



DEMOCRITUS
UNIVERSITY
OF THRACE

DEPARTMENT OF
ELECTRICAL & COMPUTER
ENGINEERING



Όραση Υπολογιστών

Εργασία 3



Ον/μο Φοιτητή: Μπιτζίδης Χαράλαμπος

AM: 57424

Περιεχόμενα:

- Σκοπός Εργασίας
- Παρουσίαση του σετ δεδομένων εκπαίδευσης
- Παραγωγή οπτικού λεξικού βασισμένη στο μοντέλο Bag of Visual Words
- Εξαγωγή περιγραφέα σε κάθε εικόνα εκπαίδευσης ‘**imagedb**’ με βάση το μοντέλο BOVW
- Ταξινόμηση μιας εικόνας κάνοντας χρήση των δυο παρακάτω ταξινομητών :
 - α. Του αλγορίθμου k-NN
 - β. Του σχήματος one-versus-all όπου για κάθε κλάση εκπαιδεύεται ένας SVM ταξινομητής.
- Αξιολόγηση του συστήματος.
- Εξέταση κλάσεων, όπου υπάρχει ιδιαίτερη αποτυχία και δικαιολόγηση των λόγων αυτής

Σκοπός Εργασίας

Η συγκεκριμένη εργασία είναι η τρίτη από μια σειρά εργασιών που θα δοθούν στο πλαίσιο του μαθήματος Όραση Υπολογιστών και έχει ως στόχο την ταξινόμηση πολλαπλών κλάσεων (multi-class classification) χρησιμοποιώντας δεδομένα εικόνων.

Το σετ δεδομένων (dataset) αποτελείται από δύο βάσεις εικόνων, αυτές των φακέλων ‘**imagedb**’ και ‘**imagedb_test**’, αντίστοιχα. Η πρώτη ‘**imagedb**’ θα χρησιμοποιηθεί για την εκπαίδευση του συστήματος και η δεύτερη ‘**imagedb_test**’ για την δοκιμή και την αξιολόγησή του. Οι βάσεις αυτές αποτελούν υποσύνολο της βάσης εικόνων *Caltech 256 Image Dataset*.

Στην εργασία αυτή το ζητούμενο ήταν να αναπτυχθεί μια μεθοδολογία η οποία:

1. Θα παρήγαγε ένα οπτικό λεξικό (visual vocabulary) με βάση το μοντέλο Bag of Visual Words (BOVW). Η δημιουργία του λεξικού να γινόταν με τη χρήση του αλγορίθμου K-Means χρησιμοποιώντας όλες τις εικόνες του συνόλου εκπαίδευσης ‘**imagedb**’.
2. Θα εξήγαγε περιγραφέα για κάθε εικόνα εκπαίδευσης ‘**imagedb**’ με βάση το μοντέλο BOVW χρησιμοποιώντας το λεξικό που προέκυψε κατά το βήμα 1. Για το βήμα αυτό δεν ήταν δυνατή η χρήση της σχετικής κλάσης της OpenCV (cv.BOWImgDescriptorExtractor).
3. Με βάση το αποτέλεσμα του βήματος 2, θα υλοποιούσε την λειτουργία ταξινόμησης μιας εικόνας κάνοντας χρήση των δυο παρακάτω ταξινομητών :
 - a. Του αλγορίθμου k-NN χωρίς τη χρήση της σχετικής OpenCV συνάρτησης (cv.ml.KNearest_create()).
 - b. Του σχήματος one-versus-all όπου για κάθε κλάση εκπαιδεύεται ένας SVM ταξινομητής.
4. Θα αξιολογούσε το συστήμα:
5. Χρησιμοποιώντας το σύνολο δοκιμής (imagedb_test), να μετράει την ακρίβεια του συστήματος (και στις δύο περιπτώσεις ταξινομητών) που εκφραζόταν ως το ποσοστό των επιτυχών ταξινομήσεων, συνολικά και ανά κλάση. Κατά την αξιολόγηση θα έλεγχε την επίδραση των εμπλεκομένων παραμέτρων, όπως ο αριθμός των οπτικών λέξεων (Βήμα 1), ο αριθμός των πλησιέστερων γειτόνων (Βήμα 3a) και ο τύπος του πυρήνα (kernel) του SVM (Βήμα 3b).

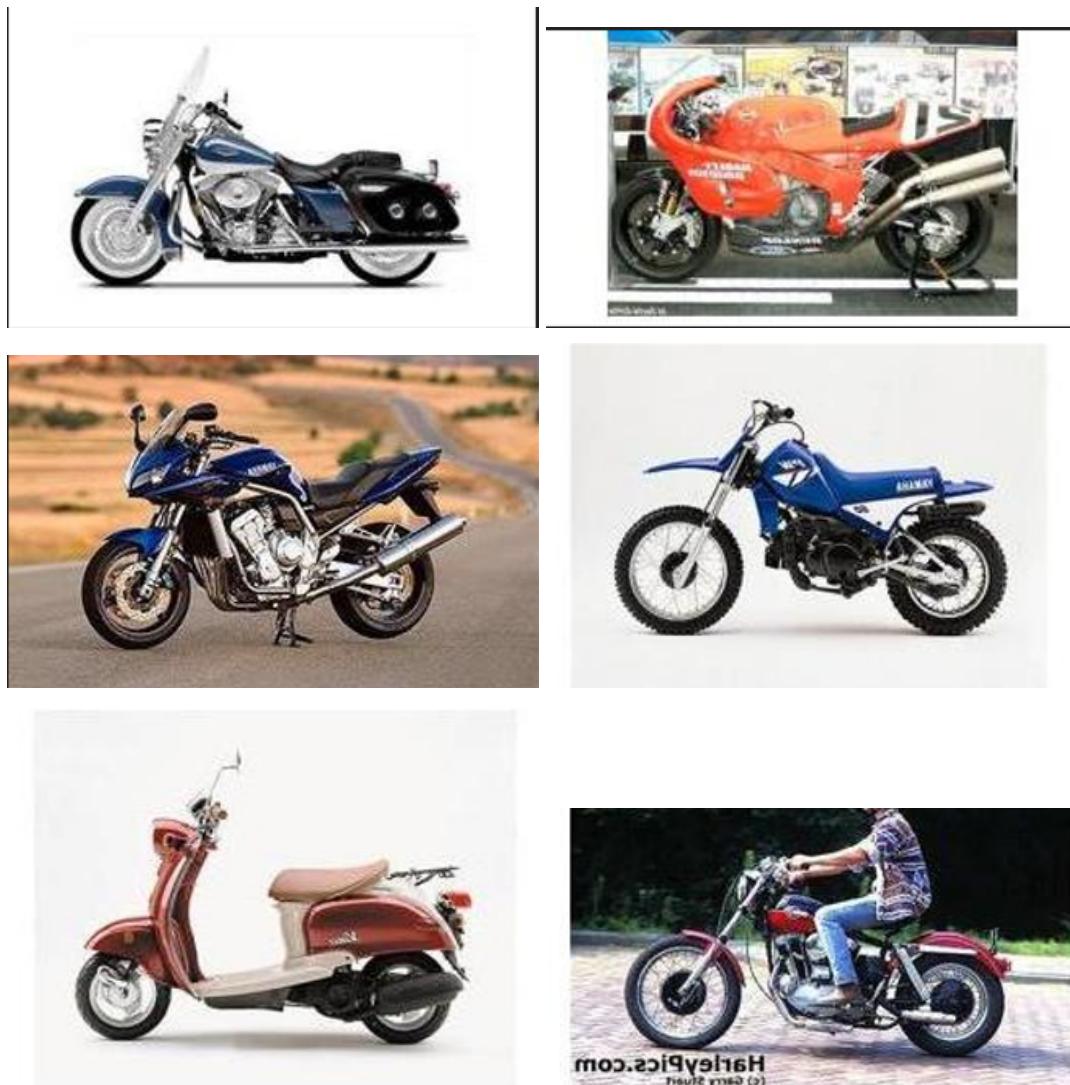
Παρουσίαση του σετ δεδομένων εκπαίδευσης

Το πρόβλημα προς επίλυση της παρούσας εργασίας ήταν η ταξινόμηση πολλαπλών κλάσεων (**multi-class classification**). Σκοπός του τελικού κώδικα ήταν να ταξινομήσει την κάθε εικόνα του σετ δοκιμής δεδομένων (**testing set**) σε μια από τις πέντε κλάσεις που υπήρχαν στα δεδομένα:

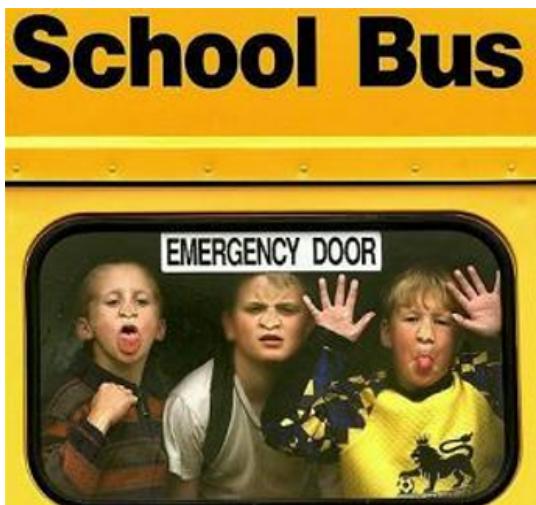
- motorbike
- school-bus
- touring-bike
- airplane
- car

Κάποιες εικόνες από το σετ εκπαίδευσης για τις παραπάνω κλάσεις φαίνονται παρακάτω:

Εικόνες από το σετ εκπαίδευσης της κλάσης motorbike:



Εικόνες από το σετ εκπαίδευσης της κλάσης school-bus:



Εικόνες από το σετ εκπαίδευσης της κλάσης touring-bike:



Εικόνες από το σετ εκπαίδευσης της κλάσης airplane:



Εικόνες από το σετ εκπαίδευσης της κλάσης car:



Σχόλια - Παρατηρήσεις:

Επιλέχθηκαν να παρουσιαστούν οι συγκεκριμένες εικόνες ως ένα δείγμα των εικόνων εκπαίδευσης για την κάθε κλάση, έτσι ώστε να φανεί η μεγάλη ποικιλομορφία που υπήρχε στα δεδομένα. Υπάρχουν δείγματα τα οποία φαίνονται να αποκλίνουν από λίγο έως πολύ στις διάφορες κλάσεις από τις ιδανικές εικόνες κλάσεων (πχ το επιθυμητό αντικείμενο της κλάσης, το οποίο να απεικονίζεται ολόκληρο χωρίς άλλα αντικείμενα γύρω του) και αυτό κάνει το πρόβλημα της ταξινόμησης πιο γενικό (λαμβάνει περισσότερους τύπους-είδη της κλάσης προς μελέτη υπόψη).

Παραγωγή οπτικού λεξικού βασισμένη στο μοντέλο Bag of Visual Words.

Εισαγωγή του μοντέλου BOVW:

Το μοντέλο Bag of Visual Words είναι ένα μοντέλο, το οποίο χρησιμοποιεί χαρακτηριστικά στοιχεία από το σετ δεδομένων (στην περίπτωση μας από εικόνες), έτσι ώστε να πάρει αποφάσεις με σκοπό την επίλυση προβλημάτων, όπως η ταξινόμηση (**classification**), η ανάκτηση εικόνων βάση περιεχομένου (**Content-Based Image Retrieval** ή **CBIR**). Στόχος του μοντέλου αυτού είναι να κωδικοποιήσει την κάθε εικόνα του σετ δεδομένων σε ένα διάνυσμα χαρακτηριστικών (**features**)

Αυτά τα χαρακτηριστικά στοιχεία μιας εικόνας που χρησιμοποιούνται, είναι σημεία ενδιαφέροντος, τα οποία είναι αντιπροσωπευτικά για την κάθε εικόνα. Τα σημεία ενδιαφέροντος αυτά ωστόσο δεν είναι τα σημεία που προκύπτουν από τους περιγραφείς SIFT, SURF, ή από άλλους αντίστοιχους περιγραφείς τοπικών χαρακτηριστικών (**local feature extractor**), διότι τα σημεία ενδιαφέροντος θα πρέπει να έχουν μια κοινή βάση αληθείας. Για τον λόγο αυτό τα σημεία ενδιαφέροντος πρέπει να έχουν προκύψει από έναν περιγραφέα καθολικών χαρακτηριστικών (**global feature extractor**).

Τα καθολικά αυτά χαρακτηριστικά δημιουργούνται για την κάθε εικόνα από το λεξικό (συνήθως χρησιμοποιείται με τους όρους **vocabulary** ή **dictionary**). Η δημιουργία του λεξικού αυτού είναι απαραίτητη, προκειμένου να δημιουργήσει τον κανόνα με βάση τον οποίο η κάθε εικόνα να μετατραπεί στο αντίστοιχο διάνυσμα καθολικών χαρακτηριστικών.

Δημιουργία οπτικού λεξικού:

Το οπτικό λεξικό είναι ένα σύνολο χαρακτηριστικών (**global features**), τα οποία εξάγονται από το σύνολο των χαρακτηριστικών (**local features**) όλων των εικόνων εκπαίδευσης. Στο λεξικό θα πρέπει να έχουν αντιστοιχιστεί με αντιπροσωπευτικό τρόπο, τα χαρακτηριστικά που έχουν υπολογιστεί στις εικόνες.

Η αντιστοίχιση των local features σε global features, πραγματοποιείται με την βοήθεια του αλγόριθμου K-Means. Ο αλγόριθμος K-Means, είναι ένας τύπος μάθησης χωρίς επίβλεψη (**unsupervised learning**), ο οποίος χρησιμοποιείται σε δεδομένα χωρίς ετικέτες (**unlabeled data**). Ο σκοπός του αλγορίθμου είναι να ομοδοποιήσει τα δεδομένα σε τόσες συστάδες (**clusters**), όσες ορίζει η παράμετρος **K** (στον κώδικα αναφέρεται ως `nump_of_words`). Η ομαδοποίηση των δεδομένων (στην περίπτωσή μας εικόνες) γίνεται με βάση την ομοιότητα των global features των εικόνων αυτών.

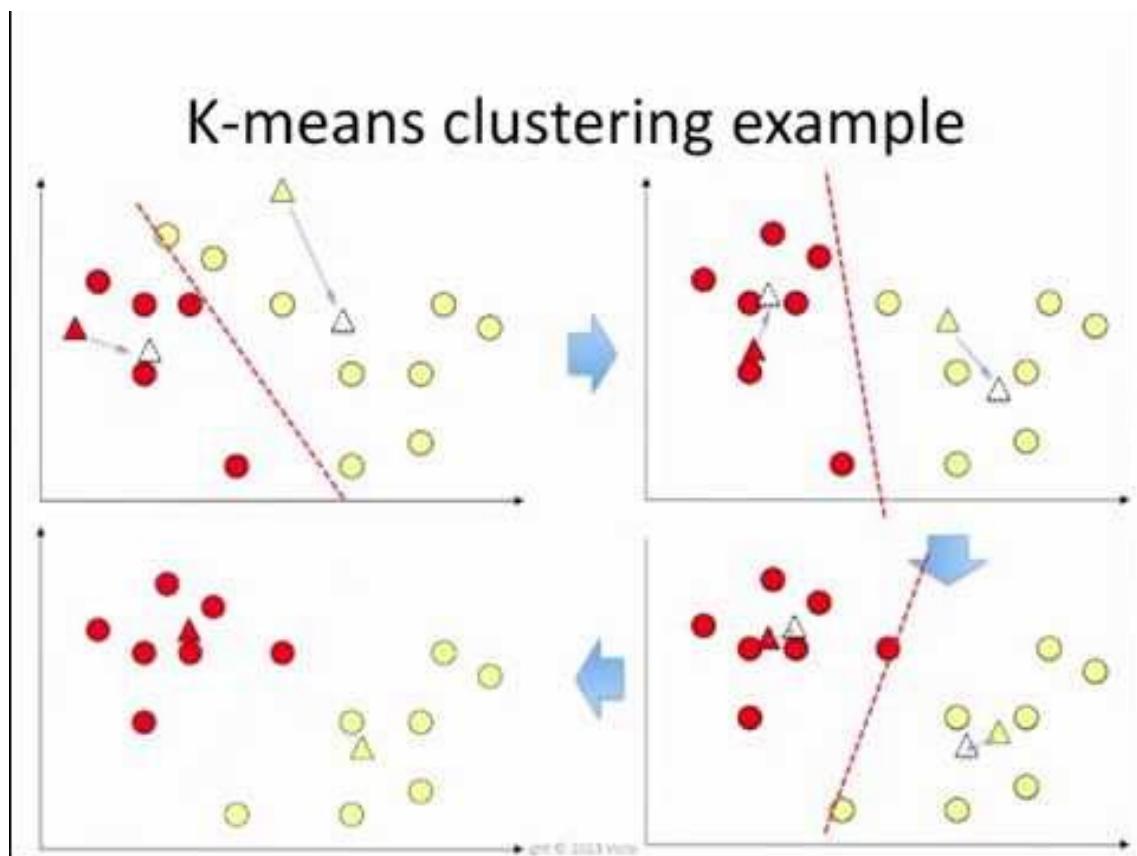
Η διαδικασία δημιουργίας του οπτικού λεξικού υλοποιείται από την συνάρτηση `create_vocabulary()` με την παρακάτω διαδικασία:

- Εξαγωγή τοπικών χαρακτηριστικών από την κάθε εικόνα του σετ εκπαίδευσης δεδομένων (**training set**). Έτσι προκύπτουν τα διανύσματα χαρακτηριστικών (**local features**) 128 διαστάσεων στην περίπτωση του SIFT.
- Εφαρμογή του αλγορίθμου K-Means:
 - Επιλογή της παραμέτρου του αριθμού των **global features** (πχ 50).
 - Διαχωρισμός όλων των **local features** που προέκυψαν από όλες τις εικόνες εκπαίδευσης σε τόσες συστάδες (**clusters**), όσες και ο αριθμός των global features που επιλέχθηκε για τον αλγόριθμο. Αυτό συμβαίνει σε έναν χώρο 128 διαστάσεων.
- Τέλος προκύπτουν οι συστάδες (πχ 50). Ο αριθμός αυτός αναφέρεται είτε με το όνομα συστάδες είτε με τον όρο λέξεις (**words**), είτε με τον όρο global features. Το λεξικό προκύπτει να είναι ένας πίνακας χαρακτηριστικών με διαστάσεις:
 - πλήθος γραμμών: ο αριθμός των συστάδων που επιλέχθηκε
 - πλήθος στηλών: 128 (το μήκος των local features που προέκυψαν από τον περιγραφέα SIFT)

Ο διαχωρισμός των local features σε συστάδες (αλγόριθμος K-Means) ακολουθεί τα παρακάτω βήματα:

1. Τυχαία τοποθέτηση των συστάδων στον 128-διάστατο χώρο.
2. Αντιστοίχιση όλων των local features στον πιο κοντινό cluster.
3. Νέα τοποθέτηση των συστάδων στον 128-διάστατο χώρο στο σημείο που προκύπτει από τον μέσο όρο των σημείων που έχουν αντιστοιχιστεί στον κάθε cluster στο προηγούμενο βήμα (Βήμα 2).
4. Επανάληψη των Βημάτων 2 και 3 έως ότου η μετατόπιση των συστάδων να μην είναι μεγαλύτερη από κάποιο κατώφλι (στον κώδικα χρησιμοποιείται με μορφή ποσοστού της τάξης του 10%)

Παρακάτω φαίνεται μια εικόνα η οποία περιέχει τα βήματα του αλγορίθμου K-Means σε διαδοχικές στιγμές:



Τελικά προκύπτει ένα σύνολο από συστάδες τα οποία έχουν συντεταγμένες τον μέσο όρο όλων των σημείων που περιέχονται στην αντίστοιχη ομάδα.

Εξαγωγή περιγραφέα σε κάθε εικόνα εκπαίδευσης ‘**imagedb**’ με βάση το μοντέλο BOVW.

Για την επίλυση του προβλήματος της ταξινόμησης, είναι απαραίτητη η αντιστοίχιση των local features σε global features στις εικόνες του σετ εκπαίδευσης. Για να πραγματοποιηθεί αυτό, είναι αναγκαία η συνεισφορά του λεξικού που υλοποιήθηκε στο προηγούμενο βήμα. Το λεξικό θα αντιστοιχίσει τα τοπικά χαρακτηριστικά της κάθε εικόνας σε ένα ιστόγραμμα. Το ιστόγραμμα αυτό θα δείχνει την κατανομή των χαρακτηριστικών στοιχείων που προέκυψαν ως συστάδες από τον αλγόριθμο K-Means.

Με το πέρας αυτής της αντιστοίχισης, τα δεδομένα εκπαίδευσης από εικόνες έχουν μετατραπεί σε διανύσματα χαρακτηριστικών με διαστάσεις:

- πλήθος γραμμών: 1
- πλήθος στηλών: ο αριθμός των συστάδων που επιλέχθηκε στον αλγόριθμο K-Means.

Το διάνυσμα αυτό των global features ονομάζεται επίσης και λέξη (**word**).

Η παραπάνω αντιστοίχιση υλοποιείται με την βοήθεια της συνάρτησης *create_index()*, στην οποία επιπλέον δημιουργούνται και δύο ακόμη μεταβλητές:

- **image_paths**: Μια λίστα από strings όπου περιέχει τα paths όλων των εικόνων του σετ εκπαίδευσης.
- **image_paths_test**: Μια λίστα από strings όπου περιέχει τα paths όλων των εικόνων του σετ δοκιμής.

Η αντιστοίχιση των εικόνων **εκπαίδευσης** σε global features πραγματοποιείται με την ακόλουθη διαδικασία:

1. Εξαγωγή των local features των εικόνων με την βοήθεια του περιγραφέα SIFT.
2. Κωδικοποίηση των local features σε global features με την βοήθεια της συνάρτησης *encode_bovw_descriptor()*
3. Τέλος όλα τα global features των εικόνων εκπαίδευσης αποθηκεύονται σε έναν πίνακα για μια καλύτερη διαχείριση και επεξεργασία.

Η συνάρτηση *encode_bovw_descriptor()*, όπως αναφέρθηκε, μετατρέπει ένα local feature σε ένα global feature. Αυτό πραγματοποιείται ως ακολούθως :

- Δημιουργία του ιστογράμματος των λέξεων του λεξικού.
- Επανάληψη για κάθε ένα keypoint του local feature της εικόνας δοκιμής:
 - Υπολογισμός της απόστασης του keypoint αυτού από όλες τις λέξεις του λεξικού (**matching**).
 - Εύρεση της λέξης που έχει την μικρότερη απόσταση (το καλύτερο matching) από το εν λόγω keypoint.
 - Ανανέωση του ιστογράμματος, αυξάνοντας την τιμή του πλήθους της λέξης με την ελάχιστη απόσταση
- Κανονικοποίηση των τιμών του ιστογράμματος, διαιρώντας όλες τις τιμές του με το άθροισμα των keypoints της εικόνας δοκιμής.

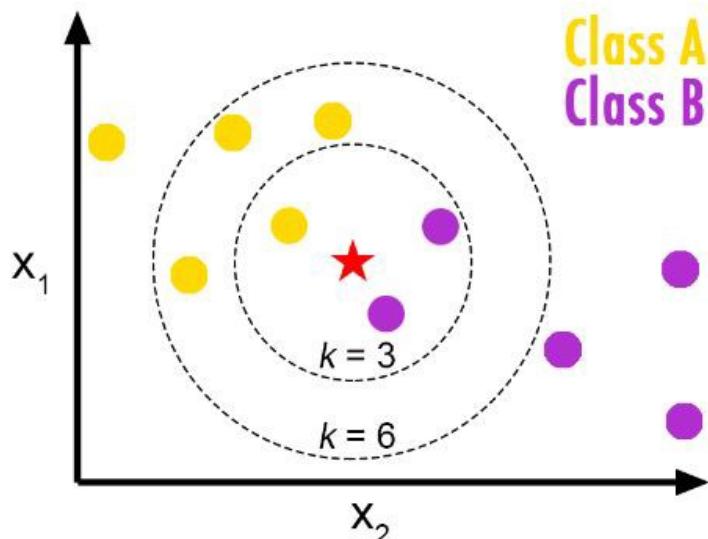
Ταξινόμηση μιας εικόνας κάνοντας χρήση του αλγορίθμου k-NN.

Ο αλγόριθμος k Nearest Neighbours (kNN), είναι ένας αλγόριθμος εποπτευόμενης μηχανικής μάθησης (**supervised machine learning**), ο οποίος λύνει προβλήματα ταξινόμησης (**classification**) και παλινδρόμησης (**regression**).

Ο αλγόριθμος του kNN, αποφασίζει για την κλάση μιας εικόνας δοκιμής, λαμβάνοντας υπόψη την κατανομή των **k** πιο κοντινών γειτόνων του global feature της εικόνας αυτής, στον χώρο των global features του οπτικού λεξικού (πχ αν έχουμε 50 λέξεις, τότε η απόφαση λαμβάνεται στον 50-διαστατο χώρο) σύμφωνα με την ευκλείδεια απόσταση.

Στην παρακάτω εικόνα φαίνεται η διαδικασία ταξινόμησης ενός σημείου δοκιμής για τις περιπτώσεις:

- $k = 3$, ταξινομείται στην κλάση B
- $k = 6$, ταξινομείται στην κλάση A



Στον κώδικα, για την υλοποίηση του KNN αλγορίθμου, χρησιμοποιήθηκε η συνάρτηση `knn_classifier()`, της οποίας ορίσματα ήταν:

- `q_bow_desc`: Το διάνυσμα χαρακτηριστικών της εικόνας δοκιμής
- `bow_descs`: Ο πίνακας διανυσμάτων χαρακτηριστικών των εικόνων εκπαίδευσης
- `img_paths`: Η λίστα με τα paths των εικόνων εκπαίδευσης
- `k`: Η παράμετρος του πλήθους των γειτόνων

Στον κώδικα η κάθε εικόνα δοκιμής που περνάει από τον αλγόριθμο, σχετίζεται με τις λέξεις του λεξικού και κατα πόσο αυτή μοιάζει με αυτές. Προκειμένου να γίνει αυτό δημιουργείται ένας ειδος ιστογράμματος, το οποίο έχει τόσες θέσεις όσες και οι κλάσεις, δηλαδή πέντε (στον κώδικα το ιστόγραμμα αυτό είναι στην μεταβλητή `neighbours`). Στο τέλος του kNN αλγορίθμου, η κάθε εικόνα δοκιμής προκύπτει να έχει ένα ιστόγραμμα γειτόνων, όπου σε

κάθε θέση του έχει το πλήθος των γειτόνων της αντίστοιχης κλάσης. Έτσι τελικά μπορεί να προκύψει απόφαση για την κλάση στην οποία θα ταξινομηθεί αυτή η εικόνα δοκιμής.

Τα βήματα του αλγορίθμου kNN που υλοποιήθηκαν ήταν:

- Δημιουργία μιας λίστας για την αποθήκευση των ευκλείδειων αποστάσεων του global feature δοκιμής από όλα τα global feature του λεξικού.
- Υπολογισμός της ευκλείδειας αυτής απόστασης
- Επανάληψη για k φορές:
 - Εύρεση ελάχιστης απόστασης, δηλαδή κοντινότερου γείτονα
 - Εύρεση της κλάσης του κοντινότερου γείτονα
 - Αύξηση του πλήθους της αντίστοιχης κλάσης, δηλαδή μεταβολή του ιστογράμματος των γειτόνων της εικόνας δοκιμής
 - Αντικατάσταση της ελάχιστης απόστασης με έναν πολύ μεγάλο αριθμό (\inf), έτσι ώστε να μην βρεθεί στην επόμενη επανάληψη ως η ίδια ελάχιστη τιμή
- Μετά από την συμπλήρωση του ιστογράμματος γειτόνων, υπολογίζεται η κλάση που προβλέπεται από την θέση που έχει τους περισσότερους γείτονες

Η συνάρτηση `knn_classifier()` επιστρέφει το ιστόγραμμα `neighbours` και την κλάση πρόβλεψης `class_pred`.

Παρατήρηση:

Ο κώδικας προβλέπει την κλάση της εικόνας δοκιμής, υπολογίζοντας την θέση μεγίστου στην κατανομή γειτόνων. Υπάρχει περίπτωση ωστόσο το μέγιστο να υπάρχει σε περισσότερες από μια θέση του ιστογράμματος. Μια πιθανή βελτίωση θα μπορούσε να είναι να υπολογίζει ο κώδικας τις αποστάσεις όλων των γειτόνων που αντιστοιχούν στις κλάσεις που βρέθηκε η μέγιστη κατανομή ιστογράμματος, και να επιλέγει την κλάση που είχε την μικρότερη κατα μέσο όρο απόσταση των γειτόνων από το σημείο δοκιμής.

Για την δοκιμή του αλγορίθμου kNN, ακολουθήθηκαν τα παρακάτω βήματα (στο κύριο μέρος του κώδικα και όχι σε συνάρτηση αυτή την φορά):

- Επανάληψη για κάθε εικόνα δοκιμής:
 - Εξαγωγή της αναμενόμενης κλάσης της εικόνας (πραγματική κλάση)
 - Εξαγωγή local features της εικόνας
 - Εξαγωγή global features της εικόνας
 - Πρόβλεψη της κλάσης από την συνάρτηση `knn_classifier()`.

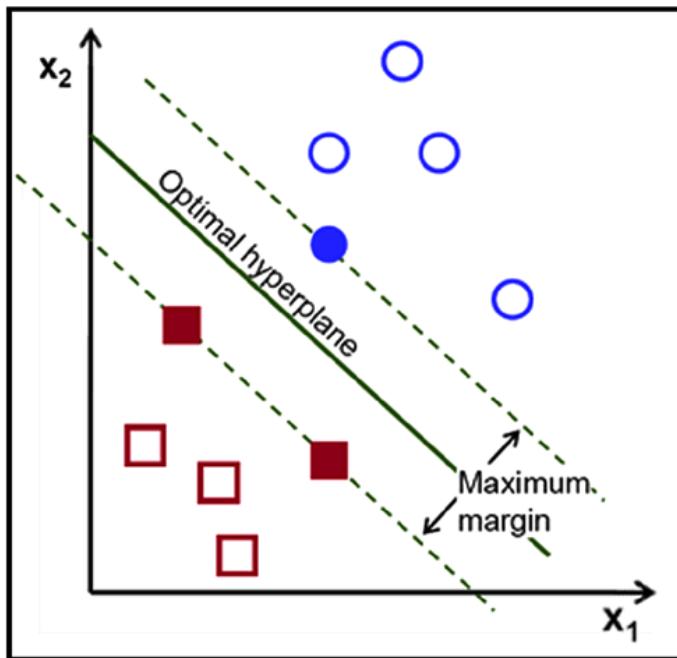
Επειδή το σετ δεδομένων δοκιμής “image_db_test”, αποτελούνταν από εικόνες οι οποίες ήταν γνωστό σε ποιες κλάσεις ανήκαν, ο αλγόριθμος δοκιμής kNN, εκτός από τις κλάσεις πρόβλεψης, επέστρεφε και τις πραγματικές κλάσεις των εικόνων δοκιμής.

Για κάθε εικόνα δοκιμής, τόσο η μεταβλητή που υπολογίζει την πραγματική κλάση της εικόνας, όσο και η μεταβλητή που υπολογίζει την προβλεπόμενη κλάση, και οι δύο αποθηκεύοντουσαν σε μια λίστα ξεχωριστά, όπου κρατούσε τις τιμές αυτές. Ο λόγος για αυτές τις λίστες ήταν η μαζική αξιολόγηση της ακρίβειας του kNN ταξινομητή σε όλο το εύρος των εικόνων δοκιμής κάθε φορά.

Ταξινόμηση μιας εικόνας κάνοντας χρήση του σχήματος one-versus-all όπου για κάθε κλάση εκπαιδεύεται ένας SVM ταξινομητής.

Ο αλγόριθμος Support Vector Machine (SVM), είναι ένας αλγόριθμος εποπτευόμενης μηχανικής μάθησης (**supervised machine learning**), ο οποίος λύνει προβλήματα ταξινόμησης (**classification**) και παλινδρόμησης (**regression**). Στην παρούσα εργασία λύνει το πρόβλημα της ταξινόμησης. Επειδή ο SVM μπορεί να λύσει μόνο προβλήματα δυαδικής ταξινόμησης και επειδή το συγκεκριμένο πρόβλημα προς επίλυση είναι η ταξινόμηση πολλαπλών κλάσεων, χρησιμοποιείται η μέθοδος **one-versus-all**.

Ο αλγόριθμος SVM προϋποθέτει το στάδιο της εκπαίδευσης (training). Κατά την εκπαίδευση ο ταξινομητής υπολογίζει ένα υπερεπίπεδο που διαχωρίζει τις δύο κλάσεις του προβλήματος δυαδικής ταξινόμησης. Στο μοντέλο BOVW προσπαθεί να διαχωρίσει τις λέξεις που ανήκουν στις δύο διαφορετικές κλάσεις. Οι λέξεις αυτές αφού έχουν τόσα στοιχεία, όσα και ο αριθμός των λέξεων που επιλέχθηκε για την κατασκευή του οπτικού λεξικού (δηλαδή ο αριθμός των global features ή των οπτικών λέξεων) βρίσκονται σε έναν χώρο που έχει τόσες διαστασεις (όσες και ο αριθμός των global features). Το υπερ-επίπεδο που υπολογίζεται έχει τόσες διαστάσεις. Το υπερ-επίπεδο αυτό χωρίζει με τέτοιο τρόπο τις λέξεις των δύο κλάσεων, έτσι ώστε η απόσταση των ακριανών λέξεων (αυτές που είναι κοντά στο υπερ-επίπεδο) να είναι η μέγιστη δυνατή. Αυτό μπορεί να φανεί εύκολα και στην παρακάτω εικόνα:



Μετά το στάδιο της εκπαίδευσης, προκειμένου να ταξινομήσει ο ταξινομητής τις εικόνες δοκιμής, έπειται το στάδιο της δοκιμής (testing). Στο στάδιο αυτό το εκπαιδευμένο SVM υπολογίζει εάν η εικόνα δοκιμής βρίσκεται από την μία πλευρά του υπερ-επιπέδου ή από την άλλη για να μπορέσει να προβλέψει την κλάση της. Επίσης είναι δυνατός ο υπολογισμός της απόστασης της λέξης της εικόνας δοκιμής στον πολυδιάστατο αυτό χώρο από το υπερ-επίπεδο διαχωρισμού, δίνοντας έτσι μια παραπάνω πληροφορία για το πόσο σίγουρη

ήταν η πρόβλεψη του μοντέλου SVM για την κλάση της εικόνας δοκιμής. Εάν η απόσταση αυτή είναι μεγάλη τότε υπάρχει μεγαλύτερη σιγουριά για το αποτέλεσμα του μοντέλου. Αυτό δεν συμβαίνει όταν η απόσταση αυτή είναι μικρή. Μια λέξη η οποία βρίσκεται κοντά στο υπερ-επίπεδο διαχωρισμού, δεν δίνει την εγγύηση ότι ανήκει στην εν λόγω κλάση καθώς απέχει κατα λίγο από την άλλη κλάση.

Στην συγκεκριμένη τεχνική δημιουργούνται τόσοι ταξινομητές SVM, όσες είναι και οι διαφορετικές κλάσεις του προβλήματος. Στην περίπτωσή μας υπάρχουν πέντε κλάσεις (motorbike, school-bus, touring-bike, airplane, car), άρα δημιουργούνται και πέντε ταξινομητές SVM. Ο κάθε ένας ταξινομητής λύνει το πρόβλημα δυαδικής ταξινόμησης μιας εικόνας, δηλαδή προβλέπει εάν η εικόνα ανήκει στην αντίστοιχη κλάση του ταξινομητή, είτε στην κλάση που περιέχει όλες τις άλλες κλάσεις του προβλήματος. Για παράδειγμα ο ταξινομητής SVM για την κλάση motorbike, προβλέπει εάν η εικόνα δοκιμής ανήκει στην μια κλάση με τις μοτοσυκλέτες είτε στην άλλη κλάση (επειδή ο SVM ταξινομητής έχει εκπαιδευτεί στο σετ δεδομένων των πέντε κλάσεων, σε αυτήν την κλάση βρίσκονται όλες οι υπόλοιπες κλάσεις: school-bus, touring-bike, airplane, car).

Από αυτούς τους πέντε ταξινομητές προκύπτει για κάθε μια εικόνα δοκιμής, η απόσταση της λέξης της εν λόγω εικόνας από το υπερ-επίπεδο διαχωρισμού. Από αυτές τις πέντε αποστάσεις λαμβάνεται η απόφαση της πρόβλεψης της κλάσης για την εικόνα δοκιμής. Επιλέγεται η κλάση που το αντίστοιχο SVM μοντέλο είχε την μεγαλύτερη απόσταση στην μεριά της κλάσης από το υπερεπίπεδο διαχωρισμού για το κάθε πρόβλημα δυαδικής ταξινόμησης.

Συγκεκριμένα στον κώδικα, για την εκπαίδευση του SVM μοντέλου, χρησιμοποιήθηκε η συνάρτηση *svm_classifier_training()*, της οποίας ορίσματα ήταν:

- *img_paths*: Η λίστα με τα paths των εικόνων εκπαίδευσης
- *svm_kernel_type*: Η παράμετρος του τύπου του πυρήνα του SVM

Επειδή, ο αλγόριθμος SVM είναι ένας αλγόριθμος εποπτευόμενης μάθησης, όπως αναφέρθηκε παραπάνω, κατα την εκπαίδευση του μοντέλου χρησιμοποιείται και η πληροφορία του σε ποια κλάση βρίσκονται οι εικόνες. Η πληροφορία αυτή υπάρχει στην μεταβλητή *img_paths*, η οποία περιέχει όλα τα paths των εικόνων εκπαίδευσης. Στην συνέχεια η πληροφορία των κλάσεων που αρχικά υπήρχε με την μορφή string, μετατρέπεται σε δυαδική πληροφορίες με την βοήθεια νέας μεταβλητής *px labels* motorbike για την κλάση motorbike. Η μεταβλητή αυτή αποτελείται από τιμές 0 και 1 που κωδικοποιούν την μη ύπαρξη ή την ύπαρξη μιας εικόνας μοτοσυκλέτας αντίστοιχα στην εν λόγω εικόνα εκπαίδευσης της λίστας motorbike.

Αντίστοιχα εκπαιδεύονται και οι υπόλοιποι τέσσερις ταξινομητές, χρησιμοποιώντας και τα αντίστοιχα labels:

- *svm_schoolbus* για την ταξινόμηση εικόνων δοκιμής αν σε αυτές απεικονίζεται σχολικό λεωφορείο ή όχι
- *svm_bike* για την ταξινόμηση εικόνων δοκιμής αν σε αυτές απεικονίζεται ποδήλατο ή όχι
- *svm_airplane* για την ταξινόμηση εικόνων δοκιμής αν σε αυτές απεικονίζεται αεροπλάνο ή όχι

- `svm_car` για την ταξινόμηση εικόνων δοκιμής αν σε αυτές απεικονίζεται αυτοκίνητο ή όχι

Τέλος οι πέντε παραπάνω ταξινομητές SVM, αποθηκεύονται και επιστρέφονται από την συνάρτηση `svm_classifier_training()` και είναι έτοιμοι για την δοκιμή και την αξιολόγηση του σε δεδομένων δοκιμής.

Μετά από την εκπαίδευση, έπειται το βήμα της δοκιμής. Συγκεκριμένα στον κώδικα, για την δοκιμή του SVM μοντέλου, χρησιμοποιήθηκε η συνάρτηση `svm_classifier_testing()`, της οποίας ορίσματα ήταν:

- `img_paths_test`: Η λίστα με τα paths των εικόνων δοκιμής
- Και όλα τα μοντέλα των ταξινομητών SVM:
 - `svm_motorbike`
 - `svm_schoolbus`
 - `svm_bike`
 - `svm_airplane`
 - `svm_car`

Επειδή το σετ δεδομένων δοκιμής “`image_db_test`”, αποτελούνταν από εικόνες οι οποίες ήταν γνωστό σε ποιες κλάσεις ανήκαν, ο αλγόριθμος δοκιμής SVM, εκτός από τις κλάσεις πρόβλεψης, επέστρεφε και τις πραγματικές κλάσεις των εικόνων δοκιμής.

Τα βήματα της συνάρτησης `svm_classifier_training()`, ήταν:

- Επανάληψη για κάθε εικόνα δοκιμής:
 - Εξαγωγή της αναμενόμενης κλάσης της εικόνας (πραγματική κλάση)
 - Εξαγωγή local features της εικόνας
 - Εξαγωγή global features της εικόνας
 - Υπολογισμός απόστασης από το υπερ-επίπεδο από όλους τους εκπαιδευμένους ταξινομητές
 - Εύρεση της ελάχιστης απόστασης από το υπερ-επίπεδο
 - Πρόβλεψη της κλάσης

Για κάθε εικόνα δοκιμής, τόσο η μεταβλητή που υπολόγιζε την πραγματική κλάση της εικόνας, όσο και η μεταβλητή που υπολόγιζε την προβλεπόμενη κλάση, και οι δύο αποθηκεόντουσαν σε μια λίστα ξεχωριστά, όπου κρατούσε τις τιμές αυτές. Ο λόγος για αυτές τις λίστες ήταν η μαζική αξιολόγηση της ακρίβειας του SVM ταξινομητή σε όλο το εύρος των εικόνων δοκιμής κάθε φορά.

Αξιολόγηση του συστήματος.

Εισαγωγή:

Για την αξιολόγηση του συστήματος χρησιμοποιήθηκε το σύνολο δεδομένων δοκιμής “`imagedb_test`”. Για την αξιολόγηση του συστήματος μετρηθήκε η ακρίβεια (**accuracy**), η οποία ορίζεται ως το ποσοστό των επιτυχών ταξινομήσεων, δηλαδή το άθροισμα των σωστών προβλέψεων δια το σύνολο όλων των δοκιμών.

Η αξιολόγηση εφαρμόστηκε τόσο στον ταξινομητή kNN όσο και στον ταξινομητή SVM. Επιπλέον να σημειωθεί ότι στους δύο ταξινομητές μετρήθηκε τόσο η ακρίβεια του συνολικού σετ δεδομένων δοκιμής (το οποίο περιελάμβανε όλες τις κλάσεις του προβλήματος), όσο και οι επιμέρους ακρίβειες των επιμέρους σετ δεδομένων δοκιμής.

Κατά την αξιολόγηση του συστήματος όπως αυτή αναφέρθηκε παραπάνω, ελέγχθηκε η επίδραση κάποιων παραμέτρων στην ακρίβεια. Οι παράμετροι αυτοί ήταν:

- Το πλήθος των οπτικών λέξεων, με το οποίο κατασκευάζεται το οπτικό λεξικό.
- Το πλήθος των πλησιέστερων γειτόνων, το οποίο επιλέγεται στον αλγόριθμο του ταξινομητή kNN.
- Ο τύπος του πυρήνα (**kernel**), ο οποίος επιλέγεται στον αλγόριθμο του ταξινομητή SVM.

Επίσης οι επιθυμητές ακρίβειες του συστήματος υπολογίστηκαν με την βοήθεια της συνάρτησης `compute_accuracy()`, τα ορίσματα της οποίας ήταν:

- `class_pred_list`: Μια λίστα που περιείχε τις κλάσεις που προέβλεπε κάθε φορά ο κάθε ταξινομητής και περιείχε όλες τα αποτελέσματα για τις εικόνες ολόκληρου του σετ δοκιμής.
- `class_exp_list`: Μια λίστα που περιείχε τις κλάσεις όλων των εικόνων ολόκληρου του σετ δοκιμής.

Τα βήματα της συνάρτησης ήταν:

- Ορισμός νέων λιστών για την κάθε ζητούμενη ακρίβεια
- Επανάληψη για κάθε ζητούμενη ακρίβεια:
 - Επανάληψη για κάθε εικόνα της νέας λίστας της συγκεκριμένης ζητούμενης ακρίβειας.
 - Έλεγχος αν η κλάση πρόβλεψης είναι η ίδια με την πραγματική κλάση της εικόνας δοκιμής.
 - Αύξηση ενός ειδικού για την συγκεκριμένη ζητούμενη ακρίβεια μετρητή σε περίπτωση σωστής πρόβλεψης, ο οποίος κρατούσε το άθροισμα των σωστών προβλέψεων της περίπτωσης αυτής.
 - Υπολογισμός της ακρίβειας ως το κλάσμα του αθροίσματος των σωστών προβλέψεων προς το πλήθος των εικόνων δοκιμής της συγκεκριμένης ζητούμενης ακρίβειας.

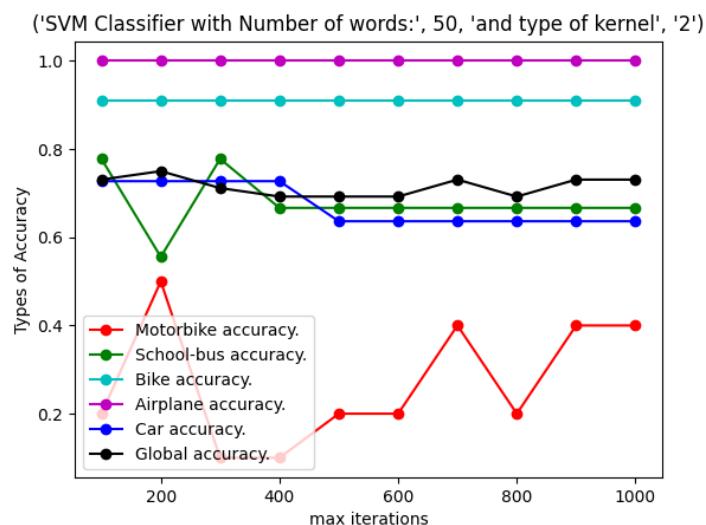
Στις επόμενες ενότητες θα παρουσιαστούν τα αποτελέσματα των ταξινομητών ποιοτικά και ποσοτικά. Στην ποσοτική προσέγγιση, θα παρουσιαστούν κανονικοποιημένα στο (0,1).

Ποιοτικά αποτελέσματα αξιολόγησης:

Παράμετρος του μέγιστου πλήθους επαναλήψεων:

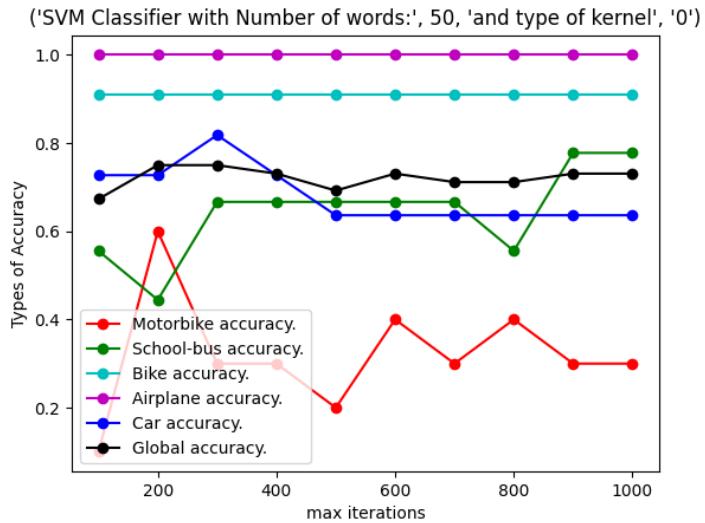
Μια επιπλέον παράμετρος για την οποία εξετάστηκε η επίδρασή της στην ακρίβεια του συστήματος, ήταν ένα από τα κριτήρια τερματισμού του αλγορίθμου SVM, και συγκεκριμένα το πλήθος των μέγιστων επαναλήψεων. Ο έλεγχος της επίδρασης της παραμέτρου αυτής στην ακρίβεια δεν ήταν ζητούμενος, ωστόσο πραγματοποιήθηκε αποσκοπώντας την βελτιστηρίση της ακρίβειας του αλγορίθμου. Με άλλα λόγια ελέγχθηκε η παράμετρος αυτή, για την βελτιστοποίηση της ακρίβειας του αλγορίθμου, έτσι ώστε ο έλεγχος για τις υπόλοιπες παραμέτρους που αναφέρθηκαν παραπάνω να ξεκινήσει από μία ήδη βελτιστοποιημένη ακρίβεια. Ο κώδικας των δοκιμών αυτής της παραμέτρου δεν αποθηκεύτηκε και άρα δεν υπάρχει στο τελικό αρχείο του κώδικα. Ωστόσο παρακάτω φαίνεται το διάγραμμα των αξιολογήσεων του συστήματος συναρτήσει της παραμέτρου αυτής. Ο έλεγχος της παραμέτρου πραγματοποιήθηκε τρεις φορές, μια για κάθε ένα τύπο πυρήνα του ταξινομητη SVM. Να σημειωθεί επίσης ότι οι δοκιμές πραγματοποιήθηκαν για συγκεκριμένο αριθμό οπτικών λέξεων και ίσο με 50.

Διάγραμμα αξιολόγησης ταξινομητή SVM συναρτήσει της παραμέτρου των μέγιστων επαναλήψεων για την περίπτωση του τύπου πυρήνα RBF:



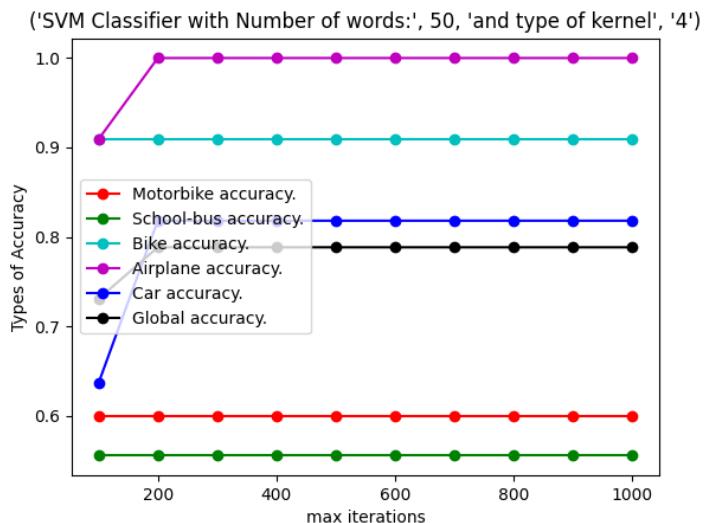
Από το παραπάνω διάγραμμα φαίνεται ότι σχεδόν για όλες τις ακρίβειες του συστήματος (motorbike_acc, bike_acc, airplane_acc, car_acc, global acc), η επιλογή της τιμής 200 για την παράμετρο των μέγιστων επαναλήψεων δίνει την μέγιστη απόδοση.

Διάγραμμα αξιολόγησης ταξινομητή SVM συναρτήσει της παραμέτρου των μέγιστων επαναλήψεων για την περίπτωση του τύπου πυρήνα LINEAR:



Από το παραπάνω διάγραμμα φαίνεται ότι σχεδόν για όλες τις ακρίβειες του συστήματος (motorbike_acc, bike_acc, airplane_acc, global acc), η επιλογή της τιμής 200 για την παράμετρο των μέγιστων επαναλήψεων δίνει την μέγιστη απόδοση.

Διάγραμμα αξιολόγησης ταξινομητή SVM συναρτήσει της παραμέτρου των μέγιστων επαναλήψεων για την περίπτωση του τύπου πυρήνα CHI2:



Από το παραπάνω διάγραμμα φαίνεται ότι για όλες τις ακρίβειες του συστήματος (motorbike_acc, bike_acc, airplane_acc, car_acc, global acc), η επιλογή των τιμών 200, 300, 400, 500, 600, 700, 800, 900, 1000 για την παράμετρο των μέγιστων επαναλήψεων δίνει την μέγιστη απόδοση.

Συμπερασματικά για όλους του τύπους πυρήνα που ελέγχθηκαν, η επιλογή της τιμής **200** για την παράμετρο των μέγιστων επαναλήψεων αποφασίστηκε ότι δίνει την μέγιστη απόδοση του ταξινομητή SVM.

Παράμετρος του πλήθους των οπτικών λέξεων:

Ο έλεγχος της επίδρασης του πλήθος των οπτικών λέξεων στην ακρίβεια του συστήματος πραγματοποιήθηκε σε συνδυασμό με τις άλλες δύο παραμέτρους, δηλαδή τον αριθμό των πλησιέστερων γειτόνων στον ταξινομητή kNN και τον τύπο του πυρηνα του SVM. Ο λόγος για αυτό είναι ότι ο αριθμός των οπτικών λέξεων επηρεάζει άμεσα το οπτικό λεξικό, το οποίο αυτό με την σειρά του επηρεάζει τους περιγραφείς των εικόνων με βάση το μοντέλο BOVW, οι οποίοι στην συνέχεια επηρεάζουν τους δύο ταξινομητες. Η ακρίβειας του συστήματος συναρτήσει της παραμέτρου του πλήθους των οπτικών λέξεων ελέγχθηκε για τις τιμές που αναγράφονται παρακάτω:

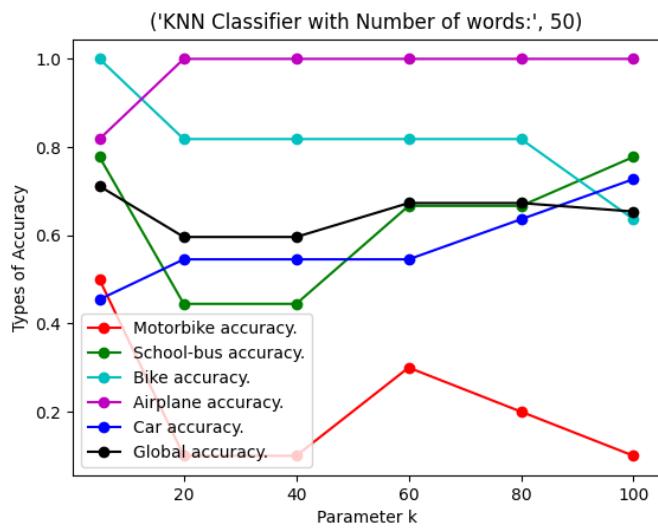
- 50
- 100
- 150

Παράμετρος του πλήθους των πλησιέστερων γειτόνων του ταξινομητή kNN:

Όπως αναφέρθηκε και παραπάνω, η επίδραση της παραμέτρου των πλησιέστερων γειτόνων του ταξινομητή kNN, ελέγχθηκε συναρτήσει του πλήθους των οπτικών λέξεων. Ανεξάρτητα από αυτήν την επιλογή, οι τιμές του πλήθους των πλησιέστερων γειτόνων για τον ταξινομητή kNN, αναγράφονται παρακάτω:

- 5
- 20
- 40
- 60
- 80
- 100

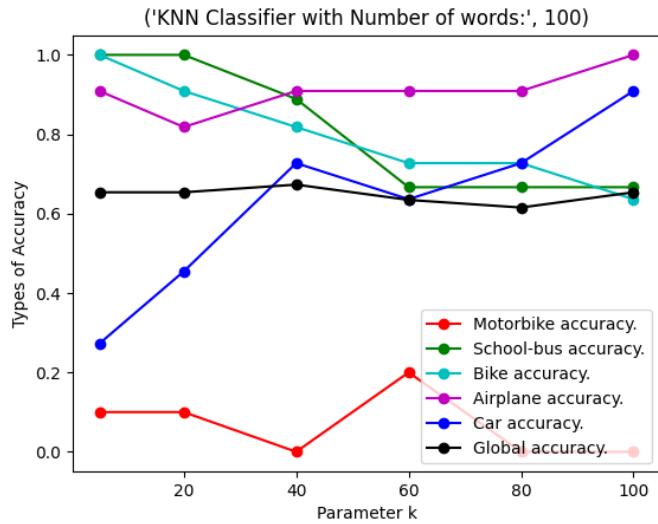
Στο παρακάτω διάγραμμα φαίνονται οι επιθυμητές ακρίβειες του συστήματος συναρτήσει των τιμών του πλήθους των πλησιέστερων γειτόνων για την περίπτωση των **50 οπτικών λέξεων**:



Από το παραπάνω διάγραμμα προκύπτουν τα εξης συμπερασματα:

- Η ακρίβεια της κλάσης motorbike παραμένει χαμηλή σχεδόν ανεξάρτητα από τις τιμές της παραμέτρου **k**.
- Η ακρίβεια της κλάσης school-bus τόσο για μικρές τιμές της παραμέτρου **k** (δηλαδή $k = 5$) όσο και για μεγάλες τιμές ($k = 60, 80, 100$) παραμένει σχετικά υψηλά.
- Η ακρίβεια της κλάσης touring-bike για μικρές τιμές της παραμέτρου **k** (δηλαδή $k = 5$) παραμένει πολύ υψηλή. Με την αύξηση του **k**, η απόδοση του ταξινομητη πεφτει.
- Η ακρίβεια της κλάσης airplane σχεδόν για όλες τις τιμές της παραμέτρου **k** παραμένει πολύ υψηλή.
- Η ακρίβεια της κλάσης car αυξάνεται με την αύξηση της παραμέτρου **k**.
- Η καθολική ακρίβεια φαίνεται να έχει την μέγιστη τιμή της στο για την τιμή του $k = 5$.

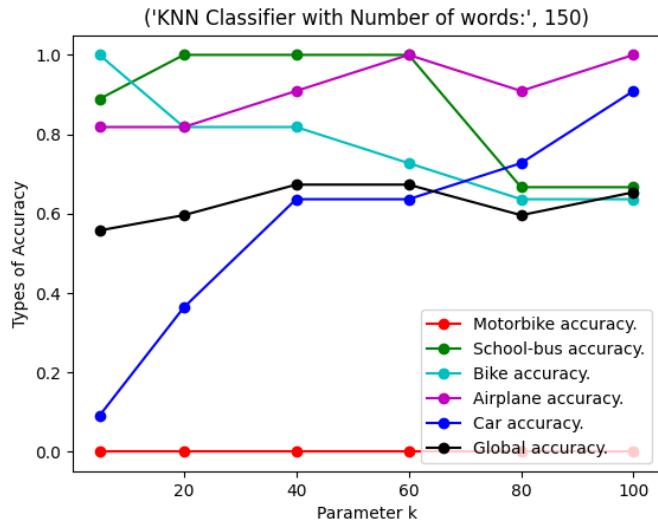
Στο παρακάτω διάγραμμα φαίνονται οι επιθυμητές ακρίβειες του συστήματος συναρτήσει των τιμών του πλήθους των πλησιέστερων γειτόνων για την περίπτωση των **100 οπτικών λέξεων**:



Από το παραπάνω διάγραμμα προκύπτουν τα εξης συμπερασματα:

- Η ακρίβεια της κλάσης motorbike παραμένει χαμηλή σχεδόν ανεξάρτητα από τις τιμές της παραμέτρου **k**.
- Η ακρίβεια της κλάσης school-bus μειώνεται με την αύξηση της παραμέτρου **k**.
- Η ακρίβεια της κλάσης touring-bike μειώνεται με την αύξηση της παραμέτρου **k**.
- Η ακρίβεια της κλάσης airplane σχεδόν για όλες τις τιμές της παραμέτρου **k** παραμένει πολύ υψηλή. Επίσης αυξάνεται με την αύξηση αύξηση της παραμέτρου **k**.
- Η ακρίβεια της κλάσης car αυξάνεται με την αύξηση της παραμέτρου **k**.
- Η καθολική ακρίβεια φαίνεται να παραμένει σταθερή ανεξάρτητα από τις τιμές της παραμέτρου **k**, με την μέγιστη τιμή αυτής στο για την τιμή του **k = 40**

Στο παρακάτω διάγραμμα φαίνονται οι επιθυμητές ακρίβειες του συστήματος συναρτήσει των τιμών του πλήθους των πλησιέστερων γειτόνων για την περίπτωση των **150 οπτικών λέξεων**:



Από το παραπάνω διάγραμμα προκύπτουν τα εξης συμπερασματα:

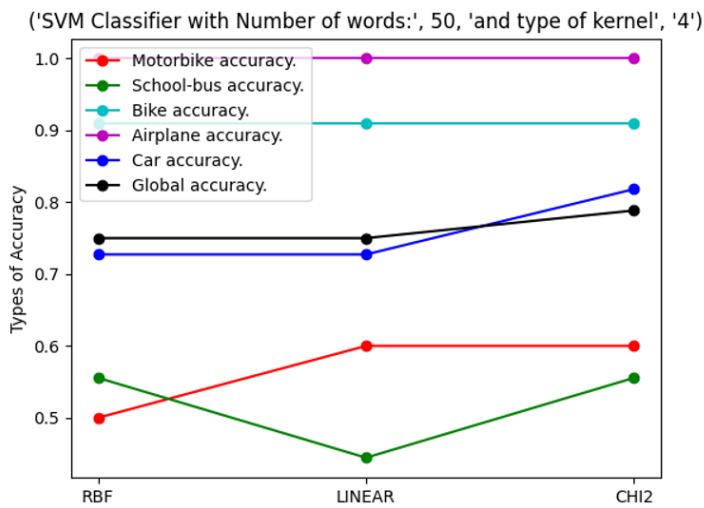
- Η ακρίβεια της κλάσης motorbike παραμένει μηδενική ανεξάρτητα από τις τιμές της παραμέτρου **k**.
- Η ακρίβεια της κλάσης school-bus μειώνεται με την αύξηση της παραμέτρου **k**, με έναν πιο αργό βαθμό σε σχέση με την περίπτωση των 100 οπτικών λέξεων.
- Η ακρίβεια της κλάσης touring-bike μειώνεται με την αύξηση της παραμέτρου **k**.
- Η ακρίβεια της κλάσης airplane σχεδόν για όλες τις τιμές της παραμέτρου **k** παραμένει πολύ υψηλή και αυξάνεται με την αύξηση αύξηση της παραμέτρου **k**.
- Η ακρίβεια της κλάσης car αυξάνεται με την αύξηση της παραμέτρου **k**. Σε σχέση με την περίπτωση των 100 οπτικών λέξεων, η ακρίβεια για τις περιπτώσεις των $k = 5$ και $k = 20$, είναι χαμηλοτερη.
- Η καθολική ακρίβεια φαίνεται να αυξάνεται με την αύξηση της παραμέτρου **k**, με την μέγιστη τιμή αυτής για τις τιμές του $k = \{40, 60\}$

Παράμετρος του τύπου πυρήνα του ταξινομητή SVM:

Όπως αναφέρθηκε και παραπάνω, η επίδραση της παραμέτρου του πυρήνα του ταξινομητή SVM, ελέγχθηκε επίσης συναρτίσει του πληθους των οπτικών λέξεων. Ανεξάρτητα από αυτήν την επιλογή, οι τιμές του πλήθους των πλησιέστερων γειτόνων για τον ταξινομητή SVM, αναγράφονται παρακάτω:

- RBF
- LINEAR
- CHI2

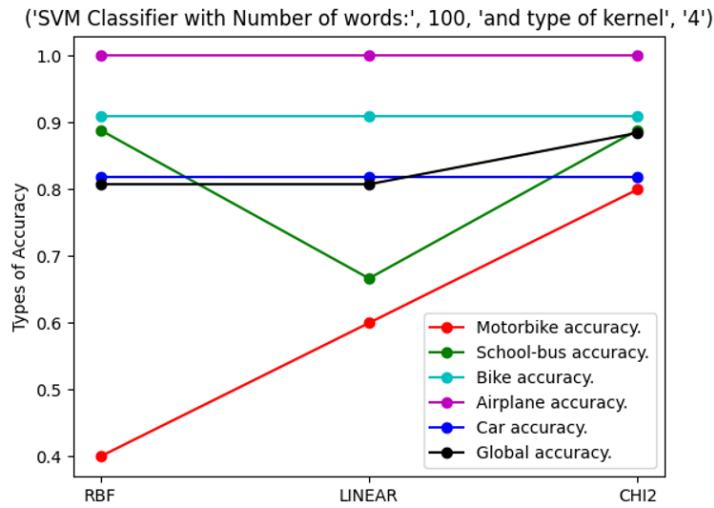
Στο παρακάτω διάγραμμα φαίνονται οι επιθυμητές ακρίβειες του συστήματος συναρτήσει του τύπου πυρήνα του ταξινομητή SVM για την περίπτωση των 50 οπτικών λέξεων:



Από το παραπάνω διάγραμμα προκύπτουν τα εξής συμπερασματα:

- Η ακρίβεια της κλάσης motorbike παραμένει στο 50% για τον τύπο RBF, ενώ στο 60% για τους τύπους LINEAR και CHI2.
- Η ακρίβεια της κλάσης school-bus παραμένει χαμηλή σχεδόν για όλους τους τύπους πυρήνα.
- Η ακρίβεια της κλάσης touring-bike παραμένει υψηλή για όλους τους τύπους πυρήνα περίπου στο 90%.
- Η ακρίβεια της κλάσης airplane παραμένει υψηλή για όλους τους τύπους πυρήνα περίπου στο 100%.
- Η ακρίβεια της κλάσης car παραμένει λίγο πιο πάνω από το 70% για τους τύπους RBF και LINEAR, ενώ περίπου στο 80% για τον τύπο CHI2.
- Η καθολική ακρίβεια φαίνεται να έχει την μέγιστη τιμή της για τον τύπο πυρήνα CHI2.

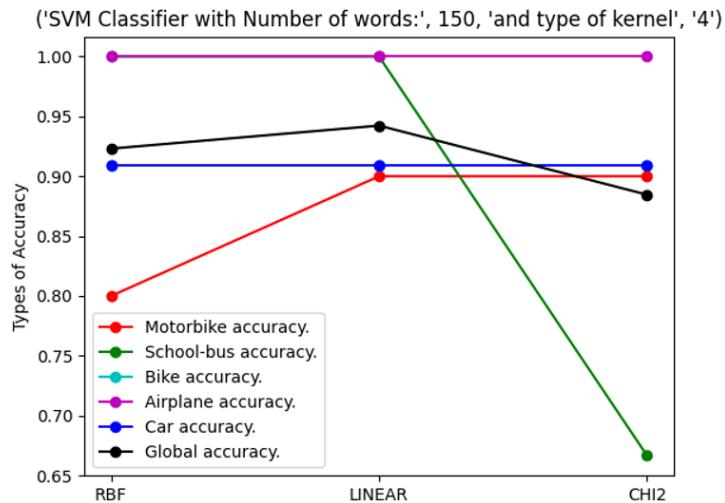
Στο παρακάτω διάγραμμα φαίνονται οι επιθυμητές ακρίβειες του συστήματος συναρτήσει του τύπου πυρήνα του ταξινομητή SVM για την περίπτωση των 100 οπτικών λέξεων:



Από το παραπάνω διάγραμμα προκύπτουν τα εξής συμπερασματα:

- Η ακρίβεια της κλάσης motorbike παραμένει περίπου στο 40% για τον τύπο RBF, περίπου στο 60% για τον τύπο LINEAR, ενώ περίπου στο 80% για τον τύπο CHI2.
- Η ακρίβεια της κλάσης school-bus παραμένει υψηλή σχεδόν για τους τύπους πυρήνα RBF και CHI2 και περίπου στο 90%, ενώ η απόδοσή του συστήματος πέφτει για την περίπτωση του τύπου πυρήνα LINEAR περίπου στο 70%.
- Η ακρίβεια της κλάσης touring-bike παραμένει υψηλή για όλους τους τύπους πυρήνα περίπου στο 90%.
- Η ακρίβεια της κλάσης airplane παραμένει υψηλή για όλους τους τύπους πυρήνα περίπου στο 100%.
- Η ακρίβεια της κλάσης car παραμένει υψηλή για όλους τους τύπους πυρήνα περίπου στο 80%.
- Η καθολική ακρίβεια φαίνεται να έχει την μέγιστη τιμή της για τον τύπο πυρήνα CHI2.

Στο παρακάτω διάγραμμα φαίνονται οι επιθυμητές ακρίβειες του συστήματος συναρτήσει του τύπου πυρήνα του ταξινομητή SVM για την περίπτωση των 150 οπτικών λέξεων:



Από το παραπάνω διάγραμμα προκύπτουν τα εξης συμπερασματα:

- Η ακρίβεια της κλάσης motorbike παραμένει περίπου στο 80% για τον τύπο RBF, περίπου στο 90% για τους τύπους LINEAR και CHI2.
- Η ακρίβεια της κλάσης school-bus παραμένει υψηλή σχεδόν για τους τύπους πυρήνα RBF και CHI2 και περίπου στο 100%, ενώ η απόδοσή του συστήματος πέφτει για την περίπτωση του τύπου πυρήνα LINEAR περίπου κάτω από το 70%.
- Η ακρίβεια της κλάσης touring-bike παραμένει υψηλή για όλους τους τύπους πυρήνα στο 100%, αν και δεν φαίνεται στο διάγραμμα.
- Η ακρίβεια της κλάσης airplane παραμένει υψηλή για όλους τους τύπους πυρήνα στο 100%.
- Η ακρίβεια της κλάσης car παραμένει υψηλή για όλους τους τύπους πυρήνα περίπου στο 90%.
- Η καθολική ακρίβεια φαίνεται να έχει την μέγιστη τιμή της για τον τύπο πυρήνα LINEAR.

Ποσοτικά αποτελέσματα αξιολόγησης:

Για την περίπτωση των 50 οπτικών λέξεων:

Αποτελέσματα του ταξινομητή kNN για τα ζητούμενα accuracies. Με υπογράμμιση σημειώνονται οι μέγιστες τιμές :

	k = 5	k = 20	k = 40	k = 60	k = 80	k = 100
motorbike	<u>0.5</u>	0.1	0.1	0.3	0.2	0.1
school-bus	<u>0.777</u>	0.444	0.444	0.666	0.666	<u>0.777</u>
touring-bike	1	<u>0.818</u>	<u>0.818</u>	<u>0.818</u>	<u>0.818</u>	0.636
airplane	0.818	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>
car	0.454	0.545	0.545	0.545	0.636	<u>0.727</u>
global	<u>0.711</u>	0.596	0.596	0.673	0.673	0.653

- Μέγιστη ακρίβεια κλάσης motorbike: 0.5
- Μέγιστη ακρίβεια κλάσης school-bus: 0.777
- Μέγιστη ακρίβεια κλάσης touring-bike: 0.818
- Μέγιστη ακρίβεια κλάσης airplane: 1
- Μέγιστη ακρίβεια κλάσης car: 0.727
- Μέγιστη καθολική ακρίβεια: 0.711

Αποτελέσματα του ταξινομητή SVM για τα ζητούμενα accuracy. Με υπογράμμιση σημειώνονται οι μέγιστες τιμές :

	SVM RBF	SVM LINEAR	SVM CHI2
motorbike	0.5	<u>0.6</u>	<u>0.6</u>
school-bus	<u>0.555</u>	0.444	<u>0.555</u>
touring-bike	<u>0.909</u>	<u>0.909</u>	<u>0.909</u>
airplane	1	1	1
car	0.727	0.727	<u>0.818</u>
global	0.75	0.75	<u>0.788</u>

- Μέγιστη ακρίβεια κλάσης motorbike: 0.6
- Μέγιστη ακρίβεια κλάσης school-bus: 0.555
- Μέγιστη ακρίβεια κλάσης touring-bike: 0.909
- Μέγιστη ακρίβεια κλάσης airplane: 1
- Μέγιστη ακρίβεια κλάσης car: 0.818
- Μέγιστη καθολική ακρίβεια: 0.788

Για την περίπτωση των 100 οπτικών λέξεων:

Αποτελέσματα του ταξινομητή kNN για τα ζητούμενα accuracy. Με υπογράμμιση σημειώνονται οι μέγιστες τιμές :

	k = 5	k = 20	k = 40	k = 60	k = 80	k = 100
motorbike	0.1	0.1	0	<u>0.2</u>	0	0
school-bus	<u>1</u>	<u>1</u>	0.888	0.666	0.666	0.666
touring-bike	<u>1</u>	0.909	0.818	0.727	0.727	0.636
airplane	0.909	0.818	0.909	0.909	0.909	<u>1</u>
car	0.272	0.454	0.727	0.636	0.727	<u>0.909</u>
global	0.653	0.653	<u>0.673</u>	0.634	0.615	0.653

- Μέγιστη ακρίβεια κλάσης motorbike: 0.2
- Μέγιστη ακρίβεια κλάσης school-bus: 1
- Μέγιστη ακρίβεια κλάσης touring-bike: 1
- Μέγιστη ακρίβεια κλάσης airplane: 1
- Μέγιστη ακρίβεια κλάσης car: 0.909
- Μέγιστη καθολική ακρίβεια: 0.673

Αποτελέσματα του ταξινομητή SVM για τα ζητούμενα accuracy. Με υπογράμμιση σημειώνονται οι μέγιστες τιμές :

	SVM RBF	SVM LINEAR	SVM CHI2
motorbike	0.4	0.5	<u>0.8</u>
school-bus	<u>0.888</u>	0.666	<u>0.888</u>
touring-bike	<u>0.909</u>	<u>0.909</u>	<u>0.909</u>
airplane	<u>1</u>	<u>1</u>	<u>1</u>
car	<u>0.818</u>	<u>0.818</u>	<u>0.818</u>
global	0.807	0.807	<u>0.884</u>

- Μέγιστη ακρίβεια κλάσης motorbike: 0.8
- Μέγιστη ακρίβεια κλάσης school-bus: 0.88
- Μέγιστη ακρίβεια κλάσης touring-bike: 0.909
- Μέγιστη ακρίβεια κλάσης airplane: 1
- Μέγιστη ακρίβεια κλάσης car: 0.818
- Μέγιστη καθολική ακρίβεια: 0.884

Για την περίπτωση των 150 οπτικών λέξεων:

Αποτελέσματα του ταξινομητή kNN για τα ζητούμενα accuracy. Με υπογράμμιση σημειώνονται οι μέγιστες τιμές :

	k = 5	k = 20	k = 40	k = 60	k = 80	k = 100
motorbike	0	0	0	0	0	0
school-bus	0.888	<u>1</u>	<u>1</u>	<u>1</u>	0.666	0.666
touring-bike	<u>1</u>	0.818	0.818	0.727	0.636	0.636
airplane	0.818	0.818	0.909	<u>1</u>	0.909	<u>1</u>
car	0.09	0.363	0.636	0.636	0.727	<u>0.909</u>
global	0.557	0.596	<u>0.673</u>	<u>0.673</u>	0.596	0.653

- Μέγιστη ακρίβεια κλάσης motorbike: 0
- Μέγιστη ακρίβεια κλάσης school-bus: 1
- Μέγιστη ακρίβεια κλάσης touring-bike: 1
- Μέγιστη ακρίβεια κλάσης airplane: 1
- Μέγιστη ακρίβεια κλάσης car: 0.909
- Μέγιστη καθολική ακρίβεια: 0.673

Αποτελέσματα του ταξινομητή SVM για τα ζητούμενα accuracy. Με υπογράμμιση σημειώνονται οι μέγιστες τιμές :

	SVM RBF	SVM LINEAR	SVM CHI2
motorbike	0.8	<u>0.9</u>	<u>0.9</u>
school-bus	<u>1</u>	<u>1</u>	0.666
touring-bike	<u>0.909</u>	<u>0.909</u>	<u>0.909</u>
airplane	<u>1</u>	<u>1</u>	<u>1</u>
car	<u>0.909</u>	<u>0.909</u>	<u>0.909</u>
global	0.923	<u>0.943</u>	0.884

- Μέγιστη ακρίβεια κλάσης motorbike: 0.9
- Μέγιστη ακρίβεια κλάσης school-bus: 1
- Μέγιστη ακρίβεια κλάσης touring-bike: 0.909
- Μέγιστη ακρίβεια κλάσης airplane: 1
- Μέγιστη ακρίβεια κλάσης car: 0.909
- Μέγιστη καθολική ακρίβεια: 0.943

Ποσοτικός σχολιασμός αποτελεσμάτων αξιολόγησης και σχολιασμός ανά κλάση:

Κλάση motorbike:

Η μέγιστη ακρίβεια για την κλάση αυτή επιτυγχάνεται για την περίπτωση των 150 οπτικών λέξεων, του ταξινομητή SVM για τους τύπους της παραμέτρου πυρήνα LINEAR και CHI2 και ισούται με 0.9.

Κλάση school-bus:

Η μέγιστη ακρίβεια για την κλάση αυτή επιτυγχάνεται για την περίπτωση των 150 οπτικών λέξεων, στους ταξινομητές kNN και SVM για τους τύπους της παραμέτρου πυρήνα RBF και LINEAR και τις τιμές του $k = \{20, 40, 60\}$ και ισούται με 1.

Κλάση touring-bike:

Η μέγιστη ακρίβεια για την κλάση αυτή επιτυγχάνεται για την περίπτωση των 100 οπτικών λέξεων, του ταξινομητή kNN για την τιμή της παραμέτρου $k=5$ και ισούται με 1. Επίσης ο ταξινομητής kNN για την περίπτωση των 150 οπτικών λέξεων σημείωσε την μέγιστη ακρίβεια για τις τιμές της παραμέτρου $k = \{20, 40, 60\}$.

Κλάση airplane:

Η μέγιστη ακρίβεια για την κλάση αυτή επιτυγχάνεται για όλες τις περιπτώσεις των οπτικών λέξεων, και στους δύο ταξινομητές, σχεδόν για όλες τις τιμές των παραμέτρων τους και ισούται με 1.

Κλάση car:

Η μέγιστη ακρίβεια για την κλάση αυτή επιτυγχάνεται για τις περιπτώσεις των 100 και 150 οπτικών λέξεων, του ταξινομητή kNN για την τιμή της παραμέτρου $k = 100$ και ισούται με 0.909. Επίσης αυτή η μέγιστη απόδοση επιτυγχάνεται και για την περίπτωση των 150 λέξεων για τον ταξινομητή SVM για όλους τους τύπους πυρήνα.

Όλες οι κλάσεις συνολικά:

Η μέγιστη καθολική ακρίβεια επιτυγχάνεται για την περίπτωση των 150 οπτικών λέξεων, του ταξινομητή SVM για τον τύπο πυρήνα LINEAR και ισούται με 0.943.

Εξέταση κλάσεων, όπου υπάρχει ιδιαίτερη αποτυχία και δικαιολόγηση των λόγων αυτής.

Ταξινομητής kNN:

Όπως φάνηκε από τα παραπάνω διαγράμματα οι κλάσεις, οι οποίες είχαν τα βέλτιστα αποτελέσματα (σχεδόν σε όλες τις περιπτώσεις ακρίβεια πάνω από 80%) ήταν:

- η κλάση airplane
- η κλάση touring-bike

Τα αποτελέσματα για τις κλάσεις αυτές οφείλονταν στα μοναδικά χαρακτηριστικά της κλάσης airplane από την μια και στην μικρη σχετικά με τις άλλες κλάσεις ποικιλομορφία της κλάσης touring-bike. Για τους λόγους αυτούς, οι δύο κατηγορίες εικόνων δοκιμής μπορούσαν να ταξινομηθούν εύστοχα από τους δύο αλγορίθμους.

Επίσης οι κλάσεις που είχαν λιγότερα βέλτιστα αποτελέσματα (ακρίβεια μεταξύ περίπου 20% και 80%) ήταν:

- η κλάση car

Τέλος οι κλάσεις οι οποίες παρουσίασαν ιδιαίτερη αποτυχία, ήταν:

- η κλάση motorbike

Ένας πιθανός λόγος για την σχετική αυτή αποτυχία, ήταν η μεγάλη ποικιλομορφία της κλάσης αυτής. Με άλλα λόγια, η κατηγορία motorbike στο σετ εκπαίδευσης είχε πολλούς και διαφορετικούς τύπους από μοτοσυκλέτες, με αποτέλεσμα τόσο τα global features που θα λάμβανε υπόψη ο αλγόριθμος kNN, όσο και το μοντέλο που προέκυπτε από τον αλγόριθμο SVM, ήταν επηρρεασμένα από αυτούς. Με την εκπαίδευση ενός σετ δεδομένων με μια τέτοια ποικιλομορφία, οι αλγόριθμοι στο πλαίσιο της δοκιμής θα μπορούσαν να βρουν αρκετά είδη από μοτοσυκλέτες, ωστόσο με μικρές πιθανότητες επιτυχίας, το οποίο δεν είναι τίποτα άλλο από την ζητουμενή ακρίβεια.

Όσον αφορά την κλάση school-bus τα αποτελέσματα της αξιολόγησης σε ποσοστά ήταν εξαρτημένα από τον αλγόριθμο και τις παραμέτρους που επιδρούσαν. Σημαντική παρατήρηση ήταν ότι η ακρίβεια της συγκεκριμένης κλάσης αυξανόταν με την αύξηση των οπτικών λέξεων. Να σημειωθεί ότι η συγκεκριμένη κλάση είχε επίσης έναν βαθμό ποικιλομορφίας τόσο στο σετ εκπαίδευσης, όσο και στο σετ δοκιμής. Αυτός είναι ενδεχομένως ένας άλλος λόγος που οι δύο ταξινομητές παρουσίασαν αυτήν την σχετική αποτυχία ανά περιπτώσεις.

Ταξινομητής SVM:

Όσον αφορά τον ταξινομητή SVM, τα αποτελέσματα ήταν σε καλύτερα ποσοστά επιτυχίας από τον kNN. Ο λόγος για αυτό ίσως ήταν η έλλειψη της παρακάτω συνθήκης:

Ο αλγόριθμος kNN επιστρέφει την κλάση που είχε τους περισσότερους γείτονες της κλάσής αυτής. Στην περίπτωση που διαφορετικές κλάσεις έχουν τον ίδιο αυτό μέγιστο αριθμό από γείτονες, ο συγκεκριμένος αλγόριθμος kNN δεν θέτει κάποιο κριτήριο επιλογής ανάμεσα στις κλάσεις αυτές, απλά επιλέγει και επιστρέφει μια από αυτές.

Ίσως για τον παραπάνω λόγο, τα αποτελέσματα του αλγορίθμου SVM να έχουν αυτήν την διαφορά στα αποτελέσματα σε σχέση με τον αλγορίθμο kNN.