

DOI: 10.11992/tis.201805002

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180928.1338.002.html>

# 基于卷积神经网络的盲文音乐识别研究

刘彪<sup>1,2</sup>, 黄蓉蓉<sup>1</sup>, 林和<sup>1</sup>, 苏伟<sup>1</sup>

(1. 兰州大学 信息科学与工程学院, 甘肃 兰州 730000; 2. 解放军 69230 部队, 新疆 乌苏 833000)

**摘要:** 盲人音乐家在交流创作的音乐作品时面临着人工转换和效率较低的问题, 信息科学技术的迅速发展为解决此类问题提供了许多解决方案。虽然目前有许多盲文音乐作品的识别方案, 但其存在识别效率低和兼容能力不足等缺点。为了避免传统方案在盲文音乐图片特征提取时过多依赖人工经验, 通过研究提出并设计了基于卷积神经网络的识别模型。在对盲文音乐图片的样例数据进行预处理之后, 通过多次反复迭代训练, 模型就可学习到盲文音乐图片中音乐符号的特征。实验结果表明, 该模型的识别有效性和较强的泛化能力为盲文音乐作品的识别提供了一种新的解决方案。

**关键词:** 机器学习; 盲文音乐识别; 卷积神经网络; 深度学习; 计算机视觉; 图像识别; 人工智能; 图像处理  
**中图分类号:** TP39   **文献标志码:** A   **文章编号:** 1673-4785(2019)01-0186-08

中文引用格式: 刘彪, 黄蓉蓉, 林和, 等. 基于卷积神经网络的盲文音乐识别研究 [J]. 智能系统学报, 2019, 14(1): 186-193.

英文引用格式: LIU Biao, HUANG Rongrong, LIN He, et al. Research on braille music recognition based on convolutional neural networks[J]. CAAI transactions on intelligent systems, 2019, 14(1): 186-193.

## Research on braille music recognition based on convolutional neural networks

LIU Biao<sup>1,2</sup>, HUANG Rongrong<sup>1</sup>, LIN He<sup>1</sup>, SU Wei<sup>1</sup>

(1. School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China; 2. No.69230 Troops of PLA, Wusu 833000, China)

**Abstract:** Blind musicians are confronted with the problems of manual conversion and low efficiency in the communication of musical works. The rapid development of information science and technology has provided many solutions to these problems. However, most of the recognition schemes for braille music works lack recognition efficiency and compatibility. In consideration of this deficiency, whereby traditional schemes rely heavily on artificial experience in braille music picture extraction, a convolution neural network-based recognition model has been developed. After preprocessing the sample data of braille music pictures through repeated iterative training, the recognition model was able learn the characteristics of music notation in braille music pictures. The experimental results showed the recognition effectiveness and strong generalization ability of the model, which provides a new idea for the recognition of braille music works.

**Keywords:** machine learning; braille music recognition; convolution neural network; deep learning; computer vision; image recognition; artificial intelligence; image processing

我国的盲人数量约为 752 万人, 是全球盲人数量最多的国家<sup>[1]</sup>。随着从事音乐工作的盲人越来越多, 盲文成为盲人与常人进行书面交流的主

要方式, 但用其创作的音乐作品仍然需要人工的转换和翻译, 所以急需一种自动盲文音乐识别方案。文献 [2] 最早使用硬件设备采集普通盲文内容的图片, 虽然可通过光学识别出相应内容, 但是缺乏与用户的交互性。文献 [3] 设计了一个单面盲文图片的智能识别系统, 包括图像采集、噪

收稿日期: 2018-05-04. 网络出版日期: 2018-09-30.

基金项目: 广西科技计划项目 (桂科 AA17204096, 桂科 AD16380076).

通信作者: 林和. E-mail: [linhe@lzu.edu.cn](mailto:linhe@lzu.edu.cn).

声消除、图像分割、特征提取和字符识别等步骤,其识别率达到94.39%,但使用了固定盲文单元大小和盲文点的位置来识别盲文符号,其可扩展性极其有限。文献[4]提出了一种用于识别双面盲文内容的高效算法,即通过对凸点和凹点的灰度值变化来区分正面点和反面点,识别率虽然比较乐观,但其局限性在于当正面点和反面点重叠较多时,对每个盲文点的质心检测和识别会出现很多歧义,其识别效果也因此受到较大影响。文献[5-7]采用传统的方法对盲文音乐符号进行识别,其中文献[7]以MusicXML作为中间桥梁来实现盲文的转换,并由此设计了从MusicXML到盲文乐谱ASCII码的自动转换软件。文献[8]提出一种基于深度学习模型识别盲文字符的方法,使用堆叠去噪自动编码器来解决盲文识别过程中特征自动提取与降维等问题,利用SDAE自动学习盲文点字的图片特征,使用Softmax分类器进行盲文识别。文献[9]提出将双面阿姆哈拉盲文图片翻译为阿姆哈拉文的系统。该系统使用方向场张量从背景中分割盲文点,使用梯度场识别正面点和反面点;对于重叠盲文点的分割是该文的核心内容,使用盲文点的属性(质心和面积)来分割重叠点并且提出了相应的正面点和反面点识别算法。文献[10]使用人工神经网络对单面的阿姆哈拉图片进行识别,采用自适应直方图均衡化和形态学操作相结合来对文件中的不同等级的噪声进行降噪处理,并在识别阶段采用人工神经网络和二进制编码相结合的形式进行盲文图像识别,识别率达到95.5%。文献[11]采用BP神经网络对英语盲文数字(0~9)图像进行识别。该文展示了BP神经网络的参数设置以及图像的特征提取过程,其训练精度达到97.1%,测试精度达到85%。文献[12]在FCM和KNN的启发下,提出了一个模糊的聚类算法和一种新的数字特征检测算法,称之为斑点检测算法,即在4个扫描方向上扫描1个盲文单元,从而生成1个八进制编码。然后模糊分类算法根据生成的八进制编码来识别相应的盲文字符,识别精度达到83%。上述这些研究存在识别过程较烦琐、识别度较低、泛化能力较差等问题。

为了使研究具有较强泛化能力和较高识别度,同时简化其识别过程,本文从盲文音乐图片的结构出发,应用卷积神经网络的识别模型,并在Tensorflow深度学习框架下进行实验。实验表明,采用卷积神经网络模型来识别盲文音乐图片具有较为乐观的结果。

## 1 识别模型

卷积神经网络(convolutional neural network, CNN)因其具有独特的结构<sup>[13-15]</sup>,在进行图像特征提取的同时,还能够提取更细节的图像信息。这样不仅解决了多数的传统神经网络中参数较多、训练缓慢的难题,而且还防止了过拟合的出现。自2016年AlphaGo战胜世界围棋冠军李世石以来,卷积神经网络再一次被推向浪潮,尤其是在计算机视觉领域的研究更是大热。卷积神经网络具有两个很重要的特点:

1) 共享权值<sup>[16-17]</sup>。传统的神经网络中每一层的权值 $w$ 仅使用一次,而当下一次使用时便会重新生成不同的权值 $w$ 。但在卷积网络中,卷积核要与图像中的每个像素值(输入向量)进行卷积运算,所以仅需要一组权值即可。当输入向量与该组权值运算完毕时,则表明该次卷积操作结束。共享权值的设计在前向传播阶段并未缩减消耗时间,但是却在某种程度上将整个模型所需要的权重参数的数量大大降低,很大程度上提高了计算机的运算性能。在进行卷积运算时,卷积核按照给定的步长 $s$ 在输入图像上从左到右、从上到下依次进行滑动,直至操作结束。相对于传统的神经网络而言,卷积神经网络的参数数量不仅降低了,而且其运算速度也在一定程度上提高了。

2) 稀疏连接<sup>[18]</sup>。为了挖掘图像空间局部关联的信息,卷积神经网络采用了通过加强神经网络中相邻层之间节点的局部连接模式,而摒弃了全连接的方式,即采用核少于输入的方式来完成。例如,如果有 $m$ 个输入 $n$ 个输出,传统的神经网络会把每个输出与每个输入进行矩阵乘法运算,则时间复杂度为 $O(m \times n)$ ,而卷积神经网络只提取其中有意义的 $k$ 个输入,其时间复杂度为 $O(k \times n)$ ,因为在实际应用中, $k$ 一般远小于 $m$ ,并更有实际意义,这一方面降低了时间复杂度,也提升了存储的效率。

## 2 盲文音乐识别方法

在传统的盲文音乐识别方法中,人工介入的工作量较大。尽管目前部分传统模型提出的算法在一定程度上能够实现对盲文图像的特征提取,但是对于以点状为基本组成且难以分辨的盲文音乐图片而言,其识别的效果并不理想,这归咎于其弱学习能力和弱适应性。而相比之下,近年来发展迅速的CNN的效果却较为乐观<sup>[19-20]</sup>,表现出较高的识别精准率和较强的泛化能力。

## 2.1 图像预处理

盲文音乐图片的分辨率因受多种因素的影响而不尽相同,为了达到良好的识别效果,还需要进行一定的图像处理过程。本文在对盲文音乐图像识别前,先对图像进行了预处理操作,包括降噪处理、倾斜校正、形态学处理、水平/垂直投影分割、点位标记、盲文单元切割等,其预处理前后的对比效果如图1所示。

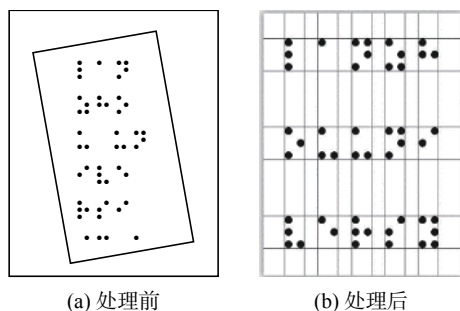


图1 盲文音乐图片倾斜校正

Fig. 1 Braille music image before tilt correction process

## 2.2 模型的构建

图2为盲文音乐图像识别的卷积神经网络模型结构。该识别模型有1个输入层,3个卷积层( $C_1$ 、 $C_3$ 、 $C_5$ ),2个池化层( $S_2$ 、 $S_4$ ),2个全连接层以及1个输出层<sup>[21]</sup>。卷积层的作用是:对盲文音乐图像进行特征提取,进而确定其周围特征参数的位置关系,实验采用多层卷积层,以便可以获得更深层次的特征图。池化层是特征映射层,其主要作用是:对输入的特征图进行压缩,简化网络计算复杂度,提取主要特征。输出层即展示盲文音乐图像的分类结果。

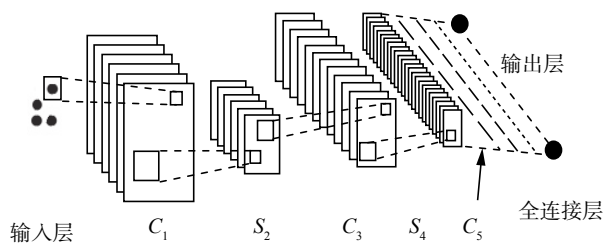


图2 识别模型结构图

Fig. 2 Identification model structure

1) 输入层。将预处理后的图像作为该模型的输入。图3为训练集中的多个样例图片,每个图像的大小为 $32 \times 64$ ,即该识别模型的输入维度为2048。



图3 训练集中的样例图片

Fig. 3 Sample images in the training set

2) $C_1$ 层。主要作用是首次进行提取图像的特征参数,它会过滤掉多余的特征。 $C_1$ 层与输入层传入的数据进行模板匹配运算,从而得到了给定数量的输出特征<sup>[22]</sup>。卷积核的作用并不是单纯地计算卷积操作后的结果,而是在给定的权重矩阵下对输入的图像数据进行匹配运算操作。若卷积核尺寸过大,网络参数会增多,网络速度会降低,过小则会影响特征提取的精度,所以卷积核的大小选择尤为关键。在此次卷积神经网络的识别模型中,每个卷积层均采用大小为 $3 \times 3$ 的卷积核,可以降低该模型的参数数量。给定输入层的维度为 $32 \times 64$ , $C_1$ 层的卷积核大小设置为 $3 \times 3$ ,并设置32个卷积核,步长为1。图4为卷积神经网络的卷积操作过程。

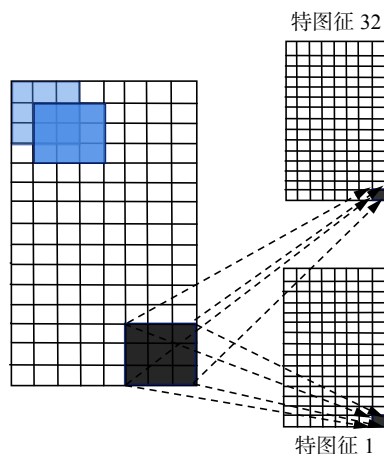


图4 卷积操作示意图

Fig. 4 Convolution operation diagram

通过该卷积操作就可获得大小为 $30 \times 62$ 的特征图。为了保证卷积前后的图像大小保持一致,在这里采用“same”卷积方式,从而保证卷积后得到特征图的大小仍为 $32 \times 64$ 。为了保证不丢失图像的边缘信息,在后面的卷积层( $C_3$ 、 $C_5$ )都采用该操作。

由于每个滤波器会得到一个特征图,经过 $C_1$ 层的卷积操作,如图5所示,图2中的盲文音乐图像可得到32个不同的特征图。某些特征图中的盲文点边界比较清晰,例如第1行第7个特征图,这说明该卷积层提取到了图像中盲文音乐符号的主干信息。其次也可以看出每个特征图的具体状态迥异,这是因为每个卷积核的权值信息不同,因而学习到的特征不同。

3) $S_2$ 层。该池化层的主要作用是对特征的具体位置进行淡化。当提取某个特征之后,只需了解该特征与其他特征的相对位置,例如图3中的盲文音乐字符,当得到上层的特征之后,就不用再考虑其字符的具体位置,只要得到下面对应的

特征,并将上下二者有效地组合在一起便是要识别的盲文音乐字符。

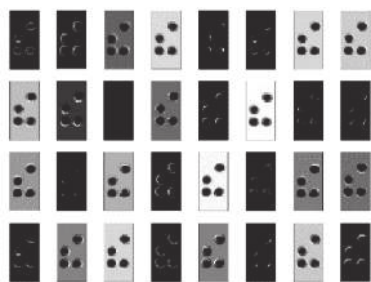


图5  $C_1$  卷积后的特征图

Fig. 5 The feature map after  $C_1$  convolution

在卷积神经网络中,特征所对应的具体位置将不再重要。对  $C_1$  层卷积后的结果进行池化操作,即得到 32 个大小为  $16 \times 32$  的结果,输出的维度相比该层输入的维度缩减了一半。该层是通过池化操作实现二次特征提取的,有效地减少了输入数据的参数数量,降低了图像的大小,简化了网络计算的复杂度。同时,它还可减弱卷积层的输出在角度变形或拉伸变化方面的敏感程度,进行主要特征提取。

如图 6 为整个池化操作过程,该操作是利用池化参数来进行最大下采样操作,池化层中的步长  $s$  设置为 1,其权值矩阵设置为  $2 \times 2$ 。在  $C_1$  层的输出特征图中,对每个特征图中进行下采样操作,根据图 6 所示此规则输出  $S_2$  层的输出结果,但是尺寸大小缩小为原来的一半<sup>[22]</sup>。

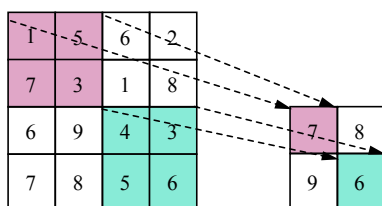


图6 下采样操作

Fig. 6 Down sampling operation

$S_2$  层的下采样操作后得到如图 7 所示的 32 个尺寸大小为  $16 \times 32$  的特征图。从图 7 中可以看出,除了图像的尺寸大小发生了变化之外,还忽略了部分盲文音乐符号的细节信息。

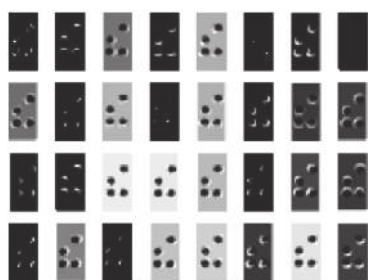


图7  $S_2$  池化操作后的特征图

Fig. 7 The feature map after  $S_2$  pooling

4)  $C_3$  层。 $C_3$  层的卷积操作建立在  $S_2$  层的基础上,其卷积操作与  $C_1$  层类似。主要是对  $S_2$  层的输出进行再卷积操作,进行特征再提取。在  $C_3$  层中,将卷积核的数目设置为 64,即可得出 64 个大小为  $14 \times 30$  且图像内容各不相同的特征图。在该层使用“same”卷积操作后,最后得到特征图的尺寸大小为  $16 \times 32$ 。值得注意的是,输入与输出的特征图虽然在尺寸上大小相同,但是最终得到的特征图数目却不一样,其中 32 个  $S_2$  层的特征图作为  $C_3$  层的输入,最终得出了如图 8 所示的 64 个  $C_3$  层的结果特征图。在图 8 中可以发现,其盲文点的边缘信息越来越突出。

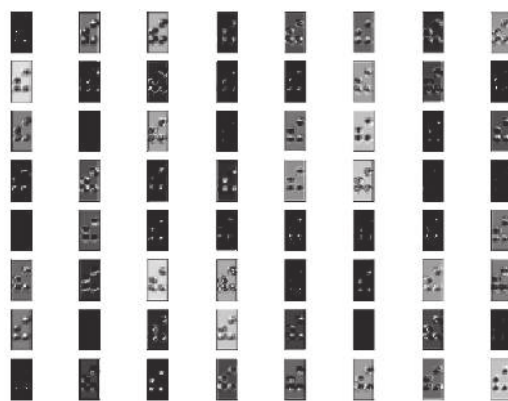


图8  $C_3$  层卷积操作后的特征图

Fig. 8 The feature map after  $C_3$  convolution operation

图 9 所展示的是  $S_2$  层与  $C_3$  层之间的关联和映射,从图中可以看出这些连接并不是一一对应的关系。卷积神经网络这样设计的优点是:消除了网络的对称性,从而让映射的数量保持在一定的范围之内;增加了特征映射的多样性,从而实现了不同特征的有效提取和传递。

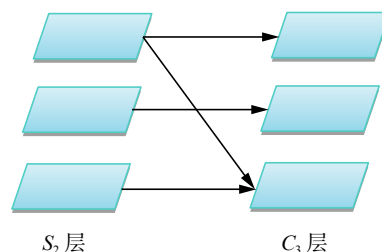


图9  $S_2$  层和  $C_3$  层的连接图

Fig. 9 Connection diagram of  $S_2$  and  $C_3$

5)  $S_4$  层。该层是识别模型最后一个池化层,  $S_4$  层是对  $C_3$  层的输出结果进行池化操作,其操作与  $S_2$  层的操作类似。通过  $S_4$  层的下采样操作得到如图 10 所示的 64 个大小为  $8 \times 16$  的特征图。

6)  $C_5$  层。 $C_5$  层是识别模型的最后一个卷积操作,作用是进一步提取图像特征。该操作与之前

的卷积操作类似,但是此处将卷积核的个数设置为128。因此得到128个大小为 $8 \times 16$ 的特征图,如图11所示。

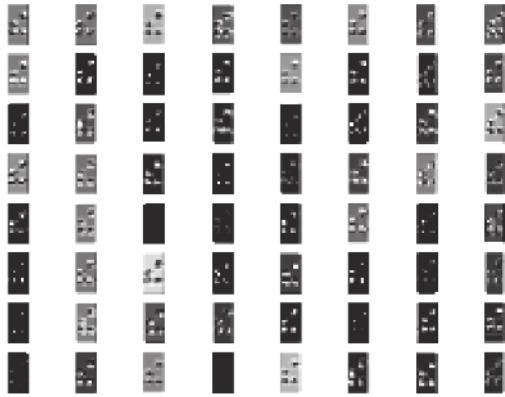


图10  $S_4$ 层操作后的特征图

Fig. 10 The feature map after  $S_4$  pool

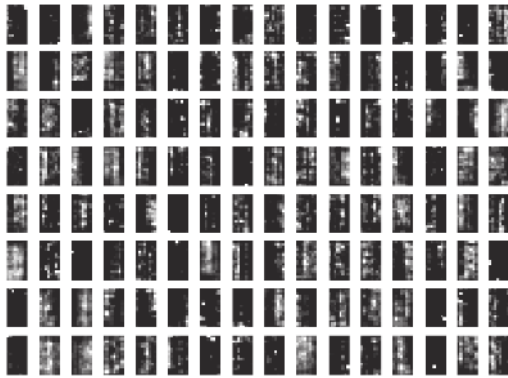


图11  $C_5$ 层卷积后的特征图

Fig. 11 The feature map after  $C_5$  convolution

7) 输出层。为了确保输出层的分类数量与实际64个盲文符号数相一致,实验的模型在最后采用了两个全连接层使得卷积层 $C_5$ 能够和输出分类关联上,即输出层有64个神经元数量。在输出结果时,本文使用softmax函数对盲文音乐符号进行分类,用dropout函数来防止过拟合现象。

### 3 图像数据集

为了增强此次盲文所训练的基于卷积神经网络识别模型的泛化能力,前期需要大量具有代表性的训练样本,供其学习并获得可观的模型。然而,由于盲文本身的局限性图像化的数据集更是相对匮乏,目前公认的数据集主要来源于:1) 线下纸质的扫描文件,主要通过现有的OCR技术对图像中的盲文字符进行提取,但这种提取质量很大程度上受限于扫描文件本身的质量(分辨率)、OCR技术的缺陷(光线敏感)、纸质盲文的规整性;2) 线上互联网共享的盲文图片,图片的分辨率、位深等属性较为可观,通常可以取得较好的识别效果;3) 电子版音乐文件,可以通过预处理

获得较高质量的盲文图片,组成较好的数据集。另外,受限于版权,实验采集到的前2种盲文音乐图片数量极其有限,主要是采用第3种盲文音乐图片作为训练集,后期的研究可以在取得相应许可后进行大规模数据的采集,以便于模型本身能够有更好的兼容性。

本文采集的6400张经过预处理后的盲文音乐符号图片(10个盲文数字、28种常用符号和26个英文大写字母)中,总共有64个盲文音乐符号,每个符号对应100张图片。对于每个盲文音乐符号,实验随机选择其中的80张作为训练集,剩余的20张作为测试集。其主要的划分情况如表1所示。

表1 训练集和测试集

Table 1 Training sets and test sets

序号	盲文符号	训练集/张	测试集/张
1	数字0~9	800	200
2	常用符号	2 240	560
3	英文字母	2 080	520

现实中电子版的盲文字符表现形式有多种,主要形式如图12所示,该图展示了测试集中的样例图片。

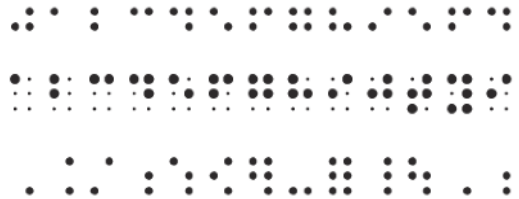


图12 测试集中的样例图片

Fig. 12 Sample images in the test sets

## 4 算法的设计

结合构建的基于卷积神经网络的盲文音乐识别模型,本节编写出相应的盲文音乐图像训练算法以及测试算法,对应的伪码描述分别如算法1和算法2所示。

### 算法1 CNN盲文识别模型训练算法

输入 TRAIN\_IMAGES 是训练集, EPOCH 是迭代次数, BATCH\_SIZE 是迭代量, TRAIN\_ERROR 是错误率;

输出  $y$  是分类结果。

1) random( $w$ );

2) random( $b$ );

3) for  $i=1$  in EPOCH:

4)  $x = \text{BATCH\_SIZE}$ ;

5)  $y_{\text{true}} = \text{batch\_y}$ ;  $x$  中每个元素对应的正确值  $y_{\text{true}}$ ;

6)  $y = \text{batch\_x} \times w + b$ ; 经过模型预测的值;  
 7)  $\text{Error} = \text{error}(\text{batch\_y}, y_{\text{true}})$  正确值与预测值的错误率;  
 8) if  $\text{Error} > \text{TRAIN\_ERROR}$   
 9) update ( $w$ ); update ( $b$ )  
 Continue step 3;  
 10) else  
 11) update ( $w$ ); update ( $b$ );  
 12) end  
 13) end  
 14) 保存训练后的 CNN 模型  $y$ , 其中  $y = x \times w + b$ ;  
 15) end  
**算法 2 CNN 盲文识别模型测试算法**  
**输入** TEST\_IMAGES 是测试数据集, TEST\_SIZE 是测试数据量;  
**输出** test\_accuracy 是正确率。  
 1) random ( $w$ ); random ( $b$ );  
 2) load 训练模型为  $y$ ;  
 3)  $\text{test\_x} = \text{TEST\_SIZE}$  的数据作为输入数据;  
 4) 获取测试图片数据  $\text{test\_x}$  与对应的标签值  $y_{\text{true}}$ ;  
 5)  $\text{test\_y} = \text{test\_x} \times w + b$ ;  
 6) 其测试正确率为  $\text{test\_accuracy} = \text{accuracy}(\text{test\_y}, y_{\text{true}})$ ;  
 7) output test accuracy

## 5 实验结果分析

在该阶段, 为了验证实验建立的卷积神经网络识别模型的有效性, 实验中的权重参数的初始化一律采用随机数。其次, 为了使得模型在训练时不过早地陷入饱和, 影响卷积神经网络的学习能力, 实验采用的随机数会比较小。

本次实验中的训练和测试都以分批迭代的方式来进行的, 分别迭代 200 次训练, 并且每 20 次输出测试数据的正确率。通过训练和测试, 正确率的变化趋势如图 13 所示, 通过分析可见在迭代 140 次后该模型的正确率走向趋于平稳, 并接近 1。这是因为: 盲文音乐符号的结构相对而言较简单 (以点状为主要特征), 且在训练和测试时使用的图片的分辨率较高, 即图像的质量较好; 其次, 识别模型对盲文音乐图片中盲文音乐符号的特征提取较为容易, 从而可以快速地进行盲文音乐符号的识别。本次实验也存在许多不足, 例如: 受限于版权, 数据集资源过小, 来源也较单

一。这些都需要在后期的研究中进行改进, 从而达到更好识别性能。

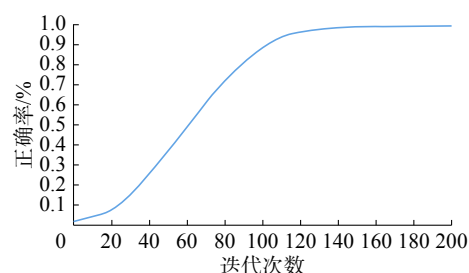


图 13 正确率与迭代次数关系

Fig. 13 The relationship between the correct rate and the number of iterations

就总体识别效果而言, 本文所训练的模型的识别准确度优于多数流行的方法, 其对比图如表 2 所示。因为本文的盲文音乐图片是基于单面形式的盲文图片, 所以此处的对比只限于单面盲文图片的识别对比。将 CNN 模型与近几年来所使用的前馈神经网络<sup>[10]</sup>、BP 神经网络<sup>[11]</sup>、模糊分类算法<sup>[12]</sup>、标准距离定位法<sup>[23]</sup>作对比, 由此可见本文中盲文音乐图像的识别正确率均高于其他方法。

表 2 正确率对比

Table 2 Correct rate comparison

序号	方法	正确率/%
1	前馈神经网络	95.5
2	BP神经网络	85
3	模糊分类算法	83
4	标准距离定位法	95.9
5	CNN识别模型	98.7

## 6 结束语

本文首次尝试了将卷积神经网络识别模型应用于盲文音乐图片。在识别过程中, 首先对盲文音乐图片进行了预处理操作, 以便于获得更好的输入图片。其次, 本文展示了通过卷积层、池化层等处理后的盲文音乐图片的特征图。在展示特征图的同时, 对识别模型的细节也进行了说明, 例如关键参数的设定。最后, 通过对实验结果进行分析, 可以发现该模型可以很好地提取到盲文音乐符号的特征, 对盲文音乐具有良好的识别效果。当然本文也存在诸多不足之处, 例如对数据集的选取较单一, 若可以采集到现实生活中纸质的盲文书籍, 那么数据集就较为充实, 该识别模型就可以广泛用到生活中, 可以减少盲文识别对人工操作的依赖性。

## 参考文献:

- [1] MARIOTTI S P. Global data on visual impairments 2010[R]. Geneva: World Health Organization, 2012: 1–17.
- [2] ANTONACOPOULOS A, BRIDSON D. A robust braille recognition system[C]//6th International Workshop on Document Analysis Systems. Florence, Italy, 2004: 533–545.
- [3] MOUSA A, HIARY H, ALOMARI R, et al. Smart braille system recognizer[J]. IJCSI international journal of computer science issues, 2013, 10(6): 52–60.
- [4] SHREEKANTH T, UDAYASHANKARA V. An algorithmic approach for double sided Braille dot recognition using image processing techniques[J]. International journal of image processing and visual communication, 2014, 2(4): 1–6.
- [5] 王倩倩. 基于 DSP 的简谱识别和盲文乐谱转换算法研究与实现 [D]. 沈阳: 东北大学, 2013.  
WANG Qianqian. Research and implementation of numbered musical notation identification and braille musical translation algorithm based on DSP[D]. Shenyang: Northeastern University, 2013.
- [6] 黄河燕, 周强, 沈鑫, 等. 基于对节记谱法的盲文乐谱自动排版方法及装置: CN105719626A[P]. 2016-06-29.  
HUANG Heyan, ZHOU Qiang, SHEN Xin, et al. Automatic braille music score typesetting method and device based on by-rhythm stave memorizing method: CN105719626A[P]. 2016-06-29.
- [7] 张志强. 视障乐谱转换软件 SunMusic 的设计实现 [D]. 兰州: 兰州大学, 2016.  
ZHANG Zhiqiang. The design and implementation of music conversion software SunMusic for visually impaired [D]. Lanzhou: Lanzhou University, 2016.
- [8] 李婷. 基于深度学习的盲文识别方法 [J]. 计算机与现代化, 2015(6): 37–40.  
LI Ting. A deep learning method for braille recognition[J]. Computer and modernization, 2015(6): 37–40.
- [9] ALI H S, ASSABIE Y. Recognition of double sided Amharic braille documents[J]. International journal of image, graphics and signal processing, 2017, 9(4): 1–9.
- [10] JIBRIL E C, MESHESHA M. Recognition of Amharic braille documents[C]//The 5th International Conference on the Advancement of Science and Technology. 2017, 15: 22–29.
- [11] WALEED M. Braille identification system using artificial neural networks[J]. Tikrit journal of pure science, 2017, 22(2): 140–145.
- [12] AL NASSIRI A, ABDULLA S. A fuzzy classification and recognition system for arabic braille segmented characters[J]. International journal of applied engineering research, 2018, 13(6): 3662–3669.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, 2012: 1097–1105.
- [14] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv: 1409.1556, 2014.
- [15] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 770–778.
- [16] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述 [J]. 计算机学报, 2017, 40(6): 1229–1251.  
ZHOU Feiyan, JIN Linpeng, DONG Jun. Review of convolutional neural network[J]. Chinese journal of computers, 2017, 40(6): 1229–1251.
- [17] 常亮, 邓小明, 周明全, 等. 图像理解中的卷积神经网络 [J]. 自动化学报, 2016, 42(9): 1300–1312.  
CHANG Liang, DENG Xiaoming, ZHOU Mingquan, et al. Convolutional neural networks in image understanding[J]. Acta automatica sinica, 2016, 42(9): 1300–1312.
- [18] 李彦冬, 郝宗波, 雷航. 卷积神经网络研究综述 [J]. 计算机应用, 2016, 36(9): 2508–2515, 2565.  
LI Yandong, HAO Zongbo, LEI Hang. Survey of convolutional neural network[J]. Journal of computer applications, 2016, 36(9): 2508–2515, 2565.
- [19] 周俊宇, 赵艳明. 卷积神经网络在图像分类和目标检测应用综述 [J]. 计算机工程与应用, 2017, 53(13): 34–41.  
ZHOU Junyu, ZHAO Yanming. Application of convolution neural network in image classification and object detection[J]. Computer engineering and applications, 2017, 53(13): 34–41.
- [20] GU Jiuxiang, WANG Zhenhua, KUEN J, et al. Recent advances in convolutional neural networks[J]. Pattern recognition, 2018, 77: 354–377.
- [21] LI Yangyan, SU Hao, QI C R, et al. Joint embeddings of shapes and images via CNN image purification[J]. ACM transactions on graphics (TOG), 2015, 34(6): 1–5.
- [22] 卢宏涛, 张秦川. 深度卷积神经网络在计算机视觉中的应用研究综述 [J]. 数据采集与处理, 2016, 31(1): 1–17.  
LU Hongtao, ZHANG Qinchuan. Applications of deep convolutional neural network in computer vision[J]. Journal of data acquisition and processing, 2016, 31(1): 1–17.
- [23] 李念峰, 董迎红, 肖志国. 基于图像处理的盲文自动识

别系统研究[J]. 制造业自动化, 2012, 34(3): 63–67.

LI Nianfeng, DONG Yinghong, XIAO Zhiguo. Study on image processing based braille automatic identification system[J]. Manufacturing automation, 2012, 34(3): 63–67.

#### 作者简介:



刘彪, 男, 1984 年生, 硕士研究生, 主要研究方向为智能软件与机器学习。



黄蓉蓉, 女, 1994 年生, 硕士研究生, 主要研究方向为人工智能与机器学习。



林和, 男, 1963 年生, 副教授, 主要研究方向为人工智能与机器学习。发表学术论文 100 余篇。

## 北京大学谭营教授主编的英文著作《群体智能：从概念到应用》由 IET 出版

日前, 由北京大学谭营教授主编的《群体智能: 从概念到应用》(英文名称: Swarm Intelligence: From Concepts to Applications) 由英国工程技术学会 (IET) 出版。该专著包括“算法原理、现有算法与方法”(Principles, current algorithms and methods)、“创新与新兴算法”(Innovation, new algorithms and methods) 和“算法应用”(Applications) 等 3 卷, 共 64 章, 汇集群体智能、进化计算、计算智能、人工智能等领域众多知名国际学者和专家的最新成果, 系统呈现了群体智能的算法原理、最新算法发展及其典型的实际应用, 充分反映了当前群体智能研究的现状和最新发展, 是群体智能研究发展进程中里程碑式的著作, 必将对领域产生积极影响。近年来, 谭营教授领导的团队在群体智能领域取得多项创新性的重要成果, 推动了相关领域研究的发展。

英国工程技术学会是工程技术领域的全球顶级专业学术学会, 总部设在伦敦, 其前身是创建于 1871 年的英国电气工程师学会 (IEE)。2006 年 3 月, IEE 与英国企业工程师学会 (IIE) 合并, 更为现名。目前, IET 在全球 150 个国家和地区拥有近 16.7 万会员, 是学会所涉及的能源电力、交通运输、信息与通信、设计与制造、建筑环境等行业和领域范围内, 欧洲规模最大、全球第二的国际专业技术学会。

相关链接: <https://www.theiet.org/resources/books/control/swarm-vol-1.cfm>

<https://www.theiet.org/resources/books/control/swarm-vol-2.cfm>

<https://www.theiet.org/resources/books/control/swarm-vol-3.cfm>