

手写数字深度特征学习与识别

陈浩翔,蔡建明,刘铿然,林秋爽,张文玲,周 涛
(华南师范大学 数学科学学院,广东 广州 510631)

摘要:深度学习中的网络结构设计、特征提取与融合是数据挖掘和模式识别理论和行业应用中的关键问题。文中以相关领域中的典型应用问题手写数字识别和权威数据库 MNIST 为实验平台(包含七万个手写数字图像),探索了深度学习网络结构的设计和特征融合问题,保证研究结果的实用性、代表性和可参考性。所给方案的步骤是:首先,设计非监督深度学习网络,进行非监督高层语义特征学习,提取深度特征(DF),探索特征的高层认知特点;其次,对手写数字数据库进行非监督多特征提取,包括 HOG(梯度方向直方图)特征、PCA(主成分分析)特征、LDA(判别分析)特征、像素分布特征、穿越次数特征和投影特征,构建手写数字典型特征库(Library of Typical Features, LTF);最后,构建深度有监督学习网络,有监督地融合深度特征 DF 和典型特征库。实验结果表明,相比于文献中的典型特征,该方案能够将手写数字识别的错误率有效降低 50%。

关键词:深度学习;特征融合;特征提取;手写数字识别;主成分分析;梯度方向直方图

中图分类号:O235

文献标识码:A

文章编号:1673-629X(2016)07-0019-05

doi:10.3969/j.issn.1673-629X.2016.07.005

Deep Learning and Recognition of Handwritten Numeral Features

CHEN Hao-xiang, CAI Jian-ming, LIU Keng-ran, LIN Qiu-shuang, ZHANG Wen-ling, ZHOU Tao
(School of Mathematics, South China Normal University, Guangzhou 510631, China)

Abstract: Network structure design, feature extraction and fusion in deep learning are key problems in data mining and pattern recognition theory and industry application. The design of deep learning network's structure and the problem of feature fusion is explored, taking handwritten numeral recognition and authoritative database MNIST, with 70 thousands of handwritten image, as the experiment platform, which guarantees the practicability, representation and reference of the research results. The solution step has been given. Firstly, the unsupervised deep learning network is designed, learning unsupervised high-level semantic features, extraction of depth features, and exploration of higher cognitive characteristics of features. Secondly, unsupervised features of handwritten database are extracted, including HOG, PCA, LDA and so on, construction of LTF. Finally, deep supervised learning network is built, fusion of deep features and the library of typical features with supervision. The result shows that this scheme can lower error rate of handwritten recognition by 50%, compared with the typical features of the present.

Key words: deep learning; feature fusion; feature extraction; handwritten numeral recognition; principal component analysis; histogram of oriented gradient

0 引言

人们能够辨识人脸、识别语音、阅读文字。这些背后都有着非常复杂的处理机制的识别行为。模式识别,这种输入原始数据并根据其类别采取相应行为的能力跟人们的生活息息相关。

在一般的模式识别处理过程中,如果要得到一个良好的分类决策结果,则其中有两个部分最为重要:特

征提取和分类决策。而近年来学术界的研究重点也放在了这两个方面。特征提取是在预处理的基础上,提取出识别目标的具体属性,易于进行分类操作;完成特征提取后,通过这些特征,分类器将感知到的目标进行有效的分类决策;最后即根据得到的分类结果,做出相应的操作反应。

深度学习通过建立一种深层非线性网络结构,实

收稿日期:2015-10-25

修回日期:2016-01-27

网络出版时间:2016-06-22

基金项目:国家自然科学基金资助项目(61273248, 61075033)

作者简介:陈浩翔(1994-),男,研究方向为人工智能与模式识别;导师:李乡儒,教授,研究方向为模式识别与数据挖掘;冯伟贞,副教授,研究方向为常微分方程应用研究。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160622.0844.028.html>

现复杂函数逼近,表征输入数据分布式表示,并体现了它对于输入样本数据的强大的本质特征的抽取能力。然而,深度结构(涉及多个非线性处理单元层)非凸目标代价函数中普遍存在的局部最小是训练困难的主要来源。Hinton 等基于深信度网(Deep Belief Networks)提出非监督贪心逐层训练算法^[1],为深度结构相关问题的优化带来了希望,随后提出多层自动编码器深层结构。

在模式识别的案例中,手写数字识别是模式识别中的一个重要的研究课题,在当今信息化的社会有非常广阔的应用。但目前对各类字体的数字识别,特别是脱机手写数字识别仍然处在发展阶段,识别效果仍然不够理想。因此,手写数字识别的研究具有重大的现实意义。

文中将以特征选择这一重要分类决策步骤作为切入点,结合深度学习的模型训练,将不同的特征融合在一起应用到深度学习的模型训练中,从而优化原始的训练算法。通过在权威数据库 MNIST^[2-4]上进行手写数字实验,分析特征融合在深层架构模型的影响与作用,并说明特征选择在深度学习模型应用过程中的有效性。

1 算法流程与架构

文中采取不同的特征提取方法,主要为主成分分析(PCA)和梯度直方图(HOG),并将其融合,输入至一个七层的深度栈式自编码网络中进行二次提取,再由一个 Softmax 分类器^[5]进行训练分类。最后利用控制变量和贪心的思想调整网络结构,以期显著降低分类的错误率。其大致流程见图 1。

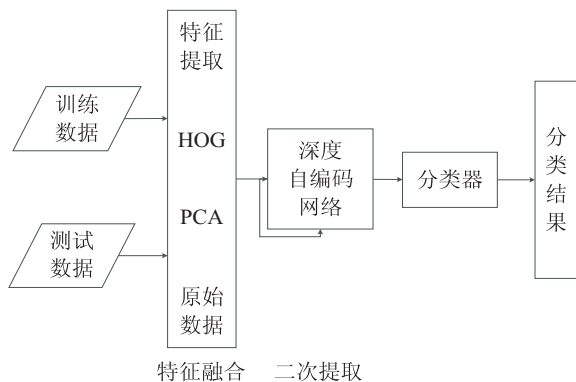


图 1 算法流程

2 特征提取方法

文中采取的特征提取算法主要有 PCA 和 HOG。

2.1 主成分分析

文中采用主成分分析法(Principal Component Analysis)降低原数据的维数以简化运算。此法在去除

噪音和冗余的同时,也保持数据的累计方差贡献率超过 90%。

PCA 处理在数学定义层面而言,是对数据做一个正交化线性变换。在本质上,采用 PCA 的目的就是在最小均方误差(MSE)的意义下,寻找最能够代表原数据的投影^[6-7]。

PCA 降维的实现过程具体如下:

(1)为了让原数据 A 的均方误差最小,向量的均值应为零,即数据应以原点为中心。因此将原数据 A 以平均值为中心移到原点得到数据 X :

$$X = A - \bar{A} \quad (1)$$

(2)协方差用于衡量两个变量的总体误差,而方差是两个变量相同的特殊情况下的协方差。计算 X 的协方差无偏估计矩阵 Y :

$$Y = \frac{1}{n-1} X X^T \quad (2)$$

由于两个变量相同,此 Y 也是方差无偏估计矩阵。

(3)计算矩阵 Y 的特征值 D 和单位正交特征向量 V ,再让特征值 D 由大到小排序。

(4)对应特征值的排序将特征向量排序后,即可得到变换矩阵 v ,令 X 乘以 v 即可得到已排序的投影数据 Z 。其中,第一个对应的就是第一主成分,第二个对应的就是次主成分,以此类推。

(5) λ_i 是协方差矩阵,也是方差矩阵的第 i 个主元的特征值。计算第 i 个主元的方差贡献率:

$$\lambda_i / \sum_{i=1}^n \lambda_i \quad (3)$$

而 R 个主元的累计方差贡献率为:

$$\sum_{i=1}^R \lambda_i / \sum_{i=1}^n \lambda_i \quad (4)$$

(6)求得累计方差贡献率超过 99% 的主成分数目 R ,然后在已排序的投影数据 Z 中取前 R 维,即可得最终的降维同时各元之间互不相关的矩阵 B 。文中 R 最后的取值为 388 维。

2.2 方向梯度直方图

方向梯度直方图(Histogram of Oriented Gradient)是法国国家计算机科学及自动控制研究所的 Dalal 等提出的一种解决人体目标检测的图像描述方法,用于提取人体的外形信息和运动信息,从而形成丰富的特征集^[8-9]。

HOG 具有以下优点:

(1)HOG 表示图像边缘(梯度)的结构特征,可以描述局部的形状信息;

(2)使用分块分单元的处理方法,让图像局部像素点之间的关系可以得到很好的表征;

(3)位置和方向空间的量化一定程度上可以抑制

平移和旋转带来的影响^[10]。

数字图像比人体图像的结构更加简单,使用 HOG 特征提取能把数字图像的主要特征囊括在特征集中。由于文中所使用的 MNIST 数据集中每个图像都存在比较大的空白区域,因此,进行 HOG 特征提取前将数据集中每个样本边缘外围的 5 行 5 列去掉,减少无用信息。

HOG 特征提取算法的实现过程具体如下:

(1) 将图像灰度化。

(2) 采用 Gamma 校正法对输入图像进行颜色空间的归一化,调节图像的对比度,降低图像局部的阴影和光照变化所造成的影响,同时抑制噪音的干扰。

由于 MNIST 中图像已经进行了灰度化以及去噪,因此不需要进行以上两步。

(3) 将图像划分为单元格 (cell), 每 2×2 个像素点组成一个单元格。

(4) 计算图像每个像素的梯度 (包括大小和方向), 捕获轮廓信息:

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y) \quad (5)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1) \quad (6)$$

其中, $G_x(x, y)$, $G_y(x, y)$, $H(x, y)$ 分别代表像素点 (x, y) 处的水平方向梯度、垂直方向梯度和像素值。像素点 (x, y) 处的梯度幅值和梯度方向分别为:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (7)$$

$$\alpha(x, y) = \tan^{-1} \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (8)$$

梯度方向分为有符号 ($0 \sim 360$) 和无符号 ($0 \sim 180$) 两种形式。文中实验采用的是无符号形式, 并将 ($0 \sim 180$) 分成 9 等份。

所以, 图像中每个单元格的梯度方向都分成 9 个方向块, 使用单元格中的梯度方向和大小对 9 个方向进行加权投影, 最后, 每个单元格产生 9 维的特征向量。

(5) 将每 2×2 个单元格组成一个块, 一个块内所有单元格的特征描述符串联起来后归一化, 便得到该块的 HOG 特征。通常使用的 HOG 结构大致有三种: 矩形 HOG、圆形 HOG 和中心环绕 HOG。它们的单位都是块 (block)。Dalal 等试验证明, 矩形的 HOG 和圆形的 HOG 检测效果基本一致, 而环绕形 HOG 效果相对差一些。因此文中使用矩形 HOG 结构进行实验。

(6) 将图像所有块的 HOG 特征串联起来就可得到图像的 HOG 特征, 共 2 304 维。

(7) 将特征进行归一化, 对特征提取后的特征向量进行特征变换。

$$\mathbf{v} \leftarrow \mathbf{v} / 255 \quad (9)$$

$$L = \text{norm}, \mathbf{v} \leftarrow \mathbf{v} / \sqrt{\|\mathbf{v}\|_2 + \varepsilon^2} \quad (10)$$

3 稀疏自编码网络

随着深度学习理念的提出, 在机器学习领域涌现了大量的相关论著。而稀疏自编码网络^[11]是其中非常重要而且实用的模型, 它可以压缩数据减少冗余与噪声、降低输入空间的维度、自主学习输入数据的特征, 而且运用贪心的思想, 稀疏自编码网络可以拥有任意的层数, 使得数据能够更进一步地压缩和提取特征。它的整个结构可以分为自编码网络、稀疏性抑制和栈式自编码。

3.1 自编码网络

自编码网络是由原始的 BP 神经网络^[12]演化而来。在原始的 BP 神经网络中, 从特征空间输入到神经网络中, 并用类别标签与输出空间来衡量误差, 用最优化理论不断求得极小值, 从而得到一个与类别标签相近的输出。但是在编码网络并不是如此, 并不用类别标签来衡量与输出空间的误差, 而是通过特征空间的输入来衡量与输出空间的误差。

把特征空间的向量 (x_1, x_2, x_3, x_4) 作为输入, 利用经过神经网络训练后的向量 (x'_1, x'_2, x'_3, x'_4) 与输入向量 (x_1, x_2, x_3, x_4) 来衡量误差。代价函数如下所示:

$$\text{Cost} = \frac{1}{2} \|\mathbf{x} - \mathbf{h}_\theta(\mathbf{x})\| \quad (11)$$

利用最优化算法不断迭代减小代价函数值, 从而最终得到一个能从原始数据中自主学习特征的特征提取的神经网络。从代数角度而言, 即从一个线性相关的向量中, 寻找出一组低维的基, 而这组基线性组合之后又能还原成原始数据, 自编码网络正是寻找了一组这样的基。

3.2 稀疏性抑制

自编码网络是简单而实用的, 但是如果隐层神经元个数很多时, 就要对一些激活度不高的神经元进行抑制, 从而达到用更低维的向量表达高维向量的要求。从机器学习的角度出发, 即是对模型复杂度加以限制。高复杂度的模型会带来训练准确率的下降, 同样也会影响到自编码网络自主学习特征的过程。加上稀疏性抑制后的自编码网络称之为稀疏自编码^[13]。

3.3 栈式自编码

神经网络很早就被提出, 但是因为局部极值、梯度弥散、数据获取等问题而构建不出深层的神经网络。直到 2006 年深度学习的提出, 才让神经网络的相关算法有了质的改变。而栈式自编码就属于深度学习理论中一种能够得到优秀深层神经网络的方法。

栈式自编码神经网络是一个由多层稀疏自编码器组成的网络。它的思想是利用逐层贪婪训练的方法, 把原来多层的神经网络剖分成一个个小的自编码网络, 每次只训练一个自编码器, 然后将前一层自编码的

输出作为其后一层自编码器的输入,最后连接一个分类器,可以是 SVM、SoftMax 等等^[14]。上述步骤是为了得到一个好的初始化深度神经网络的权重,当连接好一个分类器后,还可以用 BP 神经网络的思想,反向传播微调神经元的权重,以期得到一个分类准确率更好的栈式自编码神经网络。

4 特征融合

对于识别分类问题,数据的特征表达对于分类的效果起着至关重要的作用。文中将基于 MNIST 来探讨样本数据的原始特征(即 MNIST 特征)、PCA 特征和 HOG 特征的不同特征融合技术对于分类模型的精度影响。

4.1 特征拼接

首先,用这三种特征的训练样本分别对分类器进行训练,用测试样本测试已经训练好的分类器的分类效果,以便直观地了解这些特征本身的分类效果。结果如图 2 所示。

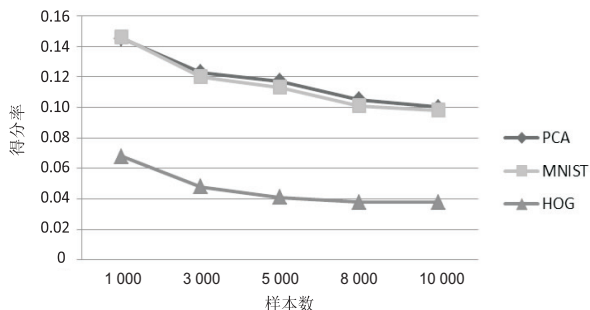


图 2 PCA 特征、样本数据的原始特征 (即 MNIST 特征)、HOG 特征的比较

可以看到,原始特征和 PCA 特征表达对于分类效果差别不大,而 HOG 的特征表达取得的分类效果远远优于原始特征和 PCA 特征。

接着,研究这两种特征融合后对于分类效果的影响,具体做法是:将特征向量 A (长度为 n) 拼接到相应的特征向量 B (长度为 m) 的尾部,组成长度为 $n + m$ 的新向量,即融合特征向量,而后将这些融合向量组成的训练样本送入分类器中训练分类器,再通过测试融合向量得到分类效果。

通过实验结果发现,PCA 特征和原始特征的融合特征的错分率处于两者之间,即分类效果比 PCA 特征高而比原始特征低,这一定程度上符合人们的主观常识:高低融合取中间;然而,PCA 和 HOG 的融合特征的分类效果却优于融合前的任何一个单独特征。这结果表明:不同的特征融合会取得不同的效果,也就是特征融合后的特征并不一定优于原特征,也不一定劣于原特征。

PCA 的全称是主成分分析,它是从数据中提取出主成分(即具有较大方差的那些维度),这些主成分用最少的维度表达了最多的信息(一般 $\geq 99\%$),在降维的同时也起到了降噪的作用,这是 PCA 的本质所在。因此可以看到,PCA 提取的特征和原始特征在本质上是没有任何差别的。

而 HOG 是方向梯度直方图,它首先将图像分成小的连通区域,把它叫细胞单元,然后采集细胞单元中各像素点的梯度或边缘的方向直方图,最后把这些直方图组合起来构成特征描述器。HOG 在拥有几何不变性和形变不变形优点的同时也存在导致维数增长的缺点。HOG 的本质在于统计梯度信息,换句话说,是统计笔画信息(因为梯度主要存在于有边缘的地方)。所以 HOG 提取的特征和原始的像素特征是完全不同的。

4.2 二次提取

上文探讨了一种特征融合,接下来探讨另外一种特征融合,即二次提取特征。

不论是 PCA 还是 HOG,它们都是人工选取特征的方法,在数据量和样本维度比较高时,人工选取的特征在一定程度上仍然存在冗余和噪声等。为此,一个自然的想法就是引入自动提取特征的方法—深度学习,使得模型能够在训练时自己从特征样本中再次学习到更简洁、更高效的特征样本。在深度学习的众多方法中,选取较为成熟的栈式稀疏自编码(Stack Auto-Encoder, SAE)作为二次特征提取的模型。

下面将在控制变量的前提下,分别用 HOG 和 PCA 作为第一次特征提取的方法,将提取到的特征通过 SAE 训练,自动提取二次特征,而后将测试用的特征送入 SoftMax 分类器测试分类效果。结果如图 3 所示。

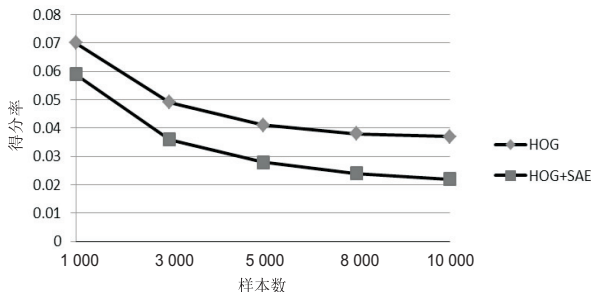


图 3 HOG 特征与 HOG+SAE 特征比较

由图 3 可以明显看到,HOG 特征在经过 SAE 二次提取特征后,所表现分类效果都有明显的提高。同样的,PCA 特征在经过 SAE 二次提取特征后,分类效果也得到明显的提高。而导致这种结果的原因,是和 SAE 的内在机理密不可分的。SAE 是由神经网络和稀疏自编码衍变而来的,其至关重要的特性就是编码和稀疏。

SAE 将输入特征(如 HOG 特征、PCA 特征)逐层编码成某个低维空间的特征,但却保留了输入特征几乎所有的信息,达到了降噪去冗的效果,但这是利用 PCA 做手工提取特征也能做到的事情,关键之处在于 SAE 的稀疏抑制。SAE 通过稀疏抑制使得尽可能多的神经元处于抑制状态,而这要求系数矩阵 W 中的元素尽可能为 0,从而抑制了基的维数。而可以通过调整稀疏性参数 ρ 来适当抑制基的维数增长,从而寻找到最佳的模型复杂度,避免过拟合,提高模型的精度和泛化能力。

4.3 特征拼接+二次提取

最后,把两种特征融合方法结合在一起,验证特征融合的优势。如图 4 所示,将比较 PCA、HOG、PCA+HOG、PCA+HOG+SAE 的分类效果。根据上面的实验,知道 PCA+HOG 融合特征的分类效果优于 PCA 特征和 HOG 特征,但是 PCA+HOG 融合特征的分类效果和 PCA+HOG+SAE 二次融合特征的分类效果是未知的,这次实验的意义在于拼接特征和二次提取的兼容性是否可以再次提高特征的表达。图 5 为实验结果示意图。

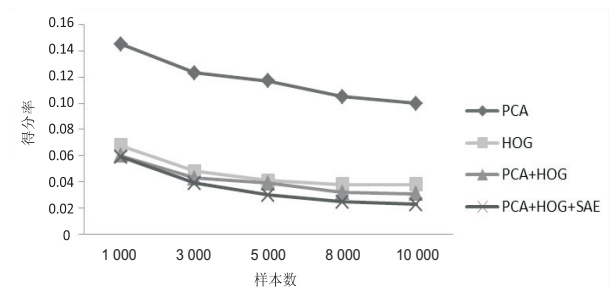


图 4 PCA 特征,HOG 特征,PCA+HOG 特征与 PCA+HOG+SAE 特征的比较

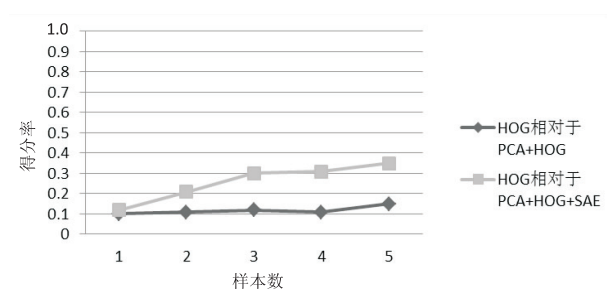


图 5 特征的相对比较

从图中可以清楚地看到,PCA 和 HOG 在拼接特征后分类效果有较为明显的提升,而后再在此基础上再做 SAE 二次特征提取后,分类结果又得到了较为明显的提升。

这说明,特征拼接和二次提取是可以兼容的,这也意味着当实际需要更高的准确率而不追求快速分类时,可以采取特征融合技术来获取更好的分类效果,提高精度。

5 结果对比

在得到一个有效的特征融合算法后,文中着重设计了深度网络的结构,利用控制变量和贪心的思想,在一定的取值区间内调整参数,设计结构,寻求最优解,最终得到了一个局部最优的网络结构和参数。再利用特征融合算法,结合调优过的深度栈式自编码网络得到一个极高的分类准确率。该准确率优于绝大部分的分类模型,其具体参数和结果见表 1。

表 1 具体参数和结果

参数	参数值
输入维度	3 088
网络层数	7
隐藏层神经元分布 *	[3 088,1 865,1 127,681,411,248,150]
最大迭代次数	200
权重惩罚系数	1e-3
稀疏惩罚系数	0.5
稀疏度	0.1
最终分类准确率/%	99.2

表中,* 为隐层神经元个数,由下式计算:

$$\text{Table}[\text{Round}[3\,088 \times (\frac{150}{3\,088})^{\frac{n}{6}}], \{n, 0, 6\}]$$

(12)

通过该模型对 MNIST 数据库进行分类,准确率达到 99.2%,错误率只有 0.80%,相比 Ranzato 和 LeCun 所得到的错误率 0.89% 和 0.85% 都有提高。

6 结束语

文中提出了一种高精度有效的有监督深度特征融合分类系统。论证了对原始数据不同表征的特征凸组合才是有效的特征融合方法,从而将 HOG、PCA 和 MNIST 三种特征有效结合在一起,并用深度自编码网络进行二次提取再进行分类。分类结果远优于仅用原始特征作为输入的稀疏自编码网络,使之错误率下降了将近 50%,证明特征融合的方法能够显著提高分类的精度,可以广泛应用于各类深度学习模型中。最后优化深度网络结构,得到了一个准确率为 99.2% 的深度特征融合分类系统,其优于绝大多数的分类模型,具有较强的实用性和可参考性。

参考文献:

[1] Hinton G E,Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science,2006,313(5786): 504-507.

[2] 边肇祺. 模式识别[M]. 北京:清华大学出版社,1988.

3.2.3 分布平均性测试

当集群的节点组数为 2 个或 3 个,进行写操作 1 000 万次时,进行数据库写操作验证,测试结果如图 5 所示。

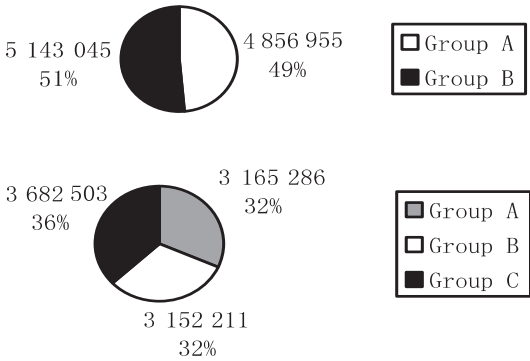


图 5 数据均匀性分布

当采用 2 个节点组时,Group 内写入的数据量分别占 51%、49%,采用 3 个节点组时,Group 内写入的数据量分别占 36%、32%、32%。节点组之间数据分布相对比较均匀,避免了数据倾斜问题的发生。

4 结束语

文中对分布式存储领域的技术展开了研究,简单介绍了 Redis 数据库及分布式系统的 CAP 理论,详细说明了普通哈希算法的不足之处。提出基于 Group 方式对物理上的存储节点进行逻辑划分,改进的一致性哈希算法提高了分布式存储系统的可靠性、可用性,并以 Redis 数据为例进行实验验证,通过增加存储节点组大大提高了集群的系统负载。

参考文献:

[1] 陆嘉恒. 分布式系统及云计算概论[M]. 北京:清华大学出版社,2013.

[2] Devine R. Design and implementation of DDH: a distributed dynamic hashing algorithm[M]//Foundations of data organization and algorithms. Berlin:Springer,1993:101-114.

[3] Kernighan B W, Ritchie D M. The C programming language[M]. [s. l.]:Prentice Hall,1988:36-40.

[4] Rivest R L. The MD4 message digest algorithm[C]//Proc of CRYPTO'90. [s. l.]:[s. n.],1991:303-311.

[5] Rivest R. The MD5 message-digest algorithm[S]. [s. l.]:IETF,1992.

[6] 张绍兰. 几类密码 Hash 函数的设计和安全性分析[D]. 北京:北京邮电大学,2011.

[7] Winternitz R S. A secure one-way hash function built from DES[C]//Proc of IEEE symposium on security and privacy. [s. l.]:IEEE Computer Society,1984:88-88.

[8] 余秦勇,陈 林,童 斌. 一种无中心的云存储架构分析[J]. 通信技术,2012,45(8):123-126.

[9] 李 正. 杂凑函数结构研究现状及新的结构设计[D]. 济南:山东大学,2010.

[10] 姜大光,奚加鹏. 分布式存储系统(OceanStore)的复制策略[J]. 计算机工程与科学,2008,30(8):144-146.

[11] 杨戬剑,林 波. 分布式存储系统中一致性哈希算法的研究[J]. 电脑知识与技术,2011,7(22):5295-5296.

[12] 赵 飞,苏 忠. 一致性哈希算法在数据库集群上的拓展应用[J]. 成都信息工程学院学报,2015,30(1):52-58.

[13] 郭 宁,张 新. 一致性哈希算法在多处进程分配的应用[J]. 计算机与现代化,2013(9):71-74.

[14] 周 瑜. 一种基于一致性 hash 算法存储资源的方法:CN, CN 103281358 A[P]. 2013.

(上接第 23 页)

[3] 边肇祺,张学工. 模式识别[M]. 第 2 版. 北京:清华大学出版社,2000.

[4] 张学工. 模式识别[M]. 第 3 版. 北京:清华大学出版社,2010.

[5] 苏高利,邓芳萍. 论基于 MATLAB 语言的 BP 神经网络的改进算法[J]. 科技通报,2003,19(2):130-135.

[6] 韩小孩,张耀辉,孙福军,等. 基于主成分分析的指标权重确定方法[J]. 四川兵工学报,2012,33(10):124-126.

[7] 齐兴敏. 基于 PCA 的人脸识别技术的研究[D]. 武汉:武汉理工大学,2007.

[8] 叶 林,陈岳林,林景亮. 基于 HOG 的行人快速检测[J]. 计算机工程,2010,36(22):206-207.

[9] 姚雪琴,李晓华,周激流. 基于边缘对称性和 HOG 的行人检测方法[J]. 计算机工程,2012,38(5):179-182.

[10] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//Proceedings of IEEE computer society confer-

ence on computer vision and pattern recognition. [s. l.]:IEEE Press,2005:886-893.

[11] Raina R, Battle A, Lee Honglak, et al. self-taught learning: transfer learning from unlabeled data[C]//Proc of international conference on machine learning. [s. l.]:[s. n.], 2007:759-766.

[12] Le Q V, Ngiam J, Coates A, et al. On optimization methods for deep learning[C]//Proc of international conference on machine learning. [s. l.]:[s. n.], 2011:265-272.

[13] Ranzato M, Huang F J, Boureau Y L, et al. Unsupervised learning of invariant feature hierarchies with applications to object recognition[C]//Proc of IEEE conference on computer vision and pattern recognition. [s. l.]:IEEE,2007:1-8.

[14] 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.