

Bag-of-keypoints approach for Tamil handwritten character recognition using SVMs

Subashini, A. And Kodikara, N.D.

Abstract

In this paper, the bag-of-keypoints approach for the off-line recognition of Tamil handwritten characters is investigated. Various pre-processing operations are performed on the digitised image to enhance the quality of an image. In the proposed method each pre-processed character image is represented by a set of local-invariant SIFT feature vectors. From a set of reference vectors, the key idea is to create a codebook for each character using K-means clustering algorithm. Then, the bag-of-keypoints are computed for the total number of character images. These features are used to train a linear support vector machine. A target character is predicted to exactly one of the twenty character classes. An average recognition rate of 81.62% on the character level has been achieved in experiments using six thousand training and two thousand testing images of twenty selected character classes. These results clearly demonstrate that the method produces good recognition accuracy on the handwritten Tamil character database and can be extended with more characters and more samples being recognised.