



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**ΜΕΘΟΔΟΙ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗΝ ΤΑΞΙΝΟΜΗΣΗ
ΔΟΡΥΦΟΡΙΚΩΝ ΥΠΕΡΦΑΣΜΑΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ**

**ΟΥΖΟΥΝΙΔΗΣ ΓΕΩΡΓΙΟΣ
ΙΟΥΛΙΟΣ 2020**



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**ΜΕΘΟΔΟΙ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗΝ ΤΑΞΙΝΟΜΗΣΗ
ΔΟΡΥΦΟΡΙΚΩΝ ΥΠΕΡΦΑΣΜΑΤΙΚΩΝ ΔΕΔΟΜΕΝΩΝ**

Τριμελής Εξεταστική Επιτροπή:

.....

Κ. Καράντζαλος

Επ. Καθηγητής Ε.Μ.Π

.....

Δ. Αργιαλάς

Καθηγητής Ε.Μ.Π

.....

Β. Καραθανάση

Καθηγητής Ε.Μ.Π

ΠΕΡΙΛΗΨΗ

Η εργασία αυτή επικεντρώνεται πάνω στην διερεύνηση, παρουσίαση, εφαρμογή και αξιολόγηση μεθόδων σχετικών με την Βαθιά Μάθηση πάνω στην ταξινόμηση υπερφασματικών δεδομένων. Πιο συγκεκριμένα παρουσιάζονται πρόσφατες μέθοδοι αιχμής που έχουν ξεχωρίσει από την βιβλιογραφία για την αποτελεσματικότητα τους και εφαρμόζονται πάνω στις δορυφορικές εικόνες του HyRANK dataset.

Οι μέθοδοι αυτοί υλοποιούνται με την χρήση των νευρωνικών δικτύων. Αρχικά παρουσιάζονται βασικά στοιχεία θεωρίας των νευρωνικών δικτύων. Αναφέρεται ο τρόπος λειτουργίας τους, τα συστατικά του μέρη καθώς και οι διάφοροι τύποι νευρωνικών δικτύων που συναντώνται στις μεθόδους Βαθιάς Μάθησης.

Στη συνέχεια αναλύεται η μεθοδολογική προσέγγιση και η αρχιτεκτονική του δικτύου της κάθε μεθόδου. Περιγράφονται λεπτομερώς τα στρώματα, η μορφή του σήματος στην είσοδο και στην πορεία του δικτύου και όλες οι δομές-λειτουργίες που εισάγονται σε αυτό.

Τέλος παρουσιάζεται η εφαρμογή των μεθόδων αυτών πάνω στις δορυφορικές υπερφασματικές εικόνες και αξιολογούνται τα αποτελέσματα. Παρατηρούνται οι προβλέψεις σε κάθε εικόνα, τα στατιστικά στοιχεία των ταξινομήσεων και εξάγονται συμπεράσματα σχετικά με την αποτελεσματικότητα και λειτουργία των μεθόδων.



ABSTRACT

This thesis focuses on exploration, presentation, implementation and evaluation of Deep Learning HSI classification methods. More precisely, these described as recent ‘state of the art methods’ are utilized on the satellite images of HyRANK dataset.

These methods are based on neural networks. Firstly, basic theory principles on neural networks are presented. This section consists of the operation, the components and the various types of neural networks.

Then the approach of each method is introduced along with the architecture of each network. The layers, the signal path and all the custom modules are presented in detail.

At the end, each method is utilized on the HSI dataset. Every prediction and statistics table is evaluated in order to evaluate their effectiveness and analyze their behavior.



ΠΕΡΙΕΧΟΜΕΝΑ

1	ΕΙΣΑΓΩΓΗ	1
1.1	Υπερφασματικά δεδομένα και ταξινόμηση	1
1.2	Αφορμή συγγραφής της εργασίας.....	4
2	ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	6
2.1	Τεχνητή Νοημοσύνη.....	6
2.2	Μηχανική Μάθηση	8
2.3	Βαθιά Μάθηση.....	9
2.4	Τεχνητά Νευρωνικά δίκτυα	10
2.4.1	Νευρώνας.....	10
2.4.2	Πλήρως συνδεδεμένο στρώμα	10
2.4.3	Συνάρτηση ενεργοποίησης.....	11
2.4.4	Βάρη και τιμή κατωφλιού.....	14
2.4.5	Συνάρτηση κόστους	15
2.4.6	Οπισθοδιάδοση	16
2.4.7	Στοχαστική καταβίβαση της παραγώγου	18
2.4.8	Συνελικτικά νεωρωνικά δίκτυα.....	19
2.4.9	Πλήρως συνελικτικά δίκτυα	23
2.4.10	Αναδραστικά νεωρωνικά δίκτυα.....	24
2.4.11	Δίκτυο Κωδικοποιητή-Αποκωδικοποιητή	28
3	ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ-ΑΛΓΟΡΙΘΜΟΙ	29
3.1	Hamida et al.	29
3.2	Lee et al.....	32
3.3	Chen et al.	35
3.4	Li et al.	36
3.5	Hu et al.....	38
3.6	He et al.	39
3.7	Mou et al.	41
3.8	FDSSC	44
3.9	Liu et al.	46
4	ΜΕΘΟΔΟΓΙΑ	47
4.1	Δεδομένα.....	47
4.2	Κανονικοποίηση των δεδομένων	50
4.3	Προσαρμογή των παραμέτρων	51

4.4	Μέθοδος disjoint	54
4.5	Επιλογή σετ εκπαίδευσης-επαλήθευσης.....	54
4.6	Πίνακας σύγχυσης - στατιστικά.....	55
4.7	Hamida et al.	56
4.8	Lee et al.....	61
4.9	Chen et al.	67
4.10	Li et al.	71
4.11	Hu et al.....	74
4.12	He et al.	77
4.13	Mou et al.	82
4.14	FDSSC	85
4.15	Liu et al.	90
4.16	Οι μέθοδοι SVM και RF	96
	4.16.1 Μηχανή Διανυσματικής Υποστήριξης (SVM)	96
	4.16.2 Μέθοδος Τυχαίων Δασών (RF)	96
5	ΑΠΟΤΕΛΕΣΜΑΤΑ	97
5.1	Hamida et al.	97
5.2	Lee et al.....	99
5.3	Chen et al.	101
5.4	Li et al.	103
5.5	Hu et al.....	105
5.6	He et al.	107
5.7	Mou et al.	109
5.8	FDSSC	111
5.9	Liu et al.	113
5.10	SVM.....	115
5.11	RF.....	117
6	ΣΥΜΠΕΡΑΣΜΑΤΑ	119
6.1	Σύγκριση με βιβλιογραφία	119
6.2	Οπτική σύγκριση προβλέψεων στις εικόνες	122
6.3	Παρατηρήσεις στην ταξινόμηση των κατηγοριών.....	124
6.4	Αρχιτεκτονική των μεθόδων.....	126
6.5	Γενική αξιολόγηση και προτάσεις	128
7	Σχετική Βιβλιογραφία-Αναφορές	131

7.1	Ξενόγλωσση Βιβλιογραφία.....	131
7.2	Ιστότοποι.....	132

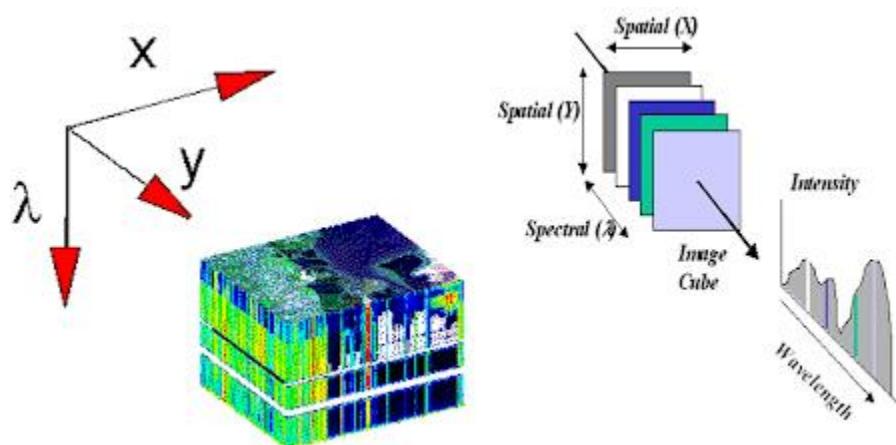


1 ΕΙΣΑΓΩΓΗ

1.1 ΥΠΕΡΦΑΣΜΑΤΙΚΑ ΔΕΔΟΜΕΝΑ ΚΑΙ ΤΑΞΙΝΟΜΗΣΗ

Η συνεχής εξέλιξη της τεχνολογίας έχει επιτρέψει στο ανθρώπινο είδος να καταφέρει μεγάλα άλματα σε όλους τους επιστημονικούς τομείς. Στον τομέα της Τηλεπισκόπησης, η ύπαρξη υπερφασματικών δεκτών έδωσε την δυνατότητα της παρατήρησης και της επεξεργασίας δεδομένων που μέχρι την κατασκευή αυτών ήταν αδύνατον να παρατηρηθούν. Οι υπερφασματικοί δέκτες συλλέγουν πληροφορίες από μέρη του ηλεκτρομαγνητικού φάσματος, οι οποίες αντιστοιχίζονται με τμήματα της Γήινης επιφάνειας.

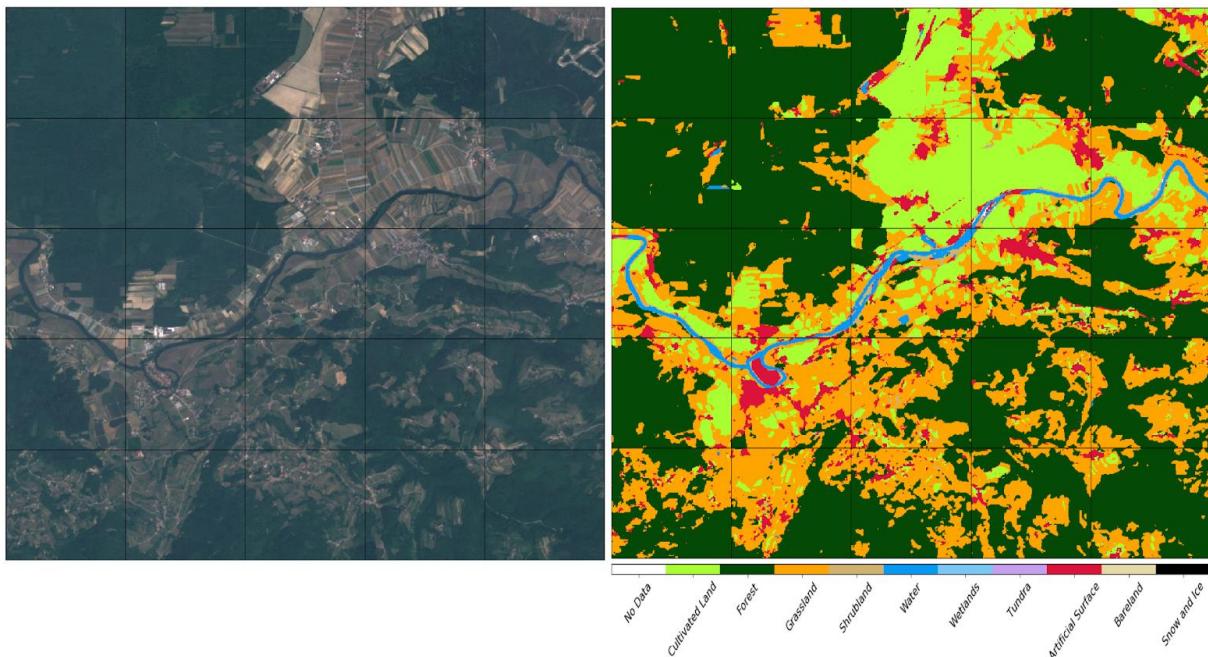
Οι πληροφορίες αυτές αποτελούν τις υπερφασματικές απεικονίσεις ή εικόνες. Αποτελούνται από διαδοχικές εικόνες της ίδιας περιοχής αναφοράς τοποθετημένες η μία πάνω στην άλλη ώστε να προκύψει ένα τρισδιάστατο διάνυσμα του οποίου η τρίτη διάσταση αναφέρεται στο ηλεκτρομαγνητικό φάσμα. Το διάνυσμα αυτό ή αλλιώς ο λεγόμενος φασματικός κύβος, περιέχει το σύνολο των τιμών της κάθε φασματικής κατηγορίας που έχει συλλέξει ο δέκτης, για κάθε εικονοστοιχείο της φυσικής επιφάνειας στην οποία αναφέρεται η εικόνα. Οι υπερφασματικές εικόνες βρίσκουν πολλές εφαρμογές σε πεδία όπως η γεωργία, η κάλυψη γης, η ορυκτολογία, η παρακολούθηση, η επιτήρηση του περιβάλλοντος, η αστρονομία κλπ. Χαρακτηριστική διαδικασία για την εξαγωγή των πληροφοριών που διαθέτουν και την χρήση τους αποτελεί η διαδικασία της ταξινόμησης.



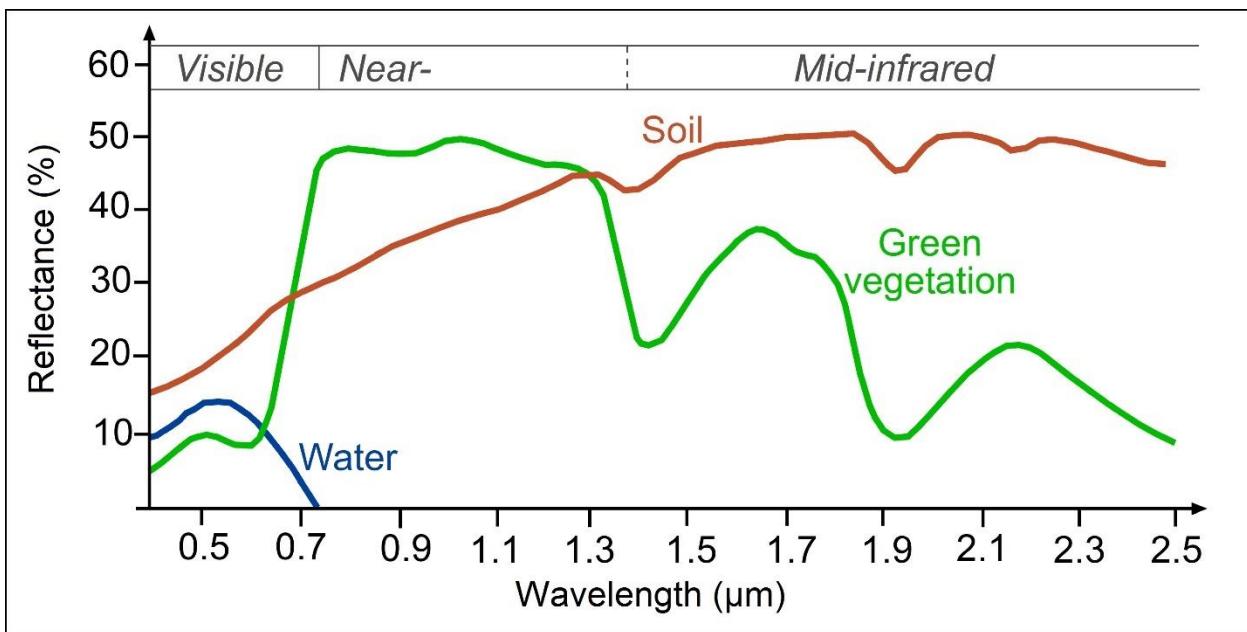
Εικόνα 1: Υπερφασματική απεικόνιση

<http://vision.gel.ulaval.ca/~jflalonde/cours/4105/h15/tps/results/projet/SASOJ/index.html>

Η ταξινόμηση υπερφασματικών εικόνων είναι η διαδικασία κατά την οποία όλα τα εικονοστοιχεία της εικόνας ταξινομούνται σε κατηγορίες. Οι κατηγορίες αυτές συνήθως αναφέρονται σε θεματικές κλάσεις ή αντικείμενα. Οι προβλέψεις των κατηγοριών βασίζονται σε ένα υπολογιστικό σύστημα που επεξεργάζεται τα εικονοστοιχεία και με διάφορες τεχνικές καταφέρνει τα διαχωρίσει και να τα ταξινομήσει στις προβλεπόμενες κατηγορίες. Ουσιαστικά, στην ταξινόμηση γίνεται μία προσπάθεια της κάθε μεθόδου να διαχωρίσει το κάθε εικονοστοιχείο με βάση την φασματική υπογραφή του. Η φασματική υπογραφή είναι «το αποτύπωμα» που αφήνει το κάθε εικονοστοιχείο στο ηλεκτρομαγνητικό φάσμα.



Εικόνα 2 :Πρόβλεψη ταξινόμησης
<https://medium.com/devseed/tracking-a-rapidly-changing-planet-bc02efe3545d>



Εικόνα 3: Φασματική υπογραφή

<http://vision.gel.ulaval.ca/~jflalonde/cours/4105/h15/tps/results/projet/SASQJ/index.html>

Υπάρχουν δυο βασικές κατηγορίες που απαρτίζουν την διαδικασία της ταξινόμησης. Η πρώτη κατηγορία ονομάζεται επιβλεπόμενη ταξινόμηση και η δεύτερη μη επιβλεπόμενη. Η διαφορά τους είναι ότι στην πρώτη τα δεδομένα που εισάγονται στον εκάστοτε αλγόριθμο περιλαμβάνουν την εικόνα αλλά και ένα υπόδειγμα των κατηγοριών από τον χρήστη. Το υπόδειγμα αυτό είναι ένα σύνολο από περιοχές εκπαίδευσης, δηλαδή ένα σύνολο από εικονοστοιχεία χαρακτηρισμένα ως προς την κατηγορία που ανήκουν. Έτσι δίνεται στον αλγόριθμο ο τελικός αριθμός των κατηγοριών που θα τελικά θα ταξινομηθούν τα εικονοστοιχεία, και τα φασματικά χαρακτηριστικά των χαρακτηρισμένων εικονοστοιχείων προκειμένου να βασιστεί ο διαχωρισμός των υπολοίπων. Η δεύτερη κατηγορία ονομάζεται μη επιβλεπόμενη ταξινόμηση. Σε αυτήν την κατηγορία ο χρήστης επιλέγει έναν αυθαίρετο τελικό αριθμό κατηγοριών στις οποίες θα ταξινομηθούν τα δεδομένα και ο αλγόριθμος διαχωρίζει τα εικονοστοιχεία με γνώμονα τον αριθμό κατηγοριών και όχι την φασματική υπογραφή προ-χαρακτηρισμένων εικονοστοιχείων.

Η βασική ομάδα των αλγορίθμων που θα εφαρμοσθούν και θα μελετηθούν στην παρούσα εργασία, ανήκει στο πεδίο της Μηχανικής Μάθησης και πιο συγκεκριμένα στην Βαθιά Μάθηση.

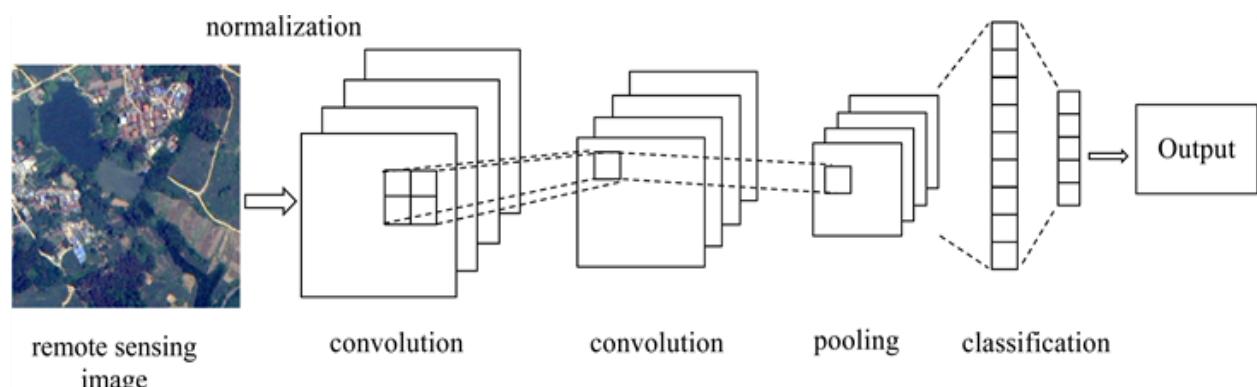
1.2 ΑΦΟΡΜΗ ΣΥΓΓΡΑΦΗΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Βασικό κίνητρο για την συγγραφή της εργασίας αποτελεί η συνεχής αναζήτηση του Τοπογράφου Μηχανικού για καλύτερες μεθόδους προσδιορισμού και ανάλυσης του χώρου της Γήινης επιφάνειας. Η σύγχρονη τεχνολογική προοπτική της επιστήμης της Τηλεπισκόπησης απαιτεί την συμπόρευση του Τοπογράφου Μηχανικού με όλες τις νέες τεχνικές, και την διαρκή παρακολούθηση και αξιολόγηση τους.

Τα τελευταία χρόνια, έχει σημειωθεί μεγάλη ανάπτυξη πάνω στον τομέα της Βαθιάς Μάθησης. Παρουσιάζονται συνεχώς νέες μέθοδοι με δυνατότητες που πριν από μερικά χρόνια φάνταζαν αδιανόητες. Οι αρχιτεκτονικές των νευρωνικών δικτύων γίνονται όλο και πολυπλοκότερες και η αποτελεσματικότητα αυξάνει. Η Βαθιά Μάθηση εφαρμόζεται στην υγεία, στην παιδεία, στην ψυχαγωγία και εισχωρεί με την πάροδο του χρόνου σε όλους τους διαθέσιμους τομείς.

Πιο συγκεκριμένα, η ανάπτυξη της Βαθιάς Μάθησης χαρακτηρίζει γενικότερα την εξέλιξη της Όρασης υπολογιστών. Η Όραση υπολογιστών χρησιμοποιεί μεθόδους Βαθιάς Μάθησης, οι οποίες όμως αναφέρονται συνήθως σε απλές εικόνες. Έτσι το ενδιαφέρον της επιστημονικής κοινότητας και όλες οι καινοτομίες γύρω από το αντικείμενο αυτό σχετίζονται κυρίως με εφαρμογές των μεθόδων αυτών σε εικόνες στο φάσμα RGB. Οι εικόνες αυτές διαθέτουν 3 μόνο φασματικά κανάλια στα οποία στηρίζονται οι μέθοδοι ώστε να τις ταξινομήσουν. Χαρακτηριστικά παραδείγματα αποτελούν ο διαγωνισμός ImageNet και τα δίκτυα Alexnet, Resnet.

Οι υπερφασματικές εικόνες διαθέτουν μεγαλύτερο εύρος δεδομένων αφού διαθέτουν πολλά περισσότερα φασματικά κανάλια από τις εικόνες RGB. Συνεπώς απαιτείται η ανάπτυξη διαφορετικών μεθόδων οι οποίες θα είναι προσαρμοσμένες στις εικόνες αυτές, και θα κατορθώνουν αποτελεσματικά να εκμεταλλευτούν αυτήν την επιπρόσθετη πληροφορία.



Εικόνα 4: Εφαρμογή δικτύου Βαθιάς Μάθησης σε εικόνα
https://www.scirp.org/html/1-2801736_89791.htm

Η παρούσα εργασία θα διερευνήσει και θα παρουσιάσει διάφορες μεθόδους Βαθιάς Μάθησης που εφαρμόζονται πάνω στις υπερφασματικές εικόνες. Οι μέθοδοι αυτοί ξεχώρισαν στην επιστημονική κοινότητα και μπορούν να χαρακτηριστούν ως σύγχρονες μέθοδοι αιχμής για την χρονική στιγμή στην οποία ολοκληρώθηκε η συγκεκριμένη εργασία. Έπειτα θα εφαρμοσθούν πραγματοποιώντας ταξινόμηση σε υπερφασματικά δεδομένα. Τελικά θα αξιολογηθούν τα αποτελέσματα της ταξινόμησης, ποσοτικά και ποιοτικά. Μέσα από την διαδικασία αύτη θα προκύψουν συμπεράσματα σχετικά με τις λειτουργίες και τις προοπτικές των μεθόδων αυτών αλλά και της χρήσης των μεθόδων Βαθιάς Μάθησης γενικότερα πάνω στην ταξινόμηση υπερφασματικής εικόνας.

Ακόμα, ένας άλλος σκοπός που θα εξυπηρετηθεί από την συγκεκριμένη εργασία είναι η εφαρμογή μεθόδων Βαθιάς Μάθησης πάνω σε δεδομένα σχετικά δυσκολότερα από αυτά που χρησιμοποιούνται συνήθως στην βιβλιογραφία. Το HyRANK dataset που θα ταξινομηθεί, σε αντίθεση με τα γνωστά dataset που χρησιμοποιούνται συνήθως στις εφαρμογές αυτές (Indian Pines, PaviaU), αποτελεί μεγαλύτερη πρόκληση για τις μεθόδους και προσδίδει στην αξιολόγηση τους ένα ρεαλιστικότερο κριτήριο της αποτελεσματικότητας τους.

2 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

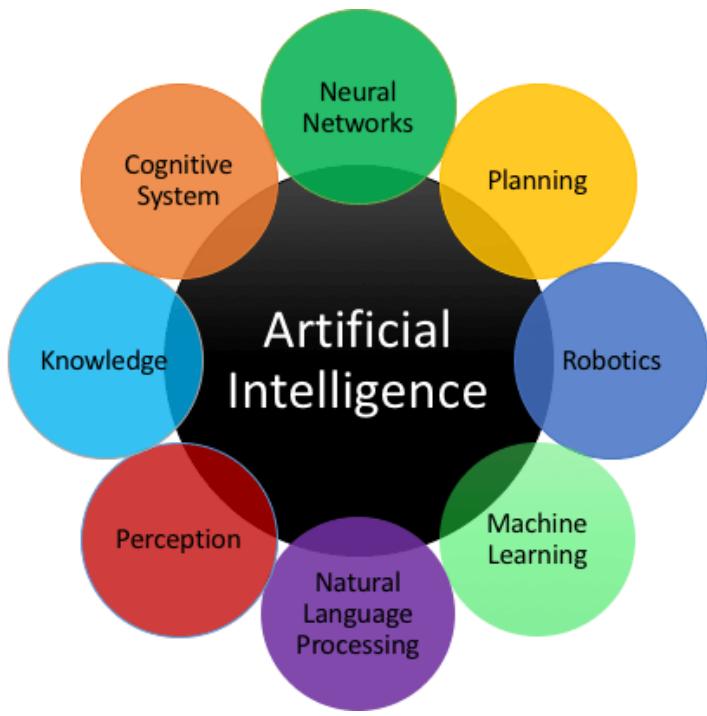
Ο όρος τεχνητή νοημοσύνη αναφέρεται στο κατόρθωμα του ανθρώπινου είδους να θέσει εφικτό οι μηχανές να προσεγγίσουν την ανθρώπινη σκέψη ή αλλιώς την «φυσική» νοημοσύνη των ανθρώπων. Αποτέλεσμα της προσέγγισης αυτής είναι οι μηχανές να αντλούν πληροφορίες από το περιβάλλον στο οποίο τοποθετούνται και να πραγματοποιούν ενέργειες προκειμένου να επιτύχουν τους σκοπούς τους.

Οι πρώτες επαφές με την τεχνητή νοημοσύνη ξεκίνησαν το καλοκαίρι του 1956 στο κολλέγιο του Dartmouth. Η επιστημονική κοινότητα ασχολούταν με την πιθανότητα να παραχθεί ένας τεχνητός ανθρώπινος εγκέφαλος. Με αφετηρία το συνέδριο στο κολλέγιο του Dartmouth, ξεκίνησε ένα μεγάλο ρεύμα ανάπτυξης της τεχνητής νοημοσύνης όπου μηχανές αποδείκνυαν γεωμετρικά θεωρήματα, μαθαίναν να μιλάνε την Αγγλική γλώσσα και διάφορες άλλες διαδικασίες. Παρόλα αυτά το 1974 ο ρυθμός ανάπτυξης μειώθηκε ακαριαία, με την περίοδο αυτή να χαρακτηρίζεται ως «ο χειμώνας της τεχνητής νοημοσύνης» μέχρι το 1980. Ένας βασικός λόγος για την στασιμότητα αυτή υπήρξε η φτωχή υπολογιστικά δυναμική της τεχνολογίας εκείνη την εποχή που δεν επέτρεπε την πλήρη αξιοποίηση της τεχνητής νοημοσύνης.

Από εκείνη την περίοδο και μετά, η εξέλιξη ήταν συνεχής. Η αύξηση της υπολογιστικής ισχύος ώθησε τις δυνατότητες και τις εφαρμογές με αποτέλεσμα η τεχνητή νοημοσύνη να ενισχύει την θέση της στο τεχνολογικό γίγνεσθαι. Από το 2011 η τρομακτική εξέλιξη των υπολογιστών σε συνδυασμό με την ύπαρξη μεγάλου όγκου δεδομένων – τα λεγόμενα Big Data – καθιέρωσαν μία νέα εποχή. Η τεχνητή νοημοσύνη εισχωρεί παντού και αποτελεί σημαντικό παράγοντα εξέλιξης σε όλη την επιστημονική κοινότητα αλλά και σαν τεχνολογικό εμπορικό προϊόν.

Ο όρος τεχνητή νοημοσύνη περιλαμβάνει μέσα του όλες τις εφαρμογές των μηχανών με τρόπο που βασίζεται στη στην ανθρώπινη σκέψη. Το μεγάλο πλήθος των εφαρμογών αυτών αλλά και των τεχνικών, προσεγγίσεων και σκοπών εξυπηρετούνται καθιστούν απαραίτητο τον κατακερματισμό του όρου σε υποενότητες. Οπότε κάτω από την σκέπη του όρου τεχνητή νοημοσύνη μπορεί κανείς να συναντήσει την επεξεργασία φυσικής γλώσσας, την ρομποτική, την όραση υπολογιστή, την βασιζόμενη συλλογιστική, μηχανική μάθηση και άλλα.

Στην συγκεκριμένη διπλωματική εργασία θα γίνει αναφορά σε σχετικές τεχνικές στην ταξινόμηση υπερφασματικών εικόνων. Θα γίνει αναφορά πάνω στην μηχανική μάθηση και στην βαθιά μάθηση που αποτελούν υποενότητες της τεχνητής νοημοσύνης.



Εικόνα 5:Τα επιμέρους αντικείμενα της Τεχνητής Νοημοσύνης
<https://www.scoro.com/blog/artificial-intelligence-predictions/>

2.2 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

Η μηχανική μάθηση είναι και αυτή μία μεγάλη υποκατηγορία της τεχνητής νοημοσύνης. Είναι η επιστήμη εκείνη που μελετά της τεχνικές που επιτρέπουν στις μηχανές να μαθαίνουν και να βελτιώνουν την διαδικασία πραγματοποίησης των στόχων τους μέσω δεδομένων που τους παρέχονται, δίχως την παρέμβαση του ανθρώπινου παράγοντα.

Ο όρος μηχανική μάθηση χρησιμοποιήθηκε αρχικά από τον Άρθουρ Σάμιουελ το 1959. Ο τομέας αυτός ξεκίνησε ως ειδικός τομέας της τεχνητής νοημοσύνης με σκοπό την αναγνώριση προτύπων. Με την πάροδο του χρόνου, η ανάπτυξη των Έμπειρων συστημάτων και άλλων πεδίων που είχαν να κάνουν με θεωρητικές προσεγγίσεις δεδομένων και αναπαραστάσεων αυτών χαρακτήρισαν την Μηχανική μάθηση σαν έναν ξεχωριστό τομέα της Τεχνητής Νοημοσύνης.

Ο τομέας αυτός ασχολείται με την επίλυση πρακτικών προβλημάτων μέσω της συνεχούς βελτιστοποίησης των τεχνικών επίλυσης βασιζόμενος σε δεδομένα. Η βελτιστοποίηση αυτή είναι το χαρακτηριστικό που διαχωρίζει την υποκατηγορία αυτήν από τις άλλες σχετικά με την τεχνητή νοημοσύνη. Ουσιαστικά ο τομέας αυτός απαντά στο ερώτημα «Πως να χτίσουμε υπολογιστικά συστήματα τα οποία βελτιώνονται αυτόμata με την εμπειρία?».

Δυο βασικές κατηγορίες προσεγγισης των τεχνικών μηχανικής μάθησης είναι η επιβλεπόμενη μάθηση και η μη επιβλεπόμενη μάθηση.

- Η επιβλεπόμενη μάθηση είναι μία προσέγγιση όπου η μηχανή δέχεται δεδομένα στα οποία είναι γνωστό το τελικό αποτέλεσμα της πρόβλεψης. Μέσω της διαδικασίας της μάθησης, η μηχανή μαθαίνει το τρόπο ώστε παίρνοντας σαν εισαγωγή το δεδομένο A να παράγει την επιθυμητή πρόβλεψη Α.
- Η μη επιβλεπόμενη μάθηση, είναι η προσέγγιση στην οποία δεν υπάρχει επιθυμητό αποτέλεσμα στα δεδομένα. Η μηχανή με διάφορες τεχνικές εξερευνά τα δεδομένα, αντιλαμβάνεται κοινά στοιχεία και προβλέπει το αποτέλεσμα.

Οι δυο αυτές προσεγγίσεις είναι γενικές και μπορούν να συνδυαστούν ενισχύοντας η μία την άλλη.

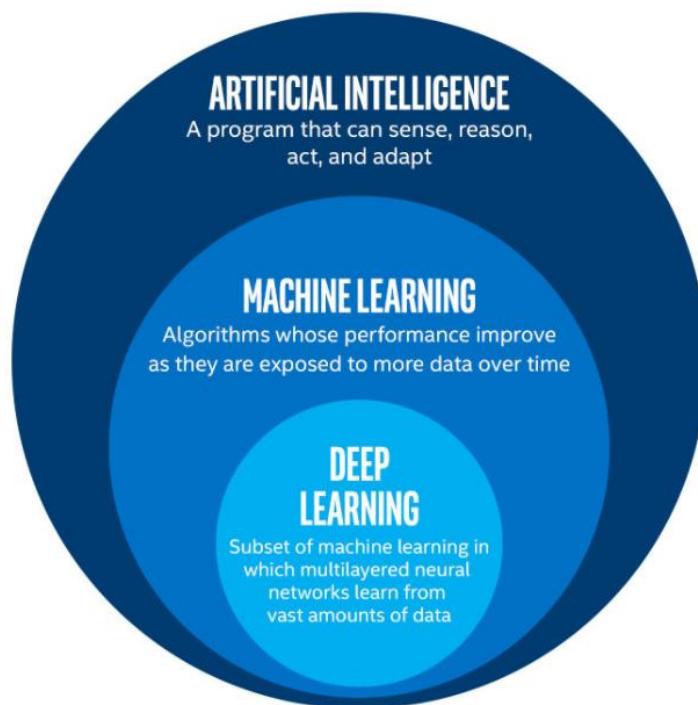
Στην φασματική ταξινόμηση οι τεχνικές επιβλεπόμενης μάθησης δέχονται ως δεδομένα εικονοστοιχείο που έχουν χαρακτηρισθεί ως προς την κατηγορία στην οποία ανήκουν από τον χρήστη. Έπειτα η μηχανή μαθαίνει να αντιλαμβάνεται αυτή τη σχέση μεταξύ εικονοστοιχείου και κατηγορίας και είναι ικανή να δεχθεί νέα αχαρακτήριστα εικονοστοιχεία και να τα ταξινομήσει στην κατηγορία που θα πρέπει να ανήκουν. Στην μη επιβλεπόμενη φασματική ταξινόμηση η μηχανή δέχεται αχαρακτήριστα εικονοστοιχεία και μαθαίνοντας τις σχέσεις μεταξύ τους δημιουργεί νέες κατηγορίες στις οποίες τα ταξινομεί αυτή. Ετσι ο χρήστης δεν παρέχει επίβλεψη στην δημιουργία των κατηγοριών, δηλαδή δεν καθορίζει αυτός τις σχέσεις εικονοστοιχείου-κατηγορίας αφού τα δεδομένα που δέχεται η μηχανή είναι αχαρακτήριστα.

2.3 ΒΑΘΙΑ ΜΑΘΗΣΗ

Η βαθιά μάθηση ανήκει στην οικογένεια της μηχανικής μάθησης. Αναφέρεται σε αλγόριθμους με περισσότερο «βάθος» ή αλλιώς πολυπλοκότερους σε σχέση με τους συμβατικούς πάνω στην μηχανική μάθηση. Αυτοί επιτρέπουν την μηχανή να αποκωδικοποιήσει και κατά συνέπεια να μάθει πιο σύνθετα χαρακτηριστικά και συσχετισμούς των δεδομένων που δέχεται ως είσοδο.

Το χαρακτηριστικό αυτό της μεγάλης και πολύπλοκης αρχιτεκτονικής έφερε την «επανάσταση της Βαθιάς Μάθησης» με αφετηρία το 2012. Από εκείνη την χρονιά και έπειτα, πολλές διαδικασίες-προκλήσεις σχετικά με την αναγνώριση προτύπων σε δεδομένα σημείωσαν ρεκόρ με την χρήση τέτοιων αρχιτεκτονικών. Η μέθοδος του Krizhevsky et al κέρδισε τον διαγωνισμό ImageNet με μεγάλη διαφορά στην ακρίβεια σε σχέση με τις ανταγωνίστριες μεθόδους μηχανικής μάθησης χρησιμοποιώντας βαθιά μάθηση.

Βασικό στοιχείο της βαθιάς μάθησης είναι τα νευρωνικά δίκτυα στα οποία θα γίνει αναφορά στη συνέχεια. Τα νευρωνικά δίκτυα χρησιμοποιούνται ευρέως στην μηχανική μάθηση στην απλή τους μορφή. Στην βαθιά μάθηση χρησιμοποιούνται με πολλά περισσότερα στρώματα και διαφορετικές προσεγγίσεις. Μερικές από αυτές τις τεχνικές θα αναλυθούν στη συνέχεια.



Εικόνα 6 :Ο ρόλος της Βαθιάς Μάθησης

<https://www.prowesscorp.com/whats-the-difference-between-artificial-intelligence-ai-machine-learning-and-deep-learning>

2.4 ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Στην προσπάθεια τη μοντελοποίησης του ανθρώπινου εγκεφάλου, δημιουργήθηκε ένα μαθηματικό σύστημα εμπνευσμένο από το νευρικό σύστημα του ανθρώπου. Το σύστημα αυτό έχει νευρώνες συνδεδεμένους μεταξύ τους σε αλληλουχία με τον ίδιο τρόπο που συνδέονται οι βιολογικοί νευρώνες του ανθρώπου. Κάθε σύνδεση μεταφέρει ένα σήμα στον επόμενο νευρώνα ή νευρώνες. Έπειτα γίνεται η επεξεργασία του σήματος και μετέπειτα μεταφορά του στους επόμενους. Το μοντέλο που περιγράφει τις διεργασίες αυτές ονομάστηκε τεχνητό νευρωνικό δίκτυο.

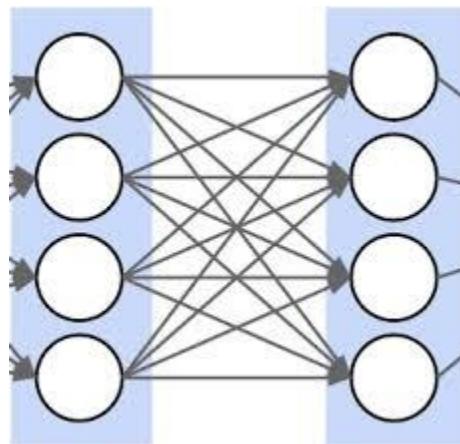
Το βασικό δομικό στοιχείο των τεχνητών νευρωνικών δικτύων είναι ο νευρώνας. Οι νευρώνες δέχονται μία ή πολλαπλές αριθμητικές εισόδους και παράγουν μία έξοδο. Η μαθηματική διαδικασία που μετατρέπει την είσοδο σε έξοδο ονομάζεται συνάρτηση ενεργοποίησης. Ακόμα, ο κάθε νευρώνας χαρακτηρίζεται από τα βάρη που ισοσταθμίζουν τις εισόδους και την τιμή κατωφλιού που προστίθεται στην είσοδο. Η βασική ιδέα του δικτύου αυτού είναι ότι τα βάρη αυτά μαζί με τις τιμές κατωφλιού είναι παράμετροι προς εκπαίδευση. Δηλαδή το δίκτυο καθώς εκπαιδεύεται μεταβάλει τις παραμέτρους αυτές με τέτοιον τρόπο ώστε να προσαρμοστούν στα δεδομένα που δίνονται ως είσοδος στο δίκτυο. Έτσι ο κάθε νευρώνας μαθαίνει να αναγνωρίζει τα δεδομένα, όπως θα αντιλαμβανόταν σιγά-σιγά τα ερεθίσματα της ζωής ένα μικρό παιδί.

2.4.1 Νευρώνας

Ο νευρώνας πήρε την ονομασία του από τον βιολογικό νευρώνα του ανθρώπινου νευρολογικού συστήματος. Λειτουργεί με τον παρόμοιο τρόπο μεταφέροντας πληροφορία από τον έναν στον άλλον.

2.4.2 Πλήρως συνδεδεμένο στρώμα

Πολλοί νευρώνες μαζί οργανώνονται σε στρώματα. Τα στρώματα αυτά πλήρως συνδεδεμένα στρώματα αφού συνδέουν τον κάθε νευρώνα με όλους τους νευρώνες του επόμενου στρώματος.



Εικόνα 7:Πλήρως συνδεδεμένο στρώμα

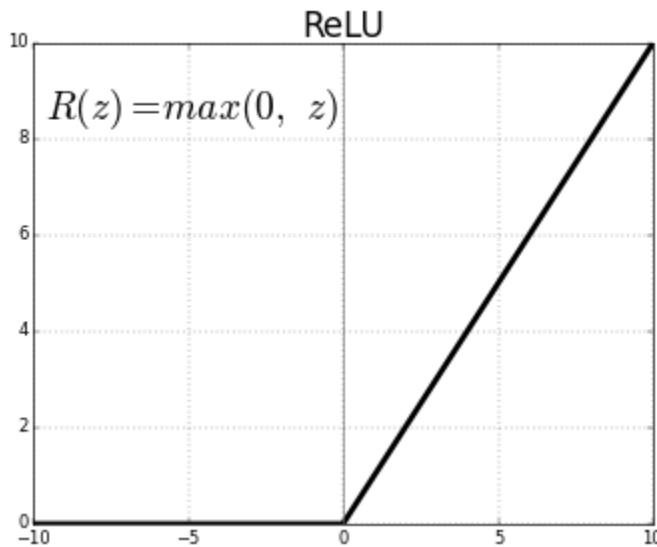
<https://www.semanticscholar.org/paper/Multi-Class-Object-Detection-from-Aerial-Images-Schweitzer-Agrawal/16d8304a1d8a5cb2288a5af90a09fd09d8b7ae09>

2.4.3 Συνάρτηση ενεργοποίησης

Η συνάρτηση ενεργοποίησης μεσολαβεί μεταξύ της εισόδου και της εξόδου του νευρώνα. Υπολογίζει με βάση την είσοδο, τα βάρη, και την τιμή κατωφλιού την τιμή της εξόδου. Όταν αναφέρεται ότι ένας νευρώνας ενεργοποιείται εννοείται ότι η έξοδος του έχει πάρει υψηλή τιμή σε αντίθεση με την απενεργοποίηση του που η έξοδος έχει χαμηλή τιμή.

Σαν μαθηματική συνάρτηση που μπορεί να πάρει διάφορες μορφές, κάθε τύπος συνάρτησης ενεργοποίησης παρουσιάζει διαφορετικές ιδιότητες επηρεάζοντας πολλές φορές την αποτελεσματικότητα του δικτύου. Ακολουθούν μερικές από τις πιο διαδεδομένες συναρτήσεις ενεργοποίησης στα νευρωνικά δίκτυα.

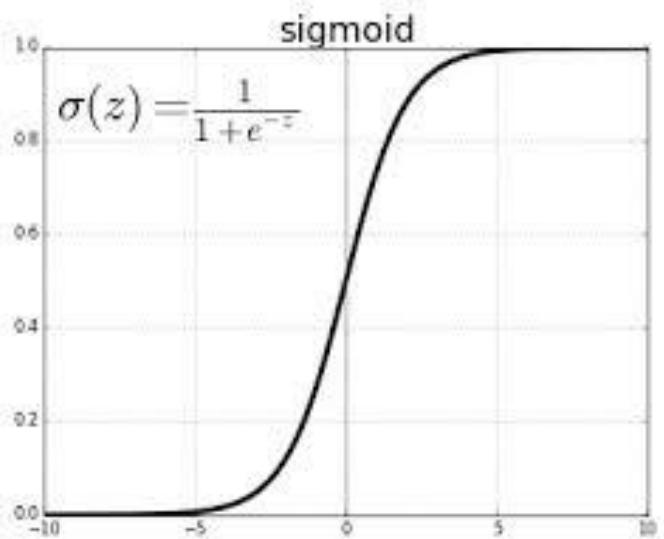
2.4.3.1 Rectified Linear Unit (ReLU)



Εικόνα 8:Η συνάρτηση ενεργοποίησης ReLU

Η ReLU είναι ίσως η πιο διαδεδομένη συνάρτηση ενεργοποίησης αυτή την στιγμή στα Τεχνητά Νευρωνικά δίκτυα. Εμφανίστηκε πρώτη φορά το 2000 από τον Hahnloser et all. Επιτρέπει ταχύτερη και ακριβέστερη μάθηση πάνω σε μεγάλο αριθμό δεδομένων σε σχέση με την σιγμοειδή συνάρτηση ή την συνάρτηση υπερβολικής εφαπτομένης. Αυτό προκύπτει από το γεγονός ότι η γραμμική μετάβαση στο θετικό μέρος του άξονα χ δεν παραμορφώνει την παράγωγο την συνάρτησης σε αντίθεση με τις προηγούμενες, με αποτέλεσμα η διαδικασία της βελτιστοποίησης μέσω της οπισθοδιάδοσης να πραγματοποιείται απροβλημάτιστα χωρίς την εμφάνιση του προβλήματος της εξαφάνισης της παραγώγου. Ακόμα συγκριτικά με τις προηγούμενες προσφέρει λιγότερο υπολογιστικό κόστος αφού υπολογίζεται εύκολα σαν γραμμική στις θετικές τιμές και μηδενική στις αρνητικές.

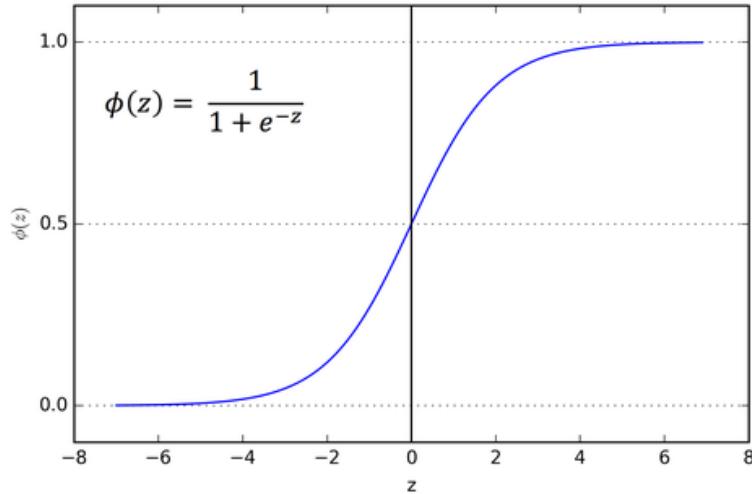
2.4.3.2 Σιγμοειδής συνάρτηση



Εικόνα 9:H σιγμοειδής συνάρτηση ενεργοποίησης

Η σιγμοειδής συνάρτηση μετατρέπει τους μεγάλους αριθμούς σε 1 και τους μικρούς σε 0. Η καμπύλη της θυμίζει το γράμμα S γι' αυτό και έχει την ονομασία αυτή. Χρησιμοποιείται αρκετά σε προβλήματα ταξινόμησης με 2 μόνο πιθανές κλάσεις διότι επιτρέπει ταχύτερη εκπαίδευση του δικτύου. Παρόλα αυτά το γεγονός ότι πολύ νωρίς στα άκρα της συνάρτησης η παραγώγος τείνει στο 0 προκαλεί το πρόβλημα της εξαφάνισης της παραγώγου με αποτέλεσμα ένα μεγάλο δίκτυο να μην μπορεί να εκπαιδευτεί σωστά.

2.4.3.3 Συνάρτηση υπερβολικής εφαπτομένης



Εικόνα 10:Η συνάρτηση ενεργοποίησης υπερβολικής εφαπτομένης

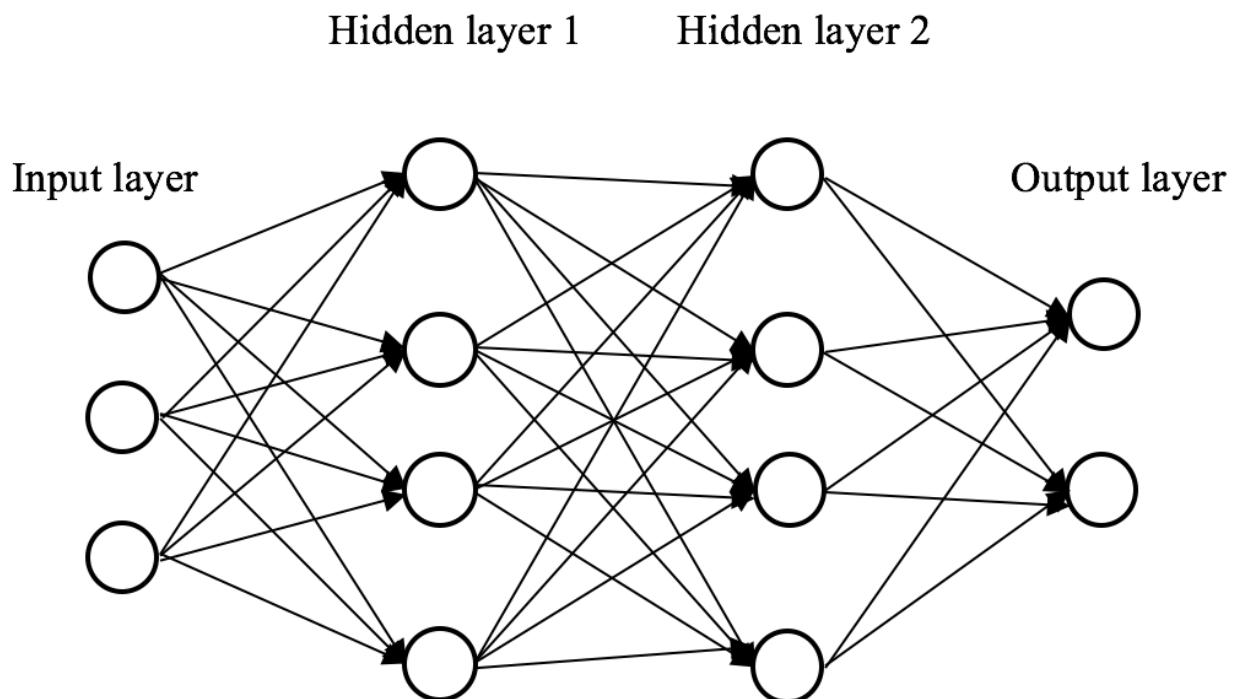
Η συνάρτηση υπερβολικής εφαπτομένης μοιάζει αρκετά με την σιγμοειδή. Μία βασική διαφορά είναι ότι η σιγμοειδής κινείται μεταξύ του χώρου $[0,1]$ ενώ η συνάρτηση υπερβολικής εφαπτομένης μεταξύ του $[-1,1]$. Το γεγονός ότι το σύνολο τιμών της είναι κεντροβαρικό προς το 0 διευκολύνει την εκπαίδευση γιατί αποτρέπει την ανανέωση των παραμέτρων σε μία μόνο κατεύθυνση κατά τη διάρκεια της οπισθοδιάδοσης.

2.4.4 Βάρη και τιμή κατωφλιού.

Τα βάρη και η τιμή του κατωφλιού σε κάθε νευρώνα είναι από τα πιο σημαντικά στοιχεία ενός νευρωνικού δικτύου. Είναι οι παράμετροι προς μάθηση, δηλαδή οι παράμετροι που οι τιμές τους θα καθορίσουν την ακρίβεια των προβλέψεων του δικτύου και γι' αυτόν τον λόγο εκπαιδεύονται. Στην αρχή του δικτύου αρχικοποιούνται προκειμένου να ξεκινήσει η διαδικασία της εκπαίδευσης. Η αρχικοποίηση αυτή πραγματοποιείται με διάφορες μεθόδους, η πιο διαδεδομένη από αυτές είναι η μέθοδος Kaiming. Οι παράμετροι προστίθενται στην κάθε παρατήρηση Y με τον παρακάτω τρόπο:

$$Y = \sum (\text{weight} * \text{input}) + \text{bias}$$

Ένα παράδειγμα τεχνητού νευρωνικού δικτύου σε ταξινόμηση εικονοστοιχείο είναι το εξής:



Εικόνα 11: Αρχιτεκτονική απλού νευρωνικού δικτύου

<https://medium.com/coinmonks/implement-back-propagation-in-neural-networks-ed09897593e7>

Στην παραπάνω εικόνα αναπαρίσταται μια τυπική μορφή ενός τεχνητού νευρωνικού δικτύου. Διακρίνονται τα επίπεδα εισόδου-εξόδου και 2 κρυφά επίπεδα. Το επίπεδο εισόδου αποτελείται από τις 3 αριθμητικές τιμές του φάσματος RGB του εικονοστοιχείου μιας εικόνας. Το επίπεδο

εξόδου αποτελείται από 2 νευρώνες που θα περιέχουν μία τιμή έκαστος, οι οποίες μπορεί να συμβολίζουν την πιθανότητα το εικονοστοιχείο να ανήκει στην κατηγορία A ή B.

Τα κρυφά στρώματα περιέχουν 4 νευρώνες το καθένα και είναι πλήρως συνδεδεμένα στρώματα. Συνολικά υπάρχουν 10 νευρώνες με $3 \times 4 + 4 \times 4 + 4 \times 2 = 52$ βάρη και 10 τιμές κατωφλιού. Κατά την διάρκεια της εκπαίδευσης του δικτύου αυτοί οι παράμετροι θα μεταβάλλονται με τέτοιον τρόπο ώστε η πρόβλεψη για το αν ανήκει το εικονοστοιχείο στην κατηγορία A ή B να έχει την μέγιστη δυνατή ακρίβεια.

2.4.5 Συνάρτηση κόστους

Η συνάρτηση κόστους είναι μία μαθηματική διαδικασία η οποία αντιστοιχίζει ένα γεγονός σε έναν πραγματικό αριθμό, συμβολίζοντας με τον αριθμό το κόστος του γεγονότος αυτού. Στην περίπτωση της ταξινόμησης υπερφασματικών εικόνων το κόστος αυτό αναπαριστά την αποχή της πρόβλεψης του δικτύου για την κατηγορία που αντιστοιχίζει στο εικονοστοιχείο, σε σχέση με την αληθινή κατηγορία στην οποία ανήκει. Αποτελεί έναν δείκτη που φανερώνει την πορεία του δικτύου. Καθώς ένα δίκτυο μαθαίνει τα δεδομένα του και βελτιώνει την αντίληψη του γι' αυτά, θα παρουσιάζει λιγότερο κόστος αφού οι σωστές προβλέψεις αυξάνονται σε αριθμό. Από μαθηματικής σκοπιάς, στην συνάρτηση αυτή βασίζεται αυτή η ικανότητα της μάθησης του δικτύου, μέσω της διαδικασίας της οπισθοδιάδοσης. Η χαρακτηριστικότερη συνάρτηση κόστους που χρησιμοποιείται στα τεχνητά νευρωνικά δίκτυα και ιδιαίτερα στις εφαρμογές της ταξινόμησης είναι ή cross-entropy loss ή λογαριθμικό κόστος.

Το λογαριθμικό κόστος έχει την μορφή :

$$Li = -\log \left(\frac{e^{fy_i}}{\sum_j e^{fj}} \right) = -fy_i + \log \sum_j e^{fj} \quad (3)$$

Οπου f, j συμβολίζονται τα στοιχεία του διανύσματος που περιέχει τις πιθανότητες για κάθε κλάση. Οι πιθανότητες αυτές εκφράζονται με τιμή μεταξύ του 0 και του 1, όπου 0 η μηδενική πιθανότητα αληθινής πρόβλεψης και 1 η μέγιστη πιθανότητα αληθινής πρόβλεψης. Κατά συνέπεια όταν η πρόβλεψη της κλάσης απέχει αρκετά από την αληθινή πρόβλεψη, το κόστος τείνει στο άπειρο αφού $\log 0$ τείνει στο άπειρο. Στην περίπτωση που η πιθανότητα είναι 1 τότε η συνάρτηση κόστους μηδενίζεται αφού $\log 1=0$.

2.4.6 Οπισθοδιάδοση

Τα τεχνητά νευρωνικά δίκτυα όπως αναφέρθηκε μαθαίνουν από τα δεδομένα που τους δίνονται ως είσοδος. Η διαδικασία της μάθησης οφείλεται στον αλγόριθμο οπισθοδιάδοσης. Ο αλγόριθμος αυτός είναι ο υπεύθυνος ώστε το δίκτυο να μαθαίνει συνεχώς αναζητώντας την βέλτιστη λύση στα προβλήματα που του ανατίθενται.

Αρχικά εξετάζεται η πορεία του σήματος ενός δεδομένου σε ένα τέτοιο δίκτυο. Το σήμα θα βρεθεί στο στρώμα εισαγωγής. Από το στρώμα εισαγωγής θα οδηγηθεί στο ένα ή περισσότερα στρώματα νευρώνων. Εκεί κάθε στοιχείο του σήματος πολλαπλασιάζεται με το αρχικοποιημένο βάρος του νευρώνα, προστίθεται σε αυτό η αρχικοποιημένη τιμή κατωφλιού και περνά από την συνάρτηση ενεργοποίησης του κάθε νευρώνα. Έπειτα σαν έξοδος ενεργοποίησης οδηγείται με την ίδια διαδικασία και στα επόμενα στρώματα έως ότου καταλήξει στο στρώμα εξόδου. Τέλος οι τιμές του στρώματος εξόδου θα χρησιμοποιηθούν για να υπολογιστεί το κόστος, μέσω της συνάρτησης κόστους. Όλη αυτή η διαδικασία ονομάζεται forward pass και αναφέρεται στην διαδικασία πρόβλεψης.

Η δυνατότητα μάθησης των τεχνιτών νευρωνικών δικτύων οφείλεται στις παραμέτρους προς μάθηση που τα απαρτίζουν. Ο κάθε νευρώνας λοιπόν είναι ένα κομμάτι «νου» του δίκτυου, εκφραζόμενο με τις τιμές του βάρους και της τιμής κατωφλιού. Οι τιμές αυτές καθορίζουν σε μεγάλο βαθμό την τελική πρόβλεψη, αφού αυτοί επηρεάζουν την έξοδο του κάθε νευρώνα. Το κλειδί σε ένα αποτελεσματικό δίκτυο είναι το σωστό σετ τιμών των παραμέτρων αυτών. Οι ιδανικές παράμετροι θα πραγματοποιούσαν ακριβείς προβλέψεις. Η προσέγγιση του σετ αυτού είναι ο σκοπός της εκπαίδευσης του δίκτυου, που ολοκληρώνεται μέσω επαναλήψεων των forward pass και backward pass που θα περιγράφει στη συνέχεια.

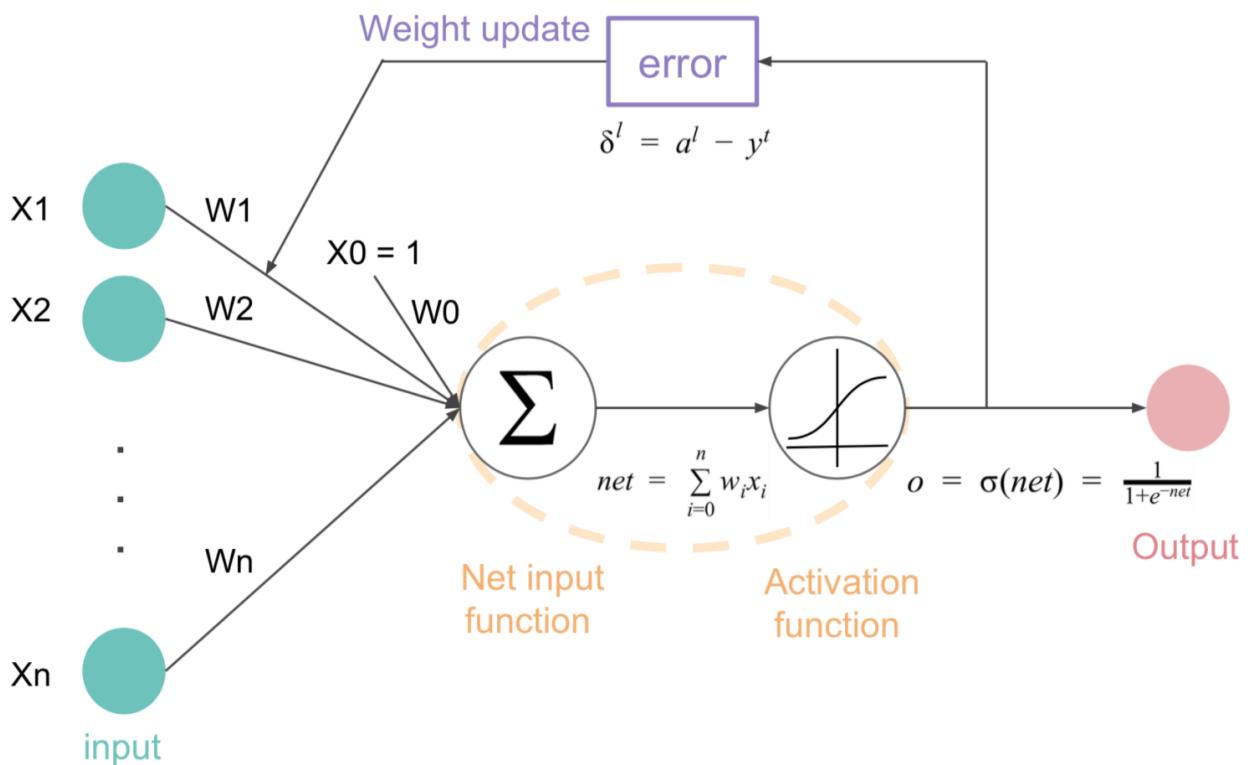
Η διαδικασία backward pass πραγματοποιείται μέσω του αλγόριθμου οπισθοδιάδοσης. Στον αλγόριθμο αυτό αφού υπολογιστεί το αποτέλεσμα της συνάρτησης κόστους, το σήμα οδηγείται στην αντίθετη πορεία από την οποία έκανε μετάβαση αρχικά, δηλαδή πίσω στο στρώμα εισόδου. Η μετάβαση αυτή γίνεται με σκοπό την τροποποίηση των βαρών και τιμών κατωφλιού του δίκτυου, ώστε με το ίδιο σήμα σε ένα νέο forward pass να πρόκυπτε λιγότερο κόστος.

Η τροποποίηση αυτή βασίζεται στην ελαχιστοποίηση της συνάρτησης του κόστους. Υπολογίζονται οι μερικές παράγωγοι της συνάρτησης του κόστους σε σχέση με την κάθε παράμετρο προς μάθηση. Οι μερικές παράγωγοι φανερώνουν την ευαισθησία του αποτελέσματος της συνάρτησης κόστους σε σχέση με τις παραμέτρους. Δείχνουν ουσιαστικά το πόσο συντέλεσε η κάθε παράμετρος στο σύνολο του κόστους του δίκτυου. Ο αλγόριθμος θα μεταβάλει τις τιμές αφαιρώντας την μερική παράγωγο από την κάθε παράμετρο στην οποία αναφέρεται ώστε η συνάρτηση κόστους να παρουσιάσει λιγότερο κόστος με το ίδιο σήμα στο forward pass.

Η μεταβολή αυτή απαιτεί και τον ορισμό ενός βήματος μεταβολής από τον χρήστη. Το βήμα αυτό ονομάζεται learning rate ή ρυθμός μάθησης. Ο ρυθμός μάθησης είναι μια σταθερά που πολλαπλασιάζεται με την μερική παράγωγο της κάθε παραμέτρου. Ένας μικρός ρυθμός μάθησης θα ελαχιστοποιεί αργά την συνάρτηση κόστους προσεγγίζοντας το ελάχιστο με καλύτερη ακρίβεια σε σχέση με έναν γρήγορο. Ο γρήγορος ρυθμός ελαχιστοποιεί ταχύτερα τη συνάρτηση δίχως τη δυνατότητα να εισχωρήσει στο ελάχιστο της, δίνοντας γρήγορα αποτελέσματα με

περιορισμένη ακρίβεια. Σε πολλές περιπτώσεις ανάλογα με την πορεία του δικτύου η σταθερά αυτή μεταβάλλεται κατά τη διάρκεια της εκπαίδευσης συνδυάζοντας ταχύτητα στα πρώτα στάδια και ακρίβεια στα μετέπειτα.

To backward pass ολοκληρώνεται με την μεταβολή των παραμέτρων. Η ολοκλήρωση του forward-backward pass σε ένα δίκτυο πάνω στο σύνολο των διαθέσιμων δεδομένων ονομάζεται epoch ή εποχή. Σε διάφορες περιπτώσεις τα διαθέσιμα δεδομένα είναι μεγάλα σε αριθμό, καθιστώντας την διαδικασία της ανεύρεσης των μερικών παραγώγων δύσκολη λόγω του αυξημένου υπολογιστικού κόστους. Σε αυτήν την περίπτωση το δίκτυο δέχεται στην είσοδο πακέτα των δεδομένων αυτών, τα λεγόμενα mini batches. Η διαδικασία της μάθησης είναι η ίδια και η εποχή ολοκληρώνεται μετά την ανανέωση των παραμέτρων από το backward pass του τελευταίου mini batch. Έπειτα η όλη διαδικασία επαναλαμβάνεται εκ νέου για τις επόμενες εποχές.



Εικόνα 12. Η λειτουργία της οπισθοδιάδοσης
<https://medium.com/coinmonks/implement-back-propagation-in-neural-networks-ed09897593e7>

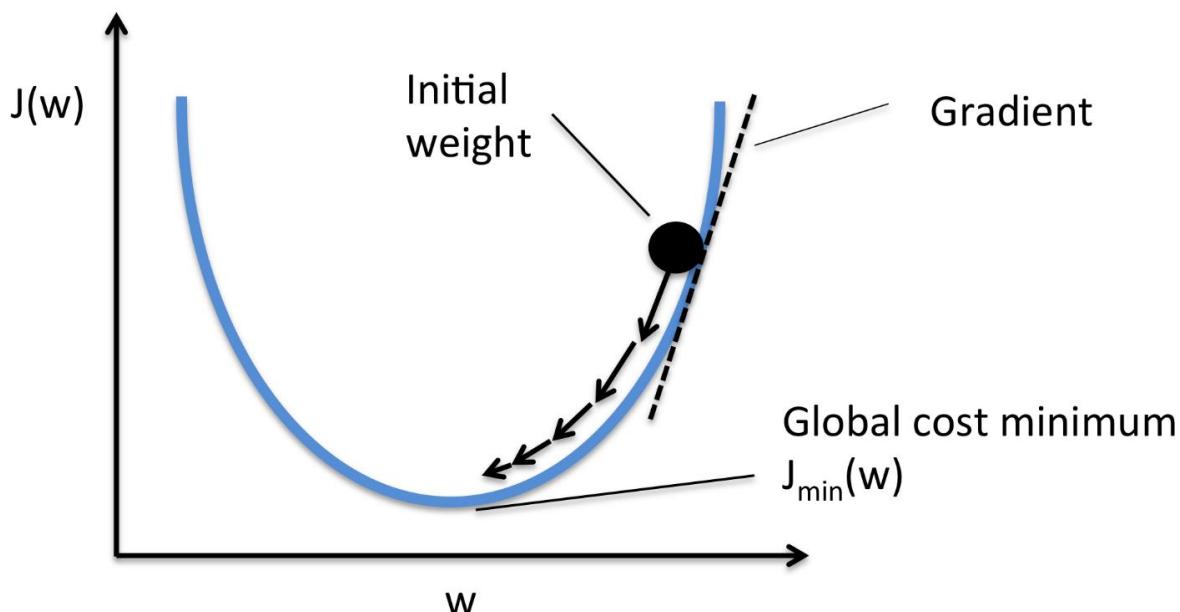
2.4.7 Στοχαστική καταβίβαση της παραγώγου

Ο αλγόριθμος οπισθοδιάδοσης είναι υπεύθυνος για τον υπολογισμό της κατεύθυνσης της μεταβολής που πρέπει να δεχθεί κάθε παράμετρος ώστε να ελαχιστοποιηθεί η συνάρτηση κόστους. Η διαδικασία που καθορίζει το πότε θα προκύψουν οι μεταβολές αυτές ονομάζεται optimization method ή μέθοδος βελτιστοποίησης..Η δημοφιλέστερη είναι η SGD (Stochastic gradient descent).

Gradient Descent ή καταβίβαση της παραγώγου ονομάζεται η διαδικασία της μεταβολής των παραμέτρων και ενημέρωσης των τιμών τους, όπως περιεγράφηκε παραπάνω. Με το πέρας της κάθε εποχής οι παράμετροι αποκτούν νέες τιμές. Η ύπαρξη των mini batches συνοδεύεται από την ύπαρξη της SGD, μιας παρόμοιας αλλά τροποποιημένης διαδικασίας. Σε αυτήν οι παράμετροι ανανεώνονται σε κάθε mini batch και όχι σε κάθε εποχή.

Το γεγονός ότι τα δεδομένα που εισάγει ο χρήστης εμφανίζουν μεγάλη συσχέτιση μεταξύ τους- αλλιώς δεν θα ήταν αποτελεσματική η διαδικασία της μάθησης- καθιστά το κάθε mini batch σε μεγάλο βαθμό αντιπροσωπευτικό δείγμα του συνόλου των δεδομένων. Οπότε η SGD προσεγγίζει σε ικανοποιητικό βαθμό την μεταβολή των παραμέτρων με αρκετά λιγότερο υπολογιστικό κόστος προσδίδοντας ταχύτητα στην μάθηση.

Υπάρχουν και άλλες μέθοδοι βελτιστοποίησης που χρησιμοποιούνται οι οποίες υπολογίζουν και άλλα στοιχεία πριν επιβάλλουν τη μεταβολή όπως μερικές παραγώγους δευτέρου βαθμού, μεταβαλλόμενους ρυθμούς μάθησης ανά παράμετρο κλπ. Μερικές από αυτούς είναι οι Adam, RMSprop, AdaGrad.

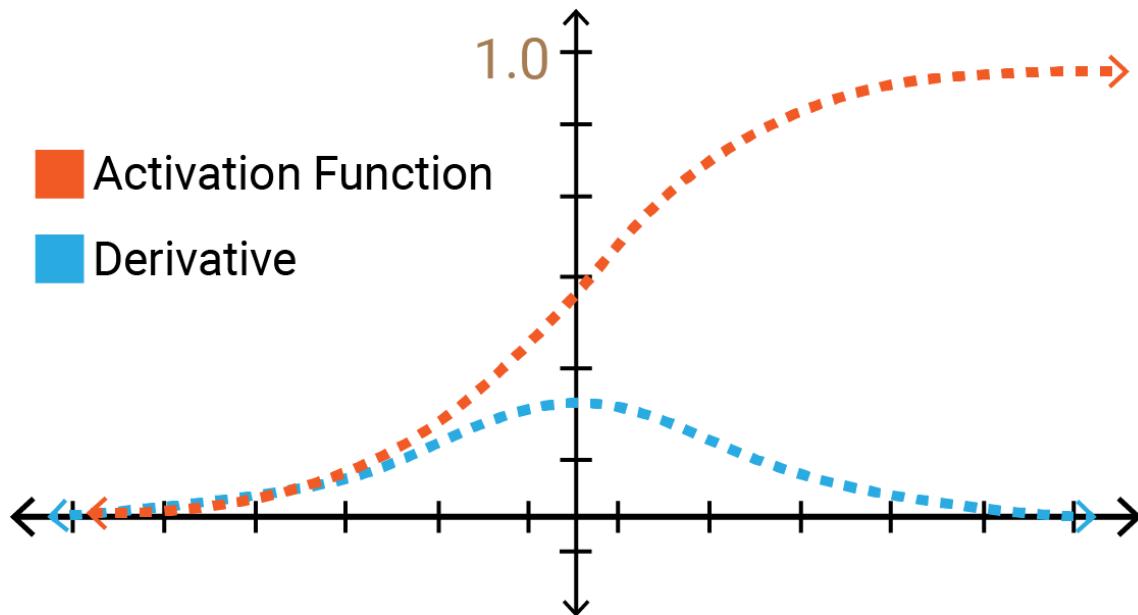


Εικόνα 13 :Διαδικασία καταβίβασης της παραγώγου

<https://medium.com/analytics-vidhya/all-about-gradient-descent-in-machine-learning-and-deep-learning-3dea4b269bf0>

2.4.7.1 Vanishing gradient problem

Στην ενότητα με τις συναρτήσεις ενεργοποίησης αναφέρθηκε το πρόβλημα της χαμένης παραγώγου. Το πρόβλημα αυτό παρουσιάζεται στα μεγάλα δίκτυα κυρίως που περιέχουν πολλά στρώματα. Κατά την διάρκεια του οπισθοδιάδοσης, όσο το σήμα οδηγείται προς το στρώμα εισόδου για να υπολογιστούν οι μερικές παράγωγοι παρατηρείται το φαινόμενο η παράγωγοι αυτοί να τείνουν στο 0. Το γεγονός αυτό εμποδίζει την ενημέρωση των παραμέτρων αφού η ποσότητα μεταβολής είναι πολύ μικρή, με αποτέλεσμα να το δίκτυο να μην εκπαιδεύεται.



Εικόνα 14: Αναπαράσταση της συνάρτησης ενεργοποίησης και της παραγώγου της
<http://socsci.uci.edu/~saberi/psych149/notes/2020psych149week7.pdf>

2.4.8 Συνελικτικά νευρωνικά δίκτυα

Το συνελικτικό νευρωνικό δίκτυο είναι ένα είδος νευρωνικού δίκτυου που έχει σαν βασικό δομικό στοιχείο τα συνελικτικά στρώματα. Η χρήση τους ξεκίνησε από 1997 (LeCun et al.) και μέχρι σήμερα χρησιμοποιούνται ευρέως στην ταξινόμηση των εικόνων, αλλά και στην αναγνώριση προτύπων και σε ζητήματα όρασης υπολογιστών γενικότερα. Αποτελεί ίσως την πιο διαδεδομένη αρχιτεκτονική νευρωνικού δίκτυου.

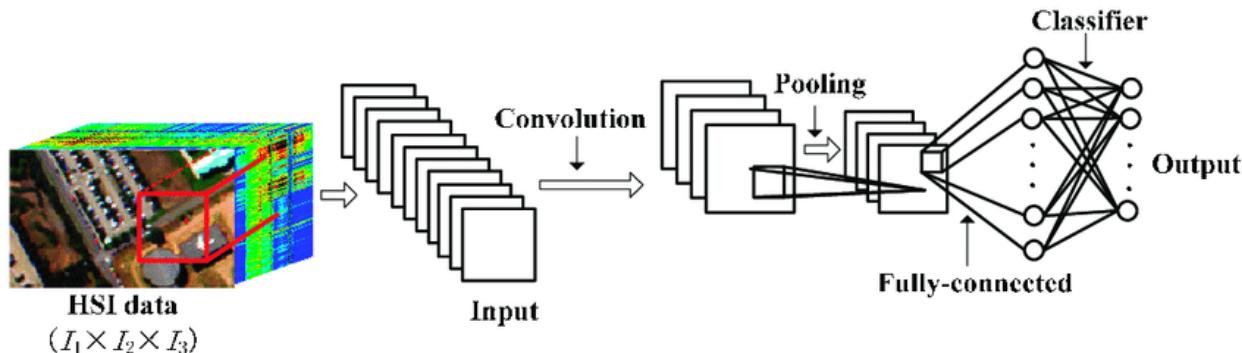
Τα δίκτυα αυτού του τύπου παρουσιάζουν πλεονεκτήματα σε σχέση με τα απλά δίκτυα στις περιπτώσεις που εφαρμόζονται. Αρχικά έχουν την δυνατότητα να μαθαίνουν πολυπλοκότερους συσχετισμούς μεταξύ των δεδομένων χωρίς να εμπλέκεται ο ανθρώπινος παράγοντας. Για παράδειγμα στην ταξινόμηση εικόνων δεν χρειάζεται να προγιγηθούν επεξεργασίες -μεταλλάξεις των δεδομένων όπως ανίχνευση ακμών κλπ. Παρόλα αυτά, η χειροκίνητη εξαγωγή στοιχείων συνδυάζεται πολλές φορές με τα δίκτυα αυτά, όπως η τεχνική PCA.

Ένα ακόμα πλεονέκτημα των συνελικτικών νευρωνικών δίκτυων σε σχέση με τα συμβατικά είναι οι μειωμένοι σε αριθμό παράμετροι προς εκπαίδευση και κατά συνέπεια τα λιγότερα υπολογιστικά κόστη. Στα συμβατικά δίκτυα όλοι οι νευρώνες του κάθε στρώματος συνδέονται με τον κάθε

νευρώνα του προηγούμενου οδηγώντας σε πολλούς παραμέτρους. Στα συνελικτικά στρώματα ο κάθε νευρώνας έχει τόσα βάρη όσα και η διάσταση του φίλτρου του στρώματος. Επίσης τα βάρη αυτά θα είναι κοινά για όλους τους νευρώνες του ίδιου στρώματος.

Στην περίπτωση της ταξινόμησης υπερφασματικών εικόνων, τα δεδομένα της εισόδου είναι μεμονωμένα εικονοστοιχεία της εικόνας ή γειτονία αυτών, που εισέρχονται στο δίκτυο σαν τρισδιάστατο διάνυσμα. Το διάνυσμα αυτό μετασχηματίζεται περνώντας από τα διάφορα συνελικτικά στρώματα του δικτύου έως ότου καταλήξει σε μονοδιάστατο διάνυσμα με στοιχεία N , που θα περιλαμβάνει τις πιθανότητες να ανήκει το κάθε στοιχείο σε μία από τις κλάσεις N .

Τα συνελικτικά στρώματα είναι τα convolution layer και pooling layer.



Εικόνα 15: Συνελικτικό νευρωνικό δίκτυο εφαρμοσμένο σε υπερφασματική εικόνα
https://www.researchgate.net/figure/Framework-of-a-three-dimensional-convolutional-neural-network-3D-CNN_fig1_327520826

2.4.8.1 Συνελικτικό στρώμα

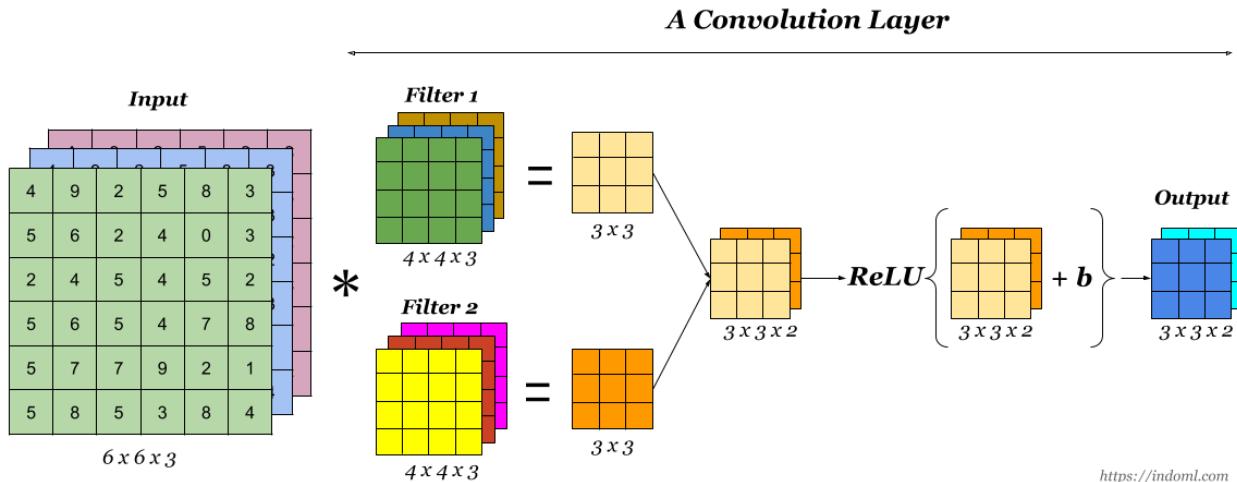
Το συνελικτικό στρώμα είναι το βασικότερο στρώμα του συνελικτικού νευρωνικού δικτύου. Είναι αυτό που προκαλεί το μεγαλύτερο υπολογιστικό κόστος αλλά έχει τη δυνατότητα να εκπαιδευτεί σε βάθος πάνω στους συσχετισμούς των δεδομένων. Το στρώμα αυτό αποτελείται από χωρικά φίλτρα που εφαρμόζονται στην εικόνα. Τα φίλτρα αυτά δρουν σε υποπεριοχές της εικόνας μέχρι να την σαρώσουν στο σύνολο της. Καθώς η εικόνα σαρώνεται, δημιουργείται ένας χάρτης χαρακτηριστικών με τα αποτελέσματα της συνέλιξης των φίλτρων. Έτσι ο χάρτης αυτός αναδεικνύει την ύπαρξη προτύπων στην εικόνα.

Τα βάρη του συνελικτικού στρώματος είναι οι τιμές των παραγόντων των φίλτρων. Συνεπώς ένα τρισδιάστατο φίλτρο $3 \times 3 \times 3$ θα περιέχει συνολικά 28 παραμέτρους, 27 βάρη του φίλτρου και μία τιμή κατωφλιού. Οι παράμετροι αυτοί είναι ίδιοι για κάθε φίλτρο που ανήκει στο στρώμα.

Τα φίλτρα εφαρμόζονται σε μία, δυο ή τρεις διαστάσεις ταυτόχρονα. Στην περίπτωση της φασματικής ταξινόμησης Υπερφασματικών εικόνων χρησιμοποιούνται δισδιάστατα φίλτρα αριθμού ίσου με τον αριθμό των φασματικών καναλιών, όπου δρα το κάθε ένα στο κανάλι που του αντιστοιχεί. Άλλη εφαρμογή του συνελικτικού στρώματος είναι η εφαρμογή ενός τρισδιάστατου φίλτρου που αντιμετωπίζει την εικόνα σαν κύβο με την φασματική διάσταση σαν διάσταση του βάθους.

Άλλη μια παράμετρος που καθορίζει την λειτουργία του φίλτρου είναι το βήμα μετακίνησης. Το κάθε φίλτρο εφαρμόζεται σε ένα εικονοστοιχείο, που ονομάζεται κεντρικό εικονοστοιχείο. Στη συνέχεια το φίλτρο μετακινείται στο επόμενο εικονοστοιχείο μέχρι να ολοκληρωθεί η εφαρμογή του σε όλα τα διαθέσιμα εικονοστοιχεία. Η απόσταση μεταξύ των διαδοχικών εικονοστοιχείων που εφαρμόζεται το φίλτρο ονομάζεται βήμα μετακίνησης ή stride.

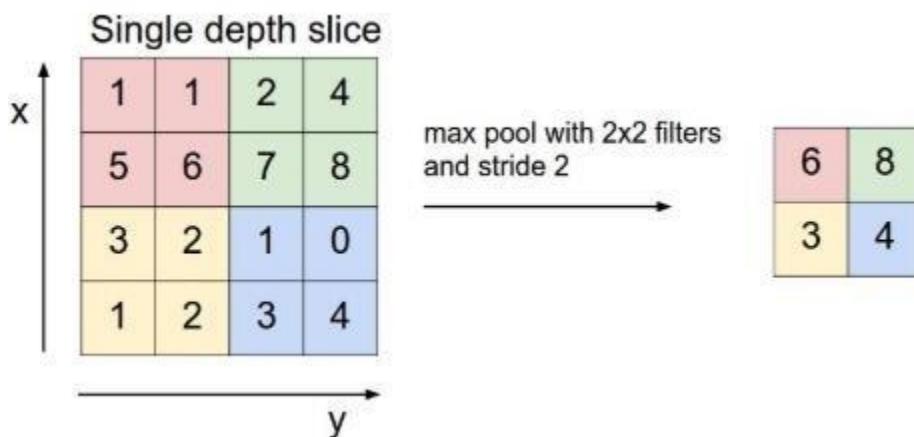
Η κατάλληλη χρήση του βήματος μετακίνησης σε συνδυασμό με την επέκταση της αρχικής εικόνας με μηδενικά κατά πλάτος και ύψος (padding) έχουν την δυνατότητα να οδηγήσουν σε χάρτη χαρακτηριστικών με όμοιες διαστάσεις με την αρχική εικόνα. Η τεχνική αυτή χρησιμοποιείται αρκετά, αφού το πέρασμα μιας εικόνας από ένα συνελικτικό στρώμα μεταβάλει τις διαστάσεις της. Πάνω στη λογική της δυνατότητας να καθοριστούν οι διαστάσεις του της εξόδου, το κάθε συνελικτικό στρώμα μπορεί να χρησιμοποιηθεί ουσιαστικά και ως στρώμα pooling.



Εικόνα 16: Συνελικτικό στρώμα και τρόπος λειτουργίας
<https://towardsdatascience.com/demystifying-convolutional-neural-networks-384785791596>

2.4.8.2 Pooling στρώμα

Τα στρώματα pooling χρησιμοποιούνται μετά από τις συνελίξεις με σκοπό την μείωση των διαστάσεων του χάρτη χαρακτηριστικών και κατά συνέπεια των παραμέτρων προς μάθηση στα επόμενα συνελικτικά στρώματα και το υπολογιστικό κόστος. Η διαδικασία αυτή δεν εισάγει καμία νέα παράμετρο. Είναι μια συνάρτηση που παίρνει τις τιμές του φίλτρου και επιστρέφει τον μέσο όρο τους, είτε την μέγιστη τιμή ανάλογα με την επιλογή. Χαρακτηριστικό παράδειγμα είναι πως η συνηθέστερη εφαρμογή pooling layer με φίλτρο 2x2 και συνάρτηση “max” μειώνει τις χωρικές διαστάσεις 50%, διατηρεί την φασματική διάσταση των εικονοστοιχείων που απέμειναν αναλλοίωτη και μειώνει τους παραμέτρους προς μάθηση κατά 75% στις επόμενες συνελίξεις.



Εικόνα 17: Στρώμα pooling και λειτουργία του
https://web.njit.edu/~usman/courses/cs732_spring19/Yanan_An_overview_of_deep_learning_in_medical_imaging_focusing_on_MRI.pdf

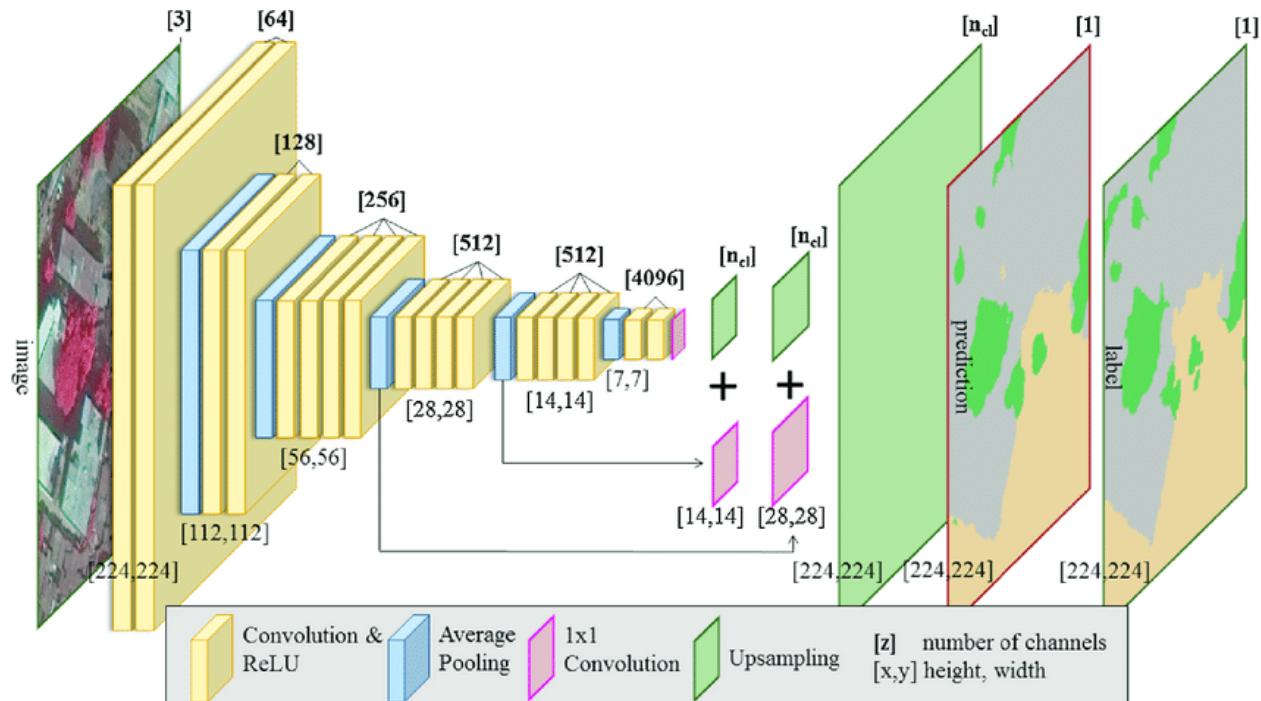
Προκειμένου να δοθεί η τελική πρόβλεψη του κάθε εικονοστοιχείου, με το πέρας των συνελικτικών και pooling στρωμάτων ακολουθεί ένα πλήρως συνδεδεμένο στρώμα ή δίκτυο το οποίο δίνει ένα τελικό διάνυσμα με τις πιθανότητες της κάθε κλάσης N για κάθε εικονοστοιχείο. Υπάρχουν όμως και περιπτώσεις δικτύων που δεν χρησιμοποιούν στο τέλος της αρχιτεκτονικής τους πλήρως συνδεδεμένα στρώματα. Αυτά τα δίκτυα ονομάζονται πλήρως συνελικτικά δίκτυα.

2.4.9 Πλήρως συνελικτικά δίκτυα

Τα πλήρως συνελικτικά δίκτυα είναι δίκτυα που περιλαμβάνουν στην αρχιτεκτονική τους μόνο συνελικτικά στρώματα και αποσιάζουν τα πλήρως συνδεδεμένα στρώματα. Το γεγονός αυτό οδηγεί σε διαφορές σε σχέση με τα τυπικά συνελικτικά δίκτυα.

Αρχικά εμφανίζουν λιγότερους παραμέτρους προς μάθηση αφού τα συνελικτικά στρώματα μοιράζονται τα βάρη στους νευρώνες του ίδιου στρώματος σε αντίθεση με τα πλήρως συνδεδεμένα δίκτυα που κάθε νευρώνας έχει τα δικά του βάρη. Ακόμα τα πλήρως συνδεδεμένα στρώματα δέχονται την είσοδο σε μορφή μονοδιάστατου πίνακα, χωρίς να εκμεταλλεύονται την χωρική πληροφορία που είχε διατηρηθεί από τις συνελίξεις που είχαν προηγηθεί. Η απουσία των πλήρως συνδεδεμένων στρωμάτων διατηρεί αυτήν την χωρική πληροφορία.

Τέλος τα πλήρως συνελικτικά δίκτυα μπορούν να δεχθούν ως είσοδο εικόνες οποιωνδήποτε διαστάσεων. Η αρχιτεκτονική των δικτύων αυτών επιτρέπει την εφαρμογή της σε πολλά προβλήματα αφού οι διαστάσεις των δεδομένων εισαγωγής δεν πρέπει να είναι καθορισμένες.



Εικόνα 18 : Πλήρες συνελικτικό δίκτυο

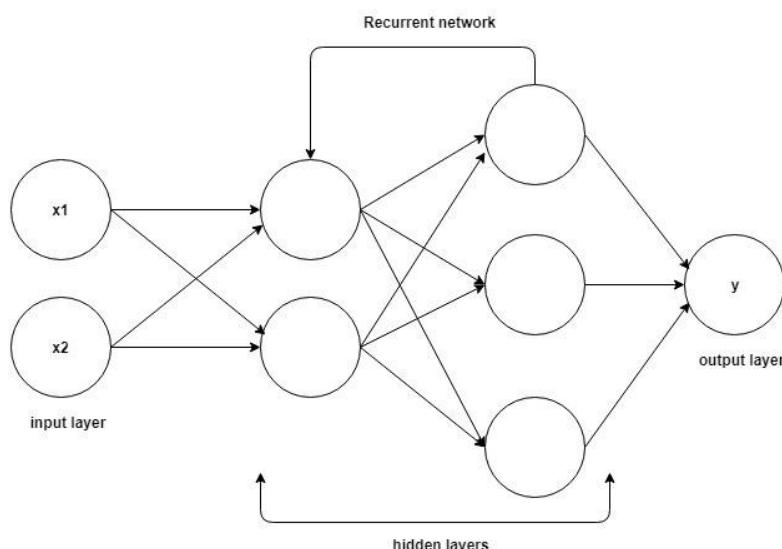
<https://www.researchgate.net/project/So2Sat-Big-Data-for-4D-Global-Urban-Mapping-1016-Bytes-from-Social-Media-to-Earth-Observation-Satellites>

2.4.10 Αναδραστικά νευρωνικά δίκτυα

Όλοι οι τύποι δικτύων που έχουν αναφερθεί μέχρι στιγμής ανήκουν στην οικογένεια των feedforward networks. Η οικογένεια αυτή χαρακτηρίζεται από το γεγονός ότι η ροή των δικτύων αυτών δεν είναι κυκλική. Το σήμα μεταφέρεται από το στρώμα εισόδου στα κρυφά στρώματα και τελικά στο στρώμα εξόδου σε μία κατεύθυνση διαδικασία του οπισθοδιάδοσης δεν μεταφέρει το σήμα αλλά υπολογίζει της παραγώγους τον παραμέτρων.

Τα αναδραστικά νευρωνικά δίκτυα ανήκουν στην κατηγορία των feedback networks. Στην κατηγορία αυτή το σήμα παίρνει διάφορες κατευθύνσεις στην ροή του δικτύου. Η διαφορά με τα feedforward networks είναι πως σε αυτά το κάθε σήμα δεν επηρεάζει την διαδικασία της μάθησης σχετικά με το επόμενο διαδοχικά σήμα προς εισαγωγή. Στα αναδραστικά νευρωνικά δίκτυα το σήμα ή επηρεάζει την πρόβλεψη του σήματος $n + 1$.

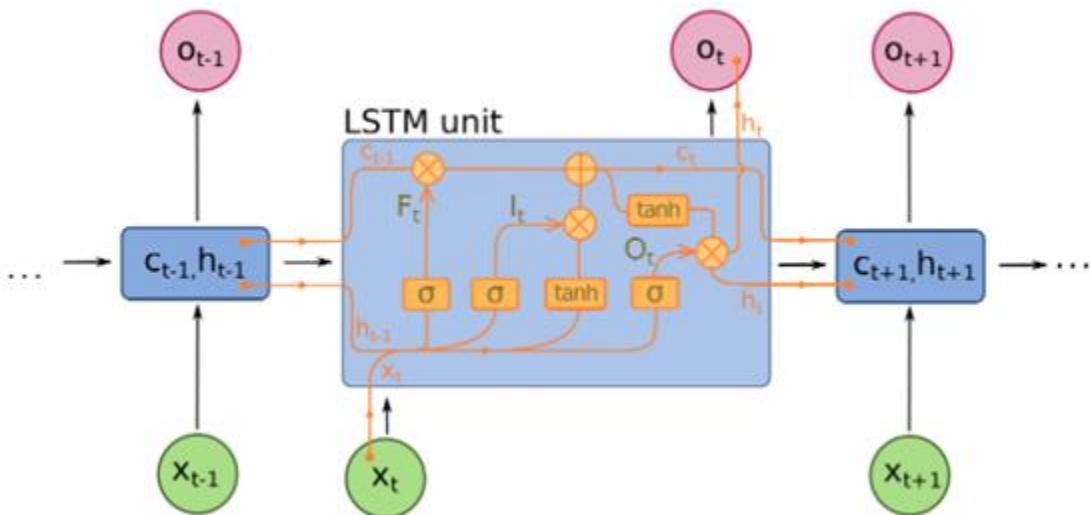
Τα δίκτυα αυτά χρησιμοποιούνται συνήθως σε προβλέψεις χρονοσειρών, στην αναγνώριση ομιλίας κλπ. Η αρχιτεκτονική τους προσπαθεί να προσεγγίσει την μνήμη του ανθρώπινου εγκέφαλου και να βασιστεί σε αυτήν για ακριβέστερη πρόβλεψη. Ένα χαρακτηριστικό παράδειγμα είναι μία διαδικασία πρόβλεψης της επόμενης λέξης μιας πρότασης όπου η πρόβλεψη θα παίρνει στοιχεία από της προηγούμενες λέξεις, την διαδικασία που ονομάζεται «κατανόηση από τα συμφραζόμενα». Τα δίκτυα που θα εξεταστούν στην διπλωματική εργασία πετυχαίνουν την προσομοίωση της μνήμης αυτής με την χρήση ειδικών στρωμάτων.



Εικόνα 19:Αναδραστικό δίκτυο
<https://mc.ai/rnn-or-recurrent-neural-network-for-noobs/>

2.4.10.1 Στρώμα Μακροπρόθεσμης-βραχυπρόθεσμης μνήμης LSTM

Τα στρώματα Μακροπρόθεσμης-βραχυπρόθεσμης μνήμης πρωτοεμφανίστηκαν το 1997 από τους Hochreiter & Schmidhuber . Είναι ιδιαίτερα στρώματα που επιτρέπουν την συγκράτηση πληροφοριών για αρκετό χρόνο κατά την διάρκεια εκπαίδευσης του δικτύου. Ο κάθε νευρώνας του αποτελείται από την κατάσταση μνήμης, την πόρτα εισόδου, την πόρτα εξόδου και την πόρτα μνήμης. Η κατάσταση μνήμης του νευρώνα είναι υπεύθυνη για την διατήρηση των στοιχείων που αποκτούνται σε κάθε πέρασμα από το στρώμα και για την τροφοδότηση με τα στοιχεία αυτά στα επόμενα περάσματα. Ουσιαστικά είναι ο χώρος αποθήκευσης της μνήμης. Η πόρτα μνήμης είναι η διαδικασία με την οποία αξιολογείται το ποια πληροφορία θα αποθηκευτεί και ποια θα ξεχασθεί από την κατάσταση μνήμης του νευρώνα. Τέλος οι πόρτες εισόδου και εξόδου είναι αυτές που καθορίζουν ποιες πληροφορίες θα προστεθούν στην είσοδο από την αποθήκη της κατάστασης μνήμης και ποιες πληροφορίες θα οδηγηθούν στην πόρτα μνήμης στο επόμενο πέρασμα αντίστοιχα. Έτσι τα στρώματα αυτά καταφέρνουν να φιλτράρουν τις σημαντικές πληροφορίες ,να πετούν τις ασήμαντες και να διατηρούν πληροφορίες που συνάντησαν αρκετά περάσματα πριν προσμοιάζοντας την μακροπρόθεσμη μνήμη του ανθρώπου.



$$i_t = \sigma(W^{(i)}x_t + U^{(i)}h_{t-1}) \quad (\text{Input gate})$$

$$f_t = \sigma(W^{(f)}x_t + U^{(f)}h_{t-1}) \quad (\text{Forget gate})$$

$$o_t = \sigma(W^{(o)}x_t + U^{(o)}h_{t-1}) \quad (\text{Output/Exposure gate})$$

Q

$$\tilde{c}_t = \tanh(W^{(c)}x_t + U^{(c)}h_{t-1}) \quad (\text{New memory cell})$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (\text{Final memory cell})$$

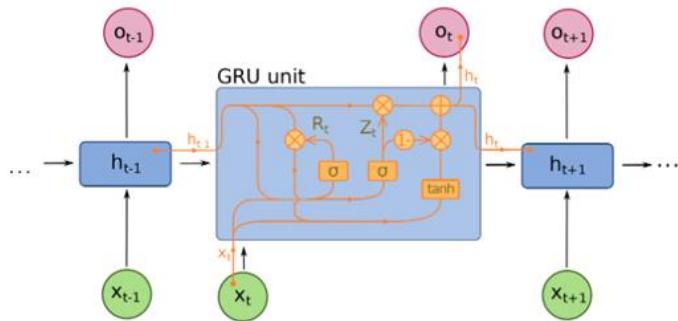
$$h_t = o_t \circ \tanh(c_t)$$

Εικόνα 20: Αρχιτεκτονική του στρώματος LSTM
https://en.wikipedia.org/wiki/Long_short-term_memory

2.4.10.2 GRU στρώμα

Το στρώμα GRU είναι μία παραλλαγή του στρώματος LSTM η οποία παρουσιάζει λιγότερα υπολογιστικά κόστη, αφού η αρχιτεκτονική του είναι πιο απλή από τον LSTM και κατά συνέπεια οι παράμετροι προς εκπαίδευση είναι λιγότεροι. Χρησιμοποιείται αρκετά διότι αν και πιο απλή αρχιτεκτονική, οι ακρίβειες που προσφέρει είναι σχεδόν αντίστοιχες με το στρώμα LSTM ή και καλύτερες σε συγκεκριμένες περιπτώσεις που έχουν να κάνουν με διαδικασίες που απαιτούν βραχυπρόθεσμη μνήμη ή λιγότερα δεδομένα μάθησης. Αυτό οφείλεται στην απλή αρχιτεκτονική του.

Το στρώμα GRU διαθέτει 2 πόρτες. Η μία πόρτα είναι η πόρτα ανανέωσης η οποία είναι υπεύθυνη να ανανεώσει τη μνήμη με το νέο σετ πληροφοριών. Ακόμα η πόρτα επαναφοράς είναι αυτή που θα καθορίσει ποια πληροφορία απορρίπτεται. Η ουσιαστική διαφορά με τα LSTM είναι ότι απουσιάζει η πόρτα εισόδου. Έτσι το σύνολο της μνήμης εκτίθεται στην είσοδο σε κάθε πέρασμα και δεν υπάρχει φίλτρο που απορρίπτει πληροφορία της μνήμης είναι σχετική με την είσοδο. Αυτό έχει σαν αποτέλεσμα την συχνότερη απόρριψη παλιών πληροφοριών από τα προηγούμενα περάσματα οδηγώντας σε βραχυπρόθεσμη μνήμη.



$$\tilde{h}_t = \tanh(r_t \circ Uh_{t-1} + Wx_t) \quad (\text{New memory})$$

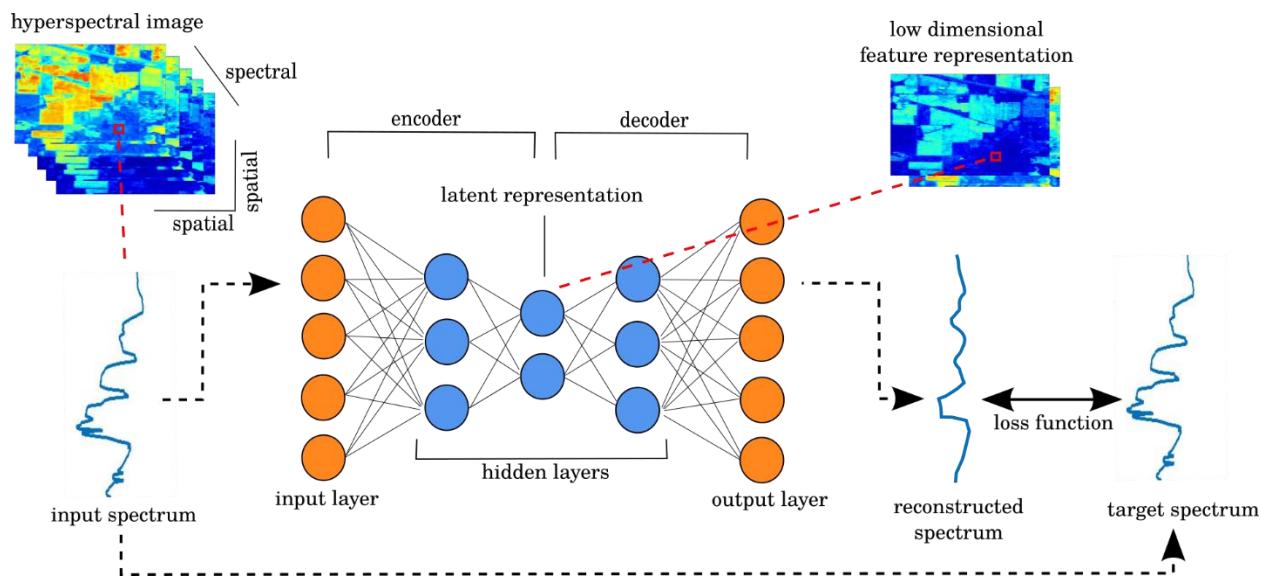
$$h_t = (1 - z_t) \circ \tilde{h}_t + z_t \circ h_{t-1} \quad (\text{Hidden state})$$

Εικόνα 21: Αρχιτεκτονική του στρώματος GRU
https://en.wikipedia.org/wiki/Gated_recurrent_unit

2.4.11 Δίκτυο Κωδικοποιητή-Αποκωδικοποιητή

Ένας άλλος τύπος νευρωνικών δικτύων είναι οι δίκτυο Κωδικοποιητή-Αποκωδικοποιητή. Χρησιμοποιούνται σε διάφορες εφαρμογές για την μείωση διαστάσεων δειγμάτων, συμπίεσης και ανασύνθεσης δεδομένων καθώς και προβλέψεις πληθυσμών κλπ. Στην φασματική ταξινόμηση χρησιμοποιούνται κυρίως για την μείωση των φασματικών διαστάσεων ή για αφαίρεση του θορύβου από τα δεδομένα.

Η αρχιτεκτονική αυτών των δικτύων απαρτίζεται από ένα ή περισσότερα πλήρως συνδεδεμένα στρώματα. Το δίκτυο δέχεται το δεδομένο και περνά από τον κωδικοποιητή και τον αποκωδικοποιητή. Στην διαδικασία της κωδικοποίησης το δίκτυο εκπαιδεύεται στην αναπαράσταση του δεδομένου με λιγότερη πληροφορία. Έπειτα εκπαιδεύεται στην αναπροσαρμογή της αναπαράστασης αυτής στην αρχική μορφή του δεδομένου. Η βελτιστοποίηση των παραμέτρων γίνεται μέσω της συνάρτησης κόστους μεταξύ της αρχικής και τελικής μορφής του δεδομένου.



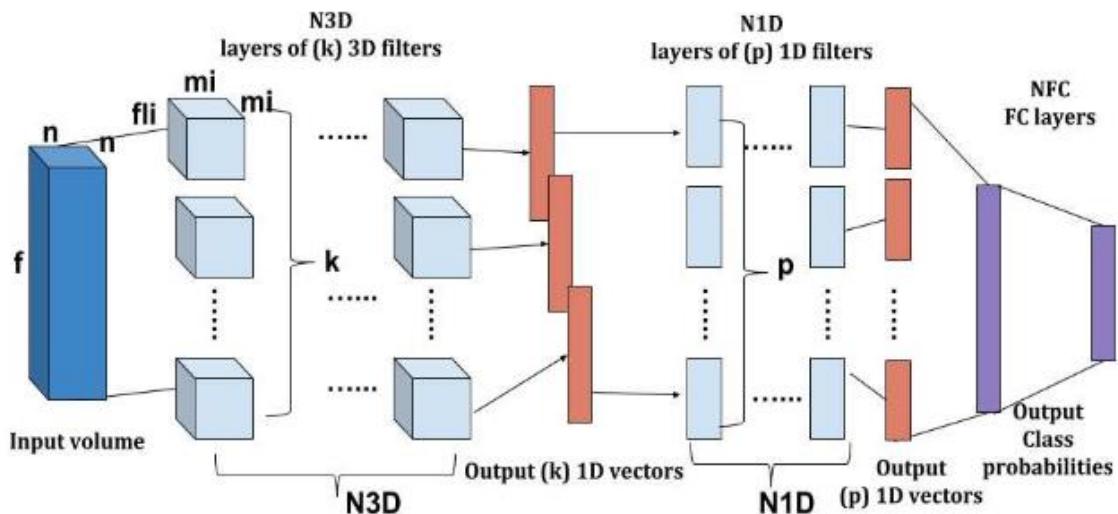
Εικόνα 22 : Αρχιτεκτονική δικτύου Κωδικοποιητή-Αποκωδικοποιητή
<https://deephyp.readthedocs.io/en/latest/>

3 ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ-ΑΛΓΟΡΙΘΜΟΙ

3.1 HAMIDA ET AL.

Το δίκτυο Hamida et al.[1] είναι ένα CNN που σχεδιάστηκε με σκοπό την ταξινόμηση υπερφασματικών εικόνων. Οι δημιουργοί του δικτύου αυτού θέλησαν να χρησιμοποιήσουν μία τρισδιάστατη προσέγγιση των δεδομένων έτσι ώστε η διαδικασία μάθησης να βασίζεται στην χωρική αλλά και φασματική πληροφορία των εικονοστοιχείο ταυτόχρονα.

Το δίκτυο αυτό δέχεται ως εισαγωγή μια γειτονιά εικονοστοιχείο και είναι της μορφής $n \times n \times f$ όπου n ο αριθμός των εικονοστοιχείο και f αριθμός των φασματικών τιμών των φασματικών καναλιών που αντιστοιχούν στο κάθε εικονοστοιχείο. Έπειτα πραγματοποιεί τρισδιάστατες συνελίξεις οι οποίες σταδιακά μαθαίνουν τούς χωρικούς-φασματικούς συσχετισμούς αλλά ταυτόχρονα αποβάλουν πληροφορία ώστε ότου προκύψει ένα μονοδιάστατο διάνυσμα. Αυτό με την σειρά του προωθείται σε άλλες μονοδιάστατες συνελίξεις μέχρι να οδηγηθεί σε ένα FCN δίκτυο για να γίνει η τελική ταξινόμηση του εικονοστοιχείου.

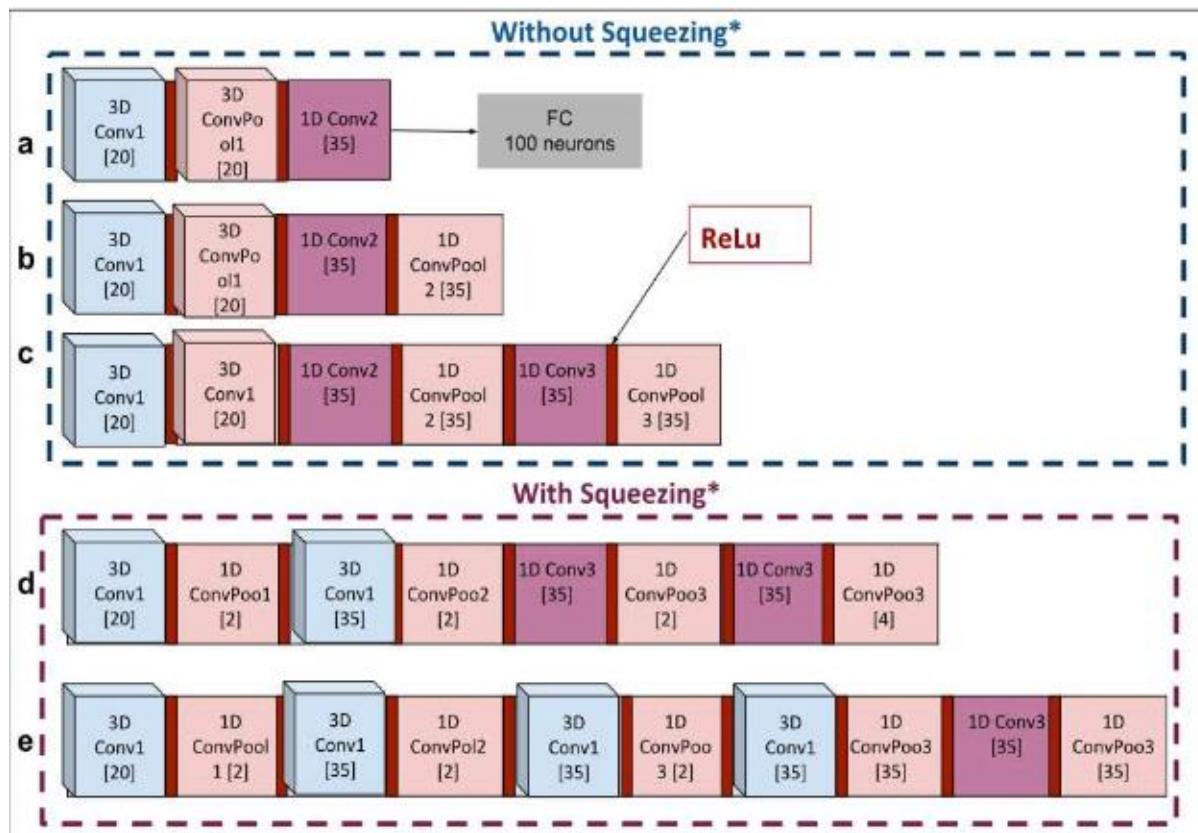


Εικόνα 23:Η βασική αρχιτεκτονική του δικτύου Hamida et al.

Καθώς αυτός είναι η βασικός κορμός του δικτύου, η δημιουργοί πραγματοποίησαν διάφορα πειράματα ώστε να προσδιορίσουν την ακριβή αρχιτεκτονική του με γνώμονα την ακρίβεια και την ταχύτητα μάθησης. Ασχολήθηκαν με 3 διαφορετικές εικόνες οι οποίες είναι γνωστές και χρησιμοποιούνται συχνά σε τέτοιου τύπου πειράματα(PaviaU,PaviaC,KSC).Πειραματίζονται με το μέγεθος της γειτονιάς των εικονοστοιχείο στην εισαγωγή, με το πόσα στρώματα είναι τα

ιδανικά αλλά και με τον αριθμό των νευρώνων σε κάθε στρώμα. Μελετούν τον ρόλο που παίζει ο αριθμός των παραμέτρων προς μάθηση του μοντέλου στην ακρίβεια του και πως μπορούν να ελαφρύνουν το μοντέλο υπολογιστικά μειώνοντας τις παραμέτρους και διατηρώντας την ακρίβεια.

Συγκρίνοντας λοιπόν την ακρίβεια της κάθε αρχιτεκτονικής και τον χρόνο εκπαίδευσης γίνεται η παραδοχή πως το δίκτυο 8 στρωμάτων (d) υπερέχει των άλλων επιλογών που φαίνονται στην παρακάτω εικόνα. Συγκεκριμένα η καλύτερη ακρίβεια σημειώθηκε με την γειτονιά των 5 εικονοστοιχείο σαν εισαγωγή.



Εικόνα 24:Οι διαφορετικές παραλλαγές στις αρχιτεκτονικές του δικτύου Hamida et al.

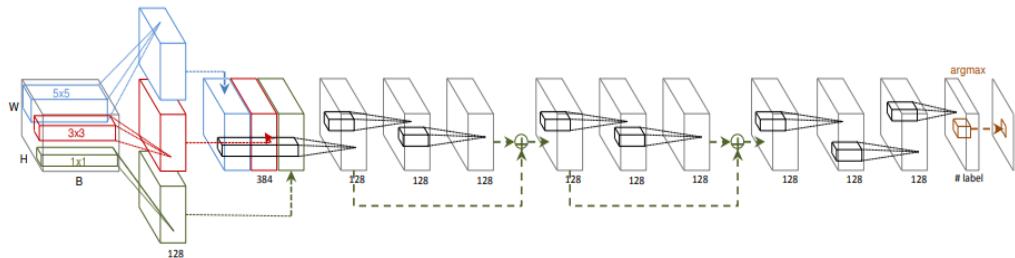
Conv Nb	3D Conv Nb	Pooling neu- ron Nb	parameter Nb	Accuracy
6	2	[2,2,35]	11123	96.1%
6	2	[2,2,4]	3931	92.8%
8	2	[2,2,35,35]	27470	98,3%
8	3	[2,2,2,4]	6862	98.9%

Εικόνα 25:Πίνακας που παρουσιάζει την ολική ακρίβεια, τον αριθμό των παραμέτρων προς εκπαίδευση για τις παραλλαγές των 6 και 8 συνελικτικών στρωμάτων

Η μέθοδος squeezing είναι ουσιαστικά μια προσπάθεια συμπίεσης του μοντέλου. Τα πειράματα έδειξαν ότι κατά τη διάρκεια της υποδειγματοληψίας, η αντικατάσταση των υπευθύνων για αυτήν στρωμάτων με συνελικτικά στρώματα που έχουν φίλτρα μιας διάστασης διατηρούν την ακρίβεια και μειώνουν τις παραμέτρους προς μάθηση. Έτσι με κατάλληλη χρήση του βήματος των φίλτρων, τα convpool στρώματα δρουν ταυτόχρονα ως συνελίξεις μαθαίνοντας πληροφορία και ως pooling στρώματα μειώνοντας τις διαστάσεις της εισαγωγής.

3.2 LEE ET AL.

Το δίκτυο Lee et al.[2] είναι ένα δίκτυο FCN το οποίο δημιουργήθηκε για την ταξινόμηση υπερφασματικών εικόνων. Στηρίζει την αποδοτικότητα του στο γεγονός ότι από την αρχή της εισόδου το δίκτυο εκπαιδεύεται στις χωρικές και φασματικές σχέσεις μεταξύ των εικονοστοιχείων. Μέσω μιας τεχνικής που αποκαλείται multi scale filter bank εξάγει αυτές τις πληροφορίες, τις μετατρέπει σε ένα μονοδιάστατο διάνυσμα και συνεχίζει με συνελίξεις μέχρι να δοθεί η πρόβλεψη. Οι συνελίξεις αυτές βασίζονται στην μέθοδο residual learning, μεγαλώνοντας το δίκτυο δίχως τα μειονεκτήματα του αυξημένου υπολογιστικού κόστους. Γενικότερα το δίκτυο αυτό σε σχέση με άλλα παρόμοια παρουσιάζει πολλούς παραμέτρους προς εκπαίδευση και καταφέρνει καλές ακρίβειες χωρίς να απαιτούνται αντίστοιχου μεγέθους δεδομένα. Οι προαναφερθείσες τεχνικές βοηθούν σε αυτό όπως και στην αποφυγή της υπερπροσαρμογής. Τέλος ένα σημαντικό πλεονέκτημα του δικτύου αυτού είναι ότι λόγο της αρχιτεκτονικής που θα δούμε παρακάτω μπορεί να εκπαιδευτεί και να εφαρμοσθεί σε κάθε τύπου εικόνα, οπότε η μέθοδος αυτή μπορεί να χρησιμοποιηθεί και για άλλες εφαρμογές πέρα από την φασματική ταξινόμηση.



Εικόνα 26 :Η αρχιτεκτονική του δικτύου Lee et al.

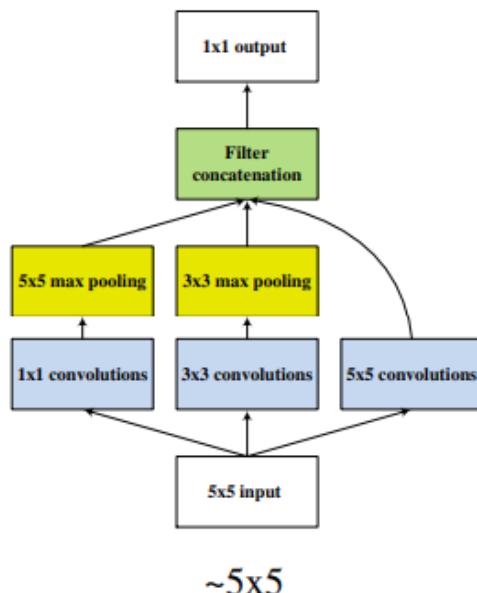
Το μέγεθος την γειτονίας των εικονοστοιχείων στην εισαγωγή είναι 5×5 . Η επιλογή αυτή σε σχέση με μικρότερες ή μεγαλύτερες γειτονίες προέκυψε βάση πειραμάτων που πραγματοποιήθηκαν σε γνωστά datasets (IndianPines, PaviaU, Salinas), έχοντας την καλύτερη ακρίβεια. Η ταυτόχρονη αξιοποίηση των χωρικών και φασματικών σχέσεων χαρίζει στο δίκτυο αρκετά περισσότερη αποτελεσματικότητα σε σχέση με την φασματική, pixel-based προσέγγιση (1×1).

TABLE X
PERFORMANCE COMPARISON OF THE PROPOSED NETWORK (IN PERCENTAGE) W.R.T. MULTI-SCALE FILTER BANKS WITH DIFFERENT CONFIGURATIONS.
 $\sim 7 \times 7$ MEANS THE MULTI-SCALE FILTER BANK CONSISTING OF 1×1 , 3×3 , 5×5 , AND 7×7 CONVOLUTION FILTERS.

Dataset	1×1	$\sim 3 \times 3$	$\sim 5 \times 5$	$\sim 7 \times 7$
Indian Pines	53.67 ± 16.63	87.37 ± 4.12	93.61 ± 0.56	93.47 ± 0.77
Salinas	50.62 ± 30.87	92.08 ± 0.77	95.07 ± 0.23	94.20 ± 0.43
University of Pavia	65.62 ± 8.18	93.59 ± 1.35	95.97 ± 0.46	95.91 ± 0.50

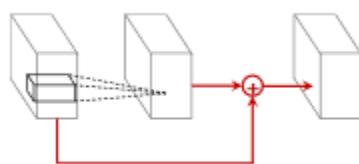
Εικόνα 27 : Πίνακας που παρουσιάζει την ολική ακρίβεια της μεθόδου πάνω στα γνωστά datasets με γειτονιές εικονοστοιχείων διάφορων μεγεθών στην είσοδο

Στο σημείο αυτό το διάνυσμα περνάει από 3 διαφορετικά φίλτρα ξεχωριστά με την διαδικασία του multi scale filter bank. Το φίλτρο 1×1 χρησιμοποιείται για να εξερευνήσει τις φασματικές σχέσεις ενώ τα 3×3 και 5×5 για τις χωρικές. Στη συνέχεια τα παράγωγα έρχονται στις ίδιες διαστάσεις μέσω στρωμάτων pooling και ενώνονται όπως στην εικόνα:



Εικόνα 28:Multi scale filter bank

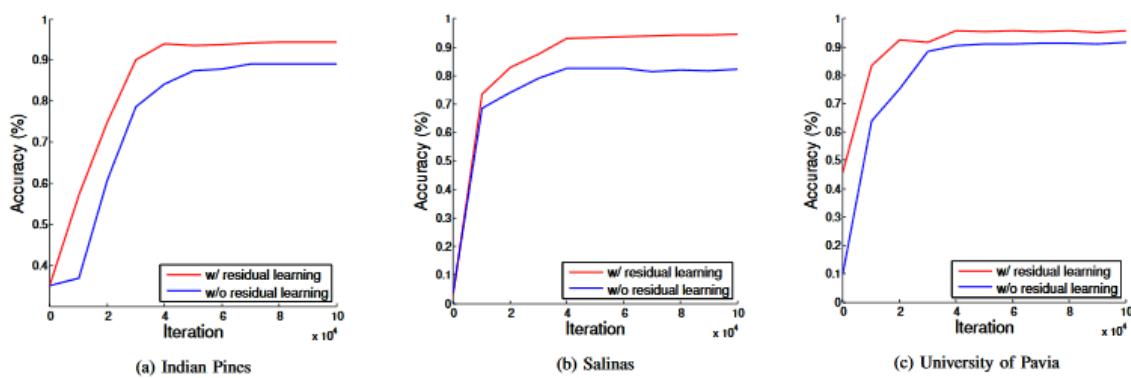
Έτσι δημιουργήθηκε ένας χάρτης χαρακτηριστικών που περιέχει πλούσιες φασματικές και χωρικές πληροφορίες για την εικόνα. Στη συνέχεια θα περάσει από διάφορα 1×1 φίλτρα με την χρήση του residual learning. Η μέθοδος αυτή χρησιμοποιεί τα διάφορα 1×1 συνελικτικά στρώματα με τέτοιον τρόπο ώστε η έξοδος της συνέλιξης να περνά σαν είσοδος στην επόμενη αλλά και σε επόμενα στρώματα, παραλείποντας ουσιαστικά την ροή. Αυτό επιτρέπει την διατήρηση των πιθανών χαμένων συσχετίσεων στην ροή και την πρόσθεση στρωμάτων με ελάχιστα υπολογιστικά κόστη, πράγμα απαραίτητο στο συγκεκριμένο δίκτυο αφού στο προηγούμενο βήμα εισάγονται πολλοί παράμετροι προς εκπαίδευση.



(a) Residual learning

Εικόνα 29: Σχηματική αναπαράσταση της δομής residual learning

Το αποτέλεσμα είναι η αύξηση της ακρίβειας σε σχέση με τα απλά συνελικτικά στρώματα. Στην παρακάτω εικόνα φαίνονται τα πειράματα που έγιναν ανάμεσα στις 2 αυτές περιπτώσεις σε γνωστά datasets. Προκύπτει το συμπέρασμα πως το residual learning καταπολεμά την ανάγκη για παραπάνω δεδομένα όταν ο αριθμός των παραμέτρων είναι υψηλός.

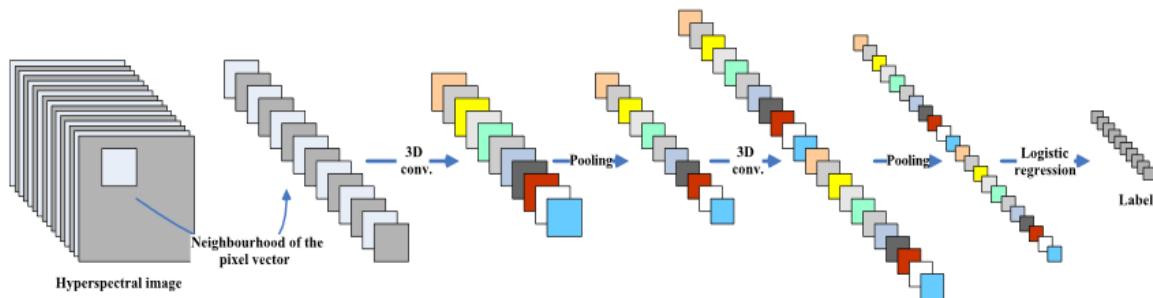


Εικόνα 30: Συγκριτικά γραφήματα εκπαίδευσης με-χωρίς την δομή residual learning

3.3 CHEN ET AL.

Το δίκτυο Chen et al.[\[3\]](#) είναι ένα CNN δίκτυο το οποίο δημιουργήθηκε για την ταξινόμηση φασματικών εικόνων. Όπως τα περισσότερα CNN δίκτυα, έτσι και εδώ γίνεται η προσπάθεια να εκπαιδευθεί από την είσοδο στις χωρικές-φασματικές σχέσεις των δεδομένων ταυτόχρονα. Απαρτίζεται από 3 συνελικτικά στρώματα καθώς και από στρώματα pool και dropout που παρεμβάλλονται

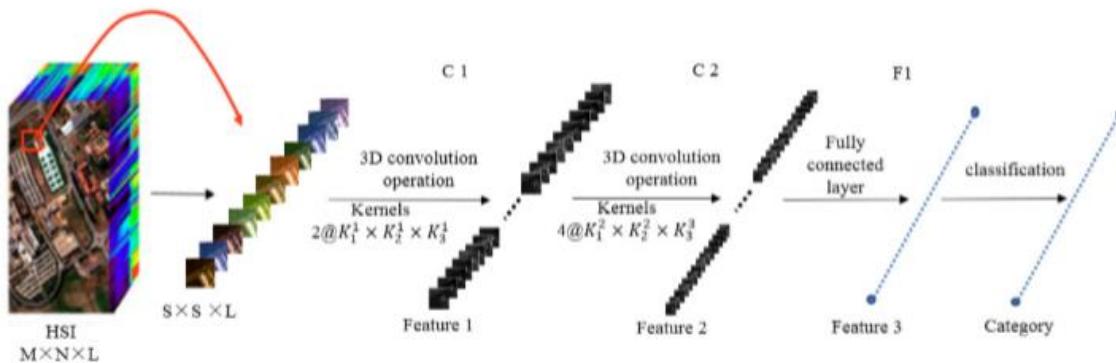
Τα συνελικτικά στρώματα του δικτύου εισάγουν πολλούς παραμέτρους προς μάθηση με επόμενο τον κίνδυνο της υπερπροσαρμογής. Τα 2 στρώματα dropout θα καταπολεμήσουν το φαινόμενο αυτό. Ενδιαφέρουσα είναι η ασυνήθιστη συγκριτικά επιλογή των μεγέθους της γειτονιάς των εικονοστοιχείων στην είσοδο, 27x27 εικονοστοιχεία.



Εικόνα 31:Η αρχιτεκτονική του δικτύου Chen et al.

3.4 LI ET AL.

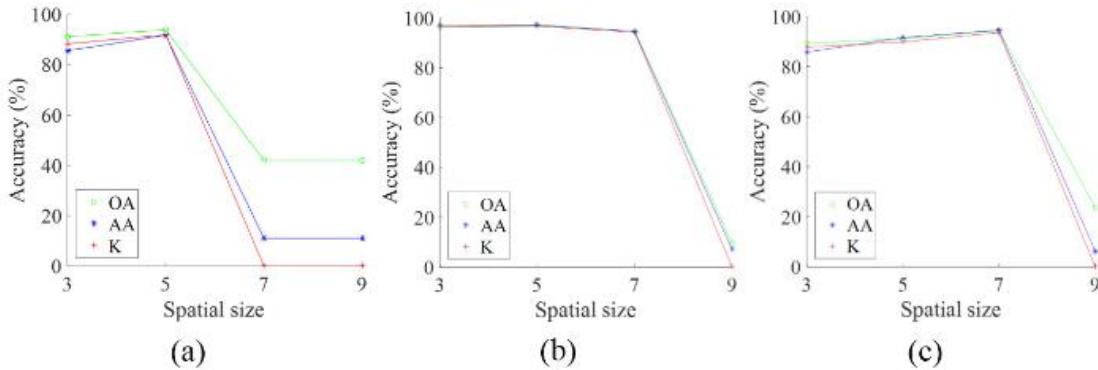
Το δίκτυο Li et al.[4] είναι ένα δίκτυο CNN που αναφέρεται στην ταξινόμηση υπερφασματικών εικόνων. Είναι ένα απλό δίκτυο που αποτελείται από 2 συνελικτικά στρώματα καταλήγοντας σε ένα FCN για το τελικό αποτέλεσμα.



Εικόνα 32:Η αρχιτεκτονική του δικτύου Li et al.

Είναι ένα σχετικά μικρό δίκτυο σε μέγεθος που καταφέρνει καλές ακρίβειες λόγω της αποδοτικής αρχιτεκτονικής του. Οι παράμετροι των στρωμάτων δεν είναι πολλοί σε σχέση με άλλα δίκτυα, όπως και ο αριθμός των φίλτρων στα στρώματα που είναι ιδιαίτερα μικρός. Ο αριθμός των στρωμάτων, των αριθμών των φίλτρων αλλά και των βαρών τους έγινε βάση πειραμάτων και γενικών συμπερασμάτων των δημιουργών. Η επιλογές τους στηρίζονται στην γρήγορη μάθηση και στην αποφυγή της υπερπροσαρμογής.

Σχετικά με την γειτονιά των εικονοστοιχείο στην είσοδο προκύπτει πως το ιδανικό μέγεθος είναι τα 5x5 εικονοστοιχεία. Τα πειράματα αναφέρονται στα γνωστά datasets(a-PaviaU,b-Botswana,c-Indian Pines)



Εικόνα 33:Γραφήματα προσδιορισμού της ακρίβειας συγκρίνοντας το μέγεθος της γειτονιάς των εικονοστοιχείων στην εισαγωγή του δικτύου

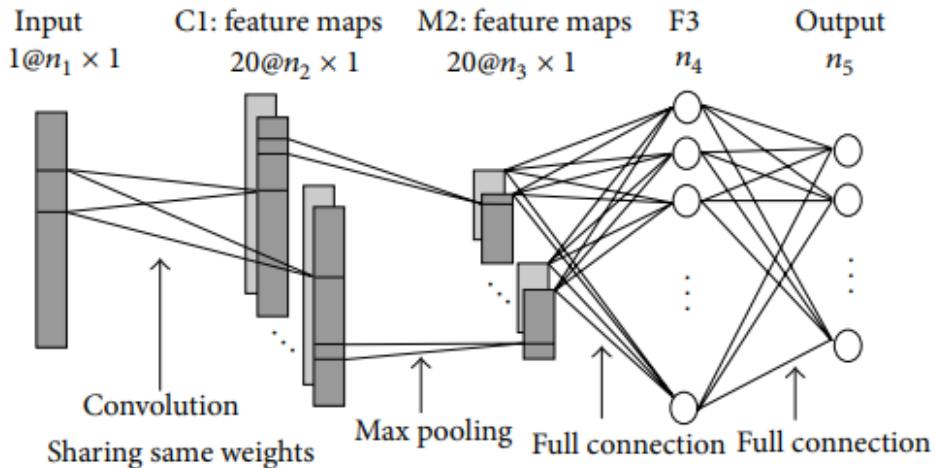
Ως αναφορά των αριθμών φίλτρων στα συνελικτικά στρώματα θεωρήθηκε ότι ένας μικρός αριθμός φίλτρων θα μειώσει τις παραμέτρους προς μάθηση και κατά συνέπεια θα αυξήσει την ακρίβεια του δικτύου μειώνοντας την πιθανότητα υπερπροσαρμογής. Έτσι επιλέγονται 2 φίλτρα για το πρώτο συνελικτικό στρώμα και 8 για το δεύτερο για το dataset Indian Pines. Αναλυτικότερα:

Layer	Kernel Size	Kernel Number	Stride	Output Size	Feature Volumes
Input	-	-	-	$5 \times 5 \times 200$	1
C1	$3 \times 3 \times 7$	2	1	$3 \times 3 \times 194$	2
C2	$3 \times 3 \times 3$	4	1	$1 \times 1 \times 192$	8
F1	-	-	-	$1 \times 1 \times 1$	128
Classification	-	-	-	$1 \times 1 \times 1$	16

Εικόνα 34:Η αρχιτεκτονική του δικτύου για το Indian Pines dataset.

3.5 HU ET AL.

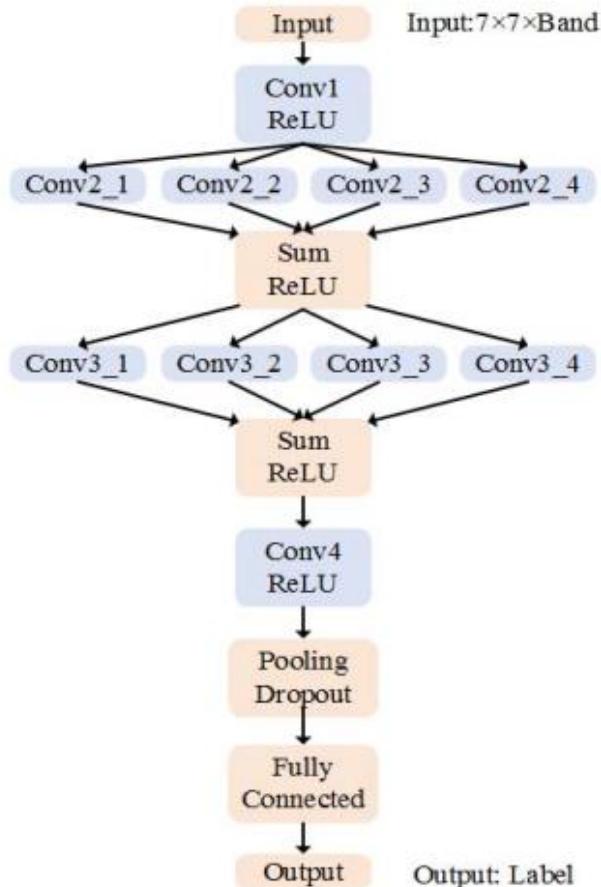
Το δίκτυο Hu et al.[\[5\]](#) είναι ένα δίκτυο CNN με στόχο την ταξινόμηση υπερφασματικών εικόνων. Είναι ένα ελαφρύ δίκτυο δίχως πολλά στρώματα ή πολλές παραμέτρους προς εκπαίδευση. Λαμβάνει υπόψιν τα φασματικά χαρακτηριστικά του κάθε εικονοστοιχείο στην εικόνα και εκπαιδεύεται σε αυτά ώστε να προβλέπει σε ποια κατηγορία ανήκουν. Σε αντίθεση με άλλα CNN δίκτυα το συγκεκριμένο εξετάζει κάθε εικονοστοιχείο ξεχωριστά και δεν ασχολείται με γειτονιές εικονοστοιχείο και τις χωρικές σχέσεις μεταξύ τους. Ακόμα στο συγκεκριμένο δίκτυο γίνεται η χρήση της υπερβολικής εφαπτομένης σαν συνάρτηση ενεργοποίησης των στρωμάτων C1 και F3.



Εικόνα 35:Η αρχιτεκτονική του δικτύου Hu et al.

3.6 HE ET AL.

Το δίκτυο He et al.[6] είναι ένα δίκτυο CNN με σκοπό την ταξινόμηση των υπερφασματικών εικόνων. Είναι ένα δίκτυο που αξιοποιεί από την αρχή τους φασματικούς αλλά και τους χωρικούς συσχετισμούς μεταξύ των εικονοστοιχείων. Χρησιμοποιεί μία ομάδα συνελίξεων που ονομάζει Multi scale 3D Convolution Block η οποία εκπαιδεύεται σε βάθος πάνω σε αυτούς τους συσχετισμούς, πριν καταλήξει σε ένα άλλο FCN δίκτυο για την τελική πρόβλεψη.



Εικόνα 36:Η αρχιτεκτονική του δικτύου He et al.

Ουσιαστικά το Multi scale 3D Convolution Block είναι μια τεχνική οι οποία ανακαλύπτει σε βάθος τις συσχετίσεις των δεδομένων στο φασματικό επίπεδο. Η επαναλαμβανόμενη χρησιμοποίηση της σε σειρά χαρίζει αποτελεσματικότητα στο δίκτυο. Επίσης πειραματικά στα δημοφιλή datasets αποδείχθηκε πως η τεχνική αυτή υπερτερεί στην ακρίβεια ενός απλού συνελικτικού 3D στρώματος.

Dataset	3D-DCNN	M3D-DCNN
Indian Pines	95.45%	97.61%
Pavia Univ	98.04%	98.49%
Salinas	96.72%	97.24%

Εικόνα 37: Συγκριτικός πίνακας ακρίβειας χρήσης-μη χρήσης Multi scale 3D Convolution Block

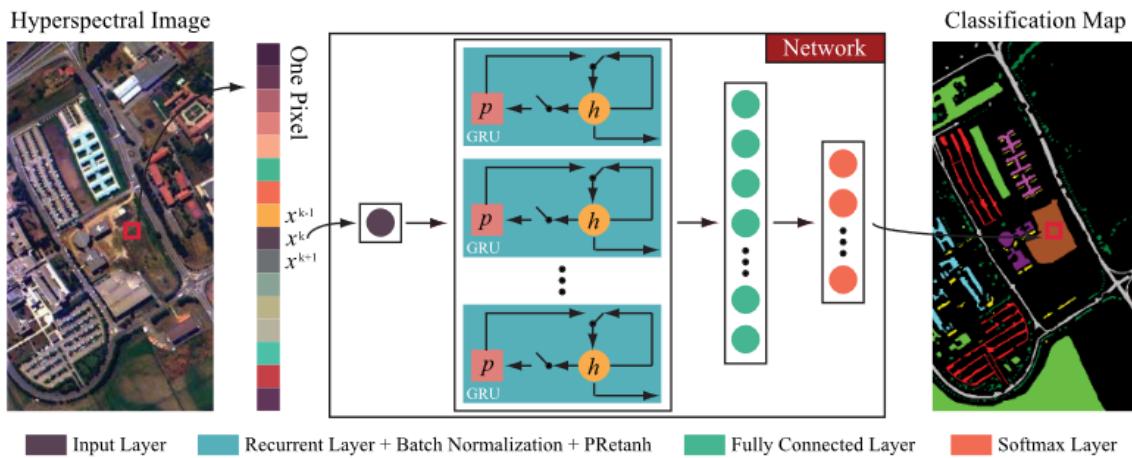
Αναλυτικότερα τα στρώματα του δικτύου:

kernel name	kernel number	kernel size H, W, B	kernel stride $\Delta(H, W, B)$
conv1	16	3,3,11	1,1,3
conv2_1		1,1,1	
conv2_2		1,1,3	
conv2_3	16	1,1,5	1,1,1
conv2_4		1,1,11	
conv3_1		1,1,1	
conv3_2		1,1,3	
conv3_3	16	1,1,5	1,1,1
conv3_4		1,1,11	
conv4	16	2,2,3	1,1,1
pooling	–	2,2,3	2,2,3

Εικόνα 38: Τα στρώματα που απαρτίζουν το δίκτυο He et al.

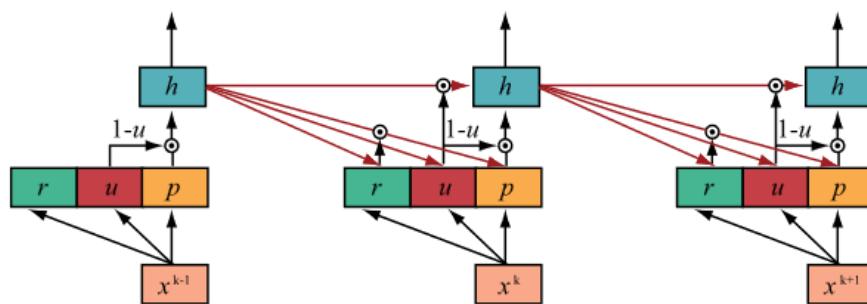
3.7 MOU ET AL.

Το δίκτυο Mou et al.^[7] είναι ένα δίκτυο RNN που σχεδιάστηκε για την ταξινόμηση υπερφασματικών εικόνων. Προσεγγίζει κάθε εικονοστοιχείο ξεχωριστά χωρίς να λαμβάνει υπόψη την γειτονιά του όπως άλλα δίκτυα. Εξετάζει την φασματική πληροφορία κατά σειρά, αναγνωρίζοντας όλες τις σχέσεις μεταξύ των φασματικών καναλιών του κάθε εικονοστοιχείο. Χρησιμοποιεί το στρώμα GRU, ένα στρώμα πιο εύχρηστο σε σχέση με τα στρώματα LSTM που χρησιμοποιούνται συνήθως στα RNN δίκτυα. Τέλος εισάγει μία νέα Συνάρτηση ενεργοποίησης, την PRTanh.



Εικόνα 39:Η αρχιτεκτονική του δικτύου Mou et al.

Το δίκτυο στην εισαγωγή δέχεται το εικονοστοιχείο με την φασματική πληροφορία των 176 καναλιών. Έπειτα γίνεται το πέρασμα από το στρώμα GRU.



Εικόνα 40:Λειτουργία του στρώματος GRU

Το στρώμα αυτό αποτελεί το βασικό δομικό στοιχείο της αρχιτεκτονικής αυτού του RNN δικτύου. Το δίκτυο εξετάζει το κάθε φασματικό κανάλι, μαθαίνει την σχέση μεταξύ των φασματικών τιμών και της κατηγορίας που ανήκει το εικονοστοιχείο. Η διαδικασία της μάθησης αυτής όμως δεν αντιμετωπίζει το κάθε φασματικό κανάλι ξεχωριστά, αλλά κατά σειρά. Δηλαδή η μνήμη του δικτύου συγκρατεί την σχέση μεταξύ τιμής-κατηγορίας του προηγούμενου καναλιού και αυτή η σχέση με την σειρά της επηρεάζει την διαδικασία μάθησης στο επόμενο κανάλι.

Έτσι προκύπτουν όλοι οι συσχετισμοί μεταξύ των φασματικών καναλιών μεταξύ τους και εξετάζεται ουσιαστικά η φασματική υπογραφή του κάθε εικονοστοιχείο ενιαία και όχι σε κοινά. Το στρώμα GRU διαθέτει 2 επιπρόσθετες πόρτες ενεργοποίησης που του επιτρέπουν τις παραπάνω ιδιότητες. Αυτές είναι υπεύθυνες για το τι θα κρατήσει στη μνήμη το στρώμα από τα προηγούμενα περάσματα δεδομένων ώστε να επηρεάσει την νέα έξοδο καθώς και στο τι θα πετάξει από την μνήμη ώστε να μην την επηρεάσει.

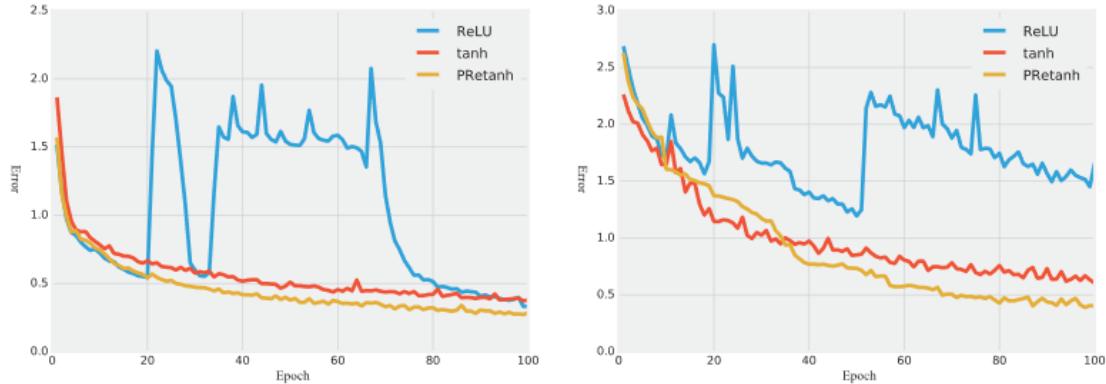
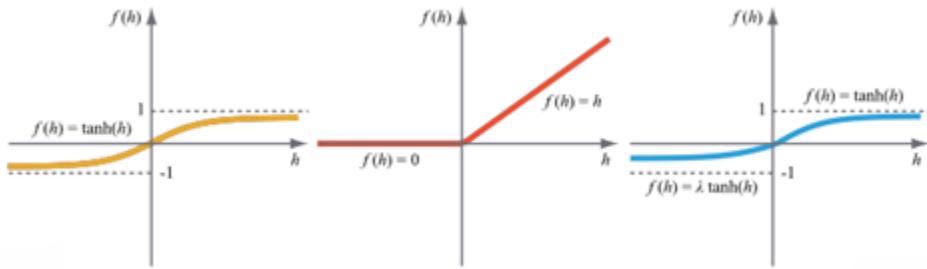
Το στρώμα GRU προτιμήθηκε από τους δημιουργούς σε σχέση με το στρώμα LSTM, διότι λύνει το πρόβλημα ‘‘vanishing gradient’’. Κατά την διάρκεια του οπισθοδιάδοσης ο υπολογισμός της παραγώγου γίνεται ολοένα και πιο δύσκολος, αφού το στρώμα LSTM κρατά στη μνήμη δεδομένα από κάθε κανάλι που περνάει. Όσο αυξάνει το βάθος της μνήμης η παράγωγος τείνει να εξαφανιστεί και δεν είναι υπολογίσιμη, με αποτέλεσμα την αδυναμία μάθησης του δικτύου. Το στρώμα GRU όπως περνά από το κάθε κανάλι διαλέγει κάθε φορά τι θα αφήσει και τι θα αποβάλει από την μνήμη του. Η εξυπνότερη αυτή διαχείριση επιτρέπει στο GRU να απαιτεί λιγότερους παραμέτρους προς μάθηση και κατά συνέπεια λιγότερα υπολογιστικά κόστη. Στην περίπτωση της ταξινόμησης υπερφασματικών δεδομένων που υπάρχουν σχετικά μικρά σε αριθμό δεδομένα το στρώμα GRU χαρίζει ταχύτητα και αποφυγή της υπερπροσαρμογής.

Η συνηθέστερη συνάρτηση είναι η συνάρτηση ReLU. Οι δημιουργοί του συγκεκριμένου δικτύου εισάγουν μία νέα συνάρτηση, την PRtanh.

$$f(h_i) = \begin{cases} \tanh(h_i), & \text{if } h_i > 0 \\ \lambda_i \tanh(h_i), & \text{if } h_i \leq 0 \end{cases}$$

Όπου λ παράμετρος προς μάθηση.

Τα πειράματα των δημιουργών έδειξαν ότι η ReLU εμφάνιζε αριθμητικά προβλήματα. Επίσης αποτελεί μια γρήγορη συνάρτηση, αφού μετατρέπει όλες τις αρνητικές τιμές σε 0, προσφέροντας ταχύτητα και εξομαλύνοντας το πρόβλημα ‘‘vanishing gradient’’. Η χρήση του στρώματος GRU όμως ήδη χαρίζει αυτές τις ιδιότητες στο δίκτυο. Έτσι αποφεύγοντας την ReLU το δίκτυο κερδίζει τις ιδιότητες μιας συνάρτησης με όρια όπως της tanh. Η πρόσθεση της παραμέτρου λ στις αρνητικές τιμές χαρίζει μεγαλύτερη διασπορά στο εύρος των τιμών διατηρώντας ένα όριο στην εξέλιξη τους.

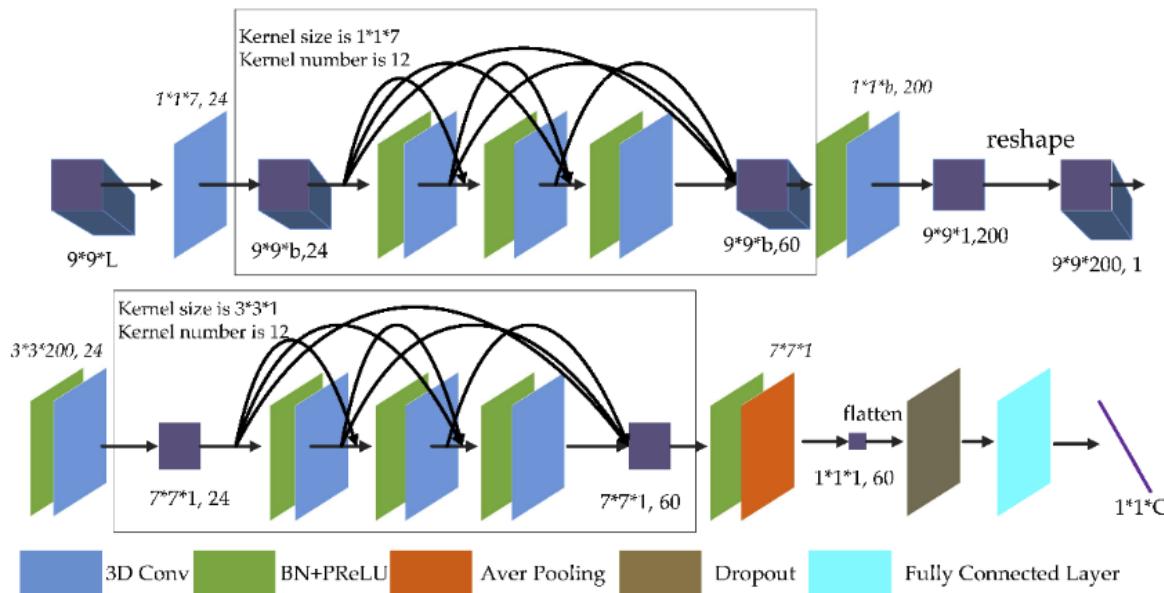


Εικόνες 41,42,43: Γραφήματα παρουσίασης και σύγκρισης της PRtanh με ReLU,tanh

Η χρήση της συνάρτησης PRtanh συνδυάζεται με την ομαλοποίηση των δεδομένων προκειμένου να μην παραμορφωθούν αρκετές αρνητικές τιμές. Στα παραπάνω πειράματα που πραγματοποίησαν οι δημιουργοί σε γνωστό dataset (PaviaU) με υψηλό learning rate της τάξης του 1.0 φαίνεται η σταθερότητα και η αποτελεσματικότητα της PRtanh σε σχέση με την ReLU και την απλή tanh.

3.8 FDSSC

Το δίκτυο FDSSC [8] είναι ένα CNN δίκτυο που δημιουργήθηκε για την ταξινόμηση υπερφασματικών δεδομένων. Είναι ένα μεγάλο δίκτυο, με πολλά στρώματα και παραμέτρους προς μάθηση. Εξετάζει την εικόνα παίρνοντας γειτονιές εικονοστοιχείο και αντλώντας φασματικές και χωρικές πληροφορίες από αυτές. Καθώς η γειτονιά περνά από τα φίλτρα μειώνονται σταδιακά οι διαστάσεις της μέχρι να καταλήξει σε ένα δίκτυο FCN για την τελική ταξινόμηση του εικονοστοιχείο. Δομικό στοιχείο του αποτελεί το Dense Spectral block και Dense Spatial block, μία ομάδα συνελικτικών στρωμάτων υπεύθυνων για την εξαγωγή των φασματικών και χωρικών σχέσεων της εικόνας.



Εικόνα 44:Η αρχιτεκτονική του δικτύου FDDSC.

Το Dense Spectral block έχει επιρροές από την τεχνική residual learning. αποτελείται από διάφορα συνελικτικά στρώματα που η έξοδος τους ταυτόχρονα περνάει ως είσοδος σε επόμενα αλλά και ως έξοδος στο τελικό στάδιο του block,όπου όλοι οι έξοδοι από τα στρώματα ενώνονται. Οπότε από την αρχή του δικτύου εξερευνιούνται σε βάθος οι φασματικές πληροφορίες της εισαγωγής. Τα στρώματα αυτά είναι παρόμοια με το πρώτο. Η διάταξη και η ροή των δεδομένων σε αυτά είναι διαφορετική γ αυτό αντιμετωπίζονται σαν Dense Spectral block.Οι διαστάσεις της εισαγωγής στην ροή του block είναι σταθερές με την χρήση κατάλληλου padding στα φίλτρα.

Το συγκεκριμένο δίκτυο είναι αρκετά μεγάλο. Συγκριτικά με τα συνήθη δεδομένα για εκπαίδευση ο αριθμός των παράμετρών αυτών είναι πολλαπλάσιος, άρα και επικίνδυνος από την άποψη του υπολογιστικού κόστους και της υπερπροσαρμογής. Έτσι χρησιμοποιούνται διάφορες τεχνικές για την αντιμετώπιση των προβλημάτων αυτών.

Αρχικά γίνεται η χρήση την συνάρτησης ενεργοποίησης ReLU. Αυτή η συνάρτηση χαρίζει ταχύτητα στο δίκτυο σε σχέση με την απλή ReLU αφού μέσω αυτής το δίκτυο μαθαίνει πιο γρήγορα. Σε συνδυασμό με την ομαλοποίηση των δεδομένων πριν από κάθε φίλτρο, το δίκτυο έχει λιγότερα υπολογιστικά κόστη. Ακόμα το στρώμα dropout ενισχύει την αποφυγή της υπερπροσαρμογής.

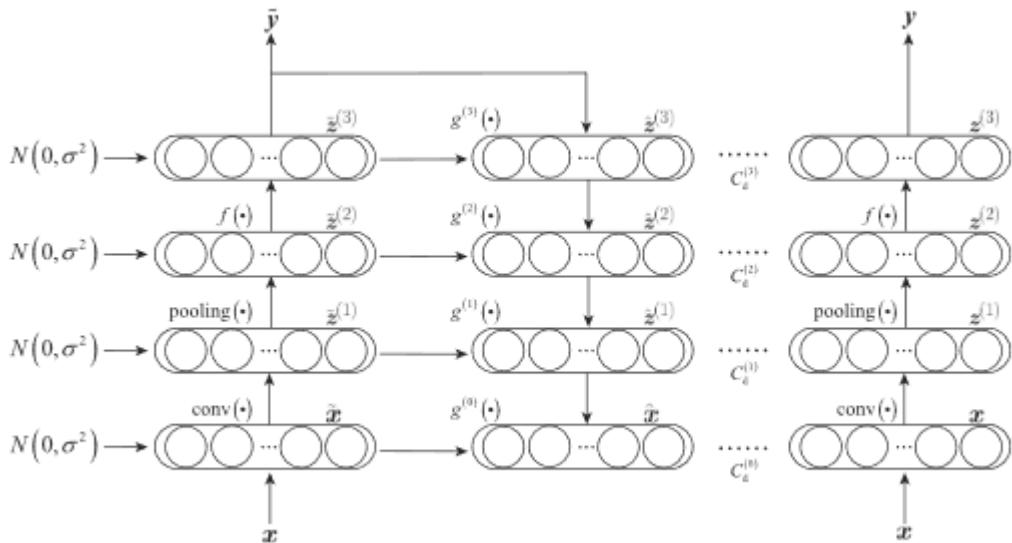
Τέλος η δομή και η φιλοσοφία των 2 block του δικτύου χαρίζουν αποτελεσματικότητα και ταχύτητα στο δίκτυο. Όπως φάνηκε και σε προηγούμενη μέθοδο, η χρήση του residual learning και των skip connection μεταξύ των στρωμάτων επιφέρει καλύτερες ακρίβειες, μειώνει την υπερπροσαρμογή και επιτρέπει τον ταχύτερο υπολογισμό της παραγώγου σε σχέση με αντίστοιχες ομάδες στρωμάτων.

3.9 LIU ET AL.

Το δίκτυο Liu et al. [9] είναι ένα νευρωνικό δίκτυο με σύνθετη αρχιτεκτονική. Αποτελείται από 2 βασικά δομικά μέρη. Στο πρώτο μέρος υπάρχει ένα τυπικό 3D CNN δίκτυο και στο δεύτερο ένα δίκτυο της μορφής Κωδικοποιητή-Αποκωδικοποιητή ώστε να ενισχυθεί η ακρίβεια του πρώτου. Τα CNN δίκτυα είναι γνωστό πως προσφέρουν αρκετά υψηλές ακρίβειες ιδιαιτέρως όταν τα δεδομένα στα οποία εκπαιδεύονται είναι μεγάλα σε αριθμό. Στην ταξινόμηση φασματικών εικόνων πολλές φορές συναντάται το πρόβλημα του μικρού αριθμού δεδομένων εκπαίδευσης και κατά συνέπεια πολλά εικονοστοιχεία αχαρακτήριστα ως προς την κατηγορία τους μένουν ανεκμετάλλευτα. Το πρόβλημα αυτό προσπαθεί να λύσει το δίκτυο Liu σαν ένα semi-supervised δίκτυο.

Semi-supervised δίκτυο είναι ένα δίκτυο που συνδυάζει τις supervised-unsupervised μεθόδους. Εκμεταλλεύεται ταυτόχρονα τα δεδομένα προς εκπαίδευση που έχει υποδείξει ο χρήστης, και τα δεδομένα που απομένουν και δεν έχουν κατηγοριοποιηθεί από αυτόν. Έτσι το δίκτυο εκπαιδεύεται σε μεγαλύτερο αριθμό δεδομένων και ο συνδυασμός των μεθόδων αυτών οδηγεί σε υψηλές ακρίβειες.

Η αρχιτεκτονική του δικτύου περιλαμβάνει στα αριστερά τον «βρόμικο» κωδικοποιητή-αποκωδικοποιητή, ο οποίος εισάγει θόρυβο στα δεδομένα και στη συνέχεια πραγματοποιεί πρόβλεψη σε αυτά. Έπειτα τα επαναφέρει στην αρχική τους κατάσταση βοηθούμενος από την χρήση residual learning. Στα δεξιά ο «καθαρός» κωδικοποιητής δίνει πρόβλεψη στα δεδομένα βοηθούμενος στην εκπαίδευση από τον «βρόμικο αποκωδικοποιητή»



Εικόνα 45: Η αρχιτεκτονική του δικτύου Liu et al.

4 ΜΕΘΟΔΟΓΙΑ

4.1 ΔΕΔΟΜΕΝΑ

Τα δεδομένα στα οποία θα χρησιμοποιηθούν οι μέθοδοι Βαθιάς Μάθησης προέρχονται από το HyRANK dataset [54]. Σκοπός του dataset αυτού είναι η διάθεση ενός υπερφασματικού dataset προς το κοινό, ώστε να δοθεί η ευκαιρία στην κοινότητα να τεστάρει και να συγκρίνει νέες μεθόδους ταξινόμησης στα υπερφασματικά δεδομένα.

Τα ήδη διαθέσιμα γνωστά dataset (Indian Pines,PaviaU,κλπ.) χαρακτηρίζονται μικρά σε μέγεθος. Συγκεκριμένα το dataset Indian Pines έχει μέγεθος 145x145 εικονοστοιχεία με 16 κλάσεις και το PaviaU 610x340 εικονοστοιχεία με 9 κλάσεις αντίστοιχα. Αυτοί οι μικροί σχετικά αριθμοί σε μέγεθος και κλάσεις μειώνονται ακόμα περισσότερο κατά την διαδικασία της ταξινόμησης, αφού οι εικόνες χωρίζονται σε σετ εκπαίδευσης-επαλήθευσης-τεσταρίσματος. Έτσι προκύπτουν πολύ λίγα χαρακτηρισμένα εικονοστοιχείο για τα σετ αυτά.

Το γεγονός ότι τα σύγχρονα δίκτυα για την ταξινόμηση υπερφασματικών εικόνων χρησιμοποιούν 3D προσεγγίσεις, αναγκάζει πολλές φορές να χρησιμοποιηθούν τα ίδια γειτονικά εικονοστοιχεία για εκπαίδευση ή για επαλήθευση-τεσταρίσμα. Ο μικρός αριθμός των εικονοστοιχείων ενισχύει το γεγονός αυτό. Ακόμα, η περίπτωση της εκπαίδευσης-επαλήθευσης-τεσταρίσματος πάνω στην ίδια εικόνα εξαιρεί από τις προκλήσεις της ταξινόμησης την παράμετρο της επιρροής της ατμόσφαιρας στις φασματικές τιμές της εικόνας.

Το HyRANK dataset δημιουργήθηκε απαντώντας στα παραπάνω προβλήματα που διέπουν τα γνωστά datasets. Αποτελείται από 5 εικόνες οι οποίες προέρχονται από τον αισθητήρα Hyperion, από τον δορυφόρο EO-1 της NASA. Οι εικόνες αναφέρονται στην χρονική στιγμή μεταξύ Μαΐου-Ιουλίου 2006. Περιέχουν 176 φασματικές κατηγορίες και 14 κλάσεις χρήσεων γης, στα πρότυπα του CORINE Land Cover. Οι ονομασίες των εικόνων είναι Dioni, Loukia, Erato, Kirkī, Nefeli. Από αυτές τις εικόνες οι πρώτες 2, δηλαδή η Dioni και η Loukia θα χρησιμοποιηθούν για την εκπαίδευση των μοντέλων και οι υπόλοιπες για το τεσταρίσμα αυτών.

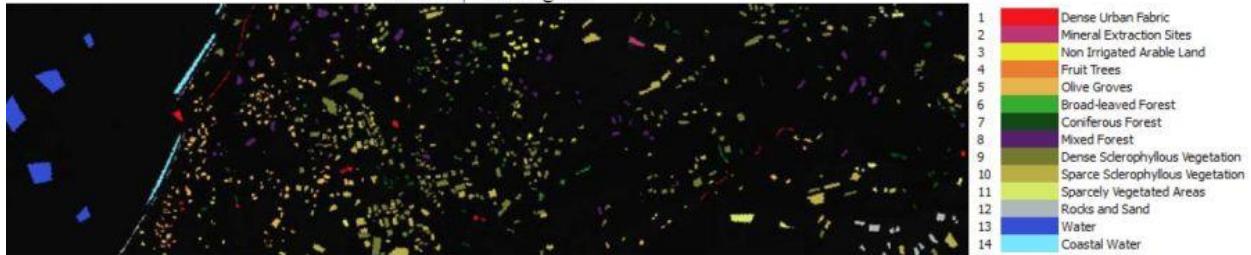
Λεπτομερέστερα τα μεγέθη των εικόνων και οι κλάσεις :

The Five Hyperion Images of HyRANK					
id	Name	Image Size	Spectral Bands	Spatial Resolution	Training/Validation
1	Dioni	250 x 1376 x 176	176	30m	<i>training</i>
2	Loukia	249 x 945 x 176	176	30m	<i>training</i>
3	Erato	241 x 1632 x 176	176	30m	<i>validation</i>
4	Nefeli	249 x 772 x 176	176	30m	<i>validation</i>
5	Kirkī	245 x 1626 x 176	176	30m	<i>validation</i>

Loukia
RGB – R:23 G:11 B:07



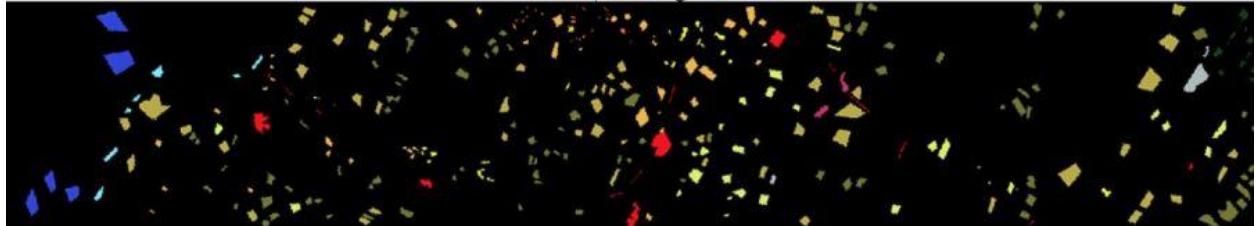
Loukia | Training Set

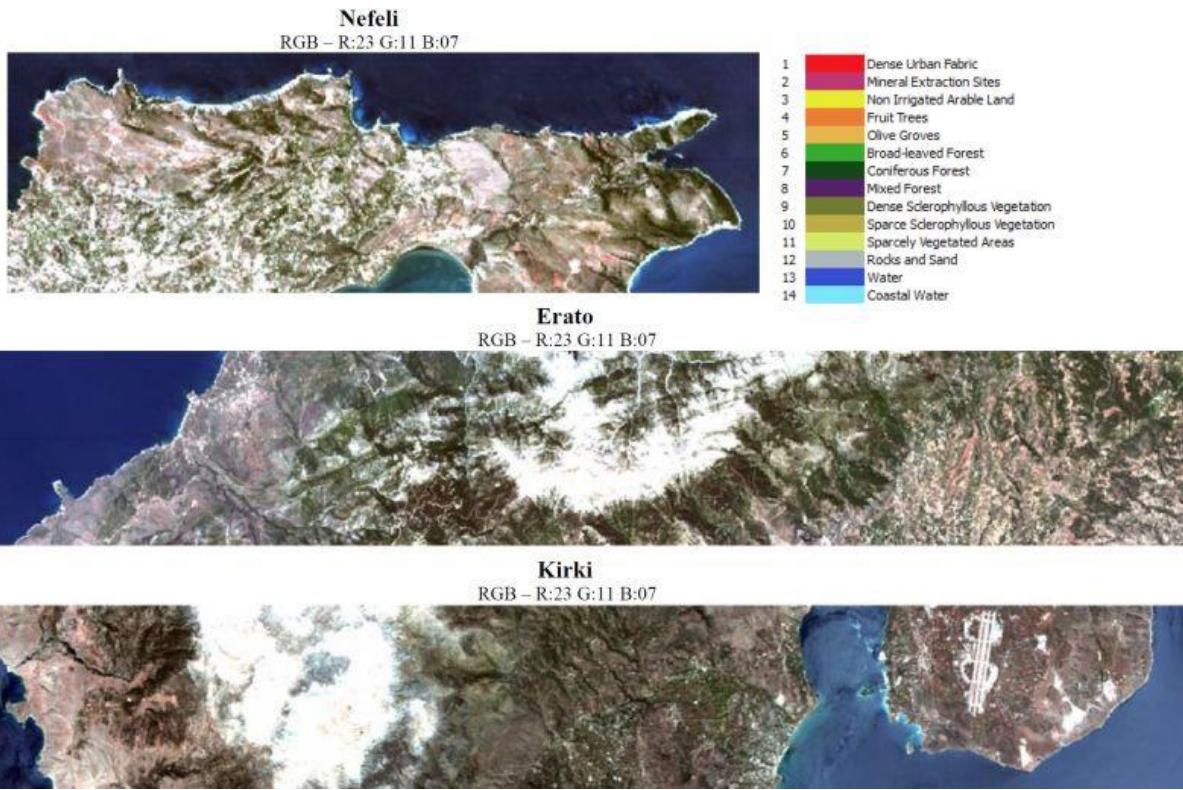


Dioni
RGB – R:23 G:11 B:07



Dioni | Training Set





Εικόνες 46,47,48: To HyRANK dataset
https://www.isprs.org/society/si/SI-2017/ISPRS-SI2017-TC3_WG4_Karantzalos_Report.pdf

4.2 ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

Σε όλα τα πειράματα οι φασματικές τιμές των εικονοστοιχείων των εικόνων θα μεταφερθούν στο διάστημα [0,1].Χρησιμοποιώντας την σχέση:

$$I = \frac{I - \min(I)}{\max(I) - \min(I)}$$

εφαρμοσμένη στο σύνολο των εικόνων, τα εύρη των τιμών θα κυμαίνονται μεταξύ του διαστήματος [0,1].Η διαδικασία της κανονικοποίησης είναι μία απαραίτητη διαδικασία που εφαρμόζεται στα δεδομένα πριν την εισαγωγή τους στα νευρωνικά δίκτυα αλλά και σε άλλους τύπους αλγορίθμων ταξινόμησης.

Χαρακτηριστικά η μέθοδος SVM χρησιμοποιεί συναρτήσεις πυρήνα που βασίζονται στην απόσταση μεταξύ των τιμών των εικονοστοιχείο. Αν το εύρος των τιμών σε μία φασματική κατηγορία είναι αναντίστοιχο με τις υπόλοιπες, αυτή η κατηγορία θα κυριαρχήσει σε σχέση με τις άλλες στην εύρεση του ιδανικού υπερεπιπέδου για τον διαχωρισμό των κατηγοριών στον χώρο. Η κανονικοποίηση εγγυάται πως η διασπορά των τιμών της κάθε φασματικής κατηγορίας θα παίξει εξίσου σημαντικό ρόλο στον ταξινομητή.

Τα νευρωνικά δίκτυα είναι σχεδιασμένα με τέτοιον τρόπο που η τελική τους ακρίβεια στην ταξινόμηση δεν επηρεάζεται από την κανονικοποίηση, αφού έχουν την δυνατότητα να μαθαίνουν τις ιδιαιτερότητες που παρουσιάζονται στα δεδομένα που εισέρχονται σε αυτά. Όμως το άνοιγμα των τιμών στο διάστημα [0,1] έχει σαν αποτέλεσμα την ταχύτερη και σταθερότερη μάθηση του δικτύου, καθώς και επιφέρει μία σχετική σταθερότητα στις τιμές των ιδανικών υπερπαραμέτρων για κάθε δίκτυο.

Αυτό συμβαίνει γιατί το συγκεκριμένο διάστημα επιτρέπει την ομαλότερη λειτουργία του αλγορίθμου της οπισθοδιάδοσης.Η διαφορά στο εύρος των τιμών που εισέρχονται στο δίκτυο οδηγεί σε διαφορετικού μεγέθους ενημερώσεις των παραμέτρων προς εκπαίδευση. Έτσι το μοντέλο δεν πλησιάζει με σταθερό βήμα το ελάχιστο της συνάρτησης κόστους.

4.3 ΠΡΟΣΑΡΜΟΓΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ

Γενικότερα έγινε προσπάθεια να διατηρηθούν οι υπερπαράμετροι που δόθηκαν στην βιβλιογραφία του κάθε δικτύου. Η διαφορετικότητα των δεδομένων εκπαίδευσης σε σχέση με τα δεδομένα που χρησιμοποιήθηκαν στην βιβλιογραφία απαιτεί την προσαρμογή αυτών για την ομαλή εκπαίδευση των δικτύων. Το μέγεθος της γειτονίας των εικονοστοιχείο που θα οδηγηθούν στην εισαγωγή, ο τύπος της συνάρτησης κόστους, η μέθοδος βελτιστοποίησης διατηρούνται όπως έχουν.

Η δυο εικόνες του σετ εκπαίδευσης (Dionis, Loukia) θα χωριστούν σε δυο νέα σετ, το νέο σετ εκπαίδευσης και το σετ επαλήθευσης. Το χώρισμα αυτό πραγματοποιείται γενικότερα στις εφαρμογές Βαθιάς Μάθησης για την επιλογή των κατάλληλων υπερπαραμέτρων. Στην συγκεκριμένη εφαρμογή το χώρισμα αυτό θα εξυπηρετήσει την προσαρμογή του ρυθμού εκπαίδευσης, του μεγέθους του πακέτου στην εισαγωγή (batch size) αλλά και την παρουσίαση των γραφημάτων του δικτύου για να γίνει η επιλογή του αριθμού των εποχών ή αλλιώς τον χρόνο εκπαίδευσης του δικτύου. Μετέπειτα θα γίνει η τελική πρόβλεψη πάνω στις 3 εικόνες του σετ τεσταρίσματος. Η Το δίκτυο θα εκπαιδευτεί εκ νέου σε όλο το σετ εκπαίδευσης, δηλαδή το σύνολο των 2 εικόνων.



Εικόνα 49: Ο διαχωρισμός των δεδομένων

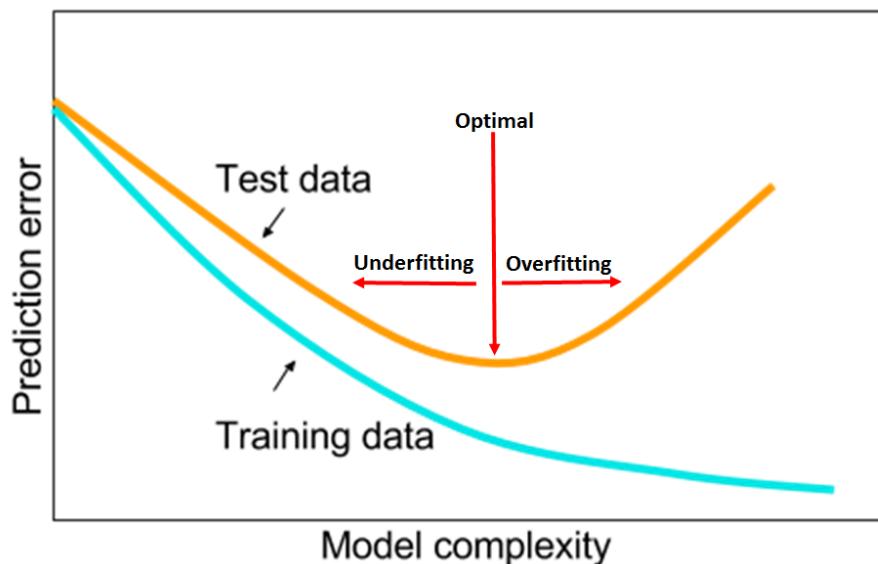
<https://www.ubuntupit.com/frequently-asked-machine-learning-interview-questions-and-answers/>

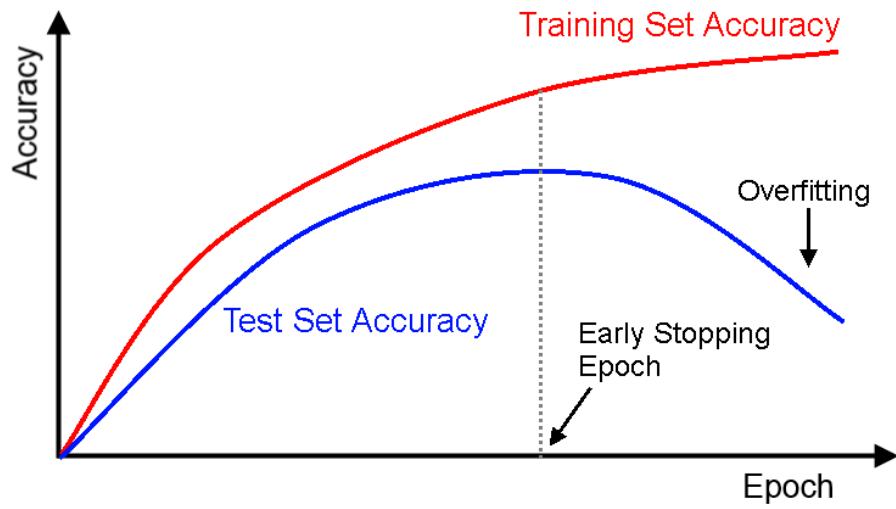
Η διαδικασία της διαλογής των δεδομένων των πειραμάτων σε δεδομένα εκπαίδευσης – επαλήθευσης και τεσταρίσματος είναι πάρα πολύ σημαντική. Η επιλογή των κατάλληλων υπερπαραμέτρων είναι άμεσα συνδεδεμένη με την αποτελεσματικότητα της κάθε μεθόδου, αφού καθορίζει την γενικότερη πορεία της μάθησης. Ένα βασικό πρόβλημα που παρουσιάζεται συχνά σε κάθε μέθοδο είναι η βεβιασμένη προσαρμογή των παραμέτρων με στόχο την υψηλότερη δυνατή ακρίβεια στο σετ τεσταρίσματος. Η πρακτική αυτή αποδεικνύεται λάθος. Η υπερπροσαρμογή αυτή οδηγεί το μοντέλο να μάθει περισσότερα πράγματα από όσα έχει σαν σκοπό, μαθαίνοντας κυρίως τον θόρυβο των δεδομένων. Οπότε πετυχαίνει μεγαλύτερη ακρίβεια στα συγκεκριμένα δεδομένα αλλά παράλληλα μειώνεται η ακρίβεια της γενίκευσης. Σε ένα νέο σετ δεδομένων τα αποτελέσματα δεν θα είναι τα αναμενόμενα με βάση το προηγούμενο σετ, και έτσι το μοντέλο θα είναι καλό στην εφαρμογή πάνω στα αρχικά δεδομένα και όχι σε γενικότερα δεδομένα.

Γι' αυτούς τους λόγους οι υπερπαράμετροι αλλά και ο χρόνος εκπαίδευσης του μοντέλου καθορίζονται παρατηρώντας την πορεία του μοντέλου μέσω των γραφημάτων πάνω στο σετ επαλήθευσης, το οποίο δημιουργείται παίρνοντας δεδομένα από το σετ εκπαίδευσης και όχι πάνω στο σετ τεσταρίσματος. Η. Τα δεδομένα του τεσταρίσματος παραμένουν κρυφά μέχρι την τελική πρόβλεψη.

Τα γραφήματα του δικτύου είναι δυο. Το πρώτο, το γράφημα κόστους, παρουσιάζει τις τιμές της συνάρτησης κόστους πάνω στο νέο σετ εκπαίδευσης και στο σετ επαλήθευσης. Καθώς εξελίσσεται η εκπαίδευση του δικτύου, η τιμές αυτές τείνουν να μειώνονται αφού η συνάρτηση κόστους ελαχιστοποιείται. Με παρόμοιο τρόπο λειτουργεί και το γράφημα της ακρίβειας. Αυτό παρουσιάζει την ακρίβεια των προβλέψεων ανά εποχή και στα δυο σετ αντίστοιχα.

Από τα γραφήματα αυτά μπορεί κανείς να παρατηρήσει την επιρροή του ρυθμού μάθησης στο μοντέλο. Η αργή μείωση της συνάρτησης κόστους φανερώνει έναν χαμηλό ρυθμό μάθησης. Η μεγάλη αυξομείωση της ακρίβειας ή και του κόστους φανερώνει έναν μεγάλο ρυθμό μάθησης με τον οποίο η ανανέωση των παραμέτρων έχει τόσο μεγάλο βήμα που δεν προσεγγίζεται η βέλτιστη λύση. Ακόμα η στιγμή όπου το κόστος των δεδομένων εκπαίδευσης συνεχίζει να ελαχιστοποιείται ενώ το αντίστοιχο κόστος των δεδομένων επαλήθευσης αυξάνει χαρακτηρίζει την αρχή της υπερπροσαρμογή του μοντέλου πάνω στα δεδομένα εκπαίδευσης. Τότε πρέπει να σταματά η εκπαίδευση του μοντέλου και να γίνεται η πρόβλεψη στα δεδομένα τεσταρίσματος.





Εικόνες 50,51 : Υπερπροσαρμογή του δικτύου μέσα από τα γραφήματα κόστους και ακρίβειας
<http://mmds-data.org/presentations/2016/s-martin.pdf>

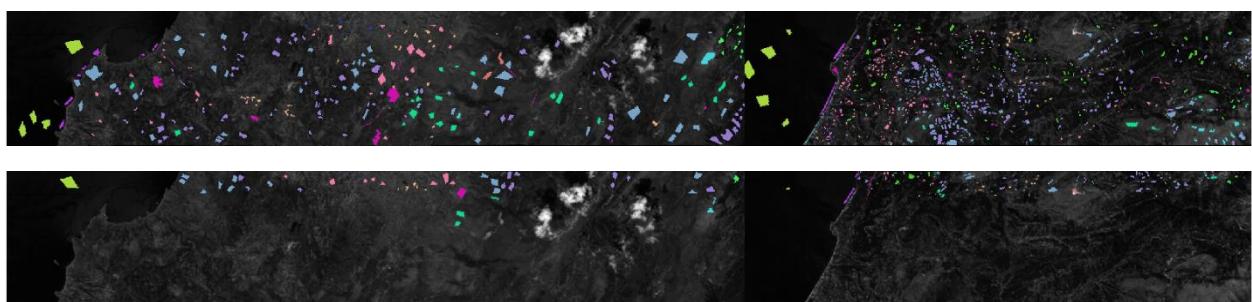
4.4 ΜΕΘΟΔΟΣ DISJOINT

Το χώρισμα αυτό θα γίνει με την μέθοδο disjoint[10]. Η μέθοδος αυτή χωρίζει τις περιοχές εκπαίδευσης με έναν τρόπο ρεαλιστικό για τους σκοπούς της επαλήθευσης. Κάθε χώρισμα αυτού του σκοπού πρέπει να γίνεται με ορισμένες προϋποθέσεις. Αρχικά θα πρέπει να διατηρείται μια ισορροπία σχετικά με την αντιπροσωπευτικότητα των κλάσεων στο κάθε σετ. Προφανώς η απουσία εικονοστοιχείο που αναφέρονται στο νερό στο σετ εκπαίδευσης δεν έχει νόημα αν πρέπει να προβλεφθούν σωστά εικονοστοιχείο νερού στο σετ επαλήθευσης, γιατί το μοντέλο δεν θα έχει εκπαιδευτεί πάνω σε αυτά για να προβλέψει.

Ένα άλλο βασικό θέμα που προκύπτει και η μέθοδος disjoint απαντά είναι η χωρική συσχέτιση των εικονοστοιχείων. Τα περισσότερα μοντέλα στην εφαρμογή που θα ακολουθήσει δέχονται ως εισαγωγή γειτονιές των εικονοστοιχείο και όχι μεμονωμένα εικονοστοιχεία. Έτσι αν έχουν επιλεγεί κάποια τυχαία εικονοστοιχείο για εκπαίδευση από μία συγκεκριμένη περιοχή εκπαίδευσης είναι πολύ πιθανό τα κοντινά εικονοστοιχείο που θα χρησιμοποιηθούν για επαλήθευση να έχουν ήδη χρησιμοποιηθεί και στην εκπαίδευση του μοντέλου. Η μέθοδος disjoint διαχωρίζει τις περιοχές εκπαίδευσης ανά οντότητα και όχι ανά εικονοστοιχείο. Οπότε για παράδειγμα από τις 10 περιοχές της εικόνας χαρακτηρισμένες ως νερό, θα χρησιμοποιηθούν οι 5 για την εκπαίδευση και οι υπόλοιπες 5 για την επαλήθευση του μοντέλου χωρίς να λαμβάνονται γειτονικά εικονοστοιχεία από την ίδια περιοχή, αντιπροσωπεύοντας ρεαλιστικότερα την συμπεριφορά του μοντέλου στο σετ τεσταρίσματος.

4.5 ΕΠΙΛΟΓΗ ΣΕΤ ΕΚΠΑΙΔΕΥΣΗΣ-ΕΠΑΛΗΘΕΥΣΗΣ

Έτσι για την αρχική εκπαίδευση και τον προσδιορισμό των υπερπαραμέτρων θα χρησιμοποιηθούν 28312 εικονοστοιχεία από τα 33527 συνολικά χαρακτηρισμένα εικονοστοιχείο των 2 εικόνων. Για το σετ επαλήθευσης θα χρησιμοποιηθούν τα υπόλοιπα 5215, δηλαδή το 5.4%. Το ποσοστό αυτό είναι σχετικά μικρό για ποσοστό σετ επαλήθευσης. Στην συγκεκριμένη εφαρμογή η τελική πρόβλεψη δίνεται από την εκπαίδευση στο σύνολο των 2 εικόνων Dioni, Loukia. Η επιλογή του ποσοστού αυτού βασίστηκε στο γεγονός ότι οι παράμετροι θα προσαρμοστούν σε ένα μοντέλο, το οποίο θα επανεκπαιδεύτει στη συνέχεια σε ένα νέο σετ με πρόσθετα δεδομένα. Συνεπώς το αρχικό σετ εκπαίδευσης με το τελικό σετ πρέπει να διαφέρουν ελάχιστα ώστε οι καθορισμένοι υπερπαράμετροι να είναι κατάλληλοι και για τα 2 σετ.



Εικόνες 52,53 : Το νέο σετ εκπαίδευσης και επαλήθευσης

4.6 ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ - ΣΤΑΤΙΣΤΙΚΑ

Προκειμένου να γίνει περισσότερο αντιληπτή η διαδικασία της ταξινόμησης των φασματικών εικόνων αλλά και η αποτελεσματικότητα των μεθόδων, χρησιμοποιούνται οι πίνακες σύγχυσης. Ένας πίνακας σύγχυσης περιέχει όλα τα στατιστικά στοιχεία τα οποία θα βοηθήσουν στην ερμηνεία τυχόν προβλημάτων της ταξινόμησης και κατά προέκταση στην επίλυση τους.

Αρχικά παράγεται ο πίνακας σύγχυσης ο οποίος περιέχει τόσες στήλες και γραμμές όσες και οι κλάσεις που ανήκουν τα εικονοστοιχείο που ταξινομούνται. Οι γραμμές συμβολίζουν τις κλάσεις που αναφέρονται στα πραγματικά εικονοστοιχείο των εικόνων, δηλαδή αυτά που έχουν χαρακτηριστεί από τον χρήστη. Οι στήλες συμβολίζουν τις κλάσεις που αναφέρονται στα ταξινομημένα εικονοστοιχείο της εικόνας, δηλαδή σε αυτά που χαρακτηρίστηκαν από το μοντέλο. Έτσι συγκρίνοντας πραγματικά και ταξινομημένα εικονοστοιχείο, προκύπτουν συμπεράσματα όπως ποια κλάση δεν ταξινομήθηκε ομαλά, σε ποιες κλάσεις κατέληξαν να ταξινομηθούν τα εικονοστοιχείο που δεν ταξινομήθηκαν ορθά κλπ. Συνδυάζοντας τις πληροφορίες αυτές με την φασματική υπογραφή της κάθε εικόνας-κατηγορίας-εικονοστοιχείο, αναλύεται σε βάθος η αποτελεσματικότητα της κάθε μεθόδου ταξινόμησης.

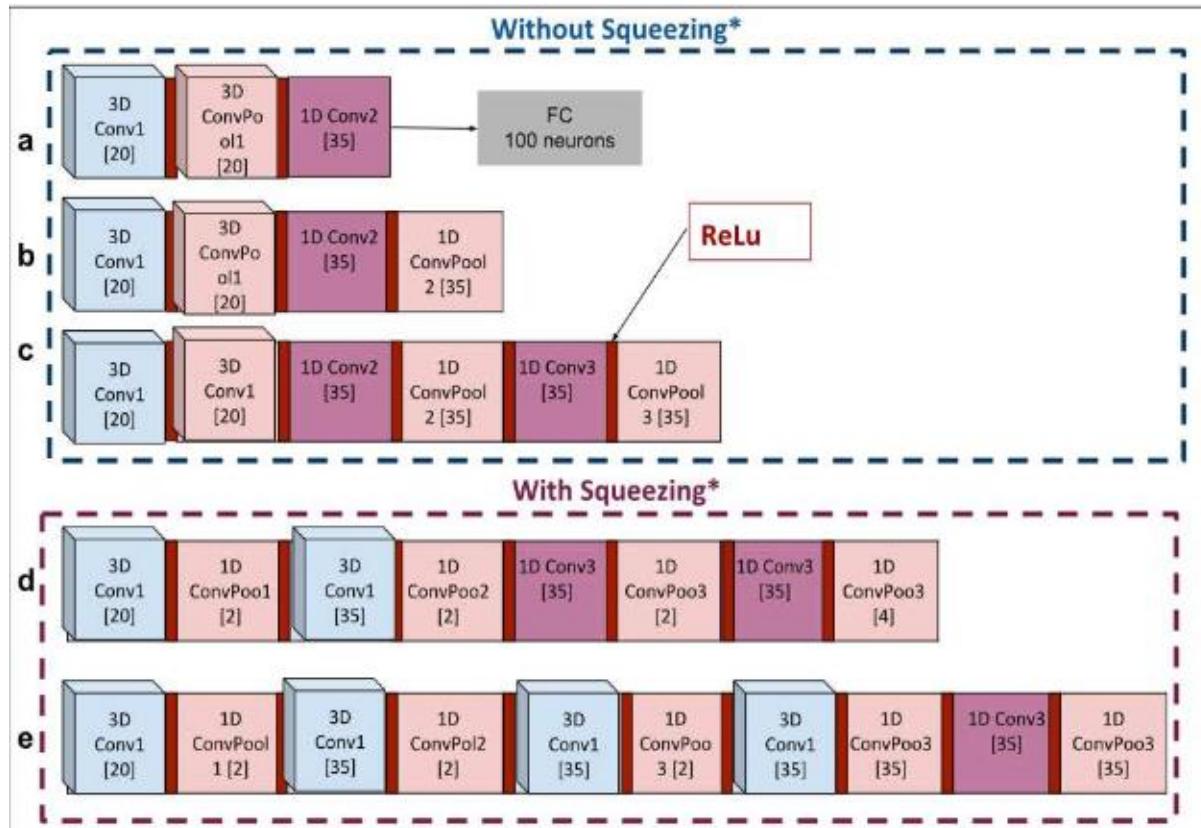
Καθώς προκύπτει η ανάγκη της σύγκρισης διαφορετικών μεθόδων μεταξύ τους, χρησιμοποιούνται κι άλλες μορφές πινάκων με γενικότερες πληροφορίες χωρίς τα λεπτομερή στοιχεία ανά κλάση. Αυτοί οι πίνακες περιέχουν τα εξής στοιχεία:

- Ολική ακρίβεια(OA), είναι το ποσοστό επί τις εκατό που εκφράζει τα εικονοστοιχεία που ταξινομήθηκαν ορθά σε σχέση με το σύνολο τους σε όλες τις κλάσεις.
- Ακρίβεια ανά κλάση, είναι η ακρίβεια που εκφράζει τα εικονοστοιχεία που ταξινομήθηκαν σωστά σε σχέση με το σύνολο των εικονοστοιχείων που ανήκουν στην κλάση αυτή μαζί με αυτά που ταξινομήθηκαν σε αυτή, ορθά ή μη ορθά.
- Ακρίβεια του δημιουργού ανά κλάση(PA), είναι η ακρίβεια από την πλευρά του ταξινομητή. Εκφράζει το ποσοστό επί τις εκατό ανάμεσα στα εικονοστοιχείο τα οποία ταξινομήθηκαν ορθά σε μία κλάση σε σχέση με το σύνολο των εικονοστοιχείων που ανήκουν στην κλάση αυτή.
- Ακρίβεια του χρήστη ανά κλάση(UA), είναι η ακρίβεια από την μεριά του παρατηρητή της ταξινομημένης εικόνας. Εκφράζει το ποσοστό επί τις εκατό ανάμεσα στα εικονοστοιχεία που ταξινομήθηκαν στην κλάση αυτή, ορθά και μη ορθά.
- F1 σκορ, είναι ο αρμονικός μέσος των ακριβειών δημιουργού και χρήστη και προέρχεται από τον τύπο:

$$F1 = (2 \times UA \times PA) / (UA + PA)$$

4.7 HAMIDA ET AL.

Το δίκτυο Hamida et al.[1] εφαρμόστηκε με βάση την αρχιτεκτονική που δημοσιεύτηκε εδώ (https://github.com/AminaBh/3D_deepLearning_for_hyperspectral_images). Η γειτονιά των εικονοστοιχείων που δέχεται σαν εισαγωγή είναι μεγέθους 5×5 όπως προτείνεται, ο ρυθμός μάθησης είναι 0,0001 όπου την στιγμή που η πορεία της εκπαίδευσης δεν είναι ανοδική στην εποχή 53 αλλάζει στην τιμή 0,00001 προκειμένου να αυξηθεί η ακρίβεια. Batch size επιλέγεται 8 και μέθοδος βελτιστοποίησης η SGD. Επιλέγεται η επικρατέστερη από τα πειράματα της βιβλιογραφίας μορφή του δικτύου με τα 8 στρώματα τα οποία είναι τα εξής :



Εικόνα 54:Οι διαφορετικές παραλλαγές στις αρχιτεκτονικές του δικτύου Hamida et al.

Εισάγεται το διάνυσμα $5 \times 5 \times 176$ όπου περνά από τα στρώματα conv1,pool1 και conv2,pool2. Αυτά τα στρώματα αποτελούνται από τρισδιάστατα φίλτρα τα οποία θα εκπαιδευθούν πάνω στους χωρικούς και φασματικούς συσχετισμούς της εισαγωγής και ταυτόχρονα με την κατάλληλη χρήση του βήματος μετακίνησης θα μετατρέψουν την εισαγωγή σε διάνυσμα 1×1 αποβαλόντας την χωρική πληροφορία.

conv1

weight $\langle 20 \times 1 \times 3 \times 3 \times 3 \rangle$

bias $\langle 20 \rangle$

pool1

weight $\langle 2 \times 20 \times 3 \times 1 \times 1 \rangle$

bias $\langle 2 \rangle$

conv2

weight $\langle 35 \times 2 \times 3 \times 3 \times 3 \rangle$

bias $\langle 35 \rangle$

pool2

weight $\langle 2 \times 35 \times 2 \times 1 \times 1 \rangle$

bias $\langle 2 \rangle$

Στη συνέχεια και πάνω στην ίδια φιλοσοφία θα ακολουθήσουν 4 αντίστοιχα στρώματα συνέλιξης εστιάζοντας στην φασματική πληροφορία, μέχρι που το διάνυσμα ανοίγεται σε μία διάσταση για να το διαδεχθεί ένα στρώμα dropout και ένα τελικό πλήρως συνδεδεμένο στρώμα.

conv3

weight $\langle 35 \times 2 \times 3 \times 1 \times 1 \rangle$

bias $\langle 35 \rangle$

pool3

weight $\langle 2 \times 35 \times 1 \times 1 \times 1 \rangle$

bias $\langle 2 \rangle$

conv4

weight $\langle 35 \times 2 \times 3 \times 1 \times 1 \rangle$

bias $\langle 35 \rangle$

pool4

weight $\langle 4 \times 35 \times 1 \times 1 \times 1 \rangle$

bias $\langle 4 \rangle$



Η πορεία της εισόδου στο μοντέλο ανά στρώμα καθώς και οι παράμετροι προς εκπαίδευση φαίνονται στον πίνακα :

Network :

Layer (type)	Output Shape	Param #
Conv3d-1	[-1, 20, 176, 3, 3]	560
Conv3d-2	[-1, 2, 88, 3, 3]	122
Conv3d-3	[-1, 35, 88, 1, 1]	1,925
Conv3d-4	[-1, 2, 45, 1, 1]	142
Conv3d-5	[-1, 35, 45, 1, 1]	245
Conv3d-6	[-1, 2, 24, 1, 1]	72
Conv3d-7	[-1, 35, 24, 1, 1]	245
Conv3d-8	[-1, 4, 12, 1, 1]	144
Dropout-9	[-1, 48]	0
Linear-10	[-1, 15]	735

Total params: 4,190

Trainable params: 4,190

Non-trainable params: 0

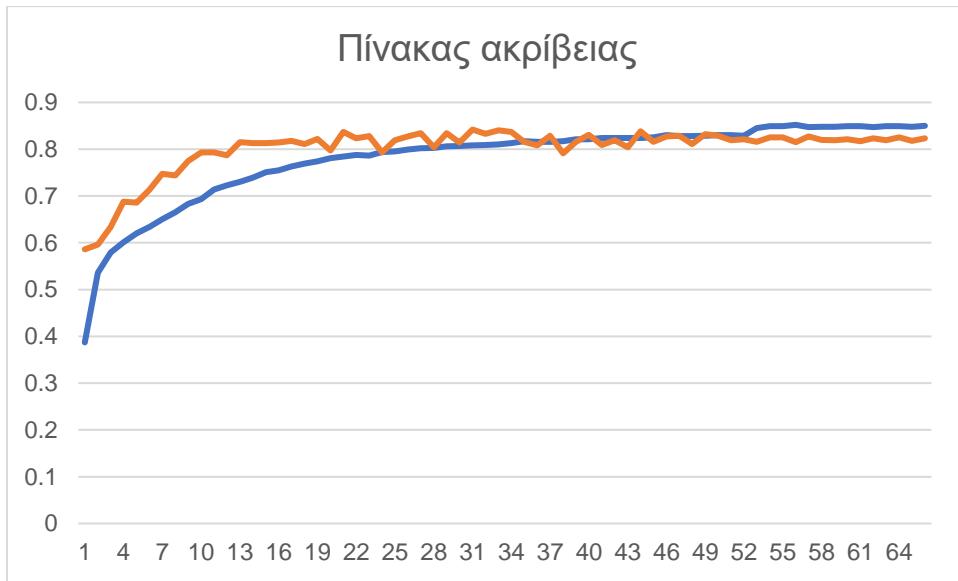
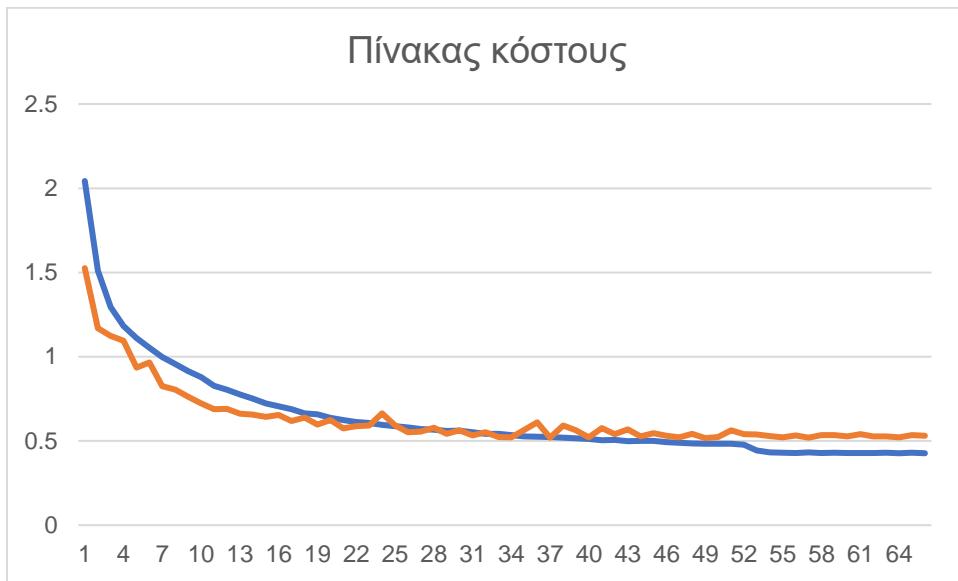
Input size (MB): 0.02

Forward/backward pass size (MB): 0.30

Params size (MB): 0.02

Estimated Total Size (MB): 0.33

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:



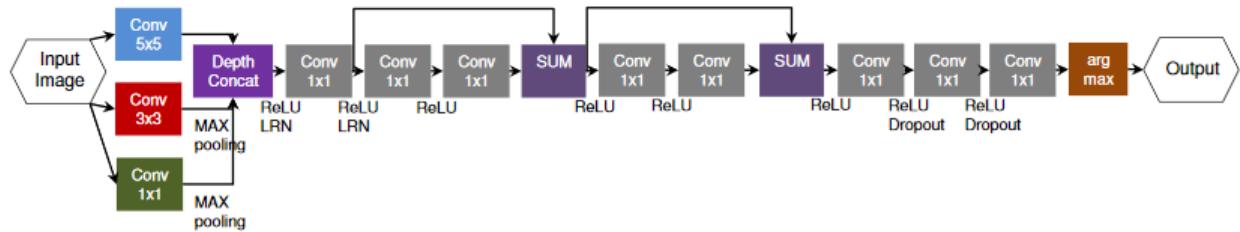
Πίνακες 1,2: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ **εκπαίδευσης-επαλήθευσης**

Με βάση τους πίνακες αυτούς, επιλέγεται η εκπαίδευση του μοντέλου έως την εποχή 18 όπου από αυτήν και έπειτα το μοντέλο δείχνει να υπερπροσαρμόζεται πάνω στο σετ εκπαίδευσης. Από τον πίνακα κόστους παρατηρείται ότι καθώς η πορεία του κόστους και των 2 σετ είναι καθοδική και υπάρχει μια σχετική διαφορά μεταξύ τους, στην εποχή 18 το κόστος του σετ εκπαίδευσης είναι πλέον στα επίπεδα του κόστους του σετ επαλήθευσης. Αυτό σημαίνει πως ενώ το κόστος εκπαίδευσης μειώνεται και επομένως το μοντέλο αντιλαμβάνεται καλύτερα το σετ εκπαίδευσης, στην περίπτωση των δεδομένων της επαλήθευσης δεν συμβαίνει το αντίστοιχο καθώς η πτωτική πορεία δεν είναι η ίδια.

Τα ίδια συμπεράσματα προκύπτουν και από τον πίνακα ακρίβειας που η ακρίβεια του σετ επαλήθευσης έχει φτάσει αρκετά νωρίς στα ανώτερα όρια της, ενώ το σετ εκπαίδευσης συνεχίζει να την αυξάνει καθ' όλη την διάρκεια της μάθησης. Προφανώς από ένα σημείο και μετά η αύξηση αυτή προέρχεται από την υπερπροσαρμογή του μοντέλου στα δεδομένα. Οπότε το μοντέλο αυτό θα εκπαιδευθεί μέχρι την εποχή 18.

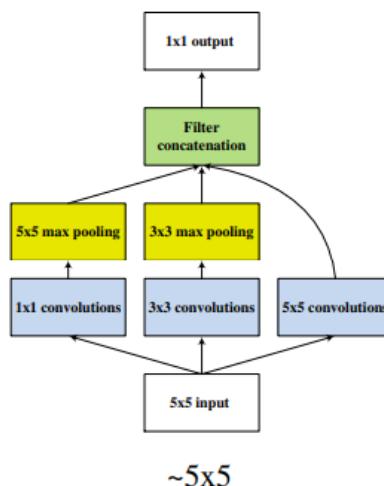
4.8 LEE ET AL.

Το δίκτυο Lee et al.[2] υλοποιήθηκε με βάση την αρχιτεκτονική που δημοσιεύτηκε. Επιλέγεται ο ρυθμός μάθησης ίσος με 0,0001, batch size ίσο με 4 και η γειτονιά των εικονοστοιχείο ίση με 5. Μέθοδος βελτιστοποίησης είναι η SGD.



Εικόνα 55 :Η αρχιτεκτονική του δικτύου Lee et al.

Στην εισαγωγή περνά το διάνυσμα $5 \times 5 \times 176$ όπου οδηγείται στα 3 ξεχωριστά συνελικτικά στρώματα που φαίνονται εδώ:



Εικόνα 56:Multi scale filter bank

conv_3x3

weight $\langle 128 \times 1 \times 176 \times 3 \times 3 \rangle$

bias $\langle 128 \rangle$

conv_1x1

weight $\langle 128 \times 1 \times 176 \times 1 \times 1 \rangle$

bias $\langle 128 \rangle$

conv_5x5

weight $\langle 128 \times 1 \times 176 \times 5 \times 5 \rangle$

bias $\langle 128 \rangle$

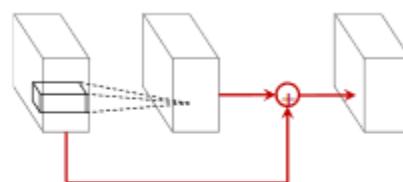
Τα στρώματα αυτά αποτελούν το multi-scale filter bank που εισάγεται στο δίκτυο. Έπειτα το διάνυσμα ενώνεται και περνά από ένα στρώμα κανονικοποίησης μέχρι να οδηγηθεί σε μορφή 1x1 στο επόμενο συνελικτικό στρώμα και έπειτα στα μπλοκ residual learning.

conv1

weight $\langle 128 \times 384 \times 1 \times 1 \rangle$

bias $\langle 128 \rangle$

Το πρώτο συνελικτικό στρώμα μετά από το multi-scale filter bank περιέχει 384 φίλτρα μεγέθους 1x1.



(a) Residual learning

Εικόνα 57: Σχηματική αναπαράσταση της δομής residual learning

conv2

weight⟨128×128×1×1⟩

bias⟨128⟩

conv3

weight⟨128×128×1×1⟩

bias⟨128⟩

Τα παραπάνω στρώματα είναι όμοιας μορφής με το conv1 με μόνη διαφορά τα 128 φίλτρα έναντι των 384 και αποτελούν το βασικό μπλοκ residual learning του δικτύου .Μετά την έξοδο από το multi-scale filter bank και το πέρασμα από το στρώμα conv1, ακολουθούν 2 residual learning blocks αποτελούμενα από τα στρώματα conv2,conv3.Αμέσως μετά, ακολουθούν τρία συνελικτικά στρώματα του ίδιου τύπου, παρόμοια με τα conv2,conv3, που πλαισιώνονται από δύο στρώματα dropout($p=0,5$) για την αποφυγή της υπερπροσαρμογής. Τέλος δίνεται η πρόβλεψη.

Η πορεία της εισαγωγής στο μοντέλο ανά στρώμα καθώς και οι παράμετροι προς εκπαίδευση φαίνονται στον πίνακα :

Network :

Layer (type)	Output Shape	Param #
Conv3d-1	[-1, 128, 1, 5, 5]	22,656
MaxPool3d-2	[-1, 128, 1, 1]	0
Conv3d-3	[-1, 128, 1, 3, 3]	202,880
MaxPool3d-4	[-1, 128, 1, 1]	0
Conv3d-5	[-1, 128, 1, 1, 1]	563,328
LocalResponseNorm-6	[-1, 384, 1, 1]	0
Conv2d-7	[-1, 128, 1, 1]	49,280
LocalResponseNorm-8	[-1, 128, 1, 1]	0
Conv2d-9	[-1, 128, 1, 1]	16,512
Conv2d-10	[-1, 128, 1, 1]	16,512
Conv2d-11	[-1, 128, 1, 1]	16,512
Conv2d-12	[-1, 128, 1, 1]	16,512
Conv2d-13	[-1, 128, 1, 1]	16,512
Dropout-14	[-1, 128, 1, 1]	0
Conv2d-15	[-1, 128, 1, 1]	16,512
Dropout-16	[-1, 128, 1, 1]	0
Conv2d-17	[-1, 15, 1, 1]	1,935

Total params: 939,151

Trainable params: 939,151

Non-trainable params: 0

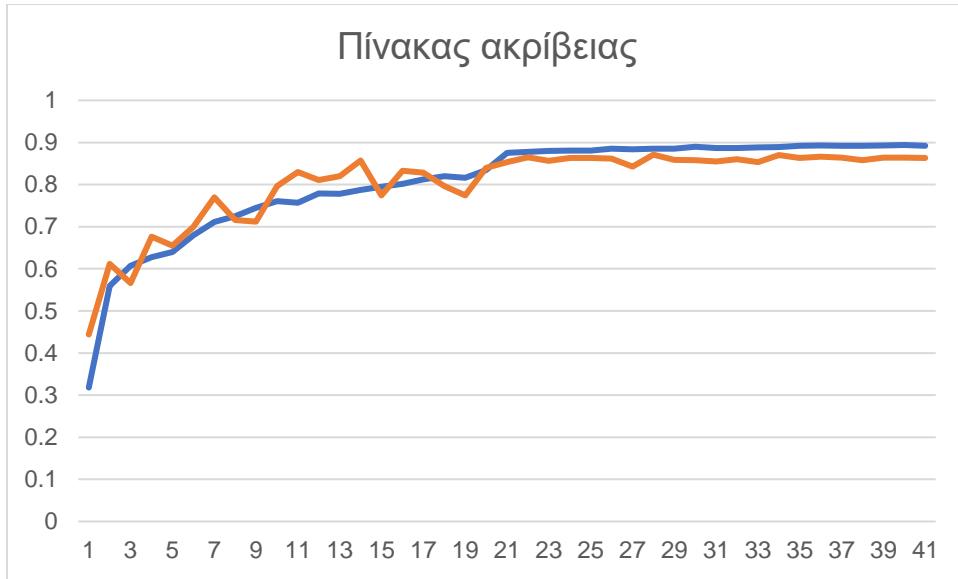
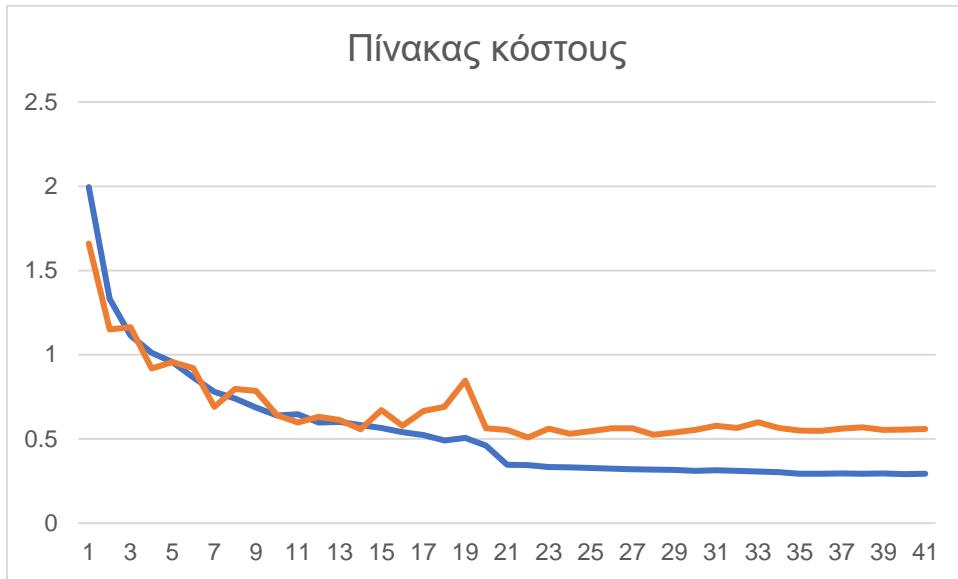
Input size (MB): 0.02

Forward/backward pass size (MB): 0.05

Params size (MB): 3.58

Estimated Total Size (MB): 3.65

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:

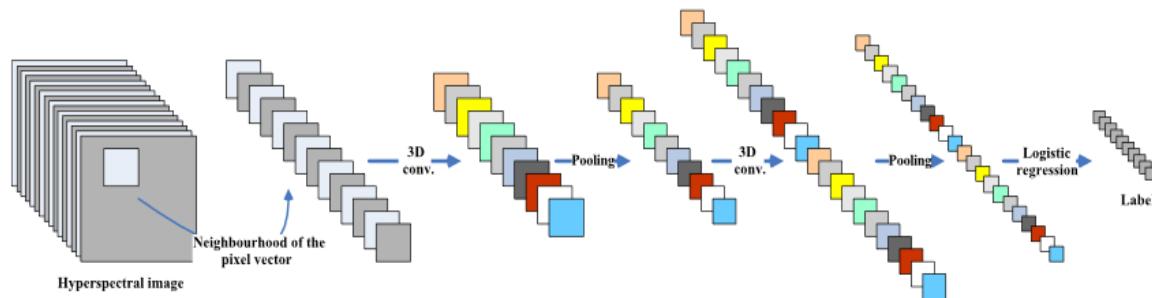


Πίνακες 3,4: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ **εκπαίδευσης-επαλήθευσης**

Από τον πίνακα κόστους παρατηρείται πως και τα 2 σετ δεδομένων παρουσιάζουν μείωση στο κόστος αλλά από την εποχή 11 και μετά το κόστος παραμένει σταθερό για το σετ επαλήθευσης καθώς το αντίστοιχο κόστος του σετ εκπαίδευσης συνεχίζει να μειώνεται. Κατά συνέπεια και στον πίνακα ακρίβειας το ποσοστό ακρίβειας του σετ επαλήθευσης μένει σταθερό από την εποχή 11 και μετά. Έτσι επιλέγεται η εκπαίδευση του μοντέλου μέχρι την στιγμή που θα ολοκληρωθεί η εποχή 11.

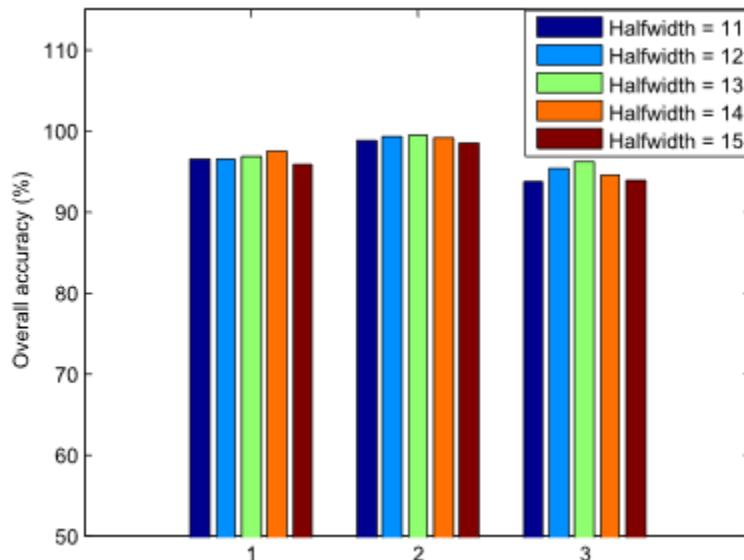
4.9 CHEN ET AL.

Το δίκτυο Chen et al.[3]. είναι ένα δίκτυο που έχει αργή εκπαίδευση ,γεγονός που οφείλεται στην μεγάλη γειτονιά των εικονοστοιχείων που δέχεται ως εισαγωγή και στα μεγάλου μεγέθους φίλτρα που εφαρμόζονται σε αυτή, εισάγοντας πολλούς παραμέτρους προς εκπαίδευση. Ο ρυθμός μάθησης επιλέγεται 0.1, batch size 200 προκειμένου να υπάρξει ταχύτερη μάθηση και η μέθοδος βελτιστοποίησης είναι η SGD.



Εικόνα 58:Η αρχιτεκτονική του δικτύου Chen et al.

Το μέγεθος της γειτονιάς των εικονοστοιχείων στην εισαγωγή επιλέγεται 27×27 .Η επιλογή αυτή είναι αυτή που προτείνεται από τα πειράματα που έγιναν στα γνωστά datasets (Indian Pines,PaviaU,KSC).

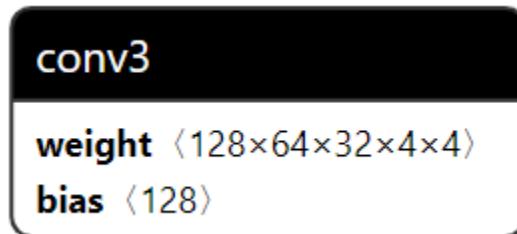


Εικόνα 59:Συγκριτικός πίνακας ακρίβειας για διαφορετικά μεγέθη γειτονιών εικονοστοιχείων στην είσοδο του δικτύου.

Αρχικά τα πρώτα συνελικτικά στρώματα είναι αυτά, στα οποία παρεμβάλλονται 2 όμοια στρώματα pooling διαστάσεων (2,2,1):



Έπειτα εισάγονται στο δίκτυο τα στρώματα dropout($p=0,5$) όπου παρεμβάλλεται ανάμεσα τους το τρίτο συνελικτικό στρώμα.



Τέλος ακολουθεί ένα πλήρως συνδεδεμένο στρώμα για να δοθεί η τελική πρόβλεψη.

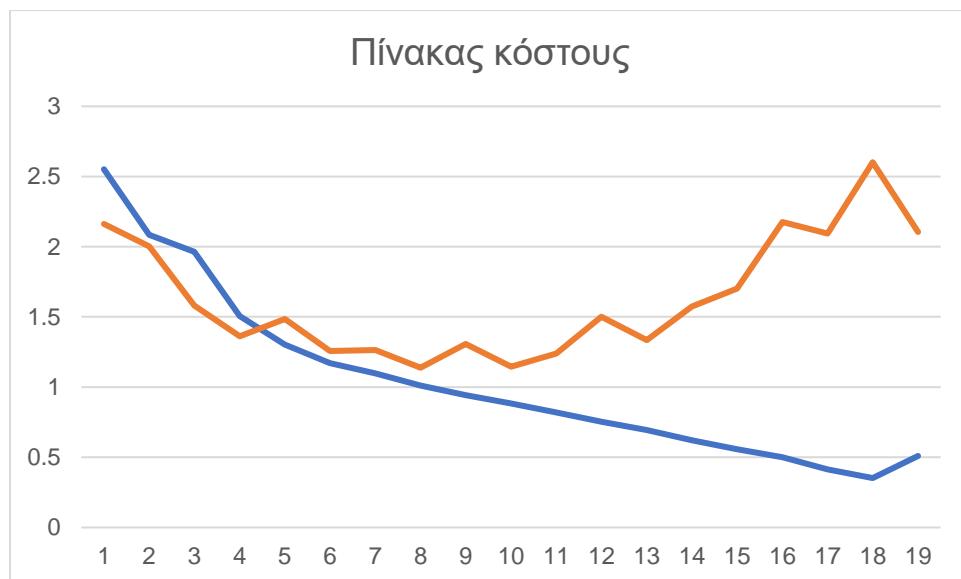


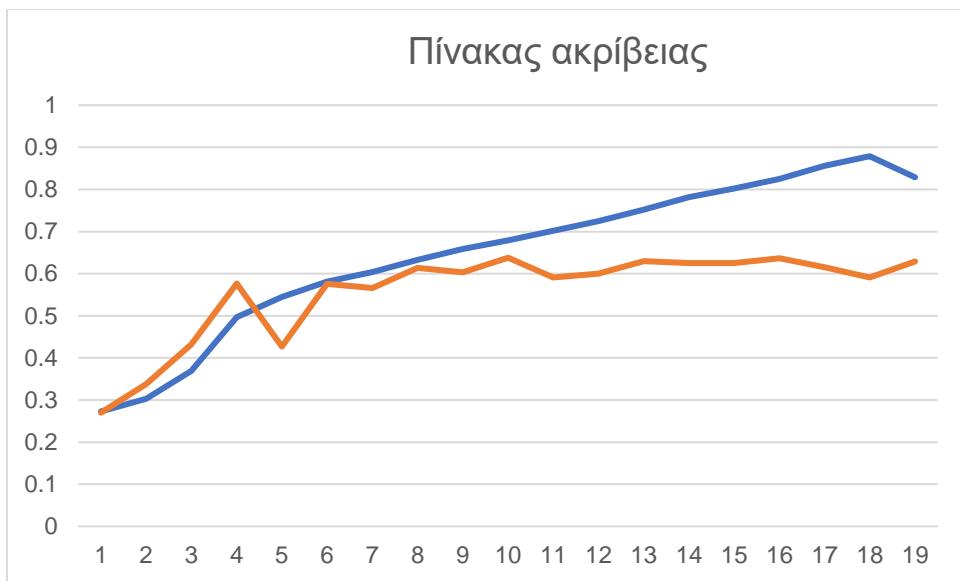
Η πορεία της εισαγωγής στο μοντέλο ανά στρώμα καθώς και οι παράμετροι προς εκπαίδευση φαίνονται στον πίνακα :

Network :

```
-----  
Layer (type)           Output Shape        Param #  
=====  
Conv3d-1               [-1, 32, 145, 24, 24]    16,416  
MaxPool3d-2            [-1, 32, 145, 12, 12]    0  
Conv3d-3               [-1, 64, 114, 8, 8]     1,638,464  
MaxPool3d-4            [-1, 64, 114, 4, 4]    0  
Dropout-5              [-1, 64, 114, 4, 4]    0  
Conv3d-6               [-1, 128, 83, 1, 1]   4,194,432  
Dropout-7              [-1, 128, 83, 1, 1]    0  
Linear-8                [-1, 15]          159,375  
-----  
Total params: 6,008,687  
Trainable params: 6,008,687  
Non-trainable params: 0  
-----  
Input size (MB): 0.49  
Forward/backward pass size (MB): 30.99  
Params size (MB): 22.92  
Estimated Total Size (MB): 54.41
```

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:



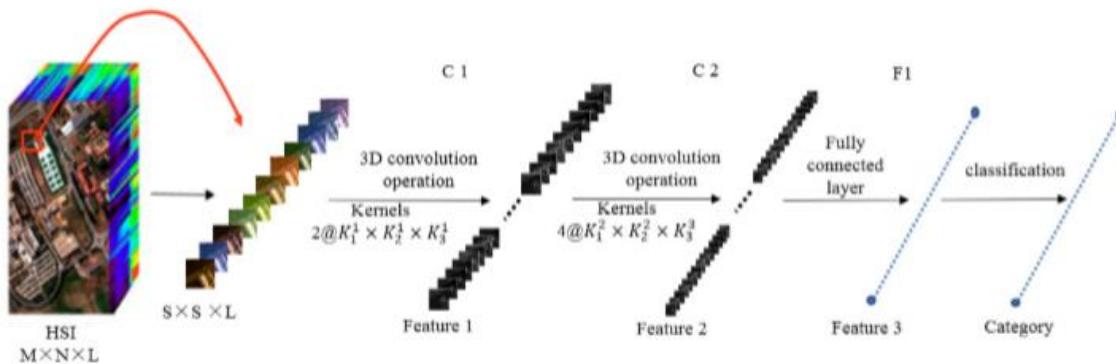


Πίνακες 5,6: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ [εκπαίδευσης-επαλήθευσης](#)

Από τους πίνακες προκύπτει πως η εποχή 8 είναι η εποχή που θα πρέπει να σταματήσει να εκπαιδεύεται ο αλγόριθμος. Το κόστος του σετ επαλήθευσης ανεβαίνει ενώ του σετ εκπαίδευσης πέφτει και η ακρίβεια διατηρείται σταθερή σε όλο το διάστημα μετά από την εποχή 8 όπου η αντίστοιχη ακρίβεια του σετ εκπαίδευσης ανεβαίνει.

4.10 LI ET AL.

Το δίκτυο Li et al.[4] υλοποιήθηκε με βάση την αρχιτεκτονική που δημοσιεύτηκε εδώ (<https://www.mdpi.com/2072-4292/9/1/67>). Ορίστηκε ρυθμός μάθησης ίσος με 0.0001, batch size στην εισαγωγή ίσο με 4 και η γειτονιά των εικονοστοιχείων ίση με 5. Μέθοδος βελτιστοποίησης είναι η SGD.



Εικόνα 60:Η αρχιτεκτονική του δικτύου Li et al.

Στην είσοδο δέχεται ένα κομμάτι τους εικόνας $5 \times 5 \times 176$ όπου 5×5 είναι η γειτονιά του κεντρικού εικονοστοιχείου και 176 τα φασματικά κανάλια. Έπειτα περνά από 16 φίλτρα (3,3,7) και στη συνέχεια από 32 φίλτρα (3,3,3) για να καταλήξει στα πλήρως συνδεδεμένα στρώματα και στην συνάρτηση κόστους.

conv1	conv2	fc
weight $\langle 16 \times 1 \times 7 \times 3 \times 3 \rangle$ bias $\langle 16 \rangle$	weight $\langle 32 \times 16 \times 3 \times 3 \times 3 \rangle$ bias $\langle 32 \rangle$	weight $\langle 15 \times 5504 \rangle$ bias $\langle 15 \rangle$

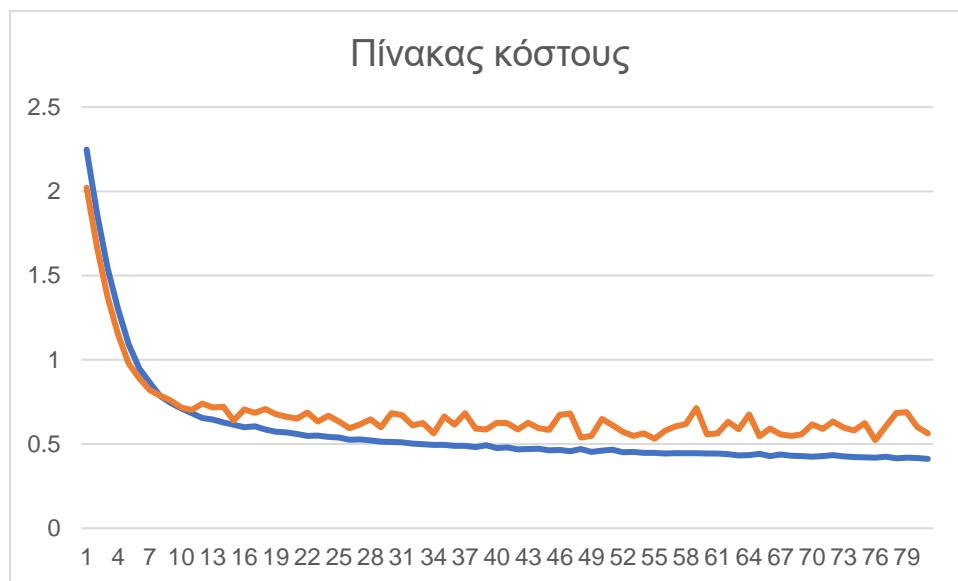
Η επιλογή των 16+32 φίλτρων προέκυψε από το γεγονός ότι τα 2+8 φίλτρα που προτείνονται για το dataset Indian Pines αναφέρονται στον μικρό αριθμό στοιχείων και γενικότερον μεγέθους του dataset. Στην περίπτωση του dataset τους συγκεκριμένης εργασίας υπάρχουν πολλά περισσότερα στοιχεία τους ταξινόμηση τα οποία απαιτούν περισσότερους παραμέτρους τους εκπαίδευση για να υπάρξει σχετικά υψηλή ακρίβεια. Άλλωστε συγκριτικά με τους τους υπόλοιπες μεθόδους που αναφέρονται στα συνελικτικά δίκτυα ο αριθμός των 32 φίλτρων είναι ασυνήθιστα μικρός.

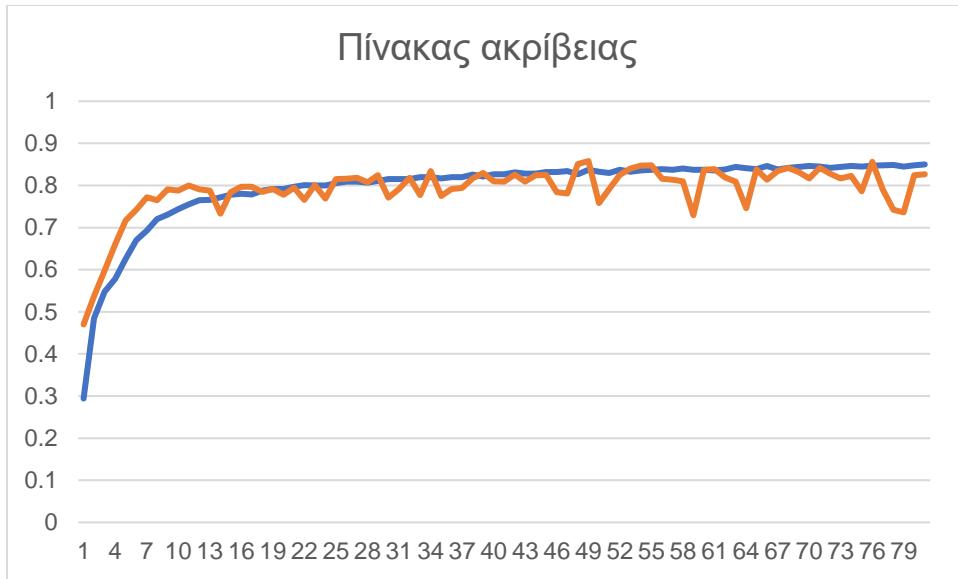
Η πορεία τους εισαγωγής στο μοντέλο ανά στρώμα καθώς και οι παράμετροι τους εκπαίδευση φαίνονται στον πίνακα :

Network :

```
-----  
Layer (type)           Output Shape        Param #  
=====              ======  =====  
Conv3d-1             [-1, 16, 172, 3, 3]    1,024  
Conv3d-2             [-1, 32, 172, 1, 1]   13,856  
Linear-3              [-1, 15]            82,575  
=====  
Total params: 97,455  
Trainable params: 97,455  
Non-trainable params: 0  
-----  
Input size (MB): 0.02  
Forward/backward pass size (MB): 0.23  
Params size (MB): 0.37  
Estimated Total Size (MB): 0.62
```

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:



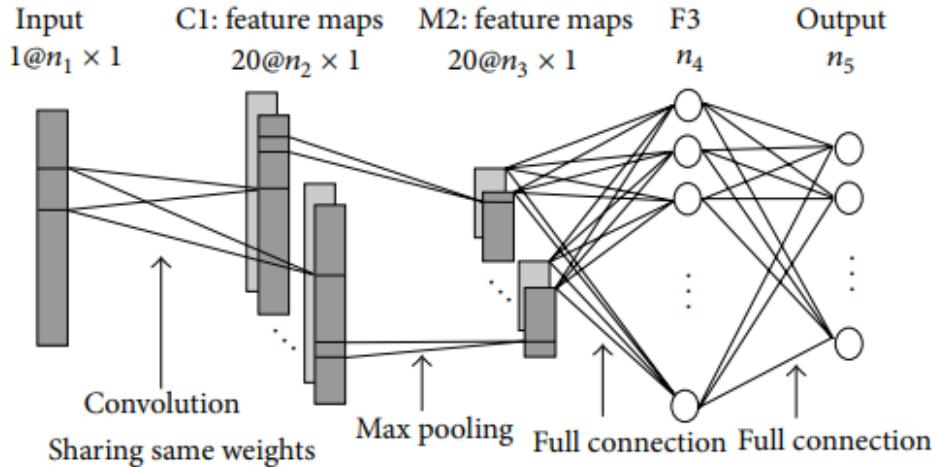


Πίνακες 7,8: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ **εκπαίδευσης-επαλήθευσης**

Από τον πίνακα κόστους προκύπτει πως από την εποχή 33 και έπειτα το κόστος του σετ επαλήθευσης παραμένει στα ίδια επίπεδα ενώ του σετ εκπαίδευσης μειώνεται. Το ίδιο συμπέρασμα προκύπτει και από τον πίνακα ακρίβειας. Οπότε το μοντέλο θα εκπαιδευθεί για 33 εποχές.

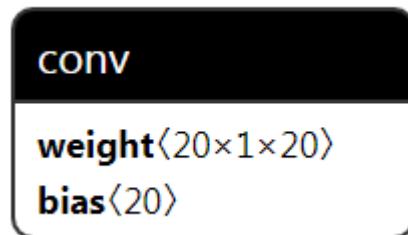
4.11 HU ET AL.

Το δίκτυο Hu et al.[5] υλοποιήθηκε με βάση την αρχιτεκτονική που δημοσιεύτηκε. Επιλέχθηκε ο ρυθμός μάθησης ίσος με 0.08, batch size στην εισαγωγή ίσο με 256. Μέθοδος βελτιστοποίησης είναι η SGD.

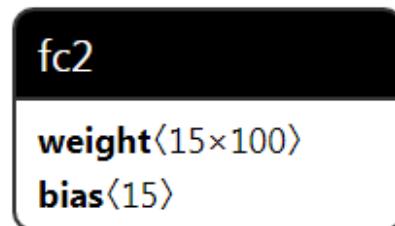


Εικόνα 61:Η αρχιτεκτονική του δικτύου Hu et al.

Η είσοδος στο δίκτυο από κάθε εικονοστοιχείο ξεχωριστά με τα φασματικά κανάλια του, δηλαδή διάνυσμα 1×176 . Αρχικά περνάει από ένα συνελικτικό στρώμα μεγέθους 1×20 .



Στη συνέχεια περνά από ένα pooling στρώμα μεγέθους 4 μειώνοντας των αριθμό των φασματικών καναλιών καταλήγοντας σε 2 πλήρως συνδεδεμένα στρώματα :



Η πορεία τους εισαγωγής στο μοντέλο ανά στρώμα καθώς και οι παράμετροι τους εκπαίδευση φαίνονται στον πίνακα :

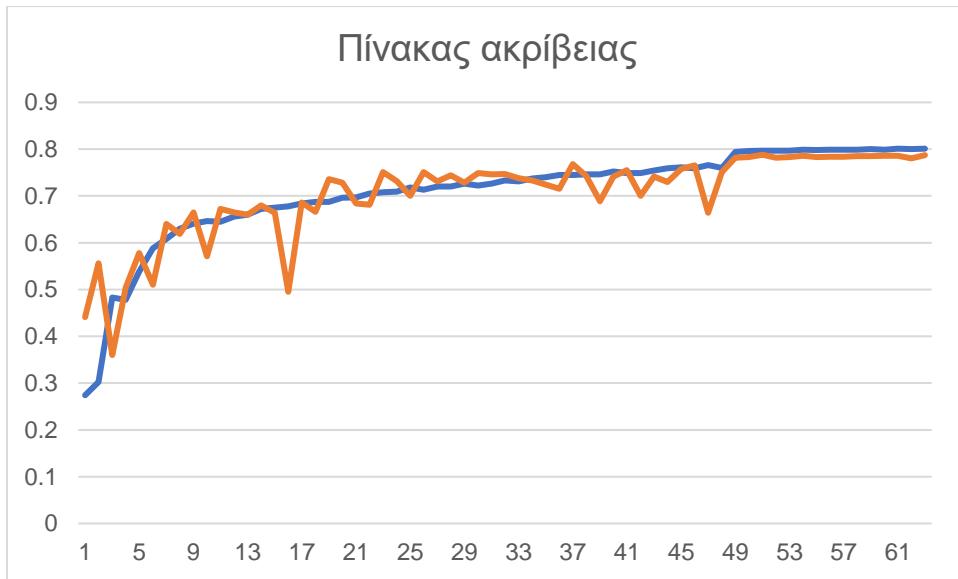
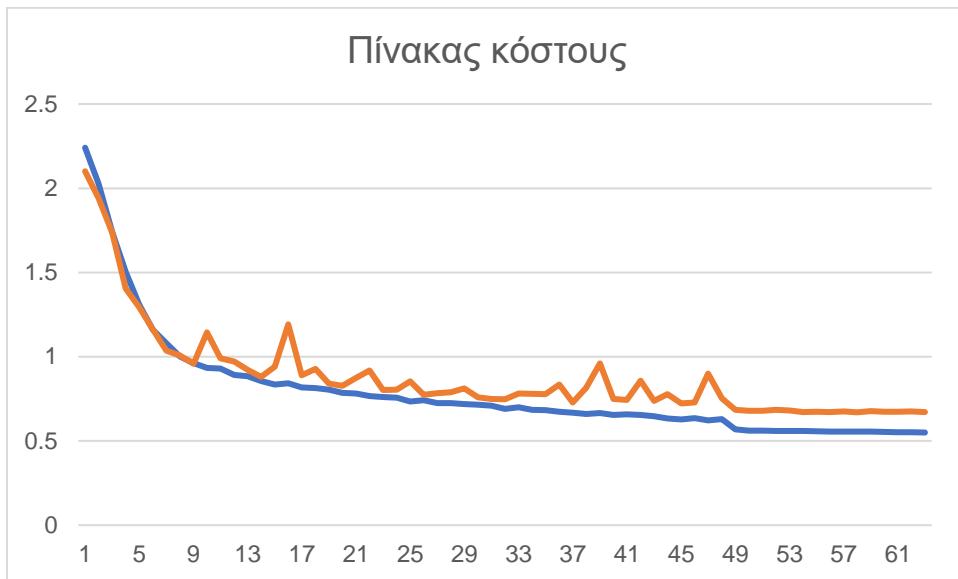
Network :

Layer (type)	Output Shape	Param #
Conv1d-1	[-1, 20, 157]	420
MaxPool1d-2	[-1, 20, 39]	0
Linear-3	[-1, 100]	78,100
Linear-4	[-1, 15]	1,515

Total params: 80,035
Trainable params: 80,035
Non-trainable params: 0

Input size (MB): 0.00
Forward/backward pass size (MB): 0.03
Params size (MB): 0.31
Estimated Total Size (MB): 0.34

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:

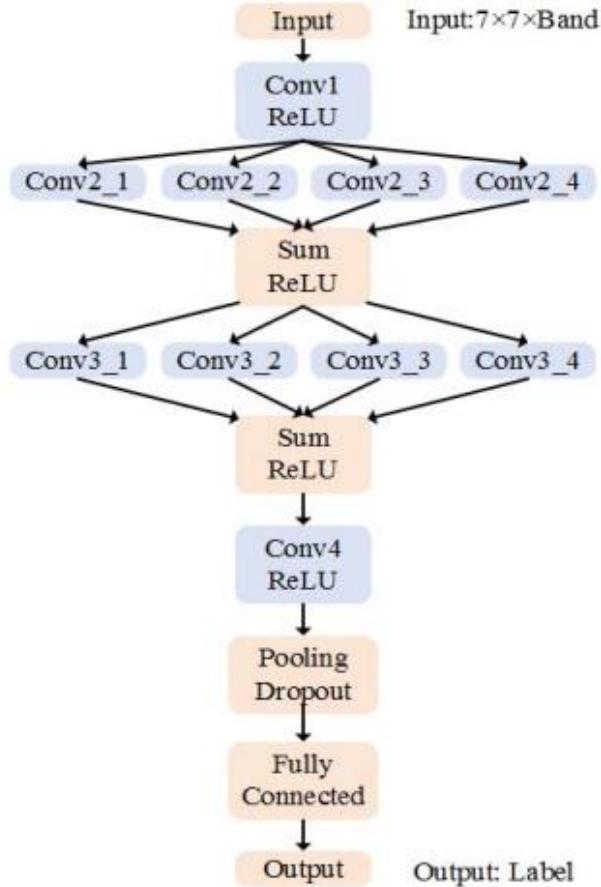


Πίνακες 8,9: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ [εκπαίδευσης](#)-[επαλήθευσης](#)

Από τον πίνακα ακρίβειας προκύπτει πως από την 21 εποχή και ύστερα η ακρίβεια του σετ επαλήθευσης σταθεροποιείται παρά την αύξηση τους αντίστοιχης ακρίβειας στο σετ εκπαίδευσης. Το ίδιο προκύπτει παρατηρώντας τον πίνακα κόστους. Επιλέγεται η εκπαίδευση για 21 εποχές.

4.12 HE ET AL.

Το δίκτυο He et al.[6] βασίστηκε στην αρχιτεκτονική που δημοσιεύτηκε. Επιλέχθηκε ο ρυθμός μάθησης ίσος με 0.0001, batch size στην εισαγωγή ίσο με 256. Μέθοδος βελτιστοποίησης είναι η SGD.



Εικόνα 62:Η αρχιτεκτονική του δικτύου He et al.

Στην είσοδο δέχεται μια γειτονία εικονοστοιχείο $7 \times 7 \times 176$ όπου 7×7 είναι το τμήμα τους εικόνας και 176 τα φασματικά κανάλια. Το πρώτο στρώμα CONV1 αποτελείται από 16 φίλτρα μεγέθους $(3,3,11)$.

conv1

weight $\langle 16 \times 1 \times 11 \times 3 \times 3 \rangle$

bias $\langle 16 \rangle$

Μετέπειτα εμφανίζεται το πρώτο σε σειρά Multi scale 3D Convolution Block. Η συγκεκριμένη ομάδα φίλτρων αποτελείται από 4 ομάδες 16 φίλτρων. Οι διαστάσεις αυτών είναι $(1,1,1), (1,1,3), (1,1,5), (1,1,7)$. Αφού αθροιστούν οι έξοδοι, το αποτέλεσμα περνά ξανά ως είσοδος σε ένα πανομοιότυπο Multi scale 3D Convolution Block.

conv2_1

weight $\langle 16 \times 16 \times 1 \times 1 \times 1 \rangle$

bias $\langle 16 \rangle$

conv2_2

weight $\langle 16 \times 16 \times 3 \times 1 \times 1 \rangle$

bias $\langle 16 \rangle$

conv2_3

weight $\langle 16 \times 16 \times 5 \times 1 \times 1 \rangle$

bias $\langle 16 \rangle$

conv2_4

weight $\langle 16 \times 16 \times 11 \times 1 \times 1 \rangle$

bias $\langle 16 \rangle$

Στη συνέχεια θα ακολουθήσει ένα συνελικτικό στρώμα με 16 φίλτρα μεγέθους $(2,2,3)$.



Στο τελευταίο στάδιο ένα στρώμα pooling (2,2,3) θα μειώσει τους διαστάσεις τους εισόδου και ένα στρώμα dropout ($p=0.6$) θα χρησιμοποιηθεί για την αποφυγή της υπερπροσαρμογής. Έτσι καταλήγει σε ένα FCN δίκτυο που θα δώσει την τελική πρόβλεψη για το εικονοστοιχείο.



Η πορεία τους εισαγωγής στο μοντέλο ανά στρώμα καθώς και οι παράμετροι τους εκπαίδευση φαίνονται στον πίνακα :

Network :

Layer (type)	Output Shape	Param #
Conv3d-1	[-1, 16, 56, 5, 5]	1,600
Conv3d-2	[-1, 16, 56, 5, 5]	272
Conv3d-3	[-1, 16, 56, 5, 5]	784
Conv3d-4	[-1, 16, 56, 5, 5]	1,296
Conv3d-5	[-1, 16, 56, 5, 5]	2,832
Conv3d-6	[-1, 16, 56, 5, 5]	272
Conv3d-7	[-1, 16, 56, 5, 5]	784
Conv3d-8	[-1, 16, 56, 5, 5]	1,296
Conv3d-9	[-1, 16, 56, 5, 5]	2,832
Conv3d-10	[-1, 16, 54, 4, 4]	3,088
Dropout-11	[-1, 13824]	0
Linear-12	[-1, 15]	207,375

Total params: 222,431

Trainable params: 222,431

Non-trainable params: 0

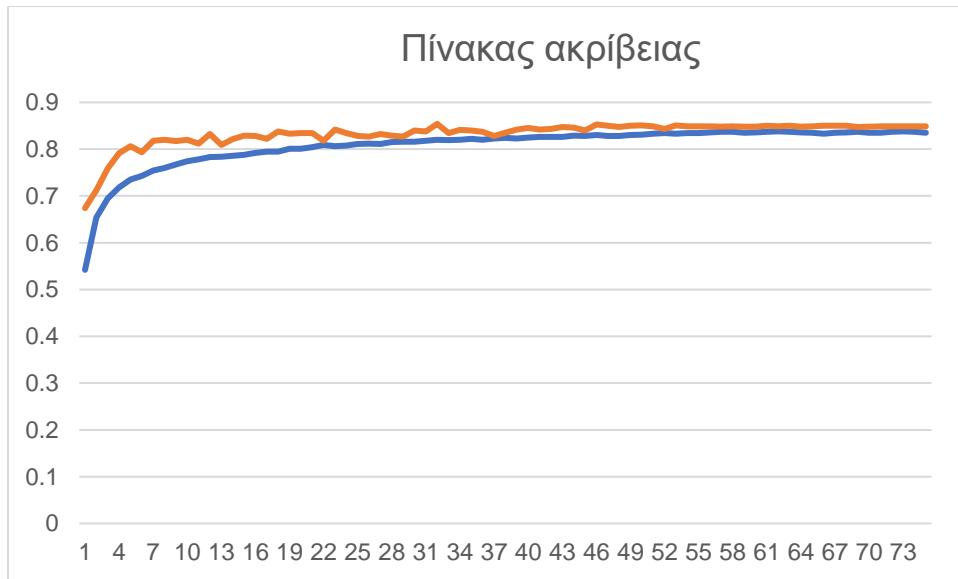
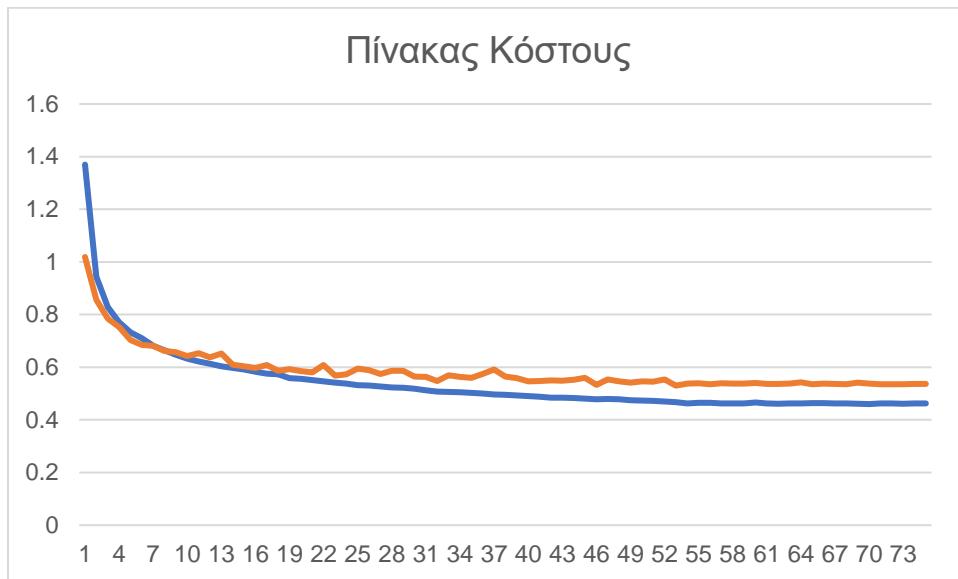
Input size (MB): 0.03

Forward/backward pass size (MB): 1.75

Params size (MB): 0.85

Estimated Total Size (MB): 2.63

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:



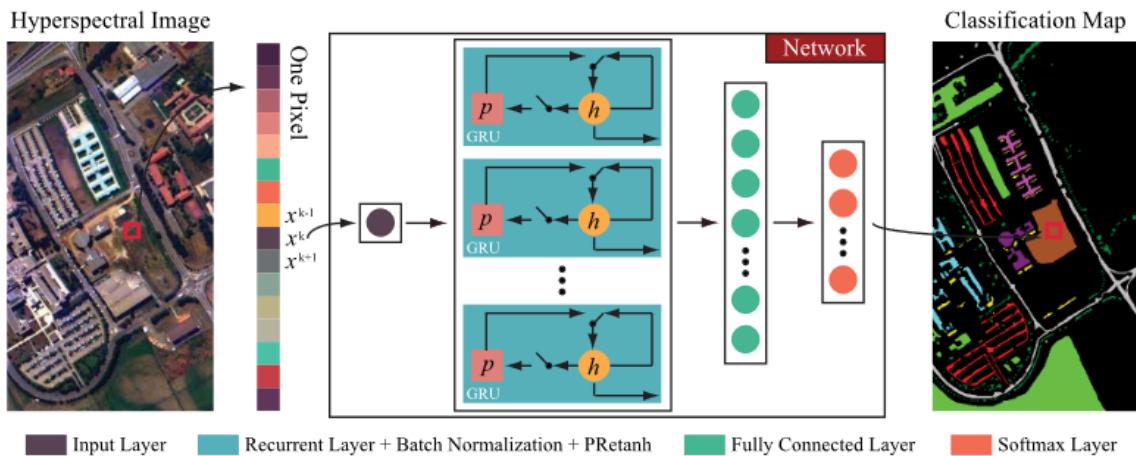
Πίνακες 10,11: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ εκπαίδευσης-επαλήθευσης

Από τους πίνακες δεν παρατηρείται το φαινόμενο τους υπερπροσαρμογής του μοντέλου. Από την εποχή 43 και έπειτα το κόστος παύει να μειώνεται και από τα 2 σετ, οπότε επιλέγεται αυτή η εποχή ως το σημείο που θα σταματήσει το μοντέλο την εκπαίδευση

4.13 MOU ET AL.

Το δίκτυο Mou et al.[7] στην εισαγωγή δέχεται ένα μοναδικό εικονοστοιχείο με την φασματική πληροφορία των 176 καναλιών. Έπειτα γίνεται το πέρασμα από το στρώμα GRU και ένα στρώμα κανονικοποίησης. Η έξοδος από το στρώμα GRU ενεργοποιείται με την PRTanh.. Τελικά ακολουθεί ένα πλήρως συνδεδεμένο στρώμα και δίνεται η πρόβλεψη. Η αρχιτεκτονική του δικτύου αναλυτικότερα βασίστηκε στην σχετική δημοσίευση.

Ο ρυθμός μάθησης είναι 0,01, το πακέτο δεδομένων στην εισαγωγή 64 και η μέθοδος βελτιστοποίησης είναι η Adadelta.



Εικόνα 63:Η αρχιτεκτονική του δικτύου Mou et al.

Η πορεία τους εισαγωγής στο μοντέλο ανά στρώμα καθώς και οι παράμετροι τους εκπαίδευση φαίνονται στον πίνακα :

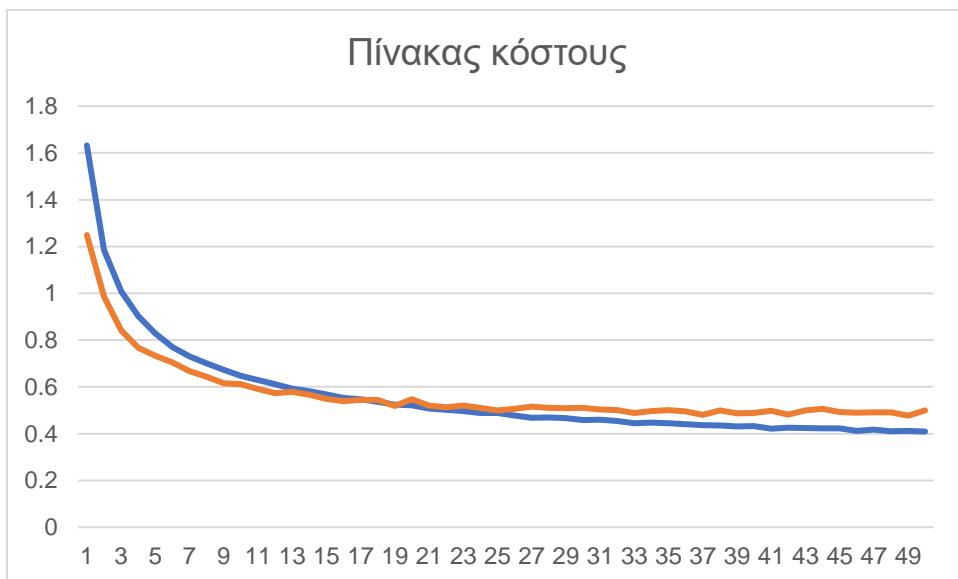
Network :

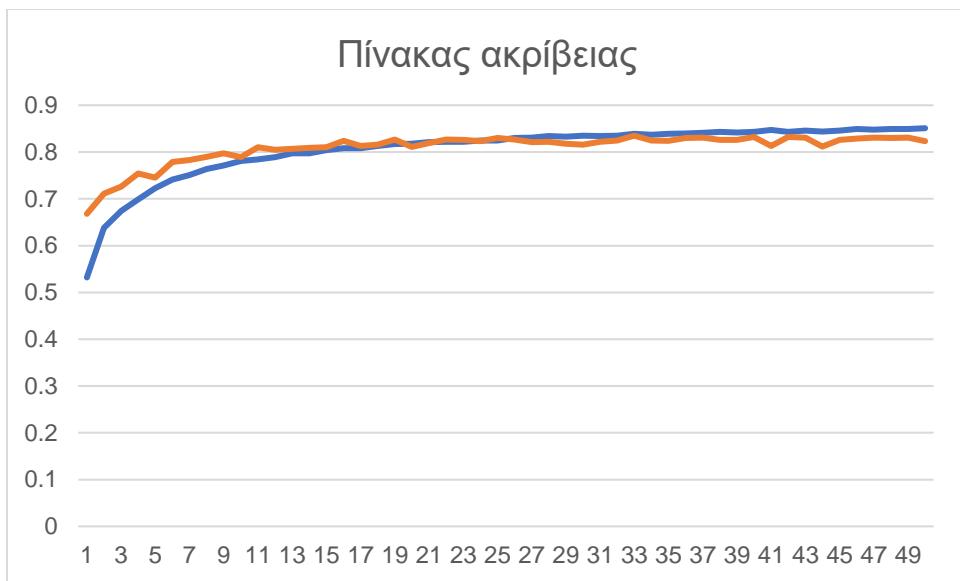
Layer (type)	Output Shape	Param #
GRU-1	[[-1, 2, 64], [-1, 2, 64]]	0
BatchNorm1d-2	[-1, 11264]	22,528
PRTanh-3	[-1, 11264]	0
Linear-4	[-1, 15]	168,975

Total params:	191,503
Trainable params:	191,503
Non-trainable params:	0

Input size (MB):	0.00
Forward/backward pass size (MB):	0.05
Params size (MB):	0.73
Estimated Total Size (MB):	0.78

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:



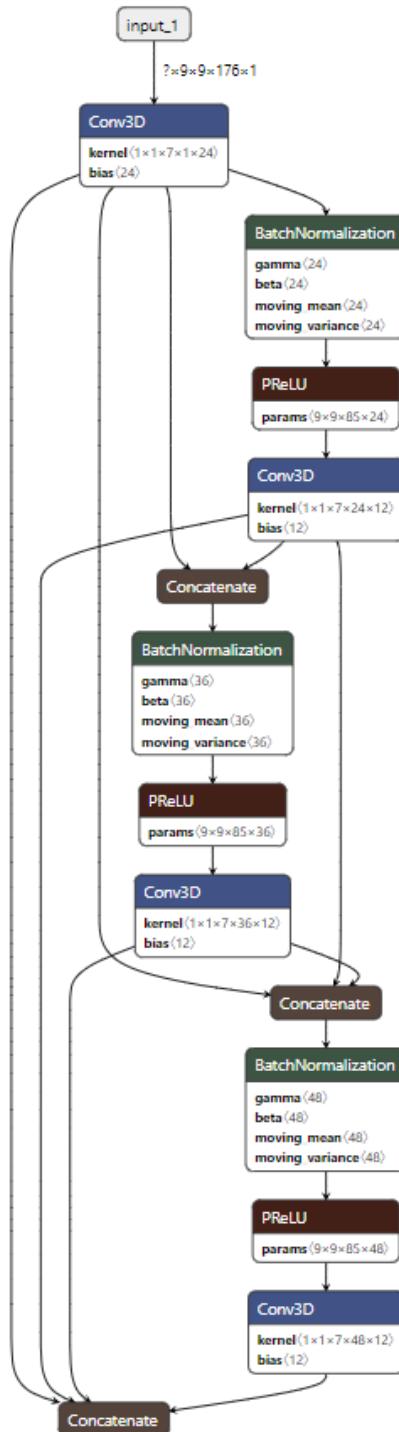


Πίνακες 12,13:Κόστος/Ακρίβεια ανά εποχή, σετ **εκπαίδευσης-επαλήθευσης**

Στο σημείο που εκπαίδευση φτάνει στην εποχή 19, παρατηρείται και από τους 2 πίνακες πως το σετ επαλήθευσης μένει στάσιμο σχετικά με την αύξηση τους ακρίβειας και μείωση του κόστους, και αφού το σετ εκπαίδευσης παρουσιάζει διαφορετική συμπεριφορά προκύπτει πως αρχίζει το μοντέλο να υπερπροσαρμόζεται. Έτσι επιλέγεται να εκπαιδευθεί το μοντέλο για 19 εποχές.

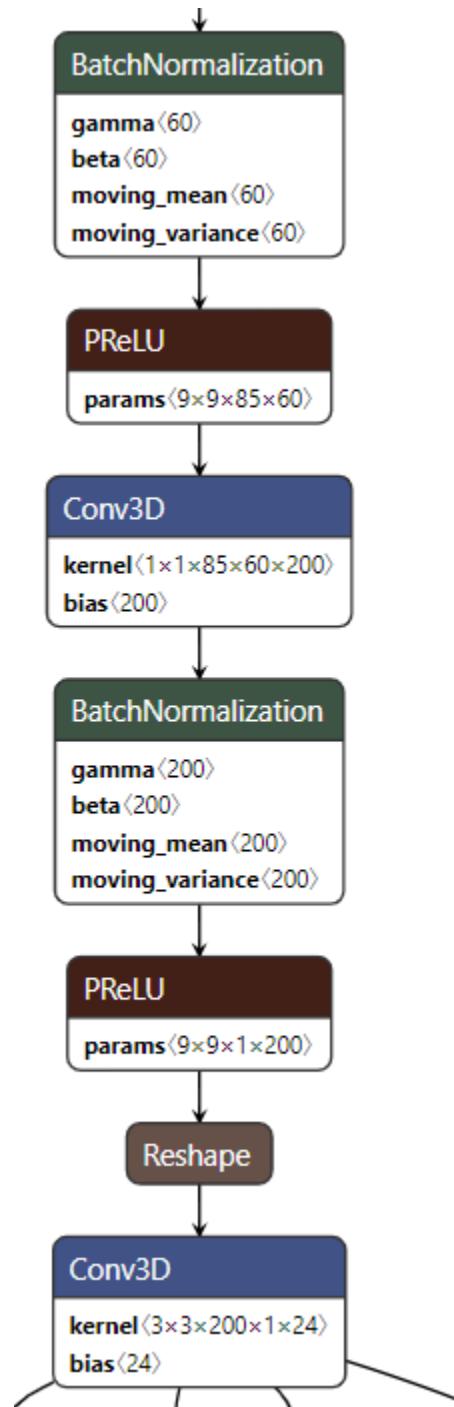
4.14 FDSSC

Το δίκτυο FDSSC [8] έχει σαν εισαγωγή την γειτονιά των 9×9 εικονοστοιχείων $\times 176$ φασματικά κανάλια. Αφού περάσει από ένα συνελικτικό στρώμα με φίλτρο $(1,1,7)$ και μειωθεί η φασματική του διάσταση, εισάγεται στο πρώτο από τα 2 μπλοκ, το Dense Spectral Block.



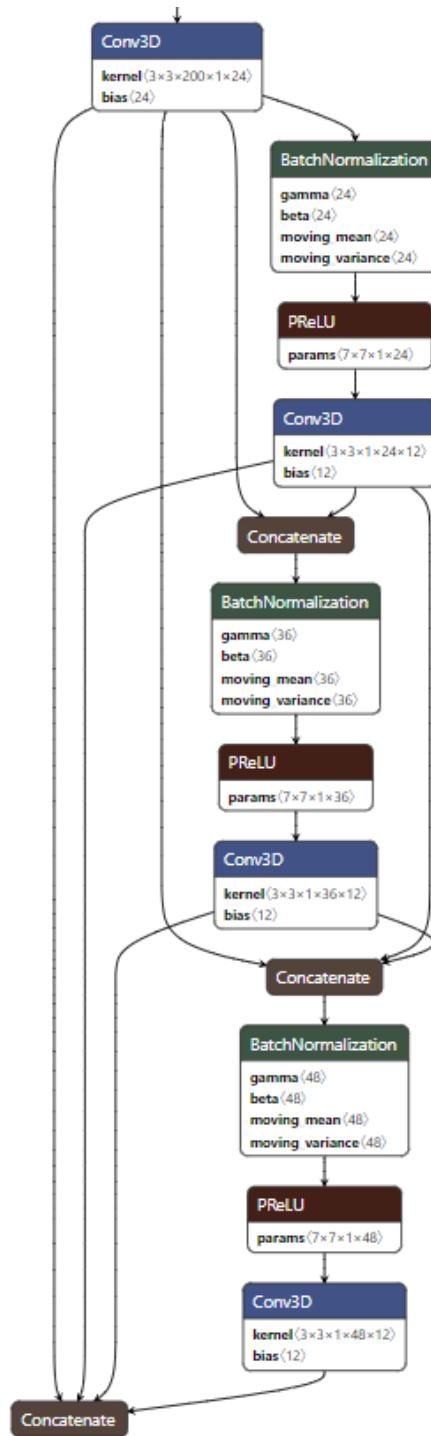
Εικόνα 64 : Τα στρώματα που απαρτίζουν το δίκτυο FDDSC

Στη συνέχεια, η ενωμένη έξοδος θα περάσει από άλλο ένα συνελικτικό στρώμα που εξαλείψει την φασματική διάσταση του, θα αναπροσαρμοστούν οι διαστάσεις του, περνώντας πλέον στο Dense Spatial block για την άντληση των χωρικών συσχετισμών.



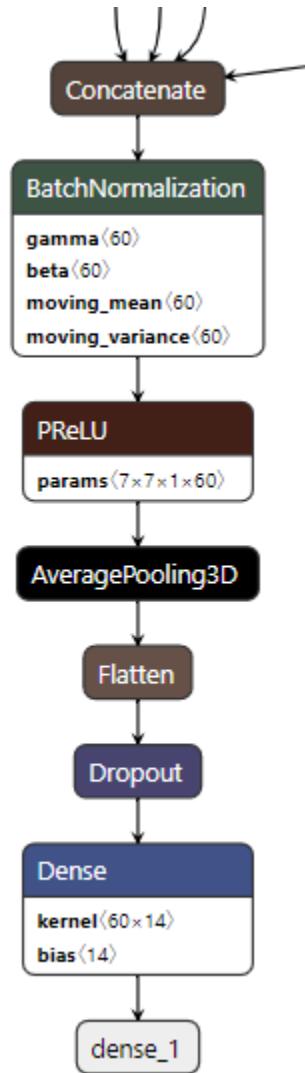
Εικόνα 65 : Τα στρώματα που απαρτίζουν το δίκτυο FDDSC

To Dense Spatial block είναι παρόμοιο με το αντίστοιχο block για την άντληση των φασματικών πληροφορίων. Αποτελείται από φίλτρα (7,7,1) σε παρόμοια διάταξη και λογική με το φασματικό, μόνο που αυτό αντλεί τους χωρικούς συσχετισμούς μεταξύ των εικονοστοιχείο.



Εικόνα 66 : Τα στρώματα που απαρτίζουν το δίκτυο FDDSC

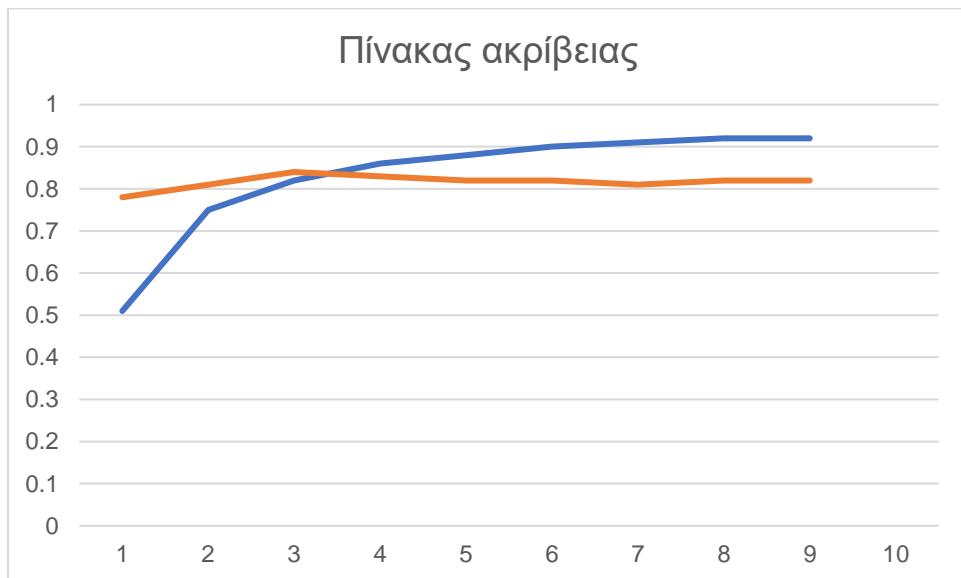
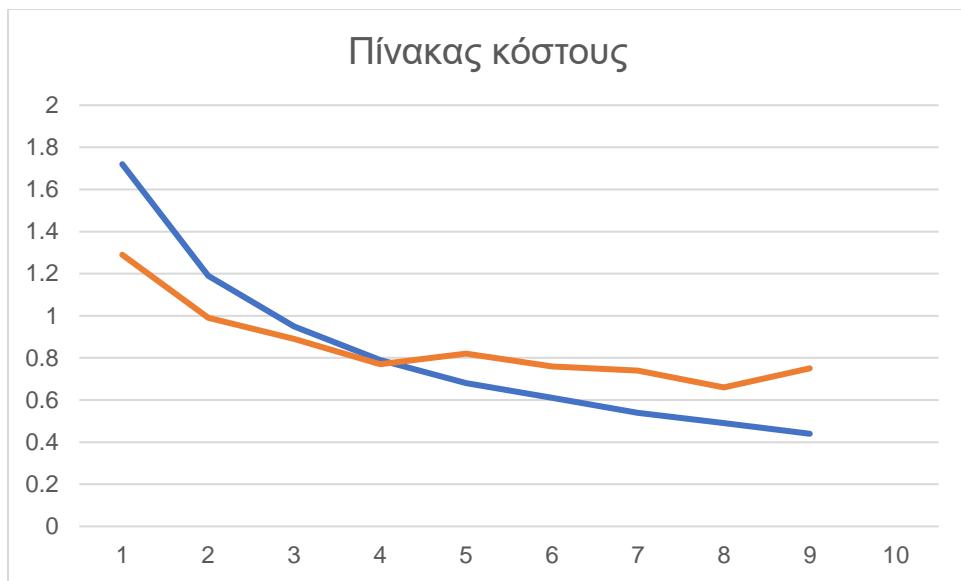
Τέλος ακολουθεί η διαδικασία του FCN δικτύου. Αφού μειωθούν περεταίρω οι διαστάσεις του διανύσματος θα ακολουθήσει ένα FC στρώμα για να δοθεί η τελική πρόβλεψη από τον SoftMax classifier.



Εικόνα 67 : Τα στρώματα που απαρτίζουν το δίκτυο FDDSC

Οι συνολικοί παράμετροι προς εκπαίδευση του δικτύου είναι 2,268,534.

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:

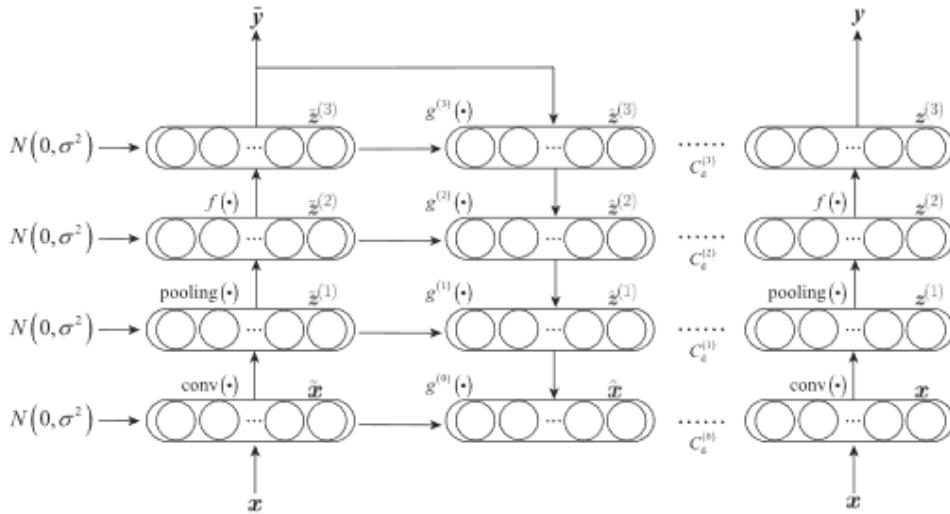


Πίνακες 14,15: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ εκπαίδευσης-επαλήθευσης

Το δίκτυο θα εκπαιδεύεται μέχρι την εποχή 4 καθώς από εκεί και έπειτα η ακρίβεια του σετ επαλήθευσης μειώνεται ενώ του σετ εκπαίδευσης αυξάνεται. Το ίδιο συμπέρασμα προκύπτει παρατηρώντας και τον πίνακα κόστους.

4.15 LIU ET AL.

Το δίκτυο Liu et al. [9] βασίστηκε στην σχετική δημοσίευση. Η αρχιτεκτονική του δικτύου αποτελείται από τον καθαρό κωδικοποιητή, τον «βρόμικο κωδικοποιητή» και τον αποκωδικοποιητή όπως ονομάζεται από τους δημιουργούς. Το ζευγάρι «βρόμικος κωδικοποιητής»-Αποκωδικοποιητής πρόκειται για ένα δίκτυο της μορφής autoencoder. Το δίκτυο αυτό ανασυνθέτοντας το σήμα που εισέρχεται θα ενισχύσει την μάθηση του καθαρού αποκωδικοποιητή, δηλαδή του CNN δικτύου που δίνει την τελική πρόβλεψη. Ο ρυθμός μάθησης είναι ίσος με 0,0001, το πακέτο των δεδομένων στην είσοδο ίσο με 128 και η μέθοδος βελτιστοποίησης είναι η SGD.



Εικόνα 68:Η αρχιτεκτονική του δικτύου Liu et al.

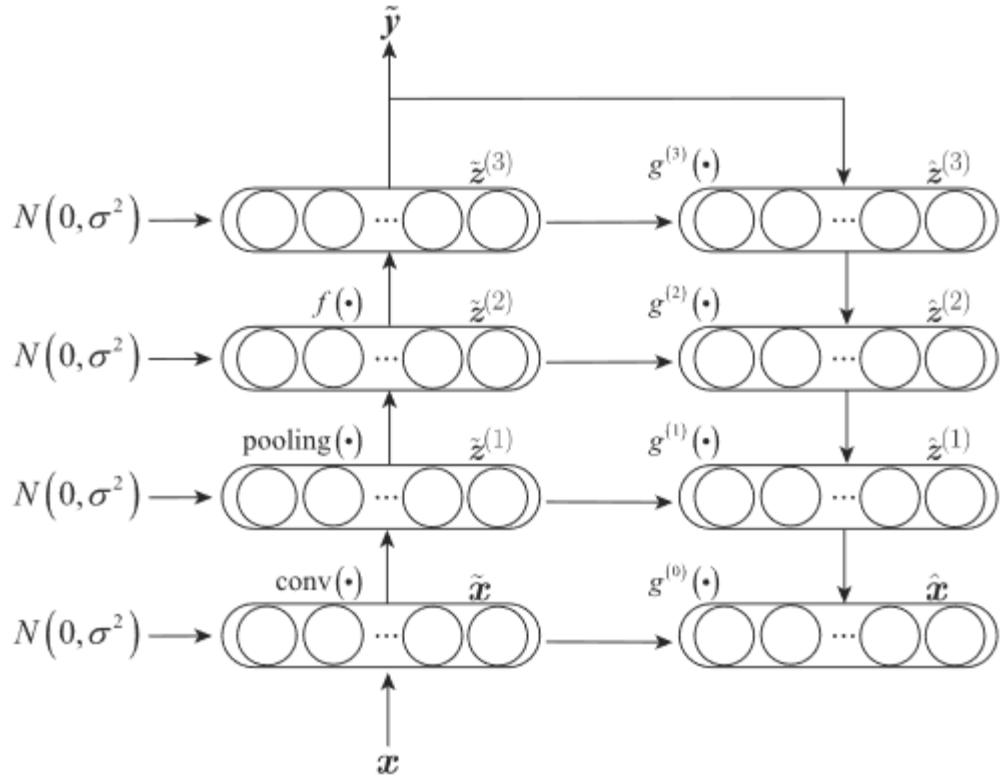
Στα αριστερά του πάνω σχήματος, φαίνεται η πορεία του σήματος x καθώς περνά από το CNN δίκτυο μέχρι να καταλήξει στην πρόβλεψη y . Το σήμα x είναι της μορφής $9 \times 9 \times 176$ όπου 9×9 η περιοχή των εικονοστοιχείων και 176 το φασματικό βάθος. Το δίκτυο αυτό είναι ένα τυπικό CNN δίκτυο που αποτελείται από ένα συνελικτικό στρώμα με φίλτρο 3×3 , ένα στρώμα pool, ένα στρώμα κανονικοποίησης και τέλος ένα πλήρως συνδεδεμένο στρώμα για να καταλήξει στην πρόβλεψη.



Η πορεία αυτή χαρακτηρίζει τον «βρόμικο κωδικοποιητή». Η ονομασία αυτή προέκυψε από το γεγονός ότι σε κάθε στρώμα της πορείας αυτής προστίθεται θόρυβος στο σήμα. Ο θόρυβος αυτός ακολουθεί την κανονική κατανομή και η πρόσθεση του έχει σαν σκοπό να εκπαιδεύσει το δίκτυο να απαλλάσσει τον θόρυβο από τα δεδομένα που εισέρχονται σε αυτό, χαρίζοντας ευρωστία και αποτελεσματικότητα.

Στη συνέχεια η πρόβλεψη για θα επιστρέψει στην αρχική μορφή του σήματος x, ανασυνθέτοντας το σήμα. Μέσα από πλήρως συνδεδεμένα στρώματα και στρώματα κανονικοποίησης το σήμα μεταλλάσσεται σταδιακά με την αντίθετη φορά έως ότου καταλήξει στην αρχική μορφή του. Αποτέλεσμα της διαδικασίας αυτής είναι η εκπαίδευση του δικτύου ώστε να αποβάλει τον θόρυβο που μπορεί να έχει η εικόνα προς ταξινόμηση, επιστρέφοντας το σήμα x «καθαρό».

Μεταξύ των 2 αυτών κατευθύνσεων παρατηρείται για άλλη μια φορά η εμφάνιση του residual learning. Όπως φαίνεται και στην εικόνα το κάθε στρώμα συνδέεται με skip connection στο αντίστοιχο στρώμα της ανασύνθεσης του σήματος. Συνεπώς η ανασύνθεση αυτή εξελίσσεται με βάση το αμέσως προηγούμενο στρώμα και το αντίστοιχο στρώμα της αρχικής πορείας του σήματος μέχρι την πρόβλεψη. Το γεγονός αυτό βελτιώνει την απαλλαγή του θορύβου από το δίκτυο και επιτρέπει την διατήρηση λεπτομερέστερων συσχετισμών που δεν διατηρήθηκαν στην πορεία των στρωμάτων, αφού επανεισέρχονται παρακάμπτοντας τα ενδιάμεσα στρώματα.



Εικόνα 69: Αναπαράσταση του δικτύου Κωδικοποιητής-Αποκωδικοποιητής.

Στη συνέχεια ο καθαρός κωδικοποιητής παρουσιάζεται με την ίδια ακριβώς αρχιτεκτονική με τον «βρόμικο». Τα στρώματα τους είναι κοινά όχι μόνο ως προς την μορφή αλλά και στις παραμέτρους προς μάθηση. Ουσιαστικά πρόκειται για 2 διαφορετικές πορείες που παίρνει το σήμα κατά την διάρκεια εκπαίδευσης του δικτύου πάνω στα ίδια στρώματα. Βασική διαφορά αποτελεί η πρόσθεση θορύβου που στην περίπτωση του καθαρού αποκωδικοποιητή απουσιάζει.

Η αποτελεσματικότητα του δικτύου βασίζεται στο γεγονός ότι οι κοινοί παράμετροι των στρωμάτων των 2 κωδικοποιητών εκπαιδεύονται ταυτοχρόνως από τα χαρακτηρισμένα και τα αχαρακτήριστα εικονοστοιχεία, σε αντίθεση με ένα τυπικό CNN. Ο καθαρός αποκωδικοποιητής λειτουργεί σαν ένα τυπικό CNN και η τελική πρόβλεψη των εικονοστοιχείο δίνεται από αυτόν. Τα στρώματα που περιλαμβάνει όμως δεν εκπαιδεύονται μόνο με βάση την σωστή πρόβλεψη του

εισερχόμενου σήματος. Η συνάρτηση κόστους του δικτύου αποτελείται από 2 επιμέρους κόστη, το supervised και το unsupervised κόστος.

Το supervised κόστος αποτελείται από το τυπικό cross-entropy loss σχετικά με την τελική πρόβλεψη από τον καθαρό αποκωδικοποιητή αλλά και από τον «βρόμικο», του οποίου η πρόβλεψη δεν χρησιμοποιείται πραγματικά αλλά με αυτόν τρόπο εγγυάται η δυνατότητα απαλλαγής του θορύβου. Τα στρώματα αυτά αναφέρονται στην ίδια κατάσταση του σήματος καθώς το σήμα ανασυντίθεται με τον ίδιο τρόπο που αποσυντίθεται έως την τελική πρόβλεψη. Η ελαχιστοποίηση των σφαλμάτων αυτών ενισχύει την δυνατότητα απαλλαγής του θορύβου περεταίρω. Την πορεία αυτή ακολουθούν όλα τα χαρακτηρισμένα εικονοστοιχείο και το supervised κόστος προκύπτει από αυτά.

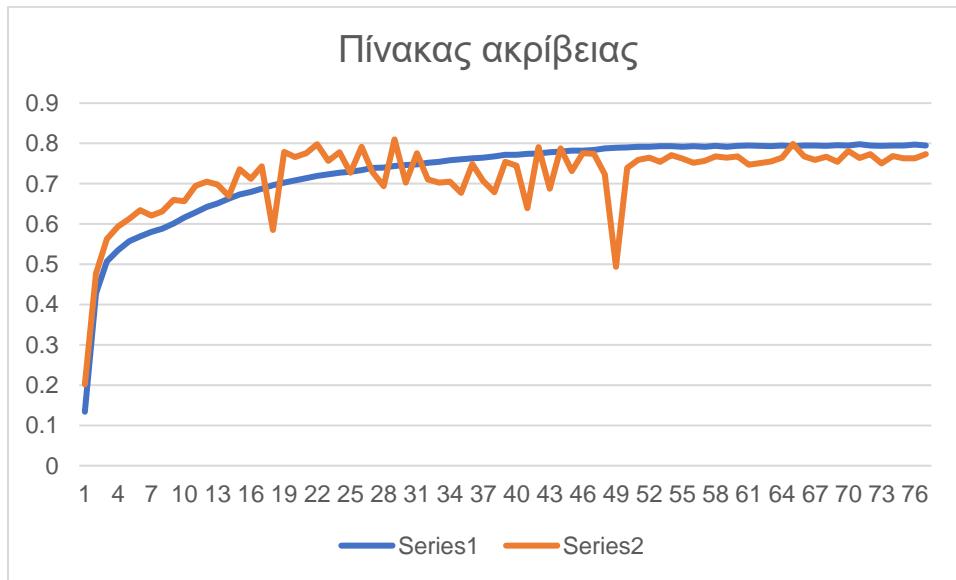
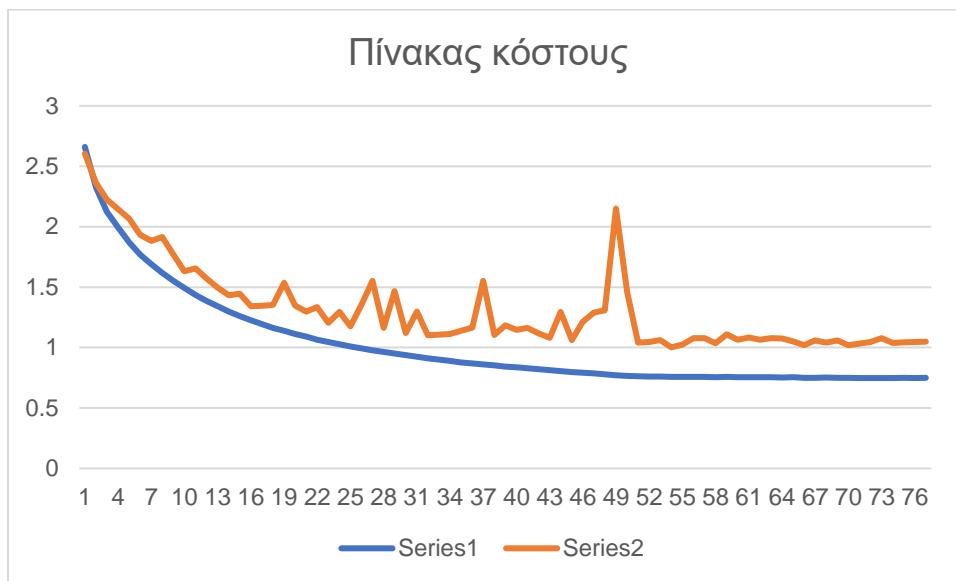
Στο unsupervised κόστος δεν υπάρχει η δυνατότητα για κόστος σχετικό με την πρόβλεψη, αφού τα εικονοστοιχεία από τα οποία προκύπτει είναι τα αχαρακτήριστα. Αποτελείται από τα αντίστοιχα μέσα τετραγωνικά σφάλματα καθαρού κωδικοποιητή-αποκωδικοποιητή των αχαρακτήριστων και χαρακτηρισμένων εικονοστοιχείων, λαμβάνοντας έναν βοηθητικό ρόλο και ενισχύοντας την μάθηση.

Network :

Layer (type)	Output Shape	Param #
<hr/>		
Conv2d-1	[-1, 80, 7, 7]	126,800
BatchNorm2d-2	[-1, 80, 7, 7]	160
MaxPool2d-3	[-1, 80, 3, 3]	0
Linear-4	[-1, 15]	10,815
Conv2d-5	[-1, 80, 7, 7]	126,800
BatchNorm2d-6	[-1, 80, 7, 7]	160
MaxPool2d-7	[-1, 80, 3, 3]	0
Linear-8	[-1, 15]	10,815
Linear-9	[-1, 720]	11,520
BatchNorm1d-10	[-1, 720]	1,440
Linear-11	[-1, 720]	519,120
BatchNorm1d-12	[-1, 720]	1,440
Linear-13	[-1, 3920]	2,826,320
BatchNorm1d-14	[-1, 3920]	7,840
Linear-15	[-1, 14256]	55,897,776
BatchNorm1d-16	[-1, 14256]	28,512
<hr/>		
Total params:	59,569,518	
Trainable params:	59,569,518	
Non-trainable params:	0	
<hr/>		
Input size (MB):	0.05	
Forward/backward pass size (MB):	0.43	
Params size (MB):	227.24	
Estimated Total Size (MB):	227.72	

Στην εικόνα παρατηρείται ο καθαρός κωδικοποιητής (1,4), ο «βρόμικος» κωδικοποιητής (5,8) και ο αποκωδικοποιητής (9,16).

Οι πίνακες του κόστους και της ακρίβειας φαίνονται εδώ, όπου το μπλε χρώμα αναφέρεται στο σετ εκπαίδευσης και το πορτοκαλί χρώμα στο σετ επαλήθευσης:



Πίνακες 16,17: Διάγραμμα κόστους/εποχή-ακρίβειας/εποχή, σετ [εκπαίδευσης-επαλήθευσης](#)

Ο πίνακας κόστους αναφέρεται στο supervised κόστος, και όχι στο σύνολο του κόστους που τείνει να μειωθεί στο δίκτυο, αφού ο σκοπός της εφαρμογής είναι η παρατήρηση του κόστους της ταξινόμησης. Οπότε ενώ το δίκτυο εκπαιδεύεται και με βάση το unsupervised κόστος ,στα διαγράμματα παρατηρείται μόνο το supervised.

Στο συγκεκριμένο μοντέλο παρατηρείται αρκετή διακύμανση στο τεστ επαλήθευσης ,η οποία δεν μπόρεσε να ομαλοποιηθεί με την κατάλληλη χρήση του ρυθμού μάθησης. Η εποχή 54 επιλέγεται ως καταλληλότερη για να σταματήσει η εκπαίδευση αφού στο σημείο αυτό σταθεροποιείται το κόστος αλλά και η ακρίβεια των 2 σετ.

4.16 ΟΙ ΜΕΘΟΔΟΙ SVM ΚΑΙ RF

4.16.1 Μηχανή Διανυσματικής Υποστήριξης (SVM)

Οι Μηχανές Διανυσματικής Υποστήριξης αποτελούν αλγόριθμους επιβλεπόμενης μάθησης που υπάγονται στην οικογένεια της Μηχανικής Μάθησης. Χρησιμοποιούνται σε διάφορες εφαρμογές ταξινόμησης δεδομένων, αναγνώρισης προτύπων κλπ. Ο αλγόριθμος αυτός προσπαθεί να βρει το βέλτιστο όριο απόφασης, ένα διάνυσμα στον N-χώρο το οποίο θα διαχωρίζει τα δείγματα σε κατηγορίες αναζητώντας την μεγαλύτερη απόσταση μεταξύ τους. Η μορφή του διανύσματος αυτού εξαρτάται από την συνάρτηση πυρήνα της μεθόδου, μπορεί να είναι γραμμική ή καμπύλη. Για τα δεδομένα της εργασίας θα χρησιμοποιηθεί η μέθοδος SVM με την χρήση της γραμμικής συνάρτησης πυρήνα.

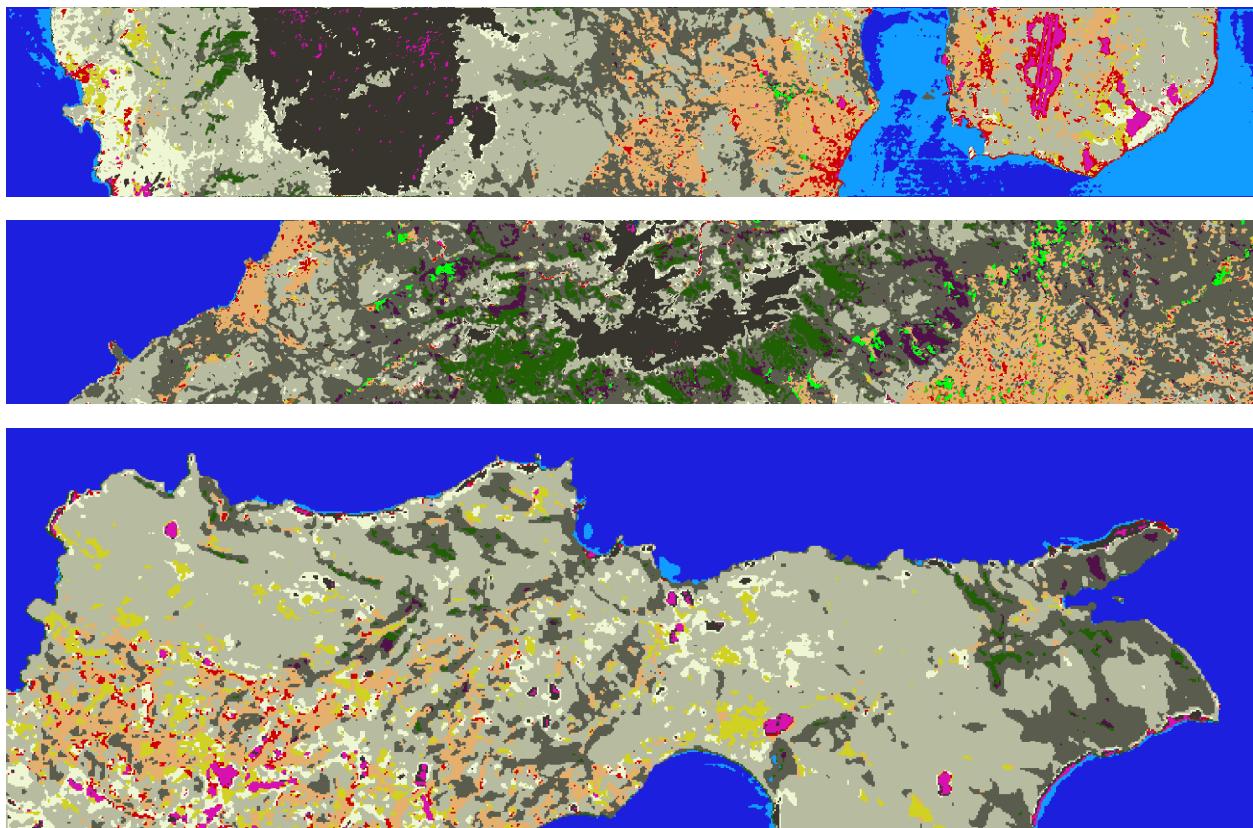
4.16.2 Μέθοδος Τυχαίων Δασών (RF)

Η μέθοδος Τυχαίων Δασών (Random forest) είναι από τις δημοφιλέστερες μεθόδους μαζί με την μέθοδο SVM στην κατηγορία τους. Είναι μία μέθοδος αρκετά γρήγορη που μπορεί να επεξεργαστεί πολλά δεδομένα στην εισαγωγή. Η ταξινόμηση με αυτήν την μέθοδο βασίζεται στην ύπαρξη δέντρων αποφάσεων. Τα δεδομένα εισέρχονται στις δομές των δέντρων αυτών και διαχωρίζονται σταδιακά μέχρι να προκύψει η κατηγορία στην οποία ανήκουν. Τελικά η συνηθέστερη πρόβλεψη επικρατεί.

5 ΑΠΟΤΕΛΕΣΜΑΤΑ

Στο παρών κεφάλαιο θα παρουσιαστούν και θα αξιολογηθούν τα αποτελέσματα των ταξινομήσεων πάνω στις υπερφασματικές εικόνες. Αρχικά θα παρουσιαστούν οι προβλέψεις των εικόνων, δηλαδή οι 3 εικόνες του σετ τεσταρίσματος ταξινομημένες στις 14 κατηγορίες του HyRANK dataset. Επειτα θα ακολουθήσουν οι πίνακες σύγχυσης και στατιστικών για κάθε κατηγορία.

5.1 HAMIDA ET AL.



Εικόνα 70 :Οι προβλέψεις στις εικόνες Kirkki,Erato,Nefeli

HAMIDA ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	36.34	80.67	33.43	50.11
2	Mineral Extraction Sites	591	97.29	32.89	32.60	49.17
3	Non-Irrigated Arable Land	1826	48.36	65.60	38.58	55.67
4	Fruit Trees	715	11.75	86.60	11.54	20.69
5	Olive Groves	5352	56.97	78.28	49.19	65.95
6	Broad-leaved Forest	610	41.97	50.00	29.56	45.63
7	Coniferous Forest	4224	62.55	59.75	44.00	61.11
8	Mixed Forest	3005	32.05	53.71	25.11	40.14
9	Dense Sclerophyllous Veg.	5824	88.19	51.17	47.88	64.76
10	Sparce Sclerophyllous Veg.	10769	80.07	69.14	58.99	74.21
11	Sparcely Vegetated Areas	2160	22.82	40.15	17.03	29.10
12	Rocks and Sand	18051	92.45	97.05	89.92	94.69
13	Water	2534	77.90	77.69	63.66	77.79
14	Coastal Water	1083	46.17	47.17	30.43	46.66
OA%						71.95
Macro-avg			56.78	63.56	41.00	55.41
Weighted-avg			71.95	74.23	58.77	71.19
Kappa %						71.92
Train time(s)		1714				
Number of parameters		4190				

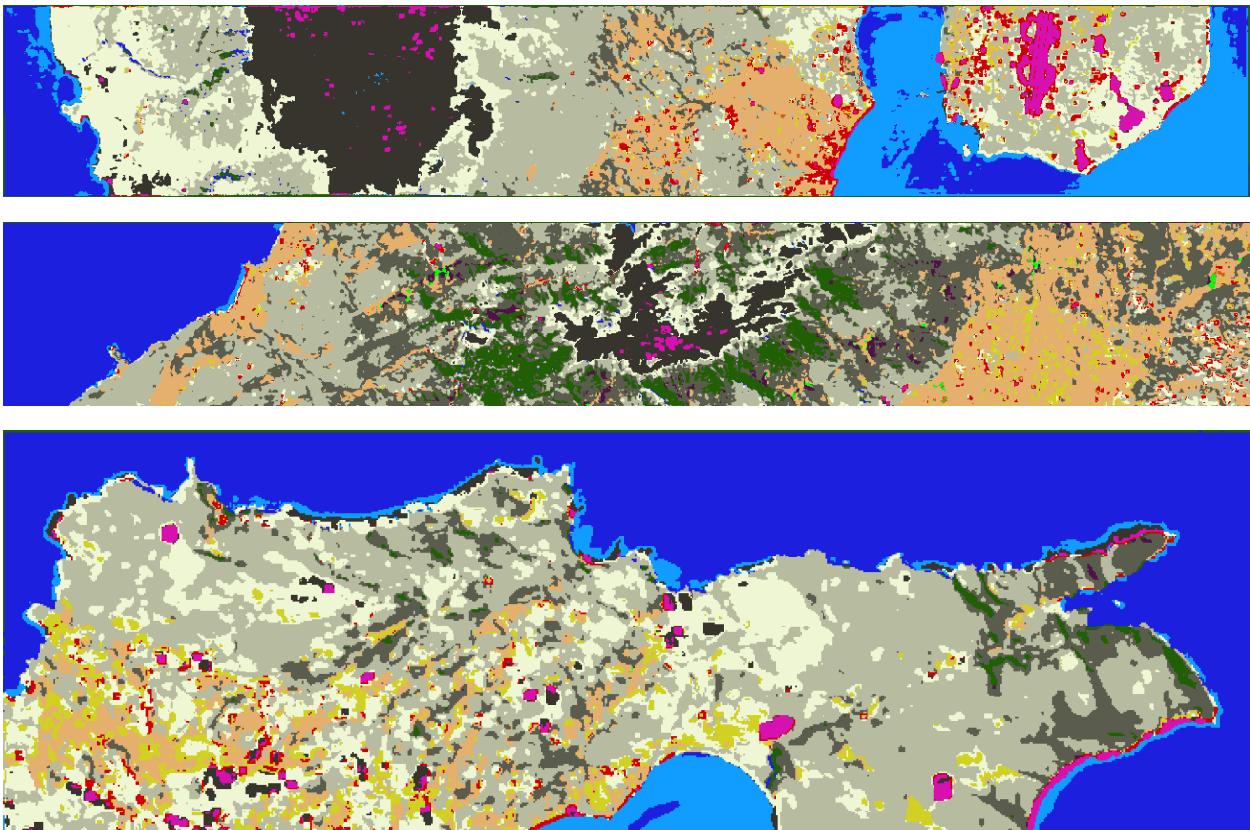
Πίνακας 18: Στατιστικά των προβλέψεων

Η μέθοδος Hamida et al. είναι από τις αποτελεσματικές μεθόδους της συγκεκριμένης εφαρμογής. Οι εικόνες δείχνουν να έχουν ταξινομηθεί σχετικά ομαλά. Παρατηρείται στην εικόνα Kirki μία σύγχυση ανάμεσα στα εικονοστοιχεία του νερού και του παράκτιου νερού, με αποτέλεσμα το παράκτιο νερό στα δεξιά της εικόνας να κυριαρχεί λανθασμένα. Το ίδιο παρατηρείται και στην εικόνα Nefeli όπου εικονοστοιχεία παράκτιου νερού περιβάλλονται από εικονοστοιχεία νερού. Το φαινόμενο αυτό επιβεβαιώνει και ο πίνακας στατιστικών με το χαμηλό σκορ που δίνει σε αυτή την κατηγορία.

Στις υπόλοιπες κατηγορίες-εικόνες δεν παρατηρείται κάποια ανωμαλία. Γενικότερα ο πίνακας στατιστικών αναδεικνύει μια ποιοτική και ισορροπημένη ταξινόμηση ανά κατηγορία, και μη χρονοβόρα. Πρόκληση αποτελεί η κατηγορία των δέντρων με φρούτα όπου η μέθοδος παρουσιάζει το χαμηλότερο σκορ μαζί με τις διάσπαρτες καλλιέργειες ενώ οι πέτρες και το χώμα ταξινομούνται αρκετά εύκολα.

Ειδικότερα η κατηγορία των δέντρων με φρούτα εμφανίζει ιδιαίτερα υψηλή ακρίβεια χρήστη αλλά χαμηλή ακρίβεια παραγωγού. Το γεγονός αυτό σημαίνει πως ενώ τα τελικά ταξινομημένα εικονοστοιχεία που αναφέρονται στην κατηγορία αυτή ανταποκρίνονται στην πραγματικότητα, η συγκεκριμένη μέθοδος δεν κατόρθωσε να προβλέψει το σύνολο των εικονοστοιχείων που αντιπροσωπεύουν την κατηγορία. Οι υπόλοιπες κατηγορίες που εμφανίζουν χαμηλή ακρίβεια παραγωγού δεν διαθέτουν μεγάλη διαφορά σε σχέση με την ακρίβεια χρήστη.

5.2 LEE ET AL.



Εικόνα 71 : Οι προβλέψεις στις εικόνες Kirki,Erato,Nefeli

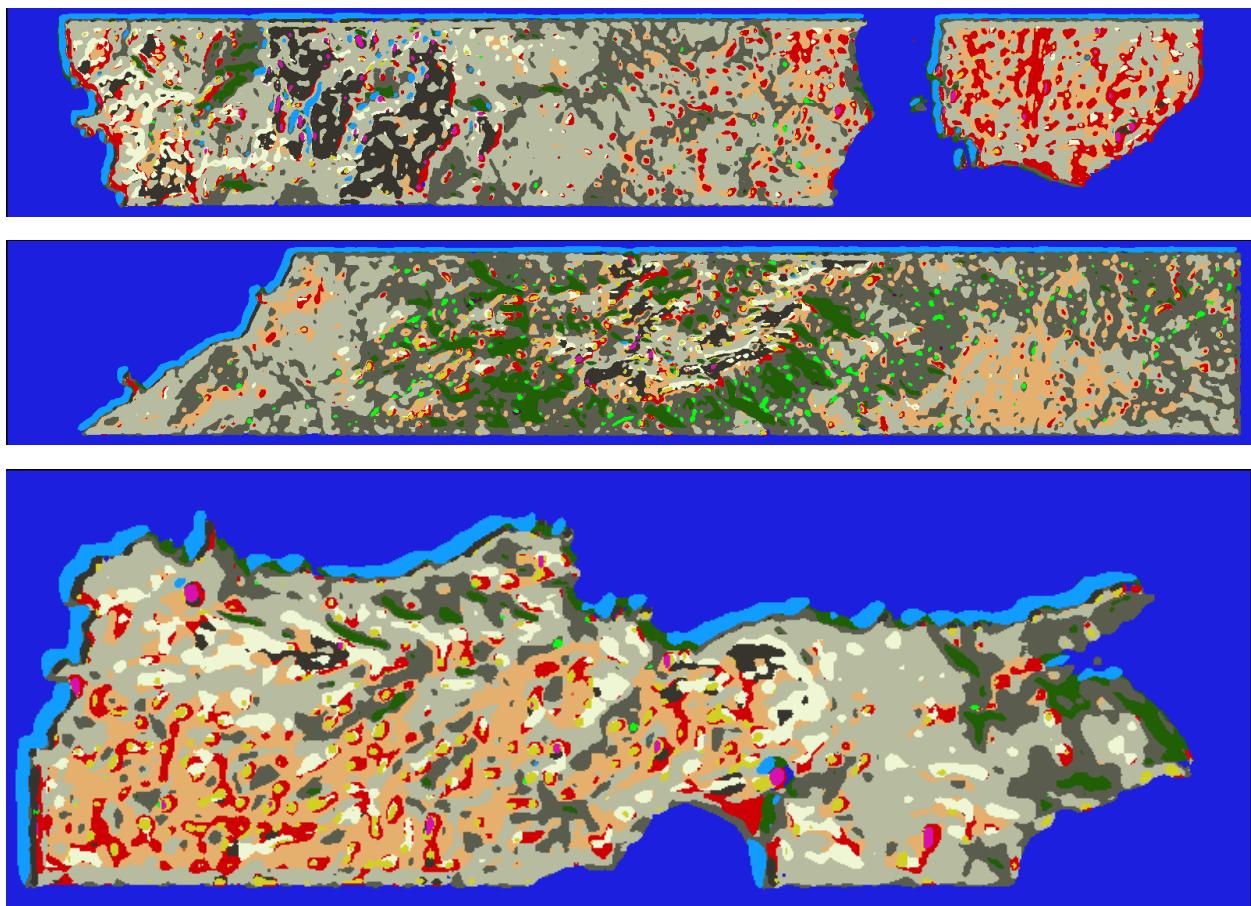
LEE ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	24.85	76.08	23.05	37.47
2	Mineral Extraction Sites	591	100.00	25.64	25.64	40.81
3	Non-Irrigated Arable Land	1826	38.94	42.32	25.44	40.56
4	Fruit Trees	715	1.40	76.92	1.39	2.75
5	Olive Groves	5352	52.84	69.64	42.95	60.09
6	Broad-leaved Forest	610	14.26	63.04	13.16	23.26
7	Coniferous Forest	4224	55.16	57.63	39.25	56.37
8	Mixed Forest	3005	8.39	50.40	7.75	14.38
9	Dense Sclerophyllous Veg.	5824	79.77	62.08	53.64	69.82
10	Sparce Sclerophyllous Veg.	10769	87.92	70.88	64.59	78.48
11	Sparcely Vegetated Areas	2160	64.58	34.99	29.36	45.39
12	Rocks and Sand	18051	93.27	97.15	90.79	95.17
13	Water	2534	76.56	84.49	67.13	80.33
14	Coastal Water	1083	80.33	57.54	50.43	67.05
OA%						71.55
Macro-avg			55.59	62.06	38.00	50.85
Weighted-avg			71.55	73.85	58.53	69.90
Kappa %						52.91
Train time(s)		801				
Number of parameters		939151				

Πίνακας 19: Στατιστικά των προβλέψεων

Η μέθοδος Lee et al. αποτελεί και αυτή μία μέθοδο αποτελεσματική και γρήγορη. Οι 2 κατηγορίες του νερού φαίνονται ταξινομημένες ομαλότερα σε σχέση με την μέθοδο Hamida et al.. Επίσης στις εικόνες δίνεται η αίσθηση περισσότερων στοιχείων της κατηγορίας 2. Χαρακτηριστικά στο κέντρο της εικόνας Erato υπάρχουν στοιχεία της κατηγορίας 2 που δεν υπάρχουν στην αντίστοιχη πρόβλεψη της εικόνας από την μέθοδο Hamida et al.. Η ακρίβεια του παραγωγού αγγίζει το 100% και φανερώνει πως η μέθοδος διαχώρισε και προέβλεψε πλήρως την κατηγορία σε σχέση με τις άλλες, αλλά οι προβλέψεις της δεν ανταποκρίνονται στην πραγματικότητα.

Η κατηγορίες 4,6,7 έχουν χαμηλά σκορ. Πιο συγκεκριμένα η μέθοδος σε αυτήν την περίπτωση δεν καταφέρνει να διαχωρίσει τις κατηγορίες μεταξύ τους οδηγώντας σε χαμηλή ακρίβεια την ταξινόμηση.

5.3 CHEN ET AL.



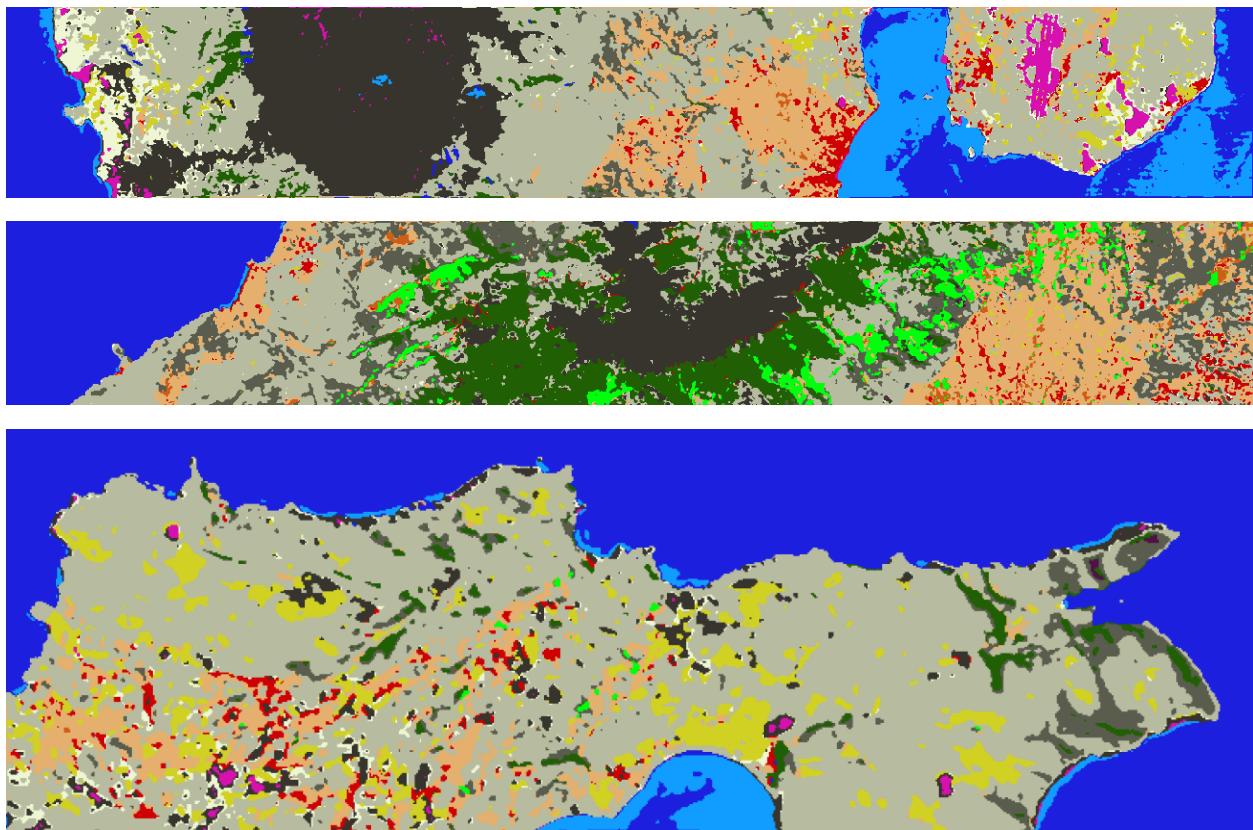
Εικόνα 72 : Οι προβλέψεις στις εικόνες Kirk,Erato,Nefeli

CHEN ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS	F1%
1	Dense Urban Fabric	2917	40.59	43.43	26.55	41.96
2	Mineral Extraction Sites	591	49.07	29.87	22.80	37.13
3	Non-Irrigated Arable Land	1826	21.30	25.15	13.04	23.07
4	Fruit Trees	715	0.00	0.00	0.00	0.00
5	Olive Groves	5352	32.66	39.54	21.78	35.77
6	Broad-leaved Forest	610	50.49	66.38	40.21	57.36
7	Coniferous Forest	4224	55.47	46.15	33.67	50.38
8	Mixed Forest	3005	0.07	8.33	0.07	0.13
9	Dense Sclerophyllous Veg.	5824	77.88	40.25	36.12	53.07
10	Sparce Sclerophyllous Veg.	10769	68.21	49.97	40.53	57.69
11	Sparcely Vegetated Areas	2160	25.74	19.41	12.44	22.13
12	Rocks and Sand	18051	61.76	95.93	60.19	75.15
13	Water	2534	100.00	84.47	84.47	91.58
14	Coastal Water	1083	52.91	60.00	39.11	56.23
OA%						55.24
Macro-avg			45.44	43.49	31.00	42.97
Weighted-avg			55.24	58.46	40.48	54.54
Kappa %						47.84
Train time(s)		16254				
Number of parameters		6008687				

Πίνακας 20: Στατιστικά των προβλέψεων

Η μέθοδος Chen et al. αποτελεί την πιο αργή μέθοδο της εφαρμογής. Η ακρίβεια της ταξινόμησης της είναι μέτρια αλλά εμφανίζει μια σχετική σταθερότητα στα επιμέρους σκορ των κλάσεων, αν εξαιρεθούν οι κατηγορίες 4,8 που όλες οι μέθοδοι παρουσιάζουν προβλήματα. Το γεγονός αυτό επιβεβαιώνεται και οπτικά καθώς οι προβλέψεις είναι «γεμάτες» και παρουσιάζουν ποικιλία στις χρήσεις γης.

Ενώ στην εικόνα Erato δεν παρατηρείται κάποιο πρόβλημα, το αεροδρόμιο στην εικόνα Kirkι χαρακτηρίζεται ως αστικός ιστός και δεν γίνεται αντιληπτό. Το παράκτιο νερό δείχνει αληθιοφανές, αλλά στα όρια της εικόνας ταξινομείται στα κενά εικονοστοιχεία χωρίς όμως να δημιουργεί πρόβλημα στην ανάγνωση της πρόβλεψης.



Εικόνα 73 : Οι προβλέψεις στις εικόνες Kirkki,Erato,Nefeli

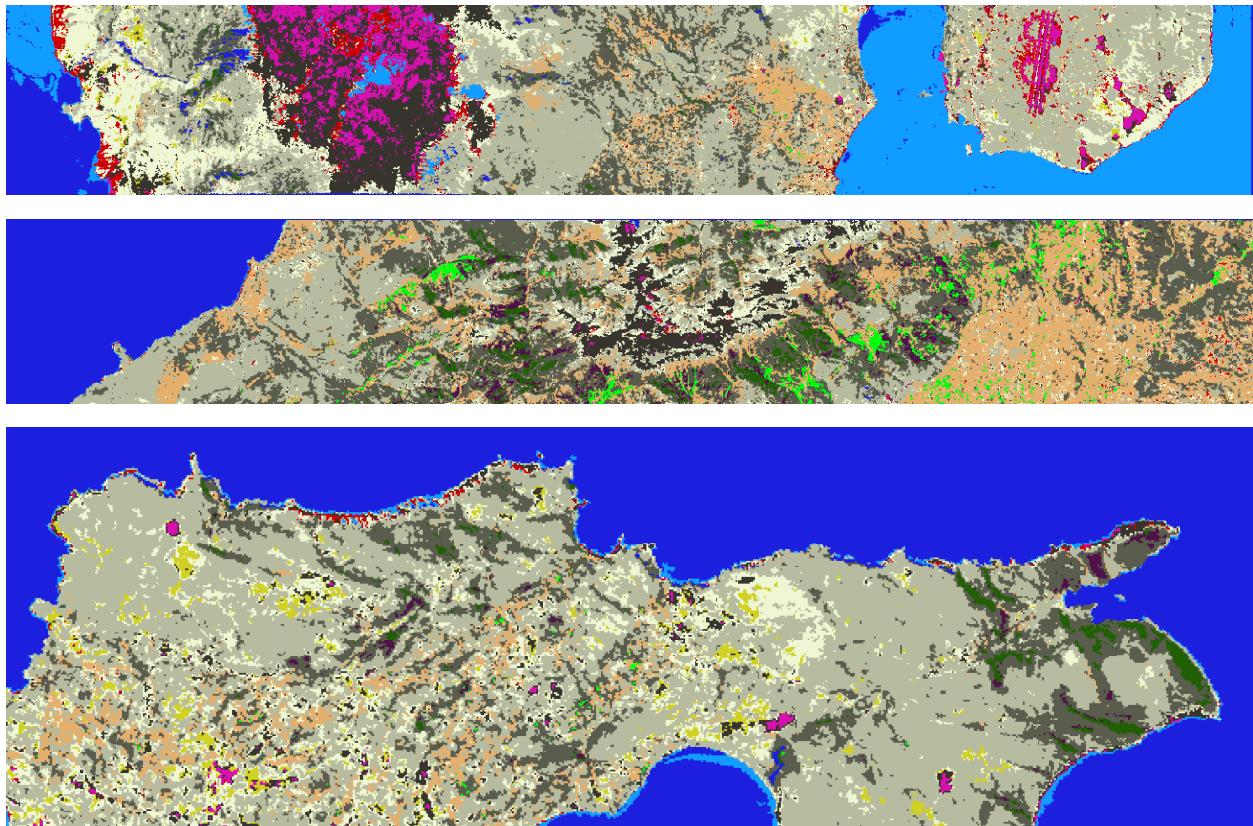
LI ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	48.71	72.65	41.16	58.32
2	Mineral Extraction Sites	591	95.94	44.82	43.99	61.10
3	Non-Irrigated Arable Land	1826	65.83	65.01	48.60	65.41
4	Fruit Trees	715	18.04	81.65	17.34	29.55
5	Olive Groves	5352	66.93	77.26	55.92	71.73
6	Broad-leaved Forest	610	90.16	48.33	45.91	62.93
7	Coniferous Forest	4224	79.10	54.57	47.69	64.59
8	Mixed Forest	3005	4.79	85.21	4.75	9.07
9	Dense Sclerophyllous Veg.	5824	72.56	77.78	60.11	75.08
10	Sparce Sclerophyllous Veg.	10769	92.00	74.06	69.58	82.06
11	Sparcely Vegetated Areas	2160	9.17	58.41	8.60	15.85
12	Rocks and Sand	18051	98.00	91.50	89.82	94.64
13	Water	2534	80.19	83.76	69.40	81.94
14	Coastal Water	1083	66.39	49.31	39.46	56.59
OA%						76.61
Macro-avg			63.42	68.88	46.00	59.20
Weighted-avg			76.61	77.76	62.82	73.78
Kappa %						68.60
Train time(s)		586				
Number of parameters		97455				

Πίνακας 21: Στατιστικά των προβλέψεων

Η μέθοδος Li et al. είναι η αποτελεσματικότερη μέθοδος που αναφέρεται σε νευρωνικό δίκτυο συγκριτικά με όλες όσες δοκιμάστηκαν. Παρουσιάζει μια ισορροπημένη συμπεριφορά ταξινομώντας τις περισσότερες κατηγορίες με υψηλή ακρίβεια. Αντιμετωπίζει την δύσκολη κατηγορία των δέντρων με φρούτα ικανοποιητικά αλλά η κατηγορία του μεικτού δάσους παρουσιάζει αρκετά χαμηλό σκορ.

Στην εικόνα Erato, ταξινομούνται πολλά εικονοστοιχείο σαν δάση με δέντρα με μεγάλα φύλλα. Ακόμα στην εικόνα Nefeli δεν παρατηρούνται τόσα πίξελ ελιάς όσα στις άλλες μεθόδους. Το ίδιο συμβαίνει και για την κατηγορία 2 που ούτε αυτή εμφανίζεται συχνά στις εικόνες. Το παράκτιο νερό για ακόμα μια φορά δείχνει να συγχέεται με το νερό, αποπροσανατολίζοντας την ταξινόμηση.

5.5 HU ET AL.



Εικόνα 74 : Οι προβλέψεις στις εικόνες Kirk, Erato, Nefeli

HU ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	8.64	31.27	0.07	13.54
2	Mineral Extraction Sites	591	93.23	8.33	0.08	15.30
3	Non-Irrigated Arable Land	1826	32.53	74.62	0.29	45.31
4	Fruit Trees	715	0.00	0.00	0.00	0.00
5	Olive Groves	5352	47.76	62.43	0.37	54.12
6	Broad-leaved Forest	610	79.84	55.66	0.49	65.59
7	Coniferous Forest	4224	42.33	63.56	0.34	50.82
8	Mixed Forest	3005	39.00	74.41	0.34	51.18
9	Dense Sclerophyllous Veg.	5824	81.18	48.55	0.44	60.76
10	Sparce Sclerophyllous Veg.	10769	81.33	71.76	0.62	76.24
11	Sparcely Vegetated Areas	2160	51.30	33.29	0.25	40.38
12	Rocks and Sand	18051	59.69	90.43	0.56	71.91
13	Water	2534	75.30	70.75	0.57	72.95
14	Coastal Water	1083	52.91	26.03	0.21	34.90
OA%						59.08
Macro-avg			53.22	50.79	0.33	46.64
Weighted-avg			59.08	68.05	0.45	60.55
Kappa %						71.23
Train time(s)		22				
Number of parameters		80035				

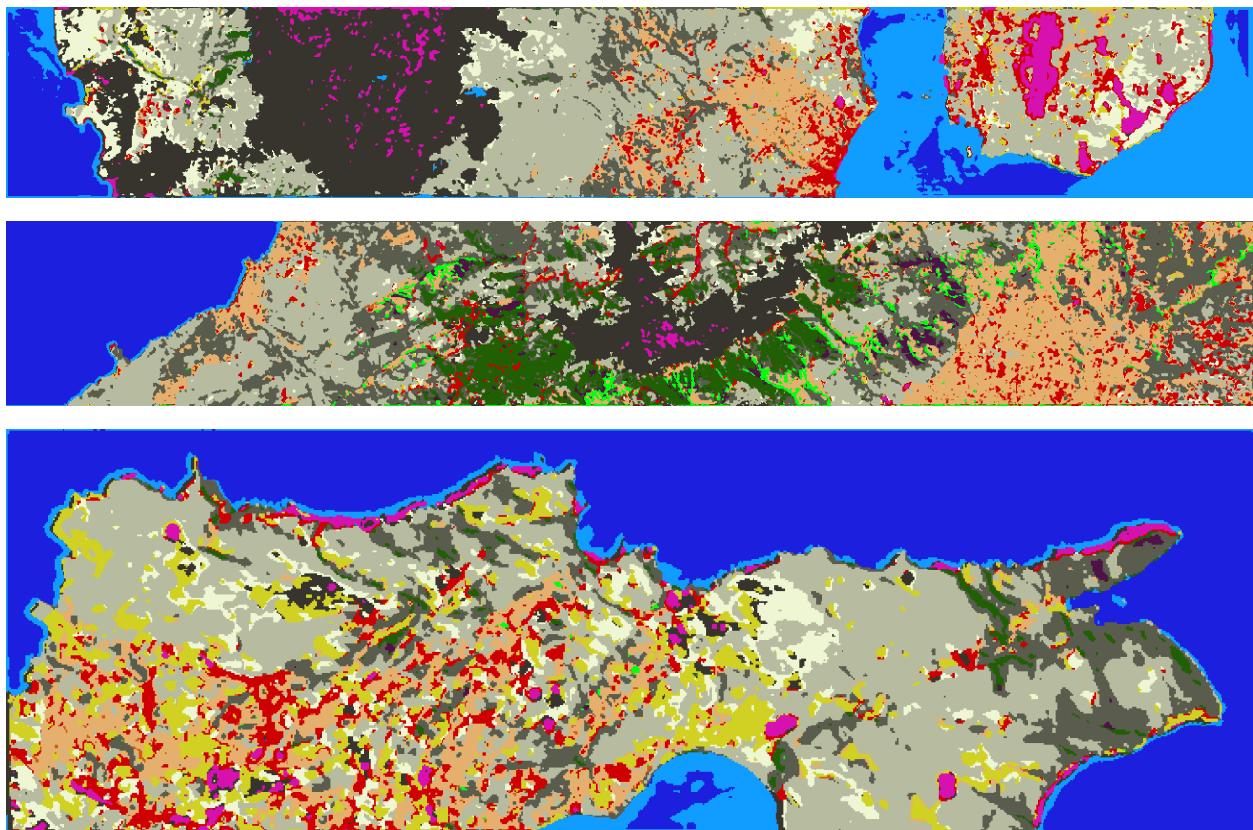
Πίνακας 22: Στατιστικά των προβλέψεων

Η μέθοδος Hu et al. αποτελεί την ταχύτερη σε εκπαίδευση μέθοδο από όλες τις μεθόδους που δοκιμάστηκαν. Στην εικόνα KirkI παρατηρείται η ύπαρξη πολλών εικονοστοιχείων της κατηγορίας 2, προς τα αριστερά της. Ο αριθμός αυτός οπτικά χαρακτηρίζεται μη φυσιολογικός. Παρόλο που η συγκεκριμένη περιοχή αποτελείται από χώμα και πέτρες σαν όρος, δεν δικαιολογείται το τόσο μεγάλο πλήθος των εικονοστοιχείων. Την οπτική παρατήρηση ενισχύει και ο πίνακας των στατιστικών που αναφέρει 8.33% ακρίβεια χρήστη, δηλαδή πολλά από τα ταξινομημένα στοιχεία δεν ανταποκρίνονται στην πραγματικότητα.

Άλλο φαινόμενο που παρατηρείται οπτικά είναι πάλι η ταξινόμηση σχετικά με τα εικονοστοιχεία του νερού. Η πρόβλεψη στην εικόνα KirkI δείχνει αφύσικη αφού το δεξιό τμήμα της περιέχει μόνο παράκτιο νερό. Στην πρόβλεψη της εικόνας Nefeli το πρόβλημα δεν είναι τόσο έντονο.

Η απουσία ταξινομημένων εικονοστοιχείων στην κατηγορία των δέντρων με φρούτα δεν αποτελεί ενδιαφέρον γεγονός αφού όλες οι μέθοδοι παρουσιάζουν χαμηλές ακρίβειες. Ενδιαφέρον αποτελεί η πολύ χαμηλή ακρίβεια στον εντοπισμό του αστικού ιστού, όπου στην θέση του ταξινομείται η κατηγορία των καλλιεργειών ελιάς και άλλες κατηγορίες. Εκτός από την οπτική παρατήρηση, ο πίνακας στατιστικών επαληθεύει το φαινόμενο αυτό με το χαμηλό ποσοστό ακρίβειας παραγωγών στην συγκεκριμένη κατηγορία, δείχνοντας ότι τα εικονοστοιχεία της κατηγορίας δεν εντοπίστηκαν τελικά.

5.6 HE ET AL.



Εικόνα 75 : Οι προβλέψεις στις εικόνες Kirk, Erato, Nefeli

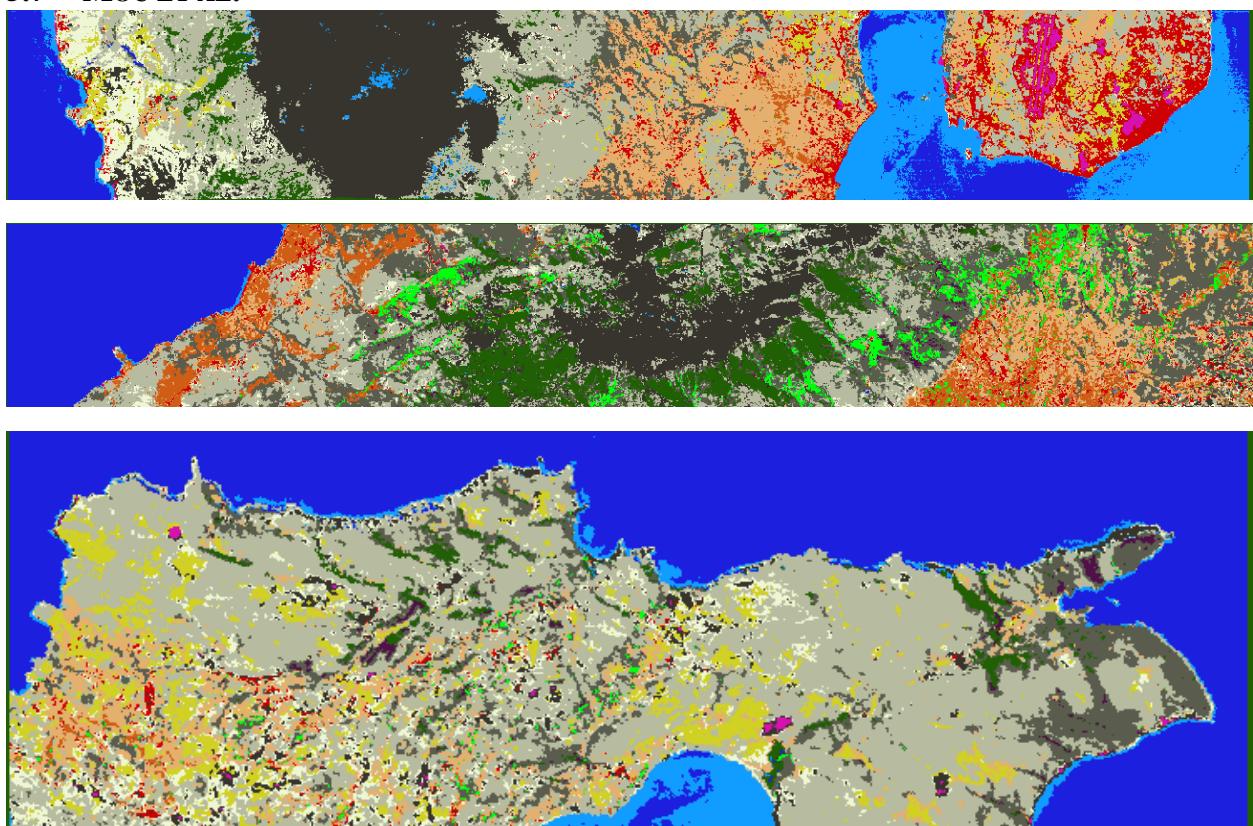
HE ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	58.00	72.37	47.49	64.40
2	Mineral Extraction Sites	591	100.00	19.33	19.33	32.39
3	Non-Irrigated Arable Land	1826	67.20	85.21	60.18	75.14
4	Fruit Trees	715	8.39	76.92	8.19	15.13
5	Olive Groves	5352	57.27	69.77	45.88	62.90
6	Broad-leaved Forest	610	75.57	76.32	61.22	75.95
7	Coniferous Forest	4224	61.81	58.29	42.86	60.00
8	Mixed Forest	3005	23.49	74.08	21.71	35.67
9	Dense Sclerophyllous Veg.	5824	81.58	65.08	56.74	72.40
10	Sparce Sclerophyllous Veg.	10769	87.28	74.99	67.60	80.67
11	Sparcely Vegetated Areas	2160	26.71	41.07	19.31	32.37
12	Rocks and Sand	18051	88.56	92.67	82.76	90.57
13	Water	2534	75.37	87.82	68.24	81.12
14	Coastal Water	1083	75.81	49.64	42.85	59.99
OA%						73.51
Macro-avg			63.36	67.40	46.00	59.91
Weighted-avg			73.51	76.20	60.48	73.15
Kappa %						63.14
Train time(s)		2324				
Number of parameters		22431				

Πίνακας 23: Στατιστικά των προβλέψεων

Η μέθοδος He et al. έχει επιτύχει προβλέψεις στις εικόνες με αρκετά πλούσιο αποτέλεσμα από την άποψη της αντιπροσωπευτικότητας των κατηγοριών. Έχει υπάρξει υπερταξινόμηση της κατηγορίας 2 σε παρόμοια κατάσταση με την μέθοδο Hu et al., με ακρίβεια χρήστη 13.33%. Γενικότερα όμως δεν παρουσιάζεται κάποιο ιδιαίτερα προβληματικό σημείο στην ταξινόμηση. Τα πάντα στις εικόνες δείχνουν φυσιολογικά. Η ταξινόμηση είναι αρκετά ποιοτική.

Η χειρότερη σε σκορ F1 κατηγορία είναι η κατηγορία των δέντρων με φρούτα όπου η ακρίβεια παραγωγού παραμένει σταθερά χαμηλή σε όλες σχεδόν τις μεθόδους.

5.7 MOU ET AL.



Εικόνα 76 : Οι προβλέψεις στις εικόνες Kirkki,Erato,Nefeli

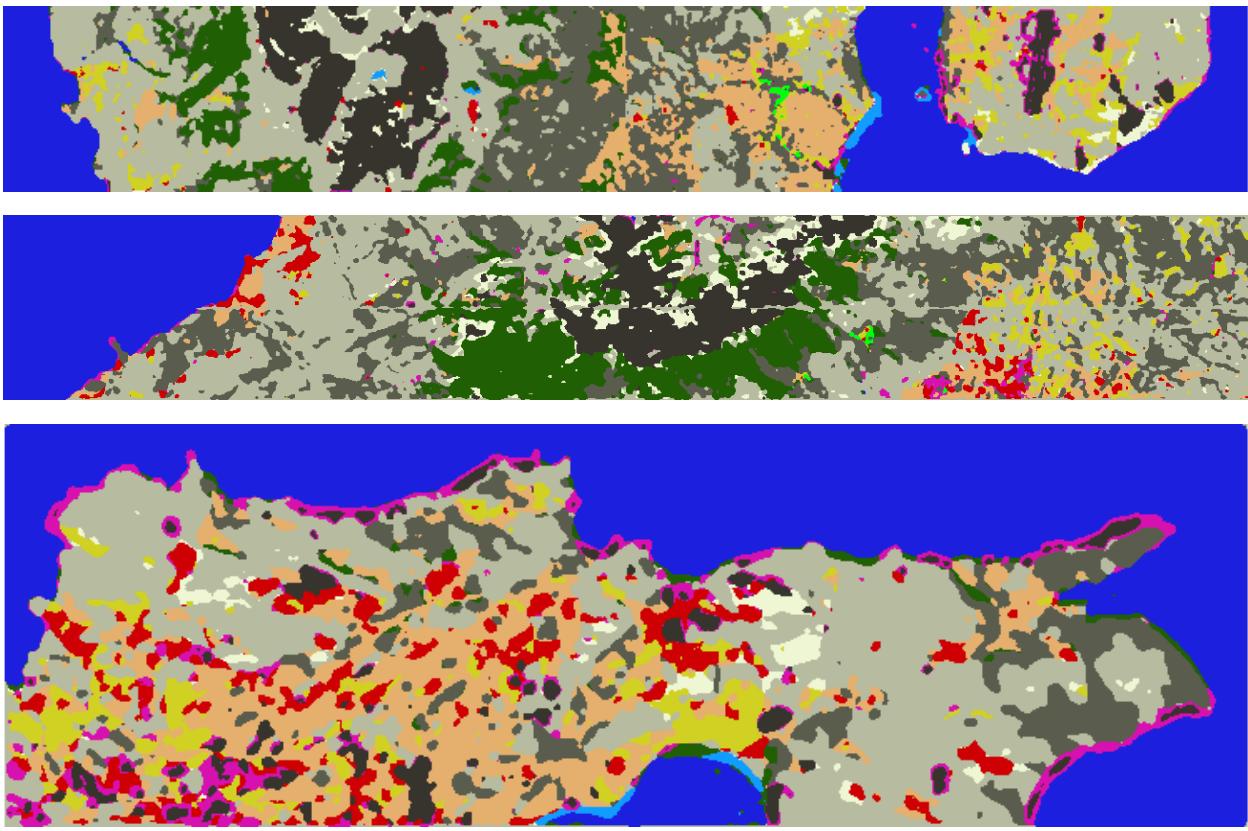
MOU ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	56.22	71.96	46.12	63.13
2	Mineral Extraction Sites	591	93.91	50.78	49.16	65.91
3	Non-Irrigated Arable Land	1826	63.96	63.62	46.83	63.79
4	Fruit Trees	715	54.55	23.61	19.73	32.95
5	Olive Groves	5352	56.99	81.31	50.39	67.01
6	Broad-leaved Forest	610	79.67	56.84	49.64	66.35
7	Coniferous Forest	4224	81.18	63.15	55.08	71.04
8	Mixed Forest	3005	11.75	72.34	11.24	20.21
9	Dense Sclerophyllous Veg.	5824	84.79	66.43	59.36	74.50
10	Sparce Sclerophyllous Veg.	10769	80.37	81.02	67.64	80.70
11	Sparcely Vegetated Areas	2160	25.28	51.36	20.40	33.88
12	Rocks and Sand	18051	96.16	92.06	88.79	94.06
13	Water	2534	76.95	83.73	66.94	80.20
14	Coastal Water	1083	65.28	36.94	30.87	47.18
OA%						75.80
Macro-avg			66.22	63.94	47.00	61.49
Weighted-avg			75.80	77.29	62.92	74.61
Kappa %						47.84
Train time(s)			127			
Number of parameters		191503				

Πίνακας 24: Στατιστικά των προβλέψεων

Η μέθοδος Mou et al. παρουσιάζει συγκρατημένα και με αρκετή ακρίβεια συγκριτικά με τις άλλες μεθόδους την κατηγορία 1. Στην εικόνα Ερατώ παρατηρείται έντονα στο δεξί τμήμα κατανομή του αστικού ιστού καθώς συνυπάρχει με τις καλλιέργειες ελιάς. Τα δέντρα με φρούτα εμφανίζουν την υψηλότερη ακρίβεια παραγωγού και κατά συνέπεια το υψηλότερο σκορ F1 σε σχέση με όλες τις μεθόδους Νευρωνικών δικτύων.

Το παράκτιο νερό αποτελεί μία προβληματική οπτικά κατηγορία όμως γενικότερα όλες οι κατηγορίες έχουν ανεκτές ακρίβειες χωρίς να ξεχωρίζει κάποιο ιδιαίτερα χαμηλό σκορ.

5.8 FDSSC



Εικόνα 78 : Οι προβλέψεις στις εικόνες Kirkki,Erato,Nefeli

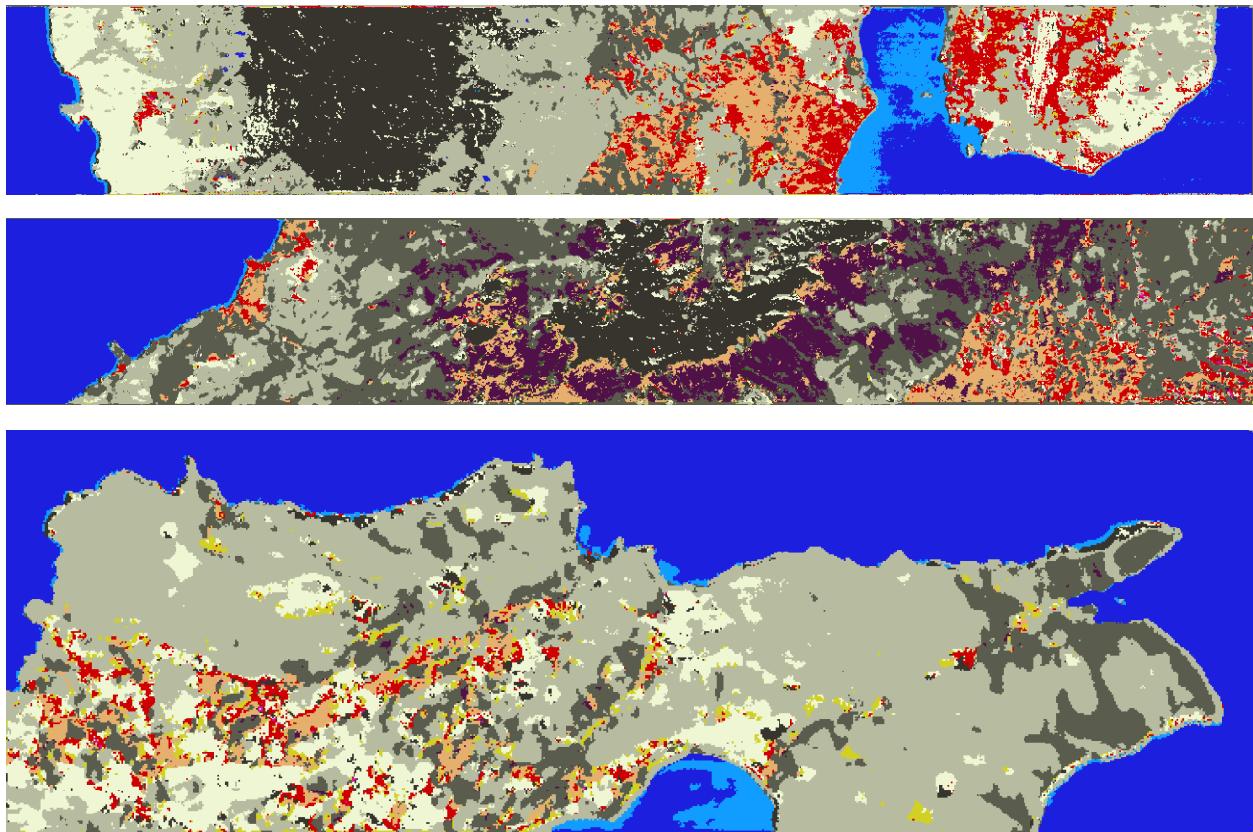
FDSSC ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS	F1%
1	Dense Urban Fabric	2917	50.98	61.65	38.70	55.81
2	Mineral Extraction Sites	591	0.51	0.65	0.29	0.57
3	Non-Irrigated Arable Land	1826	44.74	39.37	26.49	41.89
4	Fruit Trees	715	0.00	0.00	0.00	0.00
5	Olive Groves	5352	44.62	74.42	38.68	55.79
6	Broad-leaved Forest	610	10.16	71.26	9.76	17.79
7	Coniferous Forest	4224	83.00	58.77	52.45	68.81
8	Mixed Forest	3005	0.00	0.00	0.00	0.00
9	Dense Sclerophyllous Veg.	5824	77.85	46.14	40.79	57.94
10	Sparce Sclerophyllous Veg.	10769	69.51	54.35	43.89	61.00
11	Sparcely Vegetated Areas	2160	12.55	41.50	10.66	19.27
12	Rocks and Sand	18051	88.96	91.40	82.09	90.17
13	Water	2534	100.00	83.41	83.41	90.95
14	Coastal Water	1083	35.00	64.13	29.27	45.28
OA%						66.25
Macro-avg			44.13	49.08	33.00	43.23
Weighted-avg			66.25	63.97	51.19	63.41
Kappa %						59.60
Train time(s)		325				
Number of parameters		2268534				

Πίνακας 25: Στατιστικά των προβλέψεων

Σε αντίστοιχη κατάσταση με την μέθοδο Liu et al., η μέθοδος FDSSC εμφανίζει πολύ χαμηλά σκορ στις κατηγορίες 2,4. Παρόλα αυτά αξίζει να αναφερθεί πως οπτικά διατηρούνται οι λεπτομέρειες όπως το αεροδρόμιο στην εικόνα Kirkι ενώ η κατηγορία του αστικού ιστού δεν παρουσιάζεται αρκετά. Επίσης σε επίπεδο οπτικό και όχι στατιστικό το παράκτιο νερό δείχνει αληθοφανές.

Γενικότερα η συγκεκριμένη ταξινόμηση είναι ανομοιόμορφη χωρίς να παρουσιάζει συγκριτικά κάποιο πλεονέκτημα σε σχέση με τις ταξινομήσεις των άλλων μεθόδων. Ένα φαινόμενο που παρατηρείται είναι η άστοχη ταξινόμηση και στις 3 εικόνες αρκετών παράκτιων περιοχών σαν εικονοστοιχείο της κατηγορίας 1.

5.9 LIU ET AL.



Εικόνα 77 : Οι προβλέψεις στις εικόνες Kirkki,Erato,Nefeli

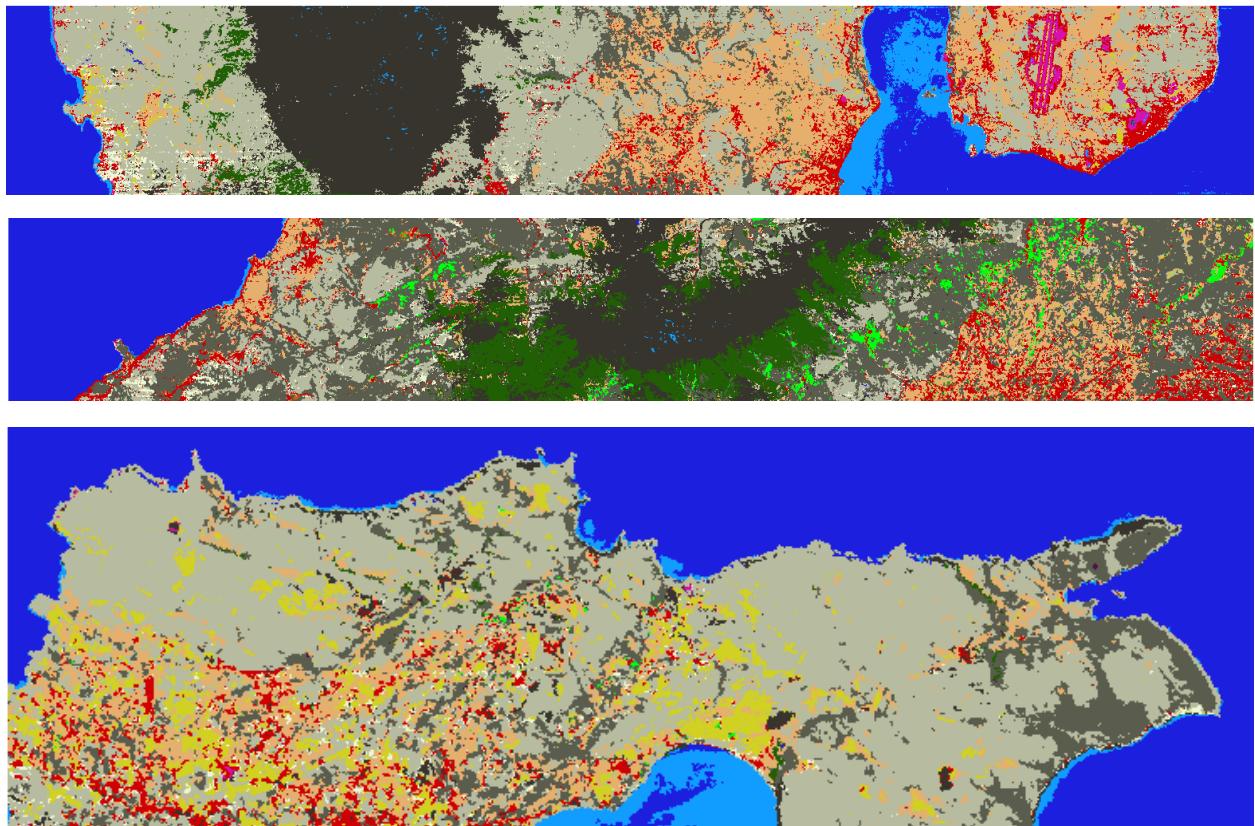
LIU ET AL.						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	57.76	53.15	38.28	55.36
2	Mineral Extraction Sites	591	0.17	2.04	0.16	0.31
3	Non-Irrigated Arable Land	1826	17.36	64.96	15.87	27.40
4	Fruit Trees	715	0.00	0.00	0.00	0.00
5	Olive Groves	5352	43.65	53.26	31.56	47.98
6	Broad-leaved Forest	610	0.00	0.00	0.00	0.00
7	Coniferous Forest	4224	0.97	95.35	0.97	1.92
8	Mixed Forest	3005	54.48	43.36	31.83	48.29
9	Dense Sclerophyllous Veg.	5824	84.05	55.89	50.53	67.13
10	Sparce Sclerophyllous Veg.	10769	84.70	73.53	64.90	78.72
11	Sparcely Vegetated Areas	2160	50.23	22.83	18.62	31.39
12	Rocks and Sand	18051	96.01	95.31	91.68	95.66
13	Water	2534	79.44	85.44	69.97	82.33
14	Coastal Water	1083	70.54	59.04	47.37	64.28
OA%						69.10
Macro-avg			45.67	50.30	33.00	42.91
Weighted-avg			69.10	71.41	55.76	65.93
Kappa %						59.60
Train time(s)		127				
Number of parameters		191503				

Πίνακας 26: Στατιστικά των προβλέψεων

Η μέθοδος Liu et al. παρουσιάζει ένα παράδειγμα υψηλής γενικής ακρίβειας που όμως ποιοτικά υστερεί. Αρχικά παρατηρώντας τις εικόνες γίνεται αμέσως αντιληπτή η απουσία χαρακτηριστικών σημείων όπως το αεροδρόμιο στα δεξιά της εικόνας Kirk. Επίσης η κατανομή του παράκτιου νερού δείχνει μη αληθοφανής.

Η συγκεκριμένη μέθοδος εμφανίζει αρκετή ανομοιομορφία στο πίνακα στατιστικών της. Οι κατηγορίες 2,4,6,7 έχουν ιδιαίτερα χαμηλά σκορ, δεν έχουν ταξινομηθεί πολλά εικονοστοιχείο στις κατηγορίες αυτές. Κατά συνέπεια δεν παρατηρείται ποικιλία στις χρήσεις γης πάνω στις εικόνες. Χαρακτηριστικό παράδειγμα αποτελεί η εικόνα Nefeli όπου το δεξί τμήμα της παρουσιάζεται πολύ φτωχό σε σχέση με τις προβλέψεις άλλων μεθόδων.

5.10 SVM



Εικόνα 79 : Οι προβλέψεις στις εικόνες Kirki,Erato,Nefeli

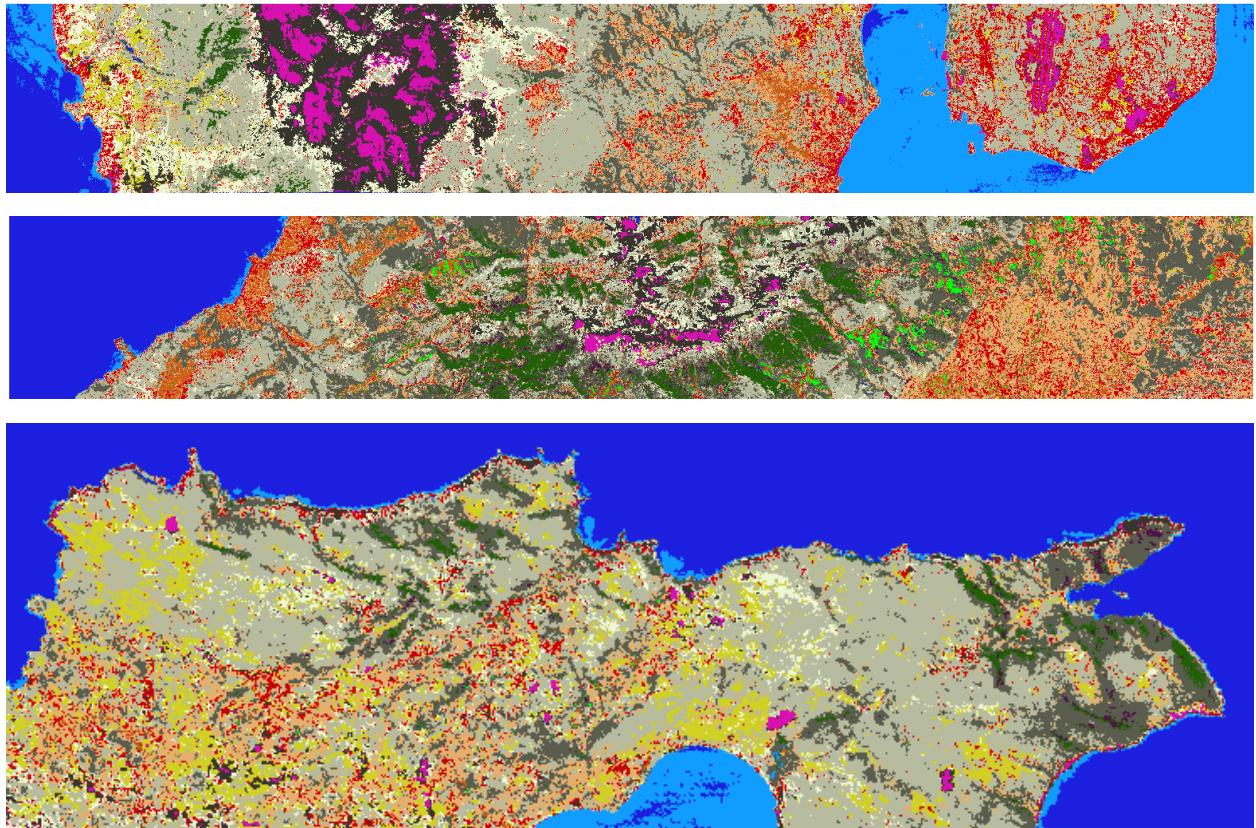
SVM						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	67.71	67.68	51.17	67.69
2	Mineral Extraction Sites	591	69.88	48.08	39.83	56.97
3	Non-Irrigated Arable Land	1826	67.69	84.77	60.35	75.27
4	Fruit Trees	715	0.56	21.05	0.55	1.09
5	Olive Groves	5352	63.30	70.12	49.85	66.54
6	Broad-leaved Forest	610	62.13	82.93	55.09	71.04
7	Coniferous Forest	4224	68.68	62.05	48.37	65.20
8	Mixed Forest	3005	0.53	84.21	0.53	1.06
9	Dense Sclerophyllous Veg.	5824	91.52	53.57	51.04	67.58
10	Sparce Sclerophyllous Veg.	10769	82.62	78.03	67.03	80.26
11	Sparcely Vegetated Areas	2160	5.60	32.79	5.02	9.57
12	Rocks and Sand	18051	98.71	94.43	93.28	96.52
13	Water	2534	99.96	87.83	87.80	93.50
14	Coastal Water	1083	67.59	77.13	56.31	72.05
OA%						76.67
Macro-avg			60.46	67.48	48.00	58.88
Weighted-avg			76.67	76.61	63.47	73.47
Kappa %						72.06
Train time(s)		10				
Number of parameters						

Πίνακας 27: Στατιστικά των προβλέψεων

Παρατηρώντας οπτικά τις προβλέψεις της μεθόδου SVM γίνεται αισθητή η ύπαρξη εικονοστοιχείων παράκτιου νερού στα αριστερά τις εικόνας Kirk. Οι προηγούμενες μέθοδοι χαρακτηρίζαν ομαδοποιημένα τους υδάτινους όγκους (χιόνι) που υπάρχουν στην περιοχή, γεγονός που η μέθοδος SVM δεν κατορθώνει επιτυχώς. Παρόλα αυτά οι εικόνες δεν παρουσιάζουν κάποιο σοβαρό πρόβλημα που να διακρίνεται οπτικά και να τις χαρακτηρίζει σε σχέση με αυτές των προηγούμενων προβλέψεων.

Οι κατηγορίες 4,8 σημείωσαν πολύ χαμηλά ποσοστά ακρίβειας παραγωγού, αναδεικνύοντας την αδυναμία της μεθόδου να διαχωρίσει τις κατηγορίες αυτές και να χαρακτηρίσει ορθά τα εικονοστοιχεία του ανήκουν

5.11 RF



Εικόνα 80 :Οι προβλέψεις στις εικόνες Erato,Kirki,Nefeli

RF						
Κατηγορία	Περιγραφή	Αριθμός εικονοστοιχείων	PA%	UA%	AC/CLASS%	F1%
1	Dense Urban Fabric	2917	54.65	49.98	35.33	52.21
2	Mineral Extraction Sites	591	93.57	4.86	4.85	9.25
3	Non-Irrigated Arable Land	1826	62.21	52.91	40.04	57.19
4	Fruit Trees	715	60.42	36.95	29.75	45.86
5	Olive Groves	5352	45.76	66.69	37.25	54.28
6	Broad-leaved Forest	610	37.87	52.62	28.24	44.04
7	Coniferous Forest	4224	57.13	64.66	43.53	60.66
8	Mixed Forest	3005	9.25	68.81	8.88	16.31
9	Dense Sclerophyllous Veg.	5824	79.19	48.80	43.26	60.39
10	Sparce Sclerophyllous Veg.	10769	72.66	69.93	55.36	71.27
11	Sparcely Vegetated Areas	2160	28.15	23.77	14.79	25.77
12	Rocks and Sand	18051	33.14	89.05	31.85	48.31
13	Water	2534	75.37	87.29	67.92	80.90
14	Coastal Water	1083	74.33	56.18	47.05	63.99
OA%						51.67
Macro-avg			55.98	55.18	35.00	49.32
Weighted-avg			51.67	69.04	38.64	54.29
Kappa %						46.12
Train time(s)			1			
Number of parameters						

Πίνακας 28: Στατιστικά των προβλέψεων

Η μέθοδος RF παρουσιάζει διαφορετική συμπεριφορά σε σχέση με όλες τις προηγούμενες μεθόδους. Σχετικά με το οπτικό κομμάτι των παρατηρήσεων γίνεται αισθητή η ύπαρξη παραπάνω εικονοστοιχείων που συμβολίζουν το παράκτιο νερό σε σχέση με τον αριθμό που θα έπρεπε, και αυτό πιστοποιείται και από την ακρίβεια του χρήστη για την κατηγορία αυτή. Το ίδιο φαίνομενο παρατηρείται και για την κατηγορία 2 που κυριαρχεί στις προβλέψεις.

Ενδιαφέρον στοιχείο είναι η υψηλή ακρίβεια που επιτυγχάνεται στην κατηγορία των δέντρων με φρούτα, αλλά και μια σχετική ομοιομορφία στα σκορ των κατηγοριών 3 έως 7 που δεν συναντήθηκε στις μεθόδους με αρχιτεκτονικές νευρωνικών δικτύων.

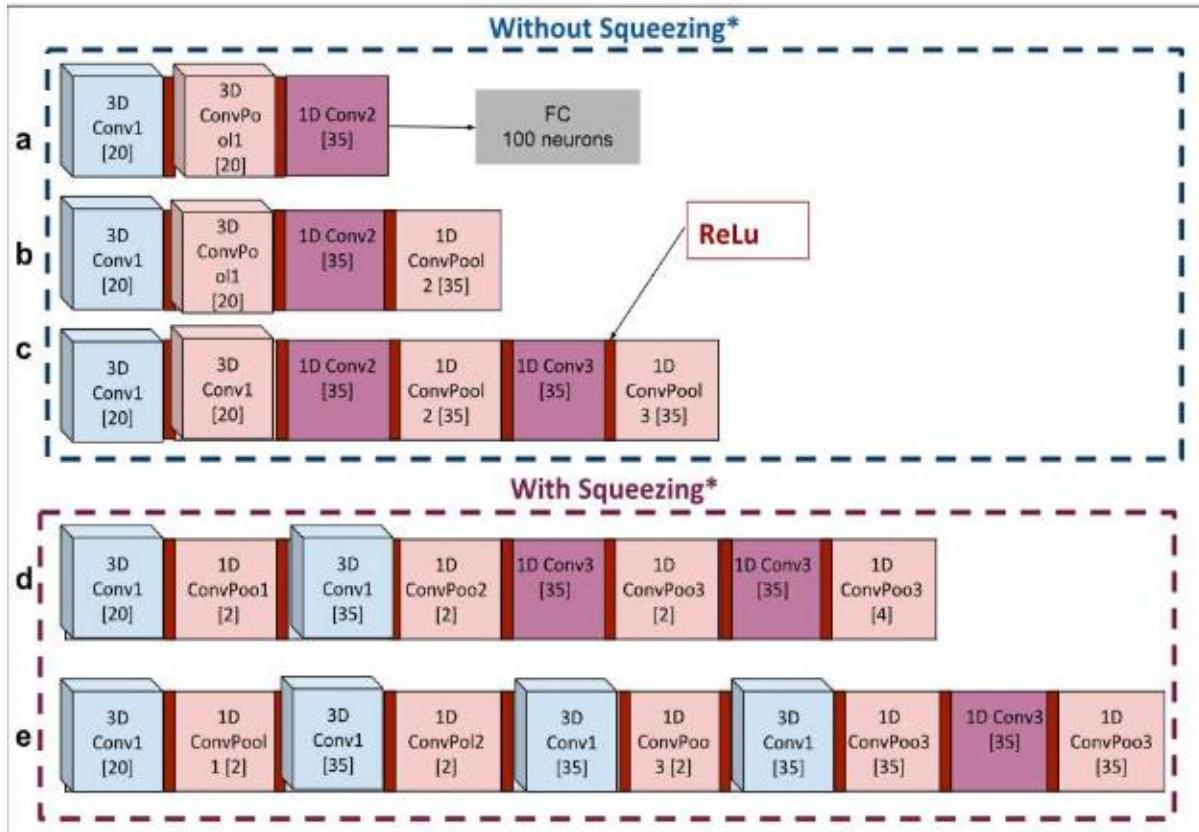
6 ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 ΣΥΓΚΡΙΣΗ ΜΕ ΒΙΒΛΙΟΓΡΑΦΙΑ

Στην συγκεκριμένη εργασία παρουσιάστηκαν διάφοροι μέθοδοι Βαθιάς Μάθησης για την ταξινόμηση υπερφασματικών δεδομένων. Οι μέθοδοι αυτοί βασίστηκαν στην αντίστοιχη βιβλιογραφία και έγινε προσπάθεια εφαρμογής τους. Η προσπάθεια αυτή προφανώς δεν μπορεί να αντικαταστήσει τις εφαρμογές των δημιουργών, και σίγουρα δεν μπορεί να αναπαραστήσει με ακριβή τρόπο τις συνθήκες στις οποίες έγιναν τα πειράματα, τις ακριβείς αρχιτεκτονικές των δικτύων κλπ. Παρόλα αυτά το πρώτο γενικό συμπέρασμα που προκύπτει από τα αποτελέσματα των ταξινομήσεων είναι πως οι υψηλές ακριβειες των μεθόδων που προέκυψαν στην βιβλιογραφία(<90%, ολικές ακριβειες ανταγωνιστικότερες από SVM) δεν επαληθεύθηκαν, παρόλο που εφαρμόστηκαν σε διαφορετικά σετ δεδομένων. Το γεγονός αυτό μπορεί να αιτιολογηθεί.

Αρχικά, τα νευρωνικά δίκτυα και πιο συγκεκριμένα τα δίκτυα Βαθιάς Μάθησης παρουσιάζουν μεγάλη ισχύ και αποτελεσματικότητα, η οποία όμως δεν επιτυγχάνεται σύντομα και απλά. Η πολυπλοκότητα των δικτύων είναι πολύ μεγάλη, οι παράμετροι που τα επηρεάζουν αρκετοί, οπότε δεν αρκεί μόνο η εφαρμογή του δικτύου στα δεδομένα. Μεγάλο ρόλο παίζει η προσαρμογή του δικτύου στο κάθε πρόβλημα ξεχωριστά, ή αλλιώς η βελτιστοποίηση του.

Στην βιβλιογραφία παρατηρήθηκε η τάση να δίνεται μία γενική κατεύθυνση της αρχιτεκτονικής των δικτύων όπως ο τύπος του δικτύου, βασική ροή των στρωμάτων, τα μεγέθη των φίλτρων κλπ. Αφού δοθεί η βασική ιδέα, οι δημιουργοί πραγματοποιούσαν πειράματα σχετικά με τον αριθμό των στρωμάτων(Hamida et al.) ,των αριθμό των φίλτρων στα συνελικτικά στρώματα(Li et al.), τα μεγέθη των φίλτρων, το μέγεθος της γειτονιάς των εικονοστοιχείων στην είσοδο του δικτύου κλπ. Οπότε όταν γίνεται αναφορά σε μια μέθοδο, ουσιαστικά γίνεται αναφορά σε μία γενική ιδέα αρχιτεκτονικής οι οποία προσαρμόστηκε στα γνωστά datasets βασιζόμενη σε πειράματα από τους δημιουργούς, με γνώμονα την επίτευξη της υψηλότερης ακριβειας. Εκτός από την δομή της αρχιτεκτονικής, μεγάλο ρόλο στην τελική αποτελεσματικότητα της μεθόδου παίζει και η επιλογή των κατάλληλων υπερπαραμέτρων. Οι δημιουργοί πραγματοποιούν πειράματα με διαφορετικές μεθόδους βελτιστοποίησης του αλγορίθμου της οπισθοδιάδοσης, διαφορετικούς ή και μεταβλητούς ρυθμούς μάθησης καθώς και διαφορετικούς τρόπους αρχικοποίησης των βαρών σε κάθε παράμετρο προς εκπαίδευση.



Εικόνα 81:Οι διαφορετικές παραλλαγές στις αρχιτεκτονικές του δικτύου Hamida et al.

Επιπρόσθετα στην προσαρμογή της κάθε μεθόδου, την τελική αποτελεσματικότητα καθορίζουν και τα δεδομένα προς ταξινόμηση. Τα γνωστά datasets αναφέρονται σε μία εικόνα που λήφθηκε μία συγκεκριμένη χρονική στιγμή. Στις υπερφασματικές εικόνες όπου η κάθε τιμή εκφράζει την τιμή της ανακλαστικότητας του εδάφους στο εκάστοτε ηλεκτρομαγνητικό εύρος φάσματος, οι τιμές αυτές επηρεάζονται από παράγοντες θορύβου όπως οι ατμοσφαιρικές διορθώσεις. Σίγουρα όταν το σύνολο των δεδομένων αναφέρεται σε μία μοναδική εικόνα τα δεδομένα παρουσιάζουν μία σχετική ομοιογένεια. Επίσης το μέγεθος των datasets αυτών είναι αρκετά μικρό με αποτέλεσμα η ακρίβεια των μεθόδων ανά κατηγορία να επαληθεύεται σε αριθμό εικονοστοιχείων που σπάνια ξεπερνά τα 100, καθιστώντας την επαλήθευση μη αντιπροσωπευτική.

Ένας άλλος καθοριστικός παράγοντας είναι η τεχνική με την οποία τα δεδομένα χωρίζονται στα σετ εκπαίδευσης και επαλήθευσης. Είναι γεγονός πως οι 3D συνελίξεις εκπαιδεύουν τα δίκτυα σε γειτονιές δεδομένων όπου είναι πολύ πιθανό να συνυπάρχουν εικονοστοιχεία και των δύο σετ. Έτσι η ακρίβεια του δικτύου επαληθεύεται πάνω σε εικονοστοιχεία που το δίκτυο έχει εκπαιδευθεί ήδη. Η χωρική συσχέτιση των δεδομένων εκπαίδευσης-επαλήθευσης στις μεθόδους που πραγματοποιούν συνελίξεις, σε συνδυασμό με το μικρό μέγεθος των εικονοστοιχείων οδηγούν σε υψηλές ακρίβειες οι οποίες όμως δεν ανταποκρίνονται σε ρεαλιστικές εφαρμογές.



Εικόνα 82: Αναπαράσταση χωρίσματος σε δεδομένων με χωρικά κριτήρια

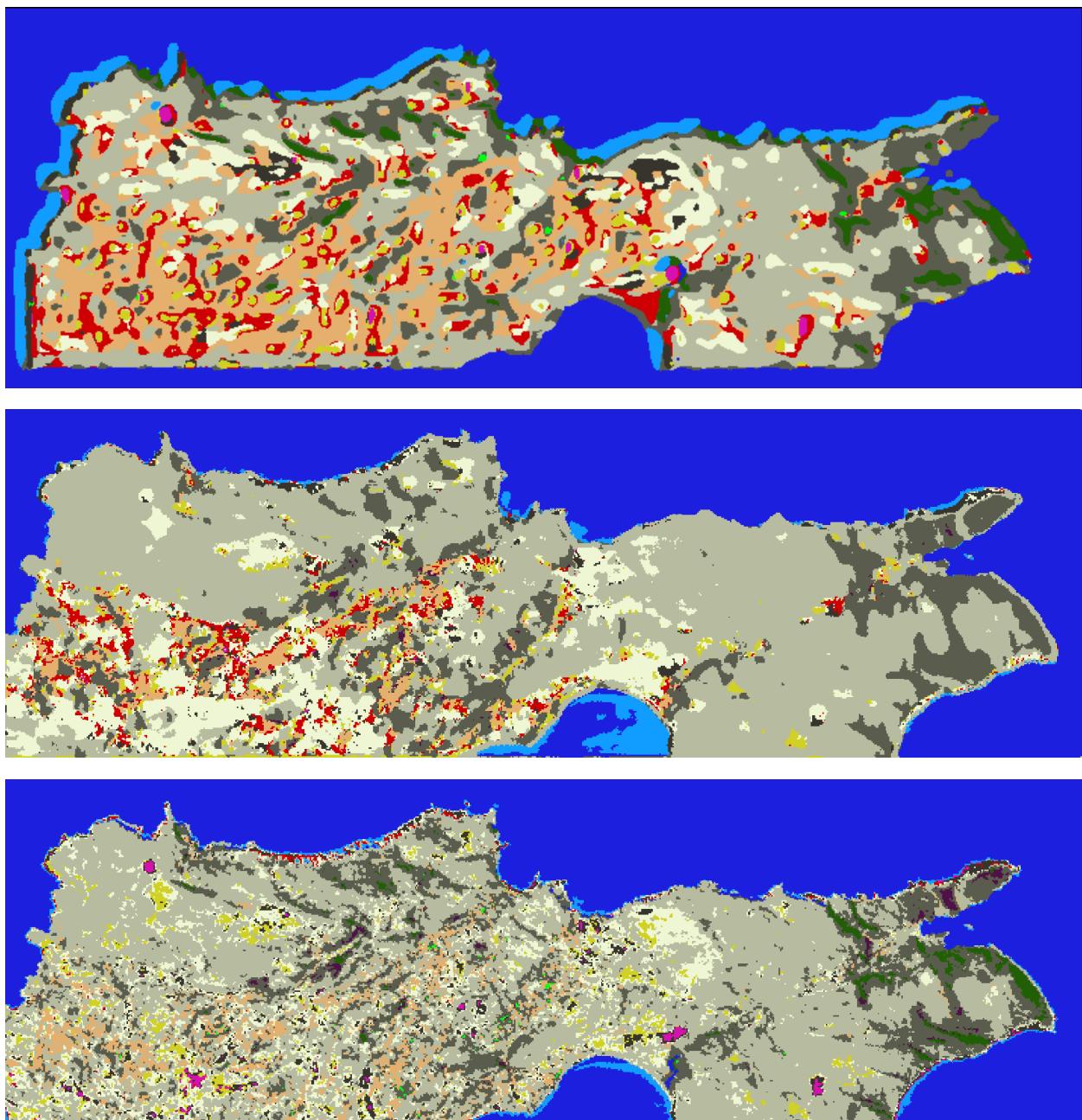
Στην παρούσα εργασία η τεχνική disjoint και το HyRANK dataset επιλέχθηκαν ώστε να καταπολεμήσουν τα παραπάνω φαινόμενα, προσδίδοντας ρεαλιστικά αποτελέσματα πάνω στις συνήθεις εφαρμογές ταξινόμησης υπερφασματικών εικόνων. Υπό αυτές τις συνθήκες είναι απίθανο να προκύψουν ακρίβειες ανάλογες της βιβλιογραφίας. Ακόμα πέρα από βασικές προσαρμογές υπερπαραμέτρων όπως ο ρυθμός μάθησης, batch size και χρόνος εκπαίδευσης, διατηρήθηκαν αυτοί που προτείνονται από τους δημιουργούς. Το ίδιο συμβαίνει και με τις τελικώς εφαρμοσμένες αρχιτεκτονικές της κάθε μεθόδου. Στην περίπτωση που σκοπός της εργασίας αυτής ήταν η βελτιστοποίηση των μεθόδων, η προσαρμογή των αρχιτεκτονικών και των υπερπαραμέτρων πάνω στα δεδομένα HyRANK θα παρουσίαζε σίγουρα υψηλότερες ακρίβειες στην ταξινόμηση.

6.2 ΟΠΤΙΚΗ ΣΥΓΚΡΙΣΗ ΠΡΟΒΛΕΨΕΩΝ ΣΤΙΣ ΕΙΚΟΝΕΣ

Οι μέθοδοι που παρουσιάστηκαν παρουσιάζουν διάφορα κοινά χαρακτηριστικά μεταξύ τους. Δομές όπως το residual learning ή τα χωρικό-φασματικά συνελικτικά μπλοκ εμφανίζονται συχνά. Άλλα δίκτυα είναι pixel-based και άλλα δέχονται γειτονίες των εικονοστοιχείων στην εισαγωγή. Το δίκτυο Lee et al. είναι ένα δίκτυο που δεν περιέχει καθόλου πλήρως συνδεδεμένα στρώματα, είναι πλήρως συνελικτικό. Επίσης παρουσιάστηκαν αρχιτεκτονικές που περιέχουν αναδραστικά στρώματα, αρχιτεκτονικές που εφαρμόζουν adversarial learning ή αλλιώς εισάγουν θόρυβο στη ροή του δικτύου και μία εφαρμογή δικτύου Κωδικοποιητή-Αποκωδικοποιητή.

Παρατηρώντας τις προβλέψεις στις εικόνες, γίνεται αμέσως αντιληπτό πώς η αίσθηση της χωρικής λεπτομέρειας στις εικόνες είναι διαφορετική ανά μέθοδο. Στις pixel-based μεθόδους όπως SVM,RF,Hu et al., οι εικόνες που παράχθηκαν απότυπώνουν τις κατηγορίες χρήσεων γης με περισσότερη χωρική ακρίβεια.. Στις μεθόδους που δέχονται γειτονίες των εικονοστοιχείων ,οι προβλέψεις που δίνονται περιγράφουν τις χρήσεις γης με έναν γενικότερο τρόπο. Χαρακτηριστικό παράδειγμα αποτελούν οι περιπτώσεις των μεθόδων Chen et al. με γειτονιά εικονοστοιχείων 27x27 και FDSSC με γειτονιά 9x9. Σε μια υποπεριοχή της εικόνας μπορεί να παρατηρηθεί η ύπαρξη διαφορετικών κλάσεων, δίνοντας διαφορετική οντότητα και σημασία στο κάθε εικονοστοιχείο Εκπαιδεύοντας το δίκτυο στους συσχετισμούς πολλών εικονοστοιχείων μαζί οι προβλέψεις γενικεύονται.

Αξίζει να αναφερθεί πως η μέθοδος Liu et al. παρότι και αυτή δέχεται γειτονία 9x9 εικονοστοιχείων στην εισαγωγή, η προβλέψεις της δεν παρουσιάζουν τα χαρακτηριστικά αυτά. Αυτό οφείλεται στην ιδιότητα απαλλαγής του θορύβου του συγκεκριμένου δικτύου, σαν δίκτυο Κωδικοποιητή-Αποκωδικοποιητή που εκπαιδεύεται στην ανασύσταση της εισαγωγής απαλλαγμένης από τον θόρυβο. Τα προηγούμενα δίκτυα δεν διαθέτουν αυτή την ιδιότητα οπότε οι μεγάλες γειτονίες προσδίδουν και αναλογικά μεγαλύτερο θόρυβο στις προβλέψεις, και έτσι εξηγείται η αίσθηση της χαμηλότερης χωρικής ακρίβειας.



Εικόνα 83: Η πρόβλεψη της εικόνας Nefeli, από της μεθόδου Chen et al., Liu et al., Hu et al.

6.3 ΠΑΡΑΤΗΡΗΣΕΙΣ ΣΤΗΝ ΤΑΞΙΝΟΜΗΣΗ ΤΩΝ ΚΑΤΗΓΟΡΙΩΝ

Η δυσκολότερη κατηγορία για επίτευξη υψηλής ακρίβειας στην ταξινόμηση αποδείχθηκε η κατηγορία 4, που αντιστοιχεί στα δέντρα με φρούτα. Σε όλες τις περιπτώσεις των μεθόδων το σκορ ήταν χαμηλό, σε πολλές μάλιστα δεν ταξινομήθηκε κανένα εικονοστοιχείο. Οι μέθοδοι RF και Mou et al. κατάφεραν να παρουσιάσουν σχετικά υψηλότερα ποσοστά ακρίβειας παραγωγού, διαχωρίζοντας επιτυχώς την συγκεκριμένη κατηγορία από τις υπόλοιπες κατηγορίες με παρόμοιες φασματικές υπογραφές. Οι υπόλοιπες μέθοδοι δεν κατάφεραν να την διαχωρίσουν, παρόλο που η ακρίβεια του χρήστη στα εικονοστοιχεία που προέβλεπαν ως δέντρα με φρούτα ήταν υψηλή.

Ακολουθεί ο πίνακας σύγχυσης των μεθόδων για την συγκεκριμένη κατηγορία, κανονικοποιημένος ώστε να εκφράζει την ακρίβεια του παραγωγού:

PA Method	A/A Acronym	1 DUF	2 MES	3 NAL	4 FRT	5 OLG	6 BLF	7 CFF	8 MXF	9 DSV	10 SSV	11 SVA	12 RCS	13 WAT	14 CWT	CLASSIF
Hamida	FRT	0.00	0.00	0.00	11.75	17.76	31.05	0.00	13.01	26.43	0.00	0.00	0.00	0.00	0.00	
Lee	FRT	0.14	0.00	0.00	1.40	91.19	7.13	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	
Li	FRT	0.28	0.00	0.00	18.04	58.88	22.24	0.00	0.00	0.56	0.00	0.00	0.00	0.00	0.00	
Hu	FRT	0.00	0.00	0.00	0.00	44.76	41.26	0.00	0.14	13.85	0.00	0.00	0.00	0.00	0.00	
He	FRT	0.28	0.00	0.00	8.39	84.06	5.31	0.00	0.00	1.96	0.00	0.00	0.00	0.00	0.00	
Mou	FRT	0.00	0.00	0.00	54.55	19.30	22.80	0.00	0.00	3.36	0.00	0.00	0.00	0.00	0.00	
Liu	FRT	0.00	0.00	0.00	0.00	72.45	0.00	0.00	13.85	13.71	0.00	0.00	0.00	0.00	0.00	
Chen	FRT	0.00	0.00	0.00	0.00	20.70	4.62	0.00	0.00	63.08	11.61	0.00	0.00	0.00	0.00	
FDDSC	FRT	2.66	13.71	1.12	0.00	25.59	3.08	5.73	0.28	41.82	6.01	0.00	0.00	0.00	0.00	
SVM	FRT	0.70	0.00	0.00	0.56	66.99	4.34	0.00	0.00	27.41	0.00	0.00	0.00	0.00	0.00	
RF	FRT	2.10	0.00	0.00	60.42	14.97	9.65	0.00	0.14	12.31	0.42	0.00	0.00	0.00	0.00	
REF																

Πίνακας 29: Ακρίβεια του παραγωγού, κατηγορία 4

Τα δέντρα με φρούτα ως κατηγορία 3 περικλείει αρκετούς τύπους δέντρων, με διαφορετικές φασματικές υπογραφές. Οι κατηγορίες 6,9 και ιδιαίτερα η 5 αποτέλεσαν συχνή λάθος πρόβλεψη της ταξινόμησης σχετικά με την κατηγορία των δέντρων με φρούτα. Γενικότερα οι μέθοδοι που περιλαμβαναν συνελικτικά φίλτρα στην αρχιτεκτονική τους καθώς και ο SVM παρουσίασαν τα παραπάνω αποτελέσματα. Οι μέθοδοι RF-Mou et al. τα πήγαν αρκετά καλύτερα.

Η λογική της γειτονιάς εικονοστοιχείων στην εισαγωγή βασίζεται στο γεγονός ότι γειτονικά στον χώρο εικονοστοιχεία μεταξύ τους εμφανίζουν κοινά φασματικά χαρακτηριστικά, και πιθανότατα να ανήκουν στην ίδια κατηγορία. Έτσι γίνεται η προσπάθεια το δίκτυο να εκπαιδευθεί πάνω στον χωρικό συσχετισμό των εικονοστοιχείων, πέρα από τον φασματικό. Δεδομένου ότι το κάθε εικονοστοιχείο αντιστοιχεί σε περιοχή 30x30 μέτρων στο έδαφος στο dataset HyRANK και ότι η κατηγορία 4 περιέχει πολλά και διαφορετικά είδη δέντρων, οι πρόσθετοι συσχετισμοί προς εκπαίδευση σε μία ήδη ετερογενή κατηγορία (φασματικά λόγω της ποικιλίας των δέντρων και της πιθανής αραιής συγκέντρωσης σε σχέση με άλλα στοιχεία της φύσης σε ένα πλαίσιο 30x30 μέτρων) απορροσανατολίζουν τις προβλέψεις. Η πληροφορία που εξάγεται από τα γειτονικά εικονοστοιχεία συσχετίζει τα δέντρα με φρούτα με άλλες κοντινές φασματικά κατηγορίες.

Η επιτυχία των RF-Mou et al., πέρα από την pixel-based προσέγγιση των μεθόδων αυτών, προέρχεται και από την λειτουργία των αρχιτεκτονικών τους. Η μέθοδος RF βασίζει την ταξινόμηση σε δέντρα αποφάσεων, χωρίς να επηρεάζεται αριθμητικά. Η μέθοδος Mou σαν αναδραστικό νευρωνικό δίκτυο δεν περιέχει φίλτρα και διατηρεί-αντλεί πληροφορίες από κάθε εικονοστοιχείο ξεχωριστά στο οποίο εκπαιδεύεται αντιμετωπίζοντας όλα τα παραπάνω

προβλήματα που αναφέρθηκαν. Η μέθοδος SVM αν και pixel-based επηρεάζεται αριθμητικά αφού προσδιορίζει αποστάσεις στον φασματικό χώρο, γι' αυτό και τα χαμηλά ποσοστά ακρίβειας.

Πάνω στην ίδια λογική παρατηρείται και η συμπεριφορά των μεθόδων στην κατηγορία του μεικτού δάσους, άλλη μια ετερογενή κατηγορία που καλύπτει ένα εύρος φασματικών υπογραφών.

PA Method	A/A Acronym	DUF	1 MES	2 NAL	3 FRT	4 OLG	5 BLF	6 CFF	7 MXF	8 DSV	9 SSV	10 SVA	11 RCS	12 WAT	13 CWT	14 CLASSIF
Hamida	MXF		0.00	0.00	0.00	0.00	0.07	0.20	53.21	32.05	14.44	0.03	0.00	0.00	0.00	0.00
Lee	MXF		0.00	0.00	0.00	0.00	0.00	0.00	47.82	8.39	43.53	0.27	0.00	0.00	0.00	0.00
Li	MXF		0.00	0.00	0.00	0.00	0.30	6.29	70.98	4.79	17.34	0.30	0.00	0.00	0.00	0.00
Hu	MXF		0.01	0.00	0.00	0.00	0.07	0.77	21.03	39.00	39.10	0.03	0.00	0.00	0.00	0.00
He	MXF		0.00	0.00	0.00	0.00	0.07	0.30	53.38	23.49	22.73	0.03	0.00	0.00	0.00	0.00
Mou	MXF		0.00	0.00	0.00	0.00	1.23	1.10	59.33	11.75	26.52	0.03	0.00	0.03	0.00	0.00
Liu	MXF		0.00	0.00	0.00	0.00	0.00	17.37	0.00	0.07	54.48	27.99	0.10	0.00	0.00	0.00
Chen	MXF		0.00	0.00	0.00	0.00	0.00	0.57	54.31	0.07	44.99	0.07	0.00	0.00	0.00	0.00
FDDSC	MXF		0.00	0.00	0.00	0.00	0.17	0.00	59.57	0.00	40.20	0.07	0.00	0.00	0.00	0.00
SVM	MXF		0.00	0.00	0.00	0.00	4.49	0.20	55.27	0.53	39.37	0.13	0.00	0.00	0.00	0.00
RF	MXF		0.00	0.00	0.00	0.10	0.33	2.83	34.24	9.25	53.08	0.17	0.00	0.00	0.00	0.00
REF																

Πίνακας 30: Ακρίβεια του παραγωγού, κατηγορία 8

Στην περίπτωση αυτή είναι εμφανής η σύγχυση της κατηγορίας του μεικτού δάσους με τις κατηγορίες 7 και 9. Η μέθοδος Liu et al. κατόρθωσε να ταξινομήσει τα περισσότερα εικονοστοιχεία που ανήκουν στην κατηγορία αφού απέφυγε την ταξινόμηση των εικονοστοιχείων στην κατηγορία 7. Οι υπόλοιπες μέθοδοι δεν κατάφεραν να διαχωρίσουν ορθά το μεικτό δάσος από τα κωνοφόρα(7) και σκληρόφυλλα δάση(9).

Σε αντίθεση, η κατηγορία που αναφέρεται στις πέτρες και στην άμμο διαχωρίστηκε επιτυχώς από τις περισσότερες μεθόδους. Η κατηγορία αυτή δεν αναφέρεται σε δάση ή καλλιέργειες και αποτελεί πιο εύκολο έργο η σωστή ταξινόμηση της, αφού δεν συνυπάρχουν πολλές άλλες κατηγορίες στην ταξινόμηση με παρόμοιες φασματικές υπογραφές . Και σε αυτήν περίπτωση παρατηρείται μία διαφορά στις ακρίβειες που επιτυγχάνει η κάθε μέθοδος που συνδέεται με την γειτονιά των εικονοστοιχείων στην εισαγωγή των δικτύων. Οι μέθοδοι Hu et al. και RF έχουν μειωμένα ποσοστά ακρίβειας σε σχέση με τις υπόλοιπες μεθόδους. Προσδιορίζουν σε μεγάλο βαθμό την υπαρξή της κατηγορίας στις προβλέψεις αλλά συγχέουν την συγκεκριμένη κατηγορία με την κατηγορία των ορυχείων.

Οι δύο αυτές κατηγορίες αναφέρονται σε όμοιας φύσης υλικά και παρουσιάζουν αρκετά κοινά χαρακτηριστικά στις φασματικές υπογραφές τους. Στην περίπτωση αυτή, οι pixel-based μέθοδοι έδειξαν αδυναμία διαχωρισμού, εκτός από την μέθοδο Mou et al. που χρησιμοποιεί αναδραστικό στρώμα και την μέθοδο SVM. Οι μέθοδοι με 3D συνελίξεις έχοντας καλύτερη πληροφόρηση μέσω των χωρικών συσχετίσεων διαχωρίσαν σωστά τις κατηγορίες.

PA Method	A/A Acronym	DUF	1 MES	2 NAL	3 FRT	4 OLG	5 BLF	6 CFF	7 MXF	8 DSV	9 SSV	10 SVA	11 RCS	12 WAT	13 CWT	14 CLASSIF	
Hamida	RCS		0.01	3.25	0.00	0.00	0.00	0.00	0.04	0.00	0.06	3.56	0.63	92.45	0.00	0.00	
Lee	RCS		0.03	6.11	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.30	93.27	0.00	0.27	
Li	RCS		0.00	0.55	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	98.00	0.07	1.31	
Hu	RCS		2.34	30.68	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.33	1.41	59.69	0.00	5.55	
He	RCS		0.00	10.13	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	88.56	0.00	1.06
Mou	RCS		0.00	0.13	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.12	96.16	0.00	3.37	
Liu	RCS		0.71	0.06	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.09	2.84	96.01	0.00	0.05	
Chen	RCS		1.21	3.26	4.23	0.11	1.93	0.13	2.03	0.01	4.08	13.87	5.61	61.76	0.02	1.74	
FDDSC	RCS		0.00	0.11	0.00	0.00	0.38	0.00	0.34	0.00	0.00	8.68	0.30	88.96	0.06	1.17	
SVM	RCS		0.01	0.01	0.03	0.00	0.00	0.00	0.02	0.00	0.00	0.14	0.14	98.71	0.00	0.93	
RF	RCS		1.04	56.35	0.43	0.00	0.02	0.00	0.00	0.00	0.03	3.07	5.90	33.14	0.00	0.02	
REF																	

Πίνακας 31: Ακρίβεια του παραγωγού, κατηγορία 12

6.4 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΩΝ ΜΕΘΟΔΩΝ

Παρατηρώντας την συμπεριφορά των μεθόδων στην ταξινόμηση αλλά και όλα τα στοιχεία που αναφέρθηκαν προηγουμένως, προκύπτουν συμπεράσματα σχετικά με τα πλεονεκτήματα-μειονεκτήματα της κάθε μεθόδου και των αρχιτεκτονικών τους. Η βέλτιστη αρχιτεκτονική σε ένα νευρωνικό δίκτυο με σκοπό την ταξινόμηση υπερφαμασματικών εικόνων πρέπει να παρουσιάζει συγκεκριμένες ιδιότητες.

Ο χρόνος εκπαίδευσης του δικτύου είναι ένα σημαντικό χαρακτηριστικό, το οποίο εκφράζει και τα υπολογιστικά κόστη. Είναι ένας παράγοντας που αν και η συνεχής πρόοδος της τεχνολογίας στην υπολογιστική ισχύ και στα frameworks Βαθιάς Μάθησης δεν τον καθιστά ανασταλτικό, πρακτικά επηρεάζει σε μεγάλο βαθμό. Ενδεικτικά σε πολλές περιπτώσεις της Γεωργίας Ακριβείας απαιτείται άμεση πρόβλεψη των χρήσεων γης και σε δεδομένα με τεράστια μεγέθη. Η χρηστικότητα της μεθόδου επίσης εξαρτάται και αυτή από τον χρόνο αφού η πρόσβαση σε σταθερούς υπολογιστές και υψηλής απόδοσης εξαρτήματα καθιστά την κάθε εφαρμογή μεθόδου πρόβλεψης πολύπλοκη διαδικασία με πολλά κόστη.

Ο αριθμός των παραμέτρων προς εκπαίδευση σε ένα δίκτυο καθορίζει σε μεγάλο βαθμό τον χρόνο εκπαίδευσης και τα υπολογιστικά κόστη. Επίσης είναι καθοριστικός παράγοντας στο φαινόμενο της υπερπροσαρμογής της μεθόδου πάνω στα δεδομένα εκπαίδευσης. Ο μικρότερος αριθμός παραμέτρων προς εκπαίδευση εγγυάται την ελάχιστη πιθανότητα για υπερπροσαρμογή και επιτρέπει στην μέθοδο να γενικεύει σε μεγαλύτερο εύρος δεδομένων. Τα πλήρως συνδεδεμένα στρώματα καθώς και οι μεγάλου μεγέθους συνελίξεις εισάγουν παραμέτρους που προσαρμόζονται στα δεδομένα εκπαίδευσης. Τέλος ζητούμενο από κάθε μέθοδο είναι μια ομοιογενείς και ποιοτική ταξινόμηση των κατηγοριών, χωρίς την ύπαρξη δύσκολων κατηγοριών που δεν ταξινομούνται και προβληματίζουν.

Συγκεκριμένες δομές στις αρχιτεκτονικές των δικτύων όπως residual learning και ξεχωριστά φασματικά-χωρικά συνελικτικά μπλοκ που ενώνονται στην συνέχεια δεν έδειξαν να προσδίδουν αυξημένη ολική ακρίβεια στο δίκτυο. Οι αρχιτεκτονικές με τυπικά συνελικτικά στρώματα αποδείχθηκαν πιο αποτελεσματικές. Οι παραπάνω δομές προσθέτουν χρόνο και υπολογιστικό κόστος στην εκπαίδευση και μαθαίνουν παραπάνω πληροφορία αυξάνοντας την ακρίβεια εκπαίδευσης, η οποία πληροφορία όμως δεν ανταποκρίνεται στις προβλέψεις και αντίθετα οδηγεί στη σύγχυση κατηγοριών όπως της κατηγορίας 4. Γενικότερα προκύπτει πως το συγκεκριμένο πρόβλημα της ταξινόμησης δεν απαιτεί βαθιές αρχιτεκτονικές με πολλά στρώματα και ειδικές τεχνικές μάθησης. Το γεγονός αυτό επιβεβαιώνει και η απόδοση της «απλής» μεθόδου SVM.

Γι' αυτό τα μοντέλα με βαθιά αρχιτεκτονική και πολλές παραμέτρους όπως Chen et al., FDSSC, Liu et al, δεν κατόρθωσαν να πετύχουν ανταγωνιστικές ακρίβειες ούτε ομογενής-ποιοτικές ταξινομήσεις. Επίσης οι μικρές γειτονιές των εικονοστοιχείων κρίνονται πιο αποτελεσματικές, σε αντίθεση με την βιβλιογραφία όπου σχεδόν σε όλες τις περιπτώσεις τα πειράματα των δημιουργών οδηγούσαν σε υψηλότερες ακρίβειες. Μεγαλύτερη γειτονιά στην εισαγωγή ισούται με μεγαλύτερη πιθανότητα εκπαίδευσης σε εικονοστοιχείο που θα χρησιμοποιηθεί στην επαλήθευση σε μικρά dataset χωρίς μελετημένη τεχνική διαχωρισμού των δεδομένων.

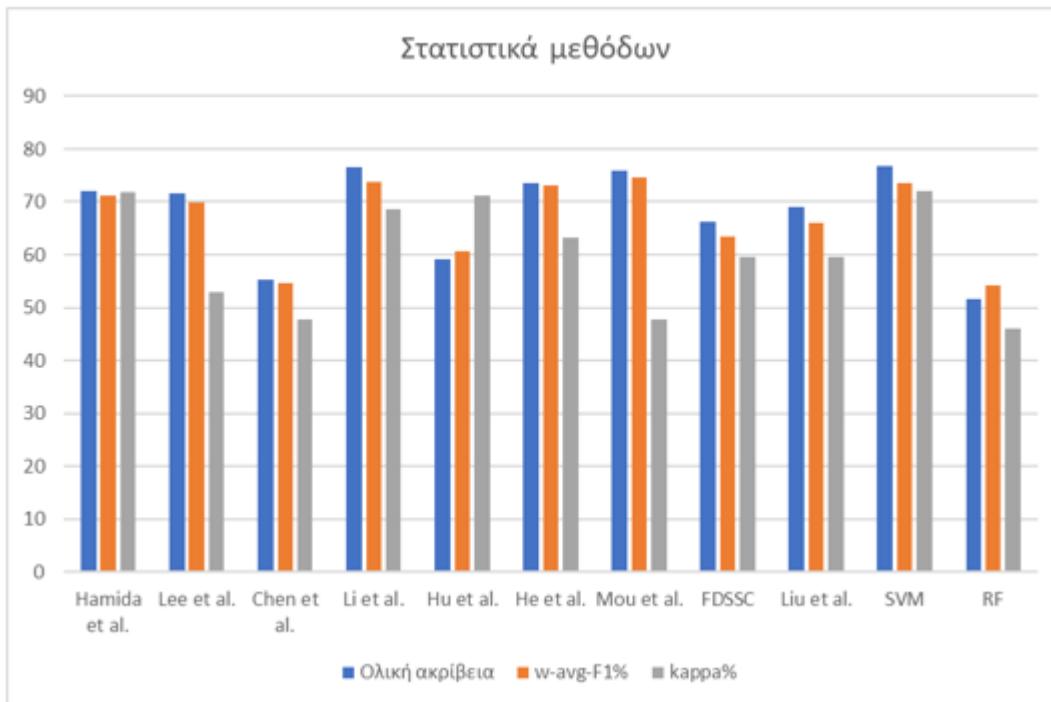
Μέθοδος	Γειτονιά εικονοστοιχείων	Ολική ακρίβεια%
Hamida et al.	5x5	71.95
Lee et al.	5x5	71.55
Chen et al.	27x27	55.24
Li et al.	5x5	76.61
Hu et al.	1x1	59.08
He et al.	7x7	73.51
Mou et al.	1x1	75.8
FDSSC	9x9	66.25
Liu et al.	9x9	69.1
SVM	1x1	76.67
RF	1x1	51.67

Πίνακας 32: Διάγραμμα με την ολική ακρίβεια των μεθόδων και μέγεθος γειτονιάς εικονοστοιχείων στην εισαγωγή

Βέβαια το HyRANK dataset διαθέτει αρκετές κατηγορίες που αναφέρονται σε φυτά και καλλιέργειες. Όπως φάνηκε και στην κατηγορία που αναφέρεται στο χώμα και τις πέτρες, η εισαγωγή στο δίκτυο γειτονιάς εικονοστοιχείων παρουσιάζει πλεονεκτήματα σε σχέση με τις pixel-based μεθόδους. Τα πλεονεκτήματα αυτά εμφανίζονται σε κατηγορίες ομοιογενείς ως προς την φασματική τους υπογραφή αλλά και ως προς την κατανομή τους στον χώρο. Οι καλλιέργειες, οι διαφορετικοί τύποι δασών απαιτούν λεπτομερέστερη αντιμετώπιση καθώς αποτελούν λεπτομερέστερες οντότητες. Η σύνθεση τους είναι πολυπλοκότερη και η διασπορά τους στον χώρο μεγαλύτερη. Κατά συνέπεια οι μικρότερες γειτονιές εικονοστοιχείων διαχωρίζονται με μεγαλύτερη επιτυχία τις οντότητες αυτές.

6.5 ΓΕΝΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΚΑΙ ΠΡΟΤΑΣΕΙΣ

Τα αποτελέσματα των ταξινομήσεων οδηγούν σε διάφορα συμπεράσματα σχετικά με την δομή και λειτουργία των μεθόδων Βαθιάς Μάθησης πάνω στην ταξινόμηση του HyRANK dataset. Επίσης δείχνουν και τη δυναμική τους απέναντι σε μεθόδους Μηχανικής Μάθησης όπως SVM και RF. Από την σκοπιά της ολικής ακρίβειας και του χρόνου παραγωγής της ταξινόμησης, η καθιερωμένη μέθοδος SVM επικρατεί των υπολοίπων, επιτυγχάνοντας την υψηλότερη ολική ακρίβεια. Ακολουθεί η μέθοδος Li et al. με διαφορά μόλις 0.06% και ποιοτικότερη συνολικά ταξινόμηση. Στη συνέχεια οι μέθοδοι που απέμειναν με την RF και Chen et al. να μην ξεπερνούν το 60% της ολικής ακρίβειας.



Πίνακας 33: Διάγραμμα με την ολική ακρίβεια, τον σταθμισμένο μέσο όρο F1 σκορ και kappa των μεθόδων

Παρόλο που επικράτησε η μέθοδος SVM, οι μέθοδοι Βαθιάς Μάθησης έδειξαν πολλές προοπτικές και ιδιότητες που οδηγούν σε καλύτερη διαχωριστική ικανότητα αρκετών κατηγοριών. Τα νευρωνικά δίκτυα έδωσαν απάντηση στην ταξινόμηση κατηγοριών που η μέθοδος SVM δεν κατάφερε να ταξινομήσει με ακρίβεια παραγωγού πάνω από 10%, όπως η κατηγορία 4, 8, 11. Ακόμα η γενικότερη εικόνα της ταξινόμησης ήταν ομοιογενής, με υψηλά σκορ στο σύνολο των κατηγοριών χωρίς αδυναμίες σε ορισμένες κατηγορίες. Δεδομένου ότι αποτελούν αρχιτεκτονικές που χρειάζονται λεπτομερή ρύθμιση ώστε να αποδώσουν αναλόγως των προοπτικών τους, κρίνονται ως αποτελεσματικότερα σε βάθος χρόνου.

Μέσα από τις μεθόδους που παρουσιάστηκαν, παρατηρήθηκαν διάφορες δομές στρωμάτων με ιδιαίτερους στόχους και φιλοσοφίες. Βασική δομή που συναντήθηκε συχνά αποτελούν τα συνελικτικά μπλοκ συνδυάζοντας την φασματική και χωρική πληροφορία. Αυτή η δομή όταν χρησιμοποιείται με μικρότερες γειτονιές εικονοστοιχείων στην εισαγωγή της τάξεως του 5x5, έδειξε να δίνει την ικανότητα στα δίκτυα να παρέχουν ποιοτικές ταξινομήσεις με αντιπροσωπευτικότητα όλων των κατηγοριών στις προβλέψεις ή να ξεχωρίζουν εύκολα ομοιογενείς κατηγορίες όπως εδαφικές(πέτρες και χώμα κλπ.) . Σε κατηγορίες δύσκολα διαχωρίσιμες που απαιτούν λεπτομερέστερη αντιμετώπιση τα μπλοκ γενικεύουν την ταξινόμηση και χάνουν τις λεπτομέρειες αυτές. Το γεγονός αυτό ενισχύουν οι μεγάλες γειτονίες των εικονοστοιχείων και τα συνελικτικά φίλτρα μεγάλου μεγέθους.

Δομές που έχουν να κάνουν με την ανάκτηση πληροφοριών του τύπου residual learning, κρίνονται ως αναποτελεσματικές όταν χρησιμοποιούνται χωρίς να εξυπηρετούν ένα βαθύτερο πλάνο όπως στην μέθοδο Liu et al. , όπου απαντούν στον θόρυβο που εισάγεται στο δίκτυο. Επιπρόσθετα τα αναδραστικά δίκτυα έδειξαν πολύ αποτελεσματικά και ταιριαστά στην ταξινόμηση υπερφασματικών δεδομένων, αφού η μέθοδος Mou et al. αποτελεί μια πολύ απλή μορφή τέτοιου δίκτυου και σημείωσε αρκετά μεγάλη ολική ακρίβεια και ποιότητα στην ταξινόμηση. Ακόμα τα δίκτυα Κωδικοποιητή-Αποκωδικοποιητή έδειξαν και αυτά χρήσιμες ιδιότητες.

Κρίνοντας το τελικό αποτέλεσμα των ταξινομήσεων, το συμπέρασμα που προκύπτει είναι πως αυτό εξαρτάται σε μεγάλο βαθμό από τα δεδομένα προς ταξινόμηση. Οι κατηγορίες που εκφράζουν οι περιοχές εκπαίδευσης, η φύση αυτών, τα φασματικά χαρακτηριστικά των εικονοστοιχείων που αναφέρονται καθώς και ο ειδικός σκοπός της ταξινόμησης ποικίλουν και επηρεάζουν τα τελικά αποτελέσματα και την αξιολόγηση τους. Οι ειδικές δομές που συναντώνται και η κάθε μέθοδος ξεχωρίστα , παρουσιάστηκαν κατάλληλες για διαφορετικούς σκοπούς.

Προκειμένου να καλύψει τις ανάγκες του ο Τοπογράφος Μηχανικός καλείται να προσδιορίσει τον σκοπό αυτό πριν χρησιμοποιήσει την κάθε μέθοδο. Σε περίπτωση που ζητούμενο είναι η παραγωγή μίας γενικής άποψης χάρτη χρήσεων γης, οι μέθοδοι με τα συνελικτικά μπλοκ κρίνονται κατάλληλοι. Σε άλλη περίπτωση όπου χρειάζεται περισσότερη ακρίβεια όπως ο εντοπισμός συγκεκριμένων καλλιεργειών, χαμηλότερες γειτονίες εικονοστοιχείων ή και εντελώς pixel-based μέθοδοι συμπεριφέρονται καλύτερα. Σε μεικτές περιπτώσεις που το πρόβλημα είναι πολύπλοκο, η χρήση residual learning και άλλων δομών δίνει λύσεις.

Στα αποτελέσματα που αναφέρθηκαν δεν παρατηρήθηκε κάποιο μειονέκτημα στην χρήση των αναδραστικών δίκτυων. Ενδιαφέρον θα αποτελούσε η αναζήτηση-σύνθεση ενός νευρωνικού δίκτυου που θα συνδύαζε αναδραστικά στρώματα και μικρές σε μέγεθος συνελίξεις. Έτσι αν το αναδραστικό στρώμα ακολουθούσε το συνελικτικό, θα διατηρούσε στη μνήμη του τους φασματικούς και χωρικούς συσχετισμούς των εικονοστοιχείων.

Άλλη μία ενδιαφέρουσα αρχιτεκτονική θα αποτελούσε ένα hybrid πλήρως συνελικτικό δίκτυο το οποίο θα είχε μεταβλητές γειτονιές εικονοστοιχείων ανάλογα με την κατηγορία στην οποία θα ανήκει το εικονοστοιχείο που εισέρχεται στο δίκτυο κατά την διάρκεια της εκπαίδευσης. Στις κατηγορίες που απαιτούν λεπτομερή διαχωρισμό, τα εικονοστοιχεία θα εισάγονται σε pixel-based μορφή ενώ στις υπόλοιπες μαζί με τα γειτονικά τους. Για τα άγνωστα εικονοστοιχεία κατά τη

διάρκεια της ταξινόμησης μπορεί να αναπτυχθεί ένα στρώμα εισόδου ή ένα ξεχωριστό δίκτυο που θα προχαρακτηρίζει το εικονοστοιχείο για να καθοριστεί η είσοδος.

7 ΣΧΕΤΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ-ΑΝΑΦΟΡΕΣ

7.1 ΞΕΝΟΓΛΩΣΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Ben Hamida, A. Benoit, P. Lambert, and A. Chokri Ben, “Deep Learning Approach for Remote Sensing Image Analysis,” in Big Data from Space (BiDS’16), Santa Cruz de Tenerife, Spain, Mar. 2016, p.133.
- [2] H. Lee and H. Kwon, “Going deeper with contextual CNN for hyper-spectral image classification ,”IEEE Transactions on Image Processing, vol. 26, no. 10, pp. 4843–4855, Oct 2017.
- [3] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, “Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks, ”IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
- [4] Y. Li, H. Zhang, and Q. Shen, “Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network”, Remote Sensing, vol. 9, no. 1, p. 67, Jan. 2017.
- [5] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, “Deep Convolutional Neural Networks for Hyperspectral Image Classification, ”Journal of Sensors, vol. 2015, p. e258619, Jul. 2015.
- [6] M. He, B. Li, and H. Chen, “Multi-scale 3d deep convolutional neural network for hyperspectral image classification,” in Proceedings of the IEEE International Conference on Image Processing (ICIP), Sept 2017,pp. 3904–3908.
- [7] L. Mou, P. Ghamisi, and X. X. Zhu, “Deep Recurrent Neural Networks for Hyperspectral Image Classification, ”IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 7, pp. 3639–3655, Jul.2017
- [8] Wang, Wenju; Dou, Shuguang; Jiang, Zhongmin; Sun, Liujie. 2018. "A Fast Dense Spectral– Spatial Convolution Network Framework for Hyperspectral Images Classification." Remote Sens. 10, no. 7: 1068.
- [9] Liu, X. Yu, P. Zhang, X. Tan, A. Yu, and Z. Xue, “A semi-supervised convolutional neural network for hyperspectral image classification, ”Remote Sensing Letters, vol. 8, pp. 839–848, 2017.
- [10] Nicolas Audebert, Bertrand Saux, Sébastien Lefèvre. Deep Learning for Classification of Hyperspectral Data: A Comparative Review. IEEE geoscience and remote sensing magazine, IEEE, 2019, 7 (2),pp.159-173. 10.1109/MGRS.2019.2912563. hal-02104998
- [11] Karantzalos, Konstantinos, Karakizi, Christina, Kandylakis, Zacharias, & Antoniou, Georgia. (2018). HyRANK Hyperspectral Satellite Dataset I (Version v001) [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.1222202>
- [12] F. Melgani and L. Bruzzone, “Classification of hyperspectral remote sensing images with support vector machines,”IEEE Transactions on Geoscience and Remote Sensing, vol. 42, no. 8, pp. 1778–1790, Aug2004.
- [13] Maltezos, E.; Doulamis, N.; Doulamis, A.; Ioannidis, C. Deep Convolutional Neural Networks for Building Extraction from Orthoimages and Dense Image Matching Point Clouds. J. Appl. Remote Sens.2017,11, 1–22.
- [14] Makantasis, K.; Karantzalos, K.; Doulamis, A.; Doulamis, N. Deep Supervised Learning for Hyperspectral Data Classification through Convolutional Neural Networks. In Proceedings

- of the 2015 IEEE International Geoscience and Remote Sensing Symposium, Milan, Italy, 26–31 July 2015; pp. 4959–4962.
- [15] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016 ;pp. 770–778.
- [16] Deng, S.; Xu, Y.; He, Y.; Yin, J.; Wu, Z. A Hyperspectral Image Classification Framework and Its Application. Inf. Sci. 2015, 299, 379–393.

7.2 ΙΣΤΟΤΟΠΟΙ

- [1] <http://cs231n.stanford.edu/slides>
- [2] <https://www.wikipedia.org/>
- [3] www.deeplearning.net
- [4] [www.ieeexplore.ieee.org](http://ieeexplore.ieee.org)
- [5] <http://cs231n.github.io/convolutional-networks/>
- [6] <https://github.com/torch/torch7/wiki/Cheatsheet>
- [7] <https://towardsai.net>
- [8] <https://towardsdatascience.com/>
- [9] <https://www.lutzroeder.com/ai/>
- [10] <https://scikit-learn.org>