



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ**  
**ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ**  
**ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ & ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ**

## **Διπλωματική Εργασία**

**Τεχνικές μηχανικής μάθησης και υπολογιστικής όρασης για επεξεργασία χειρόγραφων δειγμάτων με έμφαση στην βιομετρική ταυτοποίηση**



**Φοιτήτρια: Δημητροπούλου Μαρία**  
**ΑΜ: 50106976**

**Επιβλέπων Καθηγητής**

**Δρ. Ηλίας Ν. Ζώης**  
**Αναπληρωτής Καθηγητής**

**ΑΘΗΝΑ-ΑΙΓΑΛΕΩ, ΜΑΡΤΙΟΣ 2023**



**UNIVERSITY OF WEST ATTICA**  
**FACULTY OF ENGINEERING**  
**DEPARTMENT OF ELECTRICAL & ELECTRONICS ENGINEERING**

## **Diploma Thesis**

### **Computer Vision and pattern recognition techniques for biometric authentication.**



**Student: Dimitropoulou Maria**  
**Registration Number: 50106976**

**Supervisor**

**Elias N. Zois, Ph.D**  
**Associate Professor**

**ATHENS-EGALEO, March, 2023**

Η Διπλωματική Εργασία έγινε αποδεκτή και βαθμολογήθηκε από την εξής τριμελή επιτροπή:

Ηλίας Ζώης Αναπληρωτής Καθηγητής	Σωτήριος Καραμπέτσος Αναπληρωτής Καθηγητής	Παναγιώτης Τσιάκας Λέκτορας
(Υπογραφή)	(Υπογραφή)	(Υπογραφή)

**Copyright ©** Με επιφύλαξη παντός δικαιώματος. All rights reserved.

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ και Δημητροπούλου Μαρία,  
Μάρτιος, 2024**

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τους συγγραφείς.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν την συγγραφέα του και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις θέσεις του επιβλέποντος, της επιτροπής εξέτασης ή τις επίσημες θέσεις του Τμήματος και του Ιδρύματος.

### **ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ**

Η κάτωθι υπογεγραμμένη Δημητροπούλου Μαρία του Κωνσταντίνου, με αριθμό μητρώου 50106976 φοιτήτρια του Πανεπιστημίου Δυτικής Αττικής της Σχολής ΜΗΧΑΝΙΚΩΝ του Τμήματος ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ,

**δηλώνω υπεύθυνα ότι:**

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα 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, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διπλώματός μου.

Η Δηλούσα  
Δημητροπούλου Μαρία



Σε όσους που παρά τα εμπόδια συνεχίζουν να εργάζονται για τα όνειρά τους και γι' αυτό που αγαπούν.

Θα ήθελα αρχικά να εκφράσω τις ευχαριστίες μου προς τον καθηγητή μου για την αφοσίωσή του, τη στήριξη και την υπομονή που επέδειξε κατά τη διάρκεια της εκπόνησης της διπλωματικής μου εργασίας. Επιπλέον, θέλω να ευχαριστήσω όλους τους καθηγητές μου για την πολύτιμη γνώση που μου μετέδωσαν κατά τη διάρκεια των σπουδών μου. Επίσης, θα ήθελα να ευχαριστήσω τους συμφοιτητές μου, Μ.Α. και Π.Α., οι οποίοι με τον δικό τους τρόπο συνέβαλαν στην επιτυχή ολοκλήρωση αυτής της εργασίας. Τέλος, δεν θα μπορούσα να μην αναφέρω την στήριξη που λάμβανα από τους γονείς μου. Παρά τα εμπόδια, πάντα πίστευαν σε μένα και με στήριξαν στην επιτυχή ολοκλήρωση των σπουδών μου.

## Περίληψη

Στην παρούσα διπλωματική εργασία διερευνάται ένα σημαντικό θέμα στον τομέα της ασφάλειας και της τεχνολογίας πληροφοριών. Η επαλήθευση των χειρόγραφων υπογραφών ως βιομετρικών δεδομένων είναι ένα από τα πιο σύνθετα αλλά και ταυτόχρονα μη επεμβατικά βιομετρικά χαρακτηριστικά στον τομέα της αναγνώρισης προτύπων και της υπολογιστικής νοημοσύνης. Με τη χρήση της χειρόγραφης βάσης CEDAR, ο στόχος της συγκεκριμένης διατριβής έχει να κάνει με την εξέταση τόσο των γνησίων όσο και των πλαστών δειγμάτων χειρόγραφων υπογραφών με στόχο την αξιολόγηση και ταξινόμηση τους. Χρησιμοποιώντας μη επιβλεπόμενες μεθοδολογίες και αλγόριθμους μηχανικής μάθησης, όπως ο K-SVD και ο OMP, διερευνάται μια αραιή αναπαράσταση των στατικών εικόνων των δειγμάτων υπογραφής, η οποία είναι ένα κρίσιμο βήμα για την ακριβή ταξινόμησή τους. Ένα σύστημα επιβεβαίωσης της γνησιότητας της υπογραφής σχεδιάζεται και υλοποιείται με τη χρήση του μετασχηματισμού διχοτόμησης του αρχικού χώρου των χαρακτηριστικών. Στη συνέχεια, η χρήση ισχυρών μοντέλων Υποστήριξης Διανυσματικών Μηχανών (SVM) επιτρέπει την αξιολόγηση αυτών των χαρακτηριστικών και την ταξινόμηση των δειγμάτων υπογραφής ως γνήσια ή πλαστά με υψηλή ακρίβεια.

## Λέξεις – κλειδιά

Επεξεργασία εικόνας, υπολογιστική όραση, στατική αναγνώριση υπογραφής, αραιή αναπαράσταση, εκμάθηση λεξικού, σύστημα επιβεβαίωσης υπογραφής ανεξάρτητο του γραφέα.

## **Abstract**

In this thesis an important topic in the field of security and information technology is explored. The verification of handwritten signatures as biometric data is one of the most complex yet non-invasive biometric traits in the field of pattern recognition. With the use of the CEDAR dataset, this thesis examines both genuine and forged handwritten signature samples with the aim of evaluating and classifying them. Using unsupervised machine learning algorithms such as K-SVD and OMP, a sparse representation of the static images of the signature samples is explored, which is a crucial step for their accurate classification. A writer independent signature verifier is designed and implemented with the use of the dichotomy transform. Subsequently, the use of robust Support Vector Machine (SVM) models allows the evaluation of these features and the classification of signature samples as genuine or forged with high accuracy.

## **Keywords**

Image processing, off-line signature verification, sparse coding, dictionary learning, writer independent signature verification



## Περιεχόμενα

Κατάλογος Πινάκων .....	10
Κατάλογος Εικόνων .....	10
Αλφαβητικό Ευρετήριο .....	11
<b>ΕΙΣΑΓΩΓΗ.....</b>	<b>12</b>
Αντικείμενο της διπλωματικής εργασίας .....	12
Σκοπός και στόχοι.....	Error! Bookmark not defined.
Μεθοδολογία .....	13
Καινοτομία .....	15
<b>Δομή</b>	<b>15</b>
<b>1 ΚΕΦΑΛΑΙΟ 1<sup>ο</sup> : Εισαγωγικές έννοιες.....</b>	<b>17</b>
1.1 Ιστορικά στοιχεία και χρήση της υπογραφής.....	17
1.2 Η χειρόγραφη υπογραφή ως βιομετρικό δεδομένο της ανθρώπινης ταυτότητας. ....	18
1.3 Κατηγοριοποίηση των συστημάτων χειρόγραφων βιομετρικής ταυτοποίησης.....	19
1.3.1 Δυναμικά συστήματα επιβεβαίωσης υπογραφής.....	20
1.3.2 Στατικά συστήματα επιβεβαίωσης υπογραφής, .....	21
1.3.3 Τοπικά συστήματα επιβεβαίωσης υπογραφής (Writer Dependent) .....	22
1.3.4 Καθολικά συστήματα επιβεβαίωσης υπογραφής (Writer Independent) .....	23
1.4 Οι κατηγορίες των χειρόγραφων δειγμάτων υπογραφής .....	24
1.4.1 Η περίπτωση της επιστήμης των δικαστικών γραφολόγων .....	24
1.4.2 Η περίπτωση της επιστήμης της πληροφορικής .....	25
1.5 Τι θα δούμε στη συνέχεια. ....	26
<b>2 ΚΕΦΑΛΑΙΟ 2<sup>ο</sup>: Στοιχεία ψηφιακής επεξεργασίας εικόνας και στάδιο προεπεξεργασίας.....</b>	<b>27</b>
2.1 Ψηφιακή επεξεργασία εικόνας (Digital image processing) .....	27
2.2 Κατάτμηση εικόνας (Image segmentation) .....	27
2.3 Στάδια προεπεξεργασίας εικόνας .....	28
2.4 Μορφολογία .....	30
2.4.1 Διάβρωση (erosion) και διαστολή (dilation).....	31
2.4.2 Άνοιγμα (opening) και κλείσιμο (closing) .....	32
2.4.3 Λέπτυνση (Thinning) .....	33
<b>3 ΚΕΦΑΛΑΙΟ 3<sup>ο</sup> Στοιχεία αραιής αναπαράστασης και εκμάθησης λεξικού .....</b>	<b>35</b>
3.1 Η αραιή αναπαράσταση και η εκμάθηση λεξικού.....	35
3.1.1 Η εκμάθηση λεξικού με την μέθοδο K-SVD.....	37
3.1.2 Η αραιή αναπαράσταση με την μέθοδο OMP.....	38
3.2 Σχέση της αραιής αναπαράστασης με το θέμα της διπλωματικής εργασίας.....	41
<b>4 ΚΕΦΑΛΑΙΟ 4<sup>ο</sup> : Περιγραφή της πειραματικής διαδικασίας .....</b>	<b>42</b>
4.1 Μηχανές διανυσματικής υποστήριξης (Support vector machines) .....	42
4.2 Βάση δεδομένων .....	44
4.3 Πειραματική διαδικασία και μετρικές σφάλματος .....	45
<b>5 ΚΕΦΑΛΑΙΟ 5<sup>ο</sup> : Αποτελέσματα .....</b>	<b>49</b>
<b>6 ΚΕΦΑΛΑΙΟ 6<sup>ο</sup> : Συμπεράσματα – Προτάσεις για συνέχισης της έρευνας.....</b>	<b>53</b>
<b>Βιβλιογραφία –Αναφορές -Διαδικτυακές Πηγές.....</b>	<b>54</b>

## Κατάλογος Πινάκων

Πίνακας 1. Μέγιστες καμπύλες AUV για το Validation set, μετά από 4 επαναλήψεις. ....	49
Πίνακας 2. Συγκεντρωτικός Πίνακας μέσων τιμών AUC. ....	52

## Κατάλογος Εικόνων

Εικόνα 1. Εποπτικό overview συστήματος,.....	13
Εικόνα 2. Patch extraction.....	14
Εικόνα 3. Αραιή αναπαράσταση και εκμάθηση λεξικού. ....	15
Εικόνα 4. Ο John Hancock και η υπογραφή του. ....	18
Εικόνα 5. Δυναμικά και στατικά συστήματα, Στα δυναμικά συστήματα υπογραφής, αξίζει να σημειωθεί πως γίνεται χρήση γραφίδας και η αναπαράσταση της υπογραφής γίνεται με χρονικά μεταβαλλόμενα δεδομένα, ενώ στα στατικά συστήματα υπογραφής τα δεδομένα προέρχονται από ψηφιακές εικόνες. [2]. .....	20
Εικόνα 6. Ψηφιακά μέσα δυναμικών συστημάτων υπογραφής [6]. ....	20
Εικόνα 7. Χώρος χαρακτηριστικών τοπικών συστημάτων επιβεβαίωσης υπογραφής. ....	23
Εικόνα 8. Χώρος αποστάσεων καθολικών συστημάτων επιβεβαίωσης υπογραφής. ....	24
Εικόνα 9. Δείγμα υπογραφής προς εξέταση [15]......	25
Εικόνα 10. Παράδειγμα τυχαίας πλαστογράφησης. Δεξιά το γνήσιο δείγμα, αριστερά το τυχαίο πλαστό δείγμα. ....	26
Εικόνα 11. Αριστερά: Πλαστές υπογραφές. Δεξιά: γνήσιες υπογραφές (από τον ίδιο γραφέα). ....	26
Εικόνα 12. Αριστερά απεικονίζεται μια εικόνα σε κλίμακα του γκρι, ενώ δεξιά απεικονίζεται η ίδια εικόνα που έχει περάσει από κατωφλιοποίηση με την μέθοδο Otsu.....	29
Εικόνα 13. Παράδειγμα ιστογράμματος της προηγούμενης εικόνας.....	29
Εικόνα 14. Στάδια προεπεξεργασίας εικόνας. Η περιγραφή των σταδίων γίνεται στο κείμενο.....	30
Εικόνα 15. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του erosion. ....	31
Εικόνα 16. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του delation. ....	32
Εικόνα 17. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του opening. ....	32
Εικόνα 18. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του closing. ....	33

Εικόνα 19. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του thinning.....	34
Εικόνα 20. Ψευδοκώδικας της βασική υλοποίησης της OMP .....	40
Εικόνα 21. Ψευδοκώδικας Batch OMP. ....	40
Εικόνα 22. Αναπαράσταση λεξικού για τα γνήσια δείγματα υπογραφών. ....	41
Εικόνα 23. Παράδειγμα γραμμικού διανύσματος υποστήριξης δύο κλάσεων[34].....	42
Εικόνα 24. Παράδειγμα RBF .....	43
Εικόνα 25. Γράφημα κατανομής του RBF πυρήνα. ....	44
Εικόνα 26. Ενδεικτικά δείγματα υπογραφών CEDAR απο 5 γραφείς. Πάνω απεικονίζονται γνήσια δείγματα υπογραφής, ενώ κάτω απεικονίζονται τα πλαστά δείγματα υπογραφής.....	45
Εικόνα 27. Παράδειγμα ROC καμπύλης <sup>8</sup> .....	46
Εικόνα 28. Συνοπτικό διάγραμμα συστήματος αναγνώρισης υπογραφής. ....	48

## Αλφαβητικό Ευρετήριο

<b>AUC :</b>	Area Under Curve
<b>DW:</b>	Development Writers
<b>EER:</b>	Equal Error Rate
<b>EW:</b>	Exploitation Writers
<b>FN:</b>	False Negatives
<b>FP:</b>	False Positives
<b>FPR:</b>	False Positive Rate
<b>G:</b>	Genuine
<b>K- SVD:</b>	K-Singular Value Decomposition
<b>OMP:</b>	Orthogonal Matching Pursuit
<b>PNG:</b>	Portable Network Graphic
<b>RBF:</b>	Radial Basis Function Kernel
<b>RF:</b>	Random Forgery
<b>ROC:</b>	Receiver Operating Characteristic curve
<b>SF:</b>	Skill Forgery
<b>SR:</b>	Sparse Representation
<b>SV:</b>	Signature Verification
<b>SVM :</b>	Support Vector Machine
<b>TN:</b>	True Negatives
<b>TP:</b>	True Positive
<b>TPR:</b>	True Positive Rate
<b>WD:</b>	Writer Dependent
<b>WI:</b>	Writer Independent

## ΕΙΣΑΓΩΓΗ

Η βιομετρική αναγνώριση αποτελεί όρο ο οποίος αναφέρεται στη χρήση τόσο των φυσικών χαρακτηριστικών όσο και των χαρακτηριστικών συμπεριφοράς για τον σκοπό της αναγνώρισης της παρουσίας ή της έγκρισης-συμφωνίας ενός ανθρώπου σε ένα πλήθος διαφορετικών συναλλαγών. Τα φυσικά χαρακτηριστικά είναι εκείνα τα οποία είναι μοναδικά για την ταυτότητα ενός ανθρώπου, δεν αλλάζουν στο πέρασμα του χρόνου όπως για παράδειγμα το δακτυλικό αποτύπωμα, το DNA, η ίριδα του οφθαλμού, κλπ, και η λήψης τους ενδέχεται να εμπεριέχει ζητήματα υγείας (π.χ. laser για την ανάγνωση της ίριδας). Αντίθετα τα χαρακτηριστικά συμπεριφοράς δύναται να μεταβάλλονται με τον χρόνο ή με την διάθεση του κάθε ατόμου ή από διάφορους άλλους παράγοντες. Παραδείγματα τέτοιων χαρακτηριστικών είναι η φωνή, ο τρόπος γραφής, και ειδικότερα η χειρόγραφη υπογραφή [1], [2]. Το τελευταίο παράδειγμα, δηλαδή η χρήση χειρόγραφης υπογραφής ως μέσο βιομετρικής αναγνώρισης είναι αυτό που θα μελετηθεί στην παρούσα διπλωματική εργασία.

Η μελέτη επαλήθευσης υπογραφής είναι μια αρκετά αρχαία έννοια καθώς χρονολογείται γύρο στο 439 μ.Χ., όταν για πρώτη φορά η υπογραφή χρησιμοποιήθηκε για την πιστοποίηση γνησιότητας εγγράφων στην ρωμαϊκή αυτοκρατορία. Αρκετούς αιώνες αργότερα, το 1677 ψηφίστηκε ο πρώτος νόμος στην Αγγλία σχετικά με την επικύρωση των υπογραφών για την αποφυγή απάτης. Ακολουθώντας την ανακάλυψη του ηλεκτρονικού υπολογιστή και της συναφούς επιστήμης των υπολογιστών, δηλαδή κάπου στον 20<sup>ο</sup> αιώνα, παρατηρείται και μια στροφή για την ανάπτυξη εφαρμογών για την επαλήθευση χειρόγραφων υπογραφών που να στηρίζονται στον συνδυασμό των αρχών της δικαστικής γραφολογίας, της επιστήμης της επεξεργασίας σήματος, εικόνας καθώς και της μηχανικής και βαθιάς μάθησης. Ιστορικά, το 1965 στην Βόρεια Αμερική παρουσιάστηκε το πρώτο αυτοματοποιημένο σύστημα αναγνώρισης υπογραφής, στηριζόμενο τόσο σε δυναμικά –online χαρακτηριστικά όσο και στατικά - offline (έννοιες οι οποίες θα εξηγηθούν στα επόμενα κεφάλαια) από τους Nagel, Rosenfeld, Liu και Herbst. Από τότε και έπειτα, έχουν γίνει αρκετές μελέτες που βασίστηκαν πάνω στην αρχική ιδέα, αλλά όπως είναι φυσικό έχουν προταθεί νέες μέθοδοι και αλγόριθμοι για την αντιμετώπιση του προβλήματος αναγνώρισης και επαλήθευσης χειρόγραφων υπογραφών[2].

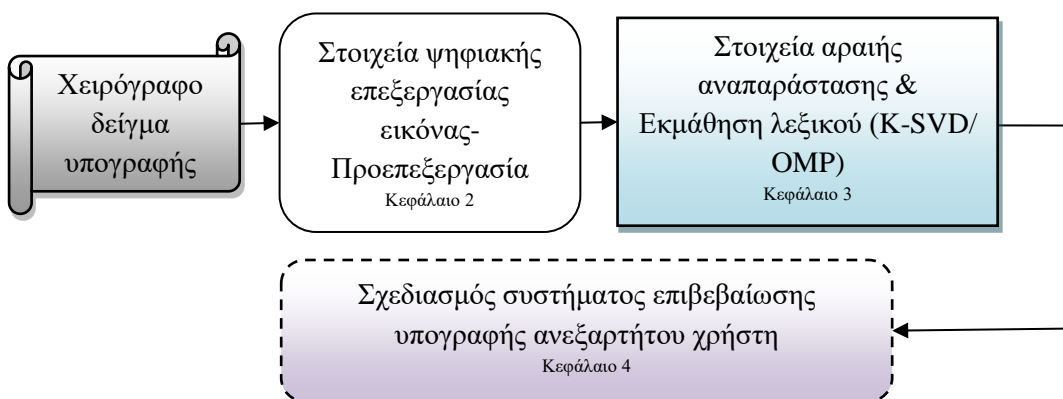
### Αντικείμενο της διπλωματικής εργασίας

Το κύριο θέμα που πραγματεύεται η διπλωματική εργασία αποτελεί η αναγνώριση της γνησιότητας εικόνων χειρόγραφων δειγμάτων, ιδιαίτερα εικόνων χειρόγραφων υπογραφών, με σύγχρονες τεχνικές επεξεργασίας εικόνας, σήματος και μηχανικής μάθησης οι οποίες θα αναφερθούν λεπτομερώς στα παρακάτω κεφάλαια. Ωστόσο για να φτάσουμε στο τελικό αποτέλεσμα η τεχνική παρουσίαση της παρούσας διπλωματική εργασία μελετάει διεξοδικά τρία βασικά στάδια. Αρχικά μελετάται το στάδιο της προετοιμασίας και προεπεξεργασίας της εικόνας, για την κατάλληλη ψηφιοποίηση της εικόνας υπό εξέταση. Στην συνέχεια, περιγράφεται το στάδιο της απεικόνισης της αρχικής ή προεπεξεργασμένης ψηφιακής εικόνας σε ένα διάνυσμα χαρακτηριστικών μέσα από την θεωρία και υλοποίηση των εννοιών της αραιής αναπαράστασης ή κωδικοποίησης. Η ανωτέρω κωδικοποίηση στηρίζεται στην ύπαρξη ενός λεξικού το οποίο μαθαίνεται επίσης από τα δεδομένα εκπαίδευσης διαμέσου μιας διαδικασίας μη επιβλεπόμενης εκμάθησης η οποία ονομάζεται *εκμάθηση λεξικού*. Το τελικό μοντέλο επιβεβαίωσης υπογραφής γίνεται με τον σχεδιασμό ενός ολικού μοντέλου ταξινόμησης, ανεξάρτητου του γραφέα, με χρήση ταξινομητών μηχανών διανυσματικής υποστήριξης (SVM). Η βιομετρική αναγνώριση χειρόγραφων υπογραφών θεωρείται ένα διαχρονικό θέμα διότι χρησιμοποιείται σε πληθώρα εφαρμογών επικαιροποίησης εγγράφων, όπως για παράδειγμα έγγραφα

τραπεζών, δημόσια έγγραφα, ακόμα και για εγκληματολογικές εφαρμογές, αντικείμενο με το οποίο συνήθως ασχολείται ο δικαστικός γραφολόγος.

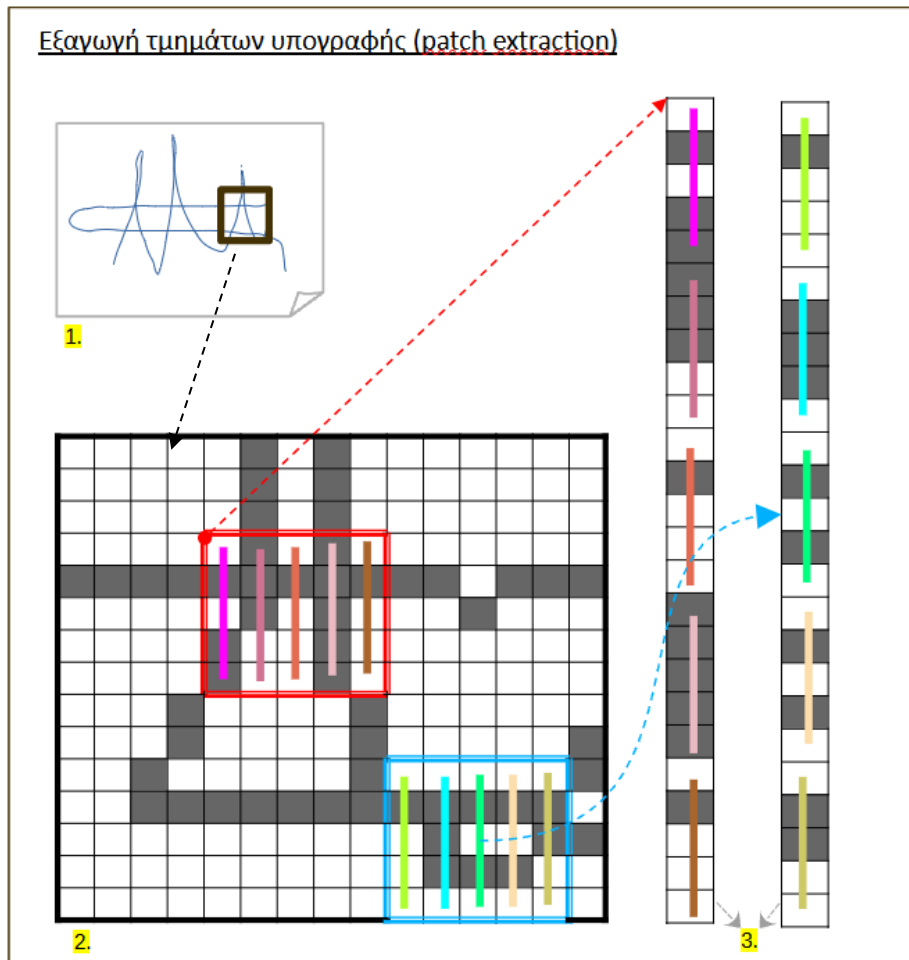
## Μεθοδολογία

Στο σημείο αυτό θα γίνει μια πλήρη ανασκόπηση όλων των σταδίων επεξεργασίας των δειγμάτων χειρόγραφης υπογραφής από το στάδιο της προεπεξεργασίας έως το στάδιο της εξαγωγής των χαρακτηριστικών. Εδώ λοιπόν θα μας απασχολήσει κυρίως το κομμάτι της επεξεργασίας εικόνας και της αραιής κωδικοποίησης μέσω της εκμάθησης του λεξικού. Ειδικότερα, αναφέρονται οι διαδικασίες α) «Εξαγωγής τμημάτων υπογραφής», β) της «Εκμάθησης λεξικού», καθώς και γ) της «Αραιής Αναπαράστασης». Αξίζει να σημειωθεί ότι στο επόμενο στάδιο (4<sup>ο</sup> κεφάλαιο) θα περιγράψουμε το προτεινόμενο σύστημα «writer independent system WI» αναγνώρισης της χειρόγραφης υπογραφής. Το εποπτικό overview του συστήματος παρουσιάζεται στην **Εικόνα 1**.



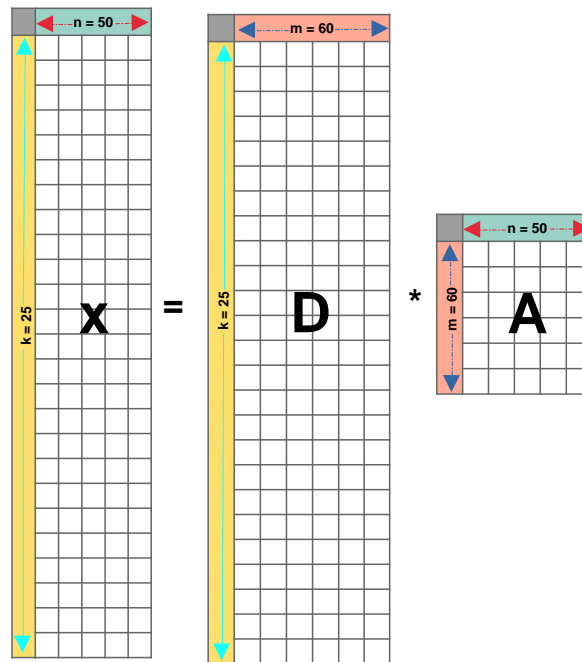
Εικόνα 1. Εποπτικό overview συστήματος.

Η προεπεξεργασία εικόνας αναφέρεται στα βήματα που ακολουθούνται προκειμένου να βελτιώσουμε την ποιότητα μιας εικόνας υπογραφής. Συγκεκριμένα, το στόχος είναι να απαλλαγούμε από οποιαδήποτε θολάδα ή θόρυβο που ενδεχομένως υπάρχει λόγω της σάρωσης της εικόνας και να επιτύχουμε καλή αντίθεση, δηλαδή να καθιστούμε τα στοιχεία της εικόνας πιο ευδιάκριτα. Σαν στόχος είναι η εξαγωγή εύκολων επεξεργάσιμων δεδομένων, με ιδιαίτερη προσοχή στις απώλειες της μη χρήσιμης πληροφορίας. Τα τυπικά βήματα προεπεξεργασίας εικόνας είναι η εξάλειψη φόντου, η αφαίρεση του θορύβου με την χρήση φίλτρων και συγκεκριμένα Gaussian φίλτρου, η καταωφλιοποίηση (*binarization*) και τέλος η λέπτυνση (*thinning*) της εικόνας. Η εξάλειψη του φόντου χρειάζεται ώστε το αποτέλεσμα του φόντου της υπογραφής να είναι λευκό και όχι γκριζο, όπως υπάρχει περίπτωση να προκύψει κατά την διαδικασία της σάρωσης. Η αποθορυβοποίηση της εικόνας γίνεται ώστε να φύγουν τυχόν μικρά, ανεπιθύμητα τυχαία στίγματα που υπάρχουν πάνω στην ασπρόμαυρη πλέον εικόνα. Η λέπτυνση (*thinning*) είναι μια τεχνική προεπεξεργασίας εικόνας η οποία αραιώνει το πάχος της γραφής της υπογραφής. Το πάχος που έχει επιλεγεί για την εκάστοτε εικόνα υπογραφής, είναι πάχος ενός pixel. Η διαδικασία αυτή βοηθά για την μετέπειτα λέπτυνση της εικόνας, όπου εκεί μας δίνονται χαρακτηριστικά για το περίγραμμα της υπογραφής [2]. Όλες αυτές οι διαδικασίες περιγράφονται με κάποια λεπτομέρεια στο επόμενο κεφάλαιο.



Εικόνα 2. Patch extraction

Στην **Εικόνα 2** περιγράφεται με έναν απλουστευμένο τρόπο η διαδικασία εξαγωγής τμημάτων υπογραφής (patch extraction). Πάνω αριστερά στην πρώτη εικόνα απεικονίζεται ένα δείγμα υπογραφής, το μαύρο επιλεγμένο κομμάτι είναι το τμήμα της υπογραφής που γίνεται μεγέθυνση για την επεξήγηση της διαδικασίας. Το 2<sup>ο</sup> σχήμα κάτω αριστερά αντιπροσωπεύει τα pixel από τα επιλεγμένα ίχνη της ανωτέρω μεγεθυμένης υπογραφής. Όπως φαίνεται πάνω στον μεγάλο πίνακα, έχουν επιλεγεί τυχαία δύο πλαίσια εξαγωγής τμημάτων (patches), ο κόκκινος και ο μπλε, μεγέθους  $5 \times 5$  το ένα. Δεξιά τα 2 αυτά πλαίσια έχουν αναπαρασταθεί ως πίνακες στήλη με τον τρόπο που φαίνεται στο παράδειγμα, δηλαδή κάτω από την πρώτη στήλη του κόκκινου πλαισίου, έχει τοποθετηθεί η 2<sup>η</sup> στήλη κοκ. , αντίστοιχα το ίδιο γίνεται για κάθε patch, δηλαδή για κάθε κομμάτι  $5 \times 5$  που δημιουργείται. Η κάθε στήλη έχει μέγεθος 25 γραμμές όπως προκύπτει από το γινόμενο των διαστάσεων του πλαισίου.



Εικόνα 3. Αραιή αναπαράσταση και εκμάθηση λεξικού.

Στην **Εικόνα 3** απεικονίζεται με γραφικό τρόπο η διαδικασία της αραιής αναπαράστασης. Στο παράδειγμα αυτό, ο πίνακας  $X$  είναι 50 στηλών  $n$  και 25 γραμμών  $k$  (ή αλλιώς patches). Το λεξικό  $D$  με τη σειρά του, αποτελείται από 60 στήλες  $m$  (ή αλλιώς atoms) και 25 γραμμές  $k$  αντίστοιχα. Ο πίνακας  $A$ , είναι ο πίνακας των αραιών συντελεστών. Το πλήθος των γραμμών  $m$  του  $A$  είναι ίσο με το πλήθος των στηλών  $m$  του λεξικού  $D$  και το πλήθος των στηλών  $n$  του  $A$  είναι ίσο με το πλήθος των στηλών  $n$  του  $X$ . Ο πίνακας διόρθωσης  $X$  αναπαρίσταται ως ένας γραμμικός συνδυασμός στοιχείων από το  $D$ , με αραιούς συντελεστές που δίνονται από το  $A$ . Αξίζει να σημειωθεί ότι το λεξικό είναι υπερπλήρες (*overcomplete*) καθώς ισχύει η σχέση  $(2 \times k) \ll m$ . Η OMP είναι μια μέθοδος αραιής αναπαράστασης. Δέχεται σαν είσοδο τον πίνακα  $X$ , δηλαδή τα patches καθώς και το λεξικό  $D$ . Η έξοδος της συνάρτησης αυτής είναι ο πίνακας αραιών συντελεστών, δηλαδή ο πίνακας  $A$  της **Εικόνα 3**.

## Καινοτομία

Η κυριότερη καινοτομία της διπλωματικής εργασίας είναι ότι ενσωματώνει για πρώτη φορά στην βιβλιογραφία τεχνικές διχοτόμησης του χώρου χαρακτηριστικών διανυσμάτων που προέρχεται από τεχνικές αραιής αναπαράστασης για την υλοποίηση ενός συστήματος επιβεβαίωσης ανεξαρτήτως του γραφέα.

## Δομή

Στο πρώτο κεφάλαιο παρέχεται μια εισαγωγή στη βιομετρική ταυτοποίηση, εξετάζοντας τη σημασία της υπογραφής ως ένα από τα βιομετρικά δεδομένα της ανθρώπινης ταυτότητας. Θα παρουσιαστούν τα ιστορικά στοιχεία καθώς και η χρήση της υπογραφής, ενώ θα κατηγοριοποιηθούν τα συστήματα χειρόγραφης βιομετρικής ταυτοποίησης, περιλαμβανομένων των δυναμικών και στατικών συστημάτων επιβεβαίωσης υπογραφής, καθώς και των συστημάτων επιβεβαίωσης υπογραφής που είναι εξαρτημένα ή ανεξάρτητα από τον συγγραφέα (Writer Dependent και Writer Independent αντίστοιχα). Στο δεύτερο κεφάλαιο, εμφανίζονται βασικά δομικά στοιχεία της ψηφιακής επεξεργασίας εικόνας, και επιπλέον θα παρουσιαστούν τεχνικές που χρησιμοποιούνται ευρέως σε αυτόν τον τομέα. Θα περιληφθούν έννοιες όπως η κατωφλιοποίηση και η κατάτμηση εικόνας, οι

μορφολογικές επεξεργασίες και άλλες τεχνικές προεπεξεργασίας εικόνας. Στο τρίτο κεφάλαιο, θα παρουσιαστεί σε μια πρώτη προσέγγιση το αντικείμενο της αραιής αναπαράστασης και την εκμάθησης λεξικού, ένα πρόβλημα βελτιστοποίησης στην υπολογιστικής όραση, το οποίο προσεγγίζεται δια μέσου των αλγορίθμων K-SVD και OMP. Επιπλέον θα επιχειρηθεί η ανάλυση της σχέσης των μεθόδων K-SVD και OMP με το θέμα της διπλωματικής εργασίας. Στο τέταρτο κεφάλαιο, θα προχωρήσουμε στην ανάλυση των μηχανών διανυσματικής υποστήριξης (SVM), συμπεριλαμβανομένης της βάσης δεδομένων και των μετρικών σφάλματος και επιπλέον, θα γίνει περιγραφή της πειραματικής διαδικασίας. Τέλος, στο πέμπτο και τελευταίο κεφάλαιο, θα προσφέρουμε μια συνολική επισκόπηση της διπλωματικής εργασίας και θα εξετάσουμε πιθανές κατευθύνσεις για περαιτέρω έρευνα και επέκταση της εργασίας αυτής.



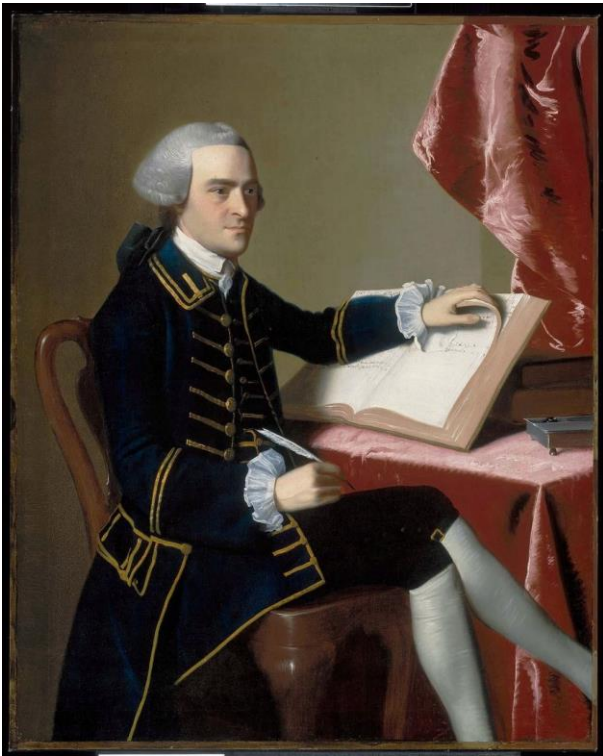
## 1 ΚΕΦΑΛΑΙΟ 1<sup>ο</sup> : Εισαγωγικές έννοιες

Στο κεφάλαιο αυτό θα γίνει μια ιστορική αναδρομή στον τρόπο χρήσης της υπογραφής ανά τα χρόνια. Θα αναφερθούν τρόποι συμφωνίας πριν την χρήση του χειρόγραφου και της γραφής γενικότερα, καθώς και σημαντικά γεγονότα και άνθρωποι που έχουν αφήσει το στίγμα τους στον τομέα της υπογραφής. Όσο αφορά την σύγχρονη ιστορία, αργότερα και την ιατρική, θα δούμε πως η γραφή, το απλό χειρόγραφο ή ακόμα και η υπογραφή, μπορούν να γίνουν σύνθετα δεδομένα της ανθρώπινης ταυτότητας, όπως για παράδειγμα είναι το δακτυλικό αποτύπωμα ή το dna. Τέλος θα γίνει μια αναφορά στις κατηγορίες των συστημάτων της βιομετρικής ταυτοποίησης, τόσο ως προς το μέσο της υπογραφής, όσο και προς την μορφή του συστήματος αναγνώρισης.

### 1.1 Ιστορικά στοιχεία και χρήση της υπογραφής

Ιστορικά, κάποιος μπορεί να παρατηρήσει ότι υπήρχαν διάφοροι τρόποι για την επισφράγιση μιας συμφωνίας μεταξύ δύο εμπλεκόμενων μερών. Ως παράδειγμα, μπορούμε να αναφέρουμε την αποκοπή και παράδοση στο έτερο μέρος της συμφωνίας μιας τούφας από τα μαλλιά ή ενός δάκτυλου. Σε άλλες πολιτιστικές παραδόσεις, η επισφράγιση μιας συμφωνίας μπορεί να περιείχε την παράδοση ενός τέκνου για γάμο, την ανταλλαγή όρκων, ή ακόμα μια πράξη έντονου συμβολισμού (π.χ. ένα χαστούκι) ώστε τα άτομα της συμφωνίας να θυμούνται την ένταση της στιγμής και άρα και την συμφωνία που επισφραγίστηκε από αυτήν [3]. Στα χρόνια που ακολούθησαν την ανακάλυψη της γραφής, η χρήση της χειρόγραφης αναγραφής του ονόματος, ως δείγμα της συμφωνίας ή της ύπαρξης ενός ή και των δύο μερών μιας συμφωνίας, αποτέλεσε κάτι που ακόμα και σήμερα χρησιμοποιείται. Παρόλη την ψηφιοποίηση της σημερινής μας εποχής, η χειρόγραφη υπογραφή χρησιμοποιείται ακόμα και τώρα για την επισφράγιση μιας συμφωνίας ή για την δήλωση της παρουσίας ενός ανθρώπου σε ένα γεγονός. Στο πέρασμα των χρόνων και καθώς η γραφή, η εκπαίδευση και η παιδεία σημείωναν ανοδική πορεία, έκαναν την εμφάνισή τους όλο και περισσότερο οι γραπτές συμφωνίες, στις οποίες η σφραγίδα και η χειρόγραφη υπογραφή σε ένα φύλλο χαρτί αποτελούσαν δείγματα μιας τυπική επικύρωσης αυτών.

Η θέσπιση των χειρόγραφων υπογραφών ως μορφή επικύρωσης συμφωνιών ποικίλουν ανά περιοχή και χρονολογία. Πολλές είναι οι υπογραφές που έχουν αφήσει το αποτύπωμά τους στην ιστορία, όπως για παράδειγμα αυτή του Αμερικάνου John Hancock στην **Εικόνα 4**, του οποίου το όνομα του στις Ηνωμένες Πολιτείες της Αμερικής (Η.Π.Α) έχει γίνει παρωνύμιο της λέξης υπογραφή. Συγκεκριμένα, ο Hancock ήταν πρόεδρος του Κογκρέσου των Η.Π.Α, όταν τον Ιούλιο του 1776 υπογράφηκε η Διακήρυξη της Ανεξαρτησίας. Σύμφωνα με τον αστικό μύθο, ο Hancock ήταν ο πρώτος που υπέγραψε το έγγραφο με το όνομά του σε τέτοιο βαθμό ώστε ο Βασιλιάς Γεώργιος της Αγγλίας να μπορεί να το διαβάσει χωρίς την χρήση των γυαλιών του. Έτσι, η υπογραφή του έγινε τόσο διάσημη καθότι ήταν μεγάλη, φανταχτερή, είχε έντονο και παχύ ίχνος και δύσκολα μπορούσε κανείς να πλαστογραφήσει [3], [4].



Εικόνα 4. Ο John Hancock<sup>1</sup> και η υπογραφή του<sup>2</sup>.

## 1.2 Η χειρόγραφη υπογραφή ως βιομετρικό δεδομένο της ανθρώπινης ταυτότητας.

Όπως έχουμε ήδη αναφέρει, η βιομετρική αναγνώριση αναφέρεται στη χρήση τόσο των φυσικών χαρακτηριστικών όσο και των χαρακτηριστικών συμπεριφοράς για την ταυτοποίηση της ύπαρξης ή της συναίνεσης ενός ανθρώπου σε μια συναλλαγή. Τα φυσικά χαρακτηριστικά είναι εκείνα τα οποία είναι μοναδικά για την ταυτότητα ενός ανθρώπου και δεν αλλάζουν ως οντότητες με το πέρασμα του χρόνου. Για παράδειγμα αναφέρεται το δακτυλικό αποτύπωμα, το DNA, η ίριδα του οφθαλμού, κλπ. Σε αντίθεση, τα χαρακτηριστικά συμπεριφοράς μπορούν να μεταβάλλονται με τον χρόνο ή με την διάθεση του κάθε ατόμου ή από διάφορους άλλους παράγοντες όπως τα ναρκωτικά ή διάφορες νευροεκφυλιστικές ασθένειες. Παραδείγματα τέτοιων χαρακτηριστικών είναι η φωνή, ο τρόπος γραφής, καθώς και η χειρόγραφη υπογραφή [1], [2], [5].

Επιπλέον των περιπτώσεων της βιομετρικής ταυτοποίησης, οι επιπτώσεις των συνθηκών υγείας και ο τρόπος με τον οποίο αυτές απεικονίζονται στο χειρόγραφο κείμενο είναι κάτι που έχει απασχολήσει εδώ και πολλά χρόνια μια κατηγορία από επιστήμονες όπως ιατροδικαστές, γραφολόγους, ψυχολόγους και νεύρο-επιστήμονες. Με την χειρόγραφη υπογραφή ως βιομετρικό δεδομένο, μπορούν, σύμφωνα με διάφορες μελέτες, να εξαχθούν διάφορα συμπεράσματα τόσο για την γνησιότητα αυτής, όσο και για την κατάσταση που βρίσκεται ή ακόμα και για τυχόν ασθένειες που παρουσιάζει ο κάθε γραφέας όπως οι ψυχιατρικές, ή οι νευρολογικές. Για παράδειγμα, το χειρόγραφο κείμενο καθώς και οι αλλαγές αυτού στον γραφικό χαρακτήρα αποτελούν εξαιρετικούς βίο-δείκτες και ως εκ τούτου μπορεί να χρησιμοποιηθούν σε εφαρμογές της ιατρικής επιστήμης για να εξαχθούν

<sup>1</sup> [John Hancock Portrait](#)

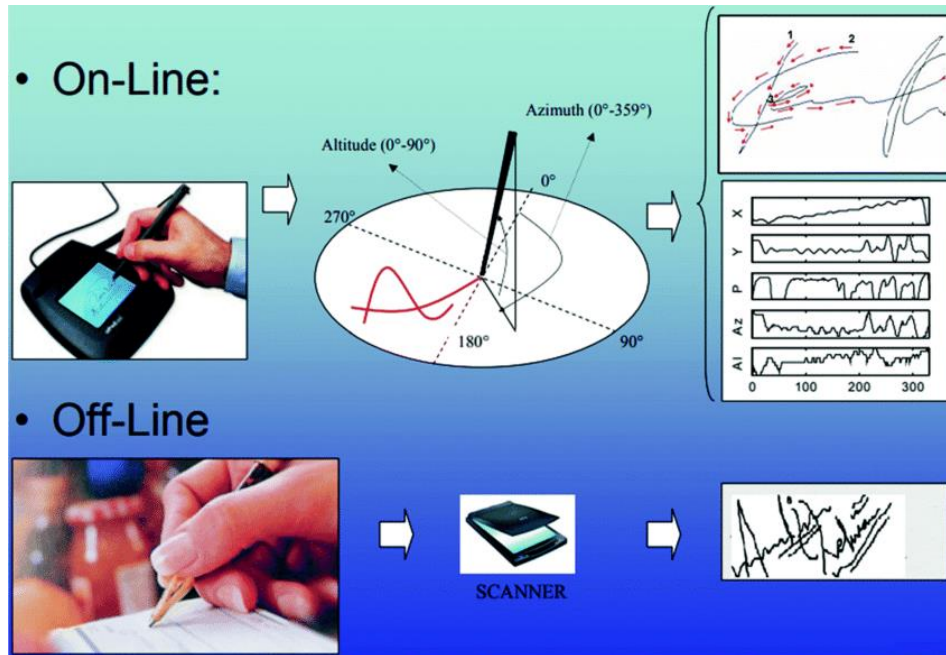
<sup>2</sup> [Signature of J. H](#)

συμπεράσματα σχετικά με το εάν ο υπογράφων κάνει χρήση ναρκωτικών, αλκοόλ ή παρουσιάζει συμπτώματα που οδηγούν σε ασθένειες της νόσου Parkinson, ή Alzheimer, δηλαδή τις λεγόμενες νευροεκφυλιστικές ασθένειες. Το πειραματικό πρωτόκολλο περιλαμβάνει την λήψη δειγμάτων με μια ψηφιακή γραφίδα πάνω σε ένα tablet, ώστε με τον τρόπο αυτό να μπορεί να γίνει ο έλεγχος των κινήσεων του χεριού. Πιο συγκεκριμένα λαμβάνονται πληροφορίες όπως η θέση της μύτης της γραφίδας στους άξονες X και Y, η θέση της γραφίδας στην επιφάνεια ή στον αέρα, η πίεση που ασκείται στο άκρο της πένας, η γωνία της πένας σε σχέση με την επιφάνειας tablet, και τέλος η σήμανση του χρόνου. Ως ένα παράδειγμα αναφέρουμε ότι κάποιες μελέτες έχουν δείξει ότι όταν η πένα βρίσκεται για περισσότερο χρόνο στον αέρα είναι πιθανό σημάδι αναγνώρισης της νόσου Parkinson ή της νόσου Alzheimer [6]. Οι ανωτέρω παρατηρήσεις υποδηλώνουν ότι ο τρόπος γραφής αποτελεί ένα πολύτιμο εργαλείο για τον χαρακτηρισμό της ποιότητας της ανθρώπινης γήρανσης, την ανίχνευση ασθενειών σε πρώιμο στάδιο, την παρακολούθηση των ασθενειών, καθώς και των αποτελεσμάτων της θεραπείας τους. Εν κατακλείδι, καθώς αυξάνεται ο αριθμός των συσκευών συλλογής και επεξεργασίας δεδομένων, η χρήση χειρόγραφου κειμένου για τον εντοπισμό και την παρακολούθηση των συνθηκών υγείας γίνεται όλο και πιο ελκυστική. Η ανάλυση γραφής σαφώς και δεν ανταγωνίζεται εγγενώς τις παραδοσιακές στρατηγικές υγείας, δύναται όμως να προσφέρει μια διαφορετική και συμπληρωματική προσέγγιση για την ανάλυση των προβλημάτων υγείας με μη επεμβατικό και οικονομικά αποδεκτό τρόπο [5].

Ένα μεγάλο πρόβλημα που υπάρχει με την χρήση των χειρόγραφων δεδομένων στην επιστήμη γενικά είναι ο Γενικός Κανονισμός Προστασίας Δεδομένων (General Data Protection Regulation- GDPR) . Ο GDPR θεσπίζει τους νέους κανόνες για την επεξεργασία δεδομένων προσωπικού χαρακτήρα εντός της ευρωπαϊκής ένωσης καθώς και για την χρήση τους εκτός των συνόρων αυτής. Στα πλαίσια του GDPR τα βιομετρικά δεδομένα θεωρούνται ειδική κατηγορία προσωπικών δεδομένων και η επεξεργασία αυτών υπόκεινται σε αυστηρούς κανόνες έως ότου ο κάθε γραφέας δώσει ρητή συγκατάθεση του [5], [7].

### **1.3 Κατηγοριοποίηση των συστημάτων χειρόγραφων βιομετρικής ταυτοποίησης.**

Στο κεφάλαιο αυτό θα γίνει μια ανασκόπηση των συστημάτων βιομετρικής ταυτοποίησης με χρήση της χειρόγραφης υπογραφής. Καθώς αυξάνεται ολοένα η χρήση της τεχνολογίας, αυξάνεται αντίστοιχα και το πλήθος των τεχνικών της βιομετρικής αναγνώρισης. Στην συνέχεια θα μας απασχολήσουν θέματα όπως η χρήση της απλής απεικόνισης της υπογραφής, για παράδειγμα σε μια εικόνα, τα λεγόμενα και ως στατικά μέσα, έως την χρήση της σε δυναμικά ψηφιακά μέσα όπως αυτά είναι μια ψηφιακή γραφίδα ή ένα tablet, τα λεγόμενα και ως δυναμικά μέσα. Παραδείγματα συστημάτων και των δύο κατηγοριών φαίνονται στην **Εικόνα 5**. Τέλος θα γίνει μια αναφορά στο τι πραγματεύονται και πως λειτουργούν τα τοπικά συστήματα και τα καθολικά συστήματα επιβεβαίωσης υπογραφής.



Εικόνα 5. Δυναμικά και στατικά συστήματα, Στα δυναμικά συστήματα υπογραφής, αξίζει να σημειωθεί πως γίνεται χρήση γραφίδας και η αναπαράσταση της υπογραφής γίνεται με χρονικά μεταβαλλόμενα δεδομένα, ενώ στα στατικά συστήματα υπογραφής τα δεδομένα προέρχονται από ψηφιακές εικόνες.<sup>3</sup> [2].

### 1.3.1 Δυναμικά συστήματα επιβεβαίωσης υπογραφής.

Στα δυναμικά συστήματα επιβεβαίωσης υπογραφής (*online*), όπως διακρίνονται στην **Εικόνα 6**, ο δυνητικός χρήστης παρέχει την υπογραφή του σε ένα κατάλληλο ψηφιακό μέσο, όπως ένα tablet, ή μια ψηφιακή γραφίδα. Στην κατηγορία αυτή, τα κύρια δεδομένα που εξετάζονται και λαμβάνονται σε πραγματικό χρόνο είναι κυρίως, η πίεση της γραφίδας στο ψηφιακό μέσο, η θέση της γραφίδας στους άξονες  $x, y$ , η γωνία που σχηματίζεται μεταξύ της γραφίδας και της επιφάνειας, ο χρόνος κατά τον οποίο η γραφίδα είναι στον αέρα, καθώς και η περιοχή που καταλαμβάνει το ίχνος της υπογραφής[5], [6].



Εικόνα 6. Ψηφιακά μέσα δυναμικών συστημάτων υπογραφής [6].

Αξίζει να σημειωθεί πως υπάρχει μεγάλη ποικιλία ηλεκτρονικών συσκευών που χρησιμοποιούνται σήμερα για την λήψη των χειρόγραφων υπογραφών. Όπως έγινε ήδη αναφορά παραπάνω, ο χρήστης έχει την δυνατότητα να υπογράψει σε ένα tablet, ή ένα έξυπνο κινητό είτε με την χρήση γραφίδας,

<sup>3</sup> [Δυναμικά και στατικά συστήματα υπογραφής](#)

είτε με το δάχτυλο του. Αντίστοιχη περίπτωση είναι και αυτή των ψηφιακών γραφίδων, όπου τόσο το υλικό κατασκευής της γραφίδας όσο και της επιφάνειας γραφής διαφέρει από συσκευή σε συσκευή. Αυτό έχει ως αποτέλεσμα να επηρεάζεται σημαντικά η αλληλεπίδραση των χρηστών με αυτά και ως εκ τούτου το τελικό αποτέλεσμα της υπογραφής να μην είναι πάντα ίδιο. Σύμφωνα με όσα αναφέρθηκαν προκύπτει το πρόβλημα της διαλειτουργικότητας, δηλαδή πρόβλημα ασυμβατότητας μεταξύ των ψηφιακών δυναμικών μέσων υπογραφής. Το πρόβλημα της διαλειτουργικότητας προκύπτει όταν τα ανωτέρω συστήματα υπογραφής χρησιμοποιούν διαφορετικά πρότυπα και αλγορίθμους για τη δημιουργία και τον έλεγχο των ψηφιακών υπογραφών. Για να λυθεί το πρόβλημα της διαλειτουργικότητας γίνεται χρήση μιας προσέγγισης δύο σταδίων. Το πρώτο είναι ένα στάδιο προεπεξεργασίας όπου τα δεδομένα από διαφορετικές συσκευές υποβάλλονται σε προεπεξεργασία για την κανονικοποίηση των σημάτων σε παρόμοιο εύρος. Το δεύτερο είναι η επιλογή των πιο ισχυρών χαρακτηριστικών για την περαιτέρω ελαχιστοποίηση των επιπτώσεων της διαλειτουργικότητας των συσκευών[5], [8].

Ένας ακόμη παράγοντας που επηρεάζει το τελικό αποτέλεσμα του ίχνους της υπογραφής και δυσκολεύει το έργο των δυναμικών συστημάτων επαλήθευσης υπογραφής, είναι η στάση του σώματος την οποία έχει ο υποψήφιος όταν δημιουργεί το ίχνος της υπογραφής στο ψηφιακό μέσο. Σε διάφορες μελέτες που έγιναν, παρατηρήθηκε πως οι συσκευές που βασίζονται σε γραφίδα είχαν καλύτερα αποτελέσματα στην απόδοση ενός βιομετρικού συστήματος όταν ο χρήστης παρέμενε σε μια καρέκλα και η συσκευή ακουμπούσε σε ένα τραπέζι, κάτι το οποίο θεωρείτε ως η πιο συνηθισμένη κατάσταση. Από την άλλη πλευρά, οι συσκευές που βασίζονται στην χρήση των δακτύλων για την δημιουργία των χειρόγραφων δειγμάτων, παρείχαν τα καλύτερα αποτελέσματα στην απόδοση στα σενάρια όπου ο χρήστης έπρεπε να χειριστεί τη συσκευή χωρίς κάποια υποστήριξη. Μια ακόμη παρατήρηση είναι πως όσο καλύτερη είναι η οπτική ανατροφοδότηση, τόσο καλύτερη είναι η απόδοση και η χρηστικότητα. Στην πραγματικότητα δε, διαπιστώθηκε ότι οι χρήστες δεν ένιωθαν άνετα όταν δεν υπήρχε οπτική ανατροφοδότηση της υπογραφής [5].

### **1.3.2 Στατικά συστήματα επιβεβαίωσης υπογραφής.**

Τα στατικά συστήματα υπογραφής (offline) είναι τεχνολογίες που βασίζονται στην αναγνώριση και επαλήθευση της υπογραφής ενός ατόμου με βάση την ψηφιακή εικόνα της. Πρόκειται λοιπόν για χαρακτηριστικά τα οποία μπορούν να μετρηθούν χωρίς την χρήση των δυναμικών χαρακτηριστικών, δηλαδή γίνεται χρήση της στατικής ψηφιακής δομής της υπογραφής. Αυτά τα χαρακτηριστικά, συνήθως περιλαμβάνουν τη γεωμετρία, τις αποστάσεις, τις γωνίες και τα διάφορα δομικά στοιχεία τα οποία προκύπτουν από την επεξεργασία της ψηφιακής εικόνας της υπογραφής. Οι διαδικασίες των στατικών μέσων υπογραφής συνήθως ακολουθούν τις τυπικές διαδικασίες ενός συστήματος Υπολογιστικής Όρασης και Αναγνώρισης Προτύπων (Computer Vision & Pattern Recognition-CVPR). Αυτά περιλαμβάνουν, την συλλογή των δεδομένων (Acquisition), την εξαγωγή των χαρακτηριστικών (feature extraction) και τέλος την επαλήθευση της υπογραφής (classification). Ποιο αναλυτικά η διαδικασία της συλλογής των δεδομένων, αφορά την περίπτωση κατά την οποία, ο χρήστης υπογράφει συνήθως πρώτα σε κάποιο μέσο - χαρτί, (σε αντίθεση με τα δυναμικά μέσα που όπως είδαμε ο χρήστης υπογράφει απ' ευθείας σε κάποιο ψηφιακό μέσο) και στη συνέχεια, η υπογραφή αυτή ψηφιοποιείται για να μετατραπεί σε ψηφιακή μορφή, δηλαδή σε μια εικόνα. Εν συνεχεία, η διαδικασία της εξαγωγής χαρακτηριστικών περιγράφει την διαδικασία κατά την οποία, τα δεδομένα της εικόνας- υπογραφής επεξεργάζονται για την εξαγωγή χαρακτηριστικών που μπορούν να αναπαραστήσουν τις δομικές ιδιότητες της υπογραφής, όπως για παράδειγμα τη γεωμετρία της. Τέλος, η διαδικασία της επαλήθευσης αντιστοιχεί στην περίπτωση κατά την οποία παρέχονται άγνωστες εικόνες υπογραφής προς ταυτοποίηση. Έτσι, κατά την διαδικασία της

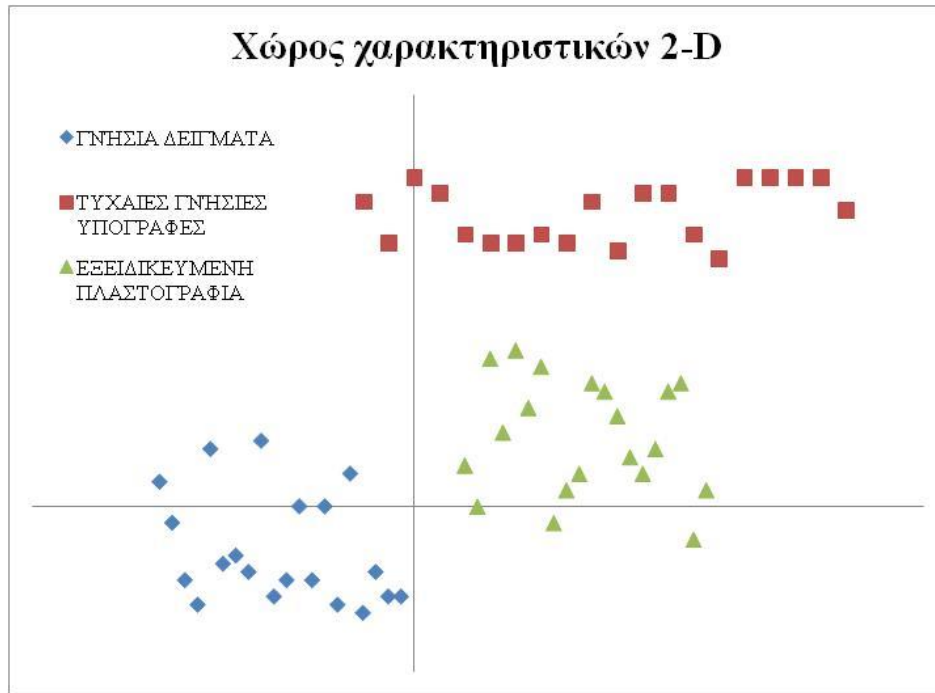
ταυτοποίησης μιας άγνωστης (query) υπογραφής-εικόνας, τα στατικά χαρακτηριστικά της συγκρίνονται με τα χαρακτηριστικά εικόνων αναφοράς τα οποία έχουν ήδη αποθηκευτεί στο σύστημα ταξινόμησης διαμέσου αλγορίθμων αναγνώρισης προτύπων ή μηχανικής μάθησης. Έτσι, το σύστημα καθορίζει εάν η άγνωστη υπογραφή που παρουσιάστηκε σε αυτό αντιστοιχεί ή όχι στον χρήστη αναφοράς. Τα στατικά μέσα επιβεβαίωσης υπογραφής είναι σχετικά απλά στην υλοποίηση και μπορούν να προσφέρουν γρήγορες επαληθεύσεις. Ωστόσο, είναι πιο ευάλωτα σε επιθέσεις απάτης και μπορεί να παρουσιάσουν προκλήσεις όταν οι υπογραφές διαφέρουν σημαντικά λόγω διακυμάνσεων στη γραφή.

Στα στατικά συστήματα επιβεβαίωσης υπογραφής, τα δεδομένα, δηλαδή οι εικόνες των χειρόγραφων υπογραφών, άρα και οι αντίστοιχες βάσεις βιομετρικών δεδομένων, αποθηκεύονται και συλλέγονται σε φύλλα χαρτιού. Έχει παρατηρηθεί το φαινόμενο, μια στατική υπογραφή να επικαλύπτει το χώρο που είναι διαθέσιμος για μια άλλη υπογραφή. Για παράδειγμα, το φύλλο χαρτιού που δίνονται οι υπογραφές συχνά είναι χωρισμένο σε πλαίσια. Σε αυτές τις περιπτώσεις συναντάται συχνά το πρόβλημα της επικάλυψης της υπογραφής από μία άλλη, καθώς δεν τηρούνται τα όρια του πλαισίου που είναι χωρισμένο το χαρτί από τους υπογράφοντες. Εξαιτίας αυτών των φαινομένων, ορισμένες υπογραφές απορρίπτονται χειροκίνητα, δηλαδή δεν συμμετέχουν κατά την ανάπτυξη των βάσεων βιομετρικών δεδομένων. Ένας ακόμη παράγοντας που περιορίζει την απόδοση ενός στατικού συστήματος αποτελεί και η σάρωση χαμηλής χωρικής (spatial) ανάλυσης, η οποία οδηγεί σε κακή ποιότητα εικόνας. Είναι σαφές ότι όσο υψηλότερη είναι η χωρική ανάλυση, τόσο περισσότερες πληροφορίες που απαιτούνται για την επεξεργασία της εικόνας μπορούν να εξαχθούν. Μια μέση χωρική ανάλυση που συναντάται στις περισσότερες βάσεις βιομετρικών δεδομένων στατικών μέσων υπογραφής είναι τα 600 dpi [5].

### 1.3.3 Τοπικά συστήματα επιβεβαίωσης υπογραφής (Writer Dependent)

Στα μοντέλα επιβεβαίωσης υπογραφής προσαρμοσμένα στον κάθε γραφέα (Writer Dependent -WD) δημιουργούνται ξεχωριστά μοντέλα εκμάθησης καθώς- επαλήθευσης για κάθε χρήστη-γραφέα, δηλαδή το σύστημα εκπαιδεύεται για κάθε ένα χρήστη ξεχωριστά. Ο χρήστης- γραφέας αυτός πρέπει να παρέχει ένα πλήθος αυθεντικών- γνήσιων υπογραφών ως αναπαράσταση της θετικής τάξης ώστε να είναι εφικτή η εκπαίδευση και η επαλήθευση του συστήματος μηχανικής μάθησης, δηλαδή του δυαδικού ταξινομητή. Από την άλλη, γίνεται εύκολα αντιληπτό ότι σε ένα πρόβλημα δυαδικής ταξινόμησης, όπως αυτό της επιβεβαίωσης υπογραφής, χρειάζεται η δημιουργία της αρνητικής τάξης, δηλαδή των λεγόμενων και ως πλαστών ή ψεύτικων υπογραφών. Για την δημιουργία της θετικής τάξης ο κάθε χρήστης χρειάζεται να παρέχει ένα πλήθος γνήσιων υπογραφών του, όπου μετά από το στάδιο της εξαγωγής χαρακτηριστικών, αυτές οι υπογραφές δίνονται σε έναν ταξινομητή. Όπως παρατηρείται στην **Εικόνα 7** τα γνήσια δείγματα ομαδοποιούνται σε μια τυχαία περιοχή στον χώρο. Αξίζει να σημειωθεί πως για κάθε ένα χρήστη ο χώρος χαρακτηριστικών που ομαδοποιούνται τα δείγματα υπογραφών είναι διαφορετικά. Για την δημιουργία της αρνητικής τάξης παρέχεται ένα σύνολο εξειδικευμένων πλαστογραφιών, και ένα σύνολο τυχαίων γνήσιων υπογραφών άλλου χρήστη, άρα όπως εύκολα γίνεται αντιληπτό η αρνητική τάξη αναπαρίσταται με 2 τρόπους, χωρίς όμως αυτό να επηρεάζει το τελικό αποτέλεσμα απόφασης. Στα τοπικά συστήματα επιβεβαίωσης υπογραφής παρατηρείται πως το κάθε σύνολο γνήσιων, εξειδικευμένων πλαστών υπογραφών ή τυχαίων υπογραφών άλλου χρήστη ομαδοποιείται ή αλλιώς συγκλίνει σε μια συγκεκριμένη περιοχή στον χώρο όπως φαίνεται στην **Εικόνα 7**. Κατά την φάση της δοκιμής του μοντέλου εισάγεται μια νέα υπογραφή, η οποία ταξινομείται ως γνήσια ή πλαστή σύμφωνα με την ομάδα που θα είναι πιο κοντά στον χώρο όπως στην περίπτωση της **Εικόνας 7**, το όριο απόφασης γνησιότητας της υπογραφής είναι στην περιοχή των μπλε δειγμάτων. Όπως είναι προφανές, για να γίνει σωστή

εκπαίδευση των μοντέλων επιβεβαίωσης υπογραφής προσαρμοσμένα στον κάθε γραφέα, θα πρέπει να δοθεί ένα μεγάλο πλήθος δειγμάτων, τόσο γνήσιων όσο και πλαστών υπογραφών, ώστε να μπορέσουν να δημιουργηθούν σωστά και με σαφήνεια οι περιοχές ενδιαφέροντος του ταξινομητή. Αρνητικά των συστημάτων αυτών είναι ο μεγάλος όγκος δειγμάτων - υπογραφών που πρέπει να δώσει ο κάθε χρήστης. Ένα ακόμη μειονέκτημα είναι ότι το σύστημα πρέπει να εκπαιδεύεται συνεχώς για κάθε ένα ξεχωριστό χρήστη καθώς όλοι οι χρήστες έχουν μοναδικά χαρακτηριστικά [9], [10], [11], [12], [13].

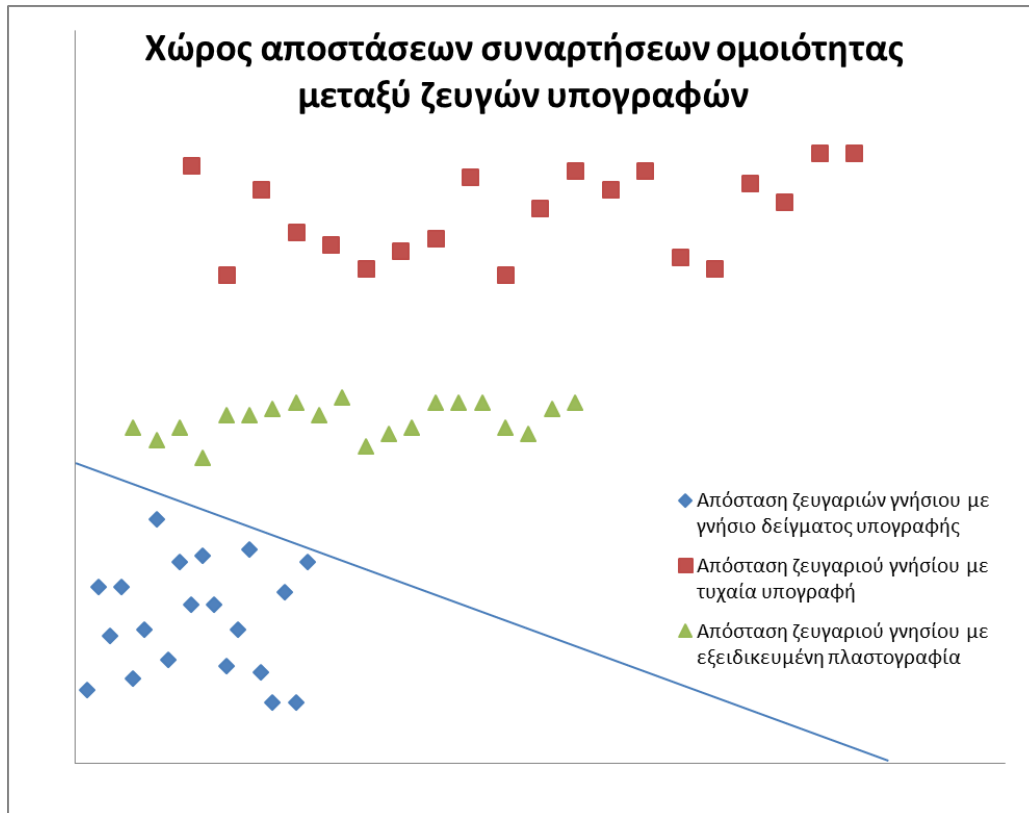


Εικόνα 7. Χώρος χαρακτηριστικών τοπικών συστημάτων επιβεβαίωσης υπογραφής.

### 1.3.4 Καθολικά συστήματα επιβεβαίωσης υπογραφής (Writer Independent)

Στα καθολικά συστήματα επιβεβαίωσης υπογραφής (**WI**), αντί το μοντέλο να εκπαιδευτεί με ένα σύνολο από γνήσιες και πλαστές υπογραφές στον χώρο των χαρακτηριστικών, εκπαιδεύεται με ομοιότητες που προκύπτουν από ζευγάρια μεταξύ γνήσιων δειγμάτων (θετική τάξη) και γνήσιων-πλαστών δειγμάτων (αρνητική τάξη). Έτσι, το μοντέλο **WI** εκπαιδεύεται ως ένα κοινό σύστημα που μαθαίνει να ξεχωρίζει όμοια από ανόμοια ζεύγη υπογραφών, όπως αυτά αναπαρίστανται από τις αποστάσεις των χαρακτηριστικών διανυσμάτων τους. Στην περίπτωση αυτή, το πλήθος των υπογραφών που χρειάζονται ανά γραφέα είναι πολύ μικρότερο σε σχέση με τα τοπικά συστήματα επιβεβαίωσης υπογραφής. Για παράδειγμα, για κάθε  $N$  το πλήθος γνήσιες υπογραφές ενός γραφέα, μπορούν να δημιουργηθούν  $\frac{N \times (N-1)}{2}$  ζεύγη, κάτι το οποίο σημαίνει ότι για  $M$  το πλήθος γραφέας που συμμετέχουν στην εκπαίδευση ενός συστήματος **WI** παράγονται  $M \times \frac{N \times (N-1)}{2}$  ζευγάρια θετικής τάξης. Αντίστοιχα αναπαρίσταται και η αρνητική τάξη, μόνο που στην περίπτωση αυτή υπάρχουν δύο τρόποι αναπαράστασης, ο ένας τρόπος είναι συγκρίνοντας τις διαφορές ενός γνήσιου δείγματος υπογραφής με μια εξειδικευμένη πλαστογραφία, ενώ ο δεύτερος τρόπος είναι συγκρίνοντας ένα γνήσιο δείγμα του γραφέα προς εξέταση με ένα τυχαίο δείγμα άλλου γραφέα. Στην περίπτωση της θετικής τάξης, οι διαφορές των δειγμάτων του ίδιου γραφέα θα είναι κοντά στην περιοχή του μηδέν, καθώς έχουμε να συγκρίνουμε δύο όμοιες υπογραφές. Από την άλλη πλευρά της αρνητικής τάξης συγκρίνοντας τα ζεύγη γνήσιων- εξειδικευμένων πλαστογραφημένων υπογραφών, η διαφορά που θα προκύψει δεν θα είναι μηδενική, αλλά κάτι που πλησιάζει στο μηδέν. Όσον αφορά τον δεύτερο τρόπο

αναπαράστασης δηλαδή των ζευγών γνήσιων υπογραφών του ατόμου προς εξέταση- γνήσιων υπογραφών τυχαίου ατόμου, η διαφορά που προκύπτει σίγουρα είναι κάτι αρκετά μεγαλύτερο του μηδενός, καθώς γίνεται σύγκριση ανόμοιων εικόνων [9], [10], [11], [12], [13].



Εικόνα 8. Χώρος αποστάσεων καθολικών συστημάτων επιβεβαίωσης υπογραφής.

## 1.4 Οι κατηγορίες των χειρόγραφων δειγμάτων υπογραφής

Στην ενότητα αυτή θα γίνει αναφορά στην εξέταση του χειρόγραφου δείγματος υπογραφής τόσο από την πλευρά της δικαστικής γραφολογίας, όσο και από την πλευρά της επιστήμης της πληροφορικής. Επίσης θα παρουσιαστούν οι κύριες διαφορές και ομοιότητες μεταξύ των ορισμών των διαφόρων μεθόδων και στις δύο αυτές επιστήμες. Οι κύριες διαφορές και ομοιότητες μεταξύ των ορισμών και των μεθόδων που χρησιμοποιούνται στις δύο αυτές επιστήμες προσφέρουν μια ευρεία και πολυποίκιλη προσέγγιση για την ανάλυση και την ερμηνεία των χειρόγραφων δειγμάτων υπογραφής, καθιστώντας τη συνεργασία μεταξύ των δύο επιστημονικών πεδίων εξαιρετικά εποικοδομητική.

### 1.4.1 Η περίπτωση της επιστήμης των δικαστικών γραφολόγων

Οι δικαστικοί γραφολόγοι δεν αντιμετωπίζουν τη βιομετρική ταυτοποίηση της υπογραφής ως ένα απλό πρόβλημα ταξινόμησης μεταξύ δύο κατηγοριών, αλλά ως ένα πρόβλημα που περιλαμβάνει πολλές παραμέτρους. Αυτό οφείλεται στο γεγονός ότι δεν περιορίζονται στην αξιολόγηση της αυθεντικότητας της υπογραφής μόνο με βάση το χειρόγραφο, αλλά εξετάζουν επίσης το περιβάλλον και τις συνθήκες που περιβάλλουν αυτήν, όπως την κατάσταση του χαρτιού **Εικόνα 9** (εάν είναι τσαλακωμένο, αν η υπογραφή καλύπτει άλλο κείμενο), το χρώμα της μελάνης, ή ακόμα και το πάχος της μύτης από το στυλό. Επιπλέον, λαμβάνουν υπόψη παράγοντες που μπορεί να επηρεάσουν τη γραφή του ατόμου, όπως η ψυχολογική και φυσική του κατάσταση, και αξιολογούν τη συνολική συμπεριφορά γραφής σε διάφορα δείγματα, προκειμένου να ανιχνεύσουν χαρακτηριστικά της κάθε



υπογραφής. Αυτή η ολιστική προσέγγιση διαφοροποιείται από την προσέγγιση που ακολουθείται στην πληροφορική, όπου συχνά η εστίαση είναι στην ανάπτυξη μοντέλων ταξινόμησης. Έτσι, οι γραφολόγοι εξετάζουν όχι μόνο την υπογραφή, αλλά και το υπόβαθρο και τις πολυδιάστατες πτυχές της [14].



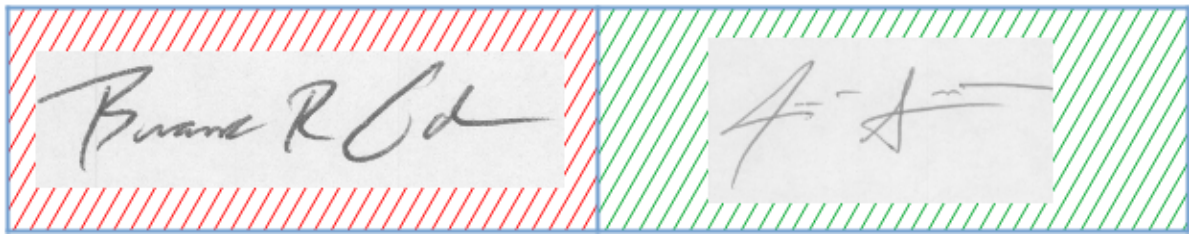
Εικόνα 9. Δείγμα υπογραφής προς εξέταση [15].

Στην περίπτωση των γραφολόγων οι κατηγορίες των πλαστογραφιών, δηλαδή της αρνητικής τάξης διακρίνονται σε τρεις υπο-κατηγορίες: α) στις μεταμφιεσμένες υπογραφές (*Disguised Signatures*), β) στις απλές πλαστογραφίες (*Simple Forgery*) και γ) στις ειδικευμένες πλαστογραφίες (*Skilled Forgery*). Η πρώτη κατηγορία, οι "μεταμφιεσμένες υπογραφές", χαρακτηρίζουν υπογραφές οι οποίες δεν είναι στην πραγματικότητα πλαστογραφημένες υπογραφές, αλλά προέρχονται από τον γνήσιο γραφέα, ο οποίος μεταμφιέζει τη δική του υπογραφή ώστε να την κάνει να φαίνεται πλαστή. Αυτό αποτελεί μια τακτική η οποία οδηγεί στην άρνηση της γνησιότητας της υπογραφής από την πλευρά του. Η δεύτερη κατηγορία είναι η "απλή πλαστογραφία", κατά την οποία ο πλαστογράφος γνωρίζει την πραγματική υπογραφή κάθε γραφέα και την πλαστογραφεί χωρίς καμία εξάσκηση ή επιπλέον φόρτο. Τέλος, η "ειδικευμένη πλαστογράφιση" σημαίνει ότι ο πλαστογράφος γνωρίζει από πριν την μορφή της πραγματικής υπογραφής κάθε γραφέα και την πλαστογραφεί μετά από επανειλημμένη εξάσκηση, δηλαδή παρατηρούμε το αποτέλεσμα μιας επαγγελματικής διαδικασίας που στην βιβλιογραφία μπορεί να εμφανιστεί και με το όνομα καλλιγραφία [14].

#### 1.4.2 Η περίπτωση της επιστήμης της πληροφορικής

Η περίπτωση της πληροφορικής και ιδιαίτερα η επεξεργασία εικόνας και η υπολογιστική όραση δεν εξετάζει το είδος του χαρτιού ή τις φθορές που μπορεί να έχει αυτό, όπως κάνουν οι δικαστικοί γραφολόγοι, παρά μόνο εξαλείφει τυχόν τέτοιες ανεπιθύμητες πληροφορίες με διάφορες τεχνικές ψηφιακής επεξεργασίας εικόνας, ώστε να πάρει μόνο το καθαρό ίχνος του δείγματος της υπογραφής. Στην περίπτωση της πληροφορικής, η πλαστογράφιση μπορεί να ταξινομηθεί σε τέσσερις κατηγορίες: α) τυχαία πλαστογράφιση (*Random Forgery*), β) απλή πλαστογράφιση (*Simple/Casual Forgery*), γ) εξομοιωμένη πλαστογράφιση (*Simulated*) και δ) εξειδικευμένη πλαστογράφιση (*Skilled Forgery*). Πιο αναλυτικά, η πρώτη περίπτωση πλαστών υπογραφών είναι η τυχαία πλαστογράφιση, όπως διακρίνεται στην **Εικόνα 10**, πρόκειται για την περίπτωση που αναφέρεται στη γνήσια υπογραφή οποιουδήποτε υπογράφοντος εκτός από τον υπογράφοντα που ελέγχεται, δηλαδή όταν ο υπογράφων είναι ένα τυχαίο πρόσωπο (ή πλαστογράφος) που προσπαθεί να αντικαταστήσει τη γνήσια υπογραφή με τη δική του τυχαία υπογραφή. Η επόμενη περίπτωση είναι η απλή πλαστογράφιση, όπου ο πλαστογράφος γνωρίζει το όνομα του αρχικού υπογράφοντος, αλλά όχι την υπογραφή. Μια άλλη περίπτωση είναι η "εξομοιωμένη πλαστογράφιση", όπου ο πλαστογράφος δεν είναι επαγγελματίας πλαστογράφος αλλά όμως έχει εξασκηθεί στην υπογραφή του αρχικού γραφέα. Τέλος, μια άλλη κατηγορία είναι η "εξειδικευμένη πλαστογράφιση", στην περίπτωση αυτή, ο επαγγελματίας πλαστογράφος, γνωρίζοντας την ταυτότητα του γραφέα, δηλαδή

το όνομα, το επώνυμο και την υπογραφή του, προσπαθεί να αντιγράψει πιστά την πρωτότυπη υπογραφή, μια τέτοια περίπτωση φαίνεται στην **Εικόνα 11**, στα αριστερά.[5], [14].



Εικόνα 10. Παράδειγμα τυχαίας πλαστογράφησης. Δεξιά το γνήσιο δείγμα, αριστερά το τυχαίο πλαστό δείγμα.



Εικόνα 11. Αριστερά: Πλαστές υπογραφές. Δεξιά: γνήσιες υπογραφές (από τον ίδιο γραφέα).

## 1.5 Τι θα δούμε στη συνέχεια.

Στην συνέχεια της παρούσας διπλωματικής θα αναφερθούμε στην επαλήθευση χειρόγραφου δείγματος υπογραφής σε στατικά συστήματα ανεξάρτητου γραφέα. Τα χειρόγραφα δείγματα υπογραφών ανήκουν στη δημοφιλή βάση δεδομένων CEDAR, η οποία περιγράφεται στην παράγραφο 4.2. Ο προγραμματισμός του προτεινόμενου συστήματος θα γίνει με το περιβάλλον καθώς και την γλώσσα προγραμματισμού MATLAB. Ειδικότερα θα γίνει επεξεργασία των δεδομένων, καθώς και εξαγωγή χαρακτηριστικών με αλγόριθμους μηχανικής μάθησης όπως ο K-SVD για την εκμάθηση ενός λεξικού και την OMP, για την αραιή αναπαράσταση, δεδομένης της ύπαρξης ενός λεξικού, έννοιες οι οποίες αναλύονται στο κεφάλαιο **Error! Reference source not found.** Τέλος ο αλγόριθμος που θα εκπαιδευτεί για να αποφανθεί την γνησιότητα ή μη της υπογραφής, είναι αυτός των μηχανών διανυσματικής υποστήριξης (support vector machines) που αναφέρεται στην παράγραφο 4.1 του κεφαλαίου 4.

## 2 ΚΕΦΑΛΑΙΟ 2<sup>ο</sup>: Στοιχεία ψηφιακής επεξεργασίας εικόνας και στάδιο προεπεξεργασίας

### 2.1 Ψηφιακή επεξεργασία εικόνας (Digital image processing)

Η ψηφιακή επεξεργασία εικόνας αναφέρεται στην διαδικασία επεξεργασίας εικόνων με χρήση ψηφιακών υπολογιστών. Αυτή η διαδικασία περιλαμβάνει την εφαρμογή διαφόρων τεχνικών και μεθόδων για τη βελτίωση, αποκατάσταση, ανάλυση, επεξεργασία και εξαγωγή πληροφοριών από ψηφιακές εικόνες με σκοπό την επίτευξη συγκεκριμένων στόχων και εφαρμογών. Η ψηφιακή επεξεργασία εικόνας χρησιμοποιείται ευρέως σε πολλούς τομείς, όπως ιατρική, ρομποτική, ασφάλεια, επικοινωνίες, αναγνώριση προτύπων και σε πολλούς ακόμα [16], [17].

Μια ασπρόμαυρη εικόνα μπορεί να οριστεί ως μια συνάρτηση δύο διαστάσεων  $f(x, y)$ , όπου  $x, y$  είναι χωρικές συντεταγμένες. Οι συντεταγμένες των εικονοστοιχείων (pixel) είναι στο επίπεδο  $x-y$ , ενώ οι διαβαθμίσεις του γκρι (gray scale) θεωρούνται ως τιμές στον άξονα  $z$  (βάθος της εικόνας). Το εικονοστοιχείο αναφέρεται στα στοιχεία που αποτελούν τη ψηφιακή εικόνα και αντιπροσωπεύουν τις διάφορες περιοχές της εικόνας με συγκεκριμένες τιμές φωτεινότητας ή χρώματος. Κάθε εικονοστοιχείο έχει μια συγκεκριμένη θέση στην εικόνα και αποτελεί τη βασική μονάδα επεξεργασίας σε ψηφιακή επεξεργασία εικόνας [16], [17].

Η επεξεργασία εικόνας και η υπολογιστική όραση αποτελούν δύο αναπόσπαστα μέρη του ευρύτερου πεδίου της υπολογιστικής επεξεργασίας εικόνας. Αν και δεν υπάρχουν αυστηρά οριοθετημένα όρια μεταξύ των δύο, μπορούμε να αναγνωρίσουμε τρία κύρια επίπεδα επεξεργασίας που αποτελούν μια χρήσιμη κατηγοριοποίηση. Τα τρία επίπεδα επεξεργασίας σε αυτό το φάσμα είναι: το χαμηλό, το μεσαίο και τέλος το υψηλό επίπεδο. Οι χαμηλού επιπέδου διαδικασίες περιλαμβάνουν πρωτόγονες λειτουργίες όπως η προεπεξεργασία εικόνας για μείωση θορύβου, ενίσχυση της αντίθεσης και την βελτίωση της ευκρίνειας της εικόνας. Οι διαδικασίες αυτές χαρακτηρίζονται από το γεγονός ότι τόσο η είσοδος όσο και η έξοδος τους είναι εικόνες. Η επεξεργασία μεσαίου επιπέδου περιλαμβάνει εργασίες όπως η κατάτμηση (διαίρεση μιας εικόνας σε περιοχές ή αντικείμενα), η περιγραφή αυτών των αντικειμένων για τη μετατροπή τους σε μια μορφή κατάλληλη για την επεξεργασία από υπολογιστή και η ταξινόμηση (αναγνώριση) των μεμονωμένων αντικειμένων. Μια διαδικασία μεσαίου επιπέδου χαρακτηρίζεται από το γεγονός ότι η είσοδος της είναι γενικά εικόνες, αλλά η έξοδος της είναι χαρακτηριστικά που εξάγονται από αυτές τις εικόνες (π.χ. ακμές, περίγραμμα και η ταυτότητα μεμονωμένων αντικειμένων). Τέλος, η επεξεργασία υψηλού επιπέδου περιλαμβάνει την "κατανόηση" μιας συλλογής αναγνωρισμένων αντικειμένων, όπως στην ανάλυση εικόνας, και, στο άλλο άκρο του φάσματος, την εκτέλεση των γνωστικών λειτουργιών που συνήθως συσχετίζονται με την όραση. Εδώ, η επεξεργασία γίνεται σε επίπεδο σύνθεσης πληροφοριών και κατανόησης του περιεχομένου της εικόνας, ενεργώντας σε επίπεδα που υπερβαίνουν την απλή αναγνώριση αντικειμένων. [17].

### 2.2 Κατάτμηση εικόνας (Image segmentation)

Η κατάτμηση εικόνας είναι η διαδικασία η οποία αναλύει μια εικόνα σε τμήματα ή αντικείμενα. Το επίπεδο της λεπτομέρειας στο οποίο γίνεται η διαίρεση της εικόνας αυτής εξαρτάται από το πρόβλημα που επιχειρείται να επιλυθεί κάθε φορά. Αυτό σημαίνει ότι η διαδικασία της κατάτμησης θα πρέπει να σταματήσει όταν τα αντικείμενα ή οι περιοχές ενδιαφέροντος σε μια εφαρμογή έχουν εντοπιστεί. Για παράδειγμα όταν πρόκειται για μια εικόνα χειρόγραφου δείγματος υπογραφής, η περιοχή ενδιαφέροντος είναι μόνο η περιοχή που βρίσκεται το ίχνος ενός χειρόγραφου δείγματος

υπογραφής. Οι περισσότεροι αλγόριθμοι κατάτμησης βασίζονται σε δύο βασικές ιδιότητες των τιμών της έντασης, δηλαδή των τιμών της ασυνέχειας καθώς και της ομοιότητας. Στην πρώτη κατηγορία, η προσέγγιση είναι να χωρίσουμε μια εικόνα βασιζόμενοι σε απότομες αλλαγές στην ένταση, όπως οι ακμές. Οι κύριες προσεγγίσεις στη δεύτερη κατηγορία βασίζονται στο να χωρίσουμε μια εικόνα σε περιοχές που είναι παρόμοιες σύμφωνα με ένα σύνολο προκαθορισμένων κριτηρίων. Οι μέθοδοι σε αυτή την κατηγορία περιλαμβάνουν την κατωφλιοποίηση (thresholding), την ανάπτυξη περιοχών (region growing), το διαχωρισμό (region splitting) και την συγχώνευση περιοχών (region merging)[17].

### 2.3 Στάδια προεπεξεργασίας εικόνας

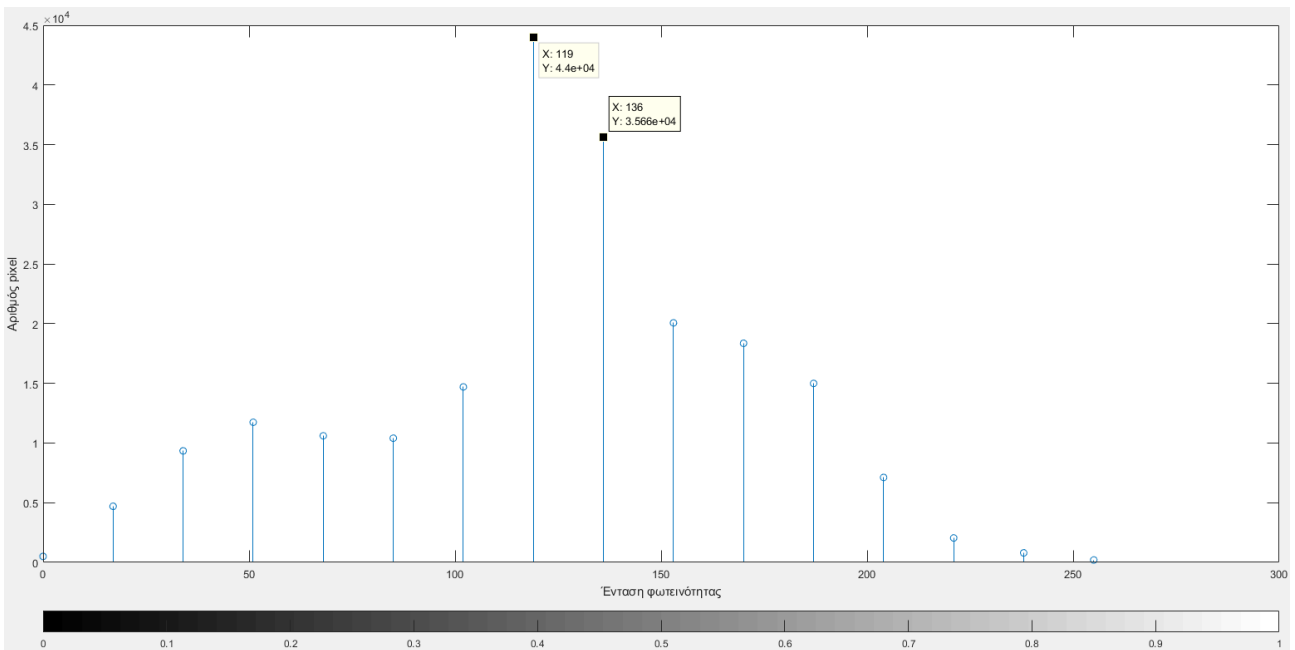
Τα στάδια της προεπεξεργασίας της εικόνας που εφαρμόζονται στην παρούσα εργασία, εξαιρουμένης της διαδικασίας εξαγωγής των patches η οποία περιεγράφηκε στην προηγούμενη παράγραφο, αποτελούν: η εξάλειψη του φόντου, η αφαίρεση του θορύβου με την χρήση φίλτρων (και συγκεκριμένα Gaussian), η κατωφλιοποίηση, και τέλος η λέπτυνση (thinning) της εικόνας.

Μια σημαντική τεχνική στον κόσμο της επεξεργασίας εικόνας είναι η κατωφλιοποίηση (thresholding), μια διαδικασία που επιτρέπει τον διαχωρισμό των διαφόρων στοιχείων μιας εικόνας, όπως αντικείμενα ή περιοχές ενδιαφέροντος, από το υπόβαθρο ή το φόντο. Κατά τη διαδικασία αυτή, εφαρμόζεται ένα κατώφλι στις τιμές έντασης των pixel, και έτσι τα pixels ταξινομούνται σε διαφορετικές κλάσεις ανάλογα με το εάν οι τιμές έντασης τους είναι υψηλότερες ή χαμηλότερες από το κατώφλι αυτό. Συνήθως, η επιλογή του κατωφλίου εξαρτάται από πολλούς παράγοντες, συμπεριλαμβανομένης της διακύμανσης των τιμών των pixel στην εικόνα. Κάποιοι από αυτούς τους παράγοντες είναι η απόσταση μεταξύ των κορυφών του ιστογράμματος της εικόνας, όπου μεγαλύτερη απόσταση μπορεί να βελτιώσει τον διαχωρισμό, το επίπεδο θορύβου στην εικόνα που επηρεάζει την ομαλότητα του ιστογράμματος, το μέγεθος των αντικειμένων σε σχέση με το υπόλοιπο φόντο, η ομοιομορφία του φωτισμού, και η ομοιομορφία των ιδιοτήτων της εικόνας. Είναι σημαντικό να κατανοήσουμε ότι οι τιμές κατωφλίου δεν είναι πάντα σταθερές και μπορεί να πρέπει να προσαρμόζονται ανάλογα με τις συνθήκες της κάθε εικόνας και την εφαρμογή που θέλουμε να επιτύχουμε. Η ακριβής επιλογή του κατωφλίου απαιτεί καλή κατανόηση του περιβάλλοντος και των συνθηκών της εικόνας, και μπορεί να είναι μια διαδικασία δοκιμής και σφάλματος μέχρι να επιτευχθεί η επιθυμητή απόδοση.

Η τεχνική κατωφλιοποίησης Otsu, η οποία πήρε το όνομά της από τον Nobuyuki Otsu, αποτελεί μια μέθοδο κατάτμησης εικόνας που χρησιμοποιείται για την αυτόματη εύρεση του κατωφλίου μιας εικόνας, ώστε να διαχωρίσει την εικόνα σε δύο διακριτές κατηγορίες (ή κλάσεις) το αντικείμενο που μας ενδιαφέρει και το φόντο. Το κατώφλι αυτό καθορίζεται ελαχιστοποιώντας τη διακύμανση της έντασης εντός της κλάσης ή, ισοδύναμα, μεγιστοποιώντας τη διακύμανση μεταξύ των κλάσεων. Η μέθοδος Otsu χρησιμοποιεί το ιστόγραμμα της εικόνας, το οποίο δείχνει πόσες φορές εμφανίζεται κάθε επίπεδο φωτεινότητας στην εικόνα. Αναλύοντας το ιστόγραμμα, η μέθοδος Otsu υπολογίζει τις πιθανότητες εμφάνισης των δύο κατηγοριών και τις μέσες τιμές των επιπέδων φωτεινότητας για κάθε κατηγορία. Να σημειωθεί πως το ιστόγραμμα μιας εικόνας είναι ένας υπολογιζόμενος μονοδιάστατος πίνακας τιμών. Στο διάγραμμα ιστογράμματος ο άξονας  $X$  αντιπροσωπεύει την κλίμακα της έντασης φωτεινότητας και έχει τιμές από 0 έως 255, ενώ ο άξονας  $Y$  αντιπροσωπεύει τον αριθμό των pixel που έχουν την δεδομένη τιμή έντασης φωτεινότητας. Η τεχνική Otsu είναι ιδιαίτερα χρήσιμη όταν η εικόνα έχει δύο ξεχωριστά και ευκρινή κορυφές στο ιστόγραμμα, που αντιπροσωπεύουν τις δύο κατηγορίες που θέλουμε να διαχωρίσουμε [18], [19].

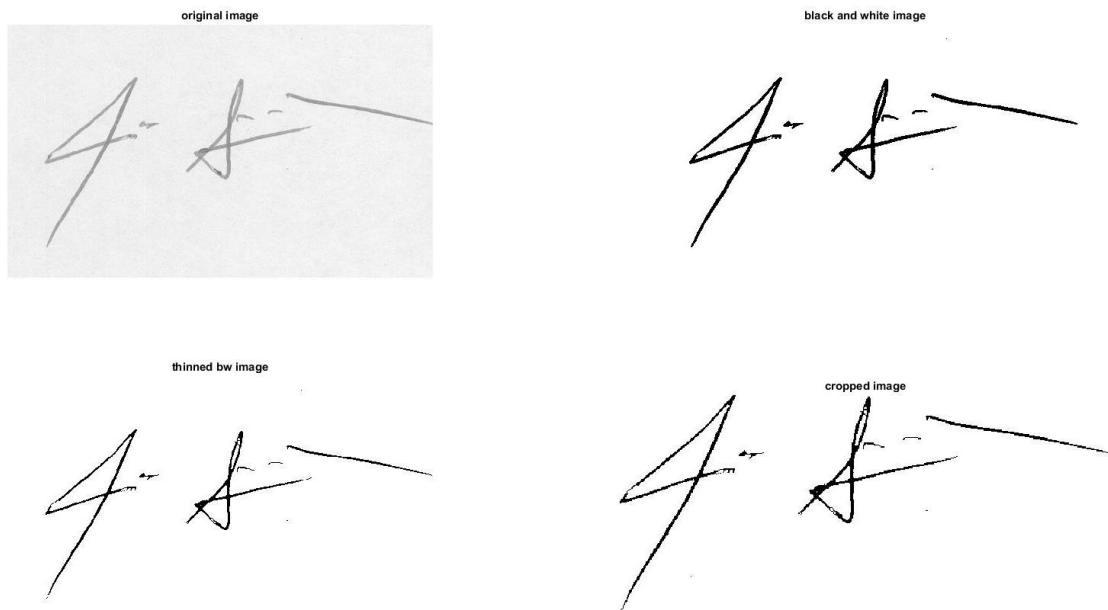


Εικόνα 12. Αριστερά απεικονίζεται μια εικόνα σε κλίμακα του γκρι, ενώ δεξιά απεικονίζεται η ίδια εικόνα που έχει περάσει από κατωφλιοποίηση με την μέθοδο Otsu.



Εικόνα 13. Παράδειγμα ιστογράμματος της προηγούμενης εικόνας.

Στην παρούσα εργασία η μέθοδος Otsu είναι το πρώτο στάδιο προεπεξεργασίας που χρησιμοποιείται για να μετατραπεί η αρχική γκρι εικόνα σε δυαδική άσπρη – μαύρη εικόνα. Να σημειωθεί πως το η γλώσσα προγραμματισμού Matlab περιέχει ενσωματωμένη την συνάρτηση (`greythresh`) η οποία υπολογίζει ένα επίπεδο κατωφλίου το οποίο μπορεί να χρησιμοποιηθεί για τη μετατροπή μιας εικόνας έντασης (`grayscale`) σε δυαδική εικόνα σύμφωνα με τη μέθοδο Otsu.



Εικόνα 14. Στάδια προεπεξεργασίας εικόνας. Η περιγραφή των σταδίων γίνεται στο κείμενο.

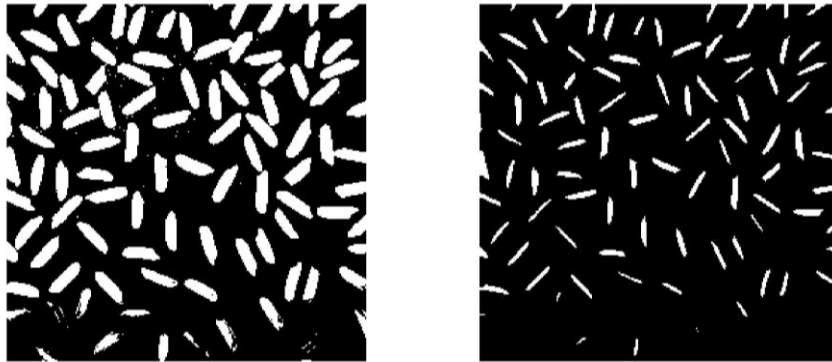
Στην **Εικόνα 14** φαίνονται τα αποτελέσματα της προεπεξεργασίας μιας γνήσιας εικόνας υπογραφής. Ειδικότερα πάνω αριστερά (*original image*) διακρίνεται η χειρόγραφη υπογραφή κατόπιν της σάρωσης, όπως παρατηρείται το χρώμα του φόντου καθώς και το ίχνος της υπογραφής είναι στην κλίμακα του γκρι. Έπειτα η εικόνα πάνω δεξιά υπογραφή (*black and white*) προκύπτει μετά την εξάλειψη του γκρι χρώματος φόντου με την χρήση του φίλτρου Otsu, και αφαίρεση τυχόν θορύβου, η εικόνα αυτή είναι της μορφής *black and white*. Η *thinned image* όπως είναι ορατό είναι η εικόνα που έχει περάσει από το στάδιο της λέπτυνσης. Τέλος η περικομμένη εικόνα (*cropped image*) αποτελεί την τελική εικόνα προς επεξεργασία, η οποία παρέχεται ως είσοδο επόμενο στάδιο, δηλαδή αυτό της εξαγωγής τμημάτων της υπογραφής.

## 2.4 Μορφολογία

Η μαθηματική μορφολογία αναδεικνύει μια ισχυρή και ενιαία μέθοδο για πολλές εφαρμογές στο πεδίο της επεξεργασίας εικόνας. Τα σύνολα στη μαθηματική μορφολογία αντιπροσωπεύουν τα αντικείμενα σε μια εικόνα. Για παράδειγμα, το σύνολο όλων των λευκών εικονοστοιχείων σε μια δυαδική εικόνα (*black and white*) μπορεί να θεωρηθεί ως μια πλήρη μορφολογική περιγραφή της εικόνας. Στις δυαδικές εικόνες, τα σύνολα που εξετάζονται είναι μέλη ενός χώρου δύο διαστάσεων ακεραίων, όπου κάθε στοιχείο ενός συνόλου είναι ένα δισδιάστατο διάνυσμα των οποίων οι συντεταγμένες είναι οι συντεταγμένες ενός λευκού ή μαύρου εικονοστοιχείου στην εικόνα. Οι μορφολογικές πράξεις εφαρμόζονται μεταξύ μιας περιοχής της εικόνας καθώς και ενός δομικού στοιχείου (*structural element*). Τα δομικά στοιχεία δημιουργούνται σε διάφορα σχήματα, αναλόγως με τα σχήματα των δομών προς επεξεργασία την κάθε φορά. Ένα σημείο του δομικού στοιχείου ονομάζεται αρχή (*origin*) και με βάση αυτού γίνονται οι κινήσεις του δομικού στοιχείου στη εικόνα. Να σημειωθεί πως η αρχή (*origin*) δεν επιβάλλεται να είναι πάντα το κέντρο του δομικού στοιχείου. Όσον αφορά τις εικόνες στην κλίμακα του γκρι, αυτές μπορούν να αναπαρασταθούν ως σύνολα όπου κάθε στοιχείο περιλαμβάνει δύο συνιστώσες, με τις πρώτες δύο συντεταγμένες να αντιστοιχούν στις συντεταγμένες ενός εικονοστοιχείου και η τρίτη συνιστώσα να αντιπροσωπεύει τη διακριτή τιμή έντασής του. Αντίστοιχα τα σύνολα σε χώρους υψηλότερης διάστασης μπορούν να περιέχουν άλλα χαρακτηριστικά εικόνας, όπως χρώμα και στοιχεία μεταβαλλόμενα με το χρόνο [17].

### 2.4.1 Διάβρωση (erosion) και διαστολή (dilation)

Δύο βασικές λειτουργίες σε ότι αφορά τη μορφολογική επεξεργασία εικόνας, είναι αυτές της διάβρωσης και της διαστολής. Χρησιμοποιούνται ευρέως σε πολλές εφαρμογές επεξεργασίας εικόνας για τη βελτίωση της ανάλυσης αντικειμένων και την εξαγωγή χαρακτηριστικών. Η διάβρωση, όπως απεικονίζεται από την **Εικόνα 15**, είναι μια μορφολογική λειτουργία που χρησιμοποιείται για τη συρρίκνωση ή την αραίωση αντικειμένων σε μια δυαδική εικόνα. Κατά την διάβρωση, ένα μικρότερο δομικό στοιχείο (structuring element) περνάει πάνω από την εικόνα και αφαιρεί τα pixels από τα άκρα των αντικειμένων, με αποτέλεσμα τη συρρίκνωσή τους. Η διάβρωση χρησιμοποιείται για την αφαίρεση μικρών λεπτομερειών ή θολών εικόνων, καθώς και για την απομόνωση των αντικειμένων από το φόντο [17].



**Εικόνα 15.** Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του erosion.

Η μαθηματική προσέγγιση της διαδικασίας της διάβρωσης μπορεί να γίνει με την ακόλουθη εξίσωση:

$$A \ominus B = \{z \mid (B)z \subseteq A\}, \quad \text{Εξ. 1}$$

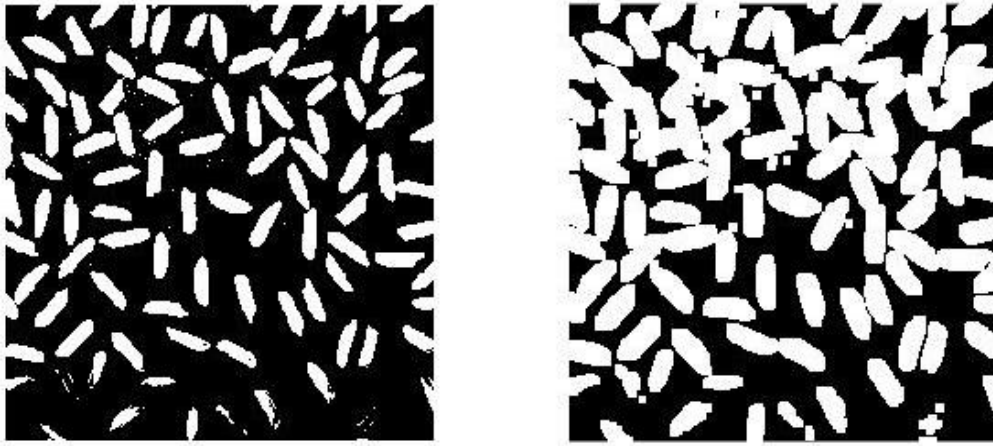
Όπου το  $A$  αναπαριστά την εικόνα πάνω στην οποία θα εφαρμοστεί η διαδικασία της διάβρωσης, το  $B$  αντιστοιχεί στο δομικό στοιχείο (structuring element) το οποίο έχει μετακινηθεί στην θέση  $z$ , το σύμβολο  $\subseteq$  σημαίνει πως το δομοστοιχείο  $B$  περιέχεται πλήρως μέσα στο  $A$ . Με άλλα λόγια, η διαδικασία erosion αφαιρεί τα σημεία από το σύνολο  $A$  που δεν είναι πλήρως καλυμμένα από το δομοστοιχείο  $B$ , συρρικνώνοντας έτσι το σύνολο.[17]

Η διαστολή, **Εικόνα 16**, είναι μια μορφολογική λειτουργία που χρησιμοποιείται για την επέκταση ή την ενίσχυση των αντικειμένων σε μια δυαδική εικόνα. Κατά τη διαστολή, το δομικό στοιχείο περνάει πάνω από την εικόνα και προσθέτει pixels στα άκρα των αντικειμένων, με αποτέλεσμα την επέκτασή τους. Η διαστολή χρησιμοποιείται για τη σύνδεση και την ενίσχυση των αντικειμένων στην εικόνα, καθώς και για τη γέμιση κενών ή τη σύνδεση απομονωμένων σημείων [17].

Η μαθηματική προσέγγιση της διαδικασίας της διαστολής δίνεται με την ακόλουθη εξίσωση:

$$A \oplus B = \{z \mid (B)_z \cap A \neq \emptyset\} \quad \text{Εξ. 2}$$

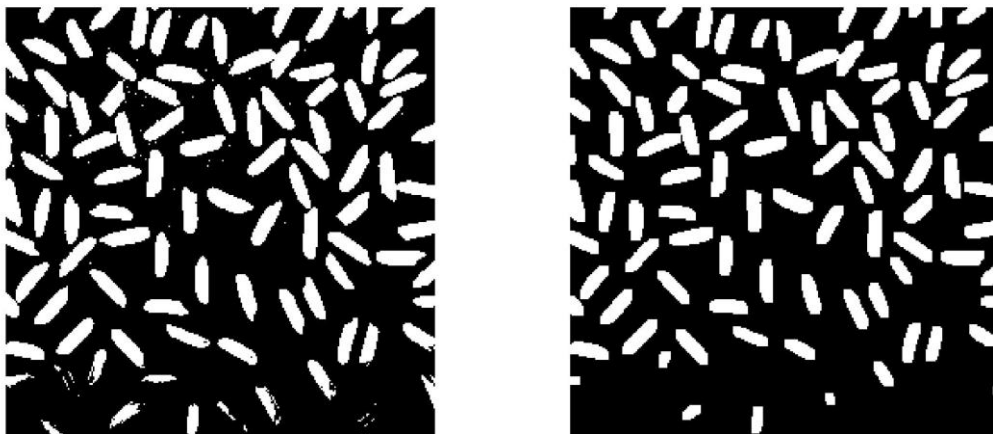
Στην παραπάνω εξίσωση, το  $(B)_z$  αναπαριστά το δομοστοιχείο  $B$  που έχει μετακινηθεί στη θέση  $z$  και το  $\cap$  σημαίνει την τομή των συνόλων. Άρα, η διαστολή ενός συνόλου  $A$  με ένα δομοστοιχείο  $B$  επεκτείνει το σύνολο  $A$  εφαρμόζοντας το δομοστοιχείο  $B$  σε κάθε σημείο του συνόλου και επισημαίνοντας τα σημεία όπου υπάρχει τουλάχιστον ένα σημείο του δομοστοιχείου που ταυτίζεται με το σύνολο  $A$  [17].



Εικόνα 16. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του dilation.

#### 2.4.2 Άνοιγμα (opening) και κλείσιμο (closing)

Δύο ακόμα βασικές μορφολογικά λειτουργίες στην επεξεργασία εικόνας όπου είναι συνδυασμός του erosion και του dilation, είναι αυτές του opening και closing. Τα φίλτρα αυτά είναι χρήσιμα στην επεξεργασία εικόνας για την βελτίωση της ανάλυσης των αντικειμένων, την εξάλειψη θολών και την επίτευξη επιθυμητών χαρακτηριστικών στην εικόνα. Ειδικότερα, το opening είναι μια μορφολογική λειτουργία που αποτελείται από τη συνδυασμένη εφαρμογή του erosion και του dilation σε μια εικόνα [17].



Εικόνα 17. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του opening.

Κατά το opening, πρώτα εφαρμόζεται το erosion για την αφαίρεση μικρών αντικειμένων ή λεπτομερειών, και στη συνέχεια εφαρμόζεται το dilation για την επαναφορά της μορφής των μεγαλύτερων αντικειμένων. Το opening χρησιμοποιείται για την αφαίρεση θολών, την απομόνωση αντικειμένων και την εξάλειψη μικρών θολών δομών.

Ο συμβολισμός του opening δίνεται από τον τελεστή « $\circ$ » και η μαθηματική εξίσωση που εκφράζει το opening είναι η ακόλουθη:

$$A \circ B = (A \ominus B) \oplus B \quad \text{Εξ. 3}$$

Στην παραπάνω εξίσωση, το « $\ominus$ » αναπαριστά τη διαδικασία της διάβρωσης (erosion) και το « $\oplus$ » αναπαριστά τη διαδικασία της διαστολής (dilation). Άρα, το opening ενός συνόλου A με ένα



δομοστοιχείο  $B$  πραγματοποιείται αρχικά με την εφαρμογή της διάβρωσης στο σύνολο  $A$  με το δομοστοιχείο  $B$ , και στη συνέχεια με τη διαδικασία της διαστολής με το ίδιο δομοστοιχείο  $B$  στο αποτέλεσμα της προηγούμενης διαδικασίας[17].

Σε αντίθεση, το closing είναι μια μορφολογική λειτουργία που αποτελείται από τη συνδυασμένη εφαρμογή του dilation και του erosion σε μια εικόνα. Κατά το closing, πρώτα εφαρμόζεται το dilation για τη σύνδεση μικρών κενών ή αντικειμένων, και στη συνέχεια εφαρμόζεται το erosion για την εξάλειψη μικρών κενών εντός των αντικειμένων. Το closing χρησιμοποιείται για τη σύνδεση αντικειμένων, την εξάλειψη μικρών θολών και την συμπλήρωση μικρών κενών στην εικόνα[17].



Εικόνα 18. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει τη διαδικασία του closing.

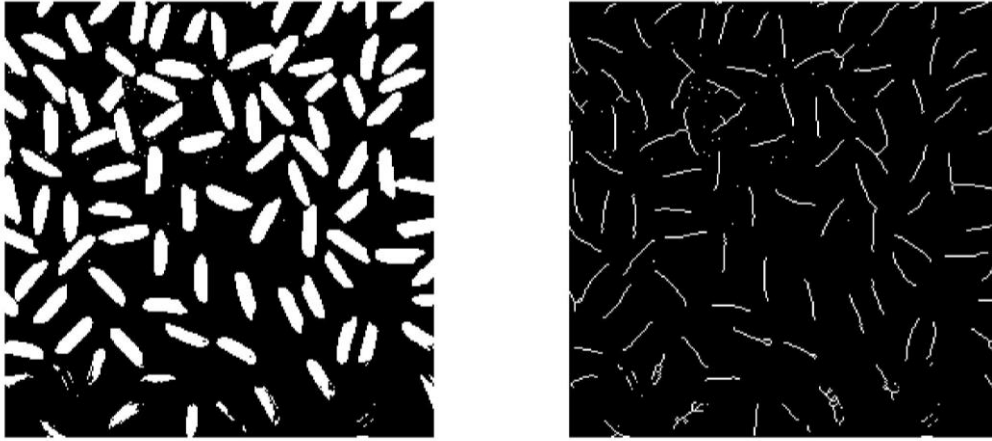
Η μαθηματική εξίσωση που διέπει το closing είναι η ακόλουθη:

$$A \cdot B = (A \oplus B) \ominus B \quad \text{Εξ. 4}$$

Έστω  $A$  το σύνολο (ή η εικόνα) πάνω στο οποίο θα εφαρμοστεί το closing, και έστω  $B$  το δομοστοιχείο (structuring element). Τότε το closing του συνόλου  $A$  με το δομοστοιχείο  $B$ , συμβολίζεται ως  $A \cdot B$ . Στην παραπάνω εξίσωση, το  $\oplus$  αναπαριστά τη διαδικασία της διαστολής (dilation) και το  $\ominus$  αναπαριστά τη διαδικασία της διάβρωσης (erosion). Άρα, το κλείσιμο ενός συνόλου  $A$  με ένα δομοστοιχείο  $B$  πραγματοποιείται αρχικά με την εφαρμογή της διαστολής στο σύνολο  $A$  με το δομοστοιχείο  $B$ , και στη συνέχεια με τη διαδικασία της διάβρωσης (erosion) με το ίδιο δομοστοιχείο  $B$  στο αποτέλεσμα της προηγούμενης διαδικασίας[17].

### 2.4.3 Λέπτυνση (Thinning)

Η λέπτυνση (thinning), Εικόνα 19. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του thinning. Εικόνα 19, είναι μια μορφολογική πράξη που χρησιμοποιείται στην επεξεργασία εικόνας για τη μείωση του πάχους των αντικειμένων σε μια δυαδική εικόνα, διατηρώντας παράλληλα το σχήμα και τη συνδεσιμότητά του. Ο στόχος της λέπτυνσης είναι η επαναληπτική αφαίρεση εικονοστοιχείων (pixel) από τα όρια των αντικειμένων έως ότου αυτά μειωθούν σε γραμμές πλάτους ενός μόνο εικονοστοιχείου. Η λέπτυνση χρησιμοποιείται συνήθως σε εφαρμογές όπως η αναγνώριση χαρακτήρων, η ανάλυση δακτυλικών αποτυπωμάτων, η επεξεργασία χειρόγραφων δειγμάτων και η επεξεργασία ιατρικών εικόνων.



Εικόνα 19. Αριστερά απεικονίζεται η αρχική εικόνα, ενώ δεξιά απεικονίζεται η εικόνα που έχει περάσει την διαδικασία του thinning.

Η μαθηματική έκφραση για τη διαδικασία της λέπτυνση" (thinning) στη μορφολογική επεξεργασία εικόνας είναι η εξής: Έστω  $A$  το σύνολο (ή η εικόνα) πάνω στο οποίο θα εφαρμοστεί η λέπτυνση. Τότε η λέπτυνση του συνόλου  $A$ , συμβολίζεται ως  $A \otimes B$  και ορίζεται ως εξής:

$$A \otimes B = A - (A \cdot B) \quad \text{εξ. 5}$$

Στην παραπάνω εξίσωση, το « $\cdot$ » αναπαριστά τη διαδικασία του κλεισίματος (closing) που περιεγράφηκε προηγουμένως, ενώ το « $-$ » αναπαριστά τη διαφορά μεταξύ των συνόλων. Άρα, η λέπτυνση ενός συνόλου  $A$  πραγματοποιείται με την αφαίρεση του αποτελέσματος του κλεισίματος του συνόλου  $A$  με ένα δομοστοιχείο  $B$  από το αρχικό σύνολο  $A$  [17].

### 3 ΚΕΦΑΛΑΙΟ 3<sup>ο</sup> Στοιχεία αραιής αναπαράστασης και εκμάθησης λεξικού

#### 3.1 Η αραιή αναπαράσταση και η εκμάθηση λεξικού.

Η αραιή αναπαράσταση (sparse representation) καθώς και η εκμάθηση λεξικού (dictionary learning) είναι τεχνικές όπου βοηθούν στην επίλυση προβλημάτων υπολογιστικής όρασης [20]. Με τον όρο αραιή αναπαράσταση νοείται η αναπαράσταση ενός σήματος με όσο το δυνατόν λιγότερους σημαντικούς συντελεστές, δηλαδή με συντελεστές οι οποίοι είναι μη μηδενικοί [21], [22]. Η συγκεκριμένη τεχνική είναι μια από τις πιο δημοφιλείς στο πεδίο της επεξεργασίας εικόνας, της υπολογιστικής όρασης και της μηχανικής μάθησης. Η αραιή αναπαράσταση είναι σημαντική σε πολλές εφαρμογές, όπως η συμπίεση δεδομένων, η ανίχνευση προτύπων, η επεξεργασία εικόνων, η αναγνώριση προτύπων κ.ά. Επιτρέπει την αντιμετώπιση του προβλήματος της υψηλής διάστασης και της πληθώρας δεδομένων μειώνοντας την πολυπλοκότητα της αναπαράστασης και εξάγοντας τα κυρίαρχα χαρακτηριστικά. Η θεωρία της αραιής αναπαράστασης επικεντρώνεται στον τρόπο εύρεσης των κατάλληλων βασικών συνιστωσών (atoms) και στον τρόπο αναπαράστασης των δεδομένων με βάση αυτές τις συνιστώσες. Υπάρχουν πολλές τεχνικές για την επίτευξη αραιών αναπαραστάσεων, όπως το Least Absolute Shrinkage and Selection Operator (LASSO), το Orthogonal Matching Pursuit (OMP), το Basis Pursuit (BP) κ.ά. Συνολικά, η αραιή αναπαράσταση αποτελεί ένα σημαντικό εργαλείο στον τομέα της επεξεργασίας σήματος και εικόνας, και επιτρέπει την αποτελεσματική αναπαράσταση και εξαγωγή χρήσιμης πληροφορίας από πολύπλοκα δεδομένα [23].

Η συνάρτηση κόστους του αλγορίθμου K-SVD αναφέρεται στη μετρική που χρησιμοποιείται για να αξιολογηθεί η ποιότητα των αναπαραστάσεων που παράγονται από τον αλγόριθμο. Συνήθως, αυτή η συνάρτηση κόστους σχετίζεται με τον βαθμό απόκλισης μεταξύ των πραγματικών δεδομένων και των αναπαραστάσεων που υπολογίζονται από τον αλγόριθμο. Είναι σημαντικό να σημειωθεί ότι η συνάρτηση κόστους μπορεί να ποικίλλει ανάλογα με τον συγκεκριμένο περιβάλλοντα χώρο εφαρμογής και τους στόχους της εκάστοτε εφαρμογής. Τρία προβλήματα βελτιστοποίησης που περιγράφουν τη συνάρτηση κόστους της αραιής αναπαράστασης είναι η συνάρτηση κανονικοποίησης ως συνάρτηση ποινής, η συνάρτηση κανονικοποίησης ως περιορισμός με βάση την τιμή του  $\mu$  (συνολικός αριθμός μη μηδενικών συντελεστών) και τέλος η τιμή σφάλματος ανακατασκευής  $\varepsilon$  [22], [24], [25].

Η συνάρτηση κανονικοποίησης σε μορφή πινάκων είναι η εξής:

$$\min_{D,A} \left\{ \frac{1}{2} \|X - DA\|_F^2 + \lambda \Psi(A) \right\} \quad \text{Εξ. 6}$$

Η συνάρτηση κανονικοποίησης (regularization) σαν περιορισμός με βάση την τιμή του  $\mu$  σε μορφή πινάκων είναι η εξής:

$$\min_{D,A} \left\{ \frac{1}{2} \|X - DA\|_F^2 \right\} \text{ s. t. } \Psi(A) \leq \mu \quad \text{Εξ. 7}$$

Αντίστοιχα χρησιμοποιώντας σαν περιορισμό την τιμή του σφάλματος ανακατασκευής  $\varepsilon$ :

$$\min_{D,A} \{ \Psi(A) \} \text{ s. t. } \|X - DA\|_F^2 \leq \varepsilon \quad \text{Εξ. 8}$$

Όπου  $\lambda$  είναι η παράμετρος κανονικοποίησης δηλαδή είναι ο συντελεστής που ρυθμίζει την συμπεριφορά του όρου  $\Psi(A)$ . Όπου  $\Psi(A)$  είναι ο όρος που περιγράφει την αραιότητα ορίζεται ως η

$l_p$ -νόρμα των συντελεστών  $a^i$ . Οι πιο κοινές μορφές είναι αυτές που βασίζονται στην  $l_0$ -νόρμα και  $l_1$ -νόρμα. Συγκεκριμένα, η  $l_p$ -νόρμα, ή αλλιώς ψευδο-νόρμα, μετρά τον αριθμό των μη μηδενικών στοιχείων, δηλαδή το επίπεδο αραιότητας στο διάνυσμα αναπαράστασης, λειτουργώντας ως άμεσος δείκτης αραιότητας. Το  $A = [a^1, a^2, a^n]$  περιγράφει τους αραιούς συντελεστές για  $n$  σήματα εισόδου.

Ο όρος  $\|X - DA\|_F^2$  εκφράζει το σφάλμα ανακατασκευής του σήματος εισόδου  $X$ , ενώ ο όρος  $\lambda\psi(A)$  εκφράζει τον περιορισμό της αραιότητας  $\lambda$ .

Η εκμάθηση λεξικού (*Dictionary Learning*) αναφέρεται σε μια διαδικασία μη εποπτευόμενης μάθησης, που στοχεύει στην ανάκτηση ενός "βέλτιστου" λεξικού για την αραιή αναπαράσταση δεδομένων. Συγκεκριμένα, στον τομέα της αραιής αναπαράστασης, το λεξικό αναφέρεται στο σύνολο των "βασικών" συνιστωσών (atoms) που χρησιμοποιούνται για να αναπαρασταθούν τα δεδομένα. Κάθε δεδομένο μπορεί να αναπαρασταθεί ως ένας γραμμικός συνδυασμός αυτών των βασικών συνιστωσών, με τον όρο ότι η αναπαράσταση θα είναι αραιή, δηλαδή περιλαμβάνει μόνο λίγες μη μηδενικές συνιστώσες. Έτσι, το λεξικό καθορίζει τον χώρο στον οποίο ζουν τα δεδομένα μας και τον τρόπο με τον οποίο αναπαρίστανται. Η διαδικασία της εκμάθησης λεξικού συνήθως περιλαμβάνει την εύρεση του βέλτιστου λεξικού, το οποίο είναι κατάλληλα προσαρμοσμένο στα δεδομένα που αναπαρίστανται. Αυτό επιτυγχάνεται με την ελαχιστοποίηση κάποιου κριτηρίου, όπως το σφάλμα ανακατασκευής των δεδομένων ή το σφάλμα της αραιής αναπαράστασης (sparsity), μέσω αλγορίθμων βελτιστοποίησης. Η εκμάθηση λεξικού έχει εφαρμογές σε πολλούς τομείς, όπως στην επεξεργασία εικόνας και σήματος, την επεξεργασία κειμένου, την επεξεργασία ήχου, την αναγνώριση προτύπων, την απομάκρυνση θορύβου και πολλά άλλα. Επιτρέπει την αποτελεσματική και συμπαγή αναπαράσταση των δεδομένων, καθιστώντας την κατάλληλη για πολύπλοκα και υψηλής διάστασης δεδομένα [23], [26], [27].

Η πρώτη φάση στην επίλυση του προβλήματος της αραιής αναπαράστασης είναι η κατασκευή του λεξικού. Αυτό το λεξικό μπορεί να είναι είτε προκαθορισμένο, δηλαδή να περιέχει ήδη ένα σύνολο διανυσμάτων που έχουν επιλεγεί εκ των προτέρων, είτε να προκύπτει από μια διαδικασία εκμάθησης, προσαρμόζοντας το περιεχόμενό του στα δεδομένα σήματα που διατίθενται. Η χρήση ενός προκαθορισμένου λεξικού είναι απλή και μπορεί να οδηγήσει σε γρήγορους αλγορίθμους, όπως ο βραχύχρονος μετασχηματισμός Fourier, αλλά η αποτελεσματικότητά του εξαρτάται από το πόσο καλά τα δεδομένα συμβιβάζονται με το λεξικό. Αντίθετα, ένα λεξικό που εκπαιδεύεται από τα ίδια τα δεδομένα μπορεί να προσαρμοστεί καλύτερα στα χαρακτηριστικά των σημάτων που περιλαμβάνουν, ξεπερνώντας τα περιορισμένα και γενικά λεξικά που είναι προκαθορισμένα. Η επιλογή του κατάλληλου λεξικού είναι ζωτικής σημασίας για την παραγωγή αραιών αναπαραστάσεων και εξαρτάται από τα χαρακτηριστικά των δεδομένων και τους στόχους της αναπαράστασης.

Μαθηματικά μιλώντας, ο όρος «λεξικό» στην αραιή αναπαράσταση είναι στην ουσία ένας πίνακας  $D \in \mathbb{R}^{m \times k}$ , δηλαδή ανήκει σε έναν αρχικό χώρο διαστάσεων  $m \times k$ . Οι στήλες του πίνακα αυτού ονομάζονται atoms και αποτελούν τα διανύσματα αναφοράς στον προαναφερόμενο υπερπλήρες χώρο. Το λεξικό αυτό λοιπόν, είθισται να είναι υπερπλήρες ( $m < k$ ), δηλαδή υπάρχουν περισσότερα άτομα για να περιγράψουν την διάσταση του χώρου του προβλήματος (δυναμική των δεδομένων) [22], [27]. Η χρήση ενός υπερπλήρους λεξικού προσφέρει τη δυνατότητα να ανακαλυφθούν ευκολότερα οι πολύπλοκες δομές και τα πρότυπα που υπάρχουν στα δεδομένα εισόδου. Αυτό

επιτρέπει στα άτομα του λεξικού να καταγράψουν χαρακτηριστικά υψηλότερου επιπέδου στα δεδομένα. Επιπλέον, οι συντελεστές αραιής αναπαράστασης δεν εξαρτώνται μόνο από το διάνυσμα εισόδου, αλλά και από τα άτομα του υπερπλήρους λεξικού. Σε αντίθεση με ένα πλήρες λεξικό όπου τα άτομα είναι συνήθως ορθοκανονικά, σε ένα υπερπλήρες λεξικό τα άτομα έχουν μοναδιαίο μέτρο και εστιάζουν σε εκείνες τις διευθύνσεις που είναι σημαντικές για το πρόβλημα που επιλύουμε[24].

### 3.1.1 Η εκμάθηση λεξικού με την μέθοδο K-SVD.

Ο K-SVD αλγόριθμος είναι η μία από τις πιο δημοφιλείς μεθόδους μηχανικής μάθησης. Στην πράξη είναι μια επαναληπτική μέθοδος η οποία εναλλάσσεται μεταξύ της αραιής κωδικοποίησης, και μιας διαδικασίας ενημέρωσης των ατόμων του λεξικού. Στόχος της μεθόδου αυτής είναι να σχεδιάσει ένα βέλτιστο υπερπλήρες λεξικό με όσο το δυνατό λιγότερους συντελεστές (αραιοί). Ως παράδειγμα χρήσης του αναφέρουμε την συμπίεση δεδομένων την αποθορυβοποίηση καθώς και την αναγνώριση του προσώπου [22], [26], [27]. Ο K-SVD εστιάζει στην ενημέρωση του λεξικού, για την οποία ενημέρωση γίνεται η υπόθεση ότι το λεξικό  $D$  καθώς και ο πίνακας αραιών συντελεστών  $A$  είναι σταθεροί. Επιλέγουμε μια στήλη  $d_k$  από το λεξικό  $D$ , επιλέγουμε την  $i$ -οστή γραμμή του πίνακα  $A$  η οποία συμβολίζεται ως  $a_i^T$ . Ο πίνακας  $E_i$  είναι ένας πίνακας σφάλματος, παρακάτω δίνεται η εξίσωση που εκφράζει τον αλγόριθμο K-SVD

$$\begin{aligned} \|X - DA\|_F^2 &= \left\| X - \sum_{j=1}^m d_j a_j^T \right\|_F^2 = \left\| \left( X - \sum_{j \neq i} d_j a_j^T \right) - d_i a_i^T \right\|_F^2 = \\ &= \|E_i - d_i a_i^T\|_F^2 \quad \text{Εξ. 9} \end{aligned}$$

Ο αλγόριθμος K-SVD έχει τέσσερις σημαντικές ιδιότητες που τον καθιστούν έναν εξαιρετικά επιτυχημένο αλγόριθμο στον τομέα της εκμάθησης προσαρμοζόμενων λεξικών. Οι ιδιότητες αυτές είναι η ευελιξία (flexibility), η απλότητα (simplicity), η αποδοτικότητα (efficiency), και τέλος μια καλά ορισμένη συνάρτηση κόστους (well-defined objective), όροι που θα αναλυθούν εκτενέστερα παρακάτω.

Αρχικά ο αλγόριθμος K-SVD έχει ευελιξία καθώς είναι ικανός να συνδυαστεί με οποιονδήποτε αλγόριθμο pursuit για την αραιή κωδικοποίηση, ανάλογα με τις ανάγκες του κάθε προβλήματος. Αυτό σημαίνει πως μπορεί να επιλεγεί ένας κατάλληλος αλγόριθμος που είτε ανταποκρίνεται στους χρονικούς περιορισμούς που έχουν τεθεί στο πρόβλημα, είτε μπορεί να χρησιμοποιηθεί σε μελλοντικό στάδιο σε συνδυασμό με το λαμβανόμενο λεξικό. Συνήθως, οι μέθοδοι σχεδιασμού υπερπλήρων προσαρμοζόμενων λεξικών διαθέτουν την παραπάνω ιδιότητα, καθώς το στάδιο της αραιής κωδικοποίησης και η ενημέρωση του λεξικού μπορούν να λειτουργήσουν ανεξάρτητα μεταξύ τους.

Ο K-SVD θεωρείται πολύ αποδοτικός αλγόριθμος λόγω της αριθμητικής του αποτελεσματικότητας και της ταχείας σύγκλισης που επιτυγχάνει. Η σύγκλιση σε τουλάχιστον ένα τοπικό ελάχιστο εγγυάται αν το μέσο τετραγωνικό σφάλμα αναπαράστασης (MSE) ακολουθεί μονοτονική μείωση κατά τη διάρκεια της διαδικασίας του αλγορίθμου. Κατά τη διάρκεια του σταδίου της ενημέρωσης του λεξικού, η ενημέρωση κάθε ατόμου  $d_\beta$  εγγυάται μία μείωση του MSE ή καμία μεταβολή του, χωρίς να παραβιάζεται ο περιορισμός αραιότητας μέσω της μεθόδου SVD. Κατά το στάδιο της

αραιής κωδικοποίησης, προσπαθούμε να ανακτήσουμε τα σήματα εκμάθησης χρησιμοποιώντας το πολύ  $T$  μη μηδενικούς συντελεστές για κάθε αναπαράσταση. Έτσι, σε κάθε βήμα αυτού του σταδίου, προσπαθούμε να μειώσουμε το συνολικό τετραγωνικό σφάλμα αναπαράστασης. Η επιτυχία αυτής της μείωσης, καθώς και η μονοτονική μείωση του MSE, εξαρτάται από την αποτελεσματικότητα των αλγορίθμων pursuit που χρησιμοποιούνται για την καλύτερη προσέγγιση της λύσης του προβλήματος. Είναι σημαντικό να σημειωθεί ότι η αποτελεσματικότητα αυτών των αλγορίθμων εξαρτάται από τους περιορισμούς αραιότητας του διανύσματος συντελεστών, οι οποίοι καθορίζονται από τα επιθυμητά χαρακτηριστικά του λεξικού. Όταν ο αριθμός των μη μηδενικών συντελεστών  $T$  είναι αρκετά μικρός σε σχέση με το μέγεθος  $n$  του λεξικού  $D$ , τότε οι αλγόριθμοι pursuit υπολογίζουν μία πολύ καλή προσέγγιση της ιδανικής λύσης.

Ο K-SVD διαθέτει την ιδιότητα της απλότητας. Αυτό αποδεικνύεται όταν λειτουργεί με ένα μόνο άτομο για την αναπαράσταση κάθε σήματος, δηλαδή όταν  $T=1$ . Σε αυτήν την περίπτωση,  $\Omega\beta = 0$  διότι ο πίνακας  $\Omega$  περιέχει μόνο τα σφάλματα που σχετίζονται με τα παραδείγματα-σήματα που χρησιμοποιούν το άτομο  $d_\beta$ , τα οποία δεν μπορούν να χρησιμοποιήσουν άλλα άτομα λόγω του  $T = 1$ . Επιπλέον, οι  $K$  ενημερώσεις των στηλών του  $D$  γίνονται πλέον ανεξάρτητες μεταξύ τους. Εάν επιπλέον επιβληθεί στον K-SVD ο μοναδικός μη μηδενικός συντελεστής για αυτό το άτομο να είναι ίσος με τη μονάδα, τότε το πρόβλημα που αντιμετωπίζει είναι το πρόβλημα της κλασσικής ομαδοποίησης. Κάτω από αυτές τις συνθήκες, το διάνυσμα αποτελείται μόνο από άσσους, δηλαδή  $\alpha_e^\beta = 1^T$ , και έτσι ο K-SVD εκτιμά, μέσω του SVD, τον καλύτερο πίνακα ενός βαθμού  $d_\beta 1^T$  που προσεγγίζει τον πίνακα  $E_e^\beta = Y_e^\beta$ . Άρα, η λύση είναι η μέση τιμή των στηλών του  $Y_e^\beta$ .

Ένα από τα κύρια πλεονεκτήματα της συνάρτησης κόστους του αλγορίθμου K-SVD είναι η καλά ορισμένη και κατανοητή φύση της. Αυτό σημαίνει ότι η συνάρτηση αυτή παρέχει μια σαφή εκτίμηση της ποιότητας της λύσης που παράγει ο αλγόριθμος. Η σαφήνεια αυτής της εκτίμησης είναι κρίσιμη, καθώς επιτρέπει στον αλγόριθμο να αποφύγει την παγίδα των τοπικών ελαχίστων και να συγκλίνει προς μια βέλτιστη λύση. Όταν οι ενέργειες του αλγορίθμου συντονίζονται με τον στόχο της ελαχιστοποίησης της συνάρτησης κόστους, αποφεύγονται οι ατέρμονες αναζητήσεις γύρω από τοπικά ελάχιστα. Αυτό σημαίνει ότι ο αλγόριθμος μπορεί να συγκλίνει πιο αποτελεσματικά προς την καλύτερη δυνατή λύση. Κατά συνέπεια, η συνάρτηση κόστους λειτουργεί ως καθοδηγητικός μηχανισμός, καθορίζοντας την κατεύθυνση προς την οποία πρέπει να κινηθεί ο αλγόριθμος προκειμένου να επιτύχει την επιθυμητή βελτιστοποίηση. Επιπλέον, η καλή ορισμένη συνάρτηση κόστους συμβάλλει στην αποτροπή της υπερεκτίμησης ή υποεκτίμησης των παραμέτρων κατά τη διάρκεια της εκτέλεσης του αλγορίθμου. Αυτό διασφαλίζει ότι ο αλγόριθμος δεν θα παρεκκλίνει από την ορθή πορεία και θα επιτύχει μια πιο αξιόπιστη και σταθερή σύγκλιση. Κατά συνέπεια, η συνάρτηση κόστους είναι ένα σημαντικό εργαλείο για την αξιολόγηση και βελτίωση της απόδοσης του αλγορίθμου K-SVD.[24].

### 3.1.2 Η αραιή αναπαράσταση με την μέθοδο OMP.

Η μέθοδος OMP (Orthogonal Matching Pursuit) αναφέρεται στην αραιή αναπαράσταση δεδομένων και είναι μια από τις τεχνικές που χρησιμοποιούνται για την εύρεση αραιών αναπαραστάσεων μέσω επιλογής επιμέρους atoms από ένα λεξικό. Αρχικά, υπάρχει ένα δεδομένο σήμα ή διάνυσμα που θέλουμε να αναπαραστήσουμε με αραιή αναπαράσταση χρησιμοποιώντας ένα λεξικό. Το λεξικό αποτελείται από μια συλλογή από "βασικές" συνιστώσες, οι οποίες είναι γνωστές. Στόχος μας είναι να επιλέξουμε ένα υποσύνολο από αυτές τις βασικές συνιστώσες, ώστε η αναπαράσταση του σήματος με βάση αυτές να είναι αραιή. Η μέθοδος OMP λειτουργεί επαναληπτικά. Αρχικά, η αναπαράσταση θέτεται σε μηδενικό διάνυσμα. Στη συνέχεια, επιλέγεται η βασική συνιστώσα που ΠΑΔΑ, Τμήμα Η&ΗΜ, Διπλωματική Εργασία, Δημητροπούλου Μαρία

έχει τη μεγαλύτερη συσχέτιση με το διάνυσμα των δεδομένων. Η συνιστώσα αυτή προστίθεται στην αναπαράσταση. Στη συνέχεια, αφαιρείται από το σήμα, και επαναλαμβάνεται η διαδικασία για να επιλεγούν οι επόμενες συνιστώσες. Η διαδικασία συνεχίζεται μέχρι να επιτευχθεί η επιθυμητή αραιή αναπαράσταση ή να επιτευχθεί κάποιο άλλο κριτήριο τερματισμού. Η μέθοδος OMP είναι ένας αποδοτικός και απλός αλγόριθμος, ιδανικός για περιπτώσεις όπου η αναπαράσταση είναι αραιή, δηλαδή οι περισσότερες συνιστώσες της είναι μηδενικές.[22], [23]

Ο OMP ανανεώνει τους συντελεστές που έχουν υπολογιστεί έως εκείνη τη στιγμή, λαμβάνοντας υπόψη το τρέχον σύνολο ατόμων πριν αφαιρεθούν για τον υπολογισμό του υπολοίπου. Για τον υπολογισμό των νέων τιμών των συντελεστών, χρησιμοποιείται ένας ψευδο-αντίστροφος πίνακας, ο οποίος συνδέεται με τη σχέση που ορίζει το σύνολο των επιλεγμένων ατόμων μέχρι εκείνη τη στιγμή. Με άλλα λόγια, ανανεώνει τους συντελεστές λαμβάνοντας υπόψη την παρούσα κατάσταση των ατόμων πριν από τις αλλαγές. Ο αλγόριθμος OMP επιλέγει σε κάθε βήμα το άτομο του λεξικού που έχει τη μεγαλύτερη σχέση με τα ήδη υπολογισμένα διανύσματα υπολοίπου. Ουσιαστικά, επιλέγει το άτομο που "ταιριάζει" καλύτερα με το υπόλοιπο του σήματος που απομένει να ανακαλυφθεί. Η ενημέρωση των υπολοίπων γίνεται με την ορθογώνια προβολή του εκτιμώμενου σήματος στο σύνολο των επιλεγμένων ατόμων μέχρι εκείνη τη στιγμή. Κάθε επαναληπτικό βήμα επιλέγει ένα άτομο που συμβάλλει στη μείωση του υπολοίπου, με βάση τη μέθοδο των ελαχίστων τετραγώνων. Ο αλγόριθμος επαναλαμβάνεται μέχρι να επιτευχθεί μια συγκεκριμένη συνθήκη τερματισμού, όπως ο αριθμός των επαναλήψεων να φτάσει σε ένα καθορισμένο όριο ή το υπόλοιπο να γίνει μικρότερο από ένα καθορισμένο επίπεδο. Το κλειδί για την απόδοση του αλγορίθμου είναι η συγκλίνουσα ικανότητά του σε μια ιδανική λύση σε λίγες επαναλήψεις [24], [25].

Στον αλγόριθμο Batch OMP, το ζητούμενο είναι η αραιή αναπαράσταση  $\gamma$ , τέτοια ώστε δεδομένου του λεξικού  $D$  και της εισόδου  $\chi$ , να επιστραφεί ο αραιός κώδικας  $\gamma$  μέσω της προσέγγισης  $\chi \approx D\gamma$ . Στην **Εικόνα 20** παρατίθεται η βασική έκδοση του ψευδοκώδικα για την OMP ενώ στην **Εικόνα 21** παρατίθεται ο ψευδοκώδικας της batch-omp, οποία αποτελεί μια πιο πρακτική υλοποίηση καθώς μπορεί να διαχειριστεί ένα μεγάλο πλήθος σημάτων [28].

---

**Algorithm** ORTHOGONAL MATCHING PURSUIT

---

- 1: Input: Dictionary  $\mathbf{D}$ , signal  $\underline{x}$ , target sparsity  $K$  or target error  $\epsilon$
  - 2: Output: Sparse representation  $\underline{\gamma}$  such that  $\underline{x} \approx \mathbf{D}\underline{\gamma}$
  - 3: Init: Set  $I := ()$ ,  $\underline{r} := \underline{x}$ ,  $\underline{\gamma} := \underline{0}$
  - 4: **while** (*stopping criterion not met*) **do**
  - 5:    $\hat{k} := \underset{k}{\text{Argmax}} |d_k^T \underline{r}|$
  - 6:    $I := (I, \hat{k})$
  - 7:    $\underline{\gamma}_I := (\mathbf{D}_I)^+ \underline{x}$
  - 8:    $\underline{r} := \underline{x} - \mathbf{D}_I \underline{\gamma}_I$
  - 9: **end while**
- 

Εικόνα 20. Ψευδοκώδικας της βασική υλοποίησης της OMP<sup>4</sup>.

---

**Algorithm** BATCH-OMP

---

- 1: Input:  $\underline{\alpha}^0 = \mathbf{D}^T \underline{x}$ ,  $\epsilon^0 = \underline{x}^T \underline{x}$ ,  $\mathbf{G} = \mathbf{D}^T \mathbf{D}$ , target error  $\epsilon$
  - 2: Output: Sparse representation  $\underline{\gamma}$  such that  $\underline{x} \approx \mathbf{D}\underline{\gamma}$
  - 3: Init: Set  $I := ()$ ,  $\mathbf{L} := [\mathbf{1}]$ ,  $\underline{\gamma} := \underline{0}$ ,  $\underline{\alpha} := \underline{\alpha}^0$ ,  $\delta^0 := 0$ ,  $n := 1$
  - 4: **while**  $\epsilon^{n-1} > \epsilon$  **do**
  - 5:    $\hat{k} := \underset{k}{\text{Argmax}} \{|\underline{\alpha}_k|\}$
  - 6:   **if**  $n > 1$  **then**
  - 7:      $\underline{w} := \text{Solve for } \underline{w} \left\{ \mathbf{L}\underline{w} = \mathbf{G}_{I, \hat{k}} \right\}$
  - 8:      $\mathbf{L} := \begin{bmatrix} \mathbf{L} & \underline{0} \\ \underline{w}^T & \sqrt{1 - \underline{w}^T \underline{w}} \end{bmatrix}$
  - 9:   **end if**
  - 10:    $I := (I, \hat{k})$
  - 11:    $\underline{\gamma}_I := \text{Solve for } \underline{c} \left\{ \mathbf{L}\mathbf{L}^T \underline{c} = \underline{\alpha}_I^0 \right\}$
  - 12:    $\underline{\beta} = \mathbf{G}_I \underline{\gamma}_I$
  - 13:    $\underline{\alpha} := \underline{\alpha}^0 - \underline{\beta}$
  - 14:    $\delta^n = \underline{\gamma}_I^T \underline{\beta}_I$
  - 15:    $\epsilon^n = \epsilon^{n-1} - \delta^n + \delta^{n-1}$
  - 16:    $n := n + 1$
  - 17: **end while**
- 

Εικόνα 21. Ψευδοκώδικας Batch OMP<sup>5</sup>.

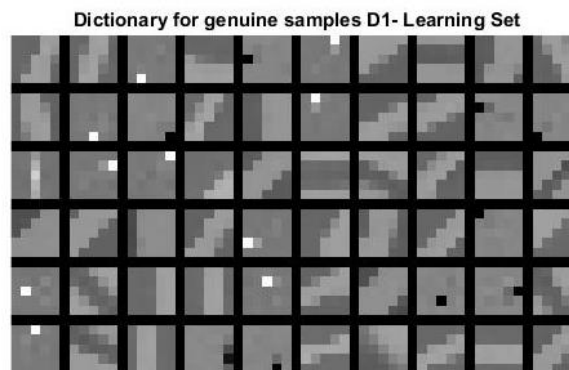
<sup>4</sup> R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," *CS Technion*, vol. 40, pp. 1-15, 2008.

<sup>5</sup> R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," *CS Technion*, vol. 40, pp. 1-15, 2008.



### 3.2 Σχέση της αραιής αναπαράστασης με το θέμα της διπλωματικής εργασίας

Στην παρούσα διπλωματική η χρήση της αραιής αναπαράστασης γίνεται για την δημιουργία διανυσμάτων χαρακτηριστικών από χειρόγραφες εικόνες υπογραφής. Ειδικότερα, η κάθε υπογραφή προς επεξεργασία δίνεται σε μια εικόνα ασπρόμαυρη, έτσι, κατά επέκταση η εικόνα αυτή αντιστοιχεί σε έναν πίνακα  $N \times M$  διαστάσεων, όσο και των pixels της κάθε εικόνας αντίστοιχα. Στο στάδιο της προεπεξεργασίας η εικόνα έχει μετατραπεί σε έναν πίνακα double με τιμές 0 έως 1. Αξίζει να σημειωθεί πως στην αναπαράσταση της εικόνας η τιμή 0 αντιστοιχεί στο λευκό, ενώ η τιμή 1, αντιστοιχεί στο μαύρο χρώμα. Όμως στο στάδιο των patches, αντιστρέφουμε τις τιμές αυτές για να γίνει εφικτή η αραιή αναπαράσταση. Σύμφωνα με την παρατήρηση αυτή, η εύρεση μη μηδενικών συντελεστών που χρειάζονται στην αραιή αναπαράσταση δεν θα μας διευκολύνει, διότι το άσπρο χρώμα που έχει το φόντο της υπογραφής αντιστοιχεί στο 1, ενώ το ίχνος της υπογραφής έχει τιμές από 0 έως 0,999. Με αποτέλεσμα να είναι αρκετά μεγάλος ο όγκος των αραιών συντελεστών και μάλιστα η περισσότερη πληροφορία είναι άχρηστη, καθώς προέρχεται από το χρώμα του φόντου. Άρα για να μπορέσει να επιτευχθεί η αραιή αναπαράσταση και να βρεθεί μόνο το ίχνος της υπογραφής, μετατρέπουμε το άσπρο φόντο της εικόνας σε μαύρο. Άρα η χρήσιμη πληροφορία πλέον αντιστοιχεί μόνο στο ίχνος της υπογραφής.



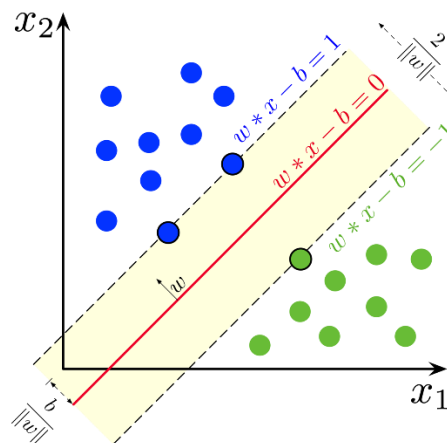
Εικόνα 22. Αναπαράσταση λεξικού για τα γνήσια δείγματα υπογραφών.

Στην **Εικόνα 22** απεικονίζεται η αναπαράσταση λεξικού για τα γνήσια δείγματα των υπογραφών, όπως αυτή έγινε στο στάδιο του learning set μετά την διαδικασία του K-SVD και OMP.

## 4 ΚΕΦΑΛΑΙΟ 4<sup>ο</sup> : Περιγραφή της πειραματικής διαδικασίας

### 4.1 Μηχανές διανυσματικής υποστήριξης (Support vector machines)

Οι Μηχανές Διανυσματικής Υποστήριξης (Support Vector Machines - SVM) αποτελούν μια κατηγορία αλγορίθμων μηχανικής μάθησης οι οποίοι χρησιμοποιούνται μεταξύ άλλων σε προβλήματα ταξινόμησης. Οι Μηχανές Διανυσματικής Υποστήριξης αποτελούν μία από τις πιο δημοφιλείς και αποτελεσματικές μεθόδους στην επεξεργασία και ανάλυση δεδομένων. Ανήκει στην κατηγορία αλγορίθμων επιβλεπόμενης μάθησης. Αυτό σημαίνει ότι κατά την εκπαίδευση του μοντέλου, τα δεδομένα πρέπει να περιλαμβάνουν τόσο τις εισόδους (χαρακτηριστικά) όσο και τις επιθυμητές εξόδους (ετικέτες) που είναι γνωστές για κάθε παράδειγμα εκπαίδευσης. Με βάση αυτές τις ετικέτες, το μοντέλο μαθαίνει να προβλέπει τις επιθυμητές εξόδους για νέες εισόδους. Η ιδέα πίσω από τα SVM είναι να βρεθεί ένα υπερεπίπεδο (ή πολλά υπερεπίπεδα, σε περίπτωση πολλαπλών κλάσεων) που διαχωρίζει τα δεδομένα των διαφορετικών κατηγοριών με τον καλύτερο δυνατό τρόπο. Ειδικότερα, επιλέγει αυτό το υπερεπίπεδο με τρόπο που τα δείγματα δεδομένων τα οποία βρίσκονται πιο κοντά σε αυτό το υπερεπίπεδο, γνωστά και ως "κοντινοί γείτονες" (nearest neighbors), να έχουν τη μέγιστη δυνατή απόσταση από αυτό. Τα δεδομένα που βρίσκονται πιο κοντά στο σημείο ενδιαφέροντος, είναι αυτά που παίζουν σημαντικό ρόλο στον καθορισμό του υπερεπιπέδου. Γενικά, στα SVM η διαδικασία αυτή οδηγεί σε μια βέλτιστη γενίκευση των μοντέλων και στην καλή απόδοση σε νέα δεδομένα [29], [30], [31], [32], [33].



Εικόνα 23. Παράδειγμα γραμμικού διανύσματος υποστήριξης δύο κλάσεων[34].

Το υπερεπίπεδο είναι το όριο απόφασης που χρησιμοποιείται για τον διαχωρισμό των διαφορετικών κλάσεων στον χώρο χαρακτηριστικών. Η μαθηματική εξίσωση που το περιγράφει το γραμμικό μοντέλο των SVM είναι της μορφής  $wx + b = 0$ . Όπου  $x$  είναι το διάνυσμα εισόδου,  $w$  είναι το διάνυσμα βαρών και  $b$  η σταθερά πόλωσης. Τα δεδομένα στοιχεία της εξίσωσης αυτής είναι μόνο το διάνυσμα εισόδου  $x$ , ενώ τα ζητούμενα είναι το διάνυσμα βαρών  $w$  καθώς και η σταθερά πόλωσης  $b$ . Δεδομένου ότι έχω γραμμικά διαχωρίσιμα δεδομένα, ο διαχωρισμός των κλάσεων περιγράφεται από τις εξισώσεις:

$$wx + b \geq +1, \text{ για ετικέτα δείγματος } t_k = +1 \quad \text{Εξ. 10}$$

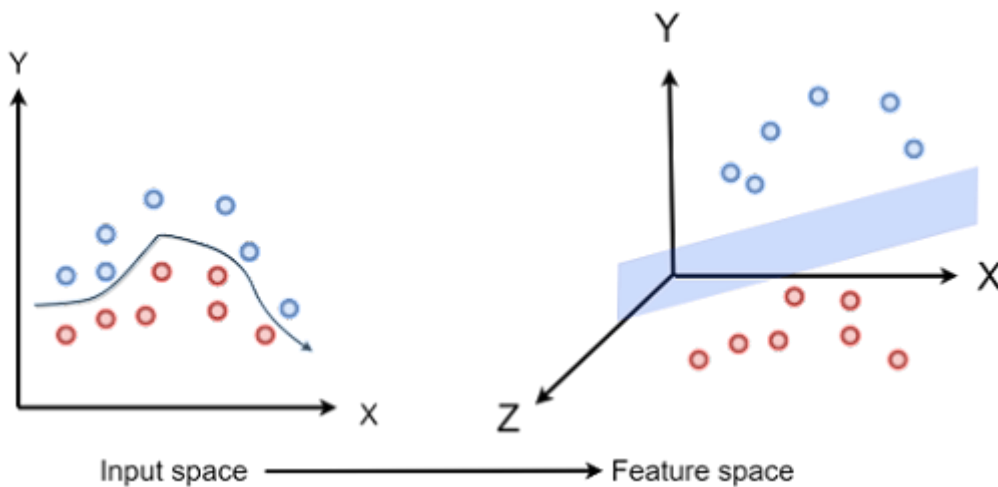
$$wx + b \leq -1, \text{ για ετικέτα δείγματος } t_k = -1 \quad \text{Εξ. 11}$$

Όπου  $k$  ο αριθμός του δείγματος. Οι ανωτέρω εξισώσεις μπορούν να ενσωματωθούν στην παρακάτω εξίσωση:

$$t_k(\mathbf{w}\mathbf{x} + \mathbf{b}) \geq +1, \text{ για } k = 1, 2, 3, \dots, N \quad \text{Εξ. 12}$$

Η βασική αποστολή των Μηχανών Διανυσματικής Στήριξης (ΜΔΣ) είναι να εντοπίσουν την επιφάνεια απόφασης με τον μέγιστο δυνατό βαθμό περιθωρίου, το οποίο ισούται με  $\frac{2}{\|\mathbf{w}\|_2}$  (Εικόνα 23). Τα διανύσματα που ικανοποιούν την ανωτέρω συνθήκη, αναφέρονται ως διανύσματα στήριξης. Αυτά τα διανύσματα βρίσκονται κοντά στην επιφάνεια λήψης απόφασης και συνεπώς, είναι εκείνα που κατηγοριοποιούνται με μεγαλύτερη δυσκολία από το σύνολο των διανυσμάτων εκπαίδευσης.

Πέραν των γραμμικών μοντέλων διανύσματος υποστήριξης υπάρχουν και τα αντίστοιχα μη γραμμικά μοντέλα διανύσματος υποστήριξης. Στην περίπτωση αυτή οι μηχανές διανυσματικής υποστήριξης χρησιμοποιούν συναρτήσεις πυρήνα για να διαχωρίσουν τα δεδομένα σε διαφορετικές κατηγορίες, εντοπίζοντας το βέλτιστο υπερεπίπεδο και μεγιστοποιώντας τις αποστάσεις μεταξύ των σημείων. Οι συναρτήσεις πυρήνα μετατρέπουν τα μη γραμμικά μοντέλα σε γραμμικά, προσφέροντας έτσι μεγαλύτερη σταθερότητα και ακρίβεια σε σχέση με τα παραδοσιακά γραμμικά μοντέλα. Στην ουσία η συνάρτηση πυρήνα είναι μια μαθηματική συνάρτηση που μετατρέπει ένα χώρο χαμηλής διάστασης σε ένα χώρο υψηλότερης διάστασης. Για παράδειγμα έστω ότι έχω μια είσοδο  $X$  που είναι δύο διαστάσεων, η συνάρτηση πυρήνα θα την μετατρέψει σε έναν χώρο τριών διαστάσεων. Δηλαδή θα αντιστοιχήσει τα δεδομένα σε έναν χώρο χαρακτηριστικών, όπου στο διάστημα αυτό θα είναι γραμμικά διαχωρίσιμα[33], [35], [36].



Εικόνα 24. Παράδειγμα RBF <sup>6</sup>

Ο πυρήνας RBF (Radial Basis Function Kernel) είναι ένας πολύ ισχυρός πυρήνας που χρησιμοποιείται στα SVM. Μπορεί να συνδυάσει πολλαπλούς πολυωνυμικούς πυρήνες διαφορετικών βαθμών για να προβάλλει τα μη γραμμικά διαχωρίσιμα δεδομένα σε χώρο υψηλότερων διαστάσεων, έτσι ώστε να μπορούν να διαχωριστούν χρησιμοποιώντας ένα υπερεπίπεδο. Για την προβολή των δεδομένων σε χώρο υψηλότερων διαστάσεων, ο πυρήνας RBF χρησιμοποιεί τη συνάρτηση που φαίνεται παρακάτω:

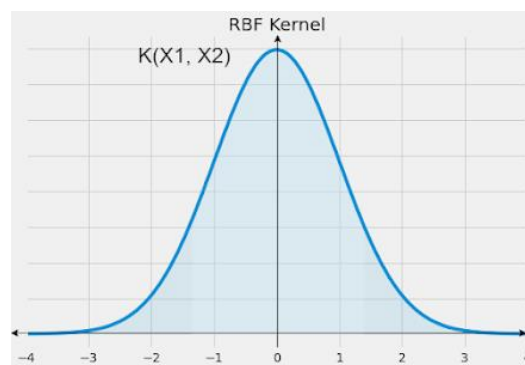
<sup>6</sup> <https://www.pycodemates.com/2022/10/the-rbf-kernel-in-svm-complete-guide.html>

$$K(X_1, X_2) = \exp\left(-\frac{\|X_1 - X_2\|^2}{2\sigma^2}\right) \quad \text{Εξ. 13}$$

Όπου  $\|X_1 - X_2\|^2$  είναι η τετραγωνική ευκλείδεια απόσταση, και  $\sigma$  είναι μια παράμετρος που καθορίζεται από τον χρήστη. Μεγαλύτερες τιμές του  $\sigma$  οδηγούν σε πιο "ευρεία" Gaussian συνάρτηση, ενώ μικρότερες τιμές οδηγούν σε πιο "απότομες" αλλαγές στην τιμή της συνάρτησης. Συνοπτικά ο όρος  $\sigma$  ελέγχει την επίδραση των κοντινών σημείων στην τελική απόφαση της μηχανής  $12\sigma^2$ , τότε αντίστοιχα η  $K(X_1, X_2) = \exp\left(-\frac{\|X_1 - X_2\|^2}{2\sigma^2}\right)$  Εξ. 13 θα γραφεί αντίστοιχα :

$$K(X_1, X_2) = \exp(-\gamma\|X_1 - X_2\|^2) \quad \text{Εξ. 14}$$

Όπως παρατηρείται η  $K(X_1, X_2) = \exp(-\gamma\|X_1 - X_2\|^2)$  Εξ. 14 έχει μια πιο απλουστευμένη μορφή, αυτή της ευκλείδειας απόστασης. Αξίζει να σημειωθεί, όπως έχει ήδη γίνει αναφορά παραπάνω, πως το γράφημα του πυρήνα RBF είναι της μορφής της καμπύλης Gauss ή αλλιώς κατανομής καμπάνας, όπως φαίνεται στην [33], [35], [36].

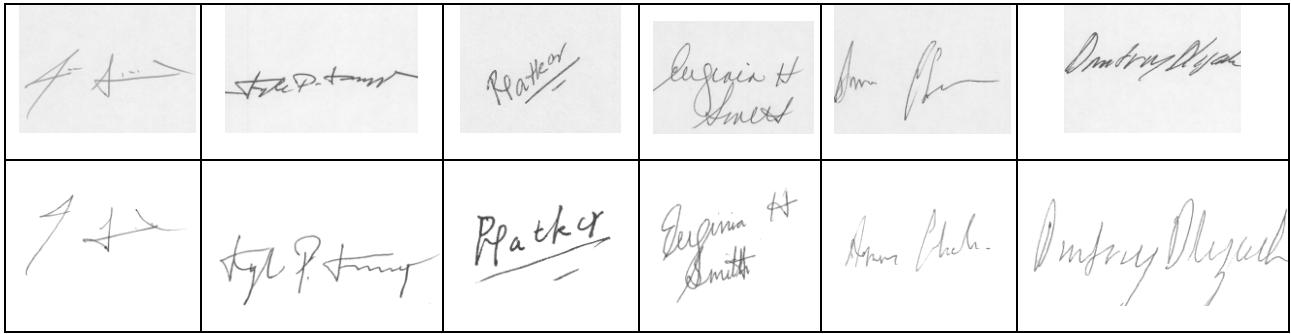


Εικόνα 25. Γράφημα κατανομής του RBF πυρήνα.<sup>7</sup>

## 4.2 Βάση δεδομένων

Στην συνέχεια θα ασχοληθούμε με ένα στατικό σύστημα επιβεβαίωσης υπογραφής, ανεξάρτητου γραφέα. Για τον λόγο αυτό, θα γίνει χρήση της βάσης δεδομένων CEDAR η οποία είναι δυτικής προέλευσης. Πρόκειται για μια βάση η οποία περιέχει χειρόγραφα δείγματα υπογραφών. Συγκεκριμένα η κάθε υπογραφή έχει δοθεί σε λευκό χαρτί με διαστάσεις  $50mm \times 50mm$  και έχει ψηφιοποιηθεί με ανάλυση 300dpi, η επέκταση των εικόνων αυτών είναι της μορφής '.png' ή Portable Network Graphic. Η βάση δεδομένων CEDAR αποτελείται από συνολικά υπογραφές από 55 γραφείς. Συγκεκριμένα το συνολικό πλήθος ανά γραφέα είναι 48 υπογραφές, όπου οι 24 εξ αυτών είναι γνήσιες (G) και οι υπόλοιπες 24 είναι πλαστές υπογραφές (SF) από τον εκάστοτε γραφέα [22], [37]. Δεν φαίνεται να υπάρχουν ή να έχουν αναφερθεί περιορισμοί για την χρήση της κατά την διάρκεια της συγγραφής της παρούσας διπλωματικής.

<sup>7</sup> <https://www.pycodermates.com/2022/10/the-rbf-kernel-in-svm-complete-guide.html>



Εικόνα 26. Ενδεικτικά δείγματα υπογραφών CEDAR από 5 γραφείς. Πάνω απεικονίζονται γνήσια δείγματα υπογραφής, ενώ κάτω απεικονίζονται τα πλαστά δείγματα υπογραφής.

Για να μελετηθούν οι τυχαίες πλαστές υπογραφές στην παρούσα διπλωματική εργασία, πραγματοποιείται η δημιουργία ενός συνόλου δεδομένων όπου συνδυάζονται γνήσιες υπογραφές από διαφορετικούς γραφείς. Αυτό σημαίνει ότι δημιουργούνται ζεύγη υπογραφών, όπου η μία πλευρά του ζεύγους ανήκει σε έναν γνήσιο γραφέα που θέλουμε να εξετάσουμε, ενώ η άλλη πλευρά ανήκει σε έναν τυχαία επιλεγμένο γνήσιο γραφέα. Αυτή η διαδικασία μίξης γνήσιων υπογραφών επιτρέπει τη δημιουργία ζευγών υπογραφών που αποτελούν αντιπροσωπευτικά δείγματα για την εκπαίδευση και αξιολόγηση μοντέλων αναγνώρισης πλαστών υπογραφών. Με αυτόν τον τρόπο, μπορούμε να διερευνήσουμε την αποτελεσματικότητα του μοντέλου μας στην αναγνώριση των τυχαίων πλαστών υπογραφών, χρησιμοποιώντας ζευγάρια υπογραφών που αποτελούνται από γνήσιες υπογραφές που προέρχονται από διαφορετικούς ανθρώπους.

### 4.3 Πειραματική διαδικασία και μετρικές σφάλματος

Οι μετρικές σφάλματος αναφέρονται σε δείκτες που χρησιμοποιούνται για να αξιολογηθεί η απόδοση ενός μοντέλου πρόβλεψης ή ταξινόμησης. Ο βασικός τρόπος αξιολόγησης είναι μέσω του συσχετισμού των προβλέψεων του μοντέλου με τις πραγματικές τιμές ή κατηγορίες των δεδομένων. Στο πλαίσιο αυτής της διπλωματικής εργασίας, οι κατηγορίες αναφέρονται στις γνήσιες και τις πλαστές υπογραφές. Οι μετρικές αυτές παρέχουν κατανοητό τρόπο να αξιολογηθεί η ακρίβεια και η απόδοση του μοντέλου σε σχέση με τα πραγματικά δεδομένα. Οι μετρικές που θα χρησιμοποιηθούν στην παρούσα εργασία είναι οι καμπύλες ROC (Receiver Operating Characteristic curve) όπου αξιολογούν την ικανότητα του μοντέλου να διαχωρίζει τις διαφορετικές κλάσεις, με τη χρήση της καμπύλης ROC και της περιοχής κάτω από αυτήν. Συγκεκριμένα η καμπύλη ROC είναι η γραφική παράσταση του ποσοστού αληθώς θετικών αποτελεσμάτων (True Positive Rate- TPR) έναντι του ποσοστού ψευδώς θετικών αποτελεσμάτων (False Positive Rate - FPR). Ειδικότερα, η τιμή AUC (Area Under Curve) κυμαίνεται από 0 έως 1 και αναπαριστά την πιθανότητα το εκπαιδευόμενο μοντέλο να αποδεχτεί ορθά ένα αυθεντικό δείγμα. Ένα μοντέλο του οποίου οι προβλέψεις είναι 100% λανθασμένες έχει AUC 0.0, ενώ ένα μοντέλο του οποίου οι προβλέψεις είναι 100% σωστές έχει AUC 1.0 [38], [39]. Παρακάτω φαίνεται ο υπολογισμός των TPR και FPR.

$$TPR = \frac{TP}{TP+FN} \quad \text{Εξ. 15} \quad \text{και} \quad FPR = \frac{FP}{FP+TN} \quad \text{Εξ. 16}$$

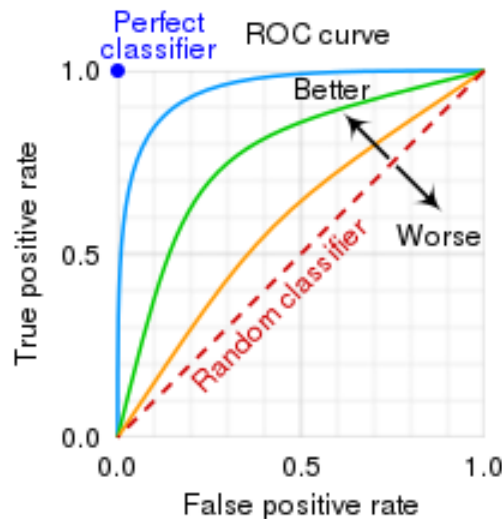
Στην αξιολόγηση των αλγορίθμων μηχανικής μάθησης, όπως οι αλγόριθμοι ταξινόμησης, συχνά χρησιμοποιούνται μετρικές όπως οι True Positives (TP), False Positives (FP), False Negatives (FN) και True Negatives (TN). Αυτές οι μετρικές προσδιορίζουν τον τρόπο με τον οποίο ένα μοντέλο αποδίδει σε σχέση με τις πραγματικές τάξεις των δεδομένων. Ειδικότερα:

*TP* (True Positives) : αληθώς θετικά, δηλαδή ένα αποτέλεσμα όπου το μοντέλο προβλέπει σωστά τη θετική τάξη. Ένα τέτοιο παράδειγμα στην παρούσα εργασία είναι εάν το μοντέλο αναγνωρίζει σωστά την γνησιότητα μιας γνήσιας υπογραφής.

*FP* (False Positives): ψευδώς θετικά, δηλαδή είναι ένα αποτέλεσμα όπου το μοντέλο προβλέπει εσφαλμένα τη θετική τάξη. Αυτό σημαίνει ότι το μοντέλο αναγνωρίζει μια κατάσταση ως θετική, ενώ στην πραγματικότητα δεν είναι. Για παράδειγμα, εάν το μοντέλο προβλέπει την πλαστή υπογραφή ως γνήσια.

*FN* (False Negatives): ψευδώς αρνητικά, δηλαδή είναι ένα αποτέλεσμα όπου το μοντέλο προβλέπει λανθασμένα την αρνητική τάξη, αντίστοιχο παράδειγμα είναι εάν το μοντέλο προβλέπει την γνήσια υπογραφή ως πλαστή.

*TN* (True Negatives): αληθώς αρνητικά, δηλαδή είναι ένα αποτέλεσμα όπου το μοντέλο προβλέπει σωστά την αρνητική τάξη. Αυτό σημαίνει ότι το μοντέλο αναγνωρίζει μια κατάσταση ως αρνητική και στη συνέχεια αποδεικνύεται ότι είναι πράγματι αρνητική, ένα αντίστοιχο παράδειγμα είναι εάν το μοντέλο προβλέψει σωστά ότι η πλαστή υπογραφή δεν είναι γνήσια.



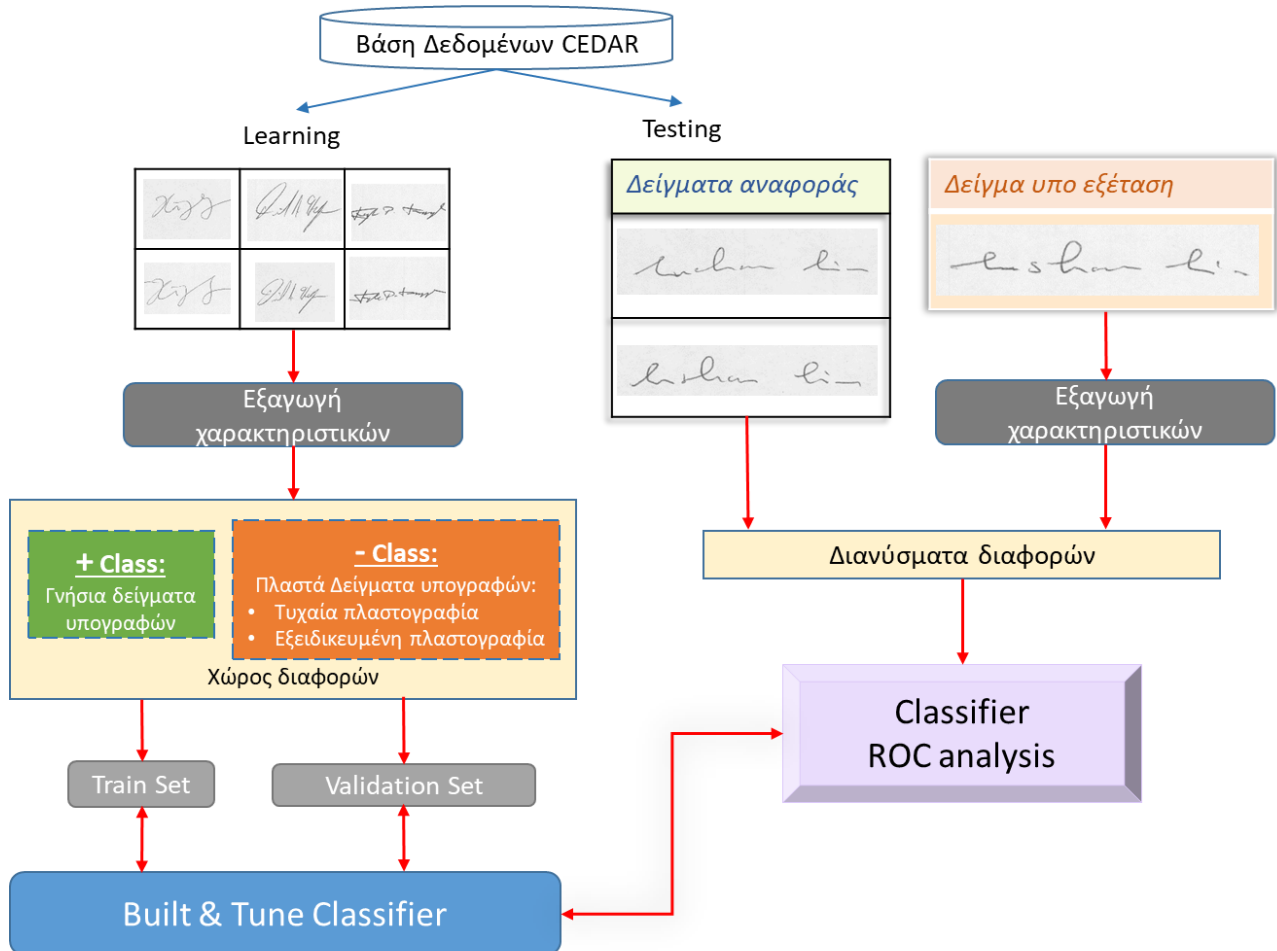
Εικόνα 27. Παράδειγμα ROC καμπύλης Error! Bookmark not defined.

Η διαδικασία που ακολουθήθηκε στο πλαίσιο της παρούσας διπλωματικής εργασίας έχει ως εξής. Αρχικά διαχωρίστηκαν τόσο το πλήθος των γραφών, όσο και το πλήθος των υπογραφών αυτών για την διαδικασία της μάθησης - εκπαίδευσης και αξιολόγησης (training & validation) και του ελέγχου (testing). Για την διαδικασία της εκπαίδευσης και αξιολόγησης επιλέχθηκαν 28 γραφείς (development writers) ενώ οι υπόλοιποι 27 χρησιμοποιήθηκαν για την διαδικασία του ελέγχου testing (exploitation writers). Από τους 28 γραφείς που επιλέχθηκαν για την διαδικασία της εκπαίδευσης και αξιολόγησης, χρησιμοποιήθηκαν για κάθε γραφέα 12 δείγματα υπογραφών για το learning set και 12 δείγματα υπογραφών για το validation set. Στην συνέχεια πραγματοποιήθηκε η αρχικοποίηση των παραμέτρων του αλγορίθμου K-SVD. Η αραιότητα που επιλέχθηκε ήταν ίση με 3, το μέγεθος του patch είναι  $5 \times 5$ . Τέλος όπως έχει γίνει ήδη αναφορά το μέγεθος του λεξικού είναι ίσο με 60.

Πριν την εξαγωγή των patches από τις εικόνες των υπογραφών του learning set, πραγματοποιείται η προεπεξεργασία τους. Αυτό περιλαμβάνει την μετατροπή τους στην κλίμακα του γκρι, τον προσδιορισμό του κατωφλίου με τη μέθοδο του Otsu για τη μετατροπή τους σε μαύρο και λευκό, καθώς και τη λέπτυνση (thinning) του ίχνους της υπογραφής. Μετά την προεπεξεργασία αυτή,

πραγματοποιείται η εξαγωγή των patches. Από την διαδικασία αυτή, ο πίνακας που θα εξαχθεί θα έχει 25 στήλες και γραμμές όσο το πλήθος των ενεργών pixels όλων των εικόνων που δόθηκαν. Ο πίνακας των patches ουσιαστικά αποτελεί τα δεδομένα που χρησιμοποιούνται ως είσοδος στον αλγόριθμο K-SVD. Αντίστοιχα αφού εφαρμοστεί ο K-SVD, θα επιστρέψει ένα υπερπλήρες κοινό λεξικό. Κατόπιν της διαδικασίας αυτής, καλείται η συνάρτηση της OMP όπου σαν παράμετροι της δίνονται τα patches, καθώς και το υπερπλήρες κοινό λεξικό, και αυτό που επιστρέφει είναι ο πίνακας των αραιών συντελεστών. Με την σειρά του ο πίνακας των αραιών συντελεστών δίνεται για την εξαγωγή των χαρακτηριστικών ειδικότερα τα χαρακτηριστικά αυτά είναι η μέση τιμή, η μέγιστη τιμή καθώς και η τυπική απόκλιση ανά atom. Τα χαρακτηριστικά αυτά θα χρησιμοποιηθούν αργότερα ώστε να επιτευχθεί ο στόχος της σύγκρισης των υπογραφών.

Η ίδια διαδικασία ακολουθείται και για τα 12 δείγματα του validation set ανά γραφέα, με την μόνη διαφορά ότι χρησιμοποιείται το ήδη υπάρχον λεξικό και άρα δεν χρειάζεται να εκτελεστεί πάλι ο K-SVD αλγόριθμος για την δημιουργία νέου λεξικού. Εκτελείται πάλι η διαδικασία της προεπεξεργασίας, εφαρμόζεται η OMP για την εύρεση των αραιών συντελεστών για το validation set και τέλος γίνεται πάλι η εξαγωγή των αντίστοιχων χαρακτηριστικών που περιεγράφηκαν από πάνω. Η διαδικασία αυτή έως τώρα αφορούσε μόνο τις γνήσιες υπογραφές των γραφέων, όμως η ίδια διαδικασία εφαρμόζεται αντίστοιχα και για τις πλαστές υπογραφές όλων των development γραφέων. Το τελικό αποτέλεσμα στο στάδιο αυτό είναι η δημιουργία τριών διαφορετικών ομάδων χαρακτηριστικών: μία για το σύνολο εκπαίδευσης, μία για το σύνολο επικύρωσης και μία για τις πλαστές υπογραφές. Ένα σημαντικό σημείο στην εργασία αυτή είναι η δημιουργία ζευγών υπογραφών γνήσιων με γνήσιες και γνήσιες με πλαστές καθώς ο υπολογισμός των διαφορών των χαρακτηριστικών τους με την χρήση του μετασχηματισμού διχοτόμησης (dichotomy transform). Ο μετασχηματισμός διχοτόμησης είναι μια τεχνική που χρησιμοποιείται για να διαχωρίσει χαρακτηριστικά ή δεδομένα σε δύο ομάδες με βάση ένα κατώφλι. Στην περίπτωση αυτή, ο μετασχηματισμός διχοτόμησης εφαρμόζεται σε διαφορετικά ζεύγη υπογραφών, προκειμένου να αναδειχθούν οι διαφορές μεταξύ γνήσιων και πλαστών υπογραφών. Αναλυτικότερα, για το learning set δημιουργούνται 66 διαφορετικά ζεύγη γνήσιων υπογραφών. Αντίστοιχα δημιουργούνται ζεύγη γνήσιων υπογραφών με τυχαίες πλαστογραφίες, δηλαδή γνήσια υπογραφή άλλου γραφέα, ενώ τέλος δημιουργούνται ζεύγη γνήσιας υπογραφής, με εξειδικευμένη πλαστογραφία. Σε όλα τα ζεύγη υπολογίζονται οι διαφορές τους. Στο σημείο αυτό τόσο για το learning set, όσο και για το validation set, καθώς γνωρίζουμε εκ των προτέρων την ταξινόμηση των υπογραφών γίνεται προσθήκη ετικετών (1, για τα γνήσια ζεύγη, -1 για τα πλαστά ζεύγη). Τα δεδομένα των διαφορών, καθώς και οι ετικέτες δίνονται σαν ορίσματα στον ταξινομητή SVM για την δημιουργία ενός υπερεπιπέδου. Τέλος ο ταξινομητής αυτός αξιολογείται με βάση τις καμπύλες ROC και την τιμή του εμβαδού της καμπύλης (Area Under Curve - AUC).

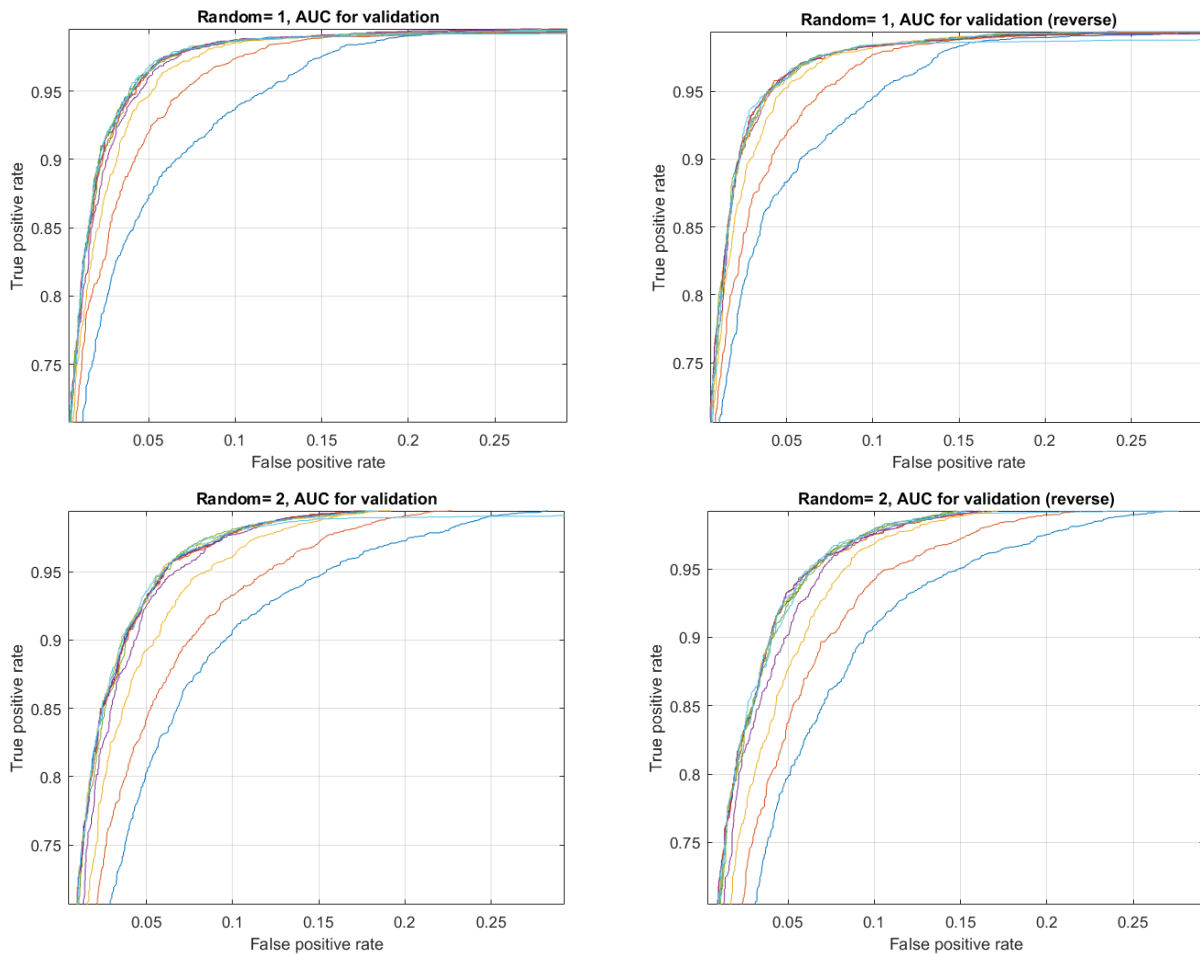


Εικόνα 28. Συνοπτικό διάγραμμα συστήματος αναγνώρισης υπογραφής.



## 5 ΚΕΦΑΛΑΙΟ 5<sup>ο</sup> : Αποτελέσματα

Εδώ παραθέτουμε μερικά από τα παραγόμενα αποτελέσματα των ROC καμπυλών για ένα πλήθος ενδεικτικών γραφέων της βάσης CEDAR (validation set) κατά την χρήση της διαδικασίας αξιολόγησης με τις μηχανές διανυσματικής υποστήριξης:

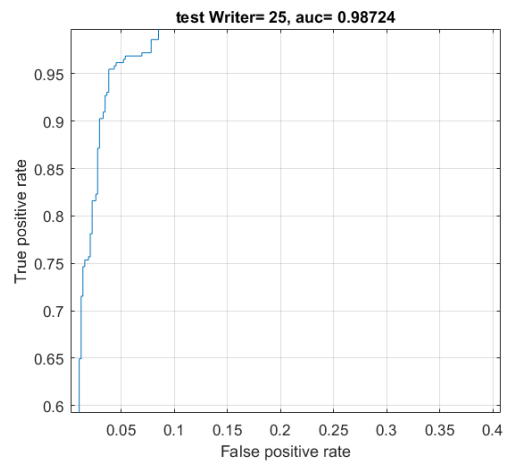
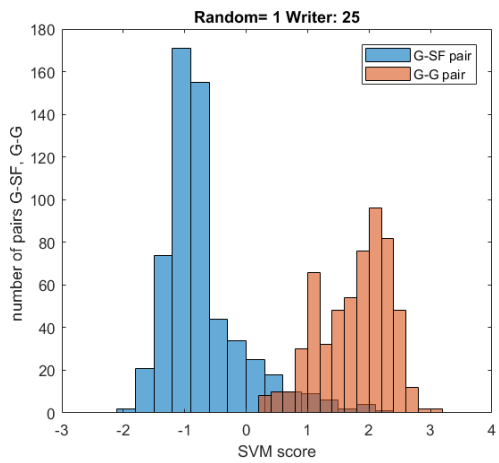
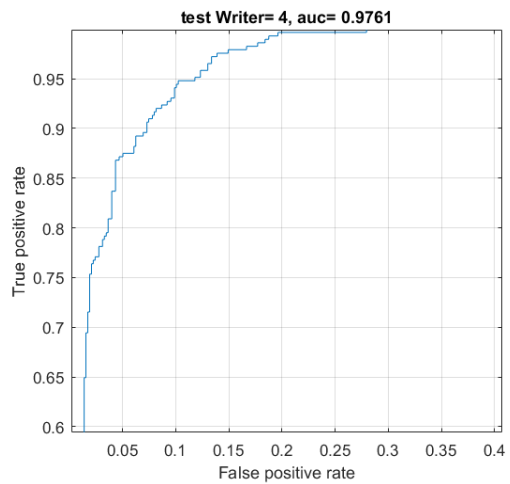
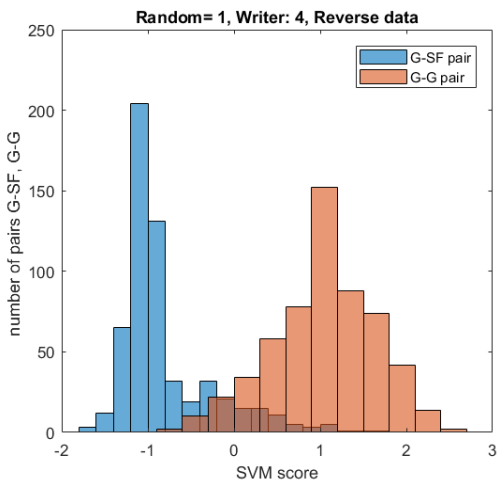
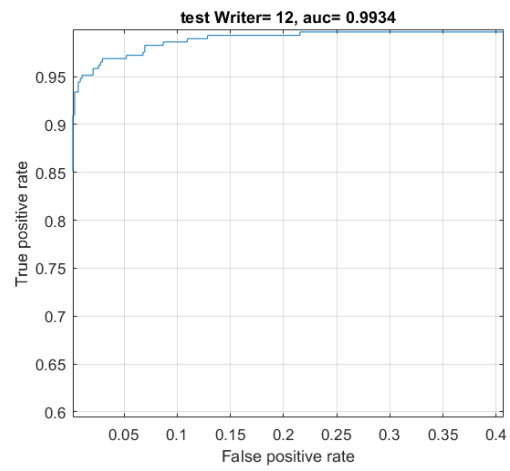
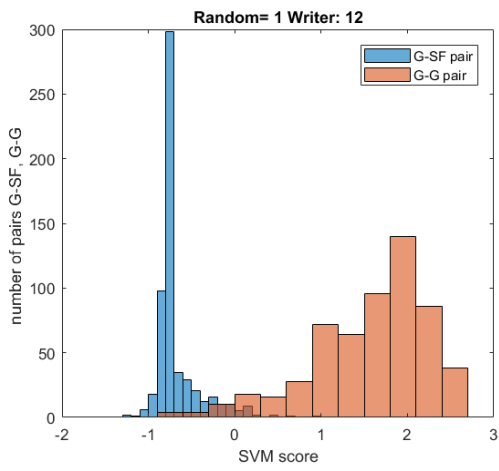


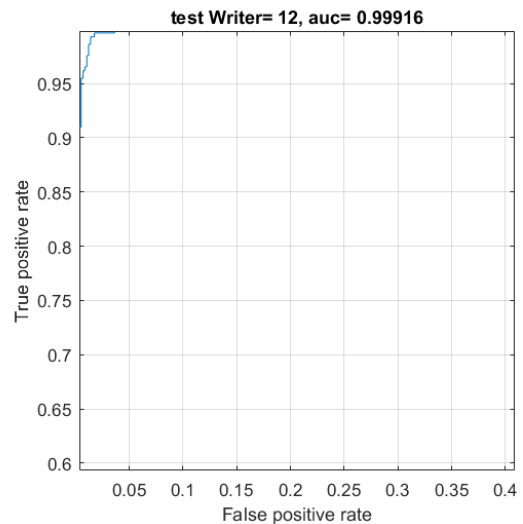
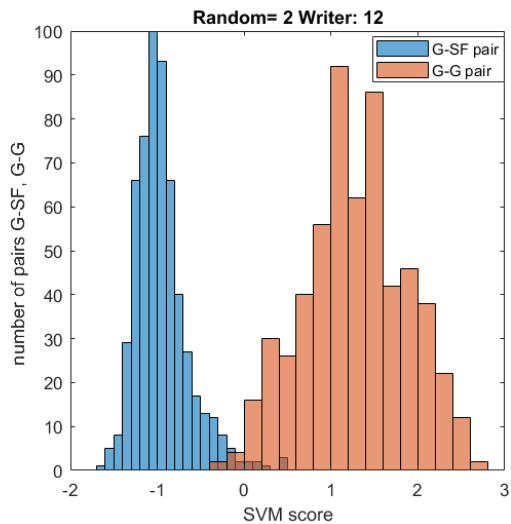
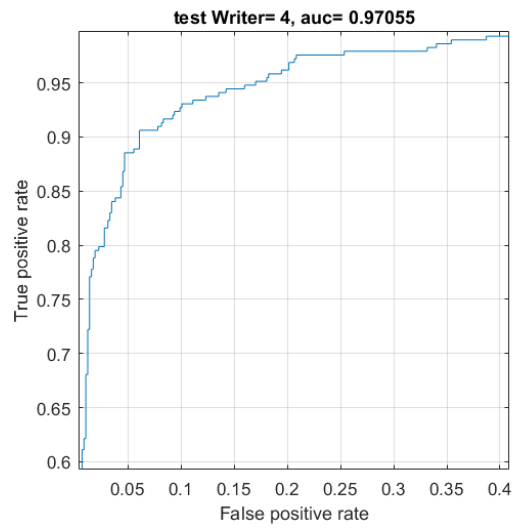
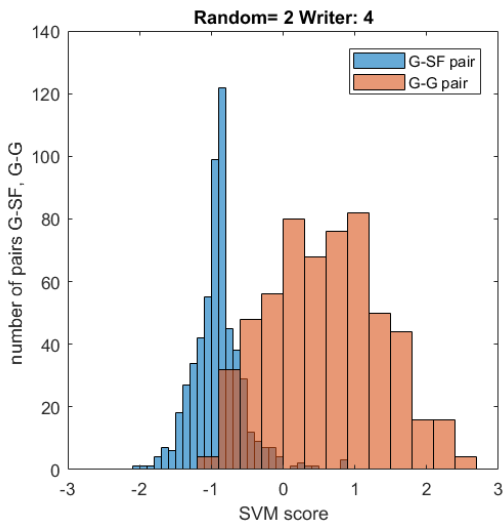
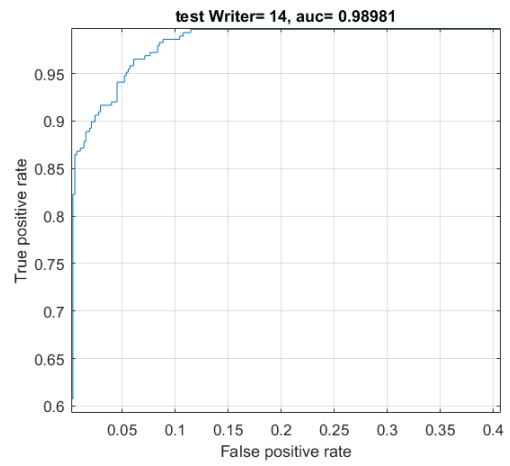
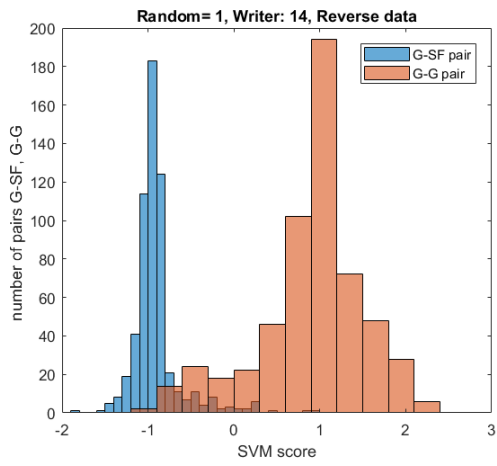
Οι τιμές που έχουν οι καμπύλες ως προς την μετρική του εμβαδού της καμπύλης (Area Under Curve - AUC) παρουσιάζονται συνοπτικά στον **Πίνακας 1**

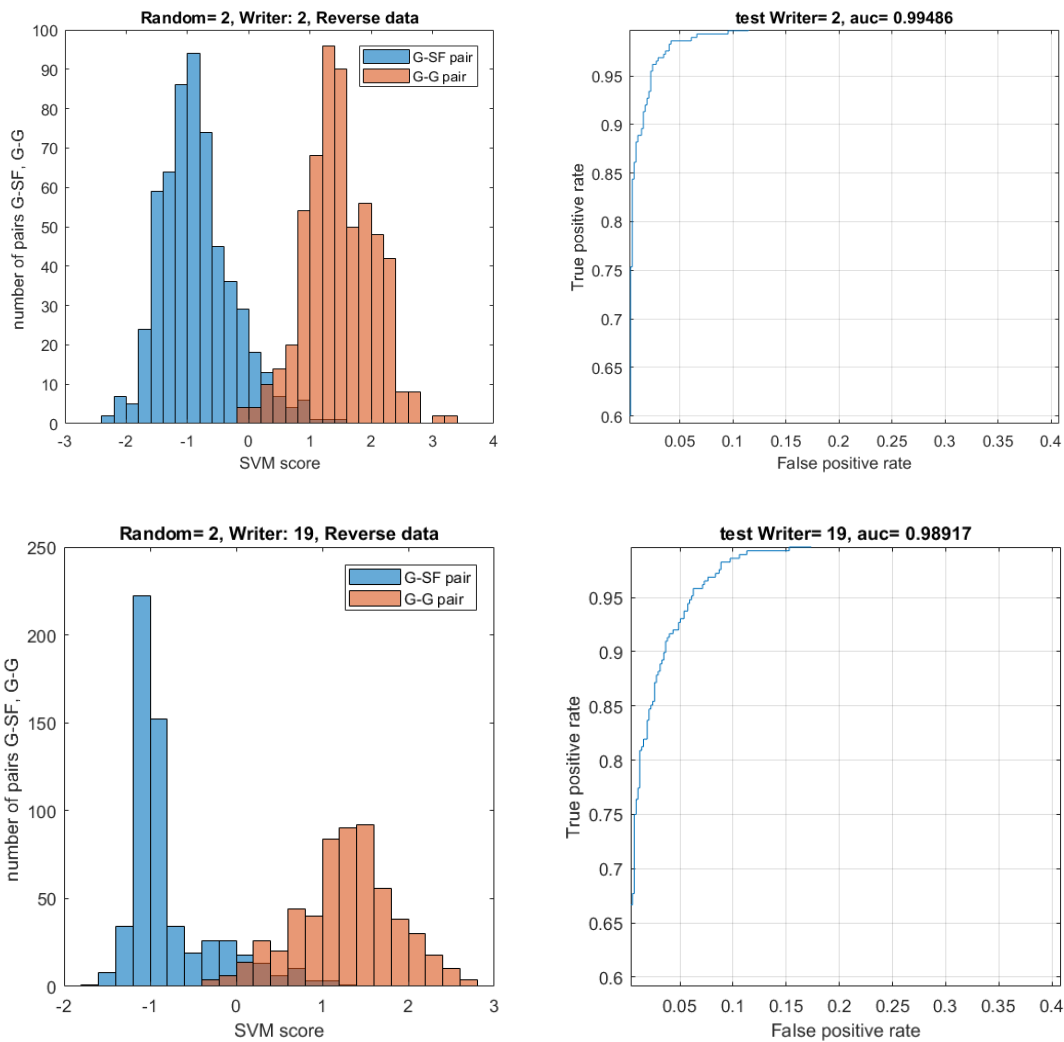
Μέγιστες τιμές AUC (Validation set)			
0.98965	0.9878	0.98631	0.98518

**Πίνακας 1.** Μέγιστες καμπύλες AUC για το testing set, μετά από 4 επαναλήψεις.

Σε ότι αφορά το κομμάτι του testing συνολικά ανά γραφέα εξετάστηκαν 576 διαφορετικά ζεύγη υπογραφών γνήσια και άλλα τόσα πλαστά. Τα πλαστά ζεύγη ανά γραφέα όπως διακρίνονται και στις παρακάτω εικόνες έχουν αρνητικές τιμές, ενώ τα γνήσια ζεύγη δειγμάτων έχουν θετικές τιμές. Τα ενδεικτικά αποτελέσματα για τα δεδομένα των τεστ, δηλαδή τα ιστογράμματα, και οι ROC καμπύλες παρουσιάζονται κάτωθι.







Η μέση τιμή που προέκυψε μετά από όλες τις επαναλήψεις στο σημείο του testing παρουσιάζονται στον **Πίνακα 2**. Να σημειωθεί πως ο γενικός μέσος όρος του εμβαδού της καμπύλης (Area Under Curve - AUC) έπειτα από όλες τις επαναλήψεις έχει την τιμή 0.9611, αυτό σημαίνει πως είναι μια αρκετά καλή τιμή διότι είναι πολύ κοντά στο 1.0.

<b>Μέση τιμή AUC</b>		
	Χωρίς ανεστραμμένα δεδομένα	Με Ανεστραμμένα δεδομένα
<b>1η Επανάληψη</b>	0.9329	0.9329
<b>2η Επανάληψη</b>	0.9903	0.9882
<b>Μέσος Όρος</b>	0.9616	0.9606

**Πίνακας 2.** Συγκεντρωτικός Πίνακας μέσω των τιμών AUC.

## 6 ΚΕΦΑΛΑΙΟ 6<sup>ο</sup> : Συμπεράσματα – Προτάσεις για συνέχισης της έρευνας

Στην παρούσα διπλωματική εργασία, επιχειρήθηκε η αξιολόγηση της αυθεντικότητας των χειρόγραφων δειγμάτων υπογραφών με την χρήση των αλγορίθμων μη επιβλεπόμενης μάθησης και αραιής αναπαράστασης K-SVD/OMP, του ταξινομητή SVM και της βάσης δεδομένων CEDAR. Το προτεινόμενο σύστημα κατατάσσετε στα συστήματα επιβεβαίωσης υπογραφής ανεξαρτήτως του χρήστη. Τα αποτελέσματα που εξήχθησαν ήταν αρκετά ικανοποιητικά, επομένως μια περαιτέρω μελέτη θα μπορούσε να γίνει με δείγματα χειρόγραφων υπογραφών από άλλες βάσεις δεδομένων τόσο on-line όσο και off-line. Μια άλλη πρόταση θα ήταν να μελετήσετε το χειρόγραφο κείμενο αντί για τις υπογραφές. Αυτή η παρατήρηση προέρχεται από το γεγονός ότι πολλά δείγματα υπογραφών δεν ήταν ένα τυχαίο μοτίβο, αλλά ήταν κείμενο με την ταυτότητα (όνομα ή επώνυμο) του υπογράφοντος.

## Βιβλιογραφία –Αναφορές -Διαδικτυακές Πηγές

- [1] Τσέλιος Κωνσταντίνος, ‘Σχεδιασμός και ανάπτυξη συστήματος βιομετρικής αναγνώρισης βασισμένο στα δακτυλικά αποτυπώματα’, Εργαστήριο Ηλεκτρονικής Τομέας Ηλεκτρονικής και Υπολογιστών Τμήμα Φυσικής Πανεπιστήμιο Πατρών. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: <https://nemertes.library.upatras.gr/server/api/core/bitstreams/0e931131-ffdf-4608-a858-1e5e45adbc82/content>
- [2] K. Bibi, S. Naz, και A. Rehman, ‘Biometric signature authentication using machine learning techniques: Current trends, challenges and opportunities’, *Multimed. Tools Appl.*, τ. 79, τχ. 1–2, σσ. 289–340, Ιανουαρίου 2020, doi: 10.1007/s11042-019-08022-0.
- [3] J. Felsenthal, ‘Give Me Your John Hancock When did we start signing our names to authenticate documents?’, SLATE. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: <https://slate.com/news-and-politics/2011/03/when-did-we-start-signing-our-names-to-authenticate-documents.html>
- [4] Wikipedia contributors, ‘John Hancock’, WIKIPEDIA- The Free Encyclopedia. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: [https://en.wikipedia.org/wiki/John\\_Hancock](https://en.wikipedia.org/wiki/John_Hancock)
- [5] M. Diaz, M. A. Ferrer, D. Impedovo, M. I. Malik, G. Pirlo, και R. Plamondon, ‘A Perspective Analysis of Handwritten Signature Technology’, *ACM Comput. Surv.*, τ. 51, τχ. 6, σσ. 1–39, Νοεμβρίου 2019, doi: 10.1145/3274658.
- [6] M. Faundez-Zanuy, J. Fierrez, M. A. Ferrer, M. Diaz, R. Tolosana, και R. Plamondon, ‘Handwriting Biometrics: Applications and Future Trends in e-Security and e-Health’, *Cogn. Comput.*, τ. 12, τχ. 5, σσ. 940–953, Σεπτεμβρίου 2020, doi: 10.1007/s12559-020-09755-z.
- [7] ‘GDPR’. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: <https://research-ethics-committee.uniwa.gr/>
- [8] R. Tolosana, R. Vera-Rodriguez, J. Ortega-Garcia, και J. Fierrez, ‘Preprocessing and Feature Selection for Improved Sensor Interoperability in Online Biometric Signature Verification’, *IEEE Access*, τ. 3, σσ. 478–489, 2015, doi: 10.1109/ACCESS.2015.2431493.
- [9] D. Bertolini, L. S. Oliveira, E. Justino, και R. Sabourin, ‘Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers’, *Pattern Recognit.*, τ. 43, τχ. 1, σσ. 387–396, Ιανουαρίου 2010, doi: 10.1016/j.patcog.2009.05.009.
- [10] E. N. Zois, A. Alexandridis, και G. Economou, ‘Writer independent offline signature verification based on asymmetric pixel relations and unrelated training-testing datasets’, *Expert Syst. Appl.*, τ. 125, σσ. 14–32, Ιουλίου 2019, doi: 10.1016/j.eswa.2019.01.058.
- [11] E. N. Zois, D. Tsourounis, I. Theodorakopoulos, A. L. Kesidis, και G. Economou, ‘A Comprehensive Study of Sparse Representation Techniques for Offline Signature Verification’, *IEEE Trans. Biom. Behav. Identity Sci.*, τ. 1, τχ. 1, σσ. 68–81, Ιανουαρίου 2019, doi: 10.1109/TBIOM.2019.2897802.
- [12] G. S. Eskander, R. Sabourin, και E. Granger, ‘Hybrid writer-independent–writer-dependent offline signature verification system’, *IET Biom.*, τ. 2, τχ. 4, σσ. 169–181, Δεκεμβρίου 2013, doi: 10.1049/iet-bmt.2013.0024.
- [13] L. G. Hafemann, R. Sabourin, και L. S. Oliveira, ‘Offline Handwritten Signature Verification - Literature Review’, στο *2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Νοεμβρίου 2017, σσ. 1–8. doi: 10.1109/IPTA.2017.8310112.
- [14] M. I. Malik και M. Liwicki, ‘From Terminology to Evaluation: Performance Assessment of Automatic Signature Verification Systems’, στο *2012 International Conference on Frontiers in Handwriting Recognition*, Bari, Italy: IEEE, Σεπτεμβρίου 2012, σσ. 613–618. doi: 10.1109/ICFHR.2012.205.
- [15] L. Alewijnse, ‘Forensic signature examination’, 2013.
- [16] Η. Ζώης, ‘Διαφάνειες διαλέξεων του μαθήματος Ψηφιακής Επεξεργασίας Εικόνας και Αναγνώρισης Προτύπων’, Τμήμα Ηλεκτρολόγων και Ηλεκτρονικών Μηχανικών, Πανεπιστήμιο Δυτικής Αττικής, 2018.
- [17] R. C. Gonzalez, R. E. Woods, και B. R. Masters, ‘Digital Image Processing, Third Edition’, *J. Biomed. Opt.*, τ. 14, τχ. 2, σ. 029901, 2009, doi: 10.1117/1.3115362.

- [18] N. Otsu, ‘A Threshold Selection Method from Gray-Level Histograms’, *IEEE Trans. Syst. Man Cybern.*, τ. Volume: 9, τχ. Issue: 1, σσ. 62–66, Ιανουαρίου 1979, doi: 10.1109/TSMC.1979.4310076.
- [19] R. C. Gonzalez, R. E. Woods, και Α. Μάργαρης, *ΨΗΦΙΑΚΗ ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ*, 4η Έκδοση. ΕΚΔΟΣΕΙΣ ΤΖΙΟΛΑ, 2020.
- [20] Z. Zhang, Y. Xu, J. Yang, X. Li, και D. Zhang, ‘A Survey of Sparse Representation: Algorithms and Applications’, *IEEE Access*, τ. 3, σσ. 490–530, 2015, doi: 10.1109/ACCESS.2015.2430359.
- [21] J. M. Giron-Sierra, *Digital Signal Processing with Matlab Examples, Volume 3*. στο Signals and Communication Technology. Singapore: Springer Singapore, 2017. doi: 10.1007/978-981-10-2540-2.
- [22] E. N. Zois, D. Tsourounis, I. Theodorakopoulos, A. L. Kesidis, και G. Economou, ‘A Comprehensive Study of Sparse Representation Techniques for Offline Signature Verification’, *IEEE Trans. Biom. Behav. Identity Sci.*, τ. 1, τχ. 1, σσ. 68–81, Ιανουαρίου 2019, doi: 10.1109/TBIOM.2019.2897802.
- [23] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York, NY: Springer New York, 2010. doi: 10.1007/978-1-4419-7011-4.
- [24] Δ. Τσουρούνης, ‘Βαθιά Αραιή Κωδικοποίηση (Deep Sparse Coding)’, Μεταπτυχιακή Διπλωματική Εργασία, Πανεπιστήμιο Πατρών, Πάτρα, 2017.
- [25] J. Mairal, ‘Sparse Modeling for Image and Vision Processing’, *Found. Trends® Comput. Graph. Vis.*, τ. 8, τχ. 2–3, σσ. 85–283, 2014, doi: 10.1561/06000000058.
- [26] M. Aharon, M. Elad, και A. Bruckstein, ‘ $\ell_1$ -SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation’, *IEEE Trans. Signal Process.*, τ. 54, τχ. 11, σσ. 4311–4322, Νοεμβρίου 2006, doi: 10.1109/TSP.2006.881199.
- [27] B. Dumitrescu και P. Irofti, *Dictionary Learning Algorithms and Applications*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-78674-2.
- [28] R. Rubinstein, M. Zibulevsky, και M. Elad, ‘Efficient Implementation of the K-SVD Algorithm using Batch Orthogonal Matching Pursuit’, 2008, [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: <https://www.researchgate.net/publication/251229200>
- [29] C.-W. Hsu, C.-C. Chang, και C.-J. Lin, ‘A Practical Guide to Support Vector Classification’.
- [30] C.-C. Chang και C.-J. Lin, ‘LIBSVM: A library for support vector machines’, *ACM Trans. Intell. Syst. Technol.*, τ. 2, τχ. 3, σσ. 1–27, Απριλίου 2011, doi: 10.1145/1961189.1961199.
- [31] C. M. Bishop, *Pattern recognition and machine learning*. στο Information science and statistics. New York: Springer, 2006.
- [32] D. Decoste, ‘Training Invariant Support Vector Machines’.
- [33] S. Haykin, *Νευρωνικά Δίκτυα και Μηχανική Μάθηση*, Τρίτη Έκδοση. Αθήνα: Παπασωτηρίου, 2010.
- [34] Wikipedia contributors, ‘Support Vector Machine’, WIKIPEDIA- The Free Encyclopedia. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: [https://en.wikipedia.org/wiki/Support\\_vector\\_machine](https://en.wikipedia.org/wiki/Support_vector_machine)
- [35] ‘The RBF kernel in SVM: A Complete Guide’, PyCodeMates. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: <https://www.pycodemates.com/2022/10/the-rbf-kernel-in-svm-complete-guide.html>
- [36] Wikipedia contributors, ‘Kernel method’, WIKIPEDIA- The Free Encyclopedia. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: [https://en.wikipedia.org/wiki/Kernel\\_method](https://en.wikipedia.org/wiki/Kernel_method)
- [37] M. M. Hameed, R. Ahmad, M. L. M. Kiah, και G. Murtaza, ‘Machine learning-based offline signature verification systems: A systematic review’, *Signal Process. Image Commun.*, τ. 93, σσ. 116139, Απριλίου 2021, doi: 10.1016/j.image.2021.116139.
- [38] ‘Classification: ROC Curve and AUC’, Google for Developers. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο: <https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc>

- [39] Wikipedia contributors, 'Receiver operating characteristic', WIKIPEDIA- The Free Encyclopedia. [Έκδοση σε ψηφιακή μορφή]. Διαθέσιμο στο:  
[https://en.wikipedia.org/wiki/Receiver\\_operating\\_characteristic](https://en.wikipedia.org/wiki/Receiver_operating_characteristic)