

# 用于字符和数字识别的若干分类方法的比较研究:实验结果

陈 蔼 祥

(广东财经大学数学与统计学院 广州 510320)

**摘 要** 分类问题是机器学习领域中的一个重要问题。给出了数字 0-9 的图象和 26 个英文大小写的图象格式的训练数据,对 SVM、NB、RT、MLP、BOOST、Knearest 6 种分类器的分类性能进行了测试和评估。实验结果表明,性能表现排名前 3 的分类器为 SVM、NB、MLP。SVM 具有更好的泛化能力,而 NB 和 MLP 对训练集的变化更为敏感;并且基于 SVM 方法的分类系统对字符和数字的识别精度达到 94.2191%,高于现有文献的结果,系统识别性能具有更全和更准确的特点。

**关键词** 机器学习,分类器,数据集,性能比较,识别精度

中图法分类号 TP301 文献标识码 A

## Comparison of Several Classification Approaches to Digit and Letter Recognition: Experimental Results

CHEN Ai-xiang

(School of Mathematics and Statistic, Guangdong University of Finance and Economics, Guangzhou 510320, China)

**Abstract** Classification is an important problem in machine learning. This paper built several image datasets consisting of ten digits and 26 upper case and 26 lower case english letter to evaluate the performance of six classifiers, SVM(Support Vector Machine), NB(Naïve Bayes), RT(Random Tree), MLP(Multi Layer Perception), BOOST, Knearest. Experimental results show SVM has better generalization ability, while NB and MLP are more sensitive to datasets. In addition, the recognition accuracy of our system based on SVM reaches to 94.2191%, which is better than many publicly reported results in the literature.

**Keywords** Machine learning, Classifier, Dataset, Performance comparsion, Recognition accuracy

## 1 引言

现实中存在着大量根据数字和字符对不同物体或对象进行识别的应用场合。例如大型港口集装箱码头需要对成百上千甚至更多的集装箱进行标号,然后将这些集装箱按照特定的要求进行堆放,以满足货物的分发和转运的相关要求;超市中所售商品一般都会列出商品名称、价格、产地等基本信息;邮政编码则是邮政系统用数字对地区进行编码,以方便其按照来源地和目的地对信笺进行快速分类;车牌号码则是一个既使用字符又使用数字的车辆编码系统,交通管理部门很容易根据特定的车牌号码定位到车辆和车主本人。在智能规划领域,上述的集装箱堆放问题已经被抽象成积木块问题和后勤规划问题,并且在规划研究领域分别被用 pddl 语言形式化描述为 block, pddl 和 logistics, pddl<sup>1)</sup> 规划问题模型而被广泛研究<sup>[1]</sup>。然而,无论是用类似 pddl 这种形式化语言,还是使用二维码或各种电子标签的方式,均是一种“机器友好”的方式获取实体世界中的信息。让机器能像人眼一样能读懂这些字符和数字,以“用户(人)友好”的方式,而不是以二维码或各种电子标签或领域专家辅助的形式来获取外界信息,这是人工

智能的一个目标。

目前大量智能终端均提供了对手写输入的支持,这表明字符和数字的识别技术已经成熟。文献[2-6]对数字识别进行了研究,文献[7-11]对字符识别进行了研究,车牌识别的研究也有不少<sup>[12,13]</sup>。已有文献中更多的是对单纯数字或单纯字符识别方面的研究,能综合识别数字和字符研究甚少。究其原因,主要是不太容易同时获得手写数字和字符作为训练的数据,大多数机器学习标准测试数据也不提供同时包含数字和字符训练的数据,例如 UCI 数据集中提供了近 20453 条有关 26 个英文字母的手写字符记录,但不包含 0-9 这 10 个数字的训练数据。实际中更多的场合需要能同时识别字符和数字。此外,UCI 数据集中的人工字符集也并非原始手写字符,而是将手写字符进行特征提取后得到的记录,而 UCI 官网上并没有公布如何将图象形式的字符进行特征提取从而得到记录集合的具体方法,换言之,我们只能在 UCI 数据集的基础上进行算法的性能测试,而如果要让系统能识别图片数据,需要自己完成特征提取部分的工作。这在某种程度上阻碍了识别技术的工程化应用。实际生活中真正通过各种智能终端的摄像头获取图片来识别对象,进行信息加工、处理、分

<sup>1)</sup> 这两个领域模型及相应的问题实例已被看作是规划研究领域的基准测试例子。

本文受国家自然科学基金(60773201),广东省自然科学基金(10451032001006140),广州市科技和信息化局应用基础研究项目(10C12140131),广东省教育厅普通高校育苗工程(LYM10081)资助。

陈蔼祥(1978-),男,博士,副教授,主要研究方向为统计机器学习、智能规划,E-mail:cax413@163.com。

析,并带来管理和效率的提升,从而提高生产力和核心竞争力的应用,尚不多见。

字符和数字的识别本质上是一个分类问题。机器学习领域已经发展出大量的性能卓越的分类算法用以解决包含字符和数字识别在内的分类问题,其中比较有竞争力的算法主要有神经网络方法<sup>[14-17]</sup>、支持向量机方法 SVM<sup>[18-21]</sup>、朴素贝叶斯方法 NB<sup>[22]</sup>、随机森林方法 RT<sup>[23,24]</sup>、k 近邻方法<sup>[25]</sup>、基于 Boost 的方法<sup>[26]</sup>。支持向量机技术最初由 Vapnik 等<sup>[18-19]</sup>提出。SVM 采用了极大间隔原理、对偶理论和核函数技巧,并以统计学习理论和最优化方法为基础,使得其成为最有效的数据分析处理工具。

本文以字符和数字的识别问题作为基准例子,通过构造不同的训练样本集,用 SVM、NB、RT、MLP、Knearest、BOOST 这 6 个分类器对这些数据集进行训练和测试,以检验这些分类器的时间和精度性能。整个测试系统实现了包含训练数据的准备、图象的特征提取、分类和识别算法的实现以及实验结果的对比等方面,本文重点介绍训练数据的准备以及分类器的实验对比结果。本文所构造的能从视频和图象中识别数字和字符的原型系统,是进一步的工程化智能机器人的一个前期基础。

第 2 节介绍本文所使用的各种数据集以及相应的生成构造方法;第 3 节介绍本文实验系统的总体框架,包括训练数据的生成流程框架以及测试系统的总体框架;第 4 节给出实验对比结果,包括不同分类器对同一数据集的时间性能和精度性能比较,同一分类器下不同数据集的精度比较,以及相应的结果分析;最后对论文工作进行了总结并指出了下一步工作。

## 2 数据集

如前文所述,目前机器学习领域已有大量关于手写体字符或数字识别的研究,但同时能识别字符和数字的研究并不多见。本文以 26 个英文字母(区分大小写)、10 个 0-9 的数字共 62 个类别的识别问题作为案例,检验 SVM、NB、RT、MLP、BOOST、Knearest 6 个分类器对该问题的分类性能。

分类问题在机器学习领域属于有监督学习范畴,而有监督学习算法首先需要已被分好类的数据集作为训练样本。目前机器学习领域已有大量的用以测试分类器性能的所谓的基准测试数据。UCI 数据集<sup>[27]</sup>提供了共 6000 个人工字符实例。但这些数据并非是人人工手写字的原始图像,而是对原始图像进行特征提取后的数据,每条实例包括类别信息共有 17 个特征。UCI 数据集官网既没有提供手写字的原始图像数据,也没有提供如何将原始图像转换成训练数据的接口<sup>1)</sup>,这意味着只能用这些数据集测试算法的性能,而无法在其数据集的基础上进一步应用。MNIST 系统<sup>[28]</sup>则是纽约大学数学科学研究所 Yann LeCun 教授开发的人工手写数字识别系统,该系统公开了 58527 幅 0-9 这 10 个数字的人工手写数字图像(这些数字经过了一些诸如旋转、平移、缩放、扭曲等数学变换),并给出了数据集中各个字段较详细的介绍。同

样,MNIST 系统没有给出如何将原始图像转换成训练集的接口,用户个人的手写数字识别系统需要自行编写转换接口。

由于公开的基准测试数据中并没有 26 个英文字母(区分大小写)的图象形式的训练样本,因此为了得到 26 个英文字母(区分大小写)的图象作为训练样本,从网上搜集 26 个英文字母的图象(如图 1 所示),并为每个字母(区分大小写)建立一个以该字母名字命名的文件夹,将图 1 中各字母图象分割成的子图保存到对应的文件夹下面。这样得到 52 个文件夹(52 个类别),共 730 个样本,平均每类 14 个样本的样本集记为 data52-730c(对应图 7 中(1))。

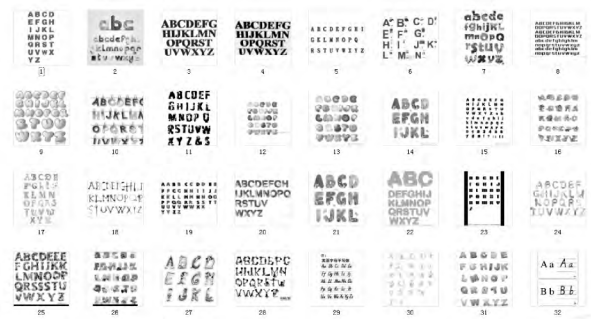


图 1 从网络上搜集的 26 个英文字母图象(部分)

为了使分类器能获取“高质量”的训练和测试样本,使分类器在不同光照、不同角度、不同距离等复杂的场景下获取的字符或数字图象均具有鲁棒的识别能力,从 data52-730c 每个类别中选取种子图象(如图 2 所示)(对应图 7 中(2)),对这些种子图像执行随机仿射、随机旋转和缩放、随机膨胀和腐蚀等一种或若干种混合变换。

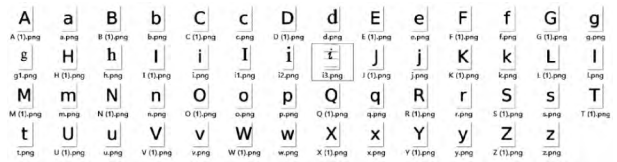
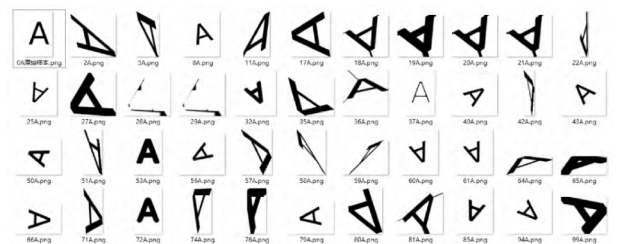


图 2 用于生成训练样本的种子图象

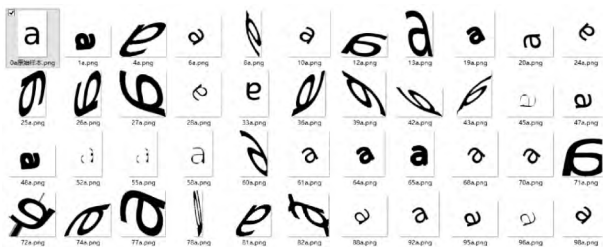
图 3 和图 4 分别给出了大写字母 A 和小写字母 a 的随机变换后的结果,其中第一个为原始样本图像。由图 3 和图 4 可以看出,经过随机变换后产生的图片产生了拉伸或扭曲的效果,但变换后这些图片仍然能被肉眼所识别。



第 1 幅图片为字母 A 的原始样本图像,其余图片均由第 1 幅图片经过仿射变换、旋转、缩放、腐蚀、膨胀这 5 种形态学变换(每一种变换的参数均随机设定)中的一种或任意几种随机混合变换后得到的结果,这些结果能被人眼所辨别

图 3

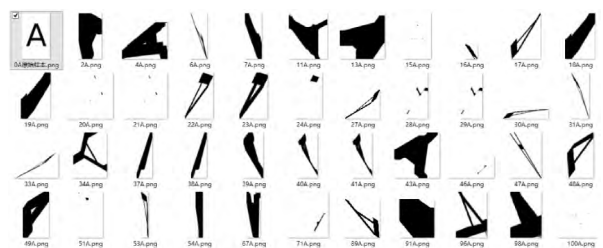
<sup>1)</sup> 本文作者曾向 UCI 数据集 Artificial Characters Data Set 提供者发邮件询问如何将图像转换成 Artificial Characters Data Set 中的训练实例,但未获任何回应。



第1幅图片为字母a的原始样本图像,其余图片均由第1幅图片经过仿射变换、旋转、缩放、腐蚀、膨胀这5种形态学变换(每一种变换的参数均随机设定)中的一种或任意几种随机混合变换后得到的结果,这些结果能为人眼所辨别

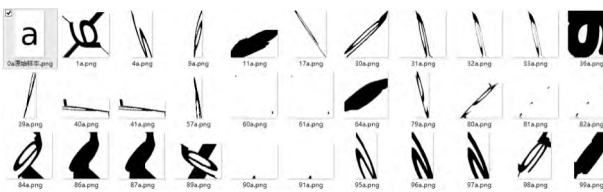
图4

由于图象变换的参数完全以随机方式选择,因此变换后得到的结果常常由于过度扭曲或缩放拉伸导致难以辨认。图5和图6分别给出了大写字母A和小写字母a被变换后难以辨别的情况。



第1幅图片为字母A的原始样本图像,其余图片均由第1幅图片经过仿射变换、旋转、缩放、腐蚀、膨胀这5种形态学变换(每一种变换的参数均随机设定)中的一种或任意几种随机混合变换后得到的结果,这些结果不易为人眼所辨别

图5



第1幅图片为字母a的原始样本图像,其余图片均由第1幅图片经过仿射变换、旋转、缩放、腐蚀、膨胀这5种形态学变换(每一种变换的参数均随机设定)中的一种随机变换或任意几种随机混合变换后得到的结果。这些结果已不易为人眼所辨别

图6

为了辅助用户能有效地产生所需要的训练样本,让分类器识别人眼所能识别的图象,本文自主设计并开发了一种交互式训练样本获取系统 CreateSamples:系统显示种子图象的每次随机变换结果,用户可根据变换的效果对当前图象选择保存、继续变换或放弃。

在 CreateSamples 系统的辅助下,得到了由图2中的种子图象随机变换后的训练集,52个类别共 43465<sup>1)</sup>个有效样本(即人眼能有效辨认),平均每个类别 835 个样本的数据集记为 data52-43465c(对应图7中(3))。

在前述两个数据集的基础上,将 data52-730c 和 data52-43465c 中的样本合并,形成了数据集 data52-43465-730c(对应图7中(4))。

此外,由于英文字母中“c”、“o”、“s”、“u”、“v”、“w”、“x”、

“z”这几个字符的大小写差别不大,预计这几个字符的大小写识别可能会对分类器的精度产生比较负面的影响。因此,我们将 data52-43465-730c 数据集中这几个字符的大小写样本合并成一类。例如将小写字母“o”文件夹“small o”下的样本拷贝到大写字母“O”文件夹“O”下,并删除文件夹“small o”。这样 52 个类别将缩减到 45 个类别,得到数据集 data45-43465-730c(对应图7中(5))。

对于 0-9 这 10 个数字的识别,使用 MNIST 系统公开的 58527 幅 0-9 这 10 个数字的人工手写数字图像,形成数据集 data10-58527d。

为了满足字符和数字的识别需要,从 data10-58527d 中的每个类别中人工抽取约 835 个样本,形成数据集 data10-8432d。然后,将 data10-8432d 和 data45-43465-730c 进行合并,得到数据集 data10-8432d-45-43465-730c(对应图7中(6))。

### 3 系统架构

本文测试系统总体架构可用图7和图8加以直观显示。图7给出了训练分类器所需数据的构造流程,其中(1)-(6)部分的数据集的构造在前一节中已有详细描述。

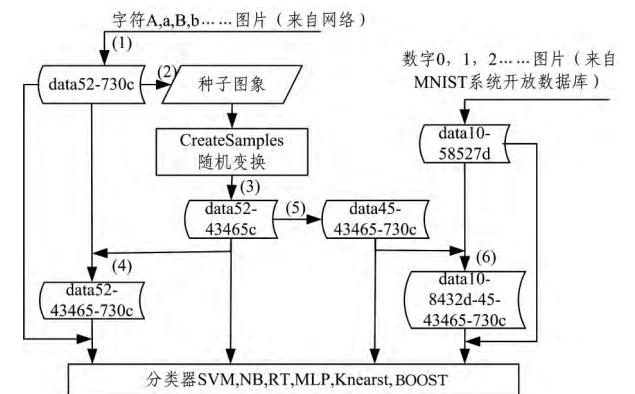


图7 训练样本生成流程

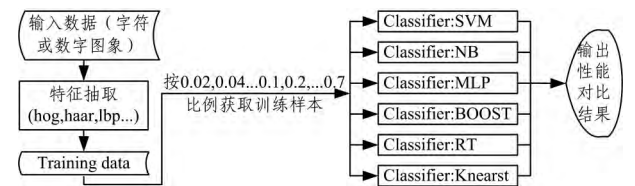


图8 测试系统总体框架

对于每个数据集,系统将逐加载图象数据,并进行特征提取,把一幅图象转换成特征向量,然后将得到的训练集按 0.02,0.04,0.06,0.08,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9 的比例将数据随机分成训练集和测试集两部分,随后将训练集作为分类器的输入对分类器进行训练。最后用测试集测试训练好的分类器,检验其分类能力。

### 4 实验结果及其分析

本文的测试系统在 Win8 + VS2013 + opencv2.4.9 环境下设计并实现。系统测试环境为 Intel(R)Core™ i3-2350 cpu @2.30GHz,6.0GB RAM。

<sup>1)</sup>之所以产生这一规模的样本作为训练数据,是因为更大的样本集会导致有些分类器不堪重负,无法在可忍耐的时间内返回结果。后文实验结果发现,对于 52627 个样本的 20% 约 1 万个训练样本,BOOST 分类器会因为内存不足而无法成功返回结果。

#### 4.1 不同分类器对同一数据集的结果

为了比较分类器的分类能力,实验中对比了SVM、RT、BOOST、MLP、Knearest、NB这6个分类器在同一个数据集上的分类情况,考察的指标主要是训练精度和测试精度这两个。可以观察到,在训练精度这个指标上,这6个分类器逼近能力的排名顺序是NB、SVM、MLP、Knearest、RT、BOOST,其中BOOST分类器在训练样本占比超过0.2时,由于内存不够(本文实验系统运行在6G内存环境下),无法正常返回结果而被排在最后。由于篇幅所限,并没有在此列出有关训练集上的数据,而只列出更关心的测试集上的详细数据。

表1—表4给出了6个分类器在数据集data10-8432d-45-43465-730c、data45-43465-730c、data52-43435c、data52-43465-730c上的测试精度。

表1 分类器在 data10-8432d-45-43465-730c 上的渐近精度(测试集)

数据集		data10-8432d-45-43465-730c					
ratio	SVM (%)	RT (%)	BOOST (%)	MLP (%)	Knearest (%)	NB (%)	
0.02	63.227	63.8524	54.4086	67.4998	56.8762	72.497	
0.04	75.5482	68.5599	65.7698	75.7081	65.3585	78.1902	
0.06	80.8252	69.1138	68.8574	79.7438	70.3546	82.3004	
0.08	83.2397	73.1559	70.8828	81.9401	72.5647	83.9095	
0.1	85.488	75.0048	70.5509	83.3631	75.0872	85.4813	
0.2	90.9095	76.3178	*	87.9993	80.8265	88.9763	
0.3	92.8964	77.5795	*	89.6247	84.2868	78.6563	
0.4	93.8614	78.9158	*	91.2117	87.1935	81.2851	
0.5	94.3596	79.2255	*	92.0063	88.5815	87.9424	
0.6	94.9586	78.8588	*	92.371	90.3095	89.4102	
0.7	94.0832	78.764	*	91.4194	87.7005	85.5105	
0.8	94.4974	79.6894	*	91.4537	88.4837	87.4113	
0.9	94.2191	80.2607	*	91.9034	89.0947	88.5899	

表2 分类器在 data45-43465-730c 上的渐近精度(测试集)

数据集		data45-43465-730c					
ratio	SVM (%)	RT (%)	BOOST (%)	MLP (%)	Knearest (%)	NB (%)	
0.02	49.7985	51.7738	32.0505	49.977	34.955	58.4367	
0.04	60.5541	59.3568	38.0589	63.7132	48.0889	69.1206	
0.06	68.465	63.4667	58.5901	69.8192	57.0176	71.8834	
0.08	73.8092	63.1015	63.4594	72.7823	57.3182	76.2395	
0.1	76.3351	67.6322	67.2534	75.6834	62.3559	77.7338	
0.2	84.261	71.9376	74.4023	82.2653	70.4237	84.3503	
0.3	86.3353	73.8973	*	85.4093	74.6617	86.4742	
0.4	90.189	76.7243	*	86.9012	77.9287	89.0452	
0.5	91.327	74.4589	*	89.0346	78.6864	89.8815	
0.6	91.4115	75.9744	*	90.084	81.1112	91.5807	
0.7	92.3307	75.8832	*	90.3908	82.0865	91.8545	
0.8	92.5456	77.866	*	90.6452	83.7133	92.4513	
0.9	93.2766	75.9622	*	90.6301	85.0659	92.6294	

表3 分类器在数据集 data52-43435c 上的渐近精度(测试集)

数据集		data52-43435c					
ratio	SVM (%)	RT (%)	BOOST (%)	MLP (%)	Knearest (%)	NB (%)	
0.02	45.1117	50.6502	23.2576	52.2266	26.5556	56.3423	
0.04	61.366	55.5185	45.8461	62.5238	44.336	67.771	
0.06	67.4705	61.7978	59.0975	67.2734	53.2436	71.5128	
0.08	72.6667	64.5742	64.0386	71.1598	57.3999	74.2848	
0.1	74.7465	65.4603	64.9774	73.5855	60.248	78.5978	
0.2	83.2566	69.2323	74.7456	79.9549	67.7494	84.7099	
0.3	87.4164	70.1171	*	84.186	72.7134	88.7365	
0.4	89.5713	74.9589	*	86.3372	75.7267	90.3928	
0.5	90.5793	75.4878	*	87.6814	77.6802	91.8568	
0.6	91.1066	74.9868	*	88.9902	79.9819	92.1452	
0.7	91.6014	77.8561	*	89.123	81.5552	93.199	
0.8	92.3304	74.4573	*	89.6427	82.7138	93.0202	
0.9	92.4549	78.6806	*	89.9233	84.5983	92.7675	

表4 分类器在数据集 data52-43465-730c 上的渐近精度(测试集)

数据集		data52-43465-730c					
ratio	SVM (%)	RT (%)	BOOST (%)	MLP (%)	Knearest (%)	NB (%)	
0.02	50.4515	48.5727	28.422	52.4586	30.6198	56.4817	
0.04	62.5052	58.0844	36.9669	63.1239	47.4158	66.5263	
0.06	66.6289	63.0342	60.3487	67.3897	53.734	70.9531	
0.08	72.9681	64.4026	62.2112	71.1869	58.0156	76.0765	
0.1	74.5446	66.1008	66.713	73.7456	60.3932	77.5259	
0.2	83.9333	71.2966	75.2976	80.916	67.6862	84.8784	
0.3	86.9304	72.5938	*	83.018	72.7797	87.9202	
0.4	88.6624	72.6154	*	85.3545	75.4118	89.5946	
0.5	90.007	74.4669	*	86.7837	77.7079	90.7747	
0.6	90.0503	75.3028	*	87.7528	79.6099	91.2636	
0.7	91.0126	76.1425	*	88.4332	81.7276	91.7662	
0.8	91.4553	76.7122	*	88.2929	82.5985	92.0682	
0.9	91.6147	76.4717	*	89.3449	84.3929	92.6826	

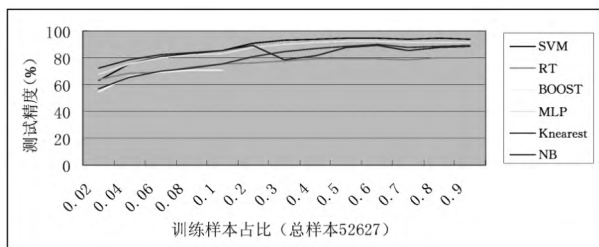


图9 分类器在 data10-8432d-45-43465-730c 的测试精度对比

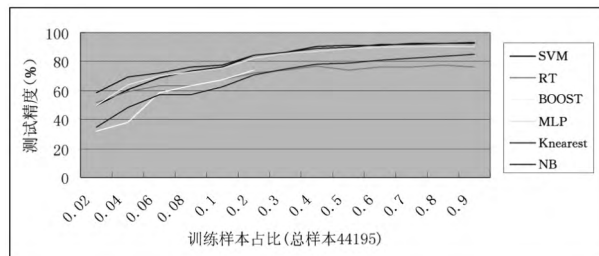


图10 分类器在 data45-43465-730c 的测试精度对比

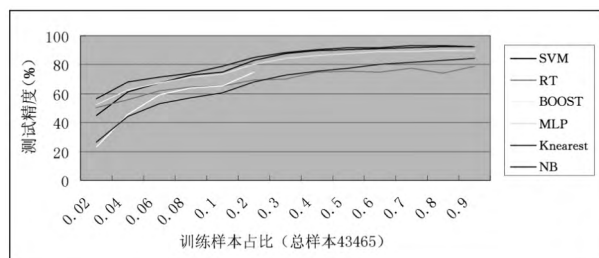


图11 分类器在 data52-43435c 的测试精度对比

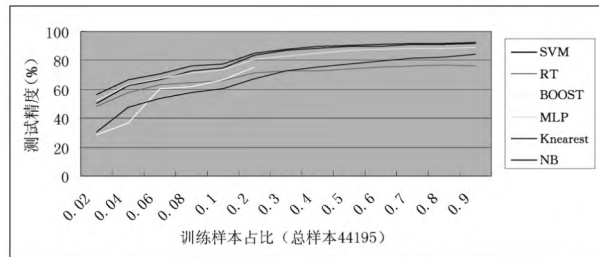


图12 分类器在 data52-43465-730c 的测试精度对比

从表1—表4和相应的图9—图12可以看出,在测试精度这个指标上,这6个分类器泛化能力排名顺序是SVM、NB、MLP、Knearest、RT、BOOST。同样,BOOST被排在最后是因为在训练样本占比超过0.2时,由于内存不够(本文实验系统运行在6G内存环境下)而无法返回结果。

不过这里有必要强调一点的就是,对于 data10-8432d-45-43465-730c、data45-43465-730c 这两个数据集,是将“c”、“o”、“s”、“u”、“v”、“w”、“x”、“z”这几个字符中大小写差别不大的字母的大小写合并作为一类得到的数据,对这类数据,SVM 泛化能力强于 NB(表 2、表 4)。但对于 data52-43435c、data52-43465-730c,其是没做类似合并的数据,此时 NB 的泛化能力则要比 SVM 强。这表明 NB 比 SVM 对数据的变化更敏感(具有更精细的识别能力)。

文献[29]利用卷积神经网络方法进行字符和数字的识别,其中使用的数据集共 44022 条记录,但其对大写字母和数字的识别率只有 89%,对小写字母的识别率也只有 89%。而本文的系统能够识别字符(大部分区分大小写)和数字,识别率达到 94.2191%。

文献[30]给出了一种基于隐马尔科夫模型的增量式学习算法用于字符和数字的识别,其中使用了 195000 个 0-9 这 10 个数字的样本,26 个英文大写字母,每类 1660 个,总共 43160 个样本作为训练数据,其对数字的最好识别率为 97.91%,对字符的识别率为 93.24%,但该文献并没有给出字符和数字混合时的识别率。

#### 4.2 同一分类器对不同数据集的结果

限于篇幅,只比较排名前 3 的分类器 SVM,NB,MLP 在不同数据集下的测试精度,结果如表 5-表 7、图 13-图 15 所示。

从表 5 可以看出,SVM 对数据集 data10-8432d-45-43465-730c(表中第一列)的测试精度总体高于另外 3 个数据集,这表明对 SVM 分类器而言,数字的加入带来了识别精度的提高。

而从表 6、表 7 则可以发现,对 MLP 和 NB 分类器而言,对数据集 data10-8432d-45-43465-730c(表中第一列)的测试精度要低于其他列,数字的加入对 MLP 和 NB 分类器的识别精度带来负面影响,这表明 SVM 的泛化能力强于 MLP 和 NB。

再看 data45-43465-730c 和 data52-43465-730c 所在列的数据,类别的合并对 SVM 和 MLP 分类器的分类性能带来了正面影响,识别率均有所提高;而对 NB 分类器的影响则不那么明显。

最后,对于数据集 data52-43435c 和 data52-43465-730c,可以看到,网络图片的加入对 SVM、MLP、NB 的渐进识别精度带来的影响都是负面的。

表 5 SVM 在不同数据集上的渐近精度(测试集)

分类器	SVM			
	data10-8432d-45-43465-730c	data45-43465-730c	data52-43435c	data52-43465-730c
0.02	63.227	49.7985	45.1117	50.4515
0.04	75.5482	60.5541	61.366	62.5052
0.06	80.8252	68.465	67.4705	66.6289
0.08	83.2397	73.8092	72.6667	72.9681
0.1	85.488	76.3351	74.7465	74.5446
0.2	90.9095	84.261	83.2566	83.9333
0.3	92.8964	86.3353	87.4164	86.9304
0.4	93.8614	90.189	89.5713	88.6624
0.5	94.3596	91.327	90.5793	90.007
0.6	94.9586	91.4115	91.1066	90.0503
0.7	94.0832	92.3307	91.6014	91.0126
0.8	94.4974	92.5456	92.3304	91.4553
0.9	94.2191	93.2766	92.4549	91.6147

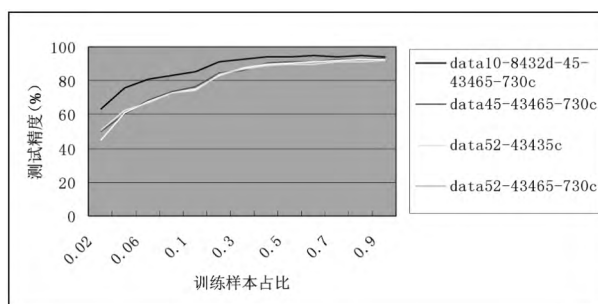


图 13 SVM 在不同数据集上的测试精度对比

表 6 MLP 在不同数据集上的渐近精度(测试集)

分类器	MLP			
	data10-8432d-45-43465-730c	data45-43465-730c	data52-43435c	data52-43465-730c
0.02	98.6616	98.6945	97.4227	98.829
0.04	98.1827	98.8593	98.0653	99.2509
0.06	97.3214	99.245	98.5678	98.9457
0.08	96.7719	98.7828	98.1587	98.5154
0.1	96.0773	98.6639	97.7057	98.257
0.2	95.1215	97.5032	96.9552	96.7547
0.3	95.1145	97.1017	96.3143	95.759
0.4	94.7097	96.4341	95.8354	95.5412
0.5	94.6045	96.3673	94.942	95.1167
0.6	94.8627	96.4823	95.0615	94.6786
0.7	94.7877	96.2004	94.7102	94.1767
0.8	94.5498	95.9863	95.0041	94.122
0.9	94.4404	95.4797	94.7178	94.3421

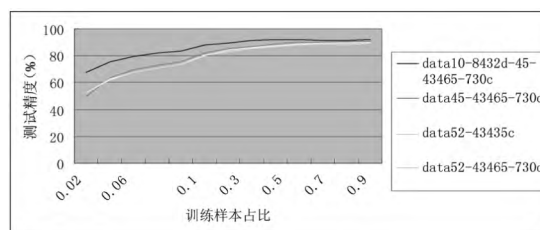


图 14 MLP 在不同数据集上的测试精度对比

表 7 NB 在不同数据集上的渐近精度(测试集)

分类器	NB			
	data10-8432d-45-43465-730c	data45-43465-730c	data52-43435c	data52-43465-730c
0.02	72.497	58.4367	56.3423	56.4817
0.04	78.1902	69.1206	67.771	66.5263
0.06	82.3004	71.8834	71.5128	70.9531
0.08	83.9095	76.2395	74.2848	76.0765
0.1	85.4813	77.7338	78.5978	77.5259
0.2	88.9763	84.3503	84.7099	84.8784
0.3	78.6563	86.4742	88.7365	87.9202
0.4	81.2851	89.0452	90.3928	89.5946
0.5	87.9424	89.8815	91.8568	90.7747
0.6	89.4102	91.5807	92.1452	91.2636
0.7	85.5105	91.8545	93.199	91.7662
0.8	87.4113	92.4513	93.0202	92.0682
0.9	88.5899	92.6294	92.7675	92.6826

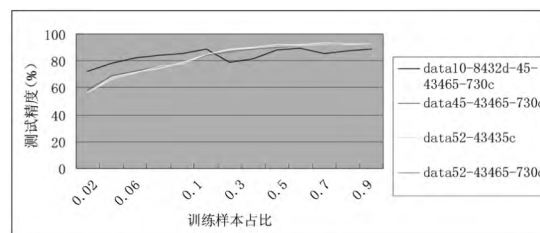


图 15 NB 在不同数据集上的测试精度对比

(下转第 121 页)

tion, 2011, 8: 114-132

- [5] Virginia E-D. Biometric identification system using a radial basis network[C]//Proc. 34<sup>th</sup> Annual IEEE Int. Carnahan Conf. on Security Technology, 2011: 47-51
- [6] Li S Z, Lu J. Face recognition using the nearest feature line method[J]. IEEE Trans. Neural Networks, 2012, 10: 439-443
- [7] Brennan V, Principe J. Face classification using a multiresolution principal component analysis[C]//Proc. IEEE Workshop Neural Network for Signal Processing, 2011: 506-515
- [8] Chen S, Cowan C F N, Grant P M. Orthogonal least squares algorithm for radial basis function network[J]. IEEE Trans. Neural Networks, 2011, 2: 302-310
- [9] Wu S Q, Er M J. Dynamic Fuzzy Neural Networks: A Novel Approach to Function Approximation[J]. IEEE Trans. Syst, Man, Cybern. Part B, 2012, 30: 358-364

- [10] Esposito A, Marinaro M, Oricchio D, et al. Approximation of continuous and discontinuous mappings by a growing neural RBF-based algorithm[J]. Neural Networks, 2013, 25: 651-665
- [11] Bors A G, Pitas I. Median radial basis function neural network[J]. IEEE Trans. Neural Networks, 2012, 23: 1351-1364
- [12] Myoody J, Darken C J. Fast Learning in network of locally-tuned processing units[J]. Neural Computation, 2011, 1: 281-294
- [13] Girosi F, Poggio T. Networks and the best approximation property[J]. Biological Cybernetics, 2012, 63: 169-176
- [14] 杨文光. 权值直接确定的三角形模糊前向神经网络[J]. 中山大学学报: 自然科学版, 2013, 52(2): 33-37
- [15] 任爱红. 模糊随机过程函数列均方差一致 Henstock 积分的可积性[J]. 中山大学学报: 自然科学版, 2012, 51(4): 41-44
- [16] Haykin S. Neural networks, a comprehensive foundation [M]. New York: Macmillan, 2012: 358-366

(上接第 106 页)

**结束语** 本文给出了一个字符和数字的图片集合,并基于此数据集设计了一套测试和评估各种分类算法的系统,对 SVM、NB、MLP、Knearest、RT、BOOST 分类器进行比较和性能评估。结果发现排名前三的分类器是 SVM、NB、MLP,其中 SVM 泛化能力最好,而 NB 以及 MLP 对数据的变化更敏感。BOOST 分类器则由于对系统内存要求高没能完成样本占比高于 0.2 后的测试,而被排在了最后。本文系统对区分大小写的 26 个英文字母以及 10 个数字的混合样本的识别率高于已有的文献报导。

本文系统只使用了一种特征提取算法——hog 特征,我们将在本文工作基础上进一步测试各种特征提取算法对分类器分类能力的影响,为将来的科研和工程应用提供坚实的基础。

### 参 考 文 献

- [1] Fox M, Long D. PDDL2. 1: An Extension to PDDL for Expressing Temporal Planning Domains[J]. Journal of Artificial Intelligence Research(JAIR), 2003, 20: 61-124
- [2] 双小川,张克. 基于统计和结构特征的手写数字识别研究[J]. 计算机工程与设计, 2012(4): 1533-1537
- [3] 陈军胜. 组合结构特征的自由手写体数字识别算法研究[J]. 计算机工程与应用, 2013(5): 179-184, 194
- [4] 杜文龙,毛学军. 手写识别技术在上网阅卷系统中的应用[J]. 电脑学习, 2010(1): 14-16
- [5] 赵健,张冬泉. 基于 OpenCV 的数字手势识别算法[J]. 计算机应用, 2013(S2): 193-196
- [6] 陶胜. 邮政编码手写体的快速识别[J]. 电脑编程技巧与维护, 2010(23): 55-58
- [7] 姜映映,敖翔,田丰,等. 基于语音和笔的手写数学公式纠错方法[J]. 计算机研究与发展, 2009(4): 689-697
- [8] 马婉婕,孙虎元,孙立娟,等. 基于神经网络集成的手写识别系统[J]. 计算机应用与软件, 2009(8): 5-7, 44
- [9] 郭星,吴建国,张义超,等. 基于汉字骨架手写识别算法分析[J]. 计算机技术与发展, 2009(7): 114-116
- [10] 陈胤子. 手写识别算法研究及在移动平台上的应用[D]. 北京: 北京邮电大学, 2011
- [11] 敖翔,王绪刚,戴国忠,等. 基于多通道融合连续手写识别纠错方法[J]. 软件学报, 2007(9): 2162-2173
- [12] 李峰. 汽车牌照自动识别系统的研究[D]. 吉林: 吉林大学, 2014
- [13] 薛丹. 车牌识别系统的算法研究[D]. 西安: 西安电子科技大学,

2013

- [14] 王西颖,戴国忠,张习文,等. 基于 HMM-FNN 模型的复杂动态手势识别[J]. 软件学报, 2008(9): 2302-2312
- [15] 商琳,王金根,姚望舒,等. 一种基于多进化神经网络的分类方法[J]. 软件学报, 2005(9): 1577-1583
- [16] 陈宇,郑德权,赵铁军. 基于 Deep Belief Nets 的中文名实体关系抽取[J]. 软件学报, 2012(10): 2572-2585
- [17] 王宪保,周德龙,王守觉. 基于仿生模式识别的构造型神经网络分类方法[J]. 计算机学报, 2007(12): 2109-2114
- [18] Cortes C, Vapnik V. Support-Vector Networks[J]. Machine Learning, 1995, 20(3): 273-297
- [19] Vapnik V, Golowich S E, Smola A J. Support Vector Method for Function Approximation[C]//Regression Estimation and Signal Processing, NIPS, 1996: 281-287
- [20] 谢娟英,谢维信. 基于特征子集区分度与支持向量机的特征选择算法[J]. 计算机学报, 2014(8): 1704-1718
- [21] 陈耀东,李仁发,李实英,等. 面向目标检测与姿态估计的联合文法模型[J]. 计算机学报, 2014(10): 2206-2217
- [22] 王双成,杜瑞杰,刘颖. 连续属性完全贝叶斯分类器的学习与优化[J]. 计算机学报, 2012(10): 2129-2138
- [23] 王爱平,万国伟,程志全,等. 支持在线学习的增量式极端随机森林分类器[J]. 软件学报, 2011(9): 2059-2074
- [24] 刘进华. 基于随机森林的语音情感识别研究[D]. 广州: 华南理工大学, 2013
- [25] 张著英,黄玉龙,王翰虎. 一个高效的 KNN 分类算法[J]. 计算机科学, 2008(3): 170-172
- [26] 查宇飞,楚瀛,王勋,等. 一种基于 Boosting 判别模型的运动阴影检测方法[J]. 计算机学报, 2007(8): 1295-1301
- [27] Bache K, Lichman M. UCI Machine Learning Repository [M]. Irvine, CA: University of California, School of Information and Computer Science, 2013
- [28] LeCun Y, Institute C, Yu N, et al. Microsoft Research, Redmond. THE MNIST DATABASE of handwritten digits[OL]. <http://yann.lecun.com/exdb/mnist/>
- [29] Pham D V. Online Handwriting Recognition Using Multi Convolution Neural Networks. Simulated Evolution and Learning [J]. Lecture Notes in Computer Science, 2012, 7673: 310-319
- [30] Pham V D. Online handwriting recognition using multi convolution neural networks[OL]. [http://www.codeproject.com/Articles/523074/Online-handwriting-recognition-using-multi-convolu-](http://www.codeproject.com/Articles/523074/Online-handwriting-recognition-using-multi-convolu-2013) 2013