样层,成为第七层卷积层,卷积核个数为64,大小为 $3 \times 3$ ,卷积步幅为1,再对第七个卷积层进行上采样,大小为 $2 \times 2$ ,之后生成第八个卷积层,卷积核个数为32,大小为 $3 \times 3$ ,卷积步幅为1。最后使用卷积核个数为1的卷积层为网络的输出,卷积核大小为 $3 \times 3$ ,网络的输出和输入图像大小一致。网络各层的正向传播公式如下:

$$F_0(x_i) = x_i \quad (3)$$

$$F_l(x_i) = \max(0, W_l * F_{l-1}(x_i) + b_l), l = 1, \dots, L-1 \quad (4)$$

$$F(x_i) = W_L * F_{L-1}(x_i) + b_L \quad (5)$$

### 3 实验结果及分析

为了验证本文笔画顺序恢复方法的有效性,使用深度学习框架Keras搭建本文提出的网络模型,并完成训练样本的处理和网络模型的训练。本文实验的运行环境为64位Windows 10系统,硬件配置为Intel(R) Core(TM) i7-6700K

CPU 4.00 GHz,内存16 GB,NVIDIA GTX1080GPU的台式电脑。

#### 3.1 手写体图像的预处理

本文使用的数据集为CASIA-OLHWDB手写体数据库<sup>[14]</sup>,该数据集包含1020人书写的在线手写中文单字样本和手写样本。其中单字样本分为3个数据库:OLHWDB1.0、OLHWDB1.1、OLHWDB1.2。OLHWDB1.0包括3866个常用的汉字(包含了3740个GB2312中常用的汉字)和171个英文数字符号,分别为420个人书写,共420套,1680258个有效样本。OLHWDB1.1包括了3755个GB2312一级汉字,171个字符,分别为300个人书写,共300套,1172907个有效样本。实验从OLHWDB1.0和OLHWDB1.1中选取了300个不同的人书写的频率较高的100个常用汉字,随机选取其中240个人书写的汉字作为训练集,其余的60个人书写的汉字为测试集。根据每个字体书写的坐标位置和书写次序计算字体相对应的热力图标签,该标签的大小与手写字体图像相同,标签中的数值部分对应着手写字体中的笔画像素部分,数值的大小代表字体书写的顺序。部分训练样本及其相应的热力图标签如图4所示。网络模型的训练过程以训练集内的所有样本的最小均方差为目标函数,以反向传播算法进行网络权重的优化和更新,为了使网络能在一定的时间内收敛,使用带动量的随机梯度学习算法进行权重的优化,训练中随机梯度算法的学习率设置为0.01,衰减率为 $1 \times 10^{-6}$ ,使用带动量的随机梯度学习算法可以使网络在权重更新时尽可能地保持权重更新之前的方向,设置随机梯度学习算法的动量项为0.9。由于训练样本较多,采用批处理的方式进行权重更新,在一定程度上可以加快网络的收敛过程,本文使用的批处理大小为16。



图4 实验中使用的训练样本和相对应的热力图标签

#### 3.2 实验结果和分析

当网络收敛到一定程度后网络损失率会变得相对较小并且在一段时间内网络损失率较为稳定,这时就认为网络已经达到收敛。使用训练集外的测试样本对训练好的模型进行测试,测试集中部分字体的恢复过程如图5所示。

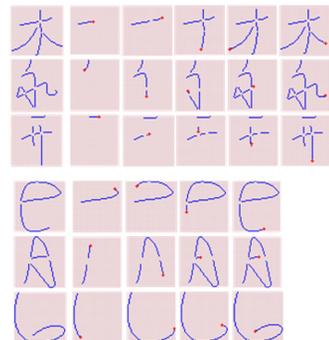


图5 测试集中部分汉字、字母和数字的笔画恢复过程

我们将测试的训练集根据字体是否被网络见过(网络见过的用 $D_o$ 表示,称为字集内,网络没有见过的用 $D_n$ 表示,称

为字集外)、字体是否是训练集中的人写的(书写者的其他字体包含在网络训练集内,用  $U_y$  表示,称为人集内,书写者书写的任何字体都不包含在网络训练集内,用  $U_n$  表示,称为人集外)分为 3 种类型:该字体被网络见过并且是人集外( $D_yU_n$ )、字体没有被网络见过并且是人集内( $D_nU_y$ )、字体没有被网络见过并且是人集外( $D_nU_n$ )。

从 OLHWDB1.0 和 OLHWDB1.1 的测试集中随机选择 60 名书写者书写的 1000 个手写体汉字作为测试样本测试网络笔画恢复的正确率,在选取的 60 名书写者中有 30 名书写者书写的其他字包含在网络训练集中,记为  $D_nU_y$ ,另外 30 名书写者书写的字并没有包含在训练集中,记为  $D_nU_n$ 。表 1 列出了使用本文算法进行笔画恢复的正确率。

表 1 1000 个手写体汉字笔画恢复的结果统计

字体笔画数	0—5 笔	5—10 笔	10 笔以上
数量	289	607	104
$D_nU_n/\%$	67.8	53.6	12.8
$D_nU_y/\%$	75.6	59.4	15.7
平均准确率/ $\%$	71.7	56.5	14.2

由表 1 可以看出,1000 个手写体测试样本中笔画数在 5—10 笔以内的字体最多,共有 607 个,其次是 5 笔以内的字体,共有 289 个,10 笔以上的汉字最少,共有 104 个。对于以上结果,我们可以看出,无论是针对  $D_nU_n$  类型的测试样本还是  $D_nU_y$  类型的测试样本,笔画恢复的准确率随着笔画数的增加逐渐下降,对于 0—5 笔的汉字,网络对笔画顺序进行恢复的平均准确率为 71.7%,5—10 笔的平均准确率为 56.5%,对于 10 笔以上的手写字体,平均准确率只有 14.2%,说明随着手写体笔画数的增加,笔画顺序恢复的难度也逐渐增加,网络恢复的准确率也随之下降,尤其是当手写体笔画数超过 10 笔时,网络进行笔画恢复的准确率急剧下降。同时还可以看出,不同类型的测试集样本对网络笔画顺序恢复的准确率也有影响,在不同的笔画数区间内, $D_nU_y$  这种类型的测试样本(这个字体属于训练集外的字体,但书写者的其他字体包含在训练集内)的准确率高于  $D_nU_n$  这种类型的测试样本(这个字体属于训练集外的字体,并且书写者书写的字体都不包含在训练集内),说明个人书写风格对网络的准确率有一定的影响,如果在训练网络时包含了特定书写者的字体,当网络对该书写者书写的其他字体进行恢复时,相对于完全陌生的书写者书写的相同字体,有较高的恢复准确率。

**结束语** 针对手写体的笔画顺序恢复问题,本文提出了一种基于端点顺序预测的手写体笔画恢复方法,该方法采用端到端的卷积神经网络模型,该模型使用经过细化、笔画片段分割等预处理后的静态手写体图像作为输入,同时输出大小相同的二维数值矩阵,该矩阵可看作手写体图像的热力图标签,利用输出结果中不同笔画片段的数值大小得到笔画片段的书写顺序,从而实现该字体笔画顺序的恢复。通过实验可以得出,本文算法能够对 0—5 笔以内的比较简单的手写字体进行有效的笔画顺序恢复,且恢复的准确率较高,但随着笔画数的增加网络的准确率有所下降,因此对于复杂字体笔画顺序的有效恢复是我们下一步重点研究的方向。

## 参 考 文 献

[1] 金连文,钟卓耀,杨钊,等.深度学习在手写汉字识别中的应用综述[J].自动化学报,2016,42(8):1125-1141.  
 [2] CORDELLA L P,STEFANO C D,MARCELLI A,et al. Writing Order Recovery from Off-Line Handwriting by Graph Trav-

ersal[J]. IEEE International Conference on Pattern Recognition 2010:1896-1899.  
 [3] DINH M,YANG H J,LEE G S,et al. Recovery of drawing order from multi-stroke English handwritten images based on graph models and ambiguous zone analysis[J]. Expert Systems with Applications,2016(64):352-364.  
 [4] LEMAIGNAN S,JACQ A,Hood D,et al. Learning by Teaching a Robot:The Case of Handwriting[J]. IEEE Robotics & Automation Magazine,2016,23(2):56-66.  
 [5] YANG M H,ZHANG K,ZHAO BQ,et al. A Robotic Writing System with Intelligent Interactive Learning Ability [C] // CHCI2017,2017.  
 [6] QIAO Y,NISHIARA M,YASUHARA M. A framework toward restoration of writing order from single-stroked handwriting image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2006(28):1724-1737.  
 [7] BOCCIGNONE G,CHIANESE A,CORDELLA L P,et al. Recovering dynamic information from static handwriting[J]. Pattern Recognition,1993,26(3):409-418.  
 [8] QIAO Y,YASUHARA M. Recovering Drawing Order from Offline Handwritten Image Using Direction Context and Optimal Euler Path[C]// IEEE International Conference on Acoustics, Speech and Signal Processing. 2006:II-II.  
 [9] NAGOYA T,FUJIOKA H. Recovering Human-Like Drawing Order from Static Handwritten Images with Double-Traced Lines[J]. Lecture Notes in Electrical Engineering,2013(253):941-948.  
 [10] NGUYEN V,BLUMENSTEIN M. Techniques for static handwriting trajectory recovery:a survey[J]. International Workshop on Document Analysis Systems,2010:463-470.  
 [11] 曹忠升,苏哲文,王元珍.一种脱机手写汉字书写顺序恢复模型[J].中国图象图形学报,2009,14(10):2074-2081.  
 [12] SHARMA A. Recovery of drawing order in handwritten digit images[C]// IEEE Second International Conference on Image Information Processing. IEEE,2014:437-441.  
 [13] NAKAI M,SHIMODAIRA H,SAGAYAMA S. Generation of hierarchical dictionary for stroke-order free Kanji handwriting recognition based on substroke HMM[C]// International Conference on Document Analysis and Recognition. IEEE,2003(1):514-518.  
 [14] ZHANG X Y,YIN F,ZHANG Y M,et al. Drawing and Recognizing Chinese Characters with Recurrent Neural Network[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2018(99):1-1.  
 [15] BHUNIA A K,BHOWMICK A,BHUNIA A K,et al. Handwriting Trajectory Recovery using End-to-End Deep Encoder-Decoder Network[C]// 2018 24th International Conference on Pattern Recognition (ICPR). Beijing,China:IEEE,2018.  
 [16] LECUN Y L,BOTTOU L,BENGIO Y,et al. Gradient-Based Learning Applied to Document Recognition[J]. Proceedings of the IEEE,1998,86(11):2278-2324.  
 [17] HE K,ZHANG X,REN S,et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. 2016:770-778.  
 [18] LECUN Y,BOSER B E,DENKER J S,et al. Handwritten digit recognition with a back-propagation network[C]// Advances in Neural Information Processing Systems. San Francisco, CA, USA,1990:396-404.