



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Ταξινόμηση έργων τέχνης ως προς το καλλιτεχνικό ρεύμα με χρήση τεχνικών
επεξεργασίας σήματος και αναγνώρισης προτύπων**

Καλλιόπη Γιαννούκου

131043

Παναγιώτα Κρούστη

131056

Επιβλέπων Καθηγητής: Χριστίνα Γεωργουλάκη, ΕΔΙΠ ΠΑΔΑ

Διπλωματική εργασία υποβληθείσα στο Τμήμα

ΑΙΓΑΛΕΩ, Οκτώβρης 2022

Πανεπιστήμιο Δυτικής Αττικής,
Τμήμα Μηχανικών Πληροφορικής και Υπολογιστών
Καλλιόπη Γιαννούκου, Παναγιώτα Κρούστη
© 2022 – Με την επιφύλαξη παντός δικαιώματος



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Η παρούσα διπλωματική εργασία παρουσιάστηκε

από τις

Καλλιόπη Γιαννούκου

131043

Παναγιώτα Κρούστη

131056

την 13/10/2022

Επιβλέπων Καθηγητής: Χριστίνα Γεωργουλάκη, ΕΔΙΠ ΠΑΔΑ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την Πέμπτη 13/10/2022.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Χριστίνα Γεωργουλάκη
ΕΔΙΠ ΠΑΔΑ

.....
Νικόλαος Βασιλάς
Καθηγητής ΠΑΔΑ

.....
Κωνσταντίνος Μαυρομμάτης
Λέκτορας ΠΑΔΑ

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Οι κάτωθι υπογεγραμμένοι Καλλιόπη Γιαννούκου του Γεωργίου, με αριθμό μητρώου 711131043 και Παναγιώτα Κρούστη του Ευστρατίου, με αριθμό μητρώου 711131056 φοιτήτριες του Πανεπιστημίου Δυτικής Αττικής της Σχολής Μηχανικών του Τμήματος Μηχανικών Πληροφορικής και Υπολογιστών, δηλώνουμε υπεύθυνα ότι:

«Είμαστε συγγραφείς αυτής της πτυχιακής/διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχαμε για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες κάναμε χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνουμε ότι αυτή η εργασία έχει συγγραφεί από εμάς αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μας, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μας ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μας».

Ο/Η Δηλών/ούσα


.....

Καλλιόπη Γιαννούκου

Ο/Η Δηλών/ούσα


.....

Παναγιώτα Κρούστη

ΠΕΡΙΛΗΨΗ**Ταξινόμηση έργων τέχνης ως προς το καλλιτεχνικό ρεύμα με χρήση
τεχνικών επεξεργασίας εικόνας και αναγνώρισης προτύπων**

Καλλιόπη Γιαννούκου, Παναγιώτα Κρούστη

Η ολοένα και αυξανόμενη ψηφιοποίηση έργων τέχνης και η επιλογή πολλών μουσείων έργων τέχνης να παρουσιάζουν τις συλλογές τους ψηφιοποιημένες στο διαδίκτυο έχουν καταστήσει την αυτοματοποιημένη ανάλυση και ταξινόμηση καλλιτεχνικών έργων πολύ χρήσιμη και σημαντική διαδικασία.

Η παρούσα διπλωματική εργασία πραγματεύεται την αυτοματοποιημένη ταξινόμηση έργων τέχνης ως προς το καλλιτεχνικό ρεύμα με χρήση τεχνικών επεξεργασίας εικόνας και αναγνώρισης προτύπων. Αρχικά, έγινε επισκόπηση των κυριότερων τάσεων στην αυτοματοποιημένη ταξινόμηση έργων ζωγραφικής. Στο πλαίσιο της εργασίας, δημιουργήσαμε βάση δεδομένων με ψηφιοποιημένα έργα τέχνης και ακολούθως εστίασαμε στην εξαγωγή κατάλληλων χαρακτηριστικών για την κωδικοποίηση του καλλιτεχνικού ρεύματος. Χρησιμοποιήσαμε ταξινομητές για τη διάκριση των πινάκων ως προς το καλλιτεχνικό ρεύμα με βάση τα επιλεγμένα χαρακτηριστικά. Τέλος, έγινε σύγκριση των χρησιμοποιούμενων ταξινομητών ως προς την ικανότητά τους να ταξινομήσουν τα συγκεκριμένα καλλιτεχνικά ρεύματα εξοπλισμένοι με τα επιλεγμένα χαρακτηριστικά.

Λέξεις κλειδιά

Τεχνητή Όραση, Επεξεργασία Εικόνας, Αναγνώριση Προτύπων, Καλλιτεχνικό ρεύμα

ABSTRACT**Artworks Style Classification based on Image Processing
and Pattern Recognition techniques**

Kalliopi Giannoukou, Panagiota Krousti

The ever-increasing digitization of artworks and the choice of many art museums to display their collections digitized online have made the automated analysis and classification of artworks a very useful and important process.

This thesis deals with the automated classification of works of art according to artistic style using image processing and pattern recognition techniques. Firstly, the main trends in automated painting classification were reviewed. As part of our research, we created a database of digitized artworks and then focused on extracting features suitable to reveal artistic style. We used classifiers to distinguish the paintings in terms of artistic style based on the selected features. Finally, a comparison was made of the used classifiers in terms of their ability to classify the specific artistic style equipped with the selected features.

Key Words

Computer Vision, Image Processing, Pattern Recognition, Art movement

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1– Αποτελέσματα με 5 καλλιτεχνικά ρεύματα και 40 Χαρακτηριστικά	106
Πίνακας 2 – Αποτελέσματα Κλασικισμού – Κυβισμού	106
Πίνακας 3 – Αποτελέσματα Κλασικισμού – Εξπρεσιονισμού.....	107
Πίνακας 4 – Αποτελέσματα Κλασικισμού – Ιμπρεσιονισμού	107
Πίνακας 5 – Αποτελέσματα Κλασικισμού – Ρεαλισμού	107
Πίνακας 6 – Αποτελέσματα Κυβισμού – Εξπρεσιονισμού.....	107
Πίνακας 7 – Αποτελέσματα Κυβισμού – Ιμπρεσιονισμού.....	108
Πίνακας 8 – Αποτελέσματα Κυβισμού – Ρεαλισμού	108
Πίνακας 9 – Αποτελέσματα Εξπρεσιονισμού – Ιμπρεσιονισμού	108
Πίνακας 10 – Αποτελέσματα Εξπρεσιονισμού – Ρεαλισμού	108
Πίνακας 11 – Αποτελέσματα Ιμπρεσιονισμού – Ρεαλισμού.....	109

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ ΚΑΙ ΕΙΚΟΝΩΝ

Εικόνα 1 - Ψηφιακή Επεξεργασία Εικόνας	5
Εικόνα 2 – Μονοχρωματική Εικόνα	7
Εικόνα 3 - Μία φυσική εικόνα και μία αντίστοιχη ψηφιακή εικόνα.	7
Εικόνα 4 – Εικόνα με 1 bit βάθος χρώματος / 2 χρώματα	8
Εικόνα 5 - Επίδραση βάθους χρώματος στην εικόνα.....	9
Εικόνα 6 - Χρωματικό Μοντέλο RGB.....	11
Εικόνα 7 - Χρωματικό μοντέλο CMYK	14
Εικόνα 8 - Ο χρωματικός Χώρος HSV	16
Εικόνα 9 – Ιστόγραμμα εικόνας	18
Εικόνα 10 - Τιμές χαρακτηριστικών με (α) μεγάλη διακλασική και (β) μικρή διακλασική απόσταση.....	21
Εικόνα 11 - Σχηματική αναπαράσταση του σκεπτικού ενός βέλτιστου υπερεπιπέδου για γραμμικά διαχωρίσιμα πρότυπα	28
Εικόνα 12 - Γεωμετρική ερμηνεία των αλγεβρικών αποστάσεων των σημείων από το βέλτιστο υπερεπίπεδο για μια περίπτωση 2 διαστάσεων.....	30
Εικόνα 13 - Υπερεπίπεδο ελαστικού περιθωρίου	31
Εικόνα 14 - The Triumph of Flora του Nicolas Poussin	39
Εικόνα 15 – The Tiger του Franz Marc	39
Εικόνα 16 – The Scream του Edvard Munch	40

Εικόνα 17 – On the beach του Eugene Boudin	41
Εικόνα 18 - Moret sur Loing, the Bridge and the Church της Jean-Baptiste Camille Corot	41
Εικόνα 19 - Το περιβάλλον εργασίας του Matlab	44
Εικόνα 20 – Πίνακας του Franz Marc (κυβισμός) και αντίστοιχο χρωματικό ιστόγραμμα.....	57
Εικόνα 21 – Πίνακας του Edgar Degas (Ιμπρεσιονισμός) και αντίστοιχο χρωματικό ιστόγραμμα.....	58
Εικόνα 22 - Classification Learner	60
Εικόνα 23 - Cross-Validation	60
Εικόνα 24 – Εκπαίδευση Ταξινομητή	61
Εικόνα 25 – Η διαδικασία του k – fold	62
Εικόνα 26 – Σχηματική αναπαράσταση των διανυσμάτων υποστήριξης.....	65
Εικόνα 27 - Σχηματική αναπαράσταση του ταξινομητή kNN.....	67
Εικόνα 28 - Ρύθμιση των παραμέτρων του KNN στο Classification Learner.....	69
Εικόνα 29 - Confussion matrix - Quadratic SVM.....	75
Εικόνα 30 - Confussion matrix- Linear SVM	76
Εικόνα 31 - Confussion matrix Cubic SVM	77
Εικόνα 32 - Confussion matrix Medium SVM	77
Εικόνα 33 - Confussion matrix Weighted KNN	78
Εικόνα 34 - Confussion matrix Coarse Gaussian SVM	79
Εικόνα 35 - Confussion matrix Gaussian Naïve Bayes	79
Εικόνα 36 - Confussion matrix Cosine KNN	80
Εικόνα 37 - Confussion matrix Medium KNN	81

Εικόνα 38 - Confussion matrix Kernel Naïve Bayes	81
Εικόνα 39 - Confussion matrix Cubic KNN	82
Εικόνα 40 - Confussion matrix Fine KNN	83
Εικόνα 41 - Confussion matrix Fine Gaussian SVM	83
Εικόνα 42 - Confussion matrix Coarse KNN	84
Εικόνα 43 - Confussion matrix- 2 Ρεύματα Linear SVM	85
Εικόνα 44 - Confussion matrix- 2 Ρεύματα Medium KNN	85
Εικόνα 45 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes	86
Εικόνα 46 - Confussion matrix- 2 Ρεύματα Linear SVM	87
Εικόνα 47 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes	87
Εικόνα 48 - Confussion matrix- 2 Ρεύματα Fine KNN	88
Εικόνα 49 - Confussion matrix- 2 Ρεύματα Medium Gaussian SVM	89
Εικόνα 50 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes.....	89
Εικόνα 51 - Confussion matrix- 2 Ρεύματα Cosine KNN	90
Εικόνα 52 - Confussion matrix- 2 Ρεύματα Quadratic SVM	91
Εικόνα 53 - Confussion matrix- 2 Ρεύματα Cosine KNN	91
Εικόνα 54 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes	92
Εικόνα 55 - Confussion matrix- 2 Ρεύματα Linear SVM	93
Εικόνα 56 - Confussion matrix- 2 Ρεύματα Cubic SVM	93
Εικόνα 57 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes	94
Εικόνα 58 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes	94
Εικόνα 59 - Confussion matrix- 2 Ρεύματα Cosine KNN	95

Εικόνα 60 - Confussion matrix- 2 Ρεύματα Linear SVM	96
Εικόνα 61 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes	96
Εικόνα 62 - Confussion matrix- 2 Ρεύματα Fine KNN	97
Εικόνα 63 - Confussion matrix- 2 Ρεύματα Linear SVM	98
Εικόνα 64 - Confussion matrix- 2 Ρεύματα Weighted KNN	98
Εικόνα 65 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes	99
Εικόνα 66 - Confussion matrix- 2 Ρεύματα Linear SVM	100
Εικόνα 67 - Confussion matrix- 2 Ρεύματα Cubic SVM	100
Εικόνα 68 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes	101
Εικόνα 69 - Confussion matrix- 2 Ρεύματα Weighted KNN.....	101
Εικόνα 70 - Confussion matrix- 2 Ρεύματα Linear SVM	102
Εικόνα 71 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes	102
Εικόνα 72 - Confussion matrix- 2 Ρεύματα Medium KNN	103
Εικόνα 73 - Confussion matrix- 2 Ρεύματα Quadratic SVM.....	104
Εικόνα 74 - Confussion matrix- 2 Ρεύματα Weighted KNN.....	104
Εικόνα 75 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes	105

ΠΕΡΙΕΧΟΜΕΝΑ

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ.....	IV
ΠΕΡΙΛΗΨΗ	V
ABSTRACT.....	VI
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ.....	VII
ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ ΚΑΙ ΕΙΚΟΝΩΝ.....	VIII
ΠΕΡΙΕΧΟΜΕΝΑ	XII
ΕΥΧΑΡΙΣΤΙΕΣ	1
ΚΕΦΑΛΑΙΟ 1 - ΕΙΣΑΓΩΓΗ.....	2
1.1 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ.....	2
1.2 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ.....	2
1.3 ΣΥΝΕΙΣΦΟΡΑ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ.....	3
ΚΕΦΑΛΑΙΟ 2 – ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	5
2.1 ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ	5
2.1.1 Ο κλάδος της Ψηφιακής Επεξεργασίας Εικόνας	5
2.1.2 Ψηφιακή Εικόνα	6
2.1.3 Βάθος Χρώματος.....	8
2.1.4 Χρωματικά Μοντέλα	9
2.1.5 Ακμές.....	17
2.1.6 Ιστόγραμμα εικόνας.....	18
2.1.7 Υφή	19

2.2	ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ	19
2.2.1	ΟΡΙΣΜΟΣ.....	19
2.2.2	ΒΑΣΙΚΑ ΣΤΑΔΙΑ ΣΧΕΔΙΑΣΜΟΥ ΣΥΣΤΗΜΑΤΟΣ ΤΑΞΙΝΟΜΗΣΗΣ	20
2.2.2.1	Στάδιο δημιουργίας και επιλογής χαρακτηριστικών	20
2.2.2.2	Στάδιο σχεδίασης του ταξινομητή.....	21
2.2.3	ΑΛΓΟΡΙΘΜΟΙ ΤΑΞΙΝΟΜΗΣΗΣ	22
2.2.3.1	Ταξινομητής Bayes	22
2.2.3.3	Ταξινομητές πλησιέστερης γειτνίασης	24
2.2.3.4	Μηχανές Διανυσμάτων Υποστήριξης (SVM).....	25
ΚΕΦΑΛΑΙΟ 3 - ΥΛΟΠΟΙΗΣΗ	34
3.1	ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ – STATE OF THE ART	34
3.2	ΜΕΘΟΔΟΛΟΓΙΑ	36
3.3	ΚΑΛΛΙΤΕΧΝΙΚΑ ΡΕΥΜΑΤΑ	38
3.4	ΛΟΓΙΣΜΙΚΟ ΠΟΥ ΕΠΙΛΕΧΘΗΚΕ - MATLAB	42
3.4.1	Τι είναι το Matlab	42
3.4.2	Περιβάλλον Εργασίας	43
3.5	ΠΕΡΙΓΡΑΦΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ	44
3.5.1	Χαρακτηριστικά προτεινόμενα από τους Icoглу, Gonsel και Sariel	45
3.5.2	Χαρακτηριστικά υφής Haralick.....	48
3.5.3	Ανίχνευση ακμών – Φίλτρο Sobel	53
3.5.4	Μέση Τιμή των καναλιών RGB & HSV	55
3.5.5	Χρωματικό Ιστόγραμμα	56

3.6	ΠΕΡΙΓΡΑΦΗ ΤΑΞΙΝΟΜΗΣΗΣ ΔΕΔΟΜΕΝΩΝ	58
3.6.1	Classification Learner.....	58
3.6.2	Cross – Validation.....	61
3.6.3	Επιλογή Ταξινομητή	63
3.6.4	Μέτρο Επίδοσης.....	70
3.7	ΠΕΙΡΑΜΑΤΑ	72
3.7.1	ΚΩΔΙΚΑΣ.....	72
3.7.2	ΑΠΟΤΕΛΕΣΜΑΤΑ.....	75
3.7.2.1	Αποτελέσματα με τα 5 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά	75
3.7.2.2	Αποτελέσματα μεταξύ 2 καλλιτεχνικών ρευμάτων	84
3.7.2.3	Σύγκριση Αποτελεσμάτων	105
	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΒΕΛΤΙΩΣΕΙΣ	111
4.1	ΣΥΜΠΕΡΑΣΜΑΤΑ	111
4.2	ΒΕΛΤΙΩΣΕΙΣ	112
	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	113

ΕΥΧΑΡΙΣΤΙΕΣ

Θα θέλαμε να ευχαριστήσουμε την επιβλέπουσα καθηγήτρια της διπλωματικής μας την κ. Χριστίνα Γεωργουλάκη για την μεγάλη βοήθεια και την καθοδήγηση που μας προσέφερε, καθώς και την άριστη συνεργασία που είχαμε.

Επιπλέον, θα θέλαμε να ευχαριστήσουμε την οικογένεια και τους φίλους μας, που μας στήριξαν καθ'όλη τη διάρκεια της εκπόνησης και της συγγραφής της παρούσας διπλωματικής εργασίας.

ΚΕΦΑΛΑΙΟ 1 - ΕΙΣΑΓΩΓΗ

1.1 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Η αλματώδης ανάπτυξη στον τομέα της μηχανικής όρασης, τα τελευταία χρόνια, οδηγεί σε ολοένα και πιο σύνθετες εφαρμογές αυτοματοποιημένης ανάλυσης και ταξινόμησης εικόνων. Μία από τις πιο απαιτητικές εφαρμογές της μηχανικής όρασης θεωρείται η ταξινόμηση πινάκων έργων τέχνης, η οποία μπορεί να περιλαμβάνει την αυθεντικοποίηση δημιουργού, τη διερεύνηση επίδρασης μεταξύ καλλιτεχνών, την ταξινόμηση έργου τέχνης με βάση το δημιουργό, το καλλιτεχνικό ρεύμα, το περιεχόμενο και άλλα πολλά.

Η ολοένα και αυξανόμενη ψηφιοποίηση έργων τέχνης και η επιλογή πολλών μουσείων έργων τέχνης να παρουσιάζουν τις συλλογές τους ψηφιοποιημένες στο διαδίκτυο έχουν καταστήσει την αυτοματοποιημένη ανάλυση και ταξινόμηση καλλιτεχνικών έργων πολύ χρήσιμη και σημαντική διαδικασία.

Η παρούσα εργασία πραγματεύεται το θέμα της αυτοματοποιημένης ταξινόμησης έργων τέχνης ως προς το καλλιτεχνικό ρεύμα που ανήκουν (Ιμπρεσιονισμός, Κλασικισμός, Κυβισμός κ.α.).

1.2 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ

Η παρούσα διπλωματική εργασία δομείται σε 4 κεφάλαια ως ακολούθως:

1. Εισαγωγή
2. Θεωρητικό Υπόβαθρο
3. Υλοποίηση
4. Συμπεράσματα και Βελτιώσεις

Στο Θεωρητικό Υπόβαθρο θα γίνει σύντομη εισαγωγή στην Ψηφιακή Επεξεργασία Εικόνων και στην Αναγνώριση Προτύπων. Επίσης, θα αναφερθούν τα βασικά στάδια σχεδιασμού ενός συστήματος ταξινόμησης.

Στο κεφάλαιο της Υλοποίησης θα παρουσιαστούν, οι τρέχουσες ερευνητικές τάσεις και προσεγγίσεις (state of the art) στο θέμα της αυτοματοποιημένης ταξινόμησης έργων τέχνης, με βάση τη διεθνή βιβλιογραφία.

Ακολούθως, θα αναφερθούν χαρακτηριστικά (features) των ψηφιοποιημένων έργων που είθισται να χρησιμοποιούνται για την ταξινόμηση έργων τέχνης αλλά και διάφοροι ταξινομητές που αποσκοπούν στην διεκπεραίωση της πολύπλοκης αυτής διαδικασίας με όσο δυνατόν μεγαλύτερη ακρίβεια. Στη συνέχεια, θα γίνει η περιγραφή του συστήματος που δημιουργήσαμε, το οποίο θα επιχειρεί να διακρίνει πίνακες ζωγραφικής ως προς διαφορετικά καλλιτεχνικά ρεύματα.

Πιο συγκεκριμένα, θα δημιουργηθεί μια βάση δεδομένων αποτελούμενη από ψηφιοποιημένα έργα τέχνης. Για το σκοπό αυτό θα χρησιμοποιηθούν ψηφιακές πινακοθήκες που διανέμουν ελεύθερα ψηφιοποιημένους πίνακες ζωγραφικής (π.χ. wikiart). Οι ψηφιοποιημένοι πίνακες θα επεξεργαστούν προκειμένου να εξαχθούν επιλεγμένα χαρακτηριστικά (features) που κωδικοποιούν το καλλιτεχνικό ρεύμα. Η επεξεργασία θα γίνει στο περιβάλλον του Matlab που παρέχει πολλά μαθηματικά εργαλεία και συναρτήσεις για την επεξεργασία εικόνας και την αναγνώριση προτύπων. Η πληροφορία που προκύπτει θα χρησιμοποιηθεί από ταξινομητή (π.χ. K-NN, Naïve Bayes, SVM). Οι πίνακες θα χωριστούν στο σύνολο εκπαίδευσης, το οποίο θα εκπαιδεύει τον ταξινομητή και στο σύνολο δοκιμής, το οποίο θα ελέγχει την ικανότητα του ταξινομητή να ταυτοποιήσει το καλλιτεχνικό ρεύμα.

Στο τελευταίο κεφάλαιο θα εξαχθούν συμπεράσματα για την ταξινομητική ικανότητα του συστήματος ταξινόμησης που υλοποιήσαμε και θα δοθούν προτάσεις για βελτίωση του συστήματος.

1.3 ΣΥΝΕΙΣΦΟΡΑ ΤΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ

Για την πραγματοποίηση της ταξινόμησης έργων τέχνης ως προς το καλλιτεχνικό ρεύμα πραγματοποιήθηκαν τα εξής:

- Δημιουργία βάσης δεδομένων ψηφιακών έργων τέχνης
- Δημιουργία συστήματος για την ταξινόμηση των έργων τέχνης το οποίο αποτελείται από
 - a) πηγαίο κώδικα για την επεξεργασία των πινάκων και την εξαγωγή κατάλληλων χαρακτηριστικών
 - b) χρήση του classification learner του Matlab για την ταξινόμηση έργων τέχνης
- Σύγκριση των ταξινομητών

- Αξιολόγηση των αποτελεσμάτων

ΚΕΦΑΛΑΙΟ 2 – ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ

2.1.1 Ο κλάδος της Ψηφιακής Επεξεργασίας Εικόνας

Η ψηφιακή επεξεργασία εικόνας είναι ένας κλάδος της επιστήμης των υπολογιστών που εφαρμόζει μια σειρά λειτουργιών σε ψηφιακές εικόνες για να επιτύχει τα επιθυμητά αποτελέσματα. Τα αποτελέσματα δίδονται συνήθως με τη μορφή ψηφιακών εικόνων. Εάν αντιμετωπίζουμε τις εικόνες ως δισδιάστατα σήματα, τότε η ψηφιακή επεξεργασία εικόνας είναι μια ειδική περίπτωση της επιστήμης της επεξεργασίας ψηφιακών σημάτων [27].



Εικόνα 1 - Ψηφιακή Επεξεργασία Εικόνας [26]

Η ψηφιακή επεξεργασία εικόνας είναι μια επιστήμη που ξεκίνησε στις αρχές της δεκαετίας του 1920 [26]. Μία από τις πρώτες εφαρμογές ήταν η βελτίωση της ποιότητας των ψηφιοποιημένων εφημερίδων που μεταδόθηκαν μέσω υποβρυχίων καλωδίων που συνδέουν το Λονδίνο και τη Νέα Υόρκη.

Στα επόμενα 30-40 χρόνια, η πρόοδος των μεθόδων επεξεργασίας εικόνας ήταν σημαντική, αλλά περιορίζεται κυρίως στην ψηφιακή μετάδοση εικόνας. Ωστόσο, στη δεκαετία του 1960, χάρη στο διαστημικό πρόγραμμα των ΗΠΑ και τη νέα υπολογιστική ισχύ που παρέχονται από υπολογιστές, η ώθηση της περιοχής στις μεθόδους επεξεργασίας ήταν εντυπωσιακή και η πρόοδος ήταν ταχεία μέχρι στιγμής. Σήμερα, η τεχνολογία επεξεργασίας εικόνας χρησιμοποιείται σε πολλές διαφορετικές εφαρμογές, όπως:

- Οπτική αναγνώριση χαρακτήρων (εκτύπωση και χειρόγραφο)
- Μηχανική-υπολογιστική όραση και ρομποτική
- Ανάλυση δεδομένων τηλεπισκόπησης (π.χ. δορυφορικές εικόνες, αεροφωτογραφίες)

- Αναγνώριση αντικειμένων
- Βιοϊατρικές εφαρμογές

Μερικοί τύποι επεξεργασίας εικόνας είναι οι εξής:

- Γεωμετρικός μετασχηματισμός: αλλαγή μεγέθους, περιστροφή, παραμόρφωση, αλλαγή προοπτικής, αλλαγή ανάλυσης (σε ψηφιακή εικόνα) ολόκληρου ή μέρους της εικόνας κ.λπ.
- Αλλαγή και διόρθωση χρώματος: αλλάζουμε την απόχρωση της εικόνας, ρυθμίζουμε τη φωτεινότητα, την αντίθεση, αλλάζουμε το χώρο χρώματος (μοντέλο), για παράδειγμα από RGB σε CMYK.
- Συμπίεση και μετατροπή της εικόνας του υπολογιστή (μετατροπή αρχείου), για παράδειγμα από τη μορφή .jpg σε μορφή .tif.
- Εφαρμογή φίλτρων για να βελτιώσουμε την ποιότητα της εικόνας ή να τονίσουμε τις δυνατότητές του (εξαλείφει το "κόκκινο μάτι" στις εικόνες του προσώπου, την ανίχνευση άκρων και την επισήμανση περιγράμματος, μείωση θορύβου κ.λπ.).
- Αναμειγνύουμε δύο ή περισσότερες εικόνες για να σχηματίσουμε μία (μοντάζ).
- Χωρίζουμε την εικόνα σε πολλές περιοχές, ο σκοπός είναι να ορίσουμε την περιοχή ενδιαφέροντος (Regions of Interest, ROI) στην εικόνα. Ενδεικτικά, η δυαδική κατάτμηση μπορεί να γίνει στο προσκήνιο και στο φόντο.
- Ανάκτηση: σκοπός είναι να εξαχθεί η "σωστή" έκδοση της εικόνας από τη θολή παραμορφωμένη εικόνα εισόδου.

2.1.2 Ψηφιακή Εικόνα

Μια εικόνα είναι το αποτέλεσμα της μετατροπής ενός σημείου σε έναν τρισδιάστατο πραγματικό χώρο σε δισδιάστατο. Ο όρος «μετασχηματισμός σημείου» σημαίνει τη μετατροπή ενός μετρήσιμου μεγέθους από ένα σημείο σε τρισδιάστατο χώρο σε ένα σημείο σε δισδιάστατο χώρο. Επομένως, χρησιμοποιώντας μαθηματικούς όρους, η εικόνα θεωρείται ένα σήμα δύο ανεξάρτητων χωρικών μεταβλητών x, y που αντιστοιχούν στις γεωμετρικές διαστάσεις του φωτεινού υποστρώματος (μήκος, πλάτος). Η εξαρτημένη μεταβλητή $z=f(x,y)$ του σήματος της

εικόνας αντιστοιχεί στις μεταβολές του φυσικού μεγέθους που περιγράφει, δηλαδή, στην ένταση της εικόνας στο συγκεκριμένο σημείο του χώρου.

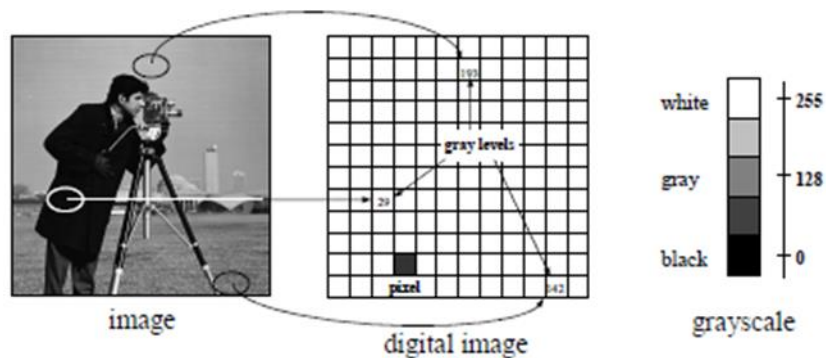
Η ένταση της εικόνας μπορεί να περιγράφει την πληροφορία μίας έγχρωμης ή μίας ασπρόμαυρης εικόνας.



Εικόνα 2 – Μονοχρωματική Εικόνα [26]

Οι χωρικές συντεταγμένες x και y και η τιμή της συνάρτησης $f(x, y)$ θεωρούνται συνεχείς (δηλαδή πραγματικοί αριθμοί με απεριόριστη ακρίβεια), επομένως δεν μπορούν να χρησιμοποιηθούν από τον υπολογιστή χωρίς ψηφιοποίηση. Μια ψηφιακή εικόνα είναι μια εικόνα $f(x, y)$ που έχει ψηφιοποιηθεί από άποψη χωρικών συντεταγμένων και φωτεινότητας.

Μια ψηφιακή εικόνα μπορεί να θεωρηθεί ως ένας πίνακας του οποίου οι δείκτες γραμμής και στήλης προσδιορίζουν ένα σημείο στην εικόνα και η τιμή του αντίστοιχου στοιχείου καθορίζει την ένταση του σημείου. Τα στοιχεία του πίνακα ονομάζονται εικονοστοιχεία ή pixel (Picture element).



Εικόνα 3 - Μία φυσική εικόνα και μία αντίστοιχη ψηφιακή εικόνα. [26]

2.1.3 Βάθος Χρώματος

Το βάθος χρώματος (color depth) ονομάζεται επίσης βάθος bit, το οποίο αναφέρεται στον αριθμό των bit που χρησιμοποιούνται για να αναπαραστήσουν το χρώμα των pixel σε μια εικόνα. Όσο μεγαλύτερο είναι το βάθος χρώματος, δηλαδή όσο περισσότερα bits χρησιμοποιεί κάθε χρώμα, τόσο περισσότερες αποχρώσεις μπορεί να περιέχει η εικόνα. Εάν το βάθος χρώματος είναι n , μπορούν να κωδικοποιηθούν 2^n αποχρώσεις [25].

Όταν το βάθος χρώματος είναι 1 bit, η εικόνα μπορεί να έχει $2^1 = 2$ διαφορετικούς τόνους. Συνήθως, αυτά τα χρώματα είναι το λευκό και το μαύρο, αλλά μπορεί να είναι οποιαδήποτε 2 χρώματα. Σε αυτήν την περίπτωση, η εικόνα ονομάζεται μονόχρωμη (monochrome ή monochromatic) ή διτονική (binary).



Εικόνα 4 – Εικόνα με 1 bit βάθος χρώματος / 2 χρώματα [47].

Κατά συνέπεια, μια εικόνα μπορεί να έχει βάθος χρώματος:

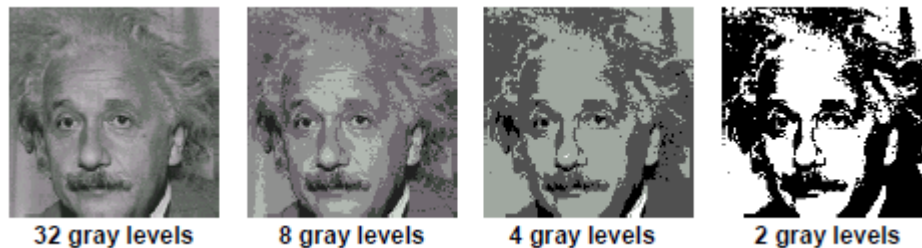
- $2^1 = 2$: Ασπρόμαυρη εικόνα (χωρίς διαβαθμίσεις γκριζου).
- $2^8 = 256$ χρώματα (ή αποχρώσεις του γκριζου).
- $2^{16} = 65536$ χρώματα. Η εικόνα με αυτό το βάθος χρώματος αναφέρεται και ως Highcolor.
- $2^{24} = 16.777.216$ χρώματα. Η εικόνα με αυτό το βάθος χρώματος αναφέρεται και ως Truecolor.

- 2^{48} χρώματα. Αυτό το βάθος χρώματος υπερβαίνει την ικανότητα του ανθρώπινου ματιού. Χρησιμοποιείται, ωστόσο, για πρακτικούς λόγους, από πολλούς σαρωτές.

Οι εικόνες που αποτελούνται από 256 χρώματα (ή λιγότερα) συνήθως αποθηκεύονται στη μνήμη του υπολογιστή ως παλέτα. Για βάθη μεγαλύτερα από 8 bit, κάθε pixel αντιπροσωπεύεται από ένα αντίστοιχο επίπεδο τριών χρωμάτων RGB (κόκκινο, πράσινο και μπλε).

Το βάθος χρώματος 16 bit συνήθως "διαίρεται" σε 5 bit για κάθε χρώμα κόκκινου και μπλε και 6 bit για το πράσινο, επειδή το ανθρώπινο μάτι είναι πιο ευαίσθητο στη διαβάθμιση του πράσινου από τα άλλα δύο χρώματα. Άλλες φορές, το 16 bit αντιπροσωπεύει οποιαδήποτε χρωματική διαφάνεια.

Στα βάθος χρώματος των 24 bits υπάρχουν 8 bits ανά βασικό χρώμα, δηλαδή $2^8 = 256$ διαβαθμίσεις κάθε βασικού χρώματος. Μερικές φορές μπορεί να χρησιμοποιηθεί και βάθος χρώματος των 32 bits. Σε αυτήν την περίπτωση τα 8 επιπλέον bits υποδεικνύουν τη συνοχή του χρώματος. [29]



Εικόνα 5 - Επίδραση βάθους χρώματος στην εικόνα [26]

2.1.4 Χρωματικά Μοντέλα

Ο σκοπός των μοντέλων χρωμάτων (ονομάζονται επίσης χρωματικοί χώροι ή συστήματα χρωμάτων) είναι να διευκολύνουν τον ορισμό των χρωμάτων με έναν τυπικό και γενικά αποδεκτό τρόπο. Ουσιαστικά, ένα χρωματικό μοντέλο ορίζεται ως η διαδικασία καθορισμού ενός συστήματος συντεταγμένων και υποχώρων μέσα στο σύστημα, όπου κάθε χρώμα αντιπροσωπεύεται από ένα σημείο. Έχουν αναπτυχθεί διάφορα χρωματικά μοντέλα για την περιγραφή των χρωμάτων σε μαθηματική μορφή, κατάλληλα για την επεξεργασία ψηφιακών

μέσων. Έτσι, καταρχήν, η εικόνα αποσυντίθεται σε εικονοστοιχεία, καθένα από τα οποία έχει ομοιόμορφο χρώμα.

Τα περισσότερα χρωματικά μοντέλα που χρησιμοποιούνται σήμερα είναι είτε προσανατολισμένα στο υλικό (όπως στην περίπτωση των εκτυπωτών και των έγχρωμων οθονών) είτε σε εφαρμογές, μία από τις οποίες είναι ο χειρισμός των χρωμάτων (για παράδειγμα, η δημιουργία για κινούμενα σχέδια). Όσον αφορά την ψηφιακή επεξεργασία εικόνας, τα πιο συχνά χρησιμοποιούμενα μοντέλα προσανατολισμένα στο υλικό είναι τα μοντέλα RGB (κόκκινο, πράσινο και μπλε) για έγχρωμες οθόνες, καθώς και μια μεγάλη κατηγορία έγχρωμων καμερών και μοντέλων CMYK (κυανό, ματζέντα, κίτρινο, μαύρο). Για την έγχρωμη εκτύπωση, το μοντέλο HSI (απόχρωση, κορεσμός, ένταση) είναι στενά συνδεδεμένο με τον τρόπο που οι άνθρωποι περιγράφουν και ερμηνεύουν τα χρώματα.

Ένα άλλο χαρακτηριστικό του μοντέλου HSI είναι ότι εξαλείφει τη σύζευξη μεταξύ του χρώματος και των πληροφοριών που σχετίζονται με την κλίμακα του γκρι στην εικόνα, καθιστώντας το κατάλληλο για την εφαρμογή πολλών τεχνικών. Υπάρχουν πολλά μοντέλα χρωμάτων που χρησιμοποιούνται σήμερα, επειδή η επιστήμη των χρωμάτων είναι ένα ευρύ πεδίο, που περιλαμβάνει πολλές εφαρμογές [9].

2.1.4.1 Χρωματικό Μοντέλο RGB

Στο χρωματικό μοντέλο RGB, κάθε χρώμα περιγράφεται από τα βασικά χρώματα (κόκκινο, πράσινο και μπλε). Το μοντέλο χρησιμοποιεί αυτά τα βασικά χρώματα για να κωδικοποιήσει όλα τα χρώματα που μπορούν να εμφανιστούν στην οθόνη (συνήθως έναν υπολογιστή).

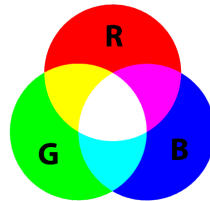
Το χρωματικό μοντέλο RGB είναι ένα προσθετικό χρωματικό μοντέλο καθώς τα τρία βασικά χρώματα προστίθενται μαζί με διάφορους τρόπους για να παράγουν μια ευρεία γκάμα χρωμάτων.

Δηλαδή, τα διαφορετικά χρώματα δημιουργούνται συνδυάζοντας τρία βασικά χρώματα σε μια συγκεκριμένη αναλογία. Τα βασικά, δευτερεύοντα χρώματα και μερικά παραδείγματα στην έκδοση 8-bit του μοντέλου παρατίθενται παρακάτω [28]:

- Μαύρο: (0,0,0)
- Λευκό: (255,255,255)

- Κόκκινο: (255,0,0)
- Πράσινο: (0,255,0)
- Μπλε: (0,0,255)
- Κίτρινο: (255,255,0)
- Γαλάζιο: (0,255,255)
- Ματζέντα: (255,0,255)
- Πορτοκαλί: (255,102,0)

Το μοντέλο βασίζεται στο καρτεσιανό σύστημα συντεταγμένων. Από την άλλη πλευρά, τα δευτερεύοντα χρώματα του μπλε, του ματζέντα και του κίτρινου αντιστοιχούν στις άλλες τρεις γωνίες. Το μαύρο αντιστοιχεί στην αρχή του συστήματος συντεταγμένων, ενώ το λευκό αντιστοιχεί στην κορυφή όσο το δυνατόν πιο μακριά από την αρχή του άξονα. Σε αυτό το μοντέλο, η κλίμακα του γκρι (δηλαδή, τα σημεία όπου το κόκκινο, το πράσινο και το μπλε έχουν την ίδια τιμή) εκτείνεται από το μαύρο στο λευκό κατά μήκος της γραμμής που συνδέει αυτά τα δύο σημεία. Όσον αφορά τα διαφορετικά χρώματα, αυτά είναι σημεία στην επιφάνεια ή στο εσωτερικό του κύβου, που περιγράφονται με τη βοήθεια διανυσμάτων ξεκινώντας από την αρχή του συστήματος συντεταγμένων. Για ευκολία στη χρήση, υποθέτουμε ότι όλες οι τιμές χρώματος έχουν κανονικοποιηθεί, έτσι ώστε ο κύβος να είναι ένας κύβος μονάδας. Με άλλα λόγια, αυτό σημαίνει ότι όλες οι τιμές των στοιχείων R, G και B ανήκουν στο εύρος τιμών [0,1] [9].



Εικόνα 6 - Χρωματικό Μοντέλο RGB

Η εικόνα που αναπαρίσταται στο χρωματικό μοντέλο RGB αποτελείται από τρεις εικόνες στοιχείων, μία για κάθε κύριο χρώμα. Όταν αυτές οι εικόνες αποστέλλονται στην οθόνη, συνδυάζονται για να σχηματίσουν μια πολύπλοκη έγχρωμη εικόνα. Ο αριθμός των bit που χρησιμοποιούνται για την αναπαράσταση κάθε pixel ονομάζεται βάθος pixel. Για παράδειγμα, μια εικόνα RGB στην οποία χρησιμοποιούνται 8 bit για να περιγράψει κάθε εικόνα που αντιστοιχεί σε κόκκινο, πράσινο και μπλε. Υπό αυτές τις συνθήκες, κάθε pixel που αντιστοιχεί

σε ένα χρώμα RGB [δηλαδή, μια τριάδα τιμών (R, G, B)] λέγεται ότι έχει βάθος 24 bit (από τον αριθμό των bit σε καθεμία από τις τρεις εικόνες επίπεδα). Για την περιγραφή των εικόνων RGB που περιγράφονται με τη βοήθεια 24-bit, χρησιμοποιείται συνήθως ο όρος πανχρωματική εικόνα. Ο συνολικός αριθμός χρωμάτων σε μια εικόνα RGB 24-bit είναι ίσος με $(2^8)^3 = 16.777.216$ [9].

2.1.4.2 Χρωματικό Μοντέλο CMY και CMYK

Τα τρία βασικά χρώματα στο CMY είναι: Κυανό (Cyan) – Ματζέντα (Magenta) – Κίτρινο (Yellow). Με τα τρία αυτά χρώματα δημιουργούνται τα δευτερογενή Κόκκινο – Πράσινο – Μπλε ως εξής:

- Κόκκινο: Ματζέντα + Κίτρινο
- Πράσινο: Κίτρινο + Κυανό
- Μπλε: Κυανό + Ματζέντα

Όπως αναφέρθηκε παραπάνω, το κυανό, το ματζέντα και το κίτρινο είναι τα δευτερεύοντα χρώματα του φωτός ή τα βασικά χρώματα των χρωστικών. Για παράδειγμα, όταν μια επιφάνεια καλυμμένη με μπλε βαφή ακτινοβολείται με λευκό φως, το κόκκινο μπορεί να μην ανακλάται από την επιφάνεια. Δηλαδή, με άλλα λόγια, το κυανό αφαιρεί το κόκκινο φως από το λευκό φως που ανακλάται στην επιφάνεια και είναι γνωστό ότι το λευκό φως αποτελείται από ίσες ποσότητες κόκκινου, πράσινου και μπλε φωτός.[28]

Αυτό το μοντέλο βασίζεται στο γεγονός ότι το τυπωμένο φόντο είναι λευκό χαρτί που αντανακλά όλα τα χρώματα (μήκη κύματος). Κάθε βασικό χρώμα στο οποίο προστίθεται μελάνι θα απορροφήσει μερικά από τα χρώματα και θα παράγει τα υπόλοιπα. Για παράδειγμα, το κίτρινο μελάνι απορροφά το μπλε και αντανακλά το πράσινο και το κόκκινο. Εδώ, ο συνδυασμός των τριών βασικών χρωμάτων δίνει μαύρο (πλήρης απορρόφηση ακτινοβολίας). Για το λόγο αυτό, το μοντέλο CMY περιγράφεται επίσης ως «αφηρημένο». Το μαύρο παράγεται επίσης από έναν συνδυασμό ενός βασικού και του αντίθετου δευτερογενούς:

- Μαύρο: Γαλάζιο + Ματζέντα + Κίτρινο
- Μαύρο: Γαλάζιο + Κόκκινο
- Μαύρο: Ματζέντα + Πράσινο

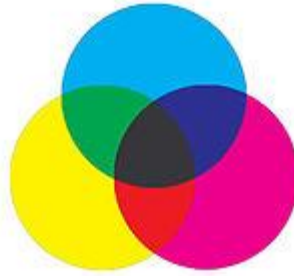
- Μαύρο: Κίτρινο + Μπλε

Ωστόσο, το μελάνι εγγενώς δεν μπορεί να αντιπροσωπεύει ένα συγκεκριμένο χρώμα μήκους κύματος (όπως τα pixel της οθόνης), αλλά ένα ευρύτερο φάσμα χρωμάτων. Το αποτέλεσμα είναι ότι ο συνδυασμός των τριών βασικών χρωμάτων δίνει καφέ αντί για μαύρο. Για το λόγο αυτό, προστέθηκε μαύρο μελάνι στο μοντέλο CMY, με αποτέλεσμα ένα χρωματικό μοντέλο CMYK (κυανό-ματζέντα-κίτρινο-μαύρο). Στην πράξη, το μοντέλο CMY δεν χρησιμοποιείται σήμερα στην εκτύπωση, αλλά χρησιμοποιείται το CMYK. Το μοντέλο CMY μπορεί να εκφραστεί ως RGB, με έναν κύβο στο σύστημα καρτεσιανού άξονα, η αρχή του άξονα είναι λευκό και ο άξονας είναι το κύριο χρώμα [28].

Όπως αναφερθήκαμε παραπάνω, οι περισσότερες συσκευές που αποθέτουν έγχρωμες χρωστικές σε χαρτί, όπως έγχρωμοι εκτυπωτές και φωτοαντιγραφικά, απαιτούν δεδομένα εισόδου CMY ή μετατρέπουν εσωτερικά χρώματα RGB σε CMY. Αυτή η μετατροπή γίνεται χρησιμοποιώντας μια απλή εξίσωση [9]:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Ας υποθέσουμε και πάλι ότι οι τιμές των χρωμάτων είναι τυποποιημένες εντός του εύρους τιμών [0,1]. Η παραπάνω εξίσωση μας λέει ότι το φως που ανακλάται από μια επιφάνεια καλυμμένη με καθαρό κυάνιο δεν περιέχει κόκκινο (δηλ. σε αυτήν την εξίσωση, είναι $C = 1 - R$). Με εντελώς παρόμοιο τρόπο, το καθαρό ματζέντα δεν αντανακλά το πράσινο και το καθαρό κίτρινο δεν αντανακλά το μπλε. Η παραπάνω εξίσωση δείχνει επίσης ότι αφαιρώντας μια μεμονωμένη τιμή CMY από τη μονάδα, η τιμή RGB μπορεί εύκολα να εξαχθεί από ένα σύνολο τιμών CMY. Όπως αναφέραμε προηγουμένως, αυτό το χρωματικό μοντέλο χρησιμοποιείται σε συνδυασμό με τη δημιουργία εκτυπώσεων σε εφαρμογές επεξεργασίας εικόνας. Επομένως, η μετατροπή από RGB σε CMY είναι σχεδόν χωρίς νόημα [9].



Εικόνα 7 - Χρωματικό μοντέλο CMYK

2.1.4.3 Χρωματικό Μοντέλο Lab

Το μοντέλο αναλύει το χρώμα κάθε pixel σε τρεις παραμέτρους L, a, b:

- Ένταση Φωτεινότητας (Luminosity).
- Η θέση χρωματικότητας μεταξύ κόκκινου και συμπληρωματικού πράσινου (a).
- Η θέση χρωματικότητας μεταξύ μπλε και συμπληρωματικού κίτρινου (b).

Στην πραγματικότητα, το μοντέλο αποτελείται από δύο υπομοντέλα, το CIE 1976 και το Hunter. Το πρώτο χρησιμοποιεί συντεταγμένες "κυβικής ρίζας" ($\sqrt[3]{x}$), ενώ το δεύτερο χρησιμοποιεί "τετραγωνική ρίζα" (\sqrt{x}). Το μοντέλο CIE είναι πιο κοινό. Η κατάταξή του είναι μία από τις κατατάξεις που πλησιάζουν περισσότερο στην ανθρώπινη αντίληψη [28].

2.1.4.4 Χρωματικό Μοντέλο HSI ή HSL

Όπως έχουμε δει προηγουμένως, η δημιουργία χρωμάτων χρησιμοποιώντας μοντέλα RGB και CMY και η εναλλαγή από το ένα μοντέλο στο άλλο είναι μια απλή διαδικασία. Όπως αναφέρθηκε προηγουμένως, αυτοί οι χρωματικοί συνδυασμοί είναι πολύ κατάλληλοι για υλοποίηση υλικού. Επιπλέον, το σύστημα RGB είναι συμβατό και όμορφο επειδή το ανθρώπινο μάτι είναι πολύ ευαίσθητο στα κύρια χρώματα του κόκκινου, του πράσινου και του μπλε, και μάλιστα σε μεγάλο βαθμό. Δυστυχώς, τα μοντέλα RGB, CMY και άλλα χρωματικά μοντέλα δεν είναι πρακτικά για ανθρώπινη ερμηνεία. Για παράδειγμα, ένα άτομο που περιγράφει το χρώμα ενός αυτοκινήτου δεν θα το περιγράψει δίνοντας το ποσοστό του κύριου χρώματος αυτού. Επιπλέον, δεν ασχολούμαστε με έγχρωμες εικόνες γιατί αποτελούνται από τρεις βασικές εικόνες, οι οποίες συνδυάζονται για να σχηματίσουν αυτή την απλή εικόνα [9].

Κάθε φορά που παρατηρούμε ένα έγχρωμο αντικείμενο, το περιγράφουμε ορίζοντας την απόχρωση (Hue), τον κορεσμό(Saturation) και τη φωτεινότητά (Luminosity) του. Η απόχρωση ή χροιά είναι ένα χαρακτηριστικό χρώμα που περιγράφει ένα καθαρό χρώμα (καθαρό κίτρινο, πορτοκαλί ή κόκκινο), ενώ ο κορεσμός μετρά τον βαθμό στον οποίο ένα καθαρό χρώμα ξεθωριάζει λόγω του λευκού φωτός. Η φωτεινότητα είναι μια βασική μέθοδος περιγραφής, η οποία δεν μπορεί να μετρηθεί στην πραγματικότητα, δίνει υπόσταση στην αχρωματική έννοια της έντασης και είναι ένας από τους βασικούς παράγοντες για την περιγραφή της αντίληψης του χρώματος. Γνωρίζουμε ότι η ένταση (δηλαδή το επίπεδο του γκρι) είναι το πιο χρήσιμο χαρακτηριστικό των μονόχρωμων εικόνων. Αυτό το ποσό είναι σαφώς μετρήσιμο και εύκολο να εξηγηθεί. Το μοντέλο που ονομάζεται χρωματικό μοντέλο HSI (απόχρωση, κορεσμός, ένταση), αφαιρεί τη συσχέτιση μεταξύ της συνιστώσας έντασης και των πληροφοριών χρώματος (δηλαδή, απόχρωση και κορεσμός) σε μια έγχρωμη εικόνα. Ως εκ τούτου, το χρωματικό μοντέλο HSI είναι ένα ιδανικό εργαλείο για την ανάπτυξη αλγορίθμων επεξεργασίας εικόνας με βάση την περιγραφή χρώματος. Αυτοί οι αλγόριθμοι είναι φυσικοί και κατανοητοί από τον άνθρωπο. Εξάλλου, οι άνθρωποι είναι οι προγραμματιστές και οι χρήστες αυτών των αλγορίθμων [9].

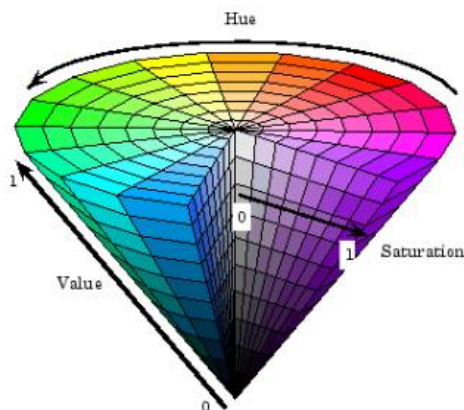
2.1.4.5 Χρωματικό Μοντέλο HSB (HSV)

Το χρωματικό μοντέλο HSB (Hue, Saturation, Brightness) αποτελείται από τα 3 εξής στοιχεία: Απόχρωση ή Χροιά (Hue), Κορεσμός ή Χρωματική καθαρότητα (Saturation) και Φωτεινότητα (Brightness). Το μοντέλο HSB αναφέρεται συχνά και ως HSV (Hue, Saturation, Value).

Η χροιά καθορίζεται από το βασικό χρώμα (π.χ. κόκκινο) της απόχρωσης. Εκφράζεται ως ένας βαθμός από 0° ως 360°. Το μοντέλο βασίζεται στα βασικά χρώματα που είναι τα πρωτεύοντα, κόκκινο, πράσινο και μπλε και τα συμπληρωματικά κυανό, πορφυρό και κίτρινο.

Ο κορεσμός είναι η καθαρότητα ή ένταση του χρώματος και αναφέρεται στο βαθμό απουσίας του λευκού φωτός από ένα χρώμα. Προσδιορίζεται ως ποσοστό %. Για παράδειγμα, ποσοστό κορεσμού 100% σημαίνει απουσία του λευκού, άρα έντονο χρώμα, ενώ ποσοστό κορεσμού 0% σημαίνει πλήρης παρουσία του λευκού, άρα το χρώμα εμφανίζεται ξεθωριασμένο. Καθορίζει, δηλαδή πόσο έντονο ή ξεθωριασμένο θα είναι το βασικό χρώμα (δηλαδή η χροιά).

Η φωτεινότητα είναι το μέτρο της φωτεινής έντασης ενός χρώματος. Χαμηλή φωτεινότητα σημαίνει σκούρο χρώμα, ενώ υψηλή φωτεινότητα σημαίνει ανοιχτό χρώμα [25].



Εικόνα 8 - Ο χρωματικός Χώρος HSV

Το χρωματικό μοντέλο HSV επινοήθηκε το 1978 από τον A. R. Smith. Το σύστημα συντεταγμένων είναι κυκλικό και τα χρώματα βρίσκονται μέσα σε έναν κώνο (Εικόνα 8). Η τιμή του H κυμαίνεται από 0° έως 360° , ενώ οι τιμές των S και V από 0 μέχρι 1 [30].

Η μετάβαση από τον χώρο RGB στον χώρο HSV γίνεται με χρήση των παρακάτω μετασχηματισμών.

Έστω r, g, b είναι οι τιμές των συνιστωσών R, G, B κανονικοποιημένες στο διάστημα [0, 1]. Θέτουμε την παράμετρο max ίση με τη μέγιστη από τις παραπάνω τιμές, και ομοίως την παράμετρο min ίση με την ελάχιστη. Τότε:

Για τον υπολογισμό της γωνίας h [0, 360) της συνιστώσας H (Hue) του χώρου HSV:

$$h = \begin{cases} 0, & \text{if } \max = \min \\ 60^\circ * \frac{g - b}{\max - \min} + 0^\circ, & \text{if } \max = r \text{ and } g \geq b \\ 60^\circ * \frac{g - b}{\max - \min} + 360^\circ, & \text{if } \max = r \text{ and } g < b \\ 60^\circ * \frac{b - r}{\max - \min} + 120^\circ, & \text{if } \max = g \\ 60^\circ * \frac{r - g}{\max - \min} + 240^\circ, & \text{if } \max = b \end{cases}$$

Για τον υπολογισμό της τιμής s [0, 1] της συνιστώσας S (Saturation) του HSV:

$$s = \begin{cases} 0, & \text{if } \max = 0 \\ \frac{\max - \min}{\max} = 1 - \frac{\min}{\max}, & \text{otherwise} \end{cases}$$

Για τον υπολογισμό της τιμής v $[0, 1]$ της συνιστώσας V (Value) του HSV:

$$v = \max$$

Οι παραπάνω εξισώσεις εφαρμόζονται σε κάθε pixel της εικόνας RGB ξεχωριστά [30].

2.1.5 Ακμές

Ακμές είναι εκείνες οι περιοχές της εικόνας στις οποίες εμφανίζονται μεγάλες μεταβολές της φωτεινότητας. Αυτές δημιουργούνται στα σημεία που αντιστοιχούν στις οριακές επιφάνειες των αντικειμένων της απεικονιζόμενης σκηνής και γι' αυτό η ανίχνευσή τους αποτελεί σημαντικό εργαλείο για την κατάτμηση της εικόνας [19].

Παρά τη θεμελιώδη σημασία των ακμών στην ψηφιακή επεξεργασία και την ανάλυση της εικόνας, δεν μπορεί να δοθεί ακριβής μαθηματικός ορισμός για αυτές. Τις δεχόμαστε ως περιοχές, στις οποίες η μεταβολή της φωτεινότητας είναι πολύ πιο έντονη από ό,τι στην υπόλοιπη εικόνα. Αυτό σημαίνει ότι στα σημεία των ακμών, η παράγωγος της φωτεινότητας, είναι σημαντικά αυξημένη. Η θέση τους, λοιπόν, εντοπίζεται υπολογίζοντας την παράγωγο σε όλα τα σημεία της εικόνας και ξεχωρίζοντας εκείνα, στα οποία αυτή έχει τις πιο μεγάλες τιμές [19].

2.1.5.1 Κατωφλίωση

Με την τεχνική αυτή η εικόνα μετατρέπεται από ασπρόμαυρη σε μονόχρωμη, σε εικόνα δηλαδή δύο τόνων. Η κατωφλίωση χρησιμοποιείται για το διαχωρισμό περιοχών της εικόνας με διαφορετικές εντάσεις φωτεινότητας. Χρησιμοποιείται επίσης για να διαχωρίσει τις ακμές από την υπόλοιπη εικόνα ύστερα από την ανίχνευσή τους [19].

2.1.5.2 Εντοπισμός περιγράμματος περιοχής εικόνας

Μετά την ανίχνευση των ακμών σε μια εικόνα είναι ανάγκη να εντοπιστούν τα περιγράμματα των περιοχών της. Αν η περιοχή που μας ενδιαφέρει μπορεί να διαχωριστεί με κατωφλίωση από την υπόλοιπη εικόνα, ο εντοπισμός του περιγράμματος επιτυγχάνεται σχετικά εύκολα. Σε

αντίθετη περίπτωση η διαδικασία εντοπισμού του περιγράμματος είναι πιο δύσκολη. Στη δεύτερη αυτή περίπτωση, το περίγραμμα της περιοχής υπολογίζεται συνδέοντας τις σημαντικές ακμές της εικόνας. Ο θόρυβος όμως της εικόνας και ο ανομοιόμορφος φωτισμός διασπών τη συνέχεια των ακμών και συγχρόνως δημιουργούν ψευδείς ακμές [19].

2.1.6 Ιστόγραμμα εικόνας

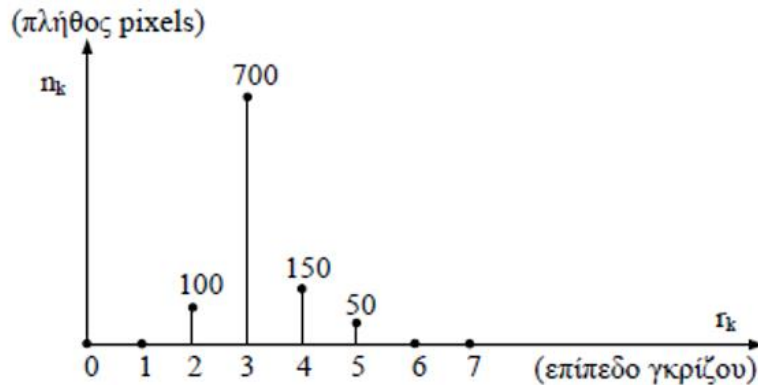
Το ιστόγραμμα (histogram) μιας εικόνας με επίπεδα γκριζου στο σύνολο $\{0, \dots, L-1\}$ δίνεται από τη διακριτή συνάρτηση:

$$P(r_k) = n_k / n \quad ; \quad k = 0, 1, \dots, L-1$$

όπου r_k είναι το k -στό επίπεδο γκριζου, n_k είναι το πλήθος των εικονοστοιχείων της εικόνας με επίπεδο γκριζου ίσο με r_k και n είναι το συνολικό πλήθος των εικονοστοιχείων της εικόνας [26]. Συνήθως, το ιστόγραμμα απεικονίζεται ως

$$h(r_k) = n_k \quad ; \quad k = 0, 1, \dots, L-1$$

όπως και στο παράδειγμα της Εικόνας 9 για $n = 1000$ pixels και $L = 8$ επίπεδα γκριζου [26].



Εικόνα 9 – Ιστόγραμμα εικόνας ($(\sum_{k=0}^7 n_k = n = 1000, L = 8)$).

Το ιστόγραμμα μας δίνει μια συνολική περιγραφή της εμφάνισης μιας εικόνας. Για παράδειγμα, αν η πλειονότητα των εικονοστοιχείων έχουν επίπεδα γκριζου κοντά στο αριστερό (δεξί) άκρο του ιστογράμματος τότε η εικόνα θα είναι σκοτεινή (φωτεινή). Αν η εικόνα έχει μικρή δυναμική περιοχή, δηλαδή αν τα επίπεδα γκριζου εκτείνονται σε ένα μικρό μέρος του $\{0, \dots, L-1\}$, τότε θα έχει συνήθως και χαμηλή χρωματική αντίθεση (contrast) [26].

2.1.7 Υφή

Μία σημαντική προσέγγιση στην περιγραφή μιας περιοχής, είναι η ποσοτική καταμέτρηση της σύστασης ή υφής της (texture). Αν και δεν υπάρχει κάποιος επίσημος ορισμός για αυτή την έννοια, διαισθητικά αυτός ο περιγραφέας μας επιτρέπει να μετρήσουμε ιδιότητες όπως είναι η ομαλότητα, η τραχύτητα και η κανονικότητα. Οι τρεις κύριες μέθοδοι που χρησιμοποιούνται στην επεξεργασία εικόνας για την περιγραφή της υφής μιας περιοχής, είναι οι στατιστικές, οι δομικές και οι φασματικές προσεγγίσεις [9].

Οι στατιστικές μέθοδοι επιτρέπουν το χαρακτηριστικό της υφής ως ομαλής, τραχιάς, κοκκώδους, κ.ο.κ. Οι δομικές τεχνικές μελετούν τις διατάξεις των πρωτογενών στοιχείων της εικόνας, όπως είναι η περιγραφή της υφής με τη βοήθεια ομοιόμορφα διατεταγμένων παράλληλων γραμμών. Τέλος, οι φασματικές τεχνικές, στηρίζονται στις ιδιότητες του φάσματος κατά Fourier και χρησιμοποιούνται κατά κύριο λόγο για την ανίχνευση ολικής περιοδικότητας σε μία εικόνα, με ταυτοποίηση στο φάσμα της στενών κορυφών υψηλής ενέργειας [9].

2.2 ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ

2.2.1 ΟΡΙΣΜΟΣ

Η αναγνώριση προτύπων είναι ένα πεδίο της επιστήμης που ασχολείται με το σχεδιασμό συστημάτων ικανών να ταξινομούν πρότυπα (patterns) σε κατηγορίες – κλάσεις (classes). Η διαδικασία αναγνώρισης προτύπων αποτελείται συνήθως από δύο βασικά στάδια. Στο πρώτο στάδιο, βασικά χαρακτηριστικά (features) παράγονται και υπολογίζονται για κάθε πρότυπο που λαμβάνει η είσοδος. Ένα βασικό χαρακτηριστικό είναι μια κατάλληλη κωδικοποίηση των πληροφοριών που περιέχονται σε κάθε πρότυπο. Με άλλα λόγια, είναι ποσοτικοποιήσεις χρήσιμων πληροφοριών που χαρακτηρίζουν κάθε πρότυπο. Στο δεύτερο στάδιο, τα πρότυπα ταξινομούνται σε μία από τις κατηγορίες του συστήματος σύμφωνα με τα χαρακτηριστικά που επιλέχθηκαν [19].

Οι τεχνικές ταξινόμησης χωρίζονται σε δύο μεγάλες κατηγορίες. Σε τεχνικές ταξινόμησης με επίβλεψη και τεχνικές ταξινόμησης χωρίς επίβλεψη. Στην πρώτη από αυτές τις δύο κατηγορίες,

οι κλάσεις είναι γνωστές εκ των προτέρων και το σύστημα ταξινόμησης εκπαιδεύεται με πολλά γνωστά πρότυπα εκπαίδευσης από κάθε κλάση. Μετά την εκπαίδευση (σχεδίαση) ενός συστήματος ταξινόμησης, αυτό μπορεί να ταξινομήσει άγνωστα μοτίβα σε κάποια κατηγορία. Στην ταξινόμηση χωρίς επίβλεψη, το πρόβλημα είναι πιο δύσκολο επειδή δεν έχουμε γνωστά πρότυπα ή δεν γνωρίζουμε εκ των προτέρων τον αριθμό των κλάσεων [19].

2.2.2 ΒΑΣΙΚΑ ΣΤΑΔΙΑ ΣΧΕΔΙΑΣΜΟΥ ΣΥΣΤΗΜΑΤΟΣ ΤΑΞΙΝΟΜΗΣΗΣ

Θα μπορούσε κανείς να πει ότι ένας ταξινομητής είναι η καρδιά ενός συστήματος ταξινόμησης. Όπως αναφέρθηκε και παραπάνω, ο σχεδιασμός ενός συστήματος ταξινόμησης μπορεί να χωριστεί σε δύο κύρια στάδια, το στάδιο δημιουργίας και επιλογής βασικών χαρακτηριστικών και το στάδιο σχεδιασμού του ταξινομητή [19].

2.2.2.1 Στάδιο δημιουργίας και επιλογής χαρακτηριστικών

Στο πρώτο στάδιο, δημιουργούνται και επιλέγονται τα βασικά χαρακτηριστικά που περιγράφουν το πρότυπο. Για παράδειγμα, αυτά τα χαρακτηριστικά μπορεί να είναι διαφορετικές ροπές περιοχής, διαφορετικά χαρακτηριστικά ιστογράμματος, χαρακτηριστικά περιγράμματος. Από τον μεγάλο αριθμό χαρακτηριστικών που μπορούν να παραχθούν, δεν περιέχουν όλα πλούσιες πληροφορίες. Για παράδειγμα, χαρακτηριστικά που έχουν περίπου την ίδια τιμή για όλες τις κλάσεις θεωρούνται ανεπαρκή για την ταξινόμηση των κλάσεων αυτών. Από όλα τα βασικά χαρακτηριστικά που μπορούν να δημιουργηθούν, θέλουμε να επιλέξουμε εκείνα των οποίων οι τιμές είναι όσο το δυνατόν διαφορετικές για πρότυπα που ανήκουν σε διαφορετικές κλάσεις. Όπως λέγεται συνήθως, με την κατάλληλη ορολογία, τα βασικά χαρακτηριστικά θα πρέπει να επιλέγονται έτσι ώστε οι τιμές που λαμβάνουν για διάφορα πρότυπα να έχουν μεγάλες αποστάσεις μεταξύ των κατηγοριών και χαμηλές αποστάσεις εντός της κατηγορίας. Στην Εικόνα 7 παρουσιάζονται δύο απλοποιημένα παραδείγματα δύο χαρακτηριστικών x_1 , x_2 . Οι αστερίσκοι υποδεικνύουν τις τιμές για τη μία κλάση και οι τελείες τις τιμές για την άλλη κλάση. Από τα παραπάνω μπορεί να συναχθεί ότι το x_1 ένα κατάλληλο, πλούσιο σε πληροφορία χαρακτηριστικό, ενώ το x_2 είναι φτωχό επειδή λαμβάνει επικαλυπτόμενες τιμές για τις δύο κατηγορίες. Η επιλογή χαρακτηριστικών δεν είναι μια απλή

διαδικασία και είναι συνήθως το αποτέλεσμα της εφαρμογής μιας σειράς μεθόδων επιλογής χαρακτηριστικών [19].



Εικόνα 10 - Τιμές χαρακτηριστικών με (α) μεγάλη διακλασική και (β) μικρή διακλασική απόσταση.

Τότε υποθέτουμε ότι έχουμε επιλέξει I βασικά χαρακτηριστικά, x_1, x_2, \dots, x_I . Με αυτά τα χαρακτηριστικά, δημιουργούμε βασικά διανύσματα χαρακτηριστικών (feature vectors), $x = (x_1, x_2, \dots, x_I)^T$ σε I -διάστατο χώρο. Επιπλέον, υποθέτουμε μια σχέση μεταξύ προτύπων και διανυσμάτων. Δηλαδή, κάθε διάνυσμα αντιστοιχεί σε ένα πρότυπο. Μια παρατήρηση που είναι ιδιαίτερα σημαντική για εμάς είναι ότι οι τιμές που λαμβάνονται από τα διαφορετικά πρότυπα δεν είναι προβλέψιμες, αλλά μάλλον έχουν έναν βαθμό τυχαιότητας [19].

2.2.2.2 Στάδιο σχεδίασης του ταξινομητή

Αφού επιλέξουμε τα χαρακτηριστικά, στο δεύτερο στάδιο σχεδιάζουμε τον ταξινομητή. Ένας ταξινομητής δεν είναι παρά μια συνάρτηση ή ένα σύνολο συναρτήσεων με ελεύθερη μεταβλητή το τυχαίο χαρακτηριστικό x που διαιρεί έναν χώρο I διαστάσεων σε περιοχές R_k , $k = 1, 2, \dots, m$. Κάθε μία από αυτές τις περιοχές αντιστοιχεί σε μία από τις m -κλάσεις (για προβλήματα δύο κλάσεων, $m = 2$). Αφού ολοκληρωθεί ο σχεδιασμός του ταξινομητή, η ταξινόμηση των άγνωστων προτύπων έχει ως εξής. Μετρά την τιμή που λαμβάνεται από το διανυσματικό χαρακτηριστικό που αντιστοιχεί στο συγκεκριμένο μοτίβο. Τα μοτίβα στη συνέχεια ταξινομούνται σε κατηγορίες που αντιστοιχούν στις περιοχές R_k του χώρου όπου βρίσκεται το διάνυσμα x [19].

Προφανώς, τέτοιες διαδικασίες περιλαμβάνουν επίσης την έννοια της εσφαλμένης ταξινόμησης. Ένας από τους στόχους της διαδικασίας σχεδιασμού ταξινόμησης είναι η ελαχιστοποίηση αυτού του σφάλματος [19].

2.2.3 ΑΛΓΟΡΙΘΜΟΙ ΤΑΞΙΝΟΜΗΣΗΣ

Οι αλγόριθμοι ταξινόμησης αντιμετωπίζουν το πρόβλημα της ταξινόμησης νέων προτύπων σε υπάρχουσες υποομάδες – κλάσεις, ενός δεδομένου συνόλου προτύπων (εποπτευόμενη μάθηση) [31].

Κάποιοι αλγόριθμοι ταξινόμησης είναι οι παρακάτω:

- Ταξινομητής μέγιστης εντροπίας
- Μπεϋζιανός ταξινομητής
- Δέντρα αποφάσεων, λίστες αποφάσεων
- Διάνυσμα υποστήριξης μηχανής (Support Vector Machines - SVM)
- Αλγόριθμοι εκτίμησης πυρήνα και K-κοντινότερου γείτονα (K-NN)
- Perceptron
- Νευρωνικά δίκτυα (π.χ. Πολυεπίπεδα Perceptron)

2.2.3.1 Ταξινομητής Bayes

Θα ξεκινήσουμε, για χάρη της απλότητας, από το πρόβλημα δύο κλάσεων, που θα συμβολίσουμε ω_1 και ω_2 αντίστοιχα. Υποθέτουμε ότι έχουμε στη διάθεσή μας ένα σύνολο N διανυσμάτων εκπαίδευσης, δηλαδή,

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}, \mathbf{x}_i \in R^l, i = 1, 2, 3, \dots, N$$

Για καθένα από τα διανύσματα εκπαίδευσης γνωρίζουμε την κλάση από την οποία προέρχεται. Η διαθεσιμότητα αυτής της πληροφορίας θα θεωρείται δεδομένη για το σχεδιασμό οποιουδήποτε ταξινομητή, όπου διαπραγματευόμαστε το πρόβλημα της ταξινόμησης.

Στη συνέχεια, υποθέτουμε ότι γνωρίζουμε τις πιθανότητες $P(\omega_1)$ και $P(\omega_2)$, τις πιθανότητες δηλαδή κάποιο χαρακτηριστικό διάνυσμα να ανήκει σε μία από τις δύο κλάσεις. Μία τέτοια πληροφορία είναι εύκολα υπολογίσιμη. Για παράδειγμα, εάν N_1 από τα διανύσματα εκπαίδευσης ανήκουν στην κλάση ω_1 και N_2 στην κλάση ω_2 ($N = N_1 + N_2$), τότε

$$P(\omega_1) \approx \frac{N_1}{N}, \quad P(\omega_2) \approx \frac{N_2}{N}$$

Μία άλλη πληροφορία που μπορεί επίσης να ποσοτικοποιηθεί, και γι' αυτό θα υποθέσουμε ότι είναι γνωστή, είναι η υπό συνθήκη πυκνότητες πιθανότητας $p(\mathbf{x}|\omega_1)$ και $p(\mathbf{x}|\omega_2)$. Οι πυκνότητες πιθανότητας (pdf) μπορούν, για παράδειγμα, να προσεγγιστούν από τα αντίστοιχα ιστογράμματα. Χρησιμοποιώντας τα N_1 διανύσματα της κλάσης ω_1 μπορούμε να προσεγγίσουμε την $p(\mathbf{x}|\omega_1)$ και με τα N_2 διανύσματα της ω_2 την $p(\mathbf{x}|\omega_2)$. Έχουμε τώρα στη διάθεσή μας όλα τα απαραίτητα υλικά για να σχεδιάσουμε τον ταξινομητή μας. Πίσω από τον σχεδιασμό αυτό υπάρχει ένα πολύ απλό και λογικό σκεπτικό. Όταν μας δοθεί ένα άγνωστο χαρακτηριστικό διάνυσμα \mathbf{x} (που δεν ανήκει προφανώς στο σύνολο των γνωστών διανυσμάτων εκπαίδευσης, X) ο ταξινομητής θα πρέπει να το ταξινομεί στην πιο πιθανή κλάση. Δεδομένης, δηλαδή, μιας μέτρησης του \mathbf{x} , που αντιστοιχεί στο άγνωστο πρότυπο, το παραπάνω σκεπτικό μας παραπέμπει στο να υπολογίσουμε τις πιθανότητες $P(\omega_1|\mathbf{x})$ και $P(\omega_2|\mathbf{x})$ και να ταξινομήσουμε το \mathbf{x} στην κλάση που αντιστοιχεί στη μεγαλύτερη τιμή από τις δύο πιθανότητες. Οι $P(\omega_i|\mathbf{x})$, $i = 1, 2$, είναι οι πιθανότητες το πρότυπο να είναι σε μια από τις δύο κλάσεις, όταν γνωρίζουμε την τιμή του \mathbf{x} , και λέγονται *a posteriori*, σε αντιθέση με τις $P(\omega_i)$, $i = 1, 2$, που λέγονται *a priori*. Τις πιθανότητες όμως $P(\omega_i|\mathbf{x})$, $i = 1, 2$, δεν τις γνωρίζουμε. Αυτό δεν είναι πρόβλημα, καθώς από τη θεωρία πιθανοτήτων γνωρίζουμε το θεώρημα Bayes που μας λέει ότι:

$$P(\omega_i)p(\mathbf{x}|\omega_i) = p(\mathbf{x})P(\omega_i|\mathbf{x})$$

Η

$$P(\omega_i|\mathbf{x}) = \frac{P(\omega_i)p(\mathbf{x}|\omega_i)}{p(\mathbf{x})}$$

Είμαστε τώρα έτοιμοι να διατυπώσουμε τον κανόνα ταξινόμησης:

Εάν η τιμή του \mathbf{x} είναι τέτοια ώστε: $P(\omega_1|\mathbf{x}) > P(\omega_2|\mathbf{x})$, ταξινομήσε το \mathbf{x} στην κλάση ω_1 .

Αλλιώς ταξινόμησε το \mathbf{x} στην κλάση ω_2 .

Από τον κανόνα αυτό καταλήγουμε στον ταξινομητή Bayes για δύο κλάσεις:

$$P(\omega_1)p(\mathbf{x}|\omega_1) > P(\omega_2)p(\mathbf{x}|\omega_2) \rightarrow \omega_1 \text{ αλλιώς } \omega_2$$

Παρατηρούμε ότι η $p(\mathbf{x})$ δεν συμπεριλαμβάνεται στους τελικούς υπολογισμούς. Αυτό οφείλεται στο ότι είναι μία θετική ποσότητα ανεξάρτητη των κλάσεων και μπορεί να απαλειφθεί από τις ανισότητες.

Η παραπάνω διαδικασία γενικεύεται και για την περίπτωση m κλάσεων. Σε αυτή την περίπτωση, το x καταχωρείται στην κλάση ω_k για την οποία η συνάρτηση $P(\omega_k)p(x|\omega_k)$ λαμβάνει τη μεγαλύτερη τιμή [19].

2.2.3.3 Ταξινομητές πλησιέστερης γειτνίασης

Ο ταξινομητής KNN (k nearest neighbor = k κοντινότερου γείτονα) είναι επίσης ένας αλγόριθμος που χρησιμοποιείται για την ταξινόμηση εικόνων. Είναι ένας αλγόριθμος μάθησης βασισμένος σε στιγμιότυπα δηλαδή η διαδικασία μάθησης αφορά απλά την αποθήκευση των δεδομένων εκπαίδευσης και θεωρείται ένας από τους πιο απλούς αλγόριθμους μάθησης. Θεωρείται εκπληκτικά ευέλικτος αλγόριθμος.

Για την ταξινόμηση του άγνωστου προτύπου ακολουθούμε τα βήματα:

- Επιλέγουμε ένα k , περιττό για την περίπτωση των δύο κλάσεων.
- Με κέντρο το x αυξάνουμε τον όγκο της περιοχής έως ότου περιλάβει συνολικά k διανύσματα εκπαίδευσης. Μετρούμε πόσα από αυτά ανήκουν στην κλάση ω_1 , ας πούμε k_1 και πόσα στην κλάση ω_2 , έστω k_2 .
- Ταξινομούμε το x στην κλάση ω_1 εάν $k_1 > k_2$ και διαφορετικά στην κλάση ω_2 . Για περισσότερες κλάσεις, ταξινομούμε το x στην κλάση που εκπροσωπείται με το μεγαλύτερο αριθμό διανυσμάτων από τα k συνολικά.

Η απλούστερη εκδοχή της μεθόδου αυτής είναι όταν $k = 1$. Με άλλα λόγια, όταν μας δίνεται το x , υπολογίζουμε την απόστασή του απ' όλα τα διανύσματα εκπαίδευσης του X , και το ταξινομούμε στην κλάση, στην οποία ανήκει το πλησιέστερο στο x από τα διανύσματα εκπαίδευσης. Αποδεικνύεται ότι, ο απλός αυτός ταξινομητής πλησιέστερου γείτονα μπορεί να οδηγήσει σε ικανοποιητικά αποτελέσματα όταν ο αριθμός N των δειγμάτων εκπαίδευσης είναι αρκετά μεγάλος. Πράγματι, αν θεωρήσουμε ότι το $N \rightarrow \infty$ και συμβολίσουμε με P_B την ελάχιστη πιθανότητα λάθους που αντιστοιχεί, βέβαια, στον Bayesian ταξινομητή, τότε η πιθανότητα λάθους του ταξινομητή πλησιέστερου γείτονα ($k = 1$), P_{NN} , φράσσεται από:

$$P_B \leq P_{NN} \leq 2P_B$$

Δηλαδή, το λάθος που θα κάνουμε θα είναι το πολύ διπλάσιο του Bayesian ταξινομητή. Για μικρές τιμές του P_B αυτό δεν είναι κακό. Για παράδειγμα, εάν $P_B = 0.005$ τότε $P_{NN} \leq 0.01$.

Λάθος της τάξης 1% δεν είναι κακό. Θα πρέπει, βέβαια, να τονιστεί ότι λάθη αυτής της τάξης αντιστοιχούν σε εύκολα προβλήματα, δηλαδή σε εύκολα διαχωρίσιμες κλάσεις.

Τέλος, για την περίπτωση όπου $k > 1$ υπάρχουν, επίσης, ενδιαφέροντα αποτελέσματα. Για $N \rightarrow \infty$ μπορεί να αποδειχθεί ότι η πιθανότητα λάθους του ταξινομητή k -πλησιέστερων γειτόνων, P_{kNN} , φράσσεται από:

$$P_B \leq P_{kNN} \leq P_B + \sqrt{2P_{NN}/k}$$

Προφανώς, για $k \rightarrow \infty$ $P_{kNN} \rightarrow P_B$. Όλα αυτά, βέβαια, απαιτούν μεγάλα N και k , που, βέβαια, στην πράξη κάθε άλλο παρά τον κανόνα αποτελούν. Ταυτόχρονα δημιουργούν πολύ υψηλές υπολογιστικές απαιτήσεις, δεδομένου ότι κάθε φορά θα πρέπει να υπολογίζονται οι πλησιέστεροι γείτονες μέσα από ένα μεγάλο αριθμό δειγμάτων [19].

Πρέπει να υπογραμμιστεί στο σημείο αυτό ότι η τεχνική KNN είναι υπερβολικά ευαίσθητη στην τιμή k δηλαδή στο πόσοι κοντινότεροι γείτονες χρησιμοποιούνται για την κατηγοριοποίηση. Σύμφωνα με μια εμπειρική μέθοδο πρέπει να ισχύει ότι :

$$k \leq \sqrt{\text{αριθμός στοιχείων εκπαίδευσης}}$$

2.2.3.4 Μηχανές Διανυσμάτων Υποστήριξης (SVM)

Η μηχανή διανυσμάτων υποστήριξης είναι μια δυαδική μηχανή με δυνατότητα μάθησης, η οποία επιδεικνύει ορισμένες εξαιρετικά κομψές ιδιότητες. Για να κατανοήσουμε πως δουλεύει μια τέτοια μηχανή, θα ήταν ίσως ευκολότερο να ξεκινήσουμε με την περίπτωση των διαχωρίσιμων προτύπων, όπως προκύπτει στο πλαίσιο της ταξινόμησης προτύπων. Σ' αυτό το πλαίσιο, η κεντρική ιδέα πίσω από τη μηχανή μπορεί να συνοψιστεί στο εξής:

Δοθέντος ενός δείγματος εκπαίδευσης, η μηχανή διανυσμάτων υποστήριξης κατασκευάζει ένα υπερεπίπεδο ως επιφάνεια απόφασης με τρόπο ώστε το περιθώριο διαχωρισμού μεταξύ θετικών και αρνητικών παραδειγμάτων να μεγιστοποιείται.

Αυτή η βασική ιδέα επεκτείνεται ώστε να καλύπτει και την πιο δύσκολη περίπτωση των μη γραμμικά διαχωρίσιμων προτύπων.

Μια ιδέα η οποία είναι θεμελιακή για την ανάπτυξη του αλγόριθμου μάθησης της μηχανής διανυσμάτων υποστήριξης είναι ο πυρήνας εσωτερικού γινομένου (inner product kernel) μεταξύ

ενός «διανύσματος υποστήριξης» x_i κι ενός διανύσματος x το οποίο αντλείται από το χώρο δεδομένων εισόδου. Το σημαντικότερο όλων είναι ότι τα διανύσματα υποστήριξης αποτελούνται από ένα μικρό υποσύνολο σημείων δεδομένων τα οποία εξάγει ο αλγόριθμος μάθησης από το ίδιο το δείγμα εκπαίδευσης. Στην πραγματικότητα, λόγω αυτής της σημαντικής ιδιότητας ο αλγόριθμος μάθησης που εμπλέκεται στην κατασκευή μιας μηχανής διανυσμάτων υποστήριξης αναφέρεται επίσης ως μέθοδο πυρήνα (kernel method). Η μέθοδος πυρήνα που παίζει βασικό ρόλο στη σχεδίαση μιας μηχανής διανυσμάτων υποστήριξης είναι βέλτιστη, με τη βελτιστότητά της να οφείλεται στην αποκαλούμενη κυρτή βελτιστοποίηση. Ωστόσο, αυτό το εξαιρετικά επιθυμητό χαρακτηριστικό της μηχανής επιτυγχάνεται με αντίτιμο την αυξημένη υπολογιστική πολυπλοκότητα.

Η μηχανή διανυσμάτων υποστήριξης μπορεί να χρησιμοποιείται για την επίλυση και των δύο τύπων προβλημάτων: ταξινόμησης μοτίβων και μη γραμμικής παλινδρόμησης. Ωστόσο, ο τομέας της επίλυσης δύσκολων προβλημάτων ταξινόμησης προτύπων είναι αυτός που οφείλεται περισσότερο από τις μηχανές διανυσμάτων υποστήριξης [40].

2.2.3.4.1 Βέλτιστο Υπερεπίπεδο Για Γραμμικά Διαχωρίσιμα Πρότυπα

Ας πάρουμε για παράδειγμα το δείγμα εκπαίδευσης $\{(x_i, d_i)\}_{i=1}^N$, όπου x_i είναι το πρότυπο εισόδου για το i -οστό παράδειγμα και d_i είναι η αντίστοιχη επιθυμητή απόκριση (έξοδος). Κατ'αρχήν, θα υποθέσουμε ότι το πρότυπο (κλάση) που αναπαρίσταται από το υποσύνολο $d_i=+1$ και το πρότυπο που αναπαρίσταται από το υποσύνολο $d_i=-1$ είναι, το ένα σε σχέση με το άλλο, «γραμμικά διαχωρίσιμα». Η εξίσωση μιας επιφάνειας απόφασης με τη μορφή ενός υπερεπίπεδου που εκτελεί το διαχωρισμό είναι:

$$w^T x + b = 0 \quad (1)$$

Όπου x είναι ένα διάνυσμα εισόδου w ένα προσαρμόσιμο διάνυσμα βαρών και b είναι μια πόλωση. Άρα, μπορούμε να γράψουμε:

$$\begin{aligned} w^T x_i + b &\geq 0 && \text{για } d_i = +1 \\ w^T x_i + b &< 0 && \text{για } d_i = -1 \end{aligned} \quad (2)$$

Η υπόθεση των γραμμικά διαχωρίσιμων προτύπων που κάνουμε εδώ αποσκοπεί στο να εξηγήσει την κεντρική ιδέα στην οποία βασίζονται οι μηχανές διανυσμάτων υποστήριξης σ' ένα αρκούντως απλό περιβάλλον.

Για ένα δεδομένο διάνυσμα βαρών w και πόλωση b , ο διαχωρισμός μεταξύ του υπερεπιπέδου που ορίζει η Εξίσωση (1) και του πλησιέστερου σημείου δεδομένων αποκαλείται περιθώριο διαχωρισμού (margin of separation) και συμβολίζεται ως ρ . Ο στόχος μιας μηχανής διανυσμάτων υποστήριξης είναι να βρει το συγκεκριμένο υπερεπίπεδο για το οποίο το περιθώριο διαχωρισμού, ρ , μεγιστοποιείται. Υπ' αυτή τη συνθήκη, η επιφάνεια απόφασης αναφέρεται ως βέλτιστο υπερεπίπεδο. Η Εικόνα 11 απεικονίζει τη γεωμετρική κατασκευή ενός βέλτιστου υπερεπιπέδου για ένα δισδιάστατο χώρο εισόδου.

Έστω ότι τα w_o και b_o συμβολίζουν τις βέλτιστες τιμές του διανύσματος βαρών και της πόλωσης, αντίστοιχα. Επομένως, το βέλτιστο υπερεπίπεδο, το οποίο αντιπροσωπεύει μια πολυδιάστατη γραμμική επιφάνεια απόφασης στο χώρο εισόδου, ορίζεται ως

$$w_o^T x + b_o = 0 \quad (3)$$

Σχέση η οποία είναι μια αναδιατύπωση της Εξ. (1). Η συνάρτηση διάκρισης:

$$g(x) = w_o^T x + b_o \quad (4)$$

Δίνει ένα αλγεβρικό μέτρο της απόστασης από το x έως το βέλτιστο υπερεπίπεδο. Για να το κατανοήσουμε μπορούμε να εκφράσουμε το x ως:

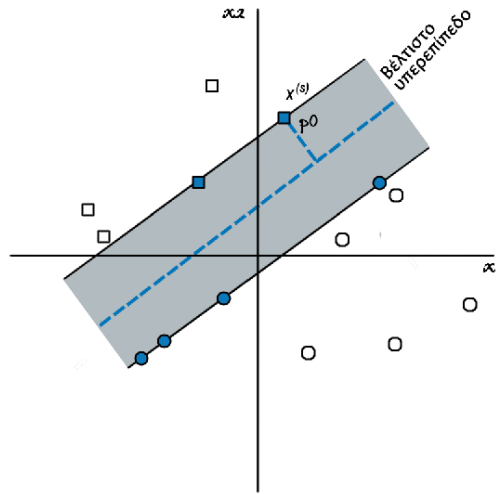
$$x = x_p + r \frac{w_o}{\|w_o\|}$$

Όπου x_p είναι η κανονική προβολή του x στο βέλτιστο υπερεπίπεδο και r είναι η επιθυμητή αλγεβρική απόσταση, η r είναι θετική εάν το x βρίσκεται στην θετική πλευρά του βέλτιστου υπερεπιπέδου και αρνητική εάν το x βρίσκεται στην αρνητική πλευρά. Εφόσον, εξ ορισμού, $g(x_p)=0$, έπεται ότι

$$g(x) = w_o^T x + b_o = r \|w_o\|$$

Η, ισοδύναμα,

$$r = \frac{g(x)}{\|w_o\|}$$



Εικόνα 11 - Σχηματική αναπαράσταση του σκεπτικού ενός βέλτιστου υπερεπίπεδου για γραμμικά διαχωρίσιμα πρότυπα: Τα σημεία δεδομένων με έγχρωμη σκίαση είναι διανύσματα υποστήριξης.

Συγκεκριμένα, η απόσταση από το σημείο αρχής (δηλ., $x=0$) έως το βέλτιστο υπερεπίπεδο δίνεται από το $b_o/\|w_o\|$. Εάν $b_o>0$, το σημείο αρχής είναι στη θετική πλευρά του βέλτιστου υπερεπίπεδου, εάν $b_o<0$, είναι στην αρνητική πλευρά. Εάν $b_o=0$, το βέλτιστο υπερεπίπεδο διέρχεται από το σημείο αρχής. Μια γεωμετρική ερμηνεία αυτών των αλγεβρικών αποτελεσμάτων δίνεται στην Εικόνα 12.

Το ζητούμενο είναι να βρεθούν οι παράμετροι w_o και b_o για το βέλτιστο υπερεπίπεδο, δοθέντος του συνόλου σύνολο εκπαίδευσης $T = \{(x_i, d_i)\}$. Υπό το πρίσμα των αποτελεσμάτων που απεικονίζεται στην Εικόνα 12, διαπιστώνουμε ότι το ζεύγος (w_o, b_o) πρέπει να ικανοποιεί τον ακόλουθο περιορισμό

$$\begin{aligned} w_o^T x_i + b_o &\geq 1 && \text{για } d_i = +1 \\ w_o^T x_i + b_o &< -1 && \text{για } d_i = -1 \end{aligned} \tag{6}$$

Εάν ισχύει η Εξίσωση (2) – δηλαδή, εάν τα πρότυπα είναι γραμμικά διαχωρίσιμα -, μπορούμε ανά πάσα στιγμή να «επανακλιμακώσουμε» τα w_o και b_o με τρόπο ώστε να ισχύει η Εξ. (6), αυτή η λειτουργία κλιμάκωσης αφήνει ανεπηρέαστη την Εξ. (3).

Τα συγκεκριμένα σημεία δεδομένων (x_i, d_i) για τα οποία η πρώτη ή η δεύτερη γραμμή της Εξ. (6) ικανοποιείται με το σύμβολο ισότητας, αποκαλούνται διανύσματα υποστήριξης (support

vectors) —σ'αυτό οφείλεται και η ονομασία των μηχανών διανυσμάτων υποστήριξης. Όλα τα υπόλοιπα παραδείγματα του δείγματος εκπαίδευσης είναι εντελώς άσχετα μεταξύ τους. Λόγω αυτής της διαφορετικότητάς τους, τα διανύσματα υποστήριξης παίζουν εξέχοντα ρόλο στη λειτουργία αυτής της κατηγορίας μηχανών μάθησης. Με εννοιολογικούς όρους, τα διανύσματα υποστήριξης είναι εκείνα τα σημεία δεδομένων που βρίσκονται πλησιέστερα στο βέλτιστο υπερεπίπεδο και, ως εκ τούτου, είναι τα πλέον δύσκολα να ταξινομηθούν. Ως τέτοια, έχουν άμεση επίδραση στη βέλτιστη θέση της επιφάνειας απόφασης.

Ας πάρουμε σαν παράδειγμα ένα διάνυσμα υποστήριξης $x^{(s)}$ για το οποίο $d^{(s)}=+1$. Τότε, εξ ορισμού, έχουμε

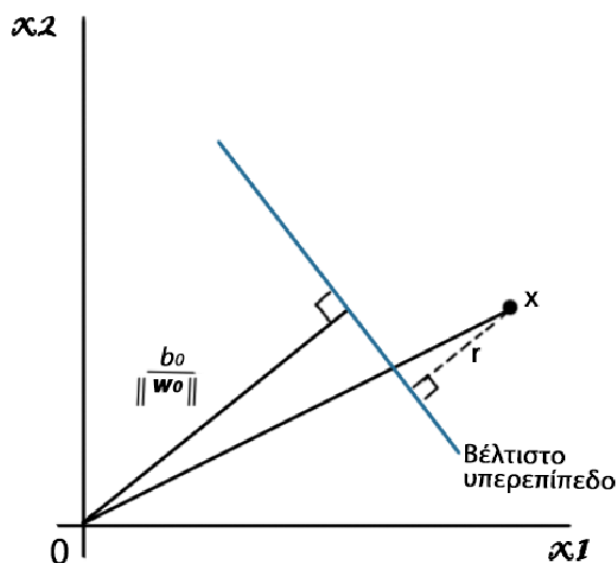
$$g(x^{(s)}) = w_o^T x^{(s)} + b_o = \mp 1 \text{ για } d^{(s)} = \mp 1 \quad (7)$$

Βάσει της Εξ. (5), η αλγεβρική απόσταση από το διάνυσμα υποστήριξης $x^{(s)}$ έως το βέλτιστο υπερεπίπεδο είναι

$$r = \frac{g(x^{(s)})}{\|w_o\|} = \begin{cases} \frac{1}{\|w_o\|} & \text{εάν } d^{(s)} = +1 \\ -\frac{1}{\|w_o\|} & \text{εάν } d^{(s)} = -1 \end{cases} \quad (8)$$

Όπου το θετικό πρόσημο υποδεικνύει ότι το $x^{(s)}$ κείται στη θετική πλευρά του βέλτιστου υπερεπιπέδου και το αρνητικό πρόσημο υποδεικνύει ότι το $x^{(s)}$ κείται στην αρνητική πλευρά του βέλτιστου υπερεπιπέδου. Έστω ότι το ρ συμβολίζει τη βέλτιστη τιμή του περιθωρίου διαχωρισμού μεταξύ των δύο κλάσεων που απαρτίζουν το δείγμα εκπαίδευσης T . Τότε η Εξ. (8),

$$\text{έπεται ότι: } \rho = 2r = \frac{2}{\|w_o\|} \quad (9)$$



Εικόνα 12 - Γεωμετρική ερμηνεία των αλγεβρικών αποστάσεων των σημείων από το βέλτιστο υπερεπίπεδο για μια περίπτωση 2 διαστάσεων.

Η Εξ. (9) δηλώνει το ακόλουθο:

Η μεγιστοποίηση του περιθωρίου διαχωρισμού μεταξύ δυαδικών κλάσεων είναι ισοδύναμη με την ελαχιστοποίηση της Ευκλείδειας νόρμας του διανύσματος βαρών w .

Συνοπτικά, το βέλτιστο υπερεπίπεδο που ορίζεται από την Εξ. (3) είναι μοναδικό, υπό την έννοια ότι το βέλτιστο διάνυσμα βαρών w_0 παρέχει το μέγιστο δυνατό διαχωρισμό μεταξύ θετικών και αρνητικών παραδειγμάτων. Αυτή η βέλτιστη συνθήκη επιτυγχάνεται ελαχιστοποιώντας την Ευκλείδεια νόρμα διανύσματος βαρών w [40].

2.2.3.4.2 Βέλτιστο Υπερεπίπεδο Για Μη Γραμμικά Διαχωρίσιμα Πρότυπα

Μέχρι τώρα, η συζήτησή μας έχει επικεντρωθεί στα γραμμικά διαχωρίσιμα πρότυπα. Σ' αυτή την ενότητα θα εξετάσουμε την πιο δύσκολη περίπτωση των μη-διαχωρίσιμων προτύπων. Δοθέντος ενός τέτοιου δείγματος δεδομένων εκπαίδευσης, δεν είναι δυνατό να κατασκευάσουμε ένα υπερεπίπεδο διαχωρισμού χωρίς να συναντήσουμε σφάλματα ταξινόμησης. Παρά ταύτα, θα

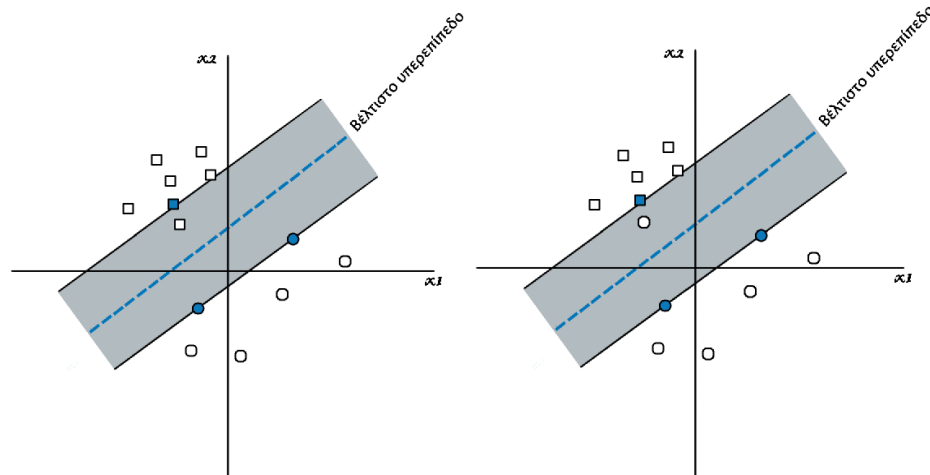
θέλαμε να βρούμε ένα βέλτιστο υπερεπίπεδο το οποίο θα ελαχιστοποιεί την πιθανότητα σφάλματος ταξινόμησης, υπολογισμένο επί του συνόλου του δείγματος εκπαίδευσης.

Το περιθώριο διαχωρισμού μεταξύ των κλάσεων λέγεται ότι είναι ελαστικό (soft) εάν υπάρχει ένα σημείο δεδομένων (x_i, d_i) το οποίο παραβιάζει την ακόλουθη συνθήκη:

$$d_i(w^T x_i + b) \geq +1, \quad i = 1, 2, \dots, N$$

Αυτή η παραβίαση μπορεί να προκύψει με έναν από δύο τρόπους:

- Το σημείο δεδομένων (x_i, d_i) εμπίπτει μέσα στην περιοχή διαχωρισμού, αλλά στη σωστή πλευρά της επιφάνειας απόφασης, όπως απεικονίζεται στην Εικόνα 13α.
- Το σημείο δεδομένων (x_i, d_i) εμπίπτει στη λάθος πλευρά της επιφάνειας απόφασης, όπως απεικονίζεται στην Εικόνα 13β.



Εικόνα 13 - Υπερεπίπεδο ελαστικού περιθωρίου (α) Το σημείο δεδομένων x_i (αναπαριστάται με ένα μικρό τετράγωνο) εμπίπτει στην περιοχή διαχωρισμού, στη σωστή πλευρά της επιφάνειας απόφασης. (β) Το σημείο δεδομένων x_i (αναπαριστάται με ένα μικρό κύκλο) εμπίπτει στη λάθος πλευρά της επιφάνειας απόφασης.

Παρατηρούμε ότι έχουμε σωστή ταξινόμηση στην πρώτη περίπτωση και λανθασμένη ταξινόμηση στη δεύτερη.

Για να “στήσουμε το σκηικό” για τη φορμαλιστική αντιμετώπιση των μη-διαχωρίσιμων σημείων δεδομένων, εισάγουμε ένα νέο σύνολο μη-αρνητικών βαθμωτών μεταβλητών, $\{\xi_i\}_{i=1}^N$,

στον ορισμό του διαχωριστικού υπερεπιπέδου (δηλ., της επιφάνειας απόφασης), όπως παρουσιάζεται εδώ:

$$d_i(w^T x_i + b) \geq +1, \quad i = 1, 2, \dots, N \quad (10)$$

Οι ξ_i αποκαλούνται μεταβλητές χαλάρωσης (slack), μετρούν την απόκλιση ενός σημείου δεδομένων από την ιδανική συνθήκη διαχωρισιμότητας προτύπων. Για $0 < \xi_i \leq 1$, το σημείο δεδομένων εμπίπτει μέσα στην περιοχή διαχωρισμού, αλλά στη σωστή πλευρά της επιφάνειας απόφασης, όπως απεικονίζεται στην Εικόνα 13α. Για $\xi_i > 1$, το σημείο εμπίπτει στη λάθος πλευρά του διαχωριστικού υπερεπιπέδου, όπως απεικονίζεται στην Εικόνα 13β. Τα διανύσματα υποστήριξης είναι εκείνα τα συγκεκριμένα σημεία δεδομένων που ικανοποιούν την Εξ. (10) ακριβώς, ακόμα κι αν $\xi_i > 0$. Επιπλέον, μπορούν να υπάρχουν διανύσματα υποστήριξης τα οποία ικανοποιούν τη συνθήκη $\xi_i = 0$. Παρατηρούμε ότι εάν ένα παράδειγμα με $\xi_i > 0$ παραληφθεί από το δείγμα εκπαίδευσης, η επιφάνεια απόφασης θα αλλάξει. Άρα, τα διανύσματα υποστήριξης ορίζονται με τον ίδιο ακριβώς τρόπο και στις δύο περιπτώσεις, γραμμικά διαχωρίσιμων και μη-διαχωρίσιμων προτύπων.

Ο στόχος μας είναι να βρούμε ένα υπερεπιπέδο διαχωρισμού για το οποίο το σφάλμα ταξινόμησης, ως μέσος όρος επί του συνόλου του δείγματος εκπαίδευσης, θα ελαχιστοποιείται. Αυτό μπορούμε να το κάνουμε ελαχιστοποιώντας τη συναρτησιακή εξίσωση

$$\Phi(\xi) = \sum_{i=1}^N I(\xi_i - 1)$$

ως προς το διάνυσμα βαρών w , υπό τον περιορισμό που περιγράφει η Εξ. (10) και τον περιορισμό στο $\|w\|^2$. Η $I(\xi)$ είναι μια συνάρτηση δείκτης (indicator function), η οποία ορίζεται ως:

$$I(\xi) = \begin{cases} 0 & \text{εάν } \xi \leq 0 \\ 1 & \text{εάν } \xi > 0 \end{cases}$$

Δυστυχώς, η ελαχιστοποίηση της $\Phi(\xi)$ ως προς το w είναι ένα πρόβλημα μη-κυρτής βελτιστοποίησης το οποίο είναι NP-πλήρες.

Για να κάνουμε αυτό το πρόβλημα βελτιστοποίησης μαθηματικώς βατό, προσεγγίζουμε τη συναρτησιακή εξίσωση $\Phi(\xi)$ γράφοντας

$$\Phi(\xi) = \sum_{i=1}^N \xi_i$$

Επιπλέον, θα απλοποιήσουμε τον υπολογισμό διατυπώνοντας τη συναρτησιακή εξίσωση που πρέπει να ελαχιστοποιηθεί ως προς το διάνυσμα βαρών w ως εξής:

$$\Phi(w, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i \quad (11)$$

Όπως και πριν, η ελαχιστοποίηση του πρώτου όρου της Εξ. (11) σχετίζεται με τη μηχανή διανυσμάτων υποστήριξης. Όσο για το δεύτερο όρο, $\sum \xi_i$, αυτός είναι ένα άνω φράγμα στον αριθμό των σφαλμάτων ελέγχου.

Η παράμετρος C ελέγχει το συμβιβασμό μεταξύ της πολυπλοκότητας της μηχανής και του αριθμού των μη-διαχωρίσιμων σημείων, συνεπώς, μπορεί να θεωρηθεί ότι είναι το αντίστροφο (reciprocal) μιας παραμέτρου η οποία αποκαλείται συνήθως παράμετρος “ομαλοποίησης”. Όταν η παράμετρος C λαμβάνει μεγάλη τιμή, αυτό σημαίνει ότι ο σχεδιαστής της μηχανής διανυσμάτων υποστήριξης έχει μεγάλη εμπιστοσύνη στην ποιότητα του δείγματος εκπαίδευσης T . Αντιστρόφως, όταν η C λαμβάνει μικρή τιμή, το δείγμα εκπαίδευσης T θεωρείται ότι είναι “θορυβώδες” και άρα θα πρέπει να δοθεί μικρότερη έμφαση σ’ αυτό.

Σε κάθε περίπτωση, η παράμετρος C πρέπει να επιλέγεται από τον χρήστη. Μπορεί να καθοριστεί πειραματικά, μέσω της στάνταρ χρήσης ενός δείγματος εκπαίδευσης (δείγμα επικύρωσης), πράγμα το οποίο είναι μια αδέξια μορφή επαναδειγματοληψίας, η χρήση της μεθόδου cross-validation (διασταυρούμενης επικύρωσης) για τη βέλτιστη επιλογή της παραμέτρου ομαλοποίησης (δηλ., $1/C$).

Σε κάθε περίπτωση, η συναρτησιακή εξίσωση $\Phi(w, \xi)$ βελτιστοποιείται ως προς τα w και $\{\xi_i\}_{i=1}^N$ υπό τον περιορισμό που περιγράφει η Εξ. (10) και τον περιορισμό $\xi_i \geq 0$. Κατ’ αυτό τον τρόπο, το τετράγωνο της νόρμας του w αντιμετωπίζεται ως ένα μέγεθος το οποίο πρέπει να ελαχιστοποιηθεί “από κοινού” ως προς τα μη-διαχωρίσιμα σημεία και όχι ως ένας περιορισμός που επιβάλλεται στην ελαχιστοποίηση του αριθμού των μη-διαχωρίσιμων σημείων.

Το πρόβλημα βελτιστοποίησης για μη-διαχωρίσιμα πρότυπα που μόλις περιγράψαμε περιλαμβάνει το πρόβλημα βελτιστοποίησης για γραμμικά διαχωρίσιμα πρότυπα ως ειδική περίπτωση. Συγκεκριμένα, θέτοντας $\xi_i = 0$ για όλα τα i στις Εξ. (10) και (11), αυτές απλοποιούνται στις αντίστοιχες μορφές για την περίπτωση των γραμμικά διαχωρίσιμων προτύπων [40].

ΚΕΦΑΛΑΙΟ 3 - ΥΛΟΠΟΙΗΣΗ

3.1 ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ – STATE OF THE ART

Η χρήση χαμηλού επιπέδου χαρακτηριστικών ψηφιοποιημένων έργων τέχνης με στόχο την εξαγωγή συμπερασμάτων δεν είναι κάτι πρωτόγνωρο. Αντίθετα, έχουν εκπονηθεί αρκετές εργασίες στο αντικείμενο και παρακάτω θα γίνει μια επισκόπηση στις εργασίες αυτές.

Μία εργασία που ασχολήθηκε με τη χρήση χαρακτηριστικών χαμηλού επιπέδου είναι αυτή των Shuqiang Jiang, Wen Gao και Weiqiang Wang [38]. Η εργασία αυτή προτείνει ένα σχέδιο για να ταξινομήσει παραδοσιακούς κινέζικους πίνακες ζωγραφικής. Ο αλγόριθμος χρησιμοποιεί τρία χαρακτηριστικά χαμηλού επιπέδου για να επιτευχθεί μία υψηλού επιπέδου ταξινόμηση. Τα χαρακτηριστικά που χρησιμοποιούνται είναι το Ιστόγραμμα Ohta, διάνυσμα συνοχής χρώματος (color coherence vector) και αυτοσυσχέτιση (auto-correlation). Στην συγκεκριμένη εργασία χρησιμοποίησαν 1254 κινέζικους πίνακες ζωγραφικής, χρησιμοποιήθηκε ο ταξινομητής SVM (Support Vector Machine) και επιτεύχθηκε 97.21% ακρίβεια.

Μία ακόμα εργασία είναι αυτή των Congcong Li και Tshuan Chen [3]. Η εργασία αυτή έχει ως στόχο την αξιολόγηση της αισθητικής οπτικής ποιότητας των ψηφιακών πινάκων ζωγραφικής. Η αξιολόγηση των πινάκων ζωγραφικής μπορεί να θεωρηθεί υποκειμενική. Ωστόσο, σε κάποιο βαθμό, ορισμένοι πίνακες θεωρούνται ότι έχουν υψηλότερη αισθητική ποιότητα από άλλους. Σε αυτή την εργασία, αντιμετώπισαν την πρόκληση αυτή ως πρόβλημα μηχανικής μάθησης, προκειμένου να αξιολογηθεί η αισθητική ποιότητα των πινάκων με βάση το οπτικό τους περιεχόμενο. Έτσι, σχεδιάστηκαν μια πληθώρα ομάδα μεθόδων εξαγωγής χαρακτηριστικών που αντιπροσωπεύουν τόσο τα καθολικά (global) χαρακτηριστικά όσο και τα τοπικά (local) χαρακτηριστικά ενός πίνακα. Έμπνευση για αυτά τα χαρακτηριστικά προήλθε από προηγούμενες γνώσεις τους στην τέχνη καθώς και μία έρευνα ερωτηματολογίου που πραγματοποιήθηκε για τη μελέτη παραγόντων που επηρεάζουν την ανθρώπινη κρίση.

Έτσι, συλλέχθηκαν πίνακες ζωγραφικής και ζήτησαν να αξιολογηθούν αυτοί οι πίνακες από μία ομάδα ανθρώπων. Στη συνέχεια, αυτοί οι πίνακες, χρησιμοποιήθηκαν και στην εκπαίδευση και στην δοκιμή των πειραμάτων τους. Τα αποτελέσματα του πειράματος δείχνουν ότι μπορούν

να ταξινομηθούν και οι πίνακες υψηλής ποιότητας και χαμηλής με απόδοση συγκρίσιμη με την ανθρώπινη κρίση. Αυτή η εργασία παρέχει ένα θέμα μηχανικής μάθησης για έρευνα σχετικά με τις αισθητικές αντιλήψεις του ανθρώπου και του υπολογιστή με βάση τα χαρακτηριστικά.

Η εργασία του Thomas Lombardi [6] προτείνει μία γενική προσέγγιση για την ταξινόμηση των καλλιτεχνικών ρευμάτων που υποστηρίζεται από τα ακόλουθα: αναγνώριση πινάκων ζωγραφικής με βάση το στυλ, προσδιόριση των βασικών σχέσεων μεταξύ των καλλιτεχνικών ρευμάτων, περιγραφή βάσει τον προσδιορισμό εγγύτητας του καλλιτεχνικού ρεύματος και την αξιολόγηση και την οπτικοποίηση των αποτελεσμάτων ταξινόμησης.

Ειδικότερα, η εργασία λαμβάνει υπόψη πολλά χρωματικά χαρακτηριστικά που δεν εφαρμόζονται στη ταξινόμηση των πινάκων, όπως γραφήματα αυτοσυσχέτισης χρώματος (color autocorrelograms) και δυναμικά χωρικά χρωματικά ιστογράμματα (dynamic spatial chromatic histograms). Η έρευνα για τα χαρακτηριστικά χρώματος έδειξε ότι η διατήρηση της συχνότητας και των χωρικών πληροφοριών του περιεχομένου χρώματος ενός πίνακα δεν βελτίωσε την ακρίβεια της ταξινόμησης. Ένας αλγόριθμος παλέτας προτείνεται για την περιγραφή του περιεχομένου χρώματος των πινάκων από τον χρωματικό χάρτη μιας εικόνας. Ο αλγόριθμος απέδωσε σε σύγκριση με άλλα χαρακτηριστικά χρώματος. Χρησιμοποιήθηκαν αρκετές εποπτευόμενες και μη εποπτευόμενες τεχνικές για ταξινόμηση, οπτικοποίηση και αξιολόγηση, συμπεριλαμβανομένου του K-NN(k – πλησιέστερος γείτονας). Οι μετρήσεις με βάση τα καλλιτεχνικά ρεύματα αποδείχθηκαν αξιόπιστες για την αξιολόγηση των αποτελεσμάτων.

Η εργασία των Jana Zujovic, Lisa Gandy, Scott Friedman, Bryan Pardo, Thrasyvoulos N. Pappas [7] περιγράφει μία προσέγγιση για την αυτόματη ταξινόμηση ψηφιακών πινάκων ως προς το καλλιτεχνικό ρεύμα. Η αυτόματη ταξινόμηση θεωρείται χρήσιμη για την οργάνωση μεγάλων ψηφιακών συλλογών. Στην συγκεκριμένη εργασία χρησιμοποιήθηκαν πίνακες ζωγραφικής από το διαδίκτυο και όχι επαγγελματικά δεδομένα υψηλής ανάλυσης. Περιλαμβάνεται μία σύγκριση με τις υπάρχουσες εξαγωγές χαρακτηριστικών και μεθόδους ταξινόμησης καθώς και μια ανάλυση της δικής τους προσέγγισης των ταξινομητών και των χαρακτηριστικών.

Στην εργασία των Guanming Lu, Zhong Gao, Danni Qin, Xin Zhao και Mengjue Liu [2] περιλαμβάνεται το πρόβλημα της επιτακτικής ανάγκης για την αποτελεσματική αναγνώριση και

ταξινόμηση των παραδοσιακών κινέζικων πινάκων ζωγραφικής. Η εργασία αυτή προτείνει ένα σύστημα αναγνώρισης με βάση το περιεχόμενο και ταξινόμηση που αντιπροσωπεύει το οπτικό περιεχόμενο των εικόνων TCP (Transmission Control Protocol) με χρωματικά χαρακτηριστικά και χαρακτηριστικά υφής. Χρησιμοποιήθηκαν 4 είδη ταξινομητών για να εκπαιδευτούν με τα βασικά χαρακτηριστικά του στυλ TCP, τα καλλιτεχνικά ρεύματα και τους ζωγράφους. Τα αποτελέσματα δείχνουν ότι ο αλγόριθμος είναι ικανός να αναγνωρίσει τις εικόνες TCP και να τις ταξινομήσει με βάση τους ζωγράφους και τα καλλιτεχνικά ρεύματα με ακρίβεια μεγαλύτερη από 85%.

Στην διπλωματική εργασία του Νικέλλη Αναστάσιου [21], δημιουργείται ένα σύστημα το οποίο επιχειρεί να ταξινομήσει εκατό πίνακες από δώδεκα διαφορετικούς δημιουργούς. Οι πίνακες, οι οποίοι είναι σε ψηφιοποιημένη μορφή, αναλύονται ώστε να εξαχθεί η χρωματική πληροφορία που μεταφέρουν και να είναι έτοιμη για χρήση από ένα νευρωνικό δίκτυο. Οι πίνακες χωρίζονται στο σύνολο εκπαίδευσης, το οποίο εκπαιδεύει το νευρωνικό δίκτυο και στο σύνολο δοκιμής, το οποίο ελέγχει την ικανότητα του νευρωνικού δικτύου να ταυτοποιήσει τους πίνακες.

Τέλος, η εργασία των Oguz Icoğlu, Bilge Gunsel, και Sanem Sariel [16] περιγράφει την αυτόματη εξαγωγή χαρακτηριστικών των πινάκων ζωγραφικής, με βάση τα καλλιτεχνικά ρεύματα όπως ο κλασικισμός, ο ιμπρεσιονισμός και ο κυβισμός. Προτείνονται 6 χαρακτηριστικά, που θα δούμε αναλυτικά παρακάτω όπου χρησιμοποιήθηκαν για τις ανάγκες αυτής της διπλωματικής εργασίας, για την αναπαράσταση περιεχομένου και φαίνεται ότι το σύνολο χαρακτηριστικών επιτρέπει την επισήμανση των καλλιτεχνικών ρευμάτων αποτελεσματικά. Οι ταξινομητές που εξετάστηκαν ήταν οι Bayesian (Μπεϋζιανός) ταξινομητής, k-NN και SVM. Το πείραμα επιτυγχάνεται με ποσοστό ακρίβειας άνω των 90%, η χαμηλότερη απόδοση παρατηρήθηκε από τον K-NN ταξινομητή.

3.2 ΜΕΘΟΔΟΛΟΓΙΑ

Η μεθοδολογία που υλοποιήθηκε για την ανάπτυξη της εργασίας είναι η εξής:

Πρώτα, αφού κατανοήσαμε το θέμα, αναζητήσαμε πληροφορίες σχετικά με τα καλλιτεχνικά ρεύματα και μελετήσαμε άρθρα σχετικά με την ταξινόμηση έργων τέχνης προκειμένου να διαμορφώσουμε μία γενική άποψη του θέματος.

Τα καλλιτεχνικά ρεύματα που χρησιμοποιήθηκαν είναι ο κλασικισμός, ο εξπρεσιονισμός, ο κυβισμός, ο ρεαλισμός και ο ιμπρεσιονισμός.

Στη συνέχεια αναζητήσαμε τη βάση δεδομένων από την οποία θα αντλήσουμε τα ψηφιοποιημένα έργα τέχνης. Αποφασίσαμε να χρησιμοποιήσουμε την ιστοσελίδα <https://www.wikiart.org/>, η οποία μας παρέχει μεγάλο αριθμό από έργα τέχνης.

Συγκεκριμένα, χρησιμοποιήθηκαν 250 ψηφιοποιημένοι πίνακες ζωγραφικής από 5 διαφορετικά καλλιτεχνικά ρεύματα, στους οποίους πραγματοποιήθηκε ισοκατανομή, δηλαδή 50 εικόνες από το καθέ καλλιτεχνικό ρεύμα.

Έπειτα μελετήσαμε σε βάθος εργασίες και βιβλία που αφορούσαν την ψηφιακή επεξεργασία εικόνας και την τεχνητή όραση ώστε να βρούμε τα κατάλληλα χαρακτηριστικά (features) τα οποία θα χρησιμοποιούσαμε.

Αφού επιλέχθηκαν τα παραπάνω, έγινε και η επιλογή κατάλληλου προγραμματιστικού περιβάλλοντος στο οποίο θα υλοποιόταν η εφαρμογή. Το λογισμικό που κρίθηκε πιο κατάλληλο για την υλοποίηση της εργασίας είναι το Matlab.

Στη συνέχεια καταγράψαμε τον κώδικα με τον οποίο επεξεργαστήκαμε τους πίνακες για να συλλέξουμε τα διάφορα χαρακτηριστικά που επιλέχθηκαν, τα εκπαιδεύσαμε με την κατάλληλη εφαρμογή του Matlab Classification Learner, όπου μας παρέχει πλήθος αλγόριθμων ταξινόμησης.

Αφού πραγματοποιήθηκαν τα πειράματα, διερευνήθηκε και ποιος ταξινομητής είχε καλύτερο ποσοστό αποτελεσμάτων.

Τέλος, πραγματοποιήθηκε η συγγραφή της διπλωματικής εργασίας.

Συνοψίζοντας τα βήματα πραγματοποίησης της διπλωματικής εργασίας είναι:

- Διατύπωση και αποσαφήνιση του θέματος
- Συλλογή ψηφιακών εικόνων
- Μελέτη για την επεξεργασία εικόνας και μηχανική όραση
- Μελέτη για την αναγνώριση προτύπων

- Μελέτη προηγούμενων εργασιών
- Επιλογή προγραμματιστικού περιβάλλοντος
- Επιλογή χαρακτηριστικών
- Επεξεργασία δεδομένων
- Εκπόνηση πειραμάτων
- Συγγραφή Διπλωματικής Εργασίας

3.3 ΚΑΛΛΙΤΕΧΝΙΚΑ ΡΕΥΜΑΤΑ

Ένα καλλιτεχνικό ρεύμα είναι μια τάση ή στυλ τέχνης με μια συγκεκριμένη κοινή φιλοσοφία ή στόχο, που ακολουθεί μια ομάδα καλλιτεχνών σε μια συγκεκριμένη χρονική περίοδο (συνήθως μήνες, χρόνια ή δεκαετίες) ή τουλάχιστον ένα καθορισμένο κίνημα μέσα σε λίγα χρόνια . Τα καλλιτεχνικά ρεύματα είναι ιδιαίτερα σημαντικά στη σύγχρονη τέχνη και κάθε διαδοχικό κίνημα θεωρείται ως ένα νέο κίνημα πρωτοπορίας [8].

Τα βασικά χαρακτηριστικά που χρησιμοποιούν οι ειδικοί για να ξεχωρίσουν τα καλλιτεχνικά ρεύματα είναι η θεματολογία, η χρήση του φωτός, η χρησιμοποιούμενη χρωματική παλέτα, η τεχνική [32].

Τα καλλιτεχνικά ρεύματα που επιλέχθηκαν είναι τα εξής:

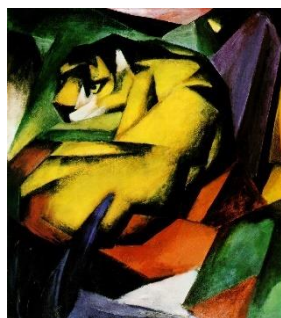
1. **Κλασικισμός**, ο οποίος γνώρισε μεγάλη άνθιση κατά τον 17^ο αιώνα. Εκφράστηκε σε όλες τις μορφές της τέχνης: την αρχιτεκτονική, τη μουσική, τη ζωγραφική, το θέατρο και τη λογοτεχνία [32]. Ο κλασικισμός ήταν μια ξεχωριστή τάση στην μπαρόκ ευρωπαϊκή ζωγραφική, ένα στυλ τέχνης στο οποίο η τήρηση αποδεκτών αισθητικών ιδεών υπερισχύει της ατομικότητας της έκφρασης. Με απλά λόγια, ήταν ένα συγκρατημένο, αρμονικό στυλ που πίστευε στην υπεροχή του σχεδιασμού, παρά στο χρώμα ή στην έκφραση. Σε γενικές γραμμές, η αισθητική του κλασικισμού ασχολήθηκε κυρίως με τη συνολική ισορροπία και αρμονία της σύνθεσης και όχι με λεπτομέρειες. Αν και η κλασική ζωγραφική είναι αρκετά ρεαλιστική, τα μεμονωμένα χαρακτηριστικά και οι υφές συχνά απλοποιούνται, δεδομένου ότι είναι δευτερεύουσας σημασίας. Κατά συνέπεια, τα στοιχεία μιας κλασικής σύνθεσης (π.χ. ανθρώπινα σώματα και πρόσωπα, χαρακτηριστικά τοπίου) έχουν παρόμοια, γενική εμφάνιση

καθώς είναι εξιδανικευμένα (δηλαδή, καθώς οι παράτυπες και ασυνήθιστες πτυχές του πραγματικού κόσμου εξομαλύνονται) [8]. Βασικός εκπρόσωπος ο Νικολά Πουσσέν.



Εικόνα 14 - The Triumph of Flora του Nicolas Poussin [8]

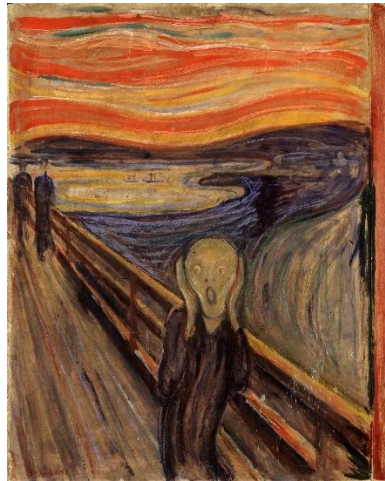
2. **Κυβισμός**, ο οποίος είναι ένα καλλιτεχνικό ρεύμα της ζωγραφικής και της γλυπτικής, στην Ευρώπη του 20^{ου} αιώνα, στον οποίο τα αντικείμενα χωρίζονται, αναλύονται, και συνθέτονται ξανά σε μια αφηρημένη μορφή [32], αντί να απεικονίζουν αντικείμενα από μία μόνο άποψη, ο καλλιτέχνης απεικονίζει το θέμα από μια πληθώρα απόψεων για να αντιπροσωπεύσει το θέμα σε ένα ευρύτερο πλαίσιο. Είναι χαρακτηριστική η παρουσίαση του θέματος με γεωμετρικούς όγκους, επίπεδα και γωνιώδεις φόρμες [8]. Διάσημος εκπρόσωπός της είναι ο Πάμπλο Πικάσο.



Εικόνα 15 – The Tiger του Franz Marc [8]

3. **Εξπρεσιονισμός**, όπου αποτελεί καλλιτεχνικό κίνημα της μοντέρνας τέχνης που αναπτύχθηκε στις αρχές του 20ού αιώνα, περίπου την περίοδο 1905-1940 και κυρίως

στο χώρο της ζωγραφικής. Η κλασική φάση του εξπρεσιονιστικού κινήματος εξαπλώθηκε σε όλη την Ευρώπη. Βασικό χαρακτηριστικό των εξπρεσιονιστών καλλιτεχνών ήταν η τάση να παραμορφώνουν την πραγματικότητα στα έργα τους, αδιαφορώντας απέναντι σε μια πιστή και αντικειμενική αναπαράσταση της. Συχνά ο εξπρεσιονισμός διακρίνεται και από μια έντονη συναισθηματική αγωνία, χαρακτηριστικά μάλιστα μπορούμε να πούμε πως ελάχιστα εξπρεσιονιστικά έργα έχουν χαρούμενη διάθεση [33]. Χαρακτηριστικός εκπρόσωπος ο Καντίνσκι.



Εικόνα 16 – The Scream του Edvard Munch [8]

4. **Ιμπρεσιονισμός**, αναπτύχθηκε στο δεύτερο μισό του 19ου αιώνα. Αν και αρχικά καλλιεργήθηκε στο χώρο της ζωγραφικής, επηρέασε τόσο τη λογοτεχνία, τον κινηματογράφο όσο και τη μουσική. Κύριο χαρακτηριστικό του ιμπρεσιονισμού στη ζωγραφική είναι τα ζωντανά χρώματα (κυρίως με χρήση των βασικών χρωμάτων), οι συνθέσεις σε εξωτερικούς χώρους, συχνά υπό ασυνήθιστες οπτικές γωνίες και η έμφαση στην αναπαράσταση του φωτός. Οι ιμπρεσιονιστές ζωγράφοι θέλησαν να αποτυπώσουν την άμεση εντύπωση (impression) που προκαλεί ένα αντικείμενο ή μια καθημερινή εικόνα [34]. Χαρακτηριστικός εκπρόσωπος είναι ο Κλωντ Μονέ.



Εικόνα 17 – On the beach του Eugene Boudin [8]

5. **Ρεαλισμός**, που είναι ένα Καλλιτεχνικό ρεύμα της ζωγραφικής την περίοδο 1850-1880. Στα μέσα του 19ου αιώνα ορισμένοι καλλιτέχνες αρχίζουν να απορρίπτουν τον αισθηματισμό του Ρομαντισμού και επιδιώκουν να ξαν' απεικονίσουν τη ζωή με ρεαλιστικό - πραγματικό τρόπο. Οι ζωγράφοι και οι λογοτέχνες του ρεαλισμού προσπαθούσαν να απεικονίσουν τη ζωή όπως ήταν, δεν επιδίωκαν να ξεφύγουν από την πραγματικότητα [35]. Κύριος εκπρόσωπος θεωρείται ο ζωγράφος Γκουστάβ Κουρμπέ.



Εικόνα 18 - Moret sur Loing, the Bridge and the Church της Jean-Baptiste Camille Corot [8]

Ο κύριος λόγος που επιλέχθηκαν τα συγκεκριμένα καλλιτεχνικά ρεύματα ήταν ο αριθμός έργων τέχνης που μπορούσαμε να χρησιμοποιήσουμε. Σε κάθε ένα από τα παραπάνω καλλιτεχνικά ρεύματα επιλέχθηκαν 50 πίνακες ζωγραφικής. Είναι χρήσιμο ένα σύνολο δεδομένων να είναι ισορροπημένο, οι κλάσεις δηλαδή να περιέχουν ίσο αριθμό δεδομένων. Αν

τα δεδομένα δεν είναι ισορροπημένα τότε ο ταξινομητής έχει μια τάση να προβλέπει πολύ συχνότερα τις κλάσεις που έχουν περισσότερα δεδομένα. Τα συγκεκριμένα ρεύματα που επιλέχθηκαν διέθεταν όλα (τουλάχιστον) 50 πίνακες στη συλλογή Wikiart.

Οι 250 πίνακες προέρχονται από την ιστοσελίδα του WikiArt, η οποία είναι μια διαδικτυακή εγκυκλοπαίδεια για πίνακες ζωγραφικής ανά τον κόσμο. Πιο συγκεκριμένα, το Wikiart είναι μία από τις δημοφιλέστερες και πληρέστερες ψηφιακές εγκυκλοπαίδειες εικαστικών τεχνών που διαθέτει πάνω από 250.000 ψηφιοποιημένα έργα τέχνης από 3000 καλλιτέχνες.

3.4 ΛΟΓΙΣΜΙΚΟ ΠΟΥ ΕΠΙΛΕΧΘΗΚΕ - MATLAB

3.4.1 Τι είναι το Matlab

Το Matlab είναι ένα πρόγραμμα σχεδιασμένο για όσους χρησιμοποιούν αριθμητικούς υπολογισμούς. Ξεκίνησε ως ένα πρόγραμμα “Εργαστηρίου Πινάκων” (“MATrix LABoratory”) και έχει αναπτυχθεί αρκετά, ώστε να είναι ένα ισχυρό εργαλείο για την εφαρμογή προσομοίωσης και μοντελοποίησης συστημάτων, προγραμματισμού, έρευνας, στην επιστήμη των μηχανικών και των επικοινωνιών [18].

Η ισχυρή, τεχνική γλώσσα που χρησιμοποιεί είναι συμπαγής και περιγραφική και επιτρέπει τη μοντελοποίηση διαφόρων συστημάτων με χρήση κώδικα.

Ο χρήστης, στο φιλικό περιβάλλον του προγράμματος, έχει τη δυνατότητα να δει τη λειτουργία και τις ιδιότητες διαφόρων σημάτων και συστημάτων, να παρεμβαίνει, να εισάγει παραμέτρους που να επηρεάζουν και καθορίζουν τον τρόπο λειτουργίας του συστήματος, ώστε να μπορεί να ερευνά, να ανακαλύψει και να κατανοήσει θέματα που αφορούν το αντικείμενο της Ψηφιακής Επεξεργασίας Εικόνας και Τεχνητής Όρασης.

Συγκεκριμένα, το περιβάλλον του Matlab περιέχει ένα σύνολο εργαλείων και βοηθητικών προγραμμάτων για την ανάπτυξη εφαρμογών επεξεργασίας και ανάλυσης εικόνας όπως η δυνατότητα εισαγωγής εικόνων και video από αρχεία και κάμερες, έτοιμες εικόνες καθώς και δυνατότητα για την ανάπτυξη αλγορίθμων και γραφικών εργαλείων.

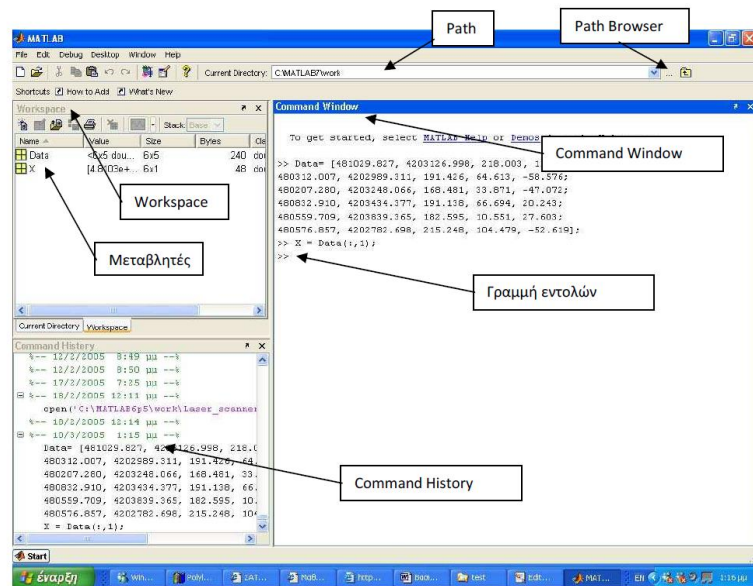
Το Matlab εμπλουτίζεται με πάνω από 80 εξειδικευμένες βιβλιοθήκες συναρτήσεων και εργαλείων (εργαλειοθήκες) για διάφορες εφαρμογές και αναλύσεις. Τα κύρια στοιχεία της επεξεργασίας εικόνας είναι:

- Εργαλειοθήκη επεξεργασίας εικόνας
- Εργαλειοθήκη απόκτησης εικόνας
- Εργαλειοθήκη Χαρτογράφησης
- Εργαλειοθήκη Επεξεργασίας Σήματος
- Εργαλειοθήκη Σχεδιασμού Φίλτρων
- Εργαλειοθήκη Ανάλυσης Κυμάτων
- Εργαλειοθήκη Βελτιστοποίησης
- Εργαλειοθήκη Στατιστικών

3.4.2 Περιβάλλον Εργασίας

Το περιβάλλον εργασίας του **Matlab** περιλαμβάνει κάποια παράθυρα, τα οποία φαίνονται στην παρακάτω εικόνα.

- Το **Command Window** είναι το παράθυρο που περιέχει τη γραμμή εντολών. Εκεί εισάγουμε τις εντολές που θέλουμε να εκτελέσουμε.
- Το **Workspace** είναι ο χώρος εργασίας στον οποίο αποθηκεύονται όλες οι μεταβλητές και οι πίνακες που δημιουργούνται μέσω των εντολών που δίνουμε στη γραμμή εντολών.
- Το **Command History**, όπου φαίνονται οι εντολές που έχουν εκτελεστεί από τη στιγμή που έγινε η εκκίνηση του προγράμματος, αλλά και εντολές που δόθηκαν κατά την προηγούμενη χρήση του προγράμματος.
- Επίσης, στο επάνω μέρος φαίνεται η γραμμή του **Path** που δείχνει τη διαδρομή του directory που βλέπει το Matlab, καθώς και κουμπί που οδηγεί στον **Path Browser**.
- Εκτός από τα παράθυρα αυτά υπάρχει και ο editor του Matlab, όπου γίνεται η συγγραφή του κώδικα των προγραμμάτων.



Εικόνα 19 - Το περιβάλλον εργασίας του Matlab

Το Matlab προσφέρει δύο εναλλακτικούς τρόπους εργασίας. Είτε μέσω του Command Window όπου ο χρήστης δίνει μεμονωμένες εντολές, μπορεί να κάνει κάποιες δοκιμές και παίρνει άμεσα κάποια αποτελέσματα, είτε μέσω scripts, προγραμμάτων δηλαδή, τα οποία ο χρήστης γράφει στο περιβάλλον του Editor.

Τέλος οι functions αποτελούν και αυτές M-Files, δηλαδή έχουν κατάληξη .m, και πρέπει να αποθηκεύονται στον ενεργό φάκελο. Η διαφορά των functions σε σχέση με τα scripts είναι ότι οι functions παίρνουν εισόδους (ορίσματα) και επιστρέφουν εξόδους. Μία function ξεκινάει πάντα (στη πρώτη γραμμή του editor) με την εξής έκφραση:

$$function[y_1, y_2, \dots, y_n] = o\sigma\mu\alpha([x_1, x_2, \dots, x_n].$$

Οι y_1, y_2, \dots, y_n είναι οι εξόδοι που επιστρέφει η function και οι x_1, x_2, \dots, x_n είναι οι εισόδοι (ορίσματα) που δέχεται [18].

3.5 ΠΕΡΙΓΡΑΦΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

Για την ταξινόμηση των πινάκων επιλέχθηκαν 40 χαρακτηριστικά χαμηλού επιπέδου. Τα χαρακτηριστικά χαμηλού επιπέδου αφορούν επιμέρους στοιχεία της εικόνας, όπως γραμμές,

ακμές, χρώματα κλπ. Σε αντίθεση με τα υψηλού επιπέδου χαρακτηριστικά που «χτίζονται» πάνω στα χαμηλού επιπέδου χαρακτηριστικά και αφορούν σχήματα, αντικείμενα, περιεχόμενο κλπ. Επίσης τα χαρακτηριστικά που χρησιμοποιήθηκαν είναι καθολικά (global) αφού για την εξαγωγή τους χρησιμοποιήθηκε ολόκληρος ο πίνακας ζωγραφικής χωρίς να γίνει κατάτμηση αυτού σε μικρότερα μέρη (τοπικά χαρακτηριστικά).

Τα διαφορετικά καλλιτεχνικά ρεύματα διαθέτουν ένα σύνολο κοινών ιδιοτήτων που μπορούν να αποτελέσουν τη βάση για την επιλογή των χαρακτηριστικών που θα είναι κατάλληλα για τη διάκριση του καλλιτεχνικού στυλ. Για παράδειγμα, με βάση την παρατήρηση ότι κάποια καλλιτεχνικά ρεύματα χαρακτηρίζονται από την προτίμηση χρωμάτων «σκούρων» ενώ άλλα χαρακτηρίζονται από τη χρήση «φωτεινών» και «ζωηρών» χρωμάτων, οι Oguz Icoглу, Bilge Günsel, και Sanem Sariel προτείνουν ως χαρακτηριστικό το ποσοστό σκούρων χρωμάτων σε ένα πίνακα.

3.5.1 Χαρακτηριστικά προτεινόμενα από τους Icoğlu, Günsel και Sariel

Αρχικά υιοθετήσαμε 6 διαφορετικά χαρακτηριστικά για την κατηγοριοποίηση των πινάκων, όπως αυτά περιγράφονται από τους Oguz Icoğlu, Bilge Günsel, και Sanem Sariel σε μία μελέτη που πραγματοποίησαν. [16]

Τα χαρακτηριστικά που προτείνουν οι Oguz Icoğlu, Bilge Günsel, and Sanem Sariel είναι τα ακόλουθα:

μ_1 : Ποσοστό σκούρων χρωμάτων.

μ_2 : Συντελεστής κλίσης που υπολογίζεται από τον χάρτη κλίσης της εικόνας.

μ_3 : Αριθμός μεγίστων στο ιστογράμμα φωτεινότητας.

μ_4 : Η χρωματική περιοχή που αντιστοιχεί στο σημείο κορυφής του ιστογράμματος της φωτεινότητας.

μ_5 : Μετά τη διαίρεση της εικόνας σε πανομοιότυπα κομμάτια, η απόκλιση του μέσου επιπέδου γκρι που αποκτάται σε κάθε κομμάτι από το μέσο επίπεδο γκρι που αποκτάται σε ολόκληρη την εικόνα.

μ_6 : Ασυμμετρία «Skew», η απόκλιση της κατανομής του γκρι επιπέδου από την κατανομή Gauss.

Οι μαθηματικές διατυπώσεις των πρώτων 6 χαρακτηριστικών είναι οι εξής:

1. Το στοιχείο φωτεινότητας (luminance) μιας έγχρωμης εικόνας που ορίζεται στον χώρο RGB υπολογίζεται από την παρακάτω εξίσωση:

$$Y = 0,299 R + 0,587 G + 0,114 B.$$

Υπάρχει πολύ υψηλή σχέση μεταξύ του στοιχείου φωτεινότητας και του επιπέδου γκρι στην εικόνα. Στην εργασία, το στοιχείο φωτεινότητας αντιπροσωπεύεται με 8 bit (επίπεδο γκρι 256) και τα pixel των οποίων η τιμή της φωτεινότητας αντιστοιχεί στο εύρος [0,64], θεωρούνται ως σκούρα εικονοστοιχεία. Το ποσοστό των σκούρων χρωμάτων ορίζεται ως:

$$\mu_1 = \frac{\text{Αριθμός σκούρων εικονοστοιχείων στην εικόνα}}{\text{Αριθμός όλων των εικονοστοιχείων της εικόνας}}$$

2. Ο Συντελεστής κλίσης ορίζεται ως:

$$\mu_2 = \text{Normalized} \left(\sum_{i=1}^r \sum_{j=1}^c \sqrt{(f_{ij})_x^2 + (f_{ij})_y^2} \right)$$

όπου το r και το c είναι το πλήθος σειρών και στηλών της εικόνας αντίστοιχα, και $(f_{ij})_x$ και $(f_{ij})_y$ είναι οι παράγωγοι πρώτης τάξης των εικονοστοιχείων (ij) της εικόνας σε κατευθύνσεις x και y αντίστοιχα.

Ειδικά, το κίνημα του «κλασικισμού» μπορεί να ταξινομηθεί αρκετά με ακρίβεια χρησιμοποιώντας αυτά τα δύο χαρακτηριστικά: μ_1 και μ_2 . Ωστόσο, αυτά τα χαρακτηριστικά, από μόνα τους, είναι ανεπαρκή για την ταξινόμηση άλλων κινημάτων τέχνης. Επιπλέον, η αλλαγή των συνθηκών φωτισμού και της ανάλυσης των εικόνων ζωγραφικής μπορεί να προκαλέσει ψευδή ταξινόμηση. Για την επίλυση τέτοιων προβλημάτων, ορίζονται πρόσθετα χαρακτηριστικά μέσα στο σύστημα:

3. Ο αριθμός μεγίστων στο ιστόγραμμα φωτεινότητας:

$$\mu_3 = \text{Number}(|H(i)| > thr),$$

όπου thr είναι η τιμή κατωφλίου που προσδιορίζεται για να αντιστοιχηθεί στο πλάτος του ιστογράμματος στο μέγιστο.

4. Η χρωματική περιοχή η οποία αντιστοιχεί στο σημείο κορυφής του ιστογράμματος της φωτεινότητας:

$$\mu_4 = \max_i |H(i)|, \quad i = 0, 1, \dots, 255,$$

όπου $H(i)$ είναι η τιμή του ιστογράμματος για το i επίπεδο του γκρι.

Η απόδοση ταξινόμησης έργων τέχνης που ανήκουν στον ιμπρεσιονισμό επαυξάνεται με τη συμπερίληψη αυτών των δύο χαρακτηριστικών (μ_3 και μ_4). Ο κύριος λόγος αυτής της αύξησης είναι ότι οι πίνακες του ιμπρεσιονισμού παράγουν ένα μέγιστο στις φωτεινές περιοχές του ιστογράμματος της φωτεινότητας. Στους κλασικιστικούς πίνακες, το μέγιστο σημείο εμφανίζεται μέσα στις σκοτεινές περιοχές και στους κυβιστικούς πίνακες, περισσότερα από ένα μέγιστα σημεία εμφανίζονται σε διάφορες περιοχές του ιστογράμματος.

5. Από την άλλη, περιλαμβάνονται τα παρακάτω χαρακτηριστικά για τη διατήρηση της απόδοσης του συστήματος ενάντια στις μεταβαλλόμενες συνθήκες φωτισμού και τις διαφορές στην ανάλυση:

$$\mu_5 = \sum_{i=1}^9 (mean_i - mean),$$

όπου το $mean_i$ αντιπροσωπεύει την τιμή φωτεινότητας του i -στου κομματιού. Ενώ το $mean$ είναι τη μέση τιμή φωτεινότητας ολόκληρης της εικόνας και υπολογίζεται από τη παρακάτω εξίσωση:

$$mean = \frac{\sum_{i=1}^r \sum_{j=1}^c f_{ij}}{rc}.$$

Στον υπολογισμό του μ_5 , οι πίνακες στο εκπαιδευτικό σετ χωρίζονται σε 9 πανομοιότυπα κομμάτια, καθένα από τα οποία διατηρεί την πρωτότυπη αναλογία των διαστάσεων του πίνακα. Για κάθε κομμάτι, η μέση τιμή χρώματος υπολογίζεται σε επίπεδο γκρι. Η απόκλιση αυτών των 9 τιμών από τη μέση τιμή χρώματος που υπολογίζεται για ολόκληρη την εικόνα δίνει το μ_5 , «Χαρακτηριστικό Τυπικής Απόκλισης της μέσης τιμής». Παράλληλα, σε πίνακες ιμπρεσιονισμού ή κλασικισμού, αυτό το χαρακτηριστικό δίνει υψηλές τιμές λόγω της τοπικής φωτεινότητας ή σημασιολογικές αλλαγές χρώματος (ανθρώπινο πρόσωπο, ουρανός... κλπ), δεν υπερβαίνει μια συγκεκριμένη τιμή για

κυβιστικούς πίνακες, αφού κάθε υποτήμημα δεν παρουσιάζει μεγάλη απόκλιση από τη συνολική μέση τιμή. Καθώς το μ_5 είναι το κριτήριο για την απόκλιση των μέσων επιπέδων του γκρι των κομματιών από ολόκληρη την εικόνα, η τιμή της δεν επηρεάζεται από αλλαγές της ανάλυσης.

6. Το έκτο χαρακτηριστικό, μ_6 , το κριτήριο της απόκλισης του επιπέδου γκρι της εικόνας από την κατανομή Gauss ορίζεται ως ακολουθεί:

$$\mu_6 = skew = \frac{\sum_{i=1}^r \sum_{j=1}^c \left(\frac{f_{ij} - mean}{\sigma}\right)^3}{rc}$$

όπου το σ ορίζεται στην παρακάτω εξίσωση.

$$\sigma = \sqrt{\frac{\sum_{i=1}^r \sum_{j=1}^c (f_{ij} - mean)^3}{rc - 1}}$$

Η ασυμμετρία δίνει ένα μέτρο της ασυμμετρίας του ιστογράμματος γύρω από τη μέση τιμή. Για συμμετρική κατανομή η ασυμμετρία είναι 0. [19]

Το χαρακτηριστικό Ασυμμετρίας "Skew" διατηρεί την αξία του, ακόμη και όταν η φωτεινότητα αυξάνεται δραστικά και βοηθά στην επιτυχή μοντελοποίηση της κατάστασης στα κλασικιστικά κινήματα όπου η χρωματική κατανομή εμφανίζει κάτι διαφορετικό από αυτό της κατανομής Gauss. Πίνακες ζωγραφικής του κυβιστικού κινήματος, γενικά, καταδεικνύουν ένα χαρακτηριστικό που ταιριάζει με την κατανομή Gauss. Ωστόσο, στον ιμπρεσιονισμό, η κατανομή χάνει ελαφρώς αυτή την ιδιότητα και εξαφανίζεται αρκετά στο κλασικιστικό κίνημα.

3.5.2 Χαρακτηριστικά υφής Haralick

Η υφή παίζει σημαντικό ρόλο στην ταξινόμηση των εικόνων και είναι σημαντικό να αναπτυχθούν χαρακτηριστικά σχετιζόμενα με την υφή. Τα επόμενα 14 χαρακτηριστικά που χρησιμοποιήθηκαν στην εργασία είναι από το άρθρο του Haralick [17] ο οποίος προτείνει την εύρεση χαρακτηριστικών υφής με στατιστικής φύσης υπολογισμούς στο χωρικό πεδίο. Θεωρεί

ότι η υφή μίας εικόνας μπορεί να αποκαλυφθεί από τη συνολική ή μέση χωρική συσχέτιση των εντάσεων του γκρι στην εικόνα.

Ο Haralick στο άρθρο του χρησιμοποιεί τον πίνακα συνεμφάνισης γκριζου επιπέδου (Grey Level Co-occurrence Matrix GLCM) [19]. Η μέθοδος GLCM είναι ένα στατιστικό μέτρο που αποτυπώνει τη χωρική κατανομή των επιπέδων του γκριζου. Υπολογίζει τις υπό συνθήκη κοινές πιθανότητες δύο ζευγαριών του γκρι να εμφανιστούν δίνοντας δύο παραμέτρους, την απόσταση μεταξύ των εικονοστοιχείων ,d, και τον προσανατολισμό –γωνία, θ. Οι δυνατές γωνίες είναι τέσσερις: 0°, 45°, 90°, και 135° .

Για παράδειγμα,

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

a) 4x4 πίνακας με 4 τιμές από το 0-3 επιπέδου του γκρι

		Grey Tone			
		0	1	2	3
Grey Tone	0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
	1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
	2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
	3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

b) Γενική μορφή πίνακα συνεμφάνιση γκριζου τόνου

$$0^\circ \quad P_H = \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

$$135^\circ \quad P_{LD} = \begin{pmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

$$90^\circ \quad P_V = \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}$$

$$45^\circ \quad P_H = \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

c) Υπολογισμός και των 4 πινάκων αποστάσεων συνεμφάνισης με γκρι τόνους

Σύμφωνα με το παραπάνω παράδειγμα πινάκων συνεμφάνισης, αρχικά στον πίνακα (a) παρουσιάζουν ένα πίνακα 4x4 με 4 τόνους του γκρι, από το 0 έως το 3. Στον πίνακα (b) δείχνει τη γενική μορφή οποιουδήποτε πίνακα συνεμφάνισης γκριζου τόνου. Για παράδειγμα, το στοιχείο στη θέση (2,1), της απόστασης $d=1$ του οριζόντιου πίνακα PH που είναι ο συνολικός αριθμός των 2 τόνων του γκρι των τιμών 2 και 1, που εμφανίστηκαν γειτονικά οριζόντια το ένα με το άλλο. Στους πίνακες (c) υπολογίζουμε τις 4 αποστάσεις 1 των πινάκων συνεμφάνισης με γκρι τόνους.

Η κατάλληλη συχνότητα κανονικοποίησης για τους πίνακες είναι εύκολα υπολογίσιμη. Όταν η σχέση με τον πλησιέστερο οριζόντιο γείτονα ($d=1, \theta=0^\circ$), θα υπάρχουν $2(N_x - 1)$ γειτονικά ζεύγη κελιών ανάλυσης σε κάθε σειρά και υπάρχουν N_y σειρές, που μας δίνουν συνολικά $2N_y(N_x - 1)$ πλησιέστερα οριζόντια ζεύγη γειτόνων. Όταν είναι ο πλησιέστερος δεξιά-διαγώνιος γείτονας ($d=1, \theta=45^\circ$), θα υπάρχουν $2(N_x - 1)$ 45° γειτονικά ζεύγη κελιών ανάλυσης σε κάθε σειρά εκτός από την πρώτη, για την οποία δεν υπάρχει κανένα ζεύγος, και υπάρχουν N_y σειρές. Αυτό μας δίνει συνολικά $2(N_y - 1)(N_x - 1)$ πλησιέστερα δεξιά-διαγώνια ζεύγη γειτόνων. Κατά συμμετρία, θα υπάρχει $2 N_x(N_y - 1)$ πλησιέστερα κατακόρυφα γειτονικά ζεύγη και $2(N_x - 1)(N_y - 1)$ πλησιέστερα αριστερά-διαγώνια γειτονικά ζεύγη.

Με βάση τον πίνακα συνεμφάνισης ο Haralick ορίζει 14 χαρακτηριστικά υφής.

Προκειμένου να παρουσιάσουμε τα Χαρακτηριστικά Υφής Στατιστικής Δεύτερης Τάξης θα πρέπει πρώτα να δώσουμε κάποιους θεμελιώδεις ορισμούς και συμβολισμούς για τους τύπους που παρουσιάζονται. Συγκεκριμένα:

- $p(i,j)$ η τιμή στη θέση (i, j) σε έναν κανονικοποιημένο GLCM, $= P(i,j)/R$. (R το συνολικό πλήθος των τιμών)
- $p_x(i)$ προκύπτει αθροίζοντας τις σειρές του $p(i, j) = \sum_{j=1}^{N_g} P(i, j)$.
- N_g ο αριθμός διακριτών επιπέδων γκρι στη κβαντισμένη εικόνα.

$$p_y(j) = \sum_{i=1}^{N_g} p(i, j).$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{i+j=k}^{N_g} p(i, j), \quad k=2,3,\dots,2N_g$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{|i-j|=k}^{N_g} p(i, j), \quad k=0,1,\dots,N_g-1.$$

Άρα τα χαρακτηριστικά υφής είναι:

1. Γωνιακή ροπή 2ης τάξης, που είναι μέτρο ομοιογένειας (Angular Second Moment):

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2.$$

2. Αντίθεση (Contrast):

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{i=1}^{N_g} p(i, j) \right\}.$$

3. Συσχέτιση (Correlation):

$$f_3 = \frac{\sum_i \sum_j (i, j) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y},$$

Όπου το μ_x , μ_y , σ_x , σ_y είναι τα μέσα και οι τυπικές αποκλίσεις των p_x και p_y .

4. Άθροισμα τετραγώνων: Διακύμανση (Sum of Square: Variance) :

$$f_4 = \sum_i \sum_j (i - \mu)^2 p(i, j).$$

5. Ροπή αντίστροφης διαφοράς (Inverse Difference Moment):

$$f_5 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j).$$

6. Μέσος όρος άθροισης (Sum Average):

$$f_6 = \sum_{i=2}^{2N_g} i p_{x+y}(i).$$

7. Άθροισμα Διακύμανσης (Sum Variance):

$$f_7 = \sum_{i=2}^{2N_g} (i - f_6)^2 p_{x+y}(i).$$

8. Άθροισμα Εντροπίας (Sum Entropy):

$$f_8 = - \sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\}.$$

9. Εντροπία (Entropy):

$$f_9 = - \sum_i \sum_j p(i, j) \log(p(i, j)).$$

10. Διαφορά Διακύμανσης (Difference Variance):

$$f_{10} = \text{variance of } p_{x-y}.$$

11. Διαφορά Εντροπίας (Difference Entropy):

$$f_{11} = - \sum_{i=0}^{N_g-1} p_{x-y}(i) \log\{p_{x-y}(i)\}.$$

12. 13. Πληροφοριακά μέτρα συσχέτισης (Information Measures of Correlation):

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2}$$

$$HXY = - \sum_i \sum_j p(i, j) \log(p(i, j))$$

Όπου HX και HY είναι οι εντροπίες των p_x και p_y , και

$$HXY1 = - \sum_i \sum_j p(i, j) \log\{p_x(i)p_y(j)\}$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(j) \log\{p_x(i)p_y(j)\}$$

14. Μέγιστος Συντελεστής Συσχέτισης (Maximal Correlation Coefficient):

$$f_{14} = (\text{Δεύτερη μεγαλύτερη ιδιοτιμή του } Q)^{1/2}$$

$$\text{Όπου } Q(i, j) = \sum_k \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)} p(i, j).$$

Αυτά τα μέτρα συσχέτισης έχουν κάποιες επιθυμητές ιδιότητες που δεν αναδεικνύονται στον ορθογώνιο συσχετισμό f_3 .

3.5.3 Ανίχνευση ακμών – Φίλτρο Sobel

Το επόμενο χαρακτηριστικό είναι σχετικά με την ανίχνευση ακμών των εικόνων. Χρησιμοποιήσαμε το φίλτρο Sobel για την ανίχνευση ακμών. Το φίλτρο Sobel υπολογίζει την διαβάθμιση έντασης της εικόνας σε κάθε pixel μέσα στην εικόνα [22]. Βρίσκει την κατεύθυνση της μεγαλύτερης αύξησης από φως προς το σκοτάδι και το ρυθμό αλλαγής προς αυτή την κατεύθυνση. Το αποτέλεσμα δείχνει πόσο απότομα ή ομαλά αλλάζει η εικόνα στο καθένα pixel, και επομένως πόσο πιθανό είναι αυτό το pixel να αντιπροσωπεύει μια ακμή. Δείχνει επίσης πώς είναι πιθανό να προσανατολιστεί αυτή η ακμή. Το αποτέλεσμα της εφαρμογής του φίλτρου σε ένα εικονοστοιχείο σε μια περιοχή σταθερής έντασης είναι ένα μηδενικό διάνυσμα. Το αποτέλεσμα της εφαρμογής του σε ένα εικονοστοιχείο σε μια ακμή είναι ένα διάνυσμα που δείχνει σε όλη την ακμή από πιο σκούρες σε φωτεινότερες τιμές.

Το φίλτρο Sobel χρησιμοποιεί δύο πίνακες 3x3. Ένα πίνακα για αλλαγές στην οριζόντια κατεύθυνση και ένα για αλλαγές στην κατακόρυφη κατεύθυνση. Οι δύο πυρήνες ενώνονται με την αρχική εικόνα για τον υπολογισμό προσεγγίσεων των παραγώγων. Αν ορίσουμε το G_x και το G_y ως δύο εικόνες που περιέχουν τις οριζόντιες και κατακόρυφες προσεγγίσεις παραγώγων αντίστοιχα, οι υπολογισμοί είναι:

$$G_x = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} * A \quad \text{και} \quad G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} * A$$

Όπου A είναι η αρχική εικόνα πηγής. Η συντεταγμένη x ορίζεται ως αύξουσα στη δεξιά κατεύθυνση και η συντεταγμένη y ορίζεται ως αύξουσα προς την κάτω κατεύθυνση.

Για να υπολογίσουμε τα G_x και G_y μετακινούμε τον κατάλληλο πίνακα (παράθυρο) από πάνω στην εικόνα εισόδου, υπολογίζοντας την τιμή για ένα pixel και στη συνέχεια μετατοπίζοντας ένα pixel προς τα δεξιά. Μόλις φτάσουμε στο τέλος της σειράς, μετακινούμαστε μέχρι την αρχή της επόμενης σειράς. Το παρακάτω παράδειγμα δείχνει τον υπολογισμό μιας τιμής του G_x :

a11	a12	a13	...	
a21	a22	a23	...	
a31	a32	a33	...	
...	

$$\text{kernel} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

b11	b12	b13	...	
b21	b22	b23	...	
b31	b32	b33	...	
...	

Εικόνα Εισόδου

Εικόνα Εξόδου (G_x)

Οι πίνακες περιέχουν θετικούς και αρνητικούς συντελεστές. Αυτό σημαίνει ότι η εικόνα εξόδου θα περιέχει θετικές και αρνητικές τιμές. Για λόγους προβολής μπορούμε:

- Να αντιστοιχίσουμε την κλίση του μηδέν σε ένα ημίτονο επίπεδο γκρι. Αυτό κάνει τις αρνητικές διαβαθμίσεις να φαίνονται πιο σκούρες και τις θετικές διαβαθμίσεις φωτεινές.
- Να χρησιμοποιήσουμε τις απόλυτες τιμές του χάρτη κλίσης (μεταξύ 0 και 255).
- Αυτό κάνει τις πολύ αρνητικές και πολύ θετικές κλίσεις να φαίνονται πιο φωτεινές. Οι πίνακες είναι ευαίσθητοι σε οριζόντιες και κάθετες μεταβάσεις. Το μέτρο μιας ακμής είναι το πλάτος και η γωνία της. Αυτά είναι εύκολα να υπολογιστούν από τα G_x και G_y .

Σε κάθε εικονοστοιχείο της εικόνας, οι προσεγγίσεις διαβάθμισης που δίνονται από το G_x και το G_y συνδυάζονται για να δώσουν το μέγεθος της κλίσης, χρησιμοποιώντας:

$$G = \sqrt{G_x^2 + G_y^2}$$

Η κατεύθυνση της κλίσης υπολογίζεται χρησιμοποιώντας:

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

3.5.4 Μέση Τιμή των καναλιών RGB & HSV

Τα επόμενα 6 χαρακτηριστικά σχετίζονται με τη χρωματική πληροφορία του πίνακα και είναι η μέση τιμή της έντασης των εικονοστοιχείων ολόκληρου του πίνακα στα επιμέρους κανάλια δύο χρωματικών μοντέλων : RGB & HSV. Η μέση τιμή, δηλαδή, των συνιστωσών Red Green Blue, Hue Saturation Value των εικονοστοιχείων των εικόνων.

Η μέση τιμή του κόκκινου, του πράσινου και του μπλε μπορεί να εκφραστεί αντίστοιχα ως:

$$m_1 = \frac{1}{MN} \sum_n \sum_m I_R(m, n).$$

$$m_2 = \frac{1}{MN} \sum_n \sum_m I_G(m, n).$$

$$m_3 = \frac{1}{MN} \sum_n \sum_m I_B(m, n).$$

Όπου Μ και Ν είναι το πλήθος γραμμών και στηλών της εικόνας, $I_R(m,n)$, $I_B(m,n)$ και $I_G(m,n)$ είναι τα κανάλια κόκκινου (red), πράσινου (green) και μπλε (blue) στα (m,n) pixel.

Η χροιά προσδιορίζει το βασικό χρώμα (π.χ. κόκκινο) της απόχρωσης. Ο κορεσμός του χρώματος που υπάρχει στους πίνακες συχνά σχετίζεται με τα αδιαφανή ή διαφανή χαρακτηριστικά, τα οποία μπορεί να εξαρτώνται από την ποσότητα νερού ή λευκής μογιός που προσθέτει ο καλλιτέχνης για να επιτύχει κάποια απόχρωση [3].

Η τιμή (ή αξία) σχετίζεται με το κατά πόσο ένα χρώμα είναι φωτεινό ή σκούρο.

Η μέση τιμή της απόχρωσης, του κορεσμού και της αξίας μπορεί να εκφραστεί αντίστοιχα ως:

$$m_4 = \frac{1}{MN} \sum_n \sum_m I_H(m, n).$$

$$m_5 = \frac{1}{MN} \sum_n \sum_m I_S(m, n).$$

$$m_6 = \frac{1}{MN} \sum_n \sum_m I_V(m, n).$$

Όπου M και N είναι το πλήθος γραμμών και στηλών της εικόνας, $I_H(m,n)$, $I_S(m,n)$ και $I_V(m,n)$ είναι η τιμή απόχρωσης (hue), κορεσμού (saturation) και αξίας (value) στα (m,n) pixel.

3.5.5 Χρωματικό Ιστόγραμμα

Τα επόμενα χαρακτηριστικά επιδιώκουν να αναδείξουν την προτίμηση που χαρακτηρίζει κάποια καλλιτεχνικά ρεύματα για χρήση συγκεκριμένων χρωμάτων. Για παράδειγμα, οι Ιμπρεσιονιστές προτιμούν τα φωτεινά, παστέλ χρώματα [8]. Πρόκειται για χαρακτηριστικά που προκύπτουν από το ιστόγραμμα της χροιάς του πίνακα.

Πιο συγκεκριμένα, τα τελευταία 12 χαρακτηριστικά σχετίζονται με τις χρωματικές επιλογές των καλλιτεχνών και υπολογίζουν τον αριθμό των pixel του ψηφιοποιημένου έργου που αντιστοιχούν σε 12 χρώματα. Η τεχνική αυτή εξαγωγής χρωματικών χαρακτηριστικών έχει προταθεί στα [20], [21]. Αρχικά υπολογίζουμε τον αριθμό pixel που απεικονίζουν το λευκό και το μαύρο. Το μαύρο αντιστοιχεί σε μηδενική τιμή (Value) σύμφωνα με το μοντέλο HSV. Ωστόσο, το μάτι αντιλαμβάνεται το μαύρο και σε ένα σύνολο άλλων τιμών κοντά στο μηδέν. Επομένως, πρέπει να βρεθεί μια τιμή που θα θεωρηθεί το ανώτερο όριο για το μαύρο χρώμα. Σύμφωνα με το [21] ως όριο λαμβάνεται η τιμή 0,1. Άρα όλα τα pixel που οι τιμές τους είναι κάτω από το 0,1 υπολογίζονται με μαύρο χρώμα. Αντίστοιχα με την ίδια λογική εντοπίζονται και τα λευκά εικονοστοιχεία. Σύμφωνα με το μοντέλο HSV, το λευκό αντιστοιχεί σε τιμή (Value) 1 και Κορεσμό (Saturation) 0. Συμφωνα με το [21] ορίζονται ως εικονοστοιχεία λευκού χρώματος όσα έχουν τιμή πάνω από 0,80 και κορεσμό κάτω από 0,25.

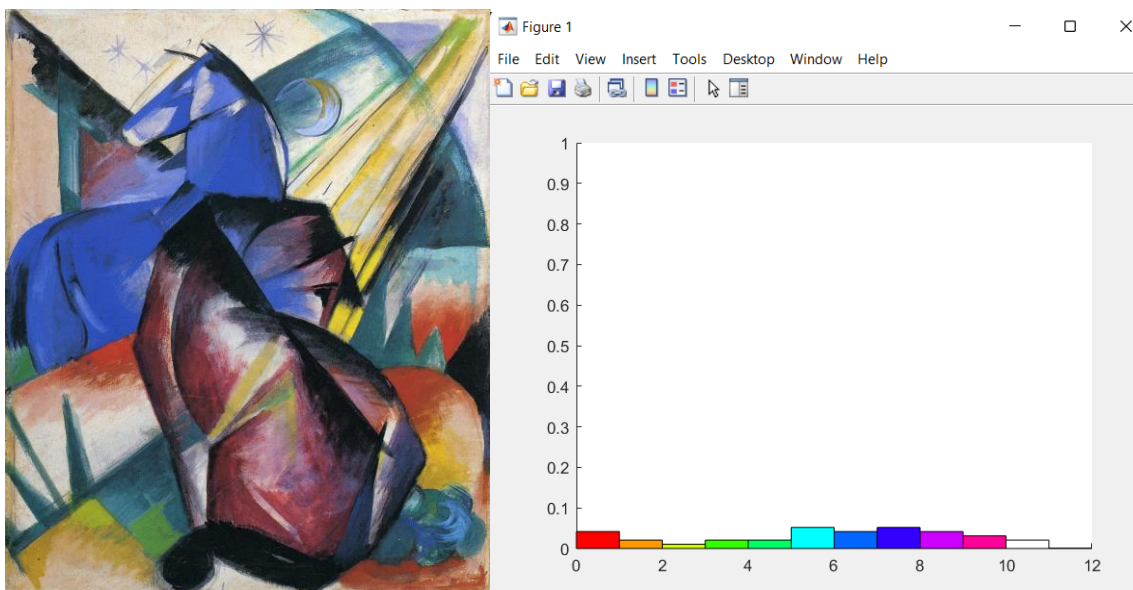
Τα υπόλοιπα χρώματα θα ταξινομηθούν βάση της τιμής της χροιάς (hue) σύμφωνα με το χρωματικό μοντέλο HSV. Το πρώτο βήμα στη διαδικασία αυτή είναι ο εντοπισμός των μάρων και λευκών εικονοστοιχείων. Προκειμένου να διαχωριστούν από τα «έγχρωμα» εικονοστοιχεία, αντιστοιχίζεται σε αυτά (τα λευκά και τα μαύρα) η τιμή -1.

Στο μοντέλο HSV, η απόχρωση παίρνει τιμές από 0^0 έως 360^0 . Ας σημειωθεί ότι στο matlab, για το μοντέλο HSV, η απόχρωση παίρνει τιμές από 0 έως 1. Ο πίνακας που θα καταρτιστεί βασίζεται σε δέκα χρώματα εκτός από το λευκό και το μαύρο. Αυτά τα χρώματα είναι κόκκινο πορτοκαλί, κίτρινο, σκούρο πράσινο, ανοιχτό πράσινο, μπλε, θαλασσί, σκούρο μπλε, σκούρο κόκκινο και μωβ. Για αυτό χωρίστηκε η περιοχή του χρώματος σε δέκα ίσα μέρη, το καθένα με

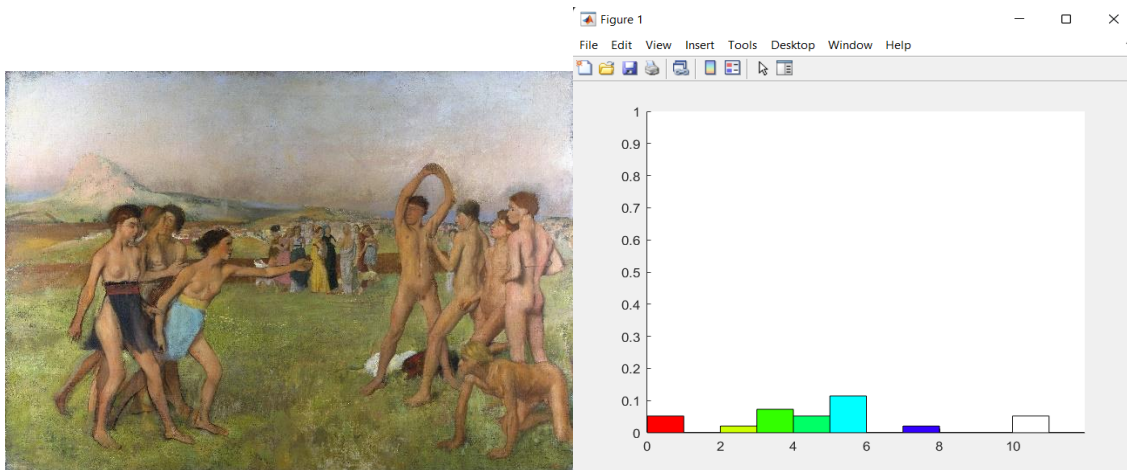
πλάτος 0,1. Στη συνέχεια, κάθε εικονοστοιχείο, ανάλογα με την τιμή της χροιάς του, ταξινομείται ως ένα από τα δέκα χρώματα. Η καταμέτρηση των εικονοστοιχείων του έργου τέχνης που αντιστοιχούν στα 12 χρώματα (μαζί με το λευκό και το μαύρο) οδηγεί στο χρωματικό ιστόγραμμα της εικόνας. Το πλήθος των εικονοστοιχείων ανά χρώμα διαιρείται με το συνολικό πλήθος εικονοστοιχείων της εικόνας για λόγους κανονικοποίησης.

Για παράδειγμα στην Εικόνα 20 βλέπουμε το έργο του Franz Marc, «Two Horses, Red and Blue», που ανήκει στο καλλιτεχνικό ρεύμα του κυβισμού και παρατηρούμε ότι τα χρώματα που υπάρχουν στην εικόνα, αντικατοπτρίζονται στο ιστογραμμά που παρατίθεται δίπλα στον πίνακα. Αναλυτικότερα τα κυρίαρχα χρώματα της εικόνας που είναι το μπλε, κόκκινο, κυανό κλπ εμφανίζονται να είναι αυτά με τα μεγαλύτερα ποσοστά και στο χρωματικό ιστόγραμμα.

Στην Εικόνα 21 βλέπουμε το έργο του Edgar Degas, «Young Spartans Exercising» που ανήκει στο καλλιτεχνικό ρεύμα του Ιμπρεσιονισμού και παρατηρούμε ότι τα κυρίαρχα χρώματα είναι το κυανό και το πράσινο. Όπως βλέπουμε και στο αντίστοιχο ιστόγραμμα αυτά είναι τα χρώματα με τα μεγαλύτερα ποσοστά.



Εικόνα 20 – Πίνακας του Franz Marc (κυβισμός) και αντίστοιχο χρωματικό ιστόγραμμα. Τα κυρίαρχα χρώματα του πίνακα ταιριάζουν με αυτά του ιστογράμματος



Εικόνα 21 – Πίνακας του Edgar Degas (Ιμπρεσιονισμός) και αντίστοιχο χρωματικό ιστόγραμμα. Στην παραπάνω εικόνα που καλλιτεχνικό ρεύμα έχει τον ιμπρεσιονισμό, παρατηρούμε ότι το κυρίαρχο χρώμα είναι το κυανό και το πράσινο. Όπως βλέπουμε και στο παραπάνω ιστόγραμμα συμπίπτει με τα χρώματα που βλέπει το ανθρώπινο μάτι.

3.6 ΠΕΡΙΓΡΑΦΗ ΤΑΞΙΝΟΜΗΣΗΣ ΔΕΔΟΜΕΝΩΝ

Αφού ολοκληρώσαμε την περιγραφή των χαρακτηριστικών, θα πρέπει αυτά τα χαρακτηριστικά να χρησιμοποιηθούν για την εκπαίδευση των ταξινομητών. Η εκπαίδευση πραγματοποιείται μέσω ενός προγράμματος του Matlab το Classification Learner το οποίο χρησιμοποιεί διάφορους ταξινομητές για την ταξινόμηση των δεδομένων βάση των χαρακτηριστικών χρησιμοποιώντας μηχανική μάθηση .

3.6.1 Classification Learner

Η εφαρμογή αυτή προσφέρει τη δυνατότητα χρήσης διαφόρων ταξινομητών για την κατηγοριοποίηση δεδομένων. Παρέχει το κατάλληλο γραφικό περιβάλλον για την οπτικοποίηση των δεδομένων, την οπτική αξιολόγηση των χαρακτηριστικών, την εκτίμηση των αποτελεσμάτων της ταξινόμησης [23] [24].

Χρησιμοποιώντας αυτήν την εφαρμογή, μπορούμε να εξερευνήσουμε τα δεδομένα μας, να επιλέξουμε χαρακτηριστικά, να εκπαιδεύσουμε τα μοντέλα και να αξιολογήσουμε τα αποτελέσματα. Υπάρχει και η δυνατότητα να εκτελέσουμε αυτοματοποιημένη εκπαίδευση για να αναζητήσουμε τον καλύτερο ταξινομητή, συμπεριλαμβανομένων δέντρων αποφάσεων (decision trees), μηχανών υποστήριξης διανυσμάτων (support vector machines), πλησιέστερων γειτόνων (nearest neighbors), ταξινομητή Bayes, νευρωνικών δικτύων (neural network) κ.α. [23]

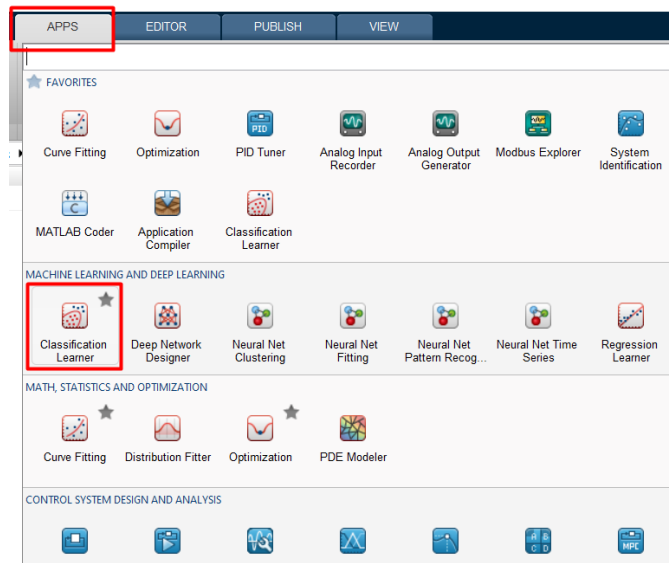
Η εκπαίδευση ενός μοντέλου στο Classification Learner αποτελείται από δύο μέρη [24]:

- Μοντέλο Εγκυρότητας (Validated Model): Εκπαιδεύουμε ένα μοντέλο με ένα σχήμα εγκυρότητας (validation scheme). Η εφαρμογή εξ ορισμού, προστατεύει από την υπερμοντελοποίηση (δηλαδή υπερχείλιση δεδομένων ή να προβλέψει αξιόπιστα μελλοντικές παρατηρήσεις) εφαρμόζοντας το Cross – Validation.
- Πλήρες Μοντέλο (Full Model): Εκπαιδεύουμε ένα μοντέλο με πλήρη δεδομένα χωρίς εγκυρότητα. Η εφαρμογή εκπαιδεύει ταυτόχρονα αυτό το μοντέλο με το μοντέλο εγκυρότητας. Ωστόσο, το μοντέλο που έχει εκπαιδευτεί με πλήρη δεδομένα δεν είναι ορατό στην εφαρμογή. Όταν επιλέγουμε έναν ταξινομητή για εξαγωγή στον χώρο εργασίας (workspace), το Classification Learner εξάγει το πλήρες μοντέλο.

Η εφαρμογή εμφανίζει τα αποτελέσματα του μοντέλου εγκυρότητας. Τα διαγνωστικά μέτρα, όπως η ακρίβεια του μοντέλου, και οι γραφικές παραστάσεις, όπως το διάγραμμα διασποράς (scatter plot) ή ο πίνακας σύγχυσης (confusion matrix), αντικατοπτρίζουν τα έγκυρα αποτελέσματα του μοντέλου. Όταν επιλέγουμε ένα μοντέλο για εξαγωγή στον χώρο εργασίας, το Classification Learner εξάγει το πλήρες μοντέλο. Επειδή το Classification Learner δημιουργεί ένα αντικείμενο μοντέλου του πλήρους μοντέλου κατά τη διάρκεια της εκπαίδευσης, δεν αντιμετωπίζουμε χρόνο καθυστέρησης κατά την εξαγωγή του μοντέλου[24].

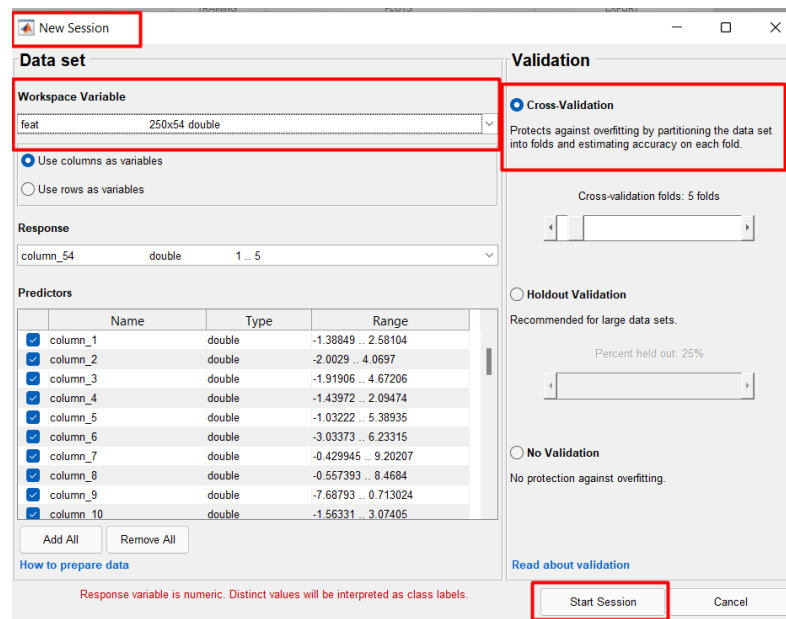
Παρακάτω παραθέτουμε τη αυτόματη εκπαίδευση του Classification Learner[24]:

1. Το Classification Learner βρίσκεται στο περιβάλλον του Matlab πάνω αριστερά και πατάμε το APPS όπως στην παρακάτω εικόνα:



Εικόνα 22 - Classification Learner

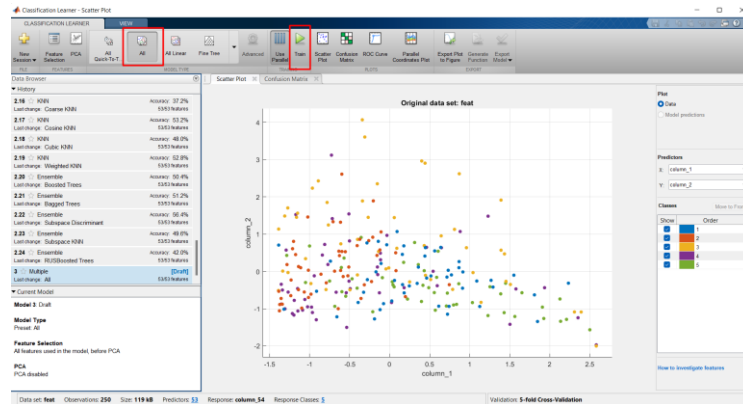
2. Έπειτα πατάμε το New Session για να μας βγάλει σε νέα καρτέλα. Τοποθετούμε όλα τα δεδομένα που περιγράψαμε και πατάμε Cross-Validation και Start Session ώστε να ξεκινήσει η εκπαίδευση.



Εικόνα 23 - Cross-Validation

Στο Cross-Validation παίρνει το σύνολο των δεδομένων και κάνει εκπαίδευση σε κάποιους από τους πίνακες που εξετάζουμε και στους περισσευόμενους τους χρησιμοποιεί σαν τεστ για να βρει σε ποιο καλλιτεχνικό ρεύμα ανήκουν. Στη συνέχεια θα μιλήσουμε αναλυτικά για το Cross - Validation.

3. Έπειτα εκπαιδεύουμε τους ταξινομητές που επιλέγουμε: SVM, KNN και Naïve Bayes.



Εικόνα 24 – Εκπαίδευση Ταξινομητή

3.6.2 Cross – Validation

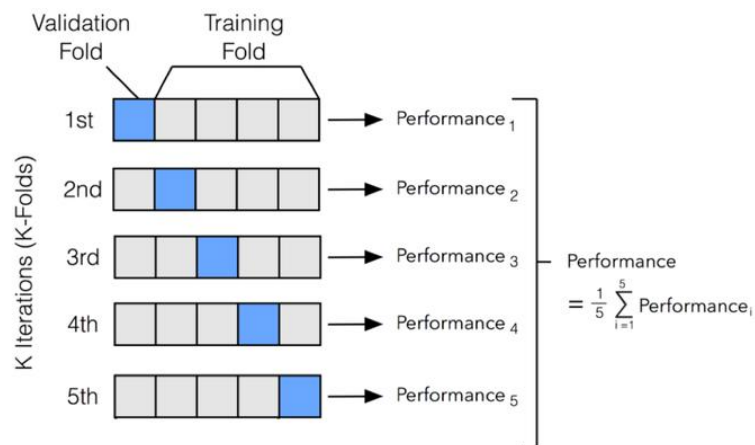
Το Cross – Validation είναι μια τεχνική αξιολόγησης μοντέλων που χρησιμοποιείται για την αξιολόγηση της απόδοσης ενός αλγορίθμου μηχανικής μάθησης για την πραγματοποίηση προβλέψεων σε νέα σύνολα δεδομένων στα οποία δεν έχει εκπαιδευτεί. Αυτό γίνεται με διαμερισμό του γνωστού συνόλου δεδομένων, χρησιμοποιώντας ένα υποσύνολο για την εκπαίδευση του αλγόριθμου και τα υπόλοιπα δεδομένα για δοκιμή [37].

Κάθε γύρος του cross-validation περιλαμβάνει τυχαία κατάτμηση του αρχικού συνόλου δεδομένων σε ένα σύνολο εκπαίδευσης και ένα σύνολο δοκιμών. Το σετ εκπαίδευσης χρησιμοποιείται στη συνέχεια για την εκπαίδευση ενός εποπτευόμενου αλγόριθμου μάθησης και το σύνολο δοκιμών χρησιμοποιείται για την αξιολόγηση της απόδοσής του. Αυτή η διαδικασία επαναλαμβάνεται πολλές φορές και το μέσο σφάλμα του cross-validation χρησιμοποιείται ως δείκτης απόδοσης (performance indicator) [37].

Το Cross – Validation είναι σημαντικό διότι όταν εκπαιδεύουμε ένα μοντέλο, χρησιμοποιώντας το Cross - Validation δεν θα υπάρξει υπερμοντελοποίηση. Η επιλογή των σετ εκπαίδευσης είναι καθοριστικής σημασίας για τη μείωση αυτού του κινδύνου. Ωστόσο, η διαίρεση του συνόλου δεδομένων για τη μεγιστοποίηση τόσο της εκμάθησης όσο και της εγκυρότητας των αποτελεσμάτων των τεστ είναι δύσκολη. Γι' αυτό το Cross – Validation είναι σημαντικό. Ουσιαστικά προσφέρει πολλές τεχνικές που διαχωρίζουν με διαφορετικό τρόπο τα διαφορετικά δεδομένα, για να βρεθεί ο καλύτερος αλγόριθμος για το μοντέλο [37].

Παρακάτω παραθέτουμε μερικές τεχνικές για το Cross – Validation [37]:

1. k – fold: Διαχωρίζει τα δεδομένα σε k τυχαία επιλεγμένα υποσύνολα περίπου ίσου μεγέθους. Ένα υποσύνολο χρησιμοποιείται για την εγκυρότητα του μοντέλου που εκπαιδεύτηκε χρησιμοποιώντας τα υπόλοιπα υποσύνολα. Αυτή η διαδικασία επαναλαμβάνεται k φορές έτσι ώστε κάθε υποσύνολο να χρησιμοποιείται ακριβώς μία φορά για εγκυρότητα. Το μέσο σφάλμα σε όλες τις k καταταμίσεις αναφέρεται ως performance. Αυτή είναι μια από τις πιο δημοφιλείς τεχνικές του cross - validation, αλλά μπορεί να χρειαστεί πολύς χρόνος για να εκτελεστεί επειδή το μοντέλο πρέπει να εκπαιδεύεται επανειλημμένα. Η παρακάτω εικόνα δείχνει τη διαδικασία.



Εικόνα 25 – Η διαδικασία του k – fold [37].

Η τεχνική αυτή είναι η τεχνική που επιλέχθηκε γι' αυτή τη διπλωματική με k=5.

2. Holdout: Διαχωρίζει τα δεδομένα τυχαία σε ακριβώς δύο υποσύνολα καθορισμένης αναλογίας για εκπαίδευση και εγκυροποίηση. Αυτή η μέθοδος εκτελεί εκπαίδευση και δοκιμές μόνο μία φορά, γεγονός που μειώνει τον χρόνο εκτέλεσης σε μεγάλα σύνολα, αλλά ερμηνεύει το αναφερόμενο σφάλμα με προσοχή σε μικρά σύνολα δεδομένων.
3. Leaveout: Χρησιμοποιώντας την προσέγγιση k-fold όπου το k είναι ίσο με τον συνολικό αριθμό των παρατηρήσεων στα δεδομένα και όλα τα δεδομένα θα χρησιμοποιηθούν μία φορά ως δοκιμαστικό σύνολο. Γνωστή και ως leave-one-out cross-validation (LOOCV).
4. Stratify: Διαχωρίζει δεδομένα έτσι ώστε τόσο τα σετ εκπαίδευσης όσο και τα σετ δοκιμών να έχουν περίπου τις ίδιες αναλογίες κατηγορίας στην απόκριση ή στο στόχο.
5. Επαναλαμβανόμενη τυχαία υποδειγματοληψία (Repeated random sub-sampling): Δημιουργεί πολλαπλές τυχαίες κατατμήσεις δεδομένων για χρήση ως σετ εκπαίδευσης και σετ δοκιμών χρησιμοποιώντας τη μεθοδολογία Monte Carlo και συγκεντρώνει αποτελέσματα σε όλες τις εκτελέσεις. Αυτή η τεχνική έχει παρόμοια ιδέα με το k-fold, αλλά κάθε σύνολο δοκιμής επιλέγεται ανεξάρτητα, πράγμα που σημαίνει ότι ορισμένα σημεία δεδομένων μπορεί να χρησιμοποιηθούν για δοκιμές περισσότερες από μία φορές.
6. Επανακατάσταση (Resubstitution): Δεν χωρίζει τα δεδομένα. Όλα τα δεδομένα χρησιμοποιούνται για την εκπαίδευση του μοντέλου. Το σφάλμα αξιολογείται συγκρίνοντας το αποτέλεσμα με τις πραγματικές τιμές. Αυτή η προσέγγιση συχνά παράγει υπερβολικά αισιόδοξες εκτιμήσεις για την απόδοση και θα πρέπει να αποφεύγεται εάν υπάρχουν επαρκή δεδομένα.

3.6.3 Επιλογή Ταξινόμητή

Όπως ειπώθηκε παραπάνω χρησιμοποιήσαμε τον Classification Learner για να εκπαιδεύσουμε αυτόματα διάφορα μοντέλα ταξινόμησης, με τα δεδομένα μας. Στην αρχή χρησιμοποιήσαμε αυτοματοποιημένη εκπαίδευση για να δοκιμάσουμε γρήγορα τα μοντέλα ταξινόμητων και στη συνέχεια εξερευνήσαμε 3 αλγόριθμους ταξινόμησης, τον Support Vector Machine, Naïve Bayes και Nearest Neighbor ταξινόμητες.

3.6.3.1 Ταξινομητές Naïve Bayes

Οι ταξινομητές Naive Bayes (απλοϊκός ταξινομητής Bayes) είναι εύκολοι στην ερμηνεία και είναι χρήσιμοι για ταξινόμηση πολλαπλών κλάσεων. Ο αλγόριθμος Naïve Bayes αξιοποιεί το θεώρημα Bayes και κάνει την υπόθεση ότι οι τιμές των προτύπων είναι ανεξάρτητες, δεδομένης της κλάσης και συνεπώς ότι η πιθανότητα της κοινής εμφάνισης των προτύπων δεδομένης μίας κλάσης είναι ίση με το γινόμενο της εμφάνισης καθενός προτύπου δεδομένης της κλάσης [39].

$$p(\mathbf{x}|\omega_i) = \prod_{j=1}^n p(x_j|\omega_i)$$

Χρησιμοποιούμε αυτούς τους ταξινομητές εάν αυτή η υπόθεση ανεξαρτησίας ισχύει για προγνωστικούς παράγοντες στα δεδομένα μας. Ωστόσο, ο αλγόριθμος εξακολουθεί να φαίνεται να λειτουργεί καλά όταν η υπόθεση ανεξαρτησίας δεν είναι έγκυρη. [39]

Δύο ταξινομητές Naïve Bayes είναι οι εξής:

- Gaussian Naive Bayes: Χρησιμοποιείται στην ταξινόμηση και υποθέτει ότι τα χαρακτηριστικά ακολουθούν μια κανονική κατανομή. [42]
- Kernel Naive Bayes: Η κατανομή «πυρήνα» είναι κατάλληλη για χαρακτηριστικά που έχουν συνεχή κατανομή. Δεν απαιτεί ισχυρή υπόθεση, όπως μια κανονική κατανομή και μπορούμε να τη χρησιμοποιήσουμε σε περιπτώσεις όπου η κατανομή των χαρακτηριστικών μπορεί να αποκλίνει από την κανονική (είναι λοξή ή να έχει πολλαπλές κορυφές). Απαιτεί περισσότερο υπολογιστικό χρόνο και περισσότερη μνήμη. Για κάθε πρόβλεψη που μοντελοποιείται με κατανομή πυρήνα, ο απλός ταξινομητής Bayes υπολογίζει μια ξεχωριστή εκτίμηση πυκνότητας πυρήνα για κάθε κλάση με βάση τα δεδομένα εκπαίδευσης για αυτήν την κλάση. Από προεπιλογή ο πυρήνας είναι ο κανονικός πυρήνας και ο ταξινομητής επιλέγει αυτόματα ένα πλάτος για κάθε κλάση και προγνωστικό. Το λογισμικό υποστηρίζει τον καθορισμό διαφορετικών πυρήνων για κάθε προγνωστικό παράγοντα και διαφορετικά πλάτη για κάθε προγνωστικό ή κλάση. [42]

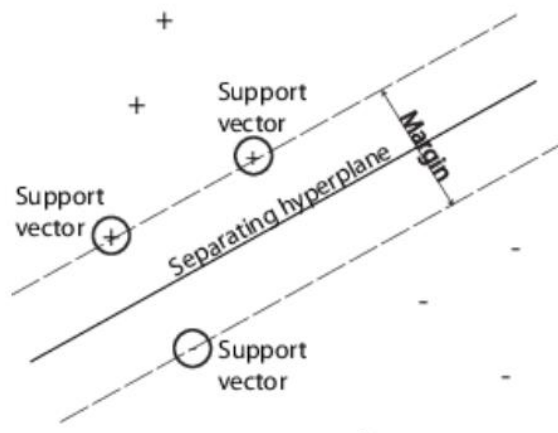
3.6.3.2 Support Vector Machine

Στο Classification Learner, μπορούμε να εκπαιδεύσουμε τον SVM όταν τα δεδομένα μας έχουν δύο ή περισσότερες κλάσεις. [39]

Ο SVM ταξινομεί τα δεδομένα βρίσκοντας το καλύτερο υπερεπίπεδο που διαχωρίζει τα σημεία δεδομένων μιας κλάσης από εκείνα μιας άλλης κλάσης. Το καλύτερο υπερεπίπεδο για έναν SVM σημαίνει αυτό με το μεγαλύτερο περιθώριο (margin) μεταξύ των δύο κλάσεων που δεν περιλαμβάνει εσωτερικά σημεία δεδομένων. [39]

Τα διανύσματα υποστήριξης είναι τα σημεία δεδομένων που βρίσκονται πιο κοντά στο διαχωριστικό υπερεπίπεδο. Αυτά τα σημεία βρίσκονται στο διαχωριστικό σύνορο των δύο κλάσεων. [39]

Η Εικόνα 26 απεικονίζει τους παραπάνω ορισμούς, με το + να υποδεικνύει σημεία δεδομένων τύπου 1 και - να υποδεικνύει σημεία δεδομένων τύπου -1.



Εικόνα 26 - Σχηματική αναπαράσταση των διανυσμάτων υποστήριξης. [39]

Παρακάτω παραθέτουμε μία σύντομη περιγραφή για τον καθένα SVM ταξινομητή.

- **Linear SVM:** Κάνει έναν απλό γραμμικό διαχωρισμό μεταξύ των κλάσεων [39]. Δηλαδή όταν ο ταξινομητής Linear SVM μπορεί εύκολα να διαχωρίσει τα δεδομένα με υπερεπίπεδο σχεδιάζοντας μια ευθεία γραμμή [41].

Kernel SVM: Στους παρακάτω ταξινομητές χρησιμοποιείται το κόλπο του πυρήνα (kernel trick). Οι πυρήνες είναι απλώς ένας έξυπνος τρόπος προσθήκης περισσότερων χαρακτηριστικών

στα δεδομένα, με την ελπίδα να γίνουν γραμμικά διαχωρίσιμα. Δηλαδή, αντί να τα προσθέτουν στην πραγματικότητα, κάτι που θα έκανε το μοντέλο αργό, κάνουν χρήση ορισμένων μαθηματικών ιδιοτήτων που μας επιτρέπουν να λάβουμε ακριβώς τα ίδια αποτελέσματα σαν να τα είχαμε προσθέσει, χωρίς να επιβραδύνει το μοντέλο. [43] Η τεχνική ταξινόμησης SVM είναι χρήσιμη όταν αντιμετωπίζεται το δίλημμα του χαμηλού χώρου μνήμης. Το SVM βρίσκει ένα υπερεπίπεδο σε πολυδιάστατο χώρο που διαιρεί τις κλάσεις με τον καλύτερο δυνατό τρόπο.

- Quadratic SVM: Σε έναν χώρο υψηλών διαστάσεων, υπάρχει μεγαλύτερη πιθανότητα τα δεδομένα να διαχωρίζονται γραμμικά. Όταν έχουμε μόνο ένα χαρακτηριστικό, το x_1 , δεν μπορούμε να διαχωρίσουμε τα δεδομένα με μία μόνο γραμμή. Η προσθήκη ενός άλλου χαρακτηριστικού, του x_2 , ίσου με το x_1 στο τετράγωνο, διευκολύνει τον διαχωρισμό των δύο κλάσεων. Η προσθήκη επιπλέον χαρακτηριστικών μπορεί να κάνει τα δεδομένα να διαχωρίζονται γραμμικά. [43]
- Cubic SVM: Εδώ κυβικός τύπος SVM χρησιμοποιείται όπου η συνάρτηση πυρήνα του ο ταξινομητής είναι κυβική και δίνεται από τον τύπο: $\kappa(x_i, x_j) = (x_i^T x_j)^3$. [44]
- Fine Gaussian SVM:

$$K(x_1, x_2) = e^{-\frac{\|x_1 - x_2\|^2}{2\sigma^2}}$$

Η παράμετρος Kernel Scale (σ) ρυθμισμένη σε $\sqrt{P}/4$, όπου P είναι ο αριθμός των χαρακτηριστικών [39].

- Medium Gaussian SVM:

$$K(x_1, x_2) = e^{-\frac{\|x_1 - x_2\|^2}{2\sigma^2}}$$

Η παράμετρος Kernel Scale ρυθμισμένη σε \sqrt{P} , όπου P είναι ο αριθμός των χαρακτηριστικών [39].

- Coarse Gaussian SVM:

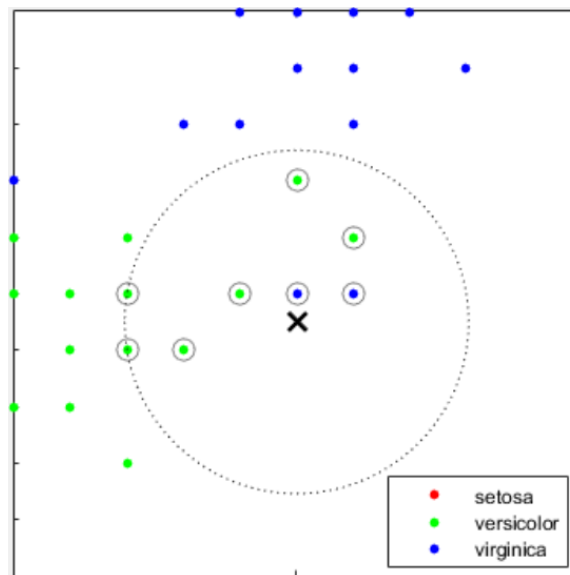
$$K(x_1, x_2) = e^{-\frac{\|x_1 - x_2\|^2}{2\sigma^2}}$$

Η παράμετρος Kernel Scale ρυθμισμένη σε $\sqrt{P} \cdot 4$, όπου P είναι ο αριθμός των χαρακτηριστικών [39].

3.6.3.3 Ταξινομητές Nearest Neighbor

Οι ταξινομητές πλησιέστερου γείτονα έχουν συνήθως καλή προγνωστική ακρίβεια σε χαμηλές διαστάσεις, αλλά μπορεί όχι σε υψηλές διαστάσεις. Έχουν υψηλή χρήση μνήμης και δεν είναι εύκολο να ερμηνευθούν. [39]

Ο ταξινομητής του k- κοντινότερου γείτονα είναι η κατηγοριοποίηση σημείων με βάση την απόστασή τους από σημεία (ή γείτονες) σε ένα σύνολο δεδομένων εκπαίδευσης και μπορεί να είναι ένας απλός αλλά αποτελεσματικός τρόπος ταξινόμησης νέων σημείων. Μπορούμε να χρησιμοποιήσουμε διάφορες μετρήσεις για να προσδιορίσουμε την απόσταση. Οι αλγόριθμοι που βασίζονται σε kNN χρησιμοποιούνται ευρέως ως κανόνες αναφοράς μηχανικής εκμάθησης. [39]



Εικόνα 27 - Σχηματική αναπαράσταση της λειτουργίας του ταξινομητή kNN. Δεδομένου ενός συνόλου X , n σημείων και μιας συνάρτησης απόστασης, η αναζήτηση k -πλησιέστερου γείτονα (kNN) μας επιτρέπει να βρούμε τα k πλησιέστερα σημεία στο X σε ένα σημείο ή ένα σύνολο σημείων. [39]

Έξι ταξινομητές KNN περιγράφονται παρακάτω με βάση την τιμή του k και την επιλεγμένη συνάρτηση απόστασης:

- Fine KNN: Χρησιμοποιεί Ευκλείδεια απόσταση, όπου είναι ένα μέτρο της πραγματικής ευθείας απόστασης μεταξύ δύο σημείων στον Ευκλείδειο χώρο. Η Ευκλείδεια απόσταση υπολογίζεται ως εξής:

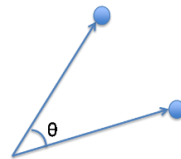
$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} . [45]$$

Ο αριθμός των γειτόνων έχει οριστεί σε 1. [39]

- Medium KNN: Χρησιμοποιεί Ευκλείδεια απόσταση. Ο αριθμός των γειτόνων έχει οριστεί σε 10. [39]
- Coarse KNN: Χρησιμοποιεί Ευκλείδεια απόσταση. Ο αριθμός των γειτόνων έχει οριστεί σε 100. [39]
- Cosine KNN: Χρησιμοποιεί μια μετρική απόστασης συνημιτόνου [39]. Η απόσταση συνημιτόνου είναι η μέτρηση της απόστασης που χρησιμοποιείται κυρίως για τον υπολογισμό της ομοιότητας μεταξύ δύο διανυσμάτων. Μετράται από το συνημίτονο της γωνίας μεταξύ δύο διανυσμάτων και καθορίζει εάν δύο διανύσματα δείχνουν προς την ίδια κατεύθυνση. Ο τύπος για την απόσταση συνημιτόνου είναι:

$$d(x, y) = (1 - \frac{x y'}{\sqrt{(xx')(yy')}})$$

$$\cos(\theta) = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| |\vec{B}|}$$



Χρησιμοποιώντας αυτόν τον τύπο μας λείει για την ομοιότητα μεταξύ των δύο διανυσμάτων και το $\cos(\theta)$ θα μας δίνει την συνημιτονική τους απόσταση.

Χρησιμοποιώντας αυτή την απόσταση παίρνουμε τιμές μεταξύ 0 και 1, όπου το 1 σημαίνει ότι τα διανύσματα είναι 100% όμοια μεταξύ τους και το 0 σημαίνει ότι δεν είναι καθόλου όμοια. [45]

Ο αριθμός των γειτόνων έχει οριστεί σε 10.

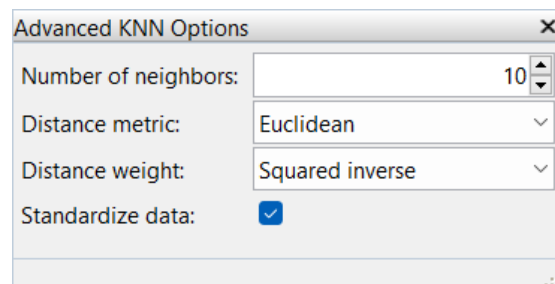
- Cubic KNN: Χρησιμοποιεί μια μετρική κυβικής απόστασης [39]. Η κυβική απόσταση (Minkowski Distance) είναι μια μέτρηση που προορίζεται για διανυσματικούς χώρους με πραγματικές τιμές. Μπορούμε να υπολογίσουμε την απόσταση Minkowski μόνο σε έναν κανονικό διανυσματικό χώρο, που σημαίνει σε ένα χώρο όπου οι αποστάσεις μπορούν να αναπαρασταθούν ως διάνυσμα που έχει μήκος και τα μήκη δεν μπορούν να είναι αρνητικά. Υπάρχουν μερικές προϋποθέσεις που πρέπει να πληροί η μέτρηση απόστασης [45]:

1. Μη αρνητικότητα: $d(x, y) \geq 0$
2. Ταυτότητα: $d(x, y) = 0$ αν και μόνο αν $x = y$
3. Συμμετρία: $d(x, y) = d(y, x)$
4. Ανισότητα τριγώνου: $d(x, y) + d(y, z) \geq d(x, z)$

$$\left(\sum_{i=1}^n |x_i - y_i|^3 \right)^{1/3}$$

Ο αριθμός των γειτόνων έχει οριστεί σε 10.

- Weighted KNN: Χρησιμοποιεί ευκλείδια απόσταση και χρησιμοποιεί ένα βάρος απόστασης $w = \frac{1}{d^2}$ [39]. Νόμος αντίστροφου τετραγώνου είναι ο νόμος που δηλώνει ότι ένα μέγεθος είναι αντιστρόφως ανάλογο με το τετράγωνο της απόστασης από την πηγή αυτού του μεγέθους [46].



Εικόνα 28 - Ρύθμιση των παραμέτρων του KNN στο Classification Learner

Ο αριθμός των γειτόνων έχει οριστεί σε 10.

3.6.4 Μέτρο Επίδοσης

Μπορούμε να ορίσουμε ως μέτρο επίδοσης μία μεταβλητή που εκφράζει ποσοτικά την αποτελεσματικότητα ή την αποδοτικότητα ή και τα δύο μαζί μιας διαδικασίας, σε σχέση με ένα στόχο. Συγκεκριμένα, το μέτρο επίδοσης, το χρησιμοποιούμε ώστε να δούμε κατά πόσο μια εικόνα, με βάση τα χαρακτηριστικά που επιλέχθηκαν, θα ταξινομηθεί σωστά. Γι' αυτό είναι σημαντική η εκτίμηση της ικανότητας πρόβλεψης του ταξινομητή για το κάθε καλλιτεχνικό ρεύμα [36].

Για να εκτιμήσουμε τις επιδόσεις ενός καλλιτεχνικού ρεύματος ενός ταξινομητή, εισάγουμε την αναγκαία ορολογία. Για την περίπτωση δυαδικής κλάσης ισχύουν οι ακόλουθοι όροι [36].:

- **Θετικές παρατηρήσεις** (positive) ονομάζονται οι παρατηρήσεις, οι οποίες ανήκουν σε μια τιμή του καλλιτεχνικού ρεύματος (πχ ρεαλισμός).
- **Αρνητικές παρατηρήσεις** (negative) ονομάζονται οι παρατηρήσεις, οι οποίες ανήκουν στην άλλη τιμή της κλάσης (πχ μη ρεαλισμός).
- **Αληθινές Θετικές Προβλέψεις** (true positive – tp) είναι το πλήθος των επιτυχών προβλέψεων για θετικές παρατηρήσεις (πχ το καλλιτεχνικό ρεύμα είναι ρεαλισμός και ο ταξινομητής προβλέπει σωστά την κλάση).
- **Αληθινές Αρνητικές Προβλέψεις** (true negative – tn) είναι το πλήθος των επιτυχημένων προβλέψεων για αρνητικές παρατηρήσεις (πχ το καλλιτεχνικό ρεύμα δεν είναι ρεαλισμός και ο ταξινομητής προβλέπει σωστά την κλάση).
- **Ψευδείς Θετικές Προβλέψεις** (false positive – fp) είναι το πλήθος των αποτυχημένων προβλέψεων για αρνητικές παρατηρήσεις (το καλλιτεχνικό ρεύμα δεν είναι ρεαλισμός, ο ταξινομητής όμως την προβλέπει ως ρεαλισμός).
- **Ψευδείς Αρνητικές Προβλέψεις** (false negative – fn) είναι το πλήθος των αποτυχημένων προβλέψεων για θετικές παρατηρήσεις (το καλλιτεχνικό ρεύμα είναι ρεαλισμός, ο ταξινομητής όμως την προβλέπει ως μη ρεαλισμός).

Ένας τρόπος παρουσίασης των επιδόσεων ανά κλάση ενός ταξινομητή είναι με τη χρήση του πίνακα σύγχυσης (confusion matrix). Ο Πίνακας Σύγχυσης είναι ένας δισδιάστατος πίνακας, όπου οι στήλες αντιστοιχούν στις προβλέψεις και οι γραμμές στις πραγματικές τιμές κλάσης.

Στα κελιά του πίνακα αναγράφονται οι αληθινές θετικές, οι αληθινές αρνητικές, οι ψευδείς θετικές και οι ψευδείς αρνητικές προβλέψεις. Στο παρακάτω πίνακα απεικονίζεται ένας Πίνακας Σύγχυσης (confusion matrix) [36].

	Πρόβλεψη Αρνητικής Κλάσης	Πρόβλεψη Θετικής Κλάσης
Πραγματική Αρνητική Κλάση	tn	fp
Πραγματική Θετική Κλάση	fn	tp

tn = true negative

tp = true positive

fn = false negative

fp = false positive

Ορισμένα πρόθετα μέτρα επίδοσης ενός ταξινομητή είναι τα ακόλουθα:

$$sensitivity = \frac{tp}{pos}$$

$$specificity = \frac{tn}{neg}$$

$$precision = \frac{tp}{tp + fp}$$

$$accuracy = sensitivity * \frac{pos}{pos + neg} + specificity * \frac{neg}{pos + neg} = \frac{tp + tn}{pos + neg}$$

Όπου pos είναι το πλήθος των θετικών παρατηρήσεων και neg είναι το πλήθος των αρνητικών παρατηρήσεων. Σύμφωνα με τα παραπάνω, η ακρίβεια (accuracy) ορίζεται ως το ποσοστό των ορθών θετικών προβλέψεων επί το ποσοστό των θετικών παρατηρήσεων συν το ποσοστό των ορθών αρνητικών προβλέψεων επί το ποσοστό των αρνητικών παρατηρήσεων ή ισοδύναμα ως το πλήθος των ορθών προβλέψεων προς το πλήθος των παρατηρήσεων [36].

3.7 ΠΕΙΡΑΜΑΤΑ

3.7.1 ΚΩΔΙΚΑΣ

Οι ψηφιοποιημένες εικόνες αποθηκεύονται ξεχωριστά – ανά ρεύμα – σε 5 ομότιτλους καταλόγους. Στη συνέχεια για την χρησιμοποίηση και οργάνωση των εικόνων από το πρόγραμμα χρησιμοποιείται η εντολή ImageDatastore, που δημιουργεί μια οργανωμένη συλλογή εικόνων. Με τη βοήθεια της εντολής αυτής οι εικόνες ετικετοποιούνται αυτόματα με βάση τον υποκατάλογο στον οποίο ανήκουν, δηλαδή classicism (κλασικισμός), cubism (κυβισμός), expressionism (εξπρεσιονισμός), impressionism (ιμπρεσιονισμός) και realism (ρεαλισμός).

Ύστερα δημιουργούμε μία συνάρτηση το featuresCalc που εκεί υπολογίζει όλα τα χαρακτηριστικά του κάθε καλλιτεχνικού ρεύματος. Η συνάρτηση αυτή δέχεται σαν όρισμα τις ψηφιοποιημένες εικόνες που ανήκουν στην εκάστοτε κλάση και επιστρέφει στην έξοδό της το διάνυσμα των χαρακτηριστικών που εξηγαγε από τους πίνακες. Καλείται 5 φορές συνολικά, μία φορά για κάθε μία ξεχωριστή κλάση, επιστρέφοντας τα χαρακτηριστικά εκείνα που αντιστοιχούν στις εικόνες της συγκεκριμένης κλάσης.

Μέσα στην συνάρτηση featuresCalc το **πρώτο** χαρακτηριστικό που υπολογίζουμε είναι το ποσοστό σκούρων χρωμάτων του κάθε πίνακα.

1. Πρώτα μετατρέπει την RGB εικόνα σε YCBCR για να βρούμε την φωτεινότητα.
2. Βρίσκουμε την φωτεινότητα.
3. Και τέλος βρίσκουμε το ποσοστό των σκούρων χρωμάτων όπως ορίζεται στο [16].

$$\text{Dark} = \text{sum}(\text{sum}(\text{LUM} < 64)) / (\text{TotalPixels});$$

Το πλήθος των εικονοστοιχείων που αντιστοιχούν στα σκούρα χρώματα διαρείται με το συνολικό πλήθος εικονοστοιχείων της εικόνας για λόγους κανονικοποίησης.

Το **δεύτερο** χαρακτηριστικό είναι ο συντελεστής κλίσης που υπολογίζεται από τον χάρτη κλίσης της εικόνας.

1. Για να το βρούμε αρχικά μετατρέπουμε την εικόνα σε κλίμακα του γκρι.
2. Έπειτα υπολογίζουμε τις κλίσεις κατεύθυνσης Gx Gy της κλίμακας του γκρι.
3. Και τέλος υπολογίζουμε τον συντελεστή κλίσης

$$\text{Sum}(\text{sum}(\text{sqrt}(Gx.^2 + Gy.^2))) / (\text{TotalPixels});$$

Το **τρίτο** χαρακτηριστικό που υπολογίζουμε είναι το σύνολο κορυφών της φωτεινότητας του ιστογράμματος της εικόνας του γκρι.

Το **τέταρτο** χαρακτηριστικό είναι η χρωματική περιοχή που αντιστοιχεί στο σημείο κορυφής του ιστογράμματος της φωτεινότητας.

Μετά τη διαίρεση της εικόνας σε πανομοιότυπα κομμάτια, το **πέμπτο** χαρακτηριστικό είναι η απόκλιση του μέσου επιπέδου γκρι που αποκτάται σε κάθε κομμάτι από το μέσο επίπεδο γκρι που αποκτάται σε ολόκληρη την εικόνα.

1. Για να το βρούμε αρχικά χωρίζουμε τον πίνακα της φωτεινότητας lum σε μικρότερους πίνακες.
2. Έπειτα βρίσκουμε την μέση τιμή της φωτεινότητας.
3. Τέλος βρίσκουμε την απόκλιση του μέσου επιπέδου γκρι.

Το **έκτο** χαρακτηριστικό είναι η ασυμμετρία «Skew», η απόκλιση της κατανομής του γκρι επιπέδου από την κατανομή Gauss.

1. Πρώτα υπολογίζουμε την μέση τιμή της κλίμακας του γκρι.
2. Έπειτα υπολογίζουμε τη διακύμανση της κλίμακας του γκρι.
3. Και τέλος υπολογίζουμε την ασυμμετρία «Skew».

Σε αυτό το σημείο ξεκινάμε την υλοποίηση της μεθόδου haralick. Υπολογίζει τις υπό συνθήκη κοινές πιθανότητες δύο ζευγαριών του γκρι να εμφανιστούν δίνοντας δύο παραμέτρους, την απόσταση μεταξύ των εικονοστοιχείων d , και τον προσανατολισμό θ (γωνία, θ). Οι δυνατές γωνίες είναι τέσσερις: 0° , 45° , 90° , και 135° .

Για τα 14 χαρακτηριστικά του Haralick υπολογίζουμε τους πίνακες συνεμφάνισης για 4 διαφορετικούς προσανατολισμούς: 0° , 45° , 90° , και 135° , ενώ και στις 4 περιπτώσεις η απόσταση των ζευγαριών θεωρείται ίση με 1 (εικονοστοιχείο). Για καθεμία από τις 4 περιπτώσεις υπολογίζονται τα 14 στατιστικά χαρακτηριστικά υφής και στη συνέχεια υπολογίζεται ο μέσος όρος των τεσσάρων.

Στα επόμενα 3 χαρακτηριστικά παίρνουμε την RGB εικόνα και υπολογίζουμε τον μέσο όρο των κόκκινων, πράσινων και μπλε αντίστοιχα. Έπειτα μετατρέπουμε την RGB εικόνα σε HSV εικόνα έτσι έχουμε άλλα 3 χαρακτηριστικά υπολογίζοντας τον μέσο όρο της απόχρωσης, του κορεσμού και της αξίας αντίστοιχα.

Στο επόμενο χαρακτηριστικό υπολογίζουμε τις ακμές με την μέθοδο Sobel. Το μήκος κάθε ακμής ορίζεται να είναι από 8 εικονοστοιχεία και πάνω.

Στο επόμενο χαρακτηριστικό υπολογίζουμε τον μέσο όρο του αριθμού των pixel των ακμών.

Στα τελευταία 12 χαρακτηριστικά υπολογίζεται ο αριθμός των εικονοστοιχείων που αντιστοιχούν στα 12 χρώματα.

Θεωρούμε το μοντέλο HSV και με βάση αυτό όσα εικονοστοιχεία έχουν Value μικρότερη του 0.1 θεωρούνται μαύρα, και όσα έχουν Value μεγαλύτερη από 0.8 και Κορεσμό μικρότερο του 0.25 είναι λευκά.

Τα υπόλοιπα χρώματα θα ταξινομηθούν βάση της τιμής της χροιάς (hue) σύμφωνα με το χρωματικό μοντέλο HSV ως εξής:

```
red = length(find(h > 0.95 | h <= 0.003))/TotalPixels;
orange = length(find(h > .003 & h <= 0.05))/ TotalPixels;
yellow = length(find(h > .05 & h <= .07))/ TotalPixels;
green = length(find(h > .07 & h <= .09))/ TotalPixels;
light_green = length(find(h > .09 & h <= .10))/
TotalPixels;
cyan = length(find(h > .10 & h <= .45))/ TotalPixels;
light_blue = length(find(h > .45 & h <= .55))/
TotalPixels;
blue = length(find(h > .55 & h <= .60))/ TotalPixels;
purple = length(find(h > .60 & h <= .75))/ TotalPixels;
magenta = length(find((h > .75) & h <= .95))/
TotalPixels;
```

Τέλος παίρνουμε όλα τα χαρακτηριστικά που βρήκαμε και για τα 5 καλλιτεχνικά ρεύματα και κάνουμε κανονικοποίηση των δεδομένων με μέσο όρο 0 και τυπική απόκλιση 1 για κάθε χαρακτηριστικό ξεχωριστά.

Τα κανονικοποιημένα χαρακτηριστικά οργανώνονται σε πίνακα 39 στηλών. Οι γραμμές του πίνακα είναι όσες και τα έργα τέχνης (250). Επιπλέον στο τέλος του πίνακα χαρακτηριστικών προστίθεται μία στήλη με τις ετικέτες των έργων. Ο πίνακας 250 επί 40 είναι η είσοδος του Classification Learner.

3.7.2 ΑΠΟΤΕΛΕΣΜΑΤΑ

Τα αποτελέσματα θα χωριστούν σε 2 κατηγορίες. Η πρώτη κατηγορία θα είναι και τα 5 καλλιτεχνικά ρεύματα και με τα 40 χαρακτηριστικά. Η δεύτερη κατηγορία θα είναι η σύγκριση δύο καλλιτεχνικών ρευμάτων μεταξύ τους.

3.7.2.1 Αποτελέσματα με τα 5 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά

Σε αυτό το κεφάλαιο θα συγκρίνουμε και τα 5 καλλιτεχνικά ρεύματα και με τα 40 χαρακτηριστικά. Ο ταξινομητής με τα καλύτερα αποτελέσματα είναι ο Quadratic SVM με ποσοστό επιτυχίας 69.6%.

Στον παρακάτω πίνακα σύγχυσης (confussion matrix) παρουσιάζονται οι συντεταγμένες 1,2,...,5 οι οποίες είναι τα εξής καλλιτεχνικά ρεύματα αντίστοιχα:

1. Κλασικισμός
2. Κυβισμός
3. Εξπρεσιονισμός
4. Ιμπρεσιονισμός
5. Ρεαλισμός

True class	1	2	3	4	5
1	39	4	1	3	3
2	4	38	1	5	2
3		2	34	11	3
4	4	4	2	29	11
5	7	2	1	6	34
	1	2	3	4	5
	Predicted class				

Εικόνα 29 - Confussion matrix - Quadratic SVM

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Quadratic SVM στον κλασικισμό αναγνώρισε 39 πίνακες και είχαμε ποσοστό ακρίβειας 78%, στον κυβισμό 38 πίνακες με ποσοστό ακρίβειας 68%, στον εξπρεσιονισμό 34 με ποσοστό ακρίβειας 68%, στον μπρεσιονισμό 29 με ποσοστό ακρίβειας 58% και στον ρεαλισμό 34 με ποσοστό ακρίβειας 68% αντίστοιχα.

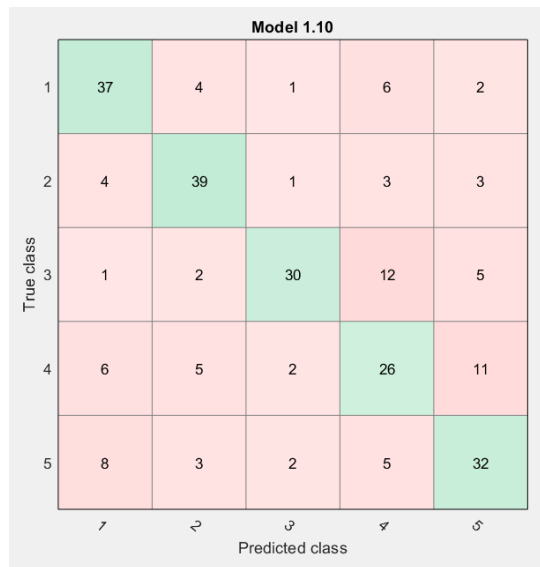
Ο επόμενος ταξινομητής είναι ο Linear SVM με ποσοστό ακρίβειας 68.8%.

True class	1	2	3	4	5
1	37	6		1	6
2	3	38		5	4
3		2	36	9	3
4	6	6	3	27	8
5	8	1	1	6	34
	1	2	3	4	5
	Predicted class				

Εικόνα 30 - Confusion matrix- Linear SVM

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Linear SVM στον κλασικισμό αναγνώρισε 37 πίνακες και είχαμε ποσοστό ακρίβειας 74%, στον κυβισμό 38 πίνακες με ποσοστό ακρίβειας 76%, στον εξπρεσιονισμό 36 με ποσοστό ακρίβειας 72%, στον μπρεσιονισμό 27 με ποσοστό ακρίβειας 54% και στον ρεαλισμό 34 με ποσοστό ακρίβειας 68% αντίστοιχα.

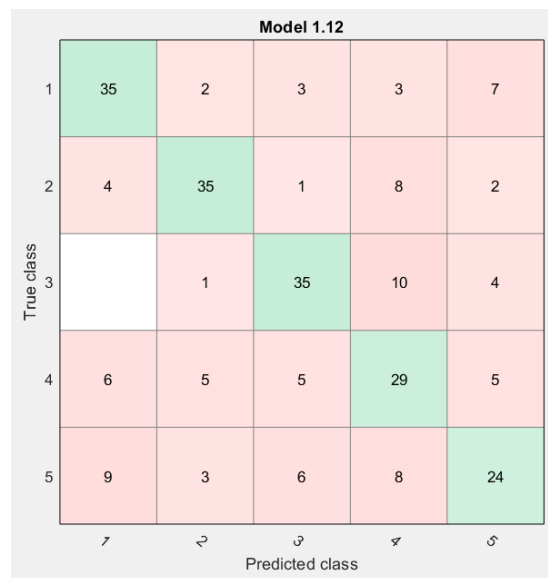
Ο επόμενος ταξινομητής είναι ο Cubic SVM με ποσοστό ακρίβειας 65.6%.



Εικόνα 31 - Confussion matrix Cubic SVM

Στον συγκεκριμένο πίνακα σύγχυσης του ταξινομητή Cubic SVM στον κλασικισμό αναγνώρισε 37 πίνακες και είχαμε ποσοστό ακρίβειας 74%, στον κυβισμό 39 πίνακες με ποσοστό ακρίβειας 78%, στον εξπρεσιονισμό 30 με ποσοστό ακρίβειας 60%, στον μπρεσιονισμό 26 με ποσοστό ακρίβειας 52% και στον ρεαλισμό 32 με ποσοστό ακρίβειας 64% αντίστοιχα.

Ο επόμενος ταξινομητής είναι ο Medium SVM με ποσοστό ακρίβειας 63.2%.



Εικόνα 32 - Confussion matrix Medium SVM

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Medium SVM στον κλασικισμό αναγνώρισε 35 πίνακες και είχαμε ποσοστό ακρίβειας 70%, στον κυβισμό 35 πίνακες με ποσοστό ακρίβειας 70%, στον εξπρεσιονισμό 35 με ποσοστό ακρίβειας 70%, στον μπρεσιονισμό 29 με ποσοστό ακρίβειας 58% και στον ρεαλισμό 24 με ποσοστό ακρίβειας 48% αντίστοιχα.

Ο επόμενος ταξινομητής είναι ο Weighted KNN με ποσοστό ακρίβειας 56.4%.

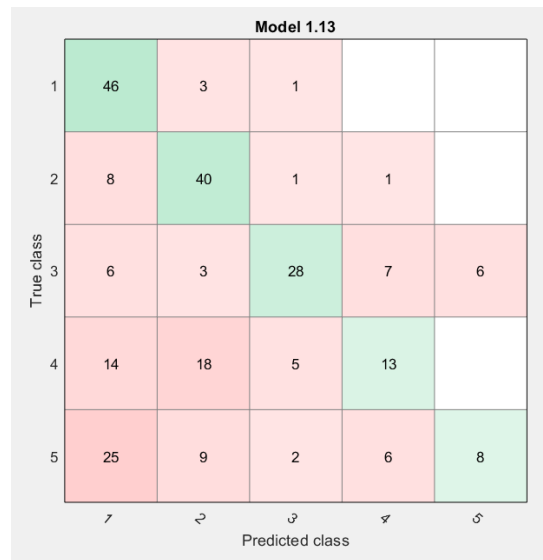
Model 1.19

1	41	4		1	4
2	5	40	1	3	1
3	6	5	21	10	8
4	11	16	2	17	4
5	16	8	1	3	22
	ζ	ς	ϑ	Ϙ	ϙ
	Predicted class				

Εικόνα 33 - Confussion matrix Weighted KNN

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Weighted KNN στον κλασικισμό αναγνώρισε 41 πίνακες και είχαμε ποσοστό ακρίβειας 82%, στον κυβισμό 40 πίνακες με ποσοστό ακρίβειας 80%, στον εξπρεσιονισμό 21 με ποσοστό ακρίβειας 42%, στον μπρεσιονισμό 17 με ποσοστό ακρίβειας 34% και στον ρεαλισμό 22 με ποσοστό ακρίβειας 44% αντίστοιχα.

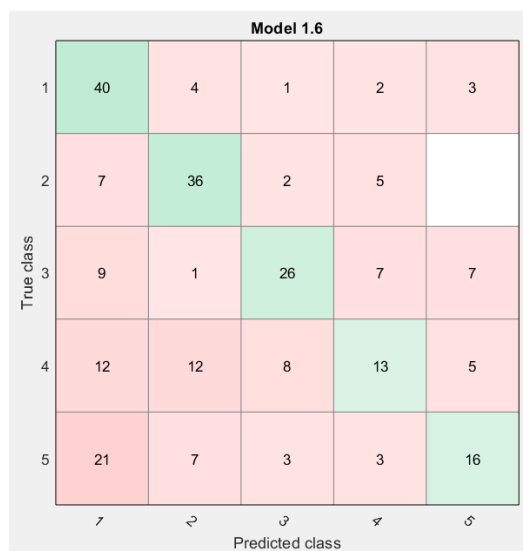
Ο επόμενος ταξινομητής είναι ο Coarse Gaussian SVM με ποσοστό ακρίβειας 54%.



Εικόνα 34- Confussion matrix Coarse Gaussian SVM

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Coarse Gaussian SVM στον κλασικισμό αναγνώρισε 46 πίνακες και είχαμε ποσοστό ακρίβειας 92%, στον κυβισμό 40 πίνακες με ποσοστό ακρίβειας 80%, στον εξπρεσιονισμό 28 με ποσοστό ακρίβειας 56%, στον μπρεσιονισμό 13 με ποσοστό ακρίβειας 26% και στον ρεαλισμό 8 με ποσοστό ακρίβειας 16% αντίστοιχα.

Ο επόμενος ταξινομητής είναι ο Gaussian Naïve Bayes με ποσοστό ακρίβειας 52.4%.



Εικόνα 35 - Confussion matrix Gaussian Naïve Bayes

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Gaussian Naïve Bayes στον κλασικισμό αναγνώρισε 40 πίνακες και είχαμε ποσοστό ακρίβειας 80%, στον κυβισμό 36 πίνακες με ποσοστό ακρίβειας 72%, στον εξπρεσιονισμό 26 με ποσοστό ακρίβειας 52%, στον ιμπρεσιονισμό 13 με ποσοστό ακρίβειας 26% και στον ρεαλισμό 16 με ποσοστό ακρίβειας 32% αντίστοιχα.

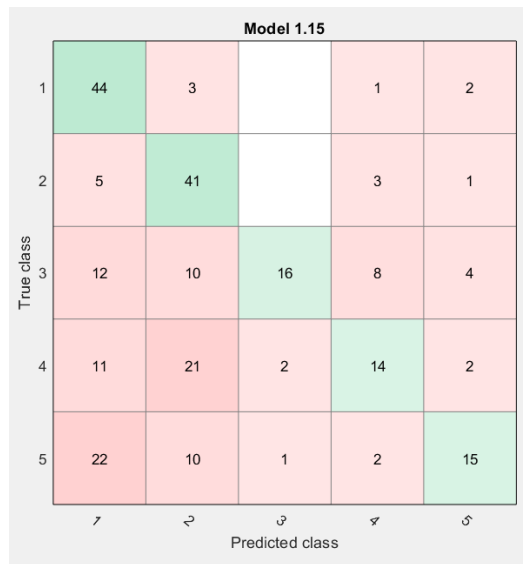
Ο επόμενος ταξινομητής έχει το ίδιο ποσοστό ακρίβειας (δηλαδή 52.4%) με τον προηγούμενο αλλά η ταξινόμηση των πινάκων είναι διαφορετική όπου είναι ο Cosine KNN.

1	33	4	2	3	8
2	3	39	3	4	1
3	5	2	30	8	5
4	8	19	8	11	4
5	11	11	3	7	18
	γ	δ	ε	ς	ζ
True class	Predicted class				

Εικόνα 36 - Confussion matrix Cosine KNN

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Cosine KNN στον κλασικισμό αναγνώρισε 33 πίνακες και είχαμε ποσοστό ακρίβειας 66%, στον κυβισμό 39 πίνακες με ποσοστό ακρίβειας 78%, στον εξπρεσιονισμό 30 με ποσοστό ακρίβειας 60%, στον ιμπρεσιονισμό 11 με ποσοστό ακρίβειας 22% και στον ρεαλισμό 18 με ποσοστό ακρίβειας 36% αντίστοιχα.

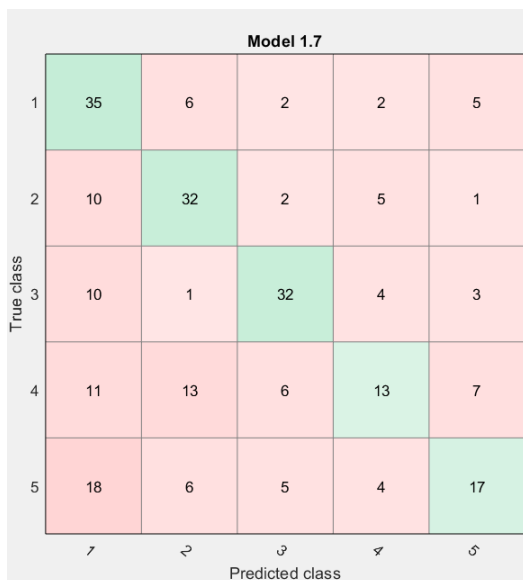
Ο επόμενος ταξινομητής είναι ο Medium KNN με ποσοστό ακρίβειας 52%.



Εικόνα 37 - Confussion matrix Medium KNN

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Medium KNN στον κλασικισμό αναγνώρισε 44 πίνακες και είχαμε ποσοστό ακρίβειας 88%, στον κυβισμό 41 πίνακες με ποσοστό ακρίβειας 82%, στον εξπρεσιονισμό 16 με ποσοστό ακρίβειας 32%, στον μπρεσιονισμό 14 με ποσοστό ακρίβειας 28% και στον ρεαλισμό 15 με ποσοστό ακρίβειας 30% αντίστοιχα.

Ο επόμενος ταξινομητής είναι ο Kernel Naïve Bayes με ποσοστό ακρίβειας 51.6%.



Εικόνα 38 - Confussion matrix Kernel Naïve Bayes

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Kernel Naïve Bayes στον κλασικισμό αναγνώρισε 35 πίνακες και είχαμε ποσοστό ακρίβειας 70%, στον κυβισμό 32 πίνακες με ποσοστό ακρίβειας 64%, στον εξπρεσιονισμό 32 με ποσοστό ακρίβειας 64%, στον μπρεσιονισμό 13 με ποσοστό ακρίβειας 26% και στον ρεαλισμό 17 με ποσοστό ακρίβειας 34% αντίστοιχα.

Ο επόμενος ταξινομητής είναι ο Cubic KNN με ποσοστό ακρίβειας 50.8%.

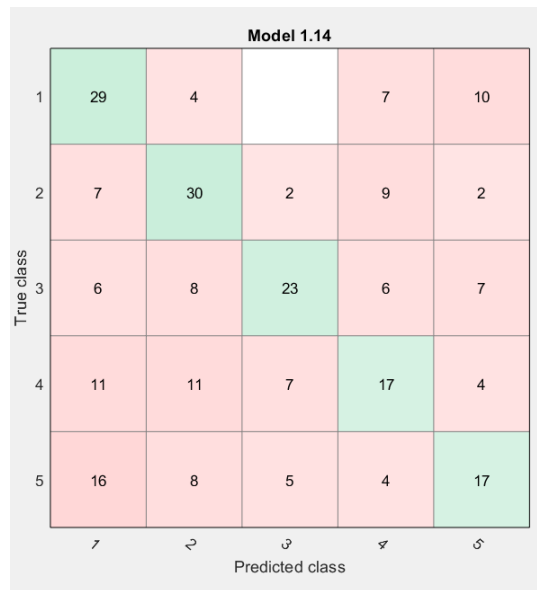
Model 1.18

True class	1	2	3	4	5
1	45	3		1	1
2	5	39	1	3	2
3	9	8	18	10	5
4	13	20	2	13	2
5	22	10	2	4	12
	γ	β	α	δ	ε
	Predicted class				

Εικόνα 39 - Confussion matrix Cubic KNN

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Cubic KNN στον κλασικισμό αναγνώρισε 45 πίνακες και είχαμε ποσοστό ακρίβειας 90%, στον κυβισμό 39 πίνακες με ποσοστό ακρίβειας 78%, στον εξπρεσιονισμό 18 με ποσοστό ακρίβειας 36%, στον μπρεσιονισμό 13 με ποσοστό ακρίβειας 26% και στον ρεαλισμό 12 με ποσοστό ακρίβειας 24% αντίστοιχα.

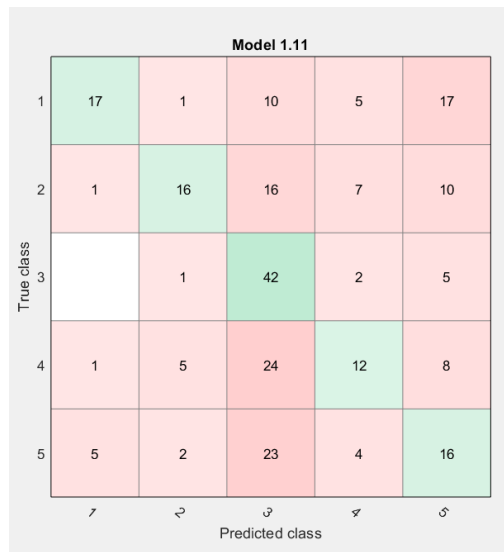
Ο επόμενος ταξινομητής είναι ο Fine KNN με ποσοστό ακρίβειας 46.4%.



Εικόνα 40 - Confussion matrix Fine KNN

Στον συγκεκριμένο πίνακα σύγχυσης του ταξινομητή Fine KNN στον κλασικισμό αναγνώρισε 29 πίνακες και είχαμε ποσοστό ακρίβειας 58%, στον κυβισμό 30 πίνακες με ποσοστό ακρίβειας 60%, στον εξπρεσιονισμό 23 με ποσοστό ακρίβειας 46%, στον μπρεσιονισμό 17 με ποσοστό ακρίβειας 34% και στον ρεαλισμό 17 με ποσοστό ακρίβειας 34% αντίστοιχα.

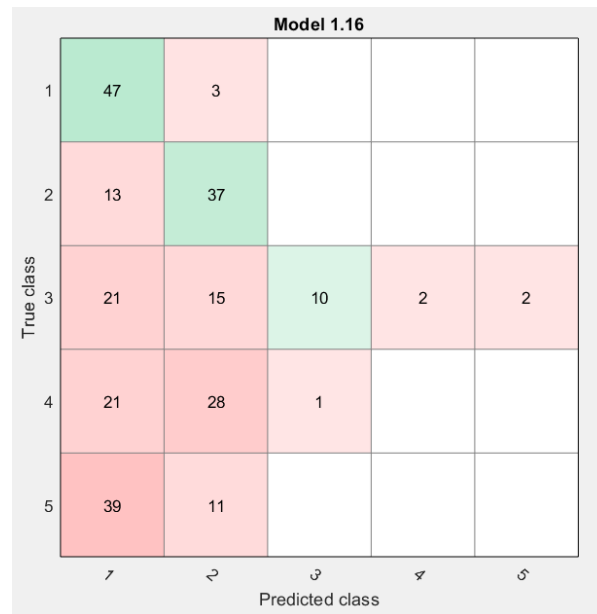
Ο επόμενος ταξινομητής είναι ο Fine Gaussian SVM με ποσοστό ακρίβειας 41.2%.



Εικόνα 41 - Confussion matrix Fine Gaussian SVM

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Fine Gaussian SVM στον κλασικισμό αναγνώρισε 17 πίνακες και είχαμε ποσοστό ακρίβειας 34%, στον κυβισμό 16 πίνακες με ποσοστό ακρίβειας 32%, στον εξπρεσιονισμό 42 με ποσοστό ακρίβειας 84%, στον ιμπρεσιονισμό 12 με ποσοστό ακρίβειας 24% και στον ρεαλισμό 16 με ποσοστό ακρίβειας 32% αντίστοιχα.

Τέλος, ο ταξινομητής με τα λιγότερο καλά αποτελέσματα είναι ο Coarse KNN με ποσοστό ακρίβειας 37.6%.



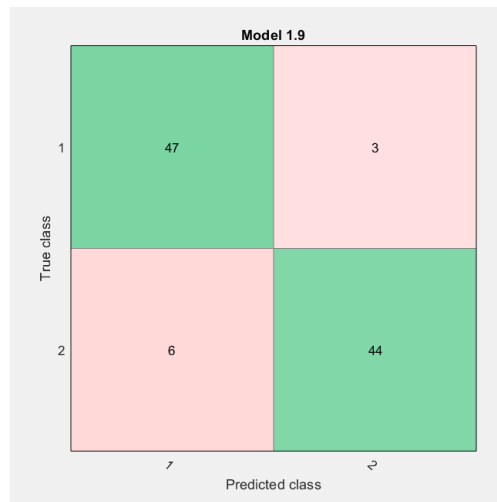
Εικόνα 42 - Confussion matrix Coarse KNN

Στον συγκεκριμένο πίνακα σύγκρισης του ταξινομητή Coarse KNN στον κλασικισμό αναγνώρισε 47 πίνακες και είχαμε ποσοστό ακρίβειας 94%, στον κυβισμό 37 πίνακες με ποσοστό ακρίβειας 74%, στον εξπρεσιονισμό 10 με ποσοστό ακρίβειας 20%, στον ιμπρεσιονισμό 0 με ποσοστό ακρίβειας 0% και στον ρεαλισμό 0 με ποσοστό ακρίβειας 0% αντίστοιχα.

3.7.2.2 Αποτελέσματα μεταξύ 2 καλλιτεχνικών ρευμάτων

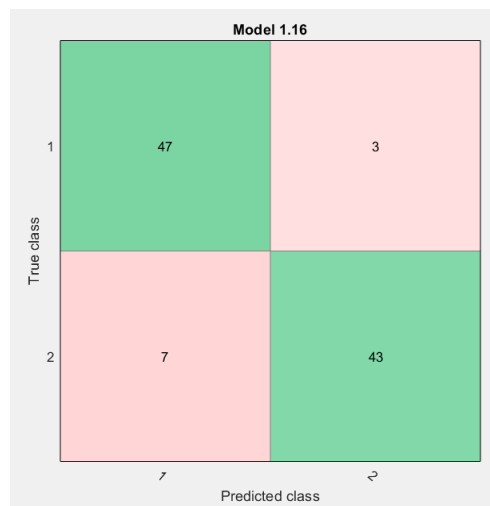
Σε αυτή την ενότητα θα συγκρίνουμε τα αποτελέσματα ανά 2 καλλιτεχνικά ρεύματα.

Πρώτα συγκρίναμε τον κλασικισμό και τον κυβισμό. Ο ταξινομητής Linear SVM είχε ποσοστό επιτυχίας 91%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, στον κλασικισμό αναγνώρισε 47 πίνακες ζωγραφικής με ποσοστό ακρίβειας 94% και στον κυβισμό 44 πίνακες με ποσοστό ακρίβειας 88% αντίστοιχα.



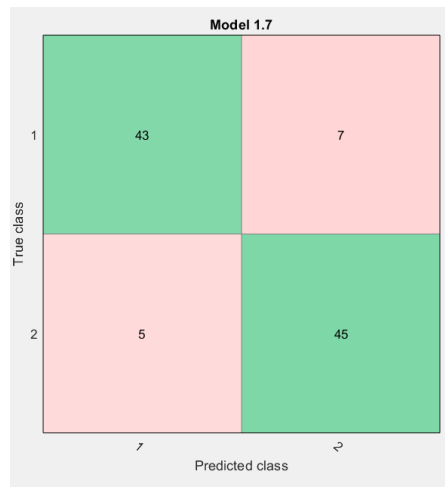
Εικόνα 43 - Confussion matrix- 2 Ρεύματα Linear SVM

Ο ταξινομητής Medium KNN είχε ποσοστό επιτυχίας 90%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, στον κλασικισμό αναγνώρισε 47 πίνακες ζωγραφικής με ποσοστό ακρίβειας 94% και στον κυβισμό 43 πίνακες με ποσοστό ακρίβειας 86% αντίστοιχα.



Εικόνα 44 - Confussion matrix- 2 Ρεύματα Medium KNN

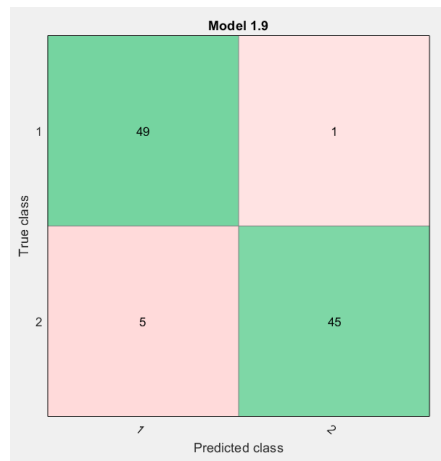
Ο ταξινομητής Gaussian Naïve Bayes είχε ποσοστό επιτυχίας 88%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κλασικισμό αναγνώρισε 43 πίνακες ζωγραφικής με ποσοστό ακρίβειας 86% και στον κυβισμό 45 πίνακες με ποσοστό ακρίβειας 90% αντίστοιχα.



Εικόνα 45 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes

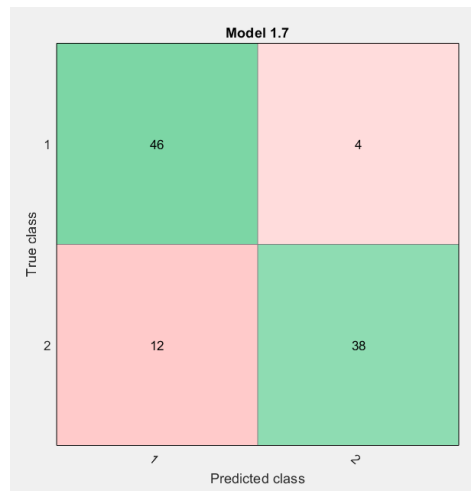
Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως Medium Gaussian SVM με ποσοστό επιτυχίας 90%, ο Quadratic SVM, ο Coarse Gaussian SVM, Cosine KNN, Coarse KNN με ποσοστό επιτυχίας 89%, ο Cubic SVM και ο Weighted KNN με ποσοστό επιτυχίας 88%, ο Kernel Naïve Bayes και ο Fine KNN με ποσοστό επιτυχίας 86%, ο Fine Gaussian SVM με 67% και τέλος ο Coarse KNN με 50%.

Έπειτα, συγκρίναμε τον κλασικισμό και τον εξπρεσιονισμό. Ο ταξινομητής Linear SVM είχε ποσοστό επιτυχίας 94%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κλασικισμό αναγνώρισε 49 πίνακες ζωγραφικής με ποσοστό ακρίβειας 98% και στον εξπρεσιονισμό 45 πίνακες με ποσοστό ακρίβειας 90% αντίστοιχα.



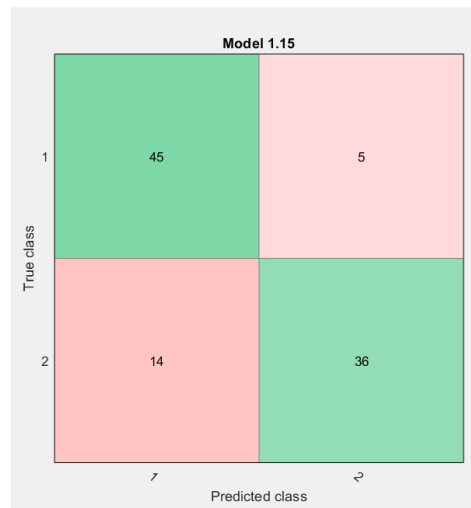
Εικόνα 46 - Confussion matrix- 2 Ρεύματα Linear SVM

Ο ταξινομητής Gaussian Naïve Bayes είχε ποσοστό επιτυχίας 84%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, στον κλασικισμό αναγνώρισε 46 πίνακες ζωγραφικής με ποσοστό ακρίβειας 92% και στον εξπρεσιονισμό 38 πίνακες με ποσοστό ακρίβειας 76% αντίστοιχα.



Εικόνα 47 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes

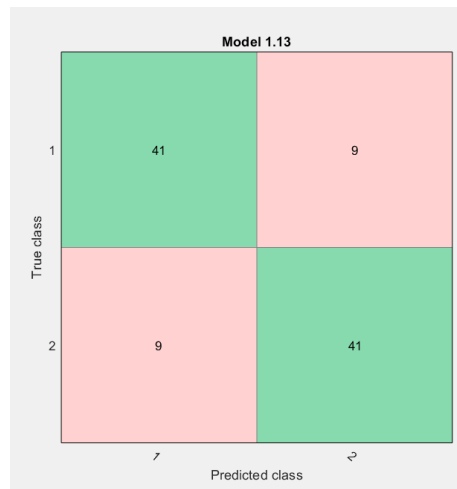
Ο ταξινομητής Fine KNN είχε ποσοστό επιτυχίας 81%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, στον κλασικισμό αναγνώρισε 45 πίνακες ζωγραφικής με ποσοστό ακρίβειας 90% και στον κυβισμό 36 πίνακες με ποσοστό ακρίβειας 72% αντίστοιχα.



Εικόνα 48 - Confussion matrix- 2 Ρεύματα Fine KNN

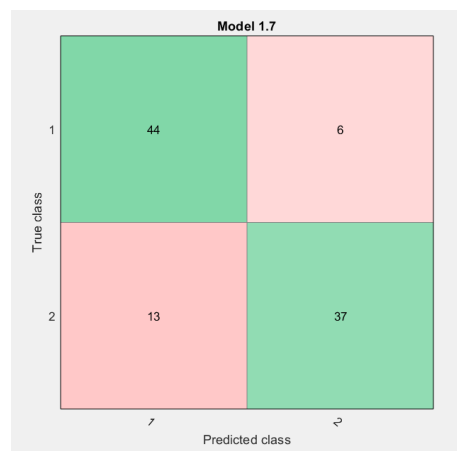
Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως Medium Gaussian SVM και ο Quadratic SVM με ποσοστό επιτυχίας 93%, ο Cubic SVM με ποσοστό επιτυχίας 87%, ο Coarse Gaussian SVM με ποσοστό επιτυχίας 86%, ο Kernel Naïve Bayes με ποσοστό επιτυχίας 82%, ο Cosine KNN και ο Wighted KNN με ποσοστό επιτυχίας 81%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 79%, ο Medium KNN με ποσοστό επιτυχίας 78%, ο Cubic KNN με ποσοστό επιτυχίας 71%, και τέλος τον Coarse KNN με 50% ποσοστό επιτυχίας.

Έπειτα, συγκρίναμε τον κλασικισμό και τον μπρεσιονισμό. Ο ταξινομητής Medium Gaussian SVM είχε ποσοστό επιτυχίας 82%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κλασικισμό αναγνώρισε 41 πίνακες ζωγραφικής με ποσοστό ακρίβειας 82% και στον μπρεσιονισμό 41 πίνακες με ποσοστό ακρίβειας 82% αντίστοιχα.



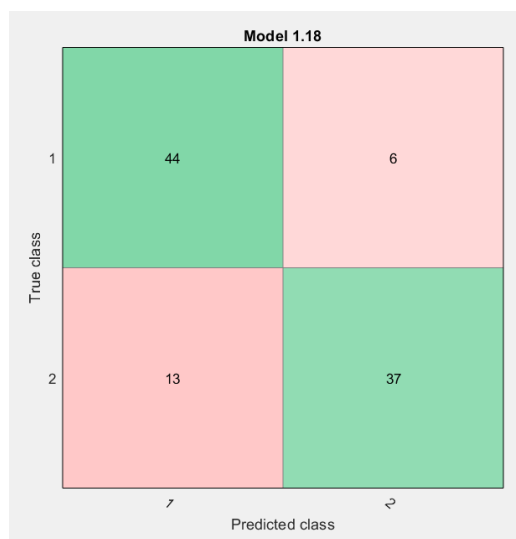
Εικόνα 49 - Confussion matrix- 2 Ρεύματα Medium Gaussian SVM

Ο ταξινομητής Gaussian Naïve Bayes είχε ποσοστό επιτυχίας 81%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, στον κλασικισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον μπρεσιονισμό 37 πίνακες με ποσοστό ακρίβειας 74% αντίστοιχα.



Εικόνα 50 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes

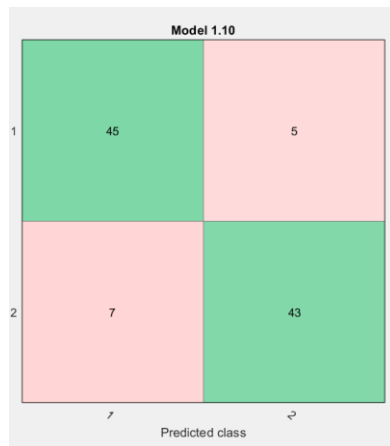
Ο ταξινομητής Cosine KNN είχε ποσοστό επιτυχίας 81%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, στον κλασικισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον μπρεσιονισμό 37 πίνακες με ποσοστό ακρίβειας 74% αντίστοιχα.



Εικόνα 51 - Confussion matrix- 2 Ρεύματα Cosine KNN

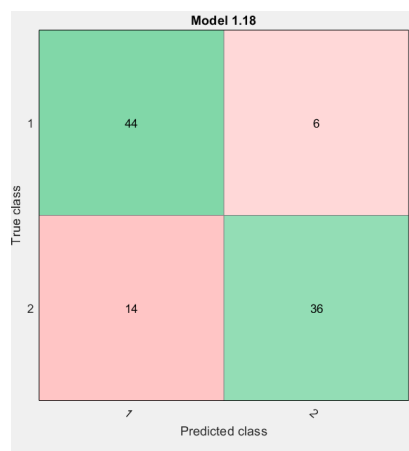
Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως Coarse Gaussian SVM με ποσοστό επιτυχίας 81%, ο Linear SVM με ποσοστό επιτυχίας 79%, ο Cubic SVM και ο Weighted KNN με ποσοστό επιτυχίας 77%, ο Quadratic SVM με ποσοστό επιτυχίας 76%, ο Kernel Naïve Bayes και ο Cubic KNN με ποσοστό επιτυχίας 75%, ο Medium KNN με ποσοστό επιτυχίας 74%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 73%, ο Fine KNN με ποσοστό επιτυχίας 70%, και τέλος με το λιγότερο ποσοστό επιτυχίας ο Coarse KNN με 50%.

Έπειτα, συγκρίναμε τον κλασικισμό και τον ρεαλισμό. Ο ταξινομητής Quadratic SVM είχε ποσοστό επιτυχίας 88%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κλασικισμό αναγνώρισε 45 πίνακες ζωγραφικής με ποσοστό ακρίβειας 90% και στον ρεαλισμό 43 πίνακες με ποσοστό ακρίβειας 86% αντίστοιχα.



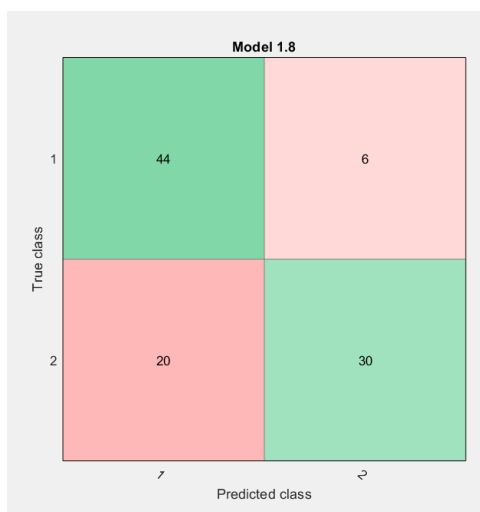
Εικόνα 52 - Confussion matrix- 2 Ρεύματα Quadratic SVM

Ο ταξινομητής Cosine KNN είχε ποσοστό επιτυχίας 80%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κλασικισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον ρεαλισμό 36 πίνακες με ποσοστό ακρίβειας 72% αντίστοιχα.



Εικόνα 53 - Confussion matrix- 2 Ρεύματα Cosine KNN

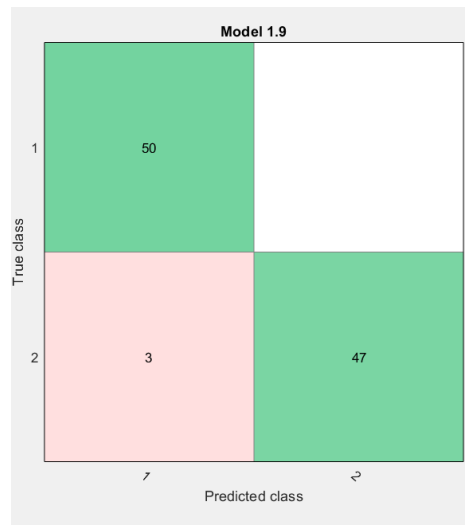
Ο ταξινομητής Kernel Naïve Bayes είχε ποσοστό επιτυχίας 74%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κλασικισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον ρεαλισμό 30 πίνακες με ποσοστό ακρίβειας 60% αντίστοιχα.



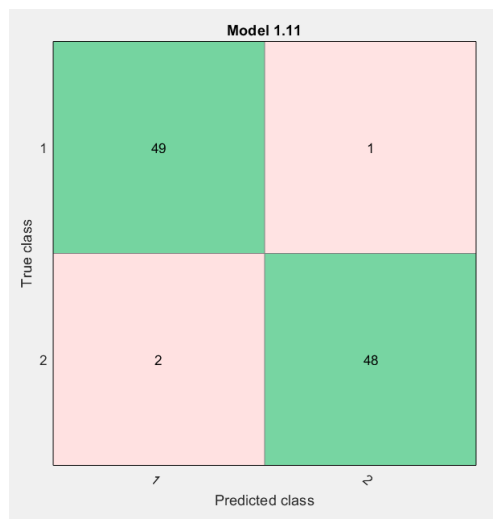
Εικόνα 54 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes

Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Linear SVM με ποσοστό επιτυχίας 86%, ο Cubic SVM και ο Coarse Gaussian SVM με ποσοστό επιτυχίας 85%, ο Weighted KNN με ποσοστό επιτυχίας 78%, ο Medium KNN με ποσοστό επιτυχίας 75%, ο Cubic KNN με ποσοστό επιτυχίας 72%, ο Gaussian Naïve Bayes με ποσοστό επιτυχίας 70%, ο Fine KNN με ποσοστό επιτυχίας 69%, ο Coarse Gaussian SVM με ποσοστό επιτυχίας 68%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 63%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

Έπειτα, συγκρίναμε τον κυβισμό και τον εξπρεσιονισμό. Ο ταξινομητής Cubic SVM και ο Linear SVM είχε ποσοστό επιτυχίας 97%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, ο Linear SVM στον κυβισμό αναγνώρισε 50 πίνακες ζωγραφικής με ποσοστό ακρίβειας 100% και στον εξπρεσιονισμό 47 πίνακες με ποσοστό ακρίβειας 94% αντίστοιχα. Αντίθετα, ο Cubic SVM στον κυβισμό αναγνώρισε 49 πίνακες ζωγραφικής με ποσοστό ακρίβειας 98% και στον εξπρεσιονισμό 48 πίνακες με ποσοστό ακρίβειας 96% αντίστοιχα.



Εικόνα 55 - Confussion matrix- 2 Ρεύματα Linear SVM

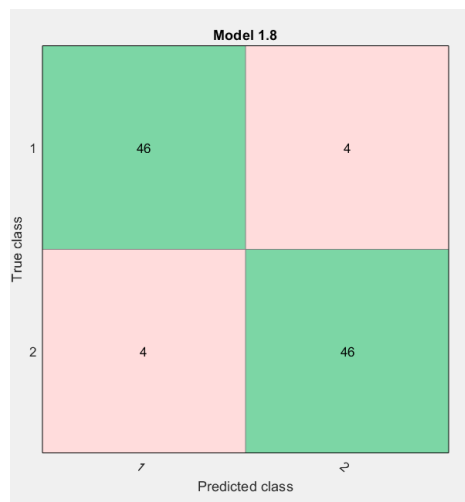


Εικόνα 56 - Confussion matrix- 2 Ρεύματα Cubic SVM

Ο ταξινομητής Kernel και ο Gaussian Naïve Bayes είχαν ποσοστό επιτυχίας 92%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κυβισμό αναγνώρισε 46 πίνακες ζωγραφικής με ποσοστό ακρίβειας 92% και στον εξπρεσιονισμό 46 πίνακες με ποσοστό ακρίβειας 92% αντίστοιχα. Το ότι αυτοί οι δύο ταξινομητές έχουν ακριβώς τα ίδια αποτελέσματα, δεν σημαίνει ότι έχουν βρει τις ίδιες εικόνες. Απλά το πλήθος των εικόνων που βρήκε είναι το ίδιο.

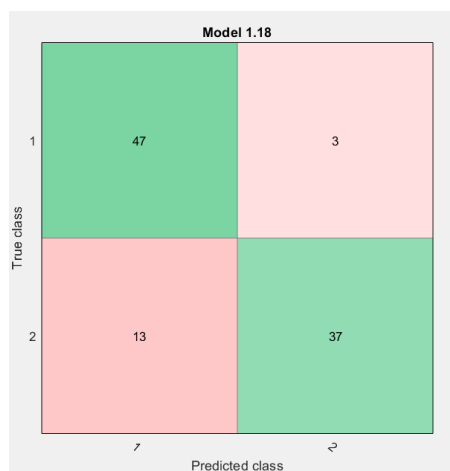


Εικόνα 57 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes



Εικόνα 58 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes

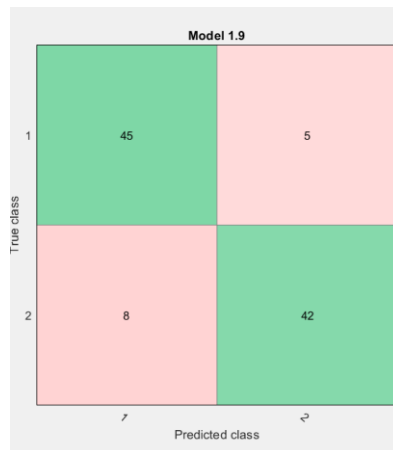
Ο ταξινομητής Cosine KNN είχε ποσοστό επιτυχίας 84%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγχυσης, στον κυβισμό αναγνώρισε 47 πίνακες ζωγραφικής με ποσοστό ακρίβειας 94% και στον εξπρεσιονισμό 37 πίνακες με ποσοστό ακρίβειας 74% αντίστοιχα.



Εικόνα 59 - Confusion matrix- 2 Ρεύματα Cosine KNN

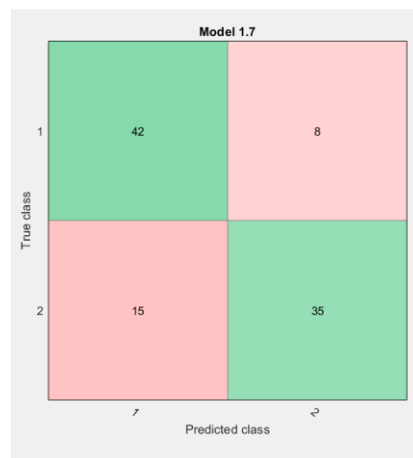
Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Quadratic SVM με ποσοστό επιτυχίας 96%, ο Medium Gaussian SVM με ποσοστό επιτυχίας 94%, ο Coarse Gaussian SVM με ποσοστό επιτυχίας 93%, ο Weighted KNN με ποσοστό επιτυχίας 83%, ο Fine KNN με ποσοστό επιτυχίας 78%, ο Medium KNN με ποσοστό επιτυχίας 77%, ο Cubic KNN με ποσοστό επιτυχίας 74%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 73%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

Έπειτα, συγκρίναμε τον κυβισμό και τον μπρεσιονισμό. Ο ταξινομητής Linear SVM είχε ποσοστό επιτυχίας 87%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκυσης, ο Linear SVM στον κυβισμό αναγνώρισε 45 πίνακες ζωγραφικής με ποσοστό ακρίβειας 90% και στον μπρεσιονισμό 42 πίνακες με ποσοστό ακρίβειας 84% αντίστοιχα.



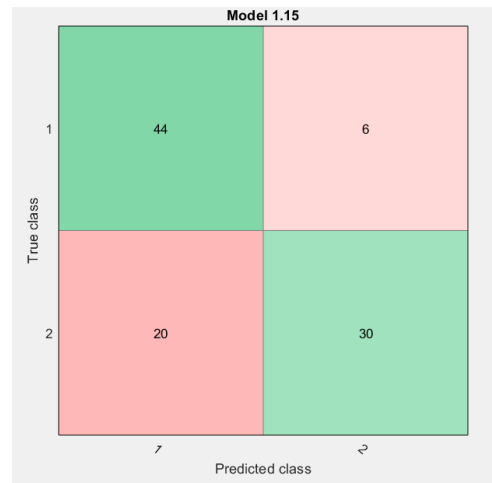
Εικόνα 60 - Confussion matrix- 2 Ρεύματα Linear SVM

Ο ταξινομητής Gaussian Naïve Bayes είχε ποσοστό επιτυχίας 77%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, ο Gaussian Naïve Bayes στον κυβισμό αναγνώρισε 42 πίνακες ζωγραφικής με ποσοστό ακρίβειας 84% και στον ιμπρεσιονισμό 35 πίνακες με ποσοστό ακρίβειας 70% αντίστοιχα.



Εικόνα 61 - Confussion matrix- 2 Ρεύματα Gaussian Naïve Bayes

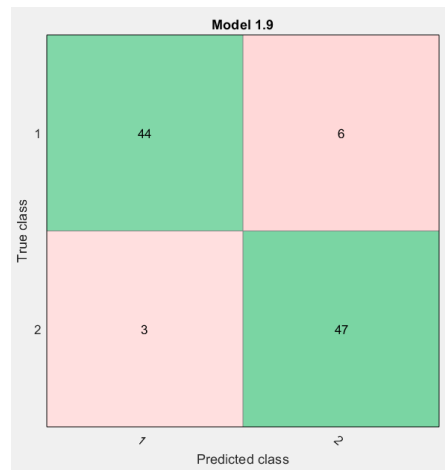
Ο ταξινομητής Fine KNN είχε ποσοστό επιτυχίας 74%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κυβισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον ιμπρεσιονισμό 30 πίνακες με ποσοστό ακρίβειας 60% αντίστοιχα.



Εικόνα 62 - Confusion matrix- 2 Ρεύματα Fine KNN

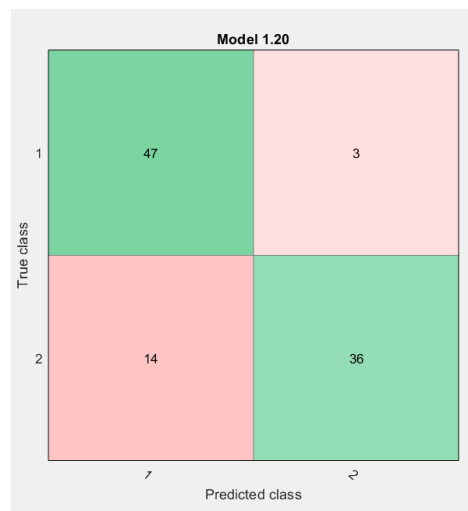
Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Quadratic SVM με ποσοστό επιτυχίας 85%, ο Medium Gaussian SVM με ποσοστό επιτυχίας 81%, ο Cubic SVM με ποσοστό επιτυχίας 80%, ο Cosine KNN με ποσοστό επιτυχίας 74%, ο Kernel Naïve Bayes με ποσοστό επιτυχίας 73%, ο Weighted KNN με ποσοστό επιτυχίας 72%, ο Medium KNN και ο Coarse Gaussian SVM με ποσοστό επιτυχίας 71%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 69%, ο Cubic KNN με ποσοστό επιτυχίας 68%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

Έπειτα, συγκρίναμε τον κυβισμό και τον ρεαλισμό. Ο ταξινομητής Linear SVM είχε ποσοστό επιτυχίας 91%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, ο Linear SVM στον κυβισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον ρεαλισμό 47 πίνακες με ποσοστό ακρίβειας 94% αντίστοιχα.



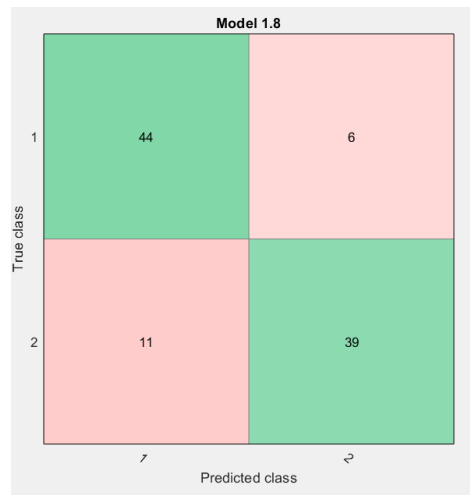
Εικόνα 63 - Confussion matrix- 2 Ρεύματα Linear SVM

Ο ταξινομητής Weighted KNN είχε ποσοστό επιτυχίας 83%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κυβισμό αναγνώρισε 47 πίνακες ζωγραφικής με ποσοστό ακρίβειας 94% και στον ρεαλισμό 36 πίνακες με ποσοστό ακρίβειας 72% αντίστοιχα.



Εικόνα 64 - Confussion matrix- 2 Ρεύματα Weighted KNN

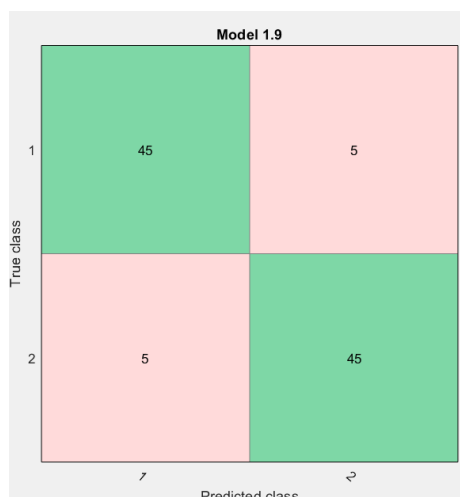
Ο ταξινομητής Kernel Naïve Bayes είχε ποσοστό επιτυχίας 83%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον κυβισμό αναγνώρισε 44 πίνακες ζωγραφικής με ποσοστό ακρίβειας 88% και στον ρεαλισμό 39 πίνακες με ποσοστό ακρίβειας 78% αντίστοιχα.



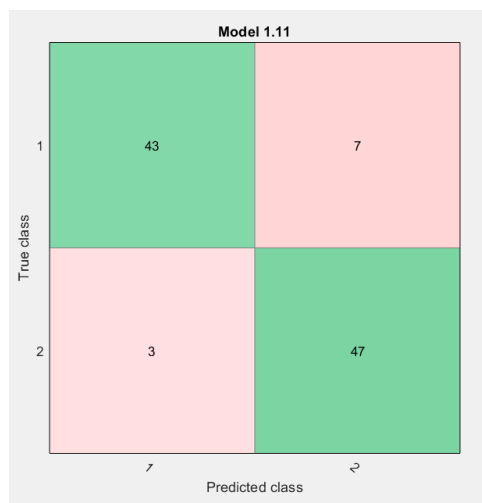
Εικόνα 65 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes

Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Medium Gaussian SVM με ποσοστό επιτυχίας 91%, ο Quadratic SVM με ποσοστό επιτυχίας 90%, ο Cubic SVM με ποσοστό επιτυχίας 87%, ο Gaussian Naïve Bayes με ποσοστό επιτυχίας 83%, ο Fine KNN με ποσοστό επιτυχίας 82%, ο Medium KNN με ποσοστό επιτυχίας 80%, ο Coarse Gaussian SVM με ποσοστό επιτυχίας 79%, ο Cubic KNN με ποσοστό επιτυχίας 78%, ο Cosine KNN με ποσοστό επιτυχίας 77%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 68%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

Έπειτα, συγκρίναμε τον εξπρεσιονισμό και τον μπρεσιονισμό. Ο ταξινομητής Cubic SVM και ο Linear SVM είχε ποσοστό επιτυχίας 90%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, ο Linear SVM στον εξπρεσιονισμό αναγνώρισε 45 πίνακες ζωγραφικής με ποσοστό ακρίβειας 90% και στον μπρεσιονισμό 45 πίνακες με ποσοστό ακρίβειας 90% αντίστοιχα. Αντίθετα, ο Cubic SVM στον εξπρεσιονισμό αναγνώρισε 43 πίνακες ζωγραφικής με ποσοστό ακρίβειας 86% και στον μπρεσιονισμό 47 πίνακες με ποσοστό ακρίβειας 94% αντίστοιχα.

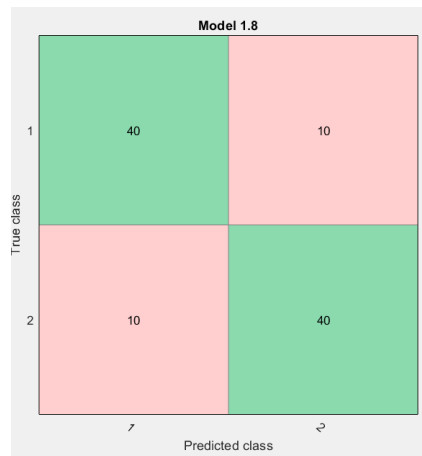


Εικόνα 66 - Confussion matrix- 2 Ρεύματα Linear SVM



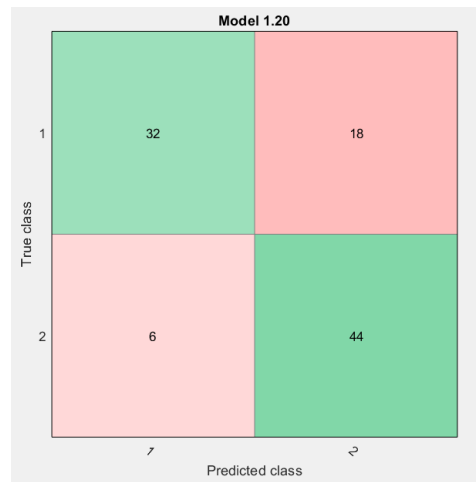
Εικόνα 67 - Confussion matrix- 2 Ρεύματα Cubic SVM

Ο ταξινομητής Kernel Naïve Bayes είχε ποσοστό επιτυχίας 80%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγχυσης, στον εξπρεσιονισμό αναγνώρισε 40 πίνακες ζωγραφικής με ποσοστό ακρίβειας 80% και στον μπρεσιονισμό 40 πίνακες με ποσοστό ακρίβειας 80% αντίστοιχα.



Εικόνα 68 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes

Ο ταξινομητής Weighted KNN είχε ποσοστό επιτυχίας 76%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγχυσης, στον εξπρεσιονισμό αναγνώρισε 32 πίνακες ζωγραφικής με ποσοστό ακρίβειας 64% και στον μπρεσιονισμό 44 πίνακες με ποσοστό ακρίβειας 88% αντίστοιχα.



Εικόνα 69 - Confussion matrix- 2 Ρεύματα Weighted KNN

Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Quadratic SVM με ποσοστό επιτυχίας 88%, ο Medium Gaussian SVM με ποσοστό επιτυχίας 85%, ο Coarse Gaussian SVM με ποσοστό επιτυχίας 74%, ο Gaussian Naïve Bayes και ο Medium KNN με ποσοστό επιτυχίας 72%, ο Fine KNN με ποσοστό επιτυχίας 71%, ο Cubic

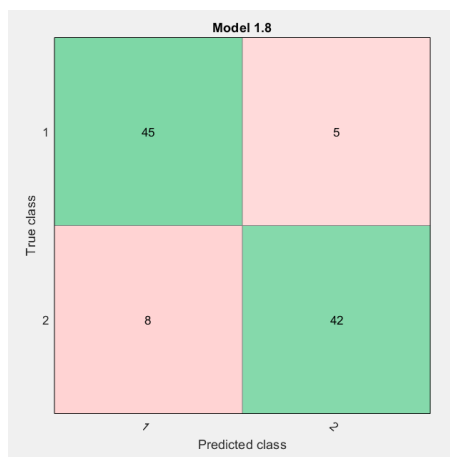
KNN και ο Cosine KNN με ποσοστό επιτυχίας 70%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 65%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

Έπειτα, συγκρίναμε τον εξπρεσιονισμό και τον ρεαλισμό. Ο ταξινομητής Linear SVM είχε ποσοστό επιτυχίας 94%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, ο Linear SVM στον εξπρεσιονισμό αναγνώρισε 45 πίνακες ζωγραφικής με ποσοστό ακρίβειας 90% και στον ρεαλισμό 49 πίνακες με ποσοστό ακρίβειας 98% αντίστοιχα.



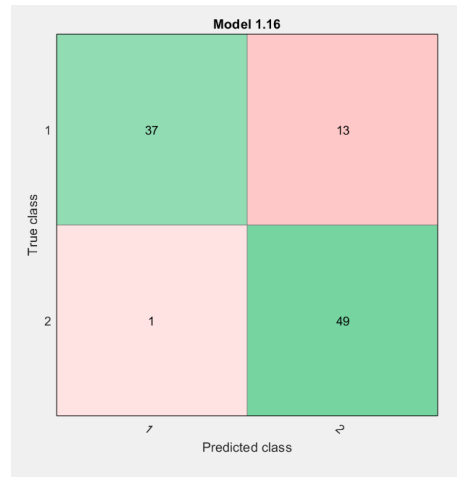
Εικόνα 70 - Confussion matrix- 2 Ρεύματα Linear SVM

Ο ταξινομητής Kernel Naïve Bayes είχε ποσοστό επιτυχίας 87%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον εξπρεσιονισμό αναγνώρισε 45 πίνακες ζωγραφικής με ποσοστό ακρίβειας 90% και στον ρεαλισμό 42 πίνακες με ποσοστό ακρίβειας 84% αντίστοιχα.



Εικόνα 71 - Confussion matrix- 2 Ρεύματα Kernel Naïve Bayes

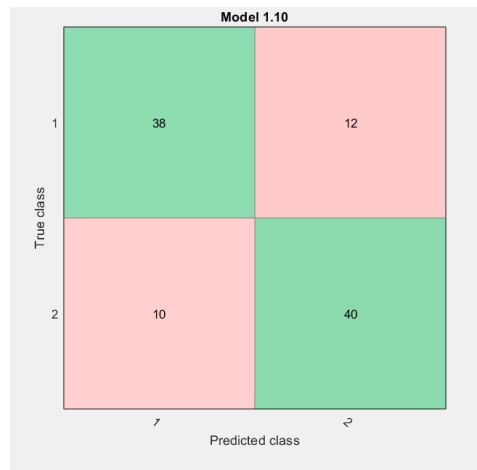
Ο ταξινομητής Medium KNN είχε ποσοστό επιτυχίας 86%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον εμπρεσιονισμό αναγνώρισε 37 πίνακες ζωγραφικής με ποσοστό ακρίβειας 74% και στον ρεαλισμό 49 πίνακες με ποσοστό ακρίβειας 98% αντίστοιχα.



Εικόνα 72 - Confussion matrix- 2 Ρεύματα Medium KNN

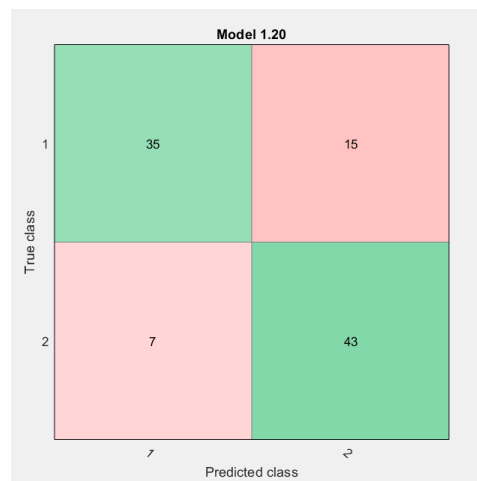
Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Quadratic SVM με ποσοστό επιτυχίας 92%, ο Cubic SVM με ποσοστό επιτυχίας 89%, ο Weighted KNN με ποσοστό επιτυχίας 85%, ο Medium Gaussian SVM με ποσοστό επιτυχίας 84%, ο Cubic KNN, ο Coarse Gaussian SVM και ο Cosine KNN με ποσοστό επιτυχίας 81%, ο Gaussian Naïve Bayes και ο Fine KNN με ποσοστό επιτυχίας 80%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 74%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

Έπειτα, συγκρίναμε τον μπρεσιονισμό και τον ρεαλισμό. Ο ταξινομητής Quadratic SVM είχε ποσοστό επιτυχίας 78%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, ο Quadratic SVM στον εμπρεσιονισμό αναγνώρισε 38 πίνακες ζωγραφικής με ποσοστό ακρίβειας 76% και στον ρεαλισμό 40 πίνακες με ποσοστό ακρίβειας 80% αντίστοιχα.



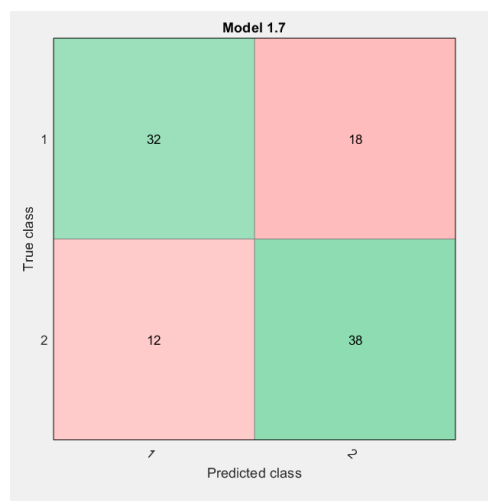
Εικόνα 73 - Confussion matrix- 2 Ρεύματα Quadratic SVM

Ο ταξινομητής Weighted KNN είχε ποσοστό επιτυχίας 78%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον μπρεσιονισμό αναγνώρισε 35 πίνακες ζωγραφικής με ποσοστό ακρίβειας 70% και στον ρεαλισμό 43 πίνακες με ποσοστό ακρίβειας 86% αντίστοιχα.



Εικόνα 74 - Confussion matrix- 2 Ρεύματα Weighted KNN

Ο ταξινομητής Gaussian Naïve Bayes είχε ποσοστό επιτυχίας 70%. Συγκεκριμένα, όπως φαίνεται και στον παρακάτω πίνακα σύγκρισης, στον μπρεσιονισμό αναγνώρισε 32 πίνακες ζωγραφικής με ποσοστό ακρίβειας 64% και στον ρεαλισμό 38 πίνακες με ποσοστό ακρίβειας 76% αντίστοιχα.



Εικόνα 75 - Confusion matrix- 2 Ρεύματα Gaussian Naïve Bayes

Παρατηρούμε ότι τα αποτελέσματα είχαν μεγάλο ποσοστό επιτυχίας στους παραπάνω ταξινομητές. Τα αποτελέσματα είναι παρόμοια και στους υπόλοιπους ταξινομητές, όπως ο Cosine KNN με ποσοστό επιτυχίας 77%, ο Medium Gaussian SVM με ποσοστό επιτυχίας 75%, ο Cubic SVM και ο Coarse Gaussian SVM με ποσοστό επιτυχίας 74%, ο Linear SVM με ποσοστό επιτυχίας 73%, ο Medium KNN με ποσοστό επιτυχίας 72%, ο Fine KNN με ποσοστό επιτυχίας 69%, ο Cubic KNN με ποσοστό επιτυχίας 67%, ο Kernel Naïve Bayes με ποσοστό επιτυχίας 66%, ο Fine Gaussian SVM με ποσοστό επιτυχίας 58%, και τέλος ο Coarse KNN με το λιγότερο ποσοστό επιτυχίας 50%.

3.7.2.3 Σύγκριση Αποτελεσμάτων

Στον παρακάτω πίνακα απεικονίζονται τα αποτελέσματα από κάθε ένα από τα 5 καλλιτεχνικά ρεύματα με 40 χαρακτηριστικά. Υπενθυμίζεται ότι για κάθε καλλιτεχνικό ρεύμα χρησιμοποιήθηκαν 50 έργα τέχνης. Στα κελιά του πίνακα αναγράφεται το πλήθος των πινάκων που ταξινομήθηκαν σωστά, για κάθε συγκεκριμένο ρεύμα και κάθε συγκεκριμένο ταξινομητή. Στην τελευταία στήλη του πίνακα αναγράφεται το ποσοστό επιτυχημένης ταξινόμησης καθενός ταξινομητή, ενώ στην τελευταία γραμμή το μέσο ποσοστό σωστής ταξινόμησης ανά καλλιτεχνικό ρεύμα.

5 Καλλιτεχνικά Ρεύματα, 40 Χαρακτηριστικά						
	Κλασικισμός	Κυβισμός	Εξπρεσιονισμός	Ιμπρεσιονισμός	Ρεαλισμός	Σύνολο
Quadratic SVM	39	38	34	29	34	69,6%
Linear SVM	37	38	34	29	34	68,8%
Cubic SVM	37	39	30	26	32	65,6%
Medium Gaussian SVM	35	35	35	29	24	63,2%
Weighted KNN	41	40	21	17	22	56,4%
Coarse Gaussian SVM	46	40	28	13	8	54,0%
Gaussian Naïve Bayes	40	36	26	13	16	52,4%
Cosine KNN	33	39	30	11	18	52,4%
Medium KNN	44	41	16	14	15	52,0%
Kernel Naïve Bayes	35	32	32	13	17	51,6%
Cubic KNN	45	39	18	13	12	50,8%
Fine KNN	29	30	23	17	17	46,4%
Fine Gaussian SVM	17	16	42	12	16	41,2%
Coarse KNN	47	37	10	0	0	37,6%
Σύνολο	74%	71%	57%	36%	41%	

Πίνακας 1 – Αποτελέσματα με 5 καλλιτεχνικά ρεύματα και 40 Χαρακτηριστικά

Όπως παρατηρούμε στον Πίνακα 1, ο Quadratic SVM ήταν ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας 69.6%. Σύμφωνα με τον Πίνακα, ο κλασικισμός είναι το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό 74%, έπειτα ο κυβισμός με 71%, ακολουθεί ο εξπρεσιονισμός με 57%, ο ρεαλισμός με 41% και τέλος ο ιμπρεσιονισμός με λιγότερο ποσοστό επιτυχίας 36%.

Έπειτα συγκρίναμε ανά δύο τα καλλιτεχνικά ρεύματα.

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κλασικισμός - Κυβισμός			
	Κλασικισμός	Κυβισμός	Σύνολο
Linear SVM	47	44	91%
Medium KNN	47	43	90%
Gaussian Naïve Bayes	43	45	88%
Σύνολο	91%	88%	

Πίνακας 2 – Αποτελέσματα Κλασικισμού – Κυβισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κλασικισμός - Εξπρεσιονισμός			
	Κλασικισμός	Εξπρεσιονισμός	Σύνολο
Linear SVM	49	45	94%
Gaussian Naïve Bayes	46	38	84%
Medium KNN	45	36	81%
Σύνολο	93%	79%	

Πίνακας 3 – Αποτελέσματα Κλασικισμού – Εξπρεσιονισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κλασικισμός - Ιμπρεσιονισμός			
	Κλασικισμός	Ιμπρεσιονισμός	Σύνολο
Medium Gaussian SVM	41	41	82%
Gaussian Naïve Bayes	44	37	81%
Cosine KNN	44	37	81%
Σύνολο	86%	77%	

Πίνακας 4 – Αποτελέσματα Κλασικισμού – Ιμπρεσιονισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κλασικισμός - Ρεαλισμός			
	Κλασικισμός	Ρεαλισμός	Σύνολο
Quadratic SVM	45	43	88%
Cosine KNN	44	36	80%
Kernel Naïve Bayes	44	30	74%
Σύνολο	89%	73%	

Πίνακας 5 – Αποτελέσματα Κλασικισμού – Ρεαλισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κυβισμός - Εξπρεσιονισμός			
	Κυβισμός	Εξπρεσιονισμός	Σύνολο
Cubic SVM	49	48	97%
Linear SVM	50	47	97%
Kernel Naïve Bayes	46	46	92%
Gaussian Naïve Bayes	46	46	92%
Cosine KNN	47	37	84%
Σύνολο	95%	90%	

Πίνακας 6 – Αποτελέσματα Κυβισμού – Εξπρεσιονισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κυβισμός - Ιμπρεσιονισμός			
	Κυβισμός	Ιμπρεσιονισμός	Σύνολο
Linear SVM	45	42	87%
Gaussian Naïve Bayes	42	35	77%
Fine KNN	44	30	74%
Σύνολο	87%	71%	

Πίνακας 7 – Αποτελέσματα Κυβισμού – Ιμπρεσιονισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Κυβισμός - Ρεαλισμός			
	Κυβισμός	Ρεαλισμός	Σύνολο
Linear SVM	44	47	91%
Weighted KNN	47	36	83%
Kernel Naïve Bayes	44	39	83%
Σύνολο	90%	81%	

Πίνακας 8 – Αποτελέσματα Κυβισμού – Ρεαλισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Εξπρεσιονισμός - Ιμπρεσιονισμός			
	Εξπρεσιονισμός	Ιμπρεσιονισμός	Σύνολο
Cubic SVM	43	47	90%
Linear SVM	45	45	90%
Kernel Naïve Bayes	40	40	80%
Weighted KNN	32	44	76%
Σύνολο	80%	88%	

Πίνακας 9 – Αποτελέσματα Εξπρεσιονισμού – Ιμπρεσιονισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Εξπρεσιονισμός - Ρεαλισμός			
	Εξπρεσιονισμός	Ρεαλισμός	Σύνολο
Linear SVM	45	49	94%
Kernel Naïve Bayes	45	42	87%
Medium KNN	37	49	86%
Σύνολο	85%	93%	

Πίνακας 10 – Αποτελέσματα Εξπρεσιονισμού – Ρεαλισμού

Ανά 2 Καλλιτεχνικά Ρεύματα με 40 Χαρακτηριστικά			
Ιμπρεσιονισμός - Ρεαλισμός			
	Ιμπρεσιονισμός	Ρεαλισμός	Σύνολο
Quadratic SVM	38	40	78%
Weighted KNN	35	43	78%
Gaussian Naïve Bayes	32	38	70%
Σύνολο	70%	81%	

Πίνακας 11 – Αποτελέσματα Ιμπρεσιονισμού – Ρεαλισμού

Στους παραπάνω 10 πίνακες, συγκρίναμε τα 5 καλλιτεχνικά ρεύματα ανά δύο. Παρατηρούμε ότι τα αποτελέσματα είναι σαφώς καλύτερα από την σύγκριση και με τα 5 καλλιτεχνικά ρεύματα, με ποσοστά επιτυχίας 78% έως και 97%. Αναλυτικότερα, στον πίνακα 2 συγκρίναμε τον κλασικισμό με τον κυβισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Linear SVM με 91%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κλασικισμός με 91%. Στον πίνακα 3 συγκρίναμε τον κλασικισμό με τον εξπρεσιονισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Linear SVM με 94%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κλασικισμός με 93%. Στον πίνακα 4 συγκρίναμε τον κλασικισμό με τον ιμπρεσιονισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Medium Gaussian SVM με 82%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κλασικισμός με 86%. Στον πίνακα 5 συγκρίναμε τον κλασικισμό με τον ρεαλισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Quadratic SVM με 88%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κλασικισμός με 89%. Στον πίνακα 6 συγκρίναμε τον κυβισμό με τον εξπρεσιονισμό, όπου οι ταξινομητές με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Linear SVM και ο Cubic SVM με 97%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κυβισμός με 95%. Στον πίνακα 7 συγκρίναμε τον κυβισμό με τον ιμπρεσιονισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Linear SVM με 87%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κυβισμός με 87%. Στον πίνακα 8 συγκρίναμε τον κυβισμό με τον ρεαλισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Linear SVM με 91%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο κυβισμός με 90%. Στον πίνακα 9 συγκρίναμε τον εξπρεσιονισμό με τον ιμπρεσιονισμό, όπου οι ταξινομητές με το μεγαλύτερο

ποσοστό επιτυχίας ήταν ο Linear SVM και ο Cubic με 90%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο ιμπρεσιονισμός με 88%. Στον πίνακα 10 συγκρίναμε τον εξπρεσιονισμό με τον ρεαλισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Linear SVM με 94%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο ρεαλισμός με 93%. Τέλος, στον πίνακα 11 συγκρίναμε τον ιμπρεσιονισμό με τον ρεαλισμό, όπου ο ταξινομητής με το μεγαλύτερο ποσοστό επιτυχίας ήταν ο Quadratic SVM με 78%, και το καλλιτεχνικό ρεύμα με το μεγαλύτερο ποσοστό ήταν ο ρεαλισμός με 81%.

Με βάση τις παραπάνω συγκρίσεις για τα ανά δύο καλλιτεχνικά ρεύματα, βλέπουμε ότι σε κάθε σύγκριση ο κλασικισμός έχει το μεγαλύτερο ποσοστό, αφού και στις 4 συγκρίσεις με τα άλλα καλλιτεχνικά ρεύματα είχε το καλύτερο ποσοστό επιτυχίας. Έπειτα, ο κυβισμός στις συγκρίσεις με τα 4 καλλιτεχνικά ρεύματα, στις 3 από τις 4 συγκρίσεις είχε τα καλύτερα αποτελέσματα.

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΒΕΛΤΙΩΣΕΙΣ

4.1 ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην διπλωματική εργασία, αναπτύχθηκε σύστημα το οποίο ταξινομεί πίνακες ζωγραφικής με βάση ποικίλα χαρακτηριστικά χαμηλού επιπέδου, όπως χαρακτηριστικά υφής Haralick, πλήθος ακμών, χρωματικά ιστογράμματα κλπ. Πιο συγκεκριμένα, στην εργασία υλοποιήθηκε σύστημα ταξινόμησης έργων ζωγραφικής σε ένα από τα πέντε καλλιτεχνικά ρεύματα: κλασικισμός, κυβισμός, εξπρεσιονισμός, ιμπρεσιονισμό και ρεαλισμό. Οι χρησιμοποιούμενοι ταξινομητές ήταν ο SVM, ο Naïve Bayes και ο KNN. Επιπλέον, το συγκεκριμένο σύστημα, με τα ίδια χαρακτηριστικά και με χρήση των ίδιων ταξινομητών χρησιμοποιήθηκε και για την ταξινόμηση (διάκριση) των καλλιτεχνικών ρευμάτων ανά δύο.

Ο ταξινομητής με το υψηλότερο ποσοστό επιτυχίας στην σύγκριση και με τα 5 καλλιτεχνικά ρεύματα αλλά και στην σύγκριση 2 καλλιτεχνικών ρευμάτων μεταξύ τους, ήταν ο SVM και συγκεκριμένα, οι Quadratic SVM και Linear SVM, με υψηλότερο ποσοστό επιτυχίας 69.6% στην περίπτωση των 5 ρευμάτων και 97% στην περίπτωση των 2 ρευμάτων. Το καλλιτεχνικό ρεύμα με τα μεγαλύτερα ποσοστά επιτυχημένης ταξινόμησης και στις δύο συγκρίσεις, ήταν ο κλασικισμός. Αντίθετα, ο ιμπρεσιονισμός ήταν το καλλιτεχνικό ρεύμα με τα μικρότερα ποσοστά επιτυχημένης ταξινόμησης.

Παρακάτω, αναφέρονται πιθανές αιτίες για τα δείγματα που δεν αναγνώρισαν οι ταξινομητές.

Αρχικά, για κάποια από τα καλλιτεχνικά ρεύματα οι διαφορές μεταξύ τους είναι μικρές, και είναι δύσκολο να εντοπιστούν ακόμη και από έναν άνθρωπο. Στις συγκρίσεις που κάναμε στο προηγούμενο κεφάλαιο, φαίνεται ότι αναγνωρίζει καλύτερα τον κλασικισμό και τον κυβισμό, αλλά με τον εξπρεσιονισμό, τον ιμπρεσιονισμό και τον ρεαλισμό, τα ποσοστά επιτυχίας μειώνονται.

Έπειτα, είναι ο αριθμός των καλλιτεχνικών ρευμάτων. Στα παραπάνω αποτελέσματα έγιναν συγκρίσεις με 5 καλλιτεχνικά ρεύματα καθώς επίσης και ανά 2 καλλιτεχνικά ρεύματα. Παρατηρήσαμε ότι όσο ο αριθμός των καλλιτεχνικών ρευμάτων αυξάνεται, αυξάνεται και το ποσοστό δυσκολίας στη σωστή ταξινόμηση. Για παράδειγμα, στα 2 καλλιτεχνικά ρεύματα, η

μέση πιθανότητα να επιλεγθεί το σωστό καλλιτεχνικό ρεύμα είναι 90%. Αντίθετα με τα 5 καλλιτεχνικά ρεύματα, το ποσοστό μειώνεται κατά 20% κ.ο.κ. Όπως είδαμε, η σύγκριση ανά 2 καλλιτεχνικά ρεύματα, είχε ποσοστό επιτυχίας έως και 97%, ενώ στα 5 καλλιτεχνικά ρεύματα το μέγιστο ποσοστό επιτυχίας έφτανε έως 69.6%. Αυτό είναι αναμενόμενο, αφού και με βάση τη διεθνή βιβλιογραφία, όταν αυξάνει ο αριθμός των συγκρινόμενων καλλιτεχνικών ρευμάτων μειώνεται η ακρίβεια ταξινόμησης.

Τέλος, άλλο ένα αίτιο είναι ο αριθμός πινάκων ανά καλλιτεχνικό ρεύμα. Για την εργασία χρησιμοποιήσαμε 50 πίνακες ανά καλλιτεχνικό ρεύμα. Αυτός ο αριθμός επιλέχθηκε διότι το πρόγραμμα με βάση και τα 40 χαρακτηριστικά έκανε αρκετό χρόνο για να ολοκληρωθεί. Έτσι αν είχαμε τους διπλάσιους πίνακες, ο χρόνος θα ήταν αρκετά μεγάλος.

4.2 ΒΕΛΤΙΩΣΕΙΣ

Τα συμπεράσματα της παραπάνω ενότητας είναι αυτά τα οποία μπορούν να οδηγήσουν μια μελλοντική έρευνα στο αντικείμενο στην επίτευξη ακόμα μεγαλύτερων ποσοστών επιτυχίας.

Σημαντικό είναι το πλήθος των δεδομένων, μιας και όσο μεγαλύτερο είναι αυτό, τόσο πιο έμπιστο και στιβαρό γίνεται το μοντέλο.

Ενδιαφέρον έχει επίσης η διερεύνηση της επίδρασης της ανάλυσης των ψηφιοποιημένων πινάκων στην απόδοση του συστήματος.

Επιπλέον, η ταξινόμηση μπορεί να διεξαχθεί και από ένα πλήθος άλλων ταξινομητών, π.χ. Νευρωνικά Δίκτυα, Δέντρα απόφασης.

Τέλος, εμείς χρησιμοποιήσαμε χαρακτηριστικά χαμηλού επιπέδου όπως ακμές, χρώματα, φωτεινότητα κ.α., ίσως αν μελλοντικά ασχοληθούμε με χαρακτηριστικά υψηλού επιπέδου, δηλαδή αναγνωρισμένα αντικείμενα, περιεχόμενο και σχήματα, θα έχουμε υψηλότερα ποσοστά επιτυχίας. Επιπλέον, για την εξαγωγή των χαρακτηριστικών αντί για global (καθολικά) να ήταν local (τοπικά), δηλαδή να γίνει κατάτμηση του πίνακα σε μικρότερα μέρη. Επιπλέον, εκτός από τα καθολικά (global) χαρακτηριστικά που υιοθετήθηκαν στην παρούσα εργασία, θα μπορούσαν να εξαχθούν και τοπικά (local) χαρακτηριστικά που θα προκύψουν από την κατάτμηση του πίνακα σε μικρότερα μέρη.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Lior Shamir, Tomasz Macura, Nikita Orlov, D. Mark Eckley, and Ilya G. Goldberg. Impressionism, Expressionism, Surrealism: Automated Recognition of Painters and Schools of Art. *ACM Trans. Applied Perc.*, 7(2), 2010.
2. G. Lu, Z. Gao, D. Qin, X. Zhao and M. Liu, "Content-Based Identifying and Classifying Traditional Chinese Painting Images," *2008 Congress on Image and Signal Processing*, 2008, pp. 570-574, doi: 10.1109/CISP.2008.477.
3. Congcong Li and Tsuhan Chen. Aesthetic Visual Quality Assessment of Paintings. *IEEE Journal of Selected Topics in Signal Processing*, 3(2):236–252, 2009.
4. Dasgupta, S. and Md. Ayaz Hossain. "Pattern recognition through brushstroke analysis: an inspection to classify van Gogh from others." (2017).
5. Rasmus Hove Johnsen, Andrea Gradecak Deep Learning Approaches to Art Style Recognition in Digital Images, Master 's Thesis, 2017 Aalborg University
6. T. E. Lombardi, "The classification of style in fine-art painting," ETD Collection for Pace University, Paper AAI3189084, 2005.
7. J. Zujovic, L. Gandy, S. Friedman, B. Pardo and T. N. Pappas, "Classifying paintings by artistic genre: an analysis of features & classifiers," *IEEE International Workshop on Multimedia Signal Processing*, Rio De Janeiro, 2009, pp. 1-5
8. WikiArt, Visual art encyclopedia, WikiArt.org - Visual Art Encyclopedia
9. Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing*, Pearson, 3rd Edition, 2007, 16-18

10. Ν. Παπαμάρκος, Ψηφιακή Επεξεργασία & Ανάλυση Εικόνας, Εκδ. Γκιούρδας, Αθήνα, 2005.
11. [Χρωματικά μοντέλα - Βικιπαίδεια \(wikipedia.org\)](http://el.wikipedia.org)
12. [Μηχανική όραση - Βικιπαίδεια \(wikipedia.org\)](http://el.wikipedia.org)
13. Computer Vision for robotic systems, An Introduction, Michael C. Fairhurst, Prentice Hall 1988
14. Computer and Machine Vision --- Theory, Algorithms, Practicalities, E. R. Davies, 2012
15. Ανάπτυξη Λογισμικού για την Αναγνώριση Οπτικών σημάτων μέσω βλεφαρίσματος, Αναστασία Αγγελίδου, Θωμαή Τσούκα, Πτυχιακή Εργασία, Επιβλέπων: Ιωάννης Καλόμοιρος, Επ. Καθηγητής, Τ.Ε.Ι. Κεντρικής Μακεδονίας, Σέρρες 2013
16. O. Icoğlu, B. Günsel and S. Sariel, "Classification and indexing of paintings based on art movements," *2004 12th European Signal Processing Conference*, 2004, pp. 749-752.
17. R. M. Haralick, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610-621, Nov. 1973, doi: 10.1109/TSMC.1973.4309314.
18. Αλέξης Παλαμίδης – Αναστασία Βελώνη, Σήματα & Συστήματα Με Matlab, Σύγχρονη Εκδοτική, Αθήνα 2008
19. Εμμανουήλ Σαγκριώτης – Σέργιος Θεοδωρίδης, Ανάλυση Εικόνας και Αναγνώριση Προτύπων, Ελληνικό Ανοικτό Πανεπιστήμιο, Τόμος Γ', Πάτρα 2003
20. Le art spy, [Le Art Spy \(rice.edu\)](http://rice.edu)
21. Νικέλλης Αναστάσιος, Ταξινόμηση Ψηφιοποιημένων Έργων Τέχνης Βάσει Χρωματικής Πληροφορίας, Πανεπιστήμιο Αιγαίου, Μυτιλήνη 2008

22. Patrice's Lectures, Edge Detection – Sobel, The University of Auckland, New Zealand, [https://www.cs.auckland.ac.nz/compsci373s1c/PatricesLectures/Edge%20detection-Sobel 2up.pdf](https://www.cs.auckland.ac.nz/compsci373s1c/PatricesLectures/Edge%20detection-Sobel%20up.pdf)
23. [Train models to classify data using supervised machine learning - MATLAB \(mathworks.com\)](#)
24. [Train Classification Models in Classification Learner App - MATLAB & Simulink \(mathworks.com\)](#)
25. Φώτης Λαζαρίνης, Πολυμέσα, Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα, [Δράση Κάλλιπος - Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα \(kallipos.gr\)](#)
26. Δρ. Ν. Βασιλάς, Ψηφιακή Επεξεργασία Εικόνας, Σημειώσεις ΤΕΙ Αθήνας 2016
27. Ι. Πήτας, Ψηφιακή Επεξεργασία Εικόνας, Εκδ. Πήτας Ιωάννης, 2010
28. [Χρωματικά μοντέλα - Βικιπαίδεια \(wikipedia.org\)](#)
29. [Χρωματικό βάθος - Βικιπαίδεια \(wikipedia.org\)](#)
30. [Ανάλυση και επεξεργασία εικόνων DICOM με τη χρήση Matlab - PDF ΔΩΡΕΑΝ Λήψη \(docplayer.gr\)](#)
31. [Αναγνώριση προτύπων - Βικιπαίδεια \(wikipedia.org\)](#)
32. <https://sevenarts.tv/wp-content/uploads/2019/11/Episode-2-Art-movements-7-Tips-GR.pdf>
33. [Εξπρεσιονισμός - Βικιπαίδεια \(wikipedia.org\)](#)
34. [Ιμπρεσιονισμός - Βικιπαίδεια \(wikipedia.org\)](#)
35. [Ρεαλισμός \(τέχνες\) - Βικιπαίδεια \(wikipedia.org\)](#)

36. Κύρκος, Ε. (2015). Επιχειρηματική ευφυΐα και εξόρυξη δεδομένων [Προπτυχιακό εγχειρίδιο]. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις.
37. [Cross-Validation - MATLAB & Simulink \(mathworks.com\)](#)
38. Shuqiang Jiang, W. Gao and Weiqiang Wang, "Classifying traditional Chinese painting images," *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, 2003, pp. 1816-1820 vol.3, doi: 10.1109/ICICS.2003.1292780.
39. [Classifier Options - MATLAB & Simulink \(mathworks.com\)](#)
40. Νευρωνικά Δίκτυα και Μηχανική Μάθηση, Simon Haykin, Πανεπιστήμιο McMaster Hamilton, Ontario, Καναδάς, Απόδοση: Ελένη Γκαγκάτσιου, Εκδόσεις Παπασωτηρίου, Αθήνα 2010
41. [SVM: Difference between Linear and Non-Linear Models - AITUDE](#)
42. [Naive Bayes Classification - MATLAB & Simulink \(mathworks.com\)](#)
43. [SVM Kernels: What Do They Actually Do? | by Michał Oleszak | Towards Data Science](#)
44. U. Jain, K. Nathani, N. Ruban, A. N. Joseph Raj, Z. Zhuang and V. G.V. Mahesh, "Cubic SVM Classifier Based Feature Extraction and Emotion Detection from Speech Signals," *2018 International Conference on Sensor Networks and Signal Processing (SNSP)*, 2018, pp. 386-391, doi: 10.1109/SNSP.2018.00081.
45. [Most Popular Distance Metrics Used in KNN and When to Use Them - KDnuggets](#)
46. [Inverse-square law - Wikipedia](#)
47. Image Processing, Analysis, and Machine Vision, Fourth Edition, Milan Sonka, Vaclav Hlavac, Roger Boyle CENGAGE Learning, 2014, ISBN-13: 978-1-133-59360-7