ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΤΕΧΝΟΛΟΓΙΕΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΚΑΙ ΤΗ ΒΙΟΛΟΓΙΑ»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ταξινόμηση εικόνων ασύρματης ενδοσκόπησης με τη χρήση της μεθόδου Bag of Words

Ιωσήφ Ν. Λούτς

Επιβλέποντες: **Περαντώνης Σταύρος,** Δ/ντής Ερευνών, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, ΕΚΕΦΕ “Δημόκριτος”

Σπύρου Ευάγγελος, Συνεργαζόμενος ερευνητής, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, ΕΚΕΦΕ “Δημόκριτος’’

ΑΘΗΝΑ

Απρίλιος 2019

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ταξινόμηση εικόνων ασύρματης ενδοσκόπησης με τη χρήση της μεθόδου bag of Words

Ιωσήφ Ν. Λούτς

Α.Μ.: ΠΙΒ0164

ΕΠΙΒΛΕΠΟΝΤΕΣ: **Περαντώνης Σταύρος,** Δ/ντής Ερευνών, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, ΕΚΕΦΕ "Δημόκριτος"

Σπύρου Ευάγγελος, Συνεργαζόμενος ερευνητής, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, ΕΚΕΦΕ "Δημόκριτος"

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: **Περαντώνης Σταύρος,** Δ/ντής Ερευνών, Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών, ΕΚΕΦΕ "Δημόκριτος"

Απρίλιος 2019

ΠΕΡΙΛΗΨΗ

Η ασύρματη ενδοσκόπηση, αποτελεί μία σχετικά καινούρια απεικονιστική μέθοδο,η οποία συμβάλλει σημαντικά στη διάγνωση ενός μεγάλου αριθμού ασθενειών του πεπτικού συστήματος. Το βασικό πρόβλημα που προκύπτει από τη μέθοδο αυτή, αφορά τον υψηλό αριθμό εικόνων που καλείται να εξετάσει ο θεράπων ιατρός προκειμένου να βγάλει τη σωστή διάγνωση. Η διαδικασία αυτή ενδέχεται να είναι αρκετά επίπονη, καθώς απαιτεί αρκετές ώρες αδιάκοπης προσοχής. Αυτή η διπλωματική εργασία, παρουσιάζει τη μέθοδο ταξινόμησης bag of words, η οποία ταξινομεί τις εικόνες αυτές σε διάφορες κατηγορίες ασθενειών, δίνοντας έτσι μία λύση στο ανωτέρω πρόβλημα. Η λογική της βασίζεται στον εντοπισμό περιοχών ενδιαφέροντος πάνω στις εικόνες, καθώς και στην εξαγωγή χαρακτηριστικών από τις περιοχές αυτές. Για την πραγματοποίηση των πειραμάτων χρησιμοποιήθηκαν 2 dataset. Αυτά ήταν το KID dataset 2, το οποίο περιέχει 8 κατηγορίες και το Normal-Abnormal σέτ, το οποίο αποτελεί μία παραλλαγή του KID dataset 2 και περιέχει μόνο 2 κατηγορίες. Τα πειράματα πραγματοποιήθηκαν σε προγραμματιστικό περιβάλλον MATLAB. Κατά την εκτέλεση τους δοκιμάστηκαν διαφορετικές μέθοδοι εξαγωγής χαρακτηριστικών, ενώ παράλληλα μετρήθηκαν τα ποσοστά ορθής ταξινόμησής των εικόνων τους για τις μεθόδους αυτές. Οι μέθοδοι εξαγωγής χαρακτηριστικών που χρησιμοποιήθηκαν ήταν όι : SIFT, RGB-SIFT, HSV-SIFT, Opponent-SIFT και HSV&Opponent SIFT (συνδυασμός). Εκτός από τις μεθόδους εξαγωγής χαρακτηριστικών, μεταξύ των πειραμάτων, μεταβάλλονταν συνεχώς και άλλες δύο μεταβλητές, διότι εξεταζόταν η επιρροή τους πάνω στα αποτελέσματα. Τα υψηλότερα αποτελέσματα τα πέτυχε η μέθοδος εξαγωγής χαρακτηριστικών HSV-SIFT, καθώς έδωσε μέγιστο ποσοστό ορθης ταξινόμησης 81.46% στο Normal-Abnormal σέτ και μέγιστο ποσοστό ορθής ταξινόμησης 47.92% στο KID dataset 2.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Μηχανική μάθηση, αναγνωριση προτύπων, υπολογιστική όραση

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Ασύρματη ενδοσκόπηση, ταξινόμηση εικόνων, μέθοδος bag of words, εξαγωγή χαρακτηριστικών, SIFT, RGB-SIFT, HSV-SIFT, Opponent-SIFT

ABSTRACT

Wireless endoscopy is a relatively new imaging method, which highly contributes to the diagnosis of many digestive system diseases. The main problem of this method, concerns the huge number of images that the clinician has to examine, in order to end up with the right diagnosis. This procedure may be very tiresome for the clinician, because it requires many hours of uninterrupted attention. This master thesis, offers a solution to the aforementioned problem, by presenting the bag of words classification method, which can classify those images in categories, according to the disease that they depict. This method is based on the detection of regions of interests in images. Once regions of interest are found, features will be extracted from those regions. Those features are the abbreviated representation of an image. Two major datasets were used for the experiments conduction. The first is KID dataset 2, which contains 8 categories and the second is the Normal-Abnormal set which contains only 2 categories. The experiments were conducted in MATLAB programming environment. During the experiments, many feature extraction methods were used and their results were evaluated. Those methods were: SIFT, RGB-SIFT HSV-SIFT, Opponent-SIFT and HSV&Opponent SIFT (Combination). Apart from those methods, two more variables were constantly changing their values between the experiments, in order to conclude if they had any impact on the results. HSV-SIFT extraction method presented the highest results, as it has reached 81.46% classification percentage in Normal-Abnormal set, and 47.92% in KID dataset 2.

SUBJECT AREA: Machine learning, pattern recognition, computer vision

KEYWORDS: Wireless endoscopy, image classification, bag of words model, feature extraction, SIFT, RGB-SIFT, HSV-SIFT, Opponent-SIFT

Στους γονείς μου και την αδερφή μου…

ΕΥΧΑΡΙΣΤΙΕΣ

Η περάτωση αυτής της πτυχιακής δεν ήταν μόνο προσωπική μου υπόθεση. Για το λόγο αυτό, δεν θα μπορούσα να μην ευχαριστήσω συγκεκριμένα άτομα, που μου παρείχαν πνευματική και υλική υποστήριξη όποτε το είχα ανάγκη. Αρχικά, θα ήθελα να ευχαριστήσω τους γονείς μου, οι οποίοι τα τελευταία χρόνια έχουν θυσιάσει τόσα πολλά, έτσι ώστε να μου παρέχουν τα απαρραίτητα εφόδια για να μπορέσω πραγματοποιήσω τα ονειρά μου. Θα ήθελα επίσης να ευχαριστήσω τον Καθηγητή κ. Περαντώνη Σταύρο, για την ανάθεση του θέματος αυτού, αλλά και για το γεγονός, ότι μέσα από την διδασκαλία του μαθήματος της αναγνώρισης προτύπων, με ενέπνευσε να ασχοληθώ ερευνητικά με το αντικείμενο αυτό. Δεν θα μπορούσα να παραλείψω και τον Μεταδιδακτορικό ερευνητή Ευάγγελο Σπύρου για την συνεργασία μας. Η συνεχής παρουσία του και οι ουσιαστικές παρατηρήσεις του, συνέβαλλαν τα μέγιστα για την περάτωση της πτυχιακής. Κυρίως όμως τον ευχαριστώ για την εμπιστοσύνη που μου έδειξε. Τέλος θα ήθελα να ευχαριστήσω όλους αυτούς, οι οποίοι στις δύσκολες στιγμές μου, με άκουσαν και με στήριξαν. Τους ευχαριστώ όλους.

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

1. ΠΕΠΤΙΚΟ ΣΥΣΤΗΜΑ ΚΑΙ ΑΣΥΡΜΑΤΗ ΕΝΔΟΣΚΟΠΗΣΗ

1.1 Εισαγωγή

Στο παρόν κεφάλαιο, πραγματοποιείται μία απλή μελέτη του πεπτικού συστήματος (Digestive System) και της ασύρματης ενδοσκόπησης (Wireless Endoscopy). Αρχικά, γίνεται μία αναφορά, στην αξία του πεπτικού συστήματος για τον ανθρώπινο οργανισμό, στα βασικά δομικά του μέρη, καθώς και στον ρόλο που επιτελεί το καθένα από αυτά. Έπειτα αναφέρονται κάποιες από τις λοιμώξεις που ενδέχεται να προσβάλλουν το πεπτικό σύστημα, ενώ παράλληλα εξετάζονται και οι συνέπειες τους. Τέλος, μελετάται η ασύρματη ενδοσκόπηση και η προσφορά της στην διάγνωση των ασθενειών του πεπτικού συστήματος.

1.2 Το πεπτικό σύστημα

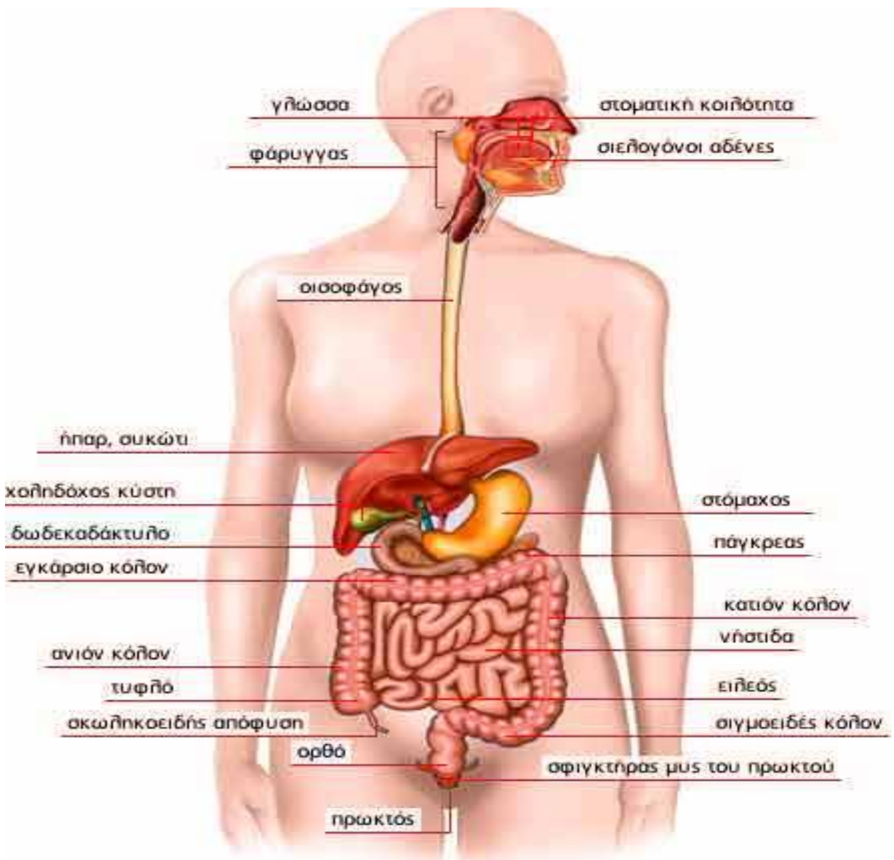
Το πεπτικό σύστημα διαδραματίζει σημαντικό ρόλο στην ομαλή λειτουργία του ανθρώπινου οργανισμού. Βασικός του σκοπός είναι η πέψη, δηλαδή η διάσπαση της τροφής με χημικό και μηχανικό τρόπο, έτσι ώστε να μετατραπεί σε θρεπτικά συστατικά, τα οποία στη συνέχεια θα απορροφηθούν από τον οργανισμό παρέχοντας του την απαραίτητη ενέργεια [1]. Τα θρεπτικά αυτά συστατικά, αξιοποιούνται επίσης και σε άλλες διεργασίες του οργανισμού, όπως η ανάπλαση κυττάρων.

Η δομή του πεπτικού συστήματος έχει ως εξής [1]:

I. **Ανώτερη γαστρεντερική οδός:** Η οποία αποτελείται από το στόμα, το φάρυγγα, τον οισοφάγο και το στομάχι.
II. **Κατώτερη γαστρεντερική οδός:** Η οποία αποτελείται από το παχύ έντερο και το λεπτό έντερο.
III. **Λοιπά όργανα που συμμετέχουν και στη διαδικασία της πέψης:** Αυτά είναι το συκώτι, πάγκρεας και χοληδόχος κύστη.

Η λειτουργία του πεπτικού συστήματος έχει ως εξής: Αρχικά η τροφή εισέρχεται στη στοματική κοιλότητα, όπου με τη βοήθεια των δοντιών και τις γλώσσας «συνθλίβεται» και αναμιγνύεται με σίελο, λαμβάνοντας με αυτόν τον τρόπο μία πιό ρευστή μορφή [2]. Στη συνέχεια κατευθύνεται στο στομάχι μέσω του οισοφάγου. Εκεί με τη βοήθεια συγκεκριμένων χημικών διεργασιών η τροφή υπόκειται σε μερική πέψη και μετατρέπεται σε χυλό. Το κύριο μέρος της πέψης λαμβάνει χώρα στο λεπτό έντερο. Το λεπτό έντερο με τη βοήθεια χημικών ουσιών που εκρίννονται από το συκώτι και το πάγκρεας, απορροφάει τα θρεπτικά συστατικά των τροφών [2],[3]. Τέλος τα άπεπτα υπολλείματα της τροφής κατευθύνονται στο παχύ έντερο, όπου και διαμένουν προσωρινά μέχρι την αποβολή τους από τον οργανισμό με τη μορφή κοπράνων [4].

Η δομή και τα όργανα του πεπτικού συστήματος, διαφαίνονται στην παρακάτω εικόνα (Εικόνα 1):



Εικόνα 1: Η δομή και τα όργανα του πεπτικού συστήματος (Πηγή: http://pemptakia2.blogspot.com/2010/10/blog-post_30.html)

1.3 Ασθένειες του πεπτικού συστήματος και η αξία έγκαιρης διάγνωσης

Όπως όλα τα συστήματα του ανθρώπινου σώματος (αναπνευστικό, καρδιαγγειακό κ.λπ), έτσι και το πεπτικό σύστημα, είναι αρκετά πιθανό να προσβληθεί από κάποια ασθένεια. Οι ασθένειες που προσβάλλουν το πεπτικό σύστημα, ενδέχεται να δημιουργούν δυσλειτουργία σε κάποιο από τα όργανα του συστήμάτος, διαταράσσοντας με αυτόν τον τρόπο την διαδικασία της πέψης. Ως εκ τούτου, προκαλούν ποίκιλα και σοβαρά σε αρκετές περιτώσεις συμπτώματα στον ανθρώπινο οργανισμό.

Αναφέρονται μερικές από τις πιο συχνά εμφανιζόμενες ασθένειες:

1. **Πολύποδες** (Polypoids): Είναι μικροί όγκοι που έχουν τη μορφή αλλοιώσεων [5]. Εμφανίζονται σε αρκετά σημεία του πεπτικού συστήματος, όπως το στομάχι, η χοληδόχος κύστη, το παχύ έντερο κ.οκ. Μερικοί από αυτούς ενδέχεται να εξελιχθούν σε καρκινικούς όγκους, γι'αυτό είναι επιτακτική η έγκαιρή διάγνωση τους [5].
2. **Αλλοιώσεις της Ηπατοπαγκρεατικής λήκυθου** (Ampulla of Vater): Η ηπατοπαγκρεατική λήκυθος, σχηματίζεται από την κοινή εκβολή του χοληδόχου και του μείζονος παγκρεατικού πόρου και βρίσκεται επί τα εντός της μείζονος δωδεκαδακτυλικής θήλης [6], [7], [8]. Μέσω αυτής εκχύνονται μέσα στο λεπτό έντερο ουσίες προερχόμενες από το πάγκρεας και το ηπαρ, οι οποίες συμβάλλουν στην πέψη. Αλλοιώσεις της ηπατοπαγκρεατικής ληκύθου, όπως κάποιο καρκίνωμα, δημιουργόυν προβλήματα στην πέψη [9].
3. **Φλεγμονές** (Inflammatory) και **αγγειακές ανωμαλίες** (Vascular): Τόσο οι φλεγμονές, όσο και οι αγγειακές ανωμαλίες, έιναι αρκετά συχνές στο πεπτικό σύστημα. Και στις δύο περιπτώσεις η έγκαιρη διάγνωση είναι αναγκαία έτσι ώστε ανάλογα με τον τύπο τους, να ακολουθηθεί και συγκεκριμένη αντιμετώπιση.

Γίνεται απόλυτα κατανοητό λοιπόν, το μέγεθος της σημασίας που αφορά την έγκαιρη και έγκυρη διάγνωση. Όσο απειλητική και αν είναι μία ασθένεια, αν εντοπισθεί στην αρχή της, τότε οι ελπίδες της αποδοτικής αντιμετώπισης της είναι πολύ περισσότερες από ότι αν η διάγνωση της γίνει σε προχωρημένο στάδιο. Η επιβίωση και η ποιότητα ζωής ενός ανθρώπου που πάσχει από κάποια ασθένεια του πεπτικού συστήματος, εξαρτάται λοιπόν από την έγκαιρη και έγκυρη διάγνωση της. Υπάρχουν πολλές μέθοδοι διάγνωσης ασθενειών του πεπτικού συστήματος. Ωστόσο, μεγάλο ενδιαφέρον παρουσιάζει η ασύρματη ενδοσκόπηση με τη χρήση ενδοσκοπικής κάψουλας (Wireless Capsule Endoscopy, WCE), η οποία δείχνει να κερδίζει όλο και περισσότερο έδαφος τα τελευταία χρόνια.

1.4 Ασύρματη ενδοσκόπηση με τη χρήση κάψουλας

1.4.1 Εισαγωγικά στοιχεία και τρόπος εξέτασης

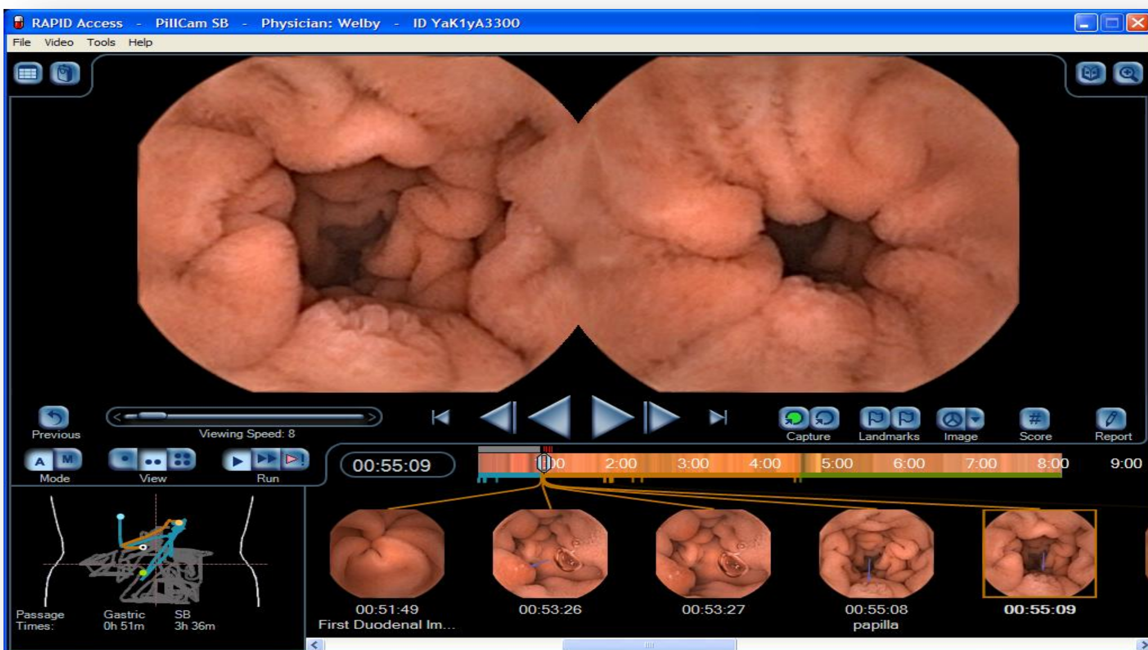
Η ασύρματη ενδοσκόπηση με τη χρήση κάψουλας (Wireless capsule endoscopy, WCE ή Video capsule endoscopy, VCE), είναι μία μη επεμβατική και ανώδυνη μέθοδος η οποία εισήχθη το 2001 στην κλινική πράξη [11]. Σκοπός της είναι να ρίξει φώς στη διάγνωση της περιοχής του λεπτού εντέρου, της οποίας ο έλεγχος παρουσιάζει πολλές δυσκολίες καθώς η συμβατική γαστροσκόπηση και η κολονοσκόπηση δεν έχουν τη δυνατότητα να την εξετάσουν εύκολα [10],[12]. Με παρόμοιες μεθόδους μπορεί επίσης να εξετασθεί η περιοχή του οισοφάγου και του παχέως εντέρου. Στις μέρες μας, συστήματα VCE κατασκευάζουν 4 εταιρείες: Η Given Images [27] που βρίσκεται στο Ισραήλ, η Olympus [28] στην Ιαπωνία, η Intromedic [29] στην Κορέα και η Chongqing jinshan science & technology στην Κίνα [30].

Το σύστημα της εξέτασης αποτελείται από 3 μέρη [13]:

- Μία μικροσκοπική και ασύρματη έγχρωμη κάμερα (Εικόνα 2), η οποία είναι τοποθετημένη μέσα σε μία κάψουλα, παρόμοιου μεγέθους με μία μεγάλη φαρμακευτική κάψουλα. Η κάμερα ενεργοποιείται και καταπίνεται από τον ασθενή με τη βοήθεια ενός ποτηριού νερού. Το μέγεθος της κάψουλας είναι 11x26mm, ενώ το βάρος της είναι 3,7 g [13]. Αξίζει να σημειωθεί ότι η κάμερα διαθέτει και πηγή φωτός.
- Ένα σύνολο 8 ηλεκτροδίων (αισθητήρων), που εφαρμόζονται σε συγκεκριμένες θέσεις στο θώρακα και στην κοιλιακή χώρα [12]. Τα ηλεκτρόδια αυτά ενώνονται με μία φορητή συσκευή όπου καταγράφονται δεδομένα. Η συσκευή αυτή μαζί με τη φορητή της μπαταρία, τοποθετούνται σε ζώνη γύρω από τη μέση του ασθενούς.
- Ο υπολογιστής που με τη βοήθεια ενός ειδικού λογισμικού,γίνεται η επεξεργασία των ληφθέντων εικόνων (Εικόνα 3).

Πρίν την εξέταση ο ασθενής οφείλει να έχει πραγματοποιήσει και την απαρραίτητη προετοιμασία. Αυτή απαιτεί την μη πρόσληψη τροφών για 8 με 12 ώρες, ενώ παράλληλα δεν λαμβάνει κανένα φάρμακο τις τελευταίες 2 ώρες [14]. Στη συνέχεια εφαρμόζονται πάνω τα ειδικά ηλεκτρόδια με τη συσκευή καταγραφής και έπειτα καταπίνεται η κάψουλα. Η κάψουλα κινείται μέσα στον ανθρώπινο οργανισμό με τις φυσιολογικές κινήσεις του πεπτικού σωλήνα. Η διάρκεια της εξέτασης είναι 8 ώρες [13]. Καθ’όλη τη διάρκεια, η κάμερα τραβάει φωτογραφίες το εσωτερικό του πεπτικού συστήματος, τις οποίες τις στέλνει στη συσκευή καταγραφής.Κατα τη διάρκεια της εξέτασης επίσης, ο ασθενής μπορεί να πραγματοποιεί τις καθημερινές τους δραστηριότητες, ωστόσο πρέπει να αποφύγει την έντονη σωματική δραστηριότητα [14]. Στο τέλος της εξέτασης τα ηλεκτρόδια αφαιρούνται και η συσκευή καταφραφής δίνεται στο θεράπων ιατρό για την επεξεργασία των εικόνων. Η κάψουλα αναμένεται να αποβληθεί από τον οργανισμό με φυσιολογικό τρόπο τα επόμενα 24ώρα.

Εικόνα 2: Η ενδοσκοπική κάψουλα Pillcam (Πηγή: http://www.medicalexpo.com/prod/given-imaging/product-75056-469484.html)



Εικόνα 3: Λογισμικό ανάλυσης και επεξεργασίας εικόνων και βίντεο ασύρματης ενδοσκόπησης (Πηγή: https://dr-delis.gr/%CE%B1%CF%83%CF%85%CF%81%CE%BC%CE%B1%CF%84%CE%AE-%CE%B5%CE%BD%CE%B4%CE%BF%CF%83%CE%BA%CF%8C%CF%80%CE%B7%CF%83%CE%B7.html)

1.4.2 Ενδείξεις και αντενδείξεις

Η ασύρματη ενδοσκόπηση ενδείκνυται στις παρακάτω περιπτώσεις [11], [13]:

- Γαστρεντερική αιμορραγία αφανούς αιτίας
- Διερεύνηση φλεγμονωδών ασθενειών του εντέρου (N.Crohn)
- Όγκοι
- Σιδηροπενική ανεμία
- Σύνδρομα πολυποδίασης
- Κοιλιοκάκη (διάγνωση, επιπλοκές)
- Νεοπλασίες λεπτού εντέρου

Ωστόσο υπάρχουν και κάποιες περιπτώσεις όπου η χρήση της ασύρματης ενδοσκόπησης είναι απαγορευτική. Οι κυρίοτερες από αυτές είναι η κύηση και η υπάρξη στενώσεων στο γαστρεντερικό σωληνα ή άλλων σοβαρών διαταραχών της κινητικότητας του πεπτικού συστήματος [11], [12]. Επίσης, αν ο ασθενής φέρει απινιδωτή ή βηματοδότη, τότε η ασύρματη ενδοσκόπηση δεν πραγματοποιείται, ή πραγματοποιείται αποκλειστικά στο χώρο του νοσοκομείου, διότι η ύπαρξη της κάψουλας ενδέχεται να επηρεάσει τη λειτουργία τους [11]. Τέλος, ταυτόχρονα με την εξέταση της ασύρματης ενδοσκόπησης δεν γίνεται να πραγματοποιηθεί και η εξέταση της μαγνητικής τομογραφίας.

1.4.3 Επιπλοκές της ασύρματης ενδοσκόπησης

Οι επιπλοκες της ασύρματης ενδοσκόπησης είναι εξαιρετικά σπάνιες. Η πιο συχνή από αυτές, είναι ο εγκλωβισμός της κάψουλας σε κάποιο σημείο του γαστρεντερικού σωλήνα. Αν συμβεί κάτι τέτοιο, τότε θα πρέπει ο ασθενής να υποβληθεί σε χειρουργική επέμβαση για την αφαίρεση της κάψουλας [10]. Αν ο γιατρός υποψιαστεί ότι υπάρχει τέτοια περίπτωση, τότε πριν την κανονική εξέταση, χρησιμοποιείται πρώτα η κάψουλα βατότητας.

Η κάψουλα βατότητας είναι μία ειδική κάψουλα, ίδιου μεγέθους με την κανονική και χωρίς κάμερα, που χορηγείται στον ασθενή πριν την κανονική εξέταση. Αν κάψουλα αυτή διασχίσει το λεπτό έντερο με ασφάλεια τότε στη συνέχεια πραγματοποιείται η εξέταση με κανονική κάψουλα [10], [13]. Αν όχι, τότε αυτή αποσυντίθεται στο έντερο με βιολογικό τρόπο και η κανονική εξέταση δεν πραγματοποιείται [13]. Με δύο λόγια ο σκοπός της κάψουλας βατότητας είναι να καταδείξει με ασφάλεια τη βατότητα του εντερικού σωλήνα.

1.4.4 Μία άλλη οπτική γωνία

Όπως γίνεται κατανοητό, από την ασύρματη ενδοσκόπηση προκύπτει ένα πολύωρο βίντεο, καθώς και ένας μεγάλος αριθμός εικόνων που συνθέτει το βίντεο αυτό. Ο γιατρός οφείλει να αφιερώσει αρκετή ώρα, παρατηρώντας με προσοχή τις εικόνες για την ανίχνευση τυχόν ανωμαλιών. Αυτό αυξάνει όμως την πιθανότητα λανθασμένης διάγνωσης, καθώς έρευνες έχουν δείξει ότι μετά από συγκεκριμένη ώρα η απόδοση της προσοχής του πέφτει. Αρκετές φορές επίσης η δουλειά αυτή, ενδέχεται να ανατεθεί σε κάποιο άλλο μέλος του ιατρικού προσωπικού (π.χ νοσοκόμα), το οποίο όμως δεν είναι εξειδικευμένο. Γενικά η ανίχνευση ανωμαλιών σε εικόνες, από όποιον και αν πραγματοποιείται, έχει περιορισμένη επιτυχία γύρω στο 40% [15].Τι θα γινόταν όμως αν αναπτυσσόταν ένα λογισμικό ταξινόμησης των εικόνων αυτών σε υγιείς και μη υγιείς; Υπάρχει τέτοια πιθανότητα; Αν ναι, πόσο αποδοτική θα είναι; Οι απαντήσεις θα δοθούν στη συνέχεια.

2. ΘΕΩΡΗΤΙΚΑ ΣΤΟΙΧΕΙΑ

2.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο, παρατέθηκαν οι απαραίτητες γνώσεις, που αφορούσαν το πεπτικό σύστημα και την ασύρματη ενδοσκόπηση. Το κεφάλαιο αυτό παρόλαυτα, κινείται σε εντελώς διαφορετική λογική από το προηγούμενο. Πλέον, η ανάλυση αφορά τα θεωρητικά εργαλεία, που θα συμβάλλουν τόσο στην απάντηση των ερωτημάτων που τέθηκαν στην τελευταία υποενότητα του προηγούμενου κεφαλάιου (υποενότητα 1.4.4), όσο και στην εξαγωγή σημαντικών συμπερασμάτων σχετικά με τη διάγνωση ασθενειών. Αρχικά, παρουσιάζεται μία θεωρητική ανάλυση της μεθόδου Bag of Words (BoW) για την ταξινόμηση εικόνων βάσει της ύπαρξης η μη αντικειμένων σ'αυτές. Στη συνέχεια εξετάζεται η μέθοδος εξαγωγής χαρακτηριστικών SIFT (Scale Invariant Feature Transform), ενώ στο τέλος, αναφέρονται τρείς μέθοδοι SIFT με χρωματική διαφοροποίηση.

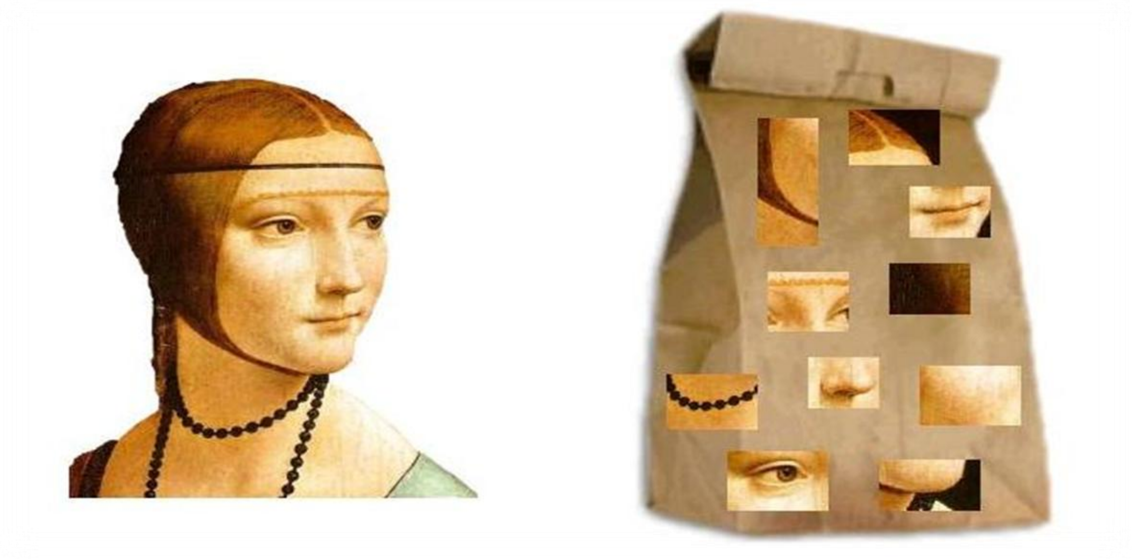
2.2 Ταξινόμηση εικόνων με τη μέθοδο Bag of Words

Η εύρεση αντικειμένων σε μία εικόνα, καθώς και η ταξινόμηση της με βάση την ύπαρξη ή όχι των αντικειμένων αυτων, είναι ένα δύσκολο ζήτημα στον τομέα της υπολογιστικής όρασης (computer vision). Ένα ανθρώπινο μάτι είναι εκπαιδευμένο στην αναγνώριση αντικειμένων σε μία εικόνα. Για έναν υπολογιστή όμως η διαδικασία αυτή είναι αρκετά πολύπλοκη και χρονοβόρα,ειδικά στην περίπτωση που η κλίμακα και η κατεύθυνση των αντικειμένων μεταβάλλονται από εικόνα σε εικόνα. Σ'αυτήν την περίπτωση η διαδικασία αναγνώρισης αντικειμένων και ταξινόμησης των εικόνων βάσει αυτών, γίνεται ακόμα πιο δύσκολη και απαιτητική.

Στην παρούσα ενότητα εξετάζεται η μέθοδος Bag of Words (BoW) για την ταξινόμηση εικόνων σε κατηγορίες. Εναλλακτικά, η μέθοδος αυτή μπορεί να βρεθεί στην βιβλιογραφία, είτε ως Bag of Features (BoF), είτε ως Bag of Visual Words. Εντύπωση προκαλεί το γεγονός, πως ενώ γίνεται αναφορά σε εικόνες, χρησιμοποιείται η λέξη «Words», η οποία στα Ελληνικά σημαίνει «λέξεις». Οπότε είναι λογικό να αναρωτηθεί κάποιος «Τι δουλειά έχουν οι λέξεις όταν συζητιέται η ταξινόμηση εικόνων;». Η απάντηση στο ερώτημα αυτό, αφορά την προέλευση της μεθόδου αυτής από την επεξεργασία και την ανάλυση κειμένων [23], [24].

Στην περίπτωση της ανάλυσης κειμένου, η ιδέα της μεθόδου Bag of Words γίνεται ιδιαίτερα σημαντική όταν υπάρχει μία βάση δεδομένων με εκατομμύρια έγγραφα κειμένου. Όταν κάποιος έχει την επιθυμία να αναζητήσει κάτι σε μία τέτοια βάση δεδομένων, χρειάζεται έναν αποδοτικό και γρήγορο τρόπο αναζήτησης. Προφανώς, δεν είναι εφικτή η σύγκριση της κάθε λέξης ενός εγγράφου, με την κάθε λέξη ενός άλλου εγγράφου, διότι κάτι τέτοιο θα ήταν εξαιρετικά χρονοβόρο. Έτσι δημιουργήθηκε η μέθοδος Bag of Words, η οποία αντιμετωπίζει κάθε έγγραφο ως ένα σάκο (Bag) σημαντικών λέξεων [24]. Με τον όρο σημαντικές λέξεις, εννοούμε τις λέξεις οι οποίες περιέχουν σημαντική πληροφορία για το κείμενο. Για παράδειγμα η λέξη «και» δεν περιέχει σημαντική πληροφορία, ενώ η λέξη «επεξεργασία», περιέχει. Μ'αυτό τον τρόπο δεν χρειάζεται να πραγματοποιηθεί αναζήτηση σ'όλες τις λέξεις του εγγράφου, αλλά μόνο στις σημαντικές. Συνεπώς η αναζήτηση γίνεται πιο αποδοτική.

Στον τομέα της ταξινόμησης εικόνων, η μέθοδος Bag of Words εφαρμόζεται ως εξής : Έστω ότι υπάρχει ένας μεγάλος αριθμός εικόνων με αντικείμενα, οι οποίες πρέπει να ταξινομηθούν σε δύο ή περισσότερες κατηγορίες, σύμφωνα με τα αντικείμενα που περιέχουν. Είναι προφανές, πως δεν θα ήταν καθόλου αποδοτικό και γρήγορο αν η κάθε εικόνα εξεταζόταν ανά pixel. Συνεπώς, με κάποιο τρόπο ανιχνεύονται κάποιες περιοχές ενδιαφέροντος (Regions of interest) στις εικόνες αυτές, οι οποίες αποτελούν τις περιοχές που περιέχουν τις χρήσιμες πληροφορίες της εικόνας. Από τις περιοχές αυτές στη συνέχεια εξάγονται κάποια διανύσματα τα οποία ονομάζονται χαρακτηριστικά (features ή feature vector) [22], [23], [24] Γι'αυτό το λόγο, όπως αναφέρθηκε και πιό πριν,η μέθοδος ονομάζεται εναλλακτικά και Bag of Features. Πλέον, οι εικόνες αντιμετωπίζονται ώς ένας σάκος που περιέχει τα χαρακτηριστικά, δηλαδή τη χρήσιμη πληροφορία [24]. Στην συνέχεια τα χαρακτηριστικά αυτά ομαδοποιούνται με μία μέθοδο ομαδοποίησης (k-means ή κάποια άλλη μέθοδο) [22]. Τα κέντρα των ομάδων που προκύπτουν κατά την ομαδοποίηση ονομάζονται οπτικές λέξεις (visual words). Αυτές οι οπτικές λέξεις σχηματίζουν ένα μεγάλο λεξικό, όπου μέσω αυτού, για την κάθε εικόνα δημιουργείται ένα ιστόγραμμα χαρακτηριστικών (feature histogram), το οποίο δείχνει τη συχνότητα εμφάνισης των οπτικών αυτών λέξεων στην εικόνα [23]. Τέλος, η ταξινόμηση των εικόνων γίνεται με βάση αυτά τα ιστογράμματα χαρακτηριστικών που προέκυψαν. Όλη αυτή η διαδικασία εξετάζεται λεπτομερώς στην ενότητα 3.3 όπου παρουσιάζονται τα βήματα του αλγορίθμου. Στην εικόνα που ακολουθεί (Εικόνα 4), μπορεί να δει κανείς την εφαρμογή της μεθόδου bag of words.



Εικόνα 4: Η λειτουργία της μεθόδου Bag of Words. Αριστερά βρίσκεται η αρχική εικόνα, ενώ στα δεξιά, βρίσκονται τα χαρακτηριστικά που έχουν εξαχθεί από αυτήν (Πηγή: https://prateekvjoshi.com/2014/08/17/image-classification-using-bag-of-words-model/)

Στην περίπτωση του μεγάλου όγκου εικόνων που συλλέγεται μέσω της ασύρματης ενδοσκόπησης, η λειτουργία της μεθόδου Bag of Words έχει ως εξής: Αρχικά, οι εικόνες πρέπει να ταξινομηθούν σε υγιείς ή μη υγιείς, σύμφωνα με τη ύπαρξη η μή, κάποιων ανωμαλιών σε αυτές. Οι περιοχές ενδιαφέροντος που ανιχνεύονται στις εικόνες, είναι οι πιθανές περιοχές ύπαρξης κάποιας ανωμαλιάς, η οποία παραπέμπει σε κάποια ασθένεια. Από τις περιοχές αυτές, εξάγονται τα χαρακτηριστικά που ομαδοποιούνται και δημιουργούν το λεξικό. Τέλος, οι εικόνες ταξινομούνται στις 2 προαναφερθείσες κατηγορίες, σύμφωνα με τα ιστογράμματα χαρακτηριστικών τους, που δημιουργούνται με βάση το λεξικό αυτό. Η διαδικασία αυτή, γίνεται να πραγματοποιηθεί και στην περίπτωση που ο γιατρός επιθυμεί να ταξινομήσει τις εικόνες, ανάλογα με την ασθένεια στην οποία παραπέμπουν. Αυτή η ταξινόμηση ωστόσο, θα πραγματοποιηθεί σε περισσότερες από 2 κατηγορίες.

2.3 Η μέθοδος εξαγωγής χαρακτηριστικών SIFT (Scale Invariant Feature Transform)

Στο αρχικό στάδιο της μεθόδου Bag of Words, ανιχνεύονται και εξάγονται χαρακτηριστικά από τις περιοχές ενδιαφέροντος των εικόνων. Ο τρόπος με τον οπόιο πραγματοποιείται η ανίχνευση και η εξαγώγη χαρακτηριστικών δεν είναι πάντα ο ίδιος, καθώς εξαρτάται από τη μέθοδο που χρησιμοποιείται. Με μία γρήγορη αναζήτηση στην βιβλιογραφία, μπορεί κάποιος να αντιληφθεί, ότι υπάρχει ένας μεγάλος αριθμος τέτοιων μεθόδων.Οι δύο πιο διαδεδομένες είναι η SIFT(Scale Invariant Feature Transform) και η SURF (Speed Up Robust Features). Αξίζει να σημειωθεί ότι η επιλογή της μεθόδου εξαγωγής χαρακτηριστικών, επηρεάζει σε σημαντικό βαθμό την απόδοση της ταξινόμησης. Σε αυτήν την ενότητα εξετάζεται ο τρόπος λειτουργίας της SIFT μεθόδου.

Όπως είναι ήδη γνωστό, η αναγνώριση ανωμαλιών σε εικόνες ασύρματης ενδοσκόπησης, είναι υψίστης σημασίας για τη σωστή ταξινόμηση των εικόνων και κατ'επέκταση τη σωστή διάγνωση. Για να γίνει σωστά η ταξινόμηση μίας εικόνας με βάση τις ανωμαλίες που πιθανώς να περιέχει, θα πρέπει τα χαρακτηριστικά που εξάγονται και με βάση τα οποία γίνεται η ταξινόμηση να είναι «καλά». Ως καλά χαρακτηριστικά ορίζονται τα χαρακτηριστικά που πληρούν τις εξής προϋποθέσεις [21]:

- Δεν μεταβάλλονται με την κλίμακα (Scale invariant)
- Δεν μεταβάλλονται με την περιστροφή (Rotation invariant)
- Δεν επηρεάζονται από τις αλλαγές της φωτεινότητας (Robust to illumination changes)
- Δεν επηρεάζονται από τις αλλαγές τις οπτικής γωνίας (Robust to viewpoint changes)

Οι περισσότεροι ανιχνευτές όπως ο Harris που χρησιμοποιείται στην ανίχνευση γωνιών, δεν πληρούν όλα τα παραπάνω κριτήρια [20]. Η μέθοδος ανίχνευσης SIFT όμως που προτάθηκε το 2004 από τον David Lowe, τα ικανοποιεί. Αυτό σημαίνει όταν αν εξαχθούν χαρακτηριστικά από μία ανωμαλία με την SIFT μέθοδο, τότε τα ίδια χαρακτηριστικά θα εξάγονταν από την ίδια ανωμαλία σε άλλη φωτογραφία, ακόμα και αν αυτή ήταν πιο μικρή (διαφορετική κλίμακα) και απεικονιζόταν από διαφορετική οπτική γωνία (διαφορετική κατεύθυνση). Συνεπώς, γίνεται κατανοητή η σημαντικότητα της μεθόδου SIFT τόσο στην ταξινόμηση, όσο και στην σύγκριση εικόνων. Τα βήματα του αλγορίθμου SIFT είναι [20]:

1. Ανίχνευση τοπικών ακρότατων στην χωροκλίμακα (Scale space extrema detection)

Αρχικά, ο αλγόριθμος ανιχνεύει τα υπόψηφια σημεία τοπικών ακρότατων στη χωροκλίμακα (scale space). Η χωροκλίμακα αφορά την απεικόνιση της ίδιας εικόνας σε διαφορετικές κλιμακες. Με αυτόν τον τρόπο, είναι δυνατόν να βρεθούν υποψήφια τοπικά ακρότατα σε πολλές κλίμακες, πράγμα που κάνει τα χαρακτηριστικά που προκύπτουν αμετάβλητα με την κλίμακα (scale invariant) [20].

Πρώτα υπολογίζεται η συνάρτηση χωροκλίμακας της εικόνας $\boldsymbol{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma})$, η οποία αποτελεί αποτέλεσμα της συνέλιξης της αρχικής εικόνας $\boldsymbol{I}(\boldsymbol{x},\boldsymbol{y})$με την Gaussian συνάρτηση $\boldsymbol{G}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma})$Δηλαδή:

$$\boldsymbol{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma}) = \boldsymbol{I}(\boldsymbol{x},\boldsymbol{y}) * \boldsymbol{G}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma}) \quad \textbf{(1.1)}$$

Στην συνέχεια υπολογίζεται η συνάρτηση Laplacian of Gaussian (LoG), η οποία είναι :

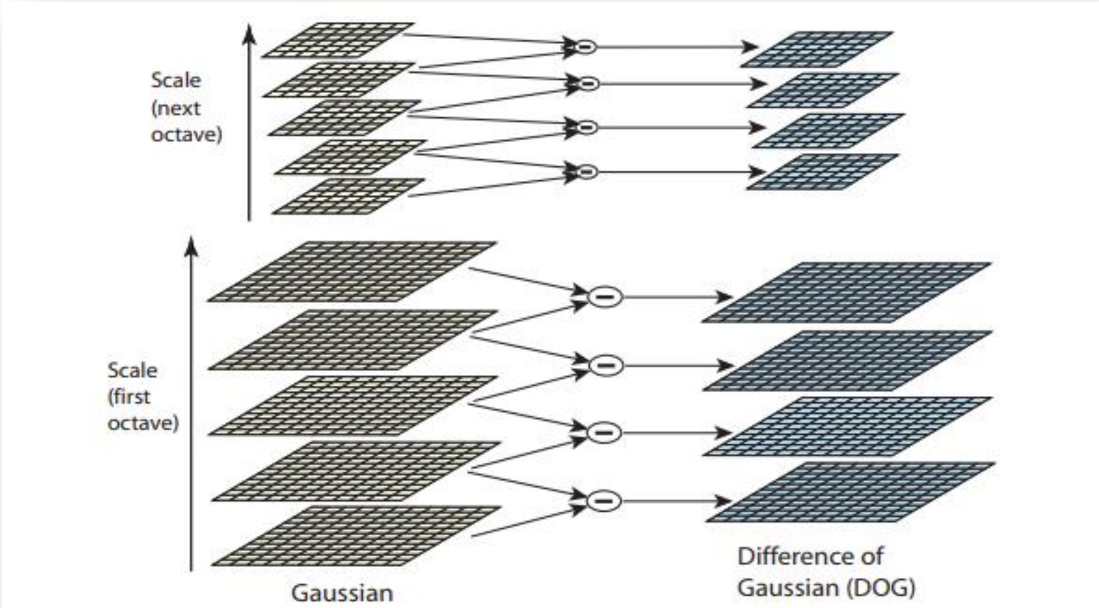
$$\boldsymbol{\sigma}^2\boldsymbol{\nabla}^2\boldsymbol{G}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma}) \quad \textbf{(1.2)}$$

Στη συνάρτηση αυτήν ανιχνεύονται τοπικά ακρότατα για πολλές τιμές της μεταβλητής $\boldsymbol{\sigma}$ η οποία αποτελεί την παράμετρο που ρυθμίζει την κλίμακα. Με αυτόν τον τρόπο, βρίσκεται ένα σύνολο από υποψήφια τοπικά ακρότατα $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma})$, όπου $\boldsymbol{x},\boldsymbol{y}$ είναι η θέση του ακρότατου, ενώ $\boldsymbol{\sigma}$ η κλίμακα στην οποία βρέθηκε.

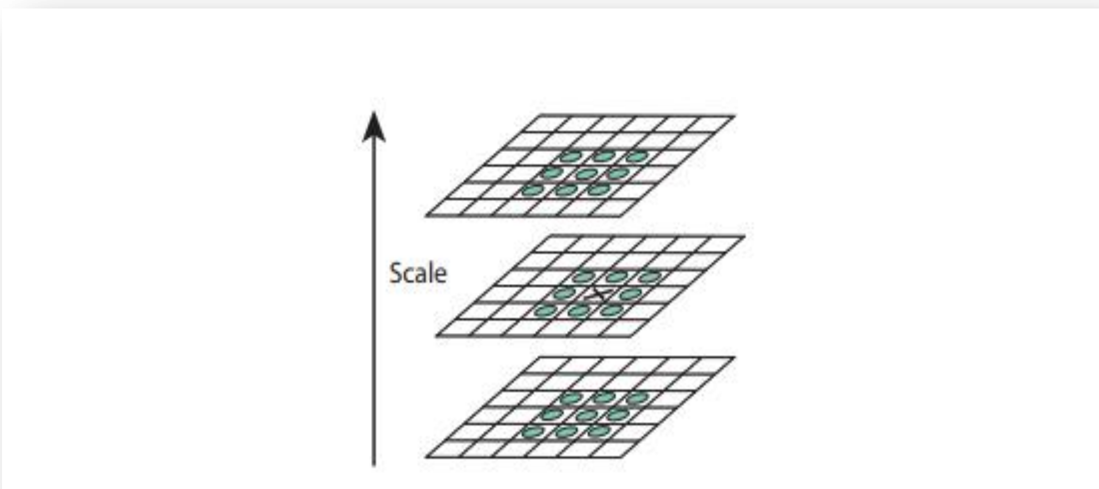
Επειδή όμως, αυτή η διαδικασία θεωρείται χρονοβόρα, για τον υπολογισμό των τοπικών ακρότατων, αξιοποιείται η συνάρτηση Difference of Gaussian (DoG) η οποία είναι [20], [21]:

$$\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma}) = \boldsymbol{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\kappa\sigma}) - \boldsymbol{L}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{\sigma}) \quad \textbf{(1.3)}$$

Η διαδικασία παραγωγής της DoG και η ανίχνευση πιθανών ακρότατων παρουσιάζονται στις εικόνες 5 και 6:



Εικόνα 5: Δημιουργία της συνάρτησης DOG σε διαφορετικές οκτάβες (Πηγή: https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html)



Εικόνα 6: Ανίχνευση πιθανού τοπικού ακρότατου (Πηγή: https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html)

2. Επιλογή σημείων ενδιαφέροντος (Keypoint localization)

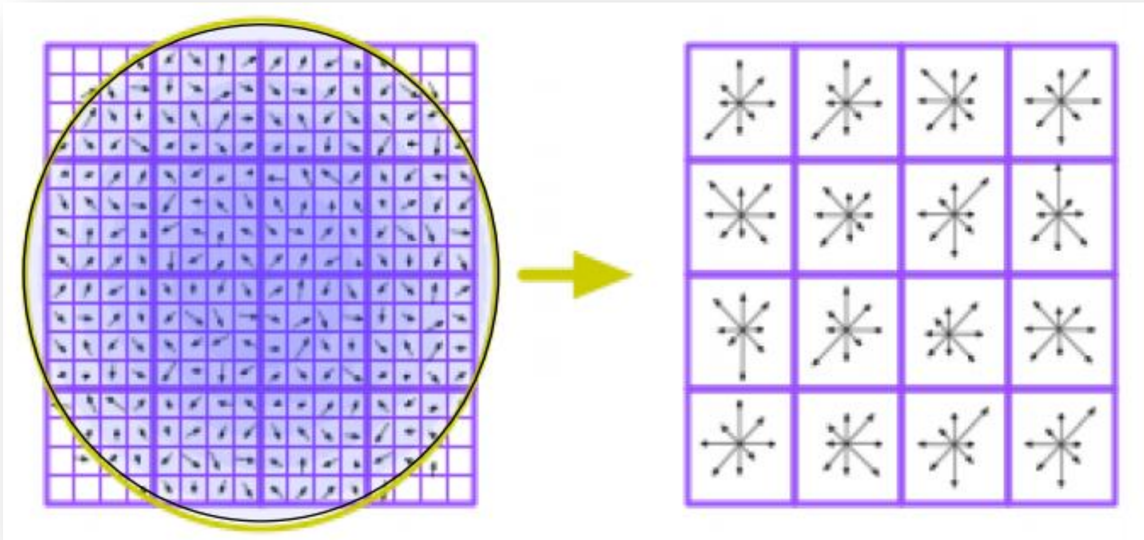
Στο προηγούμενο βήμα, ο αλγόριθμος SIFT ανίχνευσε όλα τα πιθανά ακρότατα. Αυτά τα τοπικά ακρότατα αποτελούν τα υποψήφια σημεία ενδιαφέροντος (Keypoints). Σε αυτό το βήμα συνεπώς, ο αλγόριθμος πραγματοποιεί ένα «ξεκαθάρισμα» των σημείων, έτσι ώστε να επιλεθχούν μόνο τα «δυνατά» σημεία μέσω των οποίων θα εξαχθούν τα χαρακτηριστικά στη συνέχεια. Μετά τη χρήση συγκεκριμένων μαθηματικών μεθόδων, τα σημεία που απορρίπτονται είναι αυτά που έχουν χαμηλή αντίθεση (Low-contrast) και οι άκρες (Edges) [20], [21].

3. Ανάθεση κατεύθυνσης (Orientation assignment)

Μετά τα δύο πρώτα βήματα εκτέλεσης του αλγορίθμου, έχουν αποκτηθεί πλέον τα επιθυμητά σημεία ενδιαφέροντος. Το κάθε σημείο χαρακτηρίζεται από τρείς μεταβλητές $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\sigma})$, δηλαδή την τοποθεσία του $(\boldsymbol{x}, \boldsymbol{y})$ και την κλίμακα του $\boldsymbol{\sigma}$. Σε αυτό το βήμα ο αλγόριθμος αναθέτει σε αυτά τα σημεία ενδιαφέροντος και κάποια κατεύθυνση. Κατα συνέπεια, οι μεταβλητές που χαρακτηρίζουν τα σημεία ενδιαφέροντος από τρείς πλέον γίνονται τέσσερις. Αυτές είναι οι $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\sigma}, \boldsymbol{orientation})$. Με αυτόν τον τρόπο εξασφαλίζεται ότι τα χαρακτηριστικά που θα προκύψουν στο τέλος θα ειναι αμετάβλητα και στην περιστροφή (Orientation invariant) [21].

4. Δημιουργία περιγραφέα (Keypoint descriptor)

Στο τελευταίο βήμα ο αλγόριθμος υπολογίζει τον περιγραφέα (descriptor**)** για κάθε σημείο ενδιαφέροντος. Ο περιγραφέας δημιουργείται με την επεξεργασία μίας περιοχής 16x16 γύρω από το σημείο ενδιαφέροντος. Αποτέλεσμα αυτής της επεξεργασίας είναι ένας πίνακας ιστογραμμάτων κατεύθυνσης 4x4 ο οποίος περιέχει 8 θέσεις κατευθύνσεων σε κάθε στοιχείο. Αυτό σημαίνει ότι ο περιγραφέας είναι ένα διάνυσμα 128 διαστάσεων (4*4*8 = 128) [20]. Τέλος αυτό το διάνυσμα τροποιείται έτσι ώστε να μειωθούν οι επιπτώσεις των αλλαγών φωτισμού (illumination changes). Στο δεξί σκέλος της εικόνας 7 διακρίνεται ένας περιγραφέας.



Εικόνα 7: Εξαγωγή περιγραφέα από μία περιοχή γύρω από το σημείο ενδιαφέροντος (Πηγή: https://www.researchgate.net/publication/267638940_Image_Stitching_based_on_Feature_Extraction_Techniques_A_Survey/figures?lo=1&utm_source=google&utm_medium=organic)

2.4 Χρωματικές διαφοροποιήσεις της μεθόδου SIFT

Παρά τα καλά του αποτελέσματα, ο αλγόριθμος SIFT, πέραν του μεγάλου χρόνου εκτέλεσης, παρουσιάζει και δύο ακόμη αρνητικά χαρακτηριστικά, τα οποία είναι άξια αναφοράς. Αρχικά εφαρμόζεται μόνο σε εικόνες τόνων του γκρί. Έτσι, πριν την εφαρμόγη του, μετατρέπει όλες τις εικόνες του σετ, σε εικόνες τόνων του γκρί, ανεξάρτητα του χώρου χρωμάτων (Color space) στον οποίο αυτές ανήκουν [16]. Ως εκ τούτου, ο παράγοντας του χρώματος δεν λαμβάνεται καθόλου υπόψιν κατά την ταξινόμηση των εικόνων. Αυτό όμως, επηρεάζει σε μεγάλο βαθμό την απόδοση της ταξινόμησης, καθώς το χρώμα διαδραματίζει σημαντικό ρόλο σε αυτήν, και ειδικά στην περίπτωση της αναγνώρισης ανωμαλιών σε εικόνες ασύρματης ενδοσκόπησης [17]. Το δεύτερο σημαντικό αρνητικό της μεθόδου SIFT, είναι το γεγονός πως η απόδοση της, επηρεάζεται αρνητικά από τις αλλαγές τις φωτεινότητας [18]. Αυτό σημαίνει πως όταν σε ένα σετ εικόνων υπάρχουν αλλαγές της φωτεινότητας, τότε η απόδοση της ταξινόμησης θα παρουσιάζει σημαντικές διακυμάνσεις.

Αυτά τα δύο αρνητικά χαρακτηριστικά του αλγόριθμου SIFT, μπορούν να ξεπεραστούν πλήρως, αν προσθέσουμε στο αλγόριθμό την επιπλέον πληροφορία που προσφέρει το χρώμα. Στην βιβλιογραφία, έχουν προταθεί αρκετές διαφοροποιήσεις της μεθόδου SIFT που συμπεριλαμβάνουν τη χρωματική πληροφορία. Σε αυτήν την ενότητα αναφέρονται τρείς από αυτές:

- **RGB-SIFT:** Όπως είναι γνωστό, οι RGB εικόνες αποτελούνται από τρία κανάλια: R(Red), G(Green) και B(Blue).Η λογική είναι ότι μέσω αυτών των τριών χρωμάτων, μπορούν να απεικονιστούν όλα τα χρώματα.Ο αλγόριθμος RGB-SIFT εφαρμόζεται και στα τρία κανάλια των εικόνων ανεξάρτητα, παράγοντας έτσι έναν περιγραφέα μεγέθους 128 διαστάσεων για κάθε κανάλι [18].Το μέγεθος του τελικού περιγραφέα που προκύπτει είναι ένα διάνυσμα 3*128 = 384 διαστάστάσεων [18], [19].
- **HSV-SIFT:** Οι εικόνες που ανήκουν στο χρωματικό χώρο HSV, αποτελούνται και αυτές από τρία κανάλια: H(Hue), S(Saturation) και V(Value). Η λογική είναι παρόμοια με αυτήν του RGB-SIFT. Ο HSV-SIFT, εφαρμόζεται και στα τρία κανάλια των HSV εικόνων ανεξάρτητα, παράγοντας έναν περιγραφέα 128 διαστάσεων για κάθε κανάλι [18], [19]. Και σε αυτήν την περίπτωση το μέγεθος του τελικού περιγραφέα είναι ένα διάνυσμα: 3*128 = 384 διαστάσεων [18], [19].
- **Opponent-SIFT:** Οι εικόνες που ανήκουν στο χρωματικό χώρο Opponent απότελούνται από τρία κανάλια:

(1) $$\boldsymbol{o_1} = \frac{\boldsymbol{R-G}}{\sqrt{2}} \quad \textbf{(1.4)}$$

(2) $$\boldsymbol{o_2} = \frac{\boldsymbol{R+G-2B}}{\sqrt{6}} \quad \textbf{(1.5)}$$

(3) $$\boldsymbol{o_3} = \frac{\boldsymbol{R+G+B}}{\sqrt{3}} \quad \textbf{(1.6)}$$

Τα δύο πρώτα κανάλια αναφέρονται στην πληροφορία του χρώματος, ενώ το τρίτο κανάλι, καθορίζει την ένταση του χρώματος [18], [19]. Η διαδικασία του Opponent-SIFT είναι ακριβώς ίδια με τις προηγούμενες δύο περιπτώσεις. Και σε αυτήν την περίπτωση λοιπόν, ο περιγραφέας που προκύπτει, είναι ένα διάνυσμα 3*128 = 384 διαστάσεων [18].

3. ΠΕΙΡΑΜΑΤΙΚΗ ΔΙΑΔΙΚΑΣΙΑ

3.1 Εισαγωγή

Στο πρώτο κεφάλαιο, αναφέρθηκαν οι βιολογικές γνώσεις που διέπουν το ανθρώπινο πεπτικό σύστημα, ενώ παράλληλα αναλύθηκε και εξέταση της ασύρματης ενδοσκόπησης, ως μία μέθοδος διάγνωσης των ασθενειών του συστήματος αυτού. Αναφέρθηκε επίσης, πως ο μεγάλος όγκος εικόνων που προκύπτει απο την ασύρματη ενδοσκόπηση, δημιουργεί προβλήματα σχετικά με την έγκαιρη και έγκυρη διάγνωση της κατάστασης των ασθενων. Τέλος, τέθηκαν κάποια ερωτήματα, αναφορικά με την ανάπτυξη διάφορων λογισμικών ταξινόμησης των εικόνων αυτών σε υγιείς και μη, έτσι ώστε αυτά να έχουν τη δυνατότητα να αντικαταστήσουν τον ανθρώπινο παράγοντα στη διάγνωση, ο οποίος απαιτεί περισσότερο χρόνο και ατελείωτη προσοχή.

Στο δεύτερο κεφάλαιο, αναπτύχθηκαν τα θεωρητικά εργαλεία, σύμφωνα με τα οποία μπορεί να συνθεθεί ένα λογισμικό ταξινόμησης εικόνων ασύρματης ενδοσκόπησης. Προφανώς, η ταξινόμηση θα γίνεται με βάση την ύπαρξη ανωμαλιών στις εικόνες αυτές. Οι εικόνες που περιέχουν ανωμαλίες, ταξινομούνται ως μη υγιείς καθώς έχουν αυξημένη πιθανότητα να παραπέμψουν σε κάποια ασθένεια, ενώ οι υπόλοιπες ταξινομούνται ως υγιείς. Τα εργαλεία που εξετάστηκαν, ήταν η μέθοδος ταξινόμησης Bag of Words, καθώς και η διαδεδομένη μέθοδος εξαγωγής χαρακτηριστικών από εικόνες SIFT. Στο τέλος πραγματοποιήθηκε και μία απλή αναφορά στις χρωματικές διαφοροποιήσεις της μεθόδου SIFT.

Στον παρόν κεφάλαιο, παρουσιάζεται η πειραματική διαδικασία. Δηλαδή, ο τρόπος με τον οποίο τα θεωρητικά εργαλεία που αναπτύχθηκαν στο δεύτερο κεφάλαιο, εφαρμόζονται πάνω σε πραγματικά δεδομένα εικόνων ασύρματης ενδοσκόπησης. Τα δεδομένα που χρησιμοποιήθηκαν σ’αυτήν την πειραματική διαδικασία, αποτελούν το KID dataset 2, το οποίο αναλύεται στην αρχή του κεφαλαίου αυτού. Έπειτα παρουσιάζονται αναλυτικά, τα βήματα του αλγορίθμου που χρησιμοποιήθηκε για την επεξεργασία των δεδομένων. Στη συνέχεια παρουσιάζονται τα αποτελέσματα των πειραμάτων, μέσω του εκτενούς σχολιασμού τους. Τέλος αναγράφονται τα γενικά συμπεράσματα, ενώ παράλληλα γίνεται και μία αναφορά στις μελλοντικές επεκτάσεις των πειραμάτων.

3.2 Το σέτ δεδομένων KID dataset 2

Το KID dataset 2**,** αποτελεί μέρος μίας ευρύτερης βάσης δεδομένων, που ονομάζεται KID dataset [25]. Τα δεδομένα αυτά, αποτελούνται από υψηλής ποιότητας εικόνες και βίντεο, τα οποία έχουν παρθεί μέσω ασύρματης ενδοσκόπησης από ασθενείς. Οι εικόνες είναι ταξινομημένες σε κατηγορίες ανάλογα με την ασθένεια.. Η βάση δεδομένων του KID dataset, είναι μη κερδοσπική και ανοιχτής πρόσβασης, και δημιουργήθηκε κυρίως για ερευνητικούς σκοπούς. Στόχος της, είναι η παροχή γνώσης και εκπαιδευσης σε νέους επιστήμονες, αλλά και η συμβολή της στην δημιουργία αυτόματων συστημάτων ιατρικών αποφάσεων.

Στα πειράματα που ακολουθούν, χρησιμοποιήθηκε το KID dataset 2, το οποίο αποτελείται από 2371 εικόνες, ταξινομημένες σε 8 κατηγορίες. Οι 4 από αυτές, είναι υγιείς κατηγορίες χωρίς εμφανείς ανωμαλίες, ενώ οι άλλες 4 περιέχουν ανωμαλίες που παραπέμπουν σε συγκεκριμένες ασθένειες. Οι 4 υγιεις κατηγορίες είναι:

- **Normal small bowel :** 728 εικόνες
- **Normal colon :** 169 εικόνες
- **Normal esophagus :** 282 εικόνες
- **Normal stomach :** 599 εικόνες

Οι 4 μη υγιείς κατηγορίες είναι:

- **Ampulla of vater :** 19 εικόνες
- **Inflammatory :** 227 εικόνες
- **Polypoids :** 44 εικόνες
- **Vascular :** 303 εικόνες

Επίσης, χρησιμοποιήθηκε και μία παραλλαγή του KID dataset 2, κατα την οποία οι 4 υγιείς κατηγορίες τοποθετούνται σε μία ενιαία κατηγορια που ονομάζεται «Normal», ενώ οι 4 μη υγιείς κατηγορίες τοποθετούνται σε άλλη μία ενιαία κατηγορία που ονομάζεται «Abnormal». Έτσι τα πειράματα διεξάγονται πάνω σε 2 κατηγορίες:

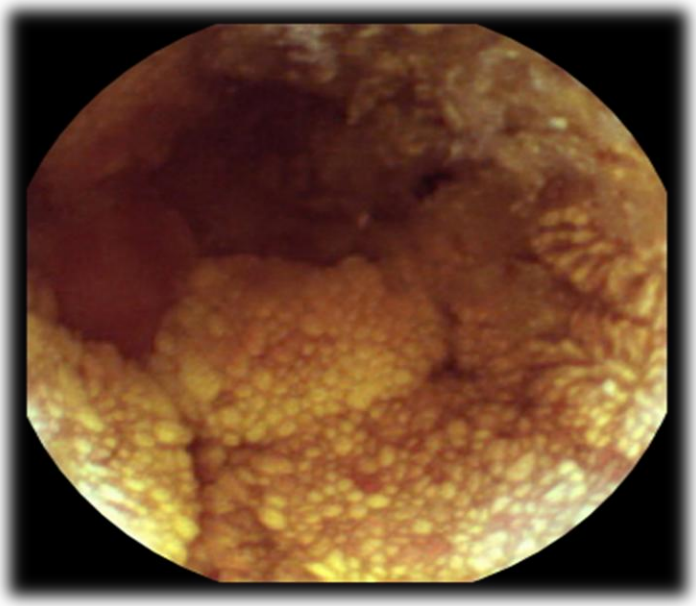
- **Normal :** 1178 εικόνες
- **Abnormal :** 593 εικόνες

Αυτή η παραλλαγή του KID dataset 2 ονομάστηκε Normal-Abnormal set.

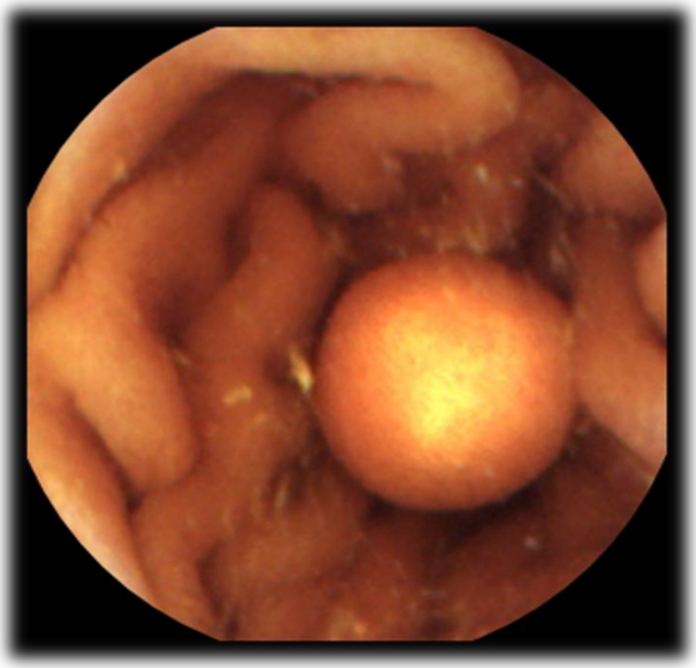
Στην συνέχεια παρατίθενται 8 φωτογραφίες από το KID dataset 2 (μία από κάθε κατηγορία), έτσι ώστε ο αναγνώστης να έχει τη δυνατότητα να κατανοήσει καλύτερα τη φύση των δεδομένων.

Εικόνα 8: Αυτή η εικόνα ανήκει στην κατηγορία Ampulla of Vater και απεικονίζει μία αλλοίωση της ηπατοπαγκρεατικής λήκυθου

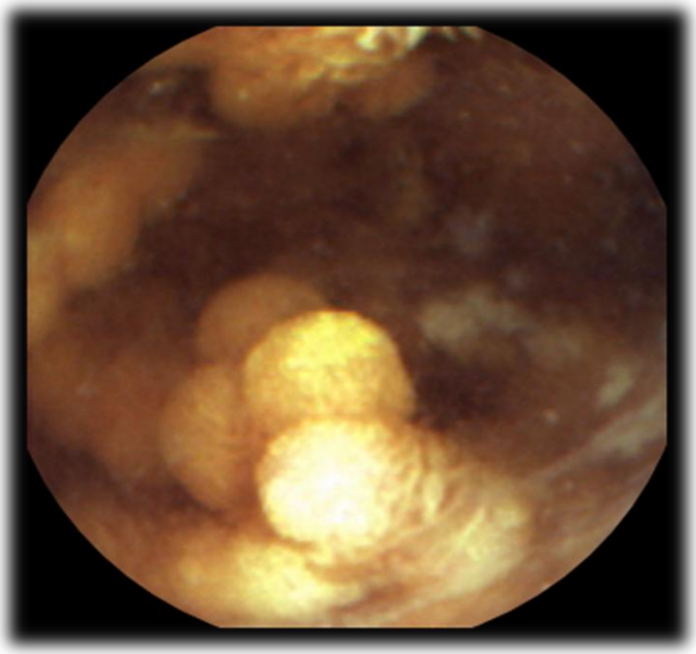
Εικόνα 9: Αυτή η εικόνα ανήκει στην κατηγορία Inflammatory και απεικονίζει μία φλεγμονή



Εικόνα 10: Αυτη η εικόνα ανήκει επίσης στην κατηγορία Inflammatory και απεικονίζει έναν διαφορετικό τύπο φλεγμονής από αυτόν της προηγούμενης εικόνας

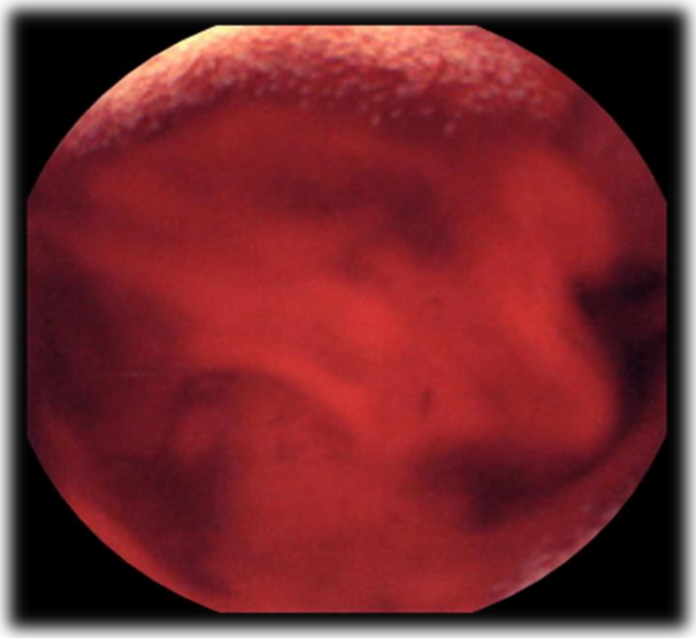


Εικόνα 11: Αυτή η εικόνα ανήκει στην κατηγορία Polypoids και απεικονιζει έναν Πολύποδα

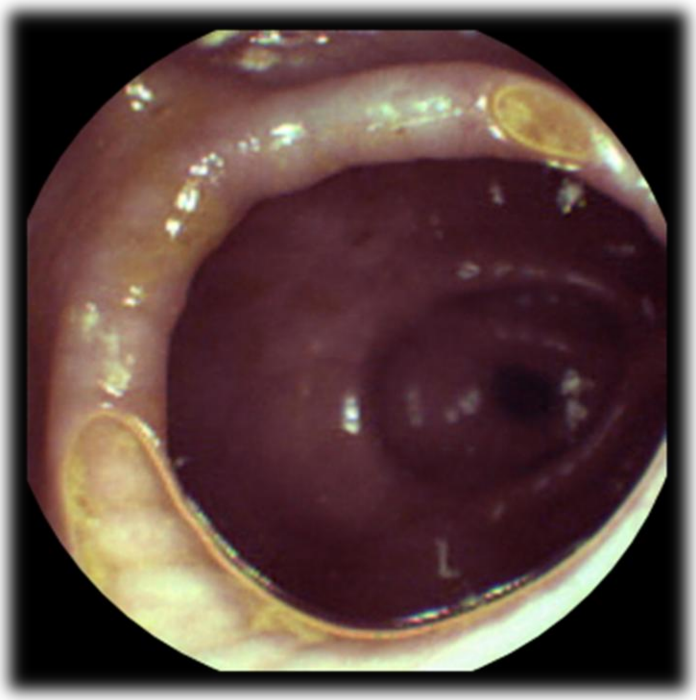


Εικόνα 12: Αυτή η εικόνα ανήκει επίσης στην κατηγορία Polypoids και απεικονίζει έναν Πολύποδα διαφορετικής μορφολογίας από τον προηγούμενο

Εικόνα 13: Αυτή η εικόνα ανήκει στην κατηγορία Vascular και απεικονίζει μία αγγειακή αλλοίωση (πάνω δεξιά)

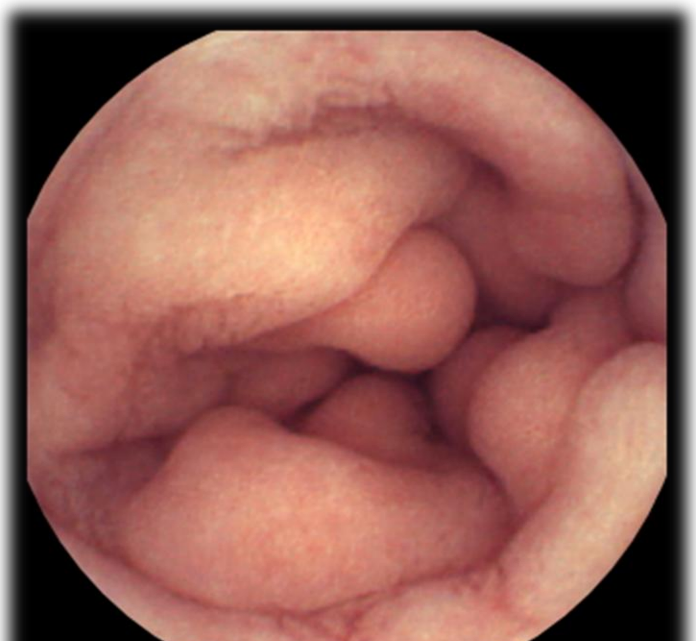


Εικόνα 14: Αυτή η εικόνα ανήκει επίσης στην κατηγορία Vascular και απεικονίζει μία εντονότερη αγγειακή αλλοίωση



Εικόνα 15: Αυτή η εικόνα ανήκει στην κατηγορία Normal-colon και απεικονίζει τμήμα του παχέος εντέρου, στο οποίο δεν παρουσιάζονται ανωμαλίες

Εικόνα 16: Αυτή η εικόνα ανήκει στην κατηγορία Normal-esophagus και απεικονίζει τμήμα του εισοφάγου, στο οποίο δεν παρουσιάζονται ανωμαλίες



Εικόνα 17: Αυτή η εικόνα ανήκει στην κατηγορία Normal-small-bowel και απεικονίζει τμήμα του λεπτού εντέρου, στο οποίο δεν παρουσιάζονται ανωμαλίες

Εικόνα 18: Αυτή η εικόνα ανήκει στην κατηγορία Normal-stomach και απεικονίζει τμήμα του στομάχου, στο οποίο δεν παρουσιάζονται ανωμαλίες

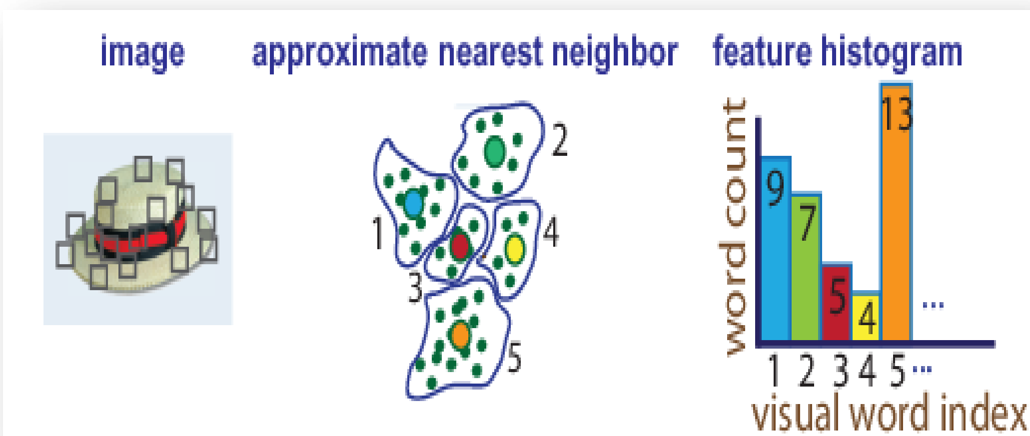
3.3 Βήματα αλγορίθμου ταξινόμησης

Καλό θα ήταν, πριν την αναφορά των αποτελεσμάτων των πειραμάτων, να πραγματοποιηθεί και μία εκτενής παρουσίαση των βημάτων του αλγορίθμου που χρησιμοποιήθηκε. Όπως επισημάνθηκε στο δεύτερο κεφάλαιο, ο αλγόριθμος που χρησιμοποιείται είναι η μέθοδος Bag of Words. Η μόνη παραλλαγή του αλγορίθμου, είναι η μέθοδος εξαγωγής χαρακτηριστικών, η οποία διαφέρει από πείραμα σε πείραμα, όπως θα τονιστεί και στις επόμενες ενότητες.

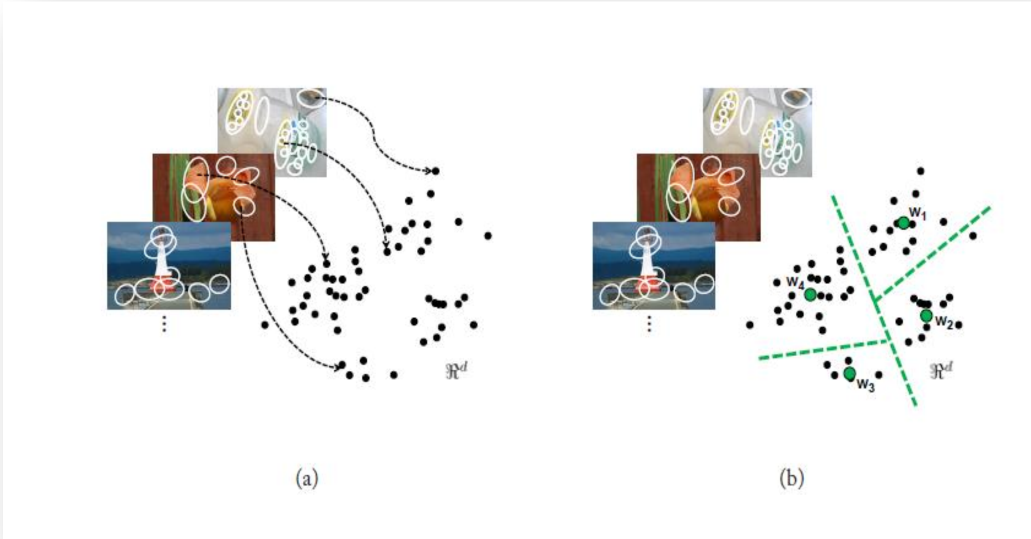
Πριν την εκτέλεση του αλγορίθμου, είναι πολύ σημαντικό να ρυθμίστουν οι παράμετροι του, έτσι ωστε τα δεδομένα του KID dataset 2 να διαχωριστούν σε δύο σετ. Ένα σετ εκπαίδευσης που θα περιέχει το 70% των εικόνων, και ένα σετ ελέγχου που θα περιέχει το υπόλοιπό 30%. Τα βήματα στη συνέχεια είναι:

1. **Εξαγωγή χαρακτηριστικών από τις εικόνες:** Αρχικά,μέσω μίας μεθόδου εξαγωγής χαρακτηριστικών (SIFT, RGB-SIFT,κ.λπ) ο αλγόριθμος εξάγει χαρακτηριστικά από τις εικόνες του σετ εκπαίδευσης. Με αυτόν τον τρόπο πλέον, όλες οι εικόνες εκπροσωπούνται από τα διανύσματα των χαρακτηριστικών (feature vectors – descriptors) τους.
2. **Ομαδοποίηση χαρακτηριστικών ή αλλιώς δημιουργία λεξικού:** Στη συνέχεια, τα χαρακτηριστικά που εξήχθησαν από τις εικόνες του σετ εκπαίδευσης, ομαδοποιούνται με τη βοήθεια της μεθόδου ομαδοποίησης k-means. Έτσι, τα χαρακτηριστικά πλέον είναι ομαδοποιημένα σε k ομάδες, όπου κάθε ομάδα αντιπροσωπεύεται από το κέντρο της. Τα κέντρα των ομάδων που προκύπτουν από την ομαδοποίηση ονομάζονται οπτικές λέξεις (visual words), γι’αυτό και η διαδικασία ομαδοποίησης μπορεί να ονομαστεί αλλιώς και δημιουργία λεξικού.
3. **Δημιουργία ιστογραμμάτων:** Σε αυτό το βήμα, ο αλγόριθμος εξάγει τα ιστογράμματα των χαρακτηριστικών των εικόνων (Feature histogram). Η εξαγωγή των ιστογραμμάτων αυτών, γίνεται με βάση τα ομαδοποιημένα χαρακτηριστικά τους στις k ομάδες (δηλαδή με βάση το λεξιλόγιο). Για παράδειγμα, αν ο αριθμός των ομάδων (δηλαδή το μέγεθος του λεξικού) είναι k = 4 (ομάδες w_1, w_2, w_3, w_4) και ο αριθμός των χαρακτηριστικών που εξήχθησαν από μία εικόνα ειναι 15, εκ των οποίων τα 5 ειναι στην ομάδα w_1, τα 4 στην ομάδα w_2, τα 2 στην ομάδα w_3 και τα 4 στην w_4, τότε το ιστόγραμμα χαρακτηριστικών της εικόνας, θα αποτελείται από 4 ράβδους (μία για κάθε ομάδα δηλάδή), όπου η ράβδος w_1 θα έχει ύψος 5, η ράβδος w_2 θα έχει ύψος 4, η w_3 ράβδος θα έχει ύψος 2 και η w_4 θα έχει ύψος 4. Γίνεται κατανοητό λοιπόν, πως στο τέλος του 3^ου^ βήματος κάθε εικόνα αντιπροσωπεύεται πλέον από το ιστόγραμμα των χαρακτηριστικών της. Στην εικόνα 19, διακρίνεται η εξαγωγή ενός ιστογράμματος χαρακτηριστικών από μία εικόνα. Από αυτό το παράδειγμα λοιπόν, μπορεί να καταλάβει κάποιος από το ιστόγραμμα, πως τα χαρακτηριστικά της εικόνας ομαδοποιήθηκαν σε 5 ομάδες. Από τις 5 ομάδες, τα 9 ανήκουν στην πρώτη ομάδα, τα 7 στη δεύτερη κ.ο.κ.

4. **Ταξινόμηση εικόνων και αξιολόγηση αποτελεσμάτων:** Στο βήμα αυτό πραγματοποιείται η ταξινόμηση των εικόνων στις κατηγορίες, ενώ παράλληλα ελέγχεται και το ποσοστό της ορθής ταξινόμησης στο σετ εκπαίδευσης. Ερευνητικά, το ποσοστό επιτυχίας στο σετ εκπαίδευσης δεν παρουσιάζει ιδιαίτερο ενδιαφέρον. Ενδιαφέρον όμως, παρουσιάζει το ποσοστό ορθής ταξινόμησης στο σετ ελέγχου, καθώς είναι πιο κοντά στα άγνωστα δεδομένα. Το ποσοστό ορθής ταξινόμησης υπολογίζεται με τον πίνακα αληθείας (confusion matrix).
5. **Επανάληψη των βημάτων 1,3,4 για το σετ ελέγχου:** Η πραγματοποίηση των 4 πρώτων βημάτων του αλγορίθμου για το σέτ εκπαίδευσης, ονομάζεται φάση εκπαίδευσης. Στη φάση εκπαίδευσης, ο ταξινομητής «μαθαίνει» τους κανόνες ταξινόμησης, μέσω της δημιουργίας του λεξικού. Γι' αυτό το λόγο, όταν επαναλαμβλανεται η εκτέλεση του αλγορίθμου για το σέτ ελέγχου δεν επαναλαμβάνεται το βήμα 2. Δηλαδή, τα ιστογράμματα στις εικόνες του σετ ελέγχου, εξάγονται με βάση το λεξικό που δημιουργήθηκε στη φάση εκπαίδευσης. Γιατί η ταξινόμηση του σετ ελέγχου, πρέπει να πραγματοποιηθεί σύμφωνα με τους κανόνες που «έμαθε» ο ταξινομητής στη φάση εκπαίδευσης. Τέλος, η απόδοση του ταξινομητή κρίνεται από το ποσοστό ορθής ταξινόμησης στο σετ ελέγχου.



Εικόνα 19: Δημιουργία ιστογράμματος από μία εικόνα. Στα δεξιά βρίσκεται η εικόνα στην οποία ανιχνευθηκαν οι περιοχές ενδιαφέροντος. Στη μέση είναι η ομαδοποίηση των χαρακτηριστικών της εικόνας με βάση το λεξικό και στα αριστερά φαίνεται το ιστόγραμμα που εξάγεται τελικώς από την εικόνα αυτή. (Πηγή: https://www.mathworks.com/help/vision/ug/image-classification-with-bag-of-visual-words.html)



Εικόνα 20 : Εξαγωγή χαρακτηριστικών και δημιουργία λεξικού (Πηγή: Kristen Grauman, Bastian Liebe (2011). 'Visual Object Recognition'. Morgan and Claypool Publishers)

3.4 Αποτελέσματα πειραμάτων

3.4.1 Αποτελέσματα πειραμάτων για το KID dataset 2

Τα αποτελέσματα που ακολουθούν, αφορούν τα πειράματα που πραγματοποιήθηκαν πάνω στο KID dataset 2, χωρίς αυτό να έχει υποστεί κάποια παραλλαγή. Αποτελούν δηλαδή πειράματα ταξινόμησης σε 8 κατηγορίες. Η κατηγορία με το μικρότερο αριθμό εικόνων, είναι η Ampulla of vater με 19 εικόνες. Γι'αυτό το λόγο κρατήθηκαν 19 εικόνες από κάθε κατηγορία. Από αυτές τις 19 εικόνες της κάθε κατηγορίας, οι 13 χρησιμοποιήθηκαν για τη φάση εκπαίδευσης, ενώ οι υπόλοιπες 6 για τη φάση ελέγχου. Πιο αναλυτικά, στα πειράματα χρησιμοποιήθηκαν συνολικά 152 εικόνες (8*19). Από αυτές τις 152 εικόνες,οι 104 (8*13), δηλαδή το 70% του συνολικού αριθμού, χρησιμοποιήθηκαν στη φάση εκπαίδευσης, ενώ οι υπόλοιπες 48 (8*6), δηλαδή το υπόλοιπο 30% χρησιμοποποιήθηκαν στη φάση ελέγχου.

Εκτός από τη μέθοδο εξαγωγής χαρακτηριστικών που αλλάζει από πείραμα σε πείραμα, μέσα στα ίδια τα πειράματα μεταβάλλονται επίσης και άλλοι δύο παράμετροι. Η πρώτη παράμετρος, είναι το μέγεθος του λεξικού και στον κώδικα ονομάζεται numWords. Η δεύτερη, είναι μία παράμετρος, η οποία καθορίζει τον αριθμό των εικόνων του σετ εκπαίδευσης, από τις οποίες εξάγονται τα χαρακτηριστικά για τη δημιουργία του λεξικού. Στο κώδικα η παράμετρος αυτή έχει το όνομα selTrainFeats. Προφανώς, σημαντικότερη επιρροή πάνω στα αποτελέσματα έχει η μέθοδος εξαγωγής χαρακτηριστικών που χρησιμοποιείται, ωστόσο η μεταβολή αυτών των δύο παραμέτρων, έχει σκοπό να διερευνήσει την πιθανή επιρροή τους πάνω στη μεταβολή των αποτελεσμάτων μέσα στην ίδια την μέθοδο.

3.4.1.1 Αποτελέσματα πειραμάτων με τη μέθοδο SIFT για το KID dataset 2

Στους παρακάτω πίνακες (Πίνακας 1 και Πίνακας 2), αναγράφονται τα αποτελέσματα των πειραμάτων του KID dataset 2, με τη χρήση της μεθόδου εξαγωγής χαρακτηριστικών SIFT, για τις διάφορες τιμές των παραμέτρων numWords και selTrainFeats.

Πίνακας 1: Αποτελέσματα SIFT για to KID dataset 2. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφεται η ελάχιστη τιμή, ενώ με πράσινο η μέγιστη

	selTrainFeats				
numWords	**104**	**90**	**80**	**70**	**60**
500	35.42	33.33	31.25	39.58	35.42
600	35.42	31.25	31.25	33.33	33.33
700	35.42	33.33	33.33	33.33	33.33
800	33.33	31.25	35.42	37.50	35.42
900	37.50	31.25	35.42	35.42	35.42
1000	33.33	31.25	29.17	35.42	33.33
1100	33.33	37.50	35.42	33.33	35.42
1200	33.33	35.42	37.50	37.50	33.33
1300	39.58	37.50	33.33	35.42	37.50
1400	33.33	31.25	37.50	33.33	39.58
1500	35.42	37.50	35.42	35.42	35.42
Mean value	**35.037**	**33.711**	**34.091**	**35.416**	**35.227**

Πίνακας 2: Αποτελέσματα SIFT για KID το dataset 2. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφεται η ελάχιστη τιμή, ενώ με πράσινο η μέγιστη

	selTrainFeats				
numWords	**50**	**40**	**30**	**20**	**10**
500	33.33	33.33	35.42	33.33	35.42
600	31.25	33.33	33.33	37.50	37.50
700	37.50	35.42	33.33	35.42	35.42
800	37.50	35.42	31.25	35.42	33.33
900	33.33	33.33	35.42	35.42	33.33
1000	35.42	31.25	33.33	33.33	35.42
1100	33.33	35.42	37.50	37.50	35.42
1200	39.58	37.50	33.33	35.42	33.33
1300	35.42	35.42	39.58	39.58	37.50
1400	33.33	37.50	37.50	33.33	35.42
1500	33.33	41.67	35.42	35.42	33.33
Mean value	**34.847**	**35.417**	**35.037**	**35.606**	**35.038**

- Η μέγιστη τιμή ποσοστού ορθής ταξινόμησης στο σέτ ελέγχου είναι 41.67% και παρουσιάζεται για το συνδυασμό μεταβλητών selTrainFeats = 40 και numWords = 1500. Δηλαδή, όταν εξάγονται χαρακτηριστικά από 40 εικόνες του σέτ εκπαίδευσης για τη δημιουργία λεξικού, ενώ το μέγεθος του λεξικού είναι 1500. Ο πίνακας αληθείας σε αυτήν την περίπτωση είναι:

Πίνακας 3: Πίνακας αληθείας για τη μέθοδο SIFT του KID dataset 2

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	4	1	0	0	0	1	0	0
Inflammatory	1	2	0	2	0	1	0	0
Normal-Colon	0	0	3	2	0	0	1	0
Normal esophagus	0	0	0	5	0	0	0	1
Normal-small-bowel	1	1	1	1	1	1	0	0
Normal-stomach	0	0	0	1	0	4	0	1
Polypoids	1	0	1	0	0	1	0	3
Vascular	0	0	0	1	3	0	1	1

Σκοπός του πίνακα αληθείας, είναι να δείξει λεπτομερώς, τον τρόπο ταξινόμησης των εικόνων του σέτ ελέγχου. Για παράδειγμα, από τον πίνακα 3 γίνεται κατανοήτο ότι από τις 6 εικόνες της κατηγορίας Ampulla of vater, οι 4 ταξινομούνται σωστά, ενώ οι υπόλοιπες 2 ταξινομούνται λάθος. Ουσιαστικά, τα στοιχεία που βρίσκονται στην κύρια διαγώνιο του πίνακα αληθείας, αντιστοιχούν στον αριθμό των εικόνων που ταξινομούνται σωστά από την κάθε κατηγορία. Κατα συνέπεια, το άθροισμα των στοιχείων της κύριας διαγωνίου, ισούται με το συνολικό αριθμό εικόνων που ταξινομήθηκε σωστά.

Άλλα σημαντικά στοιχεία που αντλούνται από τον πίνακα 3 είναι: Ο συνολικός αριθμός εικόνων που ταξινομήθηκε σωστά από το σέτ ελέγχου είναι 20/48 εικόνες. Ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες: Ampulla of vater με 4 σωστά ταξινομημένες εικόνες, Normal-esophagus με 5 σωστά ταξινομημένες εικόνες, Normal-stomach με 4 σωστά ταξινομημένες εικόνες και Normal-colon με 3 σωστά ταξινομημένες εικόνες. Από την άλλη, μη ικανοποιητική ταξινόμηση παρουσιάζουν οι κατηγορίες Vascular και Normal-small-bowel με μία σωστά ταξινομημένη εικόνα και η Polypoids με καμία σωστά ταξινομημένη εικόνα.

- Το χαμηλότερο ποσοστό ορθής ταξινόμησης στο σέτ ελέγχου, παρατηρήθηκε για το συνδυασμό των μεταβλητών selTrainFeats = 70 και numWords = 1000 και ισούται με 29.17%.
- Οι υπόλοιπες τιμές που παρατηρούνται είναι οι : 31.25%, 33.33%, 35.42%, 37.5% και 39.58%.
- Η τιμή της μεταβλητής selTrainFeats, για την οποία παρουσιάζονται τα καλύτερα κατα μέσο όρο αποτελέσματα, ισούται με 20 και έχει μέσο όρο 35.606. Ωστόσο όμως, δεν παρουσιάζονται σημαντικές διακυμάνσεις των αποτελεσμάτων με τη μεταβολή της μεταβλητής αυτής.
- Η μεταβολή του μεγέθους του λεξικού επίσης, δεν φαίνεται να επηρεάζει ιδαίτερα τα αποτελέσματα γι'αυτήν την μέθοδο.

3.4.1.2 Αποτελέσματα πειραμάτων με τη μέθοδο RGB-SIFT για το KID dataset 2

Τα αποτελέσματα που ακολουθούν, αφορούν τα πειράματα που διεξήχθησαν πάνω στο KID dataset 2 με τη μέθοδο εξαγωγής χαρακτηριστικών RGB-SIFT. Η διαφορά με την προηγούμενη μέθοδο, είναι ότι η RGB-SIFT, εκμεταλλεύεται την χρωματική πληροφορία που παρέχει ο χώρος χρωμάτων RGB. Με αυτόν τον τρόπο, κατά την ταξινόμηση των εικόνων, λαμβάνεται υπόψιν και η πληροφορία του RGB χρώματος. Τα αποτελέσματα των πειραμάτων παρουσιάζονται στους πίνακες 4 και 5 που ακολουθούν:

Πίνακας 4: Αποτελέσματα RGB-SIFT για το KID dataset 2. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**104**	**90**	**80**	**70**	**60**
500	35.42	35.42	35.42	33.33	37.50
600	35.42	33.33	33.33	31.25	35.42
700	35.42	33.33	37.50	33.33	33.33
800	37.50	39.58	37.50	33.33	35.42
900	35.42	39.58	39.58	31.25	37.50
1000	33.33	39.58	33.33	37.50	37.50
1100	35.42	37.50	35.42	33.33	33.33
1200	35.42	37.50	35.42	35.52	33.33
1300	33.33	35.42	35.42	35.42	33.33
1400	33.33	37.50	35.42	35.42	37.50
1500	39.58	39.58	33.33	31.25	37.50
Mean value	**35.417**	**37.120**	**35.606**	**33.720**	**35.605**

Πίνακας 5: Αποτελέσματα RGB-SIFT για το KID dataset 2. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**50**	**40**	**30**	**20**	**60**
500	33.33	35.42	37.50	35.42	35.42
600	39.58	35.42	37.50	35.42	37.5
700	33.33	33.33	33.33	35.42	37.50
800	37.50	35.42	35.42	35.42	33.33
900	33.33	37.50	37.50	37.50	37.50
1000	31.25	39.58	37.50	37.50	35.42
1100	39.58	39.58	33.33	35.42	35.42
1200	35.42	35.42	35.42	37.50	31.25
1300	35.42	35.42	37.50	35.42	35.42
1400	35.42	35.42	37.50	35.42	33.33
1500	37.50	35.42	33.33	33.33	35.42
Mean value	**35.605**	**36.175**	**35.984**	**35.797**	**35.228**

- Σ’αυτήν την μέθοδο, εμφανίζονται 5 τιμές ποσοστών ορθής ταξινόμησης. Αυτές είναι οι : 31.25%, 33.33%, 35.42%, 37.5% και 39.58%.
- Η μεγαλύτερη τιμή ποσοστού ορθής ταξινόμησης, είναι η 39.58% και εμφανίζεται 10 φορές.

Οι 10 πίνακες αληθείας που προκύπτουν για την μέγιστη τιμή 39.58%, παρουσιάζουν ελάχιστες διαφορές μεταξύ τους. Αυτό που έχει ενδιαφέρον σ’έναν πίνακα αληθείας όμως, είναι η κύρια διαγώνιος του, και αυτό γιατί όπως αναφέρθηκε, η κύρια διαγώνιος, δείχνει τον αριθμό των σωστά ταξινομημένων εικόνων σε κάθε κατηγορία, ενώ το άθροισμα της αποτελεί τον συνολικό αριθμό τον σωστά ταξινομημένων εικόνων. Σ’αυτήν την περίπτωση, από τους 10 πίνακες αληθείας, οι 3 παρουσιάζουν διαφορετική κατανομή στην κύρια διαγώνιο. Συνεπώς, είναι επιθυμητο, στη συνέχεια να παρουσιαστούν οι πίνακες αυτοί:

Πίνακας 6: Πρώτος πίνακας αλήθειάς για τη μέθοδο RGB-SIFT του KID dataset 2

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	4	1	0	0	0	1	0	0
Inflammatory	1	2	0	2	0	1	0	0
Normal-Colon	0	0	3	2	0	0	1	0
Normal esophagus	0	0	0	4	0	1	0	1
Normal-small-bowel	1	1	1	1	1	1	0	0
Normal-stomach	0	1	0	1	0	3	0	1
Polypoids	1	1	1	0	0	1	1	1
Vascular	0	1	0	1	2	0	1	1

Πίνακας 7: Δεύτερος πίνακας αληθείας για τη μέθοδο RGB- SIFT του KID dataset 2

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	4	1	0	0	0	1	0	0
Inflammatory	1	2	0	2	0	1	0	0
Normal-Colon	0	0	3	2	0	1	0	0
Normal esophagus	0	0	0	5	0	0	0	1
Normal-small-bowel	1	1	1	1	1	1	0	0
Normal-stomach	0	1	0	1	0	3	0	1
Polypoids	2	1	0	0	0	1	0	2
Vascular	0	1	0	1	2	0	1	1

Πίνακας 8: Τρίτος πίνακας αληθείας για τη μέθοδο RGB-SIFT του KID dataset 2

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	3	1	0	0	0	1	1	0
Inflammatory	1	2	0	2	0	1	0	0
Normal-Colon	0	0	3	2	0	0	1	0
Normal esophagus	0	0	0	5	0	0	0	1
Normal-small-bowel	1	1	1	1	1	1	0	0
Normal-stomach	0	1	0	1	0	3	0	1
Polypoids	1	1	1	0	0	1	1	1
Vascular	0	1	0	1	2	0	1	1

Στην περίπτωση του πίνακα 6, ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες Ampulla of vater και Normal esophagus, οι οποίες ταξινομούν σωστά από 4 εικόνες, όπως επίσης και οι κατηγορίες Normal-colon και Normal-stomach, οι οποίες ταξινομούν σωστά από 3 εικόνες. Αντίθετα, μη ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες Vascular, Polypoids και Norma-small bowel, οι οποίες ταξινομούν σωστά από μία εικόνα, καθώς και η κατηγορία Inflammatory που ταξινομεί σωστά 2 εικόνες. Συνολικά ταξινομούνται σωστά 19/48 εικόνες.

Για την περίπτωση του πίνακα 7, οι κατηγορίες Ampulla of vater, Normal-colon, Normal-esophagus και Normal-stomach, παρουσιάζουν ικανοποιητική ταξινόμηση, καθώς ταξινομούν σωστά 4, 3, 5 και 3 εικόνες αντίστοιχα. Από την άλλη, οι κατηγορίες Inflammatory, Normal-small-bowel, Vascular και Polypoids, δεν παρουσιάζουν ικανοποιητική ταξινόμηση, διότι ταξινομούν σωστά από 2,1,1 και 0 εικόνες αντίστοιχα. Συνολικά ταξινομούνται σωστά 19/48 εικόνες

Για τον πίνακα 8, ικανοποιητική ταξινόμηση παρουσιάζουν οι κατηγορίες Ampulla of vater, Normal-colon, Normal-stomach και Normal-esophagus. Από αυτές τις κατηγορίες, οι 3 πρώτες ταξινομούν σωστά από 3 εικόνες η κάθε μία, ενώ η Normal-esophagus, ταξινομεί σωστά 5 εικόνες. Στον αντίποδα, μη ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες Vascular, Polypoids και Normal-small-bowel, οι οποίες ταξινομούν σωστά από μία εικόνα, καθώς και η κατηγορία Inflammatory, η οποία ταξινομεί σωστά 2 εικόνες.

- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης είναι η 31.25% και εμφανίζεται 5 φορές.
- Η τιμή της μεταβλητής selTrainFeats για την οποια εμφανίζονται τα καλύτερα κατα μέσο όρο αποτελέσματα ισούται με 90 και έχει μέσο όρο 37.12.
- Και σε αυτήν τη μέθοδο, η μεταβολή μεγέθους του λεξικού δεν φαίνεται να επηρεάζει ιδιαίτερα τα αποτελέσματα.
- Με εξαίρεση τη μέγιστη και την ελάχιστη τιμή, τα απότελέσματα της μεθόδου RGB-SIFT, δεν διαφέρουν σχεδόν καθόλου από αυτά της SIFT. Συνεπώς η χρωματική πληροφορία του χρωματικού χώρου RGB δεν φαίνεται να διαδραματίζει σημαντικό ρόλο στην ταξινόμηση του KID dataset 2.

3.4.1.3 Αποτελέσματα πειραμάτων με τη μέθοδο HSV-SIFT για το KID dataset 2

Τα παρακάτω αποτελέσματα, αφορόυν πειράματα που πραγματοποιήθηκαν με τη μέθοδο εξαγωγής χαρακτηριστικών HSV-SIFT. Αυτό σημαίνει ότι κατά την ταξινόμηση του KID dataset 2, αξιοποιείται η πληροφορία που παρέχει η χρωματική συνιστώσα του χώρου χρωμάτων HSV. Οι πίνακες 9 και 10 περιέχουν τα αποτελέσματα της μεθόδου.

Πίνακας 9: Αποτελέσματα HSV-SIFT για το KID dataset 2. . Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**104**	**90**	**80**	**70**	**60**
500	45.83	41.67	39.58	37.50	43.75
600	39.58	41.67	45.83	37.50	41.67
700	41.67	41.67	41.67	39.58	39.58
800	43.75	41.67	43.75	47.92	41.67
900	43.75	41.67	43.75	41.67	41.67
1000	41.67	41.67	45.83	43.75	43.75
1100	39.58	45.83	43.75	39.58	41.67
1200	45.83	45.83	37.50	41.67	43.75
1300	37.50	37.50	41.67	47.92	41.67
1400	43.75	39.58	41.67	41.67	43.75
1500	39.58	43.75	41.67	39.58	41.67
Mean value	**42.044**	**42.046**	**42.424**	**41.667**	**42.236**

Πίνακας 10: Αποτελέσματα HSV-SIFT για το KID dataset 2. . Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**50**	**40**	**30**	**20**	**10**
500	39.58	43.75	43.75	41.67	39.58
600	47.92	45.83	43.75	39.58	45.83
700	43.75	43.75	45.83	41.67	41.67
800	45.83	41.67	45.83	39.58	43.75
900	43.75	39.58	39.58	41.67	41.67
1000	47.92	41.67	41.67	41.67	41.67
1100	43.75	41.67	39.58	41.67	43.75
1200	45.83	43.75	43.75	43.75	45.83
1300	47.92	43.75	43.75	43.75	43.75
1400	43.75	41.67	43.75	43.75	43.75
1500	43.75	43.75	41.67	43.75	39.58
Mean value	**44.886**	**42.803**	**42.991**	**42.046**	**42.802**

- Η μεγαλύτερη τιμή ποσοστού ορθής ταξινόμησης που εμφανίζεται, είναι η 47.92%, η οποία εμφανίζεται 5 φορές όπως φαίνεται στους πίνακες 9 και 10. Και στις 5 περιπτώσεις ο πίνακας αληθείας είναι:

Πίνακας 11: Πίνακας αληθείας για τη μέθοδο HSV-SIFT του KID dataset 2

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	4	1	0	0	0	1	0	0
Inflammatory	2	1	0	0	1	1	0	1
Normal-Colon	0	0	5	1	0	0	0	0
Normal esophagus	0	0	0	4	0	1	0	1
Normal-small-bowel	0	1	1	2	1	1	0	0
Normal-stomach	0	0	0	0	0	5	0	1
Polypoids	1	2	1	0	0	2	0	0
Vascular	0	0	0	0	3	0	0	3

Από τον πίνακα 11, γίνεται εύκολα κατανοητό ότι ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες : Ampulla of vater με 4 σωστά ταξινόμημένες εικόνες, Normal-colon με 5 σωστά ταξινομημένες εικόνες, Normal-esophagus με 4 σωστά ταξινομημένες εικόνες, Normal-stomach με 5 σωστά ταξινομημένες εικόνες και Vascular με 3 σωστά ταξινομημένες εικόνες. Σε αντιδιαστολή, οι κατηγορίες Inflammatory και Normal-small-bowel έχουν από μία σωστά ταξινομημένη εικόνα, ενώ η κατηγορία polypoids δεν έχει καμία σωστά ταξινομημένη εικόνα. Συνολικά ταξινομούνται σωστά 23 στις 48 εικόνες.

- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης γι'αυτήν την μέθοδο είναι η 37.5%, οποία εμφανίζεται 5 φορές.
- Άλλες τιμές που εμφανίζονται είναι οι : 39.58%, 41.67%, 43.75% και 45.83%.
- Η τιμή της μεταβλητής selTrainFeats για την οποία εμφανίζονται τα υψηλότερα κατά μέσο όρο αποτελέσματα, ισούται με 50 και έχει μέσο όρο 44.886.
- Και σ'αυτήν την μέθοδο, το μέγεθος του λεξικού, δεν φαίνεται να παίζει σημαντικό ρόλο στη διαμόρφωση των αποτελεσμάτων.

- Συγκριτικά με την μέθοδο RGB-SIFT, η μέθοδος HSV-SIFT δίνει σημαντικά καλύτερα αποτελέσματα. Αυτό αποδεικνύεται μέσα από σύγκριση των πινάκων 4 και 5 της RGB-SIFT με τους πίνακες 9 και 10 της μεθόδου HSV-SIFT. Η μέγιστη τιμή της μεθόδου RGB-SIFT είναι 39.58 % και η ελάχιστη 31.25%, ενώ οι αντίστοιχες τιμές για την μέθοδο HSV-SIFT είναι 47.92% και 37.5%. Συνεπώς, η χρωματική συνιστώσα του χώρου χρωμάτων HSV, αποτελεί σημαντική πληροφορία κατά την ταξινόμηση του KID dataset 2.

3.4.1.4 Αποτελέσματα πειραμάτων με τη μέθοδο Opponent-SIFT για το KID dataset 2

Τα παρακάτω αποτελέσματα, αφορούν τα πειράματα που πραγματοποιήθηκαν με τη χρήση της μεθόδου Opponent-SIFT. Σ' αυτήν την μέθοδο, κατα την ταξινόμηση του KID dataset 2, αξιοποιείται η χρωματική πληροφορία που παρέχεται από το χρωματικό χώρο Opponent. Τα αποτελέσματα, συνοψίζονται στους πίνακες 12 και 13 :

Πίνακας 12: Αποτελέσματα Opponent-SIFT για το KID dataset 2. . Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**104**	**90**	**80**	**70**	**60**
500	41.67	41.67	41.67	39.58	41.67
600	41.67	39.58	41.67	41.67	41.67
700	41.67	39.58	41.67	39.58	41.67
800	41.67	39.58	43.75	41.67	41.67
900	41.67	37.50	41.67	41.67	39.58
1000	41.67	37.50	41.67	41.67	37.50
1100	41.67	41.67	41.67	39.58	41.67
1200	39.58	41.67	41.67	41.67	41.67
1300	39.58	39.58	41.67	39.58	39.58
1400	39.58	37.50	37.50	39.58	39.58
1500	37.50	41.67	41.67	37.50	39.58
Mean value	**40.720**	**39.792**	**41.480**	**40.340**	**40.530**

Πίνακας 13: Αποτελέσματα Opponent-SIFT για το KID dataset 2. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**50**	**40**	**30**	**20**	**10**
500	41.67	43.75	41.67	41.67	45.83
600	41.67	39.58	41.67	43.75	43.75
700	39.58	45.83	39.58	41.67	41.67
800	39.58	43.75	41.67	41.67	41.67
900	39.58	41.67	41.67	43.75	37.50
1000	41.67	41.67	41.67	43.75	37.50
1100	39.58	39.58	41.67	41.67	41.67
1200	41.67	43.75	39.58	41.67	39.58
1300	41.67	45.83	41.67	41.67	39.58
1400	39.58	39.58	45.83	41.67	41.67
1500	39.58	41.67	41.67	41.67	41.67
Mean value	**40.530**	**42.423**	**41.668**	**42.237**	**41.099**

- Η μεγαλύτερη τιμή ποσοστού ορθής ταξινόμησης που παρατηρείται σε αυτή τη μέθοδο ισούται με 45.83 % και όπως φαίνεται και στους πίνακες 12 και13, εμφανίζεται 4 φορές. Ο 4 πίνακες αληθείας που προκύπτουν, παρουσιάζουν μικρές διαφορές μεταξύ τους, ωστόσο έχουν όλοι την ίδια κύρια διαγώνιο, γι'αυτό παρατίθεται ενδεικτικά ένας από αυτους:

Πίνακας 14: Πίνακας αληθείας για τη μέθοδο Opponent-SIFT του KID dataset 2

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	4	1	0	0	0	1	0	0
Inflammatory	1	2	0	2	0	1	0	0
Normal-Colon	0	0	4	1	0	0	1	0
Normal esophagus	0	0	0	4	0	1	1	0
Normal-small-bowel	1	1	1	0	1	2	0	0
Normal-stomach	0	0	0	0	0	5	0	1
Polypoids	2	2	1	0	0	0	0	1
Vascular	0	2	0	0	2	0	0	2

Σ'αυτή τη μέθοδο, ικανοποιητική ταξινόμηση παρουσιάζουν οι κατηγορίες Ampulla of vater, Normal-colon και Normal-esophagus, οι οποίες ταξινομούν σωστά από 4 εικόνες η κάθε μία, καθώς και η κατηγορία Normal-stomach με 5 σωστά ταξινομημένες εικόνες. Αντίθετα, μη ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες Inflammatory και vascular με 2 σωστά ταξινομημένες εικόνες, Normal-small-bowel με μία σωστά ταξινομημένη εικόνα και Polypoids με καμία σωστά ταξινομημένη εικόνα. Συνολικά ταξινομούνται σωστά 22/48 εικόνες.

- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης γι'αυτή τη μέθοδο, ισούται με 37.5 % και εμφανίζεται 9 φορές.
- Οι υπόλοιπες τιμές του ποσοστού ορθής ταξινόμησης που εμφανίζονται είναι οι : 39.58%, 41.67% και 43.75%.
- Η τιμή της μεταβλητής selTrainFeats, για την οποία εμφανίζονται υψηλότερα κατά μέσο όρο αποτελέσματα, ισούται με 40 και έχει μέσο όρο 42.423.
- Και σ'αυτήν την μέθοδο, το μέγεθος του λεξικού δεν δείχνει να διαδραματίζει καάποιο σημαντικό ρόλο στην μεταβολή των αποτελεσμάτων.

- Η μέθοδος Opponent-SIFT παράγει υψηλότερα απότελέσματα από τις μεθόδους RGB-SIFT και SIFT, ωστόσο παράγει χαμηλότερα αποτελέσματα από την HSV-SIFT. Αυτό ισχύει αν αναλογιστεί κανείς ότι το μέγιστο απότελεσμα στην μέθοδο HSV-SIFT, ισούται με 47.92 %, ενώ το αντίστοιχο μέγιστο στην Opponent-SIFT ισούται με 45.83%. Συνεπώς, η χρωματική πληροφορία του χώρου χρωμάτων Opponent, παρέχει σημαντικότερες πληροφορίες στην ταξινόμηση από την χρωματική πληροφορία του χώρου RGB, αλλά λιγότερο σημαντικές από αυτήν του HSV.

3.4.1.5 Αποτελέσματα πειραμάτων με το συνδυασμό των μεθόδων HSV-SIFT και Opponent SIFT για το KID dataset 2

Τα παρακάτω αποτελέσματα, αφορούν τα πειράματα που πραγματοποιήθηκαν με το συνδυασμό μεθόδων HSV-SIFT και Opponent-SIFT. Δεδομένου ότι οι εν λόγω μέθοδοι ξεχωριστά, δίνουν αρκετά καλά αποτελέσματα στην εφαρμογή τους πάνω στο KID dataset 2, ειναι θεμιτό να συνδυαστούν, έτσι ώστε να εξεταστεί το ενδεχόμενο της ανακάλυψης ακόμη υψηλότερων ποσοστών ορθής ταξινόμησης. Τα αποτελέσματα μπορούν να τα διακρίνει κάποιος στους πίνακες 18 και 19:

Πίνακας 15: Αποτελέσματα του συνδυασμού των μεθόδων για το KID dataset 2. . Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**104**	**90**	**80**	**70**	**60**
500	43.75	39.58	41.67	41.67	41.67
600	41.67	39.58	41.67	39.58	43.75
700	41.67	41.67	41.67	39.58	43.75
800	41.67	41.67	39.58	41.67	41.67
900	41.67	41.67	41.67	39.58	41.67
1000	39.58	41.67	39.58	39.58	43.75
1100	41.67	41.67	41.67	43.75	41.67
1200	43.75	41.67	41.67	43.75	43.75
1300	41.67	43.75	41.67	43.75	41.67
1400	41.67	43.75	41.67	41.67	41.67
1500	41.67	39.58	41.67	39.58	41.67
Mean value	**41.858**	**41.478**	**41.290**	**41.287**	**42.426**

Πίνακας 16: Αποτελέσματα του συνδυασμού των μεθόδων για το KID dataset 2. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats				
numWords	**50**	**40**	**30**	**20**	**10**
500	39.58	41.67	39.58	41.67	41.67
600	41.67	41.67	41.67	41.67	41.67
700	39.58	41.67	41.67	41.67	43.75
800	41.67	41.67	41.67	41.67	41.67
900	41.67	41.67	41.67	43.75	39.58
1000	39.58	41.67	41.67	41.67	41.67
1100	41.67	41.67	43.75	43.75	41.67
1200	41.67	41.67	43.75	41.67	39.58
1300	41.67	39.58	43.75	41.67	39.58
1400	41.67	41.67	41.67	43.75	43.75
1500	43.75	41.67	41.67	43.75	41.67
Mean value	**41.289**	**41.480**	**42.047**	**42.426**	**41.478**

- Η μεγαλύτερη τιμή ποσοστού ορθής ταξινόμησης, που εμφανίζεται γι'αυτήν τη μέθοδο ισούται με 43.75%, και εμφανίζεται 21 φορές. Και οι 21 πίνακες αληθείας που αντιπροσωπεύουν το πόσοστό αυτό, παρουσιάζουν κάποιες μικρές διαφορές μεταξύ τους. Ωστόσο, έχουν όλοι την ίδια κύρια διαγώνιο, και γι'αυτό το λόγο παρατίθεται μόνο ένας από αυτούς ως αντιπροσωπευτικό δείγμα.

Πίνακας 17: Πίνακας αληθείας για τον συνδυσμό των μεθόδων Opponent&HSV SIFT

	Ampulla of vater	Inflammatory	Normal-Colon	Normal-esophagus	Normal-small-bowel	Normal-stomach	Polypoids	Vascular
Ampulla of vater	4	1	0	0	0	1	0	0
Inflammatory	3	1	0	0	0	1	0	1
Normal-Colon	0	0	4	1	0	0	1	0
Normal esophagus	0	0	0	4	0	1	1	0
Normal-small-bowel	0	1	1	2	1	1	0	0
Normal-stomach	0	0	0	0	0	5	0	1
Polypoids	2	2	1	0	0	1	0	0
Vascular	0	0	0	0	4	0	0	2

Σ'αυτήν τη μέθοδο, ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες Ampulla of vater, Normal-colon και Normal-esophagus,οι οποίες έχουν από 4 σωστά ταξινομημένες εικόνες η κάθε μία, καθώς και η κατηγορία Normal-stomach, που έχει με 5 σωστά ταξινομημένες εικόνες. Αντίθετα, μη ικανοποιητική ταξινόμηση, παρουσιάζουν οι κατηγορίες Inflammatory και Normal-small-bowel με μία σωστά ταξινομημένη εικόνα, Vascular με 2 σωστά ταξινομημένες εικόνες και Polypoids με καμία σωστά ταξινομημένη εικόνα. Συνολικά, ταξινομούνται σωστά 21/48 εικόνες.

- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης, ισούται με 39.58% και εμφανίζεται 19 φορές. Όλα τα υπόλοιπα ποσοστά που εμφανίζονται πλήν της μέγιστης τιμής, έχουν τιμή 41.67%.
- Οι τιμές της μεταβλητής selTrainFeats, για τις οποίες εμφανίζονται υψηλότερα κατά μέσο όρο αποτελέσματα είναι οι 20 και 60 με μέσο όρο 42.426.
- Και σ'αυτήν την μέθοδο, το μέγεθος του λεξικού δεν δείχνει να διαδραματίζει καάποιο σημαντικό ρόλο στην μεταβολή των αποτελεσμάτων.
- Ο συνδυασμός των μεθόδων HSV-SIFT και Opponent-SIFT, δεν φαίνεται να παρέχει καλύτερα απότελέσματα από τις μεθόδους αυτές ξεχωριστά. Τόσο η μέθοδος Opponent-SIFT (45.83%), όσο και η μέθοδος HSV-SIFT (47.92%), παρουσιάζουν υψηλότερα ποσοστά ταξινόμησης από το 43.75% του συνδυασμού τους. Ωστόσο, αυτός ο συνδυασμός παρέχει καλύτερα απότελέσματα από τις μεθόδους RGB-SIFT και SIFT.

3.4.1.6 Συμπεράσματα για την ταξινόμηση του KID dataset 2

Μετά το πέρας των πειραμάτων με το KID dataset 2, εξάγονται κάποια σημαντικά συμπεράσματα. Τα συμπεράσματα αυτά αφορούν: Την απόδοση των μεθόδων ταξινόμησης, την ορθή ταξινόμηση της κάθε μίας από τις 8 κατηγορίες για την κάθε μέθοδο, καθώς και το αντίκτυπο των μεταβλητών selTrainFeats και numWords στη διαμόρφωση των αποτελεσμάτων.

Συμπεράσματα σχετικά με την απόδοση των μεθόδων ταξινόμησης:

Στον παρακάτω πίνακα, συνοψίζονται οι μέγιστες τιμές των 4 μεθόδων:

Πίνακας 18: Συγκεντρωτικά αποτελέσματα μέγιστων ποσοστών ταξινόμησης για το KID dataset 2. Με Bold αναγράφεται η μέγιστη τιμή ποσοστού ορθής ταξινόμησης στο σέτ ελέγχου

	SIFT	RGB-SIFT	HSV-SIFT	Opponent-SIFT	HSV&Opponent SIFT
Μέγιστο ποσοστό ορθής ταξινόμησης	41.67%	39.58%	**47.92%**	45.83%	43.75%

- Τα υψηλότερα αποτελέσματα, παρέχονται από τη μέθοδο HSV-SIFT, η οποία παρουσιάζει μέγιστο ποσοστό ταξινόμησης 47.92%.
- Ακολουθεί η μέθοδος Opponent-SIFT με μέγιστο ποσοστό 45.83%, καθώς και ο συνδυασμός των μεθόδων Opponent-SIFT με HSV-SIFT με μέγιστο ποσοστό 43.75%.
- Σειρά έχει μέθοδος SIFT που παρουσιάζει μέγιστο ποσοστό ταξινόμησης 41.67%.
- Χαμηλότερο ποσοστό από όλες της μεθόδους, φαίνεται να το παρουσιάζει η μέθοδος RGB-SIFT, καθώς το μέγιστο ποσοστό της ισούται με 39.58%.
- Συνεπώς, η χρωματική πληροφορία που παρέχει ο χώρος χρωμάτων HSV, διαδραματίζει σημαντικό ρόλο στην ταξινόμηση του KID dataset 2. Αυτό ειναι απόλυτα λογικό, αν αναλογιστεί κανείς ότι ο χρωματικός χώρος HSV, εκτός απο το γεγονός ότι χρησιμοποιείται για την παραγωγή εικόνων υψηλής ανάλυσης, παράλληλα απεικονίζει τα χρώματα και μ’έναν τρόπο εξαιρετικά κοντινό στην ανθρώπινη όραση [26].

Συμπεράσματα σχετικά με την ταξινόμηση της κάθε κατηγορίας:

- Η κατηγορία Ampulla of vater, σ’όλες τις μεθόδους, ταξινομεί σωστά 4/6 εικόνες της. Εξάίρεση αποτελεί ο πίνακας 8 της μεθόδου RGB-SIFT, στον οποίο ταξινομούνται σωστά 3/6 εικόνες.
- Η κατηγορία Inflammatory, στις μεθόδους Opponent-SIFT, SIFT και RGB-SIFT, ταξινομεί σωστά 2/6 εικόνες, ενώ στις μεθόδους HSV-SIFT και στο συνδυασμό HSV-SIFT με Opponent-SIFT, ταξινομεί σωστά 1/6 εικόνες.
- Η κατηγορία Normal-Colon, στις μεθόδους SIFT και RGB-SIFT ταξινομεί σωστά 3/6 εικόνες της. Στις μεθόδους Opponent-SIFT και στο συνδυασμό των μεθόδων HSV-SIFT με Opponent-SIFT, ταξινομεί σωστά 4/6 εικόνες. Τέλος στην μέθοδο HSV-SIFT ταξινομεί σωστά 5/6 εικόνες.
- Η κατηγορία Normal-esophagus, στη μέθοδο SIFT, καθώς και στους πίνακες 7 και 8 της μεθόδου RGB-SIFT, ταξινομεί σωστά 5/6 εικόνες. Στις υπόλοιπες κατηγορίες ταξινομεί σωστά 4/6 εικόνες.
- Η κατηγορία Normal-small-bowel, ταξινομεί σωστά 1/6 εικόνες σ’όλες τις μεθόδους.
- Η κατηγορία Normal-stomach, στη μέθοδο SIFT, ταξινομεί σωστά 4/6, ενώ στη μέθοδο RGB-SIFT, ταξινομεί σωστά 3/6 εικόνες. Στις υπόλοιπες μεθόδους ταξινομεί σωστά 5/6 εικόνες.
- Η κατηγορία Polypoids δεν ταξινομεί σωστά καμία εικόνα, με όποια μέθοδο και αν ταξινομηθεί. Εξαίρεση αποτελούν, οι πίνακες 6 και 8 της μεθόδου RGB-SIFT, όπου ταξινομεί σωστά 1/6 εικόνες.
- Η κατηγορία Vascular, με τις μεθόδους Opponent-SIFT και το συνδυασμό των μεθόδων HSV-SIFT και Opponent-SIFT, ταξινομεί σωστά 2/6 εικόνες. Παράλληλα, με τη μέθοδο HSV-SIFT, ταξινομεί σωστά 3/6 εικόνες, ενώ με τις μεθόδους SIFT και RBG-SIFT, ταξινομεί σωστά 1/6 εικόνες.

Με βάση τα παραπάνω, μπορεί να ειπωθεί ότι οι κατηγορίες Ampulla of vater, Normal-colon, Normal-esophagus και Normal-stomach ταξινομούνται ικανοποιητικά με όλες τις μεθόδους. Αντίθετα, οι κατηγορίες Normal-small-bowel, Inflammatory και Polypoids δεν ταξινομούνται ικανοποιητικά με καμία μέθοδο. Τέλος η κατηγορία Vascular, ταξινομείται ικανοποιητικά μόνο με τη μέθοδο HSV-SIFT.

Συμπεράσματα σχετικά με την επιρροή των μεταβλητών selTrainFeats και numWords:

- Και οι δύο μεταβλητές, μεταβάλλουν τα αποτελέσματα, ωστόσο δεν φαίνεται να υπάρχει κάποια άυξουσα ή φθίνουσα σχέση, η οποία συνδέει τη μεταβολή των μεταβλητών αυτών με τα αποτελέσματα. Δηλαδή, δεν μπορεί κάποιος να αποφανθεί ότι αυξάνοντας ή μειώνοντας τις τιμές των μεταβλητών αυτών, αυξάνονται ή μειώνονται τα ποσοστά ορθής ταξινόμησης.

3.4.2 Αποτελέσματα πειραμάτων για το Normal-Abnormal σέτ

Τα αποτελέσματα που ακολουθούν, αφορούν την ταξινόμηση του Normal-Abnormal σέτ για τις διάφορες μεθόδους εξαγωγής χαρακτηριστικών. Όπως διατυπώθηκε στην ενότητα 3.2, το Normal-Abnormal σέτ, δεν είναι τίποτα άλλο παρά μία παραλλαγή του KID dataset 2. Σύμφωνα με την παραλλαγή αυτή, οι εικόνες των 4 υγιών κατηριών του KID dataset 2, τοποθετούνται όλες μαζί σε μία ενιαία κατηγορία με το όνομα «Normal», ενώ οι εικόνες των 4 μη υγιών κατηγοριών, τοποθετούνται σε μία άλλη ενιαία κατηγορία με το όνομα «Abnormal». Ως εκ τούτου, το πρόβλημα της ταξινόμησης των 8 κατηγοριών, μετατρέπεται σε πρόβλημα ταξινόμησης 2 μεγάλων κατηγορίων.

Η κατηγορία Normal περιέχει 1178 εικόνες, ενώ η κατηγορία Abnormal περιέχει 593 εικόνες. Και σ’αυτήν την περίπτωση, αναζητείται αρχικά η κατηγορία με τον μικρότερο αριθμό εικόνων. Αυτή προφανώς ειναι η κατηγορία Abnormal με τις 593 εικόνες, και το λόγο αυτό, κρατούνται 593 εικόνες από την κάθε κατηγορία. Από τις 593 εικόνες της κάθε κατηγορίας, οι 415 (70%), χρησιμοποιούνται για τη φάση εκπαίδευσης, ενώ οι υπόλοιπες 178 (30%) για την φάση ελέγχου. Στο σύνολο, και από τις 2 κατηγορίες, ταξινομούνται 1186 εικόνες (2*593). Από αυτές, οι 830 (2*415), δηλαδή το 70 % χρησιμοποιείται για τη φάση εκπαίδευσης, ενώ οι υπόλοιπες 356 (2*178), δηλαδή το 30% , χρησιμοποιείται στη φάση ελέγχου.

Στα πειράματα με το Normal-Abnormal σέτ, επαναλαμβάνεται ακριβώς η ίδια διαδικασία που πραγματοποιήθηκε και στο KID dataset 2. Χρησιμοποιούνται ακριβώς οι ίδιοι μέθοδοι, παρουσιάζονται οι πίνακες κατανομής αποτελεσμάτων και οι πίνακες αλήθείας, ενώ παράλληλα εξετάζεται η επιρροή των παραμέτρων selTrainFeats και numWords στην διαμόρφωση των αποτελεσμάτων.

3.4.2.1 Αποτελέσματα πειραμάτων με τη μέθοδο SIFT για το Normal-Abnormal σέτ

Τα παρακάτω αποτελέσματα, αφορούν την εφαρμογή της μεθόδου εξαγωγής χαρακτηριστικών SIFT στην ταξινόμηση των εικόνων του Normal-Abnormal σέτ. Αξίζει να σημειωθεί, ότι η μέγιστη τιμή της μεταβλητής selTrainFeats, για την οποία ο αλγόριθμός λειτουργούσε χωρίς να δημιουργούνται προβλήματα μνήμης, ηταν η 674. Τιμές μεγαλύτερες από 674 δεν επέτρεπαν την ολοκήρωση του αλγορίθμου. Τα αποτελέσματα, σ’όλες τις περιπτώσεις αφορούν τιμές του ποσοστού ορθής ταξινόμησης στο σέτ ελέγχου. Οι πίνακες 19 και 20, συνοψίζουν τα αποτελέσματα της μεθόδου SIFT.

Πίνακας 19: Αποτελέσματα SIFT για το Normal-Abnormal σέτ. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats					
numWords	**100**	**150**	**200**	**250**	**300**	**350**
500	68.26	69.94	67.98	69.66	71.91	71.35
600	71.91	71,91	69.94	71.35	70.51	68.26
700	72.75	69.10	74.72	70.22	73.03	74.44
800	70.51	72.47	72.47	68.82	73.31	72.75
900	72.75	73.31	72.47	71.91	72.47	71.35
1000	73.88	72.47	71.91	70.79	73.60	73.60
1100	72.75	73.03	75.00	72.47	72.47	73.60
1200	72.75	73.88	72.47	71.35	73.60	71.07
1300	75.00	74.16	74.16	73.31	75.00	72.75
1400	73.31	73.88	72.19	73.60	73.31	74.72
1500	74.16	73.31	73.03	73.60	72.47	75.00
Mean value	**72.548**	**72.555**	**72.394**	**71.552**	**72.880**	**72.626**

Πίνακας 20: Αποτελέσματα SIFT για το Normal-Abnormal σέτ. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats						
numWords	**400**	**450**	**500**	**550**	**600**	**650**	**674**
500	71.35	67.70	72.47	71.35	71.35	69.94	72.75
600	69.66	71.07	72.75	72.19	71.91	71.35	73.31
700	72.19	71.91	71.35	68.54	71.35	72.47	73.31
800	73.60	71.35	71.91	71.63	73.60	69.10	71.07
900	73.31	74.16	76.40	73.60	70.79	72.47	71.35
1000	73.03	72.19	71.91	73.60	74.16	71.35	72.47
1100	74.16	74.16	73.60	74.16	71.91	71.63	71.91
1200	71.07	74.72	73.88	72.47	73.31	72.47	70.22
1300	73.03	72.19	73.03	75.28	73.31	71.91	71.91
1400	74.16	72.75	72.42	75.00	74.16	73.88	73.03
1500	72.19	73.31	72.47	75.56	76.12	72.19	72.75
Mean value	**72.522**	**72.319**	**72.926**	**73.034**	**72.906**	**71.705**	**70.878**

- Η μεγαλύτερη τιμή που εμφανίζεται σε ποσοστό ορθής ταξινόμησης ισούται με 76.4%. Ο πίνακας αληθείας γι'αυτήν την περίπτωση είναι:

Πίνακας 21: Πίνακας αληθείας για τη μέθοδο SIFT του Normal-Abnormal σέτ

	Abnormal	Normal
Abnormal	128	50
Normal	34	144

Από τον πίνακα αληθείας, εξάγονται τα εξής συμπεράσματα σχετικά με την ταξινόμηση: Στην κατηγορία Normal, από τις 178 εικόνες, οι 128 ταξινομούνται σωστά, ενώ οι υπόλοιπες 50 ταξινομούνται λάθος. Αντίστοιχα, στην κατηγορία Abnormal, από τις 178 εικόνες, οι 114 ταξινομούνται σωστά, ενώ οι υπόλοιπες 34 ταξινομούνται λάθος. Συνολικά, ταξινομούνται σωστά 272/356 εικόνες για τη μέγιστη τιμή ποσοστού ορθής ταξινόμησης της μεθόδου SIFT.

- Άλλες υψηλές τιμές ποσοστών ορθής ταξινόμησης είναι οι : 76.12%, 75.56%, 75.28% και 75%.
- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης ισούται με 67.7%.
- 13 τιμές ποσοστών ορθής ταξινόμησης παρουσιάζουν τιμές μικρότερες του 70%.
- Η τιμή της μεταβλητής selTrainFeats για την οποία παρουσιάζονται τα υψηλότερα κατα μέσο όρο αποτελέσματα, ισούται με 550 και έχει μέσο όρο 73.034.
- Το μέγεθος του λεξικού, φαίνεται ότι έως ένα βαθμό επηρεάζει τα αποτελέσματα της μεθόδου. Εύκολα παρατηρεί κάποιος πώς όλα ποσοστά ορθής ταξινόμησης που παρουσιάζουν τιμές κάτω του 70%, βρίσκονται σε τιμές λεξικού μικρότερες του 1000. Αντίθετα οι υψηλότερες τιμές ποσοστών ορθής ταξινόμησης (75.28%, 75.56%, 76.12%, 75%), τείνουν να βρίσκονται σε τιμές μεγαλύτερες του 1000.

3.4.2.2 Αποτελέσματα πειραμάτων με τη μέθοδο RGB-SIFT για το Normal-Abnormal σέτ

Τα αποτελέσματα που ακολουθούν, αφορούν την εφαρμογή της μεθόδου εξαγωγής χαρακτηριστικών RGB-SIFT στην ταξινόμηση των εικόνων του Normal-Abnormal σέτ. Όπως είναι ήδη γνωστό από προηγούμενα πειράματα, για την ταξινόμηση των εικόνων σ'αυτήν την μέθοδο, αξιοποιείται η χρωματική πληροφορία που παρέχει ο χώρος χρωμάτων RGB. Τα αποτελέσματα της μεθόδου συνοψίζονται στον πίνακα 22. Αξίζει να σημειωθεί ότι και σ'αυτήν την περίπτωση, λόγω προβλημάτων μνήμης η τιμή της μεταβλητής selTrainFeats, δεν ξεπερνάει το 220.

Πίνακας 22: Αποτελέσματα RGB-SIFT για Normal-Abnormal σέτ. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats			
numWords	**100**	**150**	**200**	**220**
500	69.38	70.22	71.07	73.31
600	72.19	72.75	69.38	69.94
700	73.31	74.44	71.63	73.03
800	73.60	74.16	75.56	73.03
900	72.75	75.28	72.47	72.19
1000	74.44	73.88	74.72	72.75
1100	75.84	70.79	75.00	73.03
1200	73.03	73.31	76.40	74.16
1300	73.88	73.31	74.72	72.47
1400	74.16	75.84	76.97	75.28
1500	74.44	74.44	76.69	74.44
Mean value	**73.365**	**73.492**	**74.055**	**73.057**

- Η μεγαλύτερη τιμή που παρατηρείται σε ποσοστό ορθής ταξινόμησης ισούται με 76.97%. Ο πίνακας αληθείας για την τιμή αυτή είναι :

Πίνακας 23: Πίνακας αληθείας για τη μέθοδο RGB- SIFT του Normal-Abnormal σέτ

	Abnormal	Normal
Abnormal	129	49
Normal	33	145

Σύμφωνα με τον πίνακα αληθείας, η ταξινόμηση για τις 2 κατηγορίες έχει ως εξής: Από την κατηγορία Abnormal, 129 εικόνες ταξινομούνται σωστά και 49 εικόνες ταξινομούνται λάθος. Από την κατηγορία Normal, 145 εικόνες ταξινομούνται σωστά και 33 εικόνες ταξινομούνται λάθος. Συνολικά, 274/356 εικόνες ταξινομούνται σωστά.

- Άλλες υψηλές τιμές ποσοστών ορθής ταξινόμησης που εμφανίζονται στη μέθοδο αυτή είναι: 76.69%, 76.4%, 75.84%, 75.56%, 75.28% και 75%.
- Στη μέθοδο SIFT, εμφανίστηκαν 13 ποσοστά ορθής ταξινόμησης με τιμές μικρότερες του 70%, στην μέθοδο RGB-SIFT εμφανίζονται μόνο 3. Αυτές είναι οι : 69.38% η οποία είναι η μικρότερη και εμφανίζεται 2 φορές και η 69.94%.
- Η τιμή της μεταβλητής selTrainFeats για την οποία εμφανίζονται τα υψηλότερα κατα μέσο όρο αποτελέσματα, ισούται με 200 και έχει μέσο όρο 74.055.
- Στη μέθοδο SIFT, η μεταβλητή selTrainFeats είχε τη δυνατότητα να λάβει τιμές μέχρι και την τιμή 674, πριν ξεκινήσουν να παρουσιάζονται προβλήματα μνήμης. Στην μέθοδο RGB-SIFT όμως, τα προβλήματα μνήμης εμφανίζονται όταν η μεταβλητή selTrainFeats λάβει τιμές μεγαλύτερες του 220. Είναι φανερό λοιπόν, ότι η τιμή της μεταβλητής για την οποία αρχίζουν τα προβλήματα μνήμης στη μέθοδο RGB-SIFT, είναι πολύ μικρότερη από την αντίστοιχη τιμή της μεθόδου SIFT. Γιατί γίνεται αυτό όμως; Η απάντηση στο ερώτημα αυτό, βρίσκεται στο γεγονός ότι όσο πολυπλοκότεροι είναι οι υπολογισμοί ενός αλγορίθμου, τόσο περισσότερη μνήμη απαιτείται για την εκτέλεση του. Στην περίπτωση αυτή, αρκεί κάποιος να θυμηθεί οτι το διάνυσμα των χαρακτηριστικών που εξάγονται από μία εικόνα με τη μέθοδο SIFT, έχει μέγεθος 128, ενώ το αντίστοιχο διάνυσμα στη μέθοδο RGB-SIFT, έχει μέγεθος 384. Συνεπώς, η μέθοδος RGB-SIFT απαιτεί περισσότερη μνήμη από την μέθοδο SIFT και γι’αυτό ο αλγόριθμος παρουσιάζει προβλήματα μνήμης σε μικρότερη τιμή της μεταβλητής selTrainFeats για τη μέθοδο αυτή. Το ίδιο ακριβώς ισχύει και για τις μεθόδους HSV-SIFT και Opponent-SIFT που ακολουθούν.
- Και σ’αυτήν τη μέθοδο, το μέγεθος του λεξικού φαίνεται να επηρεάζει έως ένα βαθμό τα αποτελέσματα που προκύπτουν. Αυτό γίνεται κατανοήτο, από το γεγονός ότι η πλειοψηφία των ποσοστών ορθής ταξινόμησης, που έχουν υψηλές τιμές, εμφανίζονται για μέγεθος λεξικού μεγαλύτερο του 1000. Αντίθετα, τα ποσοστά ορθής ταξινόμησης με τιμές μικρότερες του 70%, εμφανίζονται για μέγεθος λεξικού μικρότερο του 1000.
- Τα αποτελέσματα της μεθόδου RGB-SIFT, δεν διαφέρουν σημαντικά από αυτά της μεθόδου SIFT, καθώς η μέγιστη τιμή ποσοστού ορθής ταξινόμησης στην μέθοδο SIFT ισούται με 76.4%, ενώ η αντίστοιχη στη μέθοδο RGB-SIFT ισούται με 76.97%. Αυτό σημαίνει ότι η χρωματική πληροφορία που προσφέρει ο χώρος χρωμάτων RGB, δεν διαδραματίζει σημαντικό ρόλο στην ταξινόμηση.

3.4.2.3 Αποτελέσματα πειραμάτων με τη μέθοδο HSV-SIFT για το Normal-Abnormal σέτ

Τα αποτελέσματα που ακολουθούν, αφορούν την εφαρμογή της μεθόδου εξαγωγής χαρακτηριστικών HSV-SIFT, στην ταξινόμηση του Normal-Abnormal σέτ. Η μέθοδος αυτή, αξιοποιεί στην ταξινόμηση των εικόνων, τη χρωματική πληροφορία που παρέχεται από τον χώρο χρωμάτων HSV. Δεδομένου των όσων αναφέρθηκαν νωρίτερα, σχετικά με τα προβλήματα μνήμης και το μέγεθος της μεταβλητής selTrainFeats,η μέγιστη τιμή της μεταβλητής αυτής, σ'αυτήν την περίπτωση, ισουται με 224. Τα αποτελέσματα παρουσιάζονται στον πίνακα 24:

Πίνακας 24: Αποτελέσματα HSV-SIFT για το Normal-Abnormal σέτ. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats			
numWords	**100**	**150**	**200**	**224**
500	78.65	77.81	75.00	78.37
600	75.28	76.97	77.53	76.69
700	78.65	77.53	75.28	77.53
800	79.78	76.97	79.21	78.37
900	80.06	81.18	77.81	78.65
1000	79.49	77.81	78.09	78.37
1100	80.06	79.49	77.25	80.34
1200	80.06	79.49	76.40	79.21
1300	79.78	77.53	77.25	77.53
1400	78.93	80.06	77.53	80.62
1500	81.46	80.62	79.49	77.81
Mean value	**79.290**	**78.597**	**77.349**	**78.499**

- Συγκριτικά με τις 2 προηγούμενες μεθόδους, η HSV-SIFT παρουσιάζει εξαιρετικά υψηλά αποτελέσματα.
- Η μεγαλύτερη τιμή ποσοστού ορθής ταξινόμησης που εμφανίζεται, ισούται με 81.46%. Ο πίνακας αληθείας για την τιμή αυτή είναι:

Πίνακας 25: Πίνακας αληθείας της μεθόδου HSV-SIFT για το Normal-Abnromal σέτ

	Abnormal	Normal
Abnormal	126	52
Normal	14	164

Σύμφωνα με τον πίνακα αληθείας η ταξινόμηση έχει ως εξής: Στην κατηγορία Abnormal, ταξινομούνται σωστά 126 εικόνες, ενώ οι υπόλοιπες 52 ταξινομούνται λάθος. Στην κατηγορία Normal, 164 εικόνες ταξινομούνται σωστά, ενώ οι υπόλοιπες 14 ταξινομούνται λάθος. Συνολικά ταξινομούνται σωστά 290/356 εικόνες.

- Άλλες υψηλές τιμές που εμφανίζονται είναι : 81.18%, 80.62%, 80.34% και 80.06%.
- Συνολικά εμφανίζονται 9 ποσοστά ορθής ταξινόμησης με τιμές άνω του 80%.
- Σε αντίθεση με τις μεθόδους RGB-SIFT και SIFT, η HSV-SIFT, δεν εμφανίζει κανένα ποσοστό ορθής ταξινόμησης με τιμή χαμηλότερη του 70%.
- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης ισούται με 75%.
- Η τιμή της μεταβλητής selTrainFeats, για την οποία εμφανίζονται τα υψηλότερα κατά μέσο όρο αποτελέσματα, ισούται με 100 και έχει μέσο όρο 79.29.
- Το μέγεθος του λεξικού στη μέθοδο HSV-SIFT, επηρεάζει τα αποτελέσματα με τρόπο παρόμοιο με αυτόν των προηγούμενων μεθόδων. Οι υψηλότερες τιμές ποσοστών, έχουν την τάση να εμφανίζονται για τιμές μεγέθους λεξικού μεγαλύτερου του 1000, ενώ οι χαμηλότερες τιμές, έχουν την τάση να εμφανίζονται για τιμές λεξικού που είναι μικρότερες του 1000.
- Μέχρι στιγμής, η μέθοδος HSV-SIFT παράγει υψηλότερα αποτελέσματα συγκριτικά με τις μεθόδους SIFT και RGB-SIFT. Καμία από τις 2 προηγούμενες μεθόδους δεν παρουσίασε ποσοστό με τιμή ανώτερη του 80%, ενώ παράλληλα και στις 2 εμφανίστηκαν ποσοστά με τιμές μικρότερες του 70%. Αυτό σημαίνει ότι η χρωματική πληροφορία που παρέχει ο χώρος χρωμάτων HSV, διαδραματίζει σημαντικότερο ρόλο στην ταξινόμηση του Normal-Abnormal σέτ από αυτήν που παρέχει ο χώρος χρωμάτων RGB.

3.4.2.4 Αποτελέσματα πειραμάτων με τη μέθοδο Opponent-SIFT για το Normal-Abnormal σέτ

Τα αποτελέσματα που ακολουθούν, αφορούν τα πειράματα που πραγματοποιήθηκαν με τη μέθοδο εξαγωγής χαρακτηριστικών Opponent-SIFT για την ταξινόμηση των εικόνων του Normal-Abnormal σέτ. Όπως είναι ήδη γνωστό, για την ταξινόμηση στη μέθοδο αυτή, αξιοποιείται η χρωματική πληροφορία που παρέχει χώρος χρωμάτων Opponent. Και σε αυτήν την περίπτωση, η τιμής της μεταβλητής selTrainFeats έχει μέγιστη τιμή 224. Τα αποτελέσματα συνοψίζονται στον πίνακα 26 :

Πίνακας 26: Αποτελέσματα Opponent-SIFT για το Normal-Abnormal σέτ. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats			
numWords	**100**	**150**	**200**	**224**
500	73.31	72.75	76.12	73.03
600	74.72	74.44	75.00	75.00
700	74.72	75.56	74.44	75.28
800	75.84	75.00	71.63	78.37
900	74.16	76.4	73.31	76.4
1000	77.81	75.28	72.75	76.69
1100	75.56	75.00	76.97	75.56
1200	77.53	77.81	77.53	75.84
1300	75.56	77.25	72.47	76.97
1400	78.37	77.81	74.72	76.12
1500	77.81	75.84	76.40	77.81
Mean value	**75.944**	**75.740**	**74.667**	**76.096**

- Η μεγαλύτερη τιμή, ποσοστού ορθής ταξινόμησης που παρατηρείται ισούται με 78.37% και εμφανίζεται δύο φορές. Και στις δύο αυτές φορές, οι πίνακες αληθείας που προκύπτουν είναι διαφορετικοί. Παρακάτω παρουσιάζονται και αναλύονται οι πίνακες αυτοί.

Για την περίπτωση όπου selTrainFeats = 100 και numWords = 1400, ο πίνακας αληθείας για την τιμή 78.37 % είναι:

Πίνακας 27: Πρώτος πίνακας αληθείας της μεθόδου Opponent-SIFT για το Normal-Abnormal σέτ

	Abnormal	Normal
Abnormal	117	61
Normal	16	162

Στην περίπτωση αυτή η ταξινόμηση έχει ως εξής: Στην κατηγορία Abnormal, 117 εικόνες ταξινομούνται σωστά, ενώ 61 εικόνες ταξινομούνται λάθος. Στην κατηγορία Normal από την άλλη, 162 εικόνες ταξινομούνται σωστά, ενώ 16 εικόνες ταξινομούνται λάθος. Συνολικά ταξινομούνται σωστά 279/356 εικόνες.

Για την περίπτωση όπου selTrainFeats = 224 και numWords = 800, ο πίνακας αληθείας για τη τιμή 78.37% είναι:

Πίνακας 28: Δεύτερος πίνακας αληθείας για τη μέθοδο Opponent-SIFT του Normal-Abnormal σέτ

	Abnormal	Normal
Abnormal	120	58
Normal	19	159

Στην περίπτωση αυτή η ταξινόμηση έχει ως εξής: Στην κατηγορία Abnormal, ταξινομούνται σωστά 120 εικόνες, ενώ οι υπόλοιπες 58 ταξινομούνται λάθος. Στην κατηγορία Normal, ταξινομούνται σωστά 159 εικόνες, ενώ οι υπόλοιπες 19 ταξινομούνται λάθος. Όπως είναι λογικό και σ’αυτήν την περίπτωση, ταξινομούνται σωστά 279/356 εικόνες.

Οι δύο αυτοί πίνακες αληθείας, δεν διαφέρουν αρκετά μεταξύ τους. Προφανώς, το άθροισμα στην κύρια διαγώνιο τους είναι το ίδιο, δεδομένου ότι αναφέρονται στο ίδιο ποσοστό ορθής ταξινόμησης. Η μόνη διαφορά, είναι ότι στην περίπτωση του πίνακα 28, στην κατηγορία Abnormal, ταξινόμουνται σωστά 3 περισσότερες εικόνες, ενώ στην κατηγορία Normal 3 λίγοτερες.

- Η μικρότερη τιμή ποσοστού ορθής ταξινόμησης που εμφανίζεται, ισούται με 71.63%.
- Η τιμή της μεταβλητής selTrainFeats για την οποία εμφανίζονται τα μεγαλύτερα κατά μέσο όρο αποτελέσματα, ισούται με 224 και έχει μέσο όρο 76.096.

- Η μέθοδος Opponent-SIFT παρουσιάζει καλύτερα αποτελέσματα από τις μεθόδους SIFT και RGB-SIFT. Αυτό είναι ξεκάθαρο, αν σκεφτεί κανείς ότι σε αντίθεση με τις 2 αυτές μεθόδους, η Opponent-SIFT δεν παρουσιάζει τιμές ποσοστών ορθής ταξινόμησης κάτω του 70% , ενώ παράλληλα η μέγιστη τιμή τιμής της (78.37%), είναι μεγαλύτερη από τις μέγιστες τιμές των άλλων 2 μεθόδων (76.4 % και 76.97%).
- Παρά τα καλά της αποτελέσματα όμως, η μέθοδος Opponent-SIFT είναι κατώτερη της HSV-SIFT. Αυτό γίνεται εύκολα κατανοητό με μία σύντομη ματιά στους πίνακες 24 και 26. Το εύρος τιμών της HSV-SIFT είναι από 75% μέχρι 81.46%, ενώ το αντίστοιχό έυρος τιμών της Opponent-SIFT κυμένεται από 71.63% μέχρι 78.37%.
- Συνεπώς, η πληροφορία που παρέχει ο χώρος χρωμάτων Opponent, μπορεί να διαδραματίζει σημαντικότερο ρόλο στην ταξινόμηση του Normal-Abnormal σέτ από αυτήν του χρωματικού χώρου RGB, αλλά υστερεί μπροστά στην πληροφορία που παρέχει ο χώρος χρωμάτων HSV.
- Για την επιρροή που έχει το μέγεθος του λεξικού στην μέθοδο Opponent-SIFT για την ταξινόμηση των εικόνων, ισχύουν όλα όσα αναφέρθηκαν και στις προηγούμενες μεθόδους.

3.4.2.5 Αποτελέσματα πειραμάτων με το συνδυασμό των μεθόδων HSV-SIFT και Opponent-SIFT για το Normal-Abnormal σέτ

Τα αποτελέσματα που ακολουθούν, αφορούν τα πειράματα που πραγματοποιήθηκαν με το συνδυασμό των μεθόδων εξαγωγής χαρακτηριστικών HSV-SIFT και Opponent-SIFT. Μέχρι στιγμής, μέσω των δύο αυτών μεθόδων, εξήχθησαν τα καλύτερα αποτελέσματα στην ταξινόμηση του Normal-Abnormal σέτ. Είναι απόλυτα λογικό λοιπόν, να εξεταστεί το ενδεχόμενο της εξαγωγής ακόμη υψηλότερων αποτελεσμάτων μέσω του συνδυασμού τους. Εξαιτίας της πολυπλοκότητας των υπολογισμών που προκύπτει από το συνδυασμό των μεθόδων, καθώς και της περιορισμένης μνήμης, η τιμή της μεταβλητής selTrainFeats δεν λαμβάνει τιμές ανώτερες του 100. Τα αποτελέσματα παρουσιάζονται στον πίνακα 29:

Πίνακας 29: Αποτελέσματα του συνδυασμού των μεθόδων για το Normal-Abnormal σέτ. Τα αποτελέσματα αυτά, αφορούν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου. Με κόκκινο χρώμα αναγράφονται οι ελάχιστες τιμές, ενώ με πράσινο οι μέγιστες

	selTrainFeats
numWords	**100**
500	78.65
600	77.81
700	79.49
800	80.34
900	80.06
1000	78.09
1100	80.34
1200	80.06
1300	81.18
1400	80.06
1500	79.78
Mean value	**79.623**

- Η μεγαλύτερη τιμή ποσοστού ορθής ταξινόμησης που εμφανίζεται ισούται με 81.18%. Ο πίνακας αληθείας στην περίπτωση αυτή είναι:

Πίνακας 30: Πίνακας αληθείας του συνδυασμού των μεθόδων για το Normal-Abnormal σέτ

	Abnormal	Normal
Abnormal	124	54
Normal	13	165

Σύμφωνα με τον πίνακα αληθείας, η ταξινόμηση έχει ως εξής: Στην κατηγορία Abnormal, 124 εικόνες ταξινομούνται σωστά, ενώ 54 εικόνες ταξινομούνται λάθος. Στην κατηγορία Normal, 165 εικόνες ταξινομούνται σωστά, ενώ οι υπόλοιπες 13 ταξινομούνται λάθος. Συνολικά ταξινομούνται σωστά 289/356 εικόνες.

- Συνολικά, εκτός της μέγιστης τιμής που αναφέρθηκε, 5 ακόμη ποσοστά ορθής ταξινόμησης παρουσιάζουν τιμές μεγαλύτερες του 80%. Αυτά είναι: 80.06% που εμφανίζεται 3 φορές και 80.34% που εμφανίζεται άλλες 2.
- Η χαμηλότερη τιμή που ποσοστού ορθής ταξινόμησης που εμφανίζεται, ισούται με 77.81%.
- Ο συνδυασμός των μεθόδων HSV-SIFT και Opponent-SIFT, εμφανίζει τον υψηλότερο μέσο όρο αποτελεσμάτων, για την τιμή 100 της μεταβλητής selTrainFeats, συγκριτικά με τις προηγούμενες μεθόδους που παρουσιάστικαν. Πράγματι, η μέθοδος HSV-SIFT για selTrainFeats = 100, είχε μέσο όρο αποτελεσμάτων 79.29, ενώ ο συνδυασμός των μεθόδων HSV-SIFT και Opponent-SIFT για selTrainFeats = 100 έχει μέσο όρο 79.623.
- Φαίνεται πως ο συνδυασμός των μεθόδων HSV-SIFT και Opponent-SIFT, παράγει τα δεύτερα καλύτερα αποτελέσματα,καθώς βρίσκεται πίσω μόνο από τη μέθοδο HSV-SIFT. Πιο συγκεκριμένα, η μέθοδος HSV-SIFT, παρουσιάζει ελάχιστα χαμηλότερο μέσο όρο αποτελεσμάτων από τον συνδυασμό των μεθόδων HSV-SIFT και Opponent-SIFT. Όμως, η υψηλότερη τιμή που προκύπτει μέσω της HSV-SIFT (81.46%), είναι ελάχιστα μεγαλύτερη από αυτήν του συνδυασμού των HSV-SIFT και Opponent-SIFT (81.18%).
- Ο συνδυασμός των χρωματικών πληροφοριών, που παρέχονται από τους χώρους χρωμάτων HSV και Opponent, διαδραματίζει σημαντικότερο ρόλο στην ταξινόμηση του Normal-Abnormal σέτ, από τις πληροφορίες που παρέχουν οι χώροι χρωμάτων Opponent και RGB ξεχωριστά. Ωστόσο, η πληροφορία που παρέχει ο χώρος χρωμάτων HSV παραμένει με μικρή διαφορά η σημαντικότερη απ’όλες.
- Για την επιρροή του μεγέθους του λεξικού πάνω στα αποτελέσματα, ισχύει ότι ίσχυε και στις προηγούμενες περιπτώσεις.

3.4.2.6 Συμπεράσματα για την ταξινόμηση του Normal-Abnormal σέτ

Όπως και στο KID dataset 2, έτσι και στην περίπτωση του Normal-Abnormal σέτ, μετά το πέρας των πειραμάτων, εξάγονται κάποια συγκεντρωτικά συμπεράσματα. Και σ'αυτήν την περίπτωση, τα συμπεράσματα αυτά θα αφορούν: Την απόδοση των μεθόδων ταξινόμησης, την ορθή ταξινόμηση της κάθε μίας εκ των 2 κατηγοριών για την κάθε μέθοδο, καθώς και το αντίκτυπο των μεταβλητών selTrainFeats και numWords στη διαμόρφωση των αποτελεσμάτων.

Συμπεράσματα σχετικά με την απόδοση των μεθόδων ταξινόμησης:

Στον παρακάτω πίνακα, συνοψίζονται οι μέγιστες τιμές των 4 μεθόδων:

Πίνακας 31: Συγκεντρωτικά αποτελέσματα μέγιστων ποσοστών ταξινόμησης για το Normal-Abnormal σέτ. Με Bold αναγράφεται η μέγιστη τιμή ποσοστού ορθής ταξινόμησης στο σέτ ελέγχου

	SIFT	RGB-SIFT	HSV-SIFT	Opponent-SIFT	HSV&Opponent SIFT
Μέγιστο ποσοστό ορθής ταξινόμησης	76.4%	76.97%	**81.46%**	78.37%	81.18%

- Χαμηλότερη επίδοση στην ταξινόμηση του Normal-Abnormal σέτ, παρουσιάζει η μέθοδος SIFT με μέγιστο ποσοστό 76.4%. Ακολουθεί με ελάχιστη διαφορά (μέγιστο ποσοστό 76.97%) η μέθοδος RGB-SIFT. Αυτό σημαίνει πως η πληροφορία που παρέχει ο χώρος χρωμάτων RGB, δεν συμβάλλει σημαντικά στην αύξηση του ποσοστού ταξινόμησης.
- Στη συνέχεια, ακολουθεί η μέθοδος Opponent-SIFT, η οποία παρουσιάζει αισθητή βελτίωση στην ταξινόμηση του Normal-Abnormal σέτ, καθώς έχει μέγιστο ποσοστό 78.37%. Αυτό σημαίνει πως η χρωματική πληροφορία που παρέχει ο χώρος Opponent, διαδραματίζει σημαντικότερο ρόλο στην ταξινόμηση του Normal-Abnormal σέτ, από την πληροφορία που παρέχει ο χρωματικός χώρος RGB.
- Σειρά έχει ο συνδυασμός των μεθόδων HSV-SIFT και Opponent-SIFT, με μέγιστο ποσοστό 81.18%.
- Η μέθοδος HSV-SIFT, παρουσιάζει με ελάχιστη διαφορά της δεύτερης, την υψηλότερη επίδοση στην ταξινόμηση του Normal-Abnormal σέτ. Το μέγιστο ποσοστό σ'αυτήν την περίπτωση είναι 81.46%. Γίνεται κατανοητό λοιπόν, πώς η χρωματική πληροφορία που παρέχει ο χώρος χρωμάτων HSV, διαδραματίζει σημαντικότερο ρόλο στην ταξινόμηση του Normal-Abnormal σέτ, από τις πληροφορίες που παρέχουν οι άλλοι χώροι χρωμάτων.

Συμπεράσματα σχετικά με την ταξινόμηση των κατηγοριών:

Η ταξινόμηση για την κατηγορία Abnormal έχει ως εξής:

- Με τη μέθοδο SIFT ταξινομούνται σωστά 128/178 εικόνες.
- Με τη μέθοδο RGB-SIFT, ταξινομούνται σωστά 129/178 εικόνες.
- Με τη μέθοδο HSV-SIFT, ταξινομούνται σωστά 126/178 εικόνες.
- Με τη μέθοδο Opponent-SIFT, στην περίπτωση του πίνακα 27, ταξινομούνται σωστά 117/178 εικόνες, ενώ στην περίπτωση του πίνακα 28 ταξινομούνται σωστά 120/178.
- Τέλος, με το συνδυασμό των μεθόδων HSV-SIFT και Opponent-SIFT, ταξινομούνται σωστά 124/178 εικόνες.

Η ταξινόμηση για την κατηγορία Normal, έχει ως εξής:

- Με τη μέθοδο SIFT, ταξινομούνται σωστά 144/178 εικόνες.
- Με τη μέθοδο RGB-SIFT, ταξινομούνται σωστά 145/178 εικόνες.
- Με τη μέθοδο HSV-SIFT, ταξινομούνται σωστά 164/178 εικόνες.
- Με τη μέθοδο Opponent-SIFT, στην περίπτωση του πίνακα 27 ταξινομούνται σωστά 162/178 εικόνες, ενώ στην περίπτωση του πίνακα 28, ταξινομούνται σωστά 159/178 εικόνες.
- Τέλος με το συνδυασμό των μεθόδων HSV-SIFT και Opponent-SIFT, ταξινομούνται σωστά 165/178 εικόνες.

Η πρώτη παρατήρηση, αφορά το γεγονός ότι το Normal-Abnormal σέτ, ταξινομείται αρκετά καλύτερα από το KID dataset 2. Πράγματι, οι μέγιστες τιμές ταξινόμησης στο Normal-Abnormal σέτ κυμαίνονται από 76.4% έως 81.46%, ενώ οι αντίστοιχες τιμές του KID dataset 2 κυμαίνονται από 39.58% έως 47.92%. Γιατί γίνεται αυτό όμως;

Όπως είναι γνωστό, ένας υπολογιστής δεν έχει τη δυνατότητα να αναγνωρίζει αντικείμενα σε εικόνες όπως ένας άνθρωπος. Η μέθοδος bag of words, συμβάλλει στη αναγνώριση αντικειμένων σε εικόνες από έναν υπολογιστή, αλλά και στην ταξινόμηση των εικόνων αυτών σε κατηγορίες βάσει της ύπαρξης ή μη των αντικειμένων αυτών. Όπως για έναν άνθρωπο, έτσι και για έναν υπολογιστή, είναι αρκετά πιό εύκολη η αναγνώριση και η ταξινόμηση εικόνων σε 2 κατηγορίες, παρά σε 8. Ο αριθμός των κατηγοριών συνεπώς, αποτελεί έναν παράγοντα που επηρεάζει το ποσοστό ορθής ταξινόμησης.

Εκτός του αριθμού των κατηγοριών, ένας ακόμη παράγοντας που ενδέχεται να επηρεάσει το ποσοστό ορθής ταξινόμησης είναι η φύση του Dataset που ταξινομείται. Το KID dataset 2, αποτελείται απο πραγματικά δεδομένα, τα οποία δεν έχουν υποστεί επεξεργασία όπως τα δεδομένα άλλων dataset (π.χ Caltech 101). Ως εκ τούτου, οι κατηγορίες μεταξύ τους παρουσιάζουν μεγάλες διαφορές στον αριθμό των εικόνων τους, ενώ παράλληλα περιέχουν και εικόνες που ενδέχεται να είναι παρόμοιες με τις εικόνες άλλων κατηγοριών. Η φύση του KID dataset 2, κάνει την ταξινόμηση του ακόμη πιο δύσκολη υπόθεση, αν συνυπολογιστεί και ο παράγοντας της αναγνώρισης συγκεκριμένων ασθενειών, οι οποίες χαρακτηρίζονται από πολύ μικρές αλλοιώσεις που είναι εξαιρετικά δύσκολο να ανιχνευθούν και να ξεχωρίσουν από τους υγιείς ιστούς ή από άλλες ασθένειες.

Μία ακόμη παρατήρηση, αφορά την κατηγορία Normal, η οποία σ'όλες τις περιπτώσεις, παρουσιάζει υψηλότερη ταξινόμηση από την κατηγορία Abnormal. Αυτό είναι αρκετά λογικό, αν σκεφτεί κανείς ότι στο KID dataset 2, 3 από τις 4 Abnormal κατηγορίες δεν παρουσιάζουν ικανοποιητική ταξινόμηση. Συγκεκριμένα, οι κατηγορίες Inflammatory, Vascular και Polypoids, δεν ταξινομούνται ικανοποιητικά, ενώ η κατηγορία Ampulla of vater είναι η μόνη που παρουσιάζει ικανοποιητική ταξινόμηση. Αντίθετα, απο τις Normal κατηγορίες του KID dataset 2, η μόνη κατηγορία που δεν ταξινομείται ικανοποιητικά είναι η κατηγορία Normal-small-bowel.

Ένας ακόμη λόγος, για τον οποίο η κατηγορία Normal ταξινομείται καλύτερα από την κατηγορία Abnormal, είναι το γεγονός, πως οι 3 προαναφερθείσες Abnormal κατηγορίες που δεν ταξινομούνται ικανοποιητικά, έχουν την τάση να ταξινομούνται ως Normal κατηγορίες. Χαρακτηριστικότερο εκ των τριών, είναι το παράδειγμα, της κατηγορίας Vascular. Μια ματιά στους πίνακες 3, 11 και 17, είναι αρκετό για να κατανοήσει κάποιος το γεγονός, πως η κατηγορία Vascular, έχει σε μεγάλο βαθμό την τάση να ταξινομείται ως Normal-small-bowel. Αντίστοιχα, αλλά και σε μικρότερο βαθμό, ισχύει το ίδιο για τις κατηγορίες Inflammatory και Polypoids.

Συμπεράσματα σχετικά με την επιρροή των μεταβλητών selTrainFeats και numWords πάνω στα αποτελέσματα:

Ο αριθμός των εικόνων από τις οποίες εξάγονται χαρακτηριστικά κατά τη φάση εκπαίδευσης, δεν φαίνεται να επηρεάζει ιδιαίτερα τα αποτελέσματα, καθώς δεν υπάρχει κάποια αύξουσα η φθίνουσα σχέση που να συνδέει τη μεταβλητή selTrainFeats με το ποσοστό ορθής ταξινόμησης. Δηλαδή, δεν μπορεί να αποφανθεί κάποιος, ότι η αύξηση της εν λόγω μεταβλητής, θα οδηγήσει σε υψηλότερα ποσοστά, ούτε ότι η μείωση της θα οδηγήσει σε χαμηλότερα ποσοστά. Ωστόσο, η μεταβολή της μεταβλητής αυτής, συμβάλλει στην αναζήτηση νέων τιμών ποσοστών ταξινόμησης.

Το μέγεθος του λεξικού από την άλλη, φαίνεται να επηρεάζει τα αποτελέσματα για το Normal-Abnormal σέτ. Και σ'αυτήν την περίπτωση, δεν υπάρχει κάποια αύξουσα η φθίνουσα σχέση που να συνδέει τη μεταβλητή numWords με το ποσοστό ορθής ταξινόμησης. Εκ των αποτελεσμάτων όμως, φαίνεται ότι οι μεγάλες τιμές ποσοστών ορθής ταξινόμησης, έχουν την τάση να εμφανίζονται για υψηλές τιμές της μεταβλητής numWords, ενώ οι μικρές τιμές ποσοστών ορθής ταξινόμησης, έχουν την τάση να εμφανίζονται για μικρές τιμές της μεταβλητής numWords. Ωστόσο, αυτή η τάση δεν αποκλείει την πιθανότητα να συμβεί και το αντίθετο.

Το μέγεθος του λεξικού και ο αριθμός των εικόνων από τις οποίες εξάγονται χαρακτηριστικά κατά τη φάση εκπαίδευσης, δεν επηρεάζουν τόσο τα αποτελέσματα όμως, όσο η μέθοδος εξαγωγής χαρακτηριστικών που χρησιμοποιείται. Αυτό σημαίνει ότι δεν έχει σημασία μόνο ο αριθμός των χαρακτηριστικών που εξάγονται, αλλά και η «ποιότητα» τους. Η μέθοδος εξαγωγής χαρακτηριστικών, καθορίζει τις πληροφορίες που «αντλούνται» από τις εικόνες και κατ'επέκταση τους κανόνες ταξινόμησης. Μία μέθοδος εξαγωγής χαρακτηριστικών, ικάνη να «αντλήσει» τις κατάλληλες πληροφορίες από τις εικόνες, θα συμβάλλει σημαντικά στην αύξηση του ποσοστού ταξινόμησης. Δεν είναι τυχαίο, ότι για ίδιες τιμές των μεταβλητών selTrainFeats και numWords, αλλά για διαφορετική μέθοδο εξαγωγής χαρακτηριστικών, παρατηρούνται εντελώς διαφορετικές τιμές.

3.5 Γενικά συμπεράσματα

Πραγματοποιήθηκαν δύο μεγάλα πειράματα ταξινόμησης εικόνων ασύρματης ενδοσκόπησης. Ένα με το KID dataset 2 , το οποίο περιείχε 8 κατηγορίες και ένα με το Normal-Abnormal σέτ, το οποίο περιείχε 2 κατηγορίες. Στο κάθε πείραμα, αξιολογήθηκαν τα ποσοστά ορθής ταξινόμησης στο σέτ ελέγχου για τις μεθόδους εξαγωγής χαρακτηριστικών: SIFT, RGB-SIFT, HSV-SIFT, Opponent-SIFT και HSV & Opponent SIFT. Τόσο στο KID dataset 2, όσο και στο Normal-abnormal σέτ, καλύτερα αποτελέσματα έδωσε η μέθοδος εξαγωγής χαρακτηριστικών HSV-SIFT. Επιπλέον, αποδείχτηκε ότι το Normal-Abnormal σέτ παρουσιάζει υψηλότερα ποσοστά ορθής ταξινόμησης από το KID dataset 2. Αυτό μπορεί να οφείλεται στη δύσκολη φύση του KID dataset 2, αλλά και στο γεγονός ότι περιέχει περισσότερες κατηγορίες απ'ότι το Normal-Abnormal σέτ. Τέλος, απόδειχτηκε ότι το μέγεθος του λεξικού, καθώς και αριθμός των εικόνων από τις οποίες εξάγονται χαρακτηριστικά κατά τη φάση εκπαίδευσης, επηρεάζουν έως ένα βαθμό τα αποτελέσματα, μόνο στην περίπτωση του Normal-Abnormal σέτ.

3.6 Μελλοντικές επεκτάσεις

Το αρχικό πρόβλημα στο οποίο έπρεπε να δοθεί μία λύση, ήταν η διαχείριση του μεγάλου όγκου εικόνων που προέκυπτε από την ασύρματη ενδοσκόπηση. Μέσω της οπτικής εξέτασης των εικόνων αυτών, ο θεράπων ιατρός έπρεπε να ανιχνεύσει ανωμαλίες που παρέπεμπαν σε συγκεκριμένες ασθένειες. Η διαδικασία αυτή ωστόσο, ήταν αρκετά επίπονη, καθώς απαιτούσε πολλές ώρες αδιάκοπης προσοχής. Η λύση στο πρόβλημα αυτό δόθηκε με την αξιοποίηση μίας μεθόδου υπολογιστικής όρασης στην ιατρική. Η μέθοδος αυτή ονομάζεται bag of words και μέσω της χρήσης διαφόρων μεθόδων εξαγωγής χαρακτηριστικών, κατάφερε πετύχει ένα ικανοποιητικό ποσοστό ταξινόμησης, στη διάκριση εικόνων σε 2 κατηγορίες: υγιείς και μή (Normal-Abnormal). Είναι αυτό όμως αρκετό;

Όταν εκδηλώνονται κάποια συμπτώματα σ'έναν ασθενή, το πρώτο μέλημα του γιατρού είναι να μάθει αν ο ασθενής πάσχει από κάποια ασθένεια του πεπτικού συστήματος ή όχι. Η ταξινόμηση των εικόνων στις 2 κατηγορίες Normal-Abnormal, βοηθάει στην απάντηση του ερωτήματος αυτού, ωστόσο δεν συμβάλλει στην ακριβή διάγνωση της ασθένειας από την οποία πάσχει ο ασθενής, πράγμα το οποίο είναι ζωτικής σημασίας για τον ίδιο. Το KID dataset 2 από την άλλη, περιέχει 4 υγιείς κατηγορίες και 4 μη υγιείς κατηγορίες. Συνεπώς, ένα αντίστοιχα υψηλό ποσοστό στην ταξινόμηση του, θα συνέβαλλε στην ακριβή διάγνωση του ασθενούς σε σχετικά γρήγορο χρονικό διάστημα. Όμως, η μέθοδος bag of words, δεν κατάφερε να πετύχει υψηλό ποσοστό ταξινόμησης στις 8 αυτές κατηγορίες. Τι μπορεί να γίνει γι'αυτό;

Ο αναγνώστης δεν πρέπει να ξεχνάει ότι το KID dataset 2, είναι ένα σχετικά καινούριο dataset με δύσκολα δεδομένα. Αυτό σημαίνει ότι επιδέχεται διόρθωση και βελτίωση, έτσι ώστε οι ταξινομητές στο μέλλον να έχουν τη δυνατότητα να εκπαιδευτούν αρκετά αποδοτικά μ'αυτό. Εκτός από τις αλλαγές του dataset, στην βελτίωση του ποσοστού ταξινόμησης, μπορεί να συνεισφέρει και η δοκιμή νέων μεθόδων εξαγωγής χαρακτηριστικών. Τέτοιες μέθοδοι, μπορούν να δρούν σε διαφορετικούς χώρους χρωμάτων από αυτούς που ήδη δοκιμάστικαν. Μ'αυτόν τον τρόπο θα παρέχεται όχι μόνο η δυνατότητα της καλύτερης διάκρισης των ασθενείων μεταξύ τους, αλλά και η δυνατότητα της διάγνωσης ασθενειών, των οποίων οι αλλοιώσεις είναι δυσδιάκριτες και δεν διαφέρουν αρκετά από τον υγιή ιστό.

Προφανώς, κανένας δεν μπορεί να εγγυηθεί καλύτερα ή χειρότερα αποτελέσματα μέσω των αλλαγών αυτών. Οι νέοι επιστήμονες οφείλουν να είναι δημιουργικοί, με όρεξη και διάθεση για έρευνα, έτσι ώστε να δοκιμάζουν συνεχώς νέα πράγματα στον τομέα της πληροφορικής ιατρικής. Δεν πρέπει να ξεχνάει κανείς, πως η λύση ενός προβλήματος, οδηγεί πάντα στη δημιουργία νέων προβλημάτων. Η επίλυση αυτών των προβλημάτων όμως, κάνει την επιστήμη να προχωράει.

ΑΝΑΦΟΡΕΣ

1. https://www.mednutrition.gr/portal/ygeia/pathiseis-peptikoy/14698-katanoontas-to-peptiko-systima (προσπελάστηκε: 6/4/2019)
2. https://www.onmed.gr/ygeia/story/324835/deite-pos-leitourgei-to-peptiko-systima-vinteo (6/4/2019)
3. http://www.lib.teiher.gr/webnotes/seyp/nursingcourse/course8.pdf (προσπελάστηκε: 6/4/2019)
4. http://ebooks.edu.gr/modules/ebook/show.php/DSGL-A105/321/2155,7805/ (προσπελάστηκε: 6/4/2019)
5. https://www.care.gr/post/366/polypodes-paxeos-enterou (προσπελάστηκε: 6/4/2019)
6. http://ikee.lib.auth.gr/record/110754/files/09-01-09%20phd.pdf (προσπελάστηκε: 6/4/2019)
7. https://www.iatronet.gr/iatriko-lexiko/pagkreas.html (προσπελάστηκε: 6/4/2019)
8. https://pergamos.lib.uoa.gr/uoa/dl/frontend/file/lib/default/data/1308533/theFile (προσπελάστηκε: 6/4/2019)
9. https://emedicine.medscape.com/article/282920-overview (προσπελάστηκε: 6/4/2019)
10. https://dr-delis.gr/%CE%B1%CF%83%CF%85%CF%81%CE%BC%CE%B1%CF%84%CE%AE-%CE%B5%CE%BD%CE%B4%CE%BF%CF%83%CE%BA%CF%8C%CF%80%CE%B7%CF%83%CE%B7.html (προσπελάστηκε: 6/4/2019)
11. http://www.tzaneio.gr/wp-content/uploads/epistimonika_xronika/p12-3-3.pdf (προσπελάστηκε: 6/4/2019)
12. https://www.epege.gr/%CE%B8%CE%B1-%CE%BA%CE%AC%CE%BD%CE%B5%CF%84%CE%B5-%CE%AD%CE%BB%CE%B5%CE%B3%CF%87%CE%BF-%CE%BB%CE%B5%CF%80%CF%84%CE%BF%CF%8D-%CE%B5%CE%BD%CF%84%CE%AD%CF%81%CE%BF%CF%85-%CE%BC%CE%B5-%CE%B5%CE%BD%CE%B4/ (προσπελάστηκε: 6/4/2019)
13. https://www.kekos.gr/2014/12/11/%CE%B5%CE%BD%CE%B4%CE%BF%CF%83%CE%BA%CF%8C%CF%80%CE%B7%CF%83%CE%B7-%C2%B5%CE%B5-%CE%B1%CF%83%CF%8D%CF%81%C2%B5%CE%B1%CF%84%CE%B7-%CE%BA%CE%AC%CF%88%CE%BF%CF%85%CE%BB%CE%B1/ (προσπελάστηκε: 6/4/2019)
14. http://www.dr-med-cheilaris.gr/crbst_7.html (προσπελάστηκε: 6/4/2019)
15. M.Vasilakakis, D.Iakovidis, E.Spyrou, D.Chatzis, A.Kalaouzidis (2017) ‘Beyond Lesion Detection: Towards Semantic Interpretation of Endoscopy Videos’,Engineering applications of neural networks [online], 2 August 2017 vol.774,pp 379-390. (προσπελάστηκε: 6/4/2019)
16. M.Wafy, A.M Madbouly (2015), ‘Increase Efficiency of SURF using RGB Color Space’, (IJACSA) International Journal of Advanced Computer Science and Applications [onine], vol.6, No.8, Available from: https://www.researchgate.net/publication/326113105_Increase_Efficiency_of_SURF_using_RGB_Color_Space (προσπελάστηκε: 6/4/2019)
17. D. Iakovidis, A.Kalaouzidis (2014), ‘Automatic Lesion Detection in Wireless Capsule Endoscopy- A Simple Solution for a Complex Problem, IEEE International conference on Image Processing at Paris, France [online], October 2014, Available from: https://www.researchgate.net/publication/267041292_Automatic_Lesion_Detection_in_Wireless_Capsule_Endoscopy_-_A_Simple_Solution_for_a_Complex_Problem (προσπελάστηκε: 6/4/2019)
18. T.Rassem, B.Khoo (2011),’Object Class Recognition using Combination of Color SIFT Descriptors’ IEEE International Conference on Imaging Systems and Techniques (IST) [online], May 2011, Available from: https://www.researchgate.net/publication/215577705_Object_class_recognition_using_combination_of_color_SIFT_descriptors (προσπελάστηκε: 6/4/2019)

19. R.Kumar, Sreekumar K (2014), 'A Survey on Image Feature Descriptors', (IJCSIT) International Journal of Computer Science and Information Technologies, Vol.5 (6),2014, 7668-7673 {online], Adressed from: https://ijcsit.com/docs/Volume%205/vol5issue06/ijcsit20140506168.pdf (προσπελάστηκε: 6/4/2019)
20. https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html (προσπελάστηκε: 6/4/2019)
21. http://aishack.in/tutorials/sift-scale-invariant-feature-transform-introduction/ (προσπελάστηκε: 6/4/2019)
22. https://towardsdatascience.com/bag-of-visual-words-in-a-nutshell-9ceea97ce0fb (προσπελάστηκε: 6/4/2019)
23. http://www.cs.unc.edu/~lazebnik/spring09/lec18_bag_of_features.pdf (προσπελάστηκε: 6/4/2019)
24. https://prateekvjoshi.com/2014/08/17/image-classification-using-bag-of-words-model/ (προσπελάστηκε: 6/4/2019)
25. Koulaouzidis, D. K. Iakovidis, D. E. Yung, E. Rondonotti, U. Kopylov, J. N. Plevris, E. Toth, A. Eliakim, G. W. Johansson, W. Marlicz, and others, "KID Project: an internet-based digital video atlas of capsule endoscopy for research purposes," Endoscopy International Open, vol. 5, no. 06, pp. E477–E483, 2017;
26. https://www.lifewire.com/what-is-hsv-in-design-1078068 (προσπελάστηκε: 6/4/2019)
27. http://medtronicsolutions.medtronic.com/given-imaging
28. https://www.olympus-global.com/company/base/office.html
29. https://www.intromedic.com:549/eng/main/
30. http://english.jinshangroup.com/