

Όραση Υπολογιστών: Εργασία 3

Ζητείται να υλοποιηθεί πρόγραμμα σε Python με τη χρήση της βιβλιοθήκης OpenCV το οποίο θα αφορά στο πρόβλημα της ταξινόμησης πολλαπλών κλάσεων (multi-class classification). Το πρόγραμμα υλοποιείται με την εκτέλεση των παρακάτω βημάτων :

1. Παραγωγή οπτικού λεξικού (visual vocabulary) βασισμένη στο μοντέλο Bag of Visual Words (BOVW). Η δημιουργία του λεξικού να γίνει με τη χρήση του αλγορίθμου K-Means χρησιμοποιώντας όλες τις εικόνες του συνόλου εκπαίδευσης (imagedb_train).
2. Εξαγωγή περιγραφέα σε κάθε εικόνα εκπαίδευσης (imagedb_train) με βάση το μοντέλο BOVW χρησιμοποιώντας το λεξικό που προέκυψε κατά το βήμα 1. Για το βήμα αυτό δεν μπορείτε να χρησιμοποιήσετε τη σχετική κλάση της OpenCV (cv.BOWImgDescriptorExtractor).
3. Με βάση το αποτέλεσμα του βήματος 2, να υλοποιηθεί η λειτουργία ταξινόμησης μιας εικόνας κάνοντας χρήση των δυο παρακάτω ταξινομητών :
 - α. Του αλγορίθμου k-NN χωρίς τη χρήση της σχετικής OpenCV συνάρτησης (cv.ml.KNearest_create()).
 - β. Του σχήματος one-versus-all όπου για κάθε κλάση εκπαιδεύεται ένας SVM ταξινομητής.
4. Αξιολόγηση του συστήματος: Χρησιμοποιώντας το σύνολο δοκιμής (imagedb_test), να μετρηθεί η ακρίβεια του συστήματος (και στις δύο περιπτώσεις ταξινομητών) που εκφράζεται ως το ποσοστό των επιτυχών ταξινομήσεων. Κατά την αξιολόγηση να ελέγξετε την επίδραση των εμπλεκόμενων παραμέτρων, όπως ο αριθμός των οπτικών λέξεων (Βήμα 1), ο αριθμός των πλησιέστερων γειτόνων (Βήμα 3α) και ο τύπος του πυρήνα (kernel) του SVM (Βήμα 3β).

Σημείωση 1: Στα πλαίσια της εργασίας αυτής δίνονται δύο βάσεις εικόνων 'imagedb_train' και 'imagedb_test'. Οι βάσεις αυτές αποτελούν υποσύνολο της βάσης εικόνων 'Caltech-256' (http://www.vision.caltech.edu/Image_Datasets/Caltech256/). Η πρώτη (imagedb_train) θα χρησιμοποιηθεί για την εκπαίδευση του συστήματος (Βήμα 1, 2, 3) και η δεύτερη ('imagedb_test') για την δοκιμή και την αξιολόγησή του (Βήμα 4).

Download link: <https://vc.ee.duth.gr:6960/index.php/s/61kOkY922Mp70N7>