Multi-Task Sparse Regression Metric Learning for Heterogeneous Classification

Haotian Wu, Bin Zhou, Pengfei Zhu, Qinghua Hu, Hong Shi
College of Intelligence and Computing Tianjin University

Introduction

With the development of digital devices, industrial sensors and the ubiquitous of social networks, there are explosive growth of heterogeneous data from different sources. In the field of information technology, a web page can be described either by the text information in the web page or by the image information attached to the anchor chain of the web page. In computer vision, images are often described with different types of descriptors (such as HOG ,SIFT and LBP).

Heterogeneous data contain complementary information, which can be combined together to boost the recognition or clustering performance. To match heterogeneous data, great efforts have been devoted to learn a unified representation of the information in different sources so we can use heterogeneous data more accurately and efficiently For multi-view learning, the key challenge is how to balance the commonality and individuality of different views so that the multi-view information can be well fused.

Distance metric learning aims to learn a distance metric to measure the difference between two samples. Metric learning has been widely applied in real-world applications including face recognition, face identification ,image classification and person re-identification. Hu[1] proposed a multi-view deep metric learning(MvDML) approach by jointly learning an optimal combination of multiple distance metrics on multi-view representations.

Compared with traditional single-view data, heterogeneous data contain more information which helps improve the classification or regression performances. Existing metric learning methods ,such as ITML (Information-Theoretic Metric Learning) and GMML (Geometric Mean Metric Learning), achieve good performance in traditional single view learning tasks. However, they fail to jointly exploit the complementary information from the heterogeneous data if they are directly applied to multi-view tasks.

In this paper, we propose a novel multi-task group sparse regression metric learning (MT-SRML) for heterogeneous classification. We use a 2-degree polynomial kernel for sample pairs in each tasks to get the sample pair relationships in the feature space. Inspired by multi-task dictionary learning, MT-SRML jointly learns distance metrics from heterogeneous data and imposes a group sparse regularization item on the coefficient vectors. The proposed model aims to fuse the discriminative capabilities of different views to help improve the efficiency and accuracy of classification results. Experiments on four benchmark datasets show that MT-SRML outperforms the state-of-the art metric learning algorithms, and achieves much better performance than single view learning.

Method

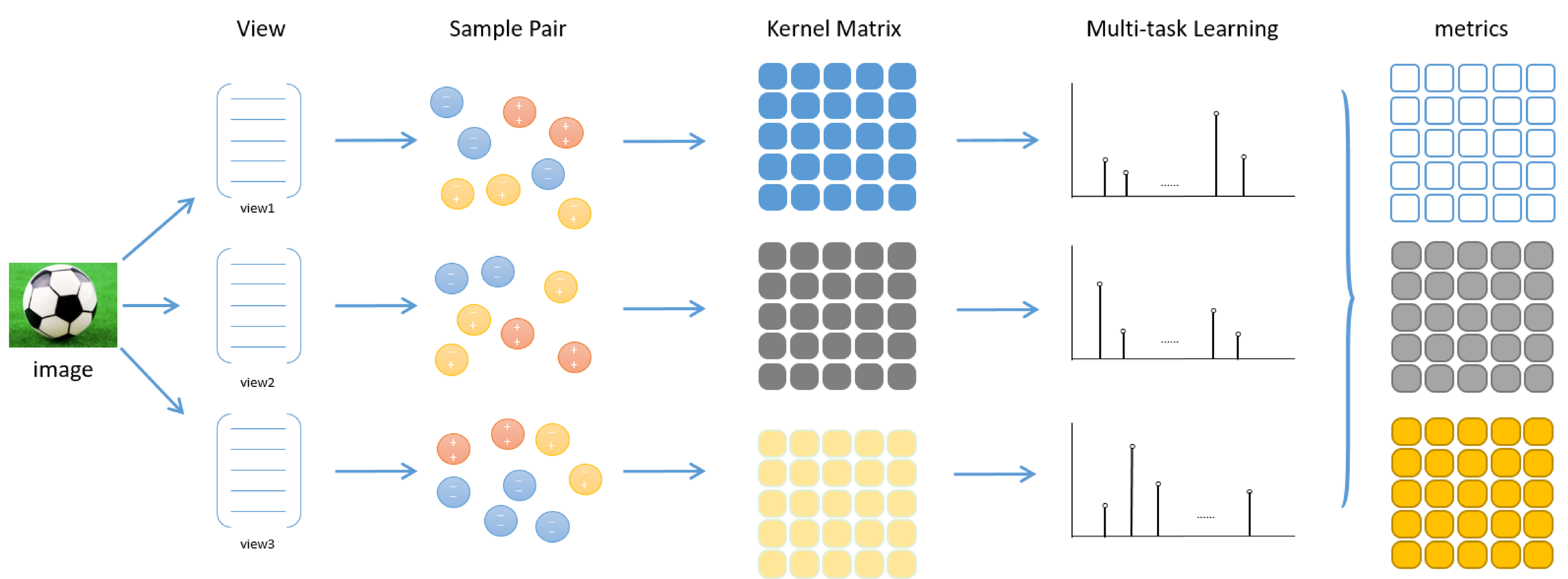

Figure 1. The flowchart of multi-task sparse regression metric learning model.

The flowchart of the proposed metric learning model is given in Figure 1. For an image, different types of feature descriptors can be extracted to construct a multi-view learning task. A multi-task sparse regression problem is solved and the regression coefficients are obtained for all views. Finally, a set of distance metrics for each view are got. The multi-task sparse regression metric learning model is formulated as follows:

$$min \sum_{i=1}^{m} \|z - a_i K_i\|_2^2 + \lambda R(a_1, a_2, \dots, a_m) \quad s.t.\, i = 1,2,\dots,M \qquad (1)$$

Here z represents the decision space information of doublets samples and we suppose sample pair kernel generated N doublets samples. a_i is the reconstruction coefficient vectors associated with each view of the heterogeneous data and $i = 1,2,\dots,M$. K is the 2-degree polynomial kernel for doublets sample pairs calculated by sample pair kernel. λ is a positive constant and R is the joint regularization item imposed on a_i .Our model is formulated as the solution to the following problem of multi-task sparse regression metric learning with l_2 mixed-norm regularization, then we rewrite (1) as:

$$min \sum_{i=1}^{m} \|z - a_i K_i\|_2^2 + \lambda \sum_{n=1}^{N} \|a^n\|_2 \quad s.t.\, i = 1,2,\dots M \qquad (2)$$

Algorithm

Algorithm 1 The algorithms of our proposed MT-SRML

Require:

Generate sample pairs (x_{i1}, x_{i2}),$i = 1, 2, \dots, N$. and kernel $\mathbf{K}$

Compute sample relation $\mathbf{z}$.

Initialization: Properly initialize $\hat{v}_{m,0}$ and $\hat{a}_{m,0}$ .Set $\alpha_0 = 1$, $t \leftarrow 0$ and $\eta \leftarrow 0.002$

repeat:

$\hat{a}_{m,t+1} = \hat{v}_{m,t} - \eta \nabla_{m,t}$

$\hat{a}_{t+1}^{n} = \left[1 - \frac{\lambda\eta}{\|\hat{a}_{t+1}^{n}\|_2}\right]_+ \hat{a}_{t+1}^{n}, \quad n = 1, \dots, N$

$\alpha_{t+1} = \frac{2}{t+2}$

$\hat{V}_{t+1} = \hat{A}_{t+1} + \frac{\alpha_{t+1}(1-\alpha_t)}{\alpha_t}(\hat{A}_{t+1} - \hat{A}_t)$

$t \leftarrow t + 1$

until convergence

Ensure:

$\mathbf{D}_m = \sum_{i=1}^{N} a_m \mathbf{T}_i, \quad m = 1, \dots, M$

The details of our method MT-SRML are given in Algorithm 1. we chose the popularly applied Accelerated Proximal Gradient(APG) model to efficiently solve problem (2).The details of the APG algorithm can be seen in paper [2].

Results

Table 1. Comparison with metric learning methods on heterogeneous data

Method	handwritten	Caltech101	MSRA	football
kNN	0.941±0.015	0.882±0.012	0.700±0.092	0.631±0.077
ITML	0.948±0.013	0.915±0.016	0.769±0.057	0.580±0.078
LMNN	0.922±0.020	0.830±0.061	0.767±0.098	0.416±0.053
LDML	0.944±0.012	0.882±0.012	0.700±0.092	0.627±0.078
GMML	0.939±0.011	0.881±0.013	0.702±0.095	0.651±0.070
HMML	0.927±0.013	0.921±0.013	0.798±0.044	0.522±0.093
EMGMML	0.839±0.012	0.919±0.007	0.802±0.030	0.702±0.088
Ours	**0.983±0.006**	**0.925±0.010**	**0.905±0.050**	**0.864±0.034**

Table 2. Comparison on each feature of heterogeneous datasets

feature	football	handwritten	MSRA	Caltech101
1	0.814	0.975	0.143	0.733
2	0.747	0.768	0.643	0.829
3	0.814	0.778	0.833	0.829
4	0.651	0.560	0.738	0.959
5	0.628	0.910	0.738	0.860
6	0.512	0.583	0.452	0.949
all	**0.864**	**0.983**	**0.905**	**0.971**

As shown in Table1, we compared our method with other 7 multi-view metric learning methods. We can see from Table1 that our method outperforms all the comparisons on these four heterogenous datasets which proves our MT-SRML method is more appropriate than the other method when dealing with heterogeneous data.

As shown in Table2, the first six rows represent 6 different features of our datasets while the last row shows the result of our method on the entirety heterogenous data. The result shows that our jointly learning method achieves a remarkable effect on heterogeneous data.

Conclusion

In this paper, we proposed a novel multi-task sparse regression metric learning (MT-SRML), which aims to jointly learn distance metrics from heterogeneous data. Metric learning is formulated as a sample pair regression task. A two-degree polynomial kernels is introduced to measure the relation of sample pairs.

Metric learning for heterogeneous data is modelled as a multi-task sparse regression problem. The proposed model jointly learns a set of distance metrics from heterogeneous data to fuse the discriminative capabilities of different views to help improve the efficiency and accuracy of classification. Experiments on four heterogeneous datasets validated the superiority of the proposed model to the state-of-the-art metric learning algorithms.

References

[1]Hu, J., Lu, J., Tan, Y.P.: Sharable and individual multi-view metric learning. IEEE Transactions on Pattern Analysis and Machine Intelligence PP(99), 1{1 (2017), doi:10.1109/TPAMI.2017.2749576

[2]Nesterov, Y.: Gradient methods for minimizing composite objective function. Core Discussion Papers **140**(1), 125{161 (2007), doi:10.1007/s10107-012-0629-5

Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grants 61502332, 61876127 and 61732011, Natural Science Foundation of Tianjin Under Grants 17JCZDJC30800, Key Scientific and Technological Support Projects of Tianjin Key R$\&$D Program 18YFZCGX00390 and 18YFZCGX00680.