

基于胶囊网络的汉字笔迹鉴定算法

doi:10.3969/j.issn.1674-7100.2018.05.009

陈 健 周 平

浙江理工大学

信息学院

浙江 杭州 310018

摘 要: 由于采集脱机汉字手写样本时忽略了书写人的心理和生理等因素对书写活动的影响,因而传统笔迹鉴定算法的泛化能力较低。针对上述问题,提出基于胶囊网络的汉字笔迹鉴定算法,并构建了跟踪采集数据集以模拟复杂背景下产生的汉字。胶囊网络构建活动向量表示特定类型的实例化参数,通过动态路由算法将活动向量路由到下一层相应的胶囊中,使下一层胶囊得到更清晰的输入信号。分别采用5种算法对HWDB1.1数据集和跟踪采集数据集进行了测试,实验结果表明:本文算法的分类准确率比其他4种算法的都高,HWDB1.1数据集和跟踪采集数据集中算法的分类准确率分别为95.82%,94.39%;本文算法具有较强的泛化性能,对训练样本数的依赖程度较低,弥补了卷积神经网络池化层的信息丢失缺陷。

关键词: 胶囊网络;笔迹鉴定;活动向量;同变性

中图分类号: TP391.12

文献标志码: A

文章编号: 1674-7100(2018)05-0051-06

0 引言

书写活动在日常生活中必不可少。经过长期的书写,成年人的书写风格趋于稳定,是一种稳定的行为特征,因而可用于身份识别。笔迹鉴定对身份识别具有重要的研究意义,在司法、金融等领域发挥着越来越重要的作用。笔迹鉴定材料分为两种:一种是在线形式,如银行的电子签名;另一种是离线形式。本文研究的对象是离线笔迹。离线笔迹鉴定样本通常具有产生背景复杂、鉴定素材少等特点,这给鉴定工作增加了难度,对鉴定算法的泛化性能也提出了更高的要求^[1-3]。目前离线笔迹鉴定主要通过对汉字手写体预处理、特征提取和分类3个步骤实现^[4]。其中,特征提取分为两种。一种是通过固定的算法提取人为指定特征,再使用SVM(support vector machine)分类器等进行分类。如:刘俊^[5]提出基于笔画笔压

特征的手写汉字笔迹鉴定方法,并采用4种基本笔画、每种笔画10个样本验证算法,10位书写人的笔迹鉴定准确率可达94%。刘宏等^[6]提出基于SVM和纹理的笔迹鉴别方法,先使用Gabor滤波器提取纹理特征,再利用SVM分类器进行分类,15位书写人的笔迹鉴定准确率达93.10%。R. Jain等^[7]使用KAS算法(K-adjacent segments)提取手写体汉字的轮廓特征模拟手写汉字,并对IAM数据集(IAMonDo-database)中300位书写人的笔记进行鉴定,分类准确率达93.90%。另一种是通过深度学习中卷积神经网络(convolutional neural network, CNN)提取特征与分类。如:Xing L. J.等^[8]提出了一种多流的卷积神经网络方法(DeepWriter),并对HWDB(casia online and offline Chinese handwriting databases)数据集中301位书写人的笔迹进行鉴定,分类准确率达98.01%。H. T. Nguyen等^[9]提出了基于卷积神经网络

收稿日期: 2018-06-12

作者简介: 陈 健(1994-),男,安徽滁州人,浙江理工大学硕士生,主要研究方向为计算机视觉,

E-mail: 2218316563@qq.com

通信作者: 周 平(1960-),男,浙江杭州人,浙江理工大学教授,博士,主要从事计算机视觉方面的研究,

E-mail: 1500064071@qq.com

的子区域级和字符级局部特征提取方法,并对 IAM 数据集中 900 位书写人的笔迹进行鉴定,分类准确率为 91.81%。Yang W. X. 等^[10]提出使用 DropStroke 数据增强方法丰富手写内容,使用深度卷积神经网络对 420 位书写人的笔迹数据集进行测试,分类准确率为 92%。

卷积神经网络通过组合低层特征形成更加抽象的高层表示特征^[11],其表征特定的特征(如汉字手写体中横的倾斜角度、勾的提笔长度、竖弯的曲率等)具有较大优势,但无法表征特征间的同变性(如横的倾斜角度、勾的提笔长度等的相对位置关系)^[12]。而胶囊网络使用活动向量表征特定类型的实例化参数,可以更好地表征脱机手写汉字的特征以及特征之间的同变性,有利于脱机手写汉字的全局特征提取^[13-14]。因此,本课题提出基于胶囊网络的汉字笔迹鉴定方法,以期对汉字笔迹鉴定提供理论参考。

1 算法原理

1.1 胶囊网络

胶囊网络的原理是,输入图像经卷积层进行原始卷积运算得到基础胶囊层,再由动态路由算法将基础胶囊路由到下一层相应的胶囊构成笔迹胶囊层,笔迹胶囊经动态路由算法得到数字胶囊层,最后使用全连接层进行图像解码与重建。每个胶囊层由多个胶囊构成,每个胶囊都是一个活动向量即实例化参数,低层胶囊与高层胶囊相连,如图 1 所示。

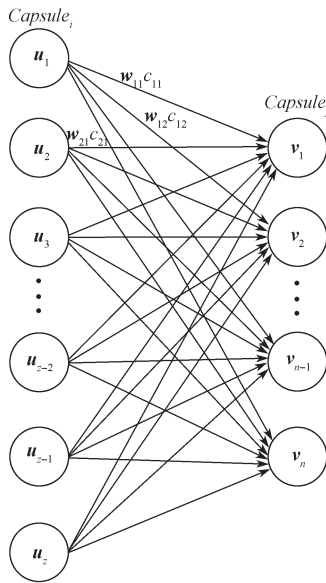


图 1 胶囊层示意图

Fig. 1 Schematic diagram of Capsule layer

动态路由是解决将低层胶囊的输出值传送给哪个高层胶囊的问题,实质上路由的数据传送是由低层胶囊的输入(预测向量)与高层胶囊的输出(激活向量)的相似度决定。如果低层胶囊的预测向量与高层胶囊的激活向量具有较高相似度,则判定两个胶囊是高度相关的。动态路由过程中,先将向量间相似度权重 b_{ij} 初始化为零并通过迭代进行更新,然后根据 b_{ij} 计算每个胶囊的动态路由耦合系数 c_{ij} ,最后通过 c_{ij} 得到 j 层胶囊的输出向量。

b_{ij} 的更新公式为

$$b_{ij} = b_{ij} + \hat{u}_{ji} \cdot v_j, \quad (1)$$

式中: \hat{u}_{ji} 为上一层胶囊的输出;

v_j 为激活函数。

b_{ij} 直接影响整个动态路由过程中的 c_{ij} 。当胶囊匹配到与自身高相似度的实体时, c_{ij} 会在动态路由迭代过程中不断地更新,即

$$c_{ij} = \frac{e^{b_{ij}}}{\sum_{k=1}^m e^{b_{ik}}}, \quad (2)$$

式中 m 为初始化相似度权重 b_{ij} 的个数。

胶囊的输出 s_j (除第一层胶囊) 是由低层胶囊输入 \hat{u}_{ji} 与 c_{ij} 得到:

$$\hat{u}_{ji} = W_{ij} u_i, \quad (3)$$

$$s_j = \sum_{i=1}^n c_{ij} \hat{u}_{ji}. \quad (4)$$

式中: u_i 是低层胶囊的输出;

W_{ij} 为胶囊所对应的权重。

为使胶囊最终的输出值在 $[0, 1]$ 之间,采用非线性函数对输出值进行压缩,即

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \times \frac{s_j}{\|s_j\|}. \quad (5)$$

最初的迭代过程中, c_{ij} 的状态是相等的,原因是低层胶囊不知道该和高层胶囊中的哪一个胶囊产生紧密的耦合状态;随着迭代过程的进行,这种均匀分布的状态很快被打破,相似度高的两个胶囊的耦合会变得紧密。动态路由算法的伪代码如下。

1: procedure DYNAMICROUTING (u_{ji}, r, l) // l 为胶囊层层数, r 为反向传播优化参数

- 2: for $capsule_i$ in layer l and result $capsule_j$ in layer $(l+1)$: $b_{ij} \leftarrow 0$
- 3: for $t=0$; $t < n$; $t++$ do
- 4: for $capsule_i$ in layer l : $c_{ij} \leftarrow e^{b_{ij}} / \sum_{k=1}^m e^{b_{ik}}$
- 5: for $capsule_j$ in layer $(l+1)$: $s_i \leftarrow \sum_{j=1}^n c_{ij} \hat{u}_{j|i}$
- 6: for $capsule_j$ in layer $(l+1)$: $v_i \leftarrow \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|}$
- 7: for $capsule_j$ in layer l and result $capsule_j$ in layer $(l+1)$: $b_{ij} \leftarrow b_{ij} + \hat{u}_{j|i} \cdot v_i$
- 8: return $V_j // V_j$ 为胶囊层输出
- 9: end procedure

胶囊网络的算法思想是由低层特征的迭代及其与高层特征的空间关系推出高层特征。汉字推理示意图如图2所示。

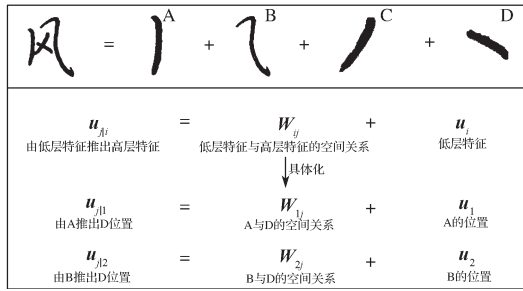


图2 胶囊网络的计算过程示意图

Fig. 2 Schematic diagram of the calculation process of the capsule network

1.2 学习模型

本文所设计的胶囊网络学习模型是将输入图像通过经典卷积得到基础胶囊, 基础胶囊经过动态路由算法逐层传递形成笔迹胶囊, 最后由全连接层重建图像。以“家”字为例的学习模型如图3所示。

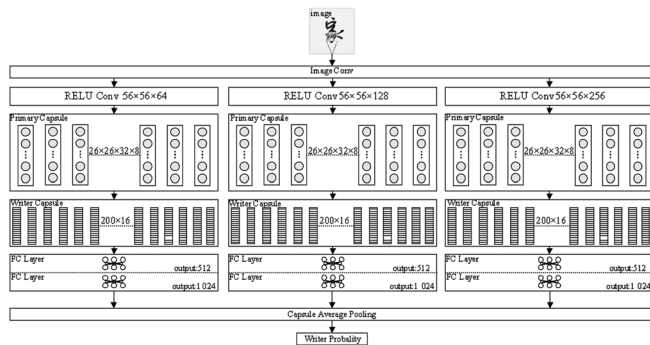


图3 学习模型设计图

Fig. 3 Learning model design

第一层是输入层, 即汉字图像通过3个经典卷积层进行卷积运算。每个经典卷积层由多个滤波器组成, 分别有64, 128, 256个滤波器。滤波器的核为 3×3 , 步长为1。卷积层的空间维度降到 56×56 。输入层的输出被传递给下一层构建胶囊。

第二层是代表最低级别多维实体的基础胶囊层。基础胶囊使用 5×5 卷积核, 步长为2, 8个输出神经元构成一个胶囊即得8维胶囊。基础胶囊层共有 $26 \times 26 \times 32$ 个胶囊。基础胶囊层的嵌套引入了新的维度, 不用通过添加层数增加网络的深度, 但在笔迹胶囊层中会添加多个新的图层^[15]。

第三层是笔迹胶囊层。笔迹胶囊由基础胶囊通过动态路由算法得到, 通过转换矩阵将8维胶囊转换为16维胶囊, 每个胶囊代表一个书写人。卷积神经网络是通过卷积产生的标量标定特征, 而笔迹胶囊层是使用带有维度的向量表示实体特征, 这对汉字笔画间存在的共性(字的整体斜率等), 以及汉字特征的相对关系有很大帮助。

第四层是全连接层。全连接层包含512, 1 024个节点的两个全连接层。该层使用网络的预测类别重构类别中的实体图像, 图像重建后, 根据三路网络的实例化参数的向量长度进行分类。

2 实验

2.1 数据集

本课题组使用HWDB1.1数据集和跟踪采集数据集分析本算法的性能。HWDB1.1数据集是由中国科学院自动化研究所于2007—2010年收集的300位书写人的脱机手写汉字样本组成, 这些样本经300 dpi扫描仪扫描后进行了切分和标注。跟踪采集数据集是目前开源的脱机手写汉字数据集, 采集时, 忽略了书写人的生理和心理状况对书写的影响, 限制200位书写人的书写时间, 每位书写人一周内每天书写50~100个汉字。这些样本经300 dpi扫描仪扫描获得彩色图像, 其后期处理过程如图4所示。文本预处理之后, 以文本二值图的外接矩形为基准, 将倾斜的文本行进行矫正, 矫正完成后再对文本进行水平方向的投影, 根据投影波形图的波峰、波谷的相对位置计算文本行的平均高度, 并统计黏连行的行数, 结合平均行高与黏连行的最小值进行分割, 得到文本行。文本分割与文本行分割的过程类似。经过筛选后, 最终得到104 800幅手写汉字图像。

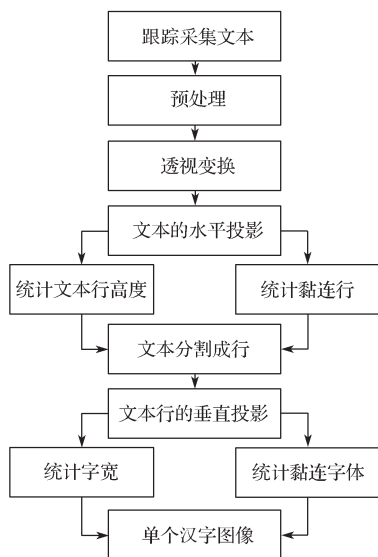


图 4 汉字采集流程图

Fig. 4 Chinese character collection process

2.2 结果与分析

本课题组将本文算法、DropStroke 算法^[10]、DeepWriter 算法^[8]、HalfDeepWriter 算法^[16]、改进的 GoogLeNet 算法^[17]对 HWDB1.1 数据集和跟踪采集数据集进行测试,分析各种算法的分类准确率和泛化性能。

HWDB1.1 数据集中,训练集和测试集按照 4:1 的比例划分,训练集包含 640 600 个样本,测试集包含 162 040 个样本。5 种算法对 HWDB1.1 数据集进行测试的结果如图 5 所示。

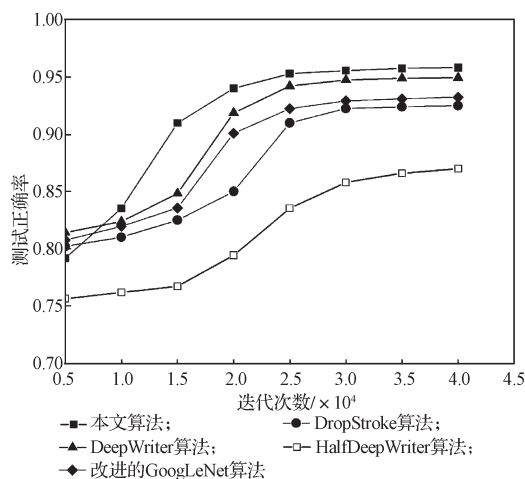


图 5 HWDB1.1 数据集的测试结果图

Fig. 5 Test result graph in HWDB1.1 data set

由图 5 可知:经过 4.0×10^4 次迭代后,5 种算法的分类准确率趋于稳定;迭代次数为 1.0×10^4 时,

本文算法的识别准确率开始快速增长,且最后分类准确率达到 95.82%,高于其他算法,这表明本文算法在计算量相对较少的情况下可以取得较高的分类准确率。

跟踪采集数据集中,训练集包含 84 200 个样本,测试集包含 20 600 个样本。利用跟踪采集数据集测试 5 种算法的泛化性能的实验结果如图 6 所示。由图可知:随着迭代次数的增加,本文算法的分类准确率逐渐提高,最后达到 94.39%,比其他算法的分类准确率都高;DropStroke 算法是使用减少汉字手写体笔画的方式以增强数据集,因而该算法的性能优于 DeepWriter 算法;对比 HWDB1.1 数据集的测试结果,本文算法的分类准确率下降了 1.43%,其他算法的分类准确率下降更明显,平均下降 8.52%。由此可知,在复杂书写背景下,DropStroke 算法、DeepWriter 算法、HalfDeepWriter 算法和改进的 GoogLeNet 算法的泛化性能较低,而本文算法在跟踪采集数据集中依然可以保持较高的分类准确率,其泛化性能较高。

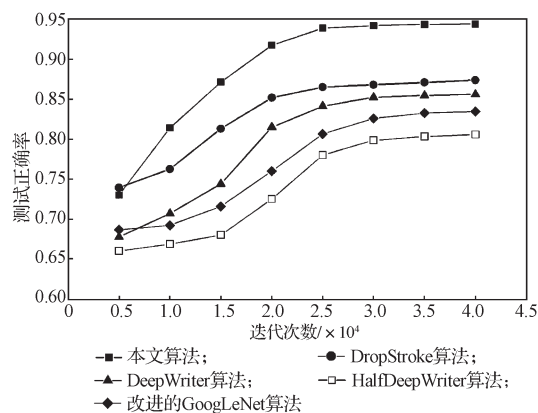


图 6 跟踪采集数据集的测试结果图

Fig. 6 Test result graph in tracking data set

本课题组分析了 5 种算法对训练样本数的敏感度。依据上述测试结果,此组实验设置迭代次数为 4.0×10^4 ,跟踪采集数据集中的训练样本数从 84 200 个依次递减。5 种算法的分类准确率如表 1 所示。

表 1 5 种算法对训练样本数的敏感度表
Table 1 Sensitivity table of five algorithms for training samples

算法名称	不同训练样本数的分类准确率				
	84 200	70 200	56 200	42 200	28 200
本文算法	94.39	93.85	92.23	91.24	89.56
DropStroke	87.08	84.31	79.03	77.62	76.39
DeepWriter	85.37	81.57	78.62	74.68	72.35
HalfDeepWriter	79.67	76.98	73.64	69.24	65.83
改进的 GoogLeNet	82.96	80.14	77.35	72.31	68.57

由表1可知,当训练样本数下降到42 200个时,DropStroke算法、DeepWriter算法、HalfDeepWriter算法和改进的GoogLeNet算法的分类准确率都有“瀑布式”的下降,而本文算法的分类准确率虽然也在下降但相对比较平稳,这表明本文算法可以使用相对较小的训练样本数取得较高的分类准确率。

3 结语

笔迹作为身份识别的一种手段有很广泛的应用前景。摹仿笔迹与被摹仿人的笔迹具有很高的相似度,这对笔迹鉴定时特征提取和分类算法提出了更高的要求,算法必须具有很强的泛化性能。因此,本课题组提出了基于胶囊网络的笔迹鉴定算法,采集样本时,考虑了书写人的心理和生理等因素,构建了跟踪采集数据集,分析了5种算法对HWDB1.1数据集和跟踪采集数据集的笔迹鉴定效果。实验结果表明:本文算法的分类准确率较高,泛化性能优于其他4种算法。

如何综合更多因素(如书写人是否自愿书写)采集样本是亟待解决的难题。因此,下一步工作中本课题组从两个方面对算法进行改善,一方面,增大分类类别数,增加样本的采样周期和复杂背景下的样本,以提高算法的泛化性能;另一方面,改进胶囊的路由算法,以降低胶囊网络对训练样本规模的依赖程度。

参考文献:

- [1] 陈如超. 当前笔迹鉴定样本存在的问题及其解决方案[J]. 中国人民公安大学学报(自然科学版), 2017(3): 34-40.
CHEN Ruchao. Problems Existing in Current Handwriting Identification Samples and Their Solutions[J]. Journal of People's Public Security University of China (Science and Technology), 2017(3): 34-40.
- [2] 董栋, 沙万中. 笔迹鉴定与笔迹心理分析比较[J]. 广西警察学院学报, 2018, 31(1): 44-49.
DONG Dong, SHA Wanzhong. Comparison of Handwriting Identification with Handwriting Psychological Analysis[J]. Journal of Guangxi Police College, 2018, 31(1): 44-49.
- [3] 高显东. 成功鉴定样本为少量正常字迹的笔迹鉴定案例分析[J]. 法制博览, 2017(27): 120.
GAO Xiandong. Successfully Identify the Sample as a Small Amount of Normal Handwriting Identification Case Analysis[J]. Legality Vision, 2017(27): 120.
- [4] AUBIN V, MORA M. A New Descriptor for Person Identity Verification Based on Handwritten Strokes Off-Line Analysis[J]. Expert Systems with Applications, 2017: 241-253. doi: 10.1016/j.eswa.2017.07.039.
- [5] 刘俊. 基于基本笔划笔压特征的手写汉字笔迹鉴定的研究[D]. 上海: 上海大学, 2003.
LIU Jun. Chinese Character Writer Identification Based on Pressure Features of Basic Strokes[D]. Shanghai: Shanghai University, 2003.
- [6] 刘宏, 李锦涛, 崔国勤, 等. 基于SVM和纹理的笔迹鉴别方法[J]. 计算机辅助设计与图形学学报, 2003, 15(12): 1479-1484.
LIU Hong, LI Jintao, CUI Guoqin, et al. Writer Identification Using Support Vector Machines and Texture Feature[J]. Journal of Computer-Aided Design and Computer Graphics, 2003, 15(12): 1479-1484.
- [7] JAIN R, DOERMANN D. Offline Writer Identification Using K-Adjacent Segments[C]//2011 International Conference on Document Analysis and Recognition. Beijing: IEEE, 2011: 769-773.
- [8] XING L J, QIAO Q. DeepWriter: A Multi-Stream Deep CNN for Text-Independent Writer Identification[C]//2016 International Conference on Frontiers in Handwriting Recognition. Shenzhen: IEEE, 2016: 584-589.
- [9] NGUYEN H T, NGUYEN C T, INO T, et al. Text-Independent Writer Identification Using Convolutional Neural Network[J]. Pattern Recognition Letters, 2018: 132-143.
- [10] YANG W X, JIN L W, LIU M F. DeepWriterID: An End-to-End Online Text-Independent Writer Identification System[J]. IEEE Intelligent Systems, 2016, 31(2): 45-53.
- [11] HAFEMANN L G, SABOURIN R, OLIVEIRA L S. Learning Features for Offline Handwritten Signature Verification Using Deep Convolutional Neural Networks[J]. Pattern Recognition, 2017, 70: 163-176.
- [12] HINTON G E, KRIZHEVSKY A, WANG S D. Transforming Auto-Encoders[C]//Artificial Neural Networks and Machine Learning. Berlin: Springer, 2011, 6791: 44-51.
- [13] SABOUR S, FROSST N, HINTON G E. Dynamic Routing Between Capsules[C]//2017 Conference on Neural Information Processing Systems. Long Beach: IEEE, 2017: 3859-3869.
- [14] XI E, BING S, JIN Y. Capsule Network Performance on Complex Data[C/OL]//2017 IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]:

- IEEE, 2017: 1–7. <https://arxiv.org/pdf/1712.03480.pdf>.
- [15] ZHANG S F, ZHAO W, WU X F, et al. Fast Dynamic Routing Based on Weighted Kernel Density Estimation[C/OL]// 2018 IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2018: 1–16. <https://arxiv.org/pdf/1805.10807.pdf>.
- [16] XING L J, QIAO Y. DeepWriter: A Multi-Stream Deep CNN for Text-Independent Writer Identification[C/OL]// International Conference on Frontiers in Handwriting Recognition. [S. l.]: IEEE, 2016: 584–589. <https://arxiv.org/pdf/1606.06472.pdf>.
- [17] 余 义. 基于内容及风格的离线手写汉字鉴定算法研究[D]. 成都: 电子科技大学, 2017.
- YU Yi. Research on Offline Handwritten Chinese Character Identification Algorithm Based on Content and Style[D]. Chengdu: University of Electronic Science and Technology of China, 2017.

Research on Chinese Character Handwriting Identification Based on Capsule Network

CHEN Jian, ZHOU Ping

(School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: Since the influence of the psychological and physiological factors of the writer on the writing activity was neglected while collecting the offline Chinese handwritten samples, the generalization ability of the traditional handwriting identification algorithm was low. A Chinese character handwriting identification algorithm based on capsule network was proposed, and a complex background of tracking the collected datasets to simulate Chinese characters was constructed. The capsule network constructed an activity vector to represent a specific type of instantiation parameter. The dynamic routing algorithm routed the activity vector to the corresponding capsule in the next layer to enable the next layer capsule to get a clearer input signal. The experimental results of five algorithms in HWDB dataset and tracking acquisition dataset showed that the classification accuracy of this algorithm was higher than that of the other four algorithms. The classification accuracy of HWDB dataset and tracking dataset algorithm were respectively 95.82% and 94.39%. The algorithm had strong generalization performance and low dependence on the number of training samples, making up for the convolutional neural network pooling layer information lost.

Keywords: CapsNets; handwriting identification; activity vector; location equivariance