

基于结构化多视图稀疏限定的 监督特征选择算法研究

史彩娟, 段昌钰, 谷志斌

(华北理工大学, 河北 唐山 063210)

摘 要: 为了有效利用多视图数据信息提升监督特征选择的性能, 构建了一种结构化多视图稀疏限定, 并基于该稀疏限定提出了一种监督特征选择方法, 即结构化多视图监督特征选择方法(SMSFS)。该方法在特征选择过程中能够同时考虑不同视图特征的重要性以及同一视图中不同特征的重要性, 从而有效的结合多视图数据信息, 提升监督特征选择的性能。SMSFS 目标函数是非凸的, 设计了一个有效的迭代算法对目标函数进行求解。将所提结构化多视图监督特征选择方法 SMSFS 应用到了图像标注任务, 在 NUS-WIDE 和 MSRA-MM2.0 图像数据库上进行了实验, 并与其他特征选择算法进行了比较, 实验结果表明该算法能够有效结合多视图数据信息, 提升特征选择性能。

关 键 词: 多视图学习; 结构化稀疏限定; 监督特征选择

中图分类号: TP 391.4

DOI: 10.11996/JGJ.2095-302X.2018061036

文献标识码: A

文章编号: 2095-302X(2018)06-1036-06

Research of Supervised Feature Selection Algorithm Based on Structured Multi-View Sparse Regularization

SHI Caijuan, DUAN Changyu, GU Zhibing

(North China University of Science and Technology, Tangshan Hebei 063210, China)

Abstract: In order to effectively utilize the multi-view data information and enhance the feature selection performance, the structured multi-view sparse regularization was constructed, based on which a novel supervised feature selection method, namely structured multi-view supervised feature selection (SMSFS), was proposed. SMSFS could simultaneously consider the importance of each view features and the importance of individual feature in each view to combine the multi-view data information effectively in the feature selection process, and then to boost the supervised feature selection performance. Because the objective function of SMSFS is non-convex, an effective iterative algorithm was proposed to solve the objective function. The proposed structured multi-view supervised feature selection method SMSFS was applied into image annotation task and extensive experiments were performed on NUS-WIDE and MSRA-MM2.0 image datasets. The proposed method SMSFS was compared with other feature selection methods and the experimental results demonstrated the effectiveness of SMSFS, which means that it could effectively utilize the multi-view data information to boost the feature selection performance.

收稿日期: 2018-04-16; 定稿日期: 2018-07-24

基金项目: 国家自然科学基金项目(61502143); 河北省自然科学基金面上项目(F2016209165); 华北理工大学杰出青年基金项目(JQ201715); 华北理工大学博士科研启动基金(201510); 河北省高等学校科学技术研究项目(QN2018115)

第一作者: 史彩娟(1977-), 女, 河北丰润人, 副教授, 博士。主要研究方向为图像处理与分析、多媒体内容分析、计算机视觉、机器学习等。

E-mail: scj-blue@163.com

Keywords: multi-view learning; structured sparse regularization; supervised feature selection

近年,随着信息技术以及计算机视觉技术的迅猛发展,图像、视频等数据被表示为不同类型的特征,如形状特征、颜色特征、纹理特征等。每一种类型的特征都是从某一特定空间对图像及视频数据进行了描述,传统上把这种由不同类型特征组成的数据称为多视图数据^[1]。然而,这些多视图数据中包含许多不相关特征,甚至是噪声,因此,如何有效去除多余特征及噪声,从而降低“维数灾难”的特征选择技术成为当前的研究热点。面对多视图数据,传统的单一视图特征选择算法已不能获得较好的特征选择效果,许多多视图特征选择方法被广泛研究和提出。其中,最直接的方法就是将多视图特征直接串接成一个长的特征向量,然后采用单一视图特征选择方法进行特征选择^[2-3]。该方法思路简单,但将多视图特征直接串接为一个长向量之后再进行特征选择,破坏了不同视图特征间的潜在关联,同时也缺乏物理解释,使特征选择的性能受到了影响。

近年来,多视图学习被广泛应用到特征选择算法之中,利用不同视图之间的互补性及一致性来提升特征选择的性能^[1,4-5]。然而,现有的多视图特征选择方法基本上是将一个视图数据看作一个整体,在特征选择过程中仅仅考虑每一个视图特征整体的重要性,而忽略了同一视图中不同特征的重要性,使特征选择性能受到影响。如果在多视图稀疏特征选择过程中,不仅考虑每一个视图的重要性,选择最具判别性的视图特征外,同时考虑同一视图中每一个特征的重要性,必将能够更好地结合多视图数据信息,从而提升特征选择的性能。文献[6]提出了组 l_1 范数(G_1 -范数),能够很好地判定同一视图中不同特征的重要性。文献[7]通过引入组 l_1 范数(G_1 -范数)提出了一种稀疏多模态学习方法来整合多视图异质特征。文献[8]基于 l_1 范数(G_1 -范数)提出了一种新的非监督深度稀疏特征选择方法,特征选择性能得到了提升。

因此,为了更好地利用多视图数据信息,本文提出了一种基于结构化多视图稀疏限定的监督特征选择算法,即结构化多视图监督特征选择(structured multi-view supervised feature selection, SMSFS)。结构化多视图稀疏限定能够使 SMSFS 在特征选择过程中不仅考虑每一个视图的重要性,

同时考虑同一视图中不同特征的重要性。SMSFS 算法被应用到了图像标注任务中,并在图像数据库 NUS-WIDE^[9]和 MSRA-MM 2.0^[10]上进行了实验,结果表明该算法优于其他特征选择算法,能够更好地结合多视图数据信息来提升特征选择的性能。

1 相关工作

本文相关工作作为多视图学习和结构化稀疏限定。

1.1 多视图学习

现实中存在许多多视图数据,如图像经常被表示为形状、颜色、纹理等不同的视觉特征,视频被表示为视觉特征及声音特征,网页被表示为网页内容及相关链接等等。每一个视图数据都具有一定的特征空间,具有特定的物理意义和统计特性。近年,多视图学习得到广泛研究,能够揭示多视图数据间的关联及有效地结合不同视图数据间的信息。根据多视图间信息结合方式的不同,可以将现有的多视图学习算法大致分成3类:①将多视图数据直接串接为一个特征向量,然后采用单一视图方法的多视图学习算法^[11-12]。该算法实现简单,但不能有效揭示不同视图特征之间的关联及补充信息,既没有考虑不同视图特征的重要性,也没有考虑同一视图中不同特征的重要性。②将每一个视图特征看成一个整体,将同一视图中的所有特征赋予相同的系数。虽然该算法能够很好的利用不同视图特征之间的关联性和互补性^[13-16],但没有考虑同一视图特征中不同特征的重要性。③在学习过程中能够同时考虑不同视图的重要性以及同一视图中不同特征的重要性^[7-8]。

1.2 结构化稀疏限定

近年来,多种结构化稀疏限定被广泛研究并被应用到特征选择算法之中。文献[17]对现有的特征选择方法进行了综述,根据所采用的两种不同的结构化稀疏限定 l_1 范数(LASSO)^[18]和 l_{1-p} ($0 < p \leq 1$)矩阵范数^[17],将特征选择方法分成了基于向量的特征选择方法和基于矩阵的特征选择方法。 l_1 范数作为惩罚项对特征选择系数施以约束,能够使被选择的特征尽可能稀疏,使训练的模型具有稀疏性和可解释性,因此被广泛应用到特征选择方法中。

对于 $l_{r,p}$ ($0 < p \leq 1$) 矩阵范数, 当 $r=2, p=1$, 即为 $l_{2,1}$ 范数^[19]。NIE 等^[20]将 $l_{2,1}$ 范数应用到了特征选择之中, 该限定能够保证在特征选择过程中充分考虑不同特征之间的关联, 联合的选取特征。WANG 和 CHEN^[21]提出了 $l_{2,p}$ ($0 < p \leq 1$) 矩阵范数, 并指出当 $p=1/2$ 时, 即 $l_{2,1/2}$ 矩阵范数具有最好的性能。文献[2]基于 $l_{2,1/2}$ 矩阵范数提出了一种图拉普拉斯半监督稀疏特征选择方法。

2 SMSFS 算法

本节提出一个新的基于结构化多视图稀疏限定的监督特征选择框架, 既结构化多视图监督特征选择 (structured multi-view supervised feature selection, SMSFS)。首先介绍结构化多视图稀疏限定, 然后提出了 SMSFS 算法, 最后给出求解目标函数的迭代算法。

设 $\mathbf{X} = [\mathbf{X}^1, \dots, \mathbf{X}^v, \dots, \mathbf{X}^m]^T \in R^{d \times n}$ 为训练图像特征矩阵, 其中, $\mathbf{X}^v = [x_1^v, x_2^v, \dots, x_n^v] \in R^{d_v \times n}$ 为第 v 个视图图像特征矩阵; $d = \sum_{v=1}^m d_v$; n 为训练数据总数; m 为视图数。 $x_i = [x_i^1, x_i^2, \dots, x_i^m]^T \in R^{\left(\sum_{v=1}^m d_v\right) \times 1}$ 为第 i 个训练图像特征, $x_i^v \in R^{d_v \times 1}$ 是其中的第 v 个视图特征向量。 $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T \in \{0, 1\}^{n \times c}$ 为训练图像的标签矩阵, 其中 c 为类别数。

2.1 结构化多视图稀疏限定

设 $\mathbf{G} \in R^{d \times c}$ 是特征选择映射矩阵, 可表示为

$$\mathbf{G} = \begin{bmatrix} g_1^1 & \dots & g_c^1 \\ \dots & \dots & \dots \\ g_1^m & \dots & g_c^m \end{bmatrix} \in R^{d \times c}, \quad \text{其中 } g_j^v \in R^{d_v \times 1},$$

$$g_j \in R^{d \times 1}, \quad (1 \leq j \leq c, 1 \leq v \leq m)。$$

映射矩阵 $\mathbf{G}$ 的组 l_1 -范数 (G_1 范数) 定义为

$$\|\mathbf{G}\|_{G_1} = \sum_{j=1}^c \sum_{v=1}^m \|g_j^v\|_2 \quad (1)$$

映射矩阵 $\mathbf{G}$ 的 $l_{2,1/2}$ 矩阵范数定义为

$$\|\mathbf{G}\|_{2,1/2} = \left(\sum_{i'=1}^d \|g^{i'}\|_2^{1/2} \right)^2 \quad (2)$$

因此, 由 $l_{2,1/2}$ 范数和 G_1 -范数构建的用于特征选择的结构化多视图稀疏限定可以表示为

$$\|\mathbf{G}\|_{2,1/2}^{1/2} + \gamma \|\mathbf{G}\|_{G_1} \quad (3)$$

结构化多视图稀疏限定中 G_1 -范数强调每一个视图的重要性, 而 $l_{2,1/2}$ 矩阵范数可以很好地考虑不同特征的重要性。因此, 结构化多视图稀疏限定能够保证在特征选择的过程中不仅考虑每一个视图的重要性, 同时考虑每一个视图中不同特征的重要性, 更好地结合多视图数据信息, 从而达到很好的特征选择性能。

2.2 SMSFS 算法

特征选择框架一般可以表示为求解下面的最小化经验误差

$$\min_G \text{loss}(\mathbf{G}^T \mathbf{X}, \mathbf{Y}) + \lambda R(\mathbf{G}) \quad (4)$$

其中, $\text{loss}(\cdot)$ 为损失函数; $\lambda R(\mathbf{G})$ 为正则项; λ 为正则化参数。

本文损失函数选择最简单的最小二乘函数, 正则项为结构化多视图稀疏限定, 因此, 本文提出的结构化多视图监督特征选择算法 SMSFS 的目标函数可以表示为

$$\arg \min_G \|\mathbf{X}^T \mathbf{G} - \mathbf{Y}\|_F^2 + \lambda \|\mathbf{G}\|_{G_1} + \mu \|\mathbf{G}\|_{2,1/2}^{1/2} \quad (5)$$

其中, $\lambda \|\mathbf{G}\|_{G_1} + \mu \|\mathbf{G}\|_{2,1/2}^{1/2}$ 为结构化多视图稀疏限定, 保证特征选择过程中不仅考虑不同视图的重要性, 同时考虑同一视图中不同特征的重要性, 更好地结合多视图数据信息, 从而提升特征选择性能; λ 和 μ 为正则化参数; $\|\mathbf{X}^T \mathbf{G} - \mathbf{Y}\|_F^2$ 为损失函数。

2.3 优化求解

由于 $l_{2,1/2}$ 范数是非凸的, 因此本文提出的结构化多视图监督特征选择框架式(5)也是非凸的, 下面提出了一个有效迭代方法来求解目标函数。

设 $\mathbf{G} = [G^1, \dots, G^d]^T$, 定义对角矩阵 $\tilde{\mathbf{D}}$ (对角元素 $\tilde{D}_{i'i'} = 1/4 \|g^{i'}\|_2^{3/2}$), 则 $\|\mathbf{G}\|_{2,1/2}^{1/2} = 4 \text{Tr}(\mathbf{G}^T \tilde{\mathbf{D}} \mathbf{G})$,

$$\|\mathbf{G}\|_{G_1} = \sum_{j=1}^c \sum_{v=1}^m \|g_j^v\|_2。$$

此时目标函数式(5)可以写成

$$\arg \min_G \text{Tr}((\mathbf{X}^T \mathbf{G} - \mathbf{Y})^T (\mathbf{X}^T \mathbf{G} - \mathbf{Y})) + \lambda \sum_{j=1}^c \text{Tr}(g_j^T \mathbf{D}_j g_j) + \mu \text{Tr}(\mathbf{G}^T \tilde{\mathbf{D}} \mathbf{G}) \quad (6)$$

其中, $\mathbf{D}_j$ ($1 \leq j \leq c$) 为块对角矩阵, 其第 v 个对角块为 $\frac{1}{2 \|g_j^v\|_2} \mathbf{I}_v$, $\mathbf{I}_v$ 是大小为 d_v 的单位矩阵。

对式(6)中的 g_j 求导, 并令等于零, 则有

$$\mathbf{X}(\mathbf{X}^T g_j - y_j) + \lambda \mathbf{D}_j g_j + 4\mu \tilde{\mathbf{D}} g_j = 0 \quad (7)$$

因此, 可得到

$$g_j = (\mathbf{X}\mathbf{X}^T + \lambda \mathbf{D}_j + 4\mu \tilde{\mathbf{D}})^{-1} \mathbf{X} y_j \quad (8)$$

算法 1. SMSFS 算法给出了求解目标函数式(5)的迭代算法。

输入: 训练图像特征矩阵为 $\mathbf{X} \in R^{d \times n}$, 训练图像标签矩阵为 $\mathbf{Y} \in R^{n \times c}$, 正则化参数 λ 和 μ 。

- ① 初始化映射矩阵 $\mathbf{G}_0 \in R^{d \times c}$ 为随机矩阵;
- ② 重复计算对角矩阵

$$\tilde{\mathbf{D}}_t = \begin{bmatrix} \frac{1}{4\|g_t^1\|_2^{3/2}} & & \\ & \dots & \\ & & \frac{1}{4\|g_t^d\|_2^{3/2}} \end{bmatrix}$$

计算块对角矩阵 $\mathbf{D}_j$ ($1 \leq j \leq c$), 其中第 v 个对角

块为 $\frac{1}{2\|g_j^v\|_2} \mathbf{I}_v$;

对于 g_j ($1 \leq j \leq c$),

$$(g_j)_t = (\mathbf{X}\mathbf{X}^T + \lambda(\mathbf{D}_j)_t + 4\mu \tilde{\mathbf{D}}_t)^{-1} \mathbf{X} y_j$$

$t=t+1$;

直到收敛;

输出: 最优化映射矩阵 $\mathbf{G} \in R^{d \times c}$ 。

3 实 验

将 SMSFS 算法应用到图像标注任务中, 并在两个大规模图像数据库 NUS-WIDE^[9]和 MSRA-MM 2.0^[10]上进行了实验。

3.1 数据库及实验设置

NUS-WIDE 数据库包含 269 648 幅图像, 81 个概念; MSRA-MM2.0 数据库包含 50 000 幅图像, 100 个概念。实验中对于两个数据库, 均选用 144 维的颜色相关图、128 维的小波纹理和 73 维(NUS-WIDE 数据库)或者 75 维(MSRA-MM2.0 数据库)的边缘方向直方图和 3 种底层视觉特征作为图像的 3 种视图特征。

实验中随机采样 3 000 个图像形成训练数据集, 剩余数据作为测试集。随机进行 10 次试验, 将 10 次结果平均作为最终的标注结果。目标函数

式(5)中的正则化参数 λ 和 μ 设定为 {0.00001, 0.001, 0.1, 1, 10, 1000, 100000}, 采用交叉验证方法记录最好结果。

实验中将算法 SMSFS 与算法 SMML^[6]和 SFUS^[3]进行了比较。采用 3 种评估准则来评价算法性能, 即平均准确率均值(mean average precision, MAP), ROC 曲线下面积 AUC 的两种情况 MicroAUC 和 MacroAUC。

3.2 与其他特征选择算法的性能比较

实验中将本文算法 SMSFS 与另外两种监督特征选择算法 SFUS 和 SMML 进行了比较, 在两个数据库上进行了实验, 结果见表 1, 最好结果用黑体标识。

表 1 不同特征选择方法性能比较

数据 库	评估准则	SFUS 算法	SMML 算法	SMSFS 算法
NUS	MAP	0.096	0.113	0.116
	MicroAUC	0.781	0.830	0.836
	MacroAUC	0.677	0.794	0.800
MSRA	MAP	0.048	0.072	0.076
	MicroAUC	0.809	0.880	0.885
	MacroAUC	0.554	0.664	0.669

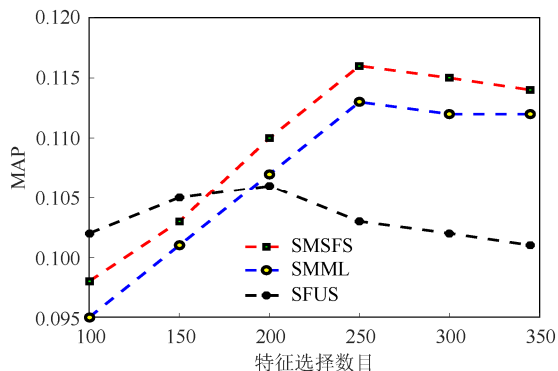
从表 1 可以看出, 无论是基于 MAP, 还是 MacroAUC 和 MicroAUC, 本文算法 SMSFS 性能优于 SFUS 和 SMML 算法, 说明 SMSFS 算法能够利用结构化多视图稀疏限定有效的结合多视图数据信息, 从而选取更具判别性的稀疏特征, 获得更好的特性选择性能。

3.3 选择特征个数的影响

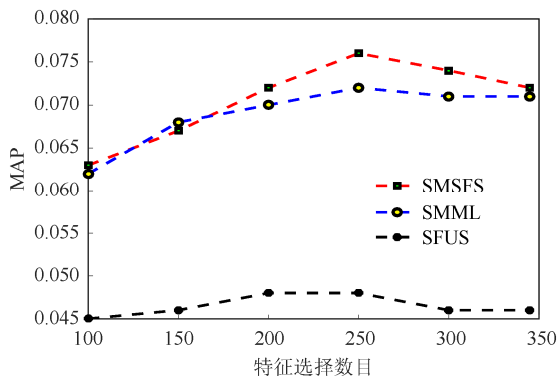
本实验验证所选特征数目对标注性能的影响。设定所选特征数目分别为 100、150、200、250 和 300 以及全部特征(对于 NUS-WIDE 数据库是 345 维, 对于 MSRA-MM 数据库是 347 维)。采用 MAP 作为评估准则, 实验结果如图 1 所示。

图 1 显示了算法 SMSFS 标注性能随着所选特征数目变化而变化的情况。从图 1 可以看出: ① 当所选用的特征数很少时(如特征数为 100)对应的 MAP 要低于不进行特征选择时(特征数为 345 或 347)的 MAP。这归因于选择的特征数很少时, 一些有用信息被丢失。② 当采用所有底层视觉特征(特征数为 345 或 347)的时候, 标注性能并不是最

好,说明最初的底层视觉特征中包含一些冗余特征,影响了标注性能。③在两个数据库上,3种算法达到最好性能时所选用的特征数目均为250,但是算法 SMSFS 性能优于另外两种算法。因此,与另外两种特征选择算法相比,本文算法 SMSFS 能够有效结合多视图数据信息,选取更具判别性的特征,从而获得最好的特征选择性能,降低噪声影响。



(a) NUS-WIDE 数据库



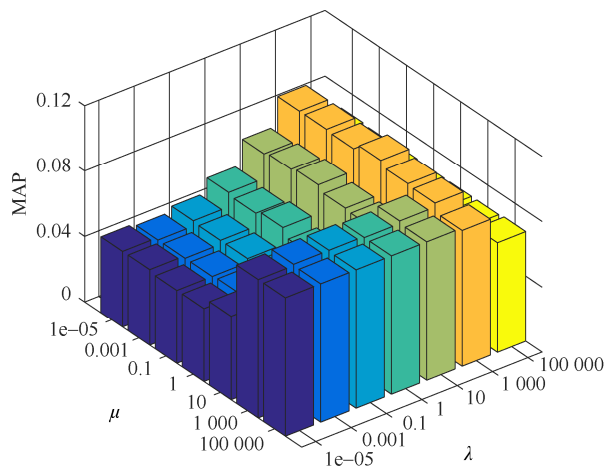
(b) MSRA-MM 数据库

图1 性能 MAP 随着选择特征数目的变化而变化情况

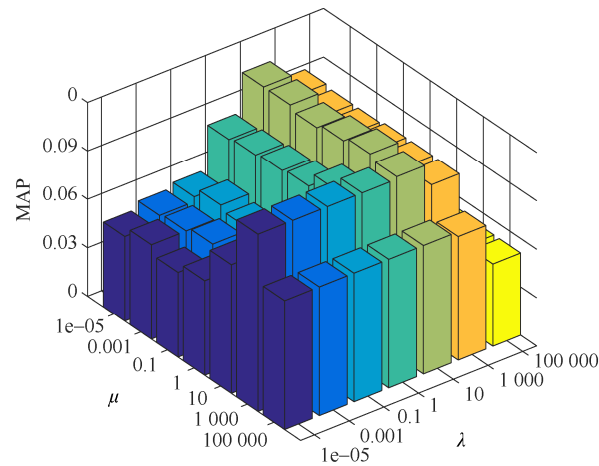
3.4 参数敏感性测试

在目标函数式(5)中存在两个正则化参数 λ 和 μ , 本节通过交叉验证的方式验证这两个正则化参数对特征选择性能的影响,其大小设定为 $\{0.00001, 0.001, 0.1, 1, 10, 1000, 10000\}$ 。采用 MAP 作为评估准则,实验结果如图2所示。

从图2可以看出,不同的 λ 和 μ 组合对应不同的 MAP 结果,该算法对于参数 λ 和 μ 比较敏感。对于 NUS-WIDE 数据库,当 $\lambda=1000$ 、 $\mu=10$ 时具有最好的性能;对于 MSRA-MM 数据库,当 $\lambda=\mu=10$ 时性能最好。



(a) NUS-WIDE 数据库



(b) MSRA-MM 数据库

图2 正则化参数敏感性分析

4 结束语

本文提出了一个新的基于结构化多视图稀疏限定的监督特征选择算法 SMSFS。本文算法能够在特征选择过程中同时考虑不同视图的重要性以及同一视图中不同特征的重要性,有效结合多视图数据信息,从而提升特征选择的性能。其目标函数是非凸的,本文设计了一个有效的迭代算法对目标函数进行求解。将该算法应用到了图像标注任务中,并在两个图像数据库上进行了实验,结果表明该算法优于现有的特征选择算法,很好地实现了特征选择。

参考文献

- [1] FENG Y F, XIAO J, ZHUANG Y T, et al. Adaptive unsupervised multi-view feature selection for visual concept recognition [C]//11th IEEE Asian Conference on Computer Vision (ACCV), 2012. New York: IEEE

- Press, 2012: 343-357.
- [2] SHI C J, RUAN Q Q, AN G Y, et al. Sparse feature selection based on graph Laplacian for web image annotation [J]. *Image and Vision Computing*, 2014, 32(3): 189-201.
 - [3] MA Z G, NIE F P, YANG Y, et al. Web image annotation via subspace-sparsity collaborated feature selection [J]. *IEEE Transactions on Multimedia*, 2012, 14(4): 1021-1030.
 - [4] SHI C J, RUAN Q Q, AN G Y, et al. Semi-supervised sparse feature selection based on multi-view laplacian regularization [J]. *Image and Vision Computing*, 2015, 33(41): 1-10.
 - [5] XU X X, LI W, XU D, et al. Co-labeling for multi-view weakly labeled learning [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016, 38(6): 1113-1125.
 - [6] WANG H, NIE F P, HUANG H, et al. Identifying disease sensitive and quantitative trait-relevant biomarkers from multidimensional heterogeneous imaging genetics data via sparse multimodal multitask learning [J]. *Bioinformatics*, 2012, 28(12): 1127-1136.
 - [7] WANG H, NIE F P, HUANG H, et al. Heterogeneous visual features fusion via sparse multimodal machine [C]//*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013. New York: IEEE Press, 2013: 3097-3102.
 - [8] CONG Y, WANG S, FAN B J, et al. UDSFS: unsupervised deep sparse feature selection [J]. *Neurocomputing*, 2016, 196: 150-158.
 - [9] CHUN T, TANG J H, HONG R, et al. NUS-WIDE: a real-world web image dataset from National University of Singapore [C]//*Proceeding of the ACM International Conference on Image and Video Retrieval (CIVR)*, 2009. New York: ACM Press, 2009: 1-9.
 - [10] LI H, WANG M, HUA X S. MSRA-MM2. 0: a large-scale web multimedia dataset [C]//*IEEE International Conference on Data Mining Workshops (ICDMW)*, 2009. New York: IEEE Press, 2009: 164-169.
 - [11] MA Z G, NIE F P, YANG Y, et al. Discriminating joint feature analysis for multimedia data understanding [J]. *IEEE Transactions on Multimedia*, 2012, 14(6): 1662-1672.
 - [12] LIU W F, TAO D C, Multiview hessian regularization for image annotation [J]. *IEEE Transactions on Image Processing*, 2013, 22(7): 2676-2687.
 - [13] WANG S P, GUO W Z. Sparse multi-graph embedding for multimodal feature representation [J]. *IEEE Transactions on Multimedia*, 2017, 19(7): 1454-1466.
 - [14] XUE X W, NIE F P, WANG S, et al. Multi-view correlated feature learning by uncovering shared component [EP/OL]. [2-18-01-10]. <https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14439>.
 - [15] JING X Y, WU F I, DONG X W, et al. Semi-supervised multi-view correlation feature learning with application to webpage classification [EP/OL]. [2018-01-20]. <https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14582>.
 - [16] SHI C J, RUAN Q Q, AN G Y, et al. Semi-supervised sparse feature selection based on multi-view laplacian regularization [J]. *Image and Vision Computing*, 2015, 41(C): 1-10.
 - [17] GUI J, SUN Z N, JI S W, et al. Feature selection based on structured sparsity: a comprehensive study [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(7): 1490-1507.
 - [18] SCHÖLKOPF B, PLATT J, HOFMANN T. Sparse multinomial logistic regression via bayesian L_1 regularisation [C]//*Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*. New York: MIT Press, 2006: 209-216.
 - [19] DING C, ZHOU D, HE X F, et al. R_1 -PCA: rotational invariant L_1 -norm principal component analysis for robust subspace factorization [C]//*ICML'06 Proceedings of the 23rd International Conference on Machine Learning (ICML)*, 2006. New York: ACM Press, 2006: 281-288.
 - [20] NIE F P, XU D, HUNG T, et al. Flexible manifold embedding: A framework for semi-supervised and unsupervised dimension reduction [J]. *IEEE Transactions on Image Processing*. 2010, 19 (7): 1921-1932.
 - [21] WANG L P, CHEN S G. $L_{2,p}$ matrix norm and its application in feature selection [EP/OL]. [2018-01-08]. https://www.researchgate.net/publication/301850380_1_2p_Matrix_Norm_and_Its_Application_in_Feature_Selection.