　附表2-1（通用）

南京邮电大学

毕业设计(论文)外文资料翻译

|  |  |
| --- | --- |
| 学 院 | 计算机学院、软件学院、网络空间安全学院 |
| 专　　业 | 计算机科学与技术 |
| 学生姓名 | 王吉凡 |
| 班级学号 | B18030510 |
| 外文出处 | Neurocomputing  Volume 136, 20 July 2014, Pages 152-161 |

附件：1.外文资料翻译译文；2.外文原文

|  |
| --- |
| 指导教师评价：  1．翻译内容与课题的结合度： □ 优 □ 良 □ 中 □ 差  2．翻译内容的准确、流畅： □ 优 □ 良 □ 中 □ 差  3．专业词汇翻译的准确性： □ 优 □ 良 □☑ 中 □ 差  4．翻译字符数是否符合规定要求：□ 符合 □ 不符合  　　　　　　　　　　　　　　　　　　　　　　　指导教师签名：  　　　　　　　　　　　　　　　　　　　　　　　　　　 2022年5月13日 |

**三种整合空间结构到有序图像回归的策略比较研究**

田青 陈松灿 谭晓阳

南京航空航天大学计算机科学与技术学院

摘要——图像通常具有特定的空间结构，相关研究表明，这些结构有助于建立更有效的图像分类算法。到目前为止，虽然已经提出了许多利用这种空间结构的解决方案，但很少有人对其进行系统总结，更不用说它们的比较研究了。另一方面，我们发现现有的面向图像的有序回归（OR）方法没有利用这种结构信息，这促使我们通过将这种空间结构嵌入到OR中来弥补比较研究的不足。最后，本文总结了三种典型的利用图像先验空间信息的策略，即嵌入结构的欧氏距离策略、用于分类器学习的结构正则化建模策略和不对图像进行矢量化的直接操作策略；更重要的是，（2）利用这些策略建立相应的ORs，对具有顺序特征的数据进行分类，并在三个评价标准下进行综合比较和分析。典型的序列图像数据集上实验结果贾菲，曼彻斯特理工和FG-NET表明，后者两种策略，总体上实现不同的增益或性能，虽然第一人不一定如预期的那样，将空间信息是否涉及直接嵌入到目标函数。

关键词：序数回归；矢量模式；矩阵模式；空间结构；欧氏距离；双线性的

1．介绍

1.1．背景

图像具有二维的固有空间结构，其中包含了显式和隐式的有利于图像分类的判别信息。例如，在人脸中，眼睛、鼻子和嘴分布在不同的区域，它们之间存在特定的几何关系。然而，目前发展起来的模式识别和机器学习算法大多是基于向量模式，进行矩阵到向量的转换过程，从而严重丢失了用于分类的有用的空间结构信息，留下了性能提升的空间。

近年来，虽然针对提高图像分类性能的方法分别制定了利用空间结构信息的策略，但仍缺乏系统的总结和比较研究。为此，本文首先对分散的相关文献进行总结，并将其分为三类；然后，我们选择当前图像分类中比较流行的一个主题，即面向图像的OR，作为比较的平台。OR是一种特殊的机器学习范式，具有分类和回归的对偶性，常用于预测标签离散但有序的场景[26，31]，如人脸年龄估计、电影评分等。除了二元性之外，选择面向图像的OR作为比较研究范式的进一步原因是（1）这些用于有序图像分类的特殊设计的OR迄今为止几乎没有利用此类空间信息，（2）基于分类和回归二元性的多指标综合评价，比单一的分类或回归指标评价更能体现在信息利用的多层面。然后，利用上述三种策略对空间信息进行补偿，开发了三种面向图像的OR变量，并在MAE、Acc和OCI三种评价标准下从回归和分类的角度进行了广泛的比较。

1.2．空间结构信息利用策略的分类

在本小节中，我们分析了现有的利用空间结构信息的分散方案，并将其归纳为以下三大类:

1.2.1．结构嵌入的欧几里德距离策略

欧几里德距离（ED）是模式识别中最常用的度量标准之一。然而，当它被用来度量两幅图像之间的相似度/距离，其中所涉及的空间结构信息没有得到充分的反映，影响了图像的分类性能。为了弥补这一损失，人们做了很多尝试[1-8]，其中参考文献[1]可以作为代表。作者开发了一个图像欧氏距离（IMED）通过嵌入空间结构的ED和应用它的手写数字图像和人脸识别与更好的性能比埃德。由于它的钝性小失真的图像和通用性能够嵌入到支持向量机等分类器，IMED可以连续扩展。如Liu等人[4]将IMED扩展到多视角性别分类，取得了更高的分类精度；Liu等人[5]进一步提出了多线性局部保持最大信息嵌入人脸识别，具有更稳定的性能。此外，Li和Lu[8]开发了一种自适应IMED （AIMED），在先验空间信息的基础上，进一步将图像的灰度知识融合到IMED中，以获得更满意的人脸和手写数字识别性能。综上所述，这些源自于IMED的方法要么被修改到不同的应用程序中，要么被嵌入到其他学习任务中，如支持向量机，以提高性能。因此，在接下来的比较研究中，我们只是采用了IMED作为基本嵌入，但它的任何有效变体都可以直接以类似的方式使用。

1.2.2．结构正规化建模策略

在这个家族中，利用空间结构的策略通常采用正则化技术来惩罚一个相关的目标函数，从而使（通过优化目标）得到的解在空间上尽可能平滑[9-13]。参考文献[9]提出的空间平滑子空间学习（SSSL）可以作为其代表，采用拉普拉斯惩罚约束投影系数为空间平滑。Zuo等人[12]更进一步，用高斯权值对拉普拉斯罚函数进行加权，实现多尺度图像平滑。Chen等人在参考文献[13]中通过再次施加拉普拉斯算子开发了一个正则化度量学习框架，并在几个基准数据集上取得了竞争人脸识别性能。从这些研究中可以很容易地发现，结构正则化建模确实可以补偿张量或矩阵到向量转换引起的空间信息损失。因此，对于面向图像的OR，我们也尝试采用这种空间正则化策略。考虑到将空间正则化[9]的这些连续策略应用到我们的问题中是微不足道的，因此在不失去一般性的前提下，我们将参考文献[9]中的空间平滑约束作为基本的正则化策略进行下面的比较研究。

1.2.3．对图像的直接操作策略

前两个家庭的策略都是向量模式导向的。虽然可以利用图像的空间结构信息，提高相关学习性能，但这些策略存在以下问题:（1）计算复杂度高；（2）所谓的“小样本问题”，即特征向量的维数高于训练集的大小，导致过拟合。因此，减轻或解决这些问题的自然方法是直接操作图像（或重塑图像）模式。沿着这条路线，已经发展了许多研究，例如参考文献[14-25]，其中Chen[14-18]和Tao[20-25]等人的作品可作为其代表。更具体地说，Chen等人通过对图像（或重塑图像）模式的双线性投影，在人脸识别和手写数字识别等分类任务中取得了与面向向量模式的对应分类任务相抗衡的性能，开发了一系列分类器，如MatMHKS[14]和MatFE+MatCD [18]；而Tao等人开发了直接操纵（高阶）张量模式的降维或分类建模，并分别将其应用于人类步态识别[20]和视觉跟踪[25]。正是对矩阵或张量作为操作单位的直接操作，使得像双线性投影（二阶张量）这样的方案比向量化的方案更能充分利用数据中所包含的固有空间结构信息。出于类似的考虑，在接下来的比较研究中，我们将图像的双向操作作为一种直接学习方案，与其他方法进行顺序学习性能的比较。

最后，我们在表1中对上述三种策略做了一个简单的比较总结。

|  |  |  |
| --- | --- | --- |
| 策略模型 | 输入模式类型 | 将样式嵌入目标 |
| 结构嵌入的欧几里德距离策略 | 向量 | 间接 |
| 结构正规化建模策略 | 向量 | 直接 |
| 对图像的直接操作策略 | 矩阵 | 直接 |

**表1.** 三种策略的比较总结。

1.3．OR的回顾

在对空间结构信息利用策略进行分类总结之后，下一步我们将以OR为研究平台，在此基础上，我们将对上述三组类别的三组图像集进行广泛的实证比较。在此之前，让我们简要回顾一下OR， OR实际上是一种特殊的学习策略，用于为序数类设计分类器，例如人类年龄估计。OR由于其回归分类的二重性和强大的能力，目前已被广泛应用于推荐系统[26]、网页排序[27]、图像检索[28]、医学图像诊断[29-30]和年龄估计[31-32]等领域。在实施过程中，人们提出了多种方法[33-44]，其中KDLOR[44]是一种著名的OR。虽然这些方法中大多数都取得了不同程度的性能，但在对图像进行操作时，几乎所有的方法都忽略了对向量化图像空间结构信息的补偿，因此，选择面向影像的OR作为研究平台，对归纳出的三种利用空间结构的类型进行比较是合理的。虽然这样一项将空间信息整合到现有OR的工作似乎微不足道，但据我们所知，目前确实还没有相关的研究。

现在，为了清楚起见，又不失一般性，我们将以参考文献[44]中提出的典型OR模型KDLOR的线性版本作为基本OR方法（这里记作LDLOR），选择IMED [1]，以SSSL[9]和双线性建模[14]作为空间结构信息利用对LDLOR进行重构的三种比较代表，得到三种改进的LDLOR版本，分别命名为IMED-LDLOR、SSSL-LDLOR和bill-LDLOR，并对多个图像基准数据集进行了一系列实验，并根据OR特定评价标准报告了对比结果。

本文的其余部分组织如下。在第2节中，我们简要回顾了一个具有代表性的OR，即LDLOR，它作为基础模型（baseline）。在第3节， 详细介绍了利用三种空间结构信息策略得到的三种重建模LDL或对应物。第四部分给出了实验结果并进行了对比分析。结论见第5节。

2．LDLOR的回顾

LDLOR是一种著名的OR，它的目的是寻找有序类的序数指标在投影后能保持良好的最佳投影方向。基于这一原理，LDLOR有两个主要特点:最大限度地扩大相邻序数类的每对均值向量之间的距离，同时最小化类内散射，这使得它不同于DA模型中使用的判别原则，如LDA[45]，因为在LDLOR上施加了数据类之间的相对顺序约束。

现在让为训练集，其中表示第个实例，为其对应的类标号，为数据集大小，为类的总数。则LDLOR可以表示为

其中，表示类内分散矩阵，表示为

其中表示第类的平均向量，为该类的集合大小。

（1）中的表达式是一个典型的二次规划（QP）问题，可以直接求解，也可以利用拉格朗日定理通过其对偶形式求解。

从其目标的制定可以看出，在应用于图像分类时，由于图像到向量的转换，也存在空间信息的丢失。

3．三种重新建模的LDLORs融合了空间结构信息

为了将图像等数据所涉及的空间结构信息应用到LDLOR中，下面我们将简要回顾IMED[1]、SSSL[9]和双线性建模[14]（作为三种空间信息利用策略的代表）的定理，然后利用它们对基本的LDLOR进行重新建模，生成新的变体:IMED-LDLOR、SSSL-LDLOR和Bil-LDLOR。

3.1．IMED–LDLOR

传统的度量方法是一种常用的度量方法，用于度量两个向量之间的相似性或距离。然而，由于它忽略了图像像素之间的空间关系，在应用于图像时，往往会产生不合理的度量结果。具体来说，让和等于2个图像，它们的矢量化版本分别为和。然后是在和之间可以写成

可见，在中，相邻像素之间的空间关系没有反映出来，因为所有像素都是用相同的权值独立处理的。但是，当图像格中的这些像素越接近时，它们对应的灰度值直观上应该越相似。换句话说， 图像通常具有灰度级的局部空间平滑性。正是基于这样的考虑，Wang 等人发明了所谓的来自ED的

和定义为矢量化图像的第和第个像素根据其在原始二维格点上的几何或空间ED的权重，一般与ED值成反比。该方法巧妙地将相邻像素之间的空间平滑关系融入到图像度量中，从而获得合理的图像度量。然而，为了使IMED是一个有效的衡量标准，G必须是半正定的，并且经常使用高斯函数，因此

分别表示矢量化后的图像中的第和第个像素。此外，由于的正半定义，可以扩展为

其中。因此，通过公式（2），x和y的图像之间的IMED实际上等于新投影的u和v之间的ED经过线性变换。

由式（2）可以发现，（1）IMED很容易以类似的方式嵌入到其他度量模型中；（2）通过线性投影变换，将相邻像素之间的空间关系嵌入到度量变换中。接下来我们将IMED嵌入到LDLOR中，生成一个空间结构信息补偿的LDLOR，即IMED-LDLOR表示为

其中表示类内分散矩阵。

其中，为第K序数类的均值向量，K为总类数，为第K个序数类的样本容量。

与（1）中的LDLOR相似，（3）中的IMED-LDLOR也可以用同样的方法求解，这里略去细节。

3.2．SSSL–LDLOR

在LDA[45-46]、LPP[46]和NPE[47]等子空间学习中，在对矢量化图像进行操作时，还需要对图像矢量化过程中产生的空间平滑信息等损失进行空间信息补偿。为此，Cai等人在文献[9]中建立了一个空间平滑学习框架，即SSSL，他们提出将空间平滑信息通过正则化的方式纳入模型学习中。具体地说，设w是一个投影向量，它的维数与向量化后的图像的维数相同，D1（D2）是二阶梯度平滑算子或矩阵在这里沿着一幅图像的行（列），表示为

式中 接下来为了描述图像空间上的整体平滑，我们引入了全局二阶梯度卷积矩阵Δ

是的单位矩阵， 表示克罗内克算子。利用卷积矩阵Δ，我们可以计算w的整个空间平滑度是

（5）的一个直观的解释是，w的项彼此越接近，ℑ的值越小，反之亦然。这样可以反映出空间的平滑性。现在，将（5）添加到（基本）LDLOR的目标函数中，我们可以得到由以下问题导出的新建模的SSSL-LDLOR

通过适当调整超参数λ的值，我们可以在适当的权衡下控制空间的平滑度。直观上，SSSL-LDLOR应该优于LDLOR。

与LDLOR相似，（6）中的SSSL-LDLOR可以直接求解，也可以通过其对偶形式求解。

3.3．Bil-LDLOR

与之前两种针对向量化图像的空间结构信息补偿策略不同，一种更为自然的方法是直接针对图像（或其重构后的矩阵模式）建立分类器。基于这样的出发点，有许多研究得到了发展[14-18]。具体来说，在参考文献[14-18]，利用双线性判别函数设计了一系列面向矩阵的分类器，取代了支持向量机（Support Vector Machines，SVMs）[48]和最小二乘支持向量机（Least Squares Support Vector Machines， LS-SVMs）[49]中的线性分类器，从而，在人脸识别中获得竞争性能。在上述研究的启发下，我们同样将双线性建模的思想引入到OR中进行图像分类，开发出相应的双线性LDLOR，即Bil-LDLOR。虽然这样的想法似乎微不足道，但据我们所知，确实还没有这样的尝试。

为了建立Bil-LDLOR，让我们定义作为一个图像，对应的双线性（判别）函数作用于，u和v分别是左右权向量。然后我们按照基本的LDLOR来建立我们的bill-LDLOR为

其中表示第k类的均值矩阵，是一个来自第k类集合的样本，其他符号的含义与（1）相同。

与（1）中基本LDLOR的推导相比，（7）中的Bil-LDLOR的目标有几个关键优点:（1）（7）中的左右权向量和可以分别由n1和n2自由变量确定，比（1）中的少多了，从而降低了VC维，易于避免过拟合，特别是当训练集的大小远小于维数时；更重要的是，（2） Bil-LDLOR可以直接对矩阵模式进行操作，避免矩阵向向量的转换，从而更理想地反映数据中所涉及的空间结构信息。因此，在对图像等结构化数据进行操作时，Bil-LDLOR应该优于基本的LDLOR。

接下来描述（7）的优化过程，由目标（7）的表述可知，它不再是联合凸，而只是双凸[50]，即对于固定的v（u）它是凸，因此，我们只能采用交替优化迭代策略来求解（7）。Bil-LDLOR的整个优化过程由两个备选优化步骤组成，描述如下:

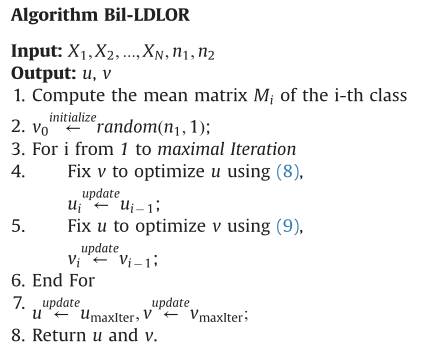
（a）固定v以优化u

这里的以及。

（b）确定u以优化v

这里，很容易发现（8）和（9）的子目标与（1）的子目标在形式上是相同的，因此这里可以直接复制（1）的实现。

下面我们在下面的算法Bil-LDLOR中列出了Bil-LDLOR完整的求解过程。



双凸优化理论[50]，我们理论上证实了Bil-LDLOR算法能够收敛局部最小值（有关详细信息，请参阅附录）和实验也证明它只是需要几种不同的迭代收敛。

3.4. 三种重新建模的LDLORs之间的总体比较

在前面的章节中，我们介绍了三种LDLOR融合空间结构信息的策略，即度量嵌入、结构正则化、空间结构正则化和空间结构正则化，还有双线性模型。在这里，我们简要总结了表2中我们重新建模的LDLORs之间的比较。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| OR方法 | 主要思想 | 凸面 | 输入模式类型 | No.的变量 |
| IMED-LDLOR | 结构嵌入的欧几里德距离策略 | 是 | 向量 |  |
| SSSL-LDLOR | 结构正规化建模策略 | 是 | 向量 |  |
| Bil-LDLOR | 双边投影矩阵 | 否 | 矩阵 |  |

**表2.** 三个重新建模的LDLORs的总体总结。分别为图像或矩阵的行和列。

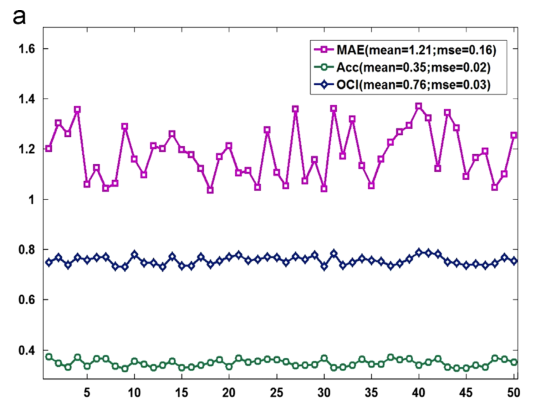
从表2可以看出，IMED-LDLOR和SSSL-LDLOR的目标都可以通过QP优化来求解，而Bil-LDLOR由于u和v的非凸性而不能求解，但仍然可以在保证收敛的前提下交替求解。

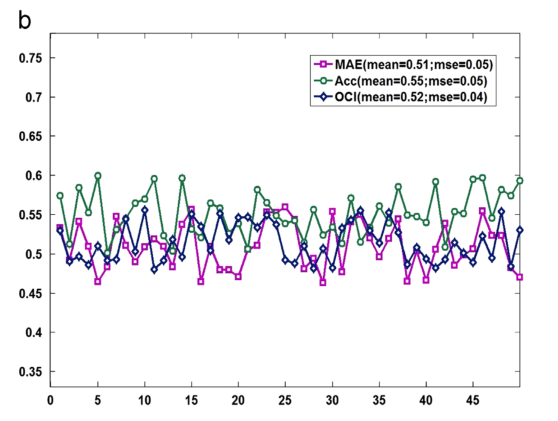
4．实验数据

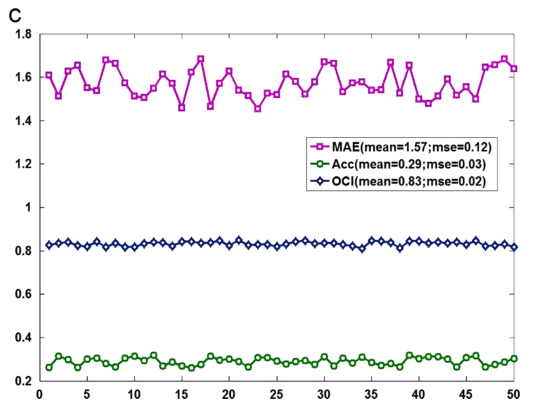
在本节中，我们在JAFFE（人类面部表情强度回归）、UMIST（人类头部姿势回归）和FG-NET（人类年龄组回归）三个基准图像数据集上进行实验，对LDLOR、IMED-LDLOR、SSSL-LDLOR和Bil-LDLOR进行实证比较，它们的类都是顺序的。为了消除图像大小对实验的影响，将所有图像裁剪并调整大小为16 ?16，直接提取原始（像素）灰度水平作为特征来表示图像。

考虑到OR在分类和回归方面的特点，我们使用最常用的标准平均绝对误差（mean absolute error，MAE）和分类精度（classification accuracy，Acc）分别评价其回归偏差和分类性能。其中，为预测结果与真实数据的平均偏差，其值越低，回归性能越好；表示分类器的分类精度，Acc值越大，分类性能越好。另一方面，考虑到OR的对偶性，我们进一步采用了最近提出的专门为OR设计的有序分类指数（OCI）评价标准[51]，OCI值越低，OR性能越好。值得注意的是，与MAE和Acc不同，能够很好地消除标记顺序类别的数值尺度对MAE和Acc的影响，从而能够更恰当地测量预测结果与地面真实值的偏差。在接下来的实验中，我们将里的参数设置为参考[51]中推荐的（1，0.75）。关于的更多细节，由于其定义的复杂性，我们在这里略去，但请参考参考文献[51]中的具体定义（7）。

在我们的实验中，在每个数据集上，我们采用了一个最接近的类均值分类器来执行最终的有序分类，并通过交叉验证（CV）报告了超过20次随机拆分的平均结果。需要注意的是，由于式（7）中Bil-LDLOR的双凸性以及算法Bil-LDLOR中v0的随机初始化，得到的解（收敛性）在v0初始值变化时并不十分稳定，如图1所示。然而，它可以进一步发现，尽管这样单一的解决方案会导致波动或不稳定的性能在测试集，一个平均的解决方案在许多相应的解决方案不同的初始化（如50在图1中，实际上10就足够了）的半（或情况）是相当稳定的。因此，在我们的实验中，对于Bil-LDLOR，所有的表格结果在一次训练数据分割后平均超过10次重复。这里值得指出的是，除了Bil-LDLOR，其余三个OR，即LDLOR、IMED-LDLOR和SSSL-LDLOR的目标都是凸的，因此可以分别得到它们各自的全局最优解，且各自都可以获得不受任何随机初始化影响的稳定性能，这意味着它们在重复过程中的平均，就像对Bil-LDLOR做的那样，总是不变的。





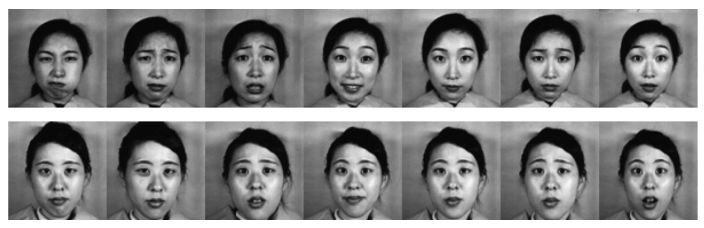


**图1.** 初始化值为时Bil-LDLOR的结果解(横轴表示一次数据分割的重复次数，总共50次)。(a)在JAFFE(#每个类为12)，(b)在UMIST(#每个类为24)，(c)在FG-NET(#每个类为18)。

除上述实验设置外，实验中涉及的所有超或权衡参数的调谐范围为{1e-5，1e-3，1e-1，1e0，1e1，1e3，1e5}。

4.1．JAFFE数据集

JAFFE原始数据集包含了10位日本女性模特的213张7个面部表情（6个基本面部表情到1个中性面部表情）的图像，并由60名日本受试者对每张图像的6个情感形容词进行了评分。在实验中，我们每类选取29个样本，这些选取的样本涵盖了从厌恶到惊讶的7个（顺序）面部表情，其中一些例子如图2所示。



**图2.** 来自JAFFE数据集的示例。

JAFFE实验结果按MAE、Acc和分别列于表3a、表3b和表3c。值得注意的是，经过t检验后，下划线加粗的结果（包括UMIST和FG-NET上的结果）与同一行的其他方法相比在统计学上最好（显著性值p=0.05）。

由表3a、表3b、表3c分别为MAE、Acc、OC的评价指标，可以看出，在JAFFE面部表情回归中，SSSL-、Bil-LDLOR（s）三项评价指标均优于基线LDLOR，尤其是SSSL- LDLOR、这部分说明使用直接的空间正则化目标或直接对图像进行操作的ORs可以优于相应的向量化版本。然而，虽然是嵌入式空间信息，但IMED-LDLOR大多表现最差（在12-24个训练样本的情况下甚至比LDLOR更差），只有在训练集较小的情况下，如，MAE表现稍好（表3a）。另一方面，随着训练样本数量从4个增加到24个（增加4个），从整体上看，所有方法的性能都在不同程度上越来越好，特别是SSSL-LDLOR和Bil-LDLOR分别取得了更显著的性能。例如，在表3b的Acc指标中，基于SSSL正则化的OR分类精度提高了约20个百分点，从0.23提高到0.42，直接OR建模的OR分类精度提高了13个百分点，从0.27提高到0.40。然而，IMED-LDLOR的Acc性能增长特别缓慢，也不太明显，仅为0.20 ~ 0.22的2个百分点，仅为SSSL-LDLOR的1 / 10。这可能是由于将空间结构信息嵌入到ED中只是针对度量，而不是针对需要优化的最终OR准则。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *4* | 1.92±0.21 | ***1.60±0.09*** | 1.78±0.18 | 1.65±0.14 |
| *8* | 1.55±0.16 | 1.48±0.07 | 1.43±0.17 | ***1.29±0.24*** |
| *12* | 1.28±0.12 | 1.44±0.05 | ***1.17±0.15*** | 1.22±0.24 |
| *16* | 1.22±0.36 | 1.46±0.24 | ***1.02±0.10*** | 1.15±0.23 |
| *20* | 1.06±0.24 | 1.40±0.06 | ***0.92±0.13*** | 1.03±0.27 |
| *24* | 0.95±0.16 | 1.36±0.06 | ***0.79±0.14*** | 0.86±0.13 |

**表3a.** LDLORs在JAFFE上的MAE比较(mean±std-dev)。每个序数类的训练样本数用斜体字表示；当实验结果相对最好但与同一行的其他结果相比没有统计学意义时，以粗体斜体表示；当实验结果与同一行的其他实验结果相比，最好的实验结果以粗体斜体下划线表示。NPer表示每个序类的训练样本数(下同)。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *4* | 0.22±0.03 | 0.20±0.04 | 0.23±0.04 | ***0.27±0.04*** |
| *8* | 0.28±0.04 | 0.22±0.04 | 0.29±0.04 | ***0.33±0.04*** |
| *12* | 0.30±0.03 | 0.21±0.03 | 0.33±0.04 | ***0.35±0.05*** |
| *16* | 0.32±0.07 | 0.20±0.04 | ***0.38±0.05*** | 0.37±0.05 |
| *20* | 0.35±0.06 | 0.23±0.04 | ***0.40±0.06*** | 0.39±0.07 |
| *24* | 0.37±0.10 | 0.22±0.06 | 0.42±0.09 | 0.40±0.04 |

**表3b.** LDLORs在JAFFE上的Acc比较(mean±std-dev)。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *4* | 0.88±0.02 | ***0.84±0.00*** | 0.87±0.02 | 0.84±0.03 |
| *8* | 0.84±0.03 | 0.82±0.02 | 0.81±0.03 | ***0.79±0.06*** |
| *12* | 0.79±0.03 | 0.81±0.02 | ***0.76±0.03*** | 0.77±0.06 |
| *16* | 0.76±0.06 | 0.80±0.04 | ***0.73±0.04*** | 0.75±0.06 |
| *20* | 0.73±0.05 | 0.79±0.02 | ***0.69±0.04*** | 0.70±0.07 |
| *24* | 0.69±0.07 | 0.77±0.02 | ***0.62±0.07*** | 0.64±0.04 |

**表3c.** LDLORs在JAFFE上的比较(mean±std-dev)。

4.2．UMIST数据集

原始的UMIST数据集由20个个体的564张图像组成。为了进行OR实验，我们从剖面图到正视图选取了6个连续有序的区间角，每个区间角关联56个样本。即选取6个有序的头姿类，每个类选取56个样本进行头姿回归。一些样品如图3所示。



**图3.**来自UMIST数据集的例子。

在UMIST上的实验结果分别列于表4a、4b、4c。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *6* | 0.74±0.11 | 0.85±0.15 | ***0.72±0.10*** | 0.81±0.14 |
| *12* | 0.61±0.08 | 1.00±0.15 | ***0.57±0.05*** | 0.59±0.09 |
| *18* | 0.58±0.07 | 1.12±0.09 | ***0.48±0.05*** | 0.54±0.06 |
| *24* | 0.58±0.08 | 1.09±0.13 | ***0.43±0.03*** | 0.50±0.06 |
| *30* | 0.60±0.05 | 1.05±0.12 | ***0.38±0.04*** | 0.49±0.06 |
| *36* | 0.67±0.11 | 1.14±0.04 | ***0.37±0.05*** | 0.46±0.05 |
| *42* | 0.85±0.15 | 1.15±0.07 | ***0.33±0.05*** | 0.45±0.07 |
| *48* | 1.26±0.29 | 1.21±0.14 | ***0.30±0.07*** | 0.42±0.07 |

**表4a.** LDLORs在UMIST上的MAE比较(means±d-dev)。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *6* | 0.43±0.05 | 0.41±0.05 | ***0.46±0.05*** | 0.44±0.05 |
| *12* | 0.49±0.05 | 0.38±0.05 | ***0.52±0.04*** | 0.52±0.05 |
| *18* | 0.51±0.04 | 0.35±0.05 | ***0.58±0.04*** | 0.54±0.06 |
| *24* | 0.51±0.05 | 0.31±0.03 | ***0.60±0.03*** | 0.56±0.04 |
| *30* | 0.49±0.03 | 0.27±0.03 | ***0.64±0.03*** | 0.57±0.04 |
| *36* | 0.45±0.04 | 0.24±0.03 | ***0.65±0.05*** | 0.59±0.04 |
| *42* | 0.39±0.07 | 0.25±0.02 | ***0.68±0.05*** | 0.60±0.06 |
| *48* | 0.32±0.08 | 0.24±0.03 | ***0.70±0.06*** | 0.62±0.06 |

**表4b.** LDLORs在UMIST上的Acc比较(means±d-dev)。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *6* | 0.64±0.05 | 0.67±0.06 | ***0.62±0.05*** | 0.67±0.05 |
| *12* | 0.57±0.05 | 0.72±0.05 | ***0.55±0.03*** | 0.56±0.05 |
| *18* | 0.56±0.04 | 0.72±0.01 | ***0.49±0.03*** | 0.53±0.06 |
| *24* | 0.56±0.05 | 0.72±0.05 | ***0.46±0.03*** | 0.51±0.04 |
| *30* | 0.57±0.03 | 0.72±0.05 | ***0.41±0.04*** | 0.50±0.04 |
| *36* | 0.60±0.05 | 0.73±0.02 | ***0.40±0.05*** | 0.48±0.04 |
| *42* | 0.67±0.06 | 0.74±0.03 | ***0.36±0.05*** | 0.46±0.06 |
| *48* | 0.77±0.08 | 0.77±0.04 | ***0.33±0.07*** | 0.43±0.06 |

**表4c.** LDLORs在UMIST上的比较(mean±std-dev)。

考察UMIST中人头位回归表4a-4c的结果可以发现（1）SSSL-LDLOR在所有评价指标中均居第一位，具有绝对的性能优势。更重要的是，随着训练集规模的增大，其优势也越来越明显，如和43时，其Acc性能在表4b约为基本LDLOR的2倍，甚至是IMED-LDLOR的3倍，说明正则化利用空间信息的效果显著；（2）的性能Bil-LDLOR directly-modeled在图像比LDLOR和IMED-LDLOR但不如SSSL-LDLOR，目击者，而空间信息的矢量化无偿口服补液盐，直接操作同样可以利用图像的空间信息，从而促进其或性能；（3）以IMED-LDLOR为主产生最差的性能，例如，平均Acc为0.31，甚至低于LDLOR的0.45。更令人惊讶的是，随着训练样本的增加，它在UMIST和JAFFE上的表现并没有像预期的那样单调增加，而是出现了明显的波动，这似乎是违反直觉的。出现这种情况，除了4.1节中分析的类似原因外，还可能是由于该数据集图像中的头部姿态未对齐。

4.3．FG-NET数据集

FG-NET数据集包含一些年龄从0到69岁的个体。在我们的实验中，我们将所有样本分为8个顺序类别，即0-1岁，2-4岁，5-8岁，9-12岁，13-16岁，17-29岁，30-43岁，44-69岁。每个类别选取43个典型样本，其中部分样本如图4所示。



**图4.** 来自FG-NET数据集的示例。

FG-NET的年龄组回归结果分别列于表5a、表5b、表5c。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *6* | 1.94±0.20 | ***1.77±0.07*** | 2.07±0.25 | 1.84±0.16 |
| *12* | 1.65±0.17 | 1.74±0.22 | 1.71±0.16 | ***1.60±0.17*** |
| *18* | 1.71±0.18 | 1.69±0.06 | ***1.51±0.12*** | 1.57±0.14 |
| *24* | 1.76±0.12 | 1.70±0.07 | ***1.41±0.08*** | 1.51±0.17 |
| *30* | 2.18±0.26 | 1.79±0.13 | ***1.30±0.10*** | 1.50±0.15 |
| *36* | 2.26±0.18 | 1.87±0.11 | ***1.27±0.12*** | 1.49±0.20 |

**表5a.** LDLORs在FG-NET上的MAE比较(means±d-dev)。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *6* | 0.19±0.03 | 0.18±0.02 | 0.19±0.03 | ***0.25±0.03*** |
| *12* | 0.22±0.03 | 0.21±0.02 | 0.21±0.04 | ***0.27±0.02*** |
| *18* | 0.22±0.03 | 0.19±0.02 | 0.24±0.04 | ***0.29±0.02*** |
| *24* | 0.22±0.03 | 0.17±0.02 | 0.27±0.03 | ***0.31±0.03*** |
| *30* | 0.19±0.04 | 0.16±0.02 | 0.27±0.02 | ***0.31±0.04*** |
| *36* | 0.20±0.07 | 0.15±0.03 | 0.28±0.03 | ***0.32±0.06*** |

**表5b.** LDLORs在FG-NET上的Acc比较(means±d-dev)。

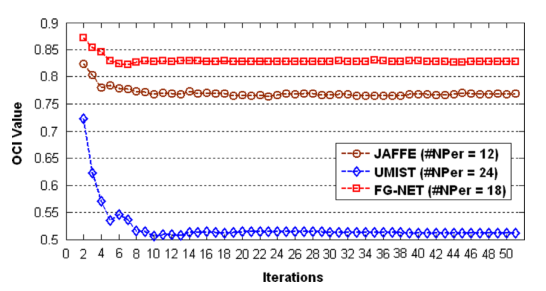
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| NPer | LDLOR | IMED-LDLOR | SSSL-LDLOR | Bil-LDLOR |
| *6* | 0.89±0.02 | ***0.86±0.01*** | 0.90±0.02 | 0.87±0.01 |
| *12* | 0.86±0.02 | 0.85±0.01 | 0.87±0.02 | ***0.84±0.02*** |
| *18* | 0.87±0.02 | 0.85±0.01 | 0.84±0.02 | ***0.83±0.01*** |
| *24* | 0.87±0.02 | 0.85±0.01 | 0.82±0.02 | ***0.81±0.02*** |
| *30* | 0.90±0.03 | 0.87±0.00 | ***0.80±0.02*** | 0.81±0.03 |
| *36* | 0.90±0.02 | 0.87±0.00 | ***0.79±0.03*** | 0.80±0.04 |

**表5c.** LDLORs在FG-NET上的比较(mean±std-dev)。

年龄组FG-NET结果的回归，我们可以发现一些提示:一方面，几乎所有的性能最好的结果都是由SSSL-LDLOR或Bil-LDLOR，特别是对Acc（表5b），后者LDLOR变体保持在大约4百分点击败了前一个。此外，无论是MAE、Acc还是OCI，SSSL-LDLOR和Bil-LDLOR的性能都得到了显著的提高，例如，随着训练样本从6个增加到36个，他们的Acc性能提高了约9个百分比（分别从0.19到0.28和0.23到0.32）。相比之下，基本的LDLOR和基于IMED的性能都没有本质上得到提高，例如Acc指数，它们的百分点分别仅提高了约3个百分点。另一方面，随着训练集规模的增大，SSSL-LDLOR和Bil-LDLOR的各项指标均出现了显著的单调改进，而基础指标和基于IMED的指标均没有出现类似的单调改进趋势。背后的原因可以在4.1和4.2节中进行类似的分析，但这里略去。

4.4．简短的总结

现在综合以上所有的实验结果和分析，我们可以发现对于图像数据集上的OR，SSSL-LDLOR和Bil-LDLOR都可以很好地利用数据中涉及的空间信息，从而显著地提高了它们的OR性能。通过分析它们的本质，我们可以发现SSSL-LDLOR和Bil-LDLOR都对OR目标施加了空间平滑约束，从而通过有意地尊重先验空间知识来提高OR性能。值得注意的是，虽然Bil-LDLOR算法的目标不是联合凸而是双凸，但使用算法Bil-LDLOR，我们可以在大约10次迭代内得到收敛解，如图5所示。



**图5.** OCI下收敛性与Bil-LDLOR交替迭代的关系

相比之下，虽然嵌入了空间信息，但IMED只是反映了在度量矫正中对空间信息的利用，而不是在OR目标优化中，因此并不能保证得到理想的结果。此外，值得注意的是，从表3a与表3b、表5a与表5b的对比来看，各评价指标之间也存在一些不一致之处。例如，以防期数等于18日，24日，30和36 FG-NET分别为人类头部姿态回归，SSSL-LDLOR相比，Bil-LDLOR的所有Acc都更好的结果。但MAE的两个结果都不占主导地位，这意味着良好的分类性能并不一定意味着回归性能好，反之亦然。因此，无论是MAE还是Acc都不够全面，无法提供评估或评价。相对于MAE和Acc来说，OCI是一个更好的OR特殊衡量指数，它的定义是指OR本质上比MAE和Acc更好，为了消除用于标记序数类的数字尺度的影响，因此，我们有偏倚地推荐OCI （）作为有序分类或回归任务的一种合理的评价指标。

5．结论

本文首先对单独提出的空间结构信息利用方法进行了系统的总结，将其分为嵌入结构的欧氏距离保持方法、结构正则化建模方法和图像直接处理方法三大类；第二，进一步做个比较其中条件，空间结构信息很少反映在现有产品口服补液盐，我们分别了使用了IMED，SSSL和双线性建模作为他们的代表，和应用他们重新建模LDLOR（或基线/基本方法）和开发相应的变体:然后分别在JAFFE、UMIST和FG-Net上对人脸表情、头姿和年龄组回归进行了充分的实验，得出结论：

（a）对图像进行直接OR建模的方法，如Bil-LDLOR，可以像2DPCA[52]一样，在一定程度上有效地保留和利用图像中所涉及的空间信息，从而获得显著的OR性能改进。

（b）基于结构正则化的OR，如SSSL-LDLOR，也可以通过将空间信息的正则化强加到它们的目标中，从而在OR性能方面获得明显的好处。

（c）结构嵌入的OR，如IMED-LDLOR，虽然嵌入了空间信息，但在提高OR性能方面通常不具有显著意义。原因是在进行目标优化时没有考虑到空间信息，进一步受到其他一些有待发现的潜在因素的影响。

（d）在OR实验中，指标MAE（用于评估回归偏差）和Acc（测量分类精度）的结果并不总是一致的，这是由于它们在优化过程中并没有捆绑在一起。鉴于OR的二元性，采用OR特异性OCI来更全面地评价OR是合理的，也是值得推荐的。此外，通过比较结果在表3c、4c和5c中w.r.t. OCI（），我们可以发现人类面部表情的困难或JAFFE和年龄或FG-NET几乎在同一水平上，都比人类头上的姿势在曼彻斯特理工回归。也就是说，人类头部姿态的回归相对于其他两个相对容易，这与人类的直觉是一致的。

本文比较研究，我们可以看到，两种策略structure-regularized和直接操作图像可以获得明显的改善或性能通过直接实施目标的空间信息分别通过直接操作或结构正规化，然而，第三类结构嵌入不能产生直观预期的性能优势，其中空间信息只是嵌入度量修改，而不是直接与OR目标相关。因此，我们可以推断空间信息是否能够提高分类器（或回归器）的性能取决于空间结构信息的嵌入方式。

感谢

国家自然科学基金面上项目（no.61170151，no.61073112）、江苏省自然科学基金面上项目（no.BK2012793）、博士学科点专项科研基金面上项目（no. 20123218110033）、江苏省研究生教育创新计划资助项目(CXLX13\_159)、中央高校基本科研业务费资助项目(NZ2013306)、江苏省青兰项目资助。

附录

根据Ref.[50]，求解双凸目标的序列在满足（1）其目标值有下界，（2）其值序列在优化过程中单调递减的条件下可以收敛。据此，我们给出了双凸Bil-LDLOR的收敛性的详细证明如下。

Bil-LDLOR的目标是有下界的。

目标（7）的函数值是有下界的，因为目标（7）等于：

其中，第一项和权衡参数C是非负的，对于给定的训练集是一个常数并且是有界限的。因此是低界限的。

在使用算法Bil-LDLOR优化过程中，Bil-LDLOR的目标值单调减小。

这儿，让 表示目标（7）第i次迭代使用算法Bil-LDLOR的函数值。为了固定v，将式（7）转化为式（8）。显然，问题（8）是一个凸二次规划w.r.t.u，因此，我们可以采用任何现成的优化方法（如SMO[53]）来获得唯一的全局最优解，其对应的函数值为，因此，；下一个固定，将式（7）转化为式（9），式（9）也是凸二次规划w.r.t.（v，ρ），正如在前一次迭代中所做的那样，我们也得到了一个全局最优解，其函数值为，因此，，最后，暗示迭代序列 减少单调了。合并是第界限的，因此，序列 求解目标（7）的方法可以收敛。

参考文献

[1] L. Wang, Y. Zhang, J. Feng, On the Euclidean distance of images, IEEE Trans.

Pattern Anal. Mach. Intell. 27 (2) (2005) 1334–1339.

[2] T. Tangkuampien, D. Suter, 3D Object Pose Inference via Kernel Principal

Component Analysis with Image Euclidean Distance, in: BMVC, 2006.

[3] J. Chen, R. Wang, S. Shan, et al. Isomap Based on the Image Euclidean Distance,

in: ICPR, 2006.

[4] J. Li, B. Lu., A framework for multi-view gender classification, Lecture Notes

Comput. Sci. 4984 (2008) 973–982.

[5] Y. Liu, Y. Liu, K. Chan., Tensor distance based multilinear locality preserved

maximum information embedding, IEEE Trans. Neural Netw. 21 (11) (2010)

1848–1854.

[6] B. Sun, J. Feng, L. Wang, Learning IMED via Shift-Invariant Transformation,

in: CVPR, 2009.

[7] W. Zuo, H. Zhang, D. zhang, et al., Post-processed LDA for face and palmprint

recognition: what is the rationale, Signal Process. 90 (2010) 2344–2352.

[8] J. Li, B. Lu, An adaptive image Euclidean distance, Pattern Recognit. 42 (2009)

349–357.

[9] D. Cai, X. He, Y. Hu, et al., Learning a Spatial Smooth Subspace for Face

Recognition, in: CVPR, 2007.

[10] S. Gu, Y. Tan, X. He., Laplacian smoothing transform for face recognition, Sci.

China 53 (12) (2010) 2415–2428.

[11] Z. Lei, S. Li, Contextual constraints based linear discriminant analysis, Pattern

Recognit. Lett. 32 (2011) 626–632.

[12] W. Zuo, L. Liu, K. Wang, et al., Spatially Smooth Subspace Face Recognition

Using LOG and DOG Penalties, in: ISNN, 2009.

[13] X. Chen, Z. Tong, H. Liu, et al., Metric Learning with Two-Dimensional

Smoothness for Visual Analysis, in: CVPR, 2012.

[14] S. Chen, Z. Wang, Y. Tian., Matrix-pattern-oriented Ho-Kashyap classifier with

regularization learning, Pattern Recognit. 40 (2007) 1533–1543.

[15] Z. Wang, S. Chen, New least squares support vector machines based on matrix

patterns, Neural Process. Lett. 26 (2007) 41–56.

[16] Z. Wang, S. Chen, Matrix-pattern-oriented least squares support vector

classifier with AdaBoost, Pattern Recognit. Lett. 29 (6) (2008) 745–753.

[17] Z. Wang, C. Zhu, D. Gao, et al., Three-fold structured classifier design based on

matrix pattern, Pattern Recognit. 46 (2013) 1532–1555.

[18] Z. Wang, S. Chen, J. Liu, et al., Pattern representation in feature extraction and

classifier design: matrix versus vector, IEEE Trans. Neural Netw. 19 (5) (2008)

758–769.

[19] Z. Zhang, T. Chow, Maximum margin multisurface support tensor machines

with application to image classification and segmentation, Expert Syst. Appl.

39 (2012) 849–860.

[20] D. Tao, X. Li, X. Wu, et al., General tensor discriminant analysis and gabor

features for gait recognition, IEEE Trans. Pattern Anal. Mach. Intell. 29 (10)

(2007) 1700–1715.

[21] D. Tao, M. Song, X. Li, et al., Bayesian tensor approach for 3-D face modeling,

IEEE Trans. Circuits Syst. Video Technol. 18 (10) (2008) 1397–1410.

[22] D. Tao, X. Li, X. Wu, et al., Tensor rank one discriminant analysis – a convergent

method for discriminative multilinear subspace selection, Neurocomputing 71

(2008) 1866–1882.

[23] J. Wen, X. Gao, Y. yuan, et al., Incremental tensor biased discriminant analysis:

a new color-based visual tracking method, Neurocomputing 73 (2010)

827–839.

[24] B. Wang, X. Gao, D. Tao, et al., A unified tensor level set for image segmentation, IEEE Trans. Syst. Man Cybern. -Part B: Cybern. 40 (3) (2010) 857–867.

[25] L. Zhang, L. Zhang, D. Tao, et al., Tensor discriminative locality alignment for

hyperspectral image spectral-spatial feature extraction, IEEE Trans. Geosci.

Remote Sens. 51 (1) (2013) 242–256.

[26] K. Lakiotaki, N. Mastsatsinis, A. Tsoukias, Multicriteria user modeling in

recommender systems, IEEE Intell. Syst. 26 (2) (2011) 64–76.

[27] T. Joachims, Optimizing Search Engineer Using Click-Through Data, in: ACM

SIGKDD, 2012.

[28] H. Wu, H.Q. Lu, S.D. Ma, A Practical SVM-Based Algorithm for Ordinal

Regression in Image Retrieval, in: ACM Multimedia, 2003.

[29] D. Zhang, Y. Wang, L. Zhou, et al., Multimodal classification of Alzheimer's

disease and mild cognitive impairment, NeuroImage 55 (2011) 856–867.

[30] K. Gray, P. Aljabar, R. Heckemann, et al., Random forest-based similarity

measures for multi-modal classification of Alzheimer's disease, NeuroImage

65 (2013) 167–175.

[31] C. Li, Q. Liu, J. Liu, et al., Learning Ordinal Discriminative Features for Age

Estimation, in: CVPR, 2012.

[32] K. Chang, C. Chen, Y. Huang, Ordinal Hyperplanes Ranker With Cost Sensitivies

for Age Estimation, in: CVPR, 2011.

[33] R. Herbrich, T. Graepel, K. Obermayer, Support Vector Learning for Ordinal

Regression, in: ICANN, 1999.

[34] W. Chu, S. Keerthi, New Approaches to Support Vector Ordinal Regression, in:

ICML, 2005.

[35] W. Chu, S. Keerthi, Support vector ordinal regression, Neural Comput. 19 (3)

(2007) 792–815.

[36] R. Herbrich, T. Graepel, K. Obermayer., Large Margin Rank Bound Arises for

Ordinal Regression, MIT Press, Cambridge, MA (2000) 115–132.

[37] A. Shashua, A. Levin, Ranking with Large Margin Principle: Two Approaches,

in: NIPS, 2003.

[38] S. Kramer, G. Widmer, B. Pfahringer, et al., Prediction of ordinal classes using

regression trees, Fundam. Inform. 47 (2001) 1–13.

[39] J. Cheng, Z. Wang, G. Pollastri, A neural network approach to ordinal

regression, in: Proceedings of the IEEE International Joint Conference on

Neural Networks, 2008.

[40] S. Fouad, P. Tino, Adaptive metric learning vector quantization for ordinal

classification, Neural Comput. 24 (2012) 2825–2851.

[41] C. Seah, I. Tsang, Y. Ong, Transductive ordinal regression, IEEE Trans. Neural

Netw. Learn. Syst. 23 (7) (2012) 1074–1086.

[42] Y. Liu, Y. Liu, K. Chan, Ordinal Regression Via Manifold Learning, in: AAAI, 2011.

[43] Y. Liu, Y. Liu, S. Zhong, et al., Semi-Supervised Manifold Ordinal Regression for

Image Ranking, in: ACM Multimedia, 2011.

[44] B. Sun, J. Li, D. Wu, et al., Kernel discriminant learning for ordinal regression,

IEEE Trans. Knowl. Data Eng. 22 (6) (2010) 906–910.

[45] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, Wiley-Interscience,

2000.

[46] X. He, S. Yan, Y. Yu, et al., Face recognition using Laplacian faces, IEEE Trans.

Pattern Anal. Mach. Intell. 27 (3) (2005) 328–340.

[47] X. He, D. Cai, S. Yan, et al., Neighborhood Preserving Embedding, in: ICCV,

2005.

[48] V. Vapnil, Statistical Learning Theory, John Wiley & Sons, New York, 1998.

[49] J. Suykens, J. Vandewalls, Least squares support vector machine classifiers,

Neural Process Lett. 9 (1999) 293–300.

[50] J. Gorski, F. Pfeuffer, K. Klamroth, Biconvex sets and optimization with

biconvex functions: a survey and extensions, Math. Methods Oper. Res. 66

(3) (2007) 373–407.

[51] J. Cardoso, R. Sousa, Measuring the performance of ordinal classification, Int. J.

Pattern Recognit. Artif. Intell. 25 (8) (2011) 1173–1195.

[52] J. Yang, D. Zhang, A. Frangi, et al., Two-dimensional PCA: a new approach to

appearance-based face representation and recognition, IEEE Trans. Pattern

Anal. Mach. Intell. 26 (1) (2004).

[53] B.E. Boser, I.M. Guyon, and V.N. Vapnik, A training algorithm for optimal

margin classifier, in: Proceedings of the 5th ACM Workshop on Computational

Learning Theory, 1992.