



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Αγρονόμων και Τοπογράφων Μηχανικών –
Μηχανικών Γεωπληροφορικής

Εργαστήριο Τηλεπισκόπησης

ΔΜΠΣ “ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ”

Μεταπτυχιακή Εργασία:

**Συγκριτική αξιολόγηση αντικειμενοστρεφών μεθόδων
χαρτογράφησης καμένων εκτάσεων και εκτίμησης του
βαθμού καταστροφής τους**

Επιβλέπων καθηγητής: Δημήτριος Αργιαλάς

Αικατερίνη Μαρία Φωτίου

Αθήνα, Φεβρουάριος 2022

| Ευχαριστίες |

Θα ήθελα να ευχαριστήσω τον καθηγητή κ. Δημήτριο Αργιαλά για την συνεχή και άμεση καθοδήγηση και βοήθεια που μου παρείχε κατά τη διάρκεια εκπόνησης αυτής της εργασίας. Επίσης, θα ήθελα να ευχαριστήσω τον κ. Πολυχρόνη Κολοκούση για τις καίριες και καθοριστικές συμβουλές που μου έδινε καθ' όλη τη διάρκεια. Τέλος, ευχαριστώ την οικογένεια μου αλλά και τις φίλες και φίλους μου που με στηρίζουν.

Π Ε Ρ Ι Λ Η Ψ Η

Το φαινόμενο των δασικών πυρκαγιών τα τελευταία χρόνια εντείνεται όλο και περισσότερο, ειδικά στις χώρες της Μεσογείου. Σε αυτό το πλαίσιο, είναι αναγκαία η έγκαιρη και αξιόπιστη αποτύπωση της κατάστασης μετά από μία πυρκαγιά προκειμένου οι αρμόδιοι φορείς να μπορούν να παρέμβουν αποτελεσματικά. Προς αυτή την κατεύθυνση, στην παρούσα εργασία αναπτύσσονται και συγκρίνονται δύο διαφορετικές αντικειμενοστρεφείς μεθοδολογίες τηλεπισκόπησης για την χαρτογράφηση καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους. Το γεγονός που μελετάται είναι η καταστροφική πυρκαγιά που ξέσπασε στο Μάτι τον Ιούλιο 2018. Τα δορυφορικά δεδομένα που χρησιμοποιήθηκαν προέρχονται από τον δορυφόρο Sentinel-2 και αφορούν σε μία λήψη πριν το γεγονός της φωτιάς (pre-fire) και μία λήψη μετά (post-fire), προκειμένου να σχηματιστούν κατάλληλοι φασματικοί δείκτες. Τα δεδομένα εισόδου και τα δεδομένα αληθείας είναι τα ίδια και για τις δύο μεθοδολογίες ώστε να είναι άμεσα συγκρίσιμες. Τα δεδομένα αληθείας προέρχονται από την υπηρεσία Copernicus Emergency Management Service (CEMS) και έχουν παραχθεί με ημι-αυτόματη μέθοδο αξιοποιώντας δορυφορικές εικόνες πολύ υψηλής ανάλυσης. Βάσει των δεδομένων αυτών ορίστηκαν και οι κλάσεις ταξινόμησης. Η πρώτη μεθοδολογία αντικειμενοστρεφούς ταξινόμησης αναπτύσσεται στη γλώσσα Python και ως μέθοδος ταξινόμησης εφαρμόζεται η επιβλεπόμενη και συγκεκριμένα ο αλγόριθμος Random Forest. Η δεύτερη μεθοδολογία υλοποιείται στο περιβάλλον του λογισμικού eCognition και αξιοποιεί τις συναρτήσεις συμμετοχής της ασαφούς λογικής για την ταξινόμηση των αντικειμένων. Από τη συνολική ποιοτική και ποσοτική αξιολόγηση των μεθοδολογιών, προκύπτει ότι και οι δύο καταφέρνουν να χαρτογραφήσουν με μεγάλη ακρίβεια τις καμένες εκτάσεις (User's Accuracy $\geq 85\%$ και Producer's Accuracy ≥ 0.87) παρόλο την αρκετά χαμηλότερη χωρική ανάλυση των δορυφορικών δεδομένων που αυτές εφαρμόστηκαν συγκριτικά με αυτά από τα οποία έχουν παραχθεί τα δεδομένα CEMS. Σχετικά με τα αποτελέσματα χαρτογράφησης του βαθμού καταστροφής τους, η ακρίβεια είναι μειωμένη και οι δύο μέθοδοι εμφανίζουν την ελάχιστη ακρίβεια στο μέτρο αξιολόγησης User's Accuracy (60%).

Λέξεις - κλειδιά: δασικές πυρκαγιές, χαρτογράφηση καμένων εκτάσεων, εκτίμηση βαθμού καταστροφής, OBIA, Random Forest, eCognition, Ασαφής Λογική

«Comparative assessment of object-based methodologies for burned area and burn severity mapping»

Aikaterini Maria Fotiou

Master Thesis

February 2022

A b s t r a c t

The phenomenon of forest fires has intensified in recent years, especially in the Mediterranean countries. In this context, the timely and reliable mapping of the post-fire situation is necessary in order for the competent bodies to be able to intervene effectively. To this end, in the present work, two different object-based remote sensing methodologies were developed and compared for the mapping of burned areas and burn severity mapping. The case study is the catastrophic fire that broke out in Mati in July 2018. The satellite data used were two Sentinel-2 acquisitions; one pre-fire acquisition and one post-fire in order to compute appropriate spectral indices. The input data and the ground truth are the same for both methodologies so that they could be comparable. The reference data was acquired through the Copernicus Emergence Management Service (CEMS) and its collection method is semi-automatic utilizing very high-resolution satellite images. Based on these data, the classification classes were defined. The first object-oriented classification methodology is developed in Python and as classification method was applied the Random Forest algorithm. The second methodology was implemented in the environment of eCognition software and utilized membership functions to classify objects. From the overall qualitative and quantitative evaluation of the methodologies, it appeared that both managed to map the burned areas with great accuracy (User's Accuracy > = 85% and Producer's Accuracy > = 0.87) despite the much lower spatial analysis of the satellite data applied compared to those from which the CEMS data were generated. Regarding the mapping results of their degree of destruction, the accuracy was reduced and both methods showed the minimum accuracy in the User's Accuracy rating measure (60%).

Key-words: wildfire, burned area mapping, burn severity mapping, OBIA, Random Forest, eCognition, Fuzzy Logic

Π ε ρ ι ε χ ό μ ε ν α

1 Εισαγωγή	11
1.1 Αντικείμενο και Στόχοι	11
1.2 Δομή Εργασίας	11
2 Θεωρητικό υπόβαθρο & βιβλιογραφική ανασκόπηση	13
2.1 Αντικειμενοστρεφής ανάλυση εικόνας.....	13
2.2 Χαρτογράφηση καμένων εκτάσεων και εκτίμηση βαθμού καταστροφής τους με μεθόδους Τηλεπισκόπησης	14
2.3 Βιβλιογραφική ανασκόπηση	16
3 Δεδομένα και προ-επεξεργασίες.....	19
3.1 Δορυφορικά δεδομένα	19
3.2 Προ-επεξεργασία δορυφορικών εικόνων.....	20
3.3 Πηγή και μέθοδος συλλογής δεδομένων αληθείας (ground truth)	23
3.4 Ποιοτικός έλεγχος δεδομένων αληθείας (ground truth).....	27
4 Μεθοδολογία	33
4.1 Ανάπτυξη αλγορίθμου αντικειμενοστρεφούς ταξινόμησης δορυφορικών εικόνων σε γλώσσα Python.....	33
4.1.1 Κατάτμηση εικόνας	35
4.1.2 Ταξινόμηση εικόνας	39
4.1.3 Αξιολόγηση ταξινόμησης	47
4.2 Αντικειμενοστρεφής ταξινόμηση δορυφορικών εικόνων με αξιοποίηση του εμπορικού λογισμικού eCognition	51
4.2.1 Κατάτμηση 1 ^{ου} επιπέδου (Level 1).....	53
4.2.2 Ταξινόμηση αντικειμένων 1 ^{ου} επιπέδου (Level 1)	55
4.2.3 Κατάτμηση 2 ^{ου} επιπέδου (Level 2)	57
4.2.4 Ταξινόμηση αντικειμένων 2 ^{ου} επιπέδου (Level 2)	58
5 Αποτελέσματα & αξιολόγηση	60
5.1 Αποτελέσματα αλγορίθμου αντικειμενοστρεφούς ταξινόμησης δορυφορικών εικόνων σε γλώσσα Python.....	60
5.1.1 Αποτελέσματα σε επίπεδο εικονοστοιχείου (pixel-based).....	67
5.1.2 Αποτελέσματα σε επίπεδο αντικειμένου (segment-based)	68
5.2 Αποτελέσματα αντικειμενοστρεφούς ταξινόμησης δορυφορικών εικόνων στο λογισμικό eCognition	69

6 Συμπεράσματα & προτάσεις.....	77
6.1 Συμπεράσματα	77
6.2 Προτάσεις.....	79
Βιβλιογραφία.....	80

Κατάλογος πινάκων

Πίνακας 1: Φασματικοί δείκτες που δημιουργήθηκαν στα πλαίσια της εργασίας.....	22
Πίνακας 2: Κανάλια του τελικού κύβου δεδομένων προς ταξινόμηση.....	22
Πίνακας 3: Ενδεικτικές περιπτώσεις τροποποίησης των δεδομένων αληθείας κατά τον ποιοτικό έλεγχό τους.....	28
Πίνακας 4: Ενδεικτικές περιπτώσεις ελέγχου των δεδομένων αληθείας της κλάσης Not Burned βάσει φασματικών χαρακτηριστικών.....	29
Πίνακας 5: Ενδεικτική περίπτωση αποκλεισμού σημειακών δεδομένων αληθείας	30
Πίνακας 6: Ενδεικτικές περιπτώσεις ελέγχου των δεδομένων αληθείας για τις κλάσεις Damaged και Destroyed βάσει φασματικών χαρακτηριστικών.	30
Πίνακας 7: Ενδεικτικές περιπτώσεις σύγχυσης μεταξύ φασματικών χαρακτηριστικών των κλάσεων Damaged και Destroyed.....	32
Πίνακας 8: Ενδεικτικές περιπτώσεις ορθής δημιουργίας αντικειμένων με τον αλγόριθμο SLIC.	37
Πίνακας 9: Ενδεικτικές περιπτώσεις αστοχιών κατά την κατάτμηση με τον αλγόριθμο SLIC.	39
Πίνακας 10: Βάρη φασματικών καναλιών στο 1 ^ο επίπεδο κατάτμησης (Level 1).....	53
Πίνακας 11: Παράμετροι κατάτμησης για το επίπεδο 1 (Level 1).	54
Πίνακας 12: Ενδεικτικά αποσπάσματα από τα αποτελέσματα της κατάτμησης 1 ^{ου} επιπέδου (Level 1).	55
Πίνακας 13: Κανόνες ασαφούς ταξινόμησης αντικειμένων του επιπέδου 1 (Level 1).	56
Πίνακας 14: Βάρη φασματικών καναλιών στο 2 ^ο επίπεδο κατάτμησης (Level 2).....	57
Πίνακας 15: Παράμετροι κατάτμησης για το επίπεδο 2 (Level 2).	57
Πίνακας 16: Ενδεικτικά αποσπάσματα από τα αποτελέσματα της κατάτμησης 2ου επιπέδου (Level 2).	58
Πίνακας 17: Κανόνες ασαφούς ταξινόμησης αντικειμένων του επιπέδου 2 (Level 2).	59
Πίνακας 18: Έκταση και ποσοστά επικάλυψης της κλάσης Not Burned (Python) με τα δεδομένα αναφοράς CEMS.....	64
Πίνακας 19: Έκταση και ποσοστά επικάλυψης της κλάσης Damaged (Python) με τα δεδομένα αναφοράς CEMS.....	64
Πίνακας 20: Έκταση και ποσοστά επικάλυψης της κλάσης Destroyed (Python) με τα δεδομένα αναφοράς CEMS.....	65
Πίνακας 21: Έλεγχος επικάλυψης ταξινομημένης κλάσης Destroyed (Python) με τα δεδομένα αναφοράς για την κλάση Destroyed.....	66
Πίνακας 22: Έκταση και ποσοστά επικάλυψης αθροιστικά για τις κλάσεις Damaged και Destroyed (Python) με τα δεδομένα αναφοράς CEMS.....	66

Πίνακας 23: Πίνακας σύγκρισης σε επίπεδο εικονοστοιχείων των αποτελεσμάτων ταξινόμησης σε περιβάλλον Python.	67
Πίνακας 24: Πίνακας Omission και Commission errors σε επίπεδο εικονοστοιχείων για τα αποτελέσματα ταξινόμησης σε περιβάλλον Python.	67
Πίνακας 25: Πίνακας σύγκρισης σε επίπεδο αντικειμένων για τα αποτελέσματα ταξινόμησης σε περιβάλλον Python.....	68
Πίνακας 26: Πίνακας Omission και Commission errors σε επίπεδο αντικείμενων για τα αποτελέσματα ταξινόμησης σε περιβάλλον Python.	68
Πίνακας 27: Πίνακας σύγκρισης σε επίπεδο εικονοστοιχείων των αποτελεσμάτων ταξινόμησης στο περιβάλλον eCognition.	73
Πίνακας 28: Πίνακας Omission και Commission errors σε επίπεδο εικονοστοιχείων για τα αποτελέσματα ταξινόμησης στο περιβάλλον eCognition.	74
Πίνακας 29: Έκταση και ποσοστά επικάλυψης της κλάσης Not Burned (eCognition) με τα δεδομένα αναφοράς CEMS.....	74
Πίνακας 30: Έκταση και ποσοστά επικάλυψης της κλάσης Damaged (eCognition) με τα δεδομένα αναφοράς CEMS.....	74
Πίνακας 31: Έκταση και ποσοστά επικάλυψης της κλάσης Destroyed (eCognition) με τα δεδομένα αναφοράς CEMS.....	75
Πίνακας 32: Έκταση και ποσοστά επικάλυψης αθροιστικά για τις κλάσεις Damaged και Destroyed (eCognition) με τα δεδομένα αναφοράς CEMS.....	75

Κ α τ ά λ ο γ ο ς ε ι κ ό ν ω ν

Εικόνα 1: Σχέση αντικειμένων και εικονοστοιχείων λαμβάνοντας υπόψη τη χωρική ανάλυση των εικόνων: (α) χαμηλή χωρική ανάλυση-εικονοστοιχεία πολύ μεγαλύτερα από αντικείμενα, (b) μέση χωρική ανάλυση-περίπου κοινό μέγεθος αντικειμένων και εικονοστοιχείων, (c) υψηλή χωρική ανάλυση: εικονοστοιχεία πολύ μικρότερα από αντικείμενα, πηγή: (Blaschke, 2010)	13
Εικόνα 2: Βαθμός καταστροφής καμένων εκτάσεων βάσει του δείκτη dNBR, όπως προτείνεται από το USGS.	15
Εικόνα 3: Φασματικά κανάλια του δορυφόρου Sentinel-2 και αντίστοιχη χωρική ανάλυση (πηγή: https://www.researchgate.net/figure/Sentinel-2-band-characteristics_tbl1_314119510).	19
Εικόνα 4: Δορυφορική εικόνα Sentinel-2 (Level-2A) μετά το γεγονός της πυρκαγιάς στο Μάτι.....	20
Εικόνα 5: Διάγραμμα ροής προ-επεξεργασιών για τη δημιουργία του τελικού σετ δεδομένων προς ταξινόμηση.	23
Εικόνα 6: Περίμετρος πυρκαγιάς και θεματικό επίπεδο “built-up”. Επισήμανση ανακρίβειας των δεδομένων “built-up”.	25
Εικόνα 7: Επεξήγηση των κλάσεων των δεδομένων Grading map από το CEMS που χρησιμοποιήθηκαν ως δεδομένα αναφοράς.....	26
Εικόνα 8: Διάγραμμα ροής αλγορίθμου στη γλώσσα Python.	34

Εικόνα 9: Σύγκριση σημειακών δεδομένων εκπαίδευσης ανά κλάση (αριστερά) και εικονοστοιχείων εκπαίδευσης (δεξιά) μετά την κανονικοποίηση του διανυσματικού αρχείου.....	42
Εικόνα 10: Μεταβλητή segments_per_class στην οποία αποθηκεύονται τα id των αντικειμένων εκπαίδευσης ανά κλάση.....	43
Εικόνα 11: Εννοιολογικό διάγραμμα ροής αλγορίθμου Multiresolution Segmentation, πηγή: (eCognition, 2014).....	53
Εικόνα 12: Αποτέλεσμα ταξινόμησης βάσει του αλγορίθμου στη γλώσσα Python (αριστερά) και δεδομένα αναφοράς από το CEMS (δεξιά).....	61
Εικόνα 13: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στο μαύρη έλλειψη επισημαίνεται η περιοχή ορθής ταξινόμησης.	61
Εικόνα 14: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης (πάνω) συγκριτικά με τα δεδομένα αναφοράς CEMS (κάτω). Στο μαύρη έλλειψη επισημαίνεται η περιοχή ορθής ταξινόμησης.	62
Εικόνα 15: Ταξινομημένη εικόνα (αριστερά) και εικόνα του δείκτη dNBR (δεξιά). Στα μωβ πλαίσια επισημαίνονται τα λανθασμένα ταξινομημένα σημεία ελέγχου.	62
Εικόνα 16: Στιγμιότυπα ταξινομημένης εικόνα με τα αντίστοιχα από την απεικόνιση των δορυφορικών δεδομένων σε έγχρωμο σύνθετο NIR-R-G. Στα μωβ πλαίσια επισημαίνονται τα λανθασμένα ταξινομημένα σημεία ελέγχου.....	63
Εικόνα 17: Αποτέλεσμα ταξινόμησης στο λογισμικό eCognition (αριστερά) και δεδομένα αναφοράς από το CEMS (δεξιά).....	69
Εικόνα 18: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στις μαύρες ελλείψεις επισημαίνονται οι περιοχές ορθής ταξινόμησης.....	70
Εικόνα 19: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά).....	70
Εικόνα 20: Ενδεικτική περίπτωση αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στο μωβ πλαίσιο επισημαίνονται σφάλματα ταξινόμησης.	71
Εικόνα 21: Ενδεικτική περίπτωση αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στο μωβ πλαίσιο επισημαίνονται σφάλματα ταξινόμησης.	71
Εικόνα 22: Σταθερότητα κλάσης και καλύτερο αποτέλεσμα ταξινόμησης αντικειμένων επιπέδου 1.....	72
Εικόνα 23: Σταθερότητα κλάσης και καλύτερο αποτέλεσμα ταξινόμησης αντικειμένων επιπέδου 2.....	72
Εικόνα 24: Πίνακες σύγχυσης σε επίπεδο εικονοστοιχείων για τη μεθοδολογία σε προγραμματιστικό περιβάλλον Python (πάνω) και στο περιβάλλον eCognition (κάτω).....	77

Στο πρώτο και εισαγωγικό κεφάλαιο, αρχικά γίνεται μία σύντομη περιγραφή του αντικείμενου της εργασίας και των επιμέρους στόχων της και τέλος παρουσιάζεται η δομή που ακολουθείται στο παρόν τεύχος.

1.1 Αντικείμενο και Στόχοι

Βασικό αντικείμενο της παρούσας μεταπτυχιακής εργασίας αποτελεί η ανάπτυξη και αξιολόγηση δύο διαφορετικών αντικειμενοστρεφών μεθόδων για τη χαρτογράφηση καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους. Ειδικότερα, αξιοποιήθηκαν δορυφορικά δεδομένα και άλλα βοηθητικά δεδομένα προκειμένου να αναπτυχθούν δύο πλήρεις μεθοδολογίες κατάτμησης, ταξινόμησης και αξιολόγησης. Τα δεδομένα είναι κοινά και για τις δύο ενώ τα εργαλεία υλοποίησης είναι διαφορετικά. Η πρώτη προσέγγιση υλοποιείται σε ανοιχτό προγραμματιστικό περιβάλλον Python και ως μέθοδο ταξινόμησης χρησιμοποιείται ο αλγόριθμος Random Forest ενώ η δεύτερη στο περιβάλλον του κλειστού λογισμικού eCognition 9 και βασίζεται στην ταξινόμηση με ασαφή λογική.

Κύριος στόχος είναι η διερεύνηση της καταλληλότητας των δύο διαφορετικών μεθοδολογιών και ο εντοπισμός τυχόν πλεονεκτημάτων και αδυναμιών. Παράλληλα με τον βασικό στόχο όμως τέθηκαν και κάποιοι επιμέρους. Αρχικά, μέσω της παρούσας εργασίας επιδιώκεται να διερευνηθεί η δυνατότητα παραγωγής έγκυρων αποτελεσμάτων χαρτογράφησης καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους με την αξιοποίηση αποκλειστικά ελεύθερα διαθέσιμων δεδομένων. Επιπλέον, επιδιώκεται η διερεύνηση των αποτελεσμάτων ταξινόμησης στην περίπτωση δημιουργίας δεδομένων αληθείας με ημι-αυτόματο τρόπο προερχόμενα από αξιόπιστα δεδομένα αναφοράς. Τέλος, κατά την ανάπτυξη των επιμέρους βημάτων των δύο μεθόδων κοινό στόχο αποτέλεσε η όσο το δυνατόν τυποποίηση των διαδικασιών προκειμένου να μπορούν εύκολα να αξιοποιηθούν και σε άλλες εφαρμογές. Ιδιαίτερα κατά την εφαρμογή αλγορίθμου στη γλώσσα Python επιδιώκεται και η εμφάνιση κατάλληλων μηνυμάτων προς τον χρήστη ώστε να έχει πλήρη εποπτεία της διαδικασίας.

1.2 Δομή Εργασίας

Η εργασία δομείται σε έξι (6) κεφάλαια συμπεριλαμβανομένου και του παρόντος και ακολουθεί τη ροή εργασιών που υλοποιήθηκαν σε πρακτικό επίπεδο για την ολοκλήρωσή της, χωρίς αυτό να σημαίνει ότι κατά την υλοποίησή της δεν υπήρχαν στοιχεία ανάδρασης.

Στο παρόν και πρώτο κεφάλαιο, γίνεται μία εισαγωγή στην εργασία περιγράφοντας το αντικείμενο, το βασικό στόχο αλλά και τους επιμέρους που τέθηκαν.

Στο δεύτερο κεφάλαιο, ορίζεται το θεωρητικό υπόβαθρο της αντικειμενοστρεφούς ανάλυσης εικόνας και της χαρτογράφησης καμένων εκτάσεων με τηλεπισκοπικές μεθόδους. Τέλος, γίνεται ανασκόπηση της βιβλιογραφίας που σχετίζονται με αυτές αλλά και γενικότερα με τη χαρτογράφηση καμένων εκτάσεων και εκτίμηση του βαθμού καταστροφής τους.

Στο κεφάλαιο 3, περιγράφονται αναλυτικά τα δορυφορικά δεδομένα που χρησιμοποιούνται, καθώς και τα δεδομένα αναφοράς από τα οποία γίνεται η συλλογή του ground truth. Επίσης, παρατίθενται αναλυτικά οι προ-επεξεργασίες που εφαρμόζονται σε αυτά.

Στο τέταρτο κεφάλαιο, παρουσιάζονται αναλυτικά οι δύο διαφορετικές αντικειμενοστρεφείς μέθοδοι για τη χαρτογράφηση καμένων εκτάσεων και κατηγοριοποίησης της καταστροφής τους.

Στο πέμπτο κεφάλαιο, παρουσιάζονται, αναλύονται και αξιολογούνται τα αποτελέσματα των δύο αντικειμενοστρεφών μεθόδων που εφαρμόστηκαν. Παρατίθενται εκτενής ποιοτική αλλά και ποσοτική αξιολόγηση.

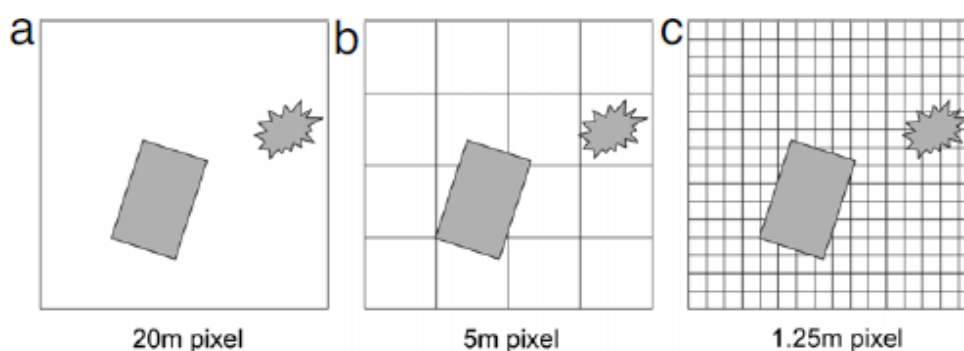
Στο έκτο και τελευταίο κεφάλαιο, βάσει των αποτελεσμάτων αλλά και της γνώσης που αποκτήθηκε από την υλοποίηση της εργασίας, εξάγονται κάποια συμπεράσματα σε σχέση με τους στόχους που έχουν τεθεί. Τέλος, γίνονται κάποιες προτάσεις για βελτίωση των αποτελεσμάτων και περαιτέρω εξέλιξη του αντικειμένου της εργασίας.

2 | Θεωρητικό υπόβαθρο & βιβλιογραφική ανασκόπηση

Στο παρόν κεφάλαιο παρουσιάζεται η λογική της αντικειμενοστρεφούς ανάλυσης εικόνας και οι τάσεις στην αξιοποίηση τηλεπισκοπικών δεδομένων και τεχνικών τόσο για την χαρτογράφηση καμένων εκτάσεων όσο και για την εκτίμηση του βαθμού καταστροφής τους. Στο πλαίσιο αυτό γίνεται επίσης ανασκόπηση στη βιβλιογραφία.

2.1 Αντικειμενοστρεφής ανάλυση εικόνας

Οι δύο μεθοδολογίες ταξινόμησης που αναπτύσσονται στην παρούσα εργασία βασίζονται στην Αντικειμενοστρεφή ανάλυση εικόνας (Object-based Image Analysis-OBIA). Σύμφωνα με αυτή τη μέθοδο το μοναδιαίο στοιχείο της εικόνας δεν είναι το μεμονωμένο εικονοστοιχείο αλλά το αντικείμενο, δηλαδή μία ομάδα εικονοστοιχείων. Τα αντικείμενα της εικόνας αποτελούν περιοχές που δημιουργούνται από ένα ή περισσότερα κριτήρια ομοιογένειας σε μία ή περισσότερες διαστάσεις. Έτσι, κάθε αντικείμενο χαρακτηρίζεται από μέση τιμή ανά κανάλι, μέγιστη και ελάχιστη τιμή, τυπική απόκλιση κ.α. περιέχοντας περισσότερες πληροφορίες από ότι τα μεμονωμένα εικονοστοιχεία (Blaschke, 2010). Πέραν της φασματικής πληροφορίας, ιδιαίτερα χρήσιμο είναι ότι τα αντικείμενα ενέχουν και χωρικές πληροφορίες όπως για παράδειγμα αποστάσεις από άλλα αντικείμενα, τοπολογικές σχέσεις κ.α. Τις δύο τελευταίες δεκαετίες, με τη μεγάλη διαθεσιμότητα δορυφορικών δεδομένων με μέτρια και υψηλή χωρική ανάλυση, η αντικειμενοστρεφής ανάλυση έχει αποκτήσει ευρεία αξιοποίηση, λαμβάνοντας υπόψη ότι όσο χαμηλότερη είναι η χωρική ανάλυση τόσο περισσότερα αντικείμενα του πραγματικού κόσμου απεικονίζονται σε ένα εικονοστοιχείο και άρα η συνένωση τους για την δημιουργία αντικείμενων χάνει την αξία της.



Εικόνα 1: Σχέση αντικειμένων και εικονοστοιχείων λαμβάνοντας υπόψη τη χωρική ανάλυση των εικόνων: (a) χαμηλή χωρική ανάλυση-εικονοστοιχεία πολύ μεγαλύτερα από αντικείμενα, (b) μέση χωρική ανάλυση-περίπου κοινό μέγεθος αντικειμένων και εικονοστοιχείων, (c) υψηλή χωρική ανάλυση: εικονοστοιχεία πολύ μικρότερα από αντικείμενα, πηγή: (Blaschke, 2010)

Σύμφωνα με τον (Blaschke, 2010) η αντικειμενοστρεφής ανάλυση εικόνας χρησιμοποιείται σε ένα ευρύ φάσμα περιβαλλοντικών εφαρμογών όπως: χαρτογράφηση κάλυψης/χρήσης γης, χαρτογράφηση δασών, βλάστησης και αστικής δόμησης, ανίχνευση αλλαγών στην κάλυψη γης καθώς και χαρτογράφηση καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους.

2.2 Χαρτογράφηση καμένων εκτάσεων και εκτίμηση βαθμού καταστροφής τους με μεθόδους Τηλεπισκόπησης

Η χαρτογράφηση καμένων εκτάσεων με μεθόδους Τηλεπισκόπησης εμπίπτει σε δύο ευρείες ερευνητικές περιοχές: η πρώτη είναι η χαρτογράφηση της καμένης έκτασης (Burned Area Mapping) και η δεύτερη η χαρτογράφηση του βαθμού καταστροφής (Burn Severity Mapping). Η Τηλεπισκόπηση παρέχει κατάλληλα δεδομένα και μεθόδους για τέτοιου είδους εφαρμογές αφού η πυρκαγιά οδηγεί σε ουσιαστικές φασματικές αλλαγές δεδομένου ότι περιορίζει τη βλάστηση εκθέτοντας το έδαφος, καταστρέφει τη χλωροφύλλη των φύλλων και αλλάζει την υγρασία πάνω και κάτω από το έδαφος (Mallinis & Koutsias, 2012). Παρατηρείται μείωση της ανάκλασης στην περιοχή του εγγύς υπέρυθρου (NIR) του φάσματος λόγω της καταστροφής των κυττάρων των φύλλων μέσα στη βλάστηση.

Πολλά προϊόντα και μεθοδολογίες έχουν αναπτυχθεί σε παγκόσμιο επίπεδο, τα οποία βασίζονται κυρίως σε δορυφορικά δεδομένα MODIS που όμως λόγω της χαμηλής χωρικής ανάλυσης δεν μπορούν να διαχειριστούν το ίδιο αξιόπιστα πυρκαγιές με σχετικά μικρή έκταση. Η ευρεία διαθεσιμότητα δορυφορικών δεδομένων με σημαντικά υψηλότερη χωρική ανάλυση (Landsat 8-2013, Sentinel 2-2015) έχει οδηγήσει σε υψηλότερα επίπεδα ακρίβειας της χαρτογράφησης καμένων εκτάσεων με μεθόδους Τηλεπισκόπησης (Belenguer-Plomer, et al., 2021). Ανάμεσα στις Τηλεπισκοπικές μεθόδους για τέτοιες εφαρμογές συγκαταλέγονται ενδεικτικά οι ταξινομήσεις βάσει κατωφλίωσης σε φασματικά κανάλια και δείκτες, ταξινομήσεις βάσει εικονοστοιχείων (pixel-based), αντικειμενοστρεφείς ταξινομήσεις όπως η παρούσα εργασία καθώς και πιο πρόσφατα ανάπτυξη νευρωνικών δικτύων.

Στο σημείο αυτό, δεδομένου ότι και για την εκπόνηση της παρούσας εργασίας έχουν αξιοποιηθεί κάποιοι φασματικοί δείκτες κατάλληλοι για χαρτογράφηση καμένων εκτάσεων, κρίνεται σκόπιμο να παρουσιαστούν κάποιοι από τους ευρέως χρησιμοποιούμενους δείκτες για τέτοιες εφαρμογές.

1. Normalized Difference Vegetation Index (NDVI)

Ο NDVI αποτελεί έναν από τους πιο διαδεδομένους φασματικούς δείκτες για τη χαρτογράφηση της βλάστησης. Έτσι λοιπόν έχει ευρεία αξιοποίηση και σε εφαρμογές για χαρτογράφηση καμένων εκτάσεων δεδομένου ότι η πυρκαγιά είτε καταστρέφει πλήρως τη βλάστηση είτε την αλλοιώνει πολύ σημαντικά. Ο δείκτης αυτός αξιοποιεί τις τιμές ανακλαστικότητας στο κόκκινο κανάλι και στο εγγύς υπέρυθρο, λαμβάνοντας υπόψη ότι η χρωστική των φύλλων, η χλωροφύλλη, απορροφά σε υψηλά ποσοστά το ορατό φως (από 0,4 έως 0,7 μm) για χρήση κατά τη φωτοσύνθεση και παράλληλα η κυτταρική δομή των φύλλων ανακλά σε μεγάλο βαθμό το εγγύς υπέρυθρο φως (από 0,7 έως 1,1 μm). Ο δείκτης αυτός λαμβάνει τιμές από -1 ως +1 με το ανώτερο όριο να αντιπροσωπεύει πολύ πυκνή βλάστηση ενώ γενικά η τιμή 0 υποδεικνύει απουσία βλάστησης. Σε εφαρμογές χαρτογράφησης καμένων εκτάσεων συνήθως υπολογίζεται ο δείκτης NDVI για δεδομένα πριν το γεγονός της πυρκαγιάς και για δεδομένα μετά και τελικά αξιοποιείται η διαφορά αυτών. Ο τύπος σχηματισμού είναι ο παρακάτω:

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$$

2. Normalized Burn Ratio (NBR)

Ο δείκτης NBR έχει παρόμοιο τύπο με τον NDVI όμως αντί για το κόκκινο κανάλι χρησιμοποιείται το μέσο υπέρυθρο (SWIR). Η υγιής βλάστηση παρουσιάζει υψηλή ανακλαστικότητα στο εγγύς υπέρυθρο και χαμηλή στο μέσο υπέρυθρο (SWIR). Στις περιοχές οι οποίες έχουν καταστραφεί από γεγονός πυρκαγιάς συμβαίνει το αντίθετο, δηλαδή οι καμένες εκτάσεις έχουν χαμηλή ανακλαστικότητα στο εγγύς υπέρυθρο (NIR) και υψηλή στο μέσο (SWIR). Ο τύπος υπολογισμού του φαίνεται παρακάτω:

$$NBR = (NIR - SWIR) / (NIR + SWIR)$$

Η διαφορά του δείκτη αυτού πριν και μετά την πυρκαγιά (dNBR) προτείνεται από το United States Geological Survey (USGS) για την κατηγοριοποίηση του βαθμού καταστροφής των καμένων εκτάσεων¹.

Severity Level	dNBR Range (scaled by 10 ³)	dNBR Range (not scaled)
Enhanced Regrowth, high (post-fire)	-500 to -251	-0.500 to -0.251
Enhanced Regrowth, low (post-fire)	-250 to -101	-0.250 to -0.101
Unburned	-100 to +99	-0.100 to +0.99
Low Severity	+100 to +269	+0.100 to +0.269
Moderate-low Severity	+270 to +439	+0.270 to +0.439
Moderate-high Severity	+440 to +659	+0.440 to +0.659
High Severity	+660 to +1300	+0.660 to +1.300

Εικόνα 2: Βαθμός καταστροφής καμένων εκτάσεων βάσει του δείκτη dNBR, όπως προτείνεται από το USGS.

3. Normalized Burn Ratio 2 (NBR2)

Πρόκειται για παραλλαγή του NBR η οποία έχει τη δυνατότητα εντοπισμού νερού στη βλάστηση και είναι ιδιαίτερα χρήσιμος και για μελέτες παρακολούθησης αποκατάστασης της βλάστησης μετά το γεγονός της πυρκαγιάς. Εξίσου συχνά χρησιμοποιείται για χαρτογράφηση καμένων εκτάσεων, όπως και η διαφορά των τιμών του πριν και μετά την πυρκαγιά (dNBR2).

$$NBR2 = (SWIR1 - SWIR2) / (SWIR1 + SWIR2)$$

4. Relativized Burn Ratio (RBR) (Parks, et al., 2014)

Δημιουργήθηκε ως εναλλακτική των δεικτών dNBR και RdNBR. Ο δείκτης αυτός βασίζεται στον NBR και ουσιαστικά συσχετίζει την διαφορά του ανάμεσα στην κατάσταση πριν και μετά το γεγονός της πυρκαγιάς με την κατάσταση πριν (pre-fire NBR). Πιο αναλυτικά, εικονοστοιχεία τα οποία αντιπροσωπεύουν χαμηλή φυτική κάλυψη πριν την πυρκαγιά θα έχουν γενικά χαμηλές τιμές dNBR ανεξάρτητα από τον βαθμό καταστροφής τους. Ο δείκτης

¹ <https://un-spider.org/advisory-support/recommended-practices/recommended-practice-burn-severity/in-detail/normalized-burn-ratio>

dNBR αδυνατεί λοιπόν πολλές φορές να ανιχνεύσει την υψηλή σοβαρότητα καταστροφής σε τέτοιες περιπτώσεις, ακόμα και αν η βλάστηση έχει εξαφανιστεί τελείως.

$$RBR = (dNBR / (pre-fireNBR + 1.001))$$

5. Burn Area Index (BAI)

Αυτός ο δείκτης υπολογίζεται από τη φασματική απόσταση κάθε εικονοστοιχείου από ένα φασματικό σημείο αναφοράς που οι πρόσφατα καμένες περιοχές τείνουν να συγκλίνουν (Chuviesco, et al., 2002). Συγκεκριμένα ο τύπος υπολογισμού του είναι:

$$BAI = 1 / ((\rho_{Cr} - \rho_r)^2 + (\rho_{Cnir} - \rho_{nir})^2),$$

Όπου ρ_{Cr} και ρ_{Cnir} είναι οι ανακλαστικότητες αναφοράς στο κόκκινο (red) και εγγύς υπέρυθρο κανάλι (NIR).

6. Burned Area Index for Sentinel-2 (Filipponi, 2018)

Ο δείκτης αυτός βασίζεται στα φασματικά κανάλια του Sentinel-2 και το καινοτόμο στοιχείο είναι ότι αξιοποιεί το Red edge κανάλι, το οποίο στοχεύει στην περιγραφή ιδιοτήτων της βλάστησης, σε συνδυασμό με λόγο καναλιών για την ανίχνευση της ραδιομετρικής απόκρισης του φάσματος SWIR.

$$BAIS2 = \left(1 - \sqrt{\frac{B06 * B07 * B8A}{B4}} \right) * \left(\frac{B12 - B8A}{\sqrt{B12 + B8A}} + 1 \right)$$

2.3 Βιβλιογραφική ανασκόπηση

Στην ενότητα αυτή παρατίθενται ενδεικτικά παραδείγματα εφαρμογών για τη χαρτογράφηση καμένων εκτάσεων (burned area mapping) και εκτίμησης του βαθμού καταστροφής τους (burn severity mapping).

Αναφορικά με μεθόδους κατωφλίωσης για χαρτογράφηση καμένων εκτάσεων αλλά και εκτίμησης του βαθμού καταστροφής, ο (Konacs, 2019) χρησιμοποιεί το φασματικό δείκτη RBR για να εξαγάγει τις καμένες εκτάσεις και στην πορεία να κατηγοριοποιήσει το βαθμό καταστροφής τους. Παράλληλα, στην εργασία αυτή παρουσιάζεται ένας νέος δείκτης, ο Burned Vegetation Index (BVI). Ο δείκτης αυτός προσδιορίζει πρώτα την υγιή βλάστηση βάσει της ανάκλασης στο κόκκινο και εγγύς υπέρυθρο κανάλι και στη συνέχεια αφαιρεί από αυτήν τα εικονοστοιχεία που προσδιορίζονται ως καμένες περιοχές βάσει της ανάκλασής τους στο εγγύς υπέρυθρο και στο μέσο υπέρυθρο. Για τον εν λόγω δείκτη έγινε ανάλυση συσχέτισης ως προς τους δείκτες RBR, dNDVI και dNBR, η οποία έδειξε ότι ο BVI έχει υψηλή συσχέτιση με αυτούς τους δείκτες που χρησιμοποιούνται ευρέως στη μελέτη και στη χαρτογράφηση καμένων περιοχών.

Οι (Polychronaki & Gitas, 2012) ανέπτυξαν μία μεθοδολογία αντικειμενοστρεφούς ταξινόμησης για τη χαρτογράφηση καμένων εκτάσεων της πυρκαγιάς στη Μεγαλόπολη το 2007, αξιοποιώντας το λογισμικό eCognition και δορυφορικές εικόνες SPOT-4. Η χρήση

αυτών των δορυφορικών δεδομένων είχε στόχο τη διερεύνηση της πιθανής δυνατότητας αξιοποίησης της τότε μελλοντικής αποστολής του δορυφόρου Sentinel-2. Η μεθοδολογία δημιουργεί δύο επίπεδα κατάτμησης και η ταξινόμηση βασίζεται σε συναρτήσεις συμμετοχής. Για την ταξινόμηση καμένων εκτάσεων αξιοποιούνται τέσσερα χαρακτηριστικά:

- Maximum difference, δηλαδή η μέγιστη διαφορά μεταξύ των μέσων τιμών κάθε αντικειμένου για όλα τα κανάλια.
- Mean value NIR
- Mean value SWIR
- Normalized Burn Ratio (NBR)

Η συνολική ακρίβεια της μεθόδου είναι 97.8%. Η ίδια μεθοδολογία εφαρμόστηκε και στην περιοχή της Πάρνηθας και του Πήλιου προκειμένου να εξεταστεί η δυνατότητα μεταφοράς της. Στις περιοχές αυτές η συνολική ακρίβεια ήταν αντίστοιχα υψηλή και συγκεκριμένα πάνω από 93% και για τις δύο.

Στην εργασία των (Kavzoglu, et al., 2016) αξιοποιείται ο αλγόριθμος Multiresolution segmentation δημιουργώντας δύο επίπεδα αντικειμένων σε συνδυασμό με ταξινόμηση βάσει συναρτήσεων συμμετοχής προκειμένου να χαρτογραφηθούν καμένες εκτάσεις στα νότια παράλια της Τουρκίας. Για το σκοπό αυτό υπολογίζονται οι δείκτες NDVI, NBR και BAI για δορυφορική λήψη Landsat ETM+ μετά το γεγονός της πυρκαγιάς. Οι δείκτες αυτοί χρησιμοποιήθηκαν σε διαφορετικούς συνδυασμούς προκειμένου να κατηγοριοποιηθούν τα αντικείμενα στις κλάσεις Burned, Slightly Burned, Non-Burned και Water. Οι συνδυασμοί αυτοί είναι οι εξής: μεμονωμένος NDVI, BAI-NBR και NDVI-NBR, με τον τελευταίο να εμφανίζεται την υψηλότερη συνολική ακρίβεια (98.37%).

Οι (Comert, et al., 2019) υλοποιούν επίσης μία μεθοδολογία αντικειμενοστρεφούς ταξινόμησης για τη χαρτογράφηση καμένων εκτάσεων χρησιμοποιώντας εικόνες Landsat 8 και επιβλεπόμενη μέθοδο ταξινόμησης. Αρχικά εφαρμόζεται ο αλγόριθμος κατάτμησης Multiresolution segmentation στο περιβάλλον του eCognition και υπολογίζονται τα χαρακτηριστικά των αντικειμένων συμπεριλαμβανομένων των μέσων τιμών καναλιών και δέκα φασματικών δεικτών για καμένες εκτάσεις (BAI, NBR, NBR2, MIRBI) και για τη βλάστηση. Για την ταξινόμηση των αντικειμένων αξιοποιείται επιβλεπόμενη ταξινόμηση και συγκεκριμένα ο αλγόριθμος Random Forest. Τα αποτελέσματα χαρακτηρίζονται από πολύ υψηλή συνολική ακρίβεια (98.7%).

Πρόσφατα, οι (Georgopoulos, et al., 2019) χρησιμοποιώντας δεδομένα Sentinel-2 πριν και μετά το γεγονός της πυρκαγιάς, ανέπτυξαν μία αντικειμενοστρεφή μεθοδολογία για τη χαρτογράφηση καμένων εκτάσεων ελαχιστοποιώντας την παρέμβαση του χρήστη. Πιο συγκεκριμένα, υπολογίζονται οι παρακάτω φασματικοί δείκτες για κάθε χρονική στιγμή και υπολογίζονται οι διαφορές τους:

- Mid-Infrared Burned Index (MIRBI)
- Normalized Burn Ration 1 (NBR2)
- Normalized Difference Infrared Index (NDII)
- Normalized Difference Water Index (NDWI)

Πραγματοποιείται κατάτμηση με τον αλγόριθμο Mean shift αξιοποιώντας μόνο τα κανάλια με την υψηλότερη χωρική ανάλυση, δηλαδή Blue, Green, Red και NIR. Από τα αντικείμενα που δημιουργούνται επιλέγονται ως αντικείμενα εκπαίδευσης μόνο το 1%, αξιοποιώντας τη

μέθοδο fuzzy C-means (FCM), στα οποία εφαρμόζονται δύο εμπειρικοί κανόνες για τον χαρακτηρισμό των αναμφίβολα καμένων και μη καμένων εκτάσεων. Οι κανόνες αυτοί αξιοποιούν τους δείκτες που έχουν υπολογιστεί σε προηγούμενο βήμα. Όσα δείγματα δεν χαρακτηρίζονται βάσει αυτών των κανόνων τότε είτε παραμένουν χωρίς χαρακτηρισμό (πλήρως αυτοματοποιημένη μέθοδος) είτε χαρακτηρίζονται χειροκίνητα από τον χρήστη (ημι-αυτόματη μέθοδος). Η μεθοδολογία αυτή εφαρμόστηκε συνολικά σε οκτώ περιοχές και τα αποτελέσματα της συνολικής ακρίβειας είναι πάνω από 93% τόσο για την πλήρως αυτοματοποιημένη εκδοχή όσο και για την ημι-αυτόματη με την τελευταία να εμφανίζει ελαφρώς καλύτερα αποτελέσματα.

Σχετικά με ακόμα πιο προηγμένες μεθόδους μηχανικής μάθησης για την χαρτογράφηση καμένων εκτάσεων, οι (Belenguer-Plomer, et al., 2021) αναπτύσσουν Συνελκτικά Νευρωνικά Δίκτυα (CNN) αξιοποιώντας δεδομένα Sentinel-1 και Sentinel-2 για την χαρτογράφηση καμένων εκτάσεων. Πιο συγκεκριμένα, ως περιοχές μελέτης ορίστηκαν δέκα περιοχές, κάθε μία από τις οποίες αντιπροσωπεύει διαφορετικά οικοσυστήματα που όμως όλα επηρεάζονται σε μεγάλο βαθμό από πυρκαγιές. Μόνο 5 από αυτές χρησιμοποιήθηκαν για τη συλλογή των δεδομένων εκπαίδευσης. Από τα πολυφασματικά δεδομένα έγινε υπολογισμός και των δεικτών NBR, NDVI, NDWI και MIRBI. Τα αποτελέσματά τους σχετικά με την αρχιτεκτονική των δικτύων δείχνουν ότι ανεξάρτητα από τα δεδομένα εισόδου (Sentinel-1, Sentine-2 ή και τα δύο), δύο κρυφές στρώσεις επιτρέπουν την μείωση του υπολογιστικού χρόνου χωρίς αυτό να συνεπάγεται μείωση της ακρίβειας χαρτογράφησης κατά τη σύγκριση με πιο βαθιά δίκτυα. Τέλος, βρέθηκε ότι η συνολική ακρίβεια στην περίπτωση που τα δεδομένα εισόδου είναι μόνο Sentinel-1 ή μόνο Sentinel-2 είναι αντίστοιχη με αυτή στην υπάρχουσα βιβλιογραφία.

Στο τρίτο κεφάλαιο περιγράφονται τα δεδομένα που χρησιμοποιήθηκαν στο σύνολο της εργασίας. Πιο συγκεκριμένα παρουσιάζονται τόσο οι δορυφορικές εικόνες στις οποίες εφαρμόστηκαν οι σχετικές μεθοδολογίες όσο και η πηγή των δεδομένων αληθείας. Και για τις δύο περιπτώσεις δεδομένων αναλύονται οι επεξεργασίες που εκτελέστηκαν προκειμένου να τροφοδοτήσουν τα μετέπειτα βήματα.

3.1 Δορυφορικά δεδομένα

Ως περιοχή μελέτης στην παρούσα εργασία έχει οριστεί το Μάτι και πιο συγκεκριμένα η ευρύτερη περιοχή (περίπου 37km²) η οποία πλήγηκε από την θανατηφόρα πυρκαγιά στις 23-07-2018. Βασική επιδίωξη της εργασίας αυτής είναι η διερεύνηση της δυνατότητας χαρτογράφησης των καμένων εκτάσεων και του βαθμού καταστροφής τους αξιοποιώντας ελεύθερα δορυφορικά δεδομένα. Λαμβάνοντας αυτό ως οδηγό, θεωρήθηκε σκόπιμο να χρησιμοποιηθούν τα δεδομένα που παρέχει ο δορυφόρος Sentinel-2, ο οποίος συνδυάζει υψηλή χωρική και χρονική ανάλυση. Πιο αναλυτικά, η πλήρης προδιαγραφή αποστολής των δίδυμων δορυφόρων (Sentinel-2A, Sentinel-2B) που πετούν στην ίδια τροχιά αλλά σταδιακά στις 180°, έχει σχεδιαστεί για να παρέχει υψηλή συχνότητα επανεπίσκεψης πέντε (5) ημερών στον Ισημερινό². Οι εικόνες από τον δορυφόρο Sentinel-2 περιέχουν δεκατρία (13) φασματικά κανάλια σε διαφορετική χωρική ανάλυση, τα οποία φαίνονται και στην εικόνα 3.

Sentinel-2 bands	Central wavelength (μm)	Resolution (m)
Band 1 – Coastal aerosol	0.443	60
Band 2 – Blue	0.490	10
Band 3 – Green	0.560	10
Band 4 – Red	0.665	10
Band 5 – Vegetation red edge	0.705	20
Band 6 – Vegetation red edge	0.740	20
Band 7 – Vegetation red edge	0.783	20
Band 8 – NIR	0.842	10
Band 8A – Vegetation red edge	0.865	20
Band 9 – Water vapour	0.945	60
Band 10 – SWIR – Cirrus	1.375	60
Band 11 – SWIR	1.610	20
Band 12 – SWIR	2.190	20

Εικόνα 3: Φασματικά κανάλια του δορυφόρου Sentinel-2 και αντίστοιχη χωρική ανάλυση (πηγή: https://www.researchgate.net/figure/Sentinel-2-band-characteristics_tbl1_314119510).

Τα δεδομένα Sentinel-2 παρέχονται προς τους χρήστες σε δύο διαφορετικά επίπεδα επεξεργασίας: το Level-1C και Level-2A. Επιγραμματικά, τα προϊόντα Level-1C έχουν υποστεί ραδιομετρικές και γεωμετρικές διορθώσεις, προβάλλονται σε παγκόσμιο σύστημα αναφοράς και οι τιμές των εικονοστοιχείων τους αντιπροσωπεύουν την ανακλαστικότητα στο ανώτερο επίπεδο της ατμόσφαιρας (Top-Of-Atmosphere reflectance). Από την άλλη, τα

² <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>

προϊόντα Level-2A πρόκεινται στην ουσία για προϊόντα Level-1C που έχουν υποστεί ορθοαναγωγή και οι τιμές τους αντιπροσωπεύουν την ανακλαστικότητα στην επιφάνεια της Γης (Bottom-Of-Atmosphere reflectance). Στα πλαίσια αυτής της εργασίας επιλέχθηκε να χρησιμοποιηθούν προϊόντα Level-2A. Συγκεκριμένα, χρησιμοποιείται μία λήψη πριν το γεγονός της πυρκαγιάς (05-07-2018) και μία λήψη μετά (09-08-2018). Οι κωδικοί αυτών είναι οι εξής:

- S2B_MSIL2A_20180705T091019_N0208_R050_T35SKC_20180705T133603
- S2A_MSIL2A_20180809T090551_N0208_R050_T35SKC_20180809T122252



Εικόνα 4: Δορυφορική εικόνα Sentinel-2 (Level-2A) μετά το γεγονός της πυρκαγιάς στο Μάτι.

3.2 Προ-επεξεργασία δορυφορικών εικόνων

Στην προηγούμενη ενότητα αναφέρθηκε ότι το επίπεδο επεξεργασίας των δορυφορικών εικόνων που επιλέχθηκαν είναι Level-2A, δηλαδή προϊόντα Bottom-Of-Atmosphere (BOA) στα οποία έχει γίνει ορθοαναγωγή. Κατά την επισκόπηση των δεδομένων παρατηρήθηκε ότι οι τιμές των εικονοστοιχείων τους ήταν σε Digital Number (DN) format και άρα πρώτο βήμα, πριν από οποιαδήποτε προ-επεξεργασία, είναι η διαίρεση των τιμών αυτών με 10000 βάσει της παρακάτω σχέσης:

$$DN = 10000 * REFLECTANCE (1)$$

Η μετατροπή αυτή, όπως και όλες οι υπόλοιπες επεξεργασίες που περιγράφονται σε αυτή την ενότητα πραγματοποιήθηκαν σε προγραμματιστικό περιβάλλον Python με την αξιοποίηση πληθώρας βιβλιοθηκών όπως η Numpy, rasterio, GDAL κ.α.

1)Επαναδειγματοληψία (Resampling)

Όπως φαίνεται και στην Εικόνα 1 τα φασματικά κανάλια των δορυφορικών δεδομένων Sentinel-2 δεν έχουν όλα την ίδια χωρική ανάλυση. Μία από τις πρώτες προ-επεξεργασίες λοιπόν είναι η επαναδειγματοληψία (resampling) προκειμένου να καλυφθεί η ανάγκη αξιοποίησης φασματικών καναλιών με διαφορετική χωρική ανάλυση. Η ανάγκη αυτή έγκειται στη δημιουργία φασματικών δεικτών κατάλληλων για χαρτογράφηση καμένων εκτάσεων, οι οποίοι βασίζονται σε φασματικά κανάλια πέραν των Blue, Green, Red και Near-Infrared που έχουν ενιαία χωρική ανάλυση 10m. Αναλυτικότερα, από τα 13 κανάλια του δορυφόρου χρησιμοποιήθηκαν μόνο τα 10 που αντιστοιχούν σε χωρική ανάλυση 10m και 20m, δηλαδή **εξαιρέθηκαν** τα παρακάτω:

- Band 1 - Coastal aerosol (60m)
- Band 9 - Water vapour (60m)
- Band 10 - SWIR Cirrus (60m)

Κατά το resampling, όλα τα κανάλια που έχουν ανάλυση 20m ανακατασκευάστηκαν έχοντας ανάλυση 10m. Αυτό έγινε αξιοποιώντας την βιβλιοθήκη Rasterio και συγκεκριμένα τη συνάρτηση Resampling. Η μέθοδος που επιλέχθηκε για την απόδοση τιμών στα εικονοστοιχεία της νέας εικόνας είναι η bilinear³.

2)Περικοπή (Clip)

Αμέσως μετά την επαναδειγματοληψία έγινε περικοπή (clip) των καναλιών με αρχική ανάλυση 10m αλλά και αυτών στα οποία εφαρμόστηκε το resampling. Πιο συγκεκριμένα, για αυτή την επεξεργασία χρησιμοποιήθηκε η βιβλιοθήκη GDAL και η συνάρτηση Warp.

3)Επίθεση φασματικών καναλιών (Layer stack)

Τα δεδομένα που προέκυψαν από τα προηγούμενα βήματα αποθηκεύτηκαν σε δύο ξεχωριστά ενιαία αρχεία GeoTIFF ανάλογα με τη χρονική στιγμή λήψης τους, δηλαδή ένα αρχείο για την δορυφορική εικόνα πριν την πυρκαγιά (pre-fire) και ένα μετά (post-fire). Καθένα από αυτά τα δύο αρχεία περιλαμβάνει δέκα (10) φασματικά κανάλια, όπως αναφέρθηκε και προηγουμένως.

4)Δημιουργία φασματικών δεικτών (Spectral indices)

Έχοντας πλέον την pre-fire και post-fire εικόνα για την περιοχή μελέτης με δέκα (10) φασματικά κανάλια και με την καλύτερη δυνατή χωρική ανάλυση βάσει των πρωταρχικών δεδομένων, δηλαδή 10m, επόμενο βήμα είναι η δημιουργία των κατάλληλων φασματικών δεικτών. Οι φασματικοί δείκτες που δημιουργούνται βασίζονται στη διεθνή βιβλιογραφία σχετικά με χαρτογράφηση καμένων εκτάσεων και burn severity και η σημασία τους έχει περιγραφεί στο Κεφάλαιο 2. Η κατασκευή τους επίσης έγινε σε προγραμματιστικό περιβάλλον Python. Οι δείκτες δημιουργήθηκαν τόσο για την pre-fire εικόνα όσο και για την

³ <https://rasterio.readthedocs.io/en/latest/topics/resampling.html#resampling-methods>

post-fire ενώ υπολογίστηκαν και διαφορές δεικτών μεταξύ των δύο χρονικών στιγμών, όπου αυτό εφαρμόζεται. Παρακάτω ακολουθεί η πλήρης λίστα με τους φασματικούς δείκτες που δημιουργήθηκαν για την παρούσα εργασία:

Πίνακας 1: Φασματικοί δείκτες που δημιουργήθηκαν στα πλαίσια της εργασίας.

Φασματικοί δείκτες	Pre-fire	Post-fire	Διαφορά δεικτών d
NDVI	✓	✓	✓
NBR	✓	✓	✓
NBR2	✓	✓	✓
RBR	-	-	✓

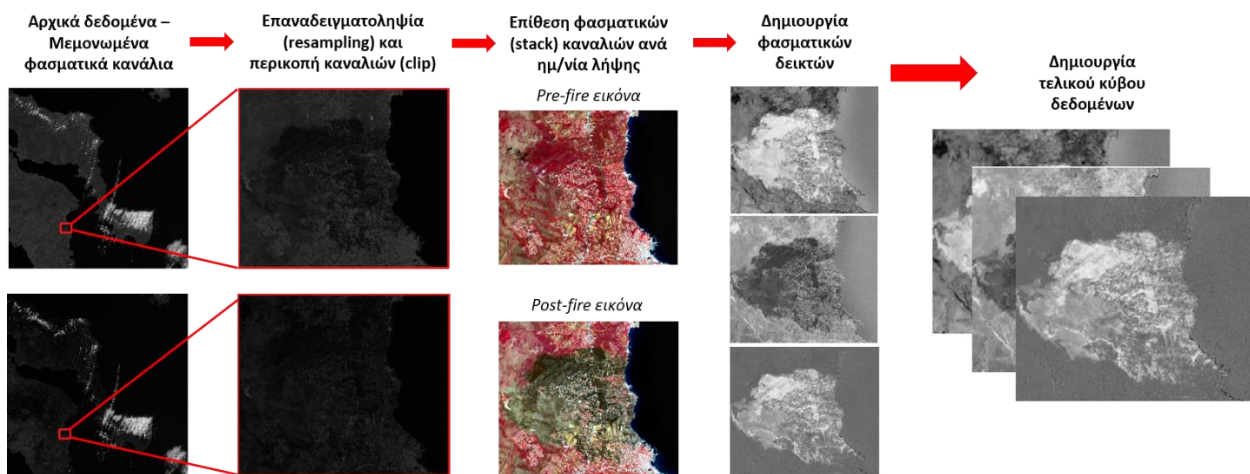
5) Δημιουργία τελικού κύβου δεδομένων

Τα τελικά δορυφορικά δεδομένα προς ταξινόμηση συνολικά αποτελούνται από 19 κανάλια. Συγκεκριμένα, περιλαμβάνουν όλα τα φασματικά κανάλια της λήψης μετά το γεγονός της πυρκαγιάς, εκτός των τριών με χωρικά ανάλυση 60m, δηλαδή της post-fire εικόνας που δημιουργήθηκε, καθώς και τους φασματικούς δείκτες που φαίνονται στον πίνακα παρακάτω. Η ενίσχυση των δεδομένων με τους δείκτες αυτούς έχει στόχο τον καλύτερη δυνατή χαρτογράφηση των καμένων εκτάσεων αλλά και του διαχωρισμού μεταξύ τους ανάλογα με τον βαθμό καταστροφής τους. Σημειώνεται ότι για να αποφασιστεί η τελική σύνθεση του κύβου δεδομένων έγιναν αρκετοί πειραματισμοί με τον αλγόριθμο ταξινόμησης στη γλώσσα Python.

Πίνακας 2: Κανάλια του τελικού κύβου δεδομένων προς ταξινόμηση.

1 Blue (post)	11 NDVI (pre-fire)
2 Green (post)	12 NBR (pre-fire)
3 Red (post-fire)	13 NDVI (post-fire)
4 Veg Red Edge (post-fire)	14 NBR (post-fire)
5 Red Edge (post-fire)	15 dNDVI
6 Red Edge (post-fire)	16 dNBR
7 NIR (post-fire)	17 NBR2 (post-fire)
8 Veg Red Edge (post-fire)	18 dNBR2
9 SWIR 1 (post-fire)	19 RBR
10 SWIR 2 (post-fire)	

Τα βήματα προ-επεξεργασίας που παρουσιάστηκαν αναλυτικά παραπάνω αποτυπώνονται στο παραστατικό διάγραμμα ροής αμέσως παρακάτω.



Εικόνα 5: Διάγραμμα ροής προ-επεξεργασιών για τη δημιουργία του τελικού σετ δεδομένων προς ταξινόμηση.

3.3 Πηγή και μέθοδος συλλογής δεδομένων αληθείας (ground truth)

Προκειμένου να αξιολογηθούν τα αποτελέσματα των δύο προσεγγίσεων που ακολουθήθηκαν είναι απαραίτητη η ύπαρξη δεδομένων αληθείας ή αλλιώς ground truth. Τα δεδομένα αληθείας πρέπει να διακρίνονται από υψηλή αξιοπιστία προκειμένου ο έλεγχος των αποτελεσμάτων να είναι και αυτός κατ' επέκταση αξιόπιστος. Πολύ συχνά σε εφαρμογές Τηλεπισκόπησης η συλλογή του ground truth γίνεται με φωτοερμηνεία και ψηφιοποίηση δορυφορικών δεδομένων πολύ υψηλής ανάλυσης (Αντωνίου, 2017) ή και με επιτόπια έρευνα/μετρήσεις (Καρακίζη, 2015). Επίσης, συχνά συναντάται η αξιοποίηση δεδομένων που έχουν παραχθεί από κρατικούς φορείς ή επιχειρησιακά προγράμματα δεδομένου ότι αυτά συνήθως παράγονται με επιτόπιες μετρήσεις/καταγραφές ή γενικότερα με πολύ αξιόπιστη μέθοδο.

Στην παρούσα εργασία, ως δεδομένα αληθείας χρησιμοποιήθηκαν διανυσματικά (vector) αρχεία που παρέχονται από την υπηρεσία Copernicus Emergency Management Service (CEMS). Το CEMS πρόκειται για μία υπηρεσία η οποία χρησιμοποιεί δορυφορικές εικόνες και άλλα γεωχωρικά δεδομένα με σκοπό να παρέχει δωρεάν υπηρεσίες χαρτογράφησης σε περιπτώσεις φυσικών καταστροφών, ανθρωπογενών καταστάσεων έκτακτης ανάγκης και ανθρωπιστικών κρίσεων για όλο τον κόσμο⁴. Η υπηρεσία αυτή ενεργοποιείται από εξουσιοδοτημένους χρήστες οι οποίοι μπορεί να είναι φορείς και οργανισμοί σε περιφερειακό, εθνικό, ευρωπαϊκό και διεθνές επίπεδο. Μερικά από τα καταστροφικά συμβάντα τα οποία καλύπτουν είναι τα εξής:

- Πλημύρες
- Σεισμοί
- Πυρκαγιές
- Εκρήξεις ηφαιστειών κ.α.

⁴ <https://emergency.copernicus.eu/mapping/ems/emergency-management-service-mapping>

Αναφορικά με τα δεδομένα αληθείας που χρησιμοποιήθηκαν αξίζει αρχικά να αναφερθεί ότι ήταν κοινά και για τις δύο διαφορετικές προσεγγίσεις της εργασίας. Στην μεν πρώτη τα δεδομένα αυτά συμμετείχαν και στη διαδικασία εκπαίδευσης (training data) και στη διαδικασία ελέγχου (test data) ενώ στη δεύτερη, η οποία βασίζεται σε κανόνες ασαφούς λογικής και άρα δεν απαιτείται διαδικασία εκπαίδευσης, αξιοποιήθηκαν μόνο στο μέρος της αξιολόγησης των αποτελεσμάτων. Τα ground truth στην προκειμένη περίπτωση είναι αποτέλεσμα της ενότητας Rapid Mapping⁵ του CEMS, σύμφωνα με την οποία παρέχονται χάρτες και σύντομες αναλύσεις εντός ωρών ή λίγων ημερών μετά το καταστροφικό γεγονός. Στην ενότητα του Rapid Mapping εντάσσονται τρεις (3) βασικές κατηγορίες παραχθέντων προϊόντων:

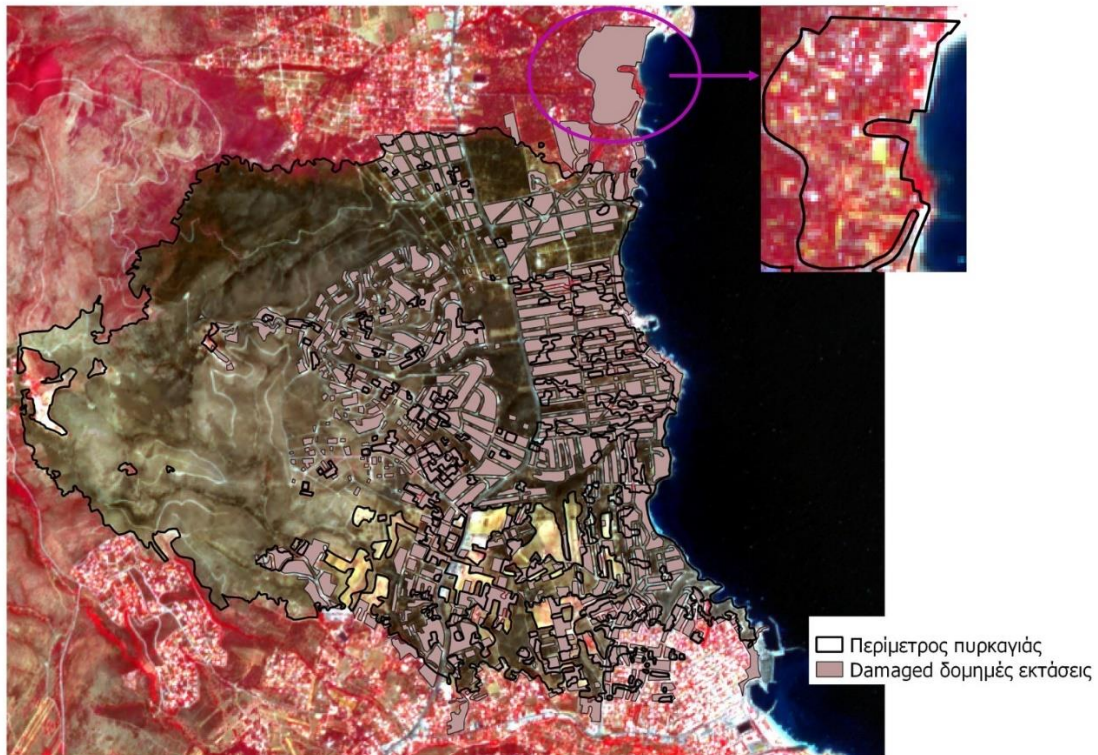
- Reference Maps: βασίζονται σε δορυφορική εικόνα πριν το γεγονός και συνήθως χρησιμοποιούνται ως βάση για συγκριτικούς σκοπούς κατά τη δημιουργία του προϊόντος μετά το γεγονός (post-emergency product).
- Delineation maps: χάρτες που αποτυπώνουν το περίγραμμα της περιοχής που έχει επηρεαστεί από το γεγονός.
- Grading maps: χάρτες που παρέχουν αξιολόγηση των επιπτώσεων του καταστροφικού γεγονότος.

Δεδομένου ότι η παρούσα εργασία δεν στοχεύει μόνο στη χαρτογράφηση καμένων εκτάσεων αλλά επιδιώκει να χαρτογραφήσει και το βαθμό καταστροφής αυτών, αξιοποιήθηκαν τα δεδομένα από το προϊόν Grading map για το Μάτι (EMSR300). Αξίζει να αναφερθεί ότι δεδομένα από το CEMS έχουν χρησιμοποιηθεί ως ground truth και σε άλλες μελέτες για χαρτογράφηση καμένων εκτάσεων και βαθμού καταστροφής από πυρκαγιά (Filipponi, 2019), (Shokouhi & Eslami, 2019), (Pinto, et al., 2021).

Πιο συγκεκριμένα, τα περιεχόμενα του προϊόντος Grading EMSR300 παρέχουν πληροφορίες σχετικά με τον βαθμό ζημιάς, τη χωρική κατανομή και έκτασή του γεγονότος της πυρκαγιάς στο Μάτι. Το προϊόν περιλαμβάνει διανυσματικά αρχεία που απεικονίζουν δεδομένα για την τοπογραφία της περιοχής, όπως ισοϋψείς καμπύλες και υδρογραφικό δίκτυο, δεδομένα για υφιστάμενες υποδομές όπως για παράδειγμα το οδικό δίκτυο καθώς και θεματικά επίπεδα που περιέχουν πληροφορία για το βαθμό καταστροφής από την πυρκαγιά. Αυτά τα θεματικά επίπεδα είναι το “natural_land_use” και το “built_up”. Στο μεν πρώτο αποτυπώνονται οι περιοχές με φυσική κάλυψη γης και ο βαθμός ζημιάς τους από την πυρκαγιά ενώ στο δεύτερο οι δομημένες εκτάσεις και ο αντίστοιχος βαθμός ζημιάς. Από αυτά τα δύο τελικά επιλέχθηκε να χρησιμοποιηθεί μόνο το “natural_land_use” για την παραγωγή του ground truth. Η επιλογή αυτή έγινε έπειτα από λεπτομερή επισκόπηση των εν λόγω θεματικών επιπέδων από την οποία παρατηρήθηκαν ορισμένα αντικρουόμενα δεδομένα μεταξύ τους. Αναλυτικότερα, όλα τα πολύγωνα του layer “built_up” χαρακτηρίζονται ως Damaged ακόμα και σε ορισμένες περιοχές που στο layer “natural_land_use” επικαλυπτόμενα τμήματα χαρακτηρίζονται ως Not Burned. Ύστερα από φωτοερμηνεία σε ορισμένα από αυτά τα τμήματα παρατηρήθηκε ότι το θεματικό επίπεδο “natural_land_use” απεικονίζει ορθότερα την κατάσταση. Επιπλέον παρατηρήθηκε ότι μία μεγάλη έκταση, η οποία βρίσκεται εκτός της περιμέτρου της πυρκαγιάς όπως αυτή αποτυπώνεται στο προϊόν Delineation, βάσει φωτοερμηνείας φαίνεται

⁵ https://emergency.copernicus.eu/mapping/sites/default/files/files/CEMS_FlyerRapidMapping_2021.pdf

ότι δεν είναι καμένη ενώ στο layer “built_up” είναι καταγεγραμμένη ως Damaged (βλ. Εικόνα 6).






Εικόνα 6: Περίμετρος πυρκαγιάς και θεματικό επίπεδο “built-up”. Επισήμανση ανακρίβειας των δεδομένων “built-up”.

Οι παρατηρήσεις αυτές, σε συνδυασμό με το γεγονός ότι η κατηγοριοποίηση του βαθμού καταστροφής για το επίπεδο “natural_land_use” έχει προκύψει από ημι-αυτόματη μέθοδο ενώ για το “built_up” με τελείως αυτόματη, οδήγησαν στον αποκλεισμό του δεύτερου από τη συλλογή ground truth. Σύμφωνα με το CEMS, κατά την ημι-αυτόματη μέθοδο παραγωγής των δεδομένων (*Semi-automatic extraction*) η ανάλυση των δορυφορικών εικόνων πραγματοποιείται με χρήση επιβλεπόμενης ταξινόμησης. Αυτή η προσέγγιση εφαρμόζεται συνήθως όπου το συμβάν παράγει αρκετά τυπικά αποτελέσματα όσον αφορά την ραδιομετρική απόκριση και λειτουργεί καλά όταν υπάρχουν διαθέσιμες εικόνες διαφορετικής χρονικής στιγμής με τον ίδιο αισθητήρα ή ανάλυση⁶. Αυτό ισχύει για τα δεδομένα CEMS που αξιοποιούνται στην παρούσα εργασία καθώς οι δορυφορικές εικόνες πριν και μετά το γεγονός της πυρκαγιάς, στις οποίες εφαρμόστηκε η ημι-αυτόματη μέθοδος, είναι και οι δύο από τον δορυφόρο Pleiades. Αναφορικά με την αυτόματη μέθοδο (*Automatic extraction*) αναφέρεται ότι χρησιμοποιείται σπάνια στο πλαίσιο των προϊόντων Rapid Mapping, συμπεριλαμβανομένων δηλαδή και των δεδομένων Grading, επειδή δεν είναι σύνηθες να εγγυάται την ομοιογένεια μεταξύ των δεδομένων πριν και μετά το συμβάν.

Έπειτα από την περιγραφή των δεδομένων αλλά και της ποιότητάς τους είναι σημαντικό να γίνει αναφορά στον βαθμό καταστροφής που αυτά αποδίδουν στα πολύγωνα αναφοράς. Οι κλάσεις λοιπόν στις οποίες κατηγοριοποιούνται οι καμένες εκτάσεις βάσει του θεματικού

⁶ <https://emergency.copernicus.eu/mapping/ems/detection-methods>

επιπέδου “natural_land_use” του προϊόντος Grading map είναι: 1)No visible damage, 2)Possibly damaged, 3)Damaged και 4)Destroyed. Τα πολύγωνα του CEMS δηλαδή, μεταξύ άλλων χαρακτηριστικών που περιλαμβάνουν, στον πίνακα περιγραφικών χαρακτηριστικών τους υπάρχει και πεδίο “damage_gra” το οποίο αντιστοιχεί στις κλάσεις αυτές.

Copernicus EMS classes		EMS-98 classes
	Possibly damaged	It refers to cases when the confidence level of the interpretation is slightly lower (e.g. bad image quality).
	Damaged	Grade 1: Negligible to slight damage Grade 2: Moderate damage Grade 3: Substantial to heavy damage
	Destroyed	Grade 4: Very heavy damage Grade 5: Destruction

Εικόνα 7: Επεξήγηση των κλάσεων των δεδομένων Grading map από το CEMS που χρησιμοποιήθηκαν ως δεδομένα αναφοράς.

Όπως φαίνεται και από την εικόνα παραπάνω, η κλάση Possibly damaged δεν σχετίζεται ευθέως με το βαθμό καταστροφής αλλά με ελαφρώς χαμηλότερο επίπεδο εμπιστοσύνης ερμηνείας/κατηγοριοποίησης της εν λόγω έκτασης στη δεδομένη εικόνα που χρησιμοποιείται για την εξαγωγή των δεδομένων. Λαμβάνοντας υπόψη αυτό αλλά και το γεγονός ότι τα δορυφορικά δεδομένα στα οποία εφαρμόζονται οι δύο μεθοδολογίες ταξινόμησης είναι διαφορετικά από αυτά τα οποία χρησιμοποιήθηκαν για την παραγωγή των προϊόντων του CEMS, αποφασίστηκε στην προκειμένη περίπτωση η κλάση Possibly damaged να συγχωνευθεί με την Damaged. Επιπλέον, στην απόφαση αυτή λήφθηκε υπόψη και η συνολικά μικρή έκταση αυτής της κλάσης. Έτσι λοιπόν, όπως θα αναφερθεί αναλυτικότερα και στο κεφάλαιο 4, τελικά οι κλάσεις ταξινόμησης που επιλέχθηκαν και σχετίζονται με την πυρκαγιά είναι τρεις (3):

1. Not_Burned, η οποία ταυτίζεται με τη No visible damage του CEMS
2. Damaged
3. Destroyed

Επειδή τα δορυφορικά δεδομένα προς ταξινόμηση απεικονίζουν και θάλασσα δημιουργήθηκε ακόμα μία κλάση ταξινόμησης που ονομάστηκε Water. Συνολικά λοιπόν οι κλάσεις ταξινόμησης, για τις οποίες προφανώς είναι αναγκαίο να συλλεχθούν και δεδομένα αληθείας, είναι τέσσερις (4).

Η διαδικασία δημιουργίας σημειακών δεδομένων αληθείας από τα πολυγωνικά δεδομένα CEMS που έχουν περιγραφεί παραπάνω, πραγματοποιήθηκε εξ ολοκλήρου στο ελεύθερο λογισμικό QGIS και ακολούθησε την ροή που περιγράφεται εν συνεχεία. Αρχικά έγινε ένωση των πολυγώνων Possibly damaged και Damaged και έπειτα διαχωρίστηκαν τα πολύγωνα του CEMS ανά κλάση (Not Burned, Damaged, Destroyed) και για κάθε μία δημιουργήθηκαν

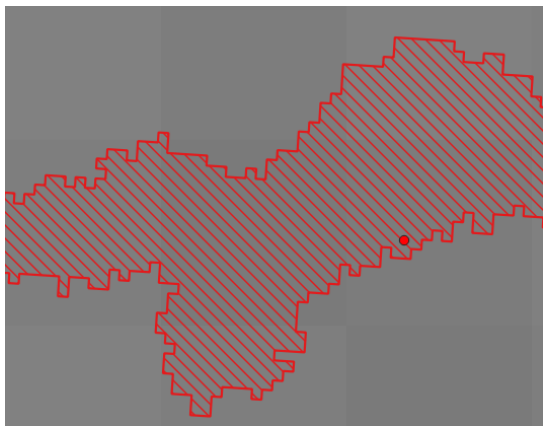
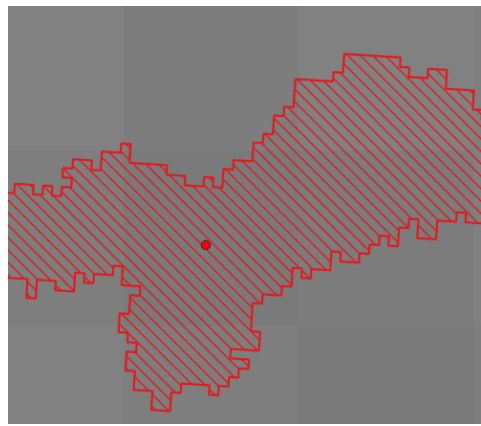
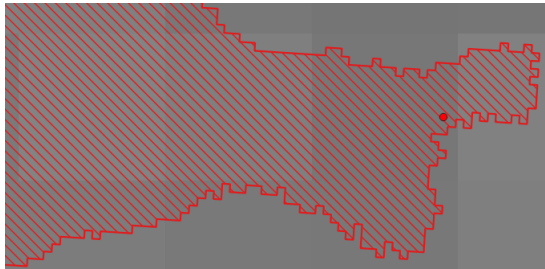
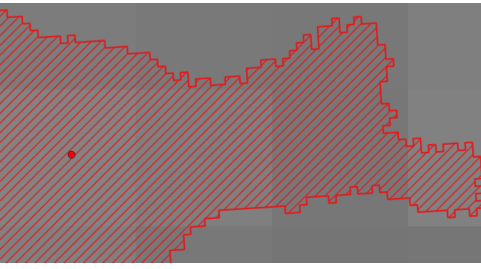
σημεία εντός των πολυγώνων της με χαρακτηριστικό πεδίο "id" στο οποίο αποδίδεται η κατηγορία κωδικοποιημένη με ακέραιο αριθμό (1: Not_Burned, 2: Damaged, 3: Destroyed). Σημειώνεται ότι η δημιουργία των σημείων έγινε με αξιοποίηση του αλγορίθμου Random points in polygons ορίζοντας διαφορετικό αριθμό σημείων εντός κάθε πολυγώνου ανάλογα με το πλήθος των πολυγώνων κάθε κλάσης και ελάχιστη απόσταση 10m, δηλαδή όσο το μέγεθος του εικονοστοιχείου των δεδομένων προς ταξινόμηση. Αυτό επιλέχθηκε ώστε τα σημεία του ground truth να αντιπροσωπεύουν διαφορετικά εικονοστοιχεία. Επειδή το εργαλείο αυτό δημιουργεί με τυχαίο τρόπο τα σημεία, στην επόμενη ενότητα προφανώς γίνεται περιγραφή του ποιοτικού ελέγχου που διεξάχθηκε και των τροποποιήσεων που έγιναν προκειμένου τα δεδομένα αληθείας να είναι όσο πιο αξιόπιστα γίνεται. Είναι σημαντικό να επισημανθεί ότι αυτή η μέθοδος παραγωγής ground truth επιλέχθηκε σε ένα πλαίσιο διερεύνησης της δυνατότητας παραγωγής ποιοτικών αποτελεσμάτων χρησιμοποιώντας ημι-αυτόματο τρόπο παραγωγής δεδομένων αληθείας, λαμβάνοντας φυσικά υπόψη την υψηλή ποιότητα των δεδομένων CEMS.

Όσον αφορά στην κλάση Water, τα σημεία συλλέχθηκαν χειροκίνητα με υπόβαθρο τα δορυφορικά δεδομένα Sentinel-2 που χρησιμοποιούνται στην εργασία. Στα σημεία αυτά στο πεδίο "id" αποδόθηκε ο αριθμός 4, υποδηλώνοντας έτσι την κλάση που αντιπροσωπεύουν.

3.4 Ποιοτικός έλεγχος δεδομένων αληθείας (ground truth)



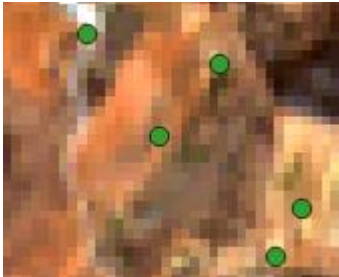

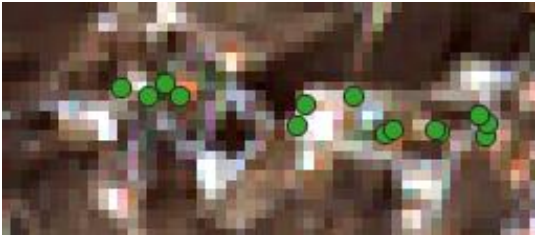
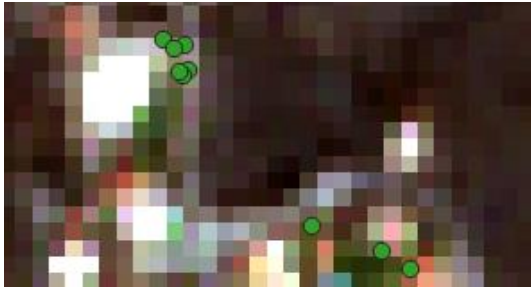
Για τη διεξαγωγή του ποιοτικού ελέγχου συνδυάστηκαν τα πολυγωνικά δεδομένα του CEMS, τα σημεία που δημιουργήθηκαν εντός αυτών στο περιβάλλον του QGIS και τα δορυφορικά δεδομένα προς ταξινόμηση. Αυτό που παρατηρήθηκε είναι ότι ορισμένα πολύγωνα είναι πολύ μικρά, ιδίως της κλάσης Destroyed η οποία αντιστοιχεί συνολικά σε αρκετά μικρή έκταση (190.631,15 m²), και σε συνδυασμό με τη μέση χωρική ανάλυση των δεδομένων του Sentinel-2 (10m), στα πολύγωνα αυτά αντιστοιχούν λίγα εικονοστοιχεία. Έτσι λοιπόν, κρίθηκε σημαντικό τα σημεία που θα χρησιμοποιηθούν ως ground truth, να αντιστοιχούν σε εικονοστοιχεία που βρίσκονται εξ ολοκλήρου εντός των πολυγώνων αναφοράς του CEMS ή έστω κατά το μεγαλύτερο μέρος τους όταν το πρώτο δεν είναι εφικτό. Αυτό αποτέλεσε και το πρώτο ζητούμενο του ποιοτικού ελέγχου. Στα παραδείγματα που ακολουθούν παρουσιάζονται ενδεικτικές περιπτώσεις από τη διαδικασία του ελέγχου που αναδεικνύουν τη χωρική διάστασή του που περιγράφηκε παραπάνω. Τα παραδείγματα επικεντρώνονται στα δεδομένα ελέγχου για την κατηγορία Destroyed καθώς εκεί αντιμετωπίστηκε το μεγαλύτερο πρόβλημα λόγω της μικρής έκτασης αυτής της κλάσης. Χαρακτηριστικό είναι ότι λαμβάνοντας υπόψη ότι το εμβαδό του εικονοστοιχείου των δεδομένων Sentinel-2 είναι 100m², η κλάση αυτή αντιπροσωπεύεται μόνο από 1906 pixels στα δεδομένα προς ταξινόμηση, δηλαδή μόλις το 0,5% όλης της εικόνας.

Πίνακας 3: Ενδεικτικές περιπτώσεις τροποποίησης των δεδομένων αληθείας κατά τον ποιοτικό έλεγχό τους.


	
<p>Στην αριστερή εικόνα παρουσιάζεται μία περίπτωση σημειακού δεδομένου ελέγχου (test) της κλάσης Destroyed που αντιστοιχεί σε εικονοστοιχείο το οποίο δεν ανήκει εξ ολοκλήρου στο αντίστοιχο πολύγωνο του CEMS. Στη δεξιά εικόνα απεικονίζεται το ίδιο σημείο μετά από τη μετακίνησή του κατά τον ποιοτικό έλεγχο ώστε να αντιστοιχεί σε εικονοστοιχείο το οποίο ανήκει εξ ολοκλήρου ή τουλάχιστον κατά το μεγαλύτερο μέρος του σε πολύγωνο αναφοράς. Το συγκεκριμένο πολύγωνο είναι επίμηκες και έτσι δεν υπήρχε κάποιο pixel που να βρίσκεται ολόκληρο εντός του πολυγώνου αναφοράς.</p>	
	
<p>Τα στιγμιότυπα παραπάνω αποτυπώνουν αντίστοιχη διόρθωση που πραγματοποιήθηκε κατά τον ποιοτικό έλεγχο με στόχο τα δεδομένα να είναι όσο πιο αξιόπιστα γίνεται. Στην περίπτωση αυτή φαίνεται ότι το σημείο μεταφέρθηκε εντός εικονοστοιχείου το οποίο ανήκει ολόκληρος εντός του πολυγώνου αναφοράς από το CEMS.</p>	

Πέραν του ζητήματος πλήρους χωρικής επικάλυψης -όταν αυτό ήταν δυνατό- των δεδομένων αληθείας με τα πολύγωνα αναφοράς, κρίθηκε σημαντικό να γίνει και ένας έλεγχος σχετικά με τα φασματικά χαρακτηριστικά των εικονοστοιχείων ανάλογα με την κλάση στην οποία ανήκουν. Επεξηγηματικά, έγινε υπέρθεση των ground truth σε κατάλληλα έγχρωμα σύνθετα και εικόνες φασματικών δεικτών προκειμένου να διαπιστωθεί ο φασματικός διαχωρισμός των εκάστοτε κλάσεων.

Πίνακας 4: Ενδεικτικές περιπτώσεις ελέγχου των δεδομένων αληθείας της κλάσης Not Burned βάσει φασματικών χαρακτηριστικών.

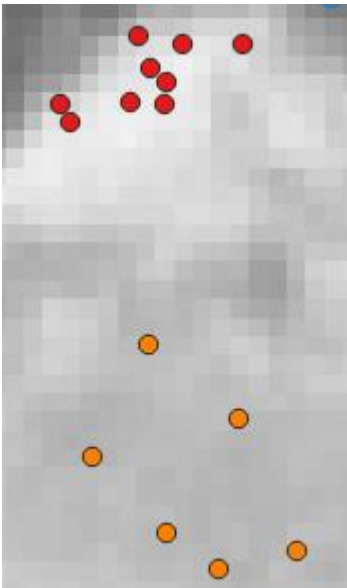
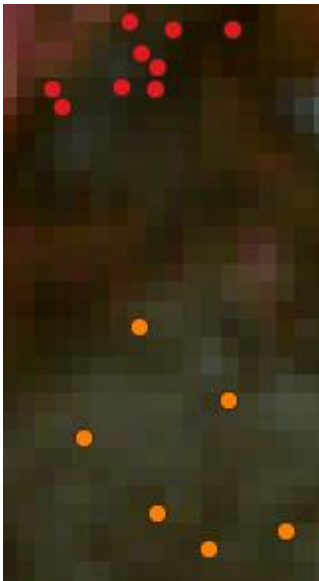
Δεδομένα αληθείας για κλάση 1: Not_Burned	
	
<p>Στα παραπάνω αποσπάσματα παρατηρούνται σημειακά δεδομένα αληθείας για την κλάση Not Burned πάνω στο έγχρωμο σύνθετο NIR-R-G της εικόνας μετά το γεγονός της πυρκαγιάς. Τα κόκκινα pixels υποδηλώνουν την ύπαρξη βλάστησης και επομένως βεβαιώνεται οπτικά ότι τα σημεία δεν αντιπροσωπεύουν καμένη έκταση.</p>	
	
<p>Στις περιπτώσεις αυτές πρόκειται επίσης για ground truth της κλάσης 1 που όμως δεν αντιπροσωπεύουν βλάστηση αλλά καλλιέργειες ή γυμνό έδαφος (χώμα). Το έγχρωμο σύνθετο πάνω στο οποίο έχουν προβληθεί είναι το φυσικό (R-G-B) και μπορεί να επιβεβαιωθεί ότι δεν πρόκειται για καμένη έκταση.</p>	
	
<p>Στα αποσπάσματα αυτά παρουσιάζονται σε φυσικό έγχρωμο σύνθετο δεδομένα αληθείας της κλάσης 1 που αντιπροσωπεύουν στοιχεία του δομημένου περιβάλλοντος καθώς στα δεδομένα αναφοράς από το CEMS υπάρχουν κάποια τέτοια λίγα τμήματα. Η υπέρθεση των δεδομένων έχει γίνει επίσης στο φυσικό έγχρωμο σύνθετο.</p>	

Πίνακας 5: Ενδεικτική περίπτωση αποκλεισμού σημειακών δεδομένων αληθείας.

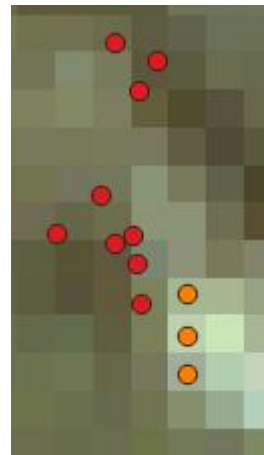

<p>Στο παραπάνω απόσπασμα φαίνονται σημειακά δεδομένα αληθείας που δημιουργήθηκαν με τον αυτόματο τρόπο που περιγράφηκε προηγουμένως και τα οποία κρίθηκαν ως μη κατάλληλα να συμμετέχουν στη διαδικασία εκπαίδευσης και ελέγχου καθώς βρίσκονται στο όριο στεριάς-θάλασσας, με ορισμένα από αυτά να βρίσκονται εντός θάλασσας. Τα σημεία αυτά διαγράφηκαν και αντί αυτών δημιουργήθηκαν με χειροκίνητο τρόπο νέα.</p>

Στον παρακάτω πίνακα επιλέχθηκε να παρουσιαστούν μερικά στιγμιότυπα από τα δεδομένα αληθείας για την κλάση Damaged συγκριτικά με αυτά της κλάσης Destroyed προκειμένου να γίνουν πιο εύκολα αντιληπτές οι διαφορές τους, δεδομένου ότι και οι δύο αντιπροσωπεύουν καμένες εκτάσεις. Η διαφοροποίηση αυτών έγκειται στο βαθμό καταστροφής που έχει επέλθει στη φυσική κάλυψη γης. Με πορτοκαλί χρώμα αποδίδονται τα δεδομένα αληθείας για την κλάση Damaged ενώ με κόκκινο για την Destroyed.

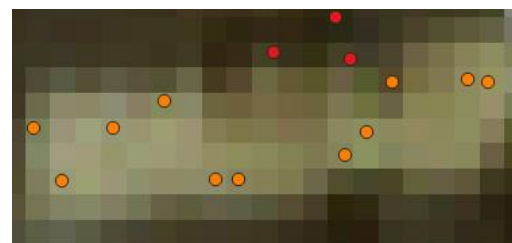
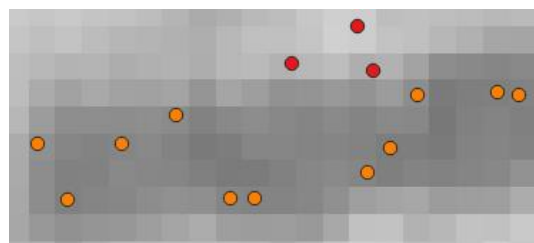
Πίνακας 6: Ενδεικτικές περιπτώσεις ελέγχου των δεδομένων αληθείας για τις κλάσεις Damaged και Destroyed βάσει φασματικών χαρακτηριστικών.

Δεδομένα αληθείας για κλάση 2: Damaged και 3: Destroyed	
	

Στο εν λόγω απόσπασμα έχει γίνει υπέρθεση των ground truth στην εικόνα του δείκτη dNBR (αριστερά). Έχει επιλεγθεί το συγκεκριμένο υπόβαθρο λαμβάνοντας υπόψη ότι ο δείκτης αυτός είναι από τους πιο ευρέως χρησιμοποιούμενους για burn severity mapping αλλά και επειδή είναι πιο εμφανής η διαφοροποίηση των εικονοστοιχείων. Στη συγκεκριμένη περίπτωση παρ' όλα αυτά, και στο έγχρωμο σύνθετο NIR-R-G φαίνεται ότι τα εικονοστοιχεία που αντιπροσωπεύουν την κλάση Destroyed είναι πιο σκούρα από αυτά της κλάσης Damaged. Αντίθετα, επειδή ο δείκτης dNBR για την κλάση Destroyed έχει υψηλότερες τιμές και λόγω του επιλεγμένου συμβολισμού, τα εικονοστοιχεία αυτής της κλάσης είναι πιο ανοιχτά από αυτά της Damaged.



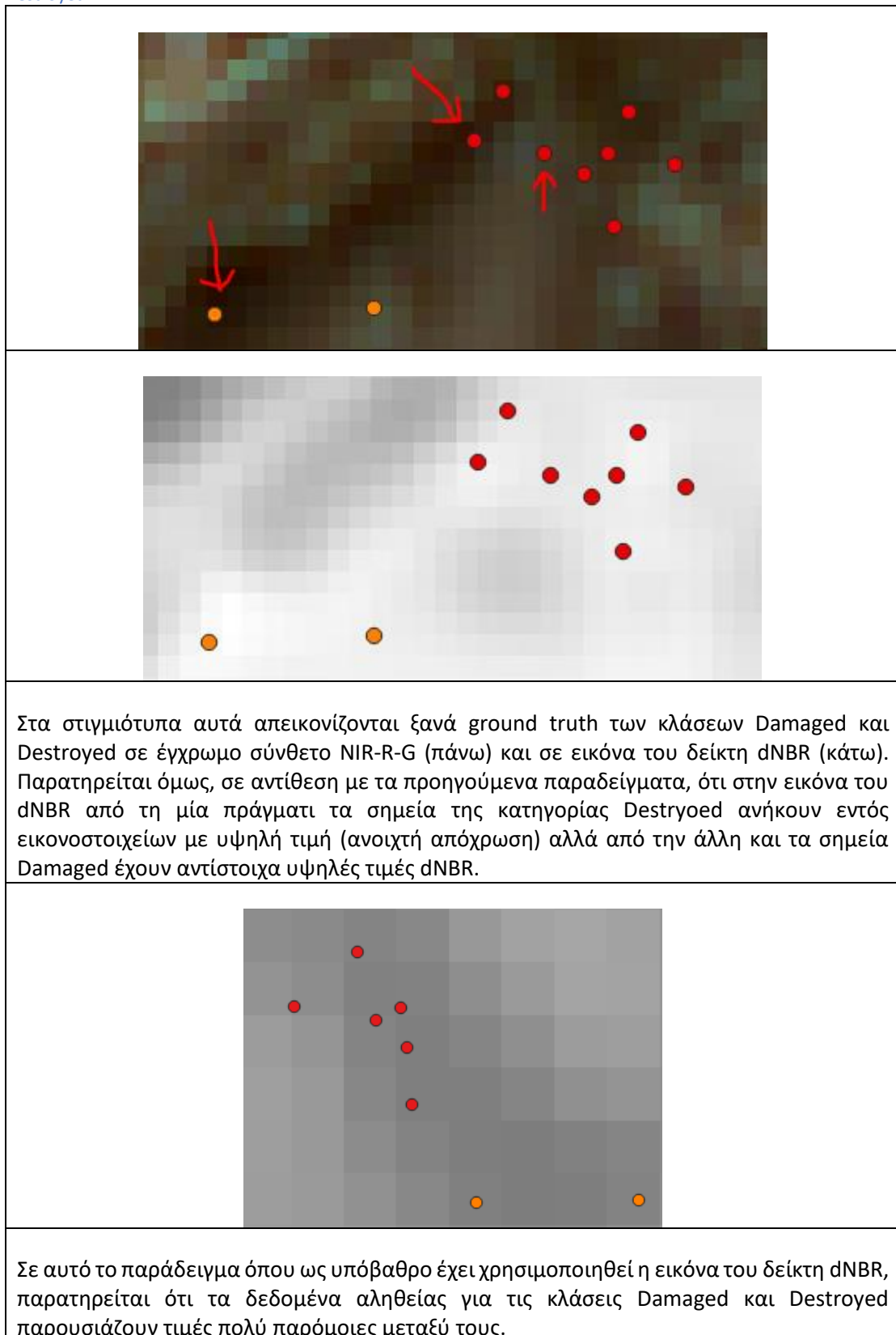
Σε αυτά τα αποσπάσματα παρατηρείται επίσης ότι τα δεδομένα αληθείας για την κλάση Damaged στην απεικόνιση του dNBR (αριστερά) αντιστοιχούν σε πιο σκούρα εικονοστοιχεία -λόγω χαμηλότερης τιμής- ενώ στο φυσικό έγχρωμο σύνθετο σε πιο ανοιχτά.



Αντίστοιχα και σε αυτή την περίπτωση είναι εμφανές ότι τα δεδομένα αληθείας για την κλάση Damaged στην περίπτωση της εικόνας του δείκτη dNBR έχουν πιο σκούρα απόχρωση, το οποίο βάσει του συμβολισμού που έχει επιλεγθεί υποδηλώνει και χαμηλότερες τιμές ενώ αντίθετα στην περίπτωση του ψευδέγχρωμου σύνθετου NIR-R-G έχουν πιο ανοιχτή απόχρωση.

Έχει αξία να αναφερθεί το γεγονός ότι παρατηρήθηκαν και ορισμένες περιπτώσεις φασματικής σύγχυσης μεταξύ των παραπάνω κατηγοριών (Damaged και Destroyed). Το φαινόμενο αυτό αναλύεται εκτενώς στη συνέχεια της εργασίας, τόσο κατά την περιγραφή των μεθοδολογιών όσο και κατά την αξιολόγηση των αποτελεσμάτων. Έτσι λοιπόν θεωρήθηκε σκόπιμο να παρουσιαστούν και τέτοια παραδείγματα καθώς η παρατήρηση αυτή είναι σημαντική για την μετέπειτα αξιολόγηση και ερμηνεία των αποτελεσμάτων και των δύο μεθοδολογιών που σχεδιάστηκαν.

Πίνακας 7: Ενδεικτικές περιπτώσεις σύγχυσης μεταξύ φασματικών χαρακτηριστικών των κλάσεων Damaged και Destroyed.



4 | Μεθοδολογία

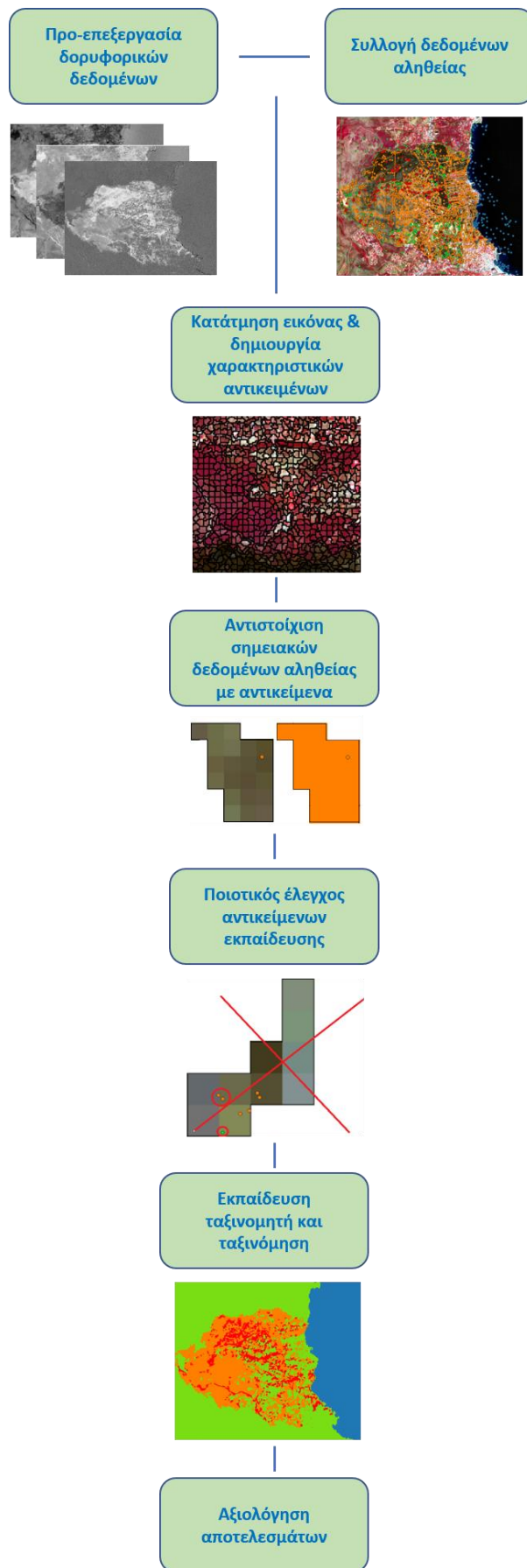
Σε αυτό το μέρος της εργασίας παρουσιάζονται αναλυτικά οι δύο διαφορετικές αντικειμενοστρεφείς μέθοδοι για τη χαρτογράφηση καμένων εκτάσεων και κατηγοριοποίησης του βαθμού καταστροφής τους. Πρώτα περιγράφεται η λύση με ανάπτυξη αλγορίθμου αντικειμενοστρεφούς ταξινόμησης σε γλώσσα Python και στη συνέχεια η αντίστοιχη μεθοδολογία που ακολουθήθηκε στο λογισμικό eCognition.

4.1 Ανάπτυξη αλγορίθμου αντικειμενοστρεφούς ταξινόμησης δορυφορικών εικόνων σε γλώσσα Python

Η open source λύση για την χαρτογράφηση καμένων εκτάσεων και την εκτίμηση του βαθμού καταστροφής τους υλοποιήθηκε αναπτύσσοντας κώδικα στη γλώσσα Python. Ο αλγόριθμος ακολουθεί τα βήματα της αντικειμενοστρεφούς ανάλυσης εικόνας (Object-Based Image Analysis – OBIA). Συνοπτικά, αφού γίνει κατάτμηση της δορυφορικής εικόνας και υπολογιστούν τα στατιστικά χαρακτηριστικά των αντικειμένων, ορίζονται τα αντικείμενα εκπαίδευσης, γίνεται εκπαίδευση του ταξινομητή και τέλος εφαρμόζεται η πρόβλεψη. Οι κλάσεις στις οποίες ταξινομείται η εικόνα είναι κοινές και στις δύο μεθόδους και όπως έχει αναφερθεί και στο προηγούμενο κεφάλαιο είναι οι εξής:

1. Not Burned
2. Damaged
3. Destroyed
4. Water

Κρίνεται σκόπιμο, προτού γίνει αναλυτική περιγραφή των επιμέρους αλγορίθμων και παραμέτρων για τα διαφορετικά βήματα υλοποίησης της εφαρμογής, να παρουσιαστεί μία εικόνα του διαγράμματος ροής (βλ. Εικόνα 8). Σημειώνεται ότι βάση για την ανάπτυξη του αλγορίθμου αποτέλεσε ο κώδικας από τον (Hafen, 2020).



Εικόνα 8: Διάγραμμα ροής αλγορίθμου στη γλώσσα Python.

4.1.1 Κατάτμηση εικόνας

Πρώτο στάδιο του αλγορίθμου σε προγραμματιστικό περιβάλλον Python είναι η κατάτμηση της τελικής δορυφορικής εικόνας όπως αυτή δημιουργήθηκε έπειτα από τις προεπεξεργασίες των πρωτογενών δορυφορικών εικόνων Sentinel-2 που περιγράφηκαν αναλυτικά στην ενότητα 3.2. Για την κατάτμηση αξιοποιήθηκε ο αλγόριθμος Simple Linear Iterative Clustering (SLIC) ο οποίος αναπτύχθηκε από τους (Achanta, et al., 2012). Συνοπτικά, ο αλγόριθμος SLIC υιοθετεί τη λογική του k-means clustering και υπολογίζει ένα μέγεθος απόστασης μεταξύ των εικονοστοιχείων λαμβάνοντας υπόψη την ευκλείδεια απόσταση της φασματικής πληροφορίας αλλά και τη χωρική απόσταση, δηλαδή την απόσταση μεταξύ των συγκρινόμενων εικονοστοιχείων. Οι υπολογισμοί αυτοί γίνονται με επαναληπτική διαδικασία ως προς τα κέντρα των αντικειμένων, τα οποία αρχικοποιούνται σε ίσες αποστάσεις μεταξύ τους βάσει του αριθμού των αντικειμένων που ορίζεται. Έτσι, στην αρχή της επαναληπτικής διαδικασίας τα αντικείμενα είναι ορθογώνια και στην πορεία, όπου τα κέντρα των αντικειμένων μετακινούνται προκειμένου να βρεθεί η βέλτιστη απόσταση, το σχήμα τους προσεγγίζει περισσότερο τα σχήματα των πραγματικών αντικειμένων που απεικονίζονται στην εικόνα. Ο αλγόριθμος αυτός επιλέχθηκε λαμβάνοντας υπόψη την πολύ γρήγορη και αρκετά αξιόπιστη παραγωγή αποτελεσμάτων.

Κατά τη χρήση του αλγορίθμου αυτού μέσω της βιβλιοθήκης `sci-kit image`⁷ (`skimage.segmentation.slic`) ο χρήστης μπορεί να ορίσει μία σειρά παραμέτρων που καθορίζουν το αποτέλεσμα. Οι παράμετροι που επιλέχθηκαν για την εφαρμογή αυτή σχετίζονται i) με τον προσεγγιστικό αριθμό των αντικειμένων που θα δημιουργηθούν (`n_segments`), ii) το συντελεστή *compactness* που καθορίζει το βάρος των γεωμετρικών χαρακτηριστικών έναντι του χρώματος, δηλαδή των φασματικών χαρακτηριστικών. Επεξηγηματικά, μεγαλύτερες τιμές της παραμέτρου αυτής δίνουν έμφαση στην χωρική απόσταση, κάνοντας τα αντικείμενα πιο τετράγωνα.

Στη συνέχεια παρατίθενται και επεξηγούνται αποσπάσματα από τον αλγόριθμο στη γλώσσα Python τα οποία αποτελούν τον πυρήνα της διαδικασίας και παράλληλα παρουσιάζονται τα αντίστοιχα αποτελέσματα αυτών που εμφανίζονται ως μηνύματα προς τον χρήστη.

```
segm = slic(img_n, n_segments = 30000, compactness = 0.12)
segments completed in: 1.9607617855072021 seconds
Number of segments created: 21634
```

Όπως φαίνεται και από το παραπάνω απόσπασμα ο αριθμός των επιθυμητών αντικειμένων ορίστηκε 30.000 ενώ η παράμετρος `compactness` πήρε τιμή 0,12. Οι τιμές αυτές επιλέχθηκαν έπειτα από αρκετούς πειραματισμούς και λαμβάνοντας υπόψη ότι η ύπαρξη πολλών αντικειμένων και άρα μικρών, θα λειτουργήσει βοηθητικά στην ταξινόμηση της κλάσης Destroyed που αποτελεί πολύ μικρή έκταση της συνολικής εικόνας. Τα αντικείμενα που τελικά δημιουργήθηκαν είναι 21.634. Αφού λοιπόν γίνει η κατάτμηση των δεδομένων εισόδου και αποθηκευτεί το αποτέλεσμα σε αρχείο GeoTIFF, επόμενο βήμα είναι ο υπολογισμός ορισμένων χαρακτηριστικών των δημιουργηθέντων αντικειμένων προκειμένου να αξιοποιηθούν στη διαδικασία της εκπαίδευσης στη συνέχεια. Για το σκοπό αυτό

⁷ <https://scikit-image.org/docs/dev/api/skimage.segmentation.html?highlight=slic#slic>

δημιουργήθηκε η παρακάτω συνάρτηση (*def segment_Features*) με την οποία για κάθε αντικείμενο υπολογίζονται συγκεκριμένα στατιστικά στοιχεία ανά φασματικό κανάλι. Αυτό υλοποιείται με χρήση της συνάρτησης `stats.describe` της βιβλιοθήκης `scipy`⁸. Η τελευταία δίνει σαν αποτέλεσμα την μέγιστη και ελάχιστη τιμή του δεδομένου εισόδου, καθώς και τη μέση τιμή, τυπική απόκλιση, ασυμμετρία (*skewness*) και κυρτότητα. Το δεδομένο εισόδου σε αυτόν τον αλγόριθμο είναι ένας πίνακας που αντιπροσωπεύει τις τιμές που παίρνει κάθε εικονοστοιχείο εντός αντικειμένου σε όλα τα κανάλια (*segment_pixels*).

```
def segment_Features(segment_pixels):
    features = []
    npixels, nbands = segment_pixels.shape
    for b in range(nbands):
        stats = scipy.stats.describe(segment_pixels[:, b])
        band_stats = list(stats.minmax) + list(stats)[2:]
        #if a segment is only 1 pixel then the variance is not applicable so
        if npixels == 1:
            band_stats[3] = 0.0
        features += band_stats
    return features
```

Η παραπάνω συνάρτηση καλείται εντός δομής επανάληψης προκειμένου να υπολογίσει για κάθε αντικείμενο αυτά τα στατιστικά χαρακτηριστικά. Κατά την ίδια επανάληψη δημιουργείται μία ακόμα μεταβλητή (*pixels_per_segment*) στην οποία αποθηκεύονται σε μορφή πίνακα τα id των αντικειμένων και ο αριθμός των εικονοστοιχείων από τα οποία αυτά αποτελούνται. Επιπλέον, σε δύο νέες λίστες αποθηκεύτηκαν τα στατιστικά χαρακτηριστικά κάθε αντικειμένου (*objects*) και τα αντίστοιχα id τους (*objects_ids*).

```
objects = []
objects_ids = []

#Initialize an empty array to store the number of pixels per segment
pixels_per_segment = np.empty((0,2), int)

for id in segment_ids:
    segment_pixels = img_n[segm == id]
    pixels_per_segment = np.append(pixels_per_segment, np.array([[id,
segment_pixels.shape[0]]]), axis=0)
    print('Pixels in segment', id, segment_pixels.shape[0])
    object_Features = segment_Features(segment_pixels)
    objects.append(object_Features)
    objects_ids.append(id)
```

Κατά την διάρκεια εκτέλεσης των εντολών αυτών εμφανίζεται μήνυμα κατά το οποίο ο χρήστης ενημερώνεται για το αντικείμενο για το οποίο γίνονται οι υπολογισμοί κάθε φορά καθώς και τον αριθμό pixels που αυτό περιέχει. Για παράδειγμα από το παρακάτω στιγμιότυπο φαίνεται ότι το αντικείμενο με id 21628 αποτελείται από 8 εικονοστοιχεία (*Pixels in segment 21628 8*). Τελικά εμφανίζεται μήνυμα για τον αριθμό των αντικειμένων

⁸ <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.describe.html>

(objects) που δημιουργήθηκαν, ο οποίος προφανώς ισούται με τον αριθμό που εμφανίζεται στο τέλος της κατάτμησης (segments), αλλά και για το σύνολο των χαρακτηριστικών για το καθένα που είναι 114 (6 στατιστικά μεγέθη X 19 κανάλια). Όπως φαίνεται και στο απόσπασμα παρακάτω, εμφανίζεται το μήνυμα “Το άθροισμα των εικονοστοιχείων ανά τμήμα ισούνται με τον συνολικό αριθμό εικονοστοιχείων της εικόνας”. Αυτό προκύπτει ως αποτέλεσμα ενός λογικού ελέγχου που έχει ενσωματωθεί στο τέλος της εν λόγω διαδικασίας προκειμένου να επαληθεύεται η ορθότητα υπολογισμού των εικονοστοιχείων ανά αντικείμενο.

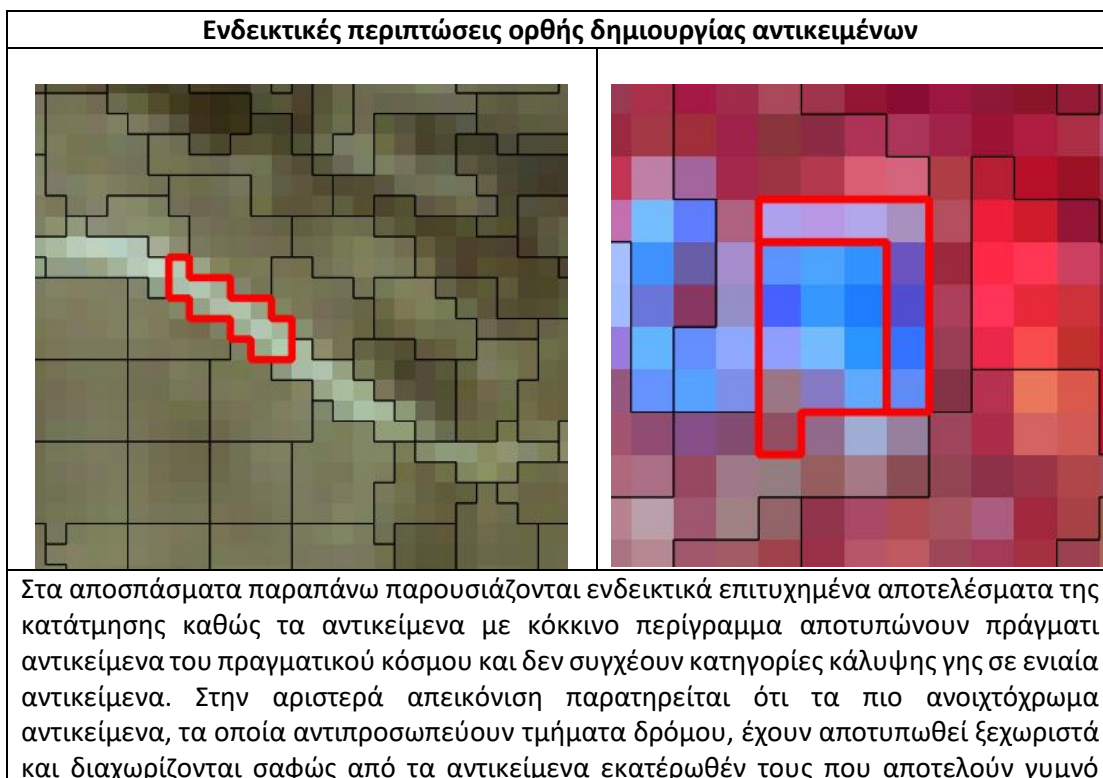
```

Pixels in segment 21627 8
Pixels in segment 21628 8
Pixels in segment 21629 13
Pixels in segment 21630 18
Pixels in segment 21631 12
Pixels in segment 21632 11
Pixels in segment 21633 11
Created 21634 objects with 114 variables in 291.8900353908539 seconds
Sum of pixels per segment equals the total number of image pixels

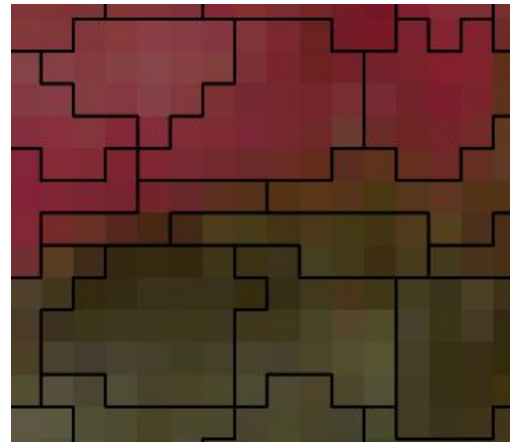
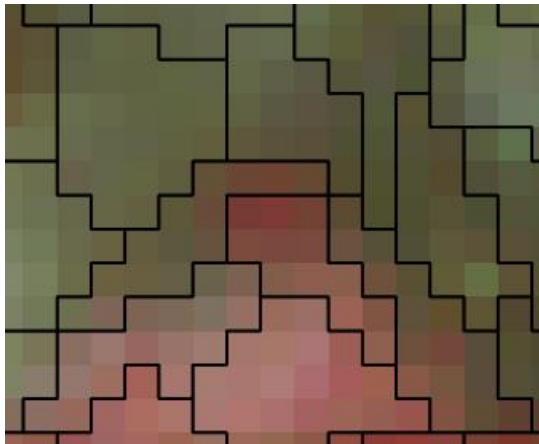
```

Στο σημείο αυτό ολοκληρώνονται οι επεξεργασίες που σχετίζονται άμεσα με την κατάτμηση των δορυφορικών δεδομένων και στην επόμενη υποενότητα αναλύεται το μέρος του αλγορίθμου που σχετίζεται με την εκπαίδευση του ταξινομητή και τελικά την ίδια την ταξινόμηση. Κρίνεται σκόπιμο στο σημείο αυτό να σχολιαστούν τα αποτελέσματα της κατάτμησης και να παρουσιαστούν ενδεικτικά παραδείγματα ορθής δημιουργίας αντικειμένων αλλά και αστοχιών κατά την κατάτμηση.

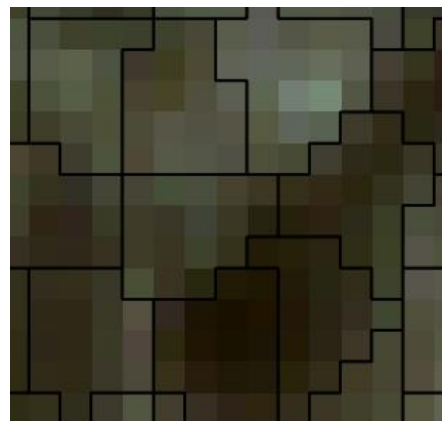
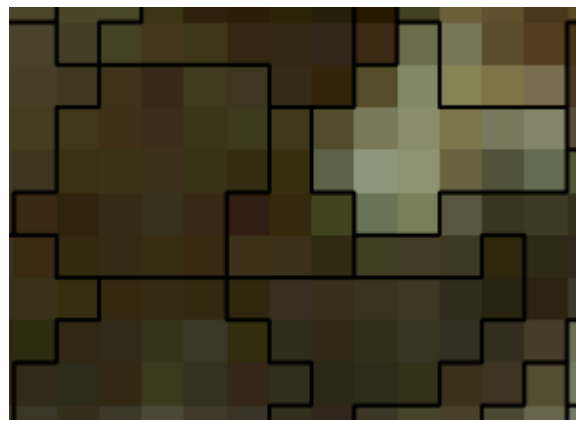
Πίνακας 8: Ενδεικτικές περιπτώσεις ορθής δημιουργίας αντικειμένων με τον αλγόριθμο SLIC.



έδαφος. Στην δεξιά εικόνα παρατηρούνται δύο αντικείμενα που αντιπροσωπεύουν πισίνα και έχουν διαχωρισθεί επαρκώς από τα γειτονικά τους εικονοστοιχεία που αντιπροσωπεύουν βλάστηση. Αυτά τα παραδείγματα αποδεικνύουν ότι πράγματι όταν παρατηρείται φασματική διαφοροποίηση των εικονοστοιχείων το σχήμα των αντικειμένων προσαρμόζεται και δεν παραμένει ορθογώνιο, όπως περιγράφηκε νωρίτερα κατά την περιγραφή του αλγορίθμου SLIC. Βέβαια παρατηρούνται αρκετά ορθογώνια αντικείμενα, κυρίως στο νοτιοδυτικό τμήμα της αριστερής εικόνας, στα οποία όμως πραγματικά δεν φαίνεται αξιοσημείωτη φασματική διαφοροποίηση.

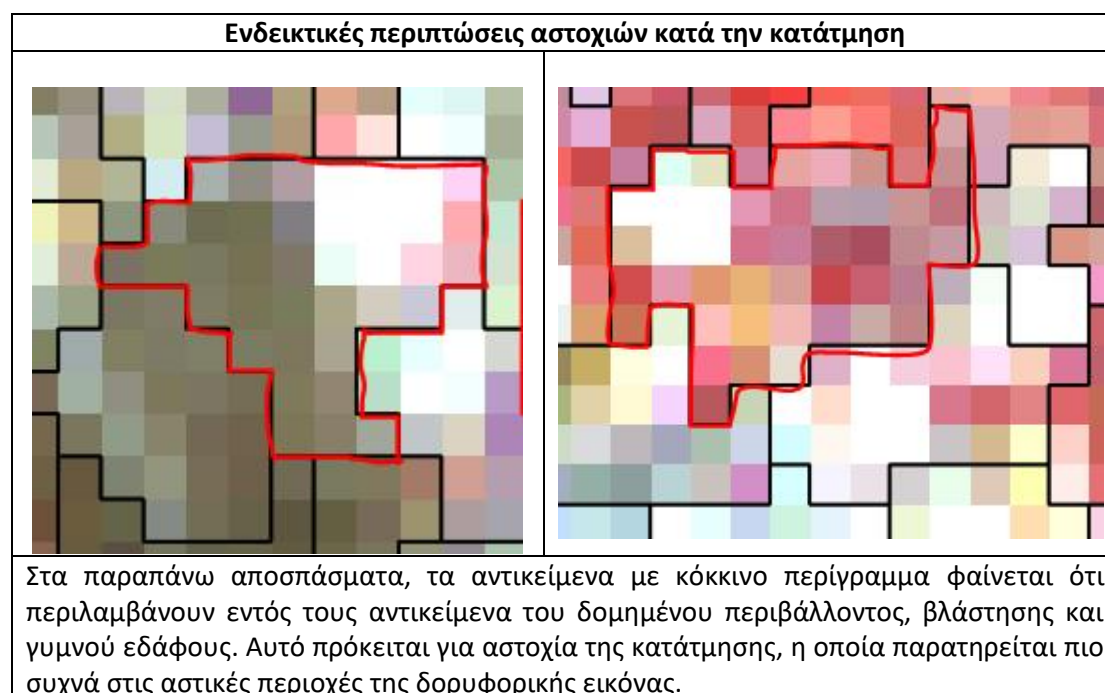


Σε αυτές τις απεικονίσεις παρουσιάζονται αποτελέσματα της κατάτμησης στα όρια καμένων εκτάσεων. Στο αριστερό στιγμιότυπο παρατηρείται ότι το βόρειο τμήμα αποτελεί καμένη εκτάση ενώ στο νότιο υπάρχει βλάστηση, την οποία υποδεικνύει το ροζ χρώμα καθώς έχει επιλεγθεί το έγχρωμο σύνθετο NIR-R-G. Φαίνεται και εδώ λοιπόν ότι το σχήμα των αντικειμένων έχει προσαρμοστεί και έτσι ο διαχωρισμός των εικονοστοιχείων σε αντικείμενα διατηρεί σε ικανοποιητικό βαθμό τα όρια των πραγματικών αντικειμένων της δορυφορικής εικόνας. Αντίστοιχη είναι και η περίπτωση της δεξιάς απεικόνισης με τη διαφορά ότι η καμένη έκταση βρίσκεται στο νότιο τμήμα ενώ η βλάστηση στο βόρειο.



Στα στιγμιότυπα παραπάνω απεικονίζονται αντικείμενα που όλα αντιπροσωπεύουν καμένες εκτάσεις, όμως φαίνεται ότι τα όρια των αντικειμένων σέβονται τη διαβάθμιση του βαθμού καταστροφής. Αυτό προκύπτει επειδή παρατηρείται ότι τα πολύ σκούρα αντικείμενα τα οποία αντιστοιχούν σε μεγαλύτερο βαθμό καταστροφής διαχωρίζονται από τα πιο ανοιχτά που αντιστοιχούν σε χαμηλότερο βαθμό.

Πίνακας 9: Ενδεικτικές περιπτώσεις αστοχιών κατά την κατάτμηση με τον αλγόριθμο SLIC.



Συνοπτικά, βάσει και των παραπάνω παραδειγμάτων σημειώνεται ότι τα αντικείμενα που δημιουργούνται καταφέρνουν να διαχωρίσουν σε ικανοποιητικό βαθμό τις καμένες εκτάσεις αλλά και μεταξύ αυτών τις διαβαθμίσεις τους. Οι μεγαλύτερες αστοχίες παρατηρούνται κυρίως στις αστικές περιοχές της εικόνας, γεγονός που δικαιολογείται εν μέρει λόγω του μεγέθους των εικονοστοιχείων που είναι 10m και έτσι δεν είναι εύκολος ο διαχωρισμός κτιρίων από το γύρω περιβάλλον. Δεδομένου ότι ο στόχος της εργασίας είναι η χαρτογράφηση των καμένων εκτάσεων και η εκτίμηση του βαθμού καταστροφής τους, κρίθηκε ότι οι αστοχίες αυτές δεν επηρεάζουν σημαντικά το αποτέλεσμα και έτσι η κατάτμηση αυτή αποτέλεσε τη βάση για την τελική ταξινόμηση της εικόνας. Είναι σημαντικό να διευκρινιστεί ότι αστοχίες παρατηρούνται και σε ορισμένες περιπτώσεις αντικειμένων καμένων εκτάσεων, αλλά αυτές περιορίζονται στην ύπαρξη μεμονωμένων “λανθασμένων” εικονοστοιχείων εντός των εν λόγω αντικειμένων.

4.1.2 Ταξινόμηση εικόνας

Πριν την εκπαίδευση του ταξινομητή και την πρόβλεψη απαραίτητο βήμα είναι ο μετασχηματισμός των δεδομένων εκπαίδευσης. Επεξηγηματικά, τα δεδομένα αυτά αρχικά είναι διαθέσιμα ως διανυσματικά αρχεία με γεωμετρία σημείου. Προκειμένου όμως να ενταχθούν στη διαδικασία εκπαίδευσης πρέπει να είναι σε αντίστοιχη μορφή με την κατατημένη εικόνα. Επομένως πρώτο βήμα προς αυτή την κατεύθυνση είναι η κανονικοποίηση (rasterization) αυτών των σημειακών διανυσματικών δεδομένων. Αφού εισαχθεί στον αλγόριθμο το αντίστοιχο αρχείο shapfile μέσω της βιβλιοθήκης Geopandas⁹, γίνεται υπολογισμός των σημείων εκπαίδευσης ανά κλάση και εμφανίζεται σχετικό μήνυμα όπως φαίνεται και παρακάτω. Το μήνυμα αυτό βοηθά στο να διατηρεί ο χρήστης μία γενική εποπτεία των δεδομένων και να μπορεί να διακρίνει τυχόν σφάλματα.

⁹ <https://geopandas.org/en/stable/>

```
#Count the training points per class from the shp
train_s = train_gdf['id'].value_counts()
train_s = train_s.sort_index(ascending = True)
print('The training samples (points) per class are:', train_s)
```

```
The training samples (points) per class are: 1    279
2     646
3     245
4     136
```

Αφού λοιπόν έχει εισαχθεί το διανυσματικό αρχείο γίνεται η κανονικοποίηση του με αξιοποίηση της βιβλιοθήκης GDAL και το παράγωγο αποθηκεύεται σε αρχείο GeoTIFF. Ιδιαίτερα σημαντικό σημείο στη διαδικασία αυτή είναι να οριστεί ορθά το πεδίο των περιγραφικών χαρακτηριστικών από το οποίο θα αποδοθεί η τιμή στα εικονοστοιχεία του νέου raster. Στην προκειμένη περίπτωση το πεδίο αυτό είναι το id και η τιμή αυτή αναπαριστά τις κλάσεις ταξινόμησης. Επίσης πρέπει να οριστεί ένα ήδη υπάρχον κανονικοποιημένο αρχείο, τα αρχικά δορυφορικά δεδομένα εν προκειμένω, βάσει του οποίου θα καθοριστεί ο γεωγραφικός μετασχηματισμός και η προβολή του νέου αυτού αρχείου. Συμπερασματικά δηλαδή, το αποτέλεσμα θα είναι μία νέα εικόνα, ίσων διαστάσεων με την αρχική, όπου κάθε εικονοστοιχείο στο οποίο αντιστοιχεί ένα σημείο εκπαίδευσης έχει τιμή ίση με την αντίστοιχη κλάση.

```
#Rasterize the train shp
train_ds = ogr.Open(train_fn)
#Crate a Layer containing the training data
lyr = train_ds.GetLayer()
#driverTiff = gdal.GetDriverByName('GTiff')
#Set the directory and the filename for the new raster file
target_fn = 'D:/BACK_UP_LAPTOP/GEOINFORMATICS/Thesis/...'

#New raster: rasterized train.shp
target_ds = driverTiff.Create(target_fn, img.RasterXSize, img.RasterYSize, 1,
gdal.GDT_UInt16)
target_ds.SetGeoTransform(img.GetGeoTransform())
target_ds.SetProjection(img.GetProjection())
gdal.RasterizeLayer(target_ds, [1], lyr, options = ["ATTRIBUTE=id"])
data = target_ds.GetRasterBand(1).ReadAsArray()
target_ds.GetRasterBand(1).WriteArray(data)
target_ds = None
```

Μετά την ολοκλήρωση αυτού του βήματος έχει ενσωματωθεί στον αλγόριθμο ένας έλεγχος κατά τον οποίο υπολογίζονται η ελάχιστη και η μέγιστη τιμή του raster αρχείου των δεδομένων εκπαίδευσης προκειμένου να διαπιστωθεί αν η μετατροπή έχει γίνει σωστά. Το μήνυμα που εμφανίζεται στον χρήστη είναι το παρακάτω:

```
min 0 max 4 mean 0.006855791962174941
[0 1 2 3 4]
```


Να σημειωθεί ότι σε όσα εικονοστοιχεία δεν εμπίπτει σημειακό δεδομένο εκπαίδευσης αυτά λαμβάνουν μηδενική τιμή. Ένα φαινόμενο το οποίο παρατηρήθηκε στην εφαρμογή αυτή και είναι γενικά πιθανό να συμβεί σε αντίστοιχες, είναι ότι κάποια σημεία εκπαίδευσης αντιστοιχούν στο ίδιο εικονοστοιχείο της εικόνας. Μάλιστα, οι πιθανότητες εμφάνισης του φαινομένου αυτού σε περιπτώσεις αυτόματης συλλογής δεδομένων εκπαίδευσης και ιδίως όταν η δορυφορική εικόνα δεν είναι πολύ υψηλής ανάλυσης, είναι αυξημένες. Κάτι τέτοιο μπορεί να οδηγήσει σε σημαντική μείωση των δεδομένων εκπαίδευσης και έτσι κρίθηκε σημαντικό να ενταχθεί στον αλγόριθμο ο υπολογισμός της ποσοτικής διαφοράς μεταξύ των διανυσματικών δεδομένων εκπαίδευσης και των κανονικοποιημένων. Έτσι, αν παρατηρηθεί ότι κατά το rasterization “χάνονται” πολλά δεδομένα τότε ο χρήστης μπορεί να παρέμβει και να τροποποιήσει το αρχικά διανυσματικά αρχείο μετακινώντας τα σημεία σε μεγαλύτερες αποστάσεις μεταξύ τους ώστε να αντιστοιχούν σε διαφορετικά εικονοστοιχεία των δορυφορικών δεδομένων.

Για να υλοποιηθεί αυτό πρέπει αρχικά να υπολογιστούν τα σημεία εκπαίδευσης από το shapfile και τα εικονοστοιχεία εκπαίδευσης ανά κλάση από το νέο κανονικοποιημένο αρχείο. Ο πρώτος υπολογισμός έχει ήδη πραγματοποιηθεί (*train_s*) κατά την εισαγωγή του shapfile των δεδομένων εκπαίδευσης όπως περιγράφεται στην αρχή της υποενότητας. Για τον δεύτερο υπολογισμό δημιουργείται μία δομή επανάληψης όπου για κάθε εικονοστοιχείο των rasterized δεδομένων εκπαίδευσης ελέγχεται η τιμή του και αν είναι διάφορη του μηδενός (0) τότε καταχωρείται σε μία λίστα (*train_pix_list*). Η λίστα ομαδοποιείται ανά τιμή, δηλαδή ανά κλάση, και τελικά μετατρέπεται σε πίνακα (*train_pixels_per_class*) προκειμένου στη συνέχεια να υπολογιστεί η διαφορά (*pts_pix_dif_class*).

```
#Calculate the training pixels per class (from rasterized shp)
rasterized_fn = target_fn
r = gdal.Open(rasterized_fn)

train_pixels = r.GetRasterBand(1).ReadAsArray()
#Get the rows and columns of the rasterized training samples (same size as the initial dataset)
rows = train_pixels.shape[0]
cols = train_pixels.shape[1]
#Iterate over the rasterized training samples to count the training pixels per class
s = 0
train_pix_list = []
for r in range(rows):
    for c in range(cols):
        s += 1
        if train_pixels[r,c] != 0:
            train_pix_list.append(train_pixels[r,c])

results = {value: len(list(freq)) for value, freq in groupby(sorted(train_pix_list))}

data = list(results.items())
train_pixels_per_class = np.array(data)

for i in range(train_pixels_per_class.shape[0]):
    print('The training pixels per class:', i+1, 'are', train_pixels_per_class[i,1])
```

```
#Calculate if there is any difference of the initial point samples and the rasterized
pts_pix_dif = np.subtract(train_pts_per_class[:, 1], train_pixels_per_class[:, 1])
pts_pix_dif_class = np.column_stack((class_ids, pts_pix_dif))
```

```
The training pixels per class: 1 are 219
The training pixels per class: 2 are 527
The training pixels per class: 3 are 245
The training pixels per class: 4 are 136
```

Στην εικόνα παρακάτω παρουσιάζονται το πλήθος των σημειακών δεδομένων εκπαίδευσης ανά κλάση συγκριτικά με αυτό των εικονοστοιχείων εκπαίδευσης μετά την κανονικοποίηση. Παρατηρείται ότι για τις κλάσεις Not Burned (1) και Damaged (2) υπάρχει μείωση των δεδομένων εκπαίδευσης ενώ τα δεδομένα για τις κλάσεις Destroyed (3) και Water (4) μένουν αναλλοίωτα. Κρίθηκε ότι η μείωση δεν είναι ιδιαίτερα σημαντική αφενός για την κλάση Damaged λόγω του ήδη μεγάλου αριθμού δεδομένων (από 646 έμειναν 527) αλλά και της επάρκειας του συνολικού αριθμού εικονοστοιχείων εκπαίδευσης όπως υπολογίζεται στο επόμενο στάδιο κατά το οποίο γίνεται αντιστοίχιση pixels και αντικειμένων.

train_pts_per_class		train_pixels_per_class	
1	279	1	219
2	646	2	527
3	245	3	245
4	136	4	136

Εικόνα 9: Σύγκριση σημειακών δεδομένων εκπαίδευσης ανά κλάση (αριστερά) και εικονοστοιχείων εκπαίδευσης (δεξιά) μετά την κανονικοποίηση του διανυσματικού αρχείου.

Ο μετασχηματισμός των δεδομένων εκπαίδευσης δεν έχει ολοκληρωθεί. Λαμβάνοντας υπόψη ότι μετά την κατάτμηση το μοναδιαίο στοιχείο της δορυφορικής εικόνας δεν είναι το εικονοστοιχείο αλλά το αντικείμενο, πρέπει τα εικονοστοιχεία εκπαίδευσης να μετατραπούν σε αντικείμενα εκπαίδευσης. Έτσι, τα στατιστικά χαρακτηριστικά τους θα μπορούν να αξιοποιηθούν στην εκπαίδευση του ταξινομητή. Για να γίνει αυτό δημιουργείται μία δομή επανάληψης εντός της οποίας κάθε αντικείμενο στο οποίο περιλαμβάνεται εικονοστοιχείο εκπαίδευσης αποθηκεύεται σε ένα σετ ανά κλάση. Τελικά δηλαδή δημιουργείται μία μεταβλητή (*segments_per_class*) η οποία περιέχει ανά κλάση τα id των αντικειμένων τα οποία θα χρησιμοποιηθούν για την εκπαίδευση όπως φαίνεται στην εικόνα 10.

```
#Part 5: Correspond the training data with segments and classes
#Exclude the segments that represent multiple classes
ground_truth = train_pixels
segments_per_class = {}
part5time = time.time()

for klass in class_ids:
    segments_of_class = segm[ground_truth == klass]
    segments_per_class[klass] = set(segments_of_class)
    print('Training segments for class', klass, ':', len(segments_of_class))
```

segments_per_class - Dictionary (4 elements)

set	169	{14851, 3590, 2056, 14857, 14347, 4109, 9742, 10770, 15380, 8213, ...}
set	372	{16393, 14346, 12305, 6162, 18452, 18453, 14385, 16437, 10299, 8265, ...}
set	159	{10756, 8202, 16395, 9741, 16916, 4631, 16926, 8737, 10797, 11822, ...}
set	136	{13313, 8710, 19975, 10764, 14868, 14869, 13334, 21016, 11298, 16424, ...}

Εικόνα 10: Μεταβλητή `segments_per_class` στην οποία αποθηκεύονται τα `id` των αντικειμένων εκπαίδευσης ανά κλάση.

Με την ίδια λογική που κατά την μετατροπή των δεδομένων εκπαίδευσης από διανυσματικό αρχείο σε κανονικοποιημένο “χάθηκαν” σημεία λόγω αντιστοίχισης στο ίδιο pixel, έτσι και εδώ κάποια εικονοστοιχεία εκπαίδευσης είναι πιθανό να αντιστοιχούν στο ίδιο αντικείμενο. Αν τα εικονοστοιχεία αντιπροσωπεύουν την ίδια κλάση τότε αυτό δεν επηρεάζει τη διαδικασία εκπαίδευσης αλλά αν εντός του ίδιου αντικειμένου εμπεριέχονται δεδομένα εκπαίδευσης από διαφορετικές κλάσεις τότε αυτά πρέπει να εξαιρούνται. Για να γίνει αυτό, δημιουργούνται δύο νέες μεταβλητές (*intersection*, *accum*) και για κάθε κλάση τα αντικείμενα εκπαίδευσής της (*segments_per_class*) στοιβάζονται στην μεταβλητή *accum*. Στην μεταβλητή *intersection* αποθηκεύεται όποιο αντικείμενο εκπαίδευσης έχει επικάλυψη με τη μεταβλητή *accum*, δηλαδή όποιο αντικείμενο έχει ήδη περάσει τον έλεγχο άρα ανήκει και σε άλλη κλάση. Έτσι, αν η μεταβλητή *intersection* παραμείνει κενή τότε κανένα αντικείμενο εκπαίδευσης δεν αντιπροσωπεύει παραπάνω από μία κλάσεις. Αν όμως αποθηκευτούν σε αυτή `id` αντικειμένων σημαίνει ότι τα αντικείμενα αυτά αντιπροσωπεύουν δύο διαφορετικές κλάσεις και άρα ο χρήσης πρέπει να ενημερώνεται ώστε στη συνέχεια να εξαιρεί τα αντικείμενα αυτά από την εκπαίδευση.

```
#Check if training segments represent multiple segments
intersection = set()
accum = set()

for class_segments in segments_per_class.values():
    print(class_segments)
    intersection |= accum.intersection(class_segments)
    accum |= class_segments
assert len(intersection) == 0, "Segment(s) represent multiple classes."
print(len(intersection), "segments represent more than one class and are going to be
excluded from the training.")
```

Το μήνυμα που εμφανίζεται στον χρήστη είναι το παρακάτω και όπως φαίνεται υπολογίζεται και ο αριθμός των αντικειμένων που πρέπει να εξαιρεθούν από τη διαδικασία εκπαίδευσης. Στην περίπτωση που δεν υπάρχει επικάλυψη μεταξύ αντικειμένων εκπαίδευσης διαφορετικών κλάσεων τότε δεν εμφανίζεται κάποιο μήνυμα. Σημειώνεται ότι για την επιλογή της τελικής κατάτμησης από τους διάφορους πειραματισμούς που έγιναν το αποτέλεσμα αυτού του ελέγχου αποτέλεσε σημαντικό κριτήριο δεδομένου ότι η συνύπαρξη σημείων εκπαίδευσης διαφορετικής κλάσης εντός ίδιου αντικειμένου υποδηλώνει μεγαλύτερη ετερογένεια αντικειμένων.

```
AssertionError: Segment(s) represent multiple classes.
```

```
In [78]: print(len(intersection), "segments represent more than one class and are going to be excluded from the training.")  
6 segments represent more than one class and are going to be excluded from the training.
```

Επόμενο βήμα λοιπόν είναι η απομάκρυνση των αντικειμένων αυτών από τη μεταβλητή **segments_per_class**. Δημιουργείται και εδώ μία επαναληπτική δομή που διατρέχει όλα τα id των αντικειμένων εκπαίδευσης για κάθε κλάση, δηλαδή τα στοιχεία του **segments_per_class** και ελέγχει αν υπάρχουν στη μεταβλητή **intersection** που περιγράφηκε προηγουμένως. Όποιο id περιέχεται και στο **intersection** σημαίνει ότι αντιπροσωπεύει περισσότερες από μία κλάσεις και έτσι διαγράφεται. Κατά την επανάληψη εμφανίζεται μήνυμα που προσδιορίζει ποια κλάση ελέγχεται τη δεδομένη στιγμή και ποια αντικείμενα εκπαίδευσης αυτής θα διαγραφούν.

```
#Exclude the segment_ids that represent more than one class from the segments_per_class
```

```
for klass in class_ids:  
    print('Loop for segments in class:', klass)  
    for segment_id in segments_per_class[klass].copy():  
        if segment_id in intersection:  
            print(klass, segment_id)  
            segments_per_class[klass].discard(segment_id)
```

```
Loop for segments in class: 1  
1 18014  
1 15099  
1 19783  
1 19923  
Loop for segments in class: 2  
2 14511  
2 11529  
2 19783  
2 19923  
2 18014  
2 15099  
Loop for segments in class: 3  
3 14511  
3 11529  
Loop for segments in class: 4
```

Από το παραπάνω μήνυμα φαίνεται ότι για τη δεδομένη εφαρμογή υπήρχαν τέσσερα αντικείμενα που περιείχαν δεδομένα εκπαίδευσης για τις κλάσεις Not Burned (1) και Damaged (2) και αντίστοιχα δύο για τις κλάσεις Damaged (2) και Destroyed (3). Στο τέλος αυτής της διαδικασίας τυπώνεται μήνυμα προς το χρήστη στο οποίο αναγράφεται για κάθε κλάση ο τελικός αριθμός των αντικειμένων εκπαίδευσης.

```
for klass in class_ids:  
    print('Training segments after quality control for class ', klass, ':',  
        len(segments_per_class[klass]))
```

```
Training segments after quality control for class 1 : 165
Training segments after quality control for class 2 : 366
Training segments after quality control for class 3 : 157
Training segments after quality control for class 4 : 136
```

Προκειμένου ο χρήστης να γνωρίζει το συνολικό πλήθος εικονοστοιχείων των αντικειμένων εκπαίδευσης δημιουργήθηκε μία επανάληψη κατά την οποία καταχωρείται σε μία μεταβλητή (*total_training_pixels*) το άθροισμα αυτό. Η πληροφορία του αριθμού των εικονοστοιχείων έχει καταγραφεί για όλα τα αντικείμενα σε προηγούμενο βήμα στην μεταβλητή *pixels_per_segment*. Μετά από αυτό υπολογίζεται τι ποσοστό της συνολικής εικόνας προς ταξινόμηση αντιπροσωπεύουν συνολικά τα εικονοστοιχεία εκπαίδευσης. Αυτή είναι μία χρήσιμη πληροφορία προς τον χρήστη προκειμένου να γνωρίζει αν τα δεδομένα εκπαίδευσης είναι επαρκή ή όχι. Όπως φαίνεται και παρακάτω, στην παρούσα εργασία ο συνολικός αριθμός των εικονοστοιχείων εκπαίδευσης αντιστοιχεί στο 4,18% της εικόνας.

```
#Calculate total number of pixels for the training segments
total_training_pixels = 0
```

```
for klass in class_ids:
    for segment_id in segments_per_class[klass]:
        total_training_pixels += pixels_per_segment[segment_id, 1]
print('The total number of pixels of the training segments is:',
total_training_pixels, 'which corresponds to the',
(total_training_pixels/img_pixels*100), '% of the image pixels')
```

```
The total number of pixels of the training segments is: 15564 which corresponds to the 4.181173436492585 % of the image pixels
```

Το επόμενο που πρέπει να γίνει είναι να αποδοθεί στα αντικείμενα εκπαίδευσης η κλάση την οποία αντιπροσωπεύει καθένα από αυτά. Ζητούμενο δηλαδή είναι η δημιουργία μίας εικόνας αντίστοιχης με το αποτέλεσμα της κατάτμησης αλλά αντί για την πληροφορία του id στα εικονοστοιχεία των αντικειμένων εκπαίδευσης να αποδίδεται η κλάση ταξινόμησης ενώ στα υπόλοιπα αντικείμενα η τιμή 0. Έτσι, αρχικά δημιουργείται ένα αντίγραφο της κατατμημένης εικόνας (*train_img*). Παράλληλα ορίζεται ένα κατώφλι (*threshold*) του οποίου η τιμή είναι μεγαλύτερη κατά ένα από το πλήθος των αντικειμένων που προέκυψαν από την κατάτμηση. Για κάθε κωδικό κλάσης δημιουργείται ένας νέος κωδικός (*class_label*) ο οποίος ισούται με τον αρχικό κωδικό (δηλαδή 1 ή 2 ή 3 ή 4) συν το κατώφλι αυτό. Έπειτα, για κάθε εικονοστοιχείο της εικόνας αντίγραφου (*train_img*) το οποίο ανήκει σε αντικείμενο εκπαίδευσης μεταβάλλεται η τιμή του και γίνεται ίση με το *class_label* της κλάσης που αντιπροσωπεύει. Με αυτό τον τρόπο όλα τα αντικείμενα εκπαίδευσης καταλήγουν να ξεχωρίζουν γιατί έχουν τιμές μεγαλύτερες από όλα τα υπόλοιπα αντικείμενα. Έτσι, σε επόμενο βήμα, σε όλα τα pixels που έχουν τιμή μικρότερη από το κατώφλι αποδίδεται μηδενική τιμή ενώ σε όσα έχουν μεγαλύτερη αφαιρείται η τιμή του κατωφλίου και άρα τους απομένει η αρχική κωδικοποίηση της κλάσης στην οποία ανήκουν.

```

##Part 6: Training and Classification of the segments
#Create an exact same image as the segmented
train_img = np.copy(seg)
#Threshold value in order to change only the value of the training segments in the
train_img
threshold = train_img.max() + 1

#Iterate through classes in order to assign new class label to the segments that
represent training segments of the classes
for klass in class_ids:
    class_label = threshold + klass
    print('class label', class_label)
    for segment_id in segments_per_class[klass]:
        train_img[train_img==segment_id] = class_label #class + threshold

#Segments that don't have training points in them --> 0 value
train_img[train_img <= threshold] = 0
#Segments that have training points --> class id (subtract the threshold)
train_img[train_img > threshold] -= threshold

```

Σε επόμενο στάδιο πρέπει τα αντικείμενα εκπαίδευσης να συνδεθούν με τα χαρακτηριστικά τους ώστε να τροφοδοτηθεί η διαδικασία εκπαίδευσης. Όπως περιγράφηκε στην υποενότητα 4.1.1, η μεταβλητή **objects** περιέχει για κάθε αντικείμενο τα στατιστικά χαρακτηριστικά του που έχουν υπολογιστεί με τη συνάρτηση `scipy.stats.describe`. Άρα γίνεται έλεγχος για κάθε κλάση και όσα στοιχεία του **objects** αποτελούν αντικείμενα εκπαίδευσης (**segments_per_class**) αποθηκεύονται οι τιμές τους, δηλαδή τα στατιστικά στοιχεία, στη μεταβλητή **class_train_object**. Παράλληλα στη μεταβλητή **training_lab** αποθηκεύεται ο κωδικός της εκάστοτε κλάσης τόσες φορές όσες το πλήθος των αντικειμένων **class_train_object**. Τέλος όλα τα αντικείμενα εκπαίδευσης αποθηκεύονται στη λίστα **training_obj**. Εμφανίζεται και εδώ μήνυμα που ενημερώνει τον χρήστη για το πλήθος των αντικειμένων εκπαίδευσης ανά κλάση με σκοπό να ελεγχθεί ότι αυτό είναι ίδιο με το προηγούμενο μήνυμα που εμφανίζεται μετά την εξαίρεση των αντικειμένων στα οποία εμπίπτουν σημεία εκπαίδευσης διαφορετικών κλάσεων.

```

#Create list that contains the objects for which the label is known (ground_truth)
and these will be used for training - training_obj
training_obj = []
#Create list that contains the labels of the training objects - training_lab
training_lab = []

for klass in class_ids:
    class_train_obj = [v for i, v in enumerate(objects) if segment_ids[i] in
segments_per_class[klass]]
    training_lab += [klass] * len(class_train_obj)
    training_obj += class_train_obj
    #Cross check the training objects per class (should equal the training
segments
per class after the quality control)
    print('Training objects per class', klass, len(class_train_obj))

```

```
Training objects per class 1 165
Training objects per class 2 366
Training objects per class 3 157
Training objects per class 4 136
```

Τελικό βήμα για την ταξινόμηση των δορυφορικών δεδομένων είναι η εισαγωγή του ταξινομητή Random Forest από την βιβλιοθήκη sci-kit learn και η εκπαίδευσή του χρησιμοποιώντας τις μεταβλητές **training_obj** και **training_lab** που δημιουργήθηκαν στο προηγούμενο βήμα. Καθώς οι μεταβλητές αυτές πρόκεινται για λίστες είναι απαραίτητη η μετατροπή του αποτελέσματος **pred** σε εικόνα όμοια με την κατατημένη που όμως αντί για τις τιμές των id των αντικειμένων θα αποτυπώνουν την κλάση στην οποία ταξινομήθηκε το εκάστοτε αντικείμενο.

Ο αλγόριθμος Random Forest συγκαταλέγεται στους αλγορίθμους Μηχανικής Μάθησης (Machine Learning-ML) και συγκεκριμένα στις μεθόδους ensemble καθώς χρησιμοποιεί ένα σύνολο δέντρων απόφασης (Decision Trees) για την ταξινόμηση. Κάθε δέντρο ψηφίζεται, ανάλογα με τη συμμετοχή στην κλάση του και το δέντρο αποφάσεων που λαμβάνει τις περισσότερες ψήφους χρησιμοποιείται για την πρόβλεψη της σχετικής κλάσης (Stumpf & Kerle, 2011). Ο αλγόριθμος Random Forest χρησιμοποιείται ευρέως σε εφαρμογές Τηλεπισκόπησης και γενικά χαρακτηρίζεται από απλότητα ενώ τα αποτελέσματά του είναι συνήθως πολύ καλά. Για τους λόγους αυτούς επιλέχθηκε να χρησιμοποιηθεί και στην παρούσα εφαρμογή.

```
#Import the classifier
from sklearn.ensemble import RandomForestClassifier
#Train the Random Forest Classifier
rf = RandomForestClassifier(bootstrap=False, n_jobs=-1, random_state=101)
rf.fit(training_obj, training_lab)
print('Fitting Random Forest Classifier -rf...')
#Predict
pred = rf.predict(objects)

#Create an exact same image as the segments
clf = np.copy(segms)
#Recreate the segmented image but assign the predicted labels to each segment
for segment_id, klass in zip(segment_ids, pred):
    clf[clf==segment_id] = klass
```

4.1.3 Αξιολόγηση ταξινόμησης

Σκοπός της υποενότητας αυτής δεν είναι η παρουσίαση και αξιολόγηση των αποτελεσμάτων αλλά η περιγραφή της μεθόδου υπολογισμού του πίνακα σύγχυσης και των άλλων μετρικών. Σημειώνεται ότι η παρουσίαση και ανάλυση των αποτελεσμάτων γίνεται εκτενώς στο Κεφάλαιο 5.

Η αξιολόγηση των αποτελεσμάτων αυτής της μεθόδου έγινε πρώτα σε επίπεδο εικονοστοιχείων (pixel-based) και έπειτα σε επίπεδο αντικειμένων (segment-based). Κατά την πρώτη αξιολόγηση εξετάζεται η ταξινόμηση μεμονωμένων εικονοστοιχείων. Έτσι, πρώτο βήμα είναι η κανονικοποίηση των δεδομένων ελέγχου (test data), τα οποία αρχικά είναι σε

διανυσματική μορφή. Η διαδικασία αυτή έχει ήδη περιγραφεί για τα δεδομένα εκπαίδευσης στην προηγούμενη υποενότητα. Και εδώ για λόγους εποπτείας των δεδομένων θεωρήθηκε σκόπιμο να υπολογιστεί ο αριθμός των εικονοστοιχείων ελέγχου ανά κλάση μετά το rasterization καθώς και του ποσοστού που αντιπροσωπεύουν συνολικά επί της δορυφορικής εικόνας.

```
#Calculate the test pixels from the rasterized test shp
rows = test.shape[0]
cols = test.shape[1]
s = 0
test_pix_list = []
for r in range(rows):
    for c in range(cols):
        s += 1
        if test[r,c] != 0:
            test_pix_list.append(test[r,c])

results = {value: len(list(freq)) for value, freq in groupby(sorted(test_pix_list))}

data = list(results.items())
test_pixels_per_class = np.array(data)

total_test_pixels = 0
for i in range(test_pixels_per_class.shape[0]):
    total_test_pixels += test_pixels_per_class[i,1]
    print('The test pixels per class:', i+1, 'are', test_pixels_per_class[i,1])
print('The total number of test pixels is:', total_test_pixels, 'which corresponds
to the', (total_test_pixels/img_pixels*100), '% of the image pixels')
```

```
The test pixels per class: 1 are 883
The test pixels per class: 2 are 857
The test pixels per class: 3 are 220
The test pixels per class: 4 are 220
The total number of test pixels is: 2180 which corresponds to the 0.5856436707500536 % of the image
pixels
```

Για την αξιολόγηση των αποτελεσμάτων αρχικά υπολογίζεται ο πίνακας σύγχυσης και στη συνέχεια μετρικές όπως η συνολική ακρίβεια (Overall Accuracy), η αξιοπιστία (User's Accuracy-UA), η ακρίβεια (Producer's Accuracy-PA) και ο συντελεστής kappa. Για τους υπολογισμούς αυτούς αξιοποιείται από τη βιβλιοθήκη sci-kit learn το module metrics. Επιπλέον βάσει του πίνακα σύγχυσης υπολογίστηκαν και τα σφάλματα τύπου I (Omission Errors) και τύπου II (Commission Errors) για κάθε κλάση. Τα αποτελέσματα, όπως αναφέρθηκε και προηγουμένως παρουσιάζονται στο επόμενο κεφάλαιο. Στη συνέχεια περιγράφεται ο τρόπος υπολογισμού αυτών στο προγραμματιστικό περιβάλλον της Python.

```
#Calculate the confusion matrix (pixel-based)
cm_pb = metrics.confusion_matrix(test[i], preds[i])
print(cm_pb)
#Calculate metrics pixel-based (pb)
#Calculate kappa
kappa_pb = metrics.cohen_kappa_score(test[i], preds[i])
#Calculate Overall Accuracy (OA)
oa_pb = metrics.accuracy_score(test[i], preds[i])
#Calculate Producer's Accuracy (PA) per class
```



```

pa_pb = metrics.recall_score(test[i], preds[i], average=None)
#Calculate average Producer's Accuracy
pa_pb_avg = metrics.recall_score(test[i], preds[i], average='weighted')
#Calculate User's Accuracy (UA) per class
ua_pb = metrics.precision_score(test[i], preds[i], average=None)
#Calculate average User's Accuracy (UA)
ua_pb_avg = metrics.precision_score(test[i], preds[i], average='weighted')
#Calculate F1 score per class
f1_pb = metrics.f1_score(test[i], preds[i], average=None)

```

Η αξιολόγηση των αποτελεσμάτων σε επίπεδο αντικειμένων (segment-based) συνεπάγεται τον έλεγχο ταξινόμησης ανά αντικείμενο και όχι ανά μεμονωμένο pixel. Για να γίνει αυτό πρέπει τα σημεία ελέγχου από τα δεδομένα αληθείας να αντιστοιχηθούν με τα αντικείμενα εντός των οποίων βρίσκονται. Ακριβώς αντίστοιχη διαδικασία πραγματοποιήθηκε για την συσχέτιση δεδομένων εκπαίδευσης και αντικειμένων και περιγράφηκε αναλυτικά στην υποενότητα 4.1.2. Και σε αυτή την περίπτωση πρέπει να γίνει έλεγχος για τυχόν αντικείμενα ελέγχου που αντιπροσωπεύουν περισσότερες από μία κλάσεις και αν βρεθούν τέτοια να αφαιρεθούν. Όπως φαίνεται και στο μήνυμα παρακάτω πράγματι εντοπίστηκαν εννέα (9) τέτοια αντικείμενα, τα οποία εξαιρέθηκαν από τη διαδικασία αξιολόγησης.

```

##Segment-based accuracy
#Correspond the test segments with the classes
test_segments_per_class = {}

for klass in class_ids:
    segments_of_class = segm[test == klass]
    test_segments_per_class[klass] = set(segments_of_class)
    print('Test segments for class', klass, ':', len(segments_of_class))

#Check if test segments represent multiple segments
intersection = set()
accum = set()

for class_segments in test_segments_per_class.values():
    print(class_segments)
    intersection |= accum.intersection(class_segments)
    accum |= class_segments
assert len(intersection) == 0, "Test segment(s) represent multiple classes."
print(len(intersection), "segments represent more than one class and are going to be excluded from the validation.")

```

```

AssertionError: Test segment(s) represent multiple classes.

In [94]: print(len(intersection), "segments represent more than one class and are going to be excluded from the validation.")
9 segments represent more than one class and are going to be excluded from the validation.

```

```

#Exclude test segments that represent multiple classes (If applicable)
for klass in class_ids:
    print('Loop for segments in class:', klass)
    for segment_id in test_segments_per_class[klass].copy():
        if segment_id in intersection:
            print(klass, segment_id)
            test_segments_per_class[klass].discard(segment_id)

```

Στο τέλος της διαδικασίας αντιστοίχισης δεδομένων ελέγχου και αντικειμένων, εντάχθηκε στον αλγόριθμο η εμφάνιση μηνύματος προς τον χρήστη σχετικά με τον τελικό αριθμό των αντικειμένων ελέγχου μετά και από την απομάκρυνση αυτών που περιείχαν σημεία ελέγχου διαφορετικών κλάσεων (**after quality control**).

```
#Get the training segments per class after discarding the segments that represent multiple classes
for klass in class_ids:
    print('Test segments after quality control for class ', klass, ':',
          len(test_segments_per_class[klass]))
```

```
Test segments after quality control for class 1 : 787
Test segments after quality control for class 2 : 749
Test segments after quality control for class 3 : 142
Test segments after quality control for class 4 : 219
```

Στη συνέχεια δημιουργείται μία λίστα που περιέχει για κάθε κλάση τα ταξινομημένα σε αυτή αντικείμενα (**clf_segments_per_class**). Από αυτή πρέπει να εξαιρεθούν τα αντικείμενα για τα οποία δεν είναι γνωστή η αληθής τους κλάση, δηλαδή όσα αντικείμενα δεν περιέχονται στη λίστα **test_segments_per_class**.

```
#Correspond the classified segments with the classes
clf_segments_per_class = {}

for klass in class_ids:
    segments_of_class = segm[clf == klass]
    clf_segments_per_class[klass] = set(segments_of_class)
    print('Classified segments for class', klass, ':', len(segments_of_class))
```

```
Classified segments for class 1 : 176663
Classified segments for class 2 : 84263
Classified segments for class 3 : 24683
Classified segments for class 4 : 86631
```

```
#Exclude the segments from clf_segments_per_class that are not in test_segments_per_class
#Create lists containing the segment ids of the test segments and the corresponding class
test_list_seg = []
test_list_cl = []
for klass in class_ids:
    for segment_id in test_segments_per_class[klass]:
        test_list_seg.append(segment_id)
        test_list_cl.append(klass)

#Discard the segments of "clf_segments_per_class" that are not in test_list
for klass in class_ids:
    for segment_id in clf_segments_per_class[klass].copy():
        if segment_id not in test_list_seg:
            clf_segments_per_class[klass].discard(segment_id)
```

Τελικός σκοπός είναι η δημιουργία δύο μονοδιάστατων πινάκων με τόσες εγγραφές όσες τα αντικείμενα ελέγχου όπου ο μεν πρώτος θα περιέχει τις κλάσεις των αντικειμένων βάσει των δεδομένων ελέγχου και ο δεύτερος τις κλάσεις των αντικειμένων όπως αυτά ταξινομήθηκαν. Βάσει αυτών των δύο μεταβλητών λοιπόν (*segm_true*, *segm_pred*) υπολογίζεται τελικά ο πίνακας σύγχυσης και οι άλλες μετρικές, ακριβώς όπως και στην αξιολόγηση σε επίπεδο εικονοστοιχείων.

```
#Create lists containing the segment ids of the classified segments and the corresponding class
clf_list_seg = []
clf_list_cl = []
for klass in class_ids:
    for segment_id in clf_segments_per_class[klass]:
        clf_list_seg.append(segment_id)
        clf_list_cl.append(klass)

test_arr = np.column_stack((test_list_seg, test_list_cl))
clf_arr = np.column_stack((clf_list_seg, clf_list_cl))
#Sort test_arr and clf_arr by segment id (1st column)
test_arr = test_arr[test_arr[:,0].argsort()]
clf_arr = clf_arr[clf_arr[:,0].argsort()]
#Define the inputs of the confusion_matrix function by keeping only the classes of test and classified segments
segm_true = test_arr[:,1]
segm_pred = clf_arr[:,1]
```

4.2 Αντικειμενοστρεφής ταξινόμηση δορυφορικών εικόνων με αξιοποίηση του εμπορικού λογισμικού eCognition

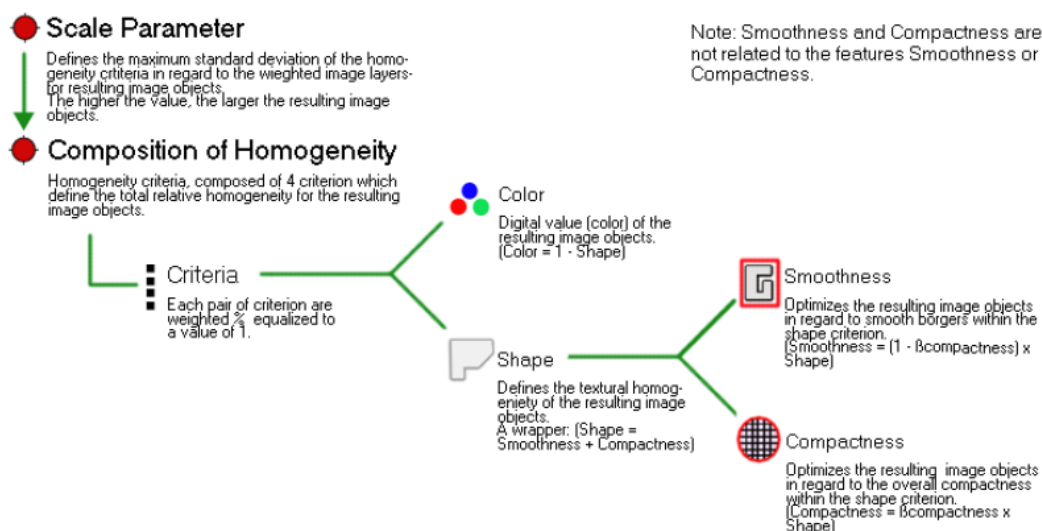
Όπως ακριβώς και στην προηγούμενη μεθοδολογία που περιγράφηκε, πρώτο βήμα για την αντικειμενοστρεφή ταξινόμηση των δορυφορικών δεδομένων στο eCognition είναι η κατάτμηση τους. Για το σκοπό αυτό αξιοποιήθηκε ο αλγόριθμος Multiresolution Segmentation που παρέχει το λογισμικό. Πρόκειται για αλγόριθμο που τοπικά, δηλαδή σε επίπεδο αντικειμένου, ελαχιστοποιεί τη μέση ετερογένειά του για μία δεδομένη ανάλυση. Ανήκει στην κατηγορία bottom-up αλγορίθμων κατάτμησης αφού το σημείο εκκίνησης είναι το μεμονωμένο εικονοστοιχείο το οποίο συνενώνεται με γειτονικά του βάσει του κριτηρίου ομοιογένειας. Το κριτήριο αυτό είναι συνδυασμός φασματικής ομοιογένειας, δηλαδή τα εικονοστοιχεία των αντικειμένων να έχουν παρόμοια φασματικά χαρακτηριστικά, και ομοιογένειας σχήματος (eCognition, 2014). Ο αλγόριθμος αυτός μπορεί να δημιουργήσει πολλαπλά επίπεδα αντικειμένων συνενώνοντας κάθε φορά αντικείμενα του προηγούμενου επιπέδου. Τα νέα αντικείμενα κάθε φορά μπορούν είτε να περιλαμβάνουν εξ ολοκλήρου το αντικείμενο του προηγούμενου επιπέδου είτε καθόλου, δηλαδή σε καμία περίπτωση δεν μπορούν να το τέμνουν. Έτσι, η μέθοδος Multiresolution Segmentation δίνει τη δυνατότητα συσχέτισης των αντικειμένων με τα αντικείμενα ανώτερου επιπέδου (υπερ-αντικείμενα) και του κατώτερο επιπέδου (υπο-αντικείμενα).

Κατά την εφαρμογή του αλγορίθμου Multiresolution Segmentation στο περιβάλλον του eCognition ο χρήστης καλείται να ορίσει μία σειρά παραμέτρων. Η πρώτη είναι ο ορισμός βαρών στα φασματικά κανάλια της εικόνας προς κατάτμηση (*Image Layer weights*). Τα βάρη

αυτά αντιστοιχούν στη σημασία του εκάστοτε καναλιού ως προς τα τελικά αποτελέσματα της κατάτμησης. Για παράδειγμα, αν η εφαρμογή σχετίζεται με την χαρτογράφηση βλάστησης και στα δεδομένα υπάρχει κανάλι που αντιπροσωπεύει τον δείκτη NDVI, τότε είναι θεμιτό σε αυτόν να δοθεί μεγαλύτερη βαρύτητα κατά την κατάτμηση και άρα να του αποδοθεί μεγαλύτερο βάρος από τα υπόλοιπα κανάλια. Επίσης, ο χρήστης έχει τη δυνατότητα να ορίσει αν στη διαδικασία της κατάτμησης θα συμμετέχει και κάποιο θεματικό επίπεδο πληροφοριών (**Thematic Layer Usage**). Πολύ σημαντική παράμετρος είναι αυτή της κλίμακας (**Scale parameter**) η οποία πρόκειται για ένα αυθαίρετο αριθμό που στην ουσία καθορίζει τη μέγιστη επιτρεπτή ετερογένεια ενός αντικειμένου. Αυτό συνεπάγεται ότι όσο μεγαλύτερη είναι η παράμετρος κλίμακας, τόσο πιο ετερογενή θα είναι τα αντικείμενα και κατ' επέκταση μεγαλύτερα αφού συνενώνονται περισσότερα εικονοστοιχεία/αντικείμενα. Επομένως η παράμετρος αυτή είναι το πεδίο στο οποίο ο χρήστης μπορεί να πειραματιστεί με το μέγεθος και άρα το πλήθος των αντικειμένων που θα προκύψουν.

Εξίσου βασική παράμετρος προκειμένου να εφαρμοστεί ο αλγόριθμος Multiresolution Segmentation είναι ο ορισμός του κριτηρίου ομοιογένειας. Αυτό γίνεται εσωτερικά από το λογισμικό υπολογίζοντας τα τέσσερα (4) κριτήρια παρακάτω, όπως φαίνεται και στην Εικόνα 11. Όλα τα παρακάτω κριτήρια εκφράζονται ως ποσοστά με το 100% να είναι η μονάδα.

- Χρώμα (**Color**): το κριτήριο αυτό καθορίζει τη βαρύτητα που δίνεται στα φασματικά χαρακτηριστικά για τη δημιουργία των αντικειμένων. Υψηλά ποσοστά οδηγούν προφανώς σε μεγάλη βαρύτητα της φασματικής πληροφορίας.
- Σχήμα (**Shape**): το κριτήριο αυτό καθορίζει τη βαρύτητα που δίνεται στο σχήμα των αντικειμένων που πρόκειται να δημιουργηθούν. Λειτουργεί συμπληρωματικά με το χρώμα, δηλαδή αν για παράδειγμα οριστεί Shape=0.3 τότε αυτόματα το Color=0.7. Έτσι η τιμή που ζητείται να οριστεί από τον χρήστη είναι μόνο μία από τα δύο παραπάνω κριτήρια και αυτό είναι το ποσοστό του κριτηρίου Shape. Σημειώνεται ότι το κριτήριο του σχήματος διαμορφώνεται βάσει των δύο παρακάτω παραμέτρων smoothness και compactness.
- Ομαλότητα (**Smoothness**): αυτή η παράμετρος σχετίζεται με τη βελτιστοποίηση των αντικειμένων ως προς την ομαλότητα των ορίων τους εντός του κριτηρίου σχήματος (shape). $Smoothness = (1 - \beta compactness) \times Shape$ (eCognition, 2014).
- Συμπαγές σχήμα (**Compactness**): το κριτήριο αυτό σχετίζεται το συμπαγές σχήμα των αντικειμένων που πρόκειται να δημιουργηθούν. Όσο υψηλότερο ποσοστό αποδοθεί στην παράμετρο αυτή τόσο πιο συμπαγές θα είναι το σχήμα των αντικειμένων. $Compactness = (\beta compactness \times Shape)$ (eCognition, 2014).



Εικόνα 11: Εννοιολογικό διάγραμμα ροής αλγορίθμου Multiresolution Segmentation, πηγή: (eCognition, 2014)

Στα πλαίσια της εργασίας αυτής και λαμβάνοντας υπόψη όλα τα παραπάνω που συνθέτουν τον τρόπο λειτουργίας της μεθόδου Multiresolution Segmentation, έγιναν πολλοί πειραματισμοί για τις διαφορετικές παραμέτρους με σημείο εκκίνησης τα δεδομένα από τη διεθνή βιβλιογραφία. Συνολικά δημιουργήθηκαν δύο (2) επίπεδα κατάτμησης τα οποία περιγράφονται στις επόμενες ενότητες.

4.2.1 Κατάτμηση 1^{ου} επιπέδου (Level 1)

Σχετικά με τα βάρη στα φασματικά κανάλια δοκιμάστηκε πληθώρα διαφορετικών επιλογών με άξονα πάντα τον διαχωρισμό καμένων εκτάσεων από τις μη καμένες αλλά και την διάκριση του βαθμού καταστροφής τους. Βάσει της εργασίας των (Mallinis, et al., 2018) βρέθηκε ότι για τις δορυφορικές εικόνες Sentinel-2 ο δείκτης dNBR έχει πολύ υψηλή συσχέτιση με επίγειες μεθόδους εκτίμησης του βαθμού καταστροφής (CBI, GeoCBI, wCBI). Στον δείκτη αυτό λοιπόν, καθώς και στους NDVI_{pre} και dNDVI, αποδόθηκε βάρος ίσο με 4, όπως φαίνεται και στον πίνακα παρακάτω. Αναφορικά με τον δείκτη NDVI πριν το γεγονός της φωτιάς (NDVI_{pre}) παρατηρήθηκε με αξιοποίηση του εργαλείου **Feature view** του λογισμικού ότι είναι ιδιαίτερα χρήσιμος για τον διαχωρισμό της κλάσης Destroyed οπότε αποφασίστηκε να δοθεί μεγάλη βαρύτητα σε αυτόν ήδη από το στάδιο της κατάτμησης. Ο δείκτης dNDVI αποτελεί σημαντικό στοιχείο για το διαχωρισμό καμένων εκτάσεων εν γένει. Σε όλα τα φασματικά κανάλια τις λήψης μετά το γεγονός της πυρκαγιάς αποδόθηκε βάρος ίσο με 1, όπως και στον δείκτη NBR_{pre} και NDVI_{post}. Στους υπόλοιπους φασματικούς δείκτες αποδόθηκε λίγο μεγαλύτερο βάρος, ίσο με 2, λαμβάνοντας υπόψη την ευρεία χρήση τους σε αντίστοιχες εφαρμογές.

Πίνακας 10: Βάρη φασματικών καναλιών στο 1^ο επίπεδο κατάτμησης (Level 1).

Blue	Green	Red	Red Edge1	Red Edge2	Red Edge3	NIR	VegRE	SWIR1	SWIR2
1	1	1	1	1	1	1	1	1	1
NDVI _{pre}	NBR _{pre}	NDVI _{post}	NBR _{post}	dNDVI	dNBR	NBR _{2post}	dNBR2	RBR	
4	1	1	2	4	4	2	2	2	

Στο πρώτο επίπεδο, το οποίο αποτέλεσε και το πιο λεπτόκοκκο επίπεδο κατάτμησης αποφασίστηκε να γίνει η ταξινόμηση της αντικειμένων στις κλάσεις Destroyed και Not Burned. Οι δύο αυτές κλάσεις επιλέχθηκαν καθώς για τη μεν πρώτη (Destroyed) τα δεδομένα αληθείας αποτελούνται από μικρά σε έκταση πολύγωνα και άρα περικλείουν λίγα εικονοστοιχεία των δορυφορικών δεδομένων προς ταξινόμηση και η δεύτερη (Not Burned) αποτελείται μεν από μεγάλα πολύγωνα αλλά παρουσιάζει και μερικά μικρά διάσπαρτα μέσα στην περίμετρο της φωτιάς. Χαρακτηριστικό είναι ότι για την κλάση Destroyed το μεγαλύτερο πολύγωνο αναφοράς από το CEMS έχει έκταση 64277m² δηλαδή 642 εικονοστοιχεία των δεδομένων προς ταξινόμηση ενώ το δεύτερο μεγαλύτερο έχει έκταση 9240m² δηλαδή μόλις 92 εικονοστοιχεία. Έτσι λοιπόν αποφασίστηκε οι κλάσεις αυτές να καθοριστούν στο επίπεδο αυτό που περιέχει και τα μικρότερα αντικείμενα.

Βάσει των παραπάνω παρατηρήσεων σε σχέση με τις εκτάσεις των κλάσεων και μετά από πολλές δοκιμές επιλέχθηκε ως παράμετρος κλίμακας (**Scale parameter**) να οριστεί το 0.35. Η παράμετρος σχήματος (**Shape**) ορίστηκε 0.1 δεδομένου ότι ο διαχωρισμός των καμένων εκτάσεων σχετίζεται σχεδόν αποκλειστικά με τα φασματικά χαρακτηριστικά και αντίστοιχα η παράμετρος **Compactness** πήρε τιμή 0 καθώς δεν επιδιώκεται η δημιουργία συμπαγών αντικειμένων.

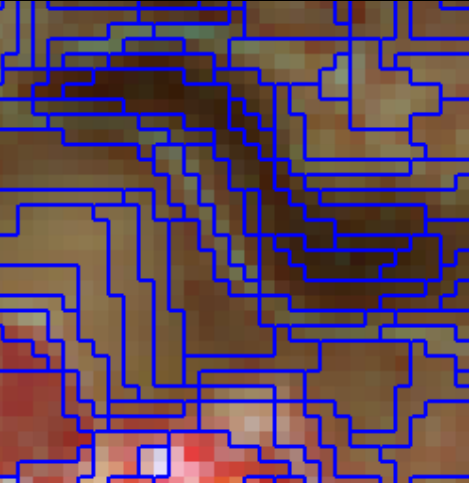
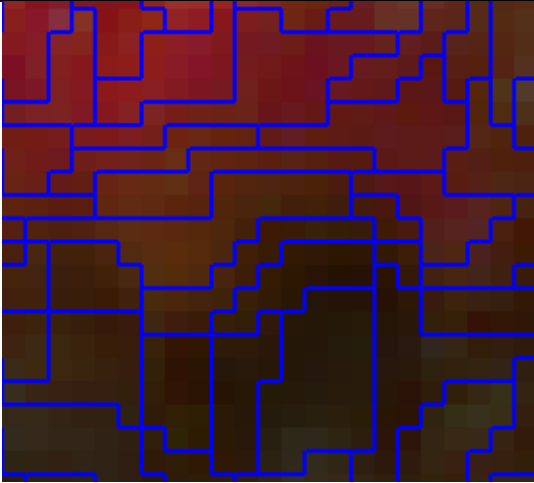
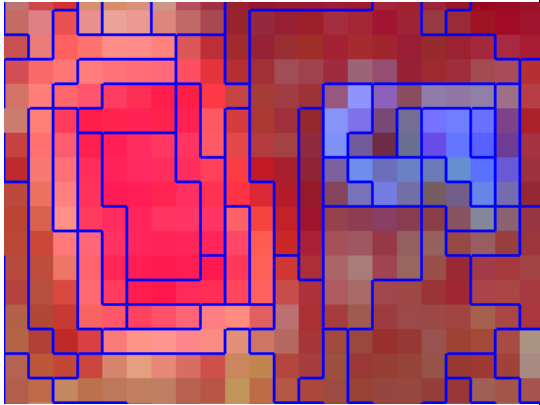
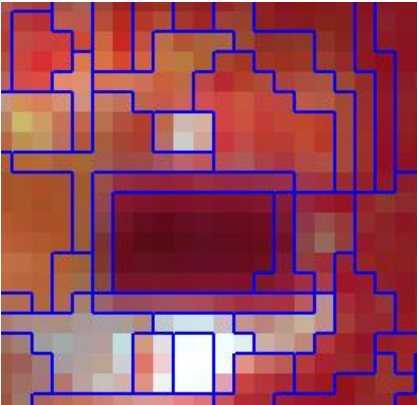
Πίνακας 11: Παράμετροι κατάτμησης για το επίπεδο 1 (Level 1).

Scale parameter	Shape	Compactness
0.35	0.1	0

Το σύνολο των αντικειμένων που δημιουργήθηκαν στο 1^ο επίπεδο κατάτμησης είναι 20761. Το πλήθος αυτό είναι αρκετά μεγάλο και κατ' επέκταση τα αντικείμενα είναι μικρά με μέσο όρο περίπου 17 εικονοστοιχεία. Αυτό είναι λογικό λόγω της πολύ μικρής παραμέτρου κλίμακας που ορίστηκε. Η ανάγκη δημιουργίας τόσο μικρών αντικειμένων σχετίζεται, όπως αναφέρθηκε και προηγουμένως, με την μέτρια χωρική ανάλυση των δορυφορικών δεδομένων και την επιδίωξη ταξινόμησης της κλάσης Destroyed και Not Burned. Η κατάτμηση της εν λόγω δορυφορικής εικόνας πραγματοποιήθηκε σε λιγότερο από 4 δευτερόλεπτα όμως γίνεται αντιληπτό ότι για επεξεργασία δεδομένων που απεικονίζουν πολύ μεγαλύτερες περιοχές, μία τόσο μικρή παράμετρος κλίμακας θα οδηγούσε σε σημαντική αύξηση στον υπολογιστικό χρόνο. Παρ' όλα αυτά, κρίθηκε ότι η δημιουργία μικρών αντικειμένων που στην πορεία θα βελτιώσουν την ταξινόμηση των κλάσεων που προαναφέρθηκαν, αποτελεί παράγοντα τυποποίησης της εν λόγω μεθοδολογίας καθώς η ομοιογένεια των αντικειμένων είναι πολύ υψηλή και μπορεί να εκμεταλλευτεί με διάφορους τρόπους στη συνέχεια.

Σε αυτό το σημείο κρίνεται σημαντικό να παρουσιαστούν μερικά στιγμιότυπα από τα αντικείμενα του επιπέδου 1 προκειμένου να αξιολογηθεί η επιλογή των παραμέτρων .

Πίνακας 12: Ενδεικτικά αποσπάσματα από τα αποτελέσματα της κατάτμησης 1^{ου} επιπέδου (Level 1).







Αποτελέσματα κατάτμησης 1 ^{ου} επιπέδου (Level 1)	
	
<p>Στα αποσπάσματα παραπάνω παρουσιάζονται στο έγχρωμο σύνθετο NIR-R-G τα αντικείμενα που δημιουργήθηκαν κατά το 1^ο επίπεδο κατάτμησης. Και τα δύο απεικονίζουν εκτάσεις καμένες και μη. Στην αριστερή εικόνα παρατηρείται ότι η έκταση με σχήμα ανάποδου S αποτελεί καμένη έκταση με μεγαλύτερο βαθμό καταστροφής από ότι η γύρω έκταση και ο διαχωρισμός αυτός φαίνεται ότι έχει αποτυπωθεί ικανοποιητικά και στα αντικείμενα που έχουν δημιουργηθεί. Στην ίδια εικόνα παρατηρείται επίσης ότι στα αντικείμενα διαχωρίζεται ικανοποιητικά και η βλάστηση από την καμένη έκταση. Στο στιγμιότυπο δεξιά φαίνεται ότι η βλάστηση (βόρειο τμήμα) διαχωρίζεται επαρκώς από την καμένη έκταση (νότιο τμήμα).</p>	
	
<p>Εδώ παρουσιάζονται επίσης στο έγχρωμο σύνθετο NIR-R-G ενδεικτικά αντικείμενα από το 1^ο επίπεδο κατάτμησης. Στο αριστερό απόσπασμα παρατηρείται ο επιτυχής διαχωρισμός μίας πισίνας σε μικρά αντικείμενα, τα οποία στην πορεία πράγματι ταξινομούνται στην κλάση Water. Επίσης φαίνεται και η δημιουργία αντικειμένων για διαφορετικές διαβαθμίσεις της βλάστησης. Αντίστοιχα στη δεξιά εικόνα παρατηρείται ότι έχουν αποτυπωθεί σε ξεχωριστό αντικείμενο στοιχεία του δομημένου περιβάλλοντος (πιθανώς κτίρια λόγω του λευκού χρώματος).</p>	

4.2.2 Ταξινόμηση αντικειμένων 1^{ου} επιπέδου (Level 1)

Για την ταξινόμηση των αντικειμένων αξιοποιήθηκε η ασαφής λογική με την εφαρμογή κανόνων που επιδιώκουν να διαχωρίσουν τα αντικείμενα στις κλάσεις βάσει των χαρακτηριστικών τους. Στην προκειμένη εφαρμογή συμπεριλήφθηκαν κανόνες που

βασίζονται μόνο στη φασματική συμπεριφορά των κλάσεων προκειμένου να επιτευχθεί όσο το δυνατόν τυποποίηση της διαδικασίας. Επεξηγηματικά, θα μπορούσαν ενδεχομένως να δομηθούν κανόνες οι οποίοι να βασίζονται στην παρατήρηση για το μικρό μέγεθος των πολυγώνων της κλάσης Destroyed. Κάτι τέτοιο όμως σχετίζεται με το συγκεκριμένο γεγονός πυρκαγιάς και τέτοιοι κανόνες δεν μπορούν να έχουν ευρεία εφαρμογή για τη μελέτη άλλων γεγονότων. Στον πίνακα που ακολουθεί παρουσιάζονται οι κανόνες στους οποίους βασίστηκε η ταξινόμηση των αντικειμένων του Level 1, η οποία αποτελεί και το βασικότερο μέρος της τελικής ταξινόμησης.

Πίνακας 13: Κανόνες ασαφούς ταξινόμησης αντικειμένων του επιπέδου 1 (Level 1).

Κανόνες ταξινόμησης αντικειμένων στην κλάση Not Burned	
Μεταβλητή	Συνάρτηση συμμετοχής
Mean dNBR	 Mean dNBR 0.115 – 0.125
NDWI	 NDWI 0 – 0.01
Κανόνες ταξινόμησης αντικειμένων στην κλάση Destroyed	
Μεταβλητή	Συνάρτηση συμμετοχής
Mean dNBR2	 Mean dNBR2 0.1 – 0.11
postNBR	 Mean postNBR -0.29 - -0.28
preNBR	 Mean preNBR 0.28 – 0.285
preNDVI	 Mean preNDVI 0.43 – 0.44

Ο πυρήνας της ταξινόμησης στην κλάση Not Burned βρίσκεται στον δείκτη dNBR. Σύμφωνα με τη διεθνή βιβλιογραφία ο δείκτης dNBR για τις μη καμένες εκτάσεις είναι μικρότερος από 0.1. Στην προκειμένη περίπτωση παρατηρείται ότι το ανώτατο όριο του δείκτη αυτού είναι λίγο υψηλότερο και συγκεκριμένα έχει οριστεί ανάμεσα στο 0.115 και 0.125. Αυτό αποφασίστηκε ύστερα από την παρατήρηση ότι με χαμηλότερες τιμές στα όρια, εκτάσεις που δεν ήταν καμένες αποκλείονταν από την κλάση Not Burned. Η χρήση κανόνα για τον δείκτη NDWI, ο οποίος δημιουργήθηκε στο περιβάλλον του eCognition, είναι επικουρική δεδομένου ότι και τα αντικείμενα που αντιπροσωπεύουν νερό (κλάση Water) παρουσιάζουν δείκτη dNBR μικρότερο από το όριο του πρώτου κανόνα. Έτσι λοιπόν, ο δείκτης NDWI καταφέρνει να αποκλείσει τα αντικείμενα νερού από την κλάση Not Burned. Σημειώνεται ότι έχει επιλεγθεί η ταξινόμηση στην κλάση Water να γίνει στα αντικείμενα του 2ου επίπεδο κατάτμησης.

Αναφορικά με τους κανόνες για την κλάση Destroyed αυτοί βασίστηκαν σε φασματικούς δείκτες που χρησιμοποιούνται ευρέως για την κατηγοριοποίηση του βαθμού καταστροφής καμένων εκτάσεων. Αρχικά, αντί για τον δείκτη dNBR2 είχε χρησιμοποιηθεί ο dNBR, ο οποίος όμως παρατηρήθηκε ότι συμπεριλάμβανε στην κλάση Destroyed περισσότερα αντικείμενα που στην πραγματικότητα ανήκουν στην κλάση Damaged. Με αφορμή αυτό και μετά από πολλούς πειραματισμούς διαπιστώθηκε ότι αρκετά αντικείμενα Damaged παρουσιάζουν αντίστοιχα υψηλές τιμές dNBR με αντικείμενα Destroyed ενώ για τον δείκτη dNBR2 σε αρκετές περιπτώσεις αυτό δεν συνέβαινε. Έτσι αποφασίστηκε να γίνει αντικατάσταση του

dNBR με τον dNBR2. Στο σημείο αυτό αξίζει να αναφερθεί ότι ο διαχωρισμός αυτών των δύο κλάσεων ήταν δύσκολος και αυτό αποτυπώνεται και στα τελικά αποτελέσματα παρακάτω. Πέραν της διαφοράς δεικτών NBR κατά τους πειραματισμούς βρέθηκε ότι και ο ίδιος ο δείκτης NBR πριν και μετά το γεγονός της πυρκαγιάς καταφέρνει να διαχωρίσει αφενός τις καμένες εκτάσεις από τις μη καμένες αλλά και το βαθμό καταστροφής σε ικανοποιητικό βαθμό (Damaged – Destroyed). Τέλος, σχετικά με τον δείκτη $NDVI_{pre}$ παρατηρήθηκε ότι η πλειοψηφία των αντικειμένων που πραγματικά ανήκουν στην κλάση Destroyed παρουσιάζουν σχετικά υψηλές τιμές πριν το γεγονός της πυρκαγιάς. Η δημιουργία κανόνα με βάση τον δείκτη αυτό οδήγησε στον μερικό αποκλεισμό αντικειμένων Damaged από την κλάση Destroyed.

4.2.3 Κατάτμηση 2^{ου} επιπέδου (Level 2)

Στο δεύτερο επίπεδο τα αντικείμενα που δημιουργήθηκαν ήταν μεγαλύτερα, αλλά είναι σημαντικό να αναφερθεί ότι από την διαδικασία συνένωσης αντικειμένων αποκλείστηκαν εκείνα τα οποία είχαν ήδη ταξινομηθεί στις δύο κλάσεις του προηγούμενου επιπέδου. Αυτό έχει στόχο την διατήρηση των αντικειμένων αυτών αναλλοίωτων ώστε μη χαθεί η φασματική διαχωρισιμότητά τους κατά τη συγχώνευση. Βάσει αυτών η παράμετρος κλίμακας (**Scale parameter**) ορίστηκε στο 1 (από 0.35 που ήταν στο επίπεδο 1) ενώ η παράμετρος σχήματος (**Shape**) και **Compactness** παρέμειναν ίδιες (0.1 και 0 αντίστοιχα). Και σε αυτό το επίπεδο δόθηκαν σε κάποιους δείκτες μεγαλύτερα βάρη ενώ σε όλα τα υπόλοιπα κανάλια το βάρος διατηρήθηκε ίσο με 1.

Πίνακας 14: Βάρη φασματικών καναλιών στο 2^ο επίπεδο κατάτμησης (Level 2).

Blue	Green	Red	Red Edge1	Red Edge2	Red Edge3	NIR	VegRE	SWIR1	SWIR2
1	1	1	1	1	1	1	1	1	1
$NDVI_{pre}$	NBR_{pre}	$NDVI_{post}$	NBR_{post}	dNDVI	dNBR	$NBR2_{post}$	dNBR2	RBR	
1	1	1	1	4	4	1	1	1	

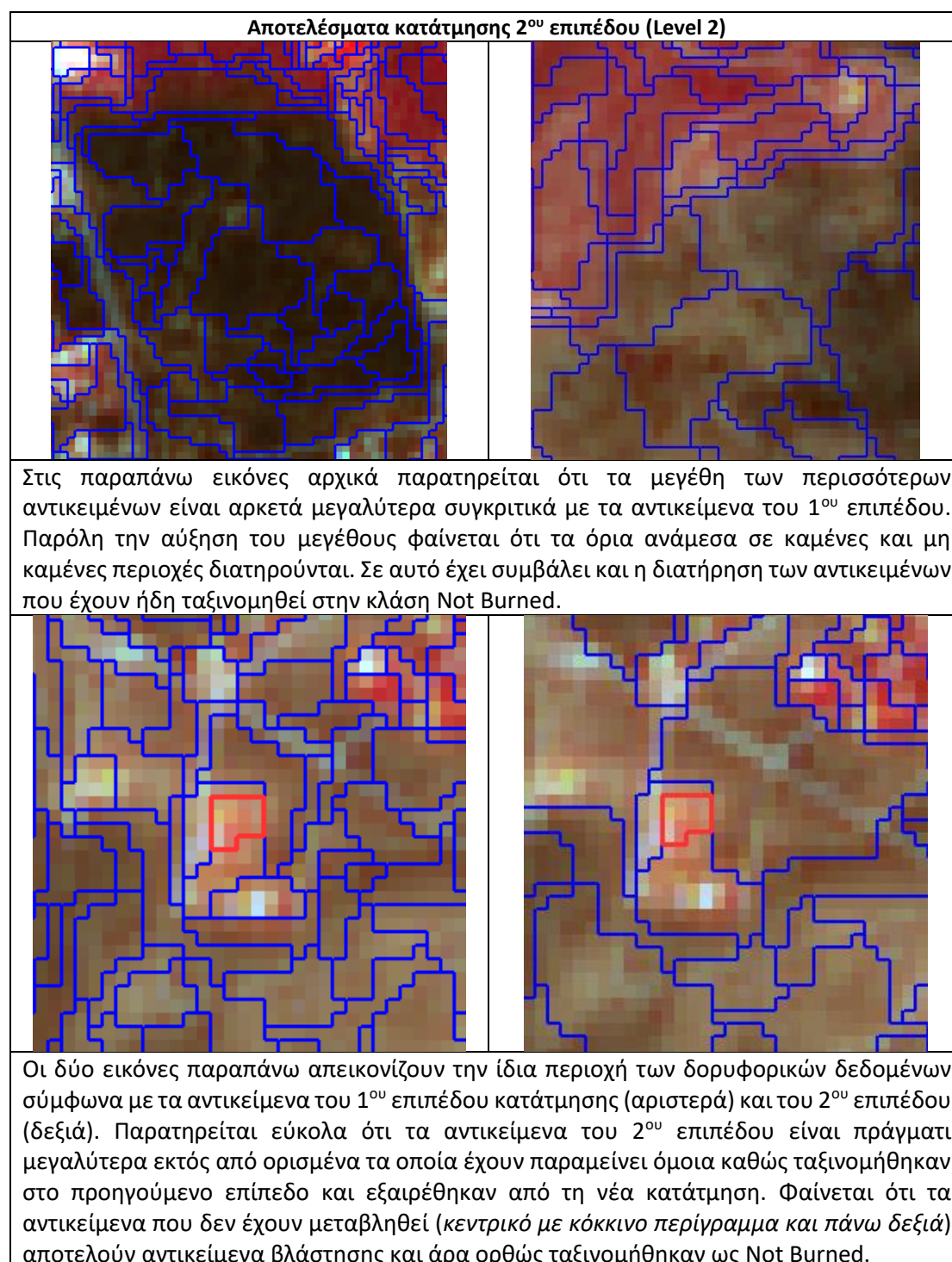
Πίνακας 15: Παράμετροι κατάτμησης για το επίπεδο 2 (Level 2).

Scale parameter	Shape	Compactness
1	0.1	0

Το πλήθος των αντικειμένων του 2^{ου} επιπέδου κατάτμησης είναι 13512. Και σε αυτό το επίπεδο η παράμετρος της κλίμακας είναι αρκετά μικρή. Αξίζει να σημειωθεί ότι δοκιμάστηκαν και μεγαλύτερες τιμές όμως παρατηρήθηκε ότι σε αυτές τις περιπτώσεις κατά την ταξινόμηση η κλάση Damaged εισχωρούσε σε ορισμένα σημεία της θάλασσας. Έτσι επιλέχθηκε η τιμή 1. Οι τιμές **Shape** και **Compactness** παρέμειναν ίδιες με τις τιμές της κατάτμησης στο 1^ο επίπεδο δεδομένου ότι η χαρτογράφηση καμένων εκτάσεων βασίζεται στα φασματικά χαρακτηριστικά των αντικειμένων και όχι στα γεωμετρικά.

Στον πίνακα παρακάτω παρατίθενται ενδεικτικά αποσπάσματα των αντικειμένων του 2^{ου} επιπέδου κατάτμησης, όπως ακριβώς έγινε και στην υποενότητα 4.2.1 για τα αντικείμενα του 1^{ου} επιπέδου.

Πίνακας 16: Ενδεικτικά αποσπάσματα από τα αποτελέσματα της κατάτμησης 2ου επιπέδου (Level 2).









4.2.4 Ταξινόμηση αντικειμένων 2^{ου} επιπέδου (Level 2)

Έχοντας ταξινομήσει τα αντικείμενα των κλάσεων Not Burned και Destroyed στο επίπεδο 1 και διατηρώντας τα αντικείμενα αυτά αναλλοίωτα κατά τη δημιουργία του επιπέδου 2, οι κανόνες για την ταξινόμηση αυτών των αντικειμένων στις αντίστοιχες κλάσεις του εν λόγω επιπέδου βασίζονται στο χαρακτηριστικό "Existence of sub objects". Εφόσον τα αντικείμενα αυτά δεν έχουν μεταβληθεί καθόλου μεταξύ των δύο επιπέδων, η τιμή που θα πάρει το

χαρακτηριστικό “Existence of sub objects” για τις δύο αυτές κλάσεις (Not Burned, Destroyed) είναι 1 (δηλαδή έχουν πλήρη επικάλυψη). Έτσι, για τις κλάσεις Not Burned 2 και Destroyed 2 ορίστηκαν κανόνες με βάση αυτό το χαρακτηριστικό, οι οποίοι κάνουν χρήση της συνάρτησης συμμετοχής με τη μορφή “Μεγαλύτερο από” και ως εύρος ορίστηκε το 0.6-0.65. Το διάστημα αυτό είναι τυχαίο, στη θέση του θα μπορούσε να είναι οποιοδήποτε διάστημα μεγαλύτερο του 0, αφού η τιμή του χαρακτηριστικού “Existence of sub objects” για αυτά τα αντικείμενα είναι πάντα 1. Στην κλάση Damaged ταξινομήθηκαν όσα αντικείμενα του 2^{ου} επιπέδου δεν ανήκουν σε καμία από τις υπόλοιπες κατηγορίες, αντιστρέφοντας το χαρακτηριστικό “Similarity to class”. Τέλος, η κλάση Water ορίστηκε με βάση το χαρακτηριστικό NDWI. Στον πίνακα που ακολουθεί παρουσιάζονται οι κανόνες στους οποίους βασίστηκε η ταξινόμηση των αντικειμένων του Level 2.

Πίνακας 17: Κανόνες ασαφούς ταξινόμησης αντικειμένων του επιπέδου 2 (Level 2).

Κανόνες ταξινόμησης αντικειμένων στην κλάση Not Burned 2	
Existence of sub objects	 Existence of sub objects Not_Burned 1 (1)
Κανόνες ταξινόμησης αντικειμένων στην κλάση Destroyed 2	
Existence of sub objects	 Existence of sub objects Destroyed 1 (1)
Κανόνες ταξινόμησης αντικειμένων στην κλάση Damaged	
Μεταβλητή	Συνάρτηση συμμετοχής
Not Similarity to class	 Destroyed 2
Not Similarity to class	 Not_Burned 2
Not Similarity to class	 Water
Κανόνες ταξινόμησης αντικειμένων στην κλάση Water	
Μεταβλητή	Συνάρτηση συμμετοχής
NDWI	 NDWI 0.01 – 0.02

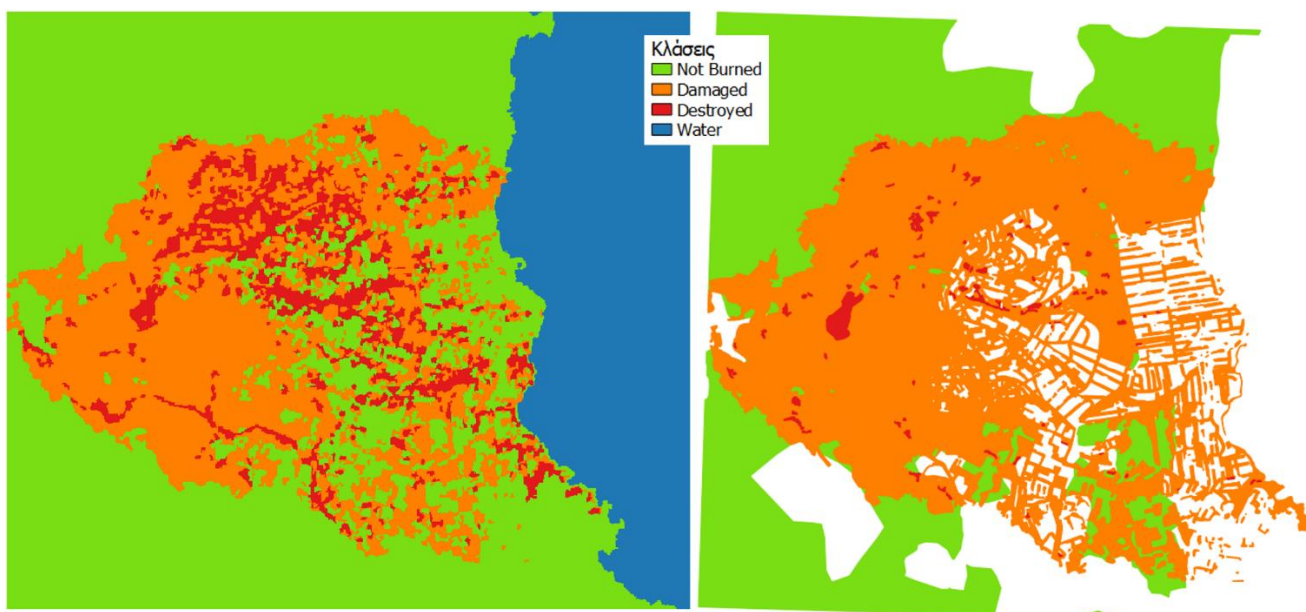
Στο κεφάλαιο αυτό παρουσιάζονται και αξιολογούνται τα αποτελέσματα που προέκυψαν από την εφαρμογή των μεθοδολογιών που παρουσιάστηκαν στο Κεφάλαιο 4.

5.1 Αποτελέσματα αλγορίθμου αντικειμενοστρεφούς ταξινόμησης δορυφορικών εικόνων σε γλώσσα Python

Στην ενότητα αυτή γίνεται παρουσίαση και αξιολόγηση των αποτελεσμάτων του αλγορίθμου αντικειμενοστρεφούς ταξινόμησης όπως αυτός περιγράφηκε στις υποενότητες 4.1.1, 4.1.2 και 4.1.3. Για την αξιολόγησή του ακολουθείται η παρακάτω δομή, η οποία εφαρμόζεται και για την εν λόγω μεθοδολογία αλλά και για την μεθοδολογία που αναπτύχθηκε στο λογισμικό eCognition:

1. Αρχικά γίνεται οπτικός έλεγχος μέσω της σύγκρισης της ταξινομημένης δορυφορικής εικόνας με τα πολυγωνικά δεδομένα αναφοράς του Copernicus EMS. Τα συμπεράσματα που προκύπτουν από εδώ είναι ποιοτικά και λειτουργούν συμπληρωματικά με την ποσοτική αξιολόγηση από τα παρακάτω βήματα.
2. Στη συνέχεια παρουσιάζεται ο πίνακας σύγχυσης και οι μετρικές για την ακρίβεια όπου αποτελούν τον πυρήνα της ποσοτικής αξιολόγησης.
3. Τέλος, υπολογίζεται για κάθε κλάση το ποσοστό επικάλυψης των αποτελεσμάτων ταξινόμησης και δεδομένων αναφοράς CEMS επί της συνολικής έκτασης τόσο των δεδομένων αναφοράς όσο και των ταξινομημένων δεδομένων.

Ο οπτικός έλεγχος καθώς και τα ποσοστά επικάλυψης είναι τα ίδια, ανεξάρτητα από το επίπεδο στο οποίο έγινε η αξιολόγηση (εικονοστοιχείο ή αντικείμενο) για αυτό και περιγράφονται εκτός της υποενότητας 5.1.1. Στην εικόνα 12 παρακάτω παρουσιάζονται η ταξινομημένη δορυφορική εικόνα (αριστερά) και τα πολυγωνικά δεδομένα αναφοράς από το CEMS (δεξιά). Αρχικά είναι σημαντικό να αναφερθεί ότι από την απεικόνιση των δεδομένων αναφοράς λείπει η κλάση Water καθώς τα δεδομένα από την υπηρεσία CEMS σχετίζονται με τις καμένες εκτάσεις και έτσι το ground για την κλάση αυτή συλλέχθηκε χειροκίνητα βάσει της δορυφορικής εικόνας. Εκτενής περιγραφή αυτού πραγματοποιήθηκε στο κεφάλαιο 3. Στο ίδιο κεφάλαιο επίσης τεκμηριώνεται και η έλλειψη δεδομένων αναφοράς για ορισμένες περιοχές της εικόνας, η οποία οφείλεται είτε στη μη αξιοποίηση του θεματικού επιπέδου “built-up” είτε σε μη επεξεργασία των περιοχών αυτών από την ίδια την υπηρεσία CEMS και άρα την μη παραγωγή δεδομένων για αυτές.



Εικόνα 12: Αποτέλεσμα ταξινόμησης βάσει του αλγορίθμου στη γλώσσα Python (αριστερά) και δεδομένα αναφοράς από το CEMS (δεξιά).

Παρατηρώντας την εικόνα παραπάνω φαίνεται εύκολα ότι η περίμετρος της καμένης έκτασης ανάμεσα στις δύο απεικονίσεις είναι πολύ παρόμοια. Επιπλέον είναι εμφανές ότι στην κλάση Destroyed (κόκκινο χρώμα) έχει ταξινομηθεί σημαντικά μεγαλύτερη έκταση σε σχέση με τα δεδομένα αναφοράς. Παρ' όλα αυτά στις περιοχές που πράγματι υπάρχουν τμήματα της κλάσης Destroyed βάσει των δεδομένων CEMS αυτά φαίνεται να έχουν εντοπιστεί και κατά την ταξινόμηση σε ένα σημαντικό ποσοστό τους. Στη συνέχεια παρουσιάζονται ενδεικτικά κάποια στιγμιότυπα από την περίμετρο των καμένων όπου έχουν ταξινομηθεί ορθά αλλά και γενικά κάποια αποσπάσματα λανθασμένης ταξινόμησης προκειμένου να εξαχθούν κάποια συμπεράσματα για τις αστοχίες της μεθοδολογίας.

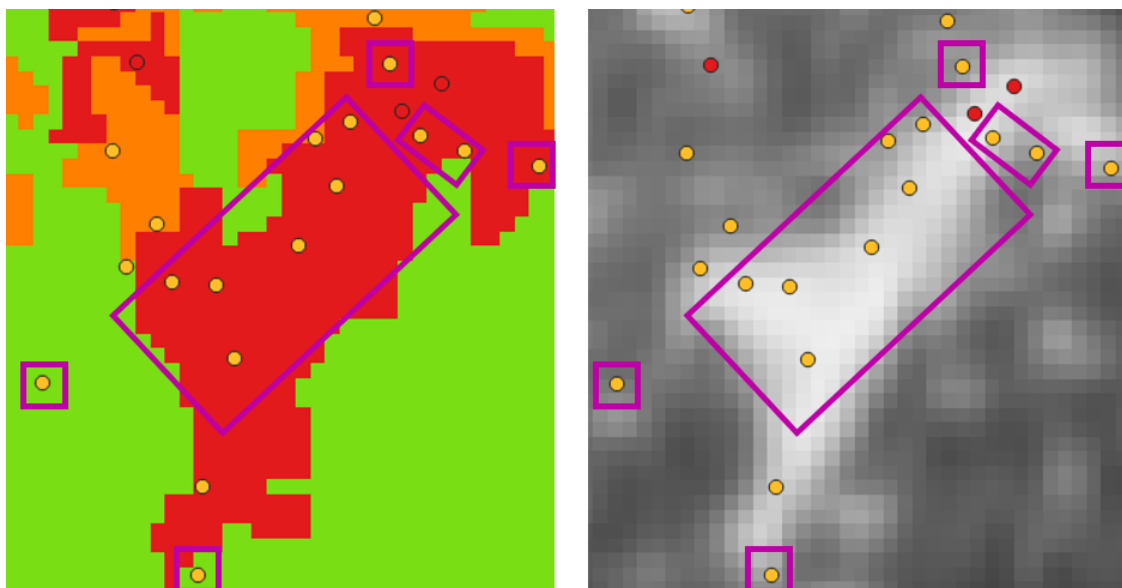


Εικόνα 13: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στο μαύρη έλλειψη επισημαίνεται η περιοχή ορθής ταξινόμησης.



Εικόνα 14: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης (πάνω) συγκριτικά με τα δεδομένα αναφοράς CEMS (κάτω). Στο μαύρη έλλειψη επισημαίνεται η περιοχή ορθής ταξινόμησης.

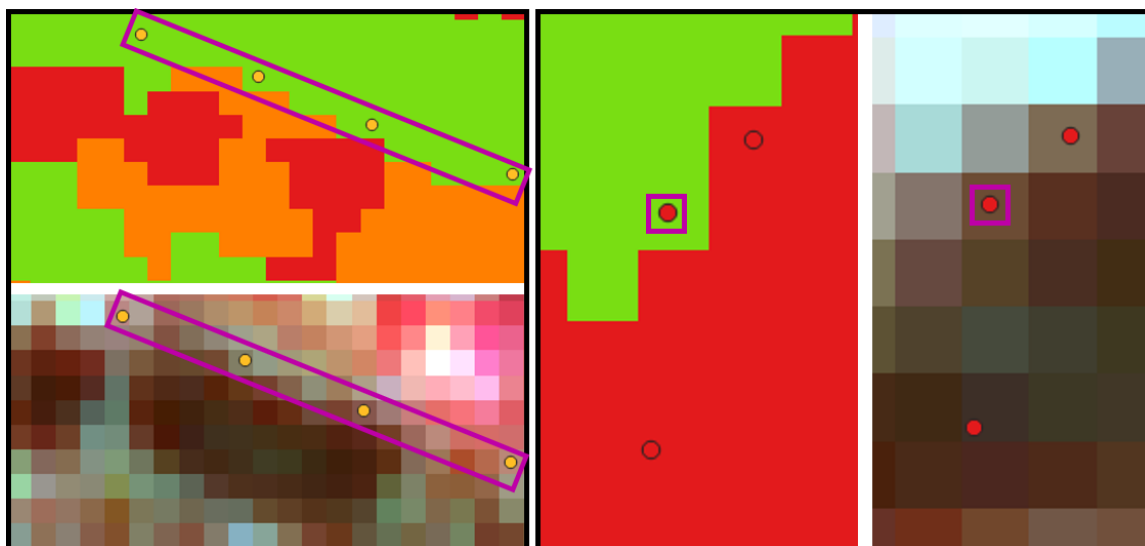
Βεβαίως παρατηρούνται και περιπτώσεις λανθασμένης ταξινόμησης όπως ενδεικτικά παρουσιάζονται στις δύο παρακάτω εικόνες.



Εικόνα 15: Ταξινομημένη εικόνα (αριστερά) και εικόνα του δείκτη dNBR (δεξιά). Στα μωβ πλαίσια επισημαίνονται τα λανθασμένα ταξινομημένα σημεία ελέγχου.

Στην παραπάνω εικόνα, στα μωβ πλαίσια επισημαίνονται σημεία ελέγχου που έχουν ταξινομηθεί λάθος. Με πράσινο (σημεία και πολύγωνα) συμβολίζεται η κλάση Not Burned, με πορτοκαλί (σημεία και πολύγωνα) η κλάση Damaged και με κόκκινο (σημεία και πολύγωνα) η Destroyed. Παρατηρούνται δύο είδη λάθος ταξινομήσεων: αρκετά σημεία Damaged που έχουν ταξινομηθεί ως Destroyed και λιγότερα της ίδιας κλάσης ως Not Burned. Όπως φαίνεται και ποσοτικά από τα παρακάτω μέτρα αξιολόγησης, η κλάση Damaged είναι

αυτή που παρουσιάζει τη μεγαλύτερη σύγχυση με τις άλλες δύο, γεγονός που εν μέρει μπορεί να δικαιολογηθεί καθώς αποτελεί την ενδιάμεση κατάσταση μεταξύ των κλάσεων Not Burned και Destroyed. Επίσης, σημαντικό ρόλο στη σύγχυση αυτή έχει και το γεγονός ότι σε αυτή την κλάση ανήκουν εκτάσεις που έχουν υποστεί ζημιές από αμελητέες ως αρκετά σοβαρές (Εικόνα 7, σελ. 26). Οι αμελητέες ζημιές είναι λογικό ότι φασματικά προσεγγίζουν περισσότερο τις μη καμένες εκτάσεις από ότι οι καταστραμμένες (κλάση Destroyed). Αντίστοιχα τμήματα Damaged που έχουν υποστεί σοβαρές ζημιές μοιάζουν φασματικά με την κλάση Destroyed. Επίσης, όπως παρατηρείται στην εικόνα του δείκτη dNBR (δεξιά απεικόνιση της Εικόνας 15) οι τιμές στις οποίες αντιστοιχούν τα ορθώς ταξινομημένα σημεία Destroyed φαίνονται όμοιες με αυτές των σημείων Damaged που όμως έχουν λανθασμένα ταξινομηθεί ως Destroyed. Όπως έχει περιγραφεί και αναλυτικότερα στο κεφάλαιο 4 υπάρχουν περιπτώσεις όπου τα όρια των δεικτών ανάμεσα στις δύο αυτές κλάσεις δεν είναι πλήρως διακριτά γεγονός που προφανώς μεταφέρεται και κατά την εκπαίδευση ενός ταξινομητή.



Εικόνα 16: Στιγμιότυπα ταξινομημένης εικόνα με τα αντίστοιχα από την απεικόνιση των δορυφορικών δεδομένων σε έγχρωμο σύνθετο NIR-R-G. Στα μωβ πλαίσια επισημαίνονται τα λανθασμένα ταξινομημένα σημεία ελέγχου.

Στις απεικονίσεις που βρίσκονται αριστερά στην Εικόνα 16 παρουσιάζονται σημεία ελέγχου της κλάσης Damaged που έχουν ταξινομηθεί ως Not Burned. Από την απεικόνιση του έγχρωμου σύνθετου NIR-R-G παρατηρείται ότι τα εν λόγω σημεία βρίσκονται στα όρια καμένης και μη καμένης έκτασης. Η ύπαρξη μη καμένης έκτασης υποδηλώνεται από τις ροζ και κόκκινες αποχρώσεις των εικονοστοιχείων. Ως εκ τούτου, πιθανή αιτία για το σφάλμα της ταξινόμησης είναι η συνύπαρξη αυτών των εικονοστοιχείων της κλάσης Damaged που βρίσκονται στα όρια με τα εικονοστοιχεία μη καμένης έκτασης. Αναφορικά με τις απεικονίσεις στο δεξιά κομμάτι της εικόνας, εκεί παρουσιάζεται ένα σημείο της κλάσης Destroyed που έχει ταξινομηθεί ως Not Burned. Να σημειωθεί ότι βάσει του πίνακα σύγχυσης που ακολουθεί στην υποενότητα 5.1.1 αυτό είναι το μοναδικό σημείο ελέγχου της κλάσης Destroyed που ταξινομήθηκε ως Not Burned ενώ το αντίθετο σφάλμα δεν παρατηρείται καθόλου. Και εδώ παρατηρείται αντίστοιχο φαινόμενο με την αριστερή απεικόνιση, δηλαδή το εικονοστοιχείο που έχει ταξινομηθεί λάθος βρίσκεται στα όρια των κλάσεων και κατά την κατάτμηση συμπεριλήφθηκε σε αντικείμενο με εικονοστοιχεία μη καμένης έκτασης. Έτσι

προφανώς τα φασματικά χαρακτηριστικά της μη καμένης έκτασης υπερίσχυσαν λόγω πλειοψηφίας και κατά την ταξινόμηση το αντικείμενο κατατάχθηκε στην κλάση Not Burned.

Σημαντικός δείκτης της ποιότητας των παραγόμενων αποτελεσμάτων είναι ο υπολογισμός του ποσοστού της κοινής έκτασης που αυτά παρουσιάζουν ανά κλάση με τα δεδομένα αναφοράς. Το ποσοστό αυτό υπολογίζεται προς την συνολική πραγματική έκταση κάθε κλάσης, δηλαδή την έκταση των δεδομένων CEMS, αλλά και προς την έκταση των κλάσεων βάσει της ταξινόμησης. Το μεν πρώτο ποσοστό υποδεικνύει πόση έκταση από τα αληθή δεδομένα κατάφερε να εντοπιστεί από τον αλγόριθμο ενώ το δεύτερο πόση έκταση από αυτή που ταξινομήθηκε σε κάθε κλάση πράγματι ανήκει σε αυτή βάσει των δεδομένων CEMS. Αναφορικά με τα ποσοστά αυτά για την μεθοδολογία που αναπτύχθηκε σε προγραμματιστικό περιβάλλον Python παρουσιάζονται στους πίνακες παρακάτω:

Πίνακας 18: Έκταση και ποσοστά επικάλυψης της κλάσης Not Burned (Python) με τα δεδομένα αναφοράς CEMS.

Κλάση 1: "Not Burned"	
Έκταση κλάσης βάσει δεδομένων CEMS (m ²)	8.903.406,684
Έκταση βάσει προϊόντων αλγορίθμου (m ²)	17.651.134,800
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων αλγορίθμου (m ²)	8.616.333,18
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	96,78 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων του αλγορίθμου (%)	48,81 %

Σχετικά με την έκταση που προέκυψε από τον αλγόριθμο για την κλάση Not_Burned, η οποία είναι σχεδόν διπλάσια από την αντίστοιχη των δεδομένων αναφοράς, είναι σημαντικό να επισημανθεί ξανά ότι τα δεδομένα από το Copernicus EMS παρουσιάζουν κάποια "κενά" σε σχέση με τη δορυφορική εικόνα που ταξινομήθηκε. Επεξηγηματικά, για κάποια μεγάλα τμήματα της εικόνας είτε δεν αντιστοιχεί κάποιο πολύγωνο από τα δεδομένα CEMS που να υποδηλώνει τον βαθμό καταστροφής τους είτε αντιστοιχεί το επίπεδο "built-up" το οποίο όμως δεν αξιοποιείται για λόγους που έχουν αναφερθεί αναλυτικά στο κεφάλαιο 3. Κατά πλειοψηφία τα τμήματα αυτά βάσει φωτοερμηνείας φαίνεται να μην έχουν επηρεαστεί από την πυρκαγιά. Άρα λοιπόν είναι λογικό να προκύπτει πολύ μεγαλύτερη έκταση για την κλάση Not Burned αφού κατά την ταξινόμηση αποδίδεται και σε αυτά τα τμήματα κάποια κλάση. Το ποσοστό των κοινών τμημάτων επί της συνολικής έκτασης των δεδομένων αναφοράς για την κλάση αυτή πλησιάζει το 100% γεγονός που υποδηλώνει ότι ο αλγόριθμος καταφέρνει να κατηγοριοποιήσει ορθά τις μη καμένες εκτάσεις. Το ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων του αλγορίθμου, το οποίο είναι χαμηλό και πλησιάζει μόλις το 50%, σχετίζεται την μη κάλυψη των δεδομένων CEMS σε τμήματα της δορυφορική εικόνας όπως αναλύθηκε στην αρχή της παραγράφου.

Πίνακας 19: Έκταση και ποσοστά επικάλυψης της κλάσης Damaged (Python) με τα δεδομένα αναφοράς CEMS.

Κλάση 2: "Damaged"	
Έκταση κλάσης βάσει δεδομένων CEMS (m ²)	10.610.595,332
Έκταση βάσει προϊόντων αλγορίθμου (m ²)	8.418.402,735

Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων αλγορίθμου (m ²)	7.276.851,97
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	68,58 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων του αλγορίθμου (%)	86,44 %

Βάσει του παραπάνω πίνακα φαίνεται ότι ο αλγόριθμος υποεκτιμά την κλάση Damaged, δηλαδή η πραγματική έκτασή της είναι μεγαλύτερη από αυτή που προέκυψε από την εφαρμογή του. Πιο αναλυτικά, ο αλγόριθμος κατάφερε να εντοπίσει ορθά το 68,58 % της έκτασης που έχει χαρακτηριστεί ως Damaged βάσει των δεδομένων αληθείας. Επίσης, βάσει των παραπάνω, το 86,44 % της έκτασης που ταξινομήθηκε ως Damaged ανήκει πράγματι σε αυτή την κατηγορία ενώ το υπόλοιπο προφανώς ανήκει σε κάποια από τις άλλες κλάσεις. Όπως επιβεβαιώνεται και από τον πίνακα σύγχυσης παρακάτω, σημαντική έκταση της εν λόγω κλάσης ταξινομείται είτε στην κλάση Not Burned είτε στη Destroyed οπότε έτσι τεκμηριώνεται το φαινόμενο υποεκτίμησής της.

Πίνακας 20: Έκταση και ποσοστά επικάλυψης της κλάσης Destroyed (Python) με τα δεδομένα αναφοράς CEMS.

Κλάση 3: “Destroyed”	
Έκταση κλάσης βάσει δεδομένων CEMS (m ²)	190631.149
Έκταση βάσει προϊόντων αλγορίθμου (m ²)	2.466.001,317
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων αλγορίθμου (m ²)	163.453,029
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	85,74 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων του αλγορίθμου (%)	6,63 %

Το υψηλό ποσοστό (85,74%) των κοινών τμημάτων επί της έκτασης των δεδομένων αληθείας για την κλάση Destroyed δείχνει ότι ο αλγόριθμος εντοπίζει τις κατεστραμμένες εκτάσεις σε αρκετά ικανοποιητικό βαθμό. Συγκρίνοντας όμως την έκταση από τα δεδομένα CEMS και από τον αλγόριθμο προκύπτει ότι ο τελευταίος ταξινομεί ως Destroyed πολύ περισσότερα τμήματα γης, είναι δηλαδή αρκετά “ευαίσθητος” στην κλάση αυτή (190.631,149 αντί 2.466.001,317). Αυτό υποδηλώνεται και από το πάρα πολύ χαμηλό ποσοστό (6,63 %) των κοινών τμημάτων επί της συνολικής έκτασης της κλάσης αυτής που προέκυψε από την ταξινόμηση. Πιο αναλυτικά, μόνο το 6,63 % της έκτασης που ταξινομήθηκε ως Destroyed ανήκει πράγματι στην κατηγορία αυτή. Η πληροφορία αυτή συνδυαστικά με τα αντίστοιχα στατιστικά για την κατηγορία Damaged από τον προηγούμενο πίνακα δείχνει ότι πιθανώς μεγάλο μέρος της κλάσης Damaged ταξινομήθηκε στην κατηγορία Destroyed. Επεξηγηματικά, το συμπέρασμα αυτό εξάγεται λαμβάνοντας υπόψη ότι η έκταση που ταξινομήθηκε από τον αλγόριθμο στην κλάση Damaged είναι μικρότερη περίπου κατά 2,1 τ.χλμ. σε σχέση με τα δεδομένα αναφοράς ενώ για την κλάση Destroyed είναι μεγαλύτερη κατά περίπου 2,3 τ.χλμ. Προκειμένου να διερευνηθεί περαιτέρω η παρατήρηση αυτή υπολογίστηκε η κοινή έκταση μεταξύ της ταξινομημένης κλάσης Destroyed και των δεδομένων αναφοράς της κλάσης Damaged.

Πίνακας 21: Έλεγχος επικάλυψης ταξινομημένης κλάσης Destroyed (Python) με τα δεδομένα αναφοράς για την κλάση Destroyed.

Έλεγχος επικάλυψης ταξινομημένης κλάσης Destroyed με τα δεδομένα αναφοράς για την κλάση Destroyed	
Έκταση κλάσης Damaged βάσει δεδομένων CEMS (m ²)	10.610.595,332
Έκταση κλάσης Destroyed βάσει προϊόντων αλγορίθμου (m ²)	2.466.001,317
Έκταση κοινών τμημάτων δεδομένων αληθείας Damaged και προϊόντων αλγορίθμου Destroyed (m ²)	1.857.666,518
Ποσοστό κοινών τμημάτων Damaged CEMS-ταξινομημένων Destroyed επί των ταξινομημένων Destroyed (%)	75,33 %
Ποσοστό κοινών τμημάτων Damaged CEMS-ταξινομημένων Destroyed επί Damaged CEMS (%)	17,51 %

Ο παραπάνω πίνακας επιβεβαιώνει την αρχική παρατήρηση ότι το μεγαλύτερο μέρος της κλάσης Destroyed που προκύπτει από τον αλγόριθμο ανήκει στην πραγματικότητα στην κλάση Damaged. Αυτή η επιβεβαίωση προκύπτει από το υψηλό ποσοστό κοινής έκτασης ανάμεσά τους (75,33%). Επίσης φαίνεται ότι το 17,51% της πραγματικής έκτασης της κλάσης Damaged ταξινομήθηκε ως Destroyed στην εφαρμογή αυτή.

Δεδομένου ότι στόχος της εργασίας είναι και η χαρτογράφηση των καμένων εκτάσεων ανεξαρτήτου του βαθμού καταστροφής αυτών, θεωρήθηκε σκόπιμο να συμπεριληφθεί ένας ακόμη πίνακας που θα αποτυπώνει τα αντίστοιχα στοιχεία συνολικά για την καμένη έκταση, δηλαδή αθροιστικά για τις κλάσεις Damaged και Destroyed.

Πίνακας 22: Έκταση και ποσοστά επικάλυψης αθροιστικά για τις κλάσεις Damaged και Destroyed (Python) με τα δεδομένα αναφοράς CEMS.

Καμένη έκταση ανεξαρτήτου βαθμού καταστροφής (Damaged & Destroyed)	
Έκταση κατηγορίας βάσει δεδομένων CEMS (m ²)	10.801.226,48
Έκταση βάσει προϊόντων αλγορίθμου (m ²)	10.884.404,052
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων αλγορίθμου (m ²)	9.323.655,776
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	86,32 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων του αλγορίθμου (%)	85,66 %

Φαίνεται λοιπόν ότι το 86,32 % της καμένης έκτασης εντοπίστηκε με επιτυχία από τον αλγόριθμο, ενώ από την έκταση που ταξινομήθηκε ως καμένη βάσει του αλγορίθμου, το 85,66% ανήκει πράγματι σε αυτή την κατηγορία. Λαμβάνοντας υπόψη συνολικά την ακρίβεια που υπολογίστηκε από τον πίνακα σύγχυσης αλλά και τα δεδομένα από τους παραπάνω πίνακες εξάγεται το συμπέρασμα ότι ο αλγόριθμος καταφέρνει να διαχωρίσει επαρκώς τις καμένες από τις μη καμένες εκτάσεις. Στο κομμάτι της κατηγοριοποίησης του βαθμού

καταστροφής όμως φαίνεται να υπερεκτιμά την έκταση των κατεστραμμένων εκτάσεων (κλάση Destroyed) εις βάρος της κλάσης Damaged.

5.1.1 Αποτελέσματα σε επίπεδο εικονοστοιχείου (pixel-based)

Όπως περιγράφηκε αναλυτικά και στο Κεφάλαιο 4, ο πίνακας σύγχυσης και οι υπόλοιπες μετρικές ακρίβειας υπολογίστηκαν με δύο διαφορετικούς τρόπους. Ο πρώτος αφορά στον υπολογισμό τους σε επίπεδο εικονοστοιχείων ενώ ο δεύτερος σε επίπεδο αντικειμένων.

Σχετικά με τον πρώτο, υπολογίστηκε συγκρίνοντας τα μεμονωμένα σημεία ελέγχου με τα αντίστοιχα σημεία της ταξινομημένης εικόνας. Τα αποτελέσματα από αυτή τη σύγκριση συνοψίζονται στον πίνακα σύγχυσης και τις μετρικές ακρίβειας παρακάτω:

Πίνακας 23: Πίνακας σύγχυσης σε επίπεδο εικονοστοιχείων των αποτελεσμάτων ταξινόμησης σε περιβάλλον Python.

		CLASSIFIED					
Κλάσεις:		Not_Burned	Damaged	Destroyed	Water	Sum	PA
TRUE DATA	Not_Damaged	809	74	0	0	883	0.92
	Damaged	137	591	129	0	857	0.69
	Destroyed	1	26	193	0	220	0.88
	Water	0	0	0	220	220	1.00
	Sum	947	691	322	220	OA:	0.83
UA		0.85	0.86	0.60	1.00	Kappa:	0.75

Πίνακας 24: Πίνακας Omission και Commission errors σε επίπεδο εικονοστοιχείων για τα αποτελέσματα ταξινόμησης σε περιβάλλον Python.

Κλάσεις:	Omission Errors (Type I)	Commission Errors (Type II)
Not_Burned	8,4%	14,6 %
Damaged	31%	14,5 %
Destroyed	12,3 %	40 %
Water	0 %	0 %

Από τον πίνακα σύγχυσης προκύπτει ότι για την κλάση Not Burned η μεγάλη πλειοψηφία των σημείων ελέγχου ταξινομήθηκε ορθά (0.92). Παρατηρείται όμως ότι υπήρξε μία υπερεκτίμηση (14,6%) της εν λόγω κλάσης καθώς ταξινομήθηκαν σε αυτή και 137 επιπλέον σημεία ελέγχου τα οποία ανήκουν στην κλάση Damaged. Από την κλάση Destroyed μόνο 1 σημείο ελέγχου ταξινομήθηκε ως Not Burned ενώ από την κλάση Water κανένα. Παρατηρείται λοιπόν ένα φαινόμενο σύγχυσης μεταξύ των κλάσεων Not Burned και Damaged. Όσον αφορά στα σφάλματα τύπου I και II (omission & commission errors) για την κλάση Not Burned είναι σε χαμηλά επίπεδα με το σφάλμα τύπου II να είναι μεγαλύτερο, γεγονός που αποτυπώνει την σχετικά μικρή υπερεκτίμησης της εν λόγω κλάσης. Αυτά σε συνδυασμό και με την επικάλυψη που παρουσιάζει η ταξινομημένη κλάση Not Burned με την αντίστοιχη των δεδομένων CEMS (96,78 %) οδηγούν στο συμπέρασμα ότι ο αλγόριθμος λειτουργεί αρκετά καλά για τον διαχωρισμό καμένων από τις μη καμένες εκτάσεις.

Για την κλάση Damaged, η οποία παρουσιάζει τη χαμηλότερη ακρίβεια, το ποσοστό των ορθά ταξινομημένων σημείων ελέγχου είναι αρκετά χαμηλότερο (0.69) και παρατηρείται ότι ένας σημαντικός αριθμός σημείων ελέγχου έχει ταξινομηθεί τόσο στην κλάση Not Burned όσο και

στην Destroyed (137 και 129 αντίστοιχα). Ήδη από την ανάλυση των αποτελεσμάτων για την κλάση Not Burned παρατηρήθηκε ένα φαινόμενο μικρής φασματικής σύγχυσης με την Damaged και εδώ φαίνεται να προστίθεται ένα αντίστοιχο και με την κλάση Destroyed. Αυτό αναδεικνύεται και από το υψηλό ποσοστό σφάλματος τύπου I (31%) της κλάσης Damaged, το οποίο ουσιαστικά δηλώνει το ποσοστό των δεδομένων αναφοράς που δεν ταξινομήθηκαν στην κλάση που αντιστοιχούν πραγματικά.

Η κλάση Destroyed από τη μία εμφανίζει υψηλό ποσοστό ορθώς ταξινομημένων σημείων ελέγχου (0.88) αλλά αρκετά χαμηλότερο ποσοστό ορθώς ταξινομημένων συνολικά (0.6). Αυτά σε συνδυασμό και με το πολύ υψηλό σφάλμα τύπου II (40%) υποδηλώνει ότι η κλάση αυτή έχει υπερεκτιμηθεί, δηλαδή έχουν ταξινομηθεί σε αυτή πολύ μεγαλύτερες εκτάσεις από την πραγματική. Άλλωστε αυτό επιβεβαιώνεται και από τα στοιχεία που παρουσιάστηκαν στον πίνακα 21 για τα ποσοστά επικάλυψης μεταξύ δεδομένων CEMS και αποτελεσμάτων ταξινόμησης.

5.1.2 Αποτελέσματα σε επίπεδο αντικειμένου (segment-based)

Για τον υπολογισμό του πίνακα σύγχυσης σε επίπεδο αντικειμένων ελέγχθηκαν τα αντικείμενα ελέγχου σε σχέση με τα ταξινομημένα αντικείμενα. Τα αντικείμενα ελέγχου έχουν προκύψει βάσει των σημειακών δεδομένων ελέγχου που περιλαμβάνουν. Αναλυτική περιγραφή γίνεται στο κεφάλαιο 4, ενότητα 4.1.3.

Πίνακας 25: Πίνακας σύγχυσης σε επίπεδο αντικειμένων για τα αποτελέσματα ταξινόμησης σε περιβάλλον Python.

	CLASSIFIED						
	Κλάσεις:	Not_Damaged	Damaged	Destroyed	Water	Sum	PA
TRUE DATA	Not_Damaged	725	62	0	0	787	0.92
	Damaged	123	510	116	0	749	0.68
	Destroyed	1	17	124	0	142	0.87
	Water	0	0	0	219	219	1.00
	Sum	849	589	240	219	OA:	0.83
UA		0.85	0.86	0.52	1.00	Kappa:	0.75

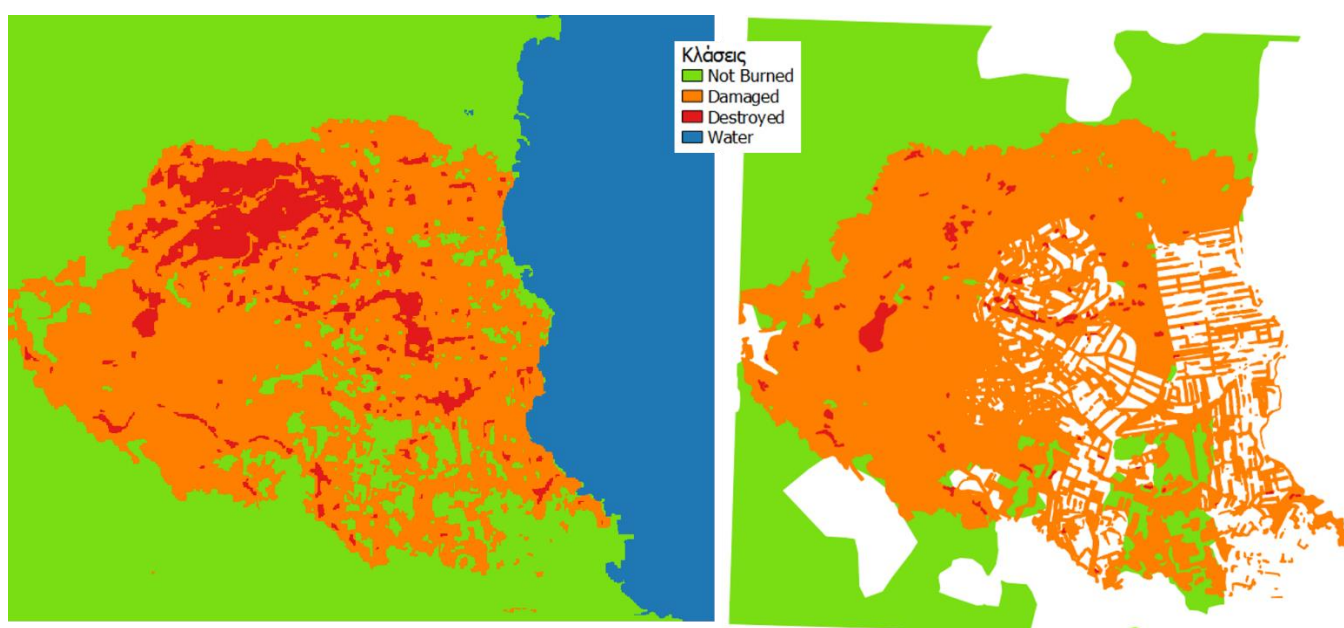
Πίνακας 26: Πίνακας Omission και Commission errors σε επίπεδο αντικειμένων για τα αποτελέσματα ταξινόμησης σε περιβάλλον Python.

Κλάσεις:	Omission Errors (Type I)	Commission Errors (Type II)
Not_Damaged	7,8 %	14,6 %
Damaged	31,9 %	13,4 %
Destroyed	12,7 %	48,3%
Water	0 %	0 %

Από τους παραπάνω πίνακες παρατηρείται ότι οι ακρίβειες των αποτελεσμάτων είτε σε επίπεδο εικονοστοιχείου είτε σε επίπεδο αντικειμένων είναι σχεδόν ίδιες.

5.2 Αποτελέσματα αντικειμενοστρεφούς ταξινόμησης δορυφορικών εικόνων στο λογισμικό eCognition

Η αξιολόγηση των αποτελεσμάτων χαρτογράφησης καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους στο περιβάλλον του λογισμικού eCognition ακολουθεί παρόμοια δομή με αυτή που χρησιμοποιήθηκε και κατά την αξιολόγηση της πρώτης μεθοδολογίας. Επεξηγηματικά, το πρώτο στάδιο είναι ο οπτικός έλεγχος των αποτελεσμάτων της ταξινόμησης συγκριτικά με τα δεδομένα αναφοράς από το CEMS. Εδώ, στο δεύτερο στάδιο αξιολόγησης προστίθενται τα εργαλεία που παρέχει το λογισμικό σχετικά με τη Σταθερότητα Ταξινόμησης (**Classification Stability**) και το Καλύτερο Αποτέλεσμα Ταξινόμησης (**Best Classification Result**). Έπειτα παρουσιάζεται ο πίνακας σύγκυσης σε επίπεδο εικονοστοιχείων (pixel-based) μαζί με τις μετρικές ακρίβειας και τέλος υπολογίζονται οι κοινές εκτάσεις κάθε κλάσης με την αντίστοιχη έκταση από τα πολυγωνικά δεδομένα του CEMS.



Εικόνα 17: Αποτέλεσμα ταξινόμησης στο λογισμικό eCognition (αριστερά) και δεδομένα αναφοράς από το CEMS (δεξιά).

Από τον οπτικό έλεγχο μεταξύ των παραγόμενων προϊόντων από την μεθοδολογία στο eCognition και των δεδομένων αναφοράς από το CEMS, είναι ήδη εμφανές ότι η περίμετρος της συνολικά καμένης έκτασης, ανεξαρτήτου βαθμού καταστροφής, είναι πολύ κοντά στην πραγματική κατάσταση. Αυτό που επίσης είναι εμφανές είναι η υπερεκτίμηση της κλάσης Destroyed. Επεξηγηματικά, από την Εικόνα 17 φαίνεται ότι τα πολύγωνα Destroyed είναι πολύ μικρότερα και λιγότερα στην πραγματικότητα από αυτά που έχουν εξαχθεί κατά τη μεθοδολογία ταξινόμησης με ασαφή λογική. Αντίστοιχο σφάλμα παρατηρείται και στα αποτελέσματα της προηγούμενης μεθοδολογίας. Η δυσκολία διαχωρισμού μεταξύ των κλάσεων Damaged και Destroyed διατυπώθηκε και παραπάνω κατά την ανάλυση των κανόνων ταξινόμησης για την κλάση Destroyed. Κατά τους διάφορους πειραματισμούς που έγιναν δεν βρέθηκαν όρια που να τις διαχωρίζουν σαφώς. Έτσι λοιπόν, έγινε προσπάθεια το τελικό αποτέλεσμα να χαρακτηρίζεται από μία ισορροπία μεταξύ των δύο κλάσεων δεδομένου ότι η δημιουργία κανόνων οι οποίοι να εντοπίζουν όλα τα αντικείμενα Destroyed

θα οδηγούσε στην ταξινόμηση ως Destroyed πολύ μεγάλου αριθμού αντικειμένων που κανονικά ανήκουν στην κλάση Damaged.

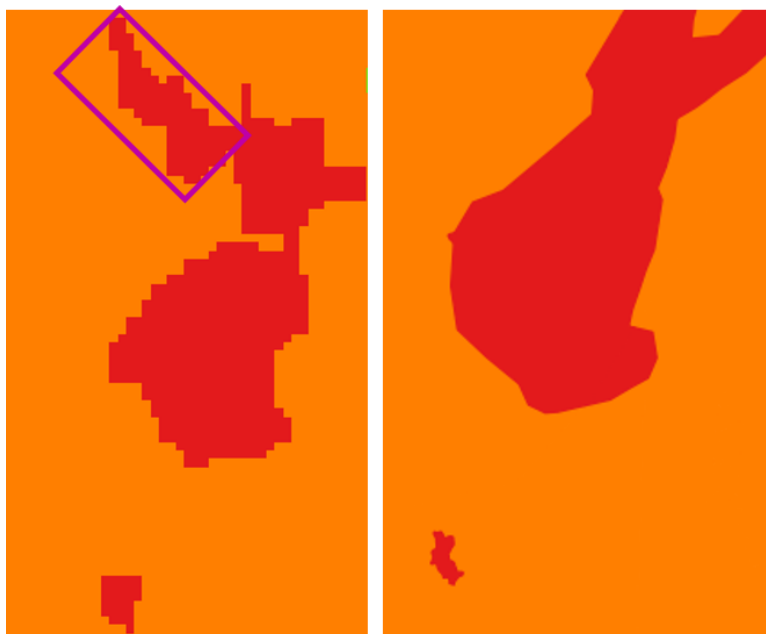


Εικόνα 18: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στις μαύρες ελλείψεις επισημαίνονται οι περιοχές ορθής ταξινόμησης.



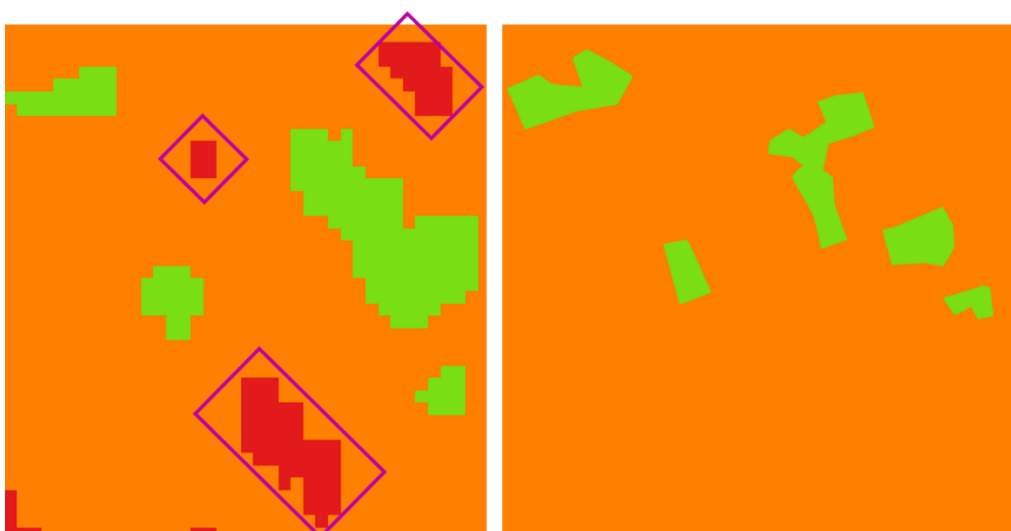
Εικόνα 19: Ενδεικτική περίπτωση ορθών αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά).

Στις εικόνες παραπάνω παρουσιάζονται ενδεικτικά κάποια αποσπάσματα στα οποία φαίνεται ότι το αποτέλεσμα της ταξινόμησης αποτυπώνει σε ικανοποιητικό βαθμό τα δεδομένα αναφοράς CEMS. Η εικόνα 18 αποτελεί τμήμα εντός της περιμέτρου της πυρκαγιάς ενώ η 19 βρίσκεται στα όρια. Παρατηρείται ότι η μεθοδολογία έχει καταφέρει να εντοπίσει ορθά ακόμα και μικρά τμήματα στα οποία η κλάση μεταβάλλεται. Το γεγονός αυτό, λαμβάνοντας υπόψη τη διαφορά της χωρικής ανάλυσης των δεδομένων από τα οποία έχουν προκύψει οι κλάσεις (Pleiades 0.5m – Sentinel-2 10m) επιβεβαιώνει την ανάγκη δημιουργίας μικρών αντικειμένων κατά το 1^ο επίπεδο ταξινόμησης και τελικά την επιτυχία της επιλογής αυτής.



Εικόνα 20: Ενδεικτική περίπτωση αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στο μωβ πλαίσιο επισημαίνονται σφάλματα ταξινόμησης.

Στην εικόνα 20 στο δεξιά απόσπασμα απεικονίζεται η μεγαλύτερη συνεχής έκταση της κλάσης Destroyed από τα δεδομένα αναφοράς. Αριστερά παρουσιάζεται το αποτέλεσμα της ταξινόμησης στο eCognition για αυτή την περιοχή. Χαρακτηριστικό είναι ότι στα δεδομένα CEMS το τμήμα αυτό έχει έκταση 64277m² και το επικαλυπτόμενο τμήμα του αποτελέσματος είναι 53812 m² δηλαδή κατά ένα μεγάλο ποσοστό έχει ταξινομηθεί ορθά. Αξίζει να σημειωθεί όμως ότι στην ίδια εικόνα, στο κάτω μέρος, φαίνεται ότι έχει αποτυπωθεί ορθά και ένα μικρό τμήμα έκτασης Destroyed. Επίσης, στο μωβ πλαίσιο αναδεικνύεται έκταση η οποία έχει ταξινομηθεί ως Destroyed ενώ από την εικόνα των δεδομένων CEMS φαίνεται ότι ανήκει στην κλάση Damaged. Αντίστοιχη περίπτωση παρουσιάζεται και στην εικόνα παρακάτω, στην οποία με μωβ πλαίσιο επισημαίνονται περιοχές οι οποίες έχουν ταξινομηθεί λάθος. Αξίζει να αναφερθεί ότι τα τμήμα Not Burned πλησιάζουν σημαντικά τα δεδομένα αναφοράς (δεξιά).



Εικόνα 21: Ενδεικτική περίπτωση αποτελεσμάτων ταξινόμησης στο eCognition (αριστερά) συγκριτικά με τα δεδομένα αναφοράς CEMS (δεξιά). Στο μωβ πλαίσιο επισημαίνονται σφάλματα ταξινόμησης.

Όπως αναφέρθηκε και στην αρχή αυτής της ενότητας, για την αξιολόγηση της εν λόγω μεθοδολογίας αξιοποιήθηκαν επιπρόσθετα κάποια εργαλεία του eCognition. Πριν παρουσιαστούν τα αποτελέσματα κρίνεται σκόπιμο να περιγραφεί η σημασία αυτών των μέτρων αξιολόγησης. Δεδομένου ότι στην μεθοδολογία αυτή έχει αξιοποιηθεί η ασαφής λογική, πρέπει να λαμβάνεται υπόψη ότι ένα αντικείμενο του οποίου οι τιμές των χαρακτηριστικών βρίσκονται εντός του εύρους της συνάρτησης συμμετοχής, τότε αυτό μπορεί να λαμβάνει βαθμό συμμετοχής σε πάνω από μία τάξεις. Σε αυτό το πλαίσιο, το μέτρο **Classification Stability** δίνει τη διαφορά μεταξύ της καλύτερης και της αμέσως επόμενης ανάθεσης κατηγορίας στα αντικείμενα κάθε κλάσης και παρουσιάζει τα στατιστικά στοιχεία τους. Το μέτρο **Best Classification Result** αντιπροσωπεύει για κάθε κλάση τον μέγιστο βαθμό συμμετοχής των αντικειμένων που ταξινομήθηκαν σε αυτή. Επομένως, τα δύο αυτά εργαλεία αξιολογούν τα αποτελέσματα ως προς τη σταθερότητα και την αξιοπιστία τους.

Επειδή το κυρίως μέρος της ταξινόμησης δομήθηκε στο 1^ο επίπεδο κατάτμησης, όπου αξιοποιήθηκαν τα φασματικά χαρακτηριστικά των αντικειμένων, επιλέχθηκε να παρουσιαστούν τα μέτρα αυτά για τις ταξινομήσεις και στα δύο επίπεδα. Είναι σημαντικό να επισημανθεί ότι η ταξινόμηση στην κλάση Damaged πραγματοποιήθηκε χρησιμοποιώντας το χαρακτηριστικό “Not similarity to class” και όχι άμεσα τα φασματικά χαρακτηριστικά των αντικειμένων.

Classification Stability

Class	Objects	Mean	StdDev	Minimum	Maximum
Not_Burned 1	9113	0.9955749	0.05255854012	0.103	1
Destroyed 1	947	0.975	0.1244780	0.1073228	1

Best Classification Result

Class	Objects	Mean	StdDev	Minimum	Maximum
Not_Burned 1	9113	0.9955749	0.05255854012	0.103	1
Destroyed 1	947	0.975	0.1244780	0.1073228	1

Εικόνα 22: Σταθερότητα κλάσης και καλύτερο αποτέλεσμα ταξινόμησης αντικειμένων επιπέδου 1.

Classification Stability

Class	Objects	Mean	StdDev	Minimum	Maximum
Not_Burned 2	9113	1	0	1	1
Damaged	2240	1	0	1	1
Destroyed 2	947	1	1.09567084e-0...	1	1
Water	1212	1	0	1	1

Best Classification Result

Class	Objects	Mean	StdDev	Minimum	Maximum
Not_Burned 2	9113	1	0	1	1
Damaged	2240	1	0	1	1
Destroyed 2	947	1	1.09567084e-0...	1	1
Water	1212	1	0	1	1

Εικόνα 23: Σταθερότητα κλάσης και καλύτερο αποτέλεσμα ταξινόμησης αντικειμένων επιπέδου 2.

Από την εικόνα 22 προκύπτει ότι η σταθερότητα και η αξιοπιστία της ταξινόμησης των αντικειμένων του 1^{ου} επιπέδου είναι πολύ υψηλές. Αυτά είναι και τα σημαντικότερα μέτρα καθώς στο εν λόγω επίπεδο η ταξινόμηση έγινε αποκλειστικά με βάση τα φασματικά χαρακτηριστικά των αντικειμένων. Οι τιμές της εικόνας 23 επίσης υποδεικνύουν μεγάλη αξιοπιστία της ταξινόμησης των αντικειμένων του επιπέδου 2, γεγονός όμως το οποίο οφείλεται στους κανόνες πάνω στους οποίους αυτή έχει δομηθεί. Επεξηγηματικά, η κλάση Damaged προκύπτει από τη μη ομοιότητα των αντικειμένων με τις υπόλοιπες ενώ οι κλάσεις Not Burned και Destroyed προκύπτουν από την ύπαρξη υπο-αντικειμένων με τις αντίστοιχες κλάσεις, κανόνες που εκ των πραγμάτων μπορεί να πάρει τιμή 0 ή 1 αφού τα αντικείμενα αυτά δεν μεταβλήθηκαν από το ένα επίπεδο στο άλλο.

Στη συνέχεια παρουσιάζεται ο πίνακας σύγκυσης, ο οποίος προέκυψε χρησιμοποιώντας ως δεδομένα αναφοράς τα ίδια σημεία που χρησιμοποιήθηκαν και κατά την αξιολόγηση της μεθοδολογίας που αναπτύχθηκε στη γλώσσα Python.

Πίνακας 27: Πίνακας σύγκυσης σε επίπεδο εικονοστοιχείων των αποτελεσμάτων ταξινόμησης στο περιβάλλον eCognition.

		CLASSIFIED					
Κλάσεις:		Not_Burned	Damaged	Destroyed	Water	Sum	PA
TRUE DATA	Not_Damaged	765	118	0	0	883	0.87
	Damaged	71	695	91	0	857	0.81
	Destroyed	0	91	129	0	220	0.61
	Water	0	0	0	220	220	1.00
	Sum	836	904	220	220	OA:	0.83
UA	0.92	0.77	0.60	1.00	Kappa:	0.75	

Από τον πίνακα 27 προκύπτει ότι για τις κλάσεις Not Burned και Damaged υψηλό ποσοστό των σημείων ελέγχου τους έχουν ταξινομηθεί ορθά (PA = 0.87 και 0.81 αντίστοιχα). Ακόμα πιο υψηλό είναι το ποσοστό των σημείων που έχουν ταξινομηθεί ως Not Burned και πράγματι ανήκουν σε αυτή την κλάση (UA = 0.92). Οι παρατηρήσεις αυτές σε συνδυασμό με το γεγονός ότι δεν παρατηρείται καμία σύγκυση μεταξύ της κλάσης Not Burned και Destroyed, οι οποίες φασματικά απέχουν κατά πολύ, οδηγούν στο συμπέρασμα ότι η κλάση Not Burned έχει στο μεγαλύτερο βαθμό ταξινομηθεί ορθά. Το αντίστροφο φαινόμενο παρατηρείται για την κλάση Damaged, δηλαδή αυτή παρουσιάζει μεγαλύτερο ποσοστό ορθώς ταξινομημένων σημείων ελέγχου (PA = 0.81) και μικρότερο ποσοστό ταξινομημένων που πράγματι ανήκουν σε αυτή (UA = 0.77). Τα ποσοστά αυτά παρατηρείται ότι είναι μειωμένα σε σχέση με αυτά της κλάσης Not Burned γεγονός που τεκμηριώνεται από τη μεγαλύτερη σύγκυση που χαρακτηρίζει την κλάση Damaged. Αναλυτικότερα, από τον πίνακα σύγκυσης προκύπτει ότι 71 και 91 σημεία που στην πραγματικότητα ανήκουν στην κλάση Damaged ταξινομήθηκαν ως Not Burned και Destroyed αντίστοιχα. Αυτό αποτυπώνεται και στο ελαφρώς αυξημένο ποσοστό του omission error (18.9%) της εν λόγω κλάσης.

Η κλάση Destroyed φαίνεται ότι παρουσιάζει τις χαμηλότερες τιμές Producer's Accuracy (PA = 0.61) και User's Accuracy (UA = 0.60). Επίσης παρουσιάζει ιδιαίτερα ψηλά omission και commission errors (41.4% και για τα δύο). Προκύπτει λοιπόν ότι η ταξινόμηση δεν έχει λειτουργήσει θετικά για την κλάση αυτή δεδομένου ότι παρατηρείται μεγάλη σύγκυση με την κλάση Damaged. Σημειώνεται ότι οι προσπάθειες που έγιναν να βελτιωθεί η ακρίβειά της οδήγησε σε μείωση της ακρίβειας για την κλάση Damaged, έτσι τελικά επιλέχθηκε μία μέση κατάσταση και για τις δύο.

Πίνακας 28: Πίνακας Omission και Commission errors σε επίπεδο εικονοστοιχείων για τα αποτελέσματα ταξινόμησης στο περιβάλλον eCognition.

Κλάσεις:	Omission Errors (Type I)	Commission Errors (Type II)
Not_Damaged	13.4 %	8.5 %
Damaged	18.9 %	23.1 %
Destroyed	41.4 %	41.4 %
Water	0 %	0 %

Πίνακας 29: Έκταση και ποσοστά επικάλυψης της κλάσης Not Burned (eCognition) με τα δεδομένα αναφοράς CEMS.

Κλάση 1: "Not_Burned"	
Έκταση κλάσης βάσει δεδομένων CEMS (m ²)	8.903.406,684
Έκταση βάσει προϊόντων eCognition (m ²)	15.485.229,225
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων eCognition (m ²)	8.618.589,35
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	96,8 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition (%)	55,66 %

Από τον παραπάνω πίνακα επιβεβαιώνεται ότι τα αποτελέσματα ταξινόμησης για την κλάση Not Burned διακρίνονται από υψηλή ακρίβεια δεδομένου ότι το ποσοστό κοινής έκτασης της ταξινομημένης αυτής κλάσης με την έκταση των δεδομένων αναφοράς είναι 96,8%. Η ερμηνεία του ποσοστού αυτού είναι παρόμοια με αυτής του μέτρου Producer's Accuracy, καθώς δηλώνει πόση έκταση των δεδομένων αναφοράς κατάφερε να βρεθεί κατά την ταξινόμηση. Από την άλλη, το ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων της ταξινόμησης μπορεί να παραλληλισθεί με το μέτρο User's Accuracy, αφού δηλώνει πόση έκταση από αυτή που ταξινομήθηκε στην εν λόγω κλάση ανήκει πράγματι σε αυτή. Παρατηρείται ότι το ποσοστό αυτό είναι χαμηλό (55,66%) και αυτό οφείλεται στην ύπαρξη τμημάτων των δεδομένων αναφοράς του CEMS που δεν έχουν χαρακτηρισμό για την κλάση τους. Το ζήτημα αυτό έχει περιγραφεί αναλυτικά στο κεφάλαιο 3 αλλά και στην αξιολόγηση των αποτελεσμάτων με την προηγούμενη μεθοδολογία.

Πίνακας 30: Έκταση και ποσοστά επικάλυψης της κλάσης Damaged (eCognition) με τα δεδομένα αναφοράς CEMS.

Κλάση 2: "Damaged"	
Έκταση κλάσης βάσει δεδομένων CEMS (m ²)	10.610.595,332
Έκταση βάσει προϊόντων eCognition (m ²)	11.148.891,337
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων eCognition (m ²)	8.410.568,841
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	79,26 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition (%)	75,44 %

Αναφορικά με την κλάση Damaged φαίνεται ότι από την πραγματική έκτασή της, η μεθοδολογία ταξινόμησης στο eCognition καταφέρνει να εντοπίσει αρκετά υψηλό ποσοστό

(79,26%). Το ποσοστό αυτό μάλιστα είναι πολύ κοντινό με την τιμή του μέτρου PA (0.81) που προέκυψε από τον πίνακα σύγκυσης. Τεκμηρίωση για το ποσοστό αυτό, το οποίο δεν θεωρείται χαμηλό, αποτελεί το φαινόμενο ταξινόμησης τμημάτων της κλάσης αυτής στην κλάση Destroyed αλλά και στην Not Burned. Παρατηρείται ότι το ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition επίσης προσεγγίζει το μέτρο UA (0.77). Οι παρατηρήσεις αυτές ενισχύουν την αξιοπιστία των μέτρων αυτών (PA, UA), δεδομένου ότι ο έλεγχος για το σύνολο της έκτασης και όχι μόνο για τα σημεία ελέγχου παράγει αντίστοιχα αποτελέσματα.

Πίνακας 31: Έκταση και ποσοστά επικάλυψης της κλάσης Destroyed (eCognition) με τα δεδομένα αναφοράς CEMS.

Κλάση 3: "Destroyed"	
Έκταση κλάσης βάσει δεδομένων CEMS (m ²)	190.631,149
Έκταση βάσει προϊόντων eCognition (m ²)	1.852.963,916
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων eCognition (m ²)	126.225,247
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	66,21 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition (%)	6,81 %

Από τον παραπάνω πίνακα αρχικά παρατηρείται ότι και τα δυο ποσοστά είναι χαμηλά, με το ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition να είναι ακραία χαμηλό (6,81%). Το τελευταίο οφείλεται στην υπερεκτίμηση της κλάσης Destroyed, δηλαδή όπως φαίνεται η πραγματική έκταση της είναι 190.631,149m² ενώ η ταξινόμηση έχει οδηγήσει στον χαρακτηρισμό ως Destroyed πολύ μεγαλύτερης έκτασης (1.852.963,916 m²). Αναφορικά με το ποσοστό της πραγματικής έκτασης που καταφέρνει να εντοπίσει ο αλγόριθμος αυτό είναι 66,21 %, λίγο μεγαλύτερο από το μέτρο PA (0.61). Το σχετικά χαμηλό αυτό ποσοστό οφείλεται στην ταξινόμηση ενός σημαντικού μέρος της κλάσης Destroyed ως Damaged, όπως φάνηκε και στον πίνακα σύγκυσης παραπάνω. Γενικά μεταξύ των δύο αυτών κλάσεων παρατηρείται σύγκυση και λαμβάνοντας υπόψη και την συνολικά μικρή έκταση της κλάσης Destroyed, αυτό οδηγεί σε σημαντική μείωση της ακρίβειάς της.

Πίνακας 32: Έκταση και ποσοστά επικάλυψης αθροιστικά για τις κλάσεις Damaged και Destroyed (eCognition) με τα δεδομένα αναφοράς CEMS.

Καμένη έκταση ανεξαρτήτου βαθμού καταστροφής (Damaged & Destroyed)	
Έκταση κατηγορίας βάσει δεδομένων CEMS (m ²)	10.801.226,48
Έκταση βάσει προϊόντων eCognition (m ²)	13.001.855,253
Έκταση κοινών τμημάτων δεδομένων αληθείας και προϊόντων eCognition (m ²)	10.257.505,344
Ποσοστό κοινών τμημάτων επί των δεδομένων CEMS (%)	94,97 %
Ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition (%)	78,89 %

Όπως και κατά την αξιολόγηση της προηγούμενης μεθοδολογίας, παρουσιάζονται και εδώ τα ποσοστά επικάλυψης της ταξινομημένης ως καμένης έκτασης, δηλαδή των κλάσεων

Damaged και Destroyed, ως προς την αντίστοιχη έκταση των δεδομένων αναφοράς. Προκύπτει ότι το ποσοστό είναι ιδιαίτερα υψηλό (94,97) αποδεικνύοντας ότι η μεθοδολογία καταφέρνει να εντοπίσει με μεγάλη ακρίβεια τις καμένες εκτάσεις. Φαίνεται όμως ότι η μεθοδολογία έχει ταξινομήσει ως καμένη έκταση συνολικά αρκετά μεγαλύτερη έκταση ($13.001.855,253\text{m}^2$ αντί $10.801.226,48\text{m}^2$) για αυτό και το χαμηλότερο ποσοστό κοινών τμημάτων επί των παραχθέντων δεδομένων eCognition. Το φαινόμενο αυτό εν μέρει οφείλεται στην μη ύπαρξη δεδομένων αναφοράς για κάποια τμήματα της εικόνας.

Αφού λοιπόν έχουν παρουσιαστεί και αναλυθεί τα αποτελέσματα και για τις δύο μεθοδολογίες χαρτογράφησης καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους, στο επόμενο και τελευταίο κεφάλαιο γίνεται η σύγκρισή τους και εξάγονται συμπεράσματα.

6 | Συμπεράσματα & προτάσεις

Στο τελευταίο αυτό κεφάλαιο συνοψίζονται τα συμπεράσματα που προκύπτουν από την εφαρμογή και αξιολόγηση των παραπάνω μεθοδολογιών, τα οποία με τη σειρά τους αξιοποιούνται για την δημιουργία προτάσεων βελτιστοποίησης και εξέλιξης της παρούσας μελέτης.

6.1 Συμπεράσματα

Στην ενότητα αυτή, αρχικά γίνεται μία σύγκριση των αποτελεσμάτων που προέκυψαν από τις δύο διαφορετικές μεθοδολογίες. Οι αναλύσεις που έγιναν στο προηγούμενο κεφάλαιο σχετικά με την ακρίβεια κάθε ταξινόμησης τροφοδοτούν τη διαδικασία σύγκρισης, η οποία με τη σειρά της, συνδυαστικά με τις αναλύσεις, οδηγούν σε μία σειρά συμπερασμάτων. Τα συμπεράσματα αυτά σχετίζονται κατά κύριο λόγο με τους στόχους της εργασίας, οι οποίοι περιγράφονται στο εισαγωγικό κεφάλαιο.

	CLASSIFIED						
	Κλάσεις:	Not_Burned	Damaged	Destroyed	Water	Sum	PA
TRUE DATA	Not_Damaged	809	74	0	0	883	0.92
	Damaged	137	591	129	0	857	0.69
	Destroyed	1	26	193	0	220	0.88
	Water	0	0	0	220	220	1.00
	Sum	947	691	322	220	OA:	0.83
UA	0.85	0.86	0.60	1.00	Kappa:	0.75	

	CLASSIFIED						
	Κλάσεις:	Not_Burned	Damaged	Destroyed	Water	Sum	PA
TRUE DATA	Not_Damaged	765	118	0	0	883	0.87
	Damaged	71	695	91	0	857	0.81
	Destroyed	0	91	129	0	220	0.61
	Water	0	0	0	220	220	1.00
	Sum	836	904	220	220	OA:	0.83
UA	0.92	0.77	0.60	1.00	Kappa:	0.75	

Εικόνα 24: Πίνακες σύγχυσης σε επίπεδο εικονοστοιχείων για τη μεθοδολογία σε προγραμματιστικό περιβάλλον Pythοn (πάνω) και στο περιβάλλον eCognition (κάτω).

Προκειμένου να διευκολυνθεί η διαδικασία σύγκρισης, στην εικόνα 24 παρουσιάζονται οι πίνακες των δύο διαφορετικών ταξινομήσεων. Οι ακρίβειες των δύο μεθόδων συνολικά είναι ίδιες. Συγκεκριμένα το Overall Accuracy είναι ίδιο και ίσο με 0,83 ενώ ο συντελεστής kappa είναι 0.75. Παρ' όλα αυτά οι ακρίβειες των κλάσεων, πέραν της κλάσης Water που είναι 100%, διαφέρουν σημαντικά. Παρατηρείται από τα αποτελέσματα και των δύο μεθοδολογιών ότι για τις κλάσεις Damaged και Destroyed, όταν βελτιώνεται η ακρίβεια της μίας μειώνεται η ακρίβεια της άλλης, γεγονός που θα αναλυθεί και στη συνέχεια.

Πιο αναλυτικά, για την κλάση Not Burned τα αποτελέσματα με την μέθοδο στο περιβάλλον της Pythοn παρουσιάζουν μεγαλύτερο True Positive (807 αντί 765 βάσει του eCognition) και άρα μεγαλύτερο Producer's Accuracy (0.92) ενώ η μέθοδος στο eCognition εμφανίζει μεγαλύτερο User's Accuracy (0.92). Στην πρώτη μεθοδολογία παρόλο που ο αριθμός των σημείων ελέγχου που ταξινομήθηκαν ορθά είναι μεγαλύτερος, παρατηρείται ότι το UA είναι χαμηλότερο από ότι στη δεύτερη καθώς έχουν ταξινομηθεί περισσότερα σημεία.

Λαμβάνοντας υπόψη και τα Omission (8.4% και 14.6%) και Commission errors (13.4% και 8.5%) αυτής της κλάσης για τις δύο μεθόδους, μπορεί να διατυπωθεί το συμπέρασμα ότι η ακρίβειά τους είναι εξίσου ικανοποιητική, με την μεν πρώτη μέθοδο να εμφανίζει υψηλότερο PA και χαμηλότερο UA ενώ η δεύτερη χαμηλότερο PA και μεγαλύτερο UA.

Σχετικά με την κλάση Damaged, η μέθοδος στο eCognition παρουσιάζει βελτιωμένα αποτελέσματα με PA=0.81 (αντί 0.69 στη Python). Και εδώ όμως, παρατηρείται να συμβαίνει το αντίθετο για το μέτρο UA, όπου στην μέθοδο με Python είναι υψηλότερο (0.86) από το αντίστοιχο με τη μέθοδο στο eCognition (0.77). Τελικά όμως, δεδομένου ότι η ταξινόμηση στο eCognition εμφανίζει μία σταθερότητα μεταξύ PA και UA οπότε κρίνεται καταλληλότερη για την εν λόγω κλάση.

Ο αλγόριθμος που αναπτύχθηκε στη Python έχει σημαντικά μεγαλύτερο PA (0.87 αντί 0.61 βάσει του eCognition) για την κλάση Destroyed ενώ το UA είναι ίδιο (0.60). Λαμβάνοντας υπόψη και τα ποσοστά επικάλυψης της ταξινομημένης κλάσης Destroyed βάσει της πρώτης μεθόδου προς την πραγματική έκταση (85.74%), το οποίο είναι σημαντικό υψηλότερο από αυτό βάσει της δεύτερης (66.21%), φαίνεται ότι η μεθοδολογία που αναπτύχθηκε στη γλώσσα Python ταξινομεί με καλύτερη ακρίβεια την κλάση αυτή.

Συμπερασματικά λοιπόν, από τη μία μεριά η μέθοδος που αναπτύχθηκε στη γλώσσα Python ταξινομεί ως Destroyed αρκετά μεγαλύτερο τμήμα από το πραγματικό και έτσι βρίσκει πράγματι το μεγαλύτερο μέρος της έκτασης Destroyed (υψηλό PA), το οποίο όμως λειτουργεί εις βάρος της κλάσης Damaged, από την οποία πολλά εικονοστοιχεία τελικά ταξινομούνται και αυτά ως Destroyed, μειώνοντας σημαντικά το PA για την κλάση Damaged. Από την άλλη η μέθοδος στο eCognition ταξινομεί μικρότερο μέρος της εικόνας ως Destroyed και παρουσιάζει PA και UA για την κλάση αυτή 0.61 και 0.6 αντίστοιχα, αλλά έτσι η κλάση Damaged έχει αυξημένη ακρίβεια.

Σχετικά με τις κλάσεις Damaged και Destroyed προκύπτει το συμπέρασμα ότι δεν είναι εύκολα διαχωρίσιμες ούτε με εφαρμογή επιβλεπόμενης ταξινόμησης ούτε με χρήση κανόνων. Αυτό σε συνδυασμό και με παρατηρήσεις από το εργαλείο “**Feature view**” οδηγεί στο συμπέρασμα ότι τ, α φασματικά χαρακτηριστικά των κλάσεων αυτών δεν είναι εύκολα διαχωρίσιμα. Πέραν της διαχρονικής διαφοράς του δείκτη NBR, δηλαδή τον dNBR, βρέθηκε ότι ο δείκτης αυτός πριν το γεγονός της φωτιάς αλλά και μετά φαίνεται χρήσιμο στον εν μέρει διαχωρισμό αυτών των δύο κλάσεων.

Όλα τα παραπάνω αφορούν στη καταλληλότητα των δύο μεθοδολογιών να αξιοποιηθούν για τη χαρτογράφηση καμένων εκτάσεων και εκτίμησης του βαθμού καταστροφής τους, η οποία αποτελεί και τον βασικό στόχο της παρούσας εργασίας. Πέραν αυτού όμως έχουν τεθεί και κάποιοι άλλοι στόχοι, όπως η διερεύνηση του αποτελέσματος στην περίπτωση συλλογής δεδομένων αληθείας με ημι-αυτόματο τρόπο αλλά και η τυποποίηση των διαδικασιών για την εύκολη αξιοποίησή τους και σε άλλες εφαρμογές.

Αναφορικά με την ημι-αυτόματη μέθοδο συλλογής του ground truth, η οποία έχει περιγραφεί αναλυτικά στο κεφάλαιο 3, φαίνεται πως δεν επηρεάζει αρνητικά τη διαδικασία ταξινόμησης στο περιβάλλον της Python. Αυτό προκύπτει λαμβάνοντας υπόψη ότι τα αποτελέσματα και των δύο μεθοδολογιών έχουν την ίδια συνολική ακρίβεια (OA=0.83) και

τον ίδιο συντελεστή k ($k=0.75$) ενώ στη μέθοδο που αναπτύχθηκε στο eCognition τα δεδομένα αληθείας δεν συμμετέχουν καθόλου ως δεδομένα εκπαίδευση στην ταξινόμηση, αφού αυτή βασίζεται σε κανόνες. Επομένως, οι όποιες αστοχίες της πρώτης μεθόδου (Pythou) δεν αποδίδονται στη συλλογή των δεδομένων εκπαίδευσης καθώς αντίστοιχες αστοχίες παρατηρούνται και κατά την δεύτερη (eCognition).

Όσον αφορά στην τυποποίηση των διαδικασιών για την εύκολη μεταφορά τους σε άλλες εφαρμογές κρίνεται ότι η μεθοδολογία στη γλώσσα Pythou υπερτερεί δεδομένου ότι ο χρήστης πρέπει απλώς να συλλέξει τα δεδομένα αληθείας και να εφαρμόσει τον αλγόριθμο στη συνέχεια. Επιπλέον, τα μηνύματα που έχουν ενταχθεί στον αλγόριθμο προσφέρουν στο χρήστη άμεση επαφή με τα στάδια που ολοκληρώνονται και τον ενημερώνουν με σκοπό να έχει πλήρη εποπτεία τυχόν σφαλμάτων κατά τη διαδικασία. Από την άλλη, η μέθοδος στο eCognition απαιτεί τον ορισμό του εύρους για τις συναρτήσεις συμμετοχής, διαδικασία η οποία απαιτεί πολλούς πειραματισμούς καθώς οι τιμές μεταβάλλονται σημαντικά ανά περιοχή αλλά και ανά δορυφορική εικόνα, ιδίως όταν πρόκειται και για διαφορετικό αισθητήρα. Πάντως, πέραν αυτού, κρίνεται ότι η ιεραρχική δομή των κλάσεων και τα χαρακτηριστικά τα οποία αξιοποιούνται για τη δημιουργία των συναρτήσεων συμμετοχής μπορούν να μεταφερθούν και σε άλλη εφαρμογή.

6.2 Προτάσεις

Λαμβάνοντας υπόψη τα συμπεράσματα παραπάνω, αλλά και την γνώση που αποκτήθηκε καθ' όλη τη διάρκεια εκπόνησης της εργασίας, ακολουθούν ορισμένες προτάσεις που πιθανώς να μπορούν να βελτιώσουν τα αποτελέσματα που προέκυψαν και δίνουν δυνατότητες επέκτασης της παρούσας εργασίας:

- ❖ Εφαρμογή ανάλυσης διαχωρισιμότητας των κλάσεων (Separability analysis) προκειμένου να επιλεγεί το βέλτιστο δυνατό σετ δεδομένων (*φασματικά κανάλια και δείκτες*) προς αντικειμενοστρεφή ταξινόμηση.
- ❖ Εφαρμογή των μεθόδων σε δορυφορικά δεδομένα πολύ υψηλής ανάλυσης.
- ❖ Τροποποίηση του αλγορίθμου στη γλώσσα Pythou προκειμένου να δέχεται ως δεδομένα εκπαίδευσης πολύγωνα τα οποία να διατηρεί αυτούσια στη διαδικασία και να μην τα ενσωματώνει στα ήδη υπάρχοντα αντικείμενα. Αξιοποίηση αλγορίθμου Region growing για την δημιουργία των πολυγώνων εκπαίδευσης.
- ❖ Τροποποίηση αλγορίθμου στη γλώσσα Pythou ώστε να υπολογίζει περισσότερα χαρακτηριστικά των αντικειμένων.
- ❖ Μελέτη και εφαρμογή πιο εξελιγμένων αλγορίθμων για την επιβλεπόμενη ταξινόμηση σε γλώσσα Pythou όπως Νευρωνικά Δίκτυα.
- ❖ Αξιοποίηση δεδομένων αναφοράς που έχουν συλλεχθεί από μετρήσεις πεδίου.
- ❖ Ο περαιτέρω διαχωρισμός των κλάσεων ώστε να αποτυπώνονται οι διαφορετικοί βαθμοί καταστροφής που περιέχει η κλάση Damaged.

Βιβλιογραφία

- Achanta, R. και συν., 2012. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 2274 - 2282.
- Belenguer-Plomer, M. A., Tanase, M. A., Chuvieco, E. & Bovolo, F., 2021. CNN-based burned area mapping using radar and optical data. *Remote Sensing of Environment*, Τόμος 260.
- Blaschke, T., 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, pp. 2-16.
- Chuvieco, E., Martín, M. P. & Palacios, A., 2002. Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination. *International Journal of Remote Sensing*, 23(23), pp. 5103-5110.
- Clewley, D. και συν., 2014. A Python-Based Open Source System for Geographic Object-Based Image Analysis (GEOBIA) Utilizing Raster Attribute Tables. *Remote Sensing*, pp. 6111-6135.
- Comert, R., Matci, D. K. & Avdan, U., 2019. OBJECT BASED BURNED AREA MAPPING WITH RANDOM FOREST ALGORITHM. *International Journal of Engineering and Geosciences (IJEG)*, Τόμος 4, pp. 078-087.
- eCognition, T., 2014. *eCognition Developer 9.0 Reference Book*. s.l.:Trimble.
- Filipponi, F., 2018. BAIS2: Burned Area Index for Sentinel-2. *2nd International Electronic Conference on Remote Sensing*, Τόμος 2.
- Filipponi, F., 2019. Exploitation of Sentinel-2 Time Series to Map Burned Areas at the National Level: A Case Study on the 2017 ItalyWildfires. *Remote Sensing*.
- Georgopoulos, N., Stavrakoudis, D. & Gitas, I. Z., 2019. OBJECT-BASED BURNED AREA MAPPING USING SENTINEL-2 IMAGERY AND SUPERVISED LEARNING GUIDED BY EMPIRICAL RULES. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 9980-9983.
- Hafen, K., 2020. *OpenSourceOptions*. [Ηλεκτρονικό]
Available at: <https://opensourceoptions.com>
- Kavzoglu, T., Erdemir, M. Y. & Tonbul, H., 2016. Evaluating performances of spectral indices for burned area mapping using object-based image analysis. *12th International symposium on spatial accuracy assessment in natural resources and environmental sciences*, 5(8).
- Kovacs, K. D., 2019. EVALUATION OF BURNED AREAS WITH SENTINEL-2 USING SNAP: THE CASE OF KINETA AND MATI, GREECE, JULY 2018. *Geographia Technica*, 14(2), pp. 20-38.
- Mallinis, G. & Koutsias, N., 2012. Comparing ten classification methods for burned area mapping in a Mediterranean environment using Landsat TM satellite data. *International Journal of Remote Sensing*, Τόμος 33, pp. 4408-4433.

Mallinis, G., Mitsopoulos, I. & Chrysafi, I., 2018. Evaluating and comparing Sentinel 2A and Landsat-8 Operational Land Imager (OLI) spectral indices for estimating fire severity in a Mediterranean pine ecosystem of Greece. *GIScience & Remote Sensing*, pp. 1-18.

Parks, S. A., Dillon, G. K. & Miller, C., 2014. A New Metric for Quantifying Burn Severity: The Relativized Burn Ratio. *Remote Sensing*, 6(3), pp. 1827-1844.

Pinto, M. M., Trigo, R. M., Trigo, I. F. & DaCamara, C. C., 2021. A Practical Method for High-Resolution Burned Area Monitoring Using Sentinel-2 and VIIRS. *Remote Sensing*.

Polychronaki, A. & Gitas, I. Z., 2012. Burned Area Mapping in Greece Using SPOT-4 HRVIR Images and Object-Based Image Analysis. *Remote Sensing*, pp. 424-438.

Shokouhi, B. V. & Eslami, M., 2019. FUZZY LOGIC BASED BURNED SEVERITY CLASSIFICATION AND MAPPING WITHLANDSAT-8 DATA. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 259-265.

Stumpf, A. & Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. *Object-oriented mapping of landslides using Random Forests*, pp. 2564-2577.

Αντωνίου, Γ., 2017. ΧΑΡΤΟΓΡΑΦΗΣΗ ΚΑΛΥΨΗΣ ΓΗΣ ΜΕ ΧΡΗΣΗ ΔΕΔΟΜΕΝΩΝ SENTINEL-2 ΑΠΟ ΠΟΛΛΑΠΛΕΣ ΗΜΕΡΟΜΗΝΙΕΣ ΜΕΣΩ ΜΗΧΑΝΩΝ ΔΙΑΝΥΣΜΑΤΙΚΗΣ ΥΠΟΣΤΗΡΙΞΗΣ. *Διπλωματική Εργασία*.

Καρακίζη, Χ., 2015. ΕΝΤΟΠΙΣΜΟΣ ΑΜΠΕΛΟΤΕΜΑΧΙΩΝ, ΕΞΑΓΩΓΗ ΓΡΑΜΜΩΝ ΦΥΤΕΥΣΗΣ ΚΑΙ ΔΙΑΧΩΡΙΣΜΟΣ ΠΟΙΚΙΛΙΩΝ ΑΜΠΕΛΟΥ ΣΕ ΔΟΡΥΦΟΡΙΚΑ ΔΕΔΟΜΕΝΑ ΠΟΛΥ ΥΨΗΛΗΣ ΑΝΑΛΥΣΗΣ. *Μεταπτυχιακή Εργασία*.

Πατεράκης, Π. Ι., 2010. Ανίχνευση και Εξαγωγή του Οδικού Δικτύου με Αντικειμενοστρεφή Ανάλυση Πολυφασματικών Απεικονίσεων Quickbird. *Διπλωματική εργασία*.