

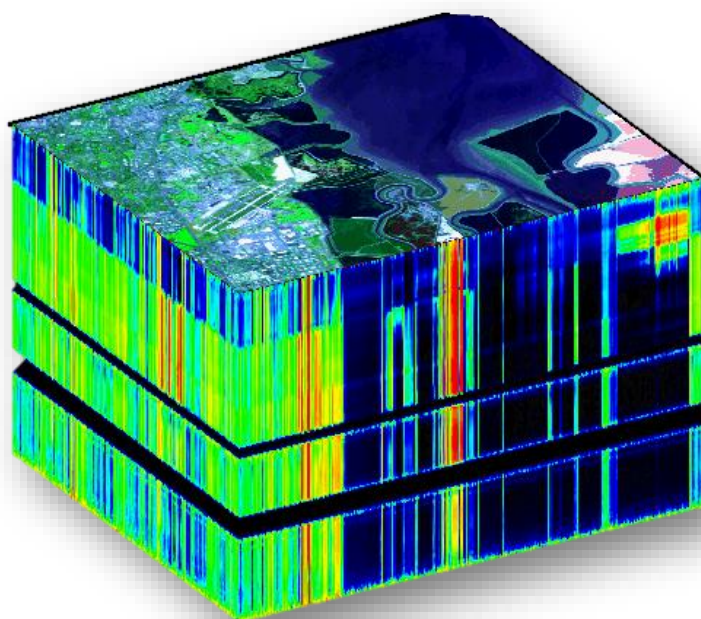


ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

Διπλωματική Εργασία

**Τεχνικές βαθιάς μηχανικής μάθησης και Συνελκτικά Νευρωνικά
Δίκτυα για την ταξινόμηση υπερφασματικών δεδομένων**

Deep learning and Convolutional Neural Networks for the classification
of hyperspectral data



Κοντόπουλος Χρήστος

Αθήνα, Οκτώβριος 2015



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

Διπλωματική Εργασία

**Τεχνικές βαθιάς μηχανικής μάθησης και Συνελικτικά Νευρωνικά
Δίκτυα για την ταξινόμηση υπερφασματικών δεδομένων**

Εγκρίθηκε από την τριμελή
εξεταστική επιτροπή την
..... 2015

.....
.....
Κ. Καράντζαλος
Επ. Καθηγητής Ε.Μ.Π

.....
.....
Δ. Αργιαλάς
Καθηγητής Ε.Μ.Π

.....
.....
Α. Δουλάμης
Επ. Καθηγητής Ε.Μ.Π

.....
.....
Χρήστος Κοντόπουλος
Διπλωματούχος Αγρονόμος και
Τοπογράφος Μηχανικός Ε.Μ.Π.

Copyright © Χρήστος Β. Κοντόπουλος, 2015
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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 εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς το συγγραφέα.

Ευχαριστίες

Αρχικά θέλω να ευχαριστήσω τον επιβλέποντα καθηγητή της διπλωματικής μου εργασίας Κωνσταντίνο Καράντζαλο. Τον ευχαριστώ ειλικρινά για όλες τις πολύτιμες συμβουλές, και την καθοδήγηση που μου έδωσε όλο αυτό το διάστημα καθώς επίσης και για τις καθοριστικές διορθώσεις και επισημάνσεις που μου υπέδειξε.

Επίσης θέλω να ευχαριστήσω βαθύτατα τον υποψήφιο διδάκτορα του Πολυτεχνείου Κρήτης, Κωνσταντίνο Μακαντάση. Οι γνώσεις του πάνω στο αντικείμενο της εργασίας αλλά και η συνεχής ανιδιοτελής βοήθειά του, αποτέλεσαν ουσιαστικό στοιχείο για την ολοκλήρωση της εργασίας.

Με την ολοκλήρωση της παρούσας εργασίας κλείνει και ο κύκλος των προπτυχιακών μου σπουδών στη Σχολή Αγρονόμων και Τοπογράφων του ΕΜΠ. Σε όλη αυτή την περίοδο αποκόμισα γνώσεις και εμπειρίες - ακαδημαϊκές και μη - οι οποίες θα με χαρακτηρίζουν ως άτομο από εδώ και στο εξής. Για αυτές τις εμπειρίες θα ήθελα να ευχαριστήσω από καρδιάς όλους ανεξαιρέτως τους φίλους και συμφοιτητές μου.

Τέλος θα ήθελα να ευχαριστήσω την οικογένειά μου, τον πατέρα μου Βασίλη, την μητέρα μου Ευδοκία και τον αδερφό μου Γιώργο. Τους ευχαριστώ για όλη την υλική και πνευματική υποστήριξη που μου προσφέρουν όλα αυτά τα χρόνια, δίνοντας μου μεταξύ άλλων και την δυνατότητα να σπουδάσω αυτό που ήθελα.

Αθήνα, Οκτώβριος 2015
Κοντόπουλος Χρήστος

ΠΕΡΙΛΗΨΗ

Η ανάπτυξη και υλοποίηση αυτοματοποιημένων αλγορίθμων μηχανικής μάθησης για την ταξινόμηση εικόνων και αναγνώριση αντικειμένων απασχολεί έντονα την επιστημονική κοινότητα για αρκετές δεκαετίες. Την τελευταία πενταετία προσεγγίσεις που βασίζονται σε αρχιτεκτονικές βαθιάς εκμάθησης (deep learning) έχουν ξεπεράσει σημαντικά τις ακρίβειες ανίχνευσης και ταξινόμησης τεχνικών με πιο απλές/ρηχές αρχιτεκτονικές. Αν και οι σχετικές εργασίες για υπερφασματικά δεδομένα είναι λίγες σε αριθμό, τα αποτελέσματα είναι και σε αυτά τα πολυδιάστατα δεδομένα ελπιδοφόρα.

Προς την κατεύθυνση αυτή, η παρούσα διπλωματική εργασία καταπιάστηκε με την διερεύνηση, μελέτη και αξιολόγηση προηγμένων αλγορίθμων επιβλεπόμενης ταξινόμησης υπερφασματικών δεδομένων. Πιο συγκεκριμένα, μελετήθηκαν και αξιολογήθηκαν τρεις τεχνικές ταξινόμησης: (i) ένα Συνελικτικό Νευρωνικό Δίκτυο βαθιάς αρχιτεκτονικής, (ii) μία μη-γραμμική Μηχανή Διανυσματικής Υποστήριξης τύπου RBF-kernel και (iii) μία γραμμική Μηχανή Διανυσματικής Υποστήριξης.

Τα πολλαπλά πειράματα για την αξιολόγηση των ταξινομητών εφαρμόστηκαν σε δύο κατηγορίες υπερφασματικών δεδομένων: (α) δεδομένα τα οποία αποτελούν αντικείμενο μελέτης και σύγκρισης στη διεθνή βιβλιογραφία και διατίθενται ελεύθερα στο διαδίκτυο και (β) δεδομένα τα οποία δημιουργήθηκαν αποκλειστικά για τους σκοπούς της εργασίας και τον εμπλουτισμό της βιβλιογραφίας, με νέα πιο σύνθετα δεδομένα αξιολόγησης σε εκτεταμένες γεωγραφικά περιοχές και αριθμό κλάσεων (>90).

Σε όλες τις περιπτώσεις τα συνελικτικά νευρωνικά δίκτυα με βαθιά μάθηση έδωσαν αποτελέσματα με τις μεγαλύτερες ακρίβειες ταξινόμησης. Τα αποτελέσματα αυτά συμφωνούν με τη βιβλιογραφία, ενώ τα αποτελέσματα στα νέα δεδομένα ελέγχου υποδεικνύουν ότι ακόμα και με 98 κλάσεις και ποσοστό δειγμάτων μάθησης στο 1%, οι ακρίβειες μπορεί να φτάσουν πάνω από 89%. Παράλληλα, η συγκριτική ανάλυση των τεχνικών ανέδειξαν τις δυνατότητες των τριών ταξινομητών τόσο σε επίπεδο ακρίβειας όσο και υπολογιστικής πολυπλοκότητας κατά την διαδικασία της εκπαίδευσής τους. Τέλος, τα συμπεράσματα και οι προτάσεις που παρατίθενται μπορούν να αποτελέσουν βοηθητικό υλικό για μελλοντικές ερευνητικές εργασίες.

ABSTRACT

The research and development of automated machine learning algorithms for image classification and object recognition, has gained significant attention from the scientific community for several decades. During the last five years deep learning based approaches have surpassed the classification accuracy of simpler/shallow architectures. Although deep learning studies on hyperspectral data are much fewer in number, results are also promising.

Towards this direction, this diploma thesis dealt with the investigation, study and evaluation of advanced supervised classification algorithms of hyperspectral data. More specifically, three classification techniques were studied and evaluated: (i) a Deep Convolution Neural Network, (ii) a non-linear Support Vector Machine, RBF-kernel type and (iii) a linear Support Vector Machine.

Several experiments were conducted for the evaluation of each classifiers based on: (a) freely available online datasets which are the subject of study and comparison to international literature and (b) data and ground truth which were created exclusively for the purposes of this work towards literature enrichment with new more complex geographical areas as well as various (>90) terrain or spectral classes.

In all cases the deep convolutional neural networks resulted to significantly lower classification error. These experimental results are consistent with the literature, while results in the new reference data indicated that even with 98 classes and training sample percentage to 1%, precision can reach a 89%. Moreover, the comparative analysis of the techniques highlights the potentialities of the three classifiers in terms of both accuracy but also computational cost during their training phase. Finally, the overall analysis and evaluation can provide valuable material for future research.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1	ΕΙΣΑΓΩΓΗ.....	10
1.1	Επιβλεπόμενη Ταξινόμηση υπερφασματικών δεδομένων.....	10
1.2	Κίνητρο και Συνεισφορά εργασίας.....	12
1.3	Δομή Εργασίας.....	13
2	ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ.....	14
2.1	Εργασίες Deep Learning και CNN.....	15
2.2	Πρόσφατες Σχετικές Εργασίες.....	19
3	ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	22
3.1	Τεχνητή Νοημοσύνη και Μηχανική Μάθηση.....	23
3.2	Νευρωνικά Δίκτυα.....	25
3.2.1	Βιολογικά Νευρωνικά Δίκτυα.....	26
3.2.2	Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ).....	28
3.3	Βασικές Αρχές Τεχνητών Νευρωνικών Δικτύων.....	29
3.3.2	Αρχιτεκτονική ΤΝΔ.....	29
3.3.2	Μοντέλο Αισθητήρα (Perceptron).....	31
3.3.3	Συνάρτηση ενεργοποίησης.....	34
3.3.4	Διαδικασία εκμάθησης (training phase).....	35
3.3.5	Αλγόριθμος Οπισθοδιάδοσης Σφάλματος (<i>backpropagation error</i>).....	37
3.4	Συνελικτικά Νευρωνικά Δίκτυα (CNN).....	43
3.5	Μηχανές Διανυσματικής Υποστήριξης (SVM).....	46
3.6	Αλγόριθμοι εφαρμογής.....	48
3.6.1	Μείωσης διάστασης δεδομένων (<i>Dimensionality Reduction</i>).....	48
3.6.1	Το Συνελικτικό Νευρωνικό Δίκτυο (CNN) εφαρμογής.....	49
3.6.2	Οι αλγόριθμοι SVM εφαρμογής.....	50
4	ΔΕΔΟΜΕΝΑ ΕΛΕΓΧΟΥ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΑ.....	52
4.2	Βιβλιογραφικά Δεδομένα.....	53
4.2	Δεδομένα SFB (San Francisco Bay) Και Βήματα Επεξεργασίας.....	67
4.2.1	Το έργο AVIRIS της NASA.....	67
4.2.2	Απόκτηση δεδομένων AVIRIS.....	70
4.2.3	Περιοχή ενδιαφέροντος και επιλεγμένες εικόνες AVIRIS.....	73
4.2.4	Στάδια προεπεξεργασίας δεδομένων.....	76
4.2.5	Ιεραρχικό Μοντέλο Κλάσεων και τελικά δεδομένα SFB.....	89

5	ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ.....	98
5.1	Δομή Κεφαλαίου	100
5.2	Παραμετροποίηση Αλγορίθμων.....	101
5.2.1	Principal Component Analysis (PCA)	101
5.2.2	Ποσοστό Ground Truth ως δεδομένα εκπαίδευσης	102
5.2.3	Αριθμός κλάσεων	102
5.2.4	Καθορισμός μέτρησης χρονικής απόδοσης ταξινομητή	102
5.2.5	Τεχνικά Χαρακτηριστικά Υπολογιστή Εργασίας.....	103
5.3	Αποτελέσματα Βιβλιογραφικών Δεδομένων	104
5.4	Αποτελέσματα Δεδομένων SFB (San Francisco Bay).....	132
5.4.1	Μονοεικονικές ταξινομήσεις	133
5.4.2	Σύνθετες πολυεικονικές ταξινομήσεις	142
6	ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ.....	148
6.1	Γενικότερα Συμπεράσματα	148
6.2	Ειδικά Συμπεράσματα	150
6.3	Προτάσεις.....	153
	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	154
	Βιβλία και Εργασίες.....	154
	Διαδικτυακές πηγές.....	156

1 ΕΙΣΑΓΩΓΗ

Στο εισαγωγικό αυτό κεφάλαιο αρχικά γίνεται μία περιγραφή του προβλήματος με το οποίο καταπιάστηκε η εργασία όπως επίσης και οι αντικειμενικοί στόχοι που τέθηκαν από αυτό. Ακολούθως, περιγράφονται το κίνητρο και η συνεισφορά της παρούσας εργασίας ενώ στο τέλος αναλύεται ο τρόπος και η λογική με την οποία αυτή έχει δομηθεί.

1.1 Επιβλεπόμενη Ταξινόμηση υπερφασματικών δεδομένων

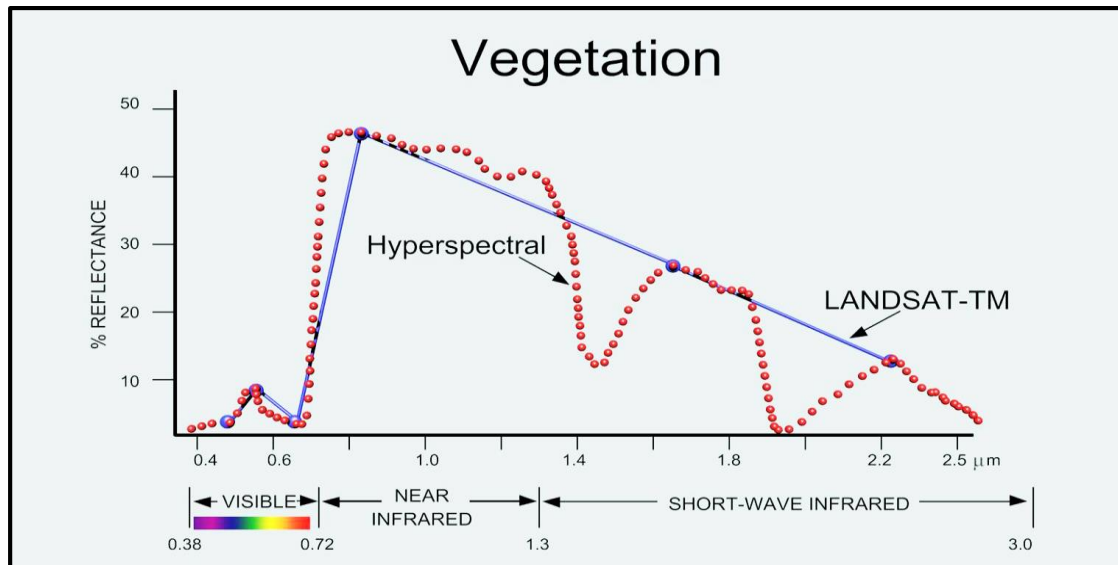
Η ταξινόμηση εικόνας είναι μία από τις πιο συνήθεις εφαρμογές στο αντικείμενο της ψηφιακής τηλεπισκόπησης και γενικότερα στον κλάδο της όρασης υπολογιστών. Σε γενικές γραμμές, ο σκοπός της κάθε ταξινόμησης εικόνας είναι η αντικατάσταση της φωτοερμηνείας των εικόνων, μέσα από αυτοματοποιημένες τεχνικές σε ένα υπολογιστικό σύστημα, για την αναγνώριση διάφορων χαρακτηριστικών (θεματικών κατηγοριών, αντικειμένων κ.α.). Πρακτικά, για την ταξινόμηση μιας ψηφιακής εικόνας, ένα υπολογιστικό σύστημα επεξεργάζεται κάθε εικονοστοιχείο της ξεχωριστά και με διάφορους ‘κανόνες απόφασης’ επιλέγεται σε ποια υποψήφια κλάση αντιστοιχεί. Βεβαίως, ο ακριβής τρόπος με τον οποίο κάθε εικονοστοιχείο ταξινομείται εξαρτάται από τον αλγόριθμο με τον οποίο δρα, ο οποίος μπορεί να βασίζεται από μια απλή στατιστική εντολή πιθανοφάνειας έως από πολυσύνθετες επαναληπτικές διαδικασίες εκμάθησης.

Σημειώνεται πως μια διαδικασία ταξινόμησης ψηφιακής εικόνας μπορεί να είναι κατά κανόνα είτε επιβλεπόμενη είτε μη επιβλεπόμενη. Στην πρώτη περίπτωση ο αλγόριθμος ταξινόμησης τροφοδοτείται με κάποια ‘ορθά’ δείγματα της εικόνας (τα οποία και ανήκουν σε γνωστή κλάση) με τα οποία ‘εκπαιδεύεται στο να ταξινομεί σωστά’. Φυσικά, αυτή η εκπαιδευτική διαδικασία μεταφράζεται ως μία επαναληπτική αναπροσαρμογή των συντελεστών και των βαρών του εκάστοτε ταξινομητή. Στην περίπτωση της μη επιβλεπόμενης ταξινόμησης πάλι, η αναπροσαρμογή αυτή γίνεται αυτόματα καθώς ο εκάστοτε ταξινομητής καλείται να αναγνωρίσει τα διάφορα χαρακτηριστικά της εικόνας χωρίς να έχει τροφοδοτηθεί με δεδομένα εισόδου. Η παρούσα εργασία καταπιάνεται με την επιβλεπόμενη ταξινόμηση υπερφασματικών δεδομένων – εικόνων.

1.Εισαγωγή

Η ύπαρξη πολυφασματικών και υπερφασματικών δεκτών, δίνει την δυνατότητα στον άνθρωπο να παρατηρεί και να επεξεργάζεται δεδομένα και φαινόμενα που επαφίονται σε μη ορατά μήκη του ηλεκτρομαγνητικού φάσματος. Με αυτόν τον τρόπο σε μία ταξινόμηση πολυφασματικής (*multispectral*) ή υπερφασματικής (*hyperspectral*) εικόνας, δίνεται η δυνατότητα στον μελετητή να συμπεριλάβει πληροφορία η οποία δεν είναι ορατή και κατά συνέπεια κατανοητή από το ανθρώπινο μάτι. Η κατανομή αυτής της πληροφορίας στους πολυφασματικούς και υπερφασματικούς δέκτες γίνεται με τη βοήθεια των καναλιών (*bands*).

Η βασική διαφοροποίηση ανάμεσα στους πολυφασματικούς και τους υπερφασματικούς δέκτες είναι ο αριθμός των καναλιών που περιλαμβάνουν και ο τρόπος με τον οποίο κατανέμεται η πληροφορία σε αυτά. Ειδικότερα ένας πολυφασματικός δέκτης έχει συνήθως μικρό αριθμό καναλιών (π.χ. έως 10) με το κάθε κανάλι να αποτυπώνει το αντανακλώμενο σήμα της ηλεκτρομαγνητικής πληροφορίας για ορισμένα όρια του ηλεκτρομαγνητικού φάσματος. Αντιθέτως, ένας υπερφασματικός δέκτης αποτελείται από ένα μεγάλο αριθμό καναλιών (π.χ. 150 και περισσότερα) στα οποία η καταγραφή της υπερφασματικής πληροφορίας γίνεται με συνεχή τρόπο, δηλαδή με σταθερό βήμα από κανάλι σε κανάλι για όλα τα διαθέσιμα μήκη κύματος. Παρακάτω (*Εικόνα 1.1*) παρουσιάζεται η φασματική υπογραφή ενός εικονοστοιχείου βλάστησης για έναν πολυφασματικό δέκτη (Landsat - TM) και για έναν υπερφασματικό.



Εικόνα 1.1: Η φασματική υπογραφή ενός εικονοστοιχείου βλάστησης, για έναν πολυφασματικό δέκτη (μπλε) και για έναν υπερφασματικό (κόκκινο)

[Πηγή: <http://blog.neosgeo.com>]

Με βάση τα παραπάνω μπορεί να γίνει εύκολα αντιληπτός ο τρόπος με τον οποίο εισέρχονται τα δεδομένα εκπαίδευσης στον κάθε ταξινομητή και άρα ο τρόπος με τον οποίο επιτελείται η διαδικασία εκμάθησης. Μία υπερφασματική εικόνα περιέχει πληροφορία τριών διαστάσεων (γραμμές, στήλες και κανάλια). Έτσι στη βασική περίπτωση ταξινόμησης, ένας αλγόριθμος εκπαιδεύεται σε όλο το βάθος της υπερφασματικής πληροφορίας και τα διάφορα εικονοστοιχεία – δείγματα κατηγοριοποιούνται βάσει αυτής. Για αυτό το λόγο ένα μεγάλο πλεονέκτημα των υπερφασματικών δεδομένων θεωρείται η περίσσεια συντελεστών και παραμέτρων που προσφέρει ο μεγάλος αριθμός καναλιών.

Τέλος, αξίζει να επισημανθεί πως το σύνολο των γνωστών και ορθά ταξινομημένων δειγμάτων που χρησιμοποιούνται ως δεδομένα εκπαίδευσης, ονομάζονται *δεδομένα Ελέγχου - Ground Truth*. Παρακάτω, τα δεδομένα Ground Truth που αναφέρονται στην εργασία είτε αποκτήθηκαν μαζί με την αντίστοιχη υπερφασματική εικόνα, είτε δημιουργήθηκαν (βλ. *Κεφάλαιο 4.2*) για να γίνει εφικτή η εφαρμογή συνθετότερων ταξινομήσεων.

Στόχος εργασίας

Στην παρούσα εργασία εφαρμόζονται τρεις προσεγγίσεις αιχμής (βλ. *Κεφάλαιο 3.6*) για την ταξινόμηση υπερφασματικών δεδομένων. Ο κυρίαρχος στόχος που αντικειμενικά τίθεται είναι η εφαρμογή των τριών αυτών προσεγγίσεων πάνω σε μία μεγάλη ποικιλία υπερφασματικών δεδομένων, με απώτερο σκοπό την ανάλυση και σύγκριση των εξαγόμενων αποτελεσμάτων.

Παράλληλα από την ανάλυση των αποτελεσμάτων βασική επιδίωξη ήταν η πλήρης περιγραφή των δυνατοτήτων και αποδόσεων των ταξινομητών, σε επίπεδο ακρίβειας και χρονικής απόδοσης. Επιπλέον, το ενδιαφέρον εστιάστηκε και στην μελέτη και ανάλυση των αποτελεσμάτων του δικτύου CNN σε σύγκριση με αυτά του ταξινομητή RBF-SVM που μέχρι και σήμερα θεωρείται μέτρο σύγκρισης στην διεθνή βιβλιογραφία για την απόδοση των αλγορίθμων. Τέλος επιδιώχθηκε και ο εμπλουτισμός της βιβλιογραφίας με νέα υπερφασματικά δεδομένα ελέγχου.

1.2 Κίνητρο και Συνεισφορά εργασίας

Βασικό κίνητρο για την μελέτη του συγκεκριμένου θέματος αποτέλεσε η σημαντική για το Τοπογράφο Μηχανικό δυνατότητα για την με ακρίβεια ανάλυση και ταξινόμηση υπερφασματικών δεδομένων σε ποικίλες εφαρμογές. Επίσης, η ραγδαία εξέλιξη των μεθόδων μηχανικής μάθησης των τελευταίων χρόνων δίνει πρόσφορο έδαφος στην επιστημονική κοινότητα για την χρήση πιο εξειδικευμένων και αποτελεσματικών εργαλείων για την ανίχνευση και εντοπισμό αντικειμένων και χαρακτηριστικών στην γήινη επιφάνεια.

Τα αποτελέσματα της εργασίας συνεισφέρουν στην αξιολόγηση προηγμένων τεχνικών μηχανικής μάθησης για την ταξινόμηση υπερφασματικών δεδομένων από άποψη ακρίβειας, ποσοστού δειγμάτων μάθησης και αλγοριθμικής απόδοσης. Επιπλέον στο πλαίσιο της διπλωματικής δημιουργήθηκαν νέα (ανοιχτά στην επιστημονική κοινότητα) δεδομένα ελέγχου που σε αντίθεση με τα έως τώρα σετ δεδομένων αναφέρονται σε μια εκτεταμένη γεωγραφικά περιοχή και έχουν έως και 98 κλάσεις. Τέλος, η παρούσα διπλωματική προτείνει συγκεκριμένη μεθοδολογία για την με ακρίβεια ταξινόμηση υπερφασματικών δεδομένων παρουσιάζοντας πολλαπλά αποτελέσματα και συγκριτικές αναλύσεις για την ευαισθησία διάφορων παραμέτρων που επηρεάζουν το αποτέλεσμα και την αποδοτικότητα των αλγορίθμων.

1.3 Δομή Εργασίας

Τα υπόλοιπα κεφάλαια της εργασίας δομούνται ως εξής:

- Στο *Κεφάλαιο 2* περιγράφονται οι πιο σύγχρονες εργασίες και συνεισφορές πάνω στο αντικείμενο της υπερφασματικής ταξινόμησης εικόνας αλλά και στη γενικότερη χρήση Συνελικτικών Νευρωνικών Δικτύων (*CNN*).
- Το *Κεφάλαιο 3* προσπαθεί να εντάξει τον αναγνώστη με όσο πιο συνοπτικό και κατανοητό τρόπο γίνεται, στις βασικές λειτουργίες και αρχιτεκτονικές των Τεχνικών Νευρωνικών Δικτύων (*ANN*), όπως επίσης και στο γενικότερο θεωρητικό υπόβαθρο στο οποίο αυτά εντάσσονται. Στο τέλος του κεφαλαίου (*ενότητες 3.4 - 3.6*) παρατίθεται μία σύντομη περιγραφή των Συνελικτικών Νευρωνικών Δικτύων (*CNN*), των Μηχανών Διανυσματικής Υποστήριξης (*SVM*) όπως επίσης και των τριών αλγορίθμων εφαρμογής της εργασίας.
- Στο *Κεφάλαιο 4* περιγράφονται όλα τα βιβλιογραφικά δεδομένα που αποκτήθηκαν για τους σκοπούς της εργασίας (*ενότητα 4.1*) και αναλύονται όλα τα δεδομένα που δημιουργήθηκαν κατά την εκπόνηση της εργασίας μαζί με απαραίτητα στάδια επεξεργασία που ακολουθήθηκαν σε αυτά (*ενότητα 4.2*).
- Στο *Κεφάλαιο 5* παρουσιάζονται αναλυτικά όλα τα αποτελέσματα που προκύπτουν στην παρούσα εργασία. Τα αποτελέσματα αυτά περιλαμβάνουν ταξινομημένες εικόνες, πίνακες αξιολόγησης χρονικής απόδοσης και ακριβειών αλλά και γραφήματα με τα οποία γίνεται ευκολότερα αντιληπτή η απόδοση του κάθε ταξινομητή στον αναγνώστη.
- Τέλος, στο *Κεφάλαιο 6* παρουσιάζονται συγκεντρωτικά τα συμπεράσματα που προκύπτουν από την ανάλυση των αποτελεσμάτων ενώ παράλληλα αναφέρονται και μερικές προτάσεις που προκύπτουν τόσο ως προβληματισμοί αλλά και ως πρακτικά προβλήματα κατά τη διεκπεραίωση της εργασίας, οι οποίες και προτείνονται για εφαρμογή σε μελλοντικές εργασίες.

2

ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Η ταξινόμηση εικόνας αποτελεί μία από τις σημαντικότερες και πιο συνήθεις εφαρμογές όσον αφορά τα υπερφασματικά δεδομένα. Τις τελευταίες δύο δεκαετίες, έχει δοκιμαστεί μία πληθώρα αλγορίθμων και προσεγγίσεων με σκοπό την καλύτερη αξιοποίηση των υπερφασματικών χαρακτηριστικών, την σύνδεσή τους με τις χωρικές οντότητες της εικόνας και εν τέλει την βελτίωση της απόδοσης στις παραγόμενες ταξινομήσεις. Οι δύο βασικότερες προσεγγίσεις ήταν τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) και οι Μηχανές Διανυσματικής Υποστήριξης (SVM).

Στις αρχές της δεκαετίας του 1990 ήταν ευρέως δεδομένο, πως η χρήση ενός απλού ΤΝΔ 3 επιπέδων θα μπορούσε να λύσει σχεδόν όλα τα προβλήματα ανάλυσης δεδομένων και εικόνων, απλώς με την κατάλληλη προσθήκη επαρκών νευρώνων στο κρυφό επίπεδο και μιας μη γραμμικής συνάρτησης ενεργοποίησης (βλ. *Κεφάλαιο 3.3*). Παρ' όλα αυτά, η μέχρι τότε ανυπαρξία συνθετότερων αρχιτεκτονικών ΤΝΔ όπως επίσης και αλγορίθμων εκπαίδευσης, αντικατέστησε σταδιακά τη χρήση των ΤΝΔ με εκείνη των SVM. Μάλιστα οι SVM, και ειδικότερα εκείνες με μη γραμμικό ταξινομητή (βλ. *Κεφάλαιο 3.5*), για πολλά χρόνια θεωρούνταν ως οι αλγόριθμοι στην αιχμή της τεχνολογίας (*state of the art*) και χρησιμοποιούνταν πάραυτα στις ταξινομήσεις υπερφασματικών εικόνων.

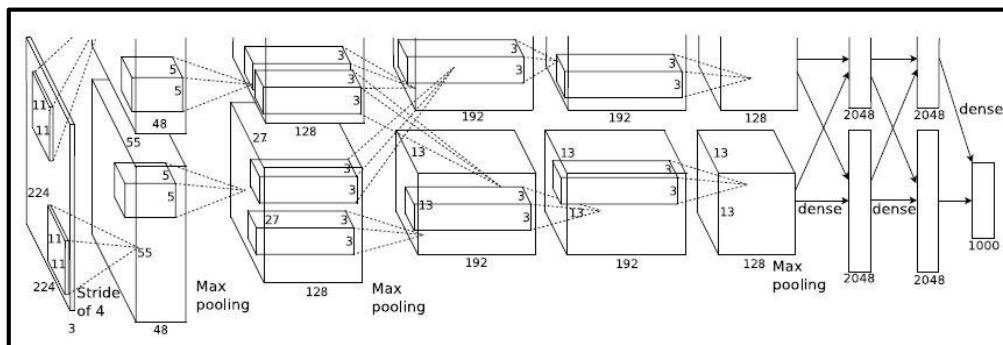
Βεβαίως από το 2006 και ύστερα όταν οι *Geoffrey E. Hinton et al.* δημοσίευσαν την εργασία τους με τίτλο *A Fast Learning Algorithm for Deep Belief Nets*, άρχισε μία σημαντική ανάπτυξη στον σχεδιασμό και τη δημιουργία ΤΝΔ με αρχιτεκτονικές βαθιάς εκμάθησης (*Deep Learning Architectures*). Ειδικότερα, οι αρχιτεκτονικές *Deep Learning* αποτέλεσαν μία από τις σημαντικότερες ανακαλύψεις την τελευταία δεκαετία στον κλάδο της Τεχνητής Νοημοσύνης (*Artificial Intelligence*) αφού πλέον είχε καταστεί δυνατή η εξαγωγή πιο σύνθετων και εξεζητημένων χαρακτηριστικών από τα υπερφασματικά δεδομένα με διάφορες μεθόδους (νευρωνικά δίκτυα, ιεραρχικά μοντέλα πιθανοφάνειας, αλγόριθμους εκμάθησης χαρακτηριστικών κ.α.). Αυτή η 'μεθοδολογική επανάσταση' προφανώς επηρέασε και επαναπροσδιόρισε τον τρόπο με τον οποίο επιτυγχάνονται ακριβέστερες και αποτελεσματικότερες ταξινομήσεις σε υπερφασματικά δεδομένα εικόνων.

Στις ακόλουθες ενότητες (2.1 και 2.2) περιγράφονται μόλις οι πιο σύγχρονες και εξειδικευμένες εργασίες, τεχνικές και συνεισφορές στον συγκεκριμένο κλάδο, με τις οποίες άλλωστε προσπαθεί να αναμετρηθεί και η παρούσα διπλωματική εργασία. Περαιτέρω αναφορά πάνω σε παλαιότερες και ίσως ξεπερασμένες μεθοδολογίες ταξινόμησης κρίνεται περιττή και ως εκ τούτου θα παραληφθεί σε ένα ούτως ή άλλως ευρύ και ερευνητικά εκτεταμένο διεπιστημονικό αντικείμενο. Έτσι, στο υποκεφάλαιο 2.1 θα αναφερθούν οι σημαντικότερες εργασίες των τελευταίων χρόνων, στον κλάδο των ΤΝΔ και των *Deep Learning* προσεγγίσεων, οι οποίες έχουν συνεισφέρει ως προς την προεπεξεργασία των δεδομένων, την αρχιτεκτονική του δικτύου και τον αλγόριθμο εκπαίδευσης. Τέλος, στο υποκεφάλαιο 2.2 και στα πλαίσια της ειδικότερης βιβλιογραφίας, εξετάζονται οι 2 σημαντικότερες δημοσιεύσεις που χρησιμοποιούν το ίδιο ή παρόμοιο Συνελικτικό Νευρωνικό Δίκτυο (*CNN*) με της παρούσας εργασίας.

2.1 Εργασίες Deep Learning και CNN

Παρακάτω, οι 2 πρώτες εργασίες αφορούν τη χρήση σύνθετων Συνελικτικών Νευρωνικών Δικτύων (*CNN*) σε εφαρμογές αναγνώρισης και ταξινόμησης αντικειμένων πάνω σε ψηφιακές εικόνες, ενώ οι 7 επόμενες αφορούν γενικότερες μεθοδολογίες και συνεισφορές πάνω σε διάφορα χαρακτηριστικά (δομή, αλγόριθμος εκπαίδευσης κλπ.) ενός ταξινομητή υπερφασματικών δεδομένων.

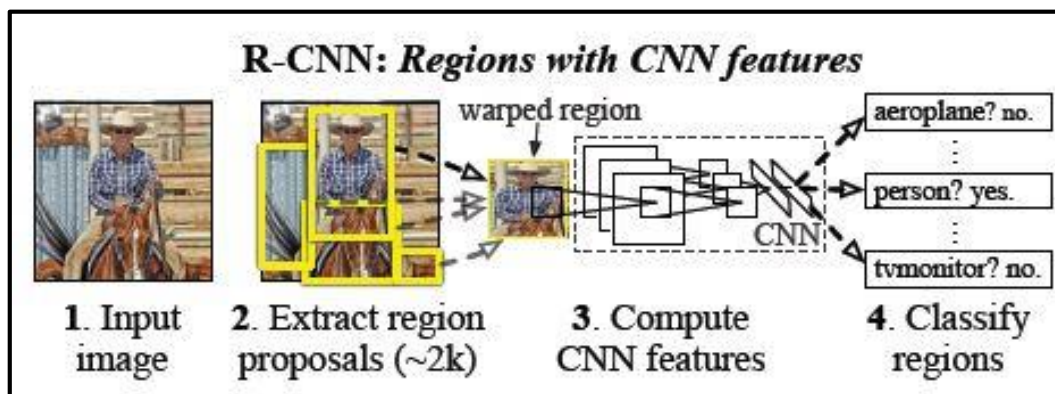
Το **2012** οι **Krizhevsky et al.** παρουσίασαν την εργασία τους με τίτλο ***ImageNet Classification with Deep Convolutional Neural Networks***. Η συγκεκριμένη εργασία καταπιάνεται με την αναγνώριση και ταξινόμηση διάφορων αντικειμένων και προτύπων πάνω σε έγχρωμες RGB εικόνες και η προτεινόμενη προσέγγιση δοκιμάστηκε πάνω στα διεθνώς βιβλιογραφικά δεδομένα συνόλου εικόνων ILSVRC-2010 και ILSVRC-2012. Η χρήση ενός Συνελικτικού δικτύου με την παράλληλη εφαρμογή επιπέδων μέγιστης συγκέντρωσης (*max pooling layers*) και την τελική σύνδεση 3 πλήρως συνδεδεμένων επιπέδων (Εικόνα 2.1), απέδωσε μακράν τα καλύτερα αποτελέσματα πάνω στα προς εξέταση δεδομένα. Το σημαντικότερο αποτέλεσμα όμως αυτής της εργασίας, είναι η παρακαταθήκη που έμεινε στην επιστημονική κοινότητα για την περαιτέρω χρήση και ανάπτυξη αρχιτεκτονικών *Deep Learning* σε περισσότερες εφαρμογές ανάλυσης δεδομένων.



Εικόνα 2.1: Το ΤΝΔ που χρησιμοποιήθηκε στην εργασία των *Krizhevsky et al.*

[Πηγή: *Krizhevsky et al. 2012*]

Πάνω στο ίδιο αντικείμενο, αυτό της αναγνώρισης προτύπων και αντικειμένων, ασχολείται και η εργασία των **Girshick et al.** του **2014** με τίτλο **'Rich feature hierarchies for accurate object detection and semantic segmentation'**. Στη δημοσίευσή τους αυτή, παρουσιάζουν αποτελέσματα κατά 30% πιο ακριβή από τις μέχρι τότε ταξινομήσεις αντικειμένων στα δεδομένα του συνόλου εικόνων *VOC 2012*. Και εδώ, η διαδικασία της εκμάθησης έγινε με ένα δίκτυο CNN ενώ η τελική πρόβλεψη της κάθε εικόνας με ένα γραμμικό ταξινομητή SVM. Η εργασία αυτή ήταν η πρώτη η οποία απέδειξε τα πολύ μεγάλα ποσοστά ακρίβειας που επιτυγχάνονται με ένα δίκτυο CNN στα ευρέως γνωστά δεδομένα *PASCAL VOC*, έναντι των παλαιότερων μεθόδων εκμάθησης.



Εικόνα 2.2: Η μεθοδολογία που ακολουθήθηκε από τους **Girshick et al.**

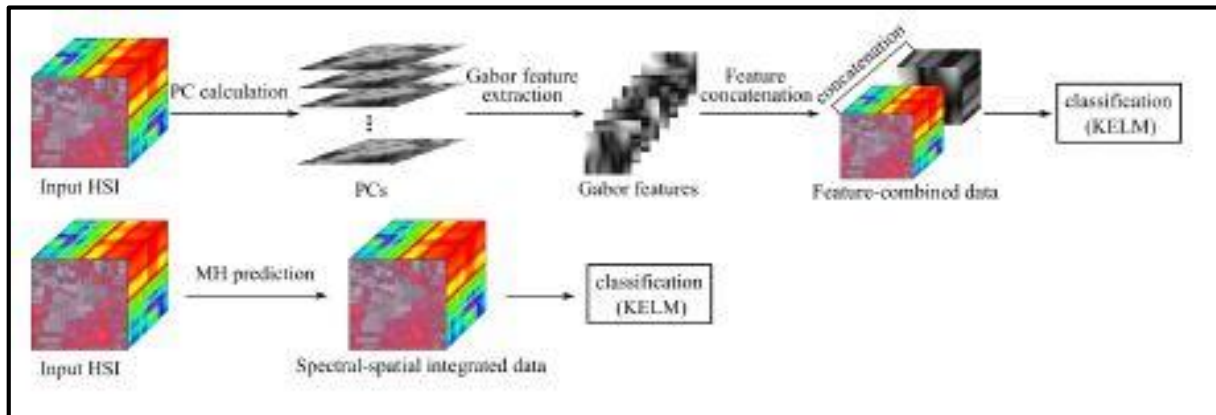
[Πηγή: **Girshick et al. 2014**]

Βεβαίως τα τελευταία χρόνια έχει καταγραφεί μεγάλη ανάπτυξη και ερευνητική κινητικότητα στις εφαρμογές που αφορούν σύνθετους αλγορίθμους και αρχιτεκτονικές και ταξινομήσεις υπερφασματικών δεδομένων. Σε αυτά τα πλαίσια το **2013** οι **Li et al.** στην εργασία τους **'Spectral-Spatial Classification of Hyperspectral Data Using Loopy Belief Propagation and Active Learning'** παρουσιάζουν μία ημι-επιβλεπόμενη μεθοδολογία ταξινόμησης οποία λαμβάνει υπόψη τόσο την φασματική (*spectral*) όσο και την χωρική (*spatial*) πληροφορία της εικόνας. Η συνεισφορά της προσέγγισης αυτής έγκειται α) στην μεθοδολογία *Ενεργής Εκμάθησης (Active Learning - AL)* που ακολουθήθηκε καθώς και β) στον τρόπο με τον οποίο υπολογίζονταν τα οριακά περιθώρια βελτίωσης του αλγορίθμου με τη μέθοδο *Loopy Belief Propagation (LBP)*. Η μέθοδος αυτή εφαρμόστηκε στα βιβλιογραφικά υπερφασματικά δεδομένα *Indian Pines* και *Pavia University*, ενώ τα παραγόμενα αποτελέσματα έφτασαν σε αποδόσεις αιχμής των συγκεκριμένων δεδομένων.

Πάνω στα ίδια δεδομένα, εφάρμοσαν την μεθοδό τους οι **Chen et al.** το **2014** στην εργασία τους με τίτλο **'Spectral-Spatial Classification of Hyperspectral Image Based on Kernel Extreme Learning Machine'**, η οποία χρησιμοποίησε ένα νευρωνικό δίκτυο ενός επιπέδου μαζί με μη γραμμική συνάρτηση ενεργοποίησης (*Kernel function*) με την ονομασία Μηχανή Ακρικής Εκμάθησης (*Extreme Learning Machine - ELM*). Πριν την εκπαίδευση όμως είχε προηγηθεί μία προεπεξεργασία των υπερφασματικών δεδομένων για την καλύτερη εξαγωγή φασματικών και χωρικών

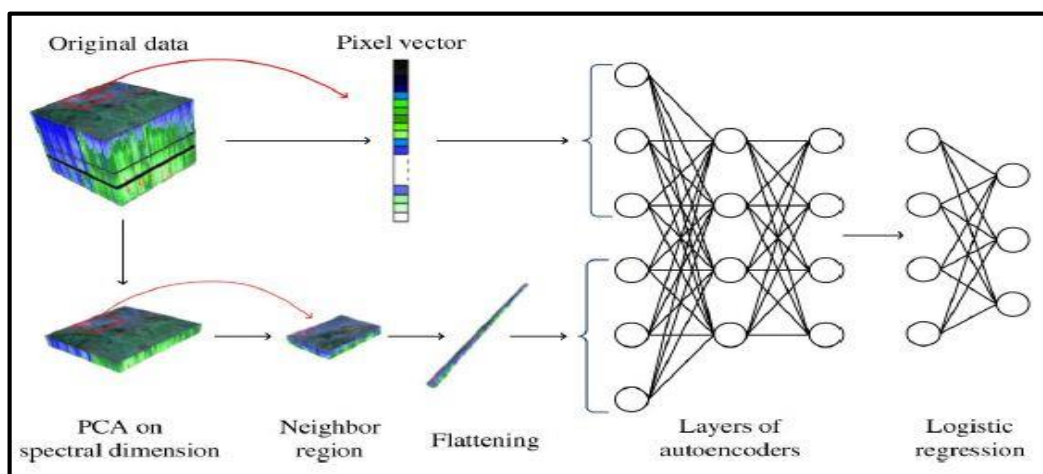
χαρακτηριστικών με τις μεθόδους των φίλτρων Gabor και της πρόβλεψης πολλαπλής υπόθεσης (*Multihypothesis – MH prediction*).

Τα αποτελέσματα που παρουσιάστηκαν απέδειξαν ότι η προσέγγιση αυτή ξεπερνά σε απόδοση τις συμβατικές μεθόδους ταξινόμησης με βάση μόνο την φασματική πληροφορία του κάθε εικονοστοιχείο (*pixel-based*).



Εικόνα 2.3: Οι δύο προσεγγίσεις που ακολουθήθηκαν κατά τους Chen et al. [Πηγή: Chen et al. 2014]

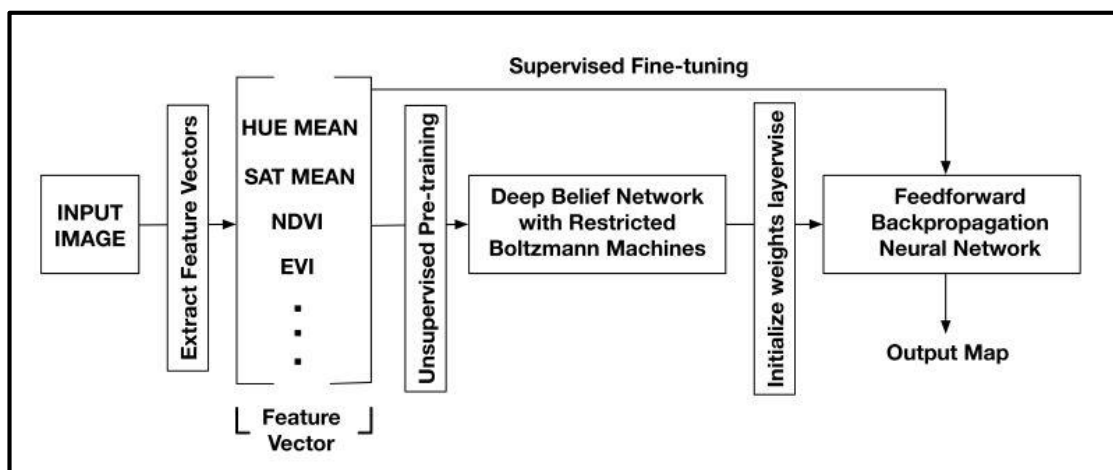
Εξάλλου, στην εργασία με τίτλο **‘Deep Learning-Based Classification of Hyperspectral Data’** που δημοσιεύτηκε το **2014** από τους **Yushi Chen et al.** εφαρμόζεται μία αρχιτεκτονική βαθιάς μάθησης (*Deep Learning*) μέσα από την οποία ταξινομούνται τα βιβλιογραφικά δεδομένα *KSC* και *Pavia University*. Στην εργασία αυτή προηγείται μία εφαρμογή της ανάλυσης PCA για την μείωση της 3^{ης} διάστασης των δεδομένων ενώ γίνεται παράλληλη εκμάθηση τόσο της φασματικής όσο και της χωρικής πληροφορίας με πολλαπλά επίπεδα *Stacked Auto – Encoders (SAE)*. Το συγκεκριμένο ΤΝΔ ολοκληρώνεται με τη χρήση ενός απλού επιπέδου παλινδρόμησης (*regression*) το οποίο κατανέμει τα κάθε εικονοστοιχεία στην κατάλληλη κλάση. Η μέθοδος αυτή συγκρίθηκε με τον ταξινομητή *RBF-SVM* και απεδείχθη ότι αποδίδει υψηλότερα ποσοστά ακρίβειας.



Εικόνα 2.4: Η μεθοδολογία που ακολουθήθηκε κατά τους Yushi Chen et al. [Πηγή: Yushi Chen et al. 2014]

Ακόμα, το 2015 οι **Xiaorui Ma et al.** στην εργασία τους με τίτλο **'Hyperspectral image classification via contextual deep learning'** παρουσιάζουν μία αρχιτεκτονική βαθιάς εκμάθησης η οποία αποδίδει υψηλά ποσοστά ακρίβειας με τη χρήση απλών ταξινομητών συνυπολογίζοντας τόσο τη χωρική όσο και την φασματική πληροφορία των υπερφασματικών δεδομένων. Η προτεινόμενη από τους συγγραφείς μέθοδος αποτελείται από 1 επίπεδο εισόδου, 3 κρυφά επίπεδα και 1 επίπεδο εξόδου. Έτσι, στα 3 αυτά κρυφά επίπεδα συνενώνονται με διαφορετικούς τρόπους η χωρικής και η υπερφασματική πληροφορία ενώ στο τελευταίο επίπεδο υλοποιείται μία πολυωνυμική παλινδρόμηση για την ταξινόμηση κάθε εικονοστοιχείου στην κατάλληλη κλάση. Από τα αποτελέσματα τα οποία παρουσιάστηκαν στην εργασία αυτή, αποδείχθηκε πως η μέθοδος αυτή μπορεί να εξάγει χωρικά και φασματικά χαρακτηριστικά υψηλής ποιότητας και ακριβώς γι' αυτό τον λόγο επιτυγχάνει αποδόσεις αιχμής, συγκρινόμενη με άλλες αρχιτεκτονικές βαθιάς εκμάθησης.

Τέλος, το 2015 οι **Saikat Basu et al.** στην εργασία τους με τίτλο **'DeepSat – A Learning framework for Satellite Imagery'** εγκαινιάζουν τόσο μία νέα προσέγγιση ταξινόμησης εικόνων καθώς επίσης και 2 νέα σετ δεδομένων ελέγχου. Το κάθε σετ δεδομένων αποτελείται από διάφορα 'αποκόμματα - patches' κατηγοριών εδάφους ενοποιημένα σε μία εικόνα με 4 κανάλια, ενώ η προτεινόμενη προσέγγιση αποτελείται από μία αρχική μη επιβλεπόμενη εκμάθηση μέσω ενός δικτύου Deep Belief (DBN) και από ένα ακόλουθο ΤΝΔ πρόσθιας τροφοδότησης με αλγόριθμο επιβλεπόμενης ταξινόμησης. Στην εργασία αυτή εφαρμόστηκαν και αξιολογήθηκαν η προτεινόμενη προσέγγιση, καθώς επίσης και διάφορες άλλες τεχνικές Deep Learning (SAE, CNN, DBN) με τα παραγόμενα αποτελέσματα να αναδεικνύουν την προτεινόμενη προσέγγιση ως αποδοτικότερη.



Εικόνα 2.5: Η προτεινόμενη προσέγγιση ταξινόμησης, κατά τους Saikat Basu et al.
[Πηγή: Saikat Basu et al. 2015]

2.2 Πρόσφατες Σχετικές Εργασίες

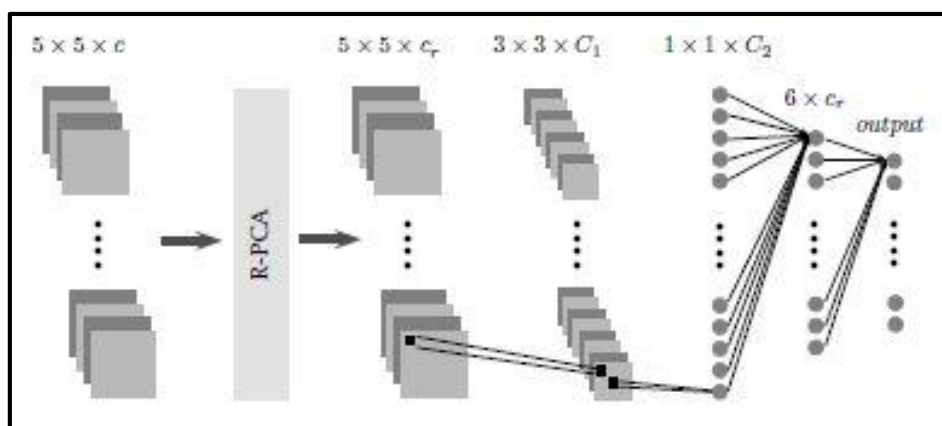
Σε αυτή την ενότητα θα αναλυθούν οι δύο βασικότερες εργασίες των τελευταίων ετών, οι οποίες καταπιάνονται με το ζήτημα της ταξινόμησης υπερφασματικών δεδομένων με τη χρήση δικτύων CNN βαθιάς εκμάθησης.

Συγκεκριμένα, το 2015 οι **Μακαντάσης κ.α.** παρουσίασαν την εργασία τους **‘Deep Supervised Learning For Hyperspectral Data Classification Through Convolutional Neural Networks’** με την οποία εγκαινιάζουν τη χρήση ενός δικτύου CNN το οποίο ακολουθείται από ένα δίκτυο πολυεπίπεδου αισθητήρα (MLP).

Η συγκεκριμένη αρχιτεκτονική προσφέρει παράλληλη αξιοποίηση της φασματικής και της χωρικής πληροφορίας μέσω των διαδοχικών συνελκτικών επιπέδων ενώ παράλληλα, με τη χρήση του αλγορίθμου οπισθοδιάδοσης σφάλματος (*backpropagation error*) στο δίκτυο MLP, αποδίδει ως μία από τις πιο αξιόπιστες μεθοδολογίες εκμάθησης και ταξινόμησης υπερφασματικών δεδομένων.

Σημειώνεται πως αντιθέτως με την συνήθη δομή των δικτύων CNN, η οποία επιτάσσει ένα επίπεδο μέγιστης συγκέντρωσης (*max-pooling layer*) έπειτα από ένα συνελκτικό, σε αυτή την εργασία δεν χρησιμοποιήθηκαν καθόλου επίπεδα μέγιστης συγκέντρωσης τα οποία συνήθως συνεισφέρουν σε εφαρμογές αναγνώρισης προτύπων, αντικειμένων κ.α.

Επιπλέον, σημειώνεται πως σε αυτή την εργασία είχε προηγηθεί μείωση στην 3^η διάσταση (υπερφασματική πληροφορία) των δεδομένων, μέσω της ανάλυσης κυρίων συνιστωσών (PCA). Τα εξαγόμενα αποτελέσματα από το προτεινόμενο δίκτυο, συγκρίθηκαν με άλλες 2 μεθόδους ταξινόμησης μέσω αλγορίθμων μηχανών διανυσματικής υποστήριξης.



Εικόνα 2.6: Η προτεινόμενη αρχιτεκτονική δικτύου CNN, κατά τους Μακαντάσης κ.α.
[Πηγή: Μακαντάσης κ.α. 2015]

Ειδικότερα η προσέγγιση αυτή απέδειξε ότι αποδίδει καλύτερα και από τον μη γραμμικό ταξινομητή RBF-SVM, ο οποίος για χρόνια θεωρείτο ως η πιο προηγμένη μέθοδος στις ταξινομήσεις υπερφασματικών εικόνων. Συγκεκριμένα στα

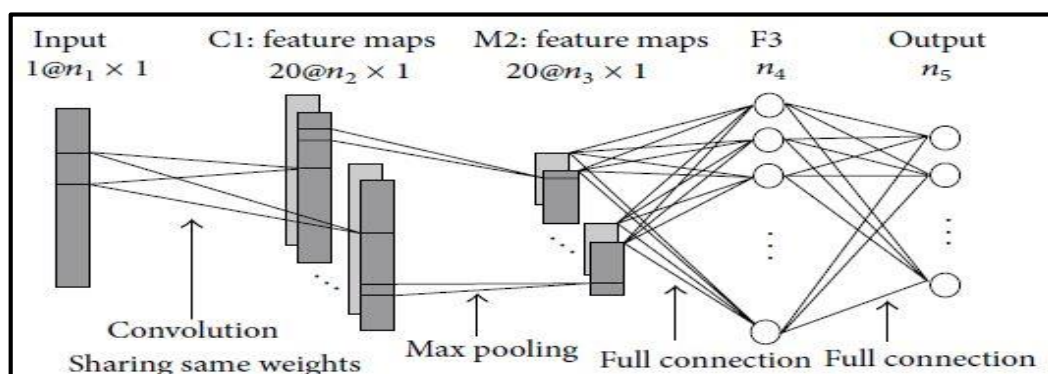
βιβλιογραφικά δεδομένα *Pavia Center*, *Pavia University*, *Indian Pines* και *Salinas* στα οποία εφαρμόστηκαν οι ταξινομητές, το δίκτυο CNN επιτυγχάνει σταθερά καλύτερες αποδόσεις από αυτές των SVM, οι οποίες παράλληλα αγγίζουν το 99% και 100% σε ποσοστά ακρίβειας.

Η τελευταία εργασία η οποία θα αναλυθεί σε αυτό το κεφάλαιο, είναι αυτή των **Wei Hu et al.** η οποία δημοσιεύτηκε το **2015** με τίτλο **'Deep Convolutional Neural Networks for Hyperspectral Image Classification'**, η οποία καταπιάνεται επίσης με την ταξινόμηση υπερφασματικών δεδομένων, μέσα από συνελκτικό νευρωνικό δίκτυο βαθιάς εκμάθησης. Εν αντιθέσει με πολλές από τις προαναφερθείσες προσεγγίσεις, η μέθοδος που περιγράφεται εδώ, λαμβάνει υπόψη μονάχα την φασματική πληροφορία της εικόνας και όχι την χωρική. Έτσι, η όποια συνέλιξη σημάτων εφαρμόζεται μέσα στο δίκτυο, αφορά το ίδιο εικονοστοιχείο κάθε φορά και τα διάφορα κανάλια που περιέχουν το σύνολο της υπερφασματικής πληροφορίας.

Συγκεκριμένα, προτείνεται μία δομή δικτύου η οποία αποτελείται από 5 επίπεδα: Ένα επίπεδο εισόδου στο οποίο ως είσοδος ορίζεται το κάθε εικονοστοιχείο επί τον αριθμό των καναλιών του, ένα συνελκτικό επίπεδο όπου υλοποιείται συνέλιξη των 20 πλησιέστερων καναλιών κάθε φορά, ένα επίπεδο μέγιστης συγκέντρωσης, ακολούθως ένα πλήρες συνδεδεμένο επίπεδο και τέλος το επίπεδο εξόδου που αντιστοιχεί στις προκύπτουσες ταξινομημένες κλάσεις (βλ. *Εικόνα 2.7*).

Σημειώνεται πως η διαδικασία εκμάθησης υλοποιήθηκε με τον αλγόριθμο οπισθοδιάδοσης σφάλματος και τη μέθοδο επικλινούς καθόδου, ενώ ο αριθμός των δεδομένων εκπαίδευσης ήταν ορισμένος στα 200 εικονοστοιχεία σε κάθε κλάση. Οι υλοποιούμενες εφαρμογές αφορούν τα βιβλιογραφικά δεδομένα *Pavia University*, *Indian Pines* και *Salinas*.

Τέλος, οι ταξινομήσεις που προκύπτουν από τη συγκεκριμένη προσέγγιση συγκρίνονται με αντίστοιχες άλλων ΤΝΔ καθώς και με του ταξινομητή RBF-SVM. Τα αποτελέσματα που παρουσιάζονται αποδεικνύουν ότι η εφαρμογή συνελκτικών επιπέδων στα διάφορα κανάλια της υπερφασματικής πληροφορίας είναι αποδοτικότερη από τις άλλες συγκρινόμενες μεθοδολογίες επιτυγχάνοντας καλύτερες ακρίβειες κατά 2% περίπου.



Εικόνα 2.7: Η προτεινόμενη αρχιτεκτονική δικτύου CNN, κατά τους Wei Hu et al.

[Πηγή: Wei Hu et al. 2015]

3

ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Στο κεφάλαιο αυτό περιγράφονται έννοιες και ορισμοί οι οποίοι ξεφεύγουν από τα στενά πλαίσια του αντικειμένου της Τηλεπισκόπησης ή της Όρασης Υπολογιστών γενικότερα. Αντιθέτως, μία προσπάθεια περιγραφής και ανάλυσης σύγχρονων αλγορίθμων και αρχιτεκτονικών Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ), μοιραία εντάσσει έννοιες που αφορούν περισσότερο το επιστημονικό αντικείμενο της Υπολογιστικής και κατ' επέκταση Τεχνητής Νοημοσύνης. (*Artificial Intelligence, Computational Intelligence*).

Παρακάτω γίνεται προσπάθεια να περιγραφούν με όσο πιο συνοπτικό τρόπο γίνεται οι βασικές αρχές λειτουργίας των ΤΝΔ καθώς επίσης οι αρχιτεκτονικές και τα χαρακτηριστικά των αλγορίθμων που εφαρμόστηκαν. Αρχικά όμως, κρίνεται απαραίτητο να γίνει αναφορά στις γενικότερες έννοιες που αφορούν τους κλάδους της Τεχνητής και Υπολογιστικής Νοημοσύνης, της Μηχανικής Μάθησης (*Machine Learning*).

3.1 Τεχνητή Νοημοσύνη και Μηχανική Μάθηση

Τεχνητή Νοημοσύνη

Η αναπαράσταση της γνώσης και η προσαρμογή των σημασιολογικών περιγραφών αποτελεί πεδίο έρευνας τις τελευταίες δεκαετίες. Οι αριθμητικοί υπολογισμοί είναι ένα πεδίο στο οποίο οι ανθρώπινες νοητικές λειτουργίες έχουν υποκατασταθεί με επιτυχία από τον Η/Υ κάτι που αναδεικνύει τη δυνατότητα υλοποίησης περαιτέρω ανεπτυγμένων συστημάτων τα οποία και θα επιτυγχάνουν μεγάλες ταχύτητες λογισμού σε ακόμα πιο εξεζητημένες εφαρμογές.

Στα πλαίσια αυτής της λογικής ξεκίνησε η πρόοδος του κλάδου της επιστήμης υπολογιστών ο οποίος ασχολείται με τη σχεδίαση και την υλοποίηση υπολογιστικών συστημάτων δηλαδή συστημάτων που επιδεικνύουν χαρακτηριστικά που σχετίζονται με τη νοημοσύνη στην ανθρώπινη συμπεριφορά και ονομάστηκε τεχνητή νοημοσύνη [Πλέρου, 2012]. Εξάλλου με τον όρο Τεχνητή Νοημοσύνη (*Artificial Intelligence*) ορίζεται η προσπάθεια ενσωμάτωσης και προσομοίωσης σε ένα υπολογιστικό σύστημα των βασικών χαρακτηριστικών της ανθρώπινης σκέψης για να επιλυθούν πρακτικά προβλήματα, [Openshaw and Openshaw, 1997].

Υπολογιστική Νοημοσύνη

Η Υπολογιστική Νοημοσύνη (*Computational Intelligence*) αποτελεί έναν από τους σημαντικότερους και πλέον εξελισσόμενους κλάδους της πληροφορικής αλλά και της Τεχνητής Νοημοσύνης γενικότερα. Στον κλάδο της Υπολογιστικής Νοημοσύνης ανήκουν και τα Τεχνητά Νευρωνικά Δίκτυα- TNN (*Artificial Neural Networks - ANN*).

Γενικότερα, ο κλάδος της Υπολογιστικής Νοημοσύνης συντίθεται από τις παρακάτω τέσσερις διακριτές τεχνολογίες, μία από τις οποίες είναι και Νευρωνικά Δίκτυα:

- Λογική της Ασάφειας (Fuzzy Logic)
- Νευρωνικά Δίκτυα (Neural Networks)
- Έμπειρα Συστήματα (Expert Systems)
- Γενετικοί Αλγόριθμοι (Genetic Algorithms)

Μηχανική Μάθηση (Machine Learning)

Η μηχανική μάθηση είναι μια περιοχή της τεχνητής νοημοσύνης η οποία αφορά αλγορίθμους και μεθόδους που επιτρέπουν τη σταδιακή 'εκμάθησή' τους σε επιμέρους εργασίες. Με αυτόν τον τρόπο καθίσταται εφικτή η κατασκευή προσαρμόσιμων προγραμμάτων υπολογιστών τα οποία λειτουργούν με βάση την αυτοματοποιημένη ανάλυση συνόλων δεδομένων και όχι τη διαίσθηση ή τη γνώση των μηχανικών κατά τη στιγμή που τα προγραμματίσαν. Εξάλλου το 1959 ο σχεδιαστής παιχνιδιών Άρθουρ Σάμουελ (Arthur Samuel) όρισε ως μηχανική μάθηση "Το πεδίο μελέτης όπου δίνει στους υπολογιστές την δυνατότητα να μαθαίνουν χωρίς να έχουν προγραμματιστεί". Η έννοια της Μηχανικής Μάθησης είναι στενά συνδεδεμένη και συνήθως επικαλύπτεται με το αντικείμενο της Υπολογιστικής Στατιστικής καθώς και οι δύο αυτοί κλάδοι καταπιάνονται με εφαρμογές όπως η ανάλυση και 'εξόρυξη' δεδομένων (*data mining*).

Οι εργασίες Μηχανικής Μάθησης χωρίζονται σε 3 βασικές κατηγορίες ανάλογα με τη φύση του 'μαθησιακού σήματος' ή/και τον εάν υπάρχει ανάδραση μέσω ενός μαθησιακού συστήματος, σε:

- **Επιβλεπόμενη Μάθηση**, κατά την οποία είναι απαραίτητο να δίνονται κάποια δεδομένα εισόδου με τα αντίστοιχα επιθυμητά αποτελέσματα –δεδομένα εξόδου, στον Η/Υ. Με αυτό τον τρόπο ο υπολογιστής εκπαιδεύεται στο να μεταβαίνει από το κάθε νέο δεδομένο εισόδου στο αντίστοιχο αποτέλεσμα.
- **Μη Επιβλεπόμενη Μάθηση**, κατά την οποία δεν παρέχονται δεδομένα εισόδου στο μαθησιακό σύστημα. Έτσι, ο υπολογιστής καλείται να μεταβεί μόνος του σε ένα 'επιθυμητό αποτέλεσμα' αναλύοντας ομοιότητες, ανομοιογένειες και πρότυπα από τα δεδομένα.
- **Ενισχυμένη Μάθηση**, κατά την οποία ένα υπολογιστικό πρόγραμμα επιδρά σε ένα δυναμικό περιβάλλον χωρίς να έχει καθοριστεί το πότε ή εάν βρίσκεται κοντά στον επιθυμητό στόχο.

Βεβαίως, παραπάνω περιγράφονται οι γενικές κατηγορίες καθώς υπάρχουν ενδιάμεσες και συνδυαστικές περιπτώσεις αλγορίθμων και εφαρμογών Μηχανικής Μάθησης (π.χ. Ημι – επιβλεπόμενη Μάθηση).

Σήμερα, οι εφαρμογές αλγορίθμων Μηχανικής Μάθησης εφαρμόζονται σε πολλούς κλάδους της οικονομικής σφαίρας (οικονομικά, ιατρική, βιοπληροφορική, χρηματιστήριο, διαδικτυακές εφαρμογές κ.α.) με την παράλληλη εξέλιξη νέων μεθόδων και πρακτικών και στο ερευνητικό επίπεδο.

Μάθηση Εις Βάθος (Deep Learning)

Ως Deep Learning, νοείται ο κλάδος της Μηχανικής Μάθησης που καταπιάνεται με πιο περίπλοκες και συνήθως μη γραμμικές αρχιτεκτονικές αλγορίθμων, έχοντας ως στόχο την εξαγωγή πιο σύνθετων και εξεζητημένων χαρακτηριστικών από τα διάφορα δεδομένα.

Η βασική διαφοροποίηση στις εφαρμογές Deep Learning, συναντάται στον αριθμό των μετασχηματισμών και αναπροσαρμογών που δέχεται ένα σήμα, από τη στιγμή της 'εισόδου' του έως τη στιγμή της 'εξόδου' του από το υπολογιστικό σύστημα. Με άλλα λόγια, το πόσο θα μετασχηματιστούν οι διάφοροι παράμετροι (ή βάρη) ενός αλγορίθμου κατά τη διαδικασία της μάθησης, έτσι ώστε να φτάσει πιο κοντά στον 'επιθυμητό' στόχο. Ο αριθμός των αναπροσαρμογών, πρακτικά μεταφράζεται στον αριθμό των διάφορων επιπέδων (layers) με τα οποία είναι δομημένος ένας Deep Learning αλγόριθμος. Τα επίπεδα αυτά δομούνται με ιεραρχικό τρόπο, ώστε οι συντελεστές των υψηλότερων επιπέδων να αναπροσαρμόζονται με βάση αυτούς των χαμηλότερων κ.ο.κ.

Οι μέθοδοι Deep Learning συνήθως συναντώνται μέσω των Τεχνητών Νευρωνικών Δικτύων – ΤΝΔ (αλλιώς Deep Neural Networks - DNN), πάνω σε διάφορες δομές και αρχιτεκτονικές (Convolutional Neural Networks, Deep Belief Networks, Convolutional Deep Belief Networks, Deep Boltzmann Machines, Stacked Auto – Encoders κ.α.).

3.2 Νευρωνικά Δίκτυα

Η μελέτη της λειτουργίας του ανθρώπινου εγκεφάλου αποτέλεσε το έναυσμα για την έρευνα και εφαρμογή μαθηματικών μοντέλων και αλγορίθμων, σε μια προσπάθεια προσομοίωσης της λειτουργίας αυτής. Η νευροεπιστήμη (neuroscience) είναι εκείνος ο κλάδος που αρχικά προσπάθησε να εξηγήσει τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου, χωρίς ωστόσο κάτι τέτοιο να είναι απόλυτα ξεκάθαρο ακόμα και σήμερα.

Πάνω σε αυτή τη λογική βασίζεται και η φιλοσοφία των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) , ως μια προσπάθεια προσομοίωσης, με την βοήθεια υπολογιστών, του ανθρώπινου νευρικού συστήματος. Εξάλλου, τα ΤΝΝ παρέχουν ένα εναλλακτικό μοντέλο το οποίο είναι εμπνευσμένο από τα βιολογικά μοντέλα σύμφωνα με το οποίο οι υπολογισμοί γίνονται παράλληλα και μαζικά και η εκπαίδευση αντικαθιστά την ανάπτυξη προγράμματος [Πλέρου Α., 2012].

3.2.1 Βιολογικά Νευρωνικά Δίκτυα

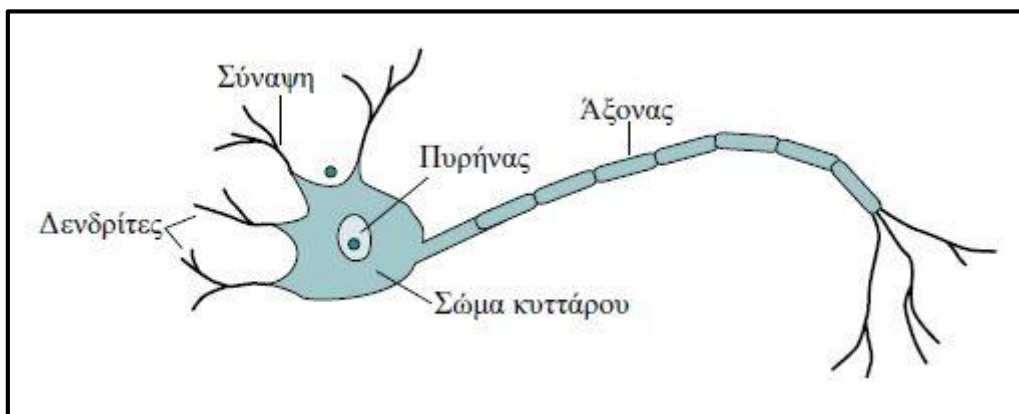
Μελετώντας τα βιολογικά νευρωνικά δίκτυα που υπάρχουν στον ανθρώπινο εγκέφαλο παρατηρούμε ότι αποτελούνται από νευρώνες.

Ως νευρώνας νοείται το μικρότερο τμήμα του εγκεφάλου που είναι ικανό να επεξεργαστεί πληροφορίες και η ύπαρξη του διαφοροποιεί τα ζώα από τα φυτά (τα φυτά δεν έχουν νευρώνες). Ένας νευρώνας λειτουργεί όπως και τα άλλα κύτταρα ενός οργανισμού και ο απόλυτος αριθμός τους σε έναν ανθρώπινο εγκέφαλο εκτιμάται περίπου στους 10^{10} .

Κάθε νευρώνας συνδέεται με πολλούς άλλους νευρώνες με τις συνδέσεις αυτές να ονομάζονται συνάψεις. Κάθε νευρώνας συμμετέχει σε περίπου 10^4 συνάψεις. Έτσι, ως νευρωνικό δίκτυο (neural net) ορίζεται ένας αριθμός νευρώνων μαζί με της διασυνδέσεις τους. Το συνολικό σύστημα των νευρωνικών δικτύων στον ανθρώπινο οργανισμό ονομάζεται Κεντρικό Νευρικό Σύστημα (Central Nervous System).

Ένας νευρώνας δεν διαφέρει σε τίποτα από τα άλλα κύτταρα του οργανισμού και αποτελείται από:

- Το σώμα του κυττάρου, στο οποίο βρίσκεται ο πυρήνας και άρα το γενετικό υλικό του οργανισμού.
- Τον άξονα ο οποίος μεταδίδει τα εξερχόμενα σήματα σε άλλους νευρώνες.
- Τους δενδρίτες, οι οποίοι αποτελούν λεπτές επεκτάσεις, και έρχονται σε επαφή με άλλους νευρώνες δεχόμενοι τα εισερχόμενα σήματα.



Εικόνα 3.1: Σχηματική απεικόνιση ενός τυπικού νευρώνα [Πηγή: Αργυράκης Π. 2001]

Αξίζει να αναφερθεί, πως η επεξεργασία στους βιολογικούς νευρώνες συντελείται πολύ πιο αργά απ' ό τι στα σύγχρονα ψηφιακά συστήματα. Όμως ο εγκέφαλος αντισταθμίζει την σχετικά αργή ταχύτητα λειτουργίας των νευρώνων, με τον πολύ μεγάλο αριθμό συνδέσεων μεταξύ των νευρώνων. Υπολογίζεται πως ο φλοιός του ανθρώπινου εγκεφάλου έχει 10^7 νευρώνες και 60^{10} συνδέσεις μεταξύ αυτών, γεγονός που είναι αδύνατον να προσομοιωθεί από κάποιο τεχνητό υπολογιστικό σύστημα, τουλάχιστον με τις μέχρι τώρα τεχνολογικές δυνατότητες.

Ενδιαφέρον παρουσιάζει επίσης ο τρόπος με τον οποίο λειτουργεί ένας νευρώνας, οποίος και μοιάζει με τον δυαδικό (*binary*) τρόπο λειτουργίας των ηλεκτρονικών υπολογιστών. Συγκεκριμένα ένας νευρώνας μπορεί να βρίσκεται είτε σε 'ενεργή' κατάσταση είτε σε 'μη ενεργή'.

Ένας νευρώνας μπορεί να παράγει ηλεκτρικό σήμα, διάρκειας της τάξης ενός msec, το οποίο και συνεχίζει την πορεία του μέσα στο νευρωνικό δίκτυο, από νευρώνα σε νευρώνα, χωρίς να ελαττωθεί καθόλου. Το ηλεκτρικό αυτό σήμα, οφείλεται στην διαφορά δυναμικού που δημιουργείται στην μεμβράνη του κυττάρου μεταξύ της εσωτερικής και εξωτερικής του επιφάνειας. Έτσι λοιπόν όταν καταφθάνουν όλα τα εκπεμπόμενα σήματα σε έναν νευρώνα, αθροίζονται τα ηλεκτρικά δυναμικά τους και υπάρχουν 2 περιπτώσεις:

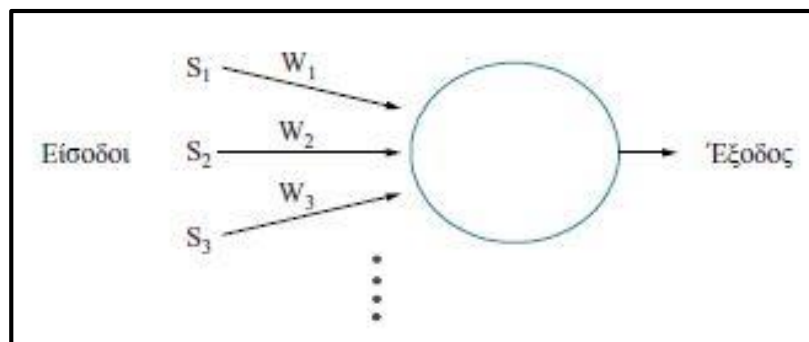
- Εάν αυτό το άθροισμα ξεπεράσει μια δεδομένη τιμή, τότε ο νευρώνας διεγείρεται, θεωρείται 'ενεργός' και εκπέμπει μέσω του άξονά του ένα ηλεκτρικό σήμα
- Εάν αυτό το άθροισμα δεν ξεπεράσει τη δεδομένη τιμή τότε ο νευρώνας παραμένει 'μη ενεργός' και το πλεονάζων δυναμικό χάνεται.

Σημειώνεται πως το εισερχόμενο σε ένα νευρώνα σήμα μπορεί να είναι διεγερτικό ή ανασταλτικό αναλόγως με το εάν βοηθάει το συνολικό άθροισμα να φτάσει τη δεδομένη τιμή 'κατώφλι' ή το αντίθετο.

3.2.2 Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)

Τα τεχνητά νευρωνικά δίκτυα αποτελούν μια προσπάθεια προσέγγισης της λειτουργίας του ανθρώπινου εγκεφάλου από μια μηχανή και έχουν την ικανότητα να εκτελούν υπολογισμούς με μαζικό παράλληλο τρόπο [Hopfield,1985].

Ένα ΤΝΔ είναι μία προσπάθεια προσομοίωσης της λειτουργίας του ανθρώπινου εγκεφάλου που σκοπό έχει την αποτελεσματικότερη, ταχύτερη και παράλληλη επεξεργασία πληροφοριών. Επί της ουσίας τα ΤΝΔ είναι μία αλληλουχία διασυνδεδεμένων Στοιχείων Επεξεργασίας (*Processing Units - Pus*) τα οποία ονομάζονται νευρώνες κατά αντιστοιχία των βιολογικών νευρώνων. Στα ΤΝΔ, οι νευρώνες – Pus συνδέονται μεταξύ τους και είναι οργανωμένοι σε επίπεδα (*layers*). Κάθε επιμέρους νευρώνας μπορεί να έχει πολλές εισόδους (*inputs*) για εισερχόμενη πληροφορία αλλά μόνο μία έξοδο (*output*) απ' όπου και εξέρχεται το σήμα εξόδου.



Εικόνα 3.2: Τυπική αναπαράσταση ενός νευρώνα, με πολλές εισόδους και μία έξοδο [Πηγή: Αργυράκης Π. 2001]

Επιπλέον χαρακτηριστικά στοιχεία ενός τεχνητού νευρώνα αποτελούν οι συνάψεις, ο αθροιστής και η συνάρτηση ενεργοποίησης. Οι *συνάψεις* αποτελούν επί της ουσίας τον τρόπο σύνδεσης των νευρώνων μεταξύ τους και χαρακτηρίζονται από το αντίστοιχο *συναπτικό βάρος* κάθε σήματος εισόδου. Ο *αθροιστής* αθροίζει όλα τα σήματα εισόδου (αφού προηγουμένως έχουν σταθμιστεί με τα αντίστοιχα βάρη) και η *συνάρτηση ενεργοποίησης* μετασχηματίζει τα σήματα εισόδου σε ένα σήμα εξόδου. Οι συναρτήσεις ενεργοποίησης μπορεί να είναι είτε γραμμικές είτε μη γραμμικές.

Πάντως, πέραν της όποιας αρχιτεκτονικής τους σχεδίασης, η βασική ομοιότητα των ΤΝΔ σε σχέση με τα βιολογικά νευρωνικά δίκτυα εντοπίζεται στον τρόπο με τον μεταχειρίζονται την 'γνώση', την οποία καλούνται να αξιοποιήσουν. Συγκεκριμένα τα ΤΝΔ λαμβάνουν πληροφορίες από κάποιο καθορισμένο περιβάλλον / δεδομένα, μέσω μιας διαδικασίας εκμάθησης (*training phase*). Επιπλέον αυτή η πληροφορία – γνώση, αποθηκεύεται στο ΤΝΔ μέσω των συνάψεων και των αντίστοιχων βαρών.

3.3 Βασικές Αρχές Τεχνητών Νευρωνικών Δικτύων

Σε αυτή την ενότητα γίνεται προσπάθεια να περιγραφούν όσο το δυνατόν πιο συνοπτικά μερικά βασικά χαρακτηριστικά της δομής, της λειτουργίας και της απόδοσης των ΤΝΔ. Εξαιτίας όμως της μεγάλης ανάπτυξης που παρατηρείται στον κλάδο αυτό (τόσο σε επίπεδο θεωρίας όσο και εφαρμογών), δεν κρίθηκε θεμιτό να παρουσιαστεί και να αναλυθεί κάθε πιθανή εκδοχή των χαρακτηριστικών αυτών. Αντ' αυτού, παρουσιάζονται οι πιο σύνηθες δομές, παράμετροι και αλγόριθμοι που εν γένει αποτελούν ένα ΤΝΔ

3.3.2 Αρχιτεκτονική ΤΝΔ

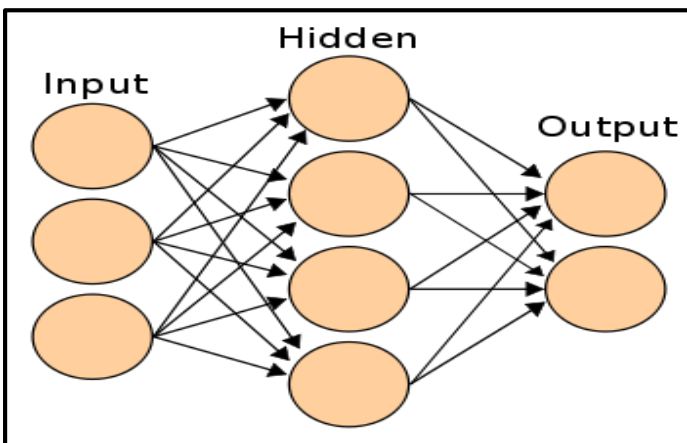
Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, οι τεχνητοί νευρώνες ενός ΤΝΔ οργανώνονται σε διάφορα επίπεδα παράλληλης διάταξης. Η αλληλουχία των παράλληλων αυτών επιπέδων είναι αυτή που αποτελεί και το σύνολο ενός ΤΝΔ. Φυσικά, σε ένα ΤΝΔ, δεν είναι όλα τα επίπεδα ίδια αλλά αντιθέτως διαφοροποιούνται ως εξής:

- Το επίπεδο εισόδου (*input layer*) αποτελεί το 1^ο επίπεδο ενός ΤΝΔ, σε αυτό εμφανίζονται τα δεδομένα εισόδου και ο αριθμός των νευρώνων του είναι ίσος με τις μεταβλητές των δεδομένων.
- Ένα κρυφό επίπεδο (*hidden layer*) συνδέει το προηγούμενο και το επόμενο κατά σειρά επίπεδο νευρώνων. Ο αριθμός των κρυφών επιπέδων που περιέχονται σε ένα ΤΝΔ, ορίζονται από τον σχεδιαστή – χρήστη και επηρεάζει το 'βάθος' και τη συνθετότητα του (μπορεί και κανένα). Ο αριθμός των νευρώνων δε, αυξάνεται πολλαπλασιαστικά ή και εκθετικά ανάλογα με τον αριθμό των κρυφών επιπέδων που επιλέγεται. Σημειώνεται πως ο αριθμός των νευρώνων σε κάθε κρυφό επίπεδο ορίζεται επίσης από τον σχεδιαστή.
- Το επίπεδο εξόδου (*output layer*) αποτελεί το τελευταίο επίπεδο ενός ΤΝΔ. Σε αυτά εμφανίζονται τα τελικά αποτελέσματα μετά το τέλος της εκμάθησης και ο αριθμός των νευρώνων του είναι ίσος με τις πιθανές μεταβλητές εξόδου των αποτελεσμάτων.

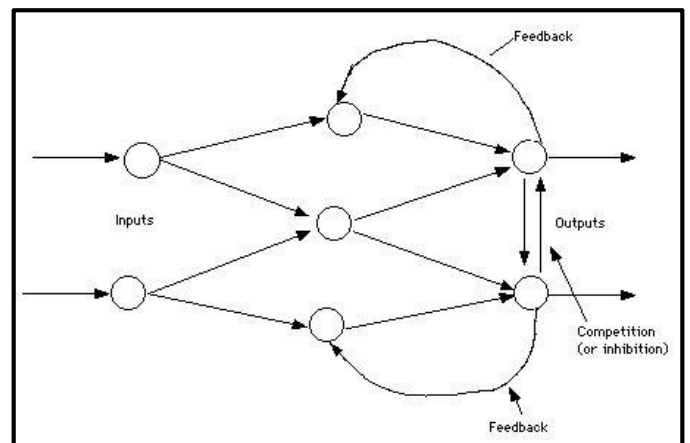
Π.χ. σε εφαρμογή ταξινόμησης υπερφασματικής εικόνας με 224 κανάλια και συνολικό αριθμό εικονοστοιχείων *Ground Truth* 10000, το επίπεδο εισόδου θα αποτελείται από 10000x224 νευρώνες (αφού το κάθε εικονοστοιχείο *Ground Truth* αποτελεί δεδομένο εκπαίδευσης και άρα μεταβλητή εισόδου), ενώ το επίπεδο εξόδου θα αποτελείται από τον αριθμό των κλάσεων *Ground Truth* στις οποίες είναι πιθανό να ταξινομηθεί κάθε εικονοστοιχείο.

Τα ΤΝΔ κατηγοριοποιούνται ανάλογα με την αρχιτεκτονική τους και τον τρόπο με τον οποίο συνδέονται οι νευρώνες μεταξύ τους:

- Σε νευρωνικά δίκτυα με πρόσθια τροφοδότηση (*feedforward*). Σε αυτή την περίπτωση το σήμα μεταφέρεται, μέσω των νευρώνων, από την είσοδο έως την έξοδο χωρίς όμως να μπορεί να συμβεί το αντίστροφο. Αυτά τα δίκτυα μπορούν με τη σειρά τους να διακριθούν σε μονοεπίπεδα δίκτυα (π.χ. *perceptron*) ή πολυεπίπεδα δίκτυα (π.χ. *multilayer perceptron*), ανάλογα με το εάν σε αυτά περιλαμβάνεται και κάποιο ή κάποια κρυφά επίπεδα (βλ. *Εικόνα 3.3*). Στο επόμενο υποκεφάλαιο θα αναλυθούν αυτές οι αρχιτεκτονικές, ως οι απλούστερες μορφές νευρωνικών δικτύων.
- Σε νευρωνικά δίκτυα με ανατροφοδότηση (*feedback*). Σε αυτή την περίπτωση τα νευρωνικά δίκτυα περιέχουν και τουλάχιστον ένα βρόγχο ανατροφοδότησης (*feedback loop*). Έτσι, τουλάχιστον ένα επίπεδο νευρώνων έχει το ρόλο του να τροφοδοτεί το σήμα εξόδου πίσω στις εισόδους όλων των άλλων νευρώνων (βλ. *Εικόνα 3.4*).



Εικόνα 3.3: Σχηματική αναπαράσταση απλού ΤΝΔ με ένα κρυφό επίπεδο και πρόσθια τροφοδότηση
[Πηγή: el.wikipedia.org]

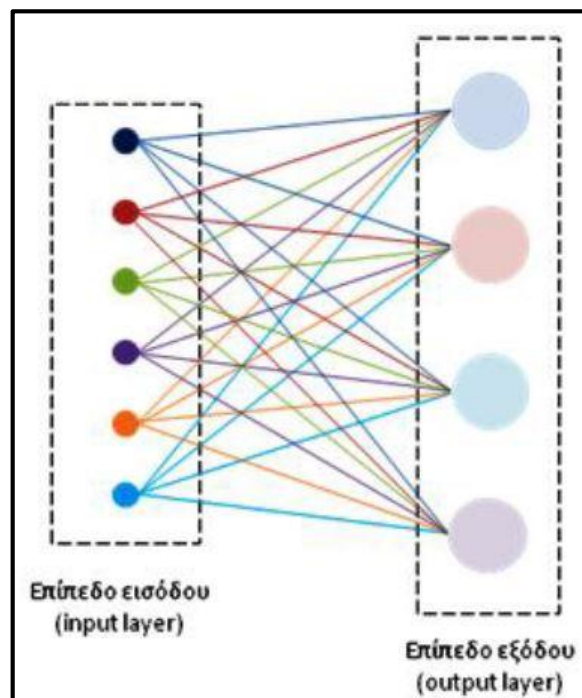


Εικόνα 3.4: Σχηματική αναπαράσταση απλού ΤΝΔ με ανατροφοδότηση σήματος
[Πηγή: www.psych.utoronto.ca]

Τέλος επισημαίνεται πως σε κάθε ΤΝΔ οι νευρώνες των γειτονικών επιμέρους επιπέδων μπορούν είτε να συνδέονται όλοι με όλους είτε να υπάρχουν πιθανές συνδέσεις που δεν υλοποιούνται. Κατά αυτό τον τρόπο, τα ΤΝΔ χαρακτηρίζονται ως πλήρως συνδεδεμένα (*fully connected*) και ως μερικώς συνδεδεμένα αντίστοιχα (*partially connected*).

3.3.2 Μοντέλο Αισθητήρα (Perceptron)

Το μοντέλο του αισθητήρα ήταν από τα πρώτα μοντέλα νευρωνικών δικτύων και αναπτύχθηκε ήδη από τη δεκαετία του 1950. Αυτό το μοντέλο αποτελείται από ένα επίπεδο εισόδου (*input layer*) και ένα επίπεδο εξόδου (*output layer*), με όλες τις επεξεργασίες να πραγματοποιούνται στο δεύτερο (βλ. *Εικόνα 3.5*). Το απλό αυτό μοντέλο χρησιμοποιεί ως συνάρτηση ενεργοποίησης τη βηματική συνάρτηση και συνήθως χρησιμοποιείται στο να διαχωρίζει ή να ταξινομεί ένα σύνολο δεδομένων σε 2 κλάσεις.



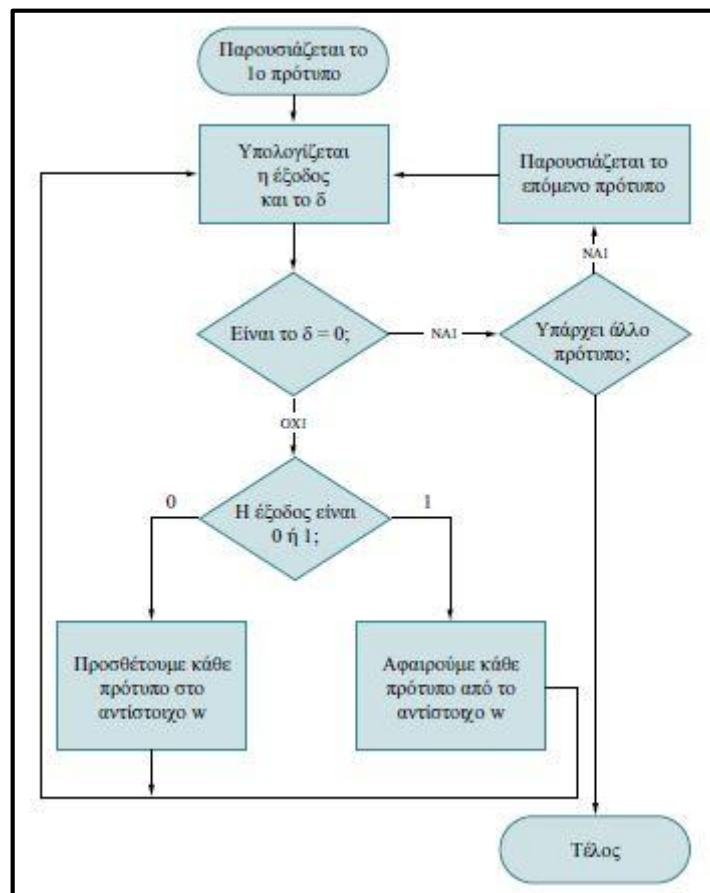
Εικόνα 3.5: Σχηματική αναπαράσταση ενός Αισθητήρα (perceptron) με 6 νευρώνες εισόδου και 4 νευρώνες εξόδου. [Πηγή: Αλεβιζάκου Ε.Γ. 2012]

Η διαδικασία εκπαίδευσης ενός αισθητήρα είναι εποπτευόμενη (ή επιβλεπόμενη), υλοποιείται με τα παραδείγματα εκπαίδευσης και αποτελούνται από τα σήματα εισόδου (*input signals*) και από το διάνυσμα των αντίστοιχων ‘επιθυμητών’ αποτελεσμάτων (*output signals*) τους. Ο σκοπός της εκπαίδευσης – εκμάθησης αυτής, είναι να υπολογιστούν οι κατάλληλες τιμές συναπτικών βαρών σε κάθε σύναψη, έτσι ώστε το δίκτυο να βρίσκει την ‘ορθή-επιθυμητή’ τιμή εξόδου για κάθε δεδομένη τιμή εισόδου.

Οι αλγόριθμοι και η διαδικασία της εκμάθησης των ΤΝΔ αναφέρονται αναλυτικά σε επόμενη ενότητα απλά εδώ αξίζει να σημειωθεί πως και στην απλή περίπτωση του αισθητήρα, η εκπαίδευση αποσκοπεί στην συνεχή και επαναληπτική διόρθωση των συναπτικών βαρών μέσω της ελαχιστοποίησης του σφάλματος-διαφοράς μεταξύ του επιθυμητού αποτελέσματος εξόδου και του πραγματικού αποτελέσματος σε μία δεδομένη επανάληψη. Στην περίπτωση του αισθητήρα και της ύπαρξης 2 πιθανών κλάσεων- αποτελεσμάτων, τα συναπτικά βάρη τροποποιούνται συνεχώς μέχρις ότου η κάθε επιμέρους διαφορά να γίνει 0. Αυτή η διαδικασία έχει καθιερωθεί να ονομάζεται Κανόνας Δέλτα (*Delta Rule*) ή κανόνας μάθησης *Widrow-Hoff*, με το Δέλτα (δ) να νοείται η διαφορά μεταξύ του επιθυμητού αποτελέσματος εξόδου και του πραγματικού αποτελέσματος. Η συνεχής τροποποίηση των συναπτικών βαρών υλοποιείται ως εξής:

- Στην περίπτωση που η πραγματική έξοδος είναι 0 ενώ η επιθυμητή 1 τότε προστίθεται η τιμή κάθε εισόδου στο αντίστοιχο συναπτικό βάρος
- Στην περίπτωση στην οποία η πραγματική έξοδος είναι 1 ενώ η επιθυμητή 0 αφαιρείται η τιμή κάθε εισόδου στο αντίστοιχο συναπτικό βάρος

Παρακάτω παρουσιάζεται σχηματικά η επαναληπτική διαδικασία εκμάθησης του αισθητήρα.



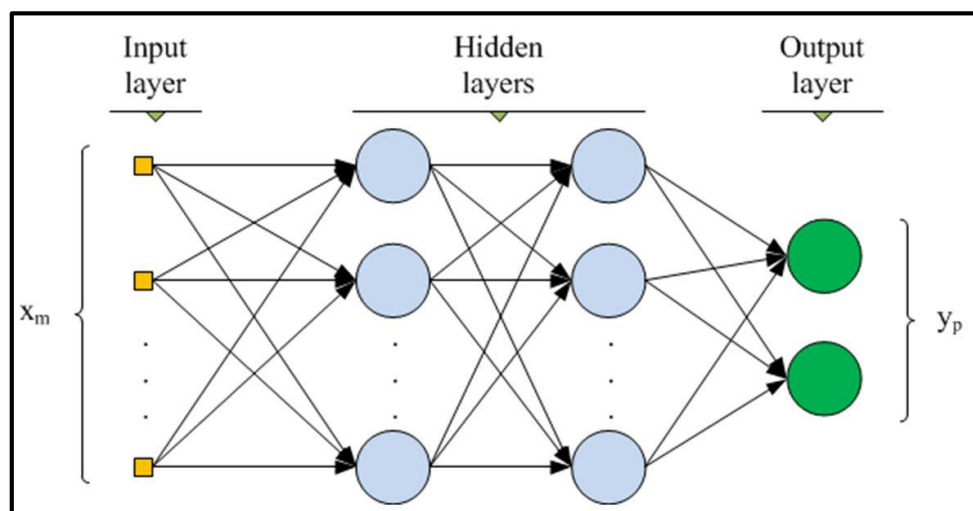
Εικόνα 3.6: Σχηματική αναπαράσταση αλγορίθμου εκπαίδευσης του αισθητήρα (perceptron) [Πηγή: Αργυράκης Π. 2001]

Το σημαντικό πρόβλημα αντιμετωπίζει ο μονοεπίπεδος πρόσθιας τροφοδότησης αισθητήρας, είναι ότι μπορεί να επιλύσει μόνο γραμμικώς διαχωρίσιμα προβλήματα. Για τον λόγο αυτόν έχουν δημιουργηθεί πιο περίπλοκες αρχιτεκτονικές, οι οποίες ανταποκρίνονται σε συνθετότερα προβλήματα, διατηρώντας όμως μερικές από τις παραπάνω αρχές που περιγράφηκαν.

Πολυεπίπεδος Αισθητήρας (Multilayer Perceptron - MLP)

Η κύρια διαφορά των αρχιτεκτονικών MLP με αυτών των απλών αισθητήρων, είναι η ύπαρξη ενός ή περισσότερων κρυφών επιπέδων. Αυτά τα επιπλέον επίπεδα νευρώνων, επιτρέπουν τη σχεδίαση και τη δημιουργία ΤΝΔ τα οποία ανταποκρίνονται σε πιο απαιτητικά προβλήματα αφού:

- Αυξάνεται κατά πολύ ο αριθμός των νευρώνων (και ο βαθμός ελευθερίας) του ΤΝΔ και άρα των συνάψεων που περιέχουν επιπλέον πληροφορία – βάρη
- Υπάρχει η δυνατότητα προσαρμογής συνθετότερων συναρτήσεων, για να μπορέσουν να περιγραφούν και στη συνέχεια να διαχωριστούν καλύτερα τα δεδομένα εισόδου. Η ροή του σήματος πραγματοποιείται πάλι από το επίπεδο εισόδου προς το επίπεδο εξόδου (*feedforward*), απλά υπάρχει η δυνατότητα εφαρμογής μη γραμμικής συνάρτησης ενεργοποίησης μεταξύ του επιπέδου εισόδου και του επόμενου κρυφού επιπέδου. Έτσι, η μετάβαση στο επίπεδο εξόδου μπορεί να επιτευχθεί με μία γραμμική συνάρτηση, όπως μία γραμμική παλινδρόμηση (*linear regression*), αφού πρώτα όμως έχει προηγηθεί ένας μη γραμμικός μετασχηματισμός σε προηγούμενο κρυφό επίπεδο.



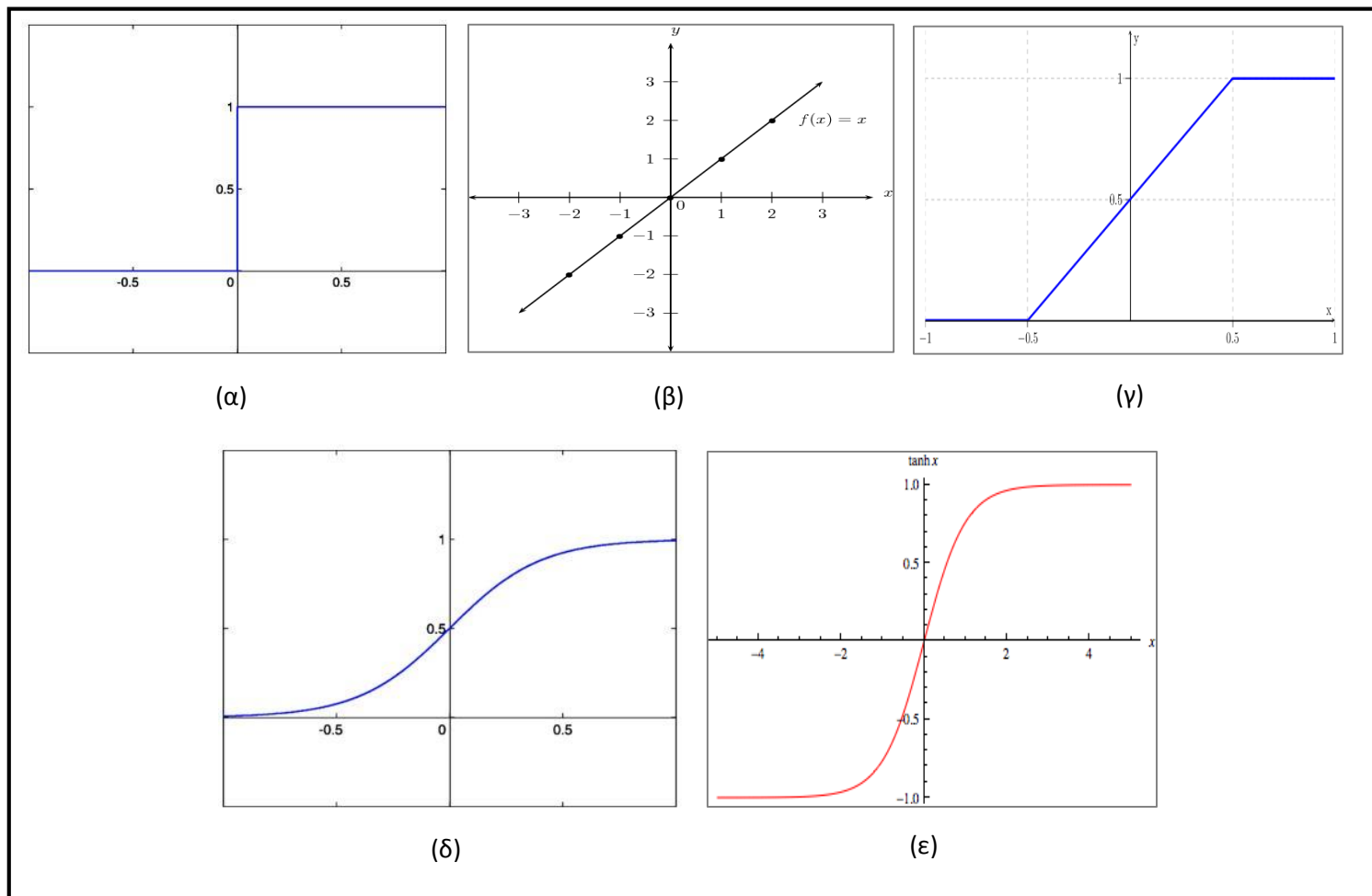
Εικόνα 3.7: Σχηματική αναπαράσταση ενός ΤΝΔ πολυεπίπεδου αισθητήρα (MLP) με δύο κρυφά επίπεδα [Πηγή: www.intechopen.com]

3.3.3 Συνάρτηση ενεργοποίησης

Υπάρχουν διάφορες συναρτήσεις ενεργοποίησης (*activation functions*) ή αλλιώς συναρτήσεις μεταφοράς (*transfer functions*) που χρησιμοποιούνται στα ΤΝΔ για τον μετασχηματισμό των συναπτικών βαρών. Φυσικά η επιλογή της όποιας συνάρτησης ενεργοποίησης εξαρτάται από την εφαρμογή στην οποία αποσκοπεί η χρήση ενός ΤΝΔ καθώς επίσης και στην αρχιτεκτονική του συγκεκριμένου δικτύου. Παρακάτω αναγράφονται μερικές συναρτήσεις ενεργοποίησης που συχνά συναντώνται σε ΤΝΔ:

- Συνάρτηση κατωφλίου (*threshold function*). Αυτή η συνάρτηση χρησιμοποιείται κυρίως στον διαχωρισμό των δεδομένων εισόδου σε 2 επιμέρους κατηγορίες. Στην περίπτωση που οι 2 αυτές κατηγορίες έχουν τιμές 0 και 1, τότε η συνάρτηση κατωφλίου ονομάζεται και βηματική συνάρτηση (*step function*) και αντίστοιχα ως συνάρτηση προσήμου (*signum function*) όταν οι τιμές είναι -1 και 1. Υπενθυμίζεται πως η βηματική συνάρτηση συναντάται στην απλή περίπτωση ενός μονοεπίπεδου αισθητήρα. Ακόμα, τονίζεται η ομοιότητα λειτουργίας αυτών των συναρτήσεων, με τον σχεδόν δυαδικό (*binary*) τρόπο λειτουργίας και εκπομπής σήματος στους βιολογικούς νευρώνες (βλ. Κεφ. 3.2.1)
- Γραμμική συνάρτηση (*Linear Function*). Αυτή η συνάρτηση χρησιμοποιείται στις περιπτώσεις εφαρμογής μιας γραμμικής παλινδρόμησης, στα επιμέρους συναπτικά βάρη των νευρώνων 2 επιπέδων.
- Ημι-γραμμική συνάρτηση (*Piecewise-Linear Function*). η Ημιγραμμική συνάρτηση είναι ένας συνδυασμός των 2 παραπάνω περιπτώσεων. Συγκεκριμένα, μέσα σε ένα δοσμένο διάστημα τιμών, λειτουργεί ως γραμμική συνάρτηση, ενώ για τις 'ακραίες' τιμές (εκτός δοσμένου διαστήματος) εκφυλίζεται σε συνάρτηση κατωφλίου.
- Σιγμοειδής συνάρτηση (*sigmoidal function*). Η σιγμοειδής συνάρτηση είναι ίσως η συνάρτηση που συναντάται πιο πολύ στα ΤΝΔ και μπορεί να χρησιμοποιηθεί αποτελεσματικά για να περιγράψει και να μετασχηματίσει μη γραμμικά φαινόμενα και δεδομένα. Συνήθως χρησιμοποιείται η λογιστική μορφή της.
- Εφαπτομενική (*tangent function*). Η εφαπτομενική είναι ένα ακόμα αξιόπιστο είδος μη γραμμικής συνάρτησης εφαρμογής που πολλές φορές χρησιμοποιείται αντί της σιγμοειδούς με τα ίδια ή αντίστοιχα αποτελέσματα.

Στην επόμενη σελίδα (βλ. *Εικόνα 3.8*), παρουσιάζονται οι γραφικές παραστάσεις των παραπάνω συναρτήσεων:



Εικόνα 3.8: Οι γραφικές παραστάσεις μερικών βασικών συναρτήσεων ενεργοποίησης. - (α) Η συνάρτηση κατωφλίου, (β) Η γραμμική συνάρτηση, (γ) Η ημιγραμμική συνάρτηση, (δ) Η λογιστική σιγμοειδής συνάρτηση, (ε) Η υπερβολική εφαπτομενική συνάρτηση [Πηγές: en.wikibooks.org, mathworld.wolfram.com]

3.3.4 Διαδικασία εκμάθησης (training phase)

Όπως αναφέρθηκε και παραπάνω για την περίπτωση του απλού αισθητήρα, έτσι και ένα οποιοδήποτε ΤΝΔ, όταν υποβάλλεται σε μια διαδικασία εκπαίδευσης ξεκινά από μία κατάσταση κατά την οποία δεν υπάρχει καμία γνώση πάνω στο πρόβλημα μελέτης. Στη συνέχεια όμως, κατά την διαδικασία εκμάθησης παρουσιάζονται τα διάφορα πρότυπα και παραδείγματα των δεδομένων με αποτέλεσμα το ΤΝΔ να εκλαμβάνει τη νέα αυτή πληροφορία ως 'γνώση', την οποία και αποθηκεύει στα συναπτικά του βάρη.

Η διαδικασία της εκμάθησης σε ένα ΤΝΔ αποτελεί τη σημαντικότερη και πιο περίπλοκη λειτουργία του. Ο τρόπος με τον οποίο επιτελείται η εκμάθηση σε ένα ΤΝΔ, καθορίζει την τελική ακρίβεια του αποτελέσματος καθώς επίσης και την χρονική του απόδοση. Η συνολική διαδικασία, επηρεάζεται από διάφορες παραμέτρους που παρακάτω θα γίνει προσπάθεια να περιγραφούν.

Μέθοδοι εκπαίδευσης

Τα διάφορα ΤΝΔ μπορούν να διαχωριστούν με βάση τον τρόπο με τον οποίο επιτελείτε η διαδικασία εκμάθησής τους. Δηλαδή, από τον τρόπο με τον οποίο το κάθε ΤΝΔ αντιλαμβάνεται τα διάφορα πρότυπα των δεδομένων και με το κατά πόσο του δίνονται κάποια 'ορθά' παραδείγματα αποτελέσματος. Αυτές οι μέθοδοι μπορούν να συνοψιστούν ως εξής:

- **Επιβλεπόμενη εκμάθηση (*supervised training*).** Κατά την επιβλεπόμενη εκμάθηση είναι απαραίτητο να τροφοδοτούνται στο ΤΝΔ κάποια 'πρότυπα εκπαίδευσης' τα οποία αποτελούν το επιθυμητό αποτέλεσμα για κάποια συγκεκριμένα δεδομένα. Με βάση αυτά τα γνωστά παραδείγματα (και τη διαφορά μεταξύ ορθής και πραγματικής εξόδου), το δίκτυο προσαρμόζει τις παραμέτρους του (συναπτικά βάρη) με στόχο τη σταδιακή μείωση του προηγούμενου σφάλματος. Η διαδικασία αυτή επαναλαμβάνεται έως ότου το σφάλμα αυτό μηδενιστεί, ή θεωρηθεί αποδεκτό. *Φυσικά, προϊόν επιβλεπόμενης εκμάθησης αποτελούν και οι εφαρμογές της συγκεκριμένης εργασίας πάνω στο πρόβλημα της επιβλεπόμενης ταξινόμησης εικόνας. Σε αυτές τις εφαρμογές δίνονται μερικά εικονοστοιχεία της προς ταξινόμηση εικόνας ως είσοδοι, μαζί με την τιμή της κλάσης στην οποία πρέπει να ταξινομηθούν. Έτσι, μέσα από την επαναληπτική διαδικασία το ΤΝΔ εκπαιδεύεται στο να ταξινομεί τα διάφορα επιμέρους εικονοστοιχεία στη σωστή κλάση κάθε φορά.*
- **Μη – Επιβλεπόμενη εκμάθηση (*unsupervised training*).** Στις εφαρμογές μη επιβλεπόμενης εκμάθησης ένα ΤΝΔ καλείται να προσεγγίσει ένα αποτέλεσμα το οποίο δεν μπορεί είναι γνωστό *a priori*, αφού δεν τροφοδοτούνται σε αυτό γνωστά παραδείγματα - αποτελέσματα. Έτσι, το ΤΝΔ προσπαθεί να συλλέξει πληροφορίες μέσα από στατιστικά μεγέθη και πρότυπα των δεδομένων, και με αυτά να οδηγηθεί στην αναπροσαρμογή των βαρών του. Τα ΤΝΔ σε αυτή τη μέθοδο, καλούνται 'κατηγοριοποιητές' και συνήθως χρησιμοποιούνται σε εφαρμογές για την κατάτμηση (*clustering*) ή την εκτίμηση πυκνότητας (*density estimation*) δεδομένων ή εικόνων.
- **Ενισχυμένη εκμάθηση (*reinforced training*).** Ούτε σε αυτή τη μέθοδο τροφοδοτείται το δίκτυο με επιθυμητά παραδείγματα – αποτελέσματα, αλλά αντιθέτως χρησιμοποιεί ένα σύνολο μέτρων επάρκειας του εξαγόμενου κάθε φορά αποτελέσματος το οποίο ονομάζεται ενισχυτικό σήμα (*reinforcement signal*). Το σήμα 'επιβραβεύει' τις ορθές συμπεριφορές του δικτύου, 'τιμωρεί' τις λανθασμένες και σταδιακά μεταβαίνει προς το επιθυμητό αποτέλεσμα.

Αλγόριθμος εκπαίδευσης

Κατά τη διαδικασία εκμάθησης και για την εκπαίδευση ενός ΤΝΔ, χρησιμοποιείται ένας αλγόριθμος διόρθωσης του σφάλματος που είναι γνωστός και ως Αλγόριθμος Σύγκλισης ή Αλγόριθμος Εκπαίδευσης. Ο αλγόριθμος εκπαίδευσης είναι αυτός που επί της ουσίας επικαθορίζει τον τρόπο με τον μεταβάλλονται οι παράμετροι (βάρη) του ΤΝΔ σε κάθε επανάληψη, καθώς επίσης και τον τρόπο με τον οποίο βελτιστοποιείται το παραγόμενο αποτέλεσμα. Παρακάτω αναγράφονται μερικοί βασικοί αλγόριθμοι εκπαίδευσης ενώ θα περιγραφεί αναλυτικότερα ο αλγόριθμος οπισθοδιάδοσης του σφάλματος, ο οποίος είναι και ο συνηθέστερος στις εφαρμογές ΤΝΔ. Οι βασικοί αυτοί αλγόριθμοι είναι:

- Ο αλγόριθμος εκπαίδευσης Hebb
- Ο αλγόριθμος εκπαίδευσης Levenberg - Marquardt
- Ο αλγόριθμος ανταγωνιστικής εκμάθησης
- Ο αλγόριθμος οπισθοδιάδοσης σφάλματος (*backpropagation error-backprop*)

3.3.5 Αλγόριθμος Οπισθοδιάδοσης Σφάλματος (*backpropagation error*)

Για πολλά χρόνια δεν υπήρχε ο αλγόριθμος εκπαίδευσης εκείνος που να μπορεί να ανταποκριθεί στις απαιτήσεις εκμάθησης πολυεπίπεδων ΤΝΔ και άρα σε πιο περίπλοκα προβλήματα και εφαρμογές. Πάνω σε αυτή τη λογική, αναπτύχθηκε ο αλγόριθμος οπισθοδιάδοσης σφάλματος, ο οποίος σήμερα χρησιμοποιείται κατά κόρον στις εφαρμογές επιβλεπόμενης εκμάθησης και ειδικά στα ΤΝΔ πρόσθιας τροφοδότησης πολλών επιπέδων (MLP). Μάλιστα η συχνή έως και δεδομένη χρήση του σε αυτού του είδους τα ΤΝΔ, τα κάνει να αναφέρονται πολλές φορές στη βιβλιογραφία και ως δίκτυα οπισθοδιάδοσης (*backprop networks*).

Τα 2 βασικά κριτήρια που πρέπει να πληρούνται για τη χρήση αυτής της μεθόδου είναι α) Το ΤΝΔ να περιέχει ένα ή και παραπάνω κρυφά επίπεδα, σε κάθε περίπτωση πάντως όχι μόνο επίπεδα εισόδου και εξόδου (μοντέλο αισθητήρα) και β) η συνάρτηση ενεργοποίησης να είναι μία μη γραμμική συνάρτηση (συνήθως η σιγμοειδής ή η εφαπτομενική). Σε γενικές γραμμές τα βασικά βήματα που ακολουθούνται για την εκμάθηση ενός ΤΝΔ με αυτή τη μέθοδο είναι τα ακόλουθα :

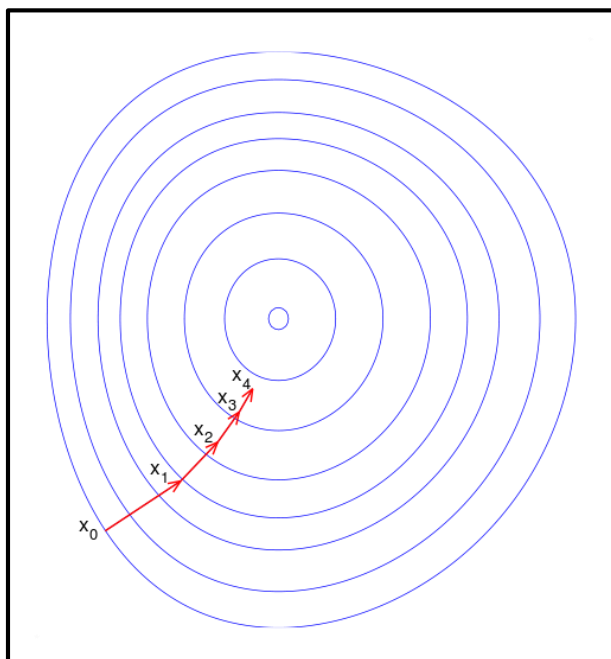
- Η διαδικασία της μάθησης ξεκινά με τις τιμές των συναπτικών βαρών να θέτονται τυχαία
- Εφόσον το αποτέλεσμα εξάγει σφάλμα (το πιο πιθανό στην 1^η επανάληψη), τα βάρη μετασχηματίζονται ώστε να ελαχιστοποιηθεί το λάθος – σφάλμα
- Η διαδικασία αυτή επαναλαμβάνεται κυκλικά πολλές φορές έως ότου το παραγόμενο σφάλμα να μηδενιστεί ή να θεωρηθεί ανεκτό

Είναι σημαντικό πάντως να επισημανθεί η διακριτότητα των βημάτων σε αυτή τη μέθοδο. Συγκεκριμένα, ο αλγόριθμος οπισθοδιάδοσης του σφάλματος αποτελείται από δύο βασικές λειτουργίες, τη μεταφορά του σήματος προς τα εμπρός (*forward pass*) και τη μεταφορά του σήματος προς τα πίσω (*backward pass*).

Στην 1^η περίπτωση το σήμα μεταβαίνει από το επίπεδο εισόδου προς το επίπεδο εξόδου μέσω όλων των πιθανών επιπέδων και νευρώνων (ανάλογα τις συνδέσεις βεβαίως). Από αυτή τη μεταφορά σήματος εξάγεται και η όποια απόκριση – έξοδος του δικτύου. Εξαιτίας όμως της διαφοράς, μεταξύ της επιθυμητής και της πραγματικής εξόδου από το δίκτυο, το σήμα διαδίδεται προς την αντίθετη κατεύθυνση (2^η περίπτωση) προκειμένου να τροποποιήσει τα συναπτικά βάρη με στόχο την ελαχιστοποίηση του σφάλματος. Η διαδικασία της ελαχιστοποίησης του σφάλματος γίνεται με την μαθηματική μέθοδο της επικλινούς καθόδου (*gradient descent*) που παρουσιάζεται παρακάτω.

Μέθοδος επικλινούς καθόδου (*gradient descent*)

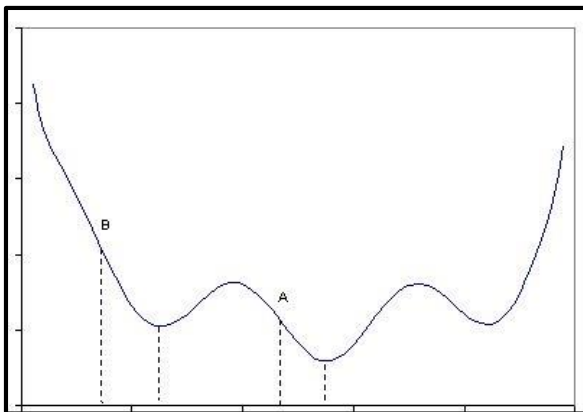
Η μέθοδος επικλινούς καθόδου, είναι μία απλή τεχνική η οποία χρησιμοποιείται για την ελαχιστοποίηση ενός μεγέθους (συνήθως του σφάλματος στις εφαρμογές των ΤΝΔ). Η μέθοδος αυτή επί της ουσίας μας επιτρέπει να κάνουμε μικρά επαναλαμβανόμενα και σταδιακά βήματα ώστε να βρεθεί το ολικό ελάχιστο μιας συνάρτησης. Αυτή η συνάρτηση συνήθως επιλέγεται να είναι μία συνάρτηση κόστους η οποία αντιπροσωπεύει την αθροισμένη πιθανότητα του να μην εξάγεται το επιθυμητό αποτέλεσμα (*Negative Log-likelihood - NLL*) σαν απόκριση από το νευρωνικό δίκτυο. Είναι εύκολα αντιληπτό ότι το παραπάνω μπορεί να μεταφραστεί ως η ‘ελαχιστοποίηση του σφάλματος’ και άρα το ζητούμενο για την διαδικασία εκμάθησης ενός ΤΝΔ.



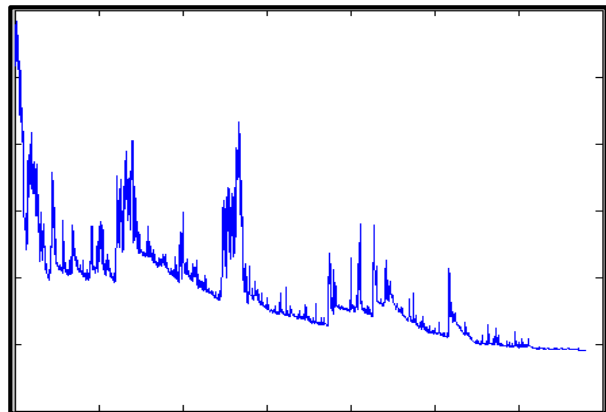
Εικόνα 3.9: Σχηματική αναπαράσταση της διαδικασίας επικλινούς καθόδου. Οι μπλε κύκλοι αναπαριστούν ίδια τιμή σφάλματος, ενώ η κόκκινη γραμμή την πορεία που ακολουθεί ο αλγόριθμος σε κάθε επιμέρους επανάληψη μεταβαίνοντας σε μικρότερο σφάλμα κάθε φορά ($x_0 - x_4$)

Εδώ πρέπει να τονιστεί πως η καμπύλη σφάλματος (όπως περιγράφεται παραπάνω - NLL) δεν είναι ποτέ κοίλη. Αυτό σημαίνει πως υπάρχει η πιθανότητα ο αλγόριθμος να βρεθεί και να 'παγιδευτεί' σε ένα τοπικό ελάχιστο και όχι στο ολικό ελάχιστο που είναι και το ζητούμενο. Αυτό το πρόβλημα συνήθως επιλύεται με τη χρήση στατιστικών μεθόδων εκπαίδευσης ή την επανεκκίνηση της διαδικασίας εκμάθησης με νέα αρχικοποίηση. Μία παράμετρος που επηρεάζει την ταχύτητα εκμάθησης όσο και το μέγεθος του τελικού σφάλματος είναι ο ρυθμός μάθησης (*learning rate*) ο οποίος παίζει και το ρόλο του βήματος πάνω στην καμπύλη σφάλματος.

Όταν η τιμή του ρυθμού μάθησης είναι μικρή, τότε η μεταβολή των βαρών σε κάθε επανάληψη είναι επίσης μικρή, το παραγόμενο σφάλμα αλλάζει λίγο και η διαδικασία της εκμάθησης αργεί περισσότερο αφού η εύρεση του τοπικού ελαχίστου καθυστερεί. Αντιθέτως, όταν η τιμή του ρυθμού μάθησης είναι μεγάλη, τα συναπτικά βάρη αλλάζουν σε μεγάλο βαθμό και υπάρχει ο κίνδυνος να μην υπάρχει επί της ουσίας 'κάθοδος' στην καμπύλη αφού η πορεία του αλγορίθμου θα εκφυλίζεται σε ταλάντωση. Σημειώνεται πάντως πως μερικές φορές και αναλόγως την εφαρμογή, η επιλογή μεγαλύτερου ρυθμού μάθησης μπορεί να 'ξεκολλήσει' την πορεία ενός αλγορίθμου από ένα τοπικό ελάχιστο το οποίο και δεν θα υπερπηδούσε με μικρότερο ρυθμό. Σε κάθε περίπτωση πάντως η πιο σωστή και ιδανική λύση για συγκεκριμένα δεδομένα και εφαρμογές, βρίσκεται με *trial-and-error*, δηλαδή με πολλαπλές δοκιμές μέχρις ότου βρούμε την ιδανική τιμή.



Εικόνα 3.10: Ενδεικτική καμπύλη κόστους με 2 τοπικά ελάχιστα και 1 τοπικό μέγιστο [Πηγή: www.kcl.ac.uk]



Εικόνα 3.11: Ενδεικτική καμπύλη ελαχιστοποίησης σφάλματος, με τη μέθοδο Stochastic Gradient Descent [Πηγή: en.wikipedia.org]

Η **μέθοδος Στοχαστικής Επικλινούς Καθόδου (Stochastic Gradient Descent - SGD)**, είναι μία παραλλαγή της παραπάνω μεθόδου. Γενικά, λειτουργεί με της ίδιες αρχές ελαχιστοποίησης του σφάλματος, με τη διαφορά πως η όλη διαδικασία πρόβλεψης του σφάλματος γίνεται πολύ γρηγορότερα. Στη βασική μορφή της μεθόδου SGD, κάτι τέτοιο επιτυγχάνεται με τη χρήση μόλις ενός δείγματος εκπαίδευσης για την κάθε επιμέρους πρόβλεψη της κλίσης.

Η παραλλαγή της μεθόδου που χρησιμοποιεί όχι ένα αλλά λίγο περισσότερα δείγματα για την πρόβλεψη της κλίσης, ονομάζεται **μέθοδος Στοχαστικής Επικλινούς Καθόδου με μικρές δέσμες (Minibatch Stochastic Gradient Descent - MSGD)**. Το σύνολο των δειγμάτων αυτών ονομάζεται *'minibatches'*. Η μέθοδος αυτή έχει αποδειχθεί ότι παράγει μικρότερες μεταβλητότητες στην κάθε επιμέρους πρόβλεψη κλίσης ενώ πετυχαίνει και καλύτερη χρήση της ιεραρχικής μνήμης στους σύγχρονους Η/Υ. Τέλος, σημειώνεται πως αυτή η μέθοδος είναι που προτείνεται για αρχιτεκτονικές ΤΝΔ τύπου Deep Learning.

Εδώ κρίνεται απαραίτητο να γίνει αναφορά στον συγκεκριμένο τρόπο με τον οποίο επιτελείται η επαναληπτική διαδικασία, στην παραπάνω μέθοδο, για τον επαναπροσδιορισμό των σφαλμάτων. Συγκεκριμένα, η επαναληπτική διαδικασία που ακολουθείται είναι η εκπαίδευση κατά εποχές (*batch training*). Στην τεχνική αυτή τα συναπτικά βάρη αλλάζουν όταν έχουν εφαρμοστεί όλα τα διανύσματα εισόδου – δεδομένα εκπαίδευσης και η παρουσίαση τους από μία φορά στο δίκτυο ονομάζεται εποχή (*epoch*). Προφανώς στην περίπτωση της MSGD η κάθε εποχή προσδιορίζεται με τον ορισμένο αριθμό δειγμάτων – *Minibatch*. Σημειώνεται πάντως, πως για διαφορετικούς αλγόριθμους και διαφορετικές εφαρμογές υπάρχουν και άλλες μέθοδοι επαναληπτικής διαδικασίας: η συνεχής εκπαίδευση (*incremental training*), η εκπαίδευση κατά τη διάρκεια λειτουργίας του συστήματος (*online training*), και η εκπαίδευση στο εργαστήριο (*offline training*). [Αλεβιζάκου Ε.Γ. 2012].

Κανονικοποίηση της συνάρτησης κόστους (loss function)

Πολλές φορές κατά τον ορισμό της συνάρτησης NLL για την ελαχιστοποίηση του σφάλματος καθορίζονται και διάφοροι παράμετροι με σκοπό την κανονικοποίηση της. Η διαδικασία αυτή πραγματοποιείται με σκοπό την αποφυγή του προβλήματος της υπερπροσαρμογής (*overfitting*). Το πρόβλημα αυτό δημιουργείται με την υπέρμετρη προσαρμογή της συνάρτησης (και άρα των βαρών) πάνω στα δεδομένα με τα οποία εκπαιδεύεται ο αλγόριθμος τη κάθε δεδομένη επανάληψη. Επί της ουσίας ο αλγόριθμος *'υπερεκπαιδεύεται'* σε κάποια δεδομένα εκπαίδευσης, δίνοντας έτσι λιγότερη βαρύτητα στα επόμενα που θα ακολουθήσουν.

Το πρόβλημα αυτό επιλύεται με τη χρήση 2 συντελεστών - όρων κανονικοποίησης, οι οποίοι συμβολίζονται L1 και L2.

Κριτήρια τερματισμού

Πιο πάνω έχει περιγραφεί ο τρόπος με τον οποίο υλοποιείται η διαδικασία εκμάθησης με τον αλγόριθμο οπισθοδιάδοσης σφάλματος κατά την οποία ο αλγόριθμος εκπαιδεύεται στο να οδηγείται σε ένα 'επιθυμητό' αποτέλεσμα. Εδώ θα προσδιοριστεί πότε και πώς τερματίζεται η διαδικασία αυτή, καθώς επίσης και το πώς ορίζεται το 'επιθυμητό' αποτέλεσμα κάθε εφαρμογής. Συγκεκριμένα παρουσιάζονται 3 κριτήρια τερματισμού:

- Πρόωρη παύση (*early stopping*). Για να γίνει κατανοητός ο τερματισμός του αλγορίθμου με πρόωρη παύση, πρέπει πρώτα να γίνει αναφορά στο σύνολο δεδομένων για επικύρωση (*validation set*). Ο ακριβής τρόπος και σκοπός με τον οποίο χωρίζονται τα δεδομένα εκπαίδευσης θα αναλυθεί στην επόμενη ενότητα, αλλά εδώ αξίζει να αναφερθεί ότι κατά τη διαδικασία της εκπαίδευσης κάποια δείγματα δεν συμμετέχουν σε αυτή ούτε στη μετέπειτα διαδικασία ελέγχου για την εξαγωγή σφάλματος. Αυτά τα δείγματα συγκροτούν το σύνολο επικύρωσης (*validation set*) τα οποία είναι εύκολο να θεωρούνται και ως 'μελλοντικά δείγματα ελέγχου'. Με τα δεδομένα επικύρωσης υπολογίζεται ένα σφάλμα επικύρωσης (*validation error*) σε κάθε εποχή το οποίο όμως είναι διαφορετικό από το σφάλμα εκπαίδευσης. Έτσι υπάρχει πιθανότητα σε κάποιο στάδιο της επαναληπτικής διαδικασίας το σφάλμα εκπαίδευσης να μειώνεται σε κάθε εποχή, ενώ το σφάλμα επικύρωσης να αυξάνεται. Σε αυτή την περίπτωση ο αλγόριθμος αντιμετωπίζει πρόβλημα υπερεκπαίδευσης (*overfitting*) και η λύση δίνεται με το να τερματιστεί η διαδικασία εκπαίδευσης. Ο τερματισμός αυτός ονομάζεται πρόωρη παύση.
- Σύγκλιση αλγορίθμου. Θεωρείται ότι υπάρχει σύγκλιση του αλγορίθμου όταν σε ένα προσδιορισμένο αριθμό εποχών δεν συναντάται μεγάλη μεταβολή στο σφάλμα. Η μεταβολή αυτή φυσικά παραμετροποιείται. Αντιστοίχως, σύγκλιση συναντάται όταν δεν υπάρχει σημαντική μεταβολή στις παραμέτρους του ΤΝΔ (συναπτικά βάρη, πόλωση).
- Περαιώση στο προκαθορισμένο όριο επαναλήψεων. Προς αποφυγή της άσκοπης παράτασης στην εκπαιδευτική διαδικασία όταν το σφάλμα μεταβάλλεται λίγο ή καθόλου, πολλές φορές τίθεται ένα προκαθορισμένο όριο επαναλήψεων, στο πέρας των οποίων ο αλγόριθμος τερματίζεται. Οι επαναλήψεις αυτές αφορούν τις εποχές (*epochs*) του αλγορίθμου, όπως περιγράφονται πιο πάνω.

Σύνολα δεδομένων εκπαίδευσης

Κατά τη διαδικασία της εκμάθησης, δεν χρησιμοποιούνται όλα τα δεδομένα Ground Truth για την άμεση εκπαίδευση του αλγορίθμου αλλά χωρίζονται σε 3 διαφορετικά σύνολα. Αυτά είναι:

- Το σύνολο εκπαίδευσης (*training set*). Τα δείγματα του συνόλου εκπαίδευσης χρησιμοποιούνται για την εκπαίδευση του αλγορίθμου και για τον προσδιορισμό των παραμέτρων σε κάθε επανάληψη. Τα δείγματα αυτά συγκροτούν δέσμες (*minibatches*) από τις οποίες προσδιορίζεται η κλίση της συνάρτησης κόστους και το σφάλμα της κάθε επανάληψης
- Το σύνολο επικύρωσης (*validation set*). Όπως περιγράφηκαν και στην προηγούμενη παράγραφο, με τα δείγματα του συνόλου επικύρωσης υπολογίζεται ένα σφάλμα επικύρωσης με το οποίο εξετάζεται η διαδικασία της εκμάθησης, ως προς το πότε θα τερματιστεί και άρα ποιο θα είναι το τελικό μοντέλο παραμέτρων.
- Το σύνολο ελέγχου (*test set*). Τα δείγματα του συνόλου δεν χρησιμοποιούνται καθόλου κατά τη διαδικασία της εκμάθησης και προσδιορισμού παραμέτρων και έτσι διατηρούνται ως ανεξάρτητα και αντιπροσωπευτικά δείγματα για τον έλεγχο της απόδοσης του αλγορίθμου. Έτσι, όταν τερματίζεται η διαδικασία εκπαίδευσης τότε τα δείγματα του συνόλου ελέγχου χρησιμοποιούνται στον προσδιορισμό του τελικού σφάλματος του αλγορίθμου.

Σημειώνεται πως τις περισσότερες φορές το σύνολο εκπαίδευσης (*training set*) είναι μεγαλύτερο από τα άλλα 2, καθώς είναι και το σημαντικότερο για τη σωστή εκπαίδευση του αλγορίθμου. Κάποια τυπικά βιβλιογραφικά ποσοστά επί του συνόλου των δεδομένων εκπαίδευσης είναι: *training set* - 70% , *validation set* - 15%, *test set* - 15% .

Φυσικά, σε μερικές εφαρμογές όπως είναι οι τηλεπισκοπικές ταξινομήσεις εικόνας, το ζητούμενο βρίσκεται και στον έλεγχο των αποδόσεων με εκπαίδευση του αλγορίθμου από λίγα δείγματα. Σε αυτές τις περιπτώσεις το σύνολο εκπαίδευσης (*training set*) ορίζεται και σε πολύ μικρότερα ποσοστά, της τάξης του 1% -5%.

3.4 Συνελικτικά Νευρωνικά Δίκτυα (CNN)

Τα τελευταία χρόνια έχει επιτευχθεί μεγάλη πρόοδος στην αντιμετώπιση προβλημάτων που αφορούν την αναγνώριση προτύπων και άλλων ζητημάτων της όρασης υπολογιστών. Σημαντική βοήθεια στην πρόοδο αυτή, έχουν δώσει και τα Συνελικτικά Νευρωνικά Δίκτυα (*Convolutional Neural Networks - CNN*).

Η σχεδίαση των νευρωνικών δικτύων CNN εμπνεύστηκε από τη βιολογία και συγκεκριμένα από τον οπτικό φλοιό τη γάτας ο οποίος διαθέτει μία περίπλοκη διάταξη κυττάρων. Η διάταξη αυτή επιβάλλει την σύνδεση κάποιων νευρωνικών κυττάρων που είναι ευαίσθητα σε συγκεκριμένες υποπεριοχές του ορατού πεδίου, τα οποία και λειτουργούν ως τοπικά 'φίλτρα' στις υποπεριοχές αυτές. Με αυτό τον τρόπο, η γάτα εκμεταλλεύεται τη μεγάλη χωρική συσχέτιση των διάφορων υποπεριοχών, με τελικό αποτέλεσμα την καλύτερη αντίληψη του χώρου.

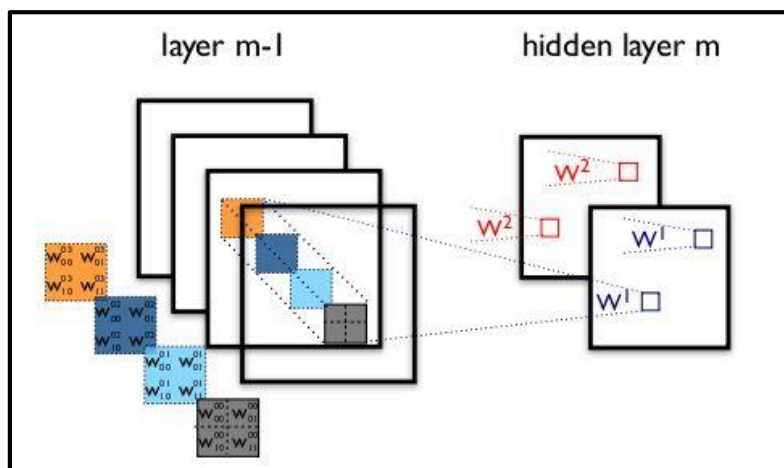
Στις προηγούμενες ενότητες αναλύθηκε η γενικότερη φιλοσοφία που διέπει τα ΤΝΔ, καθώς επίσης δόθηκε ιδιαίτερη σημασία στον τρόπο με τον οποίο λειτουργούν και εκπαιδεύονται τα δίκτυα MLP. Ουσιαστικά, τα δίκτυα CNN αποτελούν παραλλαγές των δικτύων MLP καθώς τα τελευταία επίπεδα νευρώνων ενός CNN συνήθως αποτελούν ένα δίκτυο MLP. Η διαφοροποίηση έγκειται στα επίπεδα νευρώνων που προηγούνται του MLP, που όπως και στο παραπάνω παράδειγμα του οπτικού φλοιού της γάτας, μπορούμε να τα αντιληφθούμε ως φίλτρα σε συγκεκριμένα υποσύνολα των δεδομένων μελέτης. Εξάλλου, στη βασικότερή τους μορφή, τα δίκτυα CNN μπορούν να περιγραφούν ως δίκτυα που χρησιμοποιούν πολλά αντίγραφα των ίδιων νευρώνων. Αυτό δίνει τη δυνατότητα της υπολογιστικής επεξεργασίας μεγάλων μοντέλων και δεδομένων, κρατώντας παράλληλα τον αριθμό των παραμέτρων προς προσδιορισμό σχετικά μικρό.

Τα επίπεδα νευρώνων που συνήθως χρησιμοποιούνται και προσδιορίζουν τα δίκτυα CNN χωρίζονται στις παρακάτω 2 βασικές κατηγορίες:

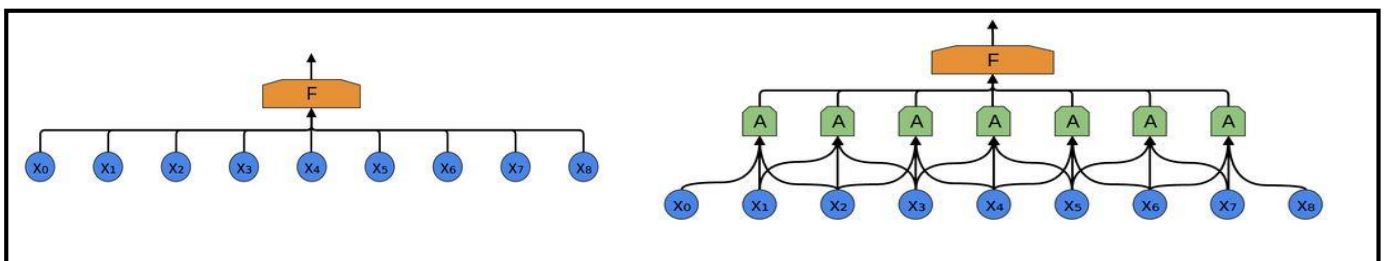
- Τα Συνελικτικά Επίπεδα (*Convolutional Layers*). Τα συνελικτικά επίπεδα αποτελούνται από ένα πλέγμα νευρώνων. Κάθε συνελικτικό επίπεδο αποτελεί είσοδο (*input*) από ένα πλέγμα του προηγούμενου επιπέδου του οποίου τα βάρη παραμένουν ίδια για όλους τους επόμενους νευρώνες του δικτύου. Κατά αυτό τον τρόπο γίνεται κατανοητό ότι το κάθε επίπεδο λειτουργεί συνελικτικά ως προς το προηγούμενο επίπεδο, με το μέγεθος του εφαρμοζόμενου πλέγματος να είναι πιθανόν διαφορετικό σε καθένα από αυτά. Στις περιπτώσεις δισδιάστατων δεδομένων, τα πλέγματα πρέπει να είναι τετραγωνικά, για τρισδιάστατα δεδομένα κυβικά κ.ο.κ.

- Τα Επίπεδα Μέγιστης Συγκέντρωσης (*Max-Pooling Layers*). Τα επίπεδα μέγιστης συγκέντρωσης τοποθετούνται συνήθως μετά από συνελκτικά επίπεδα. Γενικότερα τα επίπεδα συγκέντρωσης δέχονται ως είσοδο ένα πλέγμα από το προηγούμενο συνελκτικό επίπεδο και έχουν ως έξοδο μία μοναδική τιμή που μπορεί να είναι η μέγιστη τιμή του πλέγματος, η μέση τιμή, μία γραμμική παλινδρόμηση κ.α. Τα επίπεδα που επιστρέφουν την μέγιστη τιμή ονομάζονται επίπεδα μέγιστης συγκέντρωσης (*Max-Pooling Layers*).

Φυσικά μετά την όποια αρχιτεκτονική και αλληλουχία συνελκτικών ή μέγιστης συγκέντρωσης επιπέδων, το σύνολο των νευρώνων του τελευταίου επιπέδου γίνεται είσοδος στην ακόλουθη διάταξη MLP, ή οποιαδήποτε άλλη διάταξη πρόσθιας τροφοδότησης.

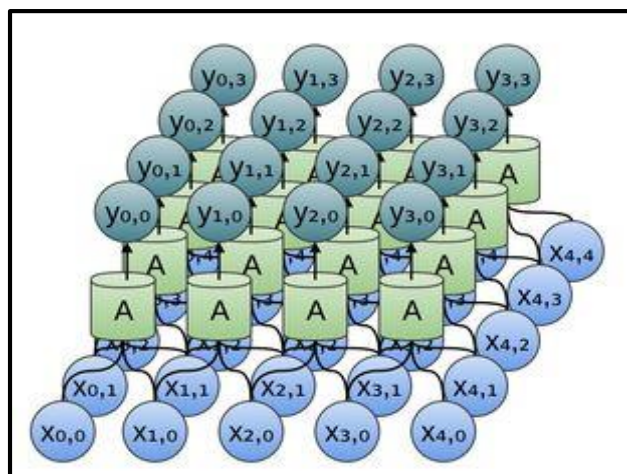


Εικόνα 3.12: Σχηματική αναπαράσταση της διάταξης δύο διαδοχικών συνελκτικών επιπέδων [Πηγή: <http://deeplearning.net>]



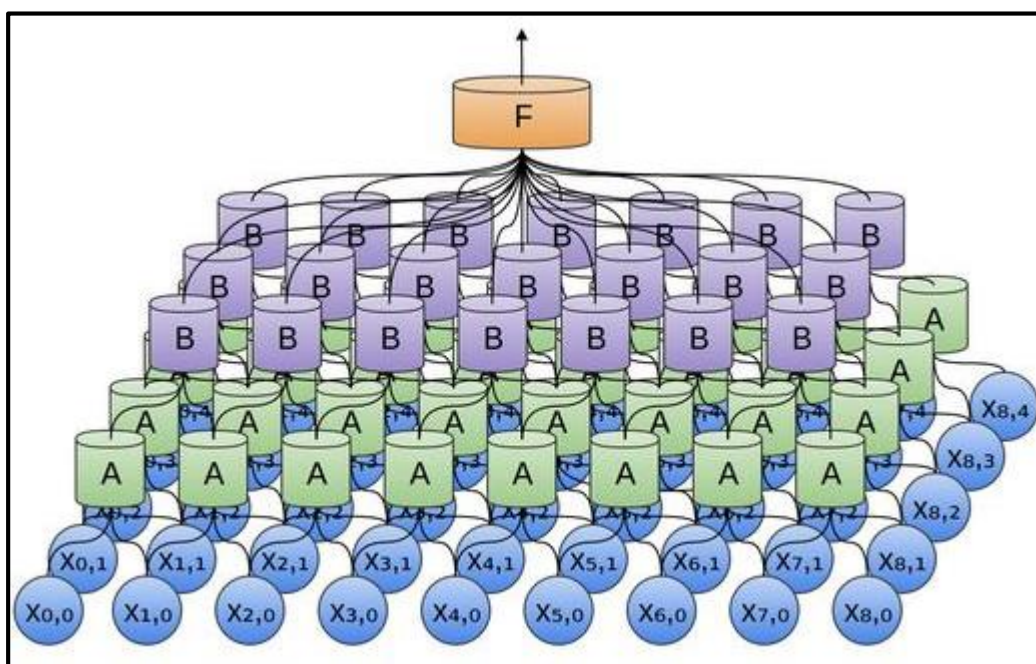
Εικόνα 3.13: Σχηματική αναπαράσταση της διάδοσης ενός σήματος μέσα σε ένα ΤΝΔ. Αριστερά κατευθείαν από τους νευρώνες σε μία συνάρτηση ενεργοποίησης για έξοδο, ενώ δεξιά με τη μεσολάβηση ενός συνελκτικού επιπέδου με πλέγμα τριών νευρώνων [Πηγή: <http://colah.github.io>]

Ειδικότερα, όταν το πρόβλημα προς αντιμετώπιση αφορά δισδιάστατα δεδομένα ή εικόνες (αναγνώριση προτύπων, ανάλυση βίντεο, ταξινομήσεις κ.α.), τότε τα συνελκτικά επίπεδα αλλά και τα πλέγματα – συνελκτικά φίλτρα – που εφαρμόζονται είναι και αυτά δισδιάστατα. Για παράδειγμα, σε μία περίπτωση τοποθέτησης συνελκτικού επιπέδου σε ένα τμήμα εικόνας μεγέθους 5x5, με πλέγμα 2x2 τότε το αποτέλεσμα εξόδου θα είναι το αποτέλεσμα της παραπάνω συνέλιξης μεγέθους 4x4 (βλ. Εικόνα 3.14).



Εικόνα 3.14: Το αποτέλεσμα ενός συνελκτικού επιπέδου με πλέγμα 2x2 σε ένα τμήμα εικόνας 5x5 [Πηγή: <http://colah.github.io>]

Τέλος παρακάτω (Εικόνα 3.15) παρουσιάζεται σχηματικά, το παράδειγμα εφαρμογής 2 διαδοχικών συνελκτικών επιπέδων με ίδιο πλέγμα (2x2) σε μία εικόνα 9x5.

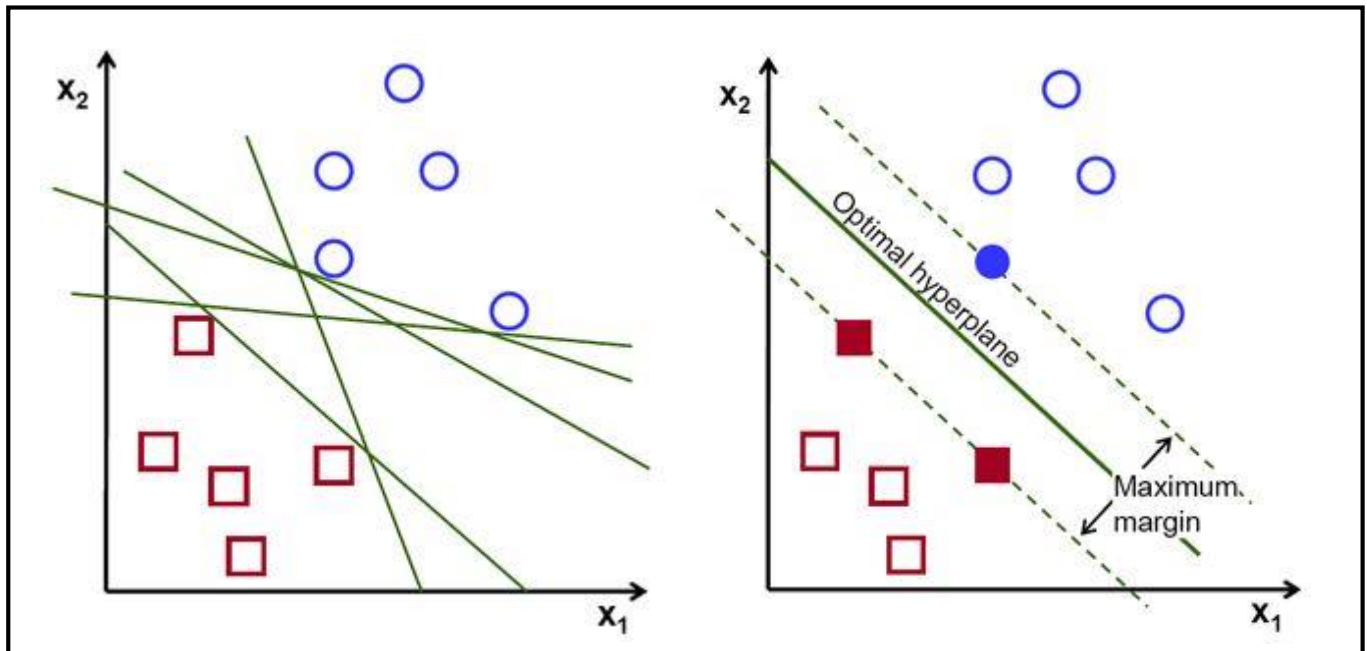


Εικόνα 3.15: Το αποτέλεσμα εφαρμογής δύο διαδοχικών συνελκτικών επιπέδων με ίδιο πλέγμα (2x2) σε μία εικόνα 9x5 [Πηγή: <http://colah.github.io>]

3.5 Μηχανές Διανυσματικής Υποστήριξης (SVM)

Παρακάτω ακολουθεί μία σύντομη περιγραφή των βασικών χαρακτηριστικών των Μηχανών Διανυσματικής Υποστήριξης. Κάτι τέτοιο άλλωστε κρίνεται απαραίτητο καθώς οι ταξινομήσεις που ακολουθούν στα επόμενα κεφάλαια, υλοποιούνται και με τη χρήση SVM για την καλύτερη σύγκριση και μελέτη της απόδοσης του δικτύου CNN.

Οι Μηχανές Διανυσματικής Υποστήριξης είναι μοντέλα επιβλεπόμενης εκμάθησης τα οποία χρησιμοποιούνται για εφαρμογές ανάλυσης δεδομένων, αναγνώρισης προτύπων, ταξινόμησης κ.α. Το βασικό χαρακτηριστικό των SVM είναι ότι εκπαιδεύονται στο να μπορούν να διαχωρίσουν τα διάφορα δεδομένα σε 2 κλάσεις, αφού πρώτα φυσικά έχουν δοθεί κάποια δεδομένα εκπαίδευσης. Συγκεκριμένα, σε μία αναπαράσταση των δειγμάτων εκπαίδευσης στο χώρο ένα μοντέλο SVM προσπαθεί να βρει και να επιλέξει το βέλτιστο 'όριο απόφασης' (*decision boundary*) μεταξύ των δεδομένων, που θα καθορίζει και το σε ποια κλάση ανήκει το κάθε επιμέρους δείγμα. Το όριο αυτό ορίζεται από τη γραμμή εκείνη που διχοτομεί το μεγαλύτερο πιθανό κενό μεταξύ των δεδομένων των 2 κατηγοριών (βλ. *Εικόνα 16*). Η απόσταση μεταξύ του ορίου απόφασης και του επιπέδου δεδομένων της μιας κατηγορίας ονομάζεται περιθώριο (*margin*), ενώ τα δείγματα κάθε κατηγορίας που βρίσκονται πλησιέστερα στο όριο απόφασης ονομάζονται διανύσματα υποστήριξης (*support vectors*).



Εικόνα 3.16: Αριστερά: Οι διάφορες πιθανές γραμμές που μπορούν να χωρίσουν τα δεδομένα σε 2 κατηγορίες. Δεξιά: Η υλοποίηση του ορίου απόφασης ως η διχοτόμος του μέγιστου δυνατού κενού μεταξύ των δεδομένων [Πηγή: <http://docs.opencv.org>]

Έτσι, κάθε νέο δείγμα που εισέρχεται στο μοντέλο, προβάλλεται στον χώρο και ανάλογα με το σε ποια πλευρά του 'κενού' ανήκει, ταξινομείται σε μία από τις δύο πιθανές κατηγορίες.

Περιπτώσεις πολλών κλάσεων

Βεβαίως, στις περισσότερες εφαρμογές ταξινόμησης και γενικότερα ανάλυσης δεδομένων, το ζητούμενο είναι ο διαχωρισμός – ταξινόμηση των διάφορων δειγμάτων σε περισσότερες από 2 υποψήφια κλάσεις. Έχουν προταθεί πολλές μέθοδοι για την αντιμετώπιση αυτού του προβλήματος με τις βασικότερες να είναι:

- Η δημιουργία νέων επιμέρους δυαδικών ταξινομητών. Σε αυτή τη περίπτωση, η κάθε κλάση εξετάζεται δυαδικά είτε μία προς όλες τις άλλες (*one-versus-all*), είτε μία προς μία (*one-versus-one*).
- Τα καθοδηγούμενα μη-κυκλικά γραφήματα (*Directed acyclic graphs SVM - DAGSVM*)
- Οι κώδικες διόρθωσης σφάλματος εξόδου.
- Η μέθοδος *Crammer and Singer* που ανάγει το πρόβλημα σε βελτιστοποίηση.

Οι μέθοδοι που εφαρμόζονται συνηθέστερα είναι αυτές της δημιουργίας νέων επιμέρους δυαδικών ταξινομητών και η μέθοδος *Crammer and Singer*.

Περιπτώσεις μη γραμμικών ταξινομήσεων

Εκτός όμως από τις εφαρμογές γραμμικών ταξινομήσεων, οι SVM μπορούν να αποδώσουν αποτελεσματικά και σε μη γραμμικές ταξινομήσεις χρησιμοποιώντας το λεγόμενο τέχνασμα του πυρήνα (*kernel trick*). Με αυτή τη τεχνική υιοθετείται η φιλοσοφία της απεικόνισης των διανυσμάτων χαρακτηριστικών σε ένα χώρο μεγαλύτερης διάστασης, όπου με μεγάλη πιθανότητα, αναμένεται οι υποψήφια κλάσεις να είναι γραμμικώς διαχωρίσιμες. Με αυτή η λογική που προκύπτει από το θεώρημα Cover, τα εσωτερικά γινόμενα στον υψηλότερης διάστασης χώρο, μπορούν να υπολογιστούν αξιοποιώντας τη σχετική συνάρτηση πυρήνα που λειτουργεί στον αρχικό, χαμηλής διάστασης χώρο. Χρησιμοποιούνται διάφορες συναρτήσεις πυρήνα, με τις βασικότερες να είναι:

- Οι πολυωνυμικές συναρτήσεις (ομοιογενής και ανομοιογενής)
- Η Γκαουσιανή ακτινική συνάρτηση βάσης (*radial basis function - RBF*)
- Η υπερβολική εφαπτομενική

3.6 Αλγόριθμοι εφαρμογής

Σε αυτό το υποκεφάλαιο γίνεται αναφορά στους 3 αλγορίθμους που επιλέχθηκαν για την υλοποίηση ταξινομήσεων υπερφασματικών εικόνων, που αποτελούν και τον βασικό σκοπό της παρούσας εργασίας. Αυτές οι 3 προσεγγίσεις, αντλούν το θεωρητικό τους υπόβαθρο και τα διάφορα ειδικά χαρακτηριστικά τους από όλες τις παραπάνω ενότητες του *Κεφαλαίου 3* και είναι:

- Ένα Συνελικτικό Νευρωνικό Δίκτυο (*CNN*)
- Μία Γραμμική Μηχανή Διανυσματικής Υποστήριξης (*Linear SVM*)
- Μία Μη Γραμμική Μηχανή Διανυσματικής Υποστήριξης, με τη χρήση της ακτινικής συνάρτησης βάσης (*RBF SVM*)

3.6.1 Μείωσης διάστασης δεδομένων (*Dimensionality Reduction*)

Πριν γίνει αναφορά στην αρχιτεκτονική και στα διάφορα χαρακτηριστικά των αλγορίθμων εφαρμογής, πρέπει να γίνει απαραίτητη αναφορά στη διαδικασία μείωσης της διάστασης των δεδομένων. Σε πολλές εφαρμογές που περιέχουν υπερφασματικά, πολυφασματικά ή οποιαδήποτε δεδομένα με μεγάλη 3^η διάσταση, πολλές φορές επιβάλλεται μία μείωση της διάστασης αυτής προκειμένου να μειώνεται το μέγεθος τους και άρα ο υπολογιστικός φόρτος του συστήματος επεξεργασίας. Παράλληλα με μία τέτοια μείωση βεβαίως, πρέπει να εξασφαλίζεται η ύπαρξη όσο το δυνατόν μεγαλύτερης πληροφορίας – σε αυτή την 3^η διάσταση - στα εναπομείναντα δεδομένα.

Η πιο συνήθης μεθοδολογία μείωσης της 3ης διάστασης των δεδομένων υπερφασματικών εικόνων, είναι η Ανάλυση Κυρίων Συνιστωσών (***Principal Component Analysis - PCA***). Η ανάλυση σε κύριες συνιστώσες είναι μία διαδικασία που χρησιμοποιεί έναν ορθογωνικό μετασχηματισμό για να μετατρέψει ένα σύνολο συσχετιζόμενων δεδομένων, σε ένα σύνολο ασυσχέτιστων γραμμικών μεταβλητών. Οι νέες μεταβλητές αυτές ονομάζονται Κύριες Συνιστώσες (***Principal Components***) και τις περισσότερες φορές διατηρούν πολύ υψηλά ποσοστά (99% και πλέον) της πληροφορίας των αρχικών δεδομένων, ανάλογα φυσικά και με τον επιλεγμένο αριθμό τους. Ο αριθμός των Κυρίων Συνιστωσών σε κάθε εφαρμογή ορίζεται από τον χρήστη αλλά σε κάθε περίπτωση είναι ίσος ή μικρότερος του αριθμού των αρχικών δεδομένων.

Στις εφαρμογές ανάλυσης υπερφασματικών ή πολυφασματικών δεδομένων η ανάλυση PCA επιβάλλεται στην 3^η διάσταση της υπερφασματικής ή πολυφασματικής πληροφορίας αντίστοιχα. Με αυτό τον τρόπο είναι δυνατόν να μειωθεί κατά πολύ ο αριθμός της διάστασης αυτής χωρίς όμως να χάνεται μεγάλο ποσοστό πληροφορίας. Ο ακριβής ορισμός του ποσοστού της διατηρημένης πληροφορίας και ο αντίστοιχος αριθμός των Κυρίων Συνιστωσών κάθε εφαρμογής αναφέρεται στο *Κεφάλαιο 5.2.1* .

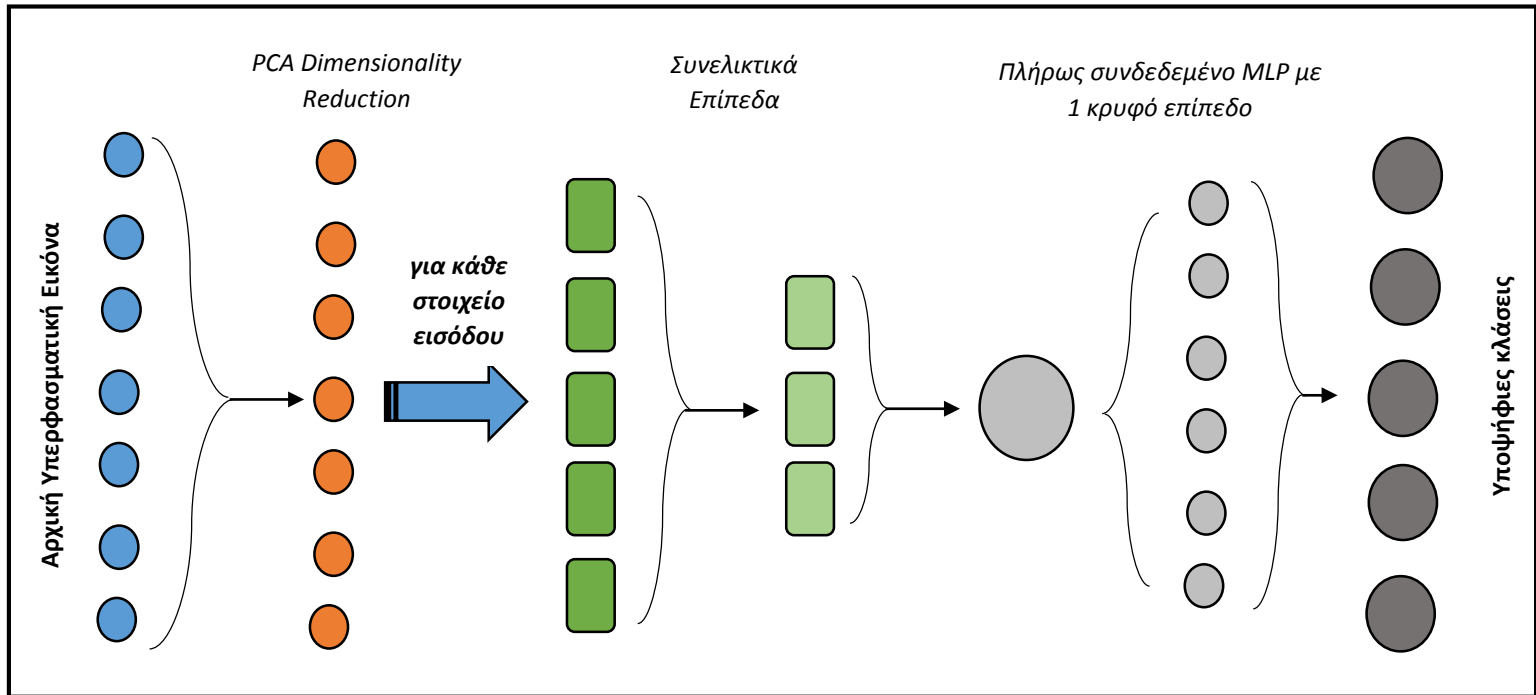
3.6.1 Το Συνελικτικό Νευρωνικό Δίκτυο (CNN) εφαρμογής

Όσον αφορά το Συνελικτικό ΤΝΔ που δημιουργήθηκε, αποτελείται από δύο διαδοχικά συνελικτικά επίπεδα (με πλέγμα 3x3 και τα δύο) τα οποία εφαρμόζονται σε παράθυρο εικόνας 5x5. Έπειτα το αποτέλεσμα της συνέλιξης αποτελεί 1 στοιχείο (όσον αφορά τις 2 διαστάσεις) το οποίο και εισέρχεται σε ένα πλήρως συνδεδεμένο δίκτυο MLP με ένα κρυφό επίπεδο (Εικόνα 3.17). Στην υπερφασματική πληροφορία δε (3^η διάσταση) αρχικά επιβλήθηκε μείωση της διάστασής της με ανάλυση PCA, ενώ ο τελικός αριθμός των κυριών συνιστωσών (components) ποικίλει στις ακόλουθες εφαρμογές. Έπειτα μέσω των διαδοχικών συνελικτικών επιπέδων, ορίστηκαν δύο συντελεστές (c_1 , c_2) οι οποίοι και καθορίζουν το τελικό μέγεθος της πληροφορίας που εισέρχεται στο ακόλουθο δίκτυο MLP. Τα διάφορα επιπλέον μοναδικά χαρακτηριστικά του δικτύου και του αλγορίθμου εκπαίδευσης παρουσιάζονται στον ακόλουθο πίνακα:

Συνελικτικά επίπεδα	Αριθμός Συνελικτικών επιπέδων	2
	Αρχικό παράθυρο εφαρμογής	5x5
	Πλέγμα (Φίλτρο) Συνελικτικών επιπέδων	3x3
	Συντελεστές 3 ^η διάστασης	$c_1=3*$ αριθμό Κ.Σ. $c_2=9*$ αριθμό Κ.Σ.
MLP	Αριθμός κρυφών επιπέδων	1
	Αριθμός νευρώνων κρυφού επιπέδου	6 * αριθμό Κ.Σ.
Training phase	Αλγόριθμος εκπαίδευση	backpropagation error
	Συνάρτηση ενεργοποίησης	Υπερβολική εφαπτομενική
	Μέθοδος ελαχιστοποίησης σφάλματος	MSGD
	Μέγεθος minibatch	20
	Ρυθμός εκμάθησης	0,3
Κριτήρια Τερματισμού	Μέγιστος αριθμός εποχών	300
	Κατώφλι προόδου	0,005%

Πίνακας 3.1: Χαρακτηριστικά στοιχεία και παράμετροι, του Συνελικτικού ΤΝΔ εφαρμογής

Τέλος, στην Εικόνα 3.17 παρουσιάζεται σχηματικά (όσο το δυνατόν πληρέστερα σε 2 διαστάσεις) το κάθε επιμέρους επίπεδο του Συνελικτικού ΤΝΔ, με τα αντίστοιχα αναγραφόμενα χαρακτηριστικά του:



Εικόνα 3.17: Σχηματική αναπαράσταση του Συνελικτικού ΤΝΔ εφαρμογής

3.6.2 Οι αλγόριθμοι SVM εφαρμογής

Όσον αφορά τους ταξινομητές εφαρμογής SVM, χρησιμοποιήθηκαν οι: *Linear SVM* και *RBF SVM* (βλ. Κεφ. 3.5). Όπως και στην περίπτωση του ΤΝΔ πριν την διαδικασία εκμάθησης, επιβλήθηκε μείωση στη διάσταση των δεδομένων εισόδου με την ανάλυση PCA. Τα αποτελέσματα που εξήχθησαν από την χρήση αυτών των ταξινομητών χρησιμοποιήθηκαν ως μέτρο σύγκρισης και αξιολόγησης της απόδοσης του συνελικτικού ΤΝΔ.

Σημειώνεται πως και οι 3 μέθοδοι ταξινόμησης υλοποιήθηκαν σε περιβάλλον ελεύθερου μη – εμπορικού λογισμικού, με τη χρήση βιβλιοθηκών ανοικτού κώδικα στην προγραμματιστική γλώσσα της Python.

4 ΔΕΔΟΜΕΝΑ ΕΛΕΓΧΟΥ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΑ

Στο παρών κεφάλαιο παρουσιάζονται αναλυτικά όλα τα δεδομένα που χρησιμοποιήθηκαν κατά την εκπόνηση της εργασίας. Επιπλέον, αναφέρονται όλα τα βήματα που ακολουθήθηκαν για την δημιουργία, την απόκτηση και την επεξεργασία αυτών των δεδομένων προκειμένου να καταστούν συμβατά με τις εφαρμογές της εργασίας.

Τα δεδομένα αυτά χωρίζονται σε 2 βασικές κατηγορίες:

- Στην 1^η κατηγορία (βλ. 4.1) εντάσσονται διάφορα υπερφασματικά δεδομένα εικόνων, τα οποία αποτελούν αντικείμενο μελέτης, εφαρμογής και σύγκρισης στη διεθνή βιβλιογραφία. Το σύνολο των δεδομένων αυτών θα αναφέρεται παρακάτω και ως 'Βιβλιογραφικά Δεδομένα'.
- Στην 2^η κατηγορία (βλ. 4.2) εντάσσονται διάφορα υπερφασματικά δεδομένα εικόνων, τα οποία και δημιουργήθηκαν αποκλειστικά για τους σκοπούς αυτής της εργασίας. Τα δεδομένα αυτά δημιουργήθηκαν προκειμένου να επιτευχθεί μεγάλη ποικιλία εφαρμογών για τις ταξινομήσεις που υλοποιήθηκαν. Το σύνολο των δεδομένων αυτών θα αναφέρεται παρακάτω και ως 'SFB Δεδομένα' (San Francisco Bay Datasets) εξαιτίας της περιοχής που αποτυπώνουν.

Σημειώνεται, πως τα όποια στάδια επεξεργασίας ακολουθήθηκαν περιγράφονται αναλυτικά στο *Κεφάλαιο 4.2*, καθώς αυτά τα στάδια εφαρμόστηκαν ως επί το πλείστον στα δεδομένα SFB.

4.2 Βιβλιογραφικά Δεδομένα

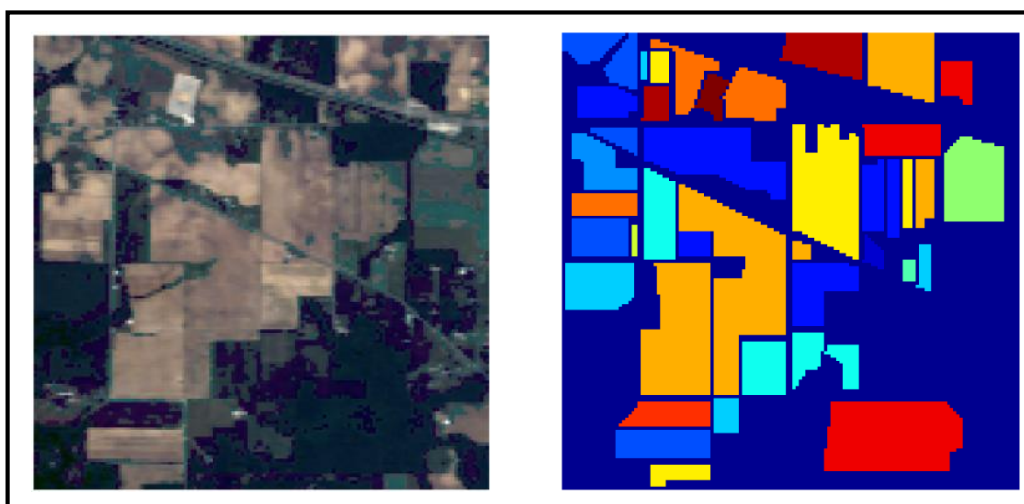
Τα βιβλιογραφικά δεδομένα, αποτελούνται συνολικά από 7 υπερφασματικές εικόνες και δεδομένα Ground Truth οι οποίες διατίθενται ελεύθερα στο διαδίκτυο. Σημειώνεται πως για 2 εικόνες (*Moffet Field Run 1* και *Moffet Field Run 3*) δεν ήταν διαθέσιμα αντίστοιχα δεδομένα Ground Truth και για αυτό χρειάστηκε να ψηφιοποιηθούν για τους σκοπούς της εργασίας.

Είναι σημαντικό να αναφερθεί πως στα δεδομένα που παρουσιάζονται σε αυτό το κεφάλαιο επιβλήθηκε ελάχιστη ή και μηδενική προεπεξεργασία καθώς κάτι τέτοιο θα τα καθιστούσε βιβλιογραφικά ασυνεπή. Έτσι, οι όποιες 'παρεμβάσεις' πάνω στις εικόνες, αφορούν εξαγωγή κάποιων καναλιών, τα οποία και υποδεικνύονται.

Παρακάτω παρουσιάζονται όλα τα βιβλιογραφικά δεδομένα που αποκτήθηκαν για τους σκοπούς της εργασίας μαζί και με τα χαρακτηριστικά στοιχεία τους.

Indian Pines

Τα υπερφασματικά δεδομένα *Indian Pines* όπως αποκτήθηκαν από το διαδίκτυο: [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral Remote Sensing Scen](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes)
[es](#) περιλαμβάνουν μία υπερφασματική εικόνα και μία εικόνα Ground Truth με 16 κλάσεις. Η υπερφασματική εικόνα είναι διαθέσιμη είτε με 224 κανάλια (το σύνολο των καναλιών AVIRIS) είτε με 200, έχοντας αφαιρέσει 24 κανάλια {[104-108], [150-163], 220} εξαιτίας της υψηλής απορρόφησης νερού. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=28, G=17, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις.



Εικόνα 4.1: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα Indian

Η υπερφασματική εικόνα αποτυπώνει αγροτική περιοχή της βορειοδυτικής Indiana των ΗΠΑ, με τα 2/3 της να καταλαμβάνουν καλλιέργειες και το 1/3 δάσος ή πολυετής φυσική βλάστηση. Επιπλέον σε αυτή την εικόνα εντοπίζονται δύο κεντρικές οδικές αρτηρίες, μία σιδηροδρομική γραμμή, μερικά μικρά κτίσματα καθώς και μικρότεροι δρόμοι. Ακόμα σημειώνεται πως καθώς η λήψη έχει παρθεί τον Ιούνιο, κάποιες καλλιέργειες είναι στα πρώτα στάδια της ανάπτυξής τους με λιγότερη από 5% εδαφοκάλυψη.

Όνομα Δεδομένων	Indian Pines
Μέγεθος Εικόνας	145x145
Δέκτης Λήψης	AVIRIS
Κάλυψη Ηλεκτρομαγνητικού φάσματος	400–2500 nm
Αριθμός Καναλιών	224 / 200
Αριθμός Κλάσεων Ground Truth	16

Πίνακας 4.1: Χαρακτηριστικά στοιχεία δεδομένων Indian Pines

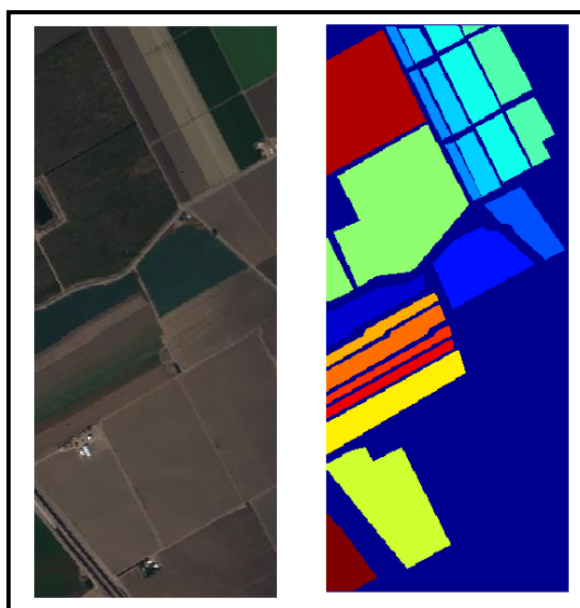
Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 16 κλάσεις, οι οποίες περιλαμβάνουν κυρίως καλλιέργειες αλλά και μερικά είδη βλάστησης και πρασίνου. Παρακάτω παρουσιάζονται οι 16 κλάσεις μαζί και με τα δείγματα (samples) που περιλαμβάνει η καθεμιά (Πίνακας 4.2.) .

#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης	#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης
1	Alfalfa	46	9	Oats	20
2	Corn-notill	1428	10	Soybean-notill	972
3	Corn-mintill	830	11	Soybean-mintill	2455
4	Corn	237	12	Soybean-clean	593
5	Grass-pasture	483	13	Wheat	205
6	Grass-trees	730	14	Woods	1265
7	Grass-pasture-mowed	28	15	Buildings-Grass-Trees-Drives	386
8	Hay-windrowed	478	16	Stone-Steel-Towers	93

Πίνακας 4.2: Οι 16 κλάσεις των δεδομένων Indian Pines

Salinas

Τα υπερφασματικά δεδομένα *Salinas* όπως αποκτήθηκαν από το διαδίκτυο: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes περιλαμβάνουν μία υπερφασματική εικόνα και μία εικόνα Ground Truth με 16 κλάσεις. Η υπερφασματική εικόνα είναι διαθέσιμη είτε με 224 κανάλια (το σύνολο των καναλιών AVIRIS) είτε με 204, έχοντας αφαιρέσει 20 κανάλια [108-112], [154-167], 224} εξαιτίας της υψηλής απορρόφησης νερού. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=28, G=17, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις.



Εικόνα 4.2: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα *Salinas*

Η υπερφασματική εικόνα αποτυπώνει καλλιεργήσιμη περιοχή στην πεδιάδα *Salinas* της California των ΗΠΑ και χαρακτηρίζεται από την υψηλή χωρική ανάλυση της (3,7 μ. ανά pixel).

Όνομα Δεδομένων	<i>Salinas</i>
Μέγεθος Εικόνας	512x217
Δέκτης Λήψης	AVIRIS
Κάλυψη Ηλεκτρομαγνητικού φάσματος	400–2500 nm
Αριθμός Καναλιών	224 / 204
Αριθμός Κλάσεων Ground Truth	16

Πίνακας 4.3: Χαρακτηριστικά στοιχεία δεδομένων *Salinas*

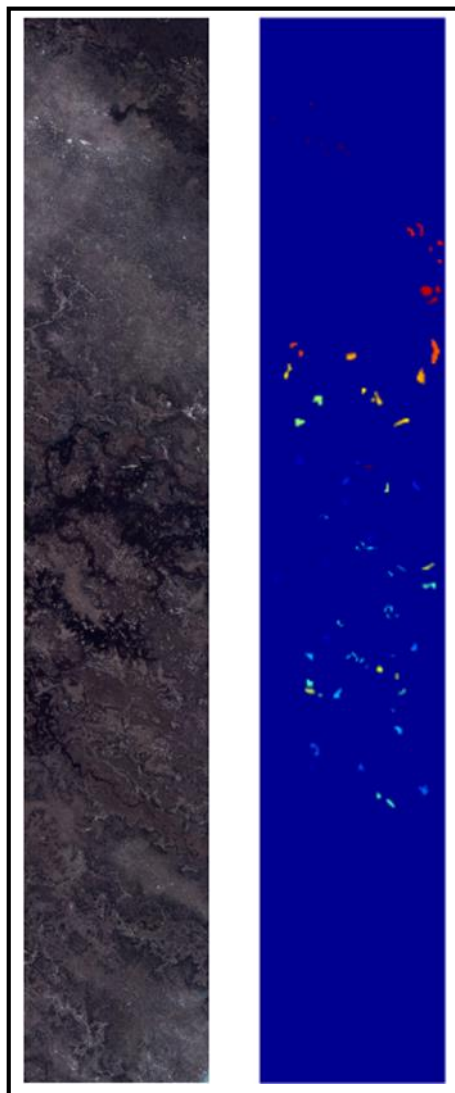
Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 16 κλάσεις, οι οποίες περιλαμβάνουν κάποια είδη καλλιεργειών της περιοχής σε διάφορα στάδια της ανάπτυξής τους. Παρακάτω παρουσιάζονται οι 16 κλάσεις μαζί και με τα δείγματα (samples) που περιλαμβάνει η καθεμιά (Πίνακας 4.4).

#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης	#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης
1	Brocoli_green_weeds_1	2009	9	Soil_vinyard_develop	6203
2	Brocoli_green_weeds_2	3726	10	Corn_senesced_green_weeds	3278
3	Fallow	1976	11	Lettuce_romaine_4wk	1068
4	Fallow_rough_plow	1394	12	Lettuce_romaine_5wk	1927
5	Fallow_smooth	2678	13	Lettuce_romaine_6wk	916
6	Stubble	3959	14	Lettuce_romaine_7wk	1070
7	Celery	3579	15	Vinyard_untrained	7268
8	Grapes_untrained	11271	16	Vinyard_vertical_trellis	1807

Πίνακας 4.4: Οι 16 κλάσεις των δεδομένων Salinas

Botswana

Τα υπερφασματικά δεδομένα *Botswana* όπως αποκτήθηκαν από το διαδίκτυο: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes περιλαμβάνουν μία υπερφασματική εικόνα και μία εικόνα Ground Truth με 14 κλάσεις. Η υπερφασματική εικόνα συλλέχθηκε στις 31/5/2001 από τον δορυφόρο EO-1 της NASA, αποτυπώνει περιοχή του δέλτα Okavango στην Botswana και έχει υποστεί προεπεξεργασία από το *UT Center for Space Research* ώστε να μετριάσουν οι επιπτώσεις των λανθασμένων σημάτων από τους ανιχνευτές (ανωμαλίες σήματος, λανθασμένη βαθμονόμηση κ.λ.π.) . Ακόμα, αφαιρέθηκαν 97 κανάλια (242 αρχικά) εξαιτίας μη βαθμονόμησης ή θορύβου και έτσι απέμειναν τα ακόλουθα 145: [10-55, 82-97, 102-119, 134-164, 187-220]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=19, G=7, B=3) καθώς και η Ground Truth εικόνα με τις 14 κλάσεις.



Εικόνα 4.3: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα Botswana

Όνομα Δεδομένων	Botswana
Μέγεθος Εικόνας	1476x256
Δέκτης Λήψης	HYPERION
Κάλυψη Ηλεκτρομαγνητικού φάσματος	400–2500 nm
Αριθμός Καναλιών	145
Αριθμός Κλάσεων Ground Truth	14

Πίνακας 4.5: Χαρακτηριστικά στοιχεία δεδομένων Botswana

Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 14 κλάσεις, οι οποίες αποτυπώνουν διάφορα είδη εδαφοκάλυψης από εποχιακούς βαλτότοπους μέχρι στερεότερα εδάφη και δάσος.

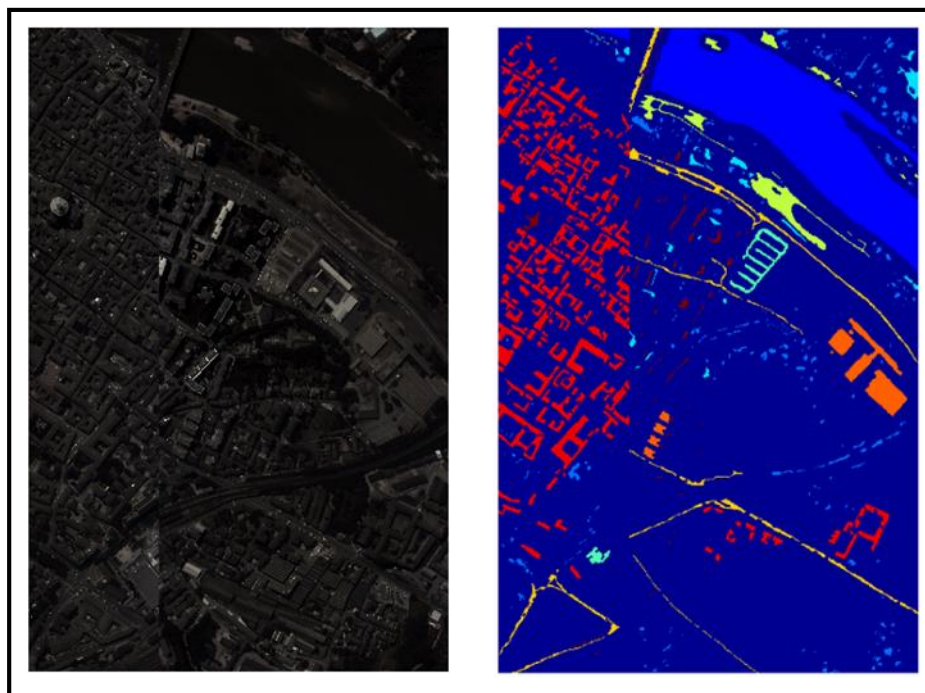
Δυστυχώς, σε αυτό το σετ Ground Truth δεν δίνονταν επακριβώς η ταυτοποίηση της κάθε κλάσης και έτσι δεν είναι γνωστό το τι κατηγορία εδάφους περιλαμβάνει η καθεμία. Παρ' όλα αυτά παρακάτω παρουσιάζεται ο αριθμός των δειγμάτων (samples) που αντιστοιχεί σε κάθε κλάση.

#	Αριθμός δειγμάτων κλάσης	#	Αριθμός δειγμάτων κλάσης
1	270	8	203
2	101	9	314
3	251	10	248
4	215	11	305
5	269	12	181
6	269	13	268
7	259	14	95

Πίνακας 4.6: Οι 14 κλάσεις των δεδομένων Botswana

Pavia Center

Τα υπερφασματικά δεδομένα *Pavia Center* όπως αποκτήθηκαν από το διαδίκτυο: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes περιλαμβάνουν μία υπερφασματική εικόνα η οποία αποτυπώνει την πόλη Ραβία της βόρειας Ιταλίας και μία εικόνα Ground Truth με 9 κλάσεις. Η υπερφασματική εικόνα είναι διαθέσιμη με 102 κανάλια (δέκτης ROSIS). Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=22, G=15, B=9) καθώς και η Ground Truth εικόνα με τις 9 κλάσεις.



Εικόνα 4.4: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα Pavia Center

Όνομα Δεδομένων	Pavia Center
Μέγεθος Εικόνας	1096x715
Δέκτης Λήψης	ROSI
Κάλυψη Ηλεκτρομαγνητικού φάσματος	430 - 837 nm
Αριθμός Καναλιών	102
Αριθμός Κλάσεων Ground Truth	9

Πίνακας 4.7: Χαρακτηριστικά στοιχεία δεδομένων Pavia Center

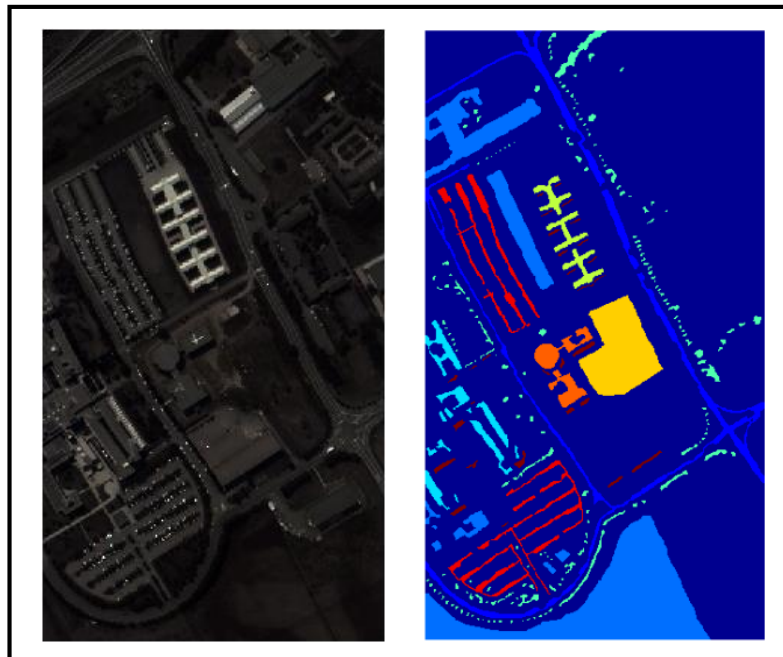
Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 9 κλάσεις, οι οποίες περιλαμβάνουν διάφορες κατηγορίες εδάφους που εντοπίζονται στην περιοχή. Παρακάτω παρουσιάζονται οι 9 κλάσεις μαζί και με τα δείγματα (samples) που περιλαμβάνει η καθεμιά (Πίνακας 4.8).

#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης
1	Water	824
2	Trees	820
3	Asphalt	816
4	Self-Blocking Bricks	808
5	Bitumen	808
6	Tiles	1260
7	Shadows	476
8	Meadows	824
9	Bare Soil	820

Πίνακας 4.8: Οι 9 κλάσεις των δεδομένων Pavia Center

Pavia University

Τα υπερφασματικά δεδομένα *Pavia Center* όπως αποκτήθηκαν από το διαδίκτυο: http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes περιλαμβάνουν μία υπερφασματική εικόνα η οποία αποτυπώνει το πανεπιστήμιο της πόλης Pavia στην βόρεια Ιταλία και μία εικόνα Ground Truth με 9 κλάσεις. Η υπερφασματική εικόνα είναι διαθέσιμη με 103 κανάλια (δέκτης ROSIS). Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=22, G=15, B=9) καθώς και η Ground Truth εικόνα με τις 9 κλάσεις.



Εικόνα 4.5: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα Pavia University

Όνομα Δεδομένων	Pavia University
Μέγεθος Εικόνας	610x340
Δέκτης Λήψης	ROSIS
Κάλυψη Ηλεκτρομαγνητικού φάσματος	430 - 837 nm
Αριθμός Καναλιών	103
Αριθμός Κλάσεων Ground Truth	9

Πίνακας 4.9: Χαρακτηριστικά στοιχεία δεδομένων Pavia University

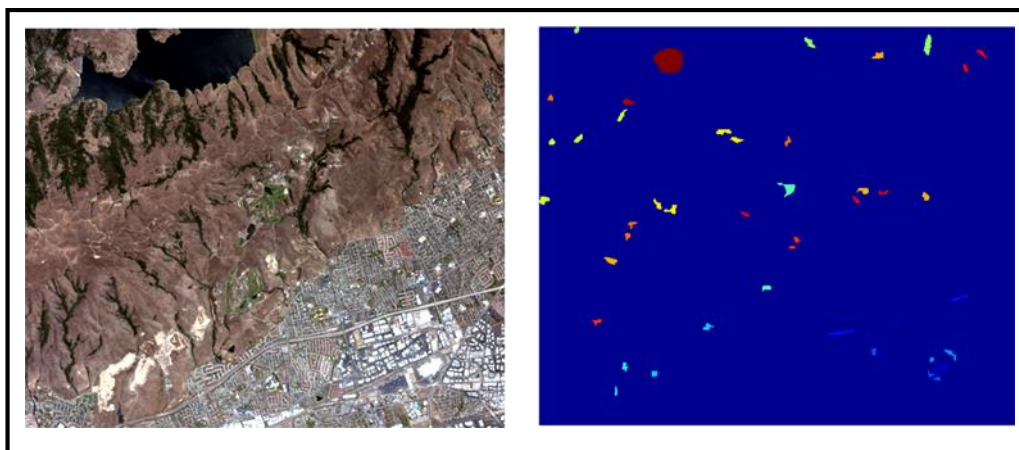
Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 9 κλάσεις, οι οποίες περιλαμβάνουν διάφορες κατηγορίες εδάφους που εντοπίζονται στην περιοχή. Παρακάτω παρουσιάζονται οι 9 κλάσεις μαζί και με τα δείγματα (samples) που περιλαμβάνει η καθεμιά (Πίνακας 4.10) .

#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947

Πίνακας 4.10: Οι 9 κλάσεις των δεδομένων Pavia University

Moffet Field Run 1

Τα υπερφασματικά δεδομένα *Moffet Field Run 1* όπως αποκτήθηκαν από το διαδίκτυο: <http://opticks.org/confluence/display/opticks/Sample+Data#SampleData-AVMOFF3> περιλαμβάνουν μία υπερφασματική εικόνα η οποία αποτυπώνει περιοχή πλησίον του κόλπου του Σαν Φρανσίσκο, στην Καλιφόρνια των ΗΠΑ. Τα δεδομένα Ground Truth ψηφιοποιήθηκαν για τους σκοπούς της εργασίας και δεν συναντώνται αλλού στη διεθνή βιβλιογραφία. Η υπερφασματική εικόνα είναι διαθέσιμη με 224 κανάλια (το σύνολο των καναλιών AVIRIS). Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=22, G=15, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις που δημιουργήθηκαν.



Εικόνα 4.6: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα Moffet Field Run 1

Όνομα Δεδομένων	Moffet Field Run 1
Μέγεθος Εικόνας	512x614
Δέκτης Λήψης	AVIRIS
Κάλυψη Ηλεκτρομαγνητικού φάσματος	400–2500 nm
Αριθμός Καναλιών	224
Αριθμός Κλάσεων Ground Truth	16

Πίνακας 4.11: Χαρακτηριστικά στοιχεία δεδομένων Moffet Field Run 1

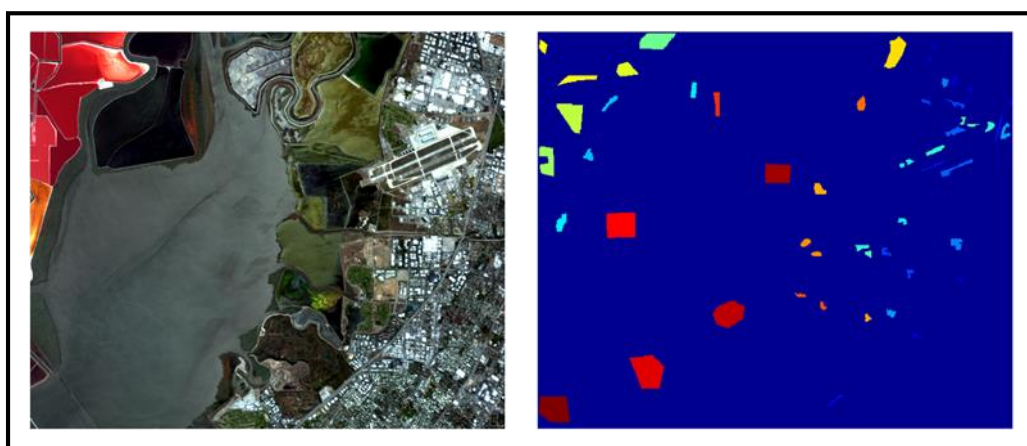
Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 16 κλάσεις, οι οποίες περιλαμβάνουν διάφορες κατηγορίες εδάφους που εντοπίζονται στην περιοχή. Παρακάτω παρουσιάζονται οι 16 κλάσεις μαζί και με τα δείγματα (samples) που περιλαμβάνει η καθεμιά (Πίνακας 4.12) .

#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης	#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης
1	Buildings	147	9	Forest trees_2	292
2	Asphalt Roads & Cement	167	10	Bare soil_1	331
3	Asphalt & Cement	116	11	Bare soil_2	329
4	Cement surface	102	12	Bare soil_3	181
5	Construction Site_1	120	13	Bare soil_4	130
6	Construction Site_2	91	14	Bare soil_5	227
7	Park / Urban green	246	15	Lake water near coast	91
8	Forest trees_1	302	16	Lake water	950

Πίνακας 4.12: Οι 16 κλάσεις των δεδομένων Moffet Field Run 1

Moffet Field Run 3

Τα υπερφασματικά δεδομένα Moffet Field Run 3 όπως αποκτήθηκαν από το διαδίκτυο: <http://opticks.org/confluence/display/opticks/Sample+Data#SampleData-AVMOFF3> περιλαμβάνουν μία υπερφασματική εικόνα η οποία αποτυπώνει περιοχή πλησίον του κόλπου του Σαν Φρανσίσκο, στην Καλιφόρνια των ΗΠΑ. Τα δεδομένα Ground Truth ψηφιοποιήθηκαν για τους σκοπούς της εργασίας και δεν συναντώνται αλλού στη διεθνή βιβλιογραφία. Η υπερφασματική εικόνα είναι διαθέσιμη με 224 κανάλια (το σύνολο των καναλιών AVIRIS). Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=22, G=15, B=9) καθώς και η Ground Truth εικόνα με τις 29 κλάσεις που δημιουργήθηκαν.



Εικόνα 4.73: Έγχρωμο σύνθετο και εικόνα Ground Truth για τα δεδομένα Moffet Field Run 3

Όνομα Δεδομένων	Moffet Field Run 3
Μέγεθος Εικόνας	512x614
Δέκτης Λήψης	AVIRIS
Κάλυψη Ηλεκτρομαγνητικού φάσματος	400–2500 nm
Αριθμός Καναλιών	224
Αριθμός Κλάσεων Ground Truth	29

Πίνακας 4.13: Χαρακτηριστικά στοιχεία δεδομένων Moffet Field Run 3

Όσον αφορά τα δεδομένα Ground Truth της εικόνας αυτά αποτελούνται από 29 κλάσεις, οι οποίες περιλαμβάνουν διάφορες κατηγορίες εδάφους που εντοπίζονται στην περιοχή. Παρακάτω παρουσιάζονται οι 29 κλάσεις μαζί και με τα δείγματα (samples) που περιλαμβάνει η καθεμιά (Πίνακας 4.14).

#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης	#	Όνομα κλάσης	Αριθμός δειγμάτων κλάσης
1	Buildings_1	46	16	Saltpond_2	813
2	Buildings_2	102	17	Saltpond_3	337
3	Buildings_3	100	18	Saltpond_4	445
4	Asphalt road & Airstrip	189	19	Pond	637
5	Cement & Soil surface	250	20	Wetland_1	225
6	Cement surface_1	329	21	Wetland_2	175
7	Cement surface_2	232	22	Wetland_3	158
8	Park / Urban green	220	23	Swamps	108
9	Low vegetation–near wetland_1	247	24	Intertidal flats	183
10	Low vegetation–near wetland_2	322	25	Coastal water_1	1170
11	Bare soil_1	139	26	Coastal water_2	1373
12	Bare soil_2	329	27	Coastal water_3	1011
13	Bare soil_3	141	28	Coastal water_4	763
14	Saline	757	29	Coastal water_5	1171
15	Saltpond_1	529			

Πίνακας 4.14: Οι 29 κλάσεις των δεδομένων Moffet Field Run 3

4.2 Δεδομένα SFB (San Francisco Bay) Και Βήματα Επεξεργασίας

Τα δεδομένα SFB (San Francisco Bay), αποτελούνται συνολικά από 7 υπερφασματικές εικόνες και τις αντίστοιχες 7 εικόνες Ground Truth τα οποία και δημιουργήθηκαν αποκλειστικά στα πλαίσια της παρούσας εργασίας και αποτυπώνουν εκτάσεις πλησίον του κόλπου του San Francisco στην California των ΗΠΑ . Ο τρόπος με τον οποίο οι εικόνες επιλέχθηκαν, αποκτήθηκαν και υπέστησαν προεπεξεργασία, περιγράφεται αναλυτικά στη συνέχεια του κεφαλαίου, όπως επίσης και ο τρόπος που υλοποιήθηκε η ψηφιοποίηση των δεδομένων Ground Truth σε συνολικά 98 κλάσεις / κατηγορίες εδάφους.

Αρχικά, δίνεται μία γενικότερη περιγραφή του έργου AVIRIS της NASA (δεδομένα, εφαρμογές, πτήσεις κ.α.) , στη συνέχεια παρουσιάζονται οι επιλεγμένες εικόνες και η μεθοδολογία απόκτησής τους, τα διάφορα στάδια προεπεξεργασίας που ακολουθήθηκαν στις εικόνες και στην ψηφιοποίηση των δεδομένων Ground Truth και τέλος το σύνολο των ψηφιοποιημένων κλάσεων σύμφωνα και με το Ιεραρχικό Μοντέλο Κλάσεων που ακολουθήθηκε.

4.2.1 Το έργο AVIRIS της NASA

Γενικά

Ο φασματογράφος AVIRIS (airborne visible/infrared imaging spectrometer) είναι ο δεύτερος κατά σειρά φασματογράφος απεικόνισης που αναπτύχθηκε από το εργαστήριο αεριοπροώθησης (Jet Propulsion Laboratory) της NASA με στόχο την επίγεια τηλεπισκόπηση. Η σχεδίαση και ανάπτυξη του AVIRIS προτάθηκε το 1983 ενώ οι πρώτες υπερφασματικές εικόνες συλλέχθηκαν το 1987. Ο φασματογράφος χρησιμοποιεί ένα οπτικό σύστημα σάρωσης και τέσσερα (4) φασματόμετρα, για την στιγμιαία απεικόνιση μιας σειράς 614-pixel σε 224 συνεχή φασματικά κανάλια-μπάντες.

Στόχος

Ο κύριος στόχος του έργου AVIRIS είναι να προσδιορίσει, να μετρήσει και να παρακολουθεί τα συστατικά της Γήινης επιφάνειας και ατμόσφαιρας. Η έρευνα με δεδομένα AVIRIS είναι κυρίως επικεντρωμένη στην κατανόηση των διαδικασιών που σχετίζονται με το παγκόσμιο περιβάλλον και την κλιματική αλλαγή. Ερευνητικοί τομείς με δεδομένα AVIRIS περιλαμβάνουν μεταξύ άλλων την οικολογία, την ωκεανογραφία, τη γεωλογία, την υδρολογία χιονιού και παγετού και τις ατμοσφαιρικές μελέτες.

<i>Είδος σαρωτή</i>	Θέαση Ναδία
<i>Εδαφοκάλυψη σάρωσης εικόνας</i>	11km (υψηλή πτήση) 1,9km (χαμηλή πτήση)
<i>Τυπικό μήκος εδαφοκάλυψης εικόνας</i>	10 – 100km
<i>Μέγεθος εδαφοψηφίδας</i>	20mX20m (υψηλή πτήση) 4mX4m (χαμηλή πτήση)
<i>Οπτικό άνοιγμα δέκτη</i>	1.0 mrad
<i>Κάλυψη ηλεκτρομαγνητικού φάσματος</i>	400 – 2500nm
<i>Αριθμός καναλιών</i>	224 συνεχή κανάλια με βήμα 10nm
<i>Χωρητικότητα εγγραφής δεδομένων</i>	10 Gigabytes δεδομένων (περίπου 850km εδαφοκάλυψης)

Πίνακας 4.15: Τεχνικά χαρακτηριστικά του φασματογράφου AVIRIS

Πλατφόρμες

Ο AVIRIS έχει πραγματοποιήσει αποστολές σε Βόρεια Αμερική, Ευρώπη και σε τμήματα της Αργεντινής και της Λατινικής Αμερικής, στους παρακάτω τέσσερις τύπους αεροσκαφών:

- *NASA's ER-2 jet*
- *Twin Otter International's turboprop*
- *Scaled Composites' Proteus*
- *NASA's WB-57*



Εικόνα 4.8: Το αεροσκάφος NASA's ER-2



Εικόνα 4.9: Το αεροσκάφος Twin Otter International's turboprop

Τεχνικά Χαρακτηριστικά	Πλατφόρμες	
	ER - 2	Twin Otter
<i>Ονομαστική ταχύτητα πτήσης</i>	730 km/h	130 km/h
<i>Ονομαστικό υψόμετρο πτήσης (από επίπεδο θάλασσας)</i>	20 km	4 km
<i>Μέγιστη καλυπτόμενη απόσταση πτήσης</i>	2200 km	560 km
<i>Μέγιστη διάρκεια πτήσης</i>	6,5 ώρες	5 ώρες

Πίνακας 4.16: Περιγραφή αερομεταφερόμενων πλατφορμών (ER - 2 και Twin Otter)

4.2.2 Απόκτηση δεδομένων AVIRIS

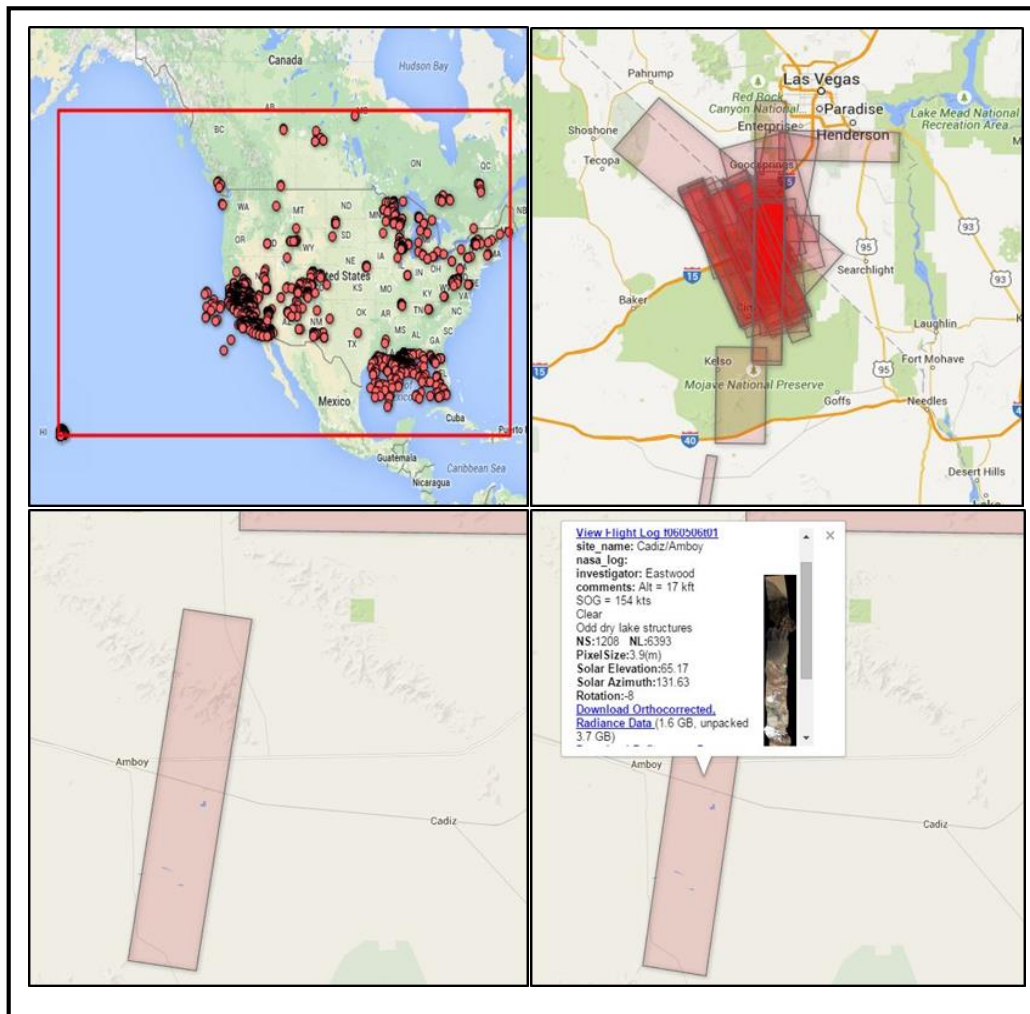
Όλα τα δεδομένα AVIRIS επεξεργάζονται από το εργαστήριο αεριοπροώθησης (JPL) της NASA. Περίπου 20 ώρες μετά τη συλλογή των δεδομένων από το σκάφος, εφαρμόζεται μία αξιολόγηση της απόδοσης του δέκτη κατά την οποία αποκαλύπτονται οι όποιες αστοχίες/ανωμαλίες, δεν ήταν δυνατόν να προβλεφθούν ή να εντοπιστούν προ και μετά την πτήση. Μέσα στις επόμενες 3 μέρες τα δεδομένα εντάσσονται στο αρχείο και παράγονται τα αντίστοιχα δεδομένα γρήγορης έγχρωμης προβολής ("Quicklook"), ενώ μέσα στο επόμενο διάστημα διατίθενται μέσω διαδικτύου με τη μορφή και τους περιορισμούς που περιγράφονται παρακάτω.

Η NASA διαθέτει όλα τα δεδομένα AVIRIS δωρεάν στους ενδιαφερόμενους είτε κατά παραγγελία (με βάση την ημερομηνία ή την περιοχή πτήσης) είτε μέσα από την διαδικτυακή σελίδα AVIRIS locator tool http://aviris.jpl.nasa.gov/alt_locator/. Από αυτή τη σελίδα ο χρήστης έχει πρόσβαση σε όλα τα δεδομένα AVIRIS από το 2006 έως και σήμερα. Όσον αφορά τις πτήσεις από το 1992 έως και το 2005 αυτές είναι διαθέσιμες για προεπισκόπηση στην κάτωθι διαδικτυακή σελίδα http://aviris.jpl.nasa.gov/locator_select.php ενώ η απόκτησή τους είναι δυνατή μόνο μέσω ηλεκτρονικής παραγγελίας.

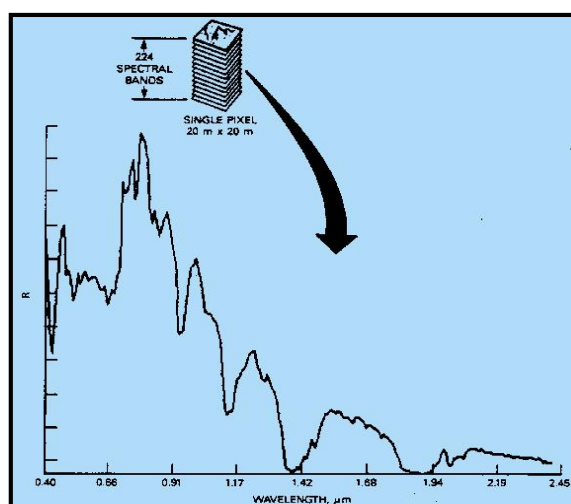
Οι μοναδικοί περιορισμοί για τα δεδομένα μέσω του AVIRIS locator tool είναι οι εξής:

- Οι διαθέσιμες πτήσεις αφορούν τις Η.Π.Α. και κάποιες περιοχές του Καναδά και του κόλπου του Μεξικού
- Από το σύνολο των πτήσεων αποκλείονται όλες οι στρατιωτικές ζώνες και βάσεις καθώς και οι περιοχές στρατιωτικών επιχειρήσεων (Military Operations Areas - MOAs)
- Οι πτήσεις να αφορούν την περίοδο 2006 έως το σήμερα

*Σημειώνεται ότι τα δεδομένα τα οποία δίνονται ελεύθερα και αναφέρονται παραπάνω, αφορούν τις τιμές ακτινοβολίας (**radiance**) που μετράει ο δέκτης σε κάθε περιοχή-pixel. Οι τιμές της ακτινοβολίας μετριοούνται σε $\mu W/(cm^2 * sr * nm)$. Οι πραγματικές τιμές αντανάκλασης (**reflectance**) κάθε κατηγορίας εδάφους προκύπτουν έπειτα από μία απαραίτητη ραδιομετρική "διόρθωση" στα δεδομένα, η οποία καθίσταται δυνατή έπειτα από επιτόπιες ραδιομετρικές μετρήσεις σε κάποιες κατηγορίες εδάφους και τη μετέπειτα επεξεργασία στο απαραίτητο λογισμικό.*



Εικόνα 4.10: Λήψεις από τη σελίδα AVIRIS locator tool της NASA, κατά την προεπισκόπηση και επιλογή πτήσεων. Πηγή: http://aviris.jpl.nasa.gov/alt_locator/



Εικόνα 4.11: Οι τιμές ανακλαστικότητας για ένα δείγμα (pixel) όπως αποτυπώνονται στο ηλεκτρομαγνητικό φάσμα Πηγή: <http://aviris.jpl.nasa.gov/aviris/spectrum.html>

Μορφή δεδομένων

Κάθε προϊόν AVIRIS ακολουθεί μία μοναδική ονομαστική κωδικοποίηση της μορφής `fygmmddttnrnrnrn_v` όπου:

f:	<i>Η αερομεταφερόμενη πτήση</i>
yy:	<i>Το έτος της πτήσης (π.χ. το 10 αναπαριστά το έτος 2010)</i>
mm:	<i>Ο μήνας της πτήσης (π.χ. το 05 αναπαριστά το μήνα Μάιο)</i>
dd:	<i>Η ημέρα της πτήσης (π.χ. το 22 αναπαριστά την 22^η ημέρα του μήνα)</i>
ttn:	<i>Κωδικοποίηση προϊόντων που έχουν εγγραφεί σε μαγνητικές ταινίες</i>
rnn:	<i>Αριθμός κύκλου ισχύος (μονίμως 00)</i>
rnn:	<i>Μοναδικός αριθμός πτήσης</i>
v:	<i>Αριθμός έκδοσης λογισμικού επεξεργασίας σήματος/ακτινοβολίας</i>

Κάθε τελικό προϊόν AVIRIS (τύπου `fygmmddttnrnrnrnrnrdn_v`) αποτελείται από τα παρακάτω αρχεία, τα οποία προσφέρονται ενιαία σε συμπιεσμένη μορφή:

*ortho.readme:	Πληροφορίες σχετικά με την ορθοαναγωγή των εικόνων
*gain:	Συντελεστές πολλαπλασιασμού για τη μετατροπή της ακτινοβολίας σε ακέραιους 16-bit
*rcc:	Συντελεστές ραδιομετρικής βαθμονόμησης
*spc:	Αρχείο φασματικής βαθμονόμησης
*eph:	Αρχείο θέσης στο σύστημα WGS-84/NAD83 UTM με συντεταγμένες x,y,z
*lonlat_eph:	Αρχείο θέσης στο σύστημα WGS-84 με γεωγραφικές συντεταγμένες - γεωγραφικό μήκος, γεωγραφικό πλάτος και υψόμετρο
*obs:	Παράμετροι σχετικές με τη γεωμετρία της παρατήρησης και τις συνθήκες φωτισμού
*obs.hdr:	Αρχείο που "ακολουθεί" και σχετίζεται με το *obs
*obs_ort:	Παράμετροι σχετικές με τη γεωμετρία της παρατήρησης και τις συνθήκες φωτισμού με χρήση του *_ort_glt "look-up table"
*obs_ort.hdr:	Αρχείο που "ακολουθεί" και σχετίζεται με το *obs_ort
*ort.plog:	Γενικές πληροφορίες επεξεργασίας των δεδομένων
*ortho.readme:	Αρχείο που αναφέρει τις πληροφορίες που περιγράφονται εδώ
*ort_glt:	Αρχείο "geometric look-up table"
*ort_glt.hdr:	Αρχείο που "ακολουθεί" και σχετίζεται με το *ort_glt
*ort_igm:	Αρχείο γεωγραφικών δεδομένων με γεωγραφικές συντεταγμένες - γεωγραφικό μήκος, γεωγραφικό πλάτος και υψόμετρο
*ort_igm.hdr:	Αρχείο που "ακολουθεί" και σχετίζεται με το *ort_igm
*ort_img:	Ορθοανοιγμένο αρχείο εικόνας με κλιμακωτή ένταση ακτινοβολίας για κάθε φατνίο σε 224 κανάλια
*ort_img.hdr:	Αρχείο που "ακολουθεί" και σχετίζεται με το *ort_img

Όπως γίνεται αντιληπτό από τις παραπάνω αναφορές, μέσω των δεδομένων AVIRIS η NASA προσφέρει ένα μεγάλο εύρος αξιόπιστων γεωγραφικών και υπερφασματικών δεδομένων που μπορεί να αξιοποιηθεί σε τομείς όπως η φωτογραμμετρία, η τηλεπισκόπηση, η γεωλογία, η ωκεανογραφία, η οικολογία, η προστασία του περιβάλλοντος κ.α.

Παρά όλα αυτά τα απαραίτητα για την εκπόνηση της συγκεκριμένη εργασίας δεδομένα, ήταν αυτά που περιείχαν ραδιομετρική πληροφορία για κάθε πτήση, δηλαδή τα 2 τελευταία αρχεία από τα πάνω αναφερόμενα (τα *fyymmddtnnpnnrnrndn_v.ort_img* και *fyymmddtnnpnnrnrndn_v.ort_img.hdr* αντίστοιχα). Όπως προαναφέρθηκε το αρχείο εικόνας τύπου **ort_img* περιέχει όλη την ραδιομετρική πληροφορία της περιοχής για 224 συνεχή φασματικά κανάλια ενώ η τιμή της ραδιομετρικής ακτινοβολίας σε κάθε φάσμα μετρείται σε $\mu W/(cm^2 * sr * nm)$. Από την άλλη, στο ακόλουθο αρχείο **ort_img.hdr* (το οποίο έχει τη μορφή αναγνώσιμου αρχείου κειμένου), περιγράφονται τα σημαντικά στοιχεία της λήψης (γραμμές, στήλες, κανάλια), τα στοιχεία βαθμονόμησης των 224 καναλιών (μήκη κύματος κάθε κέντρου καναλιού) καθώς επίσης και τις γεωγραφικές συντεταγμένες για την σωστή γεωαναφορά της εικόνας. Σημειώνεται όμως πως η μορφοποίηση και ως εκ τούτου και η ανάγνωσή των στοιχείων της γεωαναφοράς είναι συμβατή μόνο με λογισμικά τύπου ENVI γι' αυτό και δεν ήταν δυνατή η απευθείας σωστά γεωαναφερμένη προβολή τους με άλλα λογισμικά γεωγραφικών πληροφοριών (π.χ. QGIS).

4.2.3 Περιοχή ενδιαφέροντος και επιλεγμένες εικόνες AVIRIS

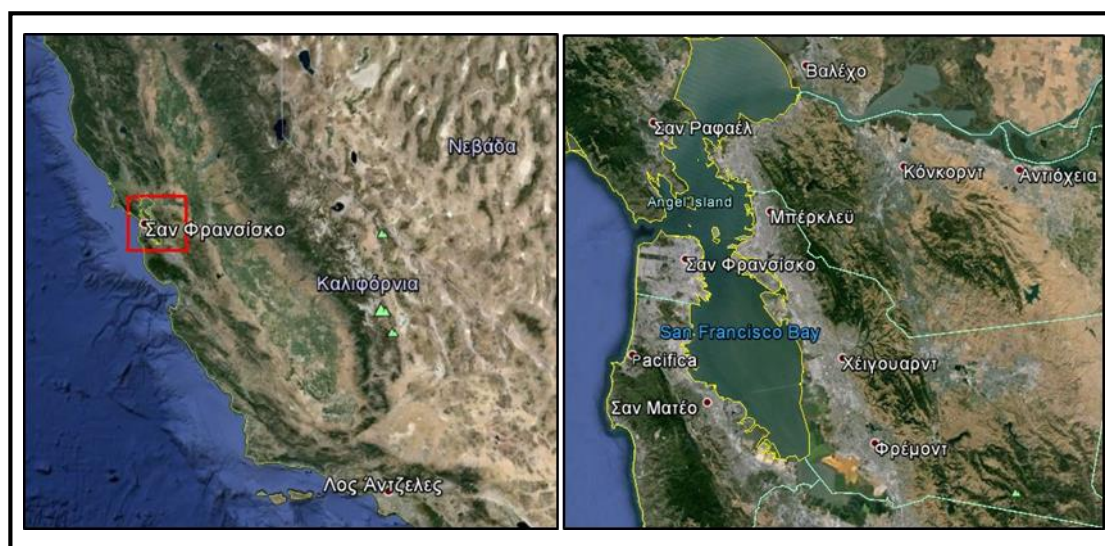
Από τη βάση δεδομένων εικόνων AVIRIS της NASA, στη συνέχεια επιλέχθηκαν τα αρχεία πτήσεων για ανάκτηση. Οι πτήσεις αυτές έπρεπε να αποτυπώνουν μία περιοχή τέτοια που να ανταποκρίνεται στους σκοπούς της εργασίας και στις ταξινομήσεις που ακολούθησαν.

Συγκεκριμένα η περιοχή μελέτης έπρεπε να καλύπτει τα παρακάτω κριτήρια:

- Να έχει μεγάλη ποικιλία ως προς τις κατηγορίες εδάφους, βλάστησης, υγρών και τεχνητών επιφανειών. Με αυτό τον τρόπο θα μπορούσαν να διαχωριστούν οι διάφορες κατηγορίες (με βάση την φασματική τους υπογραφή) για να αποτελέσουν ξεχωριστή κλάση στο ιεραρχικό μοντέλο κλάσεων/κατηγοριών .
- Να υπάρχουν αρκετές διαθέσιμες πτήσεις πάνω και γύρω από την περιοχή ενδιαφέροντος. Με αυτό τον τρόπο εξασφαλίζεται η επάρκεια σε δεδομένα, καθώς εκ των υστέρων πολλές πτήσεις κρίθηκαν ακατάλληλες (εξαιτίας εκτεταμένης νεφοκάλυψης ή κατεστραμμένων από θόρυβο καναλιών)

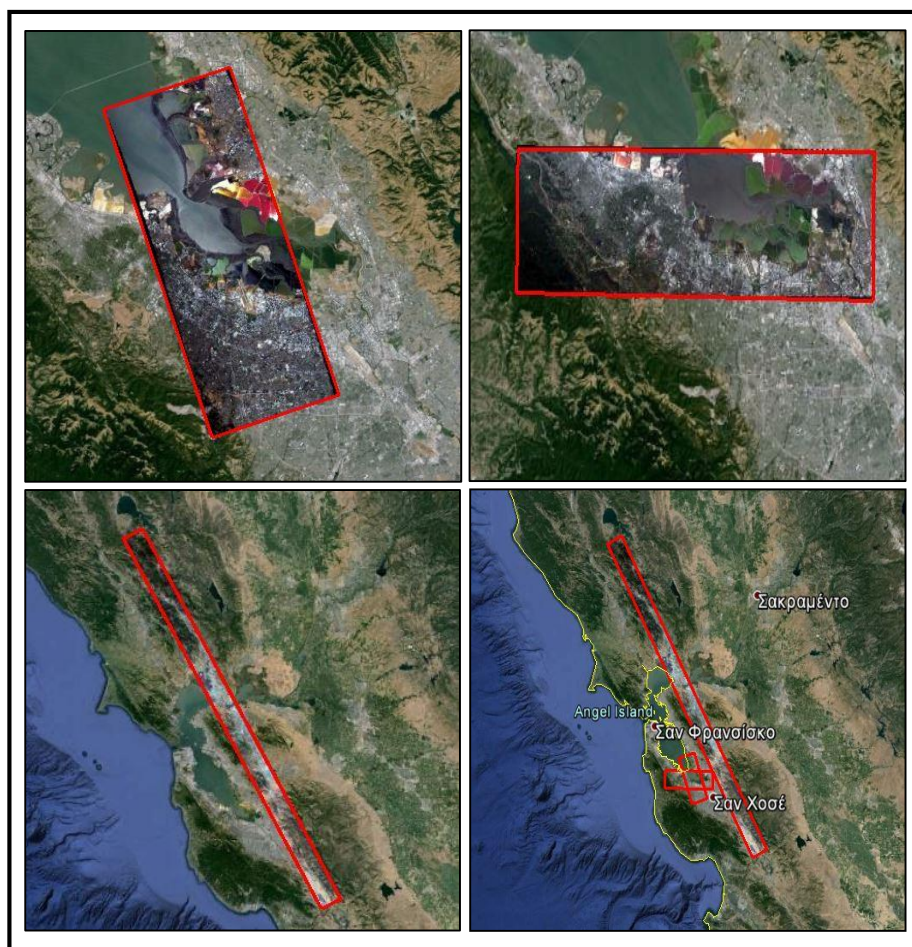
Με βάση τα παραπάνω δύο κριτήρια, επιλέχθηκε η ευρύτερη περιοχή του κόλπου του Σαν Φρανσίσκο στην Καλιφόρνια των Η.Π.Α. . Η συγκεκριμένη περιοχή κρίθηκε κατάλληλη για την δημιουργία δεδομένων που θα καλύπτουν τους σκοπούς της εργασίας, και περιλαμβάνει ένα μεγάλο κομμάτι υγρών επιφανειών του κόλπου, αραιή αλλά και πυκνή βλάστηση από τα γύρω βουνά καθώς επίσης και τμήματα αστικής δόμησης και γενικότερα τεχνητών ανθρωπογενών κατασκευών (βλ. *Εικόνα 4.12*).

Ειδικότερα η μοναδικότητα των υγροβιότοπων, των αλυκών και γενικότερα όλων των υγρών επιφανειών (είτε του κόλπου είτε της ενδοχώρας - "wetlands") έδωσε ακόμα έναν λόγο για την επιλογή αυτής της περιοχής. Τέλος, τονίζεται πως στη συγκεκριμένη περιοχή, η NASA δείχνει ιδιαίτερο ενδιαφέρον με τη μόνιμη παρακολούθηση και μελέτη σε διάφορα προγράμματα (AVIRIS, LANDSAT κ.α.)



Εικόνα 4.12: Η ευρύτερη περιοχή του κόλπου του Σαν Φρανσίσκο, από την οποία αντλήθηκαν τα δεδομένα μελέτης. Πηγή: *Google Earth*

Έπειτα από έλεγχο στη βάση δεδομένων του AVIRIS για τη συγκεκριμένη περιοχή, εντοπίστηκαν συνολικά 7 πτήσεις από τις οποίες θα μπορούσαν να αντληθούν δεδομένα. Κατόπιν ελέγχου όμως διαπιστώθηκε ότι μέσα από 3 πτήσεις θα μπορούσαν να συλλεχθούν επαρκή δεδομένα τόσο για την παραγωγή Ground Truth, όσο και για την μετέπειτα ταξινόμηση στο νευρονικό δίκτυο. Αυτές οι 3 πτήσεις πραγματοποιήθηκαν το 2008, το 2012 και το 2014 και τα τεχνικά χαρακτηριστικά τους παρουσιάζονται αναλυτικότερα στον Πίνακα 4.17:



Εικόνα 4.13: Οι επιλεγμένες πτήσεις των 2008, 2012 και 2014 ξεχωριστά και όλες μαζί, με επισημασμένες τις περιοχές που αποτυπώνουν. Πηγή: Google Earth

Τεχνικά Χαρακτηριστικά Πτήσεων και Δεδομένων	Πτήσεις		
	f080611t01p00r07rdn_c_sc01	f120507t01p00r13rdn_a_sc01	f140528t01p00r11rdn_e_sc01
Ημερομηνία πτήσης	11/6/2008	7/5/2012	28/5/2014
Διεύθυνση Πτήσης (γωνία σε σχέση με τη διεύθυνση Βοράς-Νότος)	19°	90°	29°
Μέγεθος Εδαφοψηφίδας	17m x 17m	16,7m x 16,7m	16m x 16m
Μέγεθος δεδομένων πτήσης (lines X samples)	1923 x 753	2282 x 754	15218 x 790

Πίνακας 4.17: Τεχνικά χαρακτηριστικά επιλεγμένων πτήσεων AVIRIS

4.2.4 Στάδια προεπεξεργασίας δεδομένων

Όλα τα δεδομένα SFB, έπειτα από την διαδικτυακή ανάκτησή τους, υπέστησαν μία σειρά από βήματα απαραίτητης προεπεξεργασίας προκειμένου να είναι δυνατή η εισαγωγή τους στους ταξινομητές. Τα βήματα αυτά συνοψίζονται στα παρακάτω στάδια, τα οποία και εφαρμόστηκαν με την αντίστοιχα ακόλουθη σειρά:

- Αποκοπή μικρότερων σε μέγεθος εικόνων από τα δεδομένα των πτήσεων
- Αφαίρεση καναλιών με θόρυβο. Εύρεση ελάχιστου κοινού αριθμού καναλιών για όλες τις εικόνες
- Εφαρμογή αλγορίθμου k-means για μία πρώτη κατάτμηση της εικόνας
- Έλεγχος των κλάσεων κατάτμησης για επιλογή, συγχώνευση ή διαχωρισμό
- Εξέταση των υποψηφίων κλάσεων με τις προ υπάρχουσες με βάση την υπογραφή τους
- Μορφοποίηση των δεδομένων ώστε να εισαχθούν στους ταξινομητές

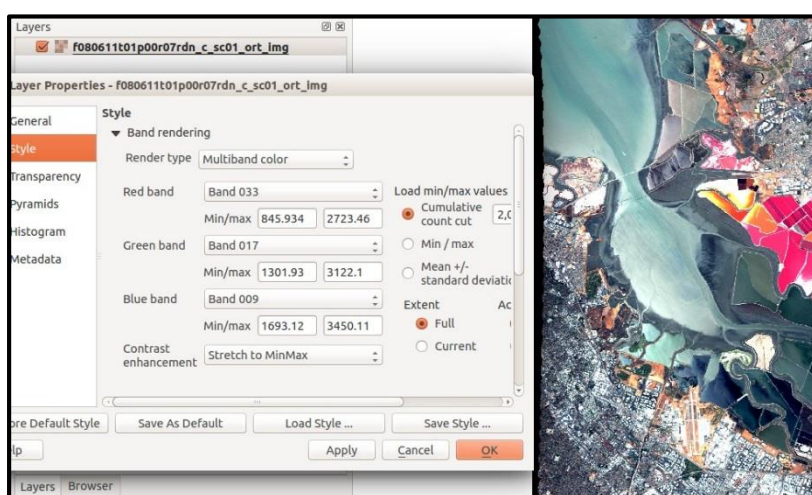
Αποκοπή μικρότερων σε μέγεθος εικόνων από τα δεδομένα των πτήσεων

Όπως αναφέρθηκε και παραπάνω, τα δεδομένα που αντλήθηκαν από τις 3 πτήσεις AVIRIS που τελικώς χρησιμοποιήθηκαν, ποικίλουν και ως προς το μέγεθος των εικόνων τους αλλά και ως προς τον όγκο που καταλαμβάνουν ως δεδομένα. Έτσι κρίθηκε απαραίτητη η περαιτέρω αποκοπή των δεδομένων αυτών, σε δεδομένα/εικόνες μικρότερου μεγέθους ώστε:

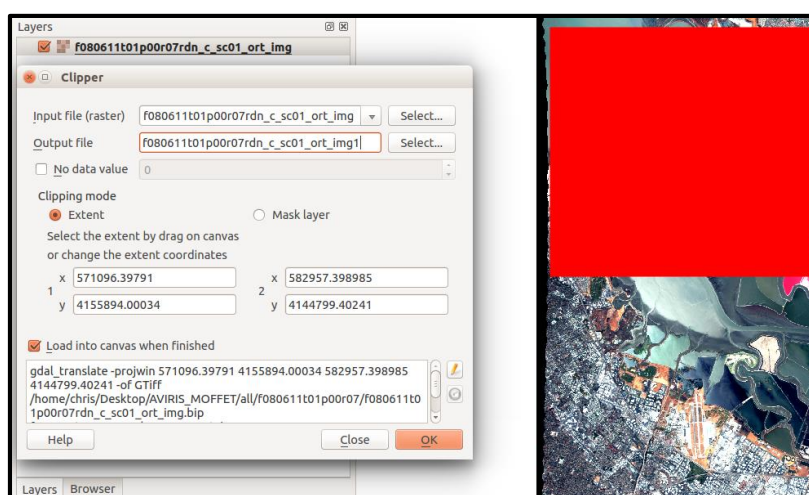
- Οι τελικές “κομμένες” εικόνες να έχουν όλες περίπου το ίδιο μέγεθος (650x550 pixels), ώστε να υπάρχει μία κοινή αναφορά ως προς τις έννοιες: *αρχικές εικόνες-δεδομένα*.
- Να είναι εύκολη, γρηγορότερη και πιο αποτελεσματική η παραγωγή ground truth, καθώς το αρχικό μέγεθος ολόκληρων των εικόνων-πτήσεων καθιστούσε απαγορευτική την όποια επεξεργασία τους. Ενδεικτικά σε μία εικόνα περίπου 650x550 pixels αντιστοιχούν περίπου 30-40 κλάσεις κατάτμησης μέσω του αλγορίθμου K-means και 15-25 τελικές αναφερόμενες κλάσεις ως ground truth. Αντιθέτως, οι 3 αρχικές εικόνες είχαν μέγεθος 1923x753, 2282x754 και 15218x790 pixels οπότε γίνεται εύκολα αντιληπτή η δυσκολία μελέτης τους.
- Να είναι δυνατή η επεξεργασία τους όσον αφορά τον υπολογιστικό φόρτο. Όπως είναι εύκολα κατανοητό, τόσο οι διαδικασίες εκμάθησης (training) και ταξινόμησης (predict) του νευρωνικού δικτύου (CNN), όσο και η διαδικασία

της αρχικής κατάτμησης των εικόνων, απαιτούν μεγάλο υπολογιστικό φόρτο και πολλές φορές ήταν αδύνατον να έρθουν εις πέρας με έναν απλό συμβατικών δυνατοτήτων ηλεκτρονικό υπολογιστή. Πάνω σε αυτή τη λογική εξάλλου, ορίστηκε και το προβλεπόμενο μέγεθος ανά νέα εικόνα στα 650x550 pixels περίπου, ως εκείνο το όριο των δεδομένων το οποίο θα μπορούσε να επεξεργαστεί απρόσκοπτα (από άποψη φυσικής μνήμης-RAM) ο ηλεκτρονικός υπολογιστής εργασίας.

Σημειώνεται πως η αποκοπή των εικόνων, έγινε στο περιβάλλον του προγράμματος QGIS, μέσω του αντίστοιχου εργαλείου αποκοπής εικόνων που διαθέτει και αφού πρώτα τα υπερφασματικά δεδομένα είχαν προβληθεί σε ένα έγχρωμο σύνθετο.



Εικόνα 4.14: Η εμφάνιση έγχρωμου RGB συνθέτου για τις εικόνες των πτήσεων μέσω του QGIS



Εικόνα 4.15: Η αποκοπή μικρότερων σε μέγεθος εικόνων μέσω του εργαλείου clipper στο QGIS

Έπειτα από την αποκοπή των νέων εικόνων, αυτές εισήχθησαν στο προγραμματιστικό περιβάλλον του MATLAB της Mathworks, στο οποίο και υλοποιήθηκαν όλα τα ακόλουθα στάδια προεπεξεργασίας των δεδομένων-εικόνων.

Αφαίρεση καναλιών με θόρυβο. Εύρεση ελάχιστου κοινού αριθμού καναλιών για όλες τις εικόνες

Η ποιότητα των προϊόντων AVIRIS έπρεπε να διασφαλιστεί σε όλα τα στάδια διεκπεραίωσης της εργασίας έτσι ώστε να είναι αξιόπιστα τόσο τα δεδομένα ground truth που δημιουργήθηκαν, όσο και τα τελικά αποτελέσματα ταξινομήσεων που προέκυψαν από τους ταξινομητές.

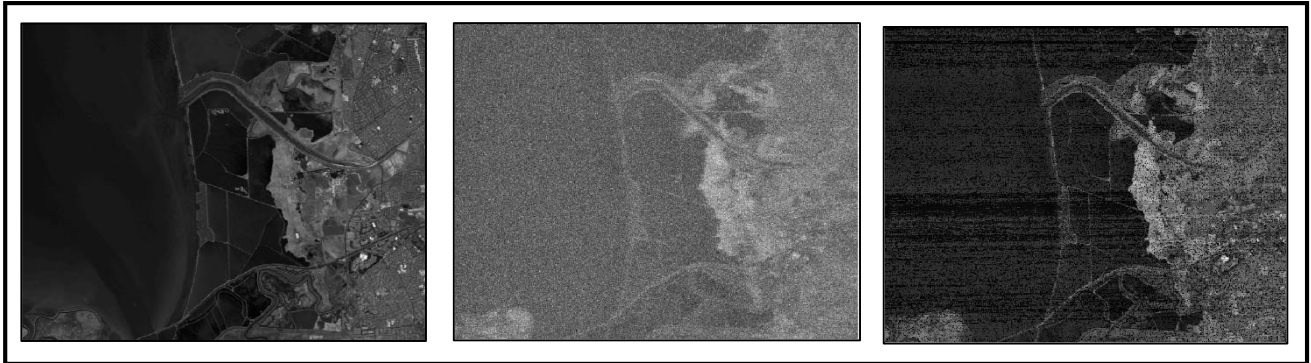
Τα δύο κύρια προβλήματα που εντοπίστηκαν όσον αφορά την αξιοπιστία των προϊόντων AVIRIS ήταν νεφοκάλυψη στις εικόνες της περιοχής μελέτης και ο θόρυβος του σήματος ο οποίος εντοπίστηκε σε διάφορα κανάλια των εικόνων. Το πρόβλημα της νεφοκάλυψης αντιμετωπίστηκε με την επιλογή των κατάλληλων "καθαρών" εικόνων κατά τη διαδικασία του πρωταρχικού ελέγχου και της μετέπειτα αποκοπής εικόνων. Όμως το πρόβλημα του θορύβου στα 224 συνεχή κανάλια του φασματογράφου, είναι πιο σύνθετο και απαιτεί την προσεκτική παρατήρηση όλων των καναλιών από τον χρήστη, τον εντοπισμό των "προβληματικών" καναλιών και την αφαίρεση αυτών από τα τελικά δεδομένα.

Κατά τη διαδικασία της παρατήρησης των καναλιών για κάθε εικόνα, παρουσιάστηκαν διάφορες περιπτώσεις "προβληματικών" ή "κατεστραμμένων" καναλιών εξαιτίας θορύβου στο σήμα της εικόνας, που μπορούν όμως να συνοψιστούν ως εξής:

- Περιπτώσεις διακριτού θορύβου
- Περιπτώσεις συνεχούς θορύβου
- Περιπτώσεις ισχνού σήματος με ελάχιστη ραδιομετρική πληροφορία εικόνας

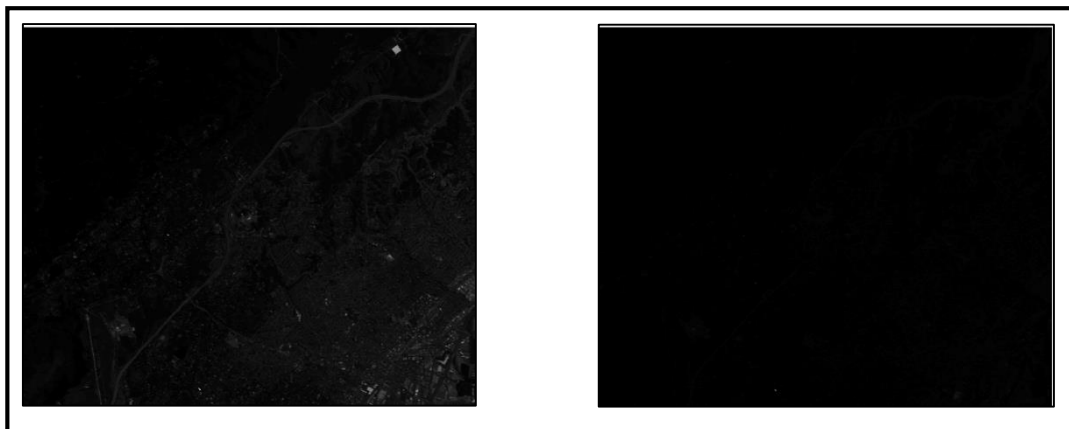
Σημειώνεται πως στις περισσότερες εικόνες εντοπίστηκαν και οι 3 περιπτώσεις αλλοίωσης της ποιότητας των δεδομένων.

Παρακάτω (Εικόνες 4.16 και 4.17) παρουσιάζονται μερικά παραδείγματα των αντίστοιχων περιπτώσεων



Εικόνα 4.16: Κανάλια εικόνας της πτήσης του 2008.

Από αριστερά: Καθαρό κανάλι, κανάλι με συνεχή θόρυβο, κανάλι με διακριτό θόρυβο

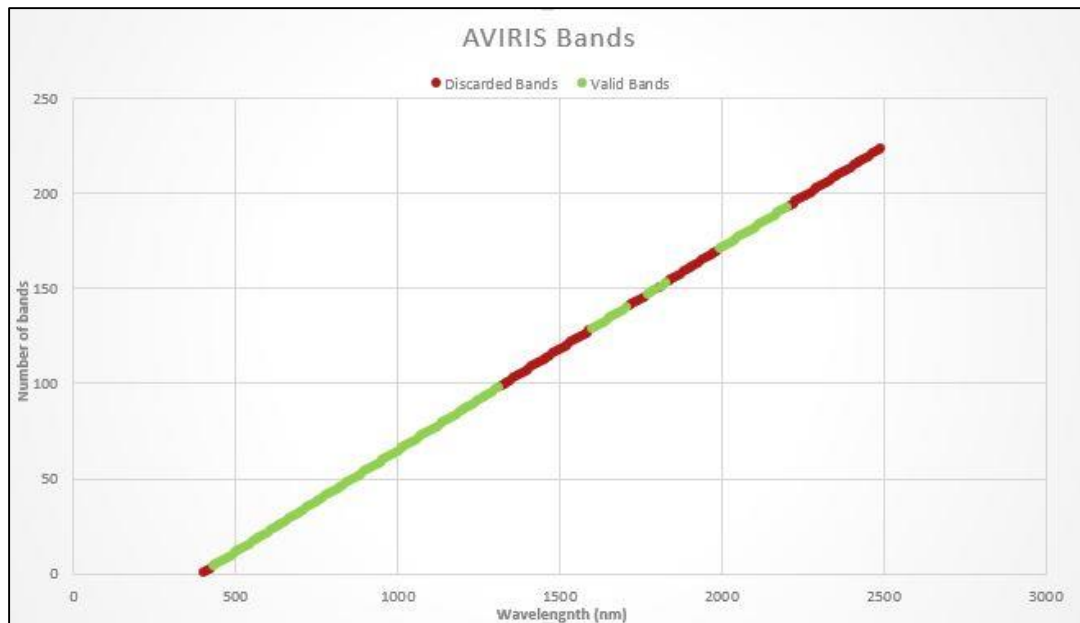


Εικόνα 4.17: Κανάλια εικόνας της πτήσης του 2012.

Αριστερά: Καθαρό κανάλι , Δεξιά: Κανάλι με ισχνό σήμα

Αφού αφαιρέθηκαν τα "προβληματικά" κανάλια για κάθε εικόνα ξεχωριστά, στη συνέχεια εντοπίστηκαν τα κοινά εναπομείναντα κανάλια για όλες τις εικόνες μελέτης (εννοώντας τα κανάλια των οποίων το σήμα προέρχεται από το ίδιο μήκος κύματος).

Έτσι, τα τελικά, κοινά, εναπομείναντα κανάλια για όλες τις εικόνες είναι συνολικά 135 και τα σήματα λήψης τους σε σχέση με τα μήκη κύματος του ηλεκτρομαγνητικού φάσματος φαίνονται στην *Εικόνα 10*.



Εικόνα 4.18: Τα τελικά, εναπομείναντα και εξαιρούμενα, κανάλια των εικόνων μελέτης σε σχέση με τα μήκη κύματος των σημάτων τους

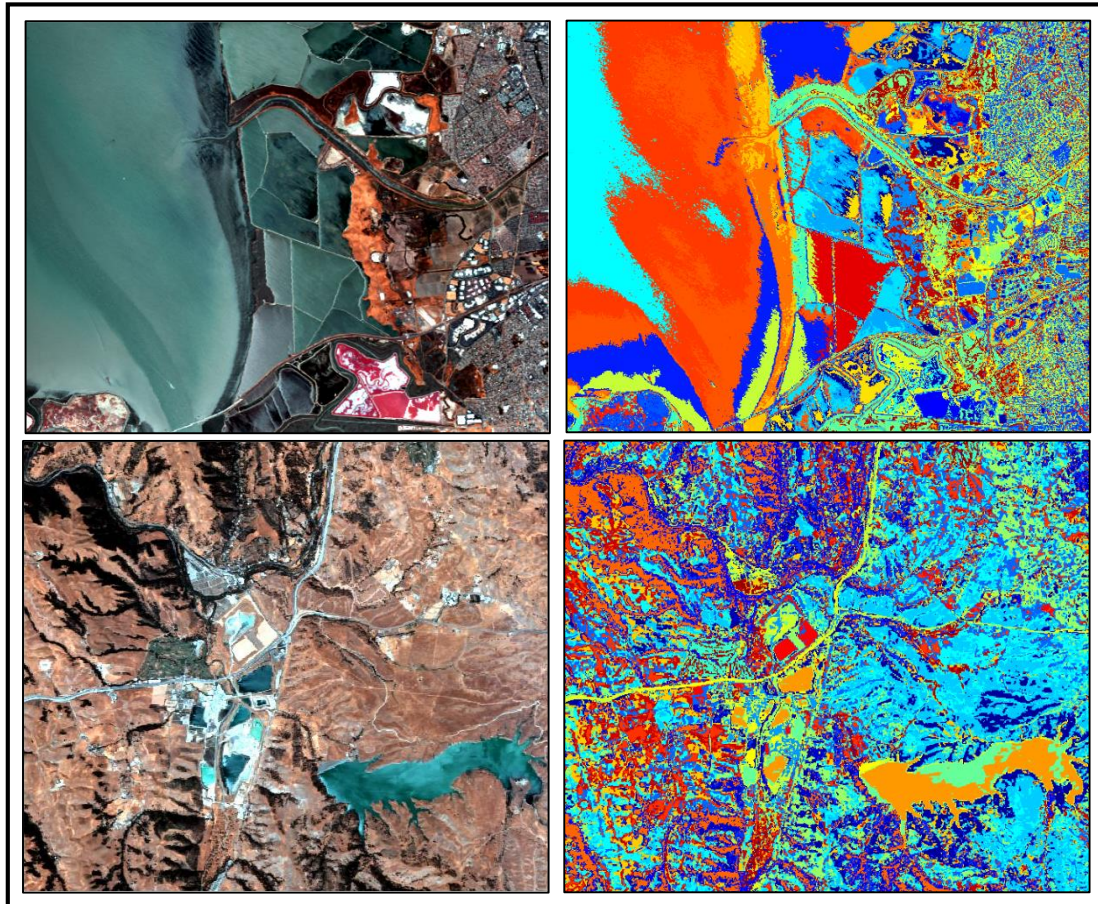
Εφαρμογή αλγορίθμου k-means για την κατάτμηση της εικόνας

Σε αυτό το στάδιο προεπεξεργασίας, σε κάθε μία εικόνα ξεχωριστά εφαρμόστηκε ο αλγόριθμος k-means ως ένα πρώτο βήμα κατάτμησης της εικόνας.

Μία αρχική κατάτμηση (εδώ με k-means) κρίθηκε απαραίτητη καθώς ο ορισμός και η δημιουργία λεπτομερών δεδομένων Ground Truth, είναι αδύνατον να επιτευχθεί μόνο με την θέαση των διάφορων καναλιών εικόνας από το χρήστη, ή από τις φασματικές υπογραφές ενός φαντίου τη φορά. Αυτό συμβαίνει καθώς είναι αρκετά δύσκολος και ασαφής ο ορισμός τόσο των ορίων, όσο και ο τελικός αριθμός κλάσεων σε μια ομοιογενή περιοχή με παραπλήσιες κατηγορίες εδάφους (π.χ. οι διαφορετικές κατηγορίες βλάστησης σε ένα ομοιογενή δάσος). Αντιθέτως, με μία πρώτη κατάτμηση της εικόνας γίνεται πιο γρήγορος και ευκολότερος τόσο ο έλεγχος των κλάσεων της κατάτμησης (clusters) μεταξύ τους όσο και έλεγχός τους με τις ήδη υλοποιημένες κλάσεις Ground Truth. *Η αναλυτική χρήση των παραγώγων του αλγορίθμου k-means για κάθε εικόνα, περιγράφεται παρακάτω μαζί με την αντίστοιχη μεθοδολογία εργασίας.*

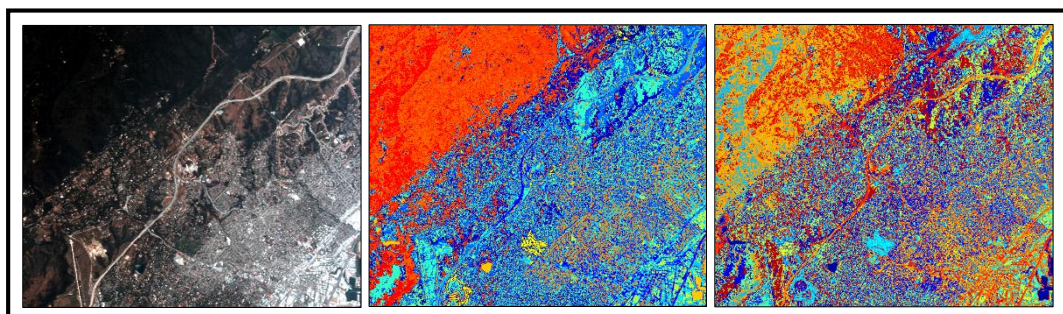
Ο αλγόριθμος κατάτμησης k-means τελικά επιλέχθηκε για την αρχική κατάτμηση της εικόνας καθώς είναι ένα αρκετά απλό, γρήγορο, αποτελεσματικό εργαλείο κατάτμησης. Συγκεκριμένα, η χρήση του αλγορίθμου μέσω της αντίστοιχης συνάρτησης στο Matlab έδινε την δυνατότητα αξιοποίησης όλης της υπερφασματικής πληροφορίας για κάθε φαντίο (σαν παράμετροι-εδώ 135 κανάλια) ενώ και ο ορισμός του επιθυμητού αριθμού κλάσεων από το χρήστη καθιστά τα παραγόμενα δεδομένα ποιοτικά επαρκή.

Ο επιλεγμένος αριθμός κλάσεων της κατάτμησης ορίστηκε ξεχωριστά για κάθε εικόνα με βάση τα μοναδικά χαρακτηριστικά της περιοχής που αποτυπώνει. Έτσι στις εικόνες που συναντάται μεγάλη ποικιλία κατηγοριών εδάφους (υγρές επιφάνειες, αστική δόμηση, περιοχές βλάστησης – *Εικόνα 4.19 πάνω*) επιλέχθηκαν περισσότερες κατηγορίες από τις εικόνες που αποτύπωναν μεγάλες περιοχές ομοιογενών χαρακτηριστικών (π.χ. βλάστηση ή γυμνό έδαφος - *Εικόνα 4.19 κάτω*). Ωστόσο, σε διάφορες εικόνες εφαρμόστηκε περισσότερες από μία φορές ο αλγόριθμος, με διαφορετικό αριθμό κλάσεων κάθε φορά, ώστε να βρεθεί ο κατάλληλος αριθμός κλάσεων κατάτμησης (βλ. *Εικόνα 4.20*). Σε κάθε περίπτωση, ο επιλεγμένος αριθμός κλάσεων της κατάτμησης στις εικόνες μελέτης κυμαίνεται από 18 έως 35.



Εικόνα 4.19: Πάνω: Εικόνα της πτήσης του 2008 με την παράγωγη κατατμημένη εικόνα k-means , 35 κλάσεων

Κάτω: Εικόνα της πτήσης του 2014 με την παράγωγη κατατμημένη εικόνα k-means , 20 κλάσεων



Εικόνα 4.20: Εικόνα πτήσης του 2012 με 2 διαφορετικά αποτελέσματα κατάτμησης. Από αριστερά: RGB σύνθετο της εικόνας, κατάτμηση με 18 κλάσεις, κατάτμηση με 30 κλάσεις

Αριθμός κλάσεων κατατημένης εικόνας	Εικόνες Μελέτης (πτήση - αριθμός)						
	2008 - 1	2008 - 2	2008 - 3	2012 - 1	2014 - 1	2014 - 2	2014 - 3
	35	30	20 - 35	18– 25– 30	30	20	25

Πίνακας 4.18: : Οι εικόνες μελέτης με τους αριθμούς κλάσεων των κατατημένων εικόνων τους

Δημιουργία κλάσεων με βάση το ιεραρχικό μοντέλο και ψηφιοποίηση της αντίστοιχης περιοχής Ground Truth στις εικόνες

Στη συνέχεια ακολούθησε το σημαντικότερο και εκτενέστερο στάδιο της προεπεξεργασίας των δεδομένων, αυτό της παραγωγής Ground Truth. Η δημιουργία δεδομένων Ground Truth και άρα κλάσεων στο ιεραρχικό μοντέλο, υλοποιήθηκε με βάση τις κατατημένες εικόνες που δημιουργήθηκαν στο προηγούμενο στάδιο όπως περιγράφεται παρακάτω.

Αρχικά, εξετάστηκαν μία προς μία οι παραχθείσες κλάσεις κατάτησης με παράλληλη θέαση του αντίστοιχου RGB έγχρωμου σύνθετου της εικόνας μελέτης. Με αυτόν τον αποφασίστηκε το εάν η κλάση κατάτησης θα είναι και υποψήφια κλάση για το Ground Truth ή όχι.

Συγκεκριμένα, μία κλάση (cluster) κατάτησης αντιμετωπίστηκε ως μία από τις παρακάτω περιπτώσεις:

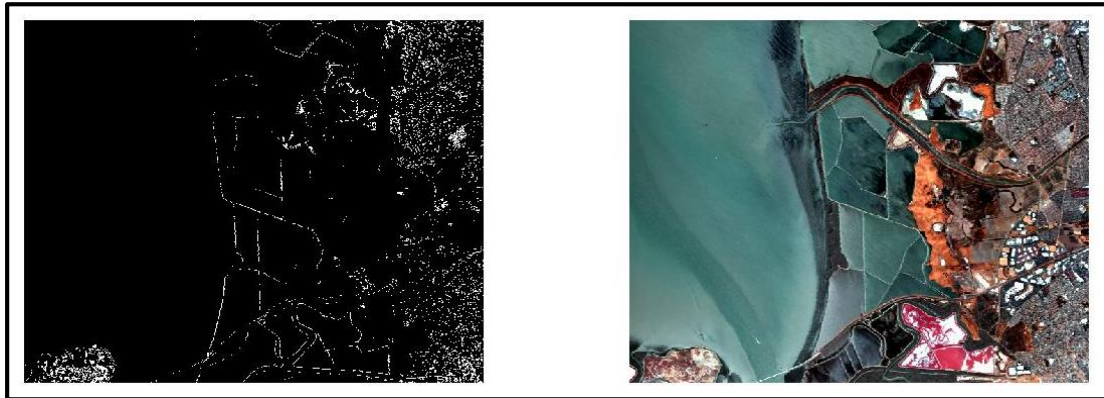
- Ορίστηκε και ως κλάση στο παραγόμενο Ground Truth
- Εξαιρέθηκε από την υπόλοιπη διαδικασία
- Διαχωρίστηκε σε δύο ή περισσότερες κλάσεις
- Συγχωνεύτηκε μαζί με άλλη κλάση σε ενιαία κλάση Ground Truth

Οι περισσότερες κλάσεις (clusters) ορίστηκαν και ως κλάσεις Ground Truth, εφόσον αναφέρονταν σε μία ομοιογενή περιοχή και παράλληλα μοναδική κατηγορία εδάφους. Αντιθέτως, οι κλάσεις (clusters) οι οποίες εξαιρέθηκαν (βλ. *Εικόνα 4.21*, *Εικόνα 4.22*) ήταν αυτές οι οποίες αποτελούνταν από διάσπαρτα και ασυνεχή pixels και από εκείνες που κρίθηκαν αδύνατον να ψηφιοποιηθούν. Τέλος παρατηρήθηκαν και κλάσεις οι οποίες είτε διαχωρίστηκαν είτε συγχωνεύτηκαν με άλλες. Οι κλάσεις οι οποίες αναφέρονταν σε παραπάνω από μία κατηγορία εδάφους διαχωρίστηκαν (βλ. *Εικόνα 4.23*) για να δημιουργηθούν 2 κλάσεις Ground Truth, ενώ οι κλάσεις που

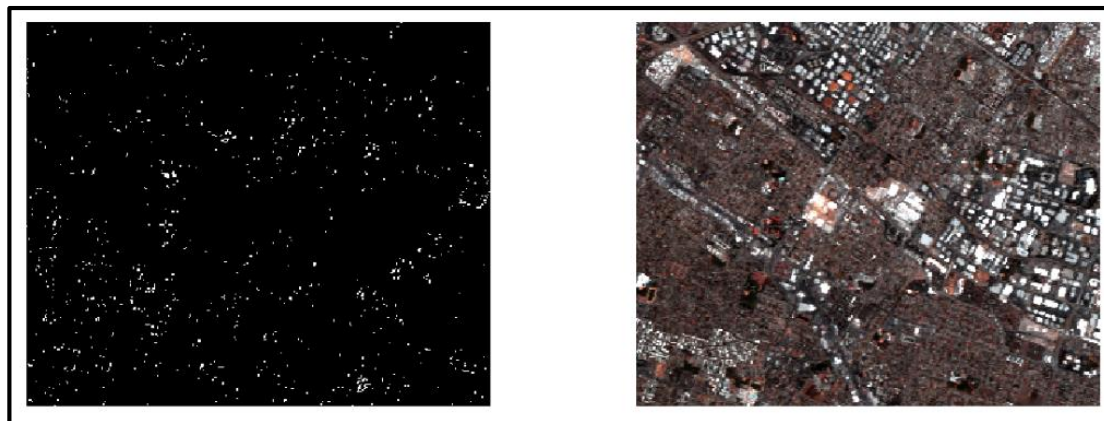
εντοπίστηκαν στην ίδια κατηγορία εδάφους της εικόνας, και που παρουσίαζαν σχεδόν ταυτόσημες φασματικές υπογραφές, συγχωνεύτηκαν.

Όπως γίνεται εύκολα αντιληπτό, οι δύο τελευταίες περιπτώσεις είναι συνήθως απόρροια εσφαλμένου καθορισμού του αριθμού κλάσεων κατά τη διαδικασία κατάτμησης με *k-means* (ορισμός μικρότερου αριθμού στην 1^η περίπτωση και μεγαλύτερου στη 2^η)

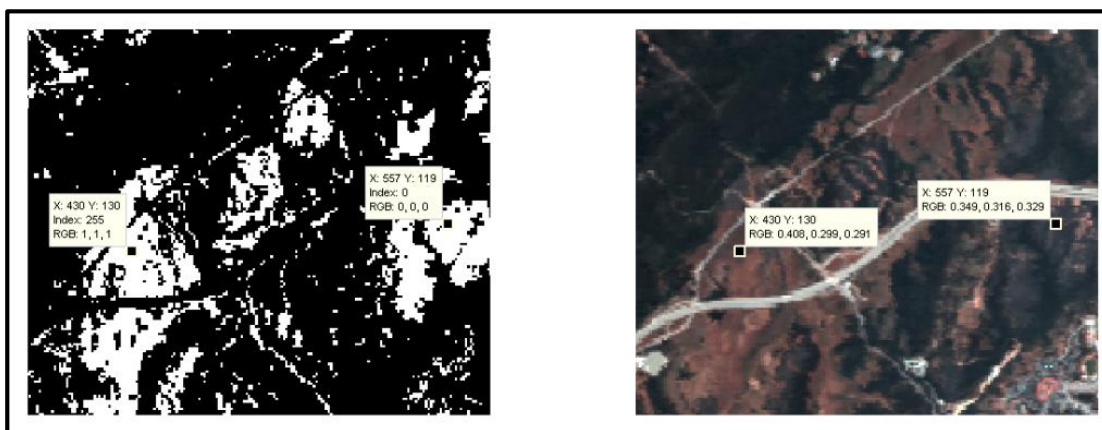
Παρακάτω παρουσιάζονται ενδεικτικά παραδείγματα για των άνωθεν αναφερόμενων περιπτώσεων.



Εικόνα 4.41: Εξάιρεση κλάσης κατάτμησης από επιλογή, εξαιτίας αδυναμίας ψηφιοποίησης



Εικόνα 4.22: Εξάιρεση κλάσης κατάτμησης από επιλογή, εξαιτίας διασποράς και ασυνέχειας των pixels



Εικόνα 4.23: Διαχωρισμός κλάσης κατάτμησης με 2 κατηγορίες εδάφους, σε 2 κλάσεις Ground Truth

Οι κλάσεις κατάτμησης που τελικά επιλέχθηκαν να οριστούν και ως κλάσεις Ground Truth, ονοματίστηκαν με βάση τη δομή του ιεραρχικού μοντέλου κλάσεων (βλ. Κεφάλαιο 4.2.4).

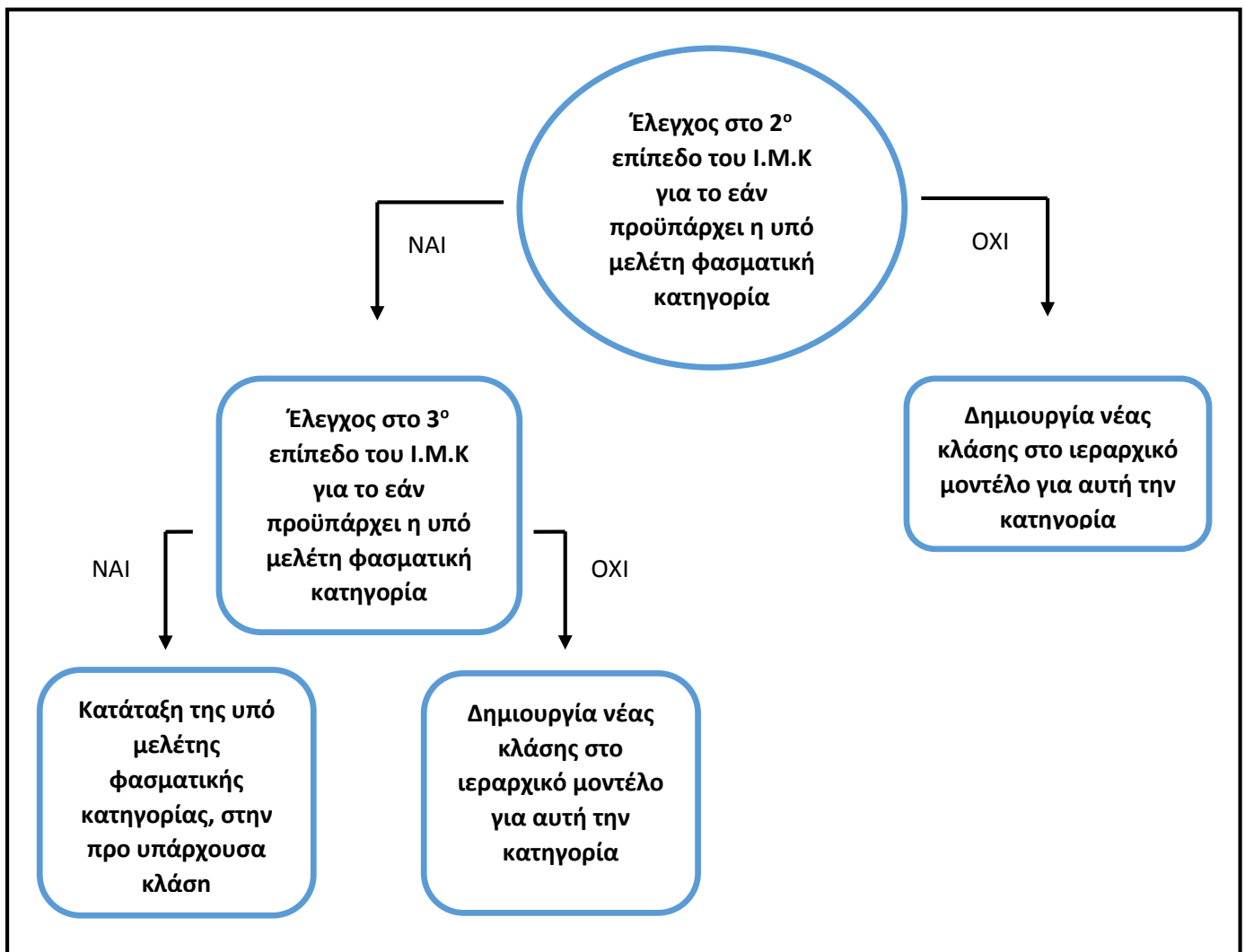
Έτσι, για την ταξινόμηση της εκάστοτε υποψήφιας κλάσης έπρεπε να γίνει έλεγχος για το κατά πόσο η συγκεκριμένη κατηγορία εδάφους (φασματική κατηγορία ακριβέστερα) προϋπάρχει στο ιεραρχικό μοντέλο. Αυτή είναι μία πολυσύνθετη διαδικασία η οποία απαιτεί τον παράλληλο έλεγχο φασματικών υπογραφών για πολλές κλάσεις, και περιγράφεται αναλυτικά εδώ.

Για κάθε υποψήφια κλάση, γίνεται έλεγχος στο κατά πόσο υπάρχει η γενικότερα αναφερόμενη κατηγορία εδάφους στο 2^ο επίπεδο του ιεραρχικού μοντέλου κλάσεων:

- Σε περίπτωση που δεν υπάρχει καμία κατηγορία 2^{ου} επιπέδου, στην οποία να μπορεί να ταξινομηθεί η υποψήφια κλάση, τότε δημιουργείται νέα κλάση στο Ιεραρχικό Μοντέλο Κλάσεων (Ι.Μ.Κ.) για την υπό μελέτη φασματική κατηγορία.
- Στην περίπτωση όπου συναντάται κατηγορία 2^{ου} επιπέδου, στην οποία να μπορεί να ταξινομηθεί η υποψήφια κλάση, γίνεται έλεγχος των φασματικών υπογραφών όλων των κατηγοριών 3^{ου} επιπέδου και:
 - Εάν βρεθεί κλάση (3^ο επίπεδο) στην οποία να συναντάται παραπλήσια/ταυτόσημη φασματική υπογραφή, η υπό μελέτη υποψήφια κλάση κατατάσσεται στην προ υπάρχουσα κλάση
 - Εάν δεν βρεθεί κλάση (3^ο επίπεδο) στην οποία να συναντάται παραπλήσια/ταυτόσημη φασματική υπογραφή, δημιουργείται νέα κλάση στο Ιεραρχικό Μοντέλο Κλάσεων (Ι.Μ.Κ.) για την υπό μελέτη φασματική κατηγορία.

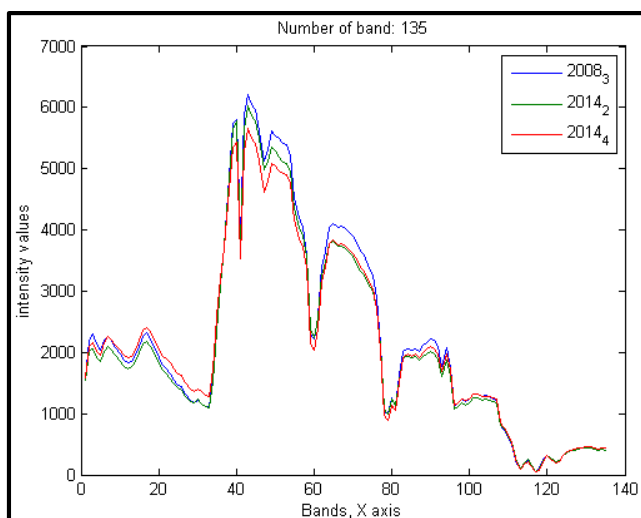
Φυσικά, το όριο της "παραπλήσιας/ταυτόσημης φασματικής υπογραφής" τέθηκε αυθαίρετα και με γνώμονα την ικανότητα των ταξινομητών να ταξινομήσει αποτελεσματικά τις εικόνες με βάση τα δεδομένα Ground Truth (π.χ. Η απόκλιση των φασματικών υπογραφών στην Εικόνα 4.24 κρίθηκε αποδεκτή, ενώ στην Εικόνα 4.25 όχι). Εξάλλου, η παραπάνω διαδικασία επαναλήφθηκε πολλές φορές και μετά την ταξινόμηση μέσω του δικτύου, προκειμένου να αποφανθεί η τελικά "σωστή" δομή και ιεραρχία των κλάσεων.

Τα παραπάνω φαίνονται σχηματικά στο παρακάτω διάγραμμα ροής (Σχήμα 1) που αποτυπώνει τη μεθοδολογία που ακολουθήθηκε.

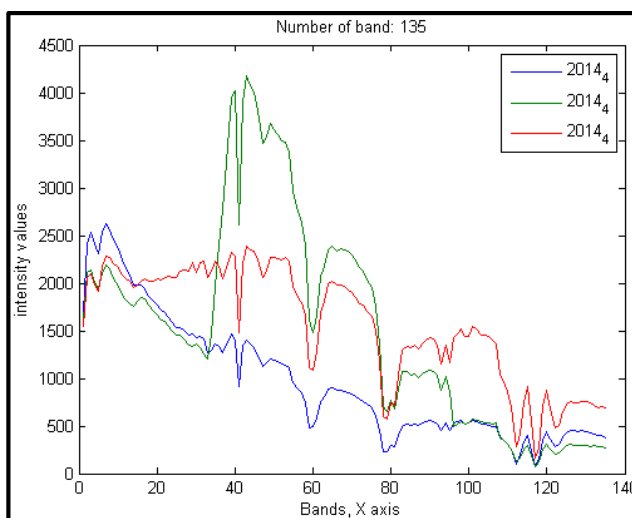


Σχήμα 4.1: Διάγραμμα ροής που αποτυπώνει τη μεθοδολογία που ακολουθήθηκε για τον ορισμό των κλάσεων στο Ιεραρχικό Μοντέλο Κλάσεων (I.M.K.)

Παρακάτω παρουσιάζονται παραδείγματα φασματικών υπογραφών δεδομένων, που είτε ορίστηκαν στην ίδια κλάση όντας σε διαφορετικές εικόνες (Εικόνα 4.24), είτε ορίστηκαν σε διαφορετικές κλάσεις όντας στην ίδια εικόνα (Εικόνα 4.25).



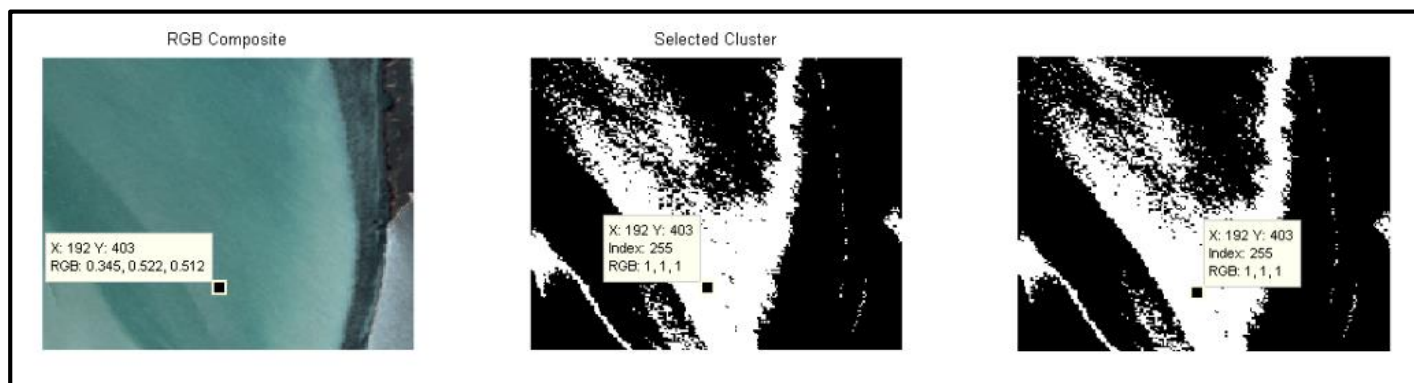
Εικόνα 4.24: Οι φασματικές υπογραφές της ίδιας κλάσης σε 3 διαφορετικές εικόνες



Εικόνα 4.25: Οι φασματικές υπογραφές 3 διαφορετικών κλάσεων (κατηγορίες εδάφους) στην ίδια εικόνα

Τέλος, αφού καθορίστηκαν οι τελικές κλάσεις που θα αποτελούσαν το Ground Truth για κάθε εικόνα ακολούθησε η διαδικασία ψηφιοποίησης τους. Πάλι μέσω του προγραμματιστικού περιβάλλοντος του MATLAB και της αντίστοιχης εντολής, υλοποιήθηκε η ψηφιοποίηση των κλάσεων με την παράλληλη θέαση ενός έγχρωμου RGB συνθέτου και της επιλεγμένης κλάσης κατάτμησης (cluster) σε binary μορφή (βλ. Εικόνα 4.26).

Σε αυτό το στάδιο δόθηκε ιδιαίτερη προσοχή ώστε η προς ψηφιοποίηση περιοχή μιας κλάσης κατάτμησης να αποτελείται από ομοιογενή γειτονικά pixels (ίδια υφή, φωτεινότητα και χρώμα στο RGB) και επίσης να αποτελείται όσο το δυνατόν από πολλά συνεχόμενα φαντίνια και όχι διάσπαρτα. Με αυτό τον τρόπο αποφεύγονται οι όποιες ασυνέχειες μεταξύ των δεδομένων της ίδιας κλάσης, καθιστώντας τα περισσότερο αξιόπιστα. Ένα τέτοιο ενδεικτικό παράδειγμα, αποτελεί η επισημασμένη περιοχή στην Εικόνα 4.26.



Εικόνα 4.265: Τα 3 παράθυρα θέασης κατά τη διαδικασία ψηφιοποίησης των δεδομένων Ground Truth για κάθε εικόνα μελέτης

Μορφοποίηση και κωδικοποίηση των δεδομένων Ground Truth

Το τελευταίο στάδιο της προεπεξεργασίας των δεδομένων, είναι αυτό της μορφοποίησης των δεδομένων Ground Truth ώστε να είναι έχουν την “σωστή” μορφή και κωδικοποίηση για την εισαγωγή τους στους ταξινομητές. Συγκεκριμένα, κατά την δημιουργία και ψηφιοποίηση των δεδομένων Ground Truth σε κάθε κλάση δόθηκε μία μοναδική τιμή που αντιπροσωπεύει και την θέση της στο Ιεραρχικό Μοντέλο Κλάσεων. Έτσι π.χ. στα pixels που ανήκουν στην πρώτη κλάση της 1^η δευτερεύουσας κατηγορίας και της 1^{ης} πρωτεύουσας κατηγορίας του Ι.Μ.Κ. δόθηκε η τιμή 111, αντίστοιχα στα pixels που ανήκουν στην πρώτη κλάση της 2^{ης} δευτερεύουσας κατηγορίας και της 2^{ης} πρωτεύουσας κατηγορίας του Ι.Μ.Κ. δόθηκε η τιμή 221 κ.ο.κ.

Όμως, μία σωστή κωδικοποίηση για τα δεδομένα Ground Truth (προκειμένου να είναι εφικτή και απρόσκοπτη η ταξινόμηση μέσω και των 3 ταξινομητών) είναι οι κλάσεις να παίρνουν τις τιμές από το 0 έως και τον συνολικό αριθμό κλάσεων της εκάστοτε ταξινόμησης. Έτσι για την ταξινόμηση μιας εικόνας με αντίστοιχα δεδομένα Ground Truth 10 κλάσεων, τα δεδομένα αυτά θα πρέπει να έχουν τις τιμές 0 (background), 1, 2, 3, 4, 5, 6, 7, 8, 9 και 10. Αυτή η διαδικασία ακολουθήθηκε τόσο στις περιπτώσεις απλών μονοεικονικών ταξινομήσεων όσο και στις συνθετότερες πολυεικονικές.

4.2.5 Ιεραρχικό Μοντέλο Κλάσεων και τελικά δεδομένα SFB

Το ιεραρχικό μοντέλο κλάσεων (IMK)

Οι παραγόμενες κλάσεις, οι οποίες δημιουργήθηκαν (*παραγωγή ground truth*) και στις οποίες εκ των υστέρων ταξινομήθηκαν τα δεδομένα μέσω των τριών αλγορίθμων/ταξινομητών, έπρεπε να ακολουθούν μία δομή εύχρηστη, ευανάγνωστη και βιβλιογραφικά συνεπή. Με μία τέτοια δομή, έγινε εύκολη η όποια εκ των υστέρων διαδικασία επανεξέτασης, συγχώνευσης, διαχωρισμού και δημιουργίας κλάσεων.

Συγκεκριμένα, το ιεραρχικό μοντέλο που δημιουργήθηκε για τους σκοπούς της εργασίας ακολουθεί τα πρότυπα του αντίστοιχου ιεραρχικού μοντέλου της Ευρωπαϊκής Υπηρεσίας Περιβάλλοντος (*European Environment Agency-EEA*), που υλοποιήθηκε και χρησιμοποιήθηκε στα πλαίσια των προγραμμάτων Corine Land Cover 2000 (*CLC 2000*) και Corine Land Cover 2006 (*CLC 2006*). Ειδικότερα, οι κατηγορίες κάλυψης γης ιεραρχούνται σε 3 επίπεδα, με 4 κατηγορίες στο 1ο επίπεδο, 15 στο 2ο και 43 στο 3ο επίπεδο που αποτελούν και το συνολικό αριθμό κατηγοριών γης του μοντέλου.

Με αντίστοιχο τρόπο δημιουργήθηκε και το ιεραρχικό μοντέλο κατηγοριών-κλάσεων της εργασίας (βλ. Πίνακες 4.19 - 4.23) που αποτελείται από 98 κλάσεις οι οποίες κατηγοριοποιούνται σε 3 επίπεδα επίσης. Φυσικά οι κλάσεις που δημιουργήθηκαν είναι προσαρμοσμένες στα χαρακτηριστικά της περιοχής μελέτης και για αυτό το λόγο υπάρχουν πολλές διαφορές με το μοντέλο της ΕΕΑ τόσο στην ονοματολογία των κατηγοριών όσο και στο περιεχόμενό τους (πολλές κατηγορίες υγρών ζωνών, έλλειψη αρκετών κατηγοριών καλλιεργειών κ.λ.π.) .

Στις επόμενες σελίδες παρουσιάζεται εποπτικά το σύνολο των παραγόμενων κλάσεων, για τα δεδομένα SFB, όπως αυτές ορίστηκαν μέσω του ιεραρχικού μοντέλου.

Level 1	Level 2	Level 3
1. Artificial Surfaces	1.1 Urban fabric	Buildings_2
		Buildings_3
		Buildings_4
		Buildings_5
	1.2 Networks	Asphalt roads & Cement_2
		Asphalt roads & Cement_3
		Airstrip
		Airstrip & rubber
	1.3 Construction sites & Materials surfaces	Construction site
		Asphalt & Cement_1
		Cement surface
		Asphalt & Soil surface
		Asphalt & Cement_2
	1.4 Artificial Vegetated Area	Park/Urban green area_1
		Park/Urban green area_2
		Park/Urban green area_3

Πίνακας 4.19: Η 1^η κατηγορία του 1^{ου} επιπέδου στο ιεραρχικό μοντέλο κλάσεων

Level 1	Level 2	Level 3
2. Vegetation & Seminatatural Areas	2.1 Forests	Forest Trees_1
		Forest Trees_2
		Forest Trees_3
		Forest Trees_4
		Forest Trees_5
		Forest Trees_6
	2.2 Herbaceous Vegetation and shrubs	Low vegetation/grass_1
		Low vegetation/grass_2
		Low vegetation/grass_3
		Low vegetation/grass_4
		Low vegetation/grass_5
		Low vegetation/grass_6
	2.3 Open spaces without vegetation	Bare soil_1
		Bare soil_2
		Bare soil_3
		Bare soil_4
		Bare soil_5
Bare soil_6		
Bare soil_7		
Bare Soil_8		
Bare Soil_9		
Bare Soil_10		
Bare Soil_11		
Bare Soil_12		
Bare Soil_13		
Bare Soil_14		
Bare soil_15		
Bare soil_16		
Bare soil_17		

Πίνακας 4.20: Η 2^η κατηγορία του 1^{ου} επιπέδου στο ιεραρχικό μοντέλο κλάσεων

Level 1	Level 2	Level 3
3. Agriculture	3.1 Tilled lands and Plantations	Tilled land
		Semi- tilled land_1
		Semi- tilled land_2
		Semi- tilled land_3
		Plantation_1
		Plantation_2
		Plantation_3
		Plantation_4
		Plantation_5
		Plantation_6
Plantation field		

Πίνακας 4.21: Η 3^η κατηγορία του 1^{ου} επιπέδου στο ιεραρχικό μοντέλο κλάσεων

Level 1	Level 2	Level 3
4. Wetlands	4.1 Maritime wetlands	Wetland_1
		Wetland_2
		Wetland_3
		Wetland_4
		Wetland_5
		Wetland_6
		Wetland_7
		Wetland_8
		Wetland_9
		Wetland_10
		Wetland_11
		Wetland_12
		Wetland_13
		Wetland_14
	4.2 Inland wetlands	Saltpond_1
		Saltpond_2
		Saltpond_3
		Saltpond_4
		Saltpond_5
		Saltpond_6
		Salt-covered Area_1
		Salt-covered Area_2
Salt-covered Area_3		
Salt-covered Area_4		
Salt-covered Area_5		

Πίνακας 4.22: Η 4^η κατηγορία του 1^{ου} επιπέδου στο ιεραρχικό μοντέλο κλάσεων

Level 1	Level 2	Level 3
5. Water Bodies	5.1 Inland water	Mini Lake / Water Suply
		Lake water_1
		Lake water_2
	5.2 Coastal Water / Lagoon	Coastal water_1
		Coastal water_2
		Coastal water_3
		Coastal water_4
		Coastal water_5
		Coastal water_6
		Coastal water_7
		Coastal water_8
		Coastal water_9
		Coastal water_10
		Coastal water_11
		Coastal water_12
Coastal water_13		

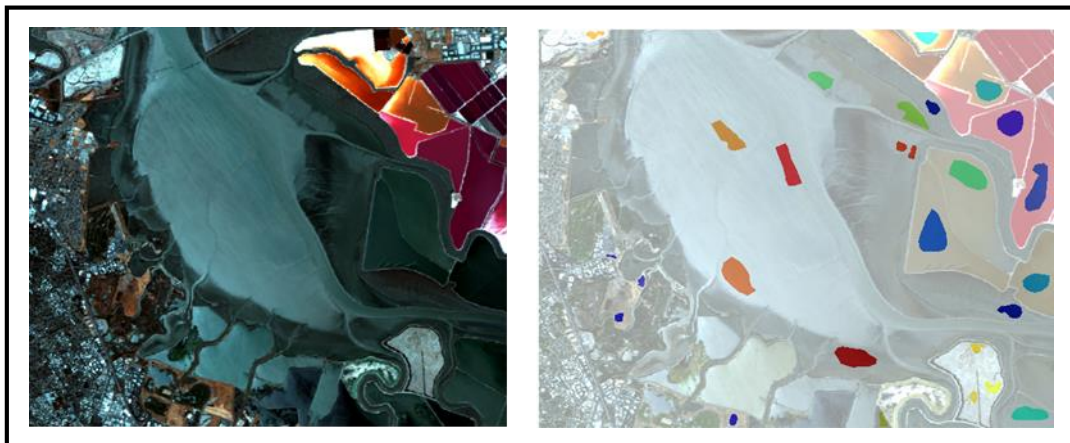
Πίνακας 4.23: Η 5^η κατηγορία του 1^{ου} επιπέδου στο ιεραρχικό μοντέλο κλάσεων

Οι τελικές υπερφασματικές εικόνες και τα δεδομένα Ground Truth

Από τις παραπάνω αναφερόμενες πτήσεις AVIRIS (του 2008, του 2012 και του 2014) επιλέχθηκαν οι περιοχές που συγκέντρωναν το μεγαλύτερο ενδιαφέρον και 'αποκόπηκαν' (βλ. 4.2.3) 7 νέες εικόνες. Παρακάτω παρουσιάζονται οι 7 αυτές νέες εικόνες που με 135 υπερφασματικά κανάλια και τα αντίστοιχα δεδομένα Ground Truth, αποτελούν το σύνολο των δεδομένων SFB.



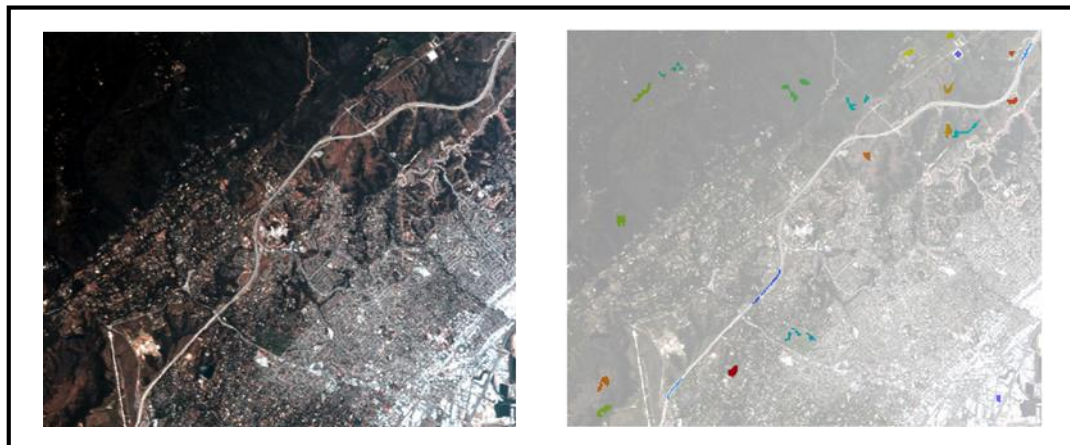
Εικόνα 4.27: Έγχρωμο σύνθετο της εικόνας 2008-1, με τα αντίστοιχα δεδομένα Ground Truth



Εικόνα 4.28: Έγχρωμο σύνθετο της εικόνας 2008-2, με τα αντίστοιχα δεδομένα Ground Truth



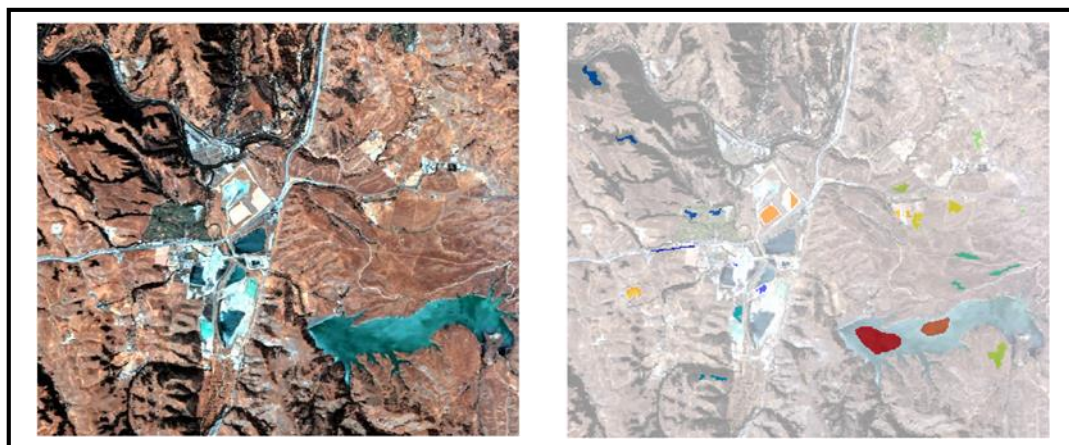
Εικόνα 4.29: Έγχρωμο σύνθετο της εικόνας 2008-3, με τα αντίστοιχα δεδομένα Ground Truth



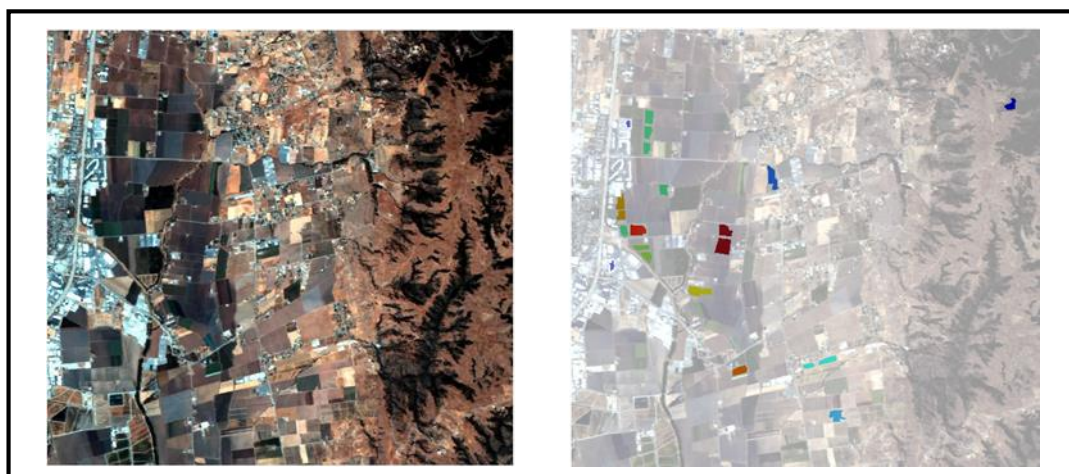
Εικόνα 4.30: Έγχρωμο σύνθετο της εικόνας 2012-1, με τα αντίστοιχα δεδομένα Ground Truth



Εικόνα 4.31: Έγχρωμο σύνθετο της εικόνας 2014-1, με τα αντίστοιχα δεδομένα Ground Truth



Εικόνα 4.32: Έγχρωμο σύνθετο της εικόνας 2014-2, με τα αντίστοιχα δεδομένα Ground Truth



Εικόνα 4.33: Έγχρωμο σύνθετο της εικόνας 2014-3, με τα αντίστοιχα δεδομένα Ground Truth

Όνομα Εικόνας (Έτος - Σειρά)	Δέκτης Λήψης	Μέγεθος Εικόνας	Αριθμός Καναλιών	Αριθμός Κλάσεων
2008 - 1	AVIRIS	516x657	135	26
2008 - 2	AVIRIS	454x591	135	22
2008 - 3	AVIRIS	544x657	135	12
2012 - 1	AVIRIS	563x676	135	14
2014 - 1	AVIRIS	589x727	135	14
2014 - 2	AVIRIS	563x624	135	14
2014 - 3	AVIRIS	529x568	135	12

Πίνακας 2.24: Τα χαρακτηριστικά στοιχεία των 7 νέων υπερφασματικών εικόνων

5

ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ

Σε αυτό το κεφάλαιο παρουσιάζονται αναλυτικά όλα τα αποτελέσματα που προέκυψαν κατά την εκπόνηση της εργασίας. Τα αποτελέσματα περιλαμβάνουν ταξινομημένες εικόνες, πίνακες αξιολόγησης χρονικής απόδοσης και ακριβειών αλλά και γραφήματα με τα οποία γίνεται ευκολότερα αντιληπτή η απόδοση του κάθε αλγορίθμου-ταξινομητή στα επιμέρους δεδομένα εικόνων (datasets).

Όπως αναφέρθηκε και σε προηγούμενο κεφάλαιο (βλ. *Κεφάλαιο 3*) στην εργασία αυτή εφαρμόστηκαν και αξιολογήθηκαν 3 διαφορετικοί ταξινομητές (με έμφαση στο Convolutional Neural Network - CNN) πάνω σε μία μεγάλη ποικιλία υπερφασματικών δεδομένων-εικόνων. Οι εφαρμογές που παρουσιάζονται παρακάτω ποικίλουν ως προς:

- Τα διαφορετικά δεδομένα στα οποία εφαρμόζονται οι ταξινομητές
- Το ποσοστό των δεδομένων Ground Truth που χρησιμοποιήθηκαν ως δεδομένα εκπαίδευσης (training samples) στην κάθε ταξινόμηση
- Τον αριθμό των κυρίων συνιστωσών (PCA) που χρησιμοποιήθηκαν στην κάθε ταξινόμηση
- Τον αριθμό των κλάσεων που χρησιμοποιήθηκαν για τις ίδιες εικόνες σε κάθε ταξινόμηση

Με βάση και τις παραπάνω παραμέτρους, τα αποτελέσματα της κάθε ταξινόμησης αξιολογούνται ως προς την ακρίβεια, την αξιοπιστία και την χρονική απόδοση που επιτυγχάνει ο κάθε ταξινομητής.

Στην επόμενη σελίδα παρουσιάζονται εποπτικά σε πίνακα (*Πίνακας 5.1*) όλες οι ταξινομήσεις που υλοποιήθηκαν για τους σκοπούς της εργασίας, μαζί και με τα επιμέρους χαρακτηριστικά της καθεμιάς.

	Datasets	CNN	RBF	LIN	Self Training Dataset	PCA reduction	N° of classes reduction
Bibliography Dataset	<i>Indian Pines</i>	✓	✓	✓	✓	✓	
	<i>Salinas</i>	✓	✓	✓	✓		
	<i>Pavia C</i>	✓	✓	✓	✓		
	<i>Pavia U</i>	✓	✓	✓	✓		
	<i>Botswana</i>	✓	✓	✓	✓	✓	
	<i>1997_1</i>	✓	✓	✓	✓		
	<i>1997_3</i>	✓	✓	✓	✓		
SFB Dataset	<i>2008_1</i>	✓	✓	✓	✓		
	<i>2014_4</i>	✓	✓	✓	✓		
	<i>SFB_all</i>	✓	✓	✓	✓	✓	✓

Πίνακας 5.1: Το σύνολο των ταξινομήσεων που υλοποιήθηκαν στα πλαίσια της παρούσας εργασίας

με:

Self Training Dataset: Ταξινόμηση κατά την οποία τα ‘δεδομένα εκπαίδευσης’ είναι ίδια με τα ‘δεδομένα πρόβλεψης’.

PCA reduction: Ταξινόμηση κατά την οποία έχει επιβληθεί μείωση στον αριθμό των κυρίων συνιστωσών.

No of classes reduction: Ταξινόμηση κατά την οποία έχει επιβληθεί μείωση στον αριθμό των κλάσεων στα δεδομένα Ground Truth.

*Σημειώνεται πως στους ακόλουθους πίνακες αποτελεσμάτων αυτού του κεφαλαίου, με τον συμβολισμό (*) θα επισημαίνονται οι χρονικές αποδόσεις στις οποίες το δίκτυο CNN έφτασε το προκαθορισμένο όριο επαναλήψεων

5.1 Δομή Κεφαλαίου

Εξαιτίας του μεγάλου αριθμού των εφαρμογών που υλοποιήθηκαν, καθώς επίσης και της πολυπλοκότητας των αλγορίθμων, στο υποκεφάλαιο 5.2 γίνεται προσπάθεια να αποσαφηνιστεί η μεθοδολογία εργασίας ως προς την παραμετροποίηση των αλγορίθμων για κάθε εφαρμογή.

Ακόμα, κρίθηκε απαραίτητο τα αποτελέσματα να χωριστούν και να παρουσιαστούν σε 2 βασικές ξεχωριστές κατηγορίες. Ο διαχωρισμός αυτός αφορά τόσο τα δεδομένα στα οποία εφαρμόστηκαν οι ταξινομητές όσο και τη διαφορετική προσέγγιση που ακολουθήσε η κάθε εφαρμογή.

Ειδικότερα, στην 1^η κατηγορία (5.3) παρουσιάζονται τα αποτελέσματα ταξινόμησης για τα βιβλιογραφικά δεδομένα στα οποία υλοποιούνται ταξινομήσεις ξεχωριστά σε κάθε υπερφασματική εικόνα.

Αντίστοιχα, στην 2^η και εκτενέστερη κατηγορία (5.4) παρουσιάζονται τα αποτελέσματα ταξινόμησης για τα δεδομένα SFB. Στο υποκεφάλαιο 5.4.1 υλοποιούνται ταξινομήσεις ξεχωριστά σε κάθε υπερφασματική εικόνα, ενώ στο υποκεφάλαιο 5.4.2 υλοποιούνται σύνθετες εφαρμογές ταξινόμησης όσον αφορά τα δεδομένα εκπαίδευσης των ταξινομητών και τις εικόνες εφαρμογής.

5.2 Παραμετροποίηση Αλγορίθμων

Όπως αναφέρθηκε παραπάνω κάθε παραχθείσα ταξινόμηση σε αυτή την εργασία διαφέρει ως προς τα χρησιμοποιούμενα δεδομένα, και τις διαφορετικές τιμές τις οποίες παίρνουν κάποιες παράμετροι των ταξινομητών. Παρακάτω αναφέρονται αναλυτικά οι τιμές αυτών των παραμέτρων για κάθε εφαρμογή καθώς επίσης παρουσιάζονται και τα τεχνικά χαρακτηριστικά του υπολογιστή στον οποίο υλοποιήθηκαν όλες οι ταξινομήσεις.

5.2.1 Principal Component Analysis (PCA)

Ο αριθμός των Κυρίων Συνιστωσών (εδώ Κ.Σ.) που επιλέχθηκε για κάθε εφαρμογή δεν είναι συγκεκριμένος. Αντιθέτως κρίθηκε σκόπιμο να μειωθεί ο αριθμός των Κ.Σ. προκειμένου να αξιολογηθούν καλύτερα οι ταξινομητές και κυρίως το νευρωνικό δίκτυο στο οποίο υλοποιήθηκαν και οι περισσότερες εφαρμογές με μειωμένο αριθμό Κ.Σ. .

Όμως, η απόδοση του κάθε ταξινομητή δεν εξαρτάται άμεσα από τον αριθμό των Κ.Σ. αλλά από το ποσοστό της πληροφορίας (των υπερφασματικών καναλιών) που περιέχεται στις Κ.Σ. της κάθε εφαρμογής. Έτσι π.χ. στα δεδομένα της εικόνας *Indian Pines* οι 10 Κ.Σ. αντιστοιχούν στο 97,21% της υπερφασματικής πληροφορίας, ενώ στα δεδομένα της εικόνας *Salinas* στο 99,97%. Μία διαφορά τέτοιας τάξης είναι σημαντική για τα δεδομένα και τις αποδόσεις των υπό δοκιμή ταξινομητών. Για αυτό το λόγο, ως δείκτης της χρησιμοποιούμενης πληροφορίας από την εκάστοτε PCA συνάρτηση τέθηκε ποσοστό και συγκεκριμένα το 99,8% της υπερφασματικής πληροφορίας. Έτσι, για όλες τις παρακάτω εφαρμογές, ασχέτως αριθμού Κ.Σ., η βέλτιστη απόδοση του κάθε ταξινομητή επετεύχθη με τουλάχιστον αυτό το ποσοστό της υπερφασματικής πληροφορίας που λειτουργεί ως 'κατώφλι'.

Σε 2 περιπτώσεις, συγκεκριμένα στα δεδομένα της εικόνας *Indian Pines* και στα δεδομένα της εικόνας *Botswana*, επετεύχθη ποσοστό περισσότερο από 99,8%, με αριθμό Κ.Σ. 55 και 25 αντίστοιχα. Σε αυτές τις 2 περιπτώσεις δεδομένων εκτός από τον επιμέρους αναφερόμενο αριθμό έγιναν δοκιμές και ταξινομήσεις με 10 Κ.Σ. .

Επιπλέον, υλοποιήθηκαν αρκετές ταξινομήσεις με μειωμένο αριθμό Κ.Σ. (και φυσικά μικρότερο του 99,8% ποσοστό) κυρίως στα δεδομένα *SFB* (βλ. 5.4.2). Σε αυτές τις περιπτώσεις οι αριθμοί των μειωμένων Κ.Σ. είναι 3, 2 και 1. Σε κάθε εφαρμογή μειωμένων Κ.Σ. πάντως, μαζί με τον αριθμό Κ.Σ. που επιλέχθηκε αναφέρεται και το αντίστοιχο ποσοστό της υπερφασματικής πληροφορίας.

5.2.2 Ποσοστό Ground Truth ως δεδομένα εκπαίδευσης

Ένας άλλος τρόπος ώστε να αξιολογηθεί η απόδοση του κάθε ταξινομητή είναι η μείωση του αριθμού δεδομένων εκπαίδευσης της ταξινόμησης. Το αποτέλεσμα της επιβλεπόμενης ταξινόμησης, με την οποία καταπιάνεται και η συγκεκριμένη εργασία, εξαρτάται σε μεγάλο βαθμό από την ποιότητα και το πλήθος των δεδομένων με τα οποία 'εκπαιδεύεται' ο ταξινομητής στην εκάστοτε κλάση. Έτσι, προκειμένου να αξιολογηθούν περαιτέρω οι αποδόσεις των 3 ταξινομητών, οι διάφορες ταξινομήσεις υλοποιήθηκαν και με διαφορετικό αριθμό δεδομένων εκπαίδευσης.

Φυσικά τα δεδομένα εκπαίδευσης που εισάγονται στον ταξινομητή, προκύπτουν από τα δεδομένα Ground Truth και έχουν διαφορετικό αριθμό για κάθε εικόνα και κλάση. Έτσι επιλέχθηκαν 6 ποσοστά, επί των δεδομένων Ground Truth κάθε εικόνας, από τα οποία προέκυψε και ο αριθμός των δεδομένων εκπαίδευσης σε κάθε ταξινόμηση. Τα εικονοστοιχεία-δείγματα που περιλαμβάνονται σε αυτό το σύνολο, επιλέγονται τυχαία και δεν ανήκουν σε συγκεκριμένη κλάση κάθε φορά. Τα ποσοστά αυτά είναι : 80%, 50%, 30%, 15%, 5%, 3%, 1% και 0,5% , με συνέπεια κάθε ταξινόμηση σε αυτή την εργασία (βλ. Πίνακας 5.1) να υλοποιείται και με τα 6 αυτά ποσοστά των δεδομένων εκπαίδευσης. Πάντως σε κάθε πίνακα αποτελεσμάτων αναγράφεται και ο απόλυτος αριθμός των δειγμάτων εκπαίδευσης.

5.2.3 Αριθμός κλάσεων

Ως επί των πλείστων, τα δεδομένα Ground Truth (και άρα οι κλάσεις) χρησιμοποιήθηκαν ολόκληρα – χωρίς περικοπές κλάσεων- στις εφαρμοζόμενες ταξινομήσεις. Όμως , όσον αφορά τα SFB δεδομένα και ειδικότερα τις ταξινομήσεις που εφαρμόστηκαν και στις 7 εικόνες SFB, επιβλήθηκε μείωση στις κλάσεις των δεδομένων. Συγκεκριμένα το σύνολο των ψηφιοποιημένων κλάσεων για τις 7 εικόνες, είναι 98. Μετά από έλεγχο των φασματικών υπογραφών, οι κλάσεις συνενώθηκαν και έτσι δημιουργήθηκαν άλλα 3 σετ δεδομένων Ground Truth, με 54 29 και 9 κλάσεις αντίστοιχα.

5.2.4 Καθορισμός μέτρησης χρονικής απόδοσης ταξινομητή

Παρακάτω, κατά την αξιολόγηση της απόδοσης των τριών ταξινομητών μελέτης παρουσιάζονται πίνακες και γραφήματα απόδοσης χρόνου. Εξάλλου, η απόδοση ενός αλγορίθμου/ταξινομητή δεν κρίνεται μόνο από την ακρίβεια της όποιας διαδικασίας που υλοποιεί αλλά και από τον χρόνο που καταλαμβάνει αυτή η διαδικασία.

Εδώ, στην συγκεκριμένη εργασία και όσον αφορά το πρόβλημα της ταξινόμησης εικόνας ως χρονική απόδοση για τους 3 ταξινομητές, θα θεωρείται το χρονικό διάστημα που καλύπτει τη διαδικασία ορισμού του αριθμού των Κυριών Συνιστωσών

(PCA reduction), τη διαδικασία εκμάθησης του αλγορίθμου (training phase), αλλά και τη διαδικασία του ελέγχου σφάλματος (testing phase). Βεβαίως, ο ταξινομητής CNN διαφέρει πολύ από τους ταξινομητές SVM, τόσο ως προς την αρχιτεκτονική του όσο και ως προς τις διάφορες παραμέτρους που μπορούν να επηρεάσουν την χρονική του απόδοση. Συγκεκριμένα, ο τρόπος με τον οποίο υλοποιείται η διαδικασία εκμάθησης μπορεί να διαφέρει από εφαρμογή σε εφαρμογή πάνω σε 2 βασικές παραμέτρους: στο κατώφλι προόδου (improvement threshold) και στον αριθμό των επαναλήψεων (epochs) στον οποίο τερματίζεται η επαναληπτική διαδικασία εκμάθησης (Αναλυτικότερα βλ. Κεφάλαιο 3.3).

Επί της ουσίας λοιπόν, η χρονική απόδοση του νευρωνικού δικτύου CNN είναι ένας συνδυασμός αυτών των δύο παραμέτρων*. Το κατώφλι προόδου (improvement threshold), καθορίζει το επιθυμητό 'βήμα' προόδου της ελαχιστοποίησης σφάλματος και φυσικά μικρότερο 'βήμα' συνεπάγεται με περισσότερες επαναλήψεις και πιο χρονοβόρα εφαρμογή. Αντίστοιχα, και εφόσον ο αλγόριθμος δεν προσεγγίζει τα επιθυμητά όρια του σφάλματος, ο αριθμός των επαναλήψεων (epochs) καθορίζει άμεσα πότε θα περαιωθεί η κάθε επιμέρους εφαρμογή. Για την συγκεκριμένη εργασία και με δεδομένο έναν ισχυρό υπολογιστή εργασίας, το 'κατώφλι προόδου' ορίστηκε ίσο με 0,005% ενώ ο αριθμός των 'εποχών' ίσος με 300.

Οι παραπάνω τιμές ορίστηκαν με εμπειρικό τρόπο και μετά από μελέτη βιβλιογραφικών πηγών, με γνώμονα πρώτα την μεγαλύτερη ελαχιστοποίηση του σφάλματος και έπειτα την χρονική απόδοση του αλγορίθμου.

** για λόγους απλούστευσης. Στην πραγματικότητα η χρονική απόδοση ενός οποιουδήποτε ΤΝΔ μπορεί να επηρεαστεί από έναν μεγάλο αριθμό παραγόντων, που είναι αδύνατο να περιγραφούν πλήρως (αρχιτεκτονική επιπέδων, αριθμός νευρώνων, ρυθμός εκμάθησης, αλγόριθμος εκπαίδευσης κ.α.)*

5.2.5 Τεχνικά Χαρακτηριστικά Υπολογιστή Εργασίας

Παραπάνω αναφέρεται η σημασία της χρονικής απόδοσης ενός αλγορίθμου που σκοπό έχει την ανάλυση δεδομένων. Η ταχύτητα όμως με την οποία περαιώνεται μια τέτοια εφαρμογή, εξαρτάται πλήρως από το υπολογιστικό σύστημα το οποίο υποβάλλεται σε αυτή. Ειδικότερα, η εφαρμογή αλγορίθμων μηχανικής μάθησης στον τομέα της όρασης υπολογιστών, έχει αποδειχθεί ότι επιφέρει μεγάλο υπολογιστικό φόρτο και ως εκ τούτου απαιτεί αξιόπιστους ηλεκτρονικούς υπολογιστές εργασίας υψηλών αποδόσεων.

Με βάση τα παραπάνω οι ταξινομήσεις που εφαρμόστηκαν κατά τη διαδικασία περαίωσης της εργασίας, υλοποιήθηκαν με τη χρήση ενός Η/Υ υψηλών αποδόσεων, του Εργαστηρίου Τηλεπισκόπησης του ΕΜΠ με τα ακόλουθα τεχνικά χαρακτηριστικά:

- Επεξεργαστής: Intel Core i7-3820 Quad core 3.6 GHz
- Φυσική μνήμη RAM: 64 GB DDR3
- Κάρτα γραφικών: NVIDIA Ge Force GTX 660 2 GB

5.3 Αποτελέσματα Βιβλιογραφικών Δεδομένων

Στο 4^ο κεφάλαιο περιγράφονται αναλυτικά κάποια υπερφασματικά δεδομένα εικόνων, τα οποία αποτελούν αντικείμενο μελέτης σύγχρονων μεθόδων και αλγορίθμων πάνω στο πρόβλημα της ψηφιακής ταξινόμησης εικόνας.

Πάνω στα δεδομένα αυτά, τα οποία αποτελούν δείκτη σύγκρισης στη διεθνή βιβλιογραφία, εφαρμόστηκαν 3 προσεγγίσεις ταξινόμησης εικόνας, τα αποτελέσματα των οποίων παρουσιάζονται παράλληλα στους παρακάτω πίνακες και γραφήματα.

Οι 2 πρώτες προσεγγίσεις ταξινομούν τα δεδομένα με τη χρήση SVM (Support Vector Machine) αλγορίθμων, ενώ στην 3^η προσέγγιση εφαρμόστηκε ένας ταξινομητής βαθιάς εκμάθησης (deep learning) μέσω ενός νευρωνικού δικτύου τύπου CNN (Convolutional Neural Network). (Αναλυτικότερα βλ. Κεφάλαιο 3)

Τα δεδομένα που χρησιμοποιούνται σε αυτό το κεφάλαιο αναγράφονται στον παρακάτω πίνακα (Πίνακας 5.2) μαζί με τον αριθμό των καναλιών τους και τον αριθμό των κλάσεων που περιέχουν τα αντίστοιχα δεδομένα Ground Truth.

Σημειώνεται πως για τις εικόνες *Moffet Field 1* και *Moffet Field 3* τα δεδομένα *Ground Truth* δημιουργήθηκαν για τους σκοπούς της εργασίας και δεν συναντώνται αλλού στην διεθνή βιβλιογραφία.

Όνομα Εικόνας	Δέκτης Λήψης	Μέγεθος Εικόνας	Αριθμός Καναλιών	Αριθμός Κλάσεων
<i>Indian Pines</i>	AVIRIS	145x145	185	16
<i>Salinas</i>	AVIRIS	512x217	185	16
<i>Botswana</i>	HYPERION	1476x256	145	14
<i>Pavia Center</i>	ROSIS	1096x715	102	9
<i>Pavia University</i>	ROSIS	610x340	103	9
<i>Moffet Field Run 1</i>	AVIRIS	512x614	135	16
<i>Moffet Field Run 3</i>	AVIRIS	512x614	135	29

Πίνακας 5.2: Χαρακτηριστικά στοιχεία των βιβλιογραφικών δεδομένων που χρησιμοποιήθηκαν στην εργασία

Indian Pines

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της υπερφασματικής εικόνας *Indian Pines* που αποτυπώνει αγροτική περιοχή της βορειοδυτικής Indiana των ΗΠΑ.

Η υπερφασματική εικόνα κατά την απόκτηση της από το διαδίκτυο: [http://www.ehu.es/ccwintco/index.php?title=Hyperspectral Remote Sensing Scenes](http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes) περιείχε 200 κανάλια μετά από εξαίρεση 24 καναλιών μεγάλης απορρόφησης νερού (224 αρχικά - με δέκτη τον AVIRIS). Παρ' όλα αυτά, εξαιρέθηκαν άλλα 15 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 185 κανάλια. Τα επιπλέον εξαιρούμενα κανάλια είναι τα: [1-2],[101-105],[141-145],[198-200]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=28, G=17, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις.



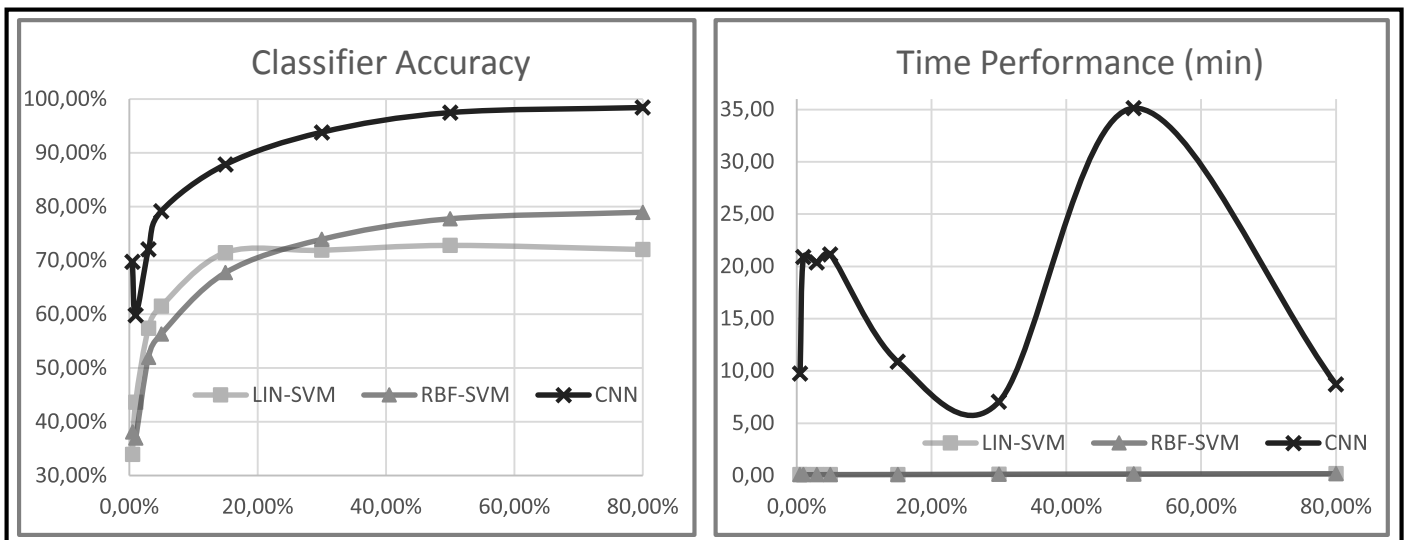
Εικόνα 5.1: Έγχρωμο σύνθετο της εικόνας *Indian Pines*, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις περιλαμβάνουν και αποτελέσματα με μείωση του αριθμού των κυρίων συνιστωσών και αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης .

Παρακάτω παρουσιάζονται τα αποτελέσματα και οι ταξινομήσεις που αφορούν τις ταξινομήσεις με 55 Κύριες Συνιστώσες και 99,81% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	51	102	307	512	1537	3075	5125	8199
	accuracy	33,92%	43,63%	57,40%	61,50%	71,46%	71,89%	72,80%	72,04%
	time (min)	0,08	0,08	0,08	0,08	0,09	0,11	0,13	0,17
RBF	accuracy	38,10%	37,01%	52,00%	56,34%	67,72%	73,92%	77,75%	78,97%
	time (min)	0,08	0,08	0,08	0,09	0,10	0,12	0,14	0,17
CNN	accuracy	69,74%	59,83%	72,03%	79,13%	87,82%	93,81%	97,46%	98,40%
	time (min)	9,76*	20,88*	20,36*	21,14*	10,88	7,07	35,14*	8,72

Πίνακας 5.3: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Indian Pines για 55 Κ.Σ.

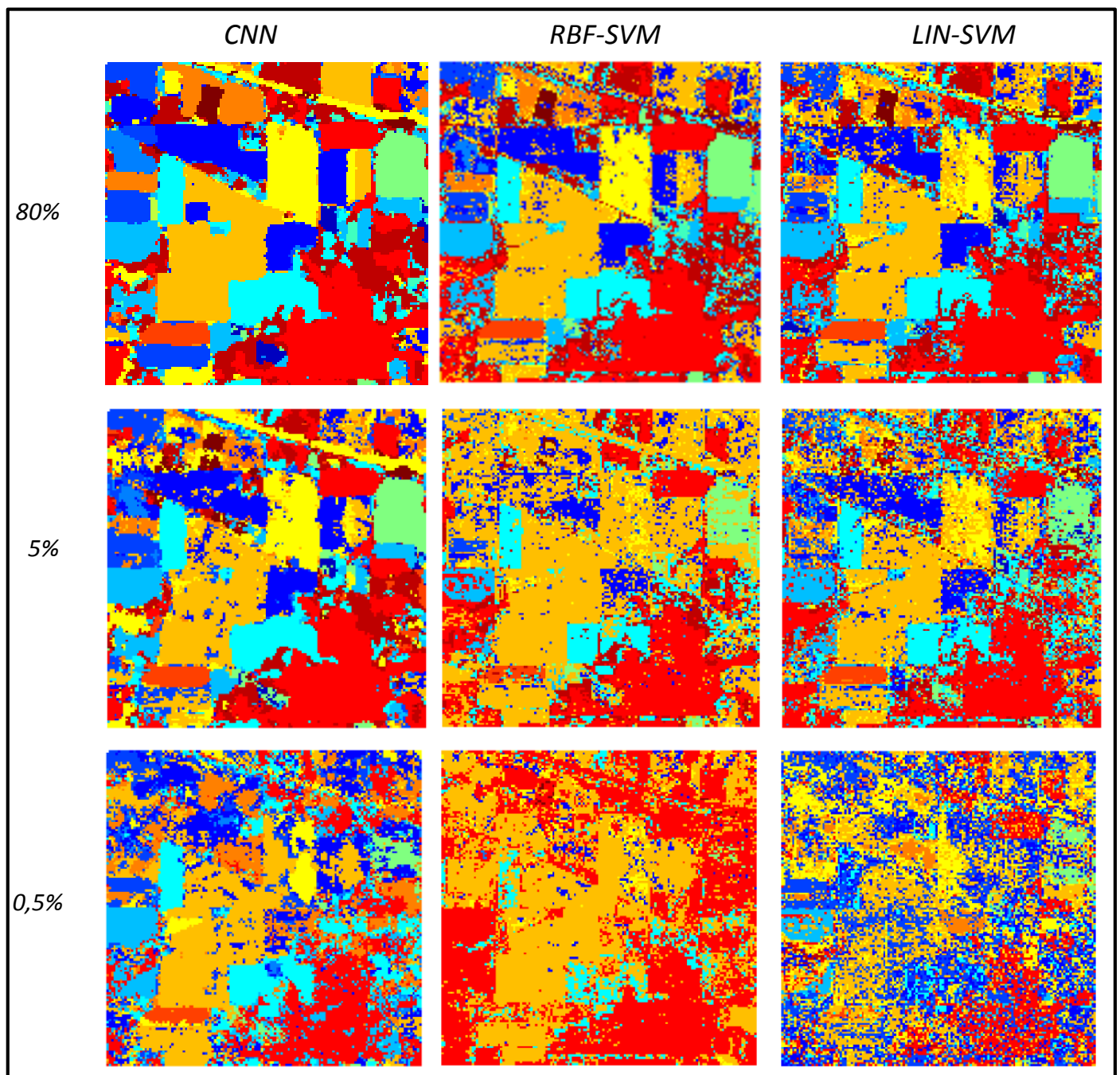


Εικόνα 5.2: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Indian Pines για 55 Κ.Σ.

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει κατά πολύ καλύτερες ακρίβειες από τους ταξινομητές SVM σε όλα τα ποσοστά δεδομένων εκπαίδευσης. Συγκεκριμένα, επιτυγχάνει καλύτερες ακρίβειες κατά περίπου 20% και από τους δύο ταξινομητές SVM σε όλες τις περιπτώσεις δεδομένων εκπαίδευσης. Επιπλέον, ο ταξινομητής CNN κυμαίνεται από περίπου 70% έως 98% σε ακρίβεια ενώ ο ταξινομητής RBF-SVM από 37% έως 79%. Σημειώνεται ακόμα, ότι οι αποδόσεις των δύο SVM ταξινομητών δεν διαφέρουν πολύ και μάλιστα σε 4 περιπτώσεις δεδομένων εκπαίδευσης (15%, 5%, 3%, 1% και 0,5%) ο Linear-SVM επιτυγχάνει καλύτερες αποδόσεις. Τέλος, παρατηρείται ότι αντιθέτως με τους άλλους δυο ταξινομητές ο Linear-SVM επιτυγχάνει περίπου το μέγιστο της απόδοσής του στο 15% του ποσοστού εκπαίδευσης.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, με αποδόσεις μερικών δευτερολέπτων. Βεβαίως στις περιπτώσεις που ο CNN ελαχιστοποίησε το σφάλμα του πριν το πέρας του ορισμένου ορίου (για τις περιπτώσεις 30% και 80%), η χρονική του απόδοση φτάνει τα 7-9 λεπτά.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για τις παραπάνω ταξινομήσεις:

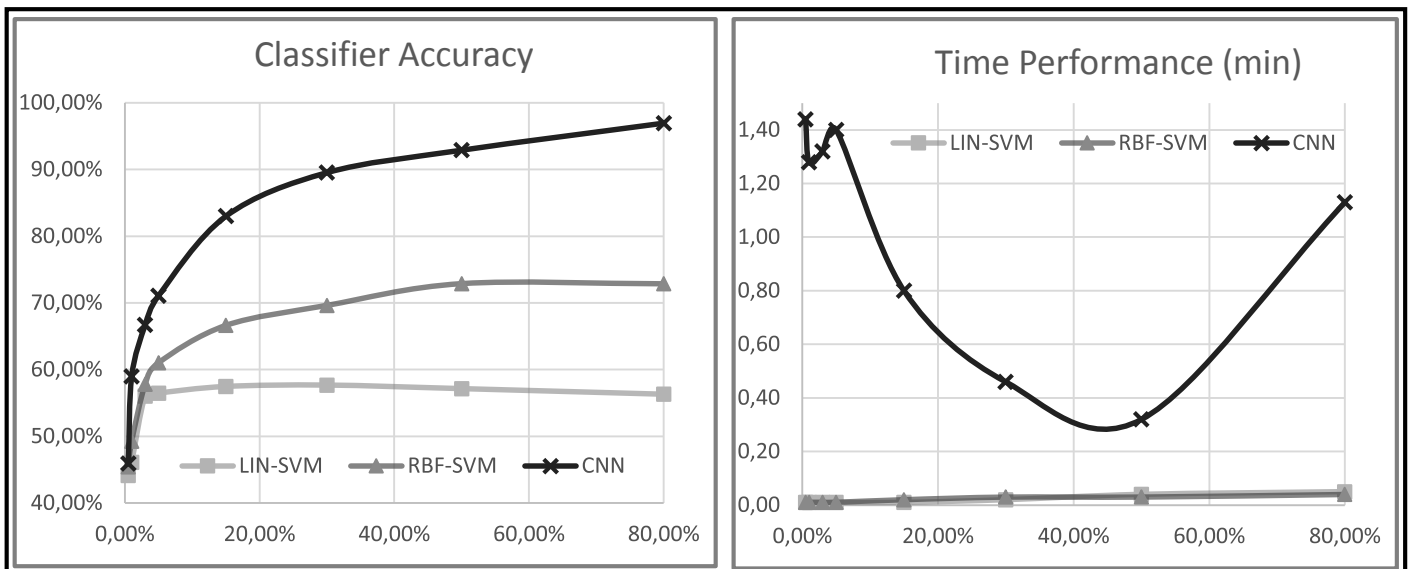


Εικόνα 5.3: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Indian Pines για 55 Κ.Σ.

Παρακάτω παρουσιάζονται τα αποτελέσματα και οι ταξινομήσεις που αφορούν τις ταξινομήσεις με 10 Κύριες Συνιστώσες και 97,21% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	51	102	307	512	1537	3075	5125	8199
	accuracy	44,15%	46,11%	56,04%	56,49%	57,49%	57,69%	57,16%	56,32%
	time (min)	0,01	0,01	0,01	0,01	0,01	0,02	0,04	0,05
RBF	accuracy	45,45%	49,26%	57,81%	61,03%	66,63%	69,62%	72,89%	72,86%
	time (min)	0,01	0,01	0,01	0,01	0,02	0,03	0,03	0,04
CNN	accuracy	45,96%	58,97%	66,70%	71,10%	83,01%	89,53%	92,90%	96,95%
	time (min)	1,44*	1,28*	1,32*	1,40*	0,80	0,46	0,32	1,13

Πίνακας 5.4: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Indian Pines για 10 Κ.Σ.

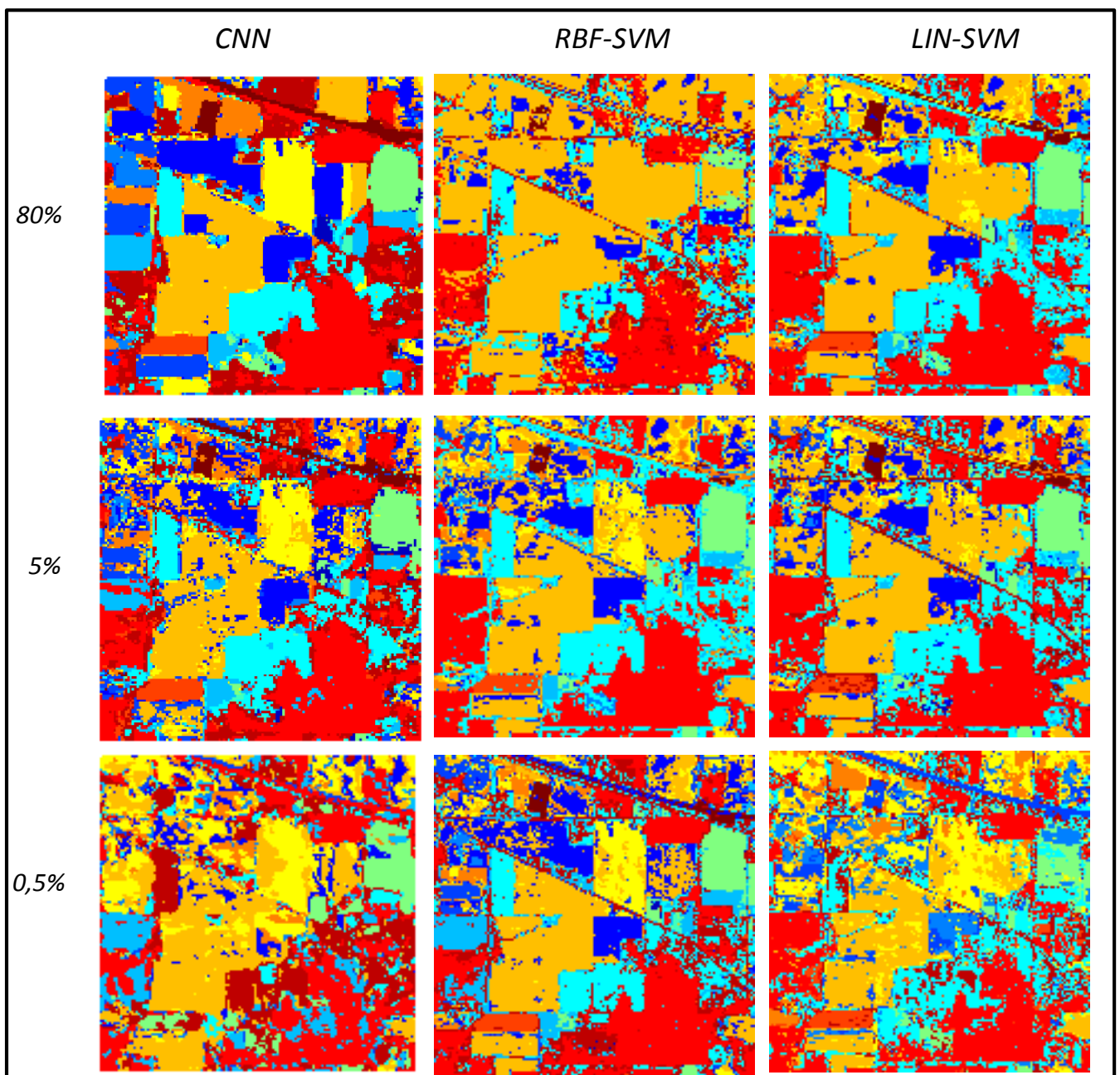


Εικόνα 5.4: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Indian Pines για 10 Κ.Σ.

Και σε αυτές τις περιπτώσεις, με μειωμένο αριθμό Κυρίων Συνιστωσών, ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους δύο ταξινομητές SVM σε όλες τις περιπτώσεις δεδομένων εκπαίδευσης. Συγκεκριμένα επιτυγχάνει ακρίβειες μεγαλύτερες κατά 20% στα μεγάλα ποσοστά δεδομένων εκπαίδευσης και περίπου 10% στα μικρότερα ποσοστά. Παράλληλα ο ταξινομητής RBF-SVM είναι ακριβέστερος του Linear-SVM για όλα τα ποσοστά εκπαίδευσης. Ακόμα, ο Linear-SVM επιτυγχάνει περίπου το μέγιστο της απόδοσής του από το 3% του ποσοστού εκπαίδευσης.

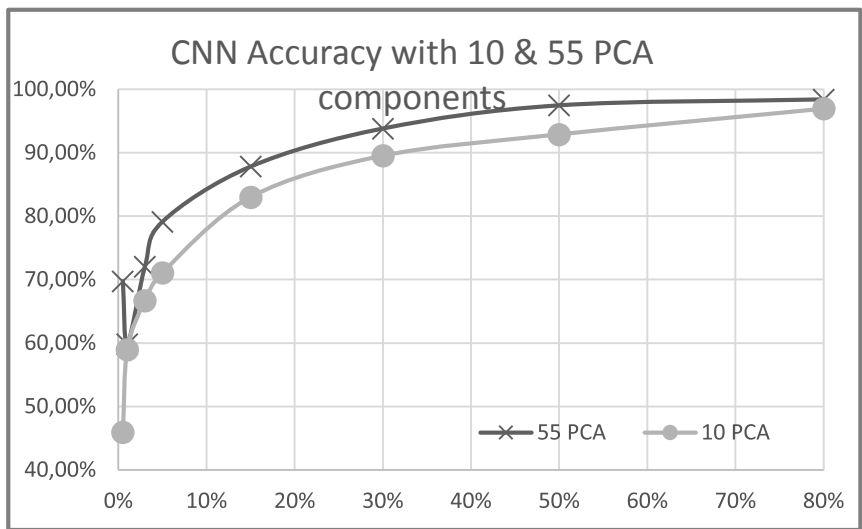
Όσον αφορά τη χρονική απόδοση, σε αυτά τα δεδομένα και με μειωμένες στις 10 τις Κ.Σ. ο CNN σε όλες τις περιπτώσεις αποδίδει σε λιγότερα από 2 λεπτά. Φυσικά, οι ταξινομητές SVM αποδίδουν και πάλι ταχύτερα, όμως η χρονική απόδοση του CNN έχει ελαττωθεί σε σχέση με την προηγούμενη εφαρμογή έως και 20 λεπτά.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για τις ταξινομήσεις με μειωμένο αριθμό Κ.Σ.:

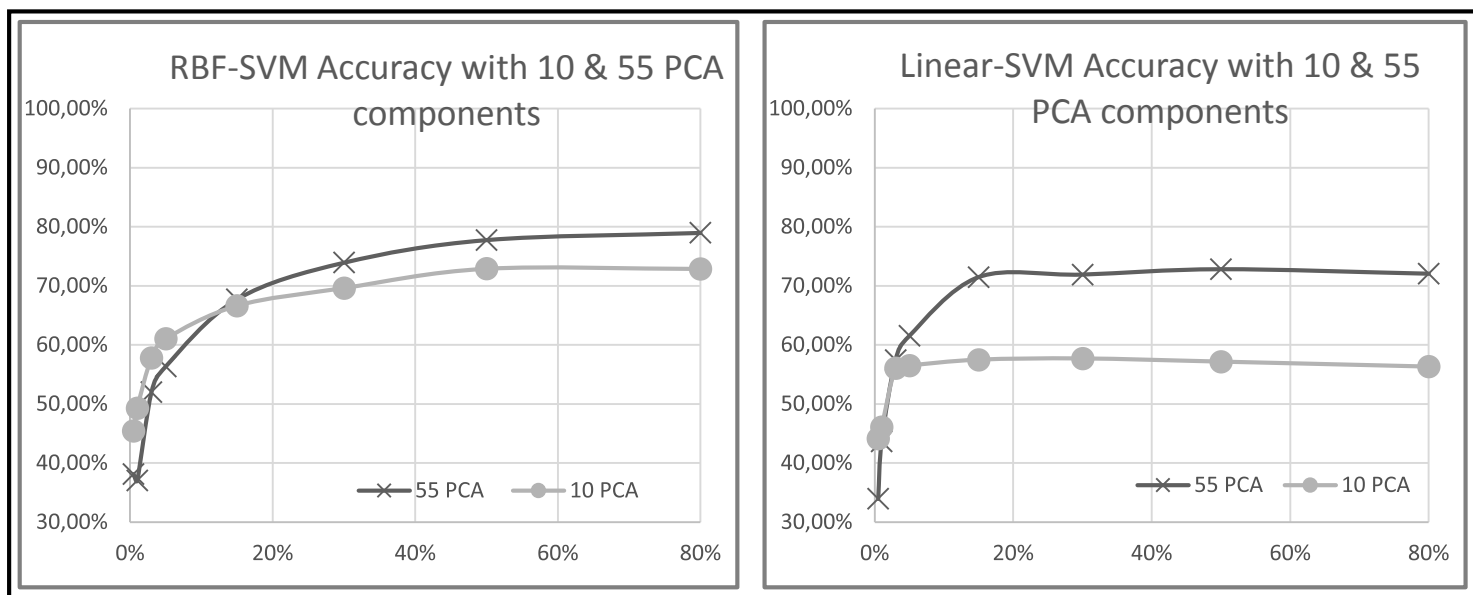


Εικόνα 5.5: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Indian Pines για 10 Κ.Σ.

Τέλος, παρουσιάζονται και τα αντίστοιχα διαγράμματα με τις συνδυαστικές αποδόσεις ακρίβειας για κάθε ταξινομητή (55 & 10 PCA):



Εικόνα 5.6: Συγκριτικό γράφημα ακριβειών του CNN για τα δεδομένα Indian Pines, με 55 Κ.Σ. και 10 Κ.Σ.



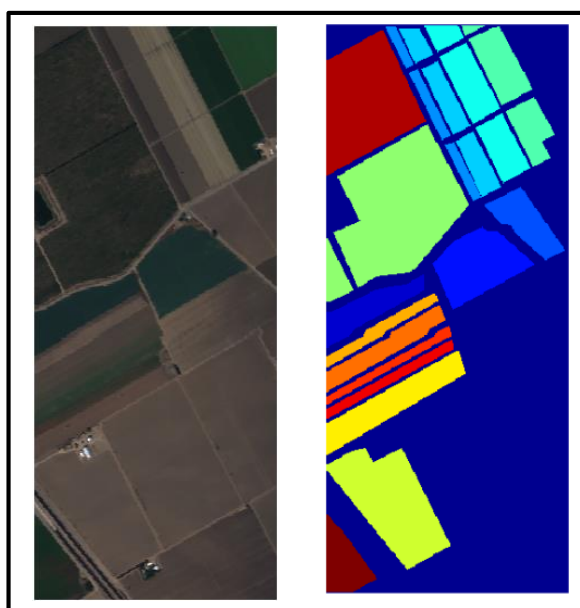
Εικόνα 5.7: Συγκριτικά γραφήματα ακριβειών των SVM για τα δεδομένα Indian Pines, με 55 Κ.Σ. και 10 Κ.Σ.

Όπως ήταν αναμενόμενο, και στις 3 περιπτώσεις ταξινομητών οι εφαρμογές με 55 Κ.Σ. αποφέρουν καλύτερα αποτελέσματα. Η μεγαλύτερη διαφορά απόδοσης συναντάται στον Linear-SVM, ενώ το νευρωνικό δίκτυο CNN επιτυγχάνει μεγάλα ποσοστά ακρίβειας και για 55 και για 10 Κ.Σ.

Salinas

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της υπερφασματικής εικόνας *Salinas* που αποτυπώνει αγροτική περιοχή της πεδιάδας Salinas στην California των ΗΠΑ.

Η υπερφασματική εικόνα κατά την απόκτηση της από το διαδίκτυο: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes περιείχε 224 κανάλια (δέκτης ο AVIRIS). Παρ'όλα αυτά, εξαιρέθηκαν 39 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 185 κανάλια. Τα 39 αυτά εξαιρούμενα κανάλια είναι τα: [1-3],[33],[97],[104-114],[150-168],[221-223]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=28, G=17, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις.

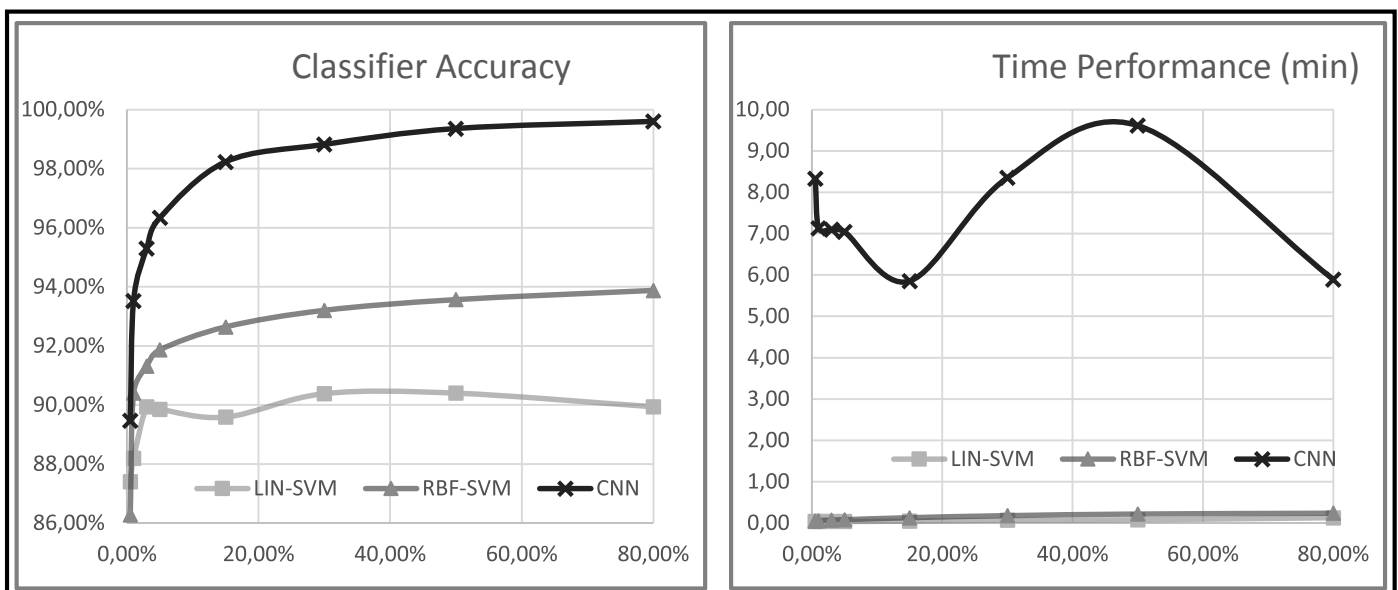


Εικόνα 5.8: Έγχρωμο σύνθετο της εικόνας Salinas, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. **Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,97% ποσοστό της υπερφασματικής πληροφορίας.**

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	271	541	1624	2706	8119	16239	27065	43303
	accuracy	87,41%	88,19%	89,94%	89,86%	89,59%	90,38%	90,40%	89,94%
	time (min)	0,04	0,04	0,04	0,04	0,05	0,07	0,08	0,13
RBF	accuracy	86,28%	90,40%	91,31%	91,86%	92,64%	93,20%	93,56%	93,88%
	time (min)	0,05	0,06	0,07	0,08	0,13	0,18	0,22	0,24
CNN	accuracy	89,47%	93,53%	95,30%	96,34%	98,23%	98,82%	99,35%	99,59%
	time (min)	8,33*	7,13*	7,10*	7,04*	5,85	8,35*	9,61*	5,89

Πίνακας 5.5: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Salinas

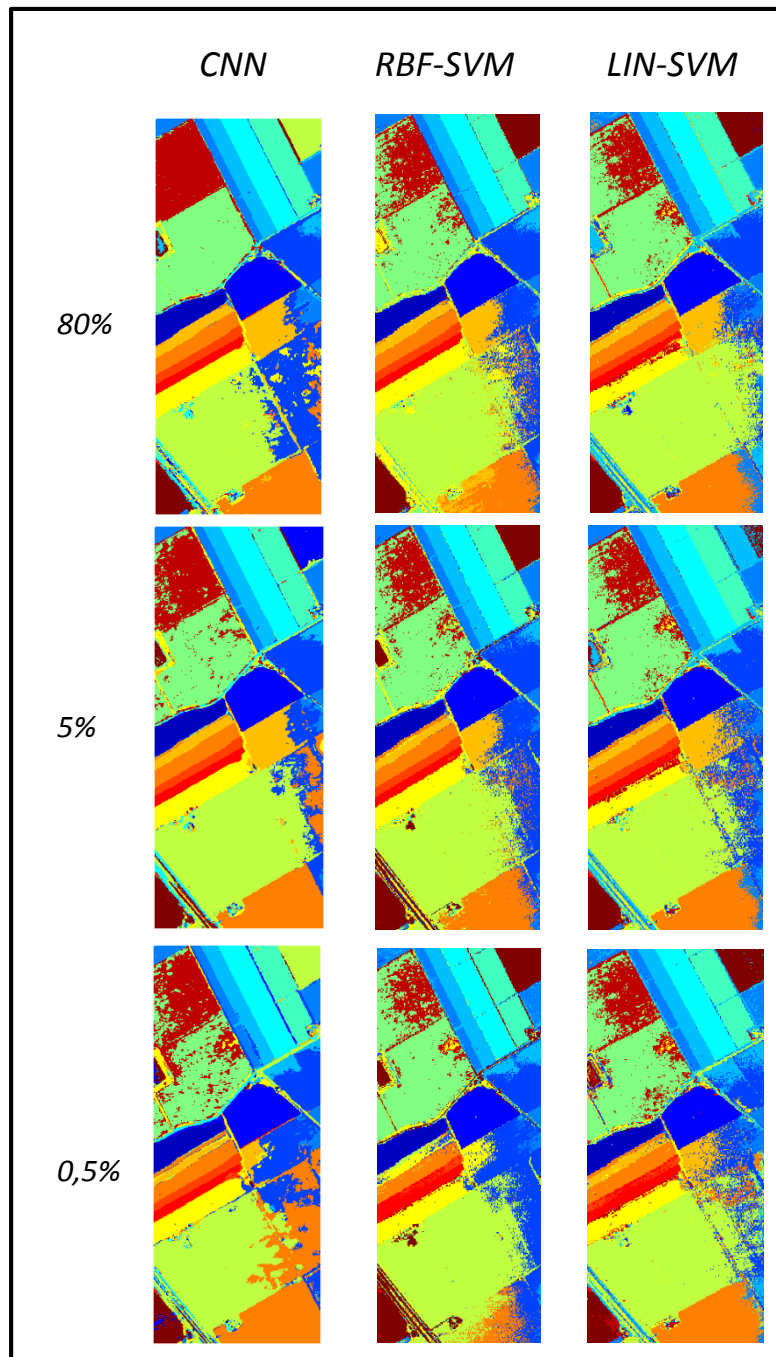


Εικόνα 5.9: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Salinas

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM σε όλα τα ποσοστά δεδομένων εκπαίδευσης. Συγκεκριμένα, επιτυγχάνει καλύτερες ακρίβειες κατά περίπου 6-5% και από τους δύο ταξινομητές SVM για τις περιπτώσεις πολλών δεδομένων εκπαίδευσης και κατά περίπου 3-4% στις υπόλοιπες. Επιπλέον, ο ταξινομητής CNN κυμαίνεται από περίπου 70% έως 99% σε ακρίβεια ενώ ο ταξινομητής RBF-SVM από 86% έως 93%. Σημειώνεται ακόμα, ότι οι αποδόσεις των δύο SVM ταξινομητών έχουν υπολογίσιμη διαφορά (2%-4%) ενώ η απόδοση του Linear-SVM φαίνεται να μην επηρεάζεται από το ποσοστό των δεδομένων εκπαίδευσης (εκτός της περίπτωσης του 0,5%).

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, με την απόδοσή τους να είναι μερικά δευτερόλεπτα, ενώ του CNN από 6 έως 10 λεπτά περίπου.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

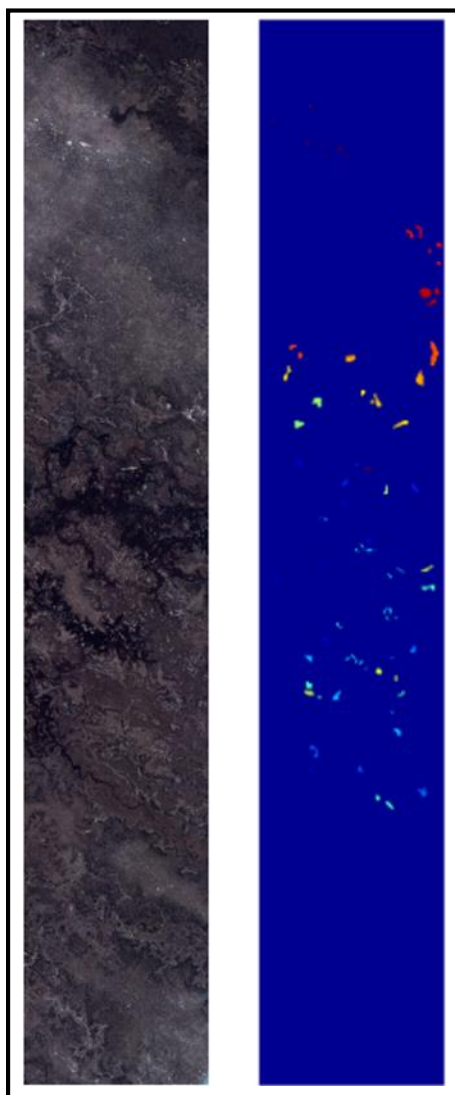


Εικόνα 5.10: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Salinas

Botswana

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της υπερφασματικής εικόνας *Botswana* που αποτυπώνει περιοχή του δέλτα Οκαναγκο στην Botswana.

Η υπερφασματική εικόνα κατά την απόκτηση της από το διαδίκτυο: [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral Remote Sensing Scen](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes)
[es](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes) περιείχε 142 κανάλια έπειτα από αφαίρεση 97 καναλιών (242 αρχικά - δέκτης ο HYPERION). Τα κανάλια που αφαιρέθηκαν είναι τα ακόλουθα 145: [10-55, 82-97, 102-119, 134-164, 187-220]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=19, G=7, B=3) καθώς και η Ground Truth εικόνα με τις 14 κλάσεις.



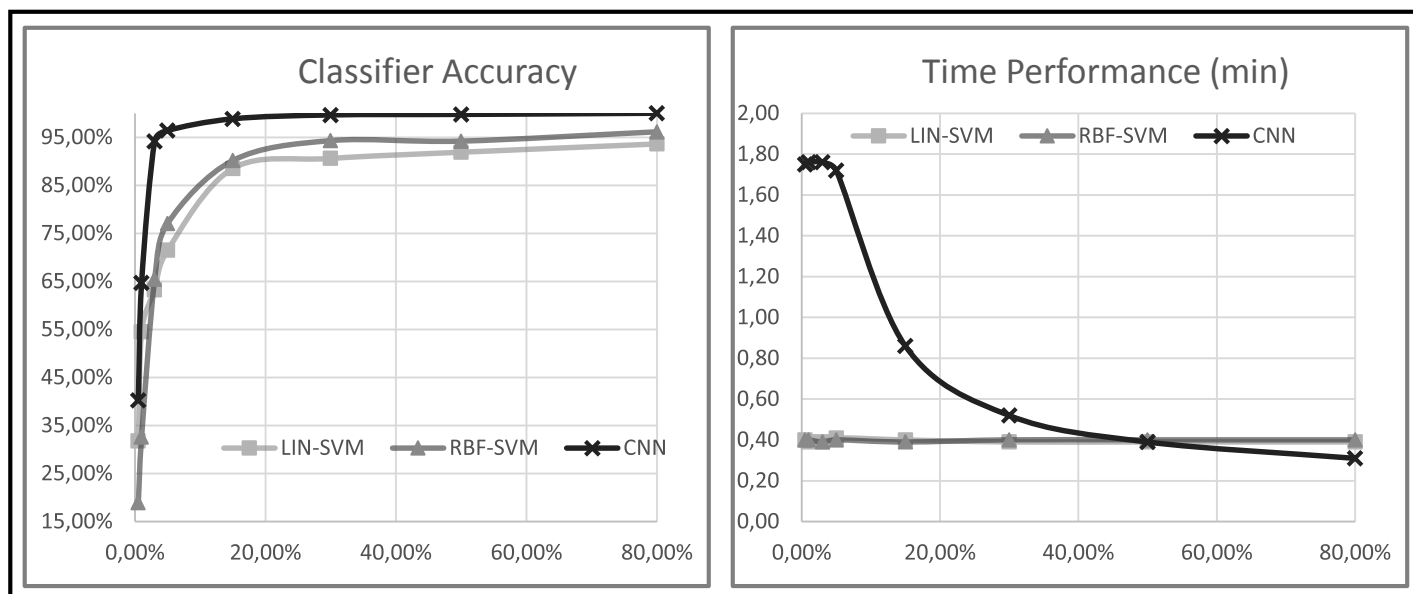
Εικόνα 5.11: Έγχρωμο σύνθετο της εικόνας *Botswana*, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις περιλαμβάνουν και αποτελέσματα με μείωση του αριθμού των κυρίων συνιστωσών και αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης.

Παρακάτω παρουσιάζονται τα αποτελέσματα και οι ταξινομήσεις που αφορούν τις ταξινομήσεις με 25 Κύριες Συνιστώσες και 99,82% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	16	32	97	162	487	974	1624	2598
	accuracy	31,82%	54,49%	63,27%	71,54%	88,55%	90,63%	91,93%	93,65%
	time (min)	0,40	0,39	0,39	0,41	0,40	0,39	0,39	0,39
RBF	accuracy	18,94%	32,57%	65,43%	77,08%	90,18%	94,31%	94,21%	96,15%
	time (min)	0,40	0,40	0,39	0,40	0,39	0,40	0,40	0,40
CNN	accuracy	40,21%	64,70%	94,16%	96,40%	98,84%	99,65%	99,75%	100,00%
	time (min)	1,75*	1,76*	1,76*	1,72*	0,86	0,52	0,39	0,31

Πίνακας 5.6: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Botswana για 25 Κ.Σ.

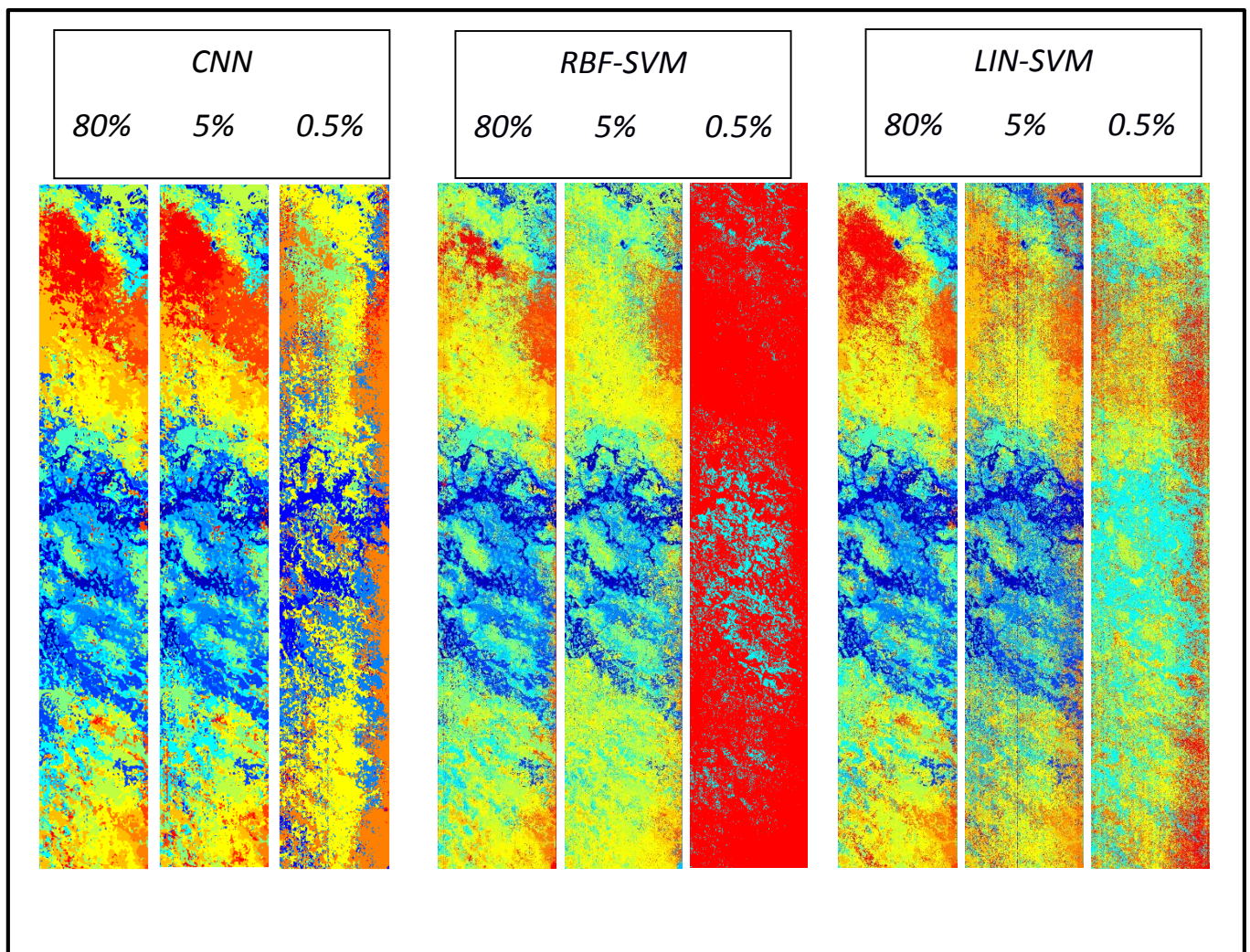


Εικόνα 5.12: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Botswana για 25 Κ.Σ.

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM σε όλα τα ποσοστά δεδομένων εκπαίδευσης. Συγκεκριμένα, επιτυγχάνει καλύτερες ακρίβειες κατά περίπου 10-6%, - από τον δεύτερο σε ακρίβεια ταξινομητή - για τα μεγάλα ποσοστά εκπαίδευσης (15-80%) και 30-15% όσον αφορά τα μικρότερα ποσοστά. Βεβαίως, το χαρακτηριστικό αυτής της εφαρμογής είναι η μεγάλη διακύμανση στις ακρίβειες που επιτυγχάνουν

όλοι οι ταξινομητές, κάτι που οφείλεται στα λίγα δείγματα εκπαίδευσης για τα μικρά ποσοστά, τα οποία και δεν επαρκούν για να επιτευχθεί αξιόπιστη εκμάθηση. Έτσι, ο CNN κυμαίνεται από 40% έως και 100% σε ακρίβεια, ενώ ο RBF ταξινομητής από 19% έως 96%. Η βασική διαφοροποίηση εντοπίζεται για τα ποσοστά εκπαίδευσης 1% και 3% όπου ο CNN παρουσιάζει κατά 30% υψηλότερα ποσοστά από τους SVM ταξινομητές, καταφέροντας έτσι να αποδίδει υψηλά ποσοστά ακρίβειας με λίγο αριθμό δειγμάτων. Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι ταχύτεροι του CNN στις περιπτώσεις που ο τελευταίος έφτασε στο προκαθορισμένο όριο επαναλήψεων. Για τα ποσοστά εκπαίδευσης 80% και 50% που δε συνέβη κάτι τέτοιο όμως, το νευρωνικό δίκτυο επιτυγχάνει χρονικές αποδόσεις ίσες και καλύτερες από αυτές των SVM.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

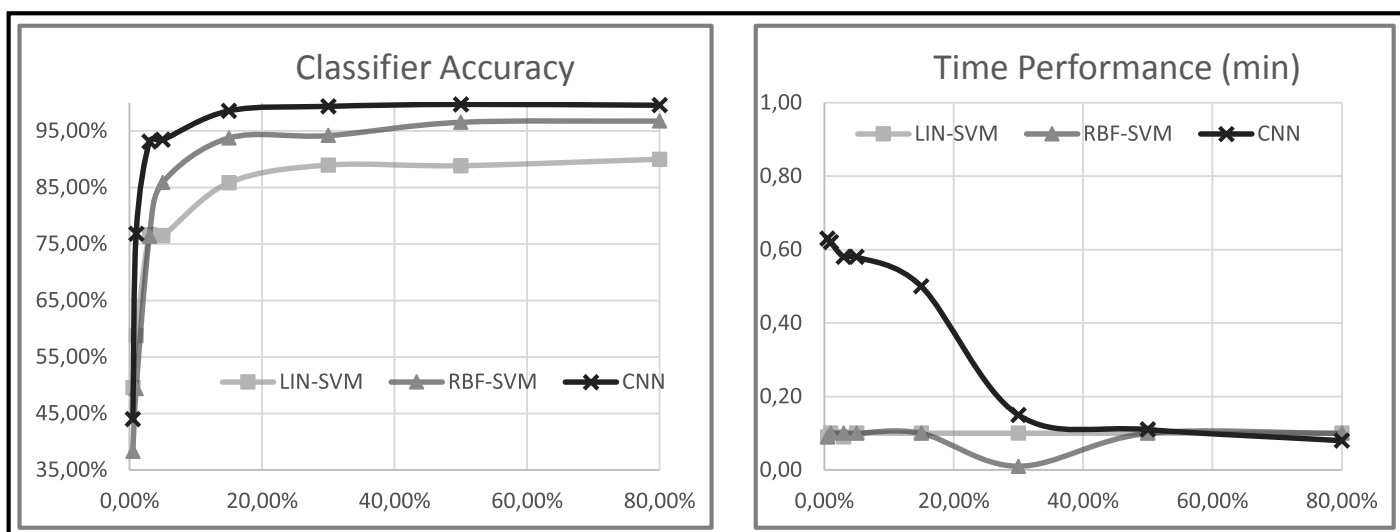


Εικόνα 5.13: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Botswana για 25 Κ.Σ.

Παρακάτω παρουσιάζονται τα αποτελέσματα και οι ταξινομήσεις που αφορούν τις ταξινομήσεις με 10 Κύριες Συνιστώσες και 99,66% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	16	32	97	162	487	974	1624	2598
	accuracy	49,55%	58,76%	76,67%	76,47%	85,83%	88,96%	88,85%	89,98%
RBF	time (min)	0,09	0,10	0,09	0,10	0,10	0,10	0,10	0,10
	accuracy	38,32%	49,39%	76,35%	85,87%	93,77%	94,19%	96,55%	96,76%
CNN	time (min)	0,09	0,10	0,10	0,10	0,10	0,01	0,10	0,10
	accuracy	43,99%	76,79%	93,08%	93,42%	98,55%	99,34%	99,69%	99,54%
	time (min)	0,63	0,62	0,58	0,58	0,50	0,15	0,11	0,08

Πίνακας 5.7: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Botswana για 10 Κ.Σ.

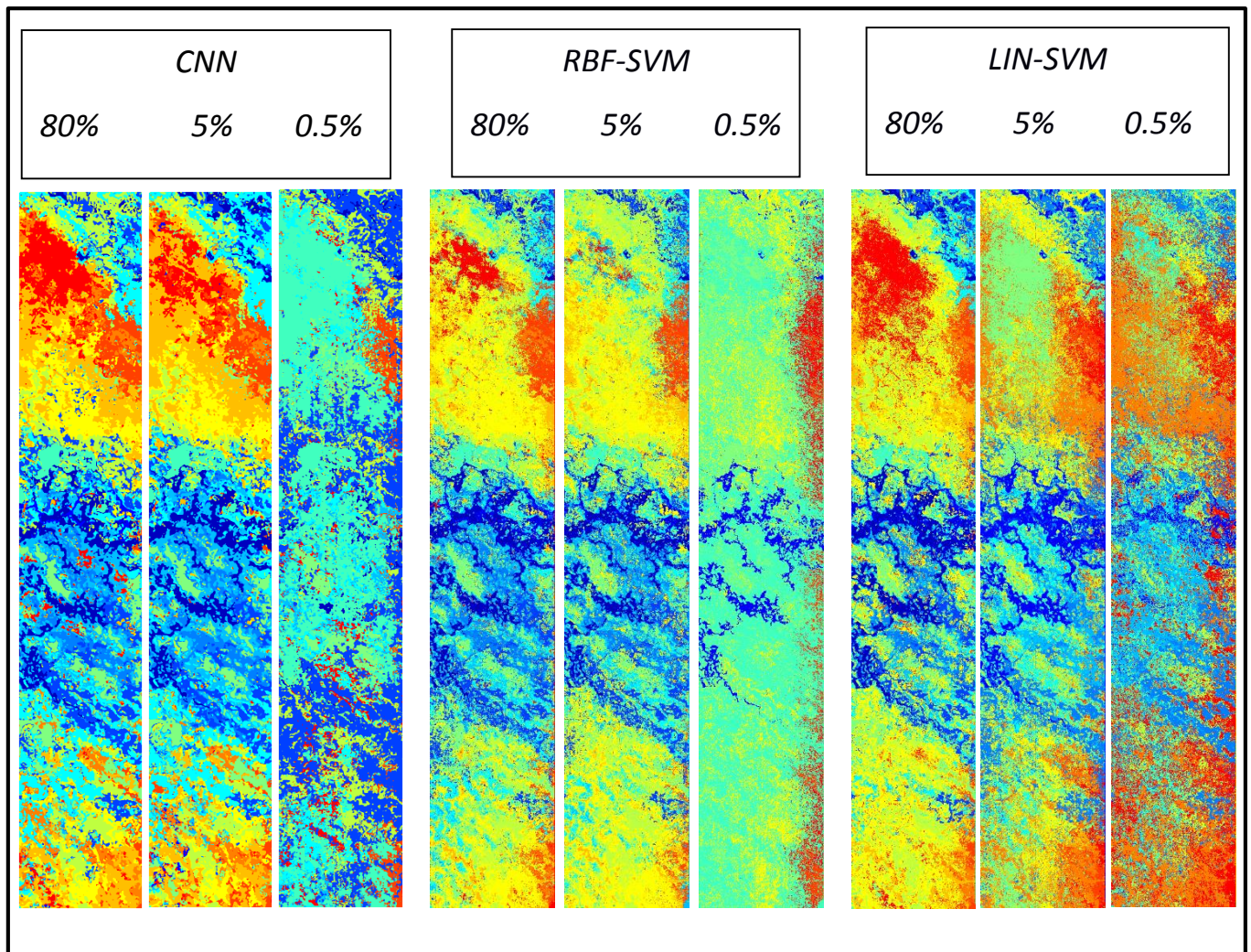


Εικόνα 5.14: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Botswana για 10 Κ.Σ.

Και σε αυτή την εφαρμογή με μειωμένο αριθμό Κ.Σ. ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM, με μικρότερες έως και μηδενικές ωστόσο αποκλίσεις σε όλα τα ποσοστά εκπαίδευσης. Ο CNN κυμαίνεται από 44% έως και 99% σε ακρίβεια, ενώ ο RBF ταξινομητής από 38% έως 97%. Και εδώ, η βασική διαφοροποίηση εντοπίζεται για τα ποσοστά εκπαίδευσης 1% και 3% όπου ο CNN παρουσιάζει κατά 20% υψηλότερα ποσοστά από τους SVM ταξινομητές, καταφέρνοντας έτσι να αποδίδει υψηλά ποσοστά ακρίβειας με λίγο αριθμό δειγμάτων.

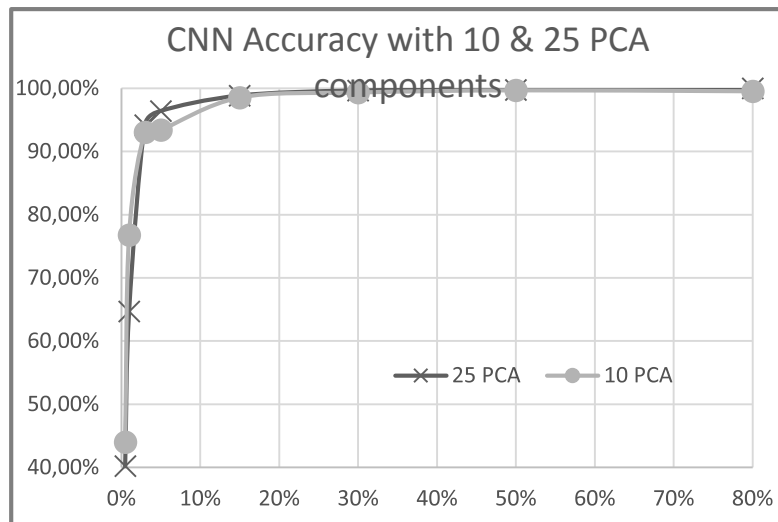
Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι ταχύτεροι του CNN στις περιπτώσεις που ο τελευταίος έφτασε στο προκαθορισμένο όριο επαναλήψεων. Για τα ποσοστά εκπαίδευσης 80%, 50% και 30% που δε συνέβη κάτι τέτοιο όμως, το νευρωνικό δίκτυο επιτυγχάνει χρονικές αποδόσεις περίπου ίσες με αυτές των SVM.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

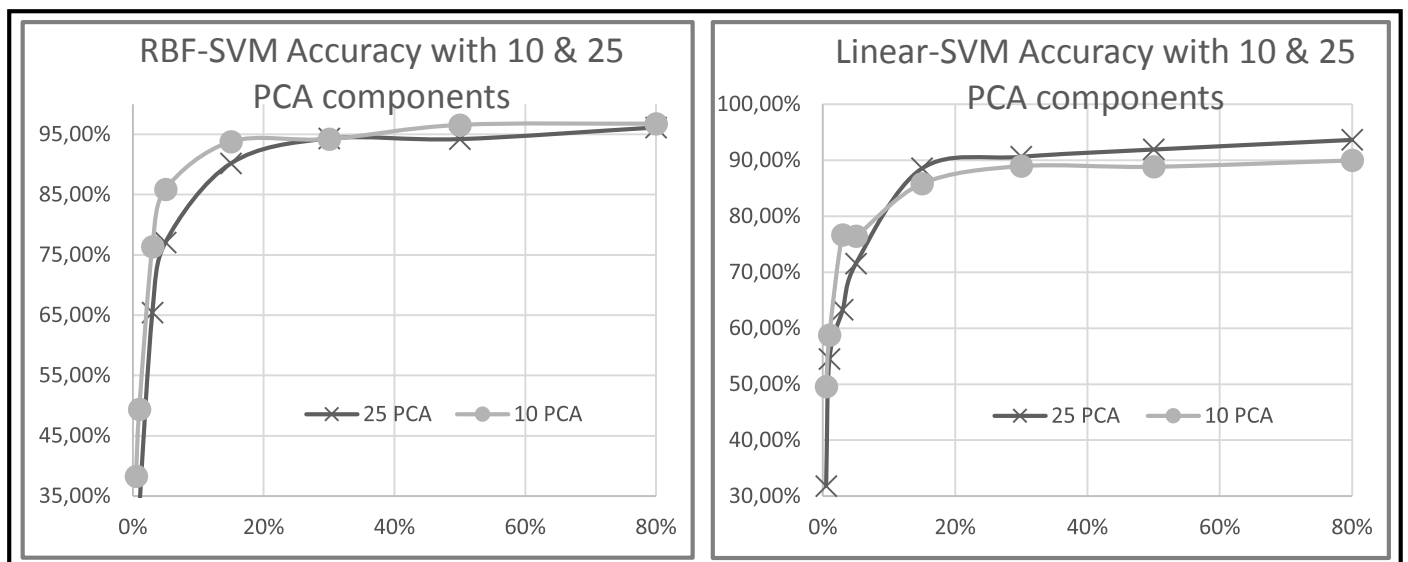


Εικόνα 5.15: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Botswana για 10 Κ.Σ.

Τέλος, παρουσιάζονται και τα αντίστοιχα διαγράμματα με τις συνδυαστικές αποδόσεις ακρίβειας για κάθε ταξινομητή (25 & 10 PCA):



Εικόνα 5.16: Συγκριτικό γράφημα ακριβειών του CNN για τα δεδομένα Botswana, με 25 Κ.Σ. και 10 Κ.Σ.



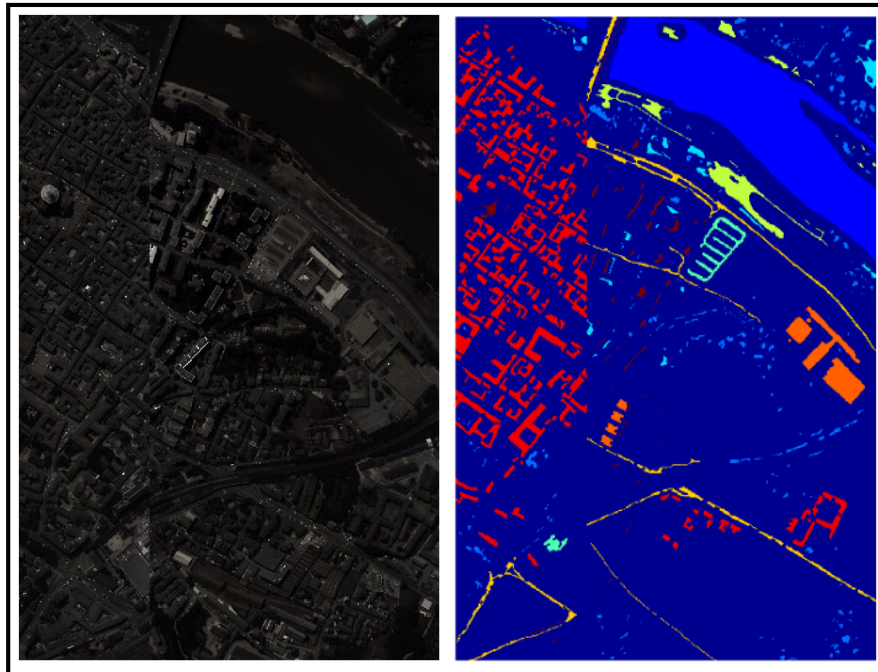
Εικόνα 5.17: Συγκριτικά γραφήματα ακριβειών των SVM για τα δεδομένα Botswana, με 25 Κ.Σ. και 10 Κ.Σ.

Όπως ήταν αναμενόμενο, και στις 3 περιπτώσεις ταξινομητών οι εφαρμογές με 25 Κ.Σ. αποφέρουν καλύτερα αποτελέσματα. Η μεγαλύτερη διαφορά απόδοσης συναντάται στον Linear-SVM, ενώ το νευρωνικό δίκτυο CNN επιτυγχάνει μεγάλα ποσοστά ακρίβειας και για 55 και για 10 Κ.Σ.

Pavia Center

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας *Pavia Center* που αποτυπώνει την πόλη Pavia της βόρειας Ιταλίας.

Η υπερφασματική εικόνα αξιοποιήθηκε με όλη την υπερφασματική πληροφορία που ήταν διαθέσιμη κατά την απόκτηση της από το διαδίκτυο: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes (102 κανάλια – με δέκτη τον ROSIS). Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=22, G=15, B=9) καθώς και η Ground Truth εικόνα με τις 9 κλάσεις.

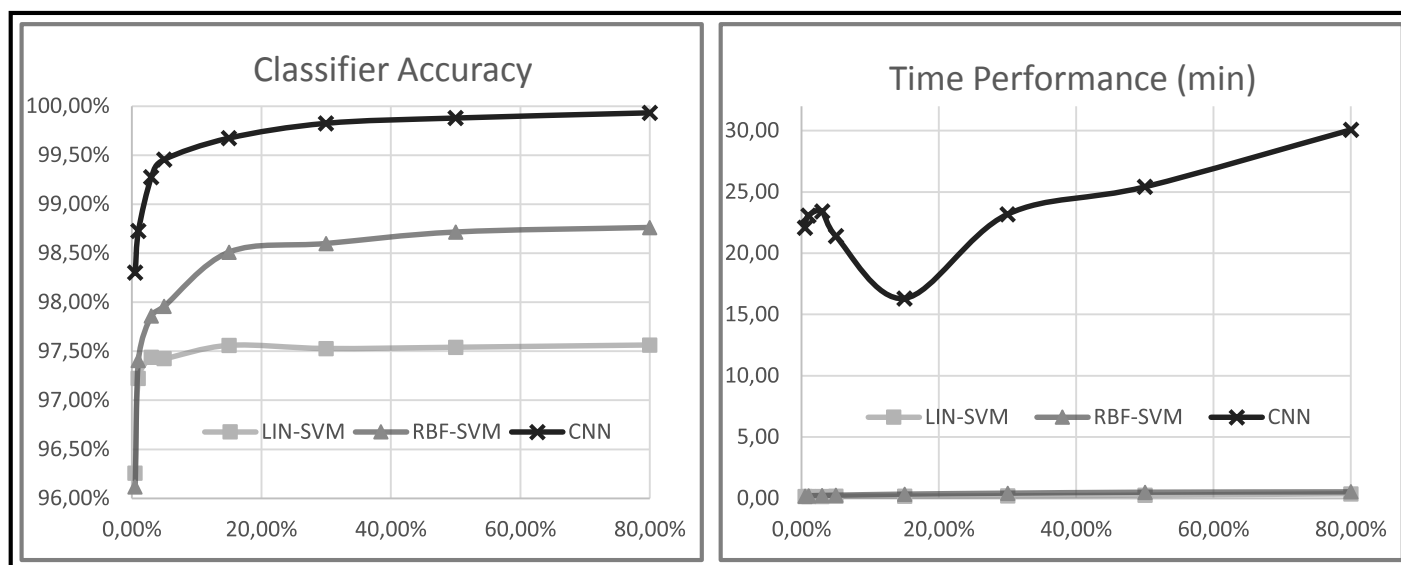


Εικόνα 5.18: Έγχρωμο σύνθετο της εικόνας Pavia Center, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,85% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	37	75	224	373	1118	2237	3728	5965
	accuracy	96,26%	97,22%	97,44%	97,42%	97,56%	97,53%	97,54%	97,56%
	time (min)	0,16	0,16	0,16	0,17	0,18	0,20	0,26	0,37
RBF	accuracy	96,12%	97,41%	97,86%	97,96%	98,51%	98,60%	98,71%	98,76%
	time (min)	0,18	0,20	0,23	0,25	0,34	0,42	0,50	0,53
CNN	accuracy	98,30%	98,73%	99,28%	99,46%	99,67%	99,83%	99,88%	99,93%
	time (min)	22,08*	23,06*	23,40*	21,40*	16,31	23,18*	25,42*	30,08*

Πίνακας 5.8: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Pavia Center

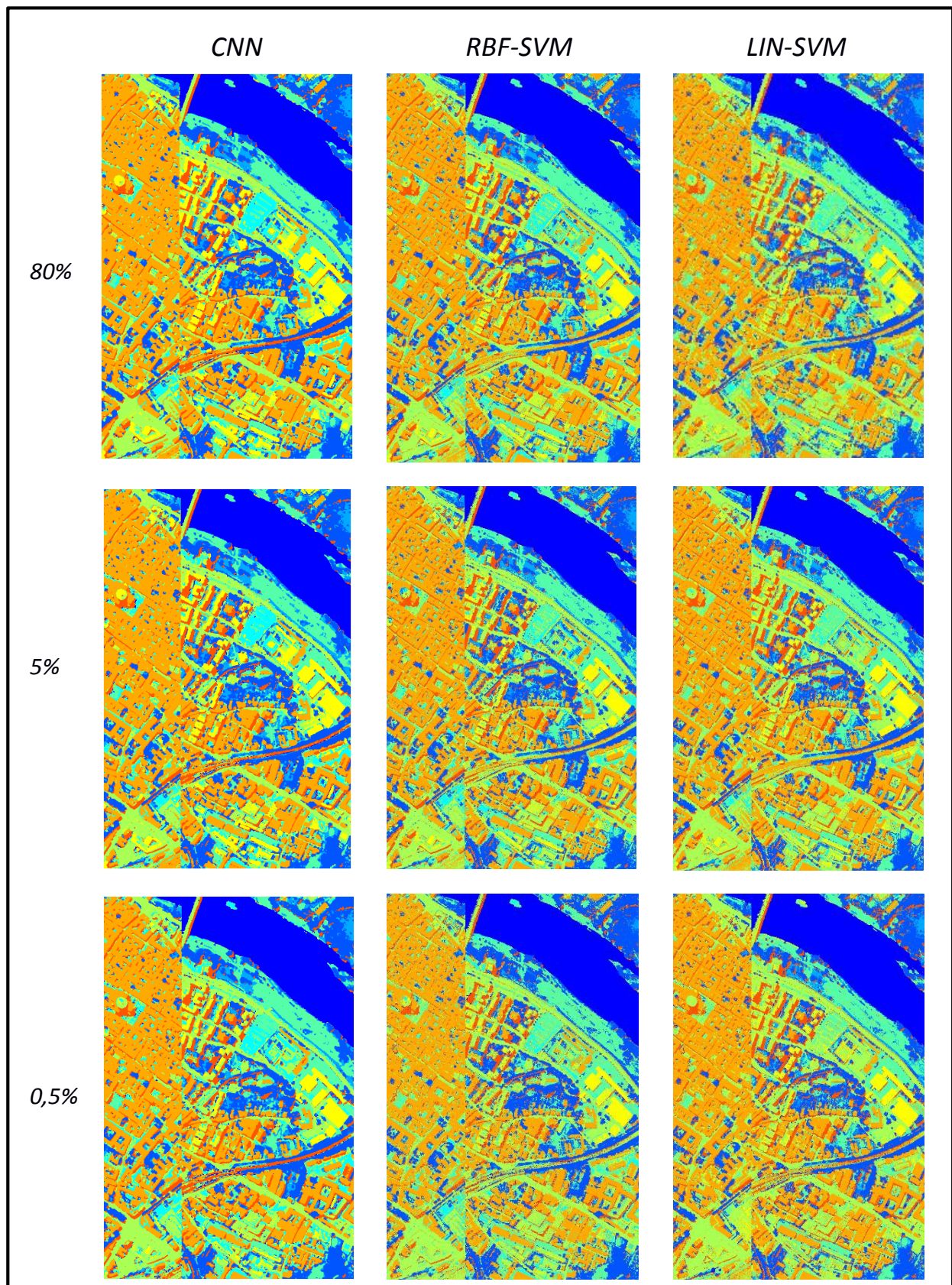


Εικόνα 5.19: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Pavia Center

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM σε όλα τα ποσοστά δεδομένων εκπαίδευσης. Παρ' όλα αυτά και οι 3 ταξινομητές πετυχαίνουν ακρίβειες πάνω από 96% ακόμα και με 0,5% ως ποσοστό εκπαίδευσης. Η συνθήκη αυτή, οφείλεται τόσο στον μεγάλο αριθμό δειγμάτων (samples) αυτού του σετ δεδομένων, όσο και στην αξιοπιστία τους. Το δεύτερο, εξάγεται ως συμπέρασμα καθώς σε καμία άλλη εφαρμογή (βιβλιογραφικών δεδομένων) δεν επιτυγχάνονται τόσο υψηλά ποσοστά ακρίβειας.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, με χρονική απόδοση τα μερικά δευτερόλεπτα. Ακόμα και στην περίπτωση που ο CNN δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (στο 15%), η χρονική απόδοσή του ξεπέρασε τα 16 λεπτά.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

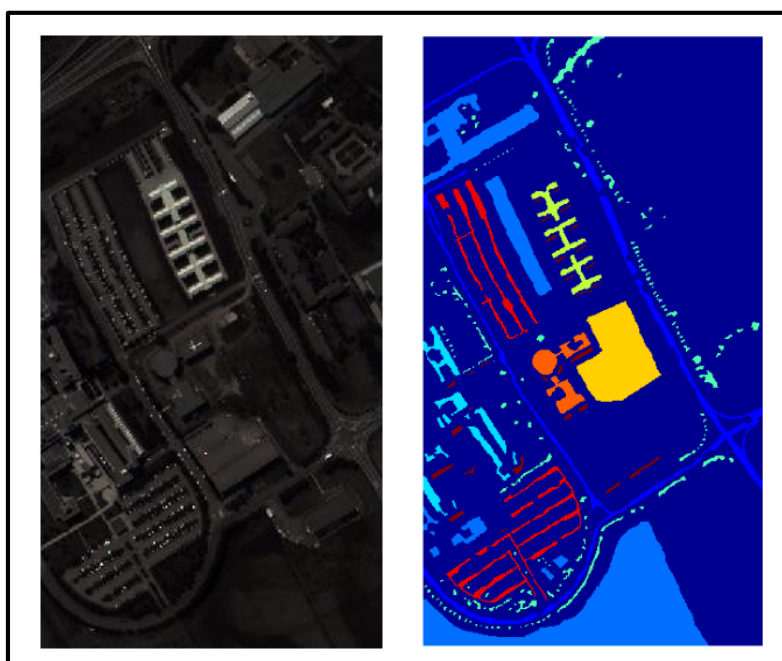


Εικόνα 5.20: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Pavia Center

Pavia University

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας *Pavia University* που αποτυπώνει το πανεπιστήμιο της πόλης Pavia στην βόρεια Ιταλία.

Η υπερφασματική εικόνα αξιοποιήθηκε με όλη την υπερφασματική πληροφορία που ήταν διαθέσιμη κατά την απόκτηση της από το διαδίκτυο: [http://www.ehu.us/ccwintco/index.php?title=Hyperspectral Remote Sensing Scenes](http://www.ehu.us/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes) (103 κανάλια – με δέκτη τον ROSIS). Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=22, G=15, B=9) καθώς και η Ground Truth εικόνα με τις 9 κλάσεις.

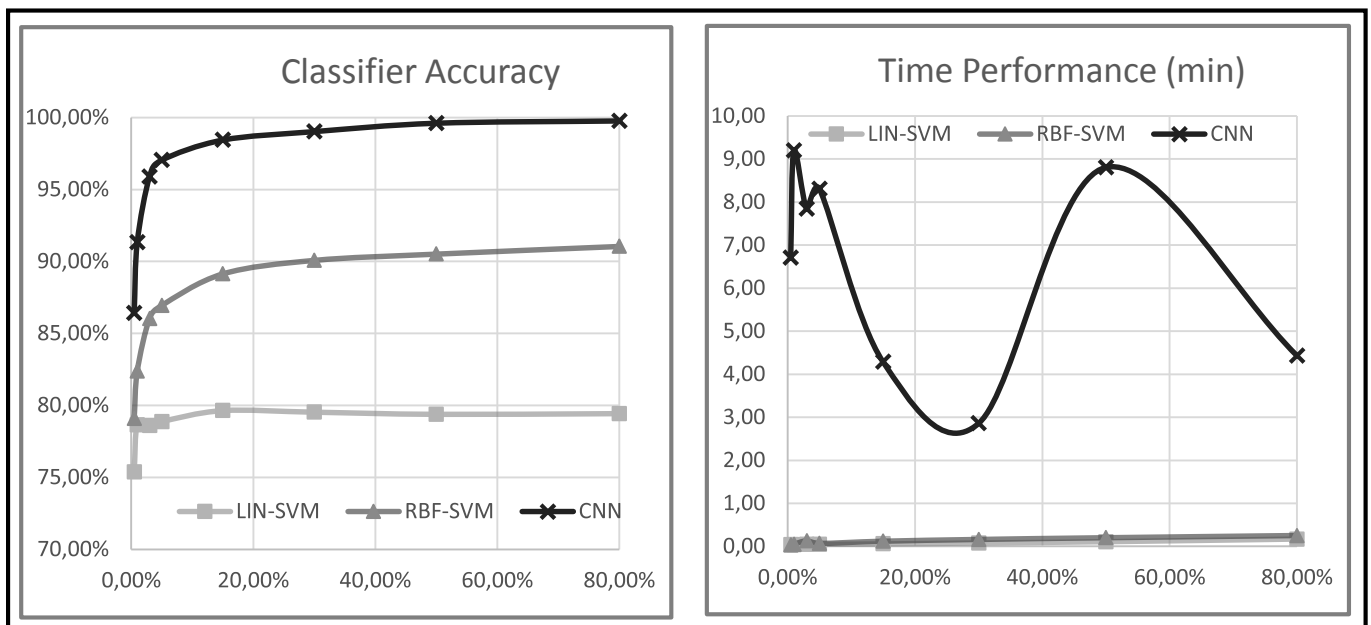


Εικόνα 5.21: Έγχρωμο σύνθετο της εικόνας Pavia University, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. **Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,81% ποσοστό της υπερφασματικής πληροφορίας.**

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	214	428	1283	2139	6416	12833	21388	34221
	accuracy	75,38%	78,64%	78,60%	78,88%	79,64%	79,53%	79,38%	79,43%
	time (min)	0,04	0,04	0,05	0,05	0,06	0,08	0,11	0,17
RBF	accuracy	79,09%	82,38%	86,04%	86,94%	89,14%	90,08%	90,51%	91,05%
	time (min)	0,03	0,05	0,13	0,07	0,12	0,16	0,20	0,25
CNN	accuracy	86,41%	91,32%	95,89%	97,04%	98,44%	99,03%	99,61%	99,77%
	time (min)	6,71*	9,20*	7,85*	8,31*	4,29	2,86	8,81*	4,43

Πίνακας 5.8: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Pavia Center

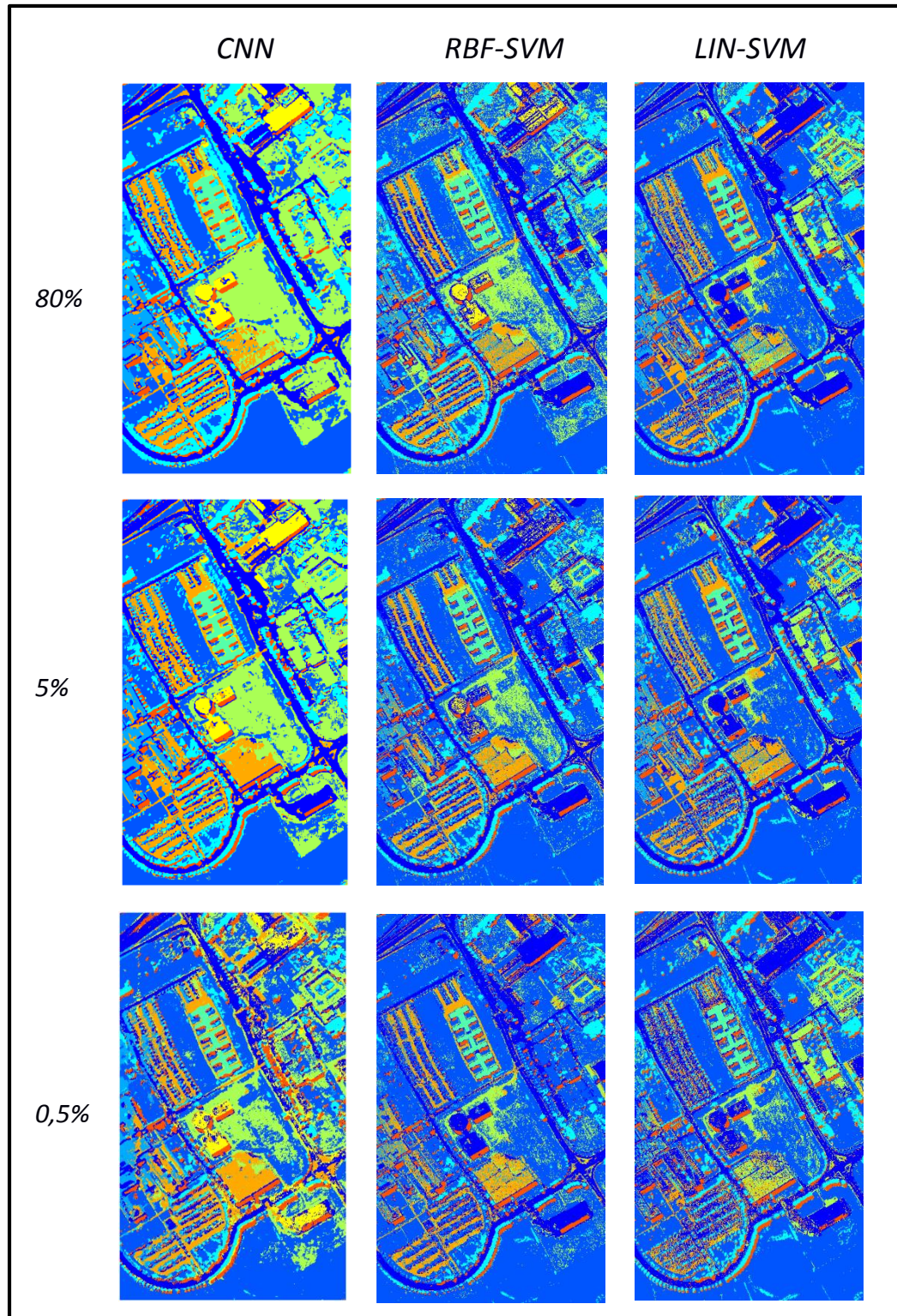


Εικόνα 5.22: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Pavia University

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει κατά πολύ καλύτερες ακρίβειες από τους ταξινομητές SVM σε όλα τα ποσοστά δεδομένων εκπαίδευσης. Συγκεκριμένα, επιτυγχάνει καλύτερες ακρίβειες κατά περίπου 10% από τον δεύτερο σε σειρά ταξινομητή για όλα τα δεδομένα εκπαίδευσης. Επιπλέον, ο ταξινομητής CNN κυμαίνεται από περίπου 86% έως 99% σε ακρίβεια ενώ ο ταξινομητής RBF-SVM από 79% έως 91%. Τέλος, παρατηρείται ότι αντιθέτως με τους άλλους δυο ταξινομητές ο Linear-SVM επιτυγχάνει εφάμιλλες αποδόσεις για όλα ποσοστά εκπαίδευσης.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, με χρονική απόδοση τα μερικά δευτερόλεπτα. Ακόμα και στις περιπτώσεις που ο CNN δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (15%, 30% και 80%), η χρονική απόδοσή του ξεπέρασε τα 2,5 λεπτά.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

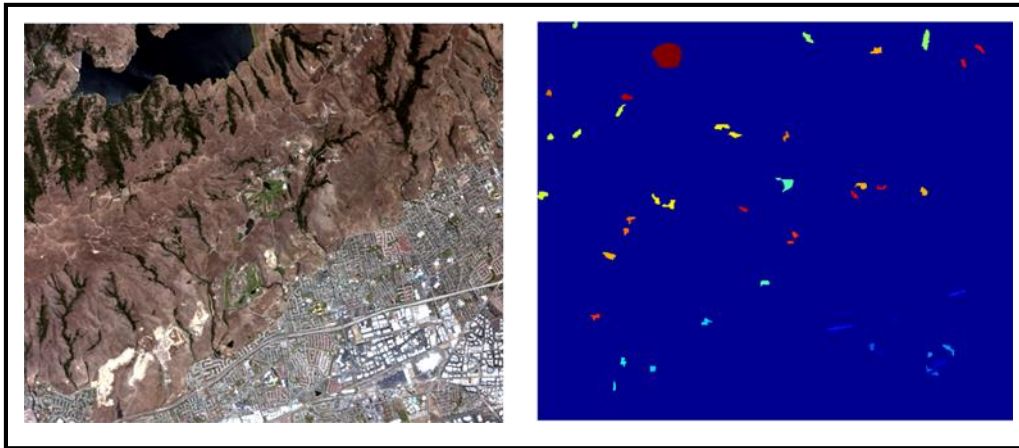


Εικόνα 5.23: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Pavia University

Moffet Field Run 1

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας *Moffet Field Run 1* που αποτυπώνει περιοχή πλησίον του κόλπου του Σαν Φρανσίσκο, στην Καλιφόρνια των ΗΠΑ..

Η υπερφασματική εικόνα κατά την απόκτηση της από το διαδίκτυο: <http://opticks.org/confluence/display/opticks/Sample+Data#SampleData-AVMOFF3> περιείχε 224 κανάλια (δέκτης ο AVIRIS). Παρ' όλα αυτά, εξαιρέθηκαν 89 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 135 κανάλια. Τα 89 αυτά εξαιρούμενα κανάλια είναι τα: [1-3],[33],[97],[104-114],[150-168],[221-223]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=28, G=17, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις.

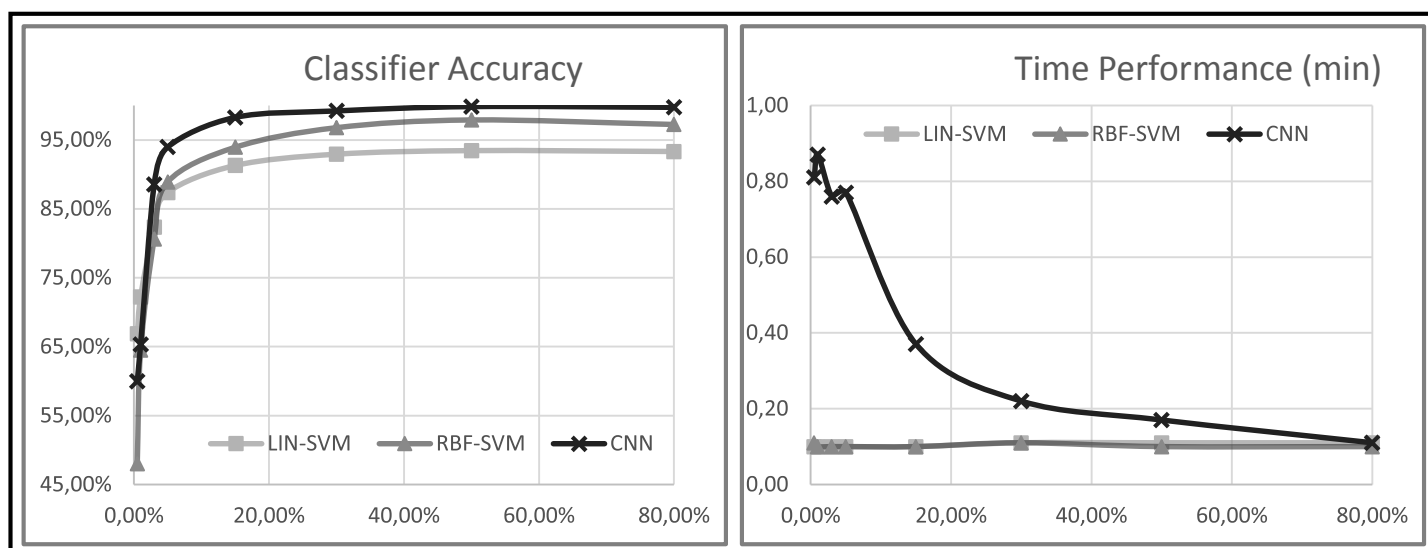


Εικόνα 5.24: Έγχρωμο σύνθετο της εικόνας Moffet Field Run 1, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,85% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	19	38	115	191	573	1147	1911	3058
	accuracy	66,86%	72,24%	82,33%	87,38%	91,29%	92,94%	93,46%	93,33%
	time (min)	0,10	0,10	0,10	0,10	0,10	0,11	0,11	0,11
RBF	accuracy	48,03%	64,53%	80,60%	88,90%	93,94%	96,79%	97,91%	97,25%
	time (min)	0,11	0,10	0,10	0,10	0,10	0,11	0,10	0,10
CNN	accuracy	60,00%	65,32%	88,53%	93,96%	98,27%	99,21%	99,84%	99,74%
	time (min)	0,81*	0,87*	0,76*	0,77*	0,37	0,22	0,17	0,11

Πίνακας 5.9: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Moffet Field Run 1

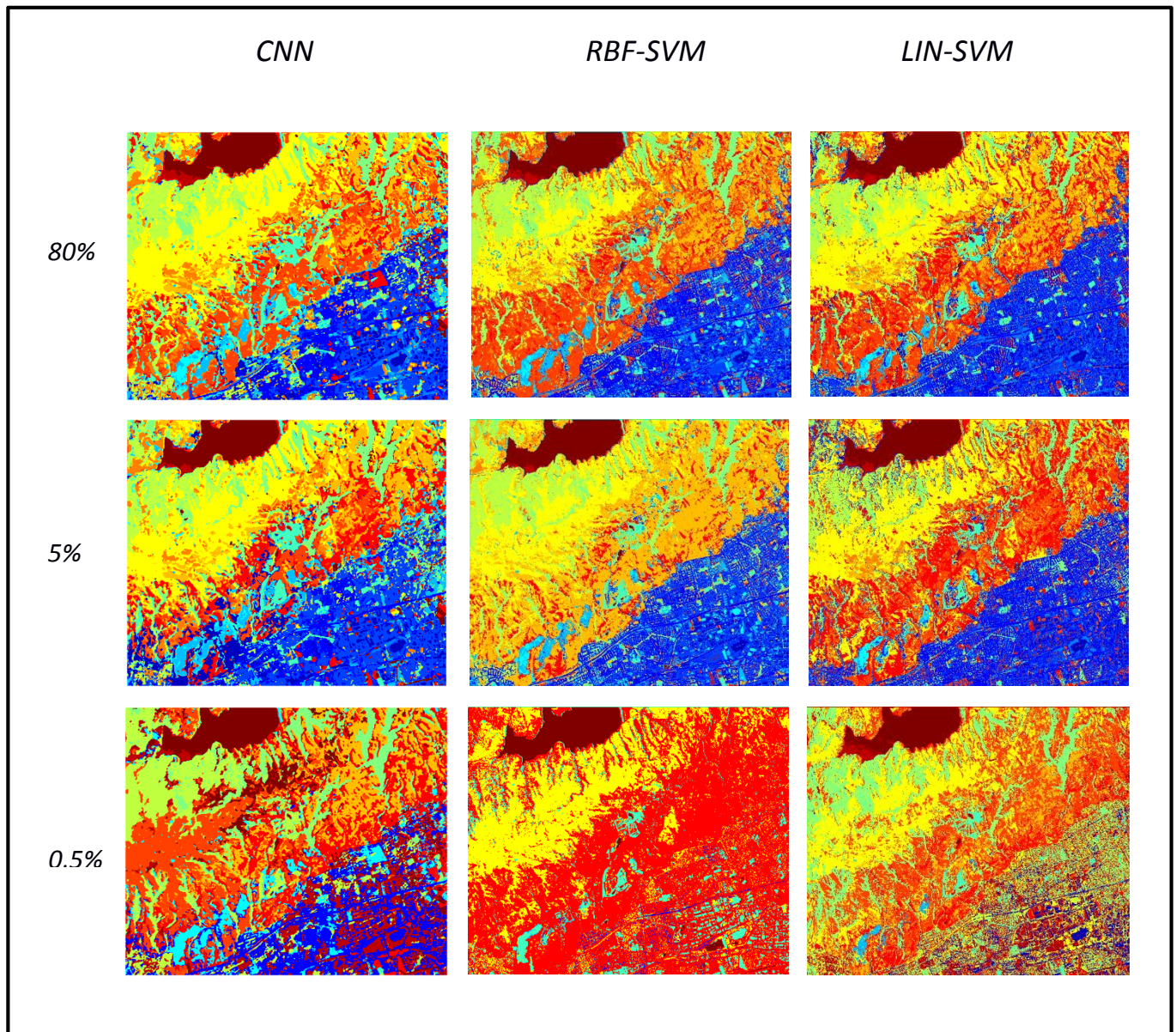


Εικόνα 5.25: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Moffet Field Run 1

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Παρ' όλα αυτά οι ταξινομητές SVM προσφέρουν και αυτοί υψηλά ποσοστά ακρίβειας κοντά στο 90% ήδη από το 5% του ποσοστού εκπαίδευσης. Ακόμα, παρατηρείται πως στα μικρά ποσοστά εκπαίδευσης (0,5% και 1%), ο LIN-SVM αποδίδει καλύτερα ακόμα και από το νευρωνικό δίκτυο κάτι που οφείλεται στον μικρό αριθμό δειγμάτων στις κλάσεις Ground Truth και ίσως στην μερικώς λανθασμένη ψηφιοποίηση των δεδομένων αυτών.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM φαίνεται να αποδίδουν ταχύτερα του CNN. Σε κάθε περίπτωση πάντως η χρονική απόδοση του ταξινομητή CNN είναι μικρότερη του λεπτού και στην περίπτωση μάλιστα του 80% ως ποσοστό εκπαίδευσης, εφάμιλλη με αυτή των SVM.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

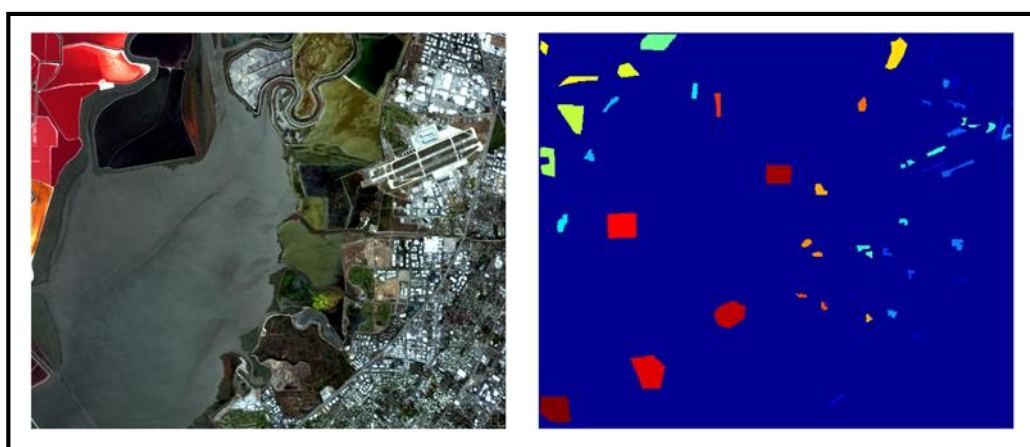


Εικόνα 5.26: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Moffet Field Run 1

Moffet Field Run 3

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας *Moffet Field Run 3* που αποτυπώνει περιοχή πλησίον του κόλπου του Σαν Φρανσίσκο , στην Καλιφόρνια των ΗΠΑ..

Η υπερφασματική εικόνα κατά την απόκτηση της από το διαδίκτυο: <http://opticks.org/confluence/display/opticks/Sample+Data#SampleData-AVMOFF3> περιείχε 224 κανάλια (δέκτης ο AVIRIS). Παρ' όλα αυτά, εξαιρέθηκαν 89 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 135 κανάλια. Τα 89 αυτά εξαιρούμενα κανάλια είναι τα: [1-3],[33],[97],[104-114],[150-168],[221-223]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=28, G=17, B=9) καθώς και η Ground Truth εικόνα με τις 16 κλάσεις.

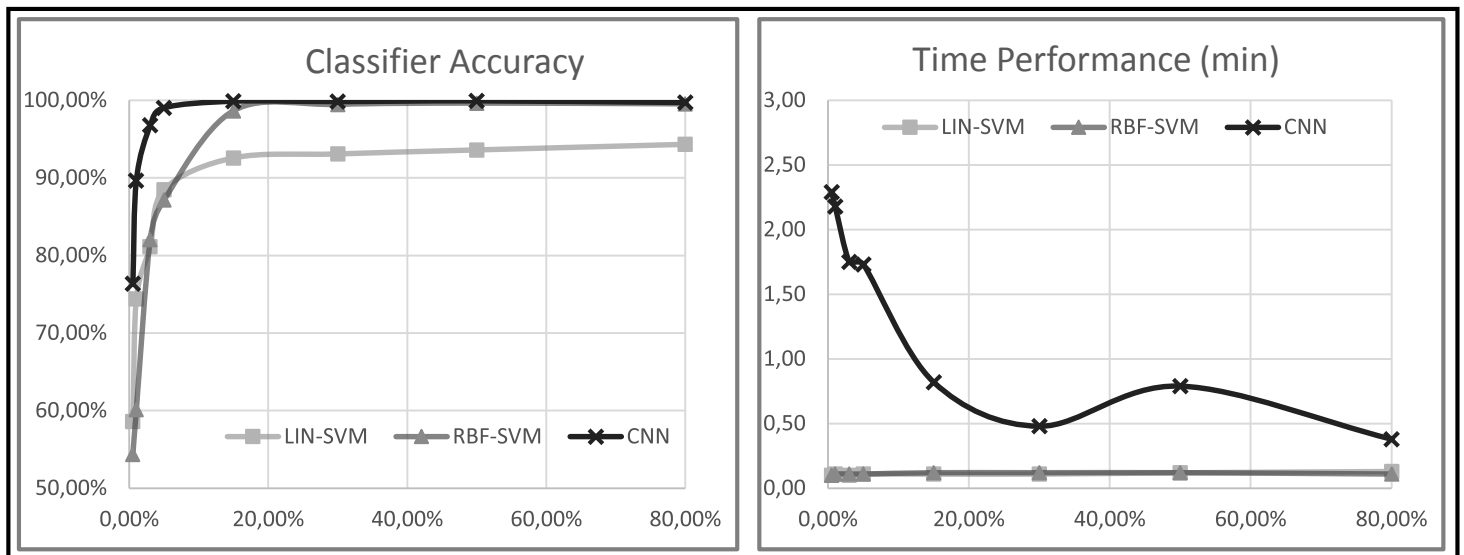


Εικόνα 5.27: Έγχρωμο σύνθετο της εικόνας Moffet Field Run 3, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,81% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	63	125	375	625	1875	3750	6251	10001
	accuracy	58,61%	74,39%	81,15%	88,48%	92,57%	93,09%	93,60%	94,32%
RBF	accuracy	54,37%	60,14%	82,03%	87,14%	98,64%	99,45%	99,62%	99,52%
	time (min)	0,10	0,11	0,10	0,11	0,11	0,11	0,12	0,13
CNN	accuracy	76,39%	89,68%	96,78%	99,04%	99,89%	99,84%	99,94%	99,72%
	time (min)	2,29*	2,18*	1,75*	1,73*	0,82	0,48	0,79	0,38

Πίνακας 5.10: Τα αποτελέσματα των ταξινομήσεων των δεδομένων Moffet Field Run 3

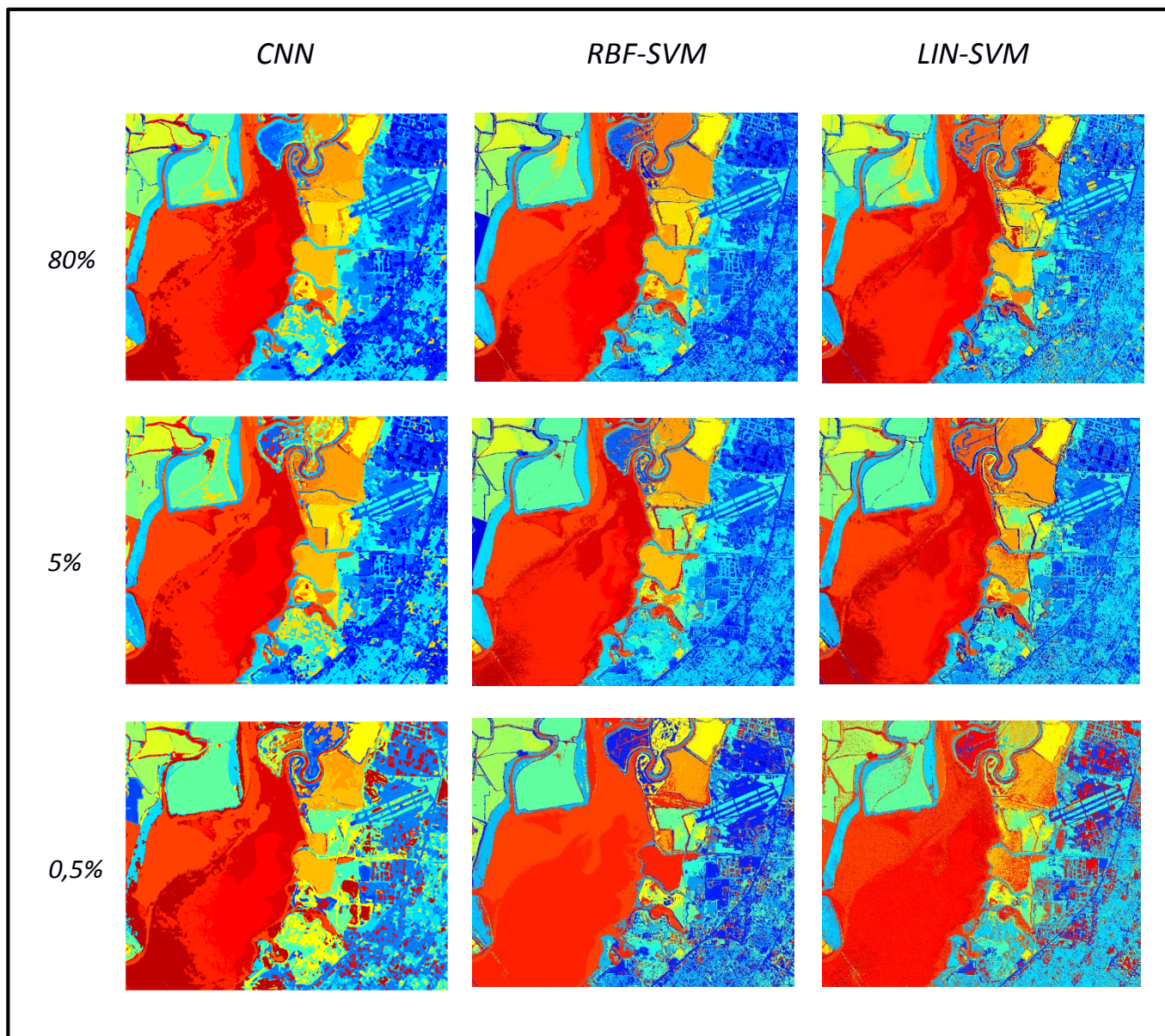


Εικόνα 5.28: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα Moffet Field Run 3

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Παρ' όλα αυτά ο ταξινομητής RBF-SVM προσφέρει εφάμιλλες ακρίβειες στα υψηλά ποσοστά εκπαίδευσης. Συγκεκριμένα, ο ταξινομητής CNN κυμαίνεται από περίπου 76% έως 99% σε ακρίβεια ενώ ο ταξινομητής RBF-SVM από 54% έως 99%.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM αποδίδουν ταχύτερα του CNN. Σε κάθε περίπτωση πάντως η χρονική απόδοση του ταξινομητή CNN είναι μικρότερη των 2,5 λεπτών και στις περιπτώσεις μάλιστα που ο αλγόριθμος δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (15%, 30%, 50% και 80%), η χρονική απόδοσή του προσεγγίζει την αντίστοιχη των SVM.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:



Εικόνα 5.29: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα Moffet Field Run 3

5.4 Αποτελέσματα Δεδομένων SFB (San Francisco Bay)

Εδώ, παρουσιάζονται και αξιολογούνται τα αποτελέσματα που προέκυψαν κατά την εκπόνηση της εργασίας και αφορούν τα δεδομένα SFB.

Όπως αναλύονται και στο 4^ο κεφάλαιο, τα δεδομένα SFB δημιουργήθηκαν αποκλειστικά για τους σκοπούς αυτής της εργασίας με στόχο την εφαρμογή ποικίλων ταξινομήσεων με διάφορες προσεγγίσεις. Τα δεδομένα αυτά, αποτελούνται συνολικά από 7 υπερφασματικές εικόνες και τις αντίστοιχες 7 εικόνες Ground Truth και αποτυπώνουν εκτάσεις πλησίον του κόλπου του San Francisco στην California των ΗΠΑ.

Το υποκεφάλαιο 5.4 χωρίζεται σε 2 ενότητες:

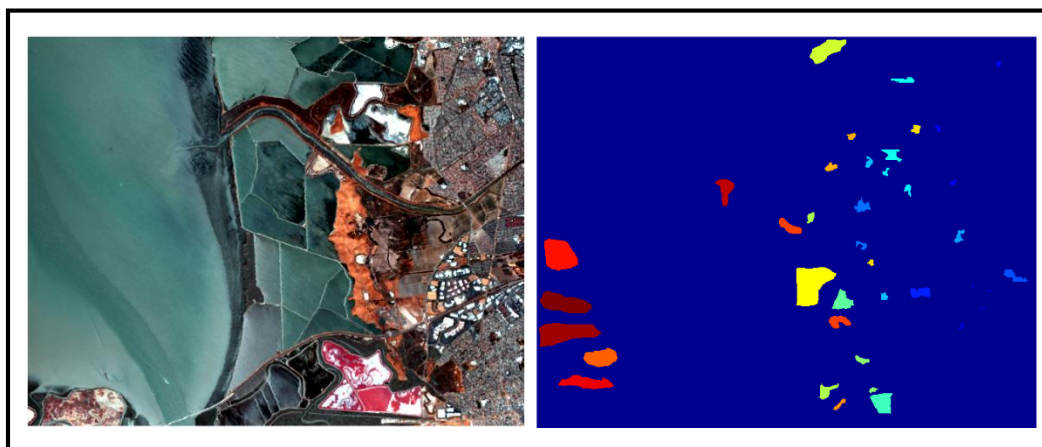
- Στην πρώτη (5.4.1) υλοποιούνται ταξινομήσεις σε 3 εικόνες των δεδομένων SFB, μονοεικονικά, με την ίδια ακριβώς προσέγγιση που ακολουθήθηκε και στο υποκεφάλαιο 5.3 . Οι 3 αυτές εικόνες (2008_1, 2012_1 και 2014_2) επιλέχθηκαν επειδή η καθεμιά αποτυπώνει τελείως διαφορετική περιοχή και κατά συνέπεια διαφορετικές κλάσεις. Με αυτόν τον τρόπο σε αυτές τις δοκιμές θα εξεταστεί η απόδοση των 3 ταξινομητών πάνω σε 3 διαφορετικές περιπτώσεις κατηγοριών εδάφους.
- Στην δεύτερη (5.4.2) υλοποιούνται ταξινομήσεις, πολυεικονικές, και στις 7 υπερφασματικές εικόνες SFB έπειτα από την ενοποίησή τους. Επιπλέον, σε αυτή την ενότητα ακολουθούν και πιο σύνθετες εφαρμογές ταξινομήσεων με μειωμένο αριθμό κλάσεων Ground Truth και μειωμένο αριθμό Κυρίων Συνιστωσών (PCA).

5.4.1 Μονοεικονικές ταξινομήσεις

2008_1

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας 2008_1 που αποτυπώνει περιοχή του κόλπου του Σαν Φρανσίσκο, στην Καλιφόρνια των ΗΠΑ. Η εικόνα αυτή ως επί των πλείστων αποτυπώνει υδάτινες επιφάνειες (εποχιακές ή μόνιμες), βάλτους, αλυκές και κάποιες κατηγορίες γυμνού εδάφους.

Η υπερφασματική εικόνα έπειτα από την αποκοπή της (βλ. 4.2.4) της, περιείχε 224 κανάλια (δέκτης ο AVIRIS). Παρ' όλα αυτά, εξαιρέθηκαν 89 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 135 κανάλια. Τα 89 αυτά εξαιρούμενα κανάλια είναι τα: [1-3],[33],[97],[104-114],[150-168],[221-223]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=32, G=13, B=7) καθώς και η Ground Truth εικόνα με 26 κλάσεις.

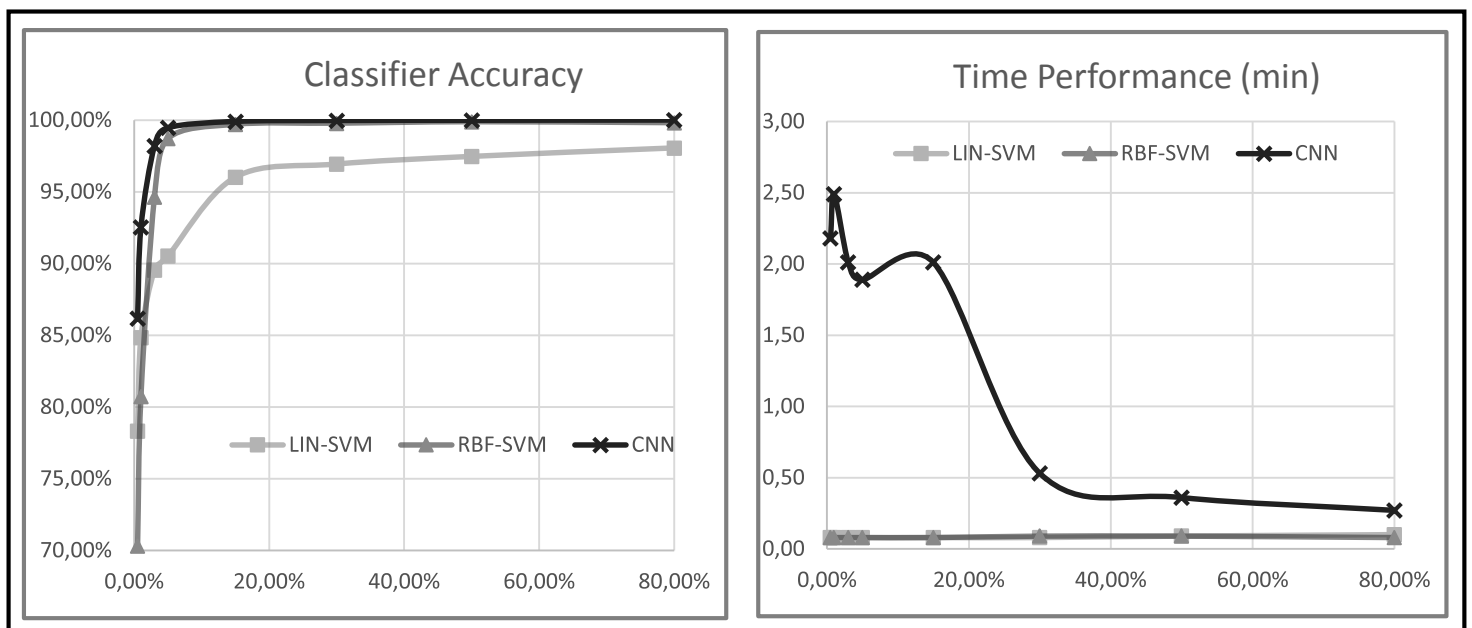


Εικόνα 5.30: Έγχρωμο σύνθετο της εικόνας 2008-1, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,94% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	68	137	410	684	2051	4102	6836	10938
	accuracy	78,32%	84,84%	89,55%	90,53%	96,02%	96,94%	97,47%	98,06%
RBF	accuracy	70,31%	80,74%	94,64%	98,71%	99,71%	99,78%	99,88%	99,82%
	time (min)	0,08	0,08	0,08	0,08	0,08	0,09	0,09	0,08
CNN	accuracy	86,16%	92,52%	98,20%	99,46%	99,91%	99,97%	99,99%	100,00%
	time (min)	2,18*	2,49*	2,01*	1,89*	2,01*	0,53	0,36	0,27

Πίνακας 5.11: Τα αποτελέσματα των ταξινομήσεων των δεδομένων 2008-1

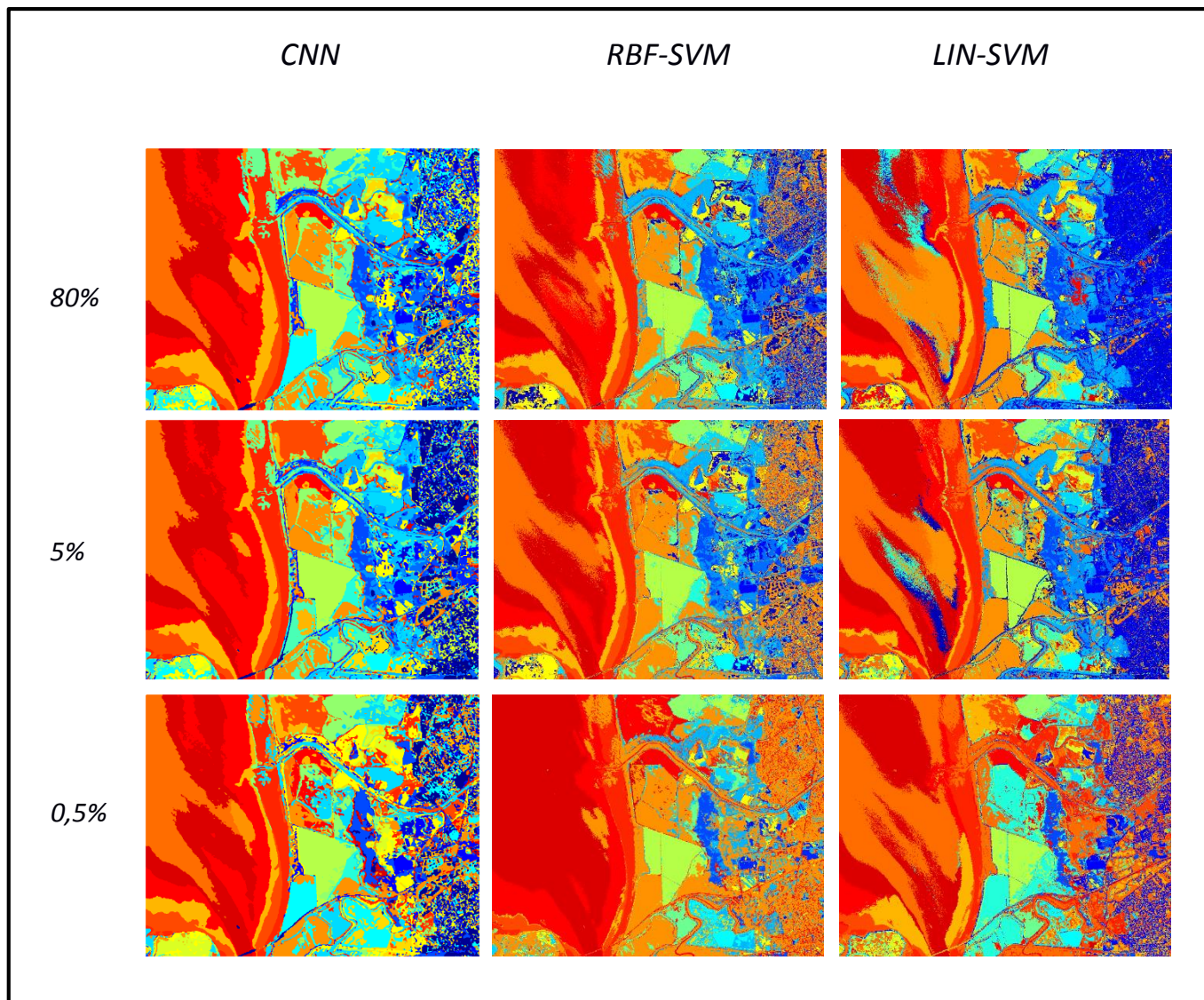


Εικόνα 5.31: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα 2008-1

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Βεβαίως, οι 2 ταξινομητές SVM και κυρίως ο RBF, επιτυγχάνουν πολύ υψηλά ποσοστά ακρίβειας ήδη από το 1% του ποσοστού εκπαίδευσης. Συγκεκριμένα, ο CNN πετυχαίνει ακρίβειες από 86% έως 100% ενώ ο RBF από 71% έως 99%.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM αποδίδουν ταχύτερα του CNN, σε μερικά μόλις δευτερόλεπτα. Σε κάθε περίπτωση πάντως η χρονική απόδοση του ταξινομητή CNN είναι μικρότερη των 2,5 λεπτών και στις περιπτώσεις μάλιστα στις οποίες ο αλγόριθμος δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (30%, 50% και 80%), η χρονική απόδοσή του προσεγγίζει την αντίστοιχη των SVM.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

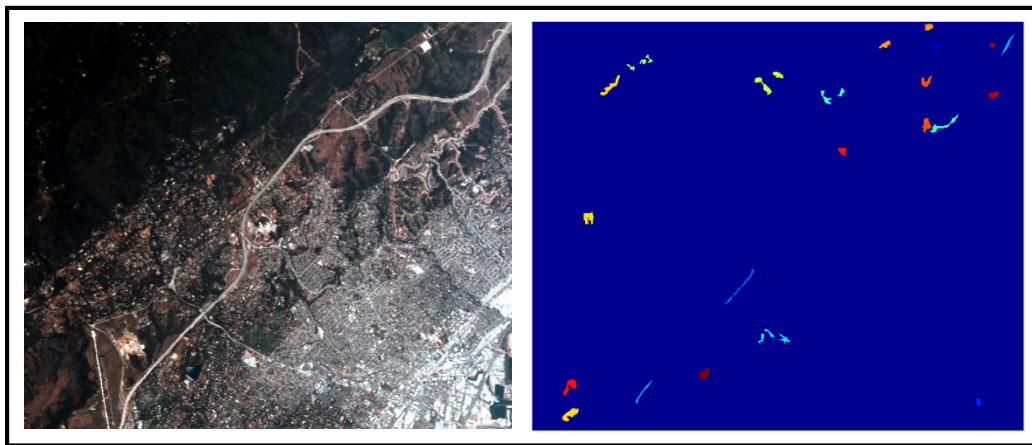


Εικόνα 5.32: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα 2008-1

2012_1

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας 2012_1 που αποτυπώνει περιοχή πλησίον του κόλπου του Σαν Φρανσίσκο, στην Καλιφόρνια των ΗΠΑ. Η εικόνα αυτή ως επί των πλείστων αποτυπώνει αστική δόμηση και οδικό δίκτυο, κατηγορίες γυμνού εδάφους και δασική έκταση.

Η υπερφασματική εικόνα έπειτα από την αποκοπή της (βλ. 4.2.4) της, περιείχε 224 κανάλια (δέκτης ο AVIRIS). Παρ' όλα αυτά, εξαιρέθηκαν 89 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 135 κανάλια. Τα 89 αυτά εξαιρούμενα κανάλια είναι τα: [1-3],[33],[97],[104-114],[150-168],[221-223]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=32, G=13, B=7) καθώς και η Ground Truth εικόνα με 14 κλάσεις.

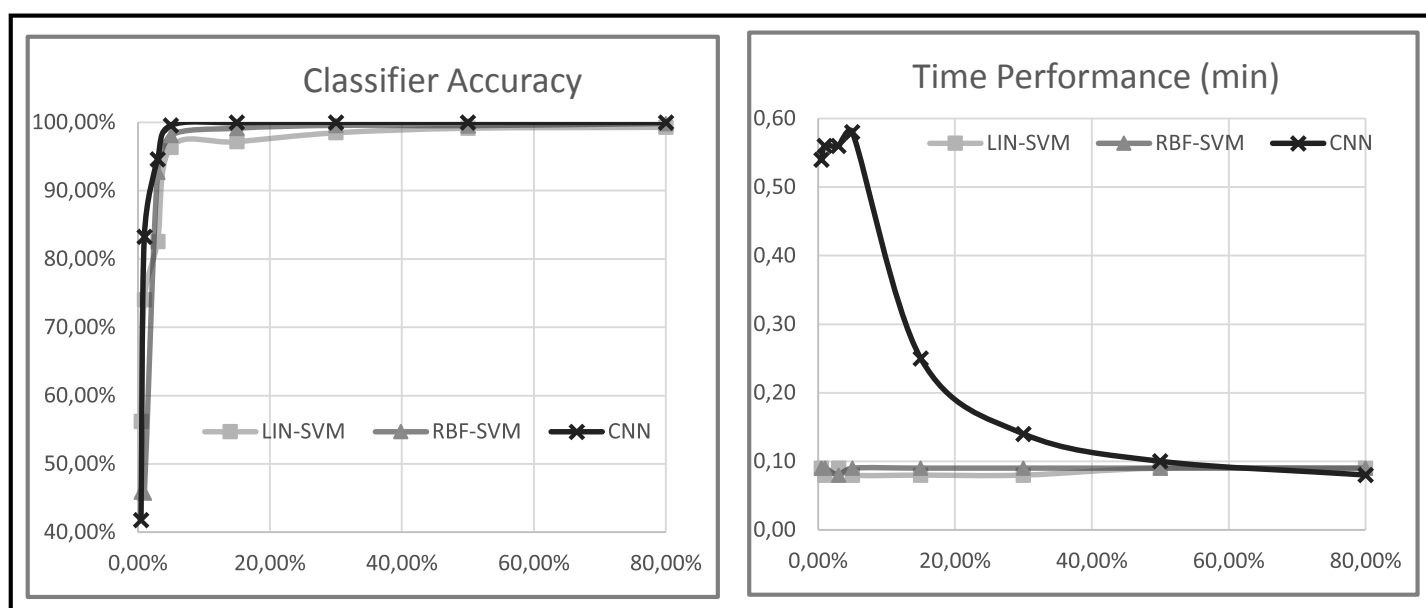


Εικόνα 5.33: Έγχρωμο σύνθετο της εικόνας 2012-1, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,94% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	14	28	85	141	424	849	1415	2263
	accuracy	56,25%	74,04%	82,58%	96,35%	97,17%	98,48%	99,15%	99,29%
RBF	accuracy	45,88%	45,71%	92,67%	98,10%	99,13%	99,55%	99,51%	99,82%
	time (min)	0,09	0,09	0,08	0,09	0,09	0,09	0,09	0,09
CNN	accuracy	41,78%	83,25%	94,60%	99,55%	99,96%	100,00%	100,00%	100,00%
	time (min)	0,54*	0,56*	0,56*	0,58*	0,25	0,14	0,10	0,08

Πίνακας 5.12: Τα αποτελέσματα των ταξινομήσεων των δεδομένων 2012-1

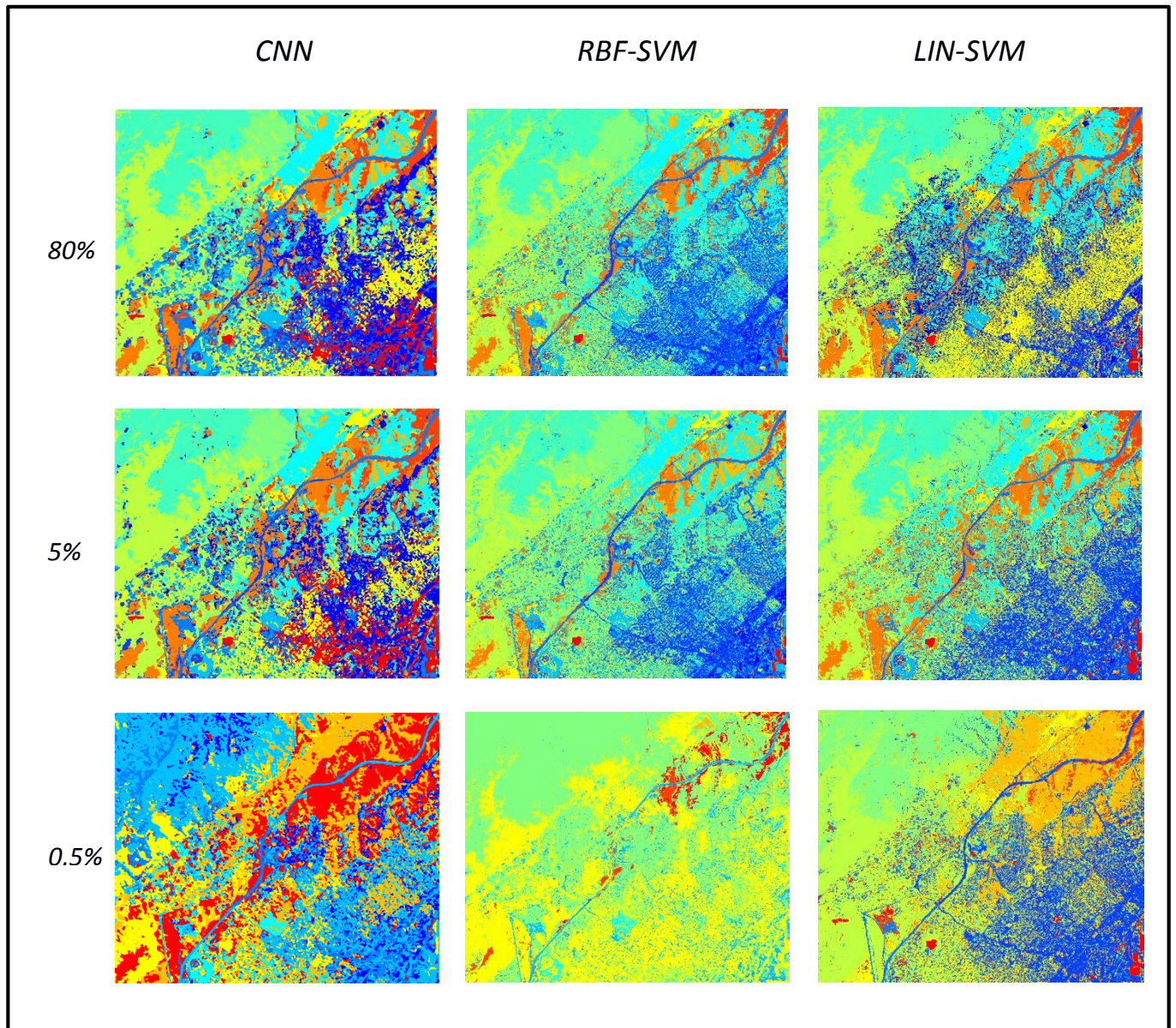


Εικόνα 5.34: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα 2012-1

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Βεβαίως, οι 2 ταξινομητές SVM και κυρίως ο RBF, επιτυγχάνουν πολύ υψηλά ποσοστά ακρίβειας ήδη από το 1% του ποσοστού εκπαίδευσης. Ειδικότερα, με το 5% του ποσοστού εκπαίδευσης και οι 3 ταξινομητές επιτυγχάνουν ακρίβειες υψηλότερες του 96%. Όπως φαίνεται και στο αριστερό γράφημα άλλωστε, οι 3 ταξινομητές έχουν παραπλήσια καμπύλη ακρίβειας.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM αποδίδουν ταχύτερα του CNN, σε μερικά μόλις δευτερόλεπτα. Παρ' όλα αυτά, σε αυτή την εφαρμογή ο ταξινομητής αποδίδει ταχύτατα και στις περιπτώσεις μάλιστα στις οποίες ο αλγόριθμος δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (15%, 30%, 50% και 80%), η χρονική απόδοσή του προσεγγίζει την αντίστοιχη των SVM.

Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:

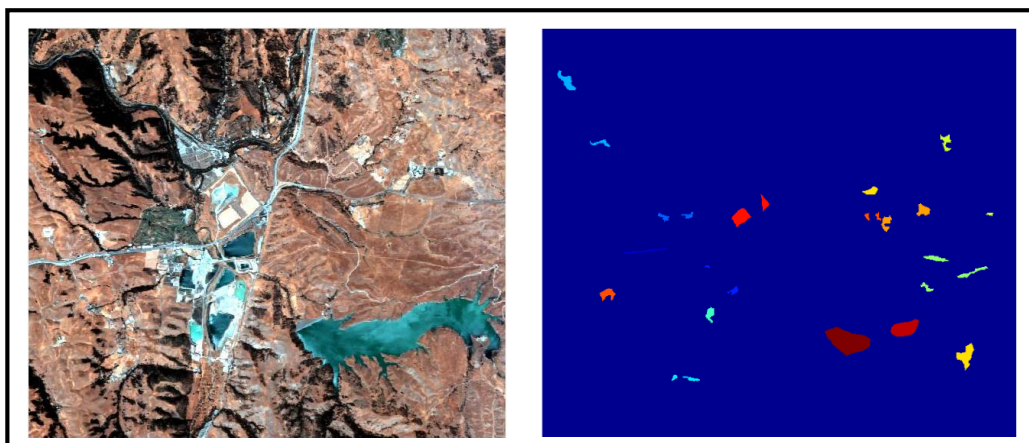


Εικόνα 5.35: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα 2012-1

2014_2

Σε αυτή την εφαρμογή εφαρμόστηκαν οι 3 ταξινομητές πάνω στα δεδομένα της εικόνας 2014_2 που αποτυπώνει περιοχή πλησίον του κόλπου του Σαν Φρανσίσκο, στην Καλιφόρνια των ΗΠΑ. Η εικόνα αυτή ως επί των πλείστων αποτυπώνει γυμνό έδαφος και χαμηλή βλάστηση, τεχνητές επιφάνειες και επιφάνειες νερού σε λίμνη.

Η υπερφασματική εικόνα έπειτα από την αποκοπή της (βλ. 4.2.4) της, περιείχε 224 κανάλια (δέκτης ο AVIRIS). Παρ'όλα αυτά, εξαιρέθηκαν 89 κανάλια με ή χωρίς θόρυβο και τελικά στα υπό δοκιμή δεδομένα απομένουν 135 κανάλια. Τα 89 αυτά εξαιρούμενα κανάλια είναι τα: [1-3],[33],[97],[104-114],[150-168],[221-223]. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της εικόνας (RGB- R=32, G=13, B=7) καθώς και η Ground Truth εικόνα με 14 κλάσεις.

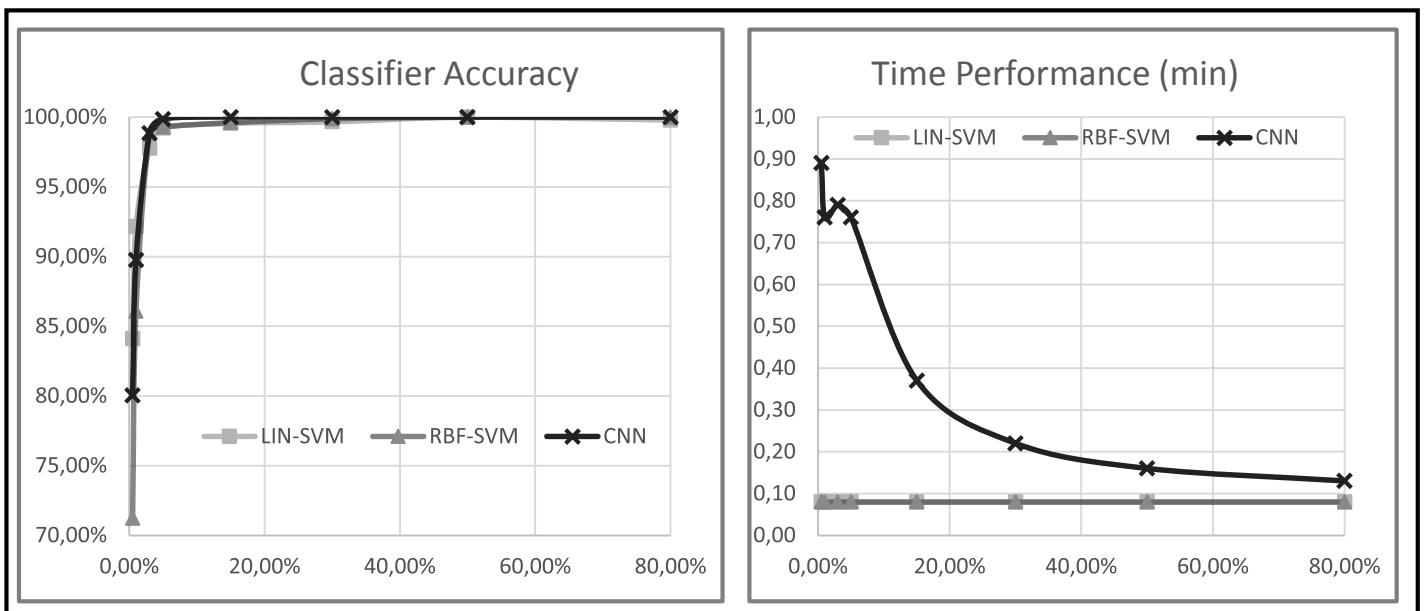


Εικόνα 5.36: Έγχρωμο σύνθετο της εικόνας 2014-2, με τα αντίστοιχα δεδομένα Ground Truth

Οι παραχθείσες ταξινομήσεις αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Σημειώνεται, πως οι ταξινομήσεις σε αυτά τα δεδομένα υλοποιήθηκαν με 10 Κύριες Συνιστώσες και 99,96% ποσοστό της υπερφασματικής πληροφορίας.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	Samples	25	51	153	255	764	1527	2545	4072
	accuracy	84,11%	92,14%	97,77%	99,21%	99,56%	99,67%	100,00%	99,80%
	time (min)	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
RBF	accuracy	71,23%	86,09%	98,85%	99,30%	99,61%	99,89%	100,00%	100,00%
	time (min)	0,08	0,08	0,08	0,08	0,08	0,08	0,08	0,08
CNN	accuracy	80,04%	89,74%	98,84%	99,86%	100,00%	100,00%	100,00%	100,00%
	time (min)	0,89*	0,76*	0,79*	0,76*	0,37	0,22	0,16	0,13

Πίνακας 5.13: Τα αποτελέσματα των ταξινομήσεων των δεδομένων 2014-2



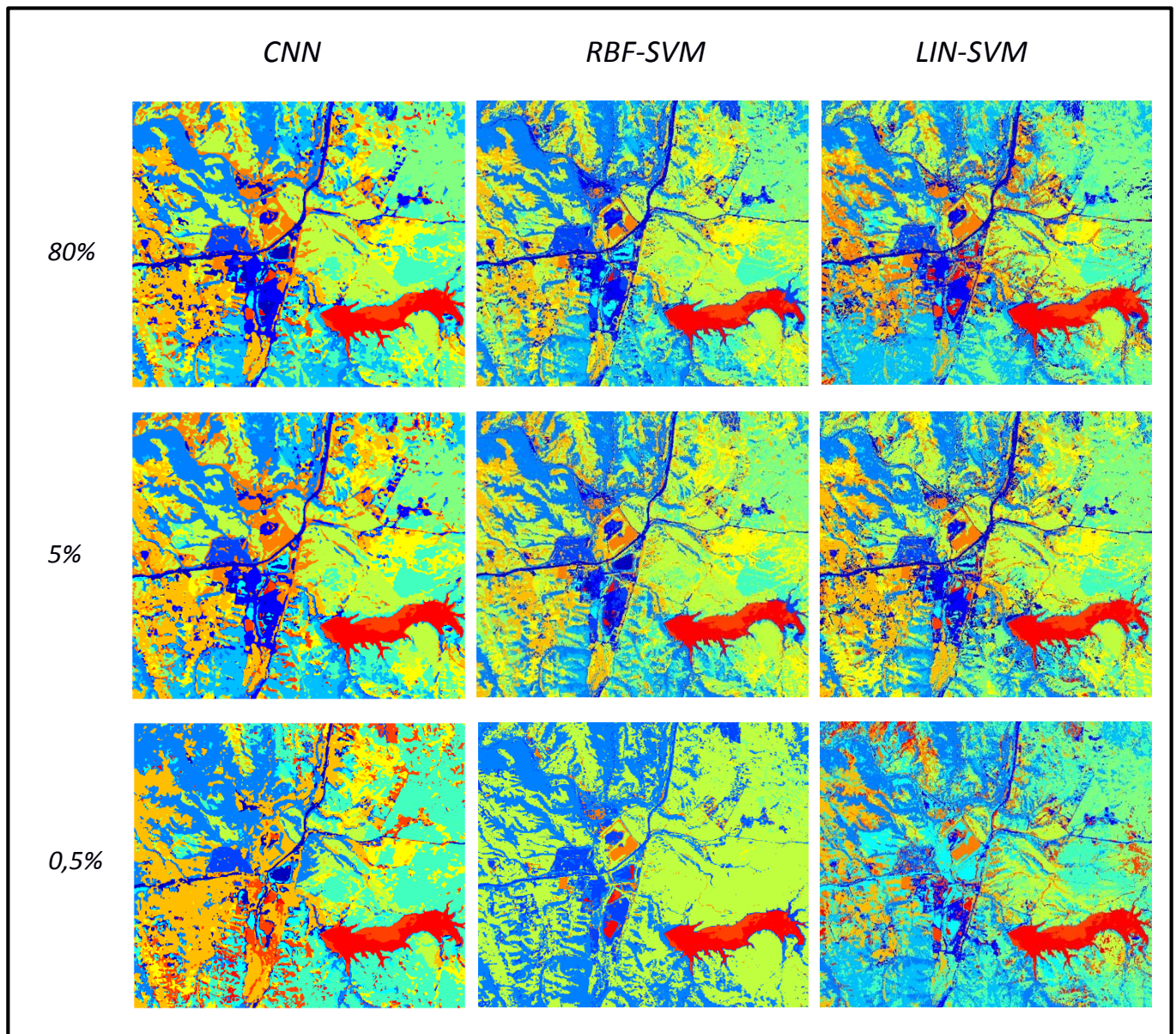
Εικόνα 5.37: Τα γραφήματα απόδοσης ακρίβειας και χρόνου, στα δεδομένα 2014-2

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Βεβαίως, και οι 2 ταξινομητές SVM επιτυγχάνουν υψηλά επίπεδα ακρίβειας σε όλα τα ποσοστά εκπαίδευσης. Συγκεκριμένα, και οι 3 ταξινομητές επιτυγχάνουν ακρίβειες κοντά ή και 100% με μόλις 3% του ποσοστού εκπαίδευσης .

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM αποδίδουν ταχύτερα του CNN, σε μερικά μόλις δευτερόλεπτα. Σε κάθε περίπτωση πάντως η χρονική απόδοση του ταξινομητή CNN είναι μικρότερη του ενός λεπτού και στις περιπτώσεις μάλιστα

στις οποίες ο αλγόριθμος δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (15%, 30%, 50% και 80%), η χρονική απόδοσή του προσεγγίζει την αντίστοιχη των SVM.

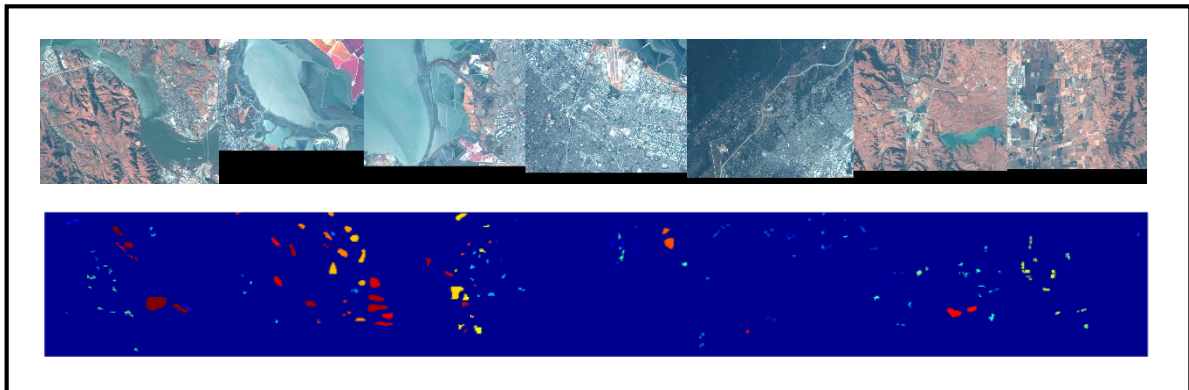
Παρακάτω, παρουσιάζονται μερικά ενδεικτικά αποτελέσματα εικόνων για αυτές τις ταξινομήσεις:



Εικόνα 5.38: Ενδεικτικά αποτελέσματα ταξινομήσεων και για τις 3 προσεγγίσεις, στα δεδομένα 2014-2

5.4.2 Σύνθετες πολυεικονικές ταξινομήσεις

Σε αυτή την ενότητα εφαρμόστηκαν οι 3 ταξινομητές πάνω στις 7 εικόνες των δεδομένων SFB ταυτόχρονα. Κάτι τέτοιο κατέστη δυνατό έπειτα από την ενοποίησή των εικόνων αυτών σε μία ενιαία. Η νέα ενιαία εικόνα – μωσαϊκό, αποτελείται από τα κοινά 135 κανάλια των υπό συνένωση εικόνων, ενώ τα ενοποιημένα δεδομένα Ground Truth περιλαμβάνουν σχεδόν όλες τις κατηγορίες εδάφους που αποτυπώνονται στις αντίστοιχες περιοχές, με συνολικά 98 κλάσεις. Παρακάτω παρουσιάζεται ένα έγχρωμο σύνθετο της ενιαίας εικόνας (RGB- R=32, G=13, B=7) καθώς και η Ground Truth εικόνα με τις 98 κλάσεις.



Εικόνα 5.39: Έγχρωμο σύνθετο της ενιαίας εικόνας SFB-all, με τα αντίστοιχα δεδομένα Ground Truth

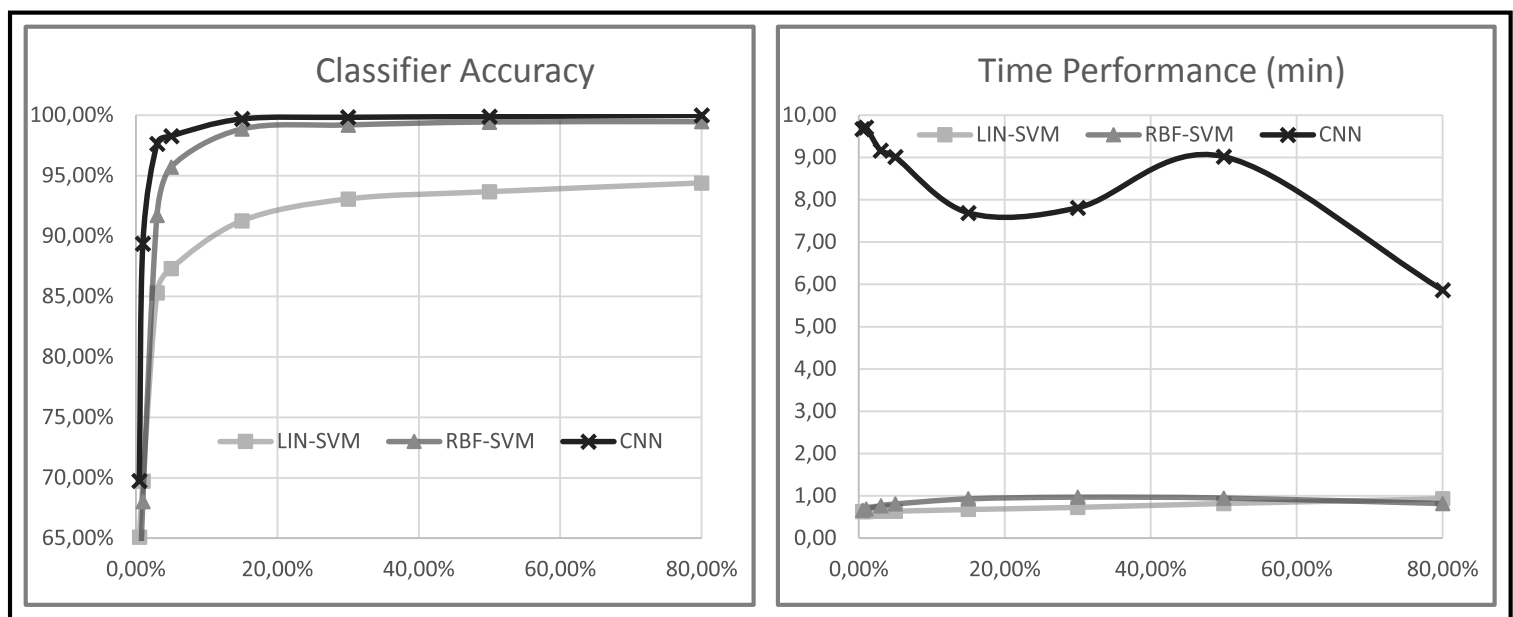
Οι παραχθείσες ταξινομήσεις αυτής της ενότητας αξιολογήθηκαν με βάση την ακρίβεια και τη χρονική απόδοσή τους σε 6 διαφορετικά ποσοστά επί των δεδομένων εκπαίδευσης, τα αποτελέσματα των οποίων παρουσιάζονται παρακάτω. Ακόμα, υλοποιήθηκαν ταξινομήσεις με μειωμένο αριθμό κυρίων συνιστωσών, καθώς επίσης και ταξινομήσεις με μειωμένο αριθμό κλάσεων έπειτα από ενοποίησή τους (54, 27 και 9).

Σημειώνεται πως για τις ακόλουθες εφαρμογές δεν θα παρουσιαστούν αποτελέσματα εικόνων ταξινομήσεων, καθώς το μέγεθος της ενιαίας εικόνας είναι τέτοιο που δεν επιτρέπει την ποιοτική αξιολόγηση των ταξινομημένων εικόνων μέσω της θέασης τους σε έντυπη μορφή.

Παρακάτω παρουσιάζονται τα αποτελέσματα που αφορούν τις ταξινομήσεις με 10 Κύριες Συνιστώσες και 99,94% ποσοστό της υπερφασματικής πληροφορίας, στα δεδομένα Ground Truth με τις 98 κλάσεις.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	accuracy	65,10%	69,72%	85,30%	87,31%	91,25%	93,06%	93,65%	94,38%
	time (min)	0,63	0,63	0,64	0,64	0,68	0,73	0,82	0,93
RBF	accuracy	56,90%	68,02%	91,72%	95,71%	98,85%	99,18%	99,42%	99,45%
	time (min)	0,66	0,69	0,77	0,81	0,93	0,97	0,95	0,82
CNN	accuracy	69,74%	89,37%	97,62%	98,27%	99,68%	99,82%	99,88%	99,98%
	time (min)	9,67*	9,71*	9,16*	9,01*	7,69	7,81	9,02*	5,86

Πίνακας 5.14: Τα αποτελέσματα των ταξινομήσεων των δεδομένων της ενιαίας εικόνας, για τις 98 κλάσεις



Εικόνα 5.40: Τα γραφήματα απόδοσης ακρίβειας και χρόνου στα δεδομένα της ενιαίας εικόνας, για τις 98 κλάσεις

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Βεβαίως, και ο ταξινομητής RBF επιτυγχάνει υψηλά επίπεδα ακρίβειας, παραπλήσια με αυτά του CNN για ποσοστό εκπαίδευσης μεγαλύτερο του 5%. Σημαντική διαφοροποίηση παρατηρείται στα αποτελέσματα του ποσοστού εκπαίδευσης 1% όπου ο CNN αποδίδει με ακρίβεια κοντά στο 90% ενώ οι SVM λιγότερο από 70%.

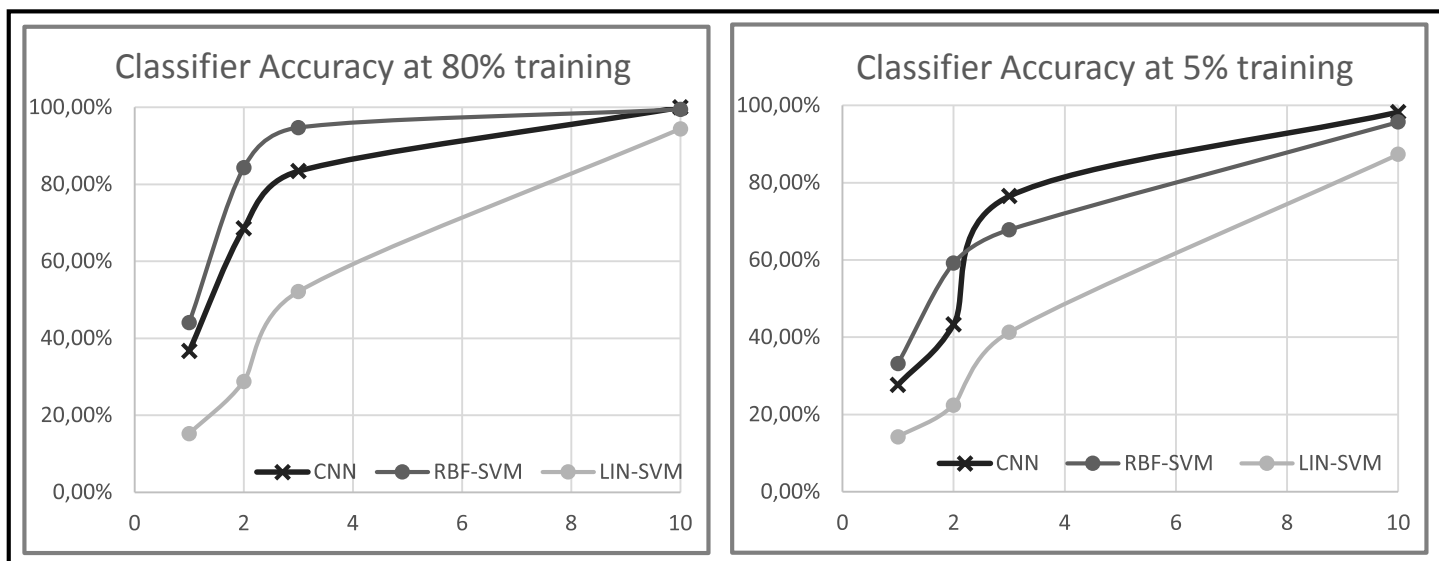
Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, αποδίδοντας σε περίπου ένα λεπτό. Στις περιπτώσεις στις οποίες ο

αλγόριθμος CNN δεν έφτασε το προκαθορισμένο όριο επαναλήψεων, η χρονική απόδοσή του κυμαίνεται από 5,5 έως 8 λεπτά.

Παρακάτω, παρουσιάζονται τα αποτελέσματα των ταξινομήσεων με μειωμένο αριθμό Κυρίων Συνιστωσών, για τα δεδομένα των 98 κλάσεων. Αυτές ορίστηκαν σε 3, 2 και 1 με το αντίστοιχο ποσοστό της υπερφασματικής πληροφορίας να είναι 98,08%, 94,20% και 72,43%. Οι εφαρμογές αυτές αφορούν και τους 3 ταξινομητές για τα ποσοστά εκπαίδευσης 80% και 5%.

	CLASSIFIER ACCURACY				
	Training percentage	Number of PCA components			
		1	2	3	10
LIN	80%	15,22%	28,79%	52,12%	94,38%
	5%	14,20%	22,41%	41,24%	87,31%
RBF	80%	44,07%	84,31%	94,71%	99,45%
	5%	33,20%	59,18%	67,82%	95,71%
CNN	80%	36,72%	68,53%	83,42%	99,98%
	5%	27,65%	43,30%	76,50%	98,27%

Πίνακας 5.15: Συγκριτικά αποτελέσματα ταξινομήσεων με 98 κλάσεις με διάφορες τιμές Κ.Σ.



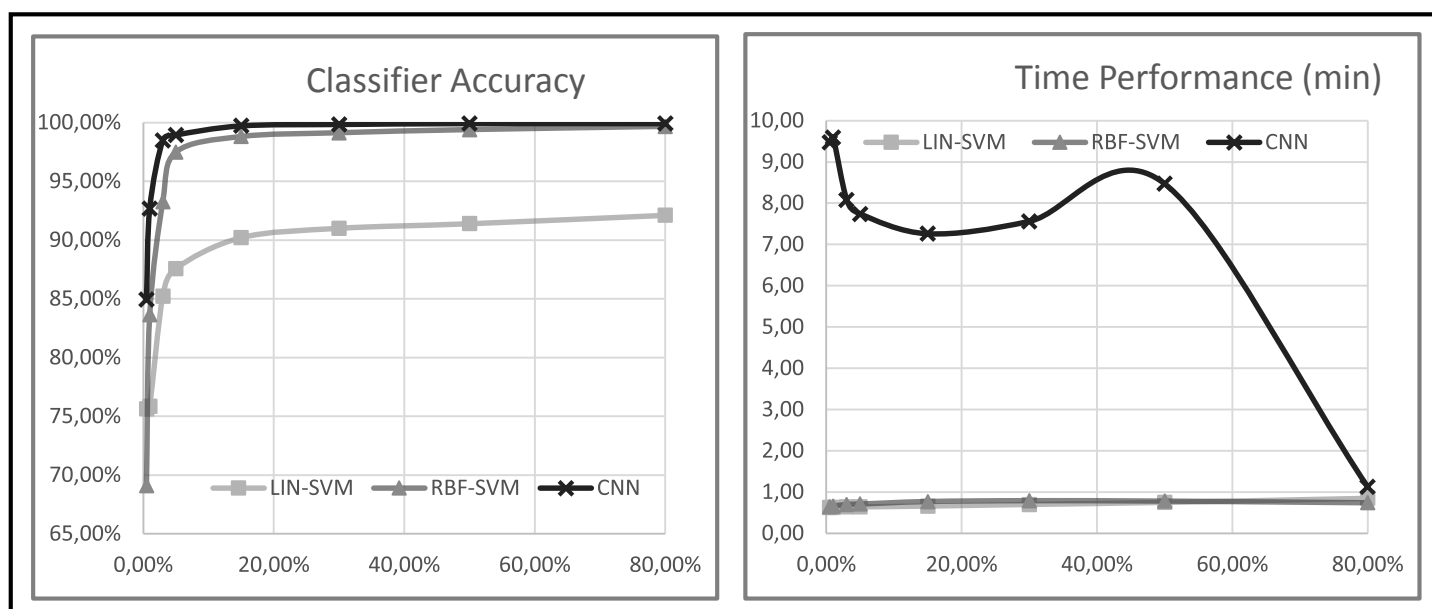
Εικόνα 5.41: Συγκριτικά γραφήματα ακριβειών των τριών ταξινομητών, για διάφορες τιμές Κ.Σ. και ποσοστού εκπαίδευσης

Όπως φαίνεται από τους πίνακες και τα διαγράμματα οι ταξινομητές CNN και RBF επιτυγχάνουν κατά πολύ καλύτερες ακρίβειες από αυτές του γραμμικού SVM. Επιπλέον, για ποσοστό εκπαίδευσης 80% ο ταξινομητής RBF αποδίδει καλύτερα στις περιπτώσεις μειωμένου αριθμού Κ.Σ. όπως επίσης και στις περιπτώσεις 1 και 2 Κ.Σ. σε ποσοστό εκπαίδευσης 5%.

Παρακάτω παρουσιάζονται τα αποτελέσματα που αφορούν τις ταξινομήσεις με 10 Κύριες Συνιστώσες και 99,94% ποσοστό της υπερφασματικής πληροφορίας, στα δεδομένα Ground Truth με τις 54 κλάσεις.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	accuracy	75,63%	75,85%	85,24%	87,57%	90,21%	91,01%	91,40%	92,11%
	time (min)	0,63	0,63	0,64	0,64	0,66	0,70	0,75	0,85
RBF	accuracy	69,09%	83,65%	93,26%	97,48%	98,81%	99,13%	99,40%	99,67%
	time (min)	0,64	0,66	0,70	0,71	0,77	0,79	0,78	0,74
CNN	accuracy	84,93%	92,69%	98,47%	98,94%	99,73%	99,85%	99,92%	99,92%
	time (min)	9,47*	9,59*	8,08*	7,74*	7,26*	7,56*	8,47*	1,13

Πίνακας 5.16: Τα αποτελέσματα των ταξινομήσεων των δεδομένων της ενιαίας εικόνας, για τις 54 κλάσεις



Εικόνα 5.42: Τα γραφήματα απόδοσης ακρίβειας και χρόνου στα δεδομένα της ενιαίας εικόνας, για τις 54 κλάσεις

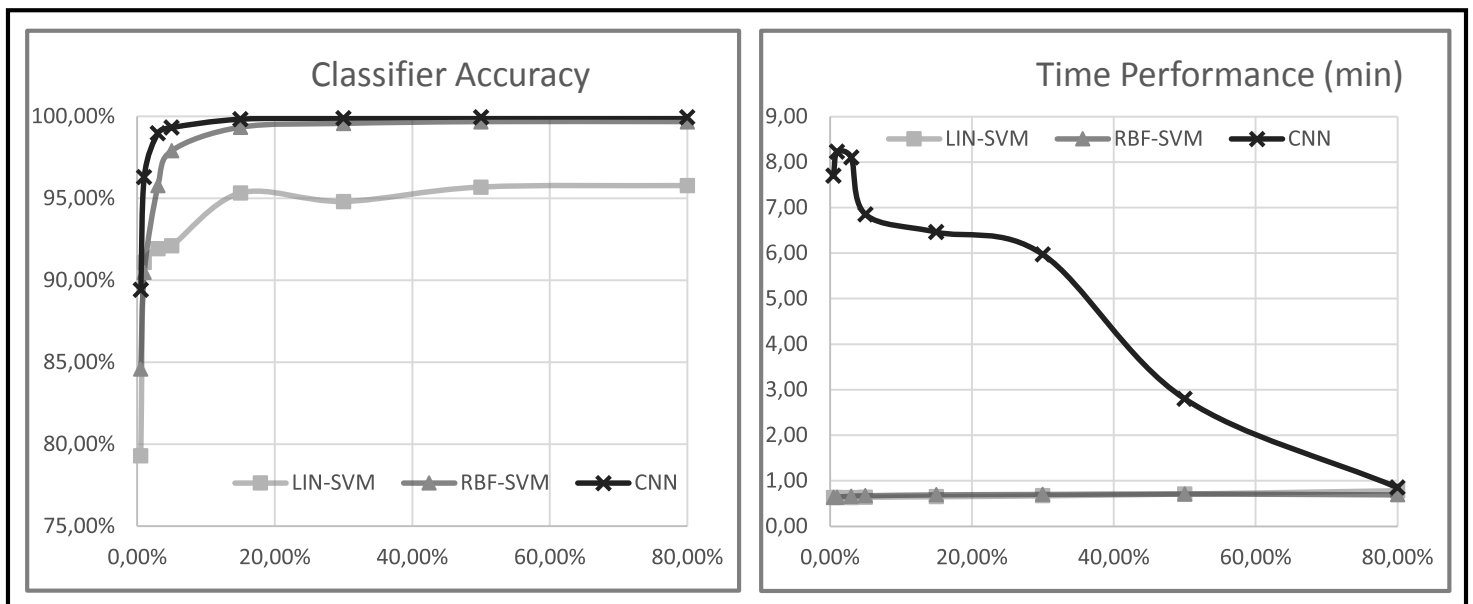
Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Βεβαίως, και ο ταξινομητής RBF επιτυγχάνει υψηλά επίπεδα ακρίβειας, παραπλήσια με αυτά του CNN, για ποσοστό εκπαίδευσης μεγαλύτερο του 5%. Σημαντική διαφοροποίηση παρατηρείται στα αποτελέσματα των ποσοστών εκπαίδευσης 0,5% και 1% όπου ο CNN αποδίδει με ακρίβεια καλύτερη κατά 15% και 9% αντίστοιχα.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, αποδίδοντας σε περίπου ένα λεπτό. Όμως, στην περίπτωση στην οποία ο αλγόριθμος CNN δεν έφτασε το προκαθορισμένο όριο επαναλήψεων (80%), η χρονική απόδοσή του προσεγγίζει αυτή των SVM.

Παρακάτω παρουσιάζονται τα αποτελέσματα που αφορούν τις ταξινομήσεις με 10 Κύριες Συνιστώσες και 99,94% ποσοστό της υπερφασματικής πληροφορίας, στα δεδομένα Ground Truth με τις 29 κλάσεις.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	accuracy	79,29%	91,10%	91,94%	92,11%	95,33%	94,81%	95,69%	95,78%
	time (min)	0,63	0,63	0,63	0,64	0,65	0,67	0,71	0,78
RBF	accuracy	84,60%	90,49%	95,78%	97,90%	99,34%	99,57%	99,68%	99,68%
	time (min)	0,64	0,64	0,66	0,67	0,69	0,70	0,71	0,69
CNN	accuracy	89,44%	96,30%	98,98%	99,32%	99,83%	99,89%	99,96%	99,96%
	time (min)	7,70*	8,23*	8,10*	6,85*	6,46*	5,97	2,80	0,85

Πίνακας 5.17: Τα αποτελέσματα των ταξινομήσεων των δεδομένων της ενιαίας εικόνας, για τις 29 κλάσεις



Εικόνα 5.43: Τα γραφήματα απόδοσης ακρίβειας και χρόνου στα δεδομένα της ενιαίας εικόνας, για τις 29 κλάσεις

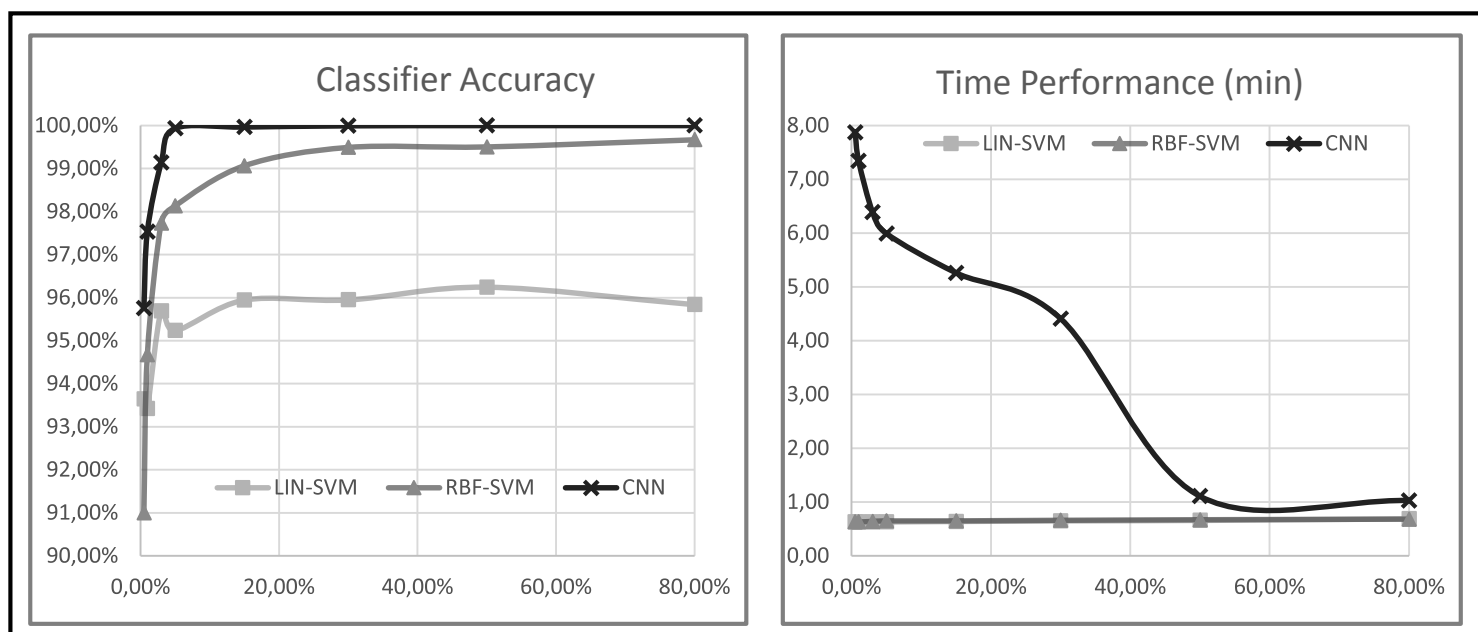
Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM. Βεβαίως, και ο ταξινομητής RBF επιτυγχάνει υψηλά επίπεδα ακρίβειας, παραπλήσια με αυτά του CNN σε όλα τα ποσοστά εκπαίδευσης. Πάντως, παρατηρείται σημαντική διαφοροποίηση στα αποτελέσματα των ποσοστών εκπαίδευσης 0,5% και 1% όπου ο CNN αποδίδει με ακρίβεια καλύτερη κατά 5% περίπου.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, αποδίδοντας σε λιγότερο από ένα λεπτό. Όμως, στην περίπτωση του 80%, στην οποία ο αλγόριθμος CNN δεν έφτασε το προκαθορισμένο όριο επαναλήψεων, η χρονική απόδοσή του προσεγγίζει αυτή των SVM.

Τέλος, παρουσιάζονται και τα αποτελέσματα που αφορούν τις ταξινομήσεις με 10 Κύριες Συνιστώσες και 99,94% ποσοστό της υπερφασματικής πληροφορίας, στα δεδομένα Ground Truth με τις 9 κλάσεις.

TRAINING		0,50%	1%	3%	5%	15%	30%	50%	80%
LIN	accuracy	93,64%	93,43%	95,69%	95,24%	95,95%	95,95%	96,25%	95,84%
	time (min)	0,63	0,63	0,63	0,63	0,64	0,65	0,66	0,69
RBF	accuracy	91,00%	94,67%	97,73%	98,13%	99,07%	99,49%	99,50%	99,67%
	time (min)	0,63	0,63	0,64	0,65	0,65	0,66	0,67	0,68
CNN	accuracy	95,75%	97,53%	99,14%	99,93%	99,97%	99,99%	100,00%	100,00%
	time (min)	7,87*	7,34*	6,39*	5,99*	5,26	4,41	1,11	1,03

Πίνακας 5.18: Τα αποτελέσματα των ταξινομήσεων των δεδομένων της ενιαίας εικόνας, για τις 9 κλάσεις



Εικόνα 5.44: Τα γραφήματα απόδοσης ακρίβειας και χρόνου στα δεδομένα της ενιαίας εικόνας, για τις 9 κλάσεις

Όπως φαίνεται από τον πίνακα και τα διαγράμματα ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες από τους ταξινομητές SVM, φτάνοντας σχεδόν στην απόλυτη ακρίβεια του 100%, ήδη από το 3% του ποσοστού εκπαίδευσης. Βεβαίως, και οι ταξινομητές SVM επιτυγχάνουν υψηλά επίπεδα ακρίβειας, άνω του 95%, για ποσοστό εκπαίδευσης μεγαλύτερο του 5%.

Όσον αφορά τη χρονική απόδοση τους, οι αλγόριθμοι SVM είναι σαφώς ταχύτεροι του CNN, αποδίδοντας σε λιγότερο από ένα λεπτό. Όμως, στις περιπτώσεις του 80% και 50%, στην οποία ο αλγόριθμος CNN δεν έφτασε το προκαθορισμένο όριο επαναλήψεων, η χρονική απόδοσή του προσεγγίζει αυτή των SVM.

6

ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ

Στο κεφάλαιο αυτό παρουσιάζονται τα συμπεράσματα που προέκυψαν κατά την εκπόνηση αυτής της εργασίας. Τα συμπεράσματα αυτά χωρίζονται σε γενικά και ειδικά. Τα γενικά συμπεράσματα καταπιάνονται με τη συνολική ποιοτική αξιολόγηση των αλγορίθμων που εφαρμόστηκαν, ενώ τα ειδικά τις πιο επιμέρους εξειδικευμένες εφαρμογές που υλοποιήθηκαν μαζί με τα όποια αποτελέσματα προέκυψαν από αυτές. Τέλος, αναφέρονται και μερικές προτάσεις που προέκυψαν τόσο ως προβληματισμοί αλλά και ως πρακτικά προβλήματα κατά τη διεκπεραίωση της εργασίας.

6.1 Γενικότερα Συμπεράσματα

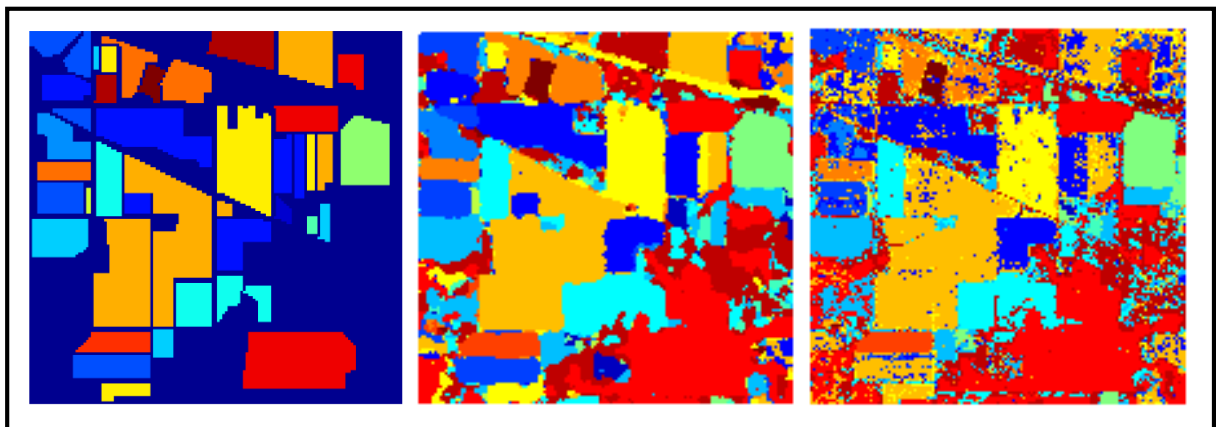
Το βασικό συμπέρασμα που εξάγεται από τα αποτελέσματα της εργασίας, είναι ότι οι μέθοδοι και οι αλγόριθμοι Machine Learning (εκμάθησης μηχανής) αποτελούν ισχυρά εργαλεία στο πρόβλημα της επιβλεπόμενης ταξινόμησης εικόνας. Η απόδοση και των 3 αλγορίθμων που εφαρμόστηκαν επιτρέπει την υλοποίηση ταξινομήσεων, από μεγάλα υπερφασματικά δεδομένα εικόνας, παρουσιάζοντας υψηλά ποσοστά ακρίβειας ακόμα και με λίγα δεδομένα εκπαίδευσης.

Επιπροσθέτως, το νευρωνικό δίκτυο CNN που χρησιμοποιήθηκε στην εργασία και αποτελεί δομή και αρχιτεκτονική Deep Learning (εκμάθησης σε βάθος), παρουσιάζει ακρίβειες που συγκρίνονται και υπερτερούν αντίστοιχων εφαρμογών στη διεθνή βιβλιογραφία. Κάτι τέτοιο άλλωστε, αναδεικνύει και η τάση των ολοένα και αυξανόμενων εφαρμογών αλγορίθμων Deep Learning στο αντικείμενο της όρασης υπολογιστών και της ταξινόμησης εικόνας ειδικότερα.

Όσον αφορά τις γενικές εφαρμογές και τα αποτελέσματα των ταξινομητών:

- Από τις δεκάδες δοκιμές ταξινομήσεων που εφαρμόστηκαν, το νευρωνικό δίκτυο CNN επιτυγχάνει, σχεδόν σε όλες, καλύτερες ακρίβειες από τις SVM. Στις μοναδικές περιπτώσεις που ένας ταξινομητής SVM υπερίσχυσε σε απόδοση, το ποσοστό εκπαίδευσης είχε τεθεί στο 0,5% είτε στο 1% οριοθετώντας έτσι την απόδοση των ταξινομητών από τα τυχαία - λίγα αναλογικά - δείγματα Ground Truth.

- Οι ταξινομήσεις που εφαρμόστηκαν πάνω στα βιβλιογραφικά δεδομένα παρουσιάζουν και τα χαμηλότερα ποσοστά ακρίβειας. Παράλληλα όμως, σε αυτά τα αποτελέσματα παρουσιάζονται και οι μεγαλύτερες διαφοροποιήσεις ανάμεσα στην απόδοση των ταξινομητών CNN και SVM. Συγκεκριμένα, στα δεδομένα της εικόνας *Indian Pines* (για 55 και 10 κύριες συνιστώσες) ο ταξινομητής CNN επιτυγχάνει καλύτερες ακρίβειες κατά περίπου 20% και από τους δύο ταξινομητές SVM σε όλες τις περιπτώσεις δεδομένων εκπαίδευσης. Αντιστοίχως, στα δεδομένα της εικόνας *Pavia University*, η διαφορά ακρίβειας από τον δεύτερο σε σειρά ταξινομητή κυμαίνεται περίπου στο 10% για όλα τα δεδομένα εκπαίδευσης (βλ. Πίνακες 5.4 – 5.3 και Πίνακας 5.8 αντίστοιχα).
- Η χρονική απόδοση του ταξινομητή CNN είναι σαφώς χειρότερη των SVM. Σχεδόν σε όλες τις εφαρμογές η διαδικασία της εκμάθησης για τις SVM καταλαμβάνει μόλις μερικά δευτερόλεπτα, ενώ η αντίστοιχη του CNN καταλαμβάνει από λίγα δευτερόλεπτα έως δεκάδες λεπτά.
- Ένα σημαντικό στοιχείο που επιτρέπει στον ταξινομητή CNN να επιτυγχάνει υψηλά ποσοστά ακρίβειας και να παράγει ορθότερα ταξινομημένες εικόνες, είναι το ότι κατά τη διαδικασία της εκμάθησης περιλαμβάνει και χωρική πληροφορία για το εκάστοτε φαντίο. Πέραν της ραδιομετρικής πληροφορίας του εκάστοτε στοιχείου εκπαίδευσης, ο ταξινομητής συνυπολογίζει και τις ραδιομετρικές τιμές των γειτονικών στοιχείων (μέσω των συνελκτικών επιπέδων), πετυχαίνοντας με αυτό τον τρόπο καλύτερη και αποτελεσματικότερη εκμάθηση για το στοιχείο αυτό (βλ. Κεφάλαιο 3.4). Στην παρακάτω εικόνα παρουσιάζονται αποτελέσματα ταξινόμησης για την εικόνα *Indian Pines* για 55 Κ.Σ. και 80% ποσοστό εκπαίδευσης . Είναι εμφανές το ποιοτικότερο και πιο ‘συμπαγή’ αποτέλεσμα ταξινόμησης για τον CNN από το αντίστοιχο του RBF.



Εικόνα 6.1: Από αριστερά: Ground Truth εικόνας *Indian Pines*, αποτέλεσμα ταξινόμησης CNN και αποτέλεσμα ταξινόμησης RBF-SVM για 55 Κ.Σ. και 80% ποσοστό εκπαίδευσης

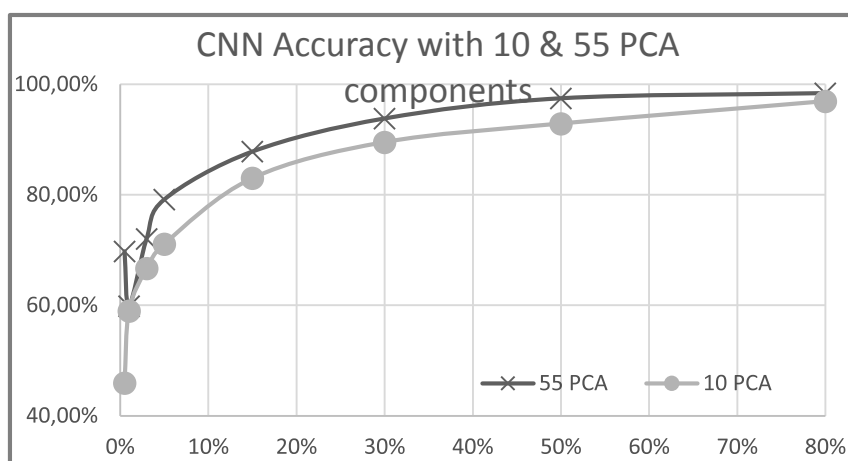
- Η διαδικασία εκμάθησης (training) και πρόβλεψης (prediction) σε αλγόριθμους machine learning, απαιτεί μεγάλο υπολογιστικό φόρτο. Παρ' όλο που ο υπολογιστής εργασίας είχε τεχνικά χαρακτηριστικά υψηλής απόδοσης (βλ. Κεφάλαιο 5.2.5), υπήρξαν περιπτώσεις που η ολοκλήρωση μιας εφαρμογής χρειαζόταν δεκάδες λεπτά. Και εάν κάτι τέτοιο αποτυπώνεται στα αποτελέσματα χρονικής απόδοσης για την εκμάθηση του ταξινομητή CNN, αναφέρεται πως ανάλογες χρονοβόρες αποδόσεις δεκάδων λεπτών συναντώνται κατά τη διαδικασία πρόβλεψης του ταξινομητή RBF-SVM.

6.2 Ειδικά Συμπεράσματα

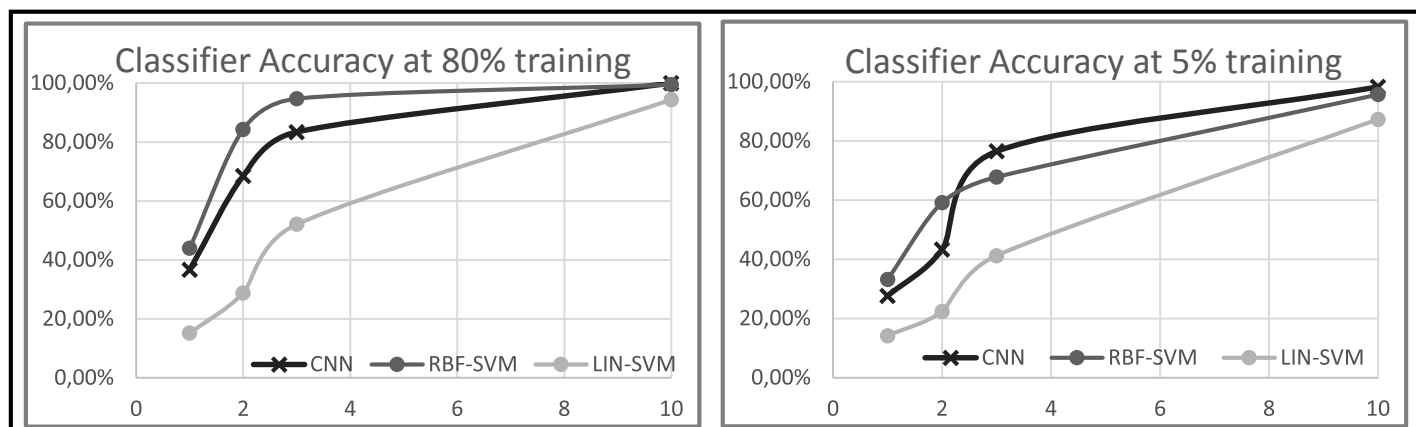
Όσον αφορά τις πιο σύνθετες εφαρμογές και τα αποτελέσματα των ταξινομητών:

- Η απόδοση της εκμάθησης του νευρωνικού δικτύου CNN εμφανίζει μεγάλες χρονικές διαφοροποιήσεις, ανάμεσα στις περιπτώσεις που ο αλγόριθμος φτάνει το προκαθορισμένο όριο επαναλήψεων και στις περιπτώσεις που τερματίζεται νωρίτερα. Ενδεικτικά είναι τα αποτελέσματα της εικόνας *Indian Pines* για 55 Κ.Σ., όπου για τα ποσοστά εκπαίδευσης 50% και 30%, ο αλγόριθμος απέδωσε σε περίπου 35 λεπτά στην 1^η περίπτωση και 7 λεπτά στην 2^η. Αυτές οι μεγάλες χρονικές διακυμάνσεις, οι οποίες συναντώνται αναλογικά σχεδόν σε όλα τα δεδομένα που δοκιμάστηκαν, εξαρτώνται σε μεγάλο βαθμό από την a priori παραμετροποίηση του αλγορίθμου (βλ. Κεφάλαιο 5.2.4) και μπορούν να εξομαλυνθούν αποφέροντας βεβαίως αντίστοιχες επιπτώσεις στις παραγόμενες ακρίβειες.
- Και οι 3 ταξινομητές αποδίδουν ταχύτερα όσο μειώνεται ο αριθμός των Κ.Σ. (Κυρίων Συνιστωσών - PCA). Αυτή η βελτίωση ταχύτητας, είναι εμφανέστερη στα αποτελέσματα του πιο χρονοβόρου ταξινομητή CNN. Ενδεικτικά, η μείωση των Κ.Σ. από 55 σε 10 στην εικόνα *Indian Pines* απέφερε χρονική μείωση της τάξης των 34 – 7 λεπτών. Στα δεδομένα της εικόνας *Botswana* πάλι, η μείωση του αριθμού των Κ.Σ. από 25 σε 10 απέφερε σχεδόν υποτριπλασιασμό της αντίστοιχης χρονικής απόδοσης, σε όλα τα ποσοστά δεδομένων εκπαίδευσης. Αυτές οι χρονικές εξαρτήσεις από τον επιλεγμένο αριθμό Κ.Σ. είναι λογικές αφού η όποια αύξηση του αριθμού αυτού επιδρά πολλαπλασιαστικά στον αριθμό των δημιουργούμενων νευρώνων και άρα στον φόρτο εργασίας του Η/Υ εργασίας.

- Οι παραγόμενες ακρίβειες και των 3 ταξινομητών εξαρτώνται από την επιλογή του αριθμού Κ.Σ. και άρα το ποσοστό της υπερφασματικής πληροφορίας που θα χρησιμοποιηθεί κατά την διαδικασία της εκμάθησης. Βεβαίως η εξάρτηση αυτή διαφέρει από εφαρμογή σε εφαρμογή και από δεδομένα σε δεδομένα. Έτσι, μείωση των Κ.Σ. από 55 σε 10 (99,81% σε 97,21%) στην εικόνα *Indian Pines* απέφερε μείωση και στις παραγόμενες ακρίβειες οι οποίες όμως δεν υπερβαίνουν το 5% στα μεγάλα ποσοστά εκπαίδευσης για τον ταξινομητή CNN (βλ. *Εικόνα 5.6*). Στα δεδομένα της εικόνας *Botswana* πάλι, η μείωση του αριθμού των Κ.Σ. από 25 σε 10 (99,82% σε 99,66%) απέφερε ελάχιστη μεταβολή στις παραγόμενες ακρίβειες. Συμπεραίνεται με αυτό τον τρόπο ότι η μεταβολή του ποσοστού της υπερφασματικής πληροφορίας που χρησιμοποιείται, επιδρά περισσότερο στην όποια μεταβολή των παραγόμενων ακριβειών. Τέλος, η αύξηση των Κ.Σ. από 1 σε 2,3 και 10 (από 72,43% σε 94,20%, 98,08% και 99,94%) στην ενιαία εικόνα *SFB all* απέφερε αύξηση και στις παραγόμενες ακρίβειες οι οποίες φαίνεται να ακολουθεί την λογαριθμική καμπύλη στα ποσοστά εκπαίδευσης 80% και 5% (βλ. *Εικόνα 5.41*).

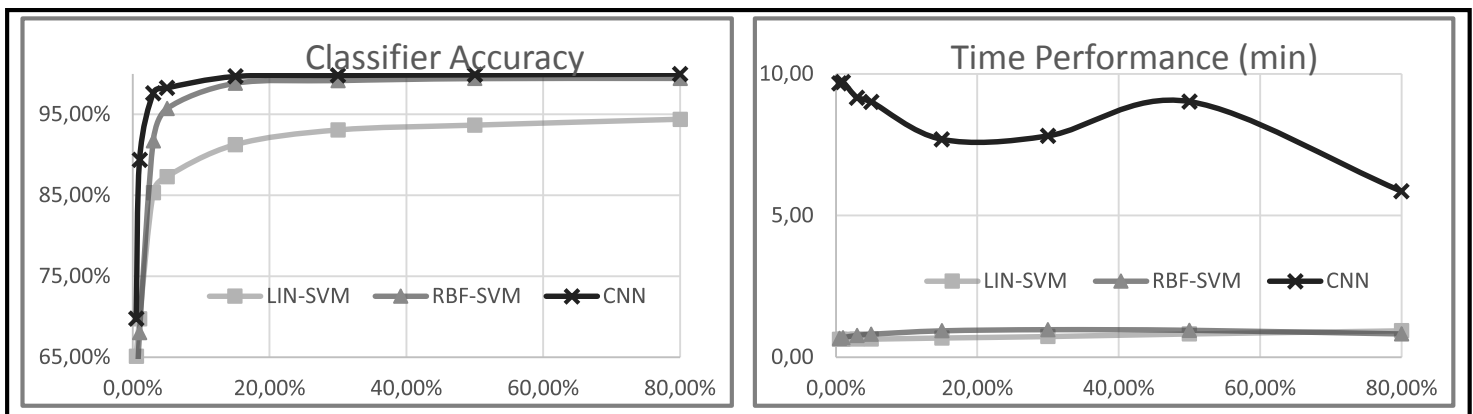


Εικόνα 6.2: Συγκριτικό γράφημα ακριβειών του CNN για τα δεδομένα *Indian Pines*, με 55 Κ.Σ. και 10 Κ.Σ.

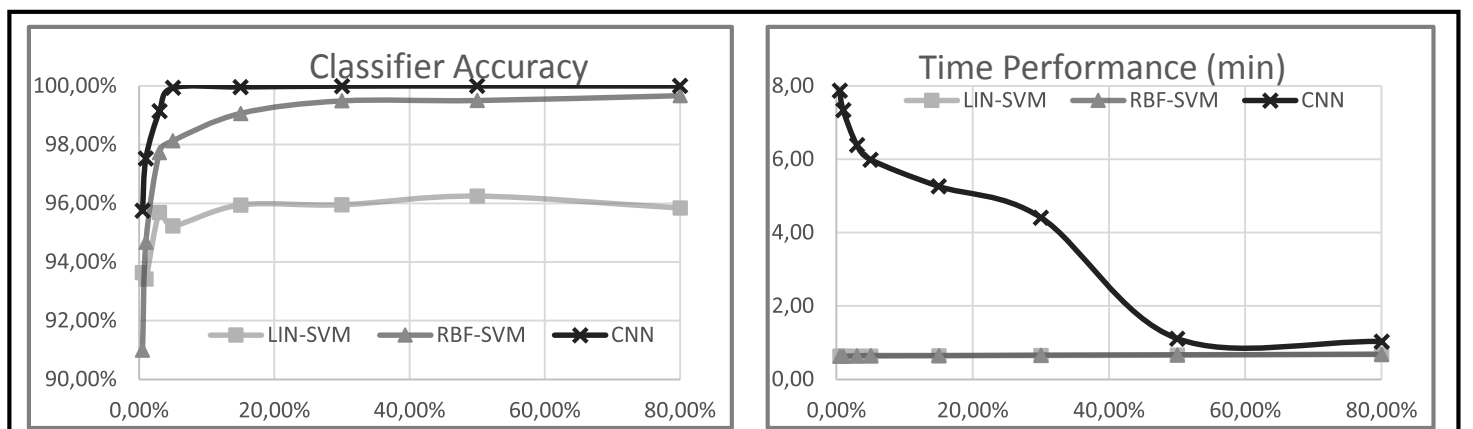


Εικόνα 6.3: Συγκριτικά γραφήματα ακριβειών των τριών ταξινομητών, για διάφορες τιμές Κ.Σ. και ποσοστού εκπαίδευσης

- Όσον αφορά την μείωση στον αριθμό των κλάσεων (από 98 σε 54, 29 και 9) που επιβλήθηκε στα δεδομένα Ground Truth της ενιαίας εικόνας *SFB all*, αυτή απέφερε μία αύξηση στα παραγόμενα ποσοστά ακρίβειας μαζί με μία παράλληλη μείωση του χρόνου απόδοσης των αλγορίθμων. Ενδεικτικά παρουσιάζονται παρακάτω τα αντίστοιχα γραφήματα αποτελεσμάτων των 98 και 9 κλάσεων (βλ. *Εικόνα 5.40* και *Εικόνα 5.44*). Η μείωση του χρόνου απόδοσης των ταξινομητών οφείλεται στην μείωση του αριθμού κλάσεων στον οποίο καλείται να 'εκπαιδευτεί' ο επιμέρους ταξινομητής που αυτομάτως συνεπάγεται σε μικρότερο φόρτο εργασίας για τον Η/Υ. Αντίστοιχα η μείωση του αριθμού των κλάσεων οδηγεί σε ευκολότερη ταξινόμηση των υπό εξέταση φατνίων στην 'σωστή' κάθε φορά κλάση. Αυτό συμβαίνει καθώς οι νέες κατηγορίες εδάφους (π.χ. 9) παρουσιάζουν μεγαλύτερες διαφορές στις φασματικές τους υπογραφές από ότι στην προηγούμενη κατάσταση (π.χ. 98). Με αυτόν τον τρόπο ένα εικονοστοιχείο που αποτυπώνει καλλιεργήσιμη έκταση λόγου χάρη, και πρόκειται να ταξινομηθεί, έχει μία υποψήφια κλάση καλλιεργείων και όχι 7 όπως στην αρχική κατάσταση.



Εικόνα 6.4: Τα γραφήματα απόδοσης ακρίβειας και χρόνου στα δεδομένα της ενιαίας εικόνας, για τις 98 κλάσεις



Εικόνα 6.5: Τα γραφήματα απόδοσης ακρίβειας και χρόνου στα δεδομένα της ενιαίας εικόνας, για τις 9 κλάσεις

6.3 Προτάσεις

Σαν προτάσεις για την υλοποίηση πληρέστερων και πιο εξειδικευμένων εφαρμογών προτείνεται:

- Ο πειραματισμός με τα διάφορα σημαντικά στοιχεία της αρχιτεκτονικής, και των συντελεστών του νευρωνικού δικτύου CNN. Τέτοια στοιχεία θα μπορούσαν να είναι: ο αριθμός των νευρώνων κάθε επιπέδου, ο αλγόριθμος εκπαίδευσης, η τιμή ορισμού για κάθε Minibatch στην περίπτωση εκπαίδευσης μέσω MSGD, η ένταξη επιπέδου μέγιστης συγκέντρωσης έπειτα από κάθε συνελκτικό επίπεδο κλπ. Τέτοιες εναλλαγές θα μπορούσαν να προσδιορίσουν την πραγματική βέλτιστη δομή και παραμετροποίηση ενός δικτύου CNN, για εφαρμογές ταξινόμησης υπερφασματικών δεδομένων.
- Η εφαρμογή και άλλων ταξινομήσεων, με εναλλακτικές τιμές στις παραμέτρους που επηρεάζουν την χρονική απόδοση του CNN. Παράμετροι όπως ο ρυθμός εκμάθησης (learning rate), το κατώφλι προόδου (improvement threshold) και ο αριθμός των επαναλήψεων έως τον τερματισμό της εκμάθησης (epochs) μπορούν να αποδώσουν διαφορετικά και σίγουρα ενδιαφέροντα αποτελέσματα τόσο σε σχέση με τη χρονική απόδοση όσο και με την παραγόμενη ακρίβεια.
- Η εφαρμογή ταξινομήσεων με δεδομένο αριθμό δειγμάτων εκπαίδευσης για κάθε κλάση. Με μία τέτοια κατάλληλη δομή εκπαίδευσης, οι αποδόσεις ακρίβειας θα παρουσιάζουν ποιοτικότερο και αναλογικό χαρακτήρα πάνω στην ορθή ή λανθασμένη ταξινόμηση της κάθε επιμέρους κλάσης.
- Η υλοποίηση συνθετότερων εφαρμογών ως προς της ενοποιημένες εικόνες. Στην παρούσα εργασία υλοποιήθηκε κοινή ταξινόμηση στις 7 εικόνες των δεδομένων SFB, έπειτα από την ενοποίησή τους. Προτείνεται η ενοποίηση εικόνων και από τα βιβλιογραφικά δεδομένα. Τέτοιες εφαρμογές θα αποφέρουν αποτελέσματα τα οποία θα μπορέσουν να αποτελέσουν δείκτη σύγκρισης στη διεθνή βιβλιογραφία καθώς κάτι παρόμοιο δεν έχει δοκιμαστεί έως τώρα.
- Η υλοποίηση συνθετότερων εφαρμογών ως προς τα δεδομένα εκπαίδευσης. Οι ταξινομήσεις που εφαρμόστηκαν σε αυτή την εργασία είχαν όλες σαν αντικείμενο, κοινά δεδομένα Ground Truth και παράγωγο εικόνας. Προτείνεται η εφαρμογή ταξινομήσεων κατά τις οποίες ο αλγόριθμος θα 'εκπαιδεύεται' σε διαφορετικά ή εν μέρει κοινά δεδομένα από αυτά που θα καλείται να ταξινομήσει.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Βιβλία και Εργασίες

- [1] Μακαντάσης Κ., Καράντζαλος Κ., Δουλάμης Α., Δουλάμης Ν., 2015. **Deep supervised learning for hyperspectral data classification through convolutional neural networks**, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015
- [2] Wei Hu, Yangyu Huang, Li Wei, Fan Zhang, Hengchao Li, 2015. **Deep Convolutional Neural Networks for Hyperspectral Image**, *Journal of Sensors*, article ID. 258619, vol. 2015
- [3] Bishop C.M., 2007. **Pattern Recognition and Machine Learning**, Springer
- [4] Yushi Chen, Zhouhan Lin, Xing Zhao, Gang Wang, Yanfeng Gu, 2014. **Deep Learning-Based Classification of Hyperspectral Data**, *IEEE journal of selected topics in applied earth observations and remote sensing*, VOL. 7, NO. 6, 2015
- [5] Chen Chen, Wei Li, Hongjun Su, Kui Liu, 2014. **Spectral-Spatial Classification of Hyperspectral Image Based on Kernel Extreme Learning Machine**, *Remote Sensing (MDPI)*, vol. 6, no. 6, pp. 5795-5814, 2014
- [6] Jun Li, Bioucas-Dias J.M., Plaza A., 2013. **Spectral–Spatial Classification of Hyperspectral Data Using Loopy Belief Propagation and Active Learning**, *IEEE transactions on geoscience and remote sensing*, vol. 51, no. 2, 2013
- [7] Xiaorui Ma, Jie Geng, Hongyu Wang, 2015. **Hyperspectral image classification via contextual deep learning**, *EURASIP Journal on Image and Video Processing*, 2015
- [8] Krizhevsky A., Sutskever I., Hinton G.E, 2012. **ImageNet Classification with Deep Convolutional Neural Networks**, *Advances in Neural Information Processing Systems Conference*, 2012
- [9] Girshick R., Donahue J., Darrell T., Malik J., 2014. **Rich feature hierarchies for accurate object detection and semantic segmentation**, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015
- [10] Basu S., Ganguly S., Mukhopadhyay S., DiBiano R., Karki M., Nemani R., 2015. **DeepSat – A Learning framework for Satellite Imagery**
- [11] Βακαλοπούλου Μ., Καράντζαλος Κ., Κομοδάκης Ν., Παραγιός Ν., 2015, **Simultaneous Registration and Change Detection in Multitemporal, Very High Resolution Remote Sensing Data**, *IEEE Computer Vision and Pattern Recognition Workshops, CVPRW'15, June 2015*, pp.1-

- [12] Theodoridis S., Koutroumbas K., 2006. **Pattern Recognition vol 3**, Academic Press
- [13] Romero A., Gatta C., Camps-Valls G., 2014. **Spectral-Spatial Classification of Hyperspectral Image Based on Kernel Extreme Learning Machine**, *IEEE Workshop on Hyperspectral Image and Signal Processing (WHISPER) conference 2014*
- [14] Tuia D., Flamary R., Courty N., 2015. **Multiclass feature learning for hyperspectral image classification: sparse and hierarchical solutions**, *ISPRS Journal of Photogrammetry and Remote Sensing, Elsevier, 2015, pp.1-14 2015*
- [15] Chen C., Li W., Tramel E.W., Cui M., Prasad S., Fowler J.E., 2014. **Spectral-spatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification**, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 4, pp. 1047-1059, 2014*
- [16] Hinton G.E., Osindero S., Teh Y.W., 2006. **A fast learning algorithm for deep belief nets**, *Journal in neural computation pages 1527-1554, 2006*
- [17] Αργιαλάς Δ., 1998, **Ψηφιακή Τηλεπισκόπηση**
- [18] Πούλος Δ., 2013. **Ανάπτυξη και αξιολόγηση αλγορίθμων για τον αυτόματο εντοπισμό σταθμευμένων αεροσκαφών σε διαχρονικά δορυφορικά δεδομένα**
- [19] Αλεβιζάκου Ε.Γ., 2012. **Η χρήση των τεχνητών νευρωνικών δικτύων στην επιστήμη της γεωδαισία με έμφαση στην πρόβλεψη κατακόρυφων μετακινήσεων**
- [20] Πλέρου Α., 2012. **Τεχνητά νευρωνικά δίκτυα προσομοίωσης του ανθρώπινου εγκεφάλου**, *Open Education - The Journal for Open and Distance Education and Educational Technology Volume 8, Number 1, 2012*
- [21] Sibi P., Allwyn Jones S., Siddarth P., 2013. **Analysis of different activation functions using back propagation neural networks**, *journal of theoretical and applied information technology, 2013*
- [22] Αργυράκης Π., 2001. **Νευρωνικά Δίκτυα και Εφαρμογές**
- [23] Πυρπιρή Θ.Χ., 2014. **Υπολογισμός του συντελεστή κυματικής μετάδοσης με χρήση νευρωνικών δικτύων & πειραματικών**
- [24] Κόλλιας Δ., 2015. **Τεχνικές Βαθιάς Μηχανικής Μάθησης και Γνώσης για Ανάλυση Συναισθήματος στην Αλληλεπίδραση Ανθρώπου**
- [25] Gege P., Bachmann S., Holzwarth S., 2008. **Airborne Remote Sensing – ROSIS Bonn 2008 Summary**
- [26] Thompson., 2013. **AVIRIS surface reflectance processing and products**
- [27] Thenkabail .P, Lyon J., Huete A., 2011. **Hyperspectral Remote Sensing of Vegetation**, *CRC press*

Διαδικτυακές πηγές

**Τελευταία επίσκεψη – Οκτώβριος 2015*

1. www.ehu.eu
2. www.aviris.jpl.nasa.gov
3. www.lars.purdue.edu
4. www.engineering.purdue.edu
5. www.spectir.com
6. www.microimages.com
7. www.escience.cn
8. www.cs.cmu.edu
9. www.isp.uv.es
10. www.engineering.purdue.edu
11. www.eo1.usgs.gov
12. www.opticks.org
13. www.exelisvis.com
14. www.hindawi.com
15. www.dl.acm.org
16. www.ieeexplore.ieee.org
17. www.deeplearning.net
18. www.colah.github.io
19. www.andrew.gibiansky.com
20. www.en.wikipedia.org
21. www.mathworks.com
22. www.sia.eionet.europa.eu
23. www.qdal.org
24. www.airbornescience.nasa.gov
25. www.engineeronadisk.com
26. www.scikit-learn.org

