نویسندگان:

T. Zare - *Signal Processing Research Group, Electrical Engineering Department, Yazd University, Yazd, Iran*.

M. T. Sadeghi - *Signal Processing Research Group, Electrical Engineering Department, Yazd University, Yazd, Iran*.

H. R. Abutalebi - *Signal Processing Research Group, Electrical Engineering Department, Yazd University, Yazd, Iran*.

J. Kittler - *Centre for Vision, Speech and Signal Processing (CVSSP), Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, Surrey*.

خلاصه مقاله:

Machine-learning solutions to classification, clustering and matching problems critically depend on the adopted metric, which in the past was selected heuristically. In the last decade, it has been demonstrated that an appropriate metric can be learnt from data, resulting in superior performance as compared with traditional metrics. This has recently stimulated a considerable interest in the topic of metric learning, especially using kernel functions, which map data to feature spaces with enhanced class separability, and implicitly define a new metric in the original feature space. The formulation of the problem of metric learning depends on the supervisory information available for the task. In this paper, we focus on semi-supervised kernel based distance metric learning where the training data set is unlabelled, with the exception of a small subset of pairs of points labelled as belonging to the same class (cluster) or different classes (clusters). The proposed method involves creating a pool of kernel functions. The corresponding kernels matrices are first clustered to remove redundancy in representation. A composite kernel constructed from the kernel clustering result is then expanded into an orthogonal set of basis functions. The mixing parameters of this expansion are then optimised using point similarity and dissimilarity information conveyed by the labels. The proposed method is evaluated on synthetic and real data sets. The results show the merit of using similarity and dissimilarity information jointly as compared to using just the similarity information, and the superiority of the proposed method over all the recently introduced metric learning approaches.

کلمات کلیدی:

Distance Metric Learning, Semi-supervised Clustering, Composite Kernels, Pairwise Similarity and Dissimilarity Constraints, Optimisation Problem

لینک ثابت مقاله در پایگاه سیویلیکا:

https://civilica.com/doc/894129