

Ταξινόμηση πολυφασματικής εικόνας WorldView-2 με τη μέθοδο της σημασιολογικής κατάτμησης εφαρμόζοντας βαθιές εκμάθησης τεχνητά νευρωνικά δίκτυα

Περίληψη

Η ταξινόμηση αποτελεί μία από τις βασικές εφαρμογές της Τηλεπισκόπησης με στόχο την εξαγωγή των τύπων εδαφικάλυψης/χρήσεων γης από μία δορυφορική εικόνα και έχουν αναπτυχθεί διάφορες μέθοδοι επιβλεπόμενης και μη επιβλεπόμενης ταξινόμησης. Στόχος της παρούσας εργασίας είναι η ταξινόμηση πολυφασματικής εικόνας απεικόνισης τμήματος της νοτιοανατολικής Λέσβου με χρήση της μεθόδου της σημασιολογικής κατάτμησης εικόνας εφαρμόζοντας βαθιάς εκμάθησης τεχνητό νευρωνικό δίκτυο. Σκοπός αυτής είναι η χαρτογράφηση της εδαφικάλυψης της περιοχής και η αξιολόγηση του μοντέλου που δημιουργήθηκε. Η μεθοδολογία βασίζεται στο συνδυασμό τεχνικών τηλεπισκόπησης και υπολογιστικής όρασης σε μια Worldview-2 (WV-2) εικόνα φθινοπωρινής λήψης του 2011. Απαραίτητη ενέργεια πριν την επεξεργασία αυτής ήταν η σύζευξη των καναλιών της, έτσι ώστε να διατηρηθεί η χωρική διακριτική ικανότητα του πανχρωματικού διακομιστή, δηλαδή 0,46εκ. και η φασματική πληροφορία του πολυφασματικού διακομιστή. Ένα βαθιάς εκμάθησης τεχνητό νευρωνικό δίκτυο δημιουργήθηκε και εκπαιδεύτηκε μέσω μιας εικόνας και δεδομένων εκπαίδευσης που καθορίστηκαν πάνω σε αυτή. Εκπαιδεύτηκαν δύο δίκτυα, ένα χρησιμοποιώντας τα οχτώ κανάλια της αρχικής εικόνα και ένα χρησιμοποιώντας μόνο τα κανάλια του ορατού φάσματος (RGB). Τα εκπαιδευμένα δίκτυα εφαρμόστηκαν στις αρχικές συνθέσεις των εικόνων για τη δημιουργία των ταξινομημένων εικόνων. Ένα σύνολο τριακοσίων τυχαίων σημείων αναφοράς συλλέχθηκαν για επαλήθευση και τα παραχθέντα προϊόντα της ταξινόμησης είναι θεματικοί χάρτες κάλυψης γης της περιοχής, γραφήματα απόδοσης ακρίβειας και απώλειας των minibatches από τη διαδικασία εκπαίδευσης του δικτύου, ιστογράμματα και πίνακες κατανομής των εικονοστοιχείων ανά κλάση, πίνακες σύγχυσης για την αξιολόγηση των ταξινομήσεων συνολικά και εικόνες συγκριτικής αξιολόγησης των κλάσεων ανά εικόνα. Το κύριο συμπέρασμα της εργασίας είναι ότι η αρχιτεκτονική δομή του δικτύου, καθώς και ο αριθμός των δεδομένων εκπαίδευσης και επαλήθευσης επηρεάζουν την ακρίβεια της ταξινόμησης.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Σημασιολογική κατάτμηση, τεχνικές βαθιές εκμάθησης, τεχνητό νευρωνικό δίκτυο, WorldView-2

Εισαγωγή

Η ευρέως εφαρμοζόμενη μέθοδος της σημασιολογικής κατάτμησης εικόνας σε διάφορους επιστημονικούς κλάδους, όπως αυτός της τεχνητής νοημοσύνης, της ιατρικής, της ρομποτικής, της δορυφορικής τηλεπισκόπησης, κ.ά., αποτελεί ένα από τα κυριότερα προβλήματα της υπολογιστικής όρασης (Garcia-Garcia et al., 2017 & Guo et. al., 2018). Σε αυτό το επίπεδο ανάλυσης της εικόνας, στόχος είναι η αντιστοίχιση κάθε εικονοστοιχείου της με μια προκαθορισμένη κλάση. Επομένως, η φύση του προβλήματος της σημασιολογικής

κατάτμησης μπορεί να θεωρηθεί ως ένα πρόβλημα ταξινόμησης και για αυτό το λόγο χρησιμοποιείται πλέον και στην Τηλεπισκόπηση. Πεδία εφαρμογής αυτής σε αυτόν τον κλάδο είναι η χαρτογράφηση της γήινης επιφάνειας, η ανίχνευση των αλλαγών, η διαχείριση κινδύνων, η αστικοποίηση, κ.ά.

Η μέθοδος της κατάτμησης εικόνας αποτελεί, παραδοσιακά, θεμελιώδη αρχή για την ταξινόμηση δορυφορικών εικόνων υψηλής ανάλυσης που βασίζεται στην ανάλυση αντικειμένου (OBIA). Τα τελευταία χρόνια, διάφορες μέθοδοι κατακερματισμού των τηλεπισκοπικών δεδομένων έχουν προταθεί, οι οποίες μπορούν να κατηγοριοποιηθούν σε τέσσερις μεγάλες κατηγορίες: (α) αυτές που βασίζονται στα εικονοστοιχεία (pixel-based), (β) αυτές που βασίζονται στις ακμές (edge-based), (γ) αυτές που βασίζονται στην περιοχή (region-based) και (δ) αυτές που βασίζονται στην γνώση (knowledge-based) (Basaeed, Bhaskar & Al-Mualla, 2016). Όλες οι προαναφερθείσες μέθοδοι ταξινόμησης – πλην της τελευταίας – βασίζονται σε ρηχές αρχιτεκτονικές, οι οποίες αποτυγχάνουν να συλλάβουν τα “ψιλά” χαρακτηριστικά των σύνθετων εικόνων κάλυψης γης. Συνεπώς, καμία από τις τρεις παραπάνω μεθόδους δεν επιτυγχάνει το υψηλό επίπεδο ακρίβειας που απαιτείται για την ταξινόμηση κάλυψης γης, πόσω μάλλον της κάλυψης αστικής γης (Huang, Zhao & Song, 2017). Επομένως, η ανάγκη εφαρμογής μεθόδων βαθιάς αρχιτεκτονικής για την εξαγωγή και κατηγοριοποίηση της πλούσιας χωρικής, φασματικής και χρονικής πληροφορίας των τεράστιων σε όγκο δορυφορικών δεδομένων είναι επιβεβλημένη (Ben Hamida et al., 2017). Παρά την πρόσφατη αύξηση του ενδιαφέροντος της επιστημονικής κοινότητας σχετικά με τα βαθιά νευρωνικά δίκτυα, η δομή τους αποτελεί αντικείμενο μελέτης από το 1990, όταν οι LeCun et al. πρότειναν μια αρχιτεκτονική νευρωνικού δικτύου για αναγνώριση του γραφικού χαρακτήρα. Έκτοτε χρειάστηκε να περάσει μια εικοσαετία μέχρι οι A. Krizhevsky, I. Sutskever & G. E. Hinton (2012) να προτείνουν την αρχιτεκτονική του AlexNet – ένα βαθύ συνελκτικό νευρωνικό δίκτυο (Convolution Neural Network – CNN) – το οποίο κέρδισε το διαγωνισμό αναγνώρισης εικόνας μεγάλης κλίμακας δεδομένων ImageNet με ακρίβεια ελέγχου 84,6%. Ακολούθησαν τα VGG-16, GooLeNet, ResNet, κ.ά., τα οποία λόγω της σπουδαιότητάς τους, χρησιμοποιούνται πλέον ως δομικά στοιχεία σε διάφορες αρχιτεκτονικές κατάτμησης (Garcia-Garcia et al., 2017).

Μερικές πρόσφατες έρευνες χρησιμοποιούν δίκτυα βαθιάς εκμάθησης για να επεξεργαστούν δορυφορικές εικόνες. Οι Papadomanolaki et al. (2016) συνέκριναν την απόδοση συγκεκριμένων μοντέλων βαθιάς αρχιτεκτονικής για την ταξινόμηση πολυφασματικών δεδομένων. Βασιζόμενοι στο DeepSat σύνολο δεδομένων, εκπάιδευσαν διαφορετικά δίκτυα (AlexNet, AlexNet-small και VGG), τα οποία τα συνέκριναν με τα Deep Belief, Autoencoders και Semiservervised δίκτυα. Η ακρίβεια της ταξινόμησης από τα AlexNet και VGG ξεπέρασε το 99,9%. Αντίστοιχα, οι Pirrotti, Sunar & Piragnolo (2016) έκαναν μια συγκριτική αξιολόγηση εννέα διαφορετικών αλγορίθμων ταξινόμησης που υπογραμμίζουν την επάρκεια και αποτελεσματικότητα των δεδομένων Sentinel-2 για χαρτογράφηση κάλυψης γης. Από την άλλη μεριά, οι Ben Hamida et al. (2017) στηριζόμενοι στις βαθιές αρχιτεκτονικές των DenseNet και SegNet χρησιμοποίησαν τη μέθοδο της σημασιολογικής κατάτμησης σε μεγάλες πολυφασματικές Sentinel-2 εικόνες για τη χαρτογράφηση της εδαφοκάλυψης μιας περιοχής ανάμεσα στη Γαλλία, την Ελβετία και την Ιταλία. Εισηγάγαν επίσης ένα καινούριο 3D DenseNet δίκτυο έτσι ώστε να επεξεργαστούν από κοινού τις χωρικές και φασματικές διαστάσεις και πρότειναν τη χρήση ενός συνόλου δεδομένων αλήθειας με θόρυβο (δηλαδή ξεπερασμένα και χαμηλής χωρικής διακριτικής ικανότητας για την εκπαίδευση και τον έλεγχο του δικτύου. Η ιδέα ήταν να εξερευνήσουν την

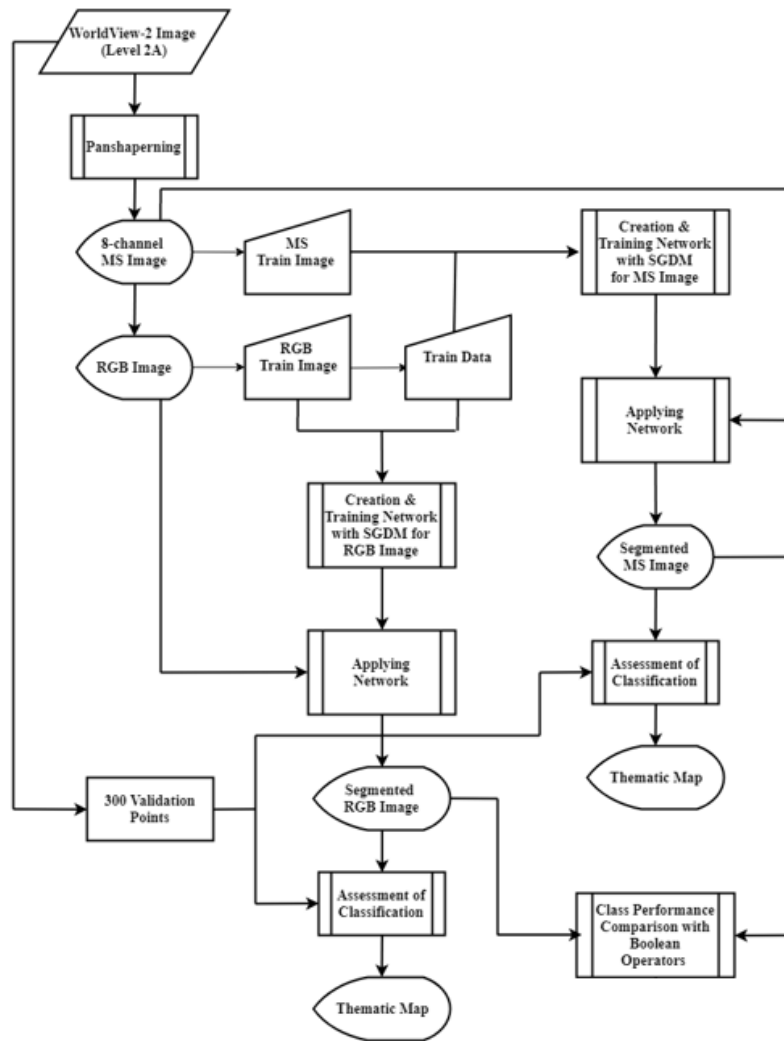
εφαρμοσιμότητα των δεδομένων αυτών σε σύγχρονους αισθητήρες και καινοτόμες μεθόδους ανάλυσης.

Ο Liebel (2018) μη έχοντας στη διάθεσή του κατάλληλο σετ δεδομένων για την εκπαίδευση ενός νευρωνικού δικτύου βαθειάς εκμάθησης για μεγάλης κλίμακας ταξινόμηση κάλυψης/χρήσης γης, χρησιμοποίησε δεδομένα εκπαίδευσης από υπάρχουσες πηγές εθελοντικής γεωγραφικής πληροφορίας, και συγκεκριμένα από το OpenStreetMap, τα οποία συνέλεξε μέσω μιας ημι-αυτόματης μεθόδου. Τα αποτελέσματα της ταξινόμησης υποδεικνύουν ότι τα CNNs είναι κατάλληλα για ταξινόμηση κάλυψης/χρήσης γης με κάποιες μικρές προσαρμογές και επίσης, είναι εύρωστα στα δεδομένα με θόρυβο. Οι López, Santos & Atzberger (2018) δημιούργησαν ένα τεχνητό νευρωνικό δίκτυο βαθειάς εκμάθησης, βασισμένο στο SegNet, για σημασιολογική κατάτμηση εικόνας Sentinel-2 MSI. Αυτό το μοντέλο ήταν ικανό να ταξινομήσει σε επίπεδο εικονοστοιχείου, πέντε διαφορετικές κλάσεις, χρησιμοποιώντας κανάλια υψηλής ανάλυσης – επομένως, το βάθος του ήταν σχετικά μικρό. Το αποτέλεσμα της ταξινόμησης (κοντά στο 90%) έδειξε ότι το εν λόγω δίκτυο μπορεί να πραγματοποιήσει κατάτμηση των κλάσεων με καλύτερες ιδιότητες, ακόμα και όταν ο αριθμός των δεδομένων εκπαίδευσης είναι μικρός. Τέλος, οι Kemker, Salvaggio & Kanan (2017) εντόπισαν το κενό στα δημοσίως διαθέσιμα δεδομένα εκπαίδευσης προερχόμενα από πολυφασματικές εικόνες και δημιούργησαν ένα σύνολο δεδομένων τα οποία συλλέχθηκαν από μη επανδρωμένο εναέριο σύστημα. Οι δεκαοχτώ κλάσεις που περιέχει χρησιμοποιήθηκαν για την αξιολόγηση μοντέλων σημασιολογικής κατάτμησης που σχεδιάστηκαν για πολυφασματικές εικόνες. Η μεγάλη χωρική μεταβλητότητα σε συνδυασμό με υψηλής ανάλυσης εικόνες, μεγάλο μέγεθος εικονοστοιχείων, μικρά και κρυφά αντικείμενα και οι άνισες κλάσεις στάθηκαν εμπόδιο στην καλή απόδοση του δικτύου.

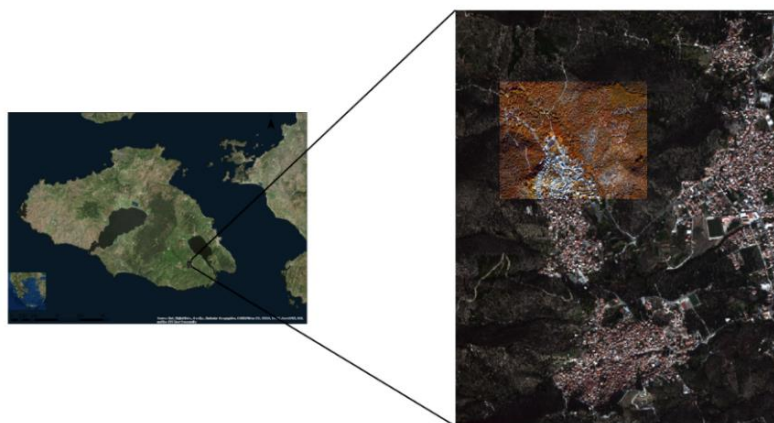
Η παρούσα εργασία παρουσιάζει τα αποτελέσματα της ταξινόμησης μιας πολυφασματικής εικόνας WV-2 του 2011, αλλά και της αντίστοιχης RGB της, με τη μέθοδο της σημασιολογικής κατάτμησης εφαρμόζοντας βαθειά εκμάθηση τεχνητό νευρωνικό δίκτυο. Συγκεκριμένα, δημιουργήθηκε, εκπαιδεύτηκε και αξιολογήθηκε ένα CNN μικρού βάθους. Τμήμα της εικόνας αποτέλεσε τη βάση εκπαίδευσης των δεδομένων και οι κλάσεις που δημιουργήθηκαν ήταν τρεις: γυμνό έδαφος, βλάστηση και δομημένο περιβάλλον.

Δεδομένα και μεθοδολογία

Η μεθοδολογία που ακολουθήθηκε βασίστηκε στην προεπεξεργασία της δορυφορικής εικόνας WorldView-2, χωρικής ανάλυσης 0,46μ. (στο παγχρωματικό) και στη συνέχεια στη δημιουργία ενός τεχνητού νευρωνικού δικτύου για την ταξινόμηση αυτής. Η Εικόνα 1, παρουσιάζει τη μεθοδολογία της παρούσας εργασίας. Ως περιοχή μελέτης επιλέχθηκε η γεωγραφική έκταση περίξ των τεσσάρων από τα έξι χωριά της Γέρας Λέσβου – Σκόπελος, Μεσαγρός, Παππάδος και Πλακάδος – που βρίσκονται στο νοτιοανατολικό τμήμα του νησιού. Τα δεδομένα που χρησιμοποιήθηκαν είναι μια δορυφορική εικόνα WorldView-2 (επίπεδο επεξεργασίας Level 2A) με οχτώ φασματικά κανάλια χωρικής διακριτικής ικανότητας 1,84μ. και ένα κανάλι στο παγχρωματικό με χωρική διακριτική ικανότητα 0,46μ. Η εν λόγω δορυφορική εικόνα λήφθηκε στις 19 Νοεμβρίου 2011 και επιλέχθηκε λόγω της καθαρότητάς της (απουσία σύννεφων). Στην Εικόνα 2 παρουσιάζεται η περιοχή μελέτης.



Εικόνα 1: Διάγραμμα μεθοδολογίας



Εικόνα 2: Περιοχή μελέτης εργασίας

Σύμφωνα με τους Li, Jing & Tang (2017) οι WV-2 εικόνες έχουν ευρεία χρήση σε διάφορα πεδία μελέτης που απαιτούν υψηλής χωρικής διακριτικής ικανότητας δεδομένα. Η σύντμηση

των οχτώ πολυφασματικών καναλιών και της παγχρωματικής μπάντας θεωρείται επιβεβλημένη για την καλύτερη ερμηνεία της γήινης επιφάνειας. Επομένως, κατά το στάδιο της προεπεξεργασίας έγινε μια προσπάθεια βελτίωσης της χωρικής ανάλυσης της δορυφορικής εικόνας, εφαρμόζοντας τη μέθοδο της σύζευξης εικόνας μεταξύ του παγχρωματικού καναλιού και των πολυφασματικών καναλιών. Αποτέλεσμα της διαδικασίας ήταν μια εικόνα οχτώ καναλιών με τη χωρική διακριτική ικανότητα του παγχρωματικού διακομιστή, δηλαδή 0,46μ. και τη φασματική πληροφορία του πολυφασματικού διακομιστή. Ο περιορισμός εισαγωγής πολυφασματικής εικόνας στο περιβάλλον εργασίας του Image Labeler του MATLAB (έκδοση 2018a) για τον καθορισμό των ετικετών (labels) στα εικονοστοιχεία, μας ώθησε να απομονώσουμε τα τρία κανάλια RGB από την 8-κάναλη πολυφασματική εικόνα. Έτσι, δημιουργήσαμε μια εικόνα εκπαίδευσης, της οποίας η χρωματική σύνθεση ήταν 7-6-3 για την ευκολότερη αναγνώριση των κλάσεων ταξινόμησης. Παράλληλα, χρειάστηκε να ενισχυθεί η χρωματική διαβάθμιση της απεικόνισης των διαφόρων καλύψεων γης στην οθόνη, μέσω της τεχνικής της γραμμικής ενίσχυσης της αντίθεσης (γραμμικό τέντωμα).

Για την εκπαίδευση του δικτύου, μια αντιπροσωπευτική 3-κάναλη εικόνα της περιοχής μελέτης, διαστάσεων 1484x1272 αποκόπηκε από τη συνολική εικόνα, έτσι ώστε να λειτουργήσει ως βάση για την αντιστοίχιση κάθε εικονοστοιχείου αυτής με μια από τις κατηγορίες ταξινόμησης. Η πλειοψηφία των εικονοστοιχείων της εικόνας κατηγοριοποιήθηκε σε μια από τις τρεις κλάσεις: 1. γυμνό έδαφος – bare, 2. βλάστηση – vegetation και 3. δομημένο περιβάλλον – urban.

Μετά τον καθορισμό των φασματικών υπογραφών, δημιουργήθηκε ένα τεχνητό νευρωνικό δίκτυο δέκα επιπέδων. Στο πρώτο επίπεδο βρίσκεται το επίπεδο εισόδου όπου εισαγάγαμε την εικόνα εκπαίδευσης (train image). Το συνελκτικό επίπεδο που ακολουθεί δέχεται ως είσοδο τα patches διαστάσεων 3x128x128 για την RGB εικόνα και 8x128x128 για την πολυφασματική εικόνα, αντίστοιχα και τα επεξεργάζεται με φίλτρα διαστάσεων 3x3. Το βήμα (stride) συνέλιξης ορίστηκε σε 1 εικονοστοιχείο, το χωρικό γέμισμα (padding) σε 1 εικονοστοιχείο για 3x3 συνελκτικά επίπεδα (για τη διατήρηση της χωρικής ανάλυσης μετά τη συνέλιξη). Το τρίτο επίπεδο εφαρμόζει τη συνάρτηση ReLU στον παράγωγο πίνακα. Το τέταρτο επίπεδο είναι είδους max pooling και εφαρμόζει φίλτρα διαστάσεων 2x2 με βήμα 1. Το πέμπτο και έκτο επίπεδο ακολουθεί την ίδια λογική (Convolutional – ReLU), ενώ το έβδομο ακολουθεί συνέλιξη προς την αντίθετη κατεύθυνση (Transposed Convolutional). Τα φίλτρα που χρησιμοποιεί για το μετασχηματισμό είναι διαστάσεων 4x4 με βήμα 1 και cropping 1. Το όγδοο επίπεδο είναι και πάλι ένα συνελκτικό επίπεδο, όπου στις ρυθμίσεις παραμέτρων χρησιμοποιήσαμε 1x1 φίλτρο συνέλιξης, το οποίο μπορεί να θεωρηθεί σαν ένας γραμμικός μετασχηματισμός των εισερχόμενων καναλιών. Στο ένατο επίπεδο εφαρμόζεται συνάρτηση ενεργοποίησης Softmax, η οποία αποδίδει δεκαδικές πιθανότητες σε κάθε κατηγορία. Το τελευταίο επίπεδο δίνει ως έξοδο την ταξινομημένη εικόνα, δηλαδή έναν τελικό πίνακα διαστάσεων 3x1, επειδή οι κατηγορίες εκπαίδευσης ήταν τρεις (Πίνακας 1).

Η εκπαίδευση του δικτύου ακολούθησε εισάγοντας σε αυτό την εικόνα (train image) και τα δεδομένα εκπαίδευσης (train data). Ως αλγόριθμος εκπαίδευσης επιλέχθηκε ο αλγόριθμος κατάβασης κλίσης με Minibatches και ορμή ορού (Stochastic Gradient Descent with Minibatches and Momentum–SGDM), ο οποίος εφαρμόστηκε επαναληπτικά σε κάθε παράδειγμα εκπαίδευσης. Με βάση αυτή την επιλογή, οι υπερπαραμέτροι που καθορίστηκαν για την εκπαίδευση του μοντέλου ήταν ο αρχικός ρυθμός μάθησης (initial learning rate) = 0,001, ο οποίος ήταν σταθερός καθ' όλη τη διάρκεια της εκπαίδευσης, ο αριθμός των εποχών = 30, το μέγεθος του υποσυνόλου των δεδομένων εκπαίδευσης που

χρησιμοποιείται για τον υπολογισμό της κλίσης του κόστους ως προς τα βάρη του μοντέλου (Minibatch size = 16) και η σταθερά κανονικοποίησης των βαρών.

Πίνακας 1. Ρυθμίσεις Παραμέτρων Δικτύου

ANALYSIS RESULT					
↑	NAME	TYPE	ACTIVATIONS	LEARNABLES	TOTAL LEARNABLES
1	imageinput 128x128x3 images with 'zerocenter' normalization	Image Input	128x128x3	-	0
2	conv_1 16 3x3x3 convolutions with stride [1 1] and padding [1 1 1 1]	Convolution	128x128x16	Weights 3x3x3x16 Bias 1x1x16	448
3	relu_1 ReLU	ReLU	128x128x16	-	0
4	maxpool 2x2 max pooling with stride [1 1] and padding [0 0 0 0]	Max Pooling	127x127x16	-	0
5	conv_2 16 3x3x16 convolutions with stride [1 1] and padding [1 1 1 1]	Convolution	127x127x16	Weights 3x3x16x16 Bias 1x1x16	2320
6	relu_2 ReLU	ReLU	127x127x16	-	0
7	transposed-conv 16 4x4x16 transposed convolutions with stride [1 1] and output cropping [1 1]	Transposed Convolution	128x128x16	Weights 4x4x16x16 Bias 1x1x16	4112
8	conv_3 3 1x1x16 convolutions with stride [1 1] and padding [0 0 0 0]	Convolution	128x128x3	Weights 1x1x16x3 Bias 1x1x3	51
9	softmax softmax	Softmax	128x128x3	-	0
10	classoutput Cross-entropy loss with 'bare', 'vegetation', and 1 other classes	Pixel Classification ...	-	-	0

Η εφαρμογή του δικτύου στην πολυφασματική εικόνα και την αντίστοιχη 3-κανάλι RGB είχε ως αποτέλεσμα την ταξινόμηση των εικόνων σύμφωνα με τις τρεις προκαθορισμένες κατηγορίες. Για την αξιολόγηση του αποτελέσματος της ταξινόμησης, τριακόσια σημεία αναφοράς, διάσπαρτα καταναμεμένα σε ολόκληρη την WV-2 εικόνα, ψηφιοποιήθηκαν και κατηγοριοποιήθηκαν στις αντίστοιχες κλάσεις των δεδομένων εκπαίδευσης. Η τυχαιότητα επιλέχθηκε ώστε να εξασφαλιστεί η αξιοπιστία της διόρθωσης.

Μέσω της παρατήρησης των πινάκων ταξινόμησης αξιολογήθηκαν τα αποτελέσματα των ταξινομήσεων. Συγκεκριμένα από τους πίνακες ταξινόμησης βρέθηκαν οι ακρίβειες του χρήστη και του παραγωγού, η συνολική ακρίβεια και ο δείκτης Kappa. Στη συνέχεια, έγινε η οπτικοποίηση των αποτελεσμάτων της ταξινόμησης κάλυψης γης. Τέλος, μέσω Boolean operators έγινε μια συγκριτική αξιολόγηση της απόδοσης κάθε μιας κλάσης ξεχωριστά ανά εικόνα.

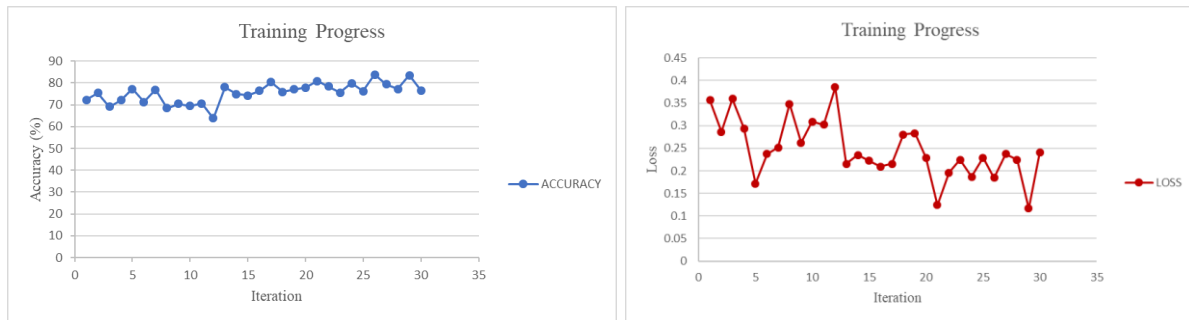
Αποτελέσματα και Συζήτηση

Ταξινόμηση RGB εικόνας

Η εκπαίδευση του δικτύου για την ταξινόμηση της RGB εικόνας διήρκεσε 229 λεπτά και 42 δευτερόλεπτα χωρίς να ενεργοποιηθεί η παράλληλη επεξεργασία και χωρίς την ύπαρξη κάρτας γραφικών. Ο κύκλος εκπαίδευσης είχε τα εξής χαρακτηριστικά: το σύνολο των εποχών είχε προκαθοριστεί στις 30. Οι επαναλήψεις που έγιναν ήταν 9.000, καθώς κάθε εποχή αποτελούνταν από 300 επαναλήψεις. Το πρόγραμμα του ρυθμού εκμάθησης ήταν σταθερό και ορίστηκε εξ αρχής σε 0,001.

Η τελική ακρίβεια του δικτύου υπολογίστηκε στο 76,6% και η απώλεια του στο 0,2404. Η μέση τιμή της ακρίβειας του δικτύου είναι 75,45%, ενώ η απώλεια είναι 0,25 περίπου. Η μέγιστη τιμή ακρίβειας που σημειώθηκε ήταν 83,79%, ενώ η ελάχιστη ήταν 63,85%. Αντίστοιχα, η μέγιστη τιμή απώλειας ήταν 0,39 περίπου, ενώ η ελάχιστη ήταν 0,12 περίπου.

Στην Εικόνα 3 παρουσιάζονται οι διακυμάνσεις της ακρίβειας και της απώλειας, αντίστοιχα της διαδικασίας εκμάθησης ανά εποχή:



Εικόνα 3: Απόδοση ακρίβειας & απώλειας δικτύου

Ο θεματικός χάρτης κάλυψης γης (Εικόνα 4) παρουσιάζει τα αποτελέσματα της ταξινόμησης της 3-κάναλης εικόνας.



Εικόνα 4: Χάρτης κάλυψης γης περιοχής μελέτης (ταξινόμηση RGB εικόνας)

Η έκταση που έχει καλύψει κάθε κλάση στην περιοχή μελέτης υπολογίστηκε από τον αριθμό των εικονοστοιχείων που έλαβαν την αντίστοιχη τιμή κατηγορίας επί το $GSD=0,46\mu$. Στον Πίνακα 2 που ακολουθεί φαίνονται αναλυτικά τα παραπάνω:

Πίνακας 2. Εδαφοκάλυψη ανά κατηγορία			
Κατηγορία	Αριθμός εικονοστοιχείων	GSD (μ)	Έκταση (τ.μ.)
Γυμνό έδαφος	985.692		453.418,32
Βλάστηση	12.600.819	0,46	5.796.377,00
Δομημένο περιβάλλον	1.833.069		843.211,70

Παρατηρούμε ότι η κατηγορία “βλάστηση” κυριαρχεί στην περιοχή με την κλάση “δομημένο περιβάλλον” να ακολουθεί, ενώ τελευταία σε κατάληψη γεωγραφικής έκτασης είναι η κατηγορία “γυμνό έδαφος”.

Με την εφαρμογή του μοντέλου ταξινόμησης που δημιουργήσαμε στην RGB εικόνα και στα δεδομένα αναφοράς εξήχθη ο πίνακας ταξινόμησης (Εικόνα 5), τον οποίο παρατηρώντας διαπιστώνουμε ότι η ταξινόμηση κατέληξε με συνολική ακρίβεια Overall Accuracy=75,7%. Υπολογίζοντας τον συντελεστή Kappa coefficient=0,47 διαπιστώνουμε ότι η ταξινόμηση της RGB εικόνας απέφυγε το 47% των σφαλμάτων που δημιουργούνται από μια τυχαία ταξινόμηση. Οι χαμηλότερες ακρίβειες παρατηρήθηκαν στην κλάση “γυμνό έδαφος”. Τα χαμηλά ποσοστά αυτής της κλάσης αιτιολογούνται κυρίως λόγω του μικρού αριθμού δεδομένων εκπαίδευσης και αναφοράς συγκριτικά με τις άλλες δυο κλάσεις. Η κλάση “βλάστηση” φαίνεται ότι ταξινομήθηκε σωστά σχεδόν πλήρως με ακρίβεια που αγγίζει το 99,5% , ενώ φτωχή απόδοση ταξινόμησης παρατηρείται και στην κλάση “δομημένο περιβάλλον” με 47,3%. Αυτό οφείλεται στην πρώτη περίπτωση λόγω υψηλού αριθμού δεδομένων εκπαίδευσης (η “βλάστηση” είναι η κυρίαρχη κλάση στην εικόνα εκπαίδευσης), ενώ στη δεύτερη περίπτωση, η σύγχυση οφείλεται στο ότι τα εικονοστοιχεία του δομημένου περιβάλλοντος είναι φασματικά όμοια με εκείνα του γυμνού εδάφους. Παρά τις αποκλίσεις που παρατηρήθηκαν, το συνολικό αποτέλεσμα του μοντέλου κρίνεται μέτριο προς ικανοποιητικό.

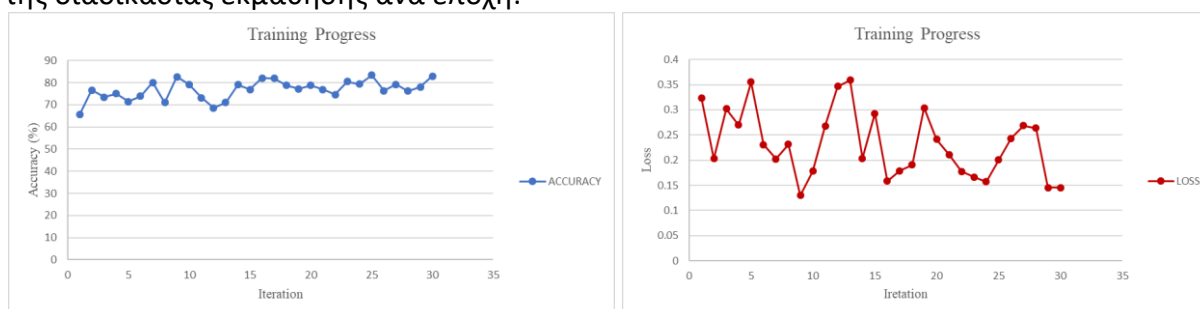
Confusion Matrix			
Output Class	1	2	3
	7 2.3%	1 0.3%	14 4.7%
	33 11.0%	185 61.7%	25 8.3%
	0 0.0%	0 0.0%	35 11.7%
Target Class			
1	17.5% 82.5%	99.5% 0.5%	47.3% 52.7%
2			
3			

Εικόνα 5: Πίνακας ταξινόμησης RGB εικόνας

Ταξινόμηση MS εικόνας

Η εκπαίδευση του δικτύου για την ταξινόμηση της πολυφασματικής εικόνας διήρκησε 176 λεπτά και 18 δευτερόλεπτα χωρίς να ενεργοποιηθεί η παράλληλη επεξεργασία και χωρίς την ύπαρξη κάρτας γραφικών. Ο κύκλος εκπαίδευσης είχε τα ίδια χαρακτηριστικά με την εκπαίδευση του δικτύου για την ταξινόμηση της RGB εικόνας.

Η τελική ακρίβεια του δικτύου υπολογίστηκε στο 83,01% και η απώλεια του στο 0,1583. Η μέση τιμή της ακρίβειας του δικτύου είναι 76,82%, ενώ η απώλεια είναι 0,23 περίπου. Η μέγιστη τιμή ακρίβειας που σημειώθηκε ήταν 83,45%, ενώ η ελάχιστη ήταν 65,57%. Αντίστοιχα, η μέγιστη τιμή απώλειας ήταν 0,36 περίπου, ενώ η ελάχιστη ήταν 0,13 περίπου. Στην Εικόνα 6 παρουσιάζονται οι διακυμάνσεις της ακρίβειας και της απώλειας, αντίστοιχα της διαδικασίας εκμάθησης ανά εποχή:



Εικόνα 6: Απόδοση ακρίβειας & απώλειας δικτύου

Αντίστοιχα, στην Εικόνα 7 παρουσιάζονται τα αποτελέσματα της ταξινόμησης της 8-κάναλης εικόνας. Επιπλέον, στον Πίνακα 3 φαίνεται η έκταση που έχει καλυφθεί ανά κλάση.

Πίνακας 3. Εδαφοκάλυψη ανά κατηγορία

Κατηγορία	Αριθμός εικονοστοιχείων	GSD (μ)	Έκταση (τ.μ.)
Γυμνό έδαφος	1.202.584	0,46	553.188,60
Βλάστηση	12.597.066		5.794.650,00
Δομημένο περιβάλλον	1.619.930		745.167,80

Παρατηρούμε ότι και στην περίπτωση της ταξινόμησης της πολυφασματικής εικόνας η κατηγορία “βλάστηση” κυριαρχεί στην περιοχή με την κλάση “δομημένο περιβάλλον” να ακολουθεί, ενώ τελευταία σε κατάληψη γεωγραφικής έκτασης είναι η κατηγορία “γυμνό έδαφος”.

Με την εφαρμογή του μοντέλου ταξινόμησης που δημιουργήσαμε στην πολυφασματική εικόνα αυτή τη φορά και στα δεδομένα αναφοράς εξήχθη ο πίνακας ταξινόμησης (Εικόνα 8). Παρατηρώντας αυτόν διαπιστώνουμε ότι η ταξινόμηση κατέληξε με συνολική ακρίβεια, Overall Accuracy=73,3%. Υπολογίζοντας τον συντελεστή Kappa coefficient=0,43 διαπιστώνουμε ότι η ταξινόμηση της MS εικόνας απέφυγε περίπου το 43% των σφαλμάτων που δημιουργούνται από μια τυχαία ταξινόμηση. Οι χαμηλότερες ακρίβειες παρατηρήθηκαν και σε αυτή την περίπτωση στην κλάση “γυμνό έδαφος” (17,5% - ποσοστό ίδιο με αυτό της ταξινόμησης της RGB εικόνας). Τα χαμηλά ποσοστά αυτής της κλάσης αιτιολογούνται, όπως προαναφέραμε, κυρίως λόγω του μικρού αριθμού δεδομένων εκπαίδευσης και αναφοράς. Η κλάση “βλάστηση” φαίνεται ότι ταξινομήθηκε σωστά σχεδόν πλήρως με ακρίβεια που αγγίζει

το 98,9% , ενώ χειρότερη απόδοση ταξινόμησης, συγκριτικά με το αντίστοιχο αποτέλεσμα στην RGB εικόνα, παρατηρείται στην κλάση “δομημένο περιβάλλον” (39,2%).



Εικόνα 7: Χάρτης κάλυψης γης περιοχής μελέτης (ταξινόμηση MS εικόνας)

Confusion Matrix				
Output Class	1	7 2.3%	2 0.7%	22 7.3% 22.6% 77.4%
	2	33 11.0%	184 61.3%	23 7.7% 76.7% 23.3%
	3	0 0.0%	0 0.0%	29 9.7% 100% 0.0%
		17.5% 82.5%	98.9% 1.1%	39.2% 60.8% 73.3% 26.7%
		Target Class		
		1	2	3

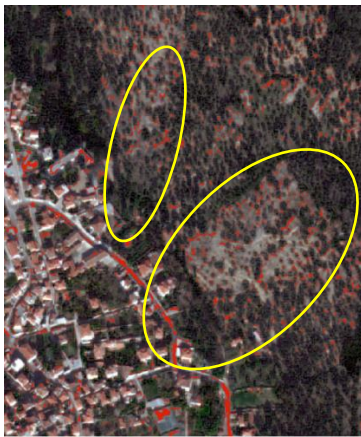
Εικόνα 8: Πίνακας ταξινόμησης MS εικόνας

Ενώ θα περίμενε κανείς μεγαλύτερη ακρίβεια ταξινόμησης της πολυφασματικής εικόνας σε σχέση με την RGB λόγω της έλλειψης ευρύ φάσματος καναλιών, τελικά το αποτέλεσμα ήταν το αντίθετο. Ωστόσο, και σε αυτή την περίπτωση το συνολικό αποτέλεσμα κρίνεται μέτριο προς ικανοποιητικό.

Θέλοντας να αποκτήσουμε μια πληρέστερη εικόνα των αποτελεσμάτων των ταξινομήσεων, εξετάσαμε πιο προσεκτικά κάποιες υποπεριοχές της περιοχής μελέτης. Μέσω της συγκριτικής αξιολόγησης της απόδοσης κάθε κλάσης ανά εικόνα, εντοπίστηκαν οι ασυμφωνίες μεταξύ των δύο μοντέλων.

Κλάση 1 – Γυμνό Έδαφος

Στην Εικόνα 9 παρουσιάζεται το γυμνό έδαφος που ταξινομήθηκε σωστά από την RGB εικόνα και λανθασμένα από την MS, ενώ στην Εικόνα 10 το αντίστροφο.



Εικόνα 9. Αποτέλεσμα ταξινόμησης RGB εικόνας

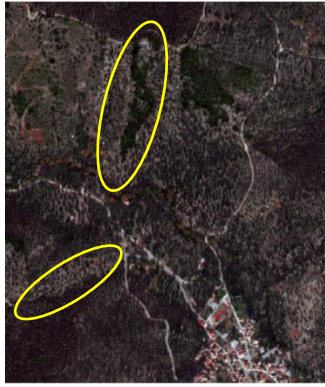


Εικόνα 10. Αποτέλεσμα ταξινόμησης MS εικόνας

Παρατηρώντας τις παραπάνω εικόνες και συγκεκριμένα τις περιοχές που οριοθετούνται μέσα στις ελλείψεις διαπιστώνουμε ότι υπάρχει ασυμφωνία μεταξύ των δύο ταξινομήσεων σε σημαντικό βαθμό. Στην αριστερή εικόνα η κάλυψη της περιοχής από την κλάση “γυμνό έδαφος” είναι μικρότερη σε σχέση με την αυτήν της δεξιάς εικόνας. Αυτό οφείλεται στο ότι στην RGB εικόνα απουσιάζει το ευρύ φάσμα καναλιών που συναντάμε στην πολυφασματική εικόνα, το οποίο βοηθάει στην καλύτερη ταξινόμηση των εικονοστοιχείων.

Κλάση 2 – Βλάστηση

Στην Εικόνα 11 παρουσιάζεται η βλάστηση που ταξινομήθηκε σωστά από την RGB εικόνα και λανθασμένα από την MS, ενώ στην Εικόνα 12 το αντίστροφο.



Εικόνα 11. Αποτέλεσμα ταξινόμησης RGB
εικόνας

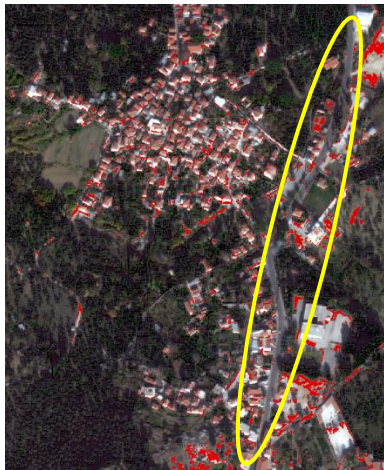


Εικόνα 12. Αποτέλεσμα ταξινόμησης MS
εικόνας

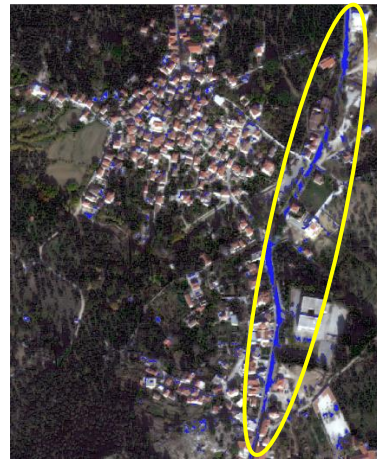
Παρατηρώντας τις παραπάνω εικόνες και συγκεκριμένα τις περιοχές που οριοθετούνται μέσα στις ελλείψεις διαπιστώνουμε ότι υπάρχει και πάλι ασυμφωνία μεταξύ των δύο ταξινομήσεων σε σημαντικό βαθμό. Στην αριστερή εικόνα η κάλυψη της περιοχής από την κλάση “βλάστηση” είναι απειροελάχιστη σε σχέση με την αυτήν της δεξιάς εικόνας. Αυτό οφείλεται στην έλλειψη υπέρυθρου καναλιού στην RGB εικόνα το οποίο βοηθάει στο να εντοπίζεται ευκολότερα η βλάστηση.

Δομημένο Περιβάλλον

Στην Εικόνα 13 παρουσιάζεται το γυμνό έδαφος που ταξινομήθηκε σωστά από την RGB εικόνα και λανθασμένα από την MS, ενώ στην Εικόνα 14 το αντίστροφο.



Εικόνα 13. Αποτέλεσμα ταξινόμησης RGB
εικόνας



Εικόνα 14. Αποτέλεσμα ταξινόμησης MS
εικόνας

Παρατηρώντας τις παραπάνω εικόνες και συγκεκριμένα την περιοχή που οριοθετείται από την έλλειψη διαπιστώνουμε ότι υπάρχει ξανά ασυμφωνία μεταξύ των δυο ταξινομήσεων. Στην αριστερή εικόνα η κάλυψη της περιοχής από την κλάση “δομημένο περιβάλλον” εμφανίζεται σε άλλα εικονοστοιχεία (κυρίως κοντά σε κτήρια) σε σχέση με την αυτήν της δεξιάς εικόνας που κυριαρχεί σε περιοχές οδικού δικτύου. Αυτό οφείλεται στη διαφορετική ένταση των εικονοστοιχείων που λαμβάνεται διαφορετικά από τους φασματικούς δέκτες και αλλιώς από τους πολυφασματικούς.

Συμπεράσματα

Η σημασιολογική κατάτμηση εικόνας στον τομέα της υπολογιστικής όρασης έχει κάνει μεγάλα άλματα προόδου λόγω των τεχνητών νευρικών δικτύων βαθιάς εκμάθησης. Σύμφωνα με τους Huang, Zhao & Song (2018) η χρήση τέτοιων αρχιτεκτονικών επιφέρει υψηλά ποσοστά ακρίβειας σε προβλήματα ταξινόμησης πολυφασματικών εικόνων. Στην περίπτωση της δικής μας μελέτης βέβαια το αποτέλεσμα δεν κρίνεται άκρως επιτυχές και θα πρέπει να εξεταστούν πιθανές βελτιώσεις τόσο στη αρχιτεκτονική δομή του δικτύου μας, όσο και στον αριθμό δεδομένων εκπαίδευσης και επαλήθευσης.

Η χρήση μιας πολυφασματικής εικόνας WorldView-2, χωρικής διακριτικής ικανότητας 0,46μ., για σημασιολογική κατάτμηση μέσω βαθιάς αρχιτεκτονικής νευρωνικών δικτύων, κρίθηκε ικανοποιητική για τη ψηφιακή ερμηνεία της εδαφοκάλυψης της περιοχής μελέτης. Όπως αναφέρουν οι Helber et al. (2015), η απόδοση των μοντέλων ταξινόμησης εξαρτάται σε μεγάλο βαθμό από τη διαθεσιμότητα υψηλής ποιότητας δεδομένων εκπαίδευσης. Στην περίπτωση μας, οι διορθώσεις που έγιναν στην εικόνα κατά το στάδιο της προ-επεξεργασίας (σύζευξη καναλιών, γραμμικό τέντωμα) συνέβαλλαν στην απόδοση των ταξινομήσεων, των οποίων η ακρίβεια κυμαίνεται γύρω στο 75%.

Από τους πίνακες ταξινόμησης προκύπτει ότι το αποτέλεσμα των ταξινομήσεων είναι μέτριο προς ικανοποιητικό. Από τη συγκριτική αξιολόγηση των κλάσεων μεμονωμένα ανά εικόνα διαπιστώσαμε ότι υπάρχει ασυμφωνία μεταξύ των δύο ταξινομήσεων σε όλες τις προς μελέτη κατηγορίες. Η πολυφασματική εικόνα απέδωσε διαφορετικό ποσοστό κάλυψης εικονοστοιχείων ανά κλάση σε σχέση με την RGB εικόνα. Αυτό οφείλεται στις παρεμφερείς εντάσεις των εικονοστοιχείων, ειδικά σε σημεία που υπάρχουν σκιές, που είχε ως αποτέλεσμα τα δίκτυα να μπερδεύουν τις κλάσεις. Αν και από την οπτικοποίηση φαίνεται ότι η ταξινόμηση της πολυφασματικής εικόνας απέδωσε πιο πιστά την πραγματικότητα, είχε άλλωστε και το δίκτυο καλύτερη απόδοση ακρίβειας (83,01% έναντι 76,6% της RGB εικόνας), οι πίνακες ταξινόμησης μαρτυρούν ότι κάτι τέτοιο δεν ισχύει.

Όπως αναφέρθηκε εξ αρχής, το δίκτυο που δημιουργήσαμε επιδέχεται βελτιώσεις. Ο πιο αποτελεσματικός τρόπος να βελτιωθεί η απόδοση ενός τεχνητού νευρωνικού δικτύου είναι η αύξηση του βάθους του, δηλαδή ο αριθμός των επιπέδων του, όπως επίσης το πλάτος του, δηλαδή ο αριθμός των μονάδων του. Ωστόσο, η βιβλιογραφία αναφέρει ότι μεγαλύτερο μέγεθος δικτύου συνεπάγεται μεγαλύτερος αριθμός παραμέτρων, το οποίο όμως κάνει το δίκτυο πιο επιρρεπές στην υπερπροσαρμογή (overfitting), ειδικά εάν το σύνολο των δεδομένων εκπαίδευσης είναι μικρού μεγέθους (Szegedy, 2015). Επομένως, αν θελήσουμε να προσθέσουμε επιπλέον επίπεδα θα πρέπει αντίστοιχα να αυξήσουμε το μέγεθος του συνόλου δεδομένων εκπαίδευσης. Έτσι, αντί μιας εικόνας διαστάσεων 1484x1272 ως εικόνα εκπαίδευσης, θα πρέπει να επιλέξουμε ολόκληρη την εικόνα και αυτόματα τα δεδομένα εκπαίδευσης θα αυξηθούν.

Η βιβλιογραφία βέβαια προτείνει να γίνεται διαμοιρασμός των συνόλων των δεδομένων σε 70% δεδομένα εκπαίδευσης, 15% δεδομένα αναφοράς και 15% δεδομένα ελέγχου πάνω στην υπό ταξινόμηση εικόνα και όλα τα σετ να εισάγονται στη διαδικασία εκπαίδευσης του μοντέλου. Στην προκειμένη περίπτωση δεν ακολουθήθηκε αυτή η τακτική. Τα δεδομένα αναφοράς συλλέχθηκαν τυχαία και τα δεδομένα ελέγχου απουσίαζαν από τη διαδικασία. Πιθανόν μια στρωματοποιημένη δειγματοληψία στα δεδομένα αναφοράς σε συνδυασμό με την είσοδο όλων των σετ δεδομένων μέσα στο μοντέλο να είχε καλύτερα αποτελέσματα.

Τέλος, εκτός από την αλλαγή βάθους του δικτύου, μια καλή πρακτική θα ήταν η αλλαγή διαστάσεων στα φίλτρα μετασχηματισμού, όπως και το βήμα τους όταν σαρώνουν την εικόνα, καθώς και τον αλγόριθμο εκπαίδευσης, τον αριθμό των εποχών, κ.λπ. Καθώς η

διαδικασία της εκπαίδευσης είναι μια διαδικασία trial and error, κρίνεται σκόπιμο να γίνουν περαιτέρω δοκιμές στο δίκτυο έτσι ώστε να λάβουμε το βέλτιστο επιθυμητό αποτέλεσμα.

Βιβλιογραφία

- Basaeed E., Bhaskar H. & Al-Mualla M., (2016). Supervised Remote Sensing Image Segmentation Using Boosted Convolutional Neural Networks. *Knowledge-Based Systems* 99, 19-27.
- Ben Hamida A., Benoit A., Lambert P., Klein L., Ben Amar C., Audebert N. & Lefevre S., (2017). Deep Learning for Semantic Segmentation of Remote Sensing Images with Rich Spectral Content. *IEEE International Geoscience and Remote Sensing Symposium*, Jul 2017, Fort Worth, United States.
- Garcia- Garcia A., Orts-Escolano S., Oprea S.O., Villena-Martinez V. & Garcia-Rodriguez J., (2017). A Review on Deep Learning Techniques Applied to Semantic Segmentation. COMBAHO Project.
- Helber P., Bischke B., Dengel A. & Borth D., (2015). EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. *Journal of Latex Class Files*, Vol. 14, No. 8, August 2015.
- Huang B., Zhao B. & Song Y., (2018). Urban Land-Use Mapping Using A Deep Convolutional Neural Network with Spatial Resolution Multispectral Remote Sensing Imagery. *Remote Sensing Environment* 214 (2018), 73-86.
- Kemker R., Salvaggio C. & Kanan C., (2017). High-Resolution Multispectral Dataset for Semantic Segmentation. *ArcXiv – Cornell University: 1703.01918v1*.
- Krizhevsky A., Sutskever I. & Hinton G.E., (2012). "Imagenet Classification with Deep Convolutional Neural Networks". *Advances in Neural Information Processing Systems*, 2012, 1097-1105.
- LeCun Y., Bengio, Y & Hinton G., (2015). Deep Learning. *Nature*, 521, 436-44.
- Li H., Jing L. and Tang Y., 2017: Assessment of Pansharpening Methods Applied to WorldView-2 Imagery Fusion. *Sensors* 2017, 17, 89.
- Liebel L., (2018). Deep Convolutional Neural Networks for Semantic Segmentation of Multispectral Sentinel-2 Satellite Imagery: An Open Data Approach to Large-Scale Land Use and Land Cover Classification. *Wissenschaftlich-Technische Jahrestagung der DGPF und PFGK18 Tagung in München – Publikationen der DGPF, Band 27*.
- Long J., Shelhamer E. & Darrell T., (2015). Fully Convolutional Networks for Semantic Segmentation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 3431-3440.
- López J., Santos S. & Atzberger C., (2018). Convolutional Neural Networks for Semantic Segmentation of Multispectral Remote Sensing Images. *IEEE 10th Latin-American Conference on Communications (LATINCOM)*, Guadalajara, 2018, pp. 1-5.
- Papadomanolaki M., Vakalopoulou M., Zagoruyko S. & Karantza K., (2016). Benchmarking Deep Learning Frameworks for the Classification of Very High-Resolution Satellite Multispectral Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. III-7, 2016. XXIII ISPRS Congress, 12-19 July 2016, Prague, Czech Republic.
- Pirrotti F., Sunar F. & Piragnolo M., (2016). Benchmark of Machine Learning Methods for Classification of a Sentinel-2 Image. *ISPRS Annals of the Photogrammetry, Remote*

Sensing and Spatial Information Sciences, Vol. III-7, 2016. XXIII ISPRS Congress, 12-19 July 2016, Prague, Czech Republic.

Simonyan K. & Zisserman A., (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. ICLR 2015

Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A., (2015). Going Deeper with Convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 2015, pp. 1-9.

Multispectral classification of WorldView-2 image based on semantic segmentation by applying deep learning convolutional neural networks

Image classification is one of the fundamental remote sensing applications aiming at land use/land cover extraction and includes supervised and unsupervised techniques. The objective of this paper is the classification of a multispectral image, based on semantic segmentation by applying deep learning artificial neural network. The aim is to create a land-cover map of the study area at the southeast part of Lesvos Island, Greece and the validation of the outputs. The methodology was based on a combination of remote sensing and computer vision techniques on a WorldView-2 (WV-2) image. The first step of the process included an image fusion in order to sharpen multispectral channels with high-resolution (0.46m) panchromatic channel. Two deep learning convolutional neural networks (deep CNN) were created and trained, one network based on eight multispectral bands of the initial image and one network based on the three bands of the visible spectrum (RGB). The trained networks were applied at the initial images for the semantic segmentation and the creation of the classified images. A set of three hundred random reference points was collected for validation of the segmentation outputs. The classification products produced by this method were the thematic maps of the study area, accuracy of performance and loss of minibatches by network training, histograms and pixel distribution tables, confusion matrices about evaluating aggregate classifications and comparative assessment images of classes per image. The main conclusion of this study is that the architectural structure of the network, as well as the number of training and verification data affect the classification accuracy.

KEY-WORDS: Semantic segmentation, deep learning techniques, convolutional neural network, classification, WordView-2