



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ
ΥΛΙΚΩΝ

Ανάπτυξη Μοντέλου Εκτίμησης της Περιεχόμενης Ποσότητας
Υδατανθράκων στα Λαμβανόμενα Γεύματα από Φωτογραφικά
Στιγμιότυπα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σαραντέα-Μαγγαλούση Μύριαμ-Γεωργία

Επιβλέπουσα : Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Οκτώβριος 2016

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ
ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Ανάπτυξη Μοντέλου Εκτίμησης της Περιεχόμενης Ποσότητας Υδατανθράκων στα Λαμβανόμενα Γεύματα από Φωτογραφικά Στιγμιότυπα

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Σαραντέα-Μαγγαλούση Μύριαμ-Γεωργία

Επιβλέπουσα : Κωνσταντίνα Σ. Νικήτα
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 21^η Οκτωβρίου 2016.

(Υπογραφή)

.....
Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

(Υπογραφή)

.....
Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....
Ανδρέας-Γεώργιος
Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2016

(Υπογραφή)

.....

Σαραντέα-Μαγγαλούση Μύριαμ-Γεωργία

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Σαραντέα Μύριαμ, 2016 – All rights reserved

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξολοκλήρου ή μέρους αυτής, για εμπορικό ή κερδοσκοπικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Ερωτήματα που αφορούν τη χρήση της εργασίας για εμπορικό- κερδοσκοπικό σκοπό πρέπει να απευθύνονται αποκλειστικά στους συγγραφείς.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτή την εργασία εκφράζουν τους συγγραφείς και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου συμπεριλαμβανόμενων Σχολών, Τομέων και Μονάδων αυτού.

Περίληψη

Η παρακολούθηση της διατροφικής αξίας των λαμβανόμενων γευμάτων μπορεί να παίξει σημαντικό ρόλο στη μείωση των κινδύνων που σχετίζονται με χρόνιες ασθένειες που εξαρτώνται από τη διατροφή. Στη νόσο του Σακχαρώδη Διαβήτη συγκεκριμένα, η εκτίμηση της ποσότητας υδατανθράκων του γεύματος είναι καθοριστικής σημασίας για το σωστό σχεδιασμό του πλάνου υποκατάστασης της ινσουλίνης του ασθενούς και κατά συνέπεια, για την καθυστέρηση ή ακόμα και την αποφυγή εμφάνισης μακροπρόθεσμων επιπλοκών. Επομένως, αυτοματοποιημένα συστήματα που υποστηρίζουν τους ασθενείς στη μέτρηση του διατροφικού περιεχομένου σε θρεπτικά συστατικά μπορούν να αποτελέσουν πολύτιμα εργαλεία για την καθημερινότητά τους.

Στην παρούσα διπλωματική εργασία παρουσιάζεται ένα σύστημα αναγνώρισης τροφίμων που αποτελείται από δύο στάδια τα οποία εκτελούν εξαγωγή κατάλληλων χαρακτηριστικών από εικόνες των τροφίμων, και στη συνέχεια ταξινομούν τα τρόφιμα σε προκαθορισμένες κατηγορίες. Το σύνολο των δεδομένων που χρησιμοποιείται αποτελείται από 1200 εικόνες τροφίμων, χωρισμένες σε έξι κατηγορίες (ψωμί, κρέας, πατάτες ρύζι, ζυμαρικά και λαχανικά). Τα χαρακτηριστικά των εικόνων εξάγονται με χρήση του αλγορίθμου Speeded Up Robust Features (SURF) σε συνδυασμό με έναν αλγόριθμο εξαγωγής χρωματικής πληροφορίας, καθώς και της τεχνικής Τοπικών Δυαδικών Προτύπων (Local Binary Patterns, LBP). Για τη μείωση της διάστασης του διανύσματος των χρησιμοποιούμενων χαρακτηριστικών, εφαρμόζεται η μέθοδος του Σάκου Χαρακτηριστικών (Bag of Features, BoF), ενώ τελικά για την εκτέλεση της ταξινόμησης χρησιμοποιείται μια δομή βασισμένη στις Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines, SVMs). Η δομή αυτή αναφέρεται στη βιβλιογραφία ως Κατευθυνόμενος Ακυκλικός Γράφος από SVMs (DAG SVM), ενώ στην παρούσα εργασία χρησιμοποιείται μια βελτιστοποίηση αυτής. Η απόδοση του συστήματος αξιολογείται σε σύγκριση με έναν ταξινομητή μηχανών υποστήριξης διανυσμάτων πολλών κλάσεων (multiclass SVM) και τελικά, το προτεινόμενο σύστημα πετυχαίνει ακρίβεια ταξινόμησης 86.75% σε σχέση με το 86.92% που πετυχαίνει ο δεύτερος ταξινομητής.

Λέξεις Κλειδιά

Σακχαρώδης Διαβήτης, Σύστημα Αναγνώρισης Τροφίμων, Εκτίμηση Υδατανθράκων, Εξαγωγή Χαρακτηριστικών, Αλγόριθμος SURF, Τοπικά Δυαδικά Πρότυπα, Σάκος Χαρακτηριστικών, Μηχανές Διανυσμάτων Υποστήριξης, Κατευθυνόμενος Ακυκλικός Γράφος από SVMs.

Η σελίδα αυτή είναι σκόπιμα λευκή.

Abstract

Dietary intake monitoring can play an important role in reducing the risk of diet related chronic diseases. In diabetes mellitus in particular, the carbohydrate estimation of the upcoming meals can be crucial for the proper treatment planning of the patient and consequently, for the delay or even the prevention of the long-term disease related complications. Therefore, automatic systems that support patients to count the nutrient contents of their meals can provide valuable tools.

In this diploma thesis, a food recognition system is presented, which consists of two modules performing feature extraction and classification of food images, respectively. The dataset used consists of 1200 food images split into six categories (bread, meat, potatoes, rice, pasta and vegetables). The image features are extracted using Speeded Up Robust Features (SURF) alongside with a color feature extraction algorithm and Local Binary Patterns (LBP) method. The Bag Of Features (BOF) model is implemented in order to reduce the features space and finally, a Support Vector Machine (SVM) based classifier is used to perform the task of classification, in a decision tree structure. The specific structure is referred as Directed Acyclic Graph using SVMs (DAG SVM) in the literature, while in the present study an optimization of this structure is used. The classification performance is evaluated against a multiclass SVM classifier and eventually, the proposed system achieves a classification accuracy of 86.75% against 86.92% achieved by the multiclass SVM.

Keywords

Diabetes Mellitus, Food Recognition System, Carbohydrate Estimation, Feature Extraction, Speeded Up Robust Features Algorithm (SURF), Local Binary Patterns (LBP), Bag of Features (BoF), Support Vector Machines (SVMs), Directed Acyclic Graph using SVMs (DAG SVM).

Η σελίδα αυτή είναι σκόπιμα λευκή.

Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής μου εργασίας, θα ήθελα να ευχαριστήσω θερμά όλους όσους συνέβαλαν με οποιοδήποτε τρόπο στην επιτυχή εκπόνησή της, και ιδιαίτερα την Καθηγήτρια Κωνσταντίνα Νικήτα για την εμπιστοσύνη που μου έδειξε, καθώς και για την ευκαιρία που μου έδωσε να ασχοληθώ με ένα θέμα εξαιρετικά ενδιαφέρον για εμένα.

Θα ήθελα επίσης να εκφράσω ένα μεγάλο ευχαριστώ στις υποψήφιες διδάκτορες Ελένη Λίτσα και Καλλιόπη Δαλακλείδη, για την υπομονή, το ενδιαφέρον και την καθοδήγηση που μου παρείχαν καθ' όλη τη διάρκεια εκπόνησης της εργασίας.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για την πολύτιμη υποστήριξη που μου προσέφεραν όλα αυτά τα χρόνια.

Η σελίδα αυτή είναι σκόπιμα λευκή.

Πίνακας περιεχομένων

Περίληψη.....	5
Ευχαριστίες.....	9
Πίνακας Περιεχομένων.....	11
Ευρετήριο Εικόνων.....	13
Ευρετήριο Πινάκων.....	15
1 Εισαγωγή.....	17
2 Σακχαρώδης Διαβήτης.....	19
2.1 Ορισμός.....	19
2.2 Ο Ρόλος της Γλυκόζης.....	19
2.3 Ομοίωση της Γλυκόζης.....	20
2.3.1 Γλυκαγόνη.....	22
2.3.2 Ινσουλίνη.....	22
2.3.3 Αμυλίνη.....	24
2.3.4 Σωματοστασίνη.....	25
2.4 Τύποι Σακχαρώδους Διαβήτη.....	25
2.4.1 Σακχαρώδης Διαβήτης Τύπου 1 (ΣΔΤ1).....	25
2.4.2 Σακχαρώδης Διαβήτης Τύπου 2 (ΣΔΤ2).....	26
2.4.3 Σακχαρώδης Διαβήτης Κύησης (ΣΔΚ).....	27
2.4.4 Άλλοι τύποι ΣΔ.....	27
2.5 Γλυκαιμικός έλεγχος.....	27
2.6 Μέτρηση υδατανθράκων σε άτομα με ΣΔ.....	29
3 Μοντέλα Αυτόματης Αξιολόγησης Τροφών.....	33
3.1 Υπολογισμός υδατανθράκων.....	33
3.2 Δυσκολίες στην αυτόματη αξιολόγηση των τροφών.....	34
3.3 Διαδικασία αυτόματης αξιολόγησης των τροφών.....	35
3.3.1 Βαθμονόμηση κάμερας.....	36

3.3.2	<i>Λήψη φωτογραφίας του γεύματος</i>	36
3.3.3	<i>Κατάτμηση της εικόνας</i>	37
3.3.4	<i>Ταξινόμηση των τροφίμων</i>	38
3.3.5	<i>Υπολογισμός ποσότητας γεύματος</i>	38
3.3.6	<i>Εξαγωγή ποσότητας υδατανθράκων</i>	39
3.4	<i>Ταξινόμηση των τροφίμων - Βιβλιογραφική επισκόπηση</i>	39
3.4.1	<i>Γενικά</i>	39
3.4.2	<i>Snap-n-Eat</i>	41
3.4.3	<i>Σύστημα αναγνώρισης τροφίμων δύο σταδίων</i>	45
3.4.4	<i>Σύγκριση των προσεγγίσεων που παρουσιάστηκαν</i>	48
4	Ανάπτυξη Συστήματος Αναγνώρισης Τροφίμων	51
4.1	<i>Σχεδίαση συστήματος</i>	51
4.2	<i>Εξαγωγή χαρακτηριστικών</i>	53
4.2.1	<i>Εισαγωγή</i>	53
4.2.2	<i>Αλγόριθμος SURF</i>	54
4.2.3	<i>Αλγόριθμος Εξαγωγής Χρωματικών Χαρακτηριστικών (ColorExtractor)</i>	55
4.2.4	<i>Σάκος Χαρακτηριστικών (Bag of Features)</i>	56
4.2.5	<i>Τοπικά Δυαδικά Πρότυπα (Local Binary Patterns)</i>	59
4.3	<i>Ταξινόμηση</i>	60
4.3.1	<i>Μηχανές Υποστήριξης Διανυσμάτων (Support Vector Machines)</i>	60
4.3.2	<i>Δυαδικά SVMs</i>	62
4.3.3	<i>DAG SVMs</i>	66
5	Αξιολόγηση Συστήματος	71
5.1	<i>Σύνολο δεδομένων</i>	71
5.2	<i>Κριτήρια αξιολόγησης</i>	72
5.3	<i>Αποτελέσματα</i>	73
5.4	<i>Αξιολόγηση αποτελεσμάτων</i>	80
6	Μελλοντική Έρευνα	83
7	Βιβλιογραφία	85

Ευρετήριο Εικόνων

Εικόνα 2.1: Μηχανισμός ρύθμισης της γλυκόζης	21
Εικόνα 2.2: Σχηματικό διάγραμμα μεταβολισμού γλυκόζης λαμβάνοντας υπόψη δύο ορμόνες ελέγχου (ινσουλίνη – γλυκαγόνη) της γλυκόζης.....	24
Εικόνα 3.1: Μετασχηματισμός του σημείου τριών διαστάσεων $P = (X, Y, Z)$ του πραγματικού κόσμου σε ένα σημείο δύο διαστάσεων $P_c = (u, v)$	36
Εικόνα 3.2: Λήψη εικόνας σε συγκεκριμένο φόντο.....	37
Εικόνα 3.3: Κατάτμηση της εικόνας για τον υπολογισμό των διαφορετικών τροφίμων.	38
Εικόνα 3.4: Σύστημα Snap ‘n Eat [19].....	42
Εικόνα 3.5: Σύγκριση SIFT και HOG χαρακτηριστικών σε διάφορες κλίμακες. [19].....	43
Εικόνα 3.6: DAG οντολογία τροφίμων	44
Εικόνα 3.7: Απόδοση ταξινόμησης με χρήση 4 διαφορετικών συνόλων χαρακτηριστικών και το συνδυασμό αυτών με χρήση της μεθόδου 5-fold cross validation.....	45
Εικόνα 3.8: Αρχιτεκτονική συστήματος που αναπτύχθηκε. [23].....	46
Εικόνα 3.9: Μήτρα σύγκρισης του προτεινόμενου συστήματος [23]	47
Εικόνα 3.10: Σύγκριση τοπικών περιγραφέων σε όρους συνολικής αξιολόγησης συστήματος. Χρησιμοποιήθηκαν 7 χρωματικοί περιγραφείς και 7 περιγραφείς υφής.....	48
Εικόνα 4.1: Σχηματικό διάγραμμα του συστήματος που αναπτύχθηκε.	52
Εικόνα 4.2: Γραφική απεικόνιση της διαδικασίας δημιουργίας των οπτικών λέξεων μέσω της μεθόδου Bag of Features.	58
Εικόνα 4.3: Γραφική απεικόνιση της διαδικασίας εξαγωγής του διανύσματος των χαρακτηριστικών που χρησιμοποιείται ως είσοδος στον ταξινομητή με χρήση του Bag of Features.	58
Εικόνα 4.4: Υπολογισμός LBP χαρακτηριστικών σε γειτονία 3x3 pixel.....	60
Εικόνα 4.5: Πρόβλημα ταξινόμησης δύο κλάσεων με χρήση SVM.	65
Εικόνα 4.6: Απόφαση της μεθόδου DDAG για ένα πρόβλημα με 4 κλάσεις.....	69
Εικόνα 5.1: Δείγματα των εικόνων του συνόλου δεδομένων μετά το πρώτο βήμα της προεπεξεργασίας. Συνολικά χρησιμοποιήθηκαν 1200 εικόνες που ανήκουν σε 6 κατηγορίες.	72
Εικόνα 5.2: Σχηματική αναπαράσταση της μεθόδου 10-fold cross validation.	73
Εικόνα 5.3: Σύγκριση αλγορίθμων εξαγωγής χαρακτηριστικών SURF και Color Extractor, με βάση τη συνολική απόδοση του συστήματος, μεταβάλλοντας τον αριθμό των visual words.	74

Εικόνα 5.4: Σύγκριση της μέγιστης απόδοσης των αλγορίθμων SURF, Color Extractor και LBP.....	74
Εικόνα 5.5: Σύγκριση των αποτελεσμάτων με χρήση συνδυασμού χαρακτηριστικών SURF, Color και LBP, με ταυτόχρονη μεταβολή του visual dictionary για τους πρώτους δύο.	76
Εικόνα 5.6: Σύγκριση αποτελεσμάτων ταξινομητών DAG SVM και Multiclass SVM, με χρήση συνδυασμού χαρακτηριστικών SURF, Color και LBP, και ταυτόχρονη μεταβολή του visual dictionary για τους πρώτους δύο.....	77

Ευρετήριο Πινάκων

Πίνακας 3.1: Σύγκριση διαφορετικών προσεγγίσεων αναγνώρισης τροφίμων.....	49
Πίνακας 5.1: Σύγκριση αλγορίθμων SURF, Color Extractor και LBP.	75
Πίνακας 5.2: Απόδοση συστήματος με χρήση συνδυασμού χαρακτηριστικών SURF, Color και LBP.	76
Πίνακας 5.3: Σύγκριση απόδοσης των αποτελεσμάτων των ταξινομητών DAG SVM και Multiclass SVM.....	77
Πίνακας 5.3: Μήτρα Σύγκρισης 1ης Επανάληψης	78
Πίνακας 5.4: Μήτρα Σύγκρισης 2ης Επανάληψης	78
Πίνακας 5.5: Μήτρα Σύγκρισης 3ης Επανάληψης	78
Πίνακας 5.6: Μήτρα Σύγκρισης 4ης Επανάληψης	78
Πίνακας 5.7: Μήτρα Σύγκρισης 5ης Επανάληψης	79
Πίνακας 5.8: Μήτρα Σύγκρισης 6ης Επανάληψης	79
Πίνακας 5.9: Μήτρα Σύγκρισης 7ης Επανάληψης	79
Πίνακας 5.10: Μήτρα Σύγκρισης 8ης Επανάληψης	79
Πίνακας 5.11: Μήτρα Σύγκρισης 9ης Επανάληψης	80
Πίνακας 5.12: Μήτρα Σύγκρισης 10ης Επανάληψης	80
Πίνακας 5.13: Μήτρα Σύγκρισης με τη μέση τιμή κάθε πεδίου στις 10 επαναλήψεις.....	80

Η σελίδα αυτή είναι σκόπιμα λευκή.

1

Εισαγωγή

Ο Σακχαρώδης Διαβήτης (ΣΔ) είναι μια μεταβολική ασθένεια η οποία χαρακτηρίζεται από αύξηση της συγκέντρωσης του σακχάρου στο αίμα (υπεργλυκαιμία) και από διαταραχή του μεταβολισμού της γλυκόζης, των λιπιδίων και των πρωτεϊνών, είτε ως αποτέλεσμα ελαττωμένης έκκρισης ινσουλίνης είτε λόγω ελάττωσης της ευαισθησίας των κυττάρων του σώματος στην ινσουλίνη. Η ινσουλίνη είναι μια ορμόνη που εκκρίνεται από το πάγκρεας και είναι απαραίτητη για τη μεταφορά της γλυκόζης που λαμβάνεται από τις τροφές, μέσα στα κύτταρα. Όταν το πάγκρεας δεν παράγει αρκετή ινσουλίνη ή η ινσουλίνη που παράγει δεν δρα σωστά, τότε η γλυκόζη που λαμβάνεται από τις τροφές δεν εισέρχεται στα κύτταρα ώστε να έχουν την απαραίτητη ενέργεια για τη λειτουργία τους και παραμένει στο αίμα με αποτέλεσμα την αύξηση των επιπέδων της και άρα την εκδήλωση της νόσου.

Ο στόχος της θεραπείας στα άτομα με ΣΔ είναι να διατηρούνται τα επίπεδα του σακχάρου αίματος όσο το δυνατόν πλησιέστερα στα φυσιολογικά, αποφεύγοντας τις οξείες επιπλοκές της νόσου, όπως την υπογλυκαιμία ή την κετοξέωση καθώς και τις χρόνιες επιπλοκές, δηλαδή τη μικροαγγειοπάθεια (νεφροπάθεια, αμφιβληστροειδοπάθεια, νευροπάθεια) και τη μακροαγγειοπάθεια (στεφανιαία νόσος, εγκεφαλικά επεισόδια και περιφερική αρτηριακή νόσος) έχοντας, όμως, καλή ποιότητα ζωής.

Στον ΣΔ τύπου 1, ο οποίος χαρακτηρίζεται από πλήρη έλλειψη ενδογενούς ινσουλίνης, η ινσουλινοθεραπεία κρίνεται απαραίτητη όχι μόνο για την καλή ρύθμιση του σακχάρου αλλά και για την ίδια την επιβίωση του ατόμου. Όμως, προκειμένου να καθοριστεί

η βέλτιστη δόση ινσουλίνης, είναι απαραίτητο να είναι γνωστή η ποσότητα και η σύνθεση της τροφής που πρόκειται να καταναλωθεί και ιδιαίτερα η περιεκτικότητα της σε υδατάνθρακες. Η κάθε τροφή έχει διαφορετικό ποσοστό υδατανθράκων, επομένως ο οργανισμός θα αποκριθεί διαφορετικά, όσον αφορά την αύξηση της γλυκόζης στο αίμα, με τη λήψη διαφορετικής τροφής. Ο υπολογισμός, όμως, των υδατανθράκων από τους διαβητικούς ασθενείς δεν είναι μια εύκολη διαδικασία.

Στόχος επομένως της συγκεκριμένης διπλωματικής εργασίας είναι η επίλυση του προβλήματος που αντιμετωπίζουν οι διαβητικοί ασθενείς σχετικά με την ακριβή μέτρηση των υδατανθράκων των γευμάτων τους. Η προσέγγιση του συγκεκριμένου προβλήματος γίνεται με την ανάπτυξη ενός μοντέλου εκτίμησης της περιεχόμενης ποσότητας υδατανθράκων μέσα από φωτογραφικά στιγμιότυπα των λαμβανόμενων γευμάτων. Τα βασικά στάδια επεξεργασίας που ακολουθεί ένα τέτοιο μοντέλο είναι τα εξής: λήψη της φωτογραφίας, διαχωρισμός των επιμέρους τροφίμων της εικόνας, ταξινόμηση των διαφορετικών τροφίμων και εκτίμηση περιεκτικότητας υδατανθράκων.

Στην παρούσα εργασία μελετάται το στάδιο της ταξινόμησης των επιμέρους τροφίμων του γεύματος, μέσω της χρήσης μεθόδων μηχανικής μάθησης για την ταξινόμησή τους και πιο συγκεκριμένα, μέσω της χρήσης μηχανών διανυσματικής υποστήριξης. Επιπλέον, γίνεται η μελέτη και η χρήση διάφορων αλγόριθμων εξαγωγής χαρακτηριστικών, προκειμένου οι εικόνες που δέχεται το μοντέλο στην είσοδό του να αναπαρασταθούν με διανύσματα κατάλληλα για την εφαρμογή των ταξινομητών.

Η οργάνωση του τόμου γίνεται σε 6 κεφάλαια.

Στο 2^ο κεφάλαιο παρουσιάζονται στοιχεία για τη νόσο του Σακχαρώδη Διαβήτη καθώς και για το φυσιολογικό μεταβολισμό της γλυκόζης.

Στο 3^ο κεφάλαιο γίνεται μια ανάλυση στα στάδια που ακολουθούν τα μοντέλα αυτόματης αξιολόγησης των τροφών, καθώς και μια βιβλιογραφική επισκόπηση σε αντίστοιχες εφαρμογές και μεθόδους.

Στο 4^ο κεφάλαιο παρουσιάζεται η μεθοδολογία που ακολουθήθηκε και οι αλγόριθμοι που χρησιμοποιήθηκαν.

Στο 5^ο κεφάλαιο παρουσιάζονται τα αποτελέσματα από την αξιολόγηση του μοντέλου που αναπτύχθηκε.

Τέλος, στο 6^ο κεφάλαιο δίνονται τα συμπεράσματα της εργασίας, καθώς και προτάσεις για μελλοντική μελέτη.

2

Σακχαρώδης Διαβήτης

2.1 Ορισμός

Ο Σακχαρώδης Διαβήτης (ΣΔ) είναι μια μεταβολική πάθηση, ένα σύνδρομο με ετερογενές και πολυπαραγοντικό υπόστρωμα, όπου χαρακτηρίζεται από διαταραχή του μεταβολισμού των υδατανθράκων, λιπών και πρωτεϊνών, η οποία οφείλεται σε έλλειψη ινσουλίνης. Η έλλειψη μπορεί να είναι πλήρης ή μερική ή σχετική. Ως σχετική χαρακτηρίζεται η έλλειψη ινσουλίνης, όταν, παρά τα αυξημένα επίπεδά της στο αίμα, δεν επαρκεί για την κάλυψη των αναγκών του μεταβολισμού λόγω παρεμπόδισης της δράσης της στους περιφερικούς ιστούς (αντίσταση στην ινσουλίνη). Η κύρια έκφραση της διαταραχής του μεταβολισμού στο ΣΔ είναι η αύξηση των επιπέδων γλυκόζης στο αίμα. [2]

2.2 Ο Ρόλος της Γλυκόζης

Ο Σακχαρώδης Διαβήτης, όπως αναφέρθηκε, χαρακτηρίζεται από αύξηση της συγκέντρωσης της γλυκόζης στο αίμα (υπεργλυκαιμία) και οφείλεται σε διαταραχή του μεταβολισμού των σακχάρων, τα οποία εισάγονται στον οργανισμό μέσω της διατροφής. Η χρόνια υπεργλυκαιμία στον ΣΔ σχετίζεται με την πρόκληση μακροχρόνιων βλαβών, δυσλειτουργίας ή ανεπάρκειας διάφορων οργάνων, και ιδιαίτερας των οφθαλμών, των

νεφρών, των νεύρων, της καρδιάς και των αγγείων, και αποτελεί παγκοσμίως μια από τις σοβαρότερες αιτίες θνησιμότητας, νοσηρότητας και αναπηρίας.

Ο ΣΔ σχετίζεται με τη λειτουργία του παγκρέατος, το οποίο είναι ένας αδένας που βρίσκεται στον οπισθοπεριτοναϊκό χώρο και παράγει σημαντικές ορμόνες. Συγκεκριμένα, από το πάγκρεας παράγονται η ινσουλίνη και η γλυκαγόνη οι οποίες ρυθμίζουν τα επίπεδα της γλυκόζης στο σώμα, καθώς και άλλες ορμόνες που συμβάλλουν στην πέψη, βοηθούν δηλαδή τον οργανισμό να αποθηκεύσει ενέργεια από τις τροφές και να ρυθμίσει το σάκχαρο του αίματος. Προκειμένου να λειτουργήσει σωστά ο οργανισμός θα πρέπει, η ενέργεια που προσλαμβάνει να ισούται με την ενέργεια που καταναλώνει. Η σημαντικότερη πηγή ενέργειας στον οργανισμό είναι η γλυκόζη, διότι η διάσπασή της οδηγεί στην παραγωγή ενώσεων με δεσμούς υψηλής χημικής ενέργειας. Επομένως, η διατήρηση των επιπέδων γλυκόζης στο αίμα κρίνεται ιδιαίτερα σημαντική.

2.3 Ομοιόσταση της Γλυκόζης

Όλοι ο ζωτικοί οργανισμοί εξασφαλίζουν την ενέργεια που χρειάζονται για να επιβιώσουν και να επιτελέσουν τις λειτουργίες τους από το περιβάλλον τους, και συγκεκριμένα από την τροφή. Οι πηγές ενέργειας του ανθρώπινου οργανισμού είναι τα σάκχαρα και τα λιπαρά οξέα, καθώς και το οξυγόνο, ενώ σε περιόδους αστίας πηγή ενέργειας είναι τα αμινοξέα.

Τα συστατικά των τροφών, σάκχαρα, λιπαρά οξέα και πρωτεΐνες δεν μπορούν να αξιοποιηθούν άμεσα και για το λόγο αυτό μετατρέπονται, με τη διαδικασία του μεταβολισμού, σε άλλες χημικές ενώσεις οι οποίες μπορούν να αξιοποιηθούν με δύο τρόπους:

1. Να οξειδωθούν και να παράγουν ενέργεια (καταβολισμός): Μέσα στο κύτταρο τα συστατικά των τροφών αντιδρούν χημικά με το οξυγόνο, με την επίδραση διαφόρων ενζύμων, με αποτέλεσμα την απελευθέρωση ενέργειας.
2. Να χρησιμοποιηθούν ως πρώτη ύλη για τη σύνθεση δομικών ή λειτουργικών μορίων που είναι απαραίτητα για τον οργανισμό (αναβολισμός).

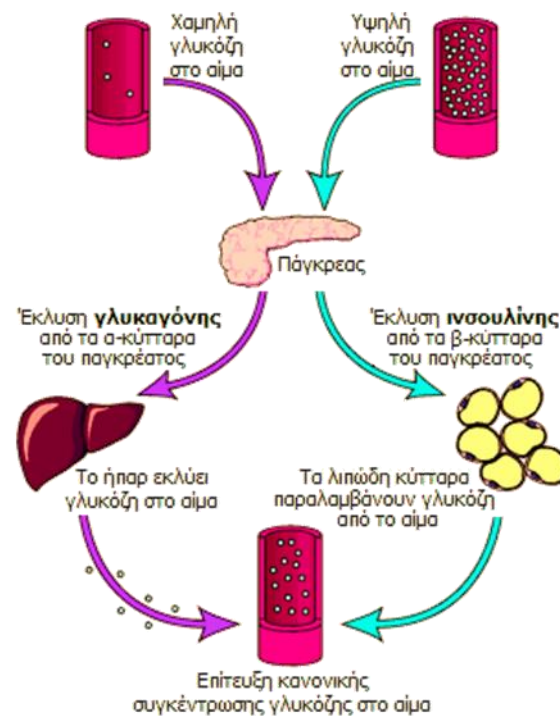
Για την ρύθμιση αυτών των δύο διαδικασιών απαιτείται η συντονισμένη δράση της ινσουλίνης και της γλυκαγόνης, οι οποίες έχουν αντίθετη δράση. Ο μοριακός λόγος των συγκεντρώσεων αυτών των ορμονών ελέγχει τον μεταβολισμό και συντελεί στη σταθεροποίηση των επιπέδων της γλυκόζης στο αίμα (ομοιόσταση της γλυκόζης).

Όταν αυξάνεται η συγκέντρωση της γλυκόζης στο αίμα μετά από ένα γεύμα, τα β-κύτταρα των νησιδίων Langerhans που βρίσκονται στο πάγκρεας εκκρίνουν περισσότερη ινσουλίνη με σκοπό να μειωθεί η γλυκόζη στο αίμα. Αντίθετα, όταν μειώνεται η συγκέντρωση της γλυκόζης στο αίμα, τα α-κύτταρα των νησιδίων Langerhans του παγκρέατος εκκρίνουν περισσότερη γλυκαγόνη προκειμένου να αυξηθεί η γλυκόζη στο αίμα.

Υπεύθυνο για την έκκριση της ινσουλίνης και της γλυκαγόνης, αλλά και άλλων αρκετών ορμονών όπως η σωματοστασίνη και το παγκρεατικό πολυπεπτίδιο, είναι το πάγκρεας. Το πάγκρεας είναι ένας αδένας ο οποίος ανήκει στο πεπτικό και στο ενδοκρινικό σύστημα. Εκτός από την έκκριση ορμονών, το πάγκρεας συμμετέχει και στη διαδικασία της πέψης, εκκρίνοντας παγκρεατικό υγρό, το οποίο περιέχει ένζυμα τα οποία βοηθούν στην απορρόφηση των θρεπτικών συστατικών.

Η νησίδα του Langerhans είναι ομάδες κυττάρων στο πάγκρεας, τα οποία παράγουν ορμόνες. Υπάρχουν, κατά κύριο λόγο, τρεις κατηγορίες τέτοιων κυττάρων:

- α-κύτταρα τα οποία παράγουν τη γλυκαγόνη
- β-κύτταρα τα οποία παράγουν την ινσουλίνη και την αμυλίνη
- δ-κύτταρα τα οποία παράγουν την σωματοστασίνη



Εικόνα 2.1: Μηχανισμός ρύθμισης της γλυκόζης

Όπως θα φανεί και στις επόμενες παραγράφους, κάθε ορμόνη έχει ένα σημαντικό ρόλο στη ρύθμιση των επιπέδων της γλυκόζης στο αίμα.

2.3.1 Γλυκαγόνη

Η γλυκαγόνη είναι μια πεπτιδική ορμόνη που αποτελείται από 29 αμινοξέα και παράγεται στο πάγκρεας από τα α-κύτταρα στη νησίδα Langerhans. Η γλυκαγόνη κινείται σε δύο άξονες όσον αφορά το μεταβολισμό της γλυκόζης. Αφενός διασπά το γλυκογόνο του ήπατος και αφετέρου αυξάνει την γλυκονεογένεση στο ήπαρ. Και οι δύο αυτές διαδικασίες αυξάνουν τη συγκέντρωση της γλυκόζης στο αίμα. Επομένως γίνεται αντιληπτό ότι η συγκεκριμένη ορμόνη έχει αντίστροφη δράση από την ινσουλίνη και ο κύριος ρόλος της είναι η αύξηση της γλυκόζης στο αίμα.

2.3.2 Ινσουλίνη

Η ινσουλίνη είναι και αυτή μία πεπτιδική ορμόνη που αποτελείται συνολικά από 51 αμινοξέα. Πιο συγκεκριμένα, αποτελείται από 2 πολυπεπτιδικές αλυσίδες οι οποίες ενώνονται μεταξύ τους με δισουλφιδικούς δεσμούς, με 21 αμινοξέα στην Α-αλυσίδα και 30 στην Β-αλυσίδα. Παράγεται στο πάγκρεας από β-κύτταρα, τα οποία είναι οργανωμένα μαζί με άλλους κυτταρικούς τύπους στις νησίδες Langerhans. Για να παραχθεί θα πρέπει να υπάρχει μεγάλη συγκέντρωση γλυκόζης στο αίμα ή αυτό να περιέχει πολλά αμινοξέα ή λιπαρά οξέα. Τότε τα β-κύτταρα δεχόμενα το ερέθισμα αντιδρούν, απελευθερώνοντας ινσουλίνη, η οποία στη συνέχεια ακολουθώντας την κυκλοφορία του αίματος δρα στα διάφορα όργανα, όπου και είναι απαραίτητη για την ομαλή λειτουργία τους.

Η ινσουλίνη εκτελεί μια σειρά από λειτουργίες σημαντικές στο μεταβολισμό τόσο των υδατανθράκων, όσο των λιπών (λιπιδίων) και των πρωτεϊνών (πρωτιδίων). Έχει την ικανότητα να αυξάνει την διαπερατότητα των κυττάρων του μυϊκού και του λιπώδη ιστού στη γλυκόζη. Η ινσουλίνη δεν επιδρά με τον ίδιο τρόπο στο ήπαρ, στο οποίο η γλυκόζη διεισδύει και χωρίς την παρουσία της. Πιο συγκεκριμένα, στους ιστούς έχουμε λήψη και καμιά φορά αποθήκευση της γλυκόζης για περαιτέρω επεξεργασία, ενώ στο ήπαρ έχουμε μόνο αποθήκευσή της. Όμως με τη διαδικασία λήψης και αποθήκευσης της γλυκόζης μειώνεται η συγκέντρωσή της στο αίμα, ενώ ταυτόχρονα τα β-κύτταρα μειώνουν την ποσότητα της ινσουλίνης που παράγουν.

Πιο συγκεκριμένα, η ινσουλίνη βοηθάει την εναπόθεση γλυκόζης στο ήπαρ με τη μορφή γλυκογόνου, το οποίο δεν είναι τίποτα άλλο από ένα σύμπλεγμα σακχάρων το οποίο και απελευθερώνει σάκχαρο κάθε φορά που έχουμε υπογλυκαιμία και η συγκέντρωση της

γλυκόζης στο αίμα είναι πολύ μικρή. Ο μηχανισμός με τον οποίο η ινσουλίνη προκαλεί πρόσληψη και εναπόθεση γλυκόζης στο ήπαρ ακολουθεί τα παρακάτω στάδια, τα οποία γίνονται ταυτόχρονα:

- Αναστολή της φωσφορυλάσης, δηλαδή του ενζύμου που προκαλεί τη διάσπαση του γλυκογόνου σε γλυκόζη.
- Αύξηση του ρυθμού πρόσληψης της γλυκόζης στο αίμα μέσω των ηπατικών κυττάρων, αυξάνοντας έτσι τη δραστηριότητα της γλυκοκινάσης. Πρόκειται για εκείνο το ένζυμο που προκαλεί την αρχική διάσπαση του γλυκογόνου σε γλυκόζη. Η προκύπτουσα γλυκόζη παγιδεύεται μέσα στα ηπατικά κύτταρα, αφού και η ημιπερατή μεμβράνη απαγορεύει προς τα έξω κίνηση.
- Αύξηση της λειτουργίας των ενζύμων που προάγουν τη σύνθεση του γλυκογόνου.

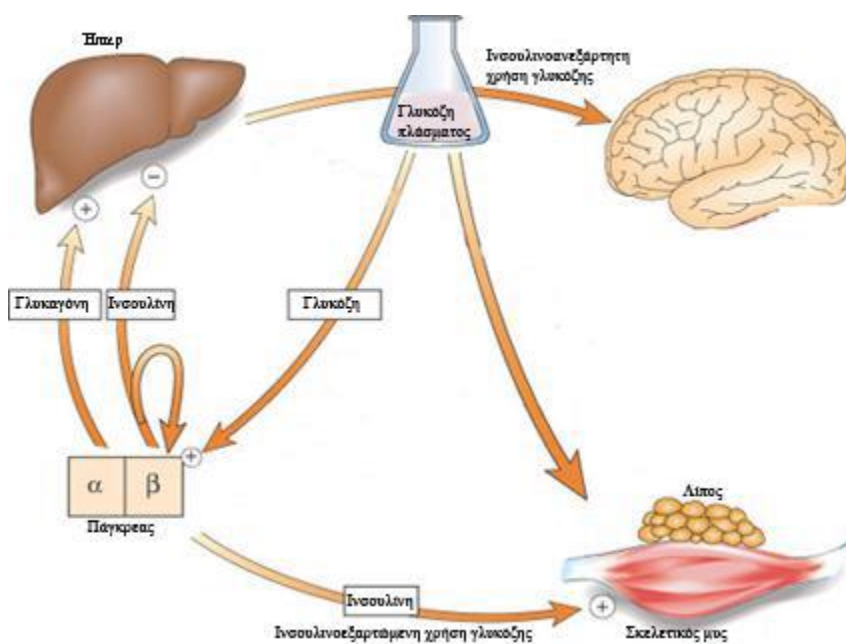
Όταν η συγκέντρωση της γλυκόζης στο αίμα αρχίζει να μειώνεται, λαμβάνουν χώρα μια σειρά από μηχανισμούς που προκαλούν την απελευθέρωσή της από το ήπαρ και την εισαγωγή της στην κυκλοφορία του αίματος. Συγκεκριμένα:

- Η ελάττωση της γλυκόζης προκαλεί μείωση της έκκρισης ινσουλίνης από το πάγκρεας.
- Αυτή η ελάττωση της ινσουλίνης προκαλεί αντιστροφή των παραπάνω φαινομένων αποθήκευσης γλυκόζης.
- Παράλληλα, ενεργοποιείται το ένζυμο φωσφορυλάση που προκαλεί τη διάσπαση του γλυκογόνου σε φωσφορική γλυκόζη.
- Τέλος, το ένζυμο φωσφατάση της γλυκόζης προκαλεί απόσπαση της φωσφορικής ρίζας από τη γλυκόζη, επιτρέποντας τη διάχυσή της στο αίμα.

Επομένως, το ήπαρ απομακρύνει με τη βοήθεια της ινσουλίνης τη γλυκόζη όταν αυτή βρίσκεται σε περίσσεια και την επαναφέρει όταν έχουμε μείωσή της κάτω από τα κανονικά επίπεδα.

Η ινσουλίνη, τέλος, εκκρίνεται όταν η συγκέντρωση της γλυκόζης στο αίμα είναι > 3 mmol/L και καθώς η συγκέντρωση της γλυκόζης αυξάνεται πέρα από αυτό το κατώφλι η έκκριση της ινσουλίνης γίνεται πιο έντονη. Όταν τα επίπεδα γλυκόζης στο αίμα είναι πάνω από 5 mmol/L τα β -κύτταρα αυξάνουν την παραγωγή ινσουλίνης. Έπειτα από ένα γεύμα, η έκκριση ινσουλίνης γίνεται σε δύο στάδια: Μια αρχική απότομη απελευθέρωση προσχηματισμένης ινσουλίνης, η οποία ακολουθείται από μία πιο αργή έκκριση που

εξαρτάται από τα επίπεδα από τα επίπεδα της γλυκόζης. Μακροχρόνια έκκριση ινσουλίνης συμβαίνει όταν τα επίπεδα γλυκόζης παραμένουν υψηλά.



Εικόνα 2.2: Σχηματικό διάγραμμα μεταβολισμού γλυκόζης λαμβάνοντας υπόψη δύο ορμόνες ελέγχου (ινσουλίνη – γλυκαγόνη) της γλυκόζης

2.3.3 Αμυλίνη

Η αμυλίνη είναι μια πεπτιδική ορμόνη που αποτελείται από 37 αμινοξέα. Βρίσκεται μαζί με την ινσουλίνη στα εκκριτικά κοκκία των β-κυττάρων του παγκρέατος, αλλά η έκκρισή της γίνεται σε πολύ μικρότερο ρυθμό από αυτή. Η έκκριση της αμυλίνης προκαλείται από τα ίδια ερεθίσματα που διεγείρουν και την έκκριση της ινσουλίνης, ενώ μελέτες έχουν δείξει ότι οι συγκεντρώσεις των δύο αυτών ορμονών στο πλάσμα αυξάνονται και ελαττώνονται παράλληλα, σε απόκριση των λαμβανόμενων γευμάτων.

Η αμυλίνη συμπληρώνει τη δράση της ινσουλίνης μέσω δύο μηχανισμών: αναστέλλει τη μεταγευματική έκκριση γλυκαγόνης και καθυστερεί την γαστρική κένωση κατά τη λήψη γεύματος. Επομένως, δημιουργεί αίσθημα κορεσμού, εμποδίζοντας με αυτό τον τρόπο την εμφάνιση μεταγευματικών υπογλυκαιμιών. Μελέτες έχουν δείξει ότι, η αμυλίνη είναι ανεπαρκής στα άτομα με ΣΔΤ1 και μειωμένης δράσης στα άτομα με ΣΔΤ2.

2.3.4 Σωματοστασίνη

Η σωματοστασίνη είναι μια ορμόνη που εκκρίνεται από τα δ-κύτταρα των νησιδίων Langerhans του παγκρέατος, καθώς επίσης και από το στομάχι και το έντερο. Η έκκρισή της διεγείρεται από την παρουσία τροφής στο γαστρεντερικό σωλήνα και από την άνοδο της στάθμης της γλυκόζης, των αμινοξέων και των λιπαρών οξέων στο αίμα.

Η σωματοστασίνη ασκεί ανασταλτική δράση στην έκκριση ινσουλίνης και γλυκαγόνης. Παράλληλα περιορίζει την κινητικότητα του εντέρου και επιβραδύνει την απορρόφηση των θρεπτικών συστατικών. Με αυτό τον τρόπο, αυξάνει τη χρονική διάρκεια της πέψης και παρακωλύει την ταχεία εξάντληση της διαθέσιμης τροφής.

2.4 Τύποι Σακχαρώδους Διαβήτη

Η ρύθμιση των επιπέδων της συγκέντρωσης της γλυκόζης στο αίμα είναι, επομένως, μια περίπλοκη διαδικασία και πραγματοποιείται σε πολλά επίπεδα. Σε περίπτωση που ο μηχανισμός της ομοιόστασης της γλυκόζης αποτύχει να κρατήσει τα επίπεδα της γλυκόζης μέσα σε φυσιολογικά πλαίσια, δημιουργείται μία παθολογική κατάσταση η οποία αποτελεί τον Σακχαρώδη Διαβήτη.

Ο Σακχαρώδης Διαβήτης έχει χρόνια πορεία και μπορεί να προκαλέσει μια σειρά σοβαρών επιπλοκών, όπως καρδιαγγειακή νόσο, χρόνια νεφρική ανεπάρκεια, βλάβες του αμφιβληστροειδούς, βλάβες των νεύρων, στυτική δυσλειτουργία κ.ά. Πρωτεύοντα ρόλο στη θεραπεία του σακχαρώδους διαβήτη παίζει η χορήγηση ινσουλίνης.

Υπάρχουν αρκετοί τύποι σακχαρώδους διαβήτη, από τους οποίους οι πιο διαδεδομένοι είναι οι παρακάτω:

2.4.1 Σακχαρώδης Διαβήτης Τύπου 1 (ΣΔΤ1)

Ο ΣΔΤ1 συχνά αποκαλείται και ινσουλινο-εξαρτώμενος ή νεανικός διαβήτης. Χαρακτηρίζεται από καταστροφή των β-κυττάρων του παγκρέατος, τα οποία είναι υπεύθυνα για την παραγωγή ινσουλίνης, με αποτέλεσμα ολική έλλειψη ή ελάχιστη έκκριση ινσουλίνης. Η ευαισθησία των κυττάρων στην ινσουλίνη είναι συνήθως φυσιολογική, ιδιαίτερα στα πρώιμα στάδια. Ο τύπος αυτός αποτελεί την κυριότερη αιτία διαβήτη σε παιδιά, μπορεί όμως να προσβάλλει και τους ενήλικες. Η καταστροφή των β-κυττάρων του παγκρέατος είναι στην πλειοψηφία των περιπτώσεων αυτοάνοσης αιτιολογίας.

Ο διαβήτης τύπου 1 εισβάλλει συνήθως απότομα και οδηγεί συχνά στην ανάπτυξη κέτωσης και διαβητικής κετοξέωσης εάν το έλλειμμα ινσουλίνης δεν καλυφθεί. Ο ασθενής είναι απόλυτα εξαρτημένος από την εξωγενή χορήγηση ινσουλίνης προκειμένου τα επίπεδα σακχάρου του αίματος να διατηρηθούν σε φυσιολογικά επίπεδα.

Στην Ευρώπη ο επιπολασμός του κυμαίνεται 0.6 - 2.6‰ και η επίπτωσή του, σε άτομα ηλικίας ≤ 15 ετών, 9.7 - 49.0/100000 ετών/κατ' έτος. Στην Ελλάδα η επίπτωση είναι 9.7/100000/κατ' έτος.

2.4.2 Σακχαρώδης Διαβήτης Τύπου 2 (ΣΔΤ2)

Ο ΣΔΤ2 συχνά αποκαλείται και μη ινσουλινο-εξαρτώμενος διαβήτης. Οφείλεται σε προοδευτική μείωση της επαρκούς ανταπόκρισης του β-κυττάρου για την έκκριση της ινσουλίνης που απαιτείται για την αντιμετώπιση των μεταβολικών αναγκών. Περιλαμβάνει όλο το φάσμα συνδυασμών, από την κατ' εξοχήν αντίσταση στην ινσουλίνη με σχετικά μικρή έλλειψη ινσουλίνης, μέχρι τη σημαντική μείωση της έκκρισης ινσουλίνης με μικρότερη αντίσταση.

Ο ΣΔΤ2 είναι η συχνότερη αιτία διαβήτη στους ενήλικες, ενώ πρόκειται για μια πολύ-παραγοντική ασθένεια. Η κληρονομικότητα είναι ένας σημαντικός παράγοντας, αλλά πολύ σημαντική επίδραση έχουν και περιβαλλοντικοί παράγοντες. Πιο συγκεκριμένα, παράγοντες που συνδέονται με τον σύγχρονο τρόπο ζωής, όπως η καθιστική ζωή και οι αλλαγές στις διατροφικές συνήθειες, έχουν ενοχοποιηθεί έντονα. Ένας τέτοιος παράγοντας που μπορεί να οδηγήσει στην ανάπτυξη διαβήτη τύπου 2 είναι η παχυσαρκία. Η παχυσαρκία προδιαθέτει στην ανάπτυξη ινσουλινο-αντοχής πιθανόν λόγω της παραγωγής από το λιπώδη ιστό ουσιών που ελαττώνουν την ευαισθησία των κυττάρων στην ινσουλίνη. Άλλοι προδιαθεσικοί παράγοντες είναι η ηλικία και το οικογενειακό ιστορικό.

Από γενετικής πλευράς, στον ΣΔΤ2 εμπλέκονται πολλά γονίδια. Επιπλέον, ο ΣΔΤ2 χαρακτηρίζεται από μεγάλη ετερογένεια καθώς διαφορετικός συνδυασμός γονιδίων εντοπίζεται σε κάθε ασθενή.

Ο τύπος αυτός του ΣΔ συναντάται σε τουλάχιστον 90% όλων των περιπτώσεων διαβήτη. Στην Ευρώπη ο επιπολασμός του κυμαίνεται 5-10%. Στην Ελλάδα έχει τριπλασιασθεί τα τελευταία 35 χρόνια και σήμερα κυμαίνεται περί το 8%.

Οι ασθενείς με ΣΔΤ2 μπορούν σε αρχικά στάδια να χειριστούν την κατάστασή τους με κατάλληλη δίαιτα σε συνδυασμό με φυσική άσκηση. Ωστόσο, οι περισσότεροι ασθενείς κάποια στιγμή θα χρειαστεί να αρχίσουν είτε φαρμακευτική αγωγή είτε χορήγηση εξωγενούς ινσουλίνης.

2.4.3 Σακχαρώδης Διαβήτης Κύησης (ΣΔΚ)

Ο ΣΔ εμφανίζεται πρώτη φορά κατά τη διάρκεια της κύησης και υποχωρεί μετά τον τοκετό. Προσβάλλει το 3-5% των κυήσεων. Όμοια με τον ΣΔΤ2, ο Σακχαρώδης Διαβήτης κύησης χαρακτηρίζεται από μειωμένη έκκριση ινσουλίνης και ελαττωμένη ευαισθησία των κυττάρων στην ινσουλίνη. Η παχυσαρκία αυξάνει την πιθανότητα εμφάνισης ΣΔ κύησης, ενώ 30-40% των γυναικών με ΣΔΚ εμφανίζουν ΣΔΤ2 αργότερα στη ζωή τους.

Ο ΣΔ κύησης είναι θεραπεύσιμος αλλά απαιτεί κατάλληλη φαρμακευτική αγωγή κατά τη διάρκεια της κύησης. Κίνδυνοι που μπορεί να παρουσιαστούν στο παιδί είναι η μακροσωμία, ανωμαλίες στο καρδιακό και στο νευρικό σύστημα και μυϊκές διαταραχές.

2.4.4 Άλλοι τύποι ΣΔ

Υπάρχουν αρκετοί τύποι διαβήτη που απλά θα αναφερθούν ονομαστικά. Αυτοί σχετίζονται με συγκεκριμένες γενετικές καταστάσεις, χειρουργεία, φάρμακα, μολύνσεις, ασθένειες του παγκρέατος και άλλες ασθένειες. Τέτοιοι τύποι ΣΔ ευθύνονται για το 1-5% όλων των περιπτώσεων διαβήτη.

Ο Όψιμος Αυτοάνοσος Διαβήτης Ενηλίκων (Latent autoimmune diabetes in adults), που είναι μια μορφή αυτοάνοσου (τύπου 1 διαβήτη) και διαγιγνώσκεται σε ενήλικες με ηλικία μεγαλύτερη από τη συνήθη ηλικία εμφάνισης του ΣΔΤ1 (συνήθως άνω των 30 ετών). Ο διαβήτης MODY (Monogenic and maturity-onset diabetes of youth), που εμφανίζεται λιγότερο συχνά και συμπεριλαμβάνει ποικίλους τύπους διαβήτη σε μη παχύσαρκα παιδιά και ενήλικες με αναγνωσμένες γενετικές μεταλλάξεις. Τέλος, ο νεογνικός διαβήτης που είναι μια σπάνια μορφή διαβήτη που διαγιγνώσκεται συνήθως στους πρώτους 6 μήνες ζωής των βρεφών.

2.5 Γλυκαιμικός έλεγχος

Όπως παρουσιάστηκε στις προηγούμενες παραγράφους βλέπει κανείς ότι, οι δύο πρώτες περιπτώσεις ΣΔ (ΣΔΤ1 και ΣΔΤ2) χαρακτηρίζονται κατά κύριο λόγο από αυξημένη συγκέντρωση της γλυκόζης στο αίμα. Τα άτομα που πάσχουν από ΣΔΤ1 πρέπει να λαμβάνουν ινσουλίνη εξωγενώς, ενώ τα άτομα με ΣΔΤ2 λαμβάνουν κατάλληλη φαρμακευτική αγωγή προκειμένου να αυξηθεί η ευαισθησία των κυττάρων στην ινσουλίνη, ενώ σε προχωρημένες περιπτώσεις με μειωμένη έκκριση ινσουλίνης χορηγείται και σε αυτά επιπλέον ινσουλίνη εξωγενώς.

Η πιο συχνής επιπλοκή από τη χορήγηση εξωγενούς ινσουλίνης είναι η πτώση της συγκέντρωσης της γλυκόζης σε επίπεδα κάτω από τα φυσιολογικά. Επομένως, το πλάνο θεραπείας του ασθενούς πρέπει να σχεδιάζεται πολύ προσεκτικά και πάντα σε συνάρτηση με τις καθημερινές συνήθειές του. Ο χρόνος λήψης ενός γεύματος καθώς και η ποσότητα και η σύνθεση αυτού, όπως επίσης η φυσική δραστηριότητα, ο χρόνος, η διάρκεια και η έντασή της, είναι παράγοντες που πρέπει να συνεκτιμηθούν προκειμένου να αποφασιστεί η κατάλληλη δόση ινσουλίνης για τον ασθενή.

Όπως προαναφέρθηκε και στην προηγούμενη ενότητα, ο ΣΔ συνδέεται με εμφάνιση μακροχρόνιων επιπλοκών, όπως η νεφροπάθεια, η περιφερική νευροπάθεια και η αμφιβληστροειδοπάθεια. Έρευνες έχουν δείξει ότι η εμφάνιση των μακροχρόνιων επιπλοκών μπορεί να καθυστερήσει ή ακόμη και να αποφευχθεί με σωστό γλυκαιμικό έλεγχο. Η σημασία του σωστού γλυκαιμικού ελέγχου καταδεικνύεται και από το γεγονός πως ο ΣΔ αποτελεί μια ιδιαίτερα διαδεδομένη νόσο, η οποία συνδέεται άμεσα με τις συνήθειες του σύγχρονου τρόπου ζωής και επομένως αναμένεται να αυξηθεί η εξάπλωσή της μέσα στα επόμενα χρόνια.

Σε παγκόσμιο επίπεδο το 2011, εκτιμάται ότι περίπου 366 εκατομμύρια άνθρωποι είχαν ΣΔ και προβλέπεται ότι μέχρι το 2030 ο αριθμός αυτός θα ανέλθει στα 552 εκατομμύρια. Το 90% περίπου των περιπτώσεων ΣΔ αφορά άτομα με ΣΔΤ2. Το κόστος από τη διαχείριση των οξέων περιστατικών που συνδέονται με το ΣΔ και κυρίως από τη διαχείριση των μακροχρόνιων επιπλοκών, είναι ιδιαίτερα υψηλό σε παγκόσμιο επίπεδο.

Τα τελευταία χρόνια η έρευνα προσανατολίζεται στη χρήση της τεχνολογίας για την υποστήριξη των ατόμων με ΣΔ στον έλεγχο της συγκέντρωσης της γλυκόζης αίματος. Μια πολλά υποσχόμενη ερευνητική περιοχή αφορά την ανάπτυξη του Τεχνητού Παγκρέατος, ενός ελεγκτή ο οποίος θα αποφασίζει αυτόνομα τη βέλτιστη δόση ινσουλίνης για κάθε χρονική στιγμή με βάση τις καταγραφές ενός αισθητήρα γλυκόζης. Ωστόσο, τόσο στην περίπτωση ενός αυτόνομου ελεγκτή όσο και στην περίπτωση του συμβατικού ελέγχου, η γνώση των πιθανών διαταραχών του συστήματος είναι ζωτικής σημασίας. Ως διαταραχές νοούνται τα γεγονότα τα οποία μπορούν να επηρεάσουν τη συγκέντρωση της γλυκόζης στο αίμα και επομένως η εκ των προτέρων γνώση τους θα ενισχύσει θετικά τον αποδοτικό έλεγχο της γλυκόζης. Οι πιο σημαντικές διαταραχές στην περίπτωση του γλυκαιμικού ελέγχου είναι η λήψη τροφής και η φυσική άσκηση.

2.6 Μέτρηση υδατανθράκων σε άτομα με ΣΔ

Η λήψη τροφής είναι ένα στοιχείο το οποίο δεν μπορεί να αγνοηθεί καθώς έχει καθοριστικό ρόλο στη διακύμανση των επιπέδων της γλυκόζης. Η τροφή αποτελείται από υδατάνθρακες, πρωτεΐνες και λίπος. Η διάσπασή της αρχίζει στο στόμα και ολοκληρώνεται στο στομάχι. Τα επιμέρους στοιχεία απορροφούνται στο έντερο, από όπου περνάνε στο αίμα και είναι διαθέσιμα για τις επιμέρους λειτουργίες του οργανισμού. Συγκεκριμένα, οι υδατάνθρακες χρησιμοποιούνται από τα κύτταρα ως βασική πηγή ενέργειας. Προκειμένου να αξιοποιηθεί η γλυκόζη του αίματος είναι απαραίτητη η ύπαρξη της ινσουλίνης η οποία κάνει τα κύτταρα διαπερατά στη γλυκόζη. Η κάθε τροφή έχει διαφορετικό ποσοστό υδατανθράκων, πρωτεϊνών και λίπους. Επομένως, ο οργανισμός θα αποκριθεί διαφορετικά, όσον αφορά την αύξηση της γλυκόζης στο αίμα, με τη λήψη διαφορετικής τροφής. Επιπλέον, για μια δεδομένη τροφή, μεγαλύτερη ποσότητα σημαίνει μεγαλύτερη κατανάλωση υδατανθράκων και συνεπώς μεγαλύτερη αύξηση της συγκέντρωσης της γλυκόζης στο αίμα. Αυτό σημαίνει πως δεν είναι δυνατό να καθοριστεί η βέλτιστη δόση ινσουλίνης αν δεν είναι γνωστή η ποσότητα και η σύνθεση της τροφής που πρόκειται να καταναλωθεί.

Η μέτρηση της περιεχόμενης ποσότητας υδατανθράκων στα γεύματα είναι πολύ σημαντική για τους ασθενείς με ΣΔ για δύο λόγους [10]. Πρώτον, τους βοηθά στο σχεδιασμό των γευμάτων και τον προγραμματισμό της φυσικής δραστηριότητας. Όσον αφορά το σχεδιασμό των γευμάτων, οι διαβητικοί μπορούν να καθορίζουν τα γεύματα με τέτοιο τρόπο ώστε να διατηρούν τη γλυκόζη του αίματος σε φυσιολογικά επίπεδα. Επίσης, μπορούν να διευρύνουν τις διατροφικές τους επιλογές περιλαμβάνοντας τροφές που πρέπει να αποφεύγονται από διαβητικούς λόγω μεγάλης περιεκτικότητας σε υδατάνθρακες. Δεύτερον, η μέτρηση των υδατανθράκων επιτρέπει την προσαρμογή των προγευματικών δόσεων ινσουλίνης στην ποσότητα των υδατανθράκων που πρόκειται να καταναλωθούν. Φυσιολογικά, στον οργανισμό η έκκριση ινσουλίνης έπεται από ένα γεύμα εξαρτάται από την ποσότητα των υδατανθράκων και γενικότερα τη σύνθεση της τροφής. Επομένως, η ποσότητα της ινσουλίνης που λαμβάνεται πριν από το γεύμα δεν μπορεί να είναι η ίδια για όλα τα γεύματα αλλά να προσαρμόζεται κατάλληλα ανάλογα με το είδος και την ποσότητα της τροφής.

Ο υπολογισμός των υδατανθράκων από τους διαβητικούς ασθενείς δεν είναι μια εύκολη διαδικασία. Πολλές φορές συστήνεται από τον θεράποντα ιατρό το ζύγισμα των τροφών. Αυτή η διαδικασία, όμως, είναι πολύ περιοριστική ως προς τον χρόνο και τον τόπο λήψης της τροφής. Η δυσκολία είναι ακόμα μεγαλύτερη στην περίπτωση που ο διαβητικός ασθενής είναι στην παιδική ηλικία, όπου απαιτείται η συνεχής επίβλεψη από κάποιον ενήλικο. Επειδή το ζύγισμα των τροφών είναι πολύ περιοριστικό, συχνά επιλέγεται η

εξαγωγή μιας εκτίμησης της ποσότητας της τροφής. Για την αντιστοίχιση της ποσότητας της τροφής με περιεχόμενους υδατάνθρακες υπάρχουν λίστες [11] οι οποίες χωρίζουν τις τροφές σε κατηγορίες και δίνουν για διάφορες τροφές την ποσότητα η οποία αντιστοιχεί σε ένα γεύμα, καθώς και την περιεκτικότητα σε υδατάνθρακες, πρωτεΐνες και λίπος για την ποσότητα αυτή.

Ωστόσο, έρευνες έχουν δείξει πως σε πολλές περιπτώσεις, οι αποκλίσεις των εκτιμήσεων από την πραγματική ποσότητα τροφής είναι σημαντικές και ικανές να οδηγήσουν τη γλυκόζη του αίματος εκτός των φυσιολογικών ορίων. Πιο συγκεκριμένα, έρευνα σε ασθενείς που υποβάλλονται σε ινσουλινοθεραπεία έδειξε πως ανακρίβειες της τάξης των 20 γραμμαρίων στην εκτίμηση των υδατανθράκων επιδρούν αρνητικά στο μεταγευματικό προφίλ γλυκόζης [12]. Αναλυτικότερα, η έρευνα έδειξε πως αν μία δόση ινσουλίνης η οποία έχει καθοριστεί με δεδομένο πως το επακόλουθο γεύμα θα περιέχει 60 γραμμάρια υδατανθράκων, ληφθεί πριν από γεύμα περιεκτικότητας 40 ή 80 γραμμαρίων υδατανθράκων, τότε οδηγεί σε μεταγευματική υπογλυκαιμία ή υπεργλυκαιμία, αντίστοιχα. Επομένως, σύμφωνα με τα αποτελέσματα της συγκεκριμένης έρευνας, οι εκτιμήσεις στην περιεκτικότητα των υδατανθράκων πρέπει να έχουν μια απόκλιση το πολύ 10 γραμμαρίων από την πραγματική τιμή.

Σε μια διαφορετική έρευνα [13], αξιολογήθηκε η ακρίβεια στον υπολογισμό των υδατανθράκων από εφήβους. Η έρευνα διεξήχθη σε 48 εφήβους ηλικίας 12 – 18 ετών με ΣΔΤ1, οι οποίοι λάμβαναν εξωγενή ινσουλίνη. Από τους συμμετέχοντες ζητήθηκε να εκτιμήσουν την περιεκτικότητα σε υδατάνθρακες σε 32 φαγητά που αποτελούν συχνή επιλογή στην συγκεκριμένη ηλικιακή κατηγορία. Η έρευνα έδειξε πως μόνο το 23% των εφήβων που συμμετείχαν στην έρευνα υπολόγισε με ακρίβεια (με μία απόκλιση των 10 γραμμαρίων) την ποσότητα των υδατανθράκων. Επίσης, η έρευνα έδειξε πως οι έφηβοι που υπολόγισαν με ακρίβεια την ποσότητα των υδατανθράκων είχαν μικρότερες τιμές HbA1C.

Σε μία μεταγενέστερη έρευνα [14], με πιο αισιόδοξα αποτελέσματα, διερευνήθηκε το κατά πόσο τα παιδιά ή οι έφηβοι και τα άτομα που τα επιβλέπουν υπολογίζουν με ακρίβεια την περιεκτικότητα των υδατανθράκων στα γεύματα. Στην έρευνα συμμετείχαν 102 παιδιά και έφηβοι, ηλικίας 8-18 ετών, με ΣΔΤ1 τα οποία λαμβάνουν εξωγενή ινσουλίνη, καθώς και τα άτομα που ήταν υπεύθυνα για τη φροντίδα τους. Σύμφωνα με τα αποτελέσματα της έρευνας, το 73% όλων των εκτιμήσεων είχε απόκλιση 10 – 15 γραμμαρίων από την πραγματική τιμή ενώ οι υπόλοιπες είχαν μεγαλύτερες αποκλίσεις.

Σε μία ακόμη έρευνα [15], μελετήθηκε η συσχέτιση μεταξύ της ακρίβειας στον υπολογισμό των υδατανθράκων από τους γονείς και του γλυκαιμικού ελέγχου στα παιδιά με ΣΔΤ1. Στην έρευνα συμμετείχαν 67 παιδιά, ηλικίας 4 – 12 ετών με ΣΔΤ1. Κατά τη διάρκεια της έρευνας, οι γονείς έκαναν εκτιμήσεις για την ποσότητα της τροφής που καταλάωναν τα

παιδιά ενώ παράλληλα υπολογιζόταν η ακρίβεια αυτών των εκτιμήσεων. Η ποιότητα του γλυκαιμικού ελέγχου εξακριβώθηκε με τη μέτρηση της HbA1C. Τα αποτελέσματα της έρευνας έδειξαν πως μεγαλύτερη ακρίβεια στην εκτίμηση των γευμάτων συσχετίζεται με χαμηλότερες τιμές HbA1C.

3

Μοντέλα Αυτόματης Αξιολόγησης Τροφών

3.1 Υπολογισμός υδατανθράκων

Όπως παρουσιάστηκε εκτενέστερα στην προηγούμενη ενότητα, είναι γεγονός ότι η εκτίμηση των υδατανθράκων του γεύματος είναι κρίσιμης σημασίας για έναν διαβητικό ασθενή. Η γνώση της περιεχόμενης ποσότητας υδατανθράκων στα γεύματα είναι πολύ σημαντική τόσο για τον σωστό προγραμματισμό των γευμάτων του ασθενούς και της φυσικής του δραστηριότητας, όσο και για την προσαρμογή των προγευματικών δόσεων ινσουλίνης, ώστε να επιτευχθεί η διατήρηση της γλυκόζης του αίματος σε φυσιολογικά επίπεδα.

Επιπλέον, έγινε φανερό μέσα από έρευνες ότι οι αποκλίσεις στην υπολογισμό των υδατανθράκων των γευμάτων από τους ίδιους τους ασθενείς που υποβάλλονται σε ινσουλινοθεραπεία σε σχέση με τις πραγματικές τιμές, είναι αρκετά σημαντικές, γεγονός που μπορεί να οδηγήσει σε λάθος αποφάσεις με συνέπεια την απορρύθμιση του μεταβολισμού γλυκόζης και την πρόκληση μεταγευματικής υπογλυκαιμίας ή υπεργλυκαιμίας.

Επομένως, η σταθερή αύξηση των ατόμων παγκοσμίως που πάσχουν από διαβήτη τύπου 1, σε συνδυασμό με την αποδεδειγμένη αδυναμία τους να εκτιμήσουν με ακρίβεια την περιεκτικότητα των υδατανθράκων της διατροφής τους, οδήγησε στην ανάγκη ανάπτυξης συστημάτων και εφαρμογών, οι οποίες θα υποστηρίζουν αυτούς τους ασθενείς στην διαδικασία καταμέτρησης των υδατανθράκων που καταναλώνουν στο γεύμα τους.

Τα τελευταία χρόνια έχει προταθεί η χρήση της τεχνολογίας και συγκεκριμένα των κινητών τηλεφώνων για αυτόματο και ακριβή υπολογισμό των περιεχόμενων υδατανθράκων στα γεύματα. Μέχρι στιγμής, έχει προταθεί στη βιβλιογραφία ένα ευρύ φάσμα εφαρμογών κινητού τηλεφώνου για την υποβοήθηση αυτής της διαδικασίας, οι οποίες κυμαίνονται από διαδραστικά ημερολόγια [27] μέχρι αισθητήρες σώματος παρακολούθησης της καθημερινής διαίτας [28]. Επιπλέον, τα σύγχρονα κινητά τηλέφωνα πλέον διαθέτουν κάμερα υψηλής ασφαλείας, σύνδεση στο διαδίκτυο, εξελιγμένους επεξεργαστές καθώς και αυξημένη χωρητικότητα. Επομένως, η αυξανόμενη αυτή υπολογιστική ισχύς των κινητών συσκευών, καθώς και η πρόοδος στον τομέα της όρασης υπολογιστών, έχουν οδηγήσει στην εμφάνιση εφαρμογών διαχείρισης της διατροφής, βασισμένες στην ανάλυση εικόνας/βίντεο.

Στις τελευταίες εφαρμογές, ένα τυπικό σενάριο χρήσης της εφαρμογής είναι το εξής: Αρχικά, ο χρήστης αποκτά μια εικόνα του επερχόμενου γεύματός του, χρησιμοποιώντας την κάμερα του κινητού του τηλεφώνου. Ακολουθεί η επεξεργασία της εικόνας, είτε τοπικά είτε σε κάποιο server με ειδικούς αλγορίθμους, προκειμένου να εξαχθεί ένα διάγραμμα χαρακτηριστικών που να περιγράφει τις οπτικές ιδιότητες της εικόνας. Τα εξαγόμενα χαρακτηριστικά, στη συνέχεια, τροφοδοτούνται σε έναν ταξινομητή, ώστε να αναγνωρισθεί η κατηγορία τροφής στην οποία ανήκει η συγκεκριμένη εικόνα και στη συνέχεια αυτό να χρησιμοποιηθεί για την εκτίμηση των υδατανθράκων. Επιπλέον, η ανάλυση αυτή πέρα από την ποσότητα των υδατανθράκων, μπορεί να εξάγει και το θερμιδικό περιεχόμενο του γεύματος καθώς και την περιεκτικότητα σε πρωτεΐνες.

Στις επόμενες παραγράφους θα γίνει αναφορά σε διάφορες εφαρμογές αυτής της κατηγορίας, στην υλοποίησή τους και κυρίως στις τεχνικές που επιλέχθηκαν για την εξαγωγή των οπτικών χαρακτηριστικών της εικόνας και την ταξινόμηση αυτής.

3.2 Δυσκολίες στην αυτόματη αξιολόγηση των τροφών

Η διατροφική ανάλυση των τροφών με τεχνικές επεξεργασίας εικόνας μέσω των κινητών τηλεφώνων είναι ιδιαίτερα ελκυστική και πολλά υποσχόμενη στη διαχείριση του ΣΔ, ωστόσο, υπάρχουν κάποιες σημαντικές δυσκολίες:

- Πολλές τροφές έχουν παρόμοια οπτικά χαρακτηριστικά και είναι δύσκολο να διακριθούν με βάση τη φωτογραφία από τον αλγόριθμο ταξινόμησης.
- Πολλές τροφές που ανήκουν στην ίδια κατηγορία παρουσιάζουν πολύ διαφορετική εμφάνιση μεταξύ τους, κάνοντας πολύ δύσκολη την αναγνώριση όλων αυτών των τροφών.

- Ένα γεύμα συνήθως αποτελείται από περισσότερα από ένα τρόφιμα. Ο διαχωρισμός των αντικειμένων δεν είναι πάντα εύκολος και ειδικά στην περίπτωση που τα τρόφιμα έχουν μη κανονικά σχήματα ή υπάρχουν επικαλύψεις.
- Ο φωτισμός κατά τη λήψη της φωτογραφίας μπορεί να επηρεάσει την κατάτμηση ή/και την αναγνώριση των τροφών.
- Υπάρχει ανάγκη για μεγάλη επεξεργαστική ισχύ για την επεξεργασία της εικόνας. Εναλλακτικά, η εικόνα μπορεί να σταλεί σε έναν κεντρικό εξυπηρετητή, προκειμένου να γίνει η ταξινόμηση των τροφών. Ωστόσο, σε αυτή την περίπτωση δημιουργείται η απαίτηση για σύνδεση του κινητού τηλεφώνου σε δίκτυο.

3.3 Διαδικασία αυτόματης αξιολόγησης των τροφών

Η διαδικασία για την αυτόματη αξιολόγηση των τροφών περιλαμβάνει τα ακόλουθα βήματα:

1. Βαθμονόμηση της κάμερας για τον υπολογισμό των παραμέτρων της κάμερας.
2. Λήψη φωτογραφίας του γεύματος.
3. Κατάτμηση της εικόνας του γεύματος στα επιμέρους στοιχεία με τεχνικές επεξεργασίας εικόνας.
4. Ταξινόμηση του κάθε τροφίμου, βασισμένη κατά κύριο λόγο σε μεθόδους μηχανικής μάθησης.
5. Υπολογισμός της ποσότητας του γεύματος.
6. Αντιστοίχιση της ποσότητας του γεύματος με την ανάλογη ποσότητα υδατανθράκων.
7. Λήψη φωτογραφίας μετά την κατανάλωση του γεύματος προκειμένου να εξαχθεί πληροφορία για την ποσότητα, ώστε να αφαιρεθεί από την αρχική εκτίμηση η ποσότητα που δεν έχει καταναλωθεί.

Στις επόμενες παραγράφους ακολουθεί μια σύντομη περιγραφή κάθε επιμέρους σταδίου της αυτόματης αξιολόγησης των τροφών, καθώς και επισκόπηση σε κάποιες βασικές τεχνικές που ακολουθούνται σε κάθε ένα από αυτά. Σε επόμενη ενότητα θα γίνει μια πιο λεπτομερής παρουσίαση του σταδίου της ταξινόμησης των τροφών, καθώς και μια βιβλιογραφική επισκόπηση σε πλήθος μεθόδων και αλγορίθμων που χρησιμοποιούνται ευρέως σε διάφορα μοντέλα ταξινόμησης.

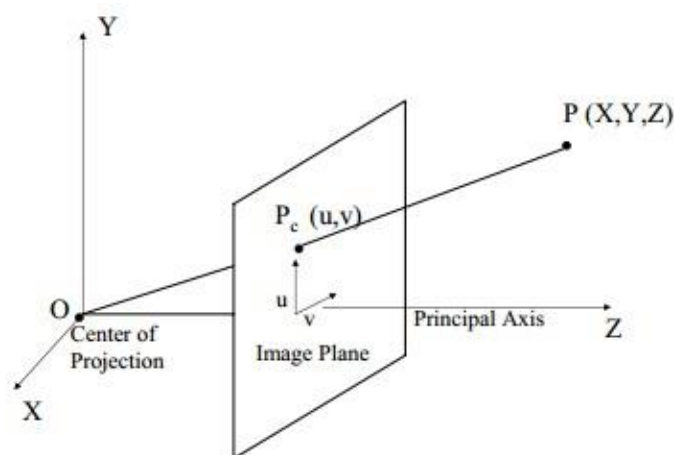
3.3.1 Βαθμονόμηση κάμερας

Για την ανακατασκευή ενός τρισδιάστατου σκηνικού στο πλάνο δύο διαστάσεων μιας κάμερας, απαιτείται ένας μετασχηματισμός. Ο μετασχηματισμός γίνεται με τη βοήθεια ενός πίνακα βαθμονόμησης ο οποίος αντιστοιχεί το σημείο τριών διαστάσεων $P = (X, Y, Z)$ του πραγματικού κόσμου σε ένα σημείο δύο διαστάσεων $P_c = (u, v)$.

Ο πίνακας βαθμονόμησης μιας κάμερας αποτελείται από παραμέτρους δύο τύπων:

- Ενδογενείς παράμετροι, οι οποίες μετασχηματίζουν τις συντεταγμένες τριών διαστάσεων σε συντεταγμένες δύο διαστάσεων μιας εικόνας.
- Εξωγενείς παράμετροι, οι οποίες περιγράφουν τη θέση της κάμερας στον κόσμο και τον προσανατολισμό της.

Οι ενδογενείς παράμετροι είναι σταθερές για μία κάμερα και επομένως αρκεί να υπολογιστούν μία φορά. Αντίθετα, οι εξωγενείς παράμετροι μεταβάλλονται σε κάθε λήψη μιας εικόνας και επομένως πρέπει να υπολογίζονται εκ νέου μετά τη λήψη κάθε εικόνας.



Εικόνα 3.1: Μετασχηματισμός του σημείου τριών διαστάσεων $P = (X, Y, Z)$ του πραγματικού κόσμου σε ένα σημείο δύο διαστάσεων $P_c = (u, v)$.

3.3.2 Λήψη φωτογραφίας του γεύματος

Το πρώτο βήμα για τον αυτόματο υπολογισμό των υδατανθράκων του γεύματος είναι η φωτογράφιση του πιάτου που περιλαμβάνει το γεύμα. Συνήθως, αυτό είναι και το μόνο στάδιο στο οποίο απαιτείται κάποια ενέργεια από το χρήστη αν και σε κάποιες εφαρμογές μπορεί να ζητηθεί επιπλέον είσοδος από το χρήστη, όπως για παράδειγμα να ζητηθεί από το χρήστη να κατονομάσει τα τρόφιμα τα οποία αποτελούν το γεύμα. Οι περισσότερες προσεγγίσεις συνιστούν τη φωτογράφιση του πιάτου πριν και μετά την κατανάλωση του

γεύματος προκειμένου τελικά να ληφθεί υπόψη μόνο ο όγκος της τροφής που καταναλώθηκε και όχι όλο το περιεχόμενο του πιάτου. Επιπλέον, κάποιες εφαρμογές επιχειρούν να συμπεράνουν τον όγκο της τροφής από μία εικόνα, ωστόσο, για μεγαλύτερη ακρίβεια μπορεί να ζητηθούν περισσότερες εικόνες (συνήθως τρεις) από το χρήστη από διαφορετικές οπτικές γωνίες γύρω από το πιάτο. Με αυτό τον τρόπο είναι δυνατό να αποφευχθεί η επικάλυψη τροφών από άλλα μέσα στο πιάτο αλλά και να γίνει πιο ακριβής υπολογισμός του όγκου.

Ένα ακόμη σημαντικό ζήτημα είναι ο ορισμός μιας μετρικής κλίμακας προκειμένου να εξαχθούν οι διαστάσεις των τροφίμων και να γίνει ο υπολογισμός του όγκου. Μία απλή λύση είναι η τοποθέτηση ενός αντικειμένου γνωστών διαστάσεων μέσα στο πλάνο λήψης της εικόνας. Η τακτική αυτή, ωστόσο, έχει το μειονέκτημα πως ο χρήστης πρέπει να έχει μαζί του το συγκεκριμένο αντικείμενο.

Άλλα ζητήματα που πρέπει να ληφθούν υπόψη κατά τη λήψη της φωτογραφίας είναι ο φωτισμός του χώρου, καθώς και ο χρωματισμός του φόντου. Σε κάποιες εφαρμογές απαιτείται η χρήση ενός συγκεκριμένου υφάσματος στο οποίο τοποθετείται το πιάτο για να ληφθεί η εικόνα. Αυτή η τακτική μπορεί να βοηθήσει και στον ορισμό της κλίμακας, καθώς και στο διαχωρισμό του πιάτου από το φόντο κατά το στάδιο της κατάτμησης.



Εικόνα 3.2: Λήψη εικόνας σε συγκεκριμένο φόντο.

3.3.3 Κατάτμηση της εικόνας

Συνήθως, ένα γεύμα περιλαμβάνει περισσότερα από ένα είδη τροφών. Ένα πρώτο βήμα για την επεξεργασία της εικόνας είναι η κατάτμηση για τον διαχωρισμό των επιμέρους τροφίμων που απεικονίζονται στην εικόνα. Η κατάτμηση είναι σημαντική και για την αναγνώριση των τροφίμων, αλλά και για τον υπολογισμό του όγκου, τα οποία αποτελούν και τα επόμενα στάδια επεξεργασίας της εικόνας. Οι αλγόριθμοι κατάτμησης μιας εικόνας

βασίζονται σε τεχνικές ανίχνευσης ακμών, τεχνικές κατωφλίου και τεχνικές βασισμένες σε περιοχές.

Οι τεχνικές ανίχνευσης ακμών εντοπίζουν τα σημεία της εικόνας στα οποία υπάρχει απότομη αλλαγή στο χρώμα και βασίζονται είτε στο ιστόγραμμα της εικόνας είτε στην πρώτη παράγωγο της εικόνας. Η μέθοδος κατωφλίου μπορεί να διαχωρίσει ανοιχτόχρωμα αντικείμενα από σκουρόχρωμο φόντο και δημιουργεί μόνο δύο κλάσεις, το φόντο και τα αντικείμενα. Στις τεχνικές βασισμένες σε περιοχές, υπάγονται οι μέθοδοι ανάπτυξης περιοχής, όπου επιλέγεται ένα σύνολο pixels γύρω από τα οποία αναπτύσσονται περιοχές ανάλογα με την ομοιότητα των γύρω pixels, και οι μέθοδοι συγχώνευσης περιοχών, όπου η αρχική εικόνα διαχωρίζεται τυχαία σε πολλές περιοχές οι οποίες στη συνέχεια συνενώνονται ανάλογα με την ομοιότητα που εμφανίζουν μεταξύ τους.



Εικόνα 3.3: Κατάτμηση της εικόνας για τον υπολογισμό των διαφορετικών τροφίμων.

3.3.4 Ταξινόμηση των τροφίμων

Μετά την κατάτμηση της εικόνας και τον εντοπισμό των διαφορετικών τροφίμων που αποτελούν το γεύμα, το επόμενο βήμα είναι η αναγνώριση του κάθε τροφίμου προκειμένου να του αποδοθεί μια ετικέτα που θα καταδεικνύει το είδος του.

Η αναγνώριση των τροφίμων μπορεί να γίνει είτε με μεθόδους μηχανικής μάθησης με χρήση ειδικά εκπαιδευμένων δικτύων, είτε με ταίριασμα χαρακτηριστικών με τα χαρακτηριστικά σημεία των τροφίμων μιας βάσης δεδομένων.

3.3.5 Υπολογισμός ποσότητας γεύματος

Η ποσοτικοποίηση του γεύματος γίνεται μέσω του υπολογισμού του όγκου των τροφίμων. Οι πρώτες προσεγγίσεις για αυτόματη εκτίμηση του όγκου των τροφών αφορούσαν τη χρήση αντιστοίχισης των τροφών με κανονικά σχήματα. Οι νεότερες προσεγγίσεις κατευθύνονται στη χρήση τρισδιάστατης ανακατασκευής τροφίμων.

3.3.6 Εξαγωγή ποσότητας υδατανθράκων

Μετά τον υπολογισμό της ποσότητας της τροφής, πρέπει να υπολογιστεί η περιεχόμενη ποσότητα υδατανθράκων. Για το σκοπό αυτό, δεδομένα για τη διατροφική πληροφορία μπορούν να αντληθούν από διάφορες βάσεις δεδομένων που είναι ελεύθερα διαθέσιμες, όπως η βάση δεδομένων τροφίμων και διατροφικής αξίας (Food and Nutrient Database for Dietary Studies - FNDDS) του Υπουργείου Γεωργίας των Ηνωμένων Πολιτειών Αμερικής, η οποία περιλαμβάνει περισσότερα από 7000 τρόφιμα με τη διατροφική τους αξία, σε συνδυασμό με πληροφορία όπως το τυπικό μέγεθος της μερίδας. Επομένως, η τελική πληροφορία που θα δίνεται ως έξοδος στο χρήστη μπορεί να μην περιορίζεται μόνο στην ποσότητα υδατανθράκων, αλλά να είναι μία πιο πλήρης διατροφική ανάλυση του γεύματος, ενημερώνοντας το χρήστη για το θερμιδικό περιεχόμενο του γεύματος, καθώς και την ποσότητα των πρωτεϊνών και των λιπών.

Επιπλέον, στην περίπτωση διαβητικών ασθενών, η εφαρμογή μπορεί, σε συνδυασμό με την πληροφορία από έναν αισθητήρα γλυκόζης και με δεδομένη την ποσότητα υδατανθράκων που πρόκειται να καταναλωθούν, να προειδοποιήσει το χρήστη για πιθανή επερχόμενη υπογλυκαιμία ή υπεργλυκαιμία.

3.4 Ταξινόμηση των τροφίμων - Βιβλιογραφική επισκόπηση

3.4.1 Γενικά

Το 2008 παρουσιάστηκε για πρώτη φορά μια εφαρμογή αναγνώρισης γευμάτων από τους Shroff et al. [29] για την ταξινόμηση εικόνων φαγητού γρήγορης εστίασης (fast food). Για κάθε κατηγορία τροφής που έχει εξαχθεί από την κατάτμηση της αρχικής εικόνας, υπολογίζεται ένα διάνυσμα από χρωματικά (κανονικοποιημένες RGB τιμές), μεγέθους, υφής (μέση τιμή, τυπική απόκλιση, εντροπία), σχήματος και context-based χαρακτηριστικά, το οποίο τροφοδοτείται σε ένα τεχνητό νευρωνικό δίκτυο (ANN) πρόσθιας τροφοδότησης. Το αποτέλεσμα που επιτεύχθηκε σε όρους ακρίβειας αναγνώρισης ήταν της τάξεως του 95%, 80%, 90% και 90% για τις ακόλουθες 4 κατηγορίες αντίστοιχα: χάμπουργκερ, τηγανιτές πατάτες, κοτομπουκιές και μηλόπιτα.

Αργότερα, το 2010, έγινε προσπάθεια για διεύρυνσης των κατηγοριών των τροφίμων και επέκτασής τους σε τρόφιμα καθημερινής κατανάλωσης. Η προσπάθεια αυτή έγινε από την ερευνητική ομάδα των Zhu et al. [30] όπου χρησιμοποιήθηκε ένα σύνολο από χρωματικά χαρακτηριστικά (ένταση των pixel και χρωματικά στοιχεία) και χαρακτηριστικά υφής (Gabor filter responses), σε συνδυασμό με μια μηχανή διανυσμάτων υποστήριξης (SVM) για την

ταξινόμηση γευμάτων σε 19 τροφικές κατηγορίες. Η μελέτη αυτή οδήγησε σε ποσοστό αναγνώρισης της τάξης του 94% για αντίγραφα τροφίμων και του 58% για εικόνες πραγματικών γευμάτων.

Μια διαφορετική προσέγγιση από πλευράς εξαγωγής χαρακτηριστικών προτάθηκε το 2012 στην εφαρμογή DietCam [31]. Στην συγκεκριμένη προσέγγιση οι ερευνητές Kong και Tan έκαναν χρήση του αλγόριθμου Scale Invariant Feature Transform (SIFT) προκειμένου να εξάγουν από τις εικόνες του συνόλου δεδομένων τα διανύσματα των χαρακτηριστικών. Η εφαρμογή DietCam προκειμένου να αναγνωρίσει τα τρόφιμα που απεικονίζονται σε μια εικόνα, προσπαθεί να ταιριάξει τα χαρακτηριστικά που εξάγονται από τον αλγόριθμο SIFT με τα χαρακτηριστικά αναφοράς συγκεκριμένων τροφίμων που έχουν καταχωρηθεί σε μια βάση δεδομένων, με τη χρήση ενός Bayesian πιθανοτικού ταξινομητή. Ωστόσο, ο αλγόριθμος SIFT αποδίδει διανύσματα χαρακτηριστικών με 128 διαστάσεις, ενώ για κάθε μια εικόνα εξάγονται εκατοντάδες τέτοιου τύπου χαρακτηριστικά. Επομένως, τα χαρακτηριστικά διανύσματα εδώ ομαδοποιούνται με τη χρήση του αλγορίθμου k-means και τροφοδοτούνται στον πιθανοτικό ταξινομητή. Η βάση δεδομένων που χρησιμοποιήθηκε αποτελείται από εικόνες φρούτων, σπιτικών φαγητών και γρήγορων γευμάτων, ενώ η απόδοση που επιτεύχθηκε είναι της τάξης του 92%, δεδομένου ότι ο αριθμός των αναφορών κάθε κατηγορίας ήταν μεγαλύτερος του 50, ενώ ο αριθμός των προς αναγνώριση τροφίμων ήταν μικρότερος του 6.

Τα τελευταία χρόνια, στον τομέα της όρασης υπολογιστών και για την αντιμετώπιση απαιτητικών προβλημάτων ταξινόμησης, έχει χρησιμοποιηθεί ευρέως το μοντέλο Bag of Features (BoF) για την περιγραφή των γενικών χαρακτηριστικών (global features) μιας εικόνας [32]. Η μέθοδος BoF προήλθε από το μοντέλο Bag of Words (BoW). Η μέθοδος BoW είναι μια δημοφιλής μέθοδος αναπαράστασης των εγγράφων στον τομέα της επεξεργασίας της φυσικής γλώσσας [33], κατά την οποία αγνοείται η σειρά εμφάνισης των λέξεων οι οποίες ανήκουν σε ένα προκαθορισμένο λεξικό, αλλά λαμβάνεται υπόψη μονάχα η συχνότητα εμφάνισής τους. Ομοίως, στην ανάλυση εικόνας, κάθε εικόνα αντιπροσωπεύεται από ένα ιστόγραμμα οπτικών λέξεων, τα οποία ορίζονται ως οι αντιπροσωπευτικοί περιγραφείς των συχνά εμφανιζόμενων οπτικών μοτίβων. Το μοντέλο BoF ταιριάζει επαρκώς στο πρόβλημα αναγνώρισης των γευμάτων, καθώς μια συγκεκριμένη κατηγορία τροφής, συνήθως, αποτελείται από ένα σύνολο διαφορετικών οπτικών στοιχείων αναμιγμένων όμως με συγκεκριμένες αναλογίες, αλλά χωρίς κάποια τυπική διάταξη στο χώρο. Επομένως, η χρήση της προσέγγισης BoF ενθαρρύνεται σε σχέση με κάποια από τις τεχνικές άμεσου ταιριάσματος των χαρακτηριστικών.

Μια διαφορετική εφαρμογή, το σύστημα FIVR [34] στην οποία ο χρήστης υποβοηθά τη διαδικασία ταξινόμησης των τροφίμων κατονομάζοντας τις τροφές που αποτελούν το γεύμα, προτάθηκε από τους Puri et al. το 2009. Με αυτό τον τρόπο δίνεται ένα μικρό σύνολο

υποψηφίων τροφίμων τα οποία πρέπει να αντιστοιχιστούν με τα τρόφιμα της εικόνας. Στη συνέχεια, επιλέγεται η χρήση δυαδικών ταξινομητών, οι οποίοι εκπαιδεύονται μεταξύ όλων των ζευγών φαγητών. Στη συγκεκριμένη προσέγγιση, η αναγνώριση βασίζεται στο συνδυασμό χρωματικών συνιστωσών, στη μέγιστη απόκριση χαρακτηριστικών στο μοντέλο ιστογράμματος *Textron* και στην επιλογή χαρακτηριστικών με χρήση του αλγόριθμου *Adaboost*. Τα *Textron* ιστογράμματα είναι παρόμοια με το μοντέλο *BoF*, αλλά χρησιμοποιούν απλούστερους περιγραφείς, έτσι ώστε να μπορούν να χρησιμοποιηθούν ιστογράμματα από όλα τα πιθανά διανύσματα χαρακτηριστικών. Με αυτό τον τρόπο, η διαδικασία ομαδοποίησης των διανυσμάτων χαρακτηριστικών μπορεί να παραλειφθεί, με αποτέλεσμα όμως το μοντέλο να λαμβάνει υπόψη λιγότερη πληροφορία και έτσι πιθανόν να μην μπορεί να ανταποκριθεί καλά σε εικόνες με μεγάλες οπτικές παραλλαγές. Επιπλέον, το προτεινόμενο σύστημα απαιτεί την παρουσία μια χρωματιστής σκακιέρας κοντά στη φωτογραφία του γεύματος, προκειμένου να μπορούν να αντιμετωπιστούν οι διάφορες συνθήκες φωτισμού. Το σύστημα πέτυχε ακρίβεια αναγνώρισης από 95% μέχρι 80%, καθώς ο αριθμός των κατηγοριών τροφίμων αυξανόταν από τις 2 στις 20.

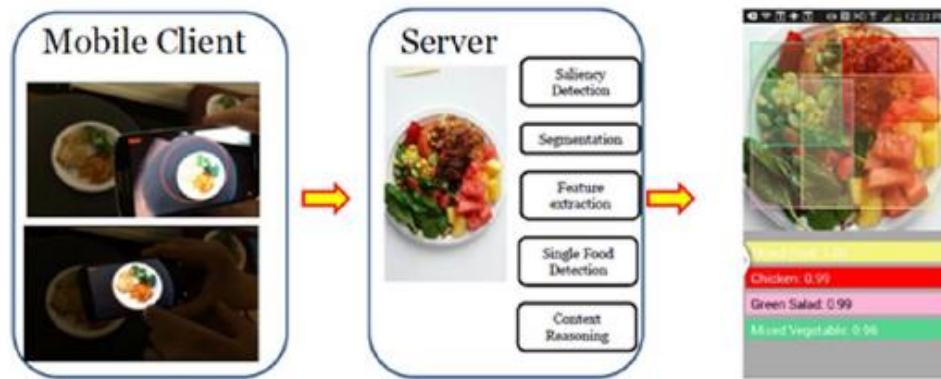
Πέραν των παραπάνω, η μελέτη της αναγνώρισης των τροφίμων με σκοπό την εκτίμηση της διατροφικής αξίας έχει τραβήξει μεγάλη προσοχή και πολλά συστήματα έχουν προταθεί τα τελευταία χρόνια. Ωστόσο, οι περισσότερες προσεγγίσεις αδυνατούν να ανταποκριθούν στην τεράστια οπτική ποικιλομορφία των τροφίμων, με αποτέλεσμα, προκειμένου να επιτύχουν ικανοποιητικά αποτελέσματα να περιορίζουν το σύνολο δεδομένων είτε με το να θεωρούν πολύ λίγες κατηγορίες τροφίμων είτε με το να εξετάζουν πολύ γενικές κατηγορίες.

Τέλος, στις επόμενες παραγράφους αξίζει να γίνει μια πιο εκτενής παρουσίαση δύο εφαρμογών που έχουν προταθεί τα τελευταία χρόνια και ανταποκρίνονται με μεγάλη ακρίβεια σε διευρυμένα σύνολα δεδομένων, με πολλές κατηγορίες τροφίμων και εικόνες που έχουν ληφθεί σε πραγματικές συνθήκες. Οι εφαρμογές αυτές είναι η *Snap-n-Eat* [19] της ερευνητικής ομάδας *Zhang et al.* καθώς και το σύστημα αναγνώρισης τροφίμων δύο σταδίων [23] της ομάδας *Antimopoulos et. al.*

3.4.2 *Snap-n-Eat*

Το 2015 προτείνεται το σύστημα *Snap-n-Eat* [19] από την ομάδα των *Zhang et al.*, το σχηματικό διάγραμμα του οποίου συνοψίζεται παρακάτω στην εικόνα 3.4. Στο σύστημα αυτό, ο χρήστης λαμβάνει μια φωτογραφία του πιάτου, η οποία περικόπτεται στο μέγεθος 400 x 400 pixel και στη συνέχεια αποστέλλεται στο server, ενεργοποιώντας έτσι τη διαδικασία αναγνώρισης. Όταν η αναγνώριση του φαγητού ολοκληρωθεί, τα αποτελέσματα

αποστέλλονται πίσω στο κινητό τηλέφωνο και εμφανίζονται στην οθόνη τα τετράγωνα οριοθέτησης του κάθε ξεχωριστού τροφίμου μαζί με τις βαθμολογίες τους. Η συνολική βαθμολογία διαρκεί γύρω στα 4 δευτερόλεπτα, συμπεριλαμβανομένου του χρόνου μετάδοσης στο δίκτυο.



Εικόνα 3.4: Σύστημα Snap 'n Eat [19]

3.4.2.1 Χαρακτηριστικά

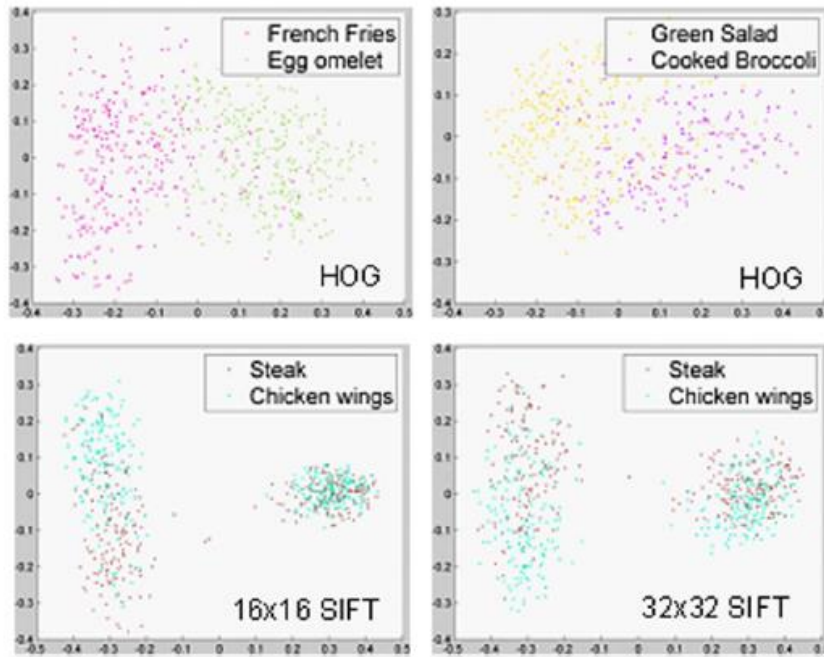
Στη συγκεκριμένη εφαρμογή, επιλέχθηκαν ως αλγόριθμοι περιγραφής των χαρακτηριστικών οι dense HOG (histogram of oriented gradients) και dense SIFT (scale-invariant feature transform). Ο λόγος για τον οποίο έγινε αυτή η επιλογή είναι ο ακόλουθος:

Καθώς τα HOG χαρακτηριστικά εξάγονται από μια εικόνα στην κλίμακα του γκρι, ουσιαστικά κωδικοποιούν πληροφορία για την υφή των τροφών διαφορετικής κατηγορίας. Επομένως, δύο ήδη τροφών με παρόμοιο χρώμα (όπως π.χ. οι τηγανιτές πατάτες και η ομελέτα που απεικονίζονται στην εικόνα 3.5 – άνω) μπορούν να διακριθούν λόγω της διαφορετικής υφής τους, όπως φαίνεται από τα HOG χαρακτηριστικά τους.

Τα χαρακτηριστικά SIFT [20] καταγράφουν παρόμοιες πληροφορίες με τα χαρακτηριστικά HOG, όμως έχουν την ιδιαιτερότητα ότι διαφορετικές κλίμακες έχουν διαφορετική διακριτική ικανότητα. Όπως απεικονίζεται και στην εικόνα 3.5 (κάτω), η μπριζόλα και οι φτερούγες κοτόπουλου ταξινομούνται και τα δύο σε δύο κλάσεις. Η μια κλάση διακρίνεται καλύτερα σε μικρότερη κλίμακα, ενώ η άλλη σε μεγαλύτερη. Αυτό, παροτρύνει την συγκεκριμένη έρευνα να λάβει υπόψη αποτελέσματα ταξινόμησης σε πολλαπλές κλίμακες.

3.4.2.2 Κωδικοποίηση χαρακτηριστικών

Η αναπαράσταση των τοπικών χαρακτηριστικών της συγκεκριμένης προσέγγισης γίνεται με χρήση τη μεθόδου Bag of Features (BoF). Αν και σε επόμενη ενότητα γίνεται εκτενής περιγραφή της μεθόδου, αξίζει εδώ να αναφερθούν τρεις σημαντικοί περιορισμοί που εντοπίστηκαν από την ομάδα ανάπτυξης της Snap-n-Eat:



Εικόνα 3.5: Σύγκριση SIFT και HOG χαρακτηριστικών σε διάφορες κλίμακες. [19]

1. Έλλειψη συνέχειας στον χώρο των χαρακτηριστικών: Το λεξικό των οπτικών λέξεων (visual dictionary) εξάγεται με έναν καθαρά παραγωγικό τρόπο, χωρίς να υπάρχει κάποια εξάρτηση μεταξύ των περιγραφέων (visual words). Επομένως, η ανάθεση ενός τοπικού περιγραφέα μεταπίπτει από τον ένα αντιπροσωπευτικό περιγραφέα στον άλλο με μικρές διαταραχές στον αρχικό χώρο των χαρακτηριστικών.
2. Αραιή αναπαράσταση: Στην αναπαράσταση BoF, κάθε τοπικός περιγραφέας συμβάλει μόνο σε μία μη μηδενική εγγραφή στο τελικό ιστόγραμμα. Αυτό, συχνά, οδηγεί στη χρήση πιο σύνθετων πυρήνων στη φάση της δοκιμής.
3. Χαμηλή διαστασιμότητα: Η διάσταση του τελικού ιστογράμματος καθορίζεται από το μέγεθος του λεξικού οπτικών λέξεων. Μεγαλύτερο μέγεθος λεξικού, συχνά, επιβαρύνει την ασυνέχεια στο χώρο των χαρακτηριστικών. Ωστόσο, αρκετές φορές για να επιτευχθεί ισχυρή ταξινόμηση απαιτείται μεγάλη διάσταση στο χώρο των χαρακτηριστικών.

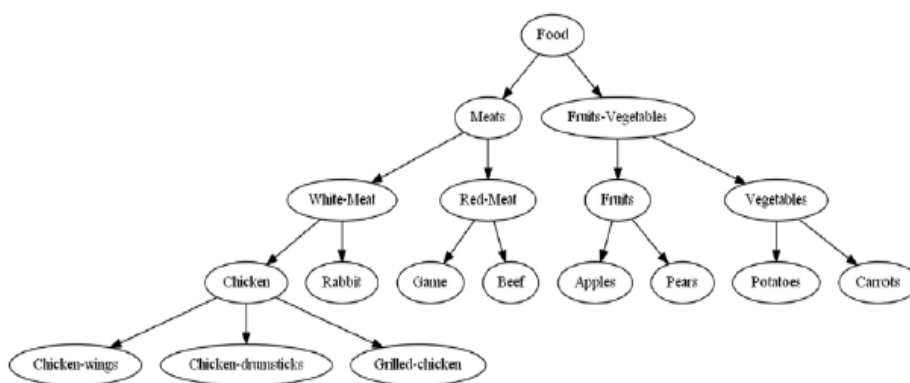
Λόγω των παραπάνω περιορισμών, αλλά κυρίως λόγω της πρόσφατης προόδου στον τομέα της οπτικής ταξινόμησης, στη συγκεκριμένη μελέτη χρησιμοποιείται η διανυσματική αναπαράσταση Fisher [21] η οποία ενισχύει την αναπαράσταση BoW (Bag of Words) μέσω της κωδικοποίησης υψηλής τάξης στατιστικών (πρώτης και προαιρετικά δεύτερης τάξης).

3.4.2.3 Ταξινόμηση

Για την ταξινόμηση επιλέγεται ένα multiclass SVM (Support Vector Machine) για κάθε κατηγορία. Για την εκμάθηση του ταξινομητή χρησιμοποιείται βαθμιαία στοχαστική κάθοδος.

Για την οργάνωση των διαφορετικών κατηγοριών τροφίμων, χρησιμοποιείται μια τροφική οντολογία (food ontology). Η οντολογία είναι ένας κατευθυνόμενος άκυκλος γράφος (Directed Acyclic Graph - DAG), όπου κάθε κόμβος αναπαριστά μια κατηγορία. Μια τέτοια τροφική οντολογία απεικονίζεται παρακάτω στην εικόνα 3.6. Αυτή η DAG οντολογία τροφίμων δίνει τη δυνατότητα οργάνωσης ενός μεγάλου αριθμού κατηγοριών τροφίμων, καθώς οργανώνει την αντιστοίχιση μεταξύ κατηγοριών τροφίμων και ταξινομητών. Για παράδειγμα, κάθε κόμβος του δέντρου της εικόνας 3.6 αντιστοιχεί σε ένα ταξινομητή και μπορεί να εμπεριέχει μια μόνο καλά καθορισμένη κατηγορία ή μια πιο γενική.

Κατά την εκμάθηση του ταξινομητή κάθε κόμβου, όλες οι εικόνες που αντιστοιχούν στην συγκεκριμένη κατηγορία αλλά και σε όλες της κατηγορίες των απογόνων του κόμβου, συμπεριλαμβάνονται στο θετικό σύνολο εικόνων, ενώ όλες οι υπόλοιπες εικόνες συμπεριλαμβάνονται στο αρνητικό σετ.



Εικόνα 3.6: DAG οντολογία τροφίμων

3.4.2.4 Αποτελέσματα

Η ομάδα που υλοποίησε την συγκεκριμένη εφαρμογή σύλλεξε το δικό της σύνολο δεδομένων, ενθαρρύνοντας τους χρήστες να ανεβάζουν σε μια πρώιμη έκδοση της εφαρμογής εικόνες των πιάτων των γευμάτων τους. Στη συνέχεια, οι εικόνες διαχωρίστηκαν χειροκίνητα

σε τροφικές κατηγορίες. Τελικά, συλλέχθηκαν 2000 εικόνες εκπαίδευσης για 15 κατηγορίες τροφίμων, με 100 έως 400 εικόνες ανά κατηγορία.

Στην εικόνα 3.7 απεικονίζονται τα αποτελέσματα της ταξινόμησης, με χρήση της μεθόδου 5-fold cross validation. Τα ποσοστά που εμφανίζονται είναι η μέση τιμή απόδοσης ταξινόμησης που επιτεύχθηκε ανά κατηγορία, άρα και το ποσοστό των εικόνων από το σύνολο αξιολόγησης που ταξινομήθηκε σωστά ανά κατηγορία.

	Pizza	Strawberry	Mixed Fruit	Burger	French Fries	Green Salad	Spaghetti	Sandwich	Steak	Chicken wings	Sushi roll	Cheerios	Egg omelet	Pancakes	Broccoli	Mean
HOG	65.3	65.7	74.2	48.9	81.9	78.2	78.6	52.5	62.2	50.9	54.5	85.7	41.4	51.3	57.7	63.3
SIFT-16	76.5	78.4	76.4	63.9	85.3	86.5	79.5	61.0	77.3	68.9	64.9	84.1	62.6	56.4	75.0	73.1
SIFT-32	70.4	78.4	68.5	61.7	86.2	85.7	75.9	59.3	72.3	72.6	62.3	93.7	58.6	62.8	65.4	71.6
LAB	42.9	88.2	52.8	36.1	53.4	82.7	35.7	44.1	66.4	57.5	63.6	44.4	47.5	38.5	73.1	55.1
Fusion	81.6	88.2	82.0	66.9	90.5	93.2	83.0	69.5	84.0	81.1	75.3	92.1	65.7	70.5	80.8	80.3

Εικόνα 3.7: Απόδοση ταξινόμησης με χρήση 4 διαφορετικών συνόλων χαρακτηριστικών και το συνδυασμό αυτών με χρήση της μεθόδου 5-fold cross validation.

3.4.3 Σύστημα αναγνώρισης τροφίμων δύο σταδίων

Το 2014 προτάθηκε από την ομάδα Anthimopoulos et al. [23] ένα σύστημα αναγνώρισης τροφίμων που αποτελείται από δύο στάδια: την περιγραφή της εικόνας του γεύματος και την ταξινόμηση της εικόνας. Το προτεινόμενο σύστημα απεικονίζεται στην εικόνα 3.8.

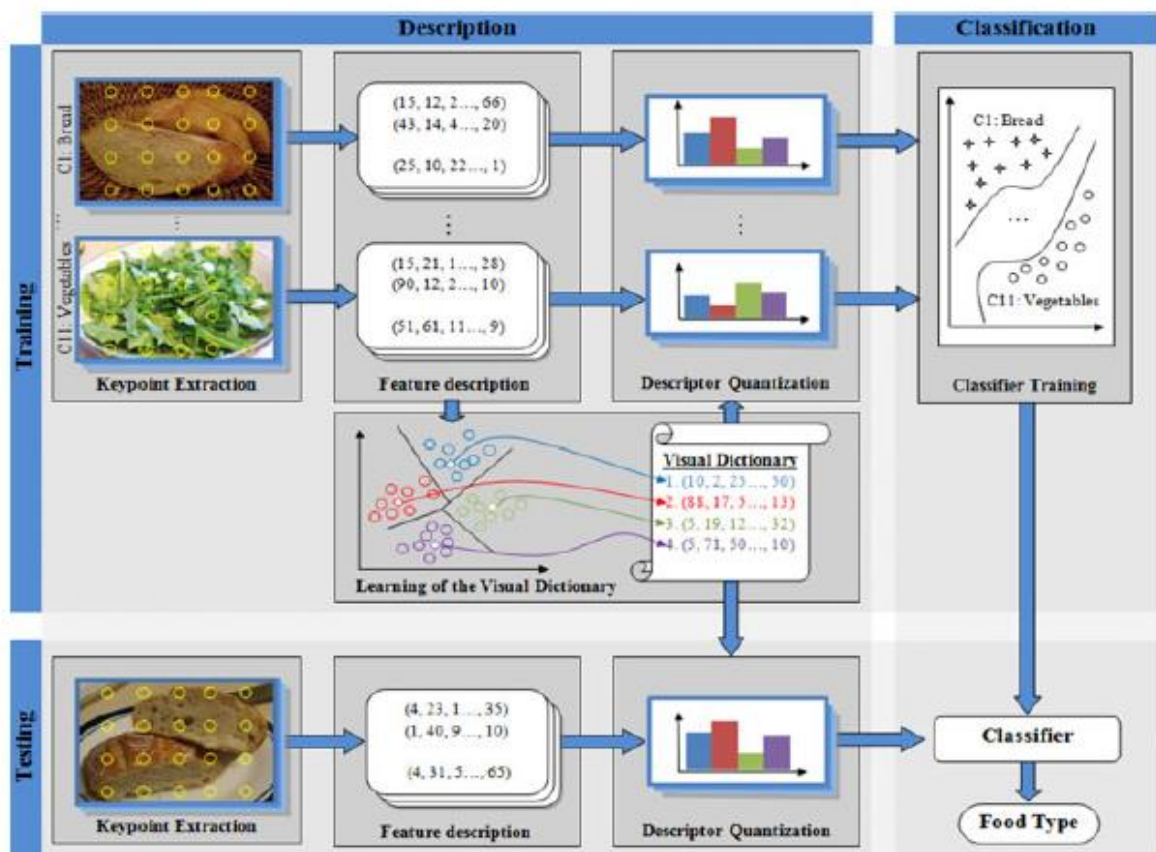
3.4.3.1 Χαρακτηριστικά και κωδικοποίηση χαρακτηριστικών

Προκειμένου να πραγματοποιηθεί η περιγραφή της εμφάνισης διαφορετικών κλάσεων τροφίμων, σε αυτή τη μελέτη όπως και στα περισσότερα πρόσφατα συστήματα αναγνώρισης τροφίμων, υιοθετήθηκε το μοντέλο Bag of Features (BoF).

Για την υλοποίηση του πρώτου σταδίου του BoF, χρησιμοποιήθηκαν τρεις διαφορετικές μέθοδοι ανίχνευσης σημείων ενδιαφέροντος: αλγόριθμοι ανίχνευσης σημείων ενδιαφέροντος και συγκεκριμένα ο SIFT [20], τυχαία δειγματοληψία και πυκνή δειγματοληψία.

Όσον αφορά την επιλογή κατάλληλου περιγραφέα, στην συγκεκριμένη προσέγγιση χρησιμοποιήθηκαν 7 περιγραφείς υφής, καθώς και 7 χρωματικοί περιγραφείς που αποτελούν παραλλαγές του αλγορίθμου ColorSIFT:

- *Παραλλαγές Color SIFT*: Ο Color SIFT υπολογίζεται με παρόμοιο τρόπο με τον SIFT, αλλά με μια μεγάλη διαφορά. Αντί να χρησιμοποιηθεί μόνο η ένταση της εικόνας, υπολογίζονται σε πολλά χρωματικά κανάλια τα ιστογράμματα της κλιμάκωσης του προσανατολισμού της εικόνας, επομένως ο προκύπτων περιγραφέας αποτελεί την συνένωση επιμέρους περιγραφών. Οι παραλλαγές αυτού που χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη είναι οι ακόλουθες: rgbSIFT, hsvSIFT, hueSIFT, opponentSIFT, cSIFT και rgSIFT.



Εικόνα 3.8: Αρχιτεκτονική συστήματος που αναπτύχθηκε. [23]

3.4.3.2 Ταξινόμηση

Προκειμένου να αναγνωριστεί ο πιο κατάλληλος ταξινομητής για το συγκεκριμένο πρόβλημα, διεξάχθηκαν πειράματα με τρεις διαφορετικές μεθόδους επιβλεπόμενης μάθησης: Μηχανές διανυσματικής υποστήριξης (Support vector machines – SVMs), Τεχνητά νευρωνικά δίκτυα (Artificial neural networks – ANNs) και τα Τυχαία Δάση (Random Forests – RFs). Στη συγκεκριμένη μελέτη χρησιμοποιήθηκαν SVMs με τρεις διαφορετικούς πυρήνες: Γραμμικός, RBF και $\exp X^2$, καθώς και δύο διαφορετικά μοντέλα ANN πρόσθιας

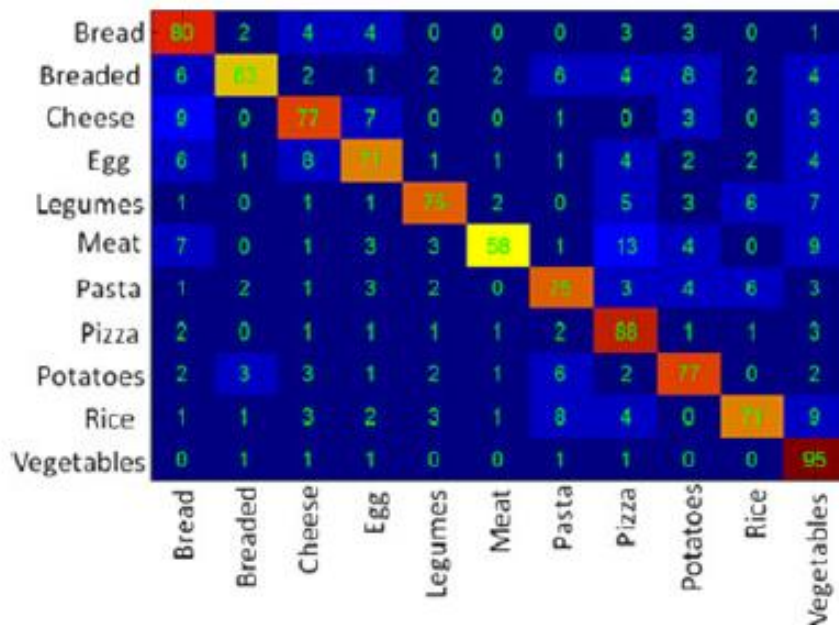
τροφοδότησης: ένα γραμμικό χωρίς κρυμμένα επίπεδα (ANNnh) και ένα μη γραμμικό με ένα κρυμμένο επίπεδο (ANNwh)[8]. Σχετικά με τα RF, χρησιμοποιήθηκε ένα RF που αποτελείται από 31 δέντρα, με κάθε διάσπαση να επιλέγει τυχαία έναν αριθμό χαρακτηριστικών ίσο με την τετραγωνική ρίζα του συνολικού αριθμού χαρακτηριστικών.

3.4.3.3 Αποτελέσματα

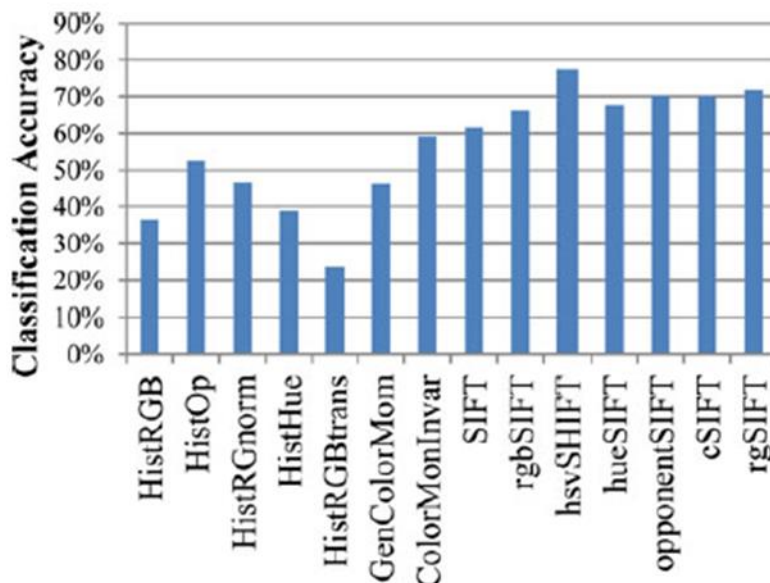
Για τις πειραματικές ανάγκες του συστήματος που αναπτύχθηκε, δημιουργήθηκε ένα σύνολο δεδομένων αποτελούμενο από 4868 έγχρωμες εικόνες από το διαδίκτυο. Συνολικά καθορίστηκαν 11 διαφορετικές κλάσεις τροφίμων, με 174 έως 677 εικόνες να ανήκουν στην κάθε κλάση. Η λήψη των εικόνων από το διαδίκτυο, οδήγησε στη δημιουργία ενός συνόλου δεδομένων, που να ανταποκρίνονται σε πραγματικές συνθήκες, με μεγάλη μεταβλητότητα εντός της κάθε κλάσης. Τέλος, οι εικόνες παρουσίασαν ένα μεγάλο εύρος συνθηκών φωτεινότητας, γωνιών παρατήρησης και διαφορετικών μερίδων τροφίμων μέσα στην ίδια κλάση.

Για την αξιολόγηση του συστήματος, εξετάστηκε η συνολική ακρίβεια αναγνώρισης (overall recognition accuracy – ORA), η οποία αντιπροσωπεύει το ποσοστό των σωστά ταξινομημένων εικόνων αξιολόγησης.

Το σύνολο των δεδομένων χωρίστηκε τυχαία σε ποσοστό 60% για τη φάση της εκπαίδευσης και το υπολειπόμενο 40% για την αξιολόγηση. Τέλος, για την αξιολόγηση του βελτιστοποιημένου BoF συστήματος, υιοθετήθηκε η προσέγγιση 5-fold cross validation. Τα αποτελέσματα που αντλήθηκαν φαίνονται στις παρακάτω εικόνες.



Εικόνα 3.9: Μήτρα σύγχυσης του προτεινόμενου συστήματος [23]



Εικόνα 3.10: Σύγκριση τοπικών περιγραφέων σε όρους συνολικής αξιολόγησης συστήματος. Χρησιμοποιήθηκαν 7 χρωματικοί περιγραφείς και 7 περιγραφείς υφής.

3.4.4 Σύγκριση των προσεγγίσεων που παρουσιάστηκαν

Τέλος, προκειμένου κανείς να έχει την δυνατότητα να αναγνωρίσει τα δυνατά σημεία, καθώς και τα ελαττώματα διάφορων αλγορίθμων στην εκτέλεση του έργου της αναγνώρισης των τροφίμων, παρουσιάζονται συνοπτικά στον πίνακα 3.1 οι μέθοδοι και οι τεχνικές που χρησιμοποιήθηκαν σε κάποιες προσεγγίσεις.

Όπως παρατηρείται, όσο αυξάνει ο αριθμός των κατηγοριών του συνόλου δεδομένων, το πλήθος των εικόνων και η οπτική ποικιλομορφία (εικόνες από πραγματικά γεύματα, διαφορετικός φωτισμός, γωνία λήψης κ.α.), η ακρίβεια ταξινόμησης του εκάστοτε συστήματος μειώνεται.

Στην επίλυση αυτού του προβλήματος, πλέον, προσανατολίζεται η ανάπτυξη εφαρμογών αναγνώρισης τροφίμων, ώστε να ανταποκρίνονται στις καθημερινές ανάγκες των μελλοντικών χρηστών τους.

Πίνακας 3.1: Σύγκριση διαφορετικών προσεγγίσεων αναγνώρισης τροφίμων.

	Σύνολο Δεδομένων	Εξαγωγή Χαρακτηριστικών	Μείωση Διάστασης	Ταξινόμηση	Αποτελέσματα
Shrof f et al.	fast food με κατηγορίες (χάμπουργκερ, τηγανιτές πατάτες, κοτομπουκιές, μηλόπιτα)	RGB τιμές, μέγεθος, υφή και σχήμα	-	τεχνητό νευρωνικό δίκτυο ANN εμπρόσθιας τροφοδότησης	95%
Zhu et al.	πραγματικά γεύματα και αντίγραφα εικόνων, 19 τροφικές κατηγορίες	χρωματικά χαρακτηριστικά, χαρακτηριστικά υφής (Gabor filter responses)	-	SVMs	πραγματικά γεύματα: 58% αντίγραφα εικόνων: 94%
Diet Cam	φρούτα, σπιτικό φαγητό, fast food	SIFT k-means (ομαδοποίηση)	-	Bayessian πιθανοτικός ταξινομητής	92% (6 εικόνες κατά την αναγνώριση)
σύστ ημα FIVR	εικόνες πραγματικών γευμάτων με 2 έως 20 κατηγορίες	συνιστώσες χρώματος, υφής, κλίμακας και επιλογή με χρήση Adaboost	Texton ιστογράμματα	SVMs	95% έως 80% με αύξηση των κατηγοριών
Snap- n-Eat	2000 εικόνες πραγματικών γευμάτων σε 15 κατηγορίες	dense HOG και dense SIFT	BoF και αναπαράστασ η Fisher	multiclass SVMs με χρήση μοντέλου τροφικής οντολογίας	85%
Anthi moro ulos et al.	4868 έγχρωμες εικόνες από το διαδίκτυο σε 11 διαφορετικές κατηγορίες	SIFT, τυχαία δειγματοληψία, πυκνή δειγματοληψία, παραλλαγές colorSIFT	BoF	SVMs, RFs, ANNnh, ANNwh	81%

4

Ανάπτυξη Συστήματος Αναγνώρισης Τροφίμων

4.1 Σχεδίαση συστήματος

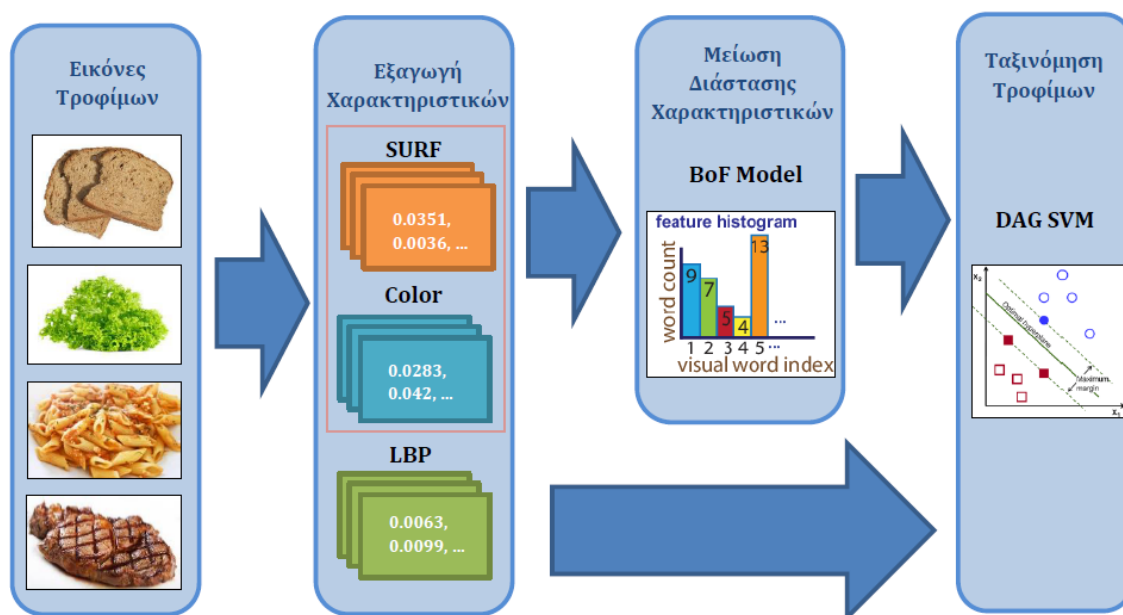
Το σύστημα που υλοποιήθηκε στην παρούσα εργασία αποτελείται από δύο στάδια σχεδιασμού και έχει ως στόχο την αναγνώριση των τροφίμων που δίνονται ως είσοδος στο σύστημα με τη μορφή εικόνων. Τα δύο επιμέρους στάδια από τα οποία αποτελείται είναι τα ακόλουθα:

1. Εξαγωγή χαρακτηριστικών
2. Ταξινόμηση

Κατά το στάδιο της εξαγωγής χαρακτηριστικών, ουσιαστικά, χρησιμοποιούνται διάφοροι αλγόριθμοι εξαγωγής και υπολογισμού χαρακτηριστικών προκειμένου να αναπαρασταθούν οι εικόνες ως διανύσματα και να εξάγουν χρήσιμη πληροφορία για την επίτευξη της μετέπειτα σωστής ταξινόμησης. Στη συγκεκριμένη εργασία, με κίνητρο την ικανότητα του αλγόριθμου SURF να καταγράφει μοτίβα χωρικής έντασης, καθώς και λόγω της σταθερότητας και της διακριτικής ικανότητας που παρέχουν τα χρωματικά και τα LBP (Local Binary Patterns) χαρακτηριστικά, εξετάζεται ένας συνδυασμός αυτών. Επιπλέον, σε αυτό το στάδιο, για τη μείωση της διάστασης του διανύσματος των χρησιμοποιούμενων χαρακτηριστικών, εφαρμόζεται η μέθοδος του Σάκου Χαρακτηριστικών (Bag of Features, BoF)

Όσον αφορά το δεύτερο στάδιο του προτεινόμενου συστήματος, χρησιμοποιείται η μέθοδος DAG SVM, η οποία αποτελεί μια μέθοδο ταξινόμησης βασισμένη στις Μηχανές Διανυσματικής Υποστήριξης, προσαρμοσμένη για προβλήματα πολλών κλάσεων. Η αξιολόγηση της μεθόδου αυτής γίνεται με σύγκρισή της με έναν ταξινομητή μηχανών υποστήριξης διανυσμάτων πολλών κλάσεων (multiclass SVM). Αναφορικά με το κίνητρο επιλογής της συγκεκριμένης τεχνικής, το DAG SVM χρησιμοποιεί την δομή ενός δέντρου απόφασης για την εκτέλεση της ταξινόμησης, επομένως, η επιλογή της κατάλληλης δομής έχει πολλά περιθώρια μελέτης με σκοπό την βελτίωση της ακρίβειας της ταξινόμησης.

Στην εικόνα 4.1 φαίνεται το σχηματικό διάγραμμα του προτεινόμενου συστήματος, στο οποίο διακρίνονται τα διάφορα στάδια σχεδίασης, καθώς και οι τεχνικές και αλγόριθμοι που χρησιμοποιήθηκαν σε καθένα από αυτά. Σε επόμενες ενότητες γίνεται μια εκτενέστερη ανάλυση όλων αυτών.



Εικόνα 4.1: Σχηματικό διάγραμμα του συστήματος που αναπτύχθηκε.

4.2 Εξαγωγή χαρακτηριστικών

4.2.1 Εισαγωγή

Προκειμένου να εφαρμοσθεί κάποια μέθοδος ταξινόμησης, όπως στην συγκεκριμένη περίπτωση τα SVMs, είναι συνήθως αναγκαίο τα δεδομένα να αναπαρασταθούν ως διανύσματα. Ειδικά, στην περίπτωση των εικόνων οι αναπαραστάσεις ονομάζονται διανύσματα περιγραφής (περιγραφείς της εικόνας), που μπορεί να αφορούν την υφή της εικόνας, το σχήμα, το χρώμα και άλλα χαρακτηριστικά αυτής. Για καθένα από τα παραπάνω χαρακτηριστικά έχουν προταθεί διάφοροι αλγόριθμοι εξαγωγής και υπολογισμού χαρακτηριστικών, καθένας από τους οποίους έχει διαφορετικά πλεονεκτήματα και γι' αυτό βρίσκουν εφαρμογή σε διαφορετικές περιπτώσεις.

Για την αναζήτηση σημείων ενδιαφέροντος σε εικόνες, τα τελευταία χρόνια έχουν χρησιμοποιηθεί ευρέως οι αλγόριθμοι SIFT (Scale-invariant feature transform) [41] και SURF (Speeded-Up Robust Features) [42], ενώ λιγότερο γνωστοί είναι άλλοι αλγόριθμοι όπως οι SPIN (Sorting points into neighborhoods) και RIFT (Rotation Invariant Feature Transform) [43]. Οι παραπάνω αλγόριθμοι εντοπίζουν στις εικόνες σημεία-κλειδιά (keypoints), και στη συνέχεια περιγράφονται με διανύσματα τα χαρακτηριστικά της εικόνας σε περιοχές (patches) με κέντρα τα σημεία-κλειδιά.

Για την περιγραφή του σχήματος, οι πιο απλοί αλγόριθμοι χρησιμοποιούν τους αλγόριθμους εντοπισμού ακμών Canny και Sobel/Prewitt, ενώ κάποιοι πιο σύνθετοι αλγόριθμοι που έχουν προταθεί είναι ο HOG (Histogram of Orientation Gradient) [44], καθώς και μια επέκτασή του ο PHOG (Pyramid Histogram of Orientation Gradients) [45].

Τέλος, για την περιγραφή της υφής μιας εικόνας, τα τελευταία χρόνια έχει χρησιμοποιηθεί ευρέως ο αλγόριθμος LBP (Local Binary Patterns) [46] λόγω της διακριτικής του ικανότητας και της χαμηλής υπολογιστικής πολυπλοκότητας.

Βέβαια, πέραν της εξαγωγής των χαρακτηριστικών της εικόνας, σημαντικό ρόλο στην ταξινόμησή της κατέχει και ο τρόπος αναπαράστασης των επιλεγμένων χαρακτηριστικών. Η πλέον συχνότερα χρησιμοποιούμενη μέθοδος αναπαράστασης των χαρακτηριστικών είναι η μέθοδος Bag of Features (BoF) [47] η σημαντικότητα της οποίας θα μελετηθεί σε επόμενη ενότητα.

Στην επόμενη ενότητα ακολουθεί η περιγραφή του αλγόριθμου SURF που επιλέχθηκε για το πρόβλημα το οποίο μελετάται, η μέθοδος BoF, ο αλγόριθμος LBP, καθώς και η ανάλυση ενός επιπλέον αλγορίθμου, του Color Extractor, του οποίου η χρήση κρίθηκε αναγκαία για το συγκεκριμένο πρόβλημα.

4.2.2 Αλγόριθμος SURF

Ο αλγόριθμος SURF [41], είναι ένας πλέον γνωστός και αποτελεσματικός αλγόριθμος ανίχνευσης και περιγραφής σημείων ενδιαφέροντος. Ο αλγόριθμος SURF προτάθηκε το 2006 από τους Bay, Tuytelaars και Van Gool.

Ο SURF είναι εν μέρει εμπνευσμένος από τον αλγόριθμο SIFT. Ο αλγόριθμος SIFT στην αρχή εξάγει από την εικόνα ορισμένα σημεία-κλειδιά ή σημεία ενδιαφέροντος όπως λέγονται, και στη συνέχεια ορίζει περιοχές ενδιαφέροντος (patches). Ο συγκεκριμένος αλγόριθμος στηρίζεται σε μια προσέγγιση διαδοχικών φίλτρων, δηλαδή αρχικά με χρήση αποτελεσματικών αλγορίθμων εντοπίζει τοποθεσίες στην εικόνα που είναι πιθανόν να αποτελούν περιοχές ενδιαφέροντος, ενώ στη συνέχεια εξετάζει με μεγαλύτερη λεπτομέρεια για να εντοπίσει με ακρίβεια τις περιοχές αυτές.

Βέβαια, η βασική έκδοση του SURF είναι πολλές φορές πιο γρήγορη από τον SIFT και οι εμπνευστές του ισχυρίζονται ότι είναι πολύ πιο ισχυρός από τον SIFT σε ότι αφορά διαφορετικούς μετασχηματισμούς εικόνας.

Ο αλγόριθμος SURF στηρίζεται στις ίδιες αρχές και βήματα με τον αλγόριθμο SIFT, και όπως αυτός, υλοποιεί έναν ανιχνευτή για τον προσδιορισμό των σημείων ενδιαφέροντος σε μια εικόνα.

Τα τρία βασικά βήματα του SURF είναι τα εξής: ο εντοπισμός σημείων ενδιαφέροντος, η περιγραφή της τοπικής γειτονιάς και το ταίριασμα.

Στο πρώτο βήμα, γίνεται ο εντοπισμός των σημείων ενδιαφέροντος σε μια εικόνα. Τα σημεία μοναδικότητας παραμένουν αμετάβλητα σε μεταβολές, όπως η αλλαγή κλίμακας μεγέθους, η περιστροφή, η αλλαγή φωτισμού ή ο θόρυβος. Έπειτα, η γειτονιά κάθε ενδιαφέροντος σημείου αναπαρίσταται από ένα χαρακτηριστικό διάνυσμα. Κάθε χαρακτηριστικό διάνυσμα χρησιμοποιείται στη διαδικασία ταύτισης κοινών σημείων ανάμεσα σε δύο εικόνες. Στο επόμενο βήμα, τα χαρακτηριστικά διανύσματα της εικόνας συγκρίνονται με αυτά μιας άλλης ώστε να βρεθούν κοινά σημεία ενδιαφέροντος. Ταυτίζοντας τα κοινά σημεία ανάμεσα σε δύο ή περισσότερες εικόνες μπορεί να υλοποιηθεί αναγνώριση και ταξινόμηση των εικόνων.

Ο αλγόριθμος SURF χρησιμοποιεί τετραγωνικά φίλτρα ως προσέγγιση της εξομάλυνσης της συνάρτησης Gauss. Το φιλτράρισμα της εικόνας με τετράγωνα είναι πολύ πιο γρήγορο αν χρησιμοποιείται το ολοκλήρωμα της εικόνας:

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j)$$

Το άθροισμα της αρχικής εικόνας μέσα σε ένα ορθογώνιο μπορεί πολύ γρήγορα να υπολογιστεί χρησιμοποιώντας την ολοκληρωμένη εικόνα, απαιτώντας υπολογισμούς στις 4 γωνίες του παραλληλογράμμου. Τα τετραγωνικά φίλτρα μεγέθους 9×9 τα οποία προσεγγίζουν την δεύτερης τάξης παράμετρο της συνάρτησης Gauss, αναπαριστούν την χαμηλότερη κλίμακα μεγέθους (υψηλότερη χωρική ανάλυση).

Ο SURF, προκειμένου να βρει τα σημεία ενδιαφέροντος, χρησιμοποιεί έναν περιγραφέα βασισμένο στον Γρήγορο Εσσιανό πίνακα (Fast Hessian Matrix), ο οποίος βασίζεται στους υπολογισμούς του Εσσιανού πίνακα σε κάθε σημείο της εικόνας. Η ορίζουσα του Εσσιανού πίνακα χρησιμοποιείται ως μέτρο της τοπικής αλλαγής γύρω από τα σημεία και τα σημεία επιλέγονται εκεί όπου η ορίζουσα αυτή είναι μέγιστη. Εκτός από την τοποθεσία, ο αλγόριθμος χρησιμοποιεί την ορίζουσα του Εσσιανού πίνακα και για τον χώρο κλιμάκωσης.

Έστω $p = (x, y)$ ένα σημείο της εικόνας I , ο Εσσιανός πίνακας στο σημείο p και με κλίμακα σ είναι:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix}$$

όπου $L_{xx}(p, \sigma)$ κ.λπ. είναι οι παράγωγοι δεύτερης τάξης της εικόνας στην κλίμακα του γκρι.

Η δημιουργία του περιγραφέα υλοποιείται σε δύο βήματα. Το πρώτο περιγράφεται από την εύρεση του αποτελεσματικού προσανατολισμού βασισμένου στη γύρω περιοχή του ενδιαφέροντος σημείου. Ο ιδανικός προσανατολισμός υπολογίζεται από τη σύγκριση των αθροισμάτων των αποκρίσεων που περιγράφονται από παράθυρα. Το μεγαλύτερο άθροισμα από τις συγκρινόμενες αντιστοιχίες δίνει τον προσανατολισμό του σημείου ενδιαφέροντος. Στη συνέχεια, για την εξαγωγή κατασκευάζεται μια τετραγωνική περιοχή (γύρω από την περιοχή των σημείων ενδιαφέροντος) και ευθυγραμμίζεται με τον προσανατολισμό που έχει επιλεγεί στο προηγούμενο βήμα. Τελικά, παράγεται το διάλυμα του περιγραφέα το οποίο συνήθως αποτελείται από 64 στοιχεία.

4.2.3 Αλγόριθμος Εξαγωγής Χρωματικών Χαρακτηριστικών (ColorExtractor)

Ο αλγόριθμος εξαγωγής χρωματικών χαρακτηριστικών ColorExtractor [48], ουσιαστικά χρησιμοποιείται για την εξαγωγή τοπικών χρωματικών χαρακτηριστικών από έγχρωμες εικόνες τύπου RGB και τελικά επιστρέφει ως μετρική την ένταση των χαρακτηριστικών αυτών.

Αρχικά, μετατρέπει τις RGB εικόνες στον χρωματικό χώρο L^*a^*b , όπου δίνεται η δυνατότητα να ποσοτικοποιηθούν πιο εύκολα οι διαφορές των χρωμάτων. Αυτό πρακτικά σημαίνει ότι οπτικά παρόμοια χρώματα στο χρωματικό πεδίο L^*a^*b θα έχουν μικρές διαφορές στις τιμές L^*a^*b που λαμβάνουν.

Στη συνέχεια, υπολογίζεται η μέση χρωματική τιμή L^*a^*b σε τετράγωνα των 4×4 pixel, η οποία στη συνέχεια χρησιμοποιείται ως η τιμή του αντίστοιχου χαρακτηριστικού. Η εικόνα πλέον μετατρέπεται σε πίνακα τριών στηλών και γραμμών ίσων με τα συνολικά χρωματικά χαρακτηριστικά.

Μετά την εξομάλυνση των χαρακτηριστικών, ακολουθεί η τεχνική της χωρικής επαύξησης (spatial augmentation), σύμφωνα με την οποία σε κάθε χρωματικό χαρακτηριστικό προσαρτάται και η θέση $[x, y]$ της εικόνας από την οποία έγινε η εξαγωγή του συγκεκριμένου χρωματικού χαρακτηριστικού. Με αυτόν τον τρόπο, η τεχνική αυτή ενσωματώνει τη χωρική διάταξη των χαρακτηριστικών μέσα σε μια εικόνα ως τμήμα του εξαγόμενου διανύσματος χαρακτηριστικών. Ως εκ τούτου, προκειμένου δύο εικόνες να έχουν παρόμοια χρωματικά χαρακτηριστικά θα πρέπει τόσο το χρώμα όσο και η χωρική κατανομή του χρώματος να είναι παρόμοια.

Τέλος, ακολουθεί η εξομάλυνση των συντεταγμένων ώστε να ανταποκρίνονται σε εικόνες διαφορετικού μεγέθους και η συνένωση των χρωματικών χαρακτηριστικών με τις χωρικές θέσεις, ενώ ως μετρική των χαρακτηριστικών τελικά επιστρέφεται η διακύμανσή τους.

4.2.4 Σάκος Χαρακτηριστικών (Bag of Features)

Η αναπαράσταση μιας εικόνας παίζει πολύ σημαντικό ρόλο στην ταξινόμησή της. Οι αλγόριθμοι SURF και Color Extractor που αναφέραμε παραπάνω, εξάγουν από κάθε εικόνα ένα πολύ μεγάλο πίνακα χαρακτηριστικών, διαφορετικού μεγέθους για κάθε εικόνα που λαμβάνουν ως είσοδο και ταυτόχρονα δύσκολα διαχειρίσιμο λόγω του μεγέθους του από τον εκάστοτε ταξινομητή. Επομένως, για την επίλυση του συγκεκριμένου προβλήματος και την αποτελεσματική αναπαράσταση μιας εικόνας χρησιμοποιείται η μέθοδος Σάκος Χαρακτηριστικών (Bag of Features - BoF).

Συγκεκριμένα, το μοντέλο Bag of Features (BoF) [47], έχει αποδειχθεί ότι έχει την ικανότητα να ανταπεξέρχεται πολύ καλά σε εικόνες που χαρακτηρίζονται από υψηλή οπτική ποικιλομορφία και έλλειψη τυπικής χωρικής διάταξης μέσα σε κάθε κλάση. Επίσης, έχει δείξει ευρωστία σε διάφορα backgrounds και σε περιπτώσεις επικάλυψης αντικειμένων μέσα σε μια εικόνα. Η εκτέλεση αυτού του αλγορίθμου βασίζεται στην εξαγωγή τοπικών χαρακτηριστικών.

Η διαδικασία εκτέλεσης της μεθόδου BoF έχει ως εξής:

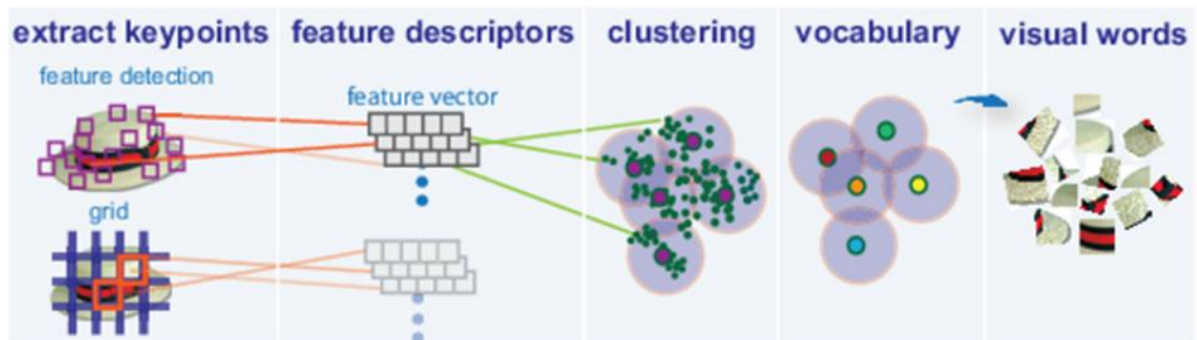
1. Αρχικά, από μια κλάση εικόνων εξάγονται τα τοπικά χαρακτηριστικά μέσω αλγορίθμων ανίχνευσης σημείων ενδιαφέροντος. Όλα τα χαρακτηριστικά των εικόνων μιας κλάσης τοποθετούνται μαζί σε ένα χώρο που λέγεται «Σάκος οπτικών λέξεων» (Bag of Visual Words).

Στην παρούσα εργασία χρησιμοποιούνται δύο αλγόριθμοι εξαγωγής χαρακτηριστικών, ο SURF και ο Color Extractor. Ένας επίσης πολύ γνωστός και λειτουργικός αλγόριθμος για τη συγκεκριμένη λειτουργία είναι και ο SIFT. Η περιγραφή των δύο αυτών αλγορίθμων έγινε στην προηγούμενη παράγραφο.

2. Έπειτα, μέσα από έναν αλγόριθμο ομαδοποίησης (εδώ χρησιμοποιήθηκε ο k-means) δημιουργούνται k αντιπροσωπευτικοί περιγραφείς (visual words). Οι αντιπροσωπευτικοί περιγραφείς αποτελούν το λεξικό οπτικών λέξεων (visual dictionary). Το μέγεθος του visual dictionary είναι ο αριθμός k στον οποίο χωρίσαμε το πλήθος όλων των περιγραφέων.
3. Στη συνέχεια, ελέγχεται ποιος αντιπροσωπευτικός περιγραφέας είναι ο πιο κοντινός για κάθε τοπικό περιγραφέα. Η απόσταση μεταξύ ενός τοπικού περιγραφέα και ενός αντιπροσωπευτικού περιγραφέα μετριέται με Ευκλείδεια απόσταση.
4. Τέλος, δημιουργείται ένα ιστόγραμμα χαρακτηριστικών για κάθε εικόνα το οποίο έχει μήκος ίσο με k και δείχνει πόσοι τοπικοί περιγραφείς της εικόνας ταιριάζουν με κάθε αντιπροσωπευτικό περιγραφέα (visual word). Έπειτα, οι τιμές των ιστογραμμάτων τροφοδοτούν τον ταξινομητή, είτε για την εκπαίδευση είτε για τη δοκιμή. Η αποτελεσματικότητα αυτού του τμήματος της διαδικασίας είναι ζωτικής σημασίας, καθώς επηρεάζει τους χρόνους επεξεργασίας τόσο της φάσης της εκπαίδευσης όσο και της δοκιμής. Η πολυπλοκότητα της κβάντωσης του περιγραφέα εξαρτάται από τις διαστάσεις του περιγραφέα και τον αριθμό των οπτικών λέξεων.

Στις εικόνες 4.1 και 4.2 απεικονίζεται γραφικά η διαδικασία που περιγράψαμε παραπάνω. Πιο συγκεκριμένα, στην εικόνα 4.1 μπορεί κανείς να παρατηρήσει τα βήματα 1 και 2 της διαδικασίας, όπου δημιουργείται ο σάκος των οπτικών λέξεων (visual words), ενώ στη συνέχεια στην εικόνα 4.2 φαίνονται τα βήματα 3 και 4 της διαδικασίας όπου

δημιουργείται το ιστόγραμμα των χαρακτηριστικών και το διάνυσμα με τις τιμές του ιστογράμματος που χρησιμοποιούνται στην εκπαίδευση και στην αξιολόγηση ταξινομητή.



Εικόνα 4.2: Γραφική απεικόνιση της διαδικασίας δημιουργίας των οπτικών λέξεων μέσω της μεθόδου Bag of Features.

Παρόλο που ο αλγόριθμος BoF έχει γνωρίσει μεγάλη επιτυχία και έχει χρησιμοποιηθεί σε πολλές εφαρμογές, έχει ένα σημαντικό μειονέκτημα. Δεν λαμβάνει καθόλου υπόψη τη χωρική αναπαράσταση της εικόνας, δηλαδή ταιριάζει τις εικόνες σύμφωνα με τα κοινά τοπικά χαρακτηριστικά, χωρίς να ενδιαφέρεται για το σε ποιο σημείο της εικόνας βρίσκονται αυτά, και έτσι αγνοεί το καθολικό θέμα της εικόνας.

Για το λόγο αυτό, επιπρόσθετα με τα χαρακτηριστικά που εξάγονται με τη μέθοδο BoW και τους αλγορίθμους SURF και Color Extractor, χρησιμοποιήσαμε και κάποια χωρικά χαρακτηριστικά που βασίζονται στην υφή της εικόνας, τα Τοπικά Δυαδικά Πρότυπα (Local Binary Patterns), τα οποία παρουσιάζονται παρακάτω.



Εικόνα 4.3: Γραφική απεικόνιση της διαδικασίας εξαγωγής του διανύσματος των χαρακτηριστικών που χρησιμοποιείται ως είσοδος στον ταξινομητή με χρήση του Bag of Features.

4.2.5 Τοπικά Δυαδικά Πρότυπα (*Local Binary Patterns*)

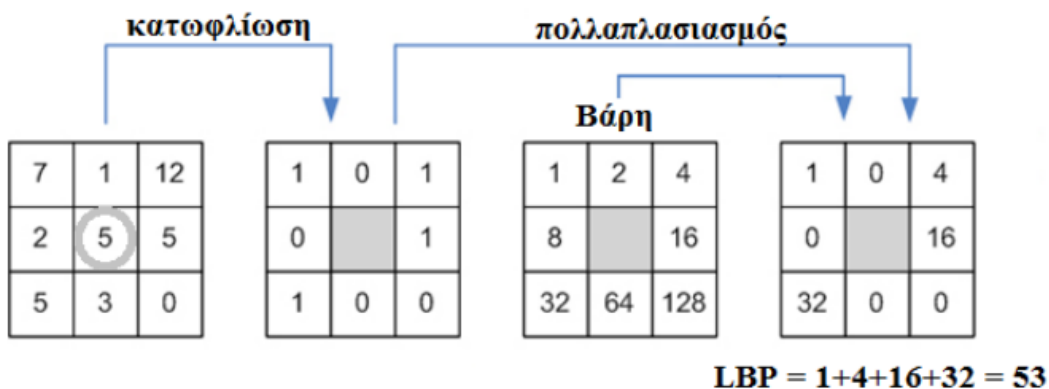
Τα Τοπικά Δυαδικά Πρότυπα (*Local Binary Patterns - LBP*) [46] προτάθηκαν το 1996 από τους Ojala κ.α. και έκτοτε έχουν χρησιμοποιηθεί ευρύτατα σε πολλές εφαρμογές ως ισχυρά και αποτελεσματικά χαρακτηριστικά υφής χάριν της διακριτικής τους ικανότητας και της χαμηλής υπολογιστικής πολυπλοκότητας. Τα LBP ουσιαστικά αποτυπώνουν και κωδικοποιούν τοπικά χαρακτηριστικά των εικόνων, τη γειτονιά, δηλαδή, γύρω από κάθε εικονοστοιχείο, και η κατανομή τους χαρακτηρίζει τις εικόνες. Η βασική ιδέα για την ανάπτυξη των LBP ήταν ότι η υφή σε δυσδιάστατες επιφάνειες μπορεί να περιγραφεί από δύο συμπληρωματικά μέτρα: τοπικά χωρικά πρότυπα (υπολογίζονται από τα LBP) και αντίθεση στην κλίμακα του γκρι (συνήθως χρησιμοποιείται συμπληρωματικά με τα LBP).

Στην αρχική τους μορφή τα LBP υπολογίζονται ως εξής:

Σε μια γειτονιά 3×3 pixel, η οποία ονομάζεται και μάσκα, η φωτεινότητα του κεντρικού εικονοστοιχείου λαμβάνεται ως τοπικό κατώφλι με αποτέλεσμα τα 8 γειτονικά εικονοστοιχεία να πάρουν την τιμή 1 ή 0 ανάλογα εάν η φωτεινότητά τους είναι μεγαλύτερη (ή ίση) ή μικρότερη από την τιμή του κατωφλίου (δηλαδή τη φωτεινότητα του κεντρικού εικονοστοιχείου). Όπως φαίνεται και στην εικόνα 4.3, το κατώφλι (κεντρικό εικονοστοιχείο) έχει τιμή 5. Όλα τα εικονοστοιχεία που έχουν φωτεινότητα μεγαλύτερη ή ίση από 5 παίρνουν την τιμή 1 ενώ όλα τα υπόλοιπα την τιμή 0. Στη συνέχεια, η τιμή κάθε εικονοστοιχείου πολλαπλασιάζεται με ένα βάρος (δύναμη του 2) ανάλογα με τη θέση του στη 3×3 γειτονιά. Για παράδειγμα, το πάνω αριστερά εικονοστοιχείο (σε σχέση με το κεντρικό) έχει βάρος $2^0=1$, το δεξιά του έχει βάρος $2^1=2$, ενώ το κάτω δεξιά έχει βάρος $2^7=128$. Τέλος, οι τιμές των οκτώ εικονοστοιχείων αθροίζονται ώστε να προκύψει η τιμή του LBP για τη συγκεκριμένη περιοχή η οποία είναι και η τιμή για το συγκεκριμένο πρότυπο (ή μονάδα υφής). Αυτή η τιμή προφανώς είναι ακέραιη, πραγματική στο διάστημα $[0, 255]$ αφού από 8 γειτονικά εικονοστοιχεία μπορούν να προκύψουν $2^8=256$ διαφορετικοί συνδυασμοί. Με άλλα λόγια, ένα σύνολο από 256 διαφορετικά πρότυπα μπορεί να προκύψει από μια γειτονιά 3×3 . Εφαρμόζοντας την παραπάνω διαδικασία σε όλη την εικόνα μπορούμε να σχηματίσουμε το ιστόγραμμα των LBP τιμών το οποίο μπορεί να ληφθεί ως το διάλυμα χαρακτηριστικών της συγκεκριμένης εικόνας. Μια χρήσιμη επέκταση του LBP είναι η ομοιόμορφη LBP, η οποία μειώνει το μήκος του αρχικού διανύσματος χαρακτηριστικών από 256 σε 59. Η επέκταση αυτή χρησιμοποιείται και στη συγκεκριμένη εργασία.

Αξίζει να αναφερθεί ότι τα LBP χαρακτηριστικά είναι ανθεκτικά σε μονοτονικές αλλαγές της κλίμακας του γκρι που οφείλονται π.χ. σε διαφορετικές συνθήκες φωτισμού.

Δηλαδή, εάν οι αρχικές τιμές φωτεινότητας αυξηθούν όλες κατά μια σταθερή τιμή τότε οι τιμές των LBP δεν μεταβάλλονται. Αντίθετα, πρέπει να τονισθεί ότι η σειρά ανάθεσης των βαρών στα εικονοστοιχεία δεν υπακούει κάποιον κανόνα και είναι στην ευχέρεια του χρήστη. Αυτό αποτελεί μειονέκτημα καθώς αλλαγή στη σειρά από τα βάρη συνεπάγεται αλλαγή στην τιμή του LBP.



Εικόνα 4.4: Υπολογισμός LBP χαρακτηριστικών σε γειτονία 3x3 pixel.

4.3 Ταξινόμηση

4.3.1 Μηχανές Υποστήριξης Διανυσμάτων (Support Vector Machines)

Οι Μηχανές Υποστήριξης Διανυσμάτων (Support Vector Machines - SVMs) είναι μια μέθοδος ταξινόμησης που αναπτύχθηκε από τον Vapnik και την ομάδα του στα AT&T Bell Labs. Τα SVMs εφαρμόστηκαν σε προβλήματα ταξινόμησης ως εναλλακτικές μέθοδοι για τα πολυστρωματικά δίκτυα. Η ταξινόμηση πετυχαίνεται από μια γραμμική ή μια μη γραμμική διαχωριστική επιφάνεια στο χώρο εισόδου του συνόλου των δεδομένων. Στόχος των SVMs είναι η ελαχιστοποίηση του σφάλματος ταξινόμησης. Τα SVMs ουσιαστικά μεταφέρουν ένα δοσμένο δυαδικό σύνολο δεδομένων σε ένα χώρο χαρακτηριστικών υψηλής διάστασης και διαχωρίζουν τις δύο κλάσεις των δεδομένων με ένα υπερεπίπεδο μεγίστου περιθωρίου.

Η χρήση των SVMs αν και έχει αποδειχθεί ως μια πολύ αποτελεσματική μέθοδος μηχανικής μάθησης για δυαδικά προβλήματα ταξινόμησης, δηλαδή για προβλήματα ταξινόμησης δύο κλάσεων, η επέκτασή της σε προβλήματα πολλών κλάσεων εξακολουθεί ακόμα να αποτελεί ένα ζήτημα που χρήζει περαιτέρω έρευνας. Εδώ πρέπει να δοθεί έμφαση στο γεγονός ότι, προκειμένου να υλοποιηθεί ένα σύστημα για την επίλυση προβλημάτων

πολλών κλάσεων με χρήση των SVMs, απαιτείται ο αριθμός των μεταβλητών του προβλήματος να είναι ανάλογος με τον αριθμό των κλάσεων αυτού. Επομένως, η επίλυση ενός τέτοιου προβλήματος είναι υπολογιστικά πολύ πιο δαπανηρή από την επίλυση ενός αντίστοιχου δυαδικού προβλήματος με τον ίδιο αριθμό δεδομένων εισόδου.

Μέχρι σήμερα, έχουν προταθεί πολλές παραλλαγές του βασικού αλγορίθμου για την επίλυση προβλημάτων πολλών κλάσεων με εφαρμογή των SVMs, όπως για παράδειγμα οι μέθοδοι Ένας-Εναντίον-Ενός (One-Against-One), Ένας-Εναντίον-Όλων (One-Against-All), Όλοι-Μαζί (All-Together) και Κατευθυνόμενος Ακυκλικός Γράφος με SVMs (Direct Acyclic Graph - DAG SVMs).

Έστω, τώρα, ότι πρέπει να επιλυθεί ένα πρόβλημα ταξινόμησης k διαφορετικών κλάσεων (πρόβλημα τάξης k). Για ένα τέτοιο πρόβλημα, η μέθοδος one-against-all κατασκευάζει k υπερεπίπεδα, όπου καθένα από αυτά έχει δημιουργηθεί με τη χρήση όλων των δεδομένων του συνόλου εκπαίδευσης. Η μέθοδος all-together επίσης κατασκευάζει k υπερεπίπεδα παρόμοια με την προηγούμενη μέθοδο, με τη διαφορά όμως ότι κάθε υπερεπίπεδο λαμβάνεται με επίλυση ενός προβλήματος βελτιστοποίησης. Η μέθοδος one-against-one κατασκευάζει $k(k-1)/2$ υπερεπίπεδα, όπου καθένα από αυτά κατασκευάζεται χρησιμοποιώντας δεδομένα δύο κλάσεων εκ των k του συνόλου εκπαίδευσης. Αν και στην τελευταία μέθοδο κατασκευάζονται $k(k-1)/2$ υπερεπίπεδα, ωστόσο ο χρόνος εκπαίδευσης αυτών είναι αρκετά μικρότερος από αυτόν των προηγούμενων μεθόδων, δεδομένου ότι κάθε SVM χρησιμοποιεί έναν μικρό αριθμό δεδομένων του συνόλου εκπαίδευσης κατά την εκμάθηση. Όσον αφορά τη μέθοδο DAG SVM, χρησιμοποιεί και αυτή $k(k-1)/2$ υπερεπίπεδα, όμως ο χρόνος δοκιμής είναι μικρότερος από εκείνον της μεθόδου one-against-one.

Σε έρευνά τους, ο Hsu και Lin [36] σύγκριναν τις προαναφερθείσες μεθόδους ως προς την επίδοσή τους και το υπολογιστικό τους κόστος και κατέληξαν στο συμπέρασμα ότι από πλευράς χρόνου εκπαίδευσης καμία μέθοδος δεν μπορεί να ανταγωνιστεί την μέθοδο one-against-one και κατ' επέκταση της μεθόδου DAG SVM, ενώ από πλευράς επίδοσης γενίκευσης (generalization performance) στατιστικά δεν φαίνεται καμία μέθοδος να υπερισχύει των υπολοίπων. Επομένως, οι μέθοδοι one-against-one και DAG SVM είναι καταλληλότερες για πρακτική χρήση σε σχέση με τις υπόλοιπες προαναφερθείσες μεθόδους.

Στην παρούσα εργασία παρουσιάζεται μια μεθοδολογία βασισμένη στη δομή των DAG SVMs, αλλά με χρήση ενός βελτιστοποιημένου αλγόριθμου επιλογής. Προκειμένου να επιλεγεί η συγκεκριμένη μεθοδολογία για τη αντιμετώπιση του προβλήματος της συγκεκριμένης εργασίας, λήφθηκαν υπόψη τα αποτελέσματα και οι επιδόσεις διάφορων αλγορίθμων ταξινόμησης πολλών κλάσεων, μερικά από αυτά παρουσιάστηκαν παραπάνω.

4.3.2 Δυαδικά SVMs

Προκειμένου να κατανοήσει κανείς πλήρως τον τρόπο λειτουργίας των SVMs, πρέπει να έχει κατανοήσει δύο έννοιες: τη δυαδικότητα και τους πυρήνες. Σε αυτή την ενότητα λοιπόν, θα γίνει μια παρουσίαση των δυαδικών SVMs (γνωστά ως binary SVMs), τα οποία αποτελούν όχι μόνο την πρωταρχική δομή των μηχανών διανυσμάτων υποστήριξης, αλλά και τη βάση στην οποία στηρίζεται η μεθοδολογία η οποία αναπτύσσεται στο υπόλοιπο της εργασίας. Πριν, βέβαια, γίνει η ανάλυση της λειτουργίας τους, πρέπει να αναφερθεί ότι τα δυαδικά SVMs ουσιαστικά χρησιμοποιούνται στο διαχωρισμό δύο και μόνο κλάσεων.

Τα SVM παράγουν ένα μη γραμμικό όριο ταξινόμησης στον αρχικό χώρο εισόδου μέσω της κατασκευής ενός γραμμικού υπερεπίπεδου σε μια μετασχηματισμένη εκδοχή του αρχικού χώρου εισόδου. Ο χώρος εισόδου μετασχηματίζεται σε έναν πολλών διαστάσεων χώρο χαρακτηριστικών με την χρήση μιας μη γραμμικής συνάρτησης.

Έστω ότι έχουμε ως δεδομένα l δείγματα εκπαίδευσης $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$, όπου $x_i \in \mathbb{R}^d$, $i = 1, 2, \dots, l$ και $y_i \in \{1, -1\}$ η ετικέτα της κλάσης. Αν τα συγκεκριμένα δείγματα εκπαίδευσης είναι γραμμικώς διαχωρίσιμα στο πεδίο των χαρακτηριστικών, τότε μπορούμε να εκφράσουμε τη συνάρτηση διαχωρισμού στην παρακάτω μορφή:

$$f(x) = \langle \varphi(x) \cdot w \rangle + b = 0$$

όπου $\varphi(x)$ είναι μια μη γραμμική συνάρτηση που περνάει την τιμή του x στο πεδίο των χαρακτηριστικών, w είναι ένα διάνυσμα βάρους και b είναι ένας σταθερός όρος. Το SVM υπολογίζει το διαχωριστικό υπερεπίπεδο των δύο κλάσεων στο σημείο όπου η απόσταση κάθε κλάσης από το υπερεπίπεδο μεγιστοποιείται.

Αν υποθέσουμε ότι, τα κοντινότερα σημεία βρίσκονται πάνω στην $f(x_i) = \pm 1$ για κάποια i . Τότε το περιθώριο ορίζεται ως:

$$\gamma = \frac{1}{\|w\|^2}$$

Επομένως, το πρόβλημα του SVM εκφράζεται από το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\min \frac{1}{2} \|w\|^2$$

το οποίο υπόκειται στον ακόλουθο περιορισμό:

$$y_i(\langle \varphi(x_i) \cdot w \rangle + b) \geq 1, \quad i = 1, 2, \dots, l.$$

Όταν τα δεδομένα εκπαίδευσης δεν είναι γραμμικώς διαχωρίσιμα στο πεδίο των χαρακτηριστικών, τότε εισάγουμε τις μεταβλητές χαλάρωσης $\xi_i (> 0)$ στις παραπάνω εξισώσεις και το πρόβλημα βελτιστοποίησης μετασχηματίζεται ως εξής:

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i$$

το οποίο υπόκειται στον παρακάτω περιορισμό:

$$y_i(\langle \varphi(x_i) \cdot w \rangle + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \text{ και } i = 1, 2, \dots, l$$

όπου C είναι μια παράμετρος που καθορίζει το trade off μεταξύ του μέγιστου περιθωρίου και του ελάχιστου σφάλματος ταξινόμησης.

Χρησιμοποιώντας τη συνάρτηση Lagrangian, λαμβάνουμε τη δυική αντικειμενική συνάρτηση, όπου αποτελεί ένα τετραγωνικό πρόβλημα προγραμματισμού που ορίζεται ως εξής:

$$\max \sum_1^l a_i - \frac{1}{2} \sum_{i,j=1}^l a_i a_j y_i y_j K(x_i, x_j)$$

με τον περιορισμό:

$$\sum_{i=1}^l a_i y_i = 0, \quad 0 \leq a_i \leq C \text{ και } i = 1, 2, \dots, l$$

όπου a_i είναι οι πολλαπλασιαστές Lagrange και $K(x_i, x_j)$ είναι ο πυρήνας που παράγει το εσωτερικό γινόμενο $\langle \varphi(x_i) \cdot \varphi(x_j) \rangle$ στο πεδίο των χαρακτηριστικών. Η συνάρτηση απόφασης υπολογίζεται στο χώρο των χαρακτηριστικών ως εξής:

$$f(x) = \sum_{i=1}^l a_i y_i K(x_i, x) + b$$

Συχνά, συμβαίνει ένα μεγάλο μέρος των τιμών του l του a_i να είναι 0. Μόνο τα σημεία που βρίσκονται πλησιέστερα στο υπερεπίπεδο, συμπεριλαμβανομένου των σημείων που κείνται στη λανθασμένη πλευρά του υπερεπιπέδου, αντιστοιχούν σε μη μηδενικές τιμές των a_i . Τα σημεία αυτά ονομάζονται διανύσματα υποστήριξης (support vectors).

Ο πυρήνας $K(x_i, x_j)$ χρησιμοποιείται για την κατασκευή του βέλτιστου υπερεπιπέδου στο χώρο των χαρακτηριστικών, χωρίς να λάβει υπόψη τον ίδιο το χώρο σε ρητή μορφή. Η απαίτηση του πυρήνα είναι να ικανοποιεί το θεώρημα Mercer. Οι πιο συνηθισμένοι τύποι πυρήνων είναι ο πολυωνυμικός, ο Gaussian, ο σιγμοειδής, ο γραμμικός και ο τετραγωνικός.

Μερικά παραδείγματα συναρτήσεων πυρήνα φαίνονται παρακάτω:

1. Πυρήνας Gauss (Ακτινικής Βάσης Συνάρτηση Πυρήνα ή σφαιρικός GRBF πυρήνας)

$$K(x, x') = \exp\left(\frac{-\|x - x'\|}{2\sigma^2}\right)$$

όπου $\sigma > 0$, είναι η παράμετρος πυρήνα γνωστή και ως εύρος Gauss.

2. Ελλειψοειδής GRBF πυρήνας

$$K(x, x') = \exp\left(-\sum_{i=1}^N \frac{\|x_i - x'_i\|}{2\sigma_i^2}\right)$$

όπου το σ_i είναι το εύρος Gauss για την i -οστή διάσταση.

3. Ο πολυωνυμικός πυρήνας

$$K(x, x') = (1 + c \langle x, x' \rangle)^d$$

όπου $d \in \mathbb{N}$ και $c > 0$.

4. Ο πυρήνας νευρωνικού δικτύου (σιγμοειδής πυρήνας)

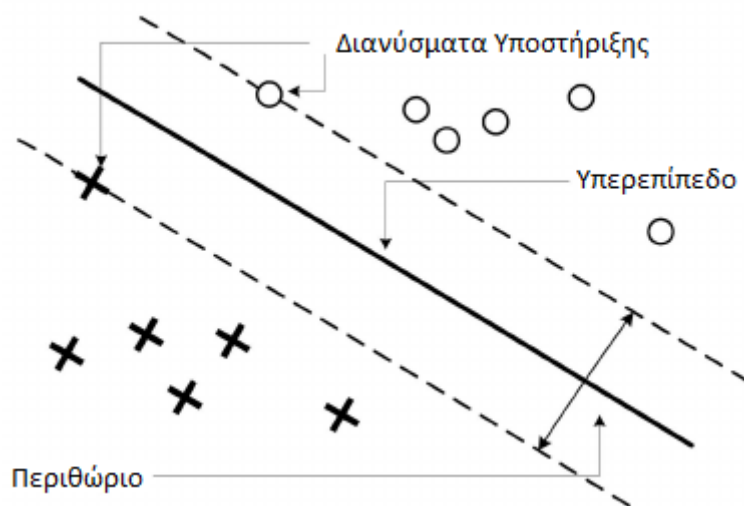
$$K(x, x') = \tanh(b \langle x, x' \rangle - c)$$

όπου $b > 0$ και $c \in \mathbb{R}$.

Οι παράμετροι στους πυρήνες, όπως είναι οι σ , c , d και b , παραπάνω, ονομάζονται παράμετροι πυρήνα. Ο πυρήνας Gauss και ο πολυωνυμικός πυρήνας είναι πυρήνες Mercer. Η κλίση b μπορεί να υπολογιστεί λαμβάνοντας οποιοδήποτε διάνυσμα υποστήριξης x_j με $a_j < C$ (τα δείγματα εκπαίδευσης x_j με μη μηδενικές μεταβλητές χαλάρωσης αντιστοιχούν σε τιμές $a_j = C$), και ως εκ τούτου:

$$b = y_j - \sum_{i=1}^l y_i a_i K(x_i, x_j)$$

Τέλος, παρακάτω στην εικόνα 4.4 απεικονίζεται γραφικά ένα πρόβλημα ταξινόμησης δύο κλάσεων με χρήση SVM.



Εικόνα 4.5: Πρόβλημα ταξινόμησης δύο κλάσεων με χρήση SVM.

4.3.3 DAG SVMs

4.3.3.1 Γενικά

Μέχρι τώρα στα δυαδικά SVMs έχουμε υποθέσει ότι τα δεδομένα είναι μόνο σε δύο κλάσεις. Πολλές, όμως, πρακτικές εφαρμογές περιλαμβάνουν περισσότερες κλάσεις, όπως για παράδειγμα, το πρόβλημα που μελετάει η συγκεκριμένη εργασία θεωρεί 6 διαφορετικές κατηγορίες γευμάτων.

Όπως προαναφέραμε, υπάρχουν πολλοί τρόποι να επεκτείνουμε τα SVMs για τέτοιες περιπτώσεις. Στη συνέχεια, θα εξεταστεί μια από αυτές τις μεθόδους η οποία βασίζεται στην μέθοδο Decision Directed Acyclic Graph (DDAG) των Platt, Cristianini, και Shawe-Taylor [37]. Προκειμένου όμως να γίνει κατανοητή η συγκεκριμένη μέθοδος, πρέπει να γίνει αναφορά στα βασικά χαρακτηριστικά των μεθόδων one-against-one και DAG SVM.

Στη μέθοδο one-against-one, αρχικά καθορίζεται η συνάρτηση απόφασης όλων των δυνατών ανά δύο συνδυασμών των κλάσεων που δίνονται. Για τον καθορισμό κάθε τέτοιας συνάρτησης, χρησιμοποιούνται μόνο τα δεδομένα εκπαίδευσης που αντιστοιχούν στις δύο κλάσεις που ανήκουν στο ζεύγος. Έτσι, στην εκπαίδευση κάθε τέτοιου ταξινομητή, τα δεδομένα εισόδου είναι αρκετά λιγότερα από το σύνολο των δεδομένων εκπαίδευσης, αλλά ο αριθμός των συναρτήσεων απόφασης αυξάνεται και αντιστοιχεί σε $k(k-1)/2$ συναρτήσεις απόφασης. Την εκπαίδευση ακολουθεί η ταξινόμηση των δεδομένων εισόδου. Για την ταξινόμηση, συνήθως χρησιμοποιείται ο αλγόριθμος «max wins». Σύμφωνα με τον αλγόριθμο «max wins», κάθε ταξινομητής δίνει μία ψήφο στην προτιμώμενη κλάση, και το τελικό αποτέλεσμα είναι η κλάση με τις περισσότερες ψήφους.

Στη μέθοδο DAG SVM χρησιμοποιείται ένας δυαδικός κατευθυνόμενος ακυκλικός γράφος, όπου κάθε κόμβος συνδέεται με το επόμενο επίπεδο με 0 ή 2 ακμές προς άλλους κόμβους. Στον κόμβο ρίζα δεν υπάρχει καμία ακμή να φτάνει σε αυτόν. Για ένα πρόβλημα k κλάσεων, ο γράφος που χρησιμοποιείται έχει $k(k-1)/2$ εσωτερικούς κόμβους και k φύλλα. Κάθε κόμβος συσχετίζεται με ένα δυαδικό SVM, ενώ τα φύλλα έχουν επισημανθεί με τα ονόματα των κλάσεων.

Προκειμένου να καθοριστεί η κλάση ενός σημείου x , ξεκινώντας από τη ρίζα, υπολογίζεται η συνάρτηση απόφασης σε αυτό τον κόμβο. Στη συνέχεια, κινείται είτε προς τα δεξιά είτε προς τα αριστερά, ανάλογα με την τιμή της εξόδου του αποτελέσματος. Ως εκ τούτου, ακολουθείται μια διαδρομή μέχρι να φτάσει στον κόμβο φύλλο που καθορίζει την κλάση του δεδομένου εισόδου. Έτσι, για ένα πρόβλημα τάξης k , ουσιαστικά υπολογίζονται $k-1$ κόμβοι απόφασης προκειμένου να καθοριστεί η κλάση της εισόδου.

Η φάση εκπαίδευσης της συγκεκριμένης μεθόδου είναι ακριβώς η ίδια με την μέθοδο one-against-one, αλλά ο χρόνος ταξινόμησης είναι πολύ μικρότερος. Το αρνητικό της μεθόδου σε σχέση με τον αλγόριθμο «max wins» είναι ότι δεν μπορεί να επιλύσει τις περιοχές που υπάρχει «ισοπαλία», και έτσι τα αποτελέσματα μπορεί να διαφέρουν ανάλογα με τη συνάρτηση απόφασης στους κόμβους με διαφορετικά ζεύγη κλάσεων.

Η μέθοδος DAG SVM, έχει πολλά κοινά με τη μέθοδο ταξινόμησης one-against-one, στην οποία ένα πρόβλημα k κλάσεων μετατρέπεται σε $k(k-1)/2$ προβλήματα δύο κλάσεων. Έρευνα που διεξάχθηκε από τον Krebel [4] έδειξε ότι, με την one-against-one τεχνική, οι αταξινομήτες περιοχές αν και μειώνονται σε σχέση με τις υπόλοιπες μεθόδους ταξινόμησης πολλών κλάσεων, εξακολουθούν να παραμένουν.

Για την επίλυση του συγκεκριμένου προβλήματος, προτάθηκε η δομή δέντρου απόφασης. Η μέθοδος αυτή παρουσιάστηκε πρώτα από τους Platt, Cristianini και Taylor [37] με το όνομα Decision Directed Acyclic Graph (DDAG) και των Kijssirikul και Ussivakul [40] με το όνομα Adaptive Directed Acyclic Graph (ADAG). Το πρόβλημα με τις μεθόδους DDAG και ADAG είναι ότι οι περιοχές γενίκευσης εξαρτώνται από τη δομή του δέντρου απόφασης.

Για την ελαχιστοποίηση του συγκεκριμένου προβλήματος, παρακάτω παρουσιάζεται ένας αλγόριθμος βελτιστοποίησης της δομής του δέντρου απόφασης, με βάση τη μέθοδο DDAG.

4.3.3.2 Βελτιστοποιημένη μέθοδος DAG SVM

Στη συγκεκριμένη μέθοδο ταξινόμησης, παρουσιάζεται ένας αλγόριθμος βελτιστοποίησης της δομής του δέντρου απόφασης του DDAG, ο οποίος προτάθηκε αρχικά από την ομάδα των Takahashi και Abe [39]. Η διαδικασία εκπαίδευσης των SVM είναι αντίστοιχη με αυτή που παρουσιάστηκε στη μέθοδο one-against-one. Δεδομένου ενός προβλήματος n κλάσεων, δημιουργούνται όλα τα δυνατά ζεύγη κλάσεων, καταλήγοντας έτσι σε $n(n-1)/2$ συνδυασμούς. Για κάθε τέτοιο ζεύγος, εκπαιδεύεται ένας ταξινομητής SVM με βάση τα δεδομένα εκπαίδευσης που ανήκουν στις δύο αυτές κλάσεις.

Προκειμένου να γίνει η εκπαίδευση των $n(n-1)/2$ ταξινομητών, έγιναν δοκιμές με διάφορους πυρήνες, μερικοί από τους οποίους ήταν ο quadratic, ο polynomial και ο rbf ή Gaussian πυρήνας. Τελικά, λόγω καλύτερης απόδοσης επιλέχθηκε να χρησιμοποιηθεί ο γραμμικός πυρήνας (linear kernel) ή αλλιώς «dot product» με αντίστοιχη συνάρτηση πυρήνα:

$$H(x, x') = x^t \cdot x'$$

Εδώ πρέπει να συμπεριληφθεί ότι, μετά την ολοκλήρωση εκπαίδευσης του κάθε ταξινομητή, υπολογίζεται επίσης η ικανότητα γενίκευσης του συγκεκριμένου SVM. Υπάρχουν πολλές μετρικές που έχουν αναπτυχθεί για την μέτρηση σφάλματος γενίκευσης. Στην παρούσα εργασία, χρησιμοποιείται η ακόλουθη εκτίμηση του σφάλματος γενίκευσης E_{ij} των κλάσεων i και j :

$$E_{ij} = \frac{SV_{ij}}{M_{ij}}$$

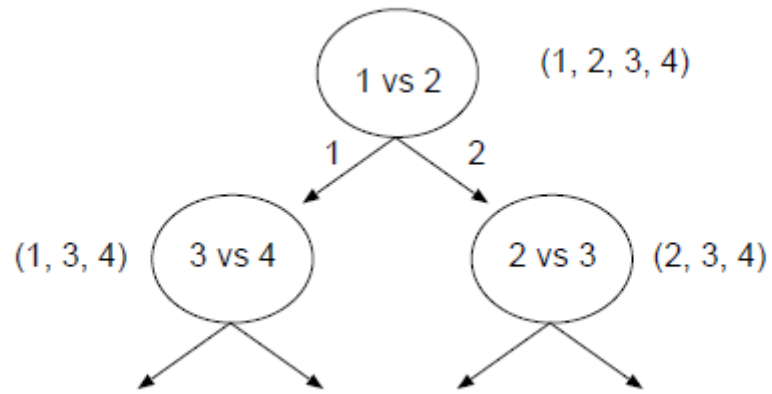
όπου, SV_{ij} είναι ο αριθμός των διανυσμάτων υποστήριξης για τις κλάσεις i και j και M_{ij} είναι ο αριθμός των δεδομένων εκπαίδευσης των κλάσεων i και j .

Η συγκεκριμένη μετρική εκτιμάται προκειμένου να χρησιμοποιηθεί σαν κριτήριο για την διαμόρφωση της δομής του δέντρου απόφασης. Η λογική στην οποία βασίζεται η συγκεκριμένη απόφαση είναι ότι, στη μέθοδο DDAG οι αταξινόμητες περιοχές αποδίδονται στις τάξεις που σχετίζονται με τους κόμβους – φύλλα του αντίστοιχου δέντρου απόφασης. Επομένως, αν επιλεχτούν να τοποθετηθούν στους ανώτερους κόμβους τα ζεύγη κλάσεων που ταξινομούνται εύκολα, οι αταξινόμητες περιοχές ανατίθενται στις κλάσεις που διαχωρίζονται δύσκολα. Αυτό σημαίνει ότι, τα ζεύγη κλάσεων που είναι δύσκολα διαχωρίσιμα ταξινομούνται τελικά από τα όρια απόφασης που έχουν καθοριστεί από αυτά τα ζεύγη.

Έτσι τελικά, ο αλγόριθμος που καθορίζει τη βέλτιστη δομή του δέντρου απόφασης ενός προβλήματος n κλάσεων είναι ο ακόλουθος:

1. Φτιάξε την αρχική λίστα: $\{1, \dots, n\}$
2. Επέλεξε το ζεύγος κλάσεων $\{i, j\}$ από τη λίστα με τη μικρότερη τιμή σφάλματος γενίκευσης E_{ij} . Αν η είσοδος x βρίσκεται στην πλευρά της κλάσης i της συνάρτησης απόφασης, διέγραψε το j από τη λίστα. Αλλιώς, διέγραψε το i .
3. Αν η λίστα του βήματος 2 έχει περισσότερα από ένα στοιχεία, τότε πήγαινε στο Βήμα 2. Αλλιώς, ταξινόμησε το x στην κλάση που αντιστοιχεί στο εναπομένον στοιχείο της λίστας και τερμάτισε τον αλγόριθμο.

Παρακάτω, στην εικόνα 4.5, απεικονίζεται η δομή της διαδικασίας που μόλις περιγράφηκε για ένα πρόβλημα τεσσάρων κλάσεων.



Εικόνα 4.6: Απόφαση της μεθόδου DDAG για ένα πρόβλημα με 4 κλάσεις.

5

Αξιολόγηση Συστήματος

5.1 Σύνολο δεδομένων

Προκειμένου να πραγματοποιηθεί η εκπαίδευση των SVMs και ο υπολογισμός της απόδοσης του συστήματος που αναπτύχθηκε, χρησιμοποιήθηκε ως είσοδος του συστήματος ένα σύνολο δεδομένων, αποτελούμενο από 1200 εικόνες τροφίμων σε μορφή jpg. Το σύνολο αυτό αποτελείται από 6 διαφορετικές κατηγορίες εικόνων, 200 εικόνες ανά κατηγορία. Οι κατηγορίες αυτές ή κλάσεις του συστήματος είναι οι ακόλουθες: Ψωμί, Κρέας, Ζυμαρικά, Πατάτες, Ρύζι και Πράσινα Λαχανικά.

Βέβαια, προτού χρησιμοποιηθεί το σύνολο των εικόνων στην είσοδο του συστήματος, προηγήθηκε το στάδιο προεπεξεργασίας. Το πρώτο βήμα της διαδικασίας αυτής αποτελείται από τη χειροκίνητη απομόνωση του μέρους της εικόνας που αποτελούσε το γεύμα αυτό καθαυτό και την τοποθέτησή του σε ένα λευκό φόντο. Δείγματα των εικόνων του συνόλου των δεδομένων έχοντας ολοκληρωθεί το πρώτο στάδιο της προεπεξεργασίας απεικονίζονται στην εικόνα 5.1. Στη συνέχεια, γίνεται χρήση της συνάρτησης `background.m` για την απομόνωση του αντικειμένου από το φόντο της εικόνας, το οποίο πήρε την τιμή $\{0, 0, 0\}$ στον χρωματικό χώρο L^*a^*b , που αντιστοιχεί στο μαύρο χρώμα. Η συνάρτηση `background.m` χρησιμοποιεί τον αλγόριθμο *k-means* με την παράμετρο *k* ίση με 2. Σε αυτό το σημείο, τα δεδομένα ήταν έτοιμα να χρησιμοποιηθούν ως είσοδος του συστήματος.



Εικόνα 5.1: Δείγματα των εικόνων του συνόλου δεδομένων μετά το πρώτο βήμα της προεπεξεργασίας. Συνολικά χρησιμοποιήθηκαν 1200 εικόνες που ανήκουν σε 6 κατηγορίες.

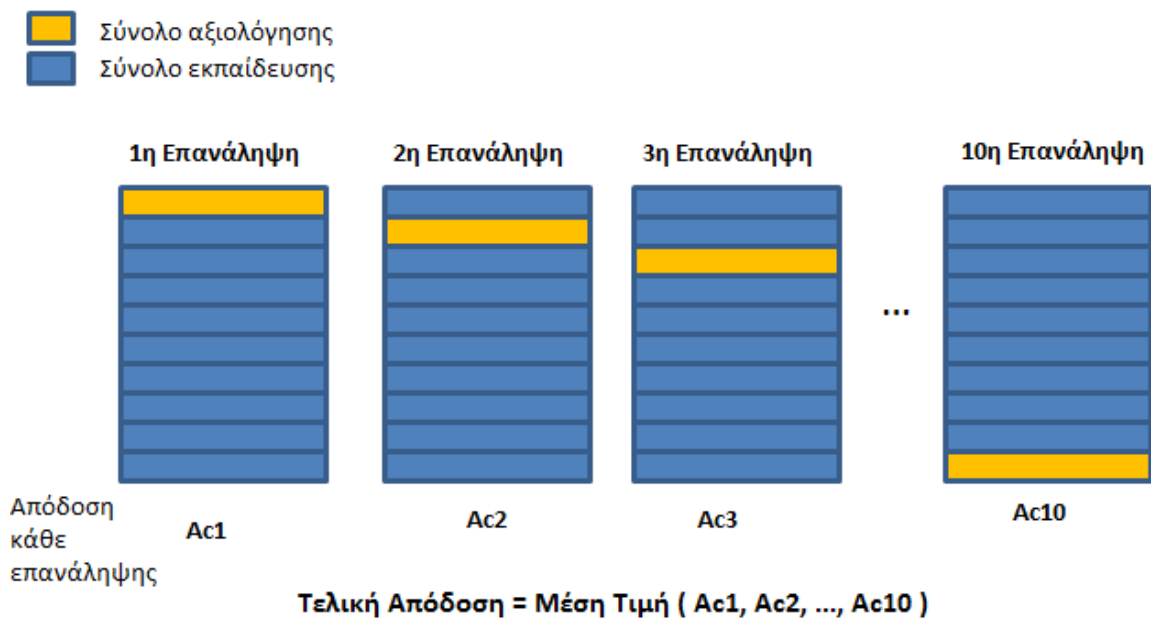
5.2 Κριτήρια αξιολόγησης

Για την αξιολόγηση του συστήματος χρησιμοποιήθηκε η μέθοδος 10-fold cross validation. Σύμφωνα με τη μέθοδο αυτή, το σύνολο των δεδομένων χωρίζεται, με τη βοήθεια μιας συνάρτησης παραγωγής τυχαίων αριθμών, σε 10 τυχαία ανεξάρτητα υποσύνολα δεδομένων. Πιο συγκεκριμένα, κάθε κατηγορία 200 εικόνων χωρίζεται τυχαία σε 10 ισάριθμες ομάδες των 20 εικόνων. Στη συνέχεια, συνδυάζοντας τις ομάδες αυτές, μια από κάθε κατηγορία, δημιουργούνται τελικά 10 ανεξάρτητα υποσύνολα (από εδώ και στο εξής folds) του συνόλου των δεδομένων, που όμως καθένα από αυτά αποτελείται από ίσο αλλά τυχαίο αριθμό εικόνων από κάθε κατηγορία.

Η διαδικασία που ακολουθεί είναι η εξής: Η μάθηση του δικτύου των SVMs, καθώς και ο έλεγχος της απόδοσης του γίνεται σε 10 επαναλήψεις, σε κάθε μία από τις οποίες χρησιμοποιούνται 9 μέρη κατά τη διαδικασία της εκπαίδευσης και 1 κατά την αξιολόγηση. Σε κάθε επανάληψη από τις 10 που πραγματοποιείται ο συγκεκριμένος κύκλος διεργασιών, επιλέγεται διαφορετικό fold για την αξιολόγηση από τις προηγούμενες φορές, με αποτέλεσμα κάθε ένα να χρησιμοποιηθεί μια μοναδική φορά στη διαδικασία της αξιολόγησης.

Τέλος, η απόδοση υπολογίζεται ως ο μέσος όρος της απόδοσης που έχει υπολογιστεί σε κάθε μια επανάληψη του παραπάνω κύκλου διεργασιών, βγάζοντας μια μέση τιμή απόδοσης του συστήματος. Η συνολική διαδικασία απεικονίζεται σχηματικά στην εικόνα 5.2.

Επισημαίνεται ότι, σε κάθε επανάληψη της διαδικασίας η απόδοση τίθεται ίση με την ακρίβεια της ταξινόμησης και υπολογίζεται ως ο λόγος των εικόνων που τοποθετήθηκαν στη σωστή κατηγορία από το σύστημα, προς το σύνολο των εικόνων που χρησιμοποιήθηκαν στη διαδικασία της αξιολόγησης, ενώ παρουσιάζεται σε ποσοστά επί τοις εκατό. Στη συγκεκριμένη εργασία, ο συνολικός αυτός αριθμός ήταν 120 σε κάθε επανάληψη.



Εικόνα 5.2: Σχηματική αναπαράσταση της μεθόδου 10-fold cross validation.

5.3 Αποτελέσματα

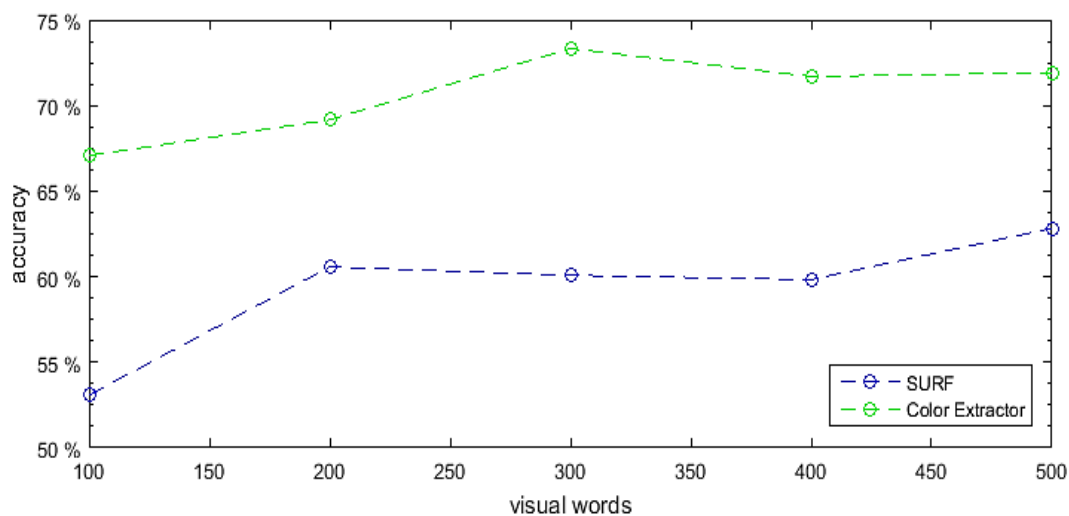
Αρχικά, γίνεται ο υπολογισμός του συνόλου των διανυσμάτων των χαρακτηριστικών με βάση τους τρεις αλγόριθμους που παρουσιάστηκαν προηγουμένως, SURF, Color Extractor και LBP Features, ενώ στη συνέχεια υπολογίζεται η απόδοση του συστήματος χρησιμοποιώντας κάθε φορά στην είσοδο ένα από τα σύνολα αυτά. Τα αποτελέσματα που λήφθηκαν με τη διαδικασία αυτή απεικονίζονται στις εικόνες 5.3 και 5.4.

Συγκεκριμένα, στην εικόνα 5.3 πραγματοποιείται σύγκριση των αποτελεσμάτων με είσοδο τα χαρακτηριστικά των αλγορίθμων SURF και Color Extractor στον ταξινομητή DAG

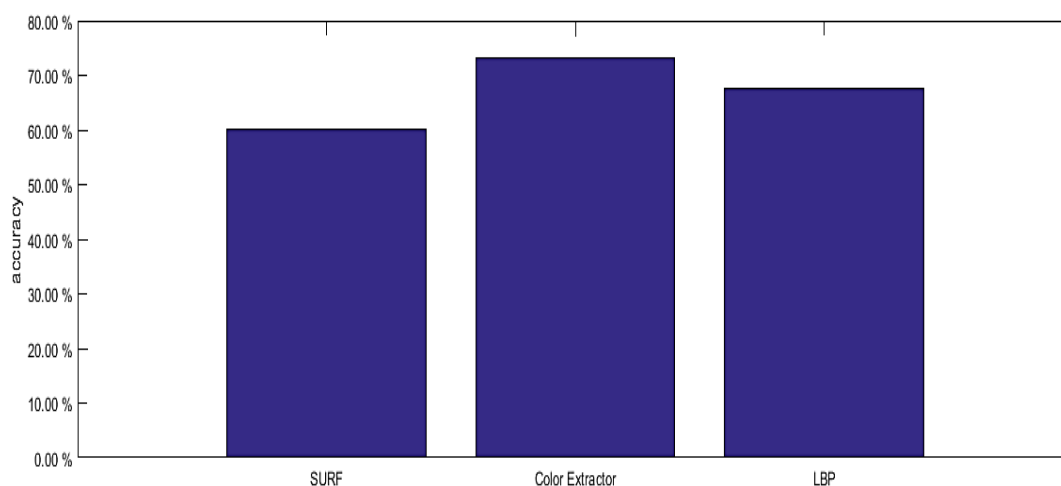
SVM, ενώ ταυτόχρονα μεταβάλλεται ο αριθμός των visual words. Οι ακριβείς τιμές απόδοσης του συστήματος φαίνονται στον πίνακα 5.1.

Στην εικόνα 5.4 απεικονίζεται παράλληλα και το αποτέλεσμα που λαμβάνεται με είσοδο τα 59 χαρακτηριστικά που εξάγονται με τη μέθοδο LBP, ενώ έχει επιλεγθεί ως μέγεθος του visual dictionary το 300 για τους άλλους δύο αλγορίθμους.

Επισημαίνεται ότι, σε αυτόν αλλά και σε όλους τους υπολογισμούς που παρουσιάζονται στη συγκεκριμένη εργασία, τα δυαδικά SVMs εκπαιδεύονται με γραμμική συνάρτηση πυρήνα, εκτός αν αναφέρεται κάτι διαφορετικό.



Εικόνα 5.3: Σύγκριση αλγορίθμων εξαγωγής χαρακτηριστικών SURF και Color Extractor, με βάση τη συνολική απόδοση του συστήματος, μεταβάλλοντας τον αριθμό των visual words.



Εικόνα 5.4: Σύγκριση της μέγιστης απόδοσης των αλγορίθμων SURF, Color Extractor και LBP.

Πίνακας 5.1: Σύγκριση αλγορίθμων SURF, Color Extractor και LBP.

Visual Words	Απόδοση Συστήματος		
	SURF	Color Extractor	LBP
100	53.08	67.08	
200	60.58	69.17	
300	60.08	73.33	59 features
400	59.83	71.17	65.78
500	62.53	71.92	

Στη συνέχεια, προκειμένου να επιτευχθεί βελτίωση της απόδοσης του συστήματος, χρησιμοποιείται ως είσοδος ένας συνδυασμός χαρακτηριστικών που εξάγονται από τους τρεις αλγορίθμους που έχουν επιλεγεί.

Για την επιλογή του κατάλληλου αριθμού των visual words που θα χρησιμοποιηθούν στη μέθοδο Bag of Features, με χρήση των αλγορίθμων SURF και Color Extractor, γίνεται μια σύγκριση διάφορων συνδυασμών χαρακτηριστικών μεταβάλλοντας τον αριθμό αυτό και αξιολογώντας το αποτέλεσμα.

Κατά τη διαδικασία αυτή, επιλέγεται κάθε φορά διαφορετικό μέγεθος του visual dictionary και χρησιμοποιείται τόσο με τον αλγόριθμο SURF όσο και με τον Color Extractor. Εν συνεχεία, στα χαρακτηριστικά αυτά, προστίθενται και το διάλυμα των 59 χαρακτηριστικών που εξάγονται μέσω της μεθόδου LBP.

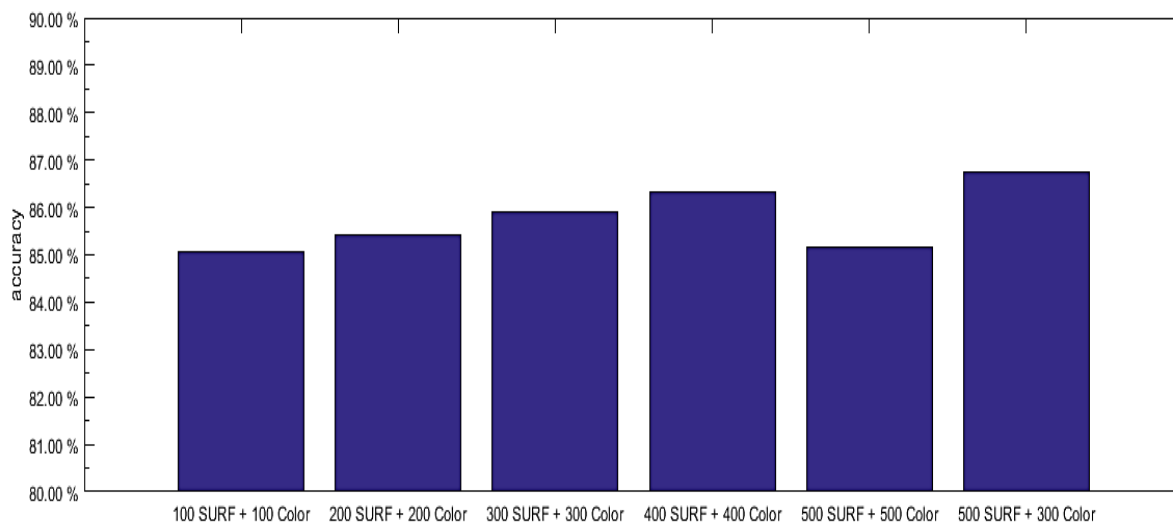
Το visual dictionary παίρνει τιμές, όπως και στα προηγούμενα αποτελέσματα, από 100 μέχρι 500 visual words, με βήμα 100. Επομένως, κάθε φορά το διάλυμα των χαρακτηριστικών που εξάγονται με τους αλγορίθμους SURF και Color Extractor έχει το ίδιο μέγεθος.

Τέλος, προκειμένου να επιτευχθεί η επιλογή του βέλτιστου συνδυασμού, γίνεται δοκιμή και με τις τιμές των visual words που στην προηγούμενη σύγκριση των αλγορίθμων SURF και Color Extractor έδωσαν τις μέγιστες τιμές απόδοσης. Αυτές αντιστοιχούν σε 500 visual words για τον αλγόριθμο SURF και σε 300 visual words για τον αλγόριθμο Color Extractor.

Τα αποτελέσματα της σύγκρισης που μόλις αναφέρθηκαν, απεικονίζονται παρακάτω στην εικόνα 5.5, ενώ οι ακριβείς τιμές της απόδοσης του συστήματος στην κάθε περίπτωση παρουσιάζονται στον πίνακα 5.2.

Επομένως, αξιολογώντας τα συγκεκριμένα αποτελέσματα, παρατηρείται ότι η βέλτιστη απόδοση επιτυγχάνεται με το σύνολο χαρακτηριστικών που προέρχεται και από τους τρεις αλγόριθμους εξαγωγής χαρακτηριστικών. Πιο συγκεκριμένα, με 500 visual words με χρήση του αλγορίθμου SURF, 300 visual words για τον αλγόριθμο Color Extractor και

φυσικά 59 επιπλέον χαρακτηριστικά LBP. Παρατηρείται ότι, η μέγιστη απόδοση που επιτεύχθηκε με χαρακτηριστικά μεμονωμένα από έναν αλγόριθμο εξαγωγής ήταν το 71.92% του αλγορίθμου Color Extractor, ενώ με τον συνδυασμό χαρακτηριστικών και από τους τρεις αλγόριθμους η απόδοση αυτή αυξάνεται στο 86,75%. Επομένως, πρόκειται για μια σημαντική αύξηση της τάξης του 15%.



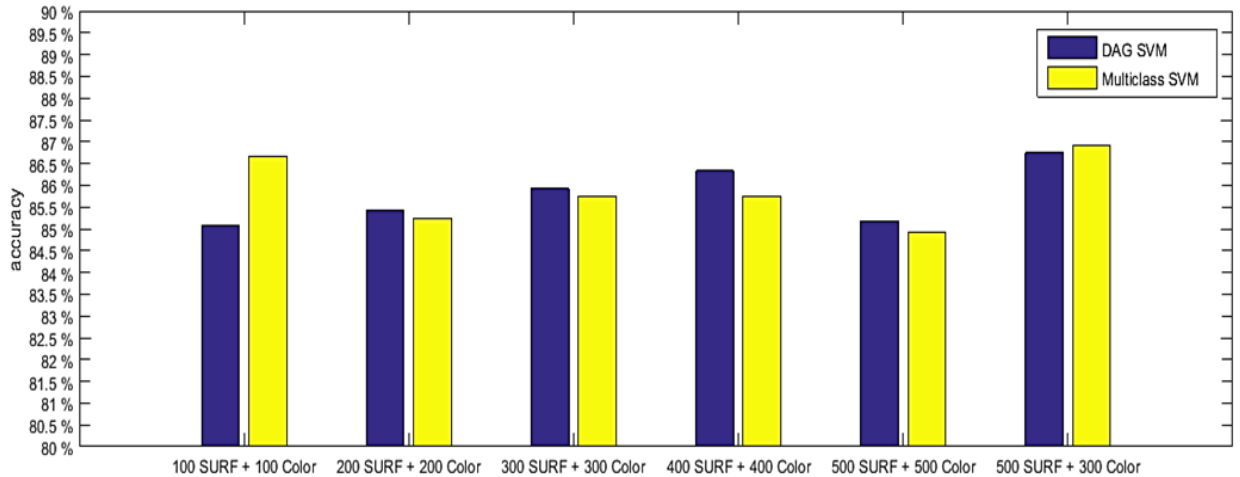
Εικόνα 5.5: Σύγκριση των αποτελεσμάτων με χρήση συνδυασμού χαρακτηριστικών SURF, Color και LBP, με ταυτόχρονη μεταβολή του visual dictionary για τους πρώτους δύο.

Πίνακας 5.2: Απόδοση συστήματος με χρήση συνδυασμού χαρακτηριστικών SURF, Color και LBP.

Visual Words	Απόδοση Συστήματος
100 SURF + 100 Color Extractor + 59 LBP	85.08
200 SURF + 200 Color Extractor + 59 LBP	85.42
300 SURF + 300 Color Extractor + 59 LBP	85.92
400 SURF + 400 Color Extractor + 59 LBP	86.66
500 SURF + 500 Color Extractor + 59 LBP	85.17
500 SURF + 300 Color Extractor + 59 LBP	86.75

Σε αυτό το σημείο, απαιτείται να παρουσιαστεί και μια σύγκριση της απόδοσης του συστήματος που αναπτύξαμε, που αποτελεί μια βελτιστοποιημένη δομή της μεθόδου DAG SVM, σε σχέση με τα αποτελέσματα που θα έδινε ένα απλό multiclass SVM από τις βιβλιοθήκες του Matlab.

Χρησιμοποιώντας κοινή είσοδο στα δύο συστήματα, υπολογίζεται η απόδοση και με τις δύο δομές ταξινομητών και τα αποτελέσματα που λαμβάνονται απεικονίζονται στην εικόνα 5.6.



Εικόνα 5.6: Σύγκριση αποτελεσμάτων ταξινομητών DAG SVM και Multiclass SVM, με χρήση συνδυασμού χαρακτηριστικών SURF, Color και LBP, και ταυτόχρονη μεταβολή του visual dictionary για τους πρώτους δύο.

Παρατηρείται ότι, με τη βελτιστοποιημένη μέθοδο DAG SVM που υλοποιήσαμε, η μέγιστη απόδοση που επιτυγχάνεται είναι 86.75 %, ενώ με ένα multiclass SVM είναι 86.92%. Ακριβέστερες μετρήσεις εμφανίζονται παρακάτω στον πίνακα 5.3. Επομένως, από πλευράς απόδοσης, τα δύο συστήματα ταξινομητών δίνουν αποτελέσματα ίδιας τάξης μεγέθους.

Πίνακας 5.3: Σύγκριση απόδοσης των αποτελεσμάτων των ταξινομητών DAG SVM και Multiclass SVM.

Visual Words	Απόδοση Συστήματος	
	DAG SVM	Multiclass SVM
100 SURF + 100 Color Extractor + 59 LBP	85.08	86,67
200 SURF + 200 Color Extractor + 59 LBP	85.42	85,25
300 SURF + 300 Color Extractor + 59 LBP	85.92	85,75
400 SURF + 400 Color Extractor + 59 LBP	86.66	86,25
500 SURF + 500 Color Extractor + 59 LBP	85.17	84,92
500 SURF + 300 Color Extractor + 59 LBP	86.75	86,92

Προκειμένου να παρατηρηθούν οι κλάσεις οι οποίες έχουν το μεγαλύτερο ποσοστό αποτυχίας κατά τη διαδικασία της ταξινόμησης, στους παρακάτω πίνακες 5.3 έως 5.12 εμφανίζονται οι Μήτρες Σύγχυσης (Confusion Matrices) που εξάχθηκαν σε κάθε μία από τις 10 επαναλήψεις της διαδικασίας 10-fold cross validation. Τα συγκεκριμένα αποτελέσματα

αφορούν το σύνολο χαρακτηριστικών που αποτελείται από 500 SURF, 300 Color και 59 LBP χαρακτηριστικά. Τα αποτελέσματα εμφανίζονται σε ποσοστά επί τοις εκατό.

Πίνακας 5.3: Μήτρα Σύγκρισης 1ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	70	5	5	5	15	0
Κρέας	5	90	0	5	0	0
Ζυμαρικά	0	10	75	5	10	0
Πατάτες	0	0	20	80	0	0
Ρύζι	5	0	0	0	95	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.4: Μήτρα Σύγκρισης 2ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	90	0	5	0	5	0
Κρέας	10	70	5	10	0	5
Ζυμαρικά	5	0	80	10	5	0
Πατάτες	0	5	5	85	5	0
Ρύζι	0	0	0	5	95	0
Λαχανικά	0	0	5	0	0	95

Πίνακας 5.5: Μήτρα Σύγκρισης 3ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	90	10	0	0	0	0
Κρέας	10	90	0	0	0	0
Ζυμαρικά	0	5	80	5	10	0
Πατάτες	0	0	0	100	0	0
Ρύζι	10	0	0	0	90	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.6: Μήτρα Σύγκρισης 4ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	80	20	0	0	0	0
Κρέας	10	85	0	5	0	0
Ζυμαρικά	0	0	95	5	0	0
Πατάτες	0	0	10	85	5	0
Ρύζι	0	0	10	10	80	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.7: Μήτρα Σύγκρισης 5ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	75	5	20	0	0	0
Κρέας	10	85	5	0	0	0
Ζυμαρικά	5	15	75	0	5	0
Πατάτες	0	5	0	95	0	0
Ρύζι	5	0	10	5	80	0
Λαχανικά	0	0	5	0	0	95

Πίνακας 5.8: Μήτρα Σύγκρισης 6ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	65	10	5	5	15	0
Κρέας	5	85	10	0	0	0
Ζυμαρικά	0	0	95	5	0	0
Πατάτες	0	0	20	75	5	0
Ρύζι	15	5	5	5	70	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.9: Μήτρα Σύγκρισης 7ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	85	5	0	0	10	0
Κρέας	0	85	0	15	0	0
Ζυμαρικά	0	0	80	15	5	0
Πατάτες	0	5	20	70	5	0
Ρύζι	5	0	0	5	90	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.10: Μήτρα Σύγκρισης 8ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	85	0	5	5	5	0
Κρέας	5	90	0	0	5	0
Ζυμαρικά	5	0	55	30	10	0
Πατάτες	5	0	0	95	0	0
Ρύζι	5	0	5	0	90	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.11: Μήτρα Σύγχυσης 9ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	80	5	10	0	5	0
Κρέας	10	85	0	5	0	0
Ζυμαρικά	5	0	80	5	0	10
Πατάτες	0	10	20	70	0	0
Ρύζι	0	0	0	0	100	0
Λαχανικά	0	0	0	0	0	100

Πίνακας 5.12: Μήτρα Σύγχυσης 10ης Επανάληψης

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	100	0	0	0	0	0
Κρέας	5	95	0	0	0	0
Ζυμαρικά	0	0	85	10	5	0
Πατάτες	0	0	5	90	5	0
Ρύζι	5	0	10	5	75	5
Λαχανικά	0	0	0	0	0	100

5.4 Αξιολόγηση αποτελεσμάτων

Στον πίνακα 5.13 που φαίνεται στη συνέχεια, απεικονίζεται η Μήτρα Σύγχυσης με στοιχεία τη μέση τιμή των 10 επαναλήψεων για κάθε αντίστοιχο πεδίο αυτού.

Παρατηρείται ότι, στο βελτιστοποιημένο σύστημα DAG SVM που αναπτύχθηκε, η επιτυχία στην ταξινόμηση της κλάσης των λαχανικών φτάνει στο 99%, ποσοστό που δείχνει την υψηλή ακρίβεια στην ταξινόμηση των τροφίμων της κλάσης αυτής.

Πίνακας 5.13: Μήτρα Σύγχυσης με τη μέση τιμή κάθε πεδίου στις 10 επαναλήψεις.

	Ψωμί	Κρέας	Ζυμαρικά	Πατάτες	Ρύζι	Λαχανικά
Ψωμί	82	6	5	1.5	5.5	0
Κρέας	7	86	2	4	0.5	0.5
Ζυμαρικά	2	3	80	9	5	1
Πατάτες	0.5	2.5	10	84.5	2.5	0
Ρύζι	5	0.5	4	3.5	86.5	0.5
Λαχανικά	0	0	1	0	0	99

Υψηλά ποσοστά, της τάξης του 85%, διακρίνονται στις κλάσεις Κρέας, Πατάτες και Ρύζι, ποσοστό όμως που για έναν διαβητικό ασθενή μπορεί να κριθεί ακατάλληλο για την επίτευξη ενός σωστού γλυκαιμικού ελέγχου.

Σε αυτό το σημείο όμως πρέπει κανείς να λάβει υπόψη του ότι, η αποτυχημένη ταξινόμηση μιας τροφή που ανήκει στην ομάδα των κλάσεων Πατάτες, Ρύζι και Ζυμαρικά, τις περισσότερες φορές έγκειται στην ταξινόμησή της σε μια από τις άλλες δύο κλάσεις αυτής της ομάδας.

Πιο συγκεκριμένα:

- Όσον αφορά τα ζυμαρικά, ένα ποσοστό 9% κατά μέση τιμή ταξινομείται στις πατάτες και ένα 5% ταξινομείται στο ρύζι. Στις υπόλοιπες κατηγορίες ταξινομείται συνολικά περίπου το 6% των τροφίμων της κλάσης.
- Όσον αφορά τις πατάτες, ένα ποσοστό 10% ταξινομείται στα ζυμαρικά και ένα 2,5% στο ρύζι, ενώ στις υπόλοιπες κλάσεις το ποσοστό αντιστοιχεί στο 3%.
- Για το ρύζι, 3,5% των τροφίμων ταξινομείται στις πατάτες και 4% στα ζυμαρικά. Εδώ έχουμε και ένα σημαντικό ποσοστό της τάξης του 5% που ταξινομείται στο ψωμί, ενώ στις υπόλοιπες κλάσεις το ποσοστό αντιστοιχεί σε 1%.

Επίσης, από τα στοιχεία του Οδηγού Διατροφής για τη ρύθμιση του Διαβήτη της Ελληνικής Διαβητολογικής Εταιρείας, βλέπουμε ότι 1/3 του φλιτζανιού ζυμαρικά και 1/3 του φλιτζανιού ρύζι (καστανό ή άσπρο) έχουν περίπου 15 γραμμάρια υδατάνθρακες, όπου την ίδια ποσότητα υδατανθράκων έχει και 1/2 του φλιτζανιού πατάτα και 1 φέτα ψωμί. Επομένως, οι τρεις αυτές κατηγορίες εκτιμάται ότι είναι πολύ κοντά σε ποσοστά υδατανθράκων ανά μερίδα γεύματος. Έχοντας αυτό σαν βάση, μπορεί να καταλήξει κανείς ότι το σφάλμα στον υπολογισμό των υδατανθράκων, αν επιλεγεί στην ταξινόμηση λανθασμένα μια άλλη κλάση αυτής της ομάδας, θα είναι σχετικά χαμηλό.

Τέλος, θα πρέπει να αναφερθεί ότι σημαντική απόκλιση στην εκτίμηση των υδατανθράκων θα υπάρξει αν τρόφιμα της ομάδας του κρέατος ταξινομηθούν σε μια από τις άλλες κλάσεις, πράγμα που σύμφωνα με τα αποτελέσματα του πίνακα 4.13 συμβαίνει σε ποσοστό 14%. Επομένως, στην ταξινόμηση του κρέατος κρίνεται απαραίτητη η βελτίωση του συγκεκριμένου συστήματος.

6

Μελλοντική Έρευνα

Τα αποτελέσματα από την αξιολόγηση του μοντέλου ήταν ιδιαίτερα ενθαρρυντικά. Ωστόσο, προκειμένου να χρησιμοποιηθεί το μοντέλο σε πραγματικά συστήματα αυτόματης αξιολόγησης τροφίμων για διαβητικούς ασθενείς, θα πρέπει να η απόδοσή του να βελτιωθεί. Μερικά σημεία τα οποία, σύμφωνα με την αξιολόγηση του συστήματος, φαίνεται να έχουν περιθώρια βελτίωσης είναι τα εξής:

- Καταρχήν, θα πρέπει να εμπλουτιστεί η βάση δεδομένων με περισσότερες εικόνες, και ειδικότερα με εικόνες που έχουν ληφθεί σε πραγματικές συνθήκες. Έτσι θα δημιουργηθεί ένα πιο αντιπροσωπευτικό σύνολο δεδομένων, πράγμα που σίγουρα θα συμβάλει στη βελτίωση της απόδοσης του συστήματος.
- Επιπρόσθετα, χρήσιμο ή ακόμα και απαραίτητο είναι να προστεθούν στο σύνολο των δεδομένων κλάσεις τροφίμων που στη παρούσα εργασία δεν έχουν συμπεριληφθεί καθόλου, όπως φρούτα και όσπρια, που αποτελούν σημαντικό μέρος της καθημερινής διαίτας.
- Σχετικά με τη δομή του ταξινομητή DAG SVM, ενδιαφέρον παρουσιάζει η περαιτέρω μελέτη της δομής του δέντρου που υλοποιείται, εξετάζοντας και άλλες

μετρικές επιλογής ταξινομητών πέρα από το σφάλμα γενίκευσης που επιλέχθηκε στην παρούσα εργασία.

- Όσον αφορά τον ταξινομητή DAG SVM, μελλοντικά μπορούν να υλοποιηθούν παράλληλα περισσότερες δομές του δέντρου απόφασης, βασισμένες σε διαφορετικές μετρικές, ενώ στη συνέχεια το τελικό αποτέλεσμα να είναι η υπερισχύουσα απόφαση μεταξύ των δομών αυτών.
- Επίσης, σχετικά με τους αλγόριθμους εξαγωγής χαρακτηριστικών που χρησιμοποιήθηκαν, αποδείχθηκε ότι καθένας τους προσθέτει χρήσιμη πληροφορία για το στάδιο της ταξινόμησης, δεδομένου ότι η απόδοση του συστήματος βελτιώνεται με χρήση του συνδυασμό αυτών σε σχέση με τη χρήση καθενός ξεχωριστά. Ωστόσο, οι συγκεκριμένοι αλγόριθμοι προσδίδουν πληροφορία σχετικά με το χρώμα, την υφή και τα σημεία ενδιαφέροντος της εικόνας. Επομένως, με την προσθήκη περισσότερων κλάσεων, θα είναι απαραίτητη η προσθήκη αλγορίθμων που εξάγουν πληροφορία για το σχήμα των αντικειμένων, ειδικά στην περίπτωση των φρούτων.
- Μελλοντικά, απαραίτητη κρίνεται και η αξιολόγηση του συστήματος που αναπτύχθηκε ως μέρος ενός ολοκληρωμένου συστήματος εκτίμησης των υδατανθράκων των λαμβανόμενων γευμάτων, το οποίο να περιλαμβάνει τόσο το στάδιο της κατάτμησης της εικόνας, όσο και το στάδιο του υπολογισμού της ποσότητας του γεύματος και της εκτίμησης της περιεκτικότητας των υδατανθράκων. Υπό αυτές τις συνθήκες, η αξιολόγηση του συστήματος θα είναι πιο κοντά στα πραγματικά δεδομένα και επομένως θα έχει μεγαλύτερη ακρίβεια.

7

Βιβλιογραφία

- [1] Ευ. Ντζιού, Χρ. Μ. Καστορίνη, Δ. Β. Παναγιωτάκος, Ανασκόπηση Διατροφικά πρότυπα και σακχαρώδης διαβήτης, Ελληνική Επιθεώρηση Διαιτολογίας-Διατροφής 2010, 1(2), 78–86
- [2] Β. Καραμάνος, Η. Μυγδάλης, Α. Σωτηρόπουλος, Π. Χριστακόπουλος, Κατευθυντήριες Οδηγίες για τη Διαχείριση του Διαβητικού Ασθενούς, Ελληνική Διαβητολογική Εταιρία, 2013
- [3] Κωνσταντία Ζαρκογιάννη, Ευφυή Συστήματα Υποστήριξης Εξατομικευμένων Ιατρικών Αποφάσεων για τη Διαχείριση του Σακχαρώδους Διαβήτη, Διδακτορική Διατριβή, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, ΕΜΠ, 2011
- [4] Ελένη Λίτσα, Ανάπτυξη εξατομικευμένου μοντέλου προσομοίωσης της κινητικής της γλυκόζης σε άτομα με Σακχαρώδη Διαβήτη τύπου I βασισμένο σε Αυτό-οργανούμενους χάρτες, Διπλωματική Εργασία, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, ΕΜΠ, 2014
- [5] Κλεοπάτρα Γ. Παπαβασιλείου, Διερεύνηση της επίδρασης των λιπιδίων των τροφίμων στις ανάγκες ταχείας ινσουλίνης για την κάλυψη του γεύματος σε άτομα με Σακχαρώδη διαβήτη τύπου 1, Μεταπτυχιακή Διατριβή, Τμήμα Επιστήμης Τροφίμων και Διατροφής του Ανθρώπου, Γεωπονικό Πανεπιστήμιο Αθηνών, 2015

- [6] L. Szablewski; Glucose Homeostasis – Mechanism and Defects; Diabetes - Damages and Treatments; 2011
- [7] American Diabetes Association Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2014-37 January
- [8] C. Patterson, G. Dahlquist, E. Gyürüs, A. Green, G. Soltész; EURODIAB Study Group. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study; 2009 Jun 13
- [9] Ν. Κατσίκη, Φ. Ηλιάδης, Α. Ζαντίδης, Τ. Διδάγγελος; Σακχαρώδης διαβήτης: Διάγνωση και ταξινόμηση; Ελληνικά Διαβητολογικά Χρονικά 23, 1: 78 – 86, 2010
- [10] T. Kawamura, The importance of carbohydrate counting in the treatment of children with diabetes, Pediatric Diabetes, vol. 8, suppl.6, pp. 57–62, 2007
- [11] M.L. Wheeler, A. Daly, A. Evert, M. Franz, P. Geil, L. A. Holzmeister, K. Kulkarni, E. Loghmani, T.A. Ross, P. Wolf; Choose Your Foods: Exchange Lists for Diabetes, Sixth Edition, 2008: Description and Guidelines for Use; Journal of the American Diabetic Association, vol. 108, no. 5, pp. 883–888, 2008.
- [12] C. E. Smart, B. R. King, P. McElduff, C. E. Collins; In children using intensive insulin therapy, a 20-g variation in carbohydrate amount significantly impacts on postprandial glycaemia; Diabetic Medicine, vol. 29, no. 7, pp. 21–24, 2012.
- [13] F.K. Bishop, D.M. Maahs, G. Spiegel, D. Owen, G.J. Klingensmith, A. Bortsov, J. Thomas, E.J. Mayer-Davis; The Carbohydrate Counting in Adolescents With Type 1 Diabetes (CCAT) Study; Diabetes Spectrum, vol. 22, no. 1, pp. 56–62, 2009.
- [14] C.E. Smart, K. Ross, J.A. Edge, B.R. King, P. McElduff, C.E. Collins; Can Children with Type 1 diabetes and their caregivers estimate the carbohydrate content of meals and snacks?; Diabetic Medicine, vol. 27, no. 3, pp. 348–353, 2010.
- [15] S.N. Mehta, L.K. Volkening, N. Quinn, L.M.B. Laffel; Impact of Carbohydrate Counting on Glycemic Control in Children With Type 1 Diabetes; Diabetes Care, vol. 32, pp. 1014–1016, 2009.
- [16] K. Zarkogianni, E. Litsa, K. Mitsis, P. Wu, C. Kaddi, C. Cheng, M. Wang, K.S. Nikita; A Review of Emerging Technologies for the Management of Diabetes Mellitus; IEEE Transactions on Biomedical Engineering; 2015 ;62(12).

- [17] L. Spanou, KV Dalakleidi, K. Zarkogianni, P. Anastasia, K.S. Nikita, V. Vasiliki, M. Alevizaki, E. Anastasiou; Ketonemia and ketonuria in gestational diabetes mellitus. *Hormones*; 2015.
- [18] K. Zarkogianni, K. Mitsis, E. Litsa, M.T. Arredondo, G. Fico, A. Fioravanti, K.S. Nikita; Comparative assessment of glucose prediction models for Patients with Type 1 Diabetes Mellitus applying sensors for glucose and physical activity monitoring; *Medical & Biological Engineering & Computing*. 2015 ;53(12).
- [19] W. Zhang, Q. Yu, B. Siddiquie, A. Divakaran, H. Sawhney; “Snap-n-Eat”: Food Recognition and Nutrition Estimation on a Smartphone; 2015.
- [20] D.G. Lowe; Distinctive image features from scale-invariant keypoints; *Computer Vision*. 2004; 60:91-110.
- [21] F. Perronnin, C. Dance; Fisher kernels on visual vocabularies for image categorization.; Paper presented at: IEEE Conference on Computer Vision and Pattern Recognition; June 2007; Minneapolis, MN.
- [22] MY Chen, YH Yang, CJ Ho, et al.; Automatic Chinese food identification and quantity estimation. Paper presented at: SIGGRAPH Asia; November 2012; Singapore.
- [23] M. Anthimopoulos, L. Gianola, L. Scarnato, P. Diem, and S. Mougiakakou; A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model; *Ieee Journal Of Biomedical And Health Informatics*, Vol. 18, No. 4, July 2014.
- [24] M. Szummer, R. W. Picard, “Indoor-outdoor image classification,” in *Proc. IEEE Int. Workshop Content-Based Access Image Video Database*, 1998, pp. 42–51.
- [25] F. Mindru, T. Tuytelaars, L. Van Gool, and T. Moons, “Moment invariants for recognition under changing viewpoint and illumination,” *Comput. Vis. Image Understanding*, vol. 94, pp. 3–27, 2004.
- [26] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1999.
- [27] M. Rossi, A. Nicolucci, P. Bartolo, D. Bruttomesso, A. Girelli, F. Ampudia, D. Kerr, A. Ceriello, L. Mayor, F. Pellegrini, D. Horwitz; “Diabetes interactive diary: A new telemedicine system enabling flexible diet and insulin therapy while improving quality of life: An open-label, international, multicenter, randomized

- study,” *Diabetes Care*, vol. 33, no. 1, pp. 109–115, 2010
- [28] O. Amft and G. Tröster, “Recognition of dietary activity events using on-body sensors,” *Artif. Intell. Med.*, vol. 42, no. 2, pp. 121–136, 2008.
- [29] G. Shroff, A. Smailagic, and D. P. Siewiorek, “Wearable context-aware food recognition for calorie monitoring,” in *Proc. 12th IEEE Int. Symp. Wearable Comput.*, 2008, pp. 119–120.
- [30] F. Zhu, M. Bosch, I. Woo, S. Y. Kim, C. J. Boushey, D. S. Ebert, and E. J. Delp, “The use of mobile devices in aiding dietary assessment and evaluation,” *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 4, pp. 756–766, Aug. 2010.
- [31] F. Kong and J. Tan, “DietCam: Automatic dietary assessment with mobile camera phones,” *Pervasive Mobile Comput.*, vol. 8, pp. 147–163, Feb. 2012.
- [32] L. Fei-Fei and P. Perona, “A bayesian hierarchical model for learning natural scene categories,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2005, vol. 2, pp. 524–531.
- [33] T. Joachims, “Text categorization with support vector machines: Learning with many relevant features,” in *Proc. 10th Eur. Conf. Mach. Learning*, 1998, pp. 137–142.
- [34] M. Puri, Z. Zhu, Q. Yu, A. Divakaran, and H. Sawhney, “Recognition and volume estimation of food intake using a mobile device,” in *Proc. Workshop Appl. Comput. Vis.*, 2009, pp. 1–8.
- [35] R. Debnath, N. Takahide, H. Takahashi; A decision based one-against-one method for multi-class support vector machine; *Pattern Anal Applic* (2004) 7: 164–175
- [36] C.W. Hsu, C.J. Lin; A comparison of methods for multiclass support vector machines; *IEEE Trans Neural Networks*; 13:415–425; 2002
- [37] J. C. Platt, N. Cristianini, J. Shawe-Taylor; Large margin DAGs for multiclass classification; *Advances in Neural Information Processing Systems 12*, pages 547–553; The MIT Press, Cambridge, MA, 2000.
- [38] U. H.-G. Kreßel; Pairwise classification and support vector machines; *Advances in Kernel Methods: Support Vector Learning*; pages 255–268; The MIT Press, Cambridge, MA, 1999.
- [39] T. Fumitake, S. Abe; Optimizing directed acyclic graph, Support vector machines; *Artificial Neural Networks in Pattern Recognition (ANNPR)*, 2003, p. 166-173.

-
- [40] M. Pontil, A. Verri; Support vector machines for 3-d object recognition; IEEE Transactions on Pattern Analysis and Machine Intelligence; 20(6):637–646, 1998.
- [41] D.G Lowe; Distinctive Image Features from Scale-Invariant Keypoints; International Journal of Computer Vision (2004) 60: 91.
- [42] H. Bay, A. Ess, T. Tuytelaars, L. Gool ; SURF: Speeded Up Robust Features; Computer Vision and Image Understanding; Volume 110 Issue 3, June 2008; Pages 346-359.
- [43] S. Lazebnik, C. Schmid, J. Ponce; A Sparse Texture Representation Using Local Affine Regions; Journal IEEE Transactions on Pattern Analysis and Machine Intelligence ; Volume 27 Issue 8, August 2005; Page 1265-1278.
- [44] N. Dalal, B. Triggs; Histograms of oriented gradients for human detection; Computer Vision and Pattern Recognition, 2005 ; CVPR 2005.
- [45] A. Bosch, A. Zisserman, X. Munoz; Representing shape with a spatial pyramid kernel; Proceedings of the 6th ACM international conference on Image and video retrieval; The Netherlands; July 2007 Pages 401-408.
- [46] T. Ojala, M. Pietikainen, T. Maenpaa; Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns; IEEE Transactions on Pattern Analysis and Machine Intelligence; Volume 24 Issue 7, July 2002; Page 971-987.
- [47] Mathworks Documentation; Image Category Classification Using Bag of Features; <http://www.mathworks.com/help/vision/examples/image-category-classification-using-bag-of-features.html>
- [48] Mathworks Documentation; Image Retrieval Using Customized Bag of Features; https://www.mathworks.com/help/vision/examples/image-retrieval-using-customized-bag-of-features.html#responsive_offcanvas