ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΔΠΜΣ «ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ»

Ανίχνευση κτιρίων με χρήση βαθιάς μάθησης

Building detection using deep learning

Διπλωματική εργασία της

Άννα Ατζαράκη

Επιβλέπων:
Δημήτρης Αργιαλάς

Αθήνα, Οκτώβριος 2022

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΔΠΜΣ «ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ»

Ανίχνευση κτιρίων με χρήση βαθιάς μάθησης

Building detection using deep learning

Διπλωματική εργασία
της

Άννα Ατζαράκη

Επιβλέπων:
Δημήτρης Αργιαλάς

Εγκρίθηκε από την τριμελή επιτροπή την 31 Οκτωβρίου 2022.

(*Υπογραφή*)	(*Υπογραφή*)	(*Υπογραφή*)
……	…….	……..……
Δημήτρης Αργιαλάς Καθηγητής Ε.Μ.Π.	Πολυχρόνης Κολοκούσης Διδάκτωρ Ε.Μ.Π.	Κωνσταντίνος Καράντζαλος Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2022

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Περίληψη 9

Abstract 11

1. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ 13
1.1. Τεχνητή νοημοσύνη 13
1.1.1. Μηχανική μάθηση 14
1.1.2. Βαθιά μάθηση 15
1.2. Ανίχνευση αντικειμένων 23
1.2.1. Ανίχνευση με βαθιά μάθηση 25
1.2.2. Μοντέλο Single-Shot Detector 27
1.3. Μεταφορά γνώσης 28
1.3.1. Τεχνική εκμάθηση μεταφοράς ResNet-34 29
1.4. Ακρίβεια ανίχνευσης αντικειμένων 29

2. ΠΕΡΙΟΧΗ ΜΕΛΕΤΗΣ 33
2.1. Γεωγραφικά χαρακτηριστικά Βρίσας 33

3. ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ 34

4. ΔΕΔΟΜΕΝΑ 35
4.1. Λήψη δεδομένων 35
4.2. Βίντεο 35
4.2.1. Α’ Βίντεο 35
4.2.2. Β’ Βίντεο 35

5. ΜΕΘΟΔΟΛΟΓΙΑ 36
5.1. Προετοιμασία συνόλου δεδομένων 36
5.1.1. Εξαγωγή εικόνων 36
5.1.1. Οπτικός έλεγχος 37
5.1.2. Ποιοτικός έλεγχος εικόνων 37
5.1.3. Προσανατολισμός φωτογραφιών 40
5.1.4. Προετοιμασία & Εισαγωγή των σημείων ελέγχου εδάφους (Γεωαναφορά) 42
5.1.5. Βελτιστοποίηση προσανατολισμού 44
5.1.6. Δημιουργία πυκνού νέφους σημείων 45
5.1.7. Εξαγωγή ορθοφωτογραφίας 47

5.2. Διαμόρφωση λειτουργικού συστήματος 49
5.3. Δημιουργία δειγμάτων εκπαίδευσης 51
5.3.1. Δημιουργία εκπαιδευτικού σχήματος 51
5.3.2. Συλλογή δειγμάτων εκπαίδευσης 52
5.3.3. Δημιουργία patch εικόνων 53
5.4. Ανίχνευση κτιρίων με χρήση μοντέλου βαθιάς μάθησης 56
5.4.1. Εκπαίδευση μοντέλου βαθιάς μάθησης 56
5.4.2. Εντοπισμός κτιρίων με χρήση μοντέλου βαθιάς μάθησης 61
5.5. Υπολογισμός ακρίβειας ανίχνευσης κτιρίων 64

6. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΑΞΙΟΛΟΓΗΣΗ 66
6.1. Αποτελέσματα εκπαίδευσης μοντέλων βαθιάς μάθησης 66
6.1.1. Αποτελέσματα εκπαίδευσης πρώτου μοντέλου 66
6.1.2. Αποτελέσματα εκπαίδευσης δεύτερου μοντέλου 68
6.1.3. Αποτελέσματα εκπαίδευσης τρίτου μοντέλου 70
6.2. Αποτελέσματα ανίχνευσης κτιρίων 72
6.2.1. Αποτελέσματα ανίχνευσης κτιρίων πρώτου μοντέλου 72
6.2.2. Αποτελέσματα ανίχνευσης κτιρίων δεύτερου μοντέλου 76
6.2.3. Αποτελέσματα ανίχνευσης κτιρίων τρίτου μοντέλου 80
6.3. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων 84
6.3.1. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων πρώτου μοντέλου 84
6.3.2. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων δεύτερου μοντέλου 88
6.3.3. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων τρίτου μοντέλου 92

7. ΣΥΖΗΤΗΣΗ - ΣΥΜΠΕΡΑΣΜΑΤΑ 96

8. ΒΙΒΛΙΟΓΡΑΦΙΑ 99

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Μερικοί ορισμοί της τεχνητής νοημοσύνης, οργανωμένοι σε τέσσερις κατηγορίες. (Russel & Norvig, 2003) 13
Εικόνα 2: Τεχνητή νοημοσύνη, μηχανική μάθηση, βαθιά μάθηση (ESRI, χ.χ.) 15
Εικόνα 3: Ένα νευρωνικό δίκτυο παραμετροποιείται από τα βάρη του. (Chollet, 2017) 17
Εικόνα 4: Μια λειτουργία απώλειας μετρά την ποιότητα της εξόδου του δικτύου. (Chollet, 2017) 18
Εικόνα 5: Η βαθμολογία απώλειας χρησιμοποιείται ως σήμα ανάδρασης για την προσαρμογή των βαρών. (Chollet, 2017) 19
Εικόνα 6: Παράδειγμα σημασιολογικής κατάτμησης. (ESRI, χ.χ.) 20
Εικόνα 7: Παράδειγμα κατάτμησης στιγμιοτύπων. (ESRI, χ.χ.) 20
Εικόνα 8: Παράδειγμα πανοπτικής κατάτμησης. (ESRI, χ.χ.) 21
Εικόνα 9: Παράδειγμα μετάφρασης εικόνας. (ESRI, χ.χ.) 21
Εικόνα 10: Παράδειγμα ανίχνευσης αλλαγών. (ESRI, χ.χ.) 22
Εικόνα 11: Παράδειγμα ανίχνευσης αντικειμένων. (ESRI, χ.χ.) 22
Εικόνα 12: Παράδειγμα ταξινόμησης εικόνας. (ESRI, χ.χ.) 23
Εικόνα 13: Διαφορά μεταξύ ταξινόμησης και ανίχνευσης αντικειμένων. (ArcGIS Developers, χ.χ.) 24
Εικόνα 14: Χρήση συνελικτικού νευρωνικού δικτύου για ανίχνευση αντικειμένων. (Pathak et al., 2018) 26
Εικόνα 15: Architecture of a convolutional neural network with a SSD detector. (Liu et al., 2016) 28
Εικόνα 16: Παράδειγμα πρόβλεψης ανίχνευσης αντικειμένου. (ESRI, χ.χ.) 30
Εικόνα 17: Ο λόγος Intersection over Union (IoU). (ESRI, χ.χ.) 32
Εικόνα 18: Μετατροπή βίντεο σε εικόνες. 37
Εικόνα 19: Εισαγωγή των εικόνων στο Agistoft Metashape Professional. 38
Εικόνα 20: Χρήση “Estimate Image Quality”. 39
Εικόνα 21: Υπολογισμός ποιότητας εικόνας. 39
Εικόνα 22: Αποκλεισμός εικόνων με ακατάλληλη ποιότητα. 40
Εικόνα 23: Χρήση “ Align Photos”. 41
Εικόνα 24: Ορισμός της παραμέτρου ακρίβειας. 41
Εικόνα 25: Αποτέλεσμα προσανατολισμού φωτογραφιών. 42

Εικόνα 26: Εντοπισμός σημείων ελέγχου στο έδαφος 43
Εικόνα 27: Χρήση "Add Marker". 43
Εικόνα 28: Τοποθέτηση σημείων ελέγχου στο μοντέλο. 44
Εικόνα 29: Χρήση "Optimize Cameras". 44
Εικόνα 30: Προκαθορισμένες επιλογές λογισμικού. 45
Εικόνα 31: Υπολογισμός σφάλματος. 45
Εικόνα 32: Χρήση "Build Dense Cloud". 46
Εικόνα 33: Ορισμός της παραμέτρου ποιότητας. 47
Εικόνα 34: Δημιουργία πυκνού νέφους σημείων για τη Βρίσα. 47
Εικόνα 35: Εξαγωγή ορθοφωτογραφίας. 48
Εικόνα 36: Ορθοφωτοχάρτης 26-03-2019. 48
Εικόνα 37: Βιβλιοθήκες βαθιάς μάθησης για ArcGIS. Walbridge et al., 2021 49
Εικόνα 38: Λήψη του πακέτου βιβλιοθηκών για το ArcGIS Pro. 50
Εικόνα 39:Εγκατάσταση του πακέτου βιβλιοθηκών για το ArcGIS Pro. 50
Εικόνα 40: Εισαγωγή της ορθοφωτογραφίας στο λογισμικό. 50
Εικόνα 41: Εργαλείο ετικετών αντικειμένων. 52
Εικόνα 42: Εισαγωγή κλάσης και καταχώρηση τιμής. 52
Εικόνα 43: Τα δείγματα εκπαίδευσης που συλλέχθηκαν. 53
Εικόνα 44: Ορισμός παραμέτρων για την δημιουργία των τσίπ. 54
Εικόνα 45: Ορισμός παραμέτρων για την δημιουργία των τσίπ. 55
Εικόνα 46: Αποτέλεσμα της εξαγωγής των patch εικόνας και τα μεταδεδομένα τους. 55
Εικόνα 47: Μέρος των patch εικόνας. 56
Εικόνα 48: Ορισμός παραμέτρων για εκπαίδευση του πρώτου (A) μοντέλου βαθιάς μάθησης. 58
Εικόνα 49: Ορισμός παραμέτρων για εκπαίδευση του δεύτερου (B) μοντέλου βαθιάς μάθησης. 59
Εικόνα 50: Ορισμός παραμέτρων για εκπαίδευση του τρίτου (C) μοντέλου βαθιάς μάθησης. 60
Εικόνα 51: Ορισμός παραμέτρων για την πρώτη εφαρμογή εντοπισμού κτιρίων του πρώτου (A) μοντέλου βαθιάς μάθησης. 64
Εικόνα 52: Υπολογισμός ακρίβειας για την πρώτη εφαρμογή εντοπισμού κτιρίων του πρώτου (A) μοντέλου βαθιάς μάθησης. 65
Εικόνα 53: Εκπαίδευση A μοντέλου βαθιάς μάθησης. 66

Εικόνα 54: Εκπαίδευση Β μοντέλου βαθιάς μάθησης. 68
Εικόνα 55: Εκπαίδευση C μοντέλου βαθιάς μάθησης. 70
Εικόνα 56: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_1. 72
Εικόνα 57: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_2. 73
Εικόνα 58: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_3. 74
Εικόνα 59: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_4. 75
Εικόνα 60: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_1. 76
Εικόνα 61: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_2. 77
Εικόνα 62: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_3. 78
Εικόνα 63: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_4. 79
Εικόνα 64: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_1. 80
Εικόνα 65: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_2. 81
Εικόνα 66: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_3. 82
Εικόνα 67: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_4. 83

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Συγκέντρωση επιπέδων που χρησιμοποιούνται για την ανίχνευση αντικειμένων. (Pathak et al., 2018) 27
Πίνακας 2: Μετρήσεις ακρίβειας πρόβλεψης της ανίχνευσης αντικειμένου. (ESRI, χ.χ.) 30
Πίνακας 3: Χαρακτηριστικά ακρίβειας πρόβλεψης της ανίχνευσης αντικειμένου. (ESRI, χ.χ.) 32
Πίνακας 4: Μεταδεδομένα Α' πτήσης βίντεο. 35
Πίνακας 5: Μεταδεδομένα Β' πτήσης βίντεο. 35
Πίνακας 6: Τα τρία μοντέλα εκπαίδευσης με διαφορετικές παραμέτρους. 57
Πίνακας 7: Εφαρμογή ανίχνευσης των τριών μοντέλων με διαφορετικές παραμέτρους. 61
Πίνακας 8: Ακρίβεια ανίχνευσης κτιρίων πειράματος Α_1 84
Πίνακας 9: Ακρίβεια ανίχνευσης κτιρίων πειράματος Α_2 85
Πίνακας 10: Ακρίβεια ανίχνευσης κτιρίων πειράματος Α_3 86
Πίνακας 11: Ακρίβεια ανίχνευσης κτιρίων πειράματος Α_4 87
Πίνακας 12: Ακρίβεια ανίχνευσης κτιρίων πειράματος Β_1 88
Πίνακας 13: Ακρίβεια ανίχνευσης κτιρίων πειράματος Β_2 89
Πίνακας 14: Ακρίβεια ανίχνευσης κτιρίων πειράματος Β_3 90
Πίνακας 15: Ακρίβεια ανίχνευσης κτιρίων πειράματος Β_4 91
Πίνακας 16: Ακρίβεια ανίχνευσης κτιρίων πειράματος C_1 92
Πίνακας 17: Ακρίβεια ανίχνευσης κτιρίων πειράματος C_2 93
Πίνακας 18: Ακρίβεια ανίχνευσης κτιρίων μοντέλου C_3. 94
Πίνακας 19: Ακρίβεια ανίχνευσης κτιρίων μοντέλου C_4. 95
Πίνακας 20: Συνολικά αποτελέσματα εκπαίδευσης. 96

Περίληψη

Οι εφευρέτες ονειρευόντουσαν από καιρό να δημιουργήσουν μηχανές που σκέφτονται. Αυτή η επιθυμία χρονολογείται από πίσω τουλάχιστον στην εποχή της αρχαίας Ελλάδας. Οι μυθικές μορφές Πυγμαλίων, Δαίδαλος, και ο Ήφαιστος μπορούν όλοι να ερμηνευθούν ως θρυλικοί εφευρέτες και η Γαλάτεια, ο Τάλως, και η Πανδώρα μπορούν όλοι να θεωρηθούν ως τεχνητή ζωή (Ovid & Martin, 2004, Sparkes, 1996; Tandy, 1997).

Τα τελευταία χρόνια, ο τομέας της τεχνητής νοημοσύνης παρουσιάζει σημαντική ανάπτυξη και ιδιαίτερα στις τεχνικές βαθιάς μάθησης οι οποίες καλούνται να αντιμετωπίσουν πληθώρα προβλημάτων ως προς την ανίχνευση αντικειμένων.

Η παρούσα εργασία εμβαθύνει στην αυτόματη ανίχνευση κτιρίων και στην ανάλυση της μελέτης αυτής για τον εντοπισμό των κτιρίων μέσω τεχνικών βαθιάς μάθησης.

Η εργασία διεκπεραιώθηκε με το λογισμικό ArcGIS Pro της εταιρίας ESRI, διεθνής προμηθευτής λογισμικού συστημάτων γεωγραφικών πληροφοριών, διαδικτυακών GIS και εφαρμογών διαχείρισης βάσεων γεωγραφικών δεδομένων. Στο περιβάλλον του λογισμικού αναπτύχθηκε μελέτη με βάση τη βαθιά μάθηση για την ανίχνευση κτιρίων. Ως δεδομένο για την επίτευξη των στόχων της εργασίας, χρησιμοποιήθηκε μια αεροφωτογραφία έτους 2019 και αποτελεί προϊόν της προπτυχιακής μου εργασίας. Δημιουργήθηκε έπειτα από δεδομένα που συλλέχθηκαν μέσω πτήσης UAS ενώ περιοχή μελέτης είναι η Βρίσα Λέσβου.

Η ανάλυση περιλάμβανε τη δημιουργία δειγμάτων εκπαίδευσης και την εκπαίδευση μοντέλων βαθιάς μάθησης με κύριο στόχο την ανίχνευση των κτιρίων της περιοχής.

Συνολικά δημιουργήθηκαν – ψηφιοποιήθηκαν 280 δείγματα εκπαίδευσης κτιρίων.

Για τις ανάγκες της εργασίας, εκπαιδεύτηκαν τρία μοντέλα με βαθιά μάθηση σύμφωνα με διαφορετικές παραμέτρους. Για την ανίχνευση των κτιρίων πραγματοποιήθηκαν δώδεκα πειράματα, τρία για κάθε μοντέλο βαθιάς μάθησης.

Πιο συγκεκριμένα η παρούσα μελέτη εστιάζει στην ορθή ανίχνευση των κτιρίων, εξετάζει τι αποτελεί κτίριο στην ορθοφωτογραφία σύμφωνα με τα δείγματα που συλλέχθηκαν.

Η επικύρωση των αποτελεσμάτων πραγματοποιήθηκε υπολογίζοντας την ακρίβεια με βάση τα δείγματα εκπαίδευσης και τα δώδεκα πειράματα - ανιχνεύσεις που εφαρμόστηκαν. Ως αποτέλεσμα η διαδικασία αυτή είχε μετρητικές ακρίβειας οι οποίες αξιολογήθηκαν και με την βοήθεια φωτοερμηνείας της αεροφωτογραφίας.

Τα αποτελέσματα της αξιολόγησης για το πρώτο μοντέλο εκπαίδευσης είχαν ικανοποιητικά αποτελέσματα καθώς το πείραμα από τα τέσσερα με την μεγαλύτερη ακρίβεια ήταν εκείνο με ακρίβεια 69%.

Το δεύτερο μοντέλο εκπαίδευσης είχε αξιόλογα αποτελέσματα καθώς και τα τέσσερα πειράματα είχαν ακρίβεια πάνω από 60% με το τέταρτο πείραμα να φτάνει 82% ακρίβεια οδηγώντας στο καλύτερο αποτέλεσμα της μελέτης.

Το μοντέλο με την χειρότερη ακρίβεια ήταν το τρίτο μοντέλο εκπαίδευσης το οποίο παρουσίασε εσφαλμένη ανίχνευση. Τα τέσσερα πειράματα είχαν ακρίβεια μικρότερη από 2% με το πρώτο και δεύτερο πείραμα να λαμβάνουν μηδενικά ποσοστά.

Συγκρίνοντας τα μοντέλα μεταξύ τους διαπιστώθηκε ότι διπλασιάζοντας το μέγεθος της παρτίδας το μοντέλο αποδίδει καλύτερα. Όσο μικρότερο μέγεθος παρτίδας δέχεται το μοντέλο βαθιάς μάθησης τόσο πιο εσφαλμένα είναι τα αποτελέσματα του, αντίστοιχα όσο μεγαλύτερο μέγεθος της παρτίδας λαμβάνει το μοντέλο τόσο πιο αποδοτικό είναι. Το μέγεθος της παρτίδας εξετάζει την ακρίβεια της εκτίμησης της διαβάθμισης σφάλματος κατά την εκπαίδευση του νευρωνικού δικτύου.
Συνοψίζοντας η παρούσα εργασία, προβάλει πως η αυτόματη ανίχνευση σε ένα χώρο με ένα συγκεκριμένο αντικείμενο ανίχνευσης μπορεί να φέρει εξαιρετικά αποτελέσματα λαμβάνοντας υπόψη συγκεκριμένες παραμέτρους.
Η μέθοδος ανίχνευσης με βαθιά μάθηση είναι μια αξιόπιστη μέθοδος όσον αφορά την ανίχνευση αντικειμένων με απαραίτητη προϋπόθεση τη μελέτη και τη παραμετροποίηση της μεθόδου όσον αφορά το λογισμικό εργασίας.

Η εργασία διακρίνεται σε έξι ενότητες. Στη πρώτη ενότητα πραγματοποιείται η ανάλυση του θεωρητικού υπόβαθρου στην ανίχνευση κτιρίων και τη βαθιά μάθηση. Στη δεύτερη ενότητα παρουσιάζεται η περιοχή μελέτης, η Βρίσα Λέσβου. Οι επόμενες τρείς ενότητες, προβάλλουν τις μεθόδους που χρησιμοποιήθηκαν, τα δεδομένα αλλά και την επεξεργασία τους. Η εργασία ολοκληρώνεται με τα αποτελέσματα και τα συμπεράσματα της μελέτης.

Λέξεις – κλειδιά: Βαθιά μάθηση, τεχνητή νοημοσύνη, συνελικτικά νευρωνικά δίκτυα, ανίχνευση αντικειμένων, ανίχνευση κτιρίων, ResNet34, SSD, δείγματα εκπαίδευσης

Abstract

Inventors have long dreamed of creating machines that think. This desire dates back to at least the time of ancient Greece. The mythical figures Pygmalion, Daedalus, and Hephaestus may all be interpreted as legendary inventors, and Galatea, Talos, and Pandora may all be regarded as artificial life. (Ovid & Martin, 2004, Sparkes, 1996; Tandy, 1997).

Over the last years, the study of artificial intelligence shows significant growth, especially regarding deep learning techniques which are called to deal with a multitude of problems in terms of object recognition.

This paper delves into the automatic building detection and the study of the results for the localization of buildings. The work was carried out using ArcGIS Pro software from ESRI (an international software supplier of geographic information systems, online GIS and geographic database management applications). Deep learning-based study for building detection was developed in this software environment.

As a source data to achieve the goals of this study an aerial photo from 2019 was used, which was a product of my undergraduate studies. It was created with data collected through a drone flight and the area of study is Vrisa in Lesvos island in Greece.

The analysis involved generating training samples and training deep learning models with the main goal of detecting building in this area.

A total of 280 training samples were created and digitized.

For the needs of this study, three deep learning models were trained according to different parameters. Twelve experiments were performed to detect the buildings, three for each deep learning model. More specifically, in this study we focus on the correct buildings detection and the notion of what constitutes a building in the orthoimagery according to the samples collected. The validation of the results was performed by calculating the accuracy based on the training samples on the twelve experiments - detections applied.

As a result, this process had accuracy metrics that were also evaluated with the help of photo interpretation of aerial photography. The process evaluation for the first training model has satisfactory results as the experiment out of the four with the highest accuracy was the one with that scored 69%.

The second training model had remarkable results as all four experiments had over 60% accuracy with the fourth experiment reaching 82% leading to the best result of the study.

The model with the worst accuracy was the third training model which showed a false detection. All four experiments had an accuracy of less than 2% with the first and second experiments receiving zero percentages.

Upon comparing the models, it was found that by doubling the batch size, better performance is achived. When the deep learning model receives a smaller batch size, the results are incorrect. However, with a bigger batch size, a more efficient result is achieved.

The batch size examines the accuracy of the estimation of error gradient during the training of the neural network training.To summarize this study, we can see how

automatic detection in a controlled space using a specific detection object can bring excellent results considering specific parameters. Deep learning detection method is reliable in object detection with the prerequisite of studying and parameterizing the method in terms of working software.

The study is divided into six sections. The first section analyzes the theoretical background in building recognition through deep learning. The second section presents the study area, Vrisa. The next three sections analyses the methods used, the data and their processing. The paper concludes with the results and conclusions of the study.

Keywords: Deep learning, artificial intelligent, Convolutional Neural Network, object detection, building detection, ResNet34, SSD, training samples

1. ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΣΚΟΠΗΣΗ

1.1. Τεχνητή νοημοσύνη

Η Τεχνητή Νοημοσύνη, που ορίζεται ευρέως (και κάπως κυκλικά), ασχολείται με την ευφυή συμπεριφορά στα αντικείμενα. Η ευφυής συμπεριφορά, με τη σειρά της, περιλαμβάνει την αντίληψη, τη συλλογιστική, τη μάθηση, την επικοινωνία και τη δράση σε πολύπλοκα περιβάλλοντα. Η τεχνητή νοημοσύνη έχει ως έναν από τους μακροπρόθεσμους στόχους της την ανάπτυξη μηχανών που μπορούν να κάνουν αυτά τα πράγματα όπως οι άνθρωποι μπορούν ή ενδεχομένως και καλύτερα. Ένας άλλος στόχος της Τεχνητής Νοημοσύνης είναι να κατανοήσει αυτό το είδος συμπεριφοράς είτε συμβαίνει σε μηχανές είτε σε ανθρώπους ή άλλα ζώα. (Nilsson, 1998)
Η τεχνητή νοημοσύνη καθιστά τις μηχανές ικανές να 'κατανοούν' το περιβάλλον τους, να επιλύουν προβλήματα και να δρουν προς την επίτευξη ενός συγκεκριμένου στόχου. Ο υπολογιστής λαμβάνει δεδομένα (ήδη έτοιμα ή συλλεγμένα μέσω αισθητήρων, π.χ. κάμερας), τα επεξεργάζεται και ανταποκρίνεται βάσει αυτών. (Ευρωπαϊκό Κοινοβούλιο, 2020)

Σύμφωνα με τους Russell και Norvig (2003), η Τεχνητή Νοημοσύνη ορίζεται και διαμορφώνεται μέσα από οκτώ εγχειρίδια όπως φαίνεται στην παρακάτω εικόνα.

Συστήματα που σκέπτονται σαν τον άνθρωπο	**Συστήματα που σκέπτονται ορθολογικά**
«Η συναρπαστική νέα προσπάθεια για να κάνουμε τους υπολογιστές να σκέπτονται … *μηχανές με νόηση*, με την πλήρη και κυριολεκτική έννοια.» (Haugeland, 1985) «[Η αυτοματοποίηση των] δραστηριοτήτων που συσχετίζουμε με την ανθρώπινη σκέψη, όπως η λήψη αποφάσεων, η επίλυση προβλημάτων, η μάθηση…» (Bellman, 1978)	«Η μελέτη των νοητικών ικανοτήτων με τη χρήση υπολογιστικών μοντέλων.» (Charniak & McDermott, 1985) «Η μελέτη των υπολογιστικών εργασιών που μας δίνουν τη δυνατότητα να αντιλαμβανόμαστε, να συλλογιζόμαστε, και να ενεργούμε.» (Winston, 1992)
Συστήματα που ενεργούν σαν τον άνθρωπο	**Συστήματα που ενεργούν ορθολογικά**
«Η τέχνη της δημιουργίας μηχανών που πραγματοποιούν λειτουργίες οι οποίες απαιτούν νοημοσύνη όταν πραγματοποιούνται από ανθρώπους.» (Kurzweil, 1990) «Η μελέτη του πώς μπορούμε να κάνουμε τους υπολογιστές να κάνουν πράγματα στα οποία, προς το παρόν, οι άνθρωπο είναι καλύτεροι.» (Rich & Knight, 1991)	«Υπολογιστική Νοημοσύνη είναι η μελέτη της σχεδίασης ευφυών πρακτόρων.» (Poole κ.α., 1998) «Η τεχνητή νοημοσύνη ασχολείται με τη ευφυή συμπεριφορά των τεχνουργημάτων.» (Nilsson, 1998)

Εικόνα 1: Μερικοί ορισμοί της τεχνητής νοημοσύνης, οργανωμένοι σε τέσσερις κατηγορίες. (Russel & Norvig, 2003)

Χαρακτηριστικά, αναφέρουν πως οι παραπάνω ορισμοί αυτοί έχουν διαφορές σε δύο κύριες διαστάσεις. Χονδρικά, οι επάνω ορισμοί ενδιαφέρονται περισσότερο για τις διαδικασίες σκέψης και τη συλλογιστική, ενώ οι κάτω ορισμοί ασχολούνται με τη συμπεριφορά. Οι ορισμοί στα αριστερά μετρούν την επιτυχία με βάση την εγγύτητα προς τις ανθρώπινες επιδόσεις, ενώ οι ορισμοί στα δεξιά τη μετρούν σε σχέση με μια ιδανική έννοια νοημοσύνης, την οποία ονομάζουν ορθολογικότητα. Ένα σύστημα είναι ορθολογικό αν κάνει «το σωστό», με δεδομένα όσα γνωρίζει. (Russell & Norvig, 2003)

1.1.1. Μηχανική μάθηση

Οι δυσκολίες που αντιμετωπίζουν τα συστήματα που βασίζονται σε σύνθετες κωδικοποιημένες γνώσεις υποδηλώνουν ότι τα συστήματα τεχνητής νοημοσύνης χρειάζονται την ικανότητα να αποκτούν τις δικές τους γνώσεις, εξάγοντας μοτίβα από ανεπεξέργαστα δεδομένα. Αυτή η δυνατότητα είναι γνωστή ως Μηχανική Μάθηση. Η εισαγωγή της μηχανικής μάθησης επέτρεψε στους υπολογιστές να αντιμετωπίσουν προβλήματα που αφορούν τη γνώση του πραγματικού κόσμου και να λάβουν αποφάσεις που φαίνονται υποκειμενικές. Ένας απλός αλγόριθμος μηχανικής μάθησης που ονομάζεται λογιστική παλινδρόμηση μπορεί να καθορίσει εάν θα συστήσει καισαρική τομή. (Mor-Yosef et al., 1990)

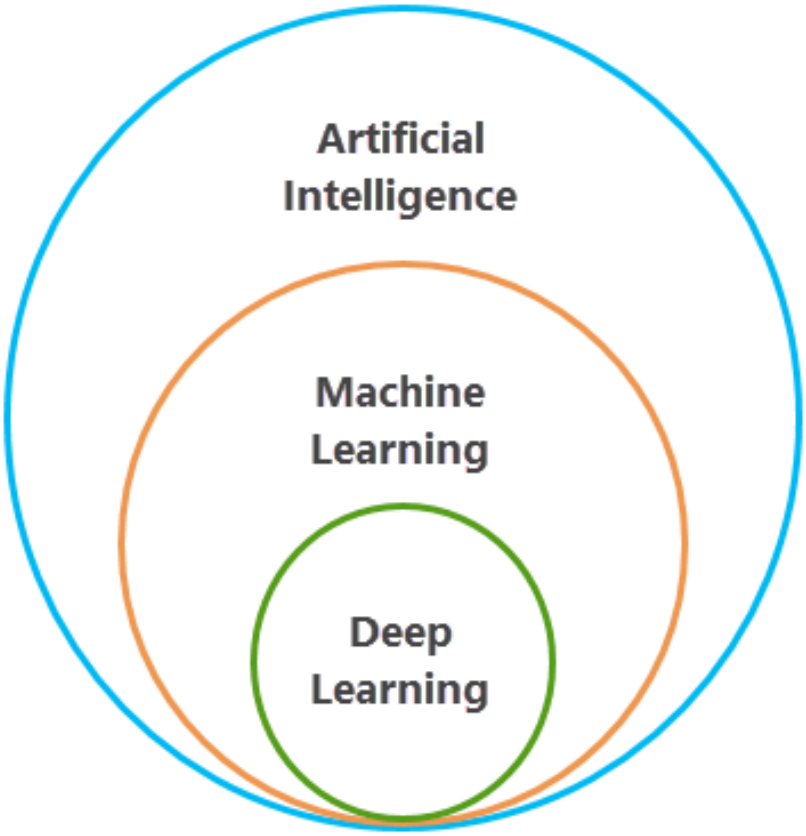
Η Μηχανική Μάθηση (Machine learning - ML) είναι ένα πεδίο της Τεχνητής Νοημοσύνης (Artificial Intelligence - AI) που δημιουργεί συστήματα με την ικανότητα να μαθαίνουν αυτόματα και να βελτιώνονται χωρίς να είναι ρητά προγραμματισμένα για τον υπολογισμό ή την λύση προβλημάτων. Οι αλγόριθμοι του ML επιτρέπουν στους υπολογιστές να εκπαιδεύονται πάνω στα δεδομένα εισόδου και χρησιμοποιούν στατιστική ανάλυση για να εξάγουν τιμές οι οποίες εμπίπτουν σε ένα συγκεκριμένο εύρος. Η διαδικασία της μάθησης ξεκινά με παρατηρήσεις, που αποτελούν παραδείγματα, ή εμπειρικά αποτελέσματα ή οδηγίες, ούτως ώστε να αναγνωριστούν πρότυπα στα δεδομένα και να ληφθούν καλύτερες αποφάσεις στο μέλλον, με βάση τα παραδείγματα που διαθέτουμε. Ο πρωταρχικός σκοπός είναι να επιτρέψουμε στους υπολογιστές να μαθαίνουν αυτόματα, χωρίς ανθρώπινη παρέμβαση ή βοήθεια, και να προσαρμόζουν τις πράξεις τους κατάλληλα. Στο ML, τα καθήκοντα ταξινομούνται γενικά σε ευρείες κατηγορίες. Οι κατηγορίες αυτές βασίζονται στον τρόπο με τον οποίο λαμβάνεται η μάθηση ή στον τρόπο με τον οποίο δίνεται ανάδραση στην εκμάθηση στο ανεπτυγμένο σύστημα. Δύο από τις πιο ευρέως υιοθετημένες μεθόδους ML είναι η επιβλεπόμενη μάθηση, η οποία εκπαιδεύει αλγορίθμους που βασίζονται στα δεδομένα εισόδου και εξόδου τα οποία επισημαίνονται (αποκτούν ετικέτες-labels) από τον άνθρωπο και η μη επιβλεπόμενη μάθηση, η οποία παρέχει τον αλγόριθμο χωρίς επισημασμένα δεδομένα, ούτως ώστε να του επιτρέψει να βρει δομή στα δεδομένα εισόδου του. (Χρονοπούλου, 2019)

Η μηχανική μάθηση είναι ένας κλάδος της τεχνητής νοημοσύνης στον οποίο τα δομημένα δεδομένα υποβάλλονται σε επεξεργασία με έναν αλγόριθμο για την επίλυση ενός προβλήματος. Τα παραδοσιακά δομημένα δεδομένα απαιτούν από ένα άτομο να επισημάνει τα δεδομένα, όπως μια εικόνα γατών και σκύλων, έτσι ώστε τα ειδικά χαρακτηριστικά για κάθε τύπο ζώου να μπορούν να γίνουν κατανοητά στον αλγόριθμο

και να χρησιμοποιηθούν για την αναγνώριση αυτών των ζώων σε άλλες εικόνες. (ESRI, χ.χ.)

1.1.2. Βαθιά μάθηση

Η βαθιά μάθηση είναι ένα υποσύνολο της μηχανικής μάθησης που χρησιμοποιεί διάφορα επίπεδα αλγορίθμων με τη μορφή νευρωνικών δικτύων. Τα δεδομένα εισόδου αναλύονται μέσω διαφορετικών επιπέδων του δικτύου, με κάθε επίπεδο να καθορίζει συγκεκριμένα χαρακτηριστικά και μοτίβα στα δεδομένα. Για παράδειγμα, εάν θέλει κάποιος να προσδιορίσει χαρακτηριστικά όπως κτίρια και δρόμους, το μοντέλο βαθιάς μάθησης μπορεί να εκπαιδευτεί με εικόνες διαφορετικών κτιρίων και δρόμων, να επεξεργαστεί τις εικόνες μέσω επιπέδων εντός του νευρικού δικτύου και, στη συνέχεια, να βρει τα αναγνωριστικά που απαιτούνται για την ταξινόμηση ενός κτιρίου ή δρόμου. (ESRI, χ.χ.)



Εικόνα 2: Τεχνητή νοημοσύνη, μηχανική μάθηση, βαθιά μάθηση (ESRI, χ.χ.)

Η βαθιά μάθηση είναι ένας τύπος μηχανικής μάθησης και τεχνητής νοημοσύνης (AI) που μιμείται τον τρόπο με τον οποίο οι άνθρωποι αποκτούν ορισμένους τύπους γνώσης. Η βαθιά μάθηση είναι ένα σημαντικό στοιχείο της επιστήμης των δεδομένων, το οποίο περιλαμβάνει στατιστική και προγνωστική μοντελοποίηση. Είναι εξαιρετικά επωφελές για τους επιστήμονες δεδομένων που είναι επιφορτισμένοι με τη συλλογή, την ανάλυση και την ερμηνεία μεγάλων ποσοτήτων δεδομένων. Η βαθιά μάθηση κάνει αυτή τη διαδικασία ταχύτερη και ευκολότερη.

Στην απλούστερη μορφή της, η βαθιά μάθηση μπορεί να θεωρηθεί ως ένας τρόπος αυτοματοποίησης της προγνωστικής ανάλυσης. Ενώ οι παραδοσιακοί αλγόριθμοι μηχανικής μάθησης είναι γραμμικοί, οι αλγόριθμοι βαθιάς μάθησης στοιβάζονται σε μια ιεραρχία αυξανόμενης πολυπλοκότητας και αφαίρεσης.
Για να κατανοήσει κανείς τη βαθιά μάθηση αρκεί να φανταστεί ένα μικρό παιδί του οποίου η πρώτη λέξη είναι σκύλος. Το μικρό παιδί μαθαίνει τι είναι και τι δεν είναι σκύλος δείχνοντας αντικείμενα και λέγοντας τη λέξη σκύλος. Ο γονιός λέει: «Ναι, αυτός είναι σκύλος» ή« Όχι, αυτός δεν είναι σκύλος». Καθώς το μικρό παιδί συνεχίζει να δείχνει αντικείμενα, συνειδητοποιεί περισσότερο τα χαρακτηριστικά που διαθέτουν όλα τα σκυλιά. Αυτό που κάνει το μικρό παιδί, χωρίς να το γνωρίζει, είναι να αποσαφηνίσει μια πολύπλοκη αφαίρεση -- την έννοια του σκύλου -- χτίζοντας μια ιεραρχία στην οποία κάθε επίπεδο αφαίρεσης δημιουργείται με γνώση που αποκτήθηκε από το προηγούμενο επίπεδο της ιεραρχίας. (Burns & Brush, 2021)

Πως λειτουργεί η βαθιά μάθηση;

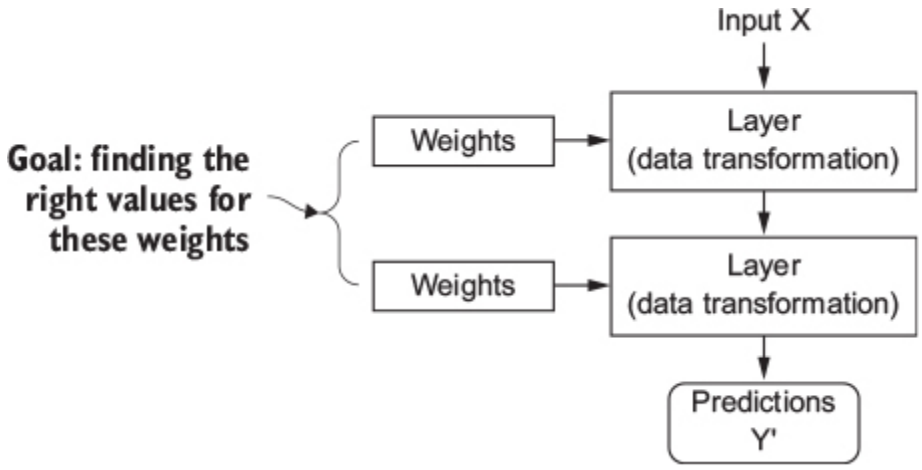
Σύμφωνα με άρθρο της IBM Cloud (2020), τα νευρωνικά δίκτυα βαθιάς μάθησης, ή τεχνητά νευρωνικά δίκτυα, προσπαθούν να μιμηθούν τον ανθρώπινο εγκέφαλο μέσω ενός συνδυασμού εισροών δεδομένων, βαρών και προκατάληψης. Αυτά τα στοιχεία συνεργάζονται για την ακριβή αναγνώριση, ταξινόμηση και περιγραφή αντικειμένων μέσα στα δεδομένα.
Τα βαθιά νευρωνικά δίκτυα αποτελούνται από πολλαπλά επίπεδα διασυνδεδεμένων κόμβων, καθένα από τα οποία βασίζεται στο προηγούμενο επίπεδο για να βελτιώσει και να βελτιστοποιήσει την πρόβλεψη ή την κατηγοριοποίηση. Αυτή η εξέλιξη των υπολογισμών μέσω του δικτύου ονομάζεται διάδοση προς τα εμπρός. Τα επίπεδα εισόδου και εξόδου ενός βαθιού νευρωνικού δικτύου ονομάζονται ορατά στρώματα. Το επίπεδο εισόδου είναι το σημείο όπου το μοντέλο βαθιάς μάθησης προσλαμβάνει τα δεδομένα για επεξεργασία και το επίπεδο εξόδου είναι το σημείο όπου γίνεται η τελική πρόβλεψη ή ταξινόμηση.
Μια άλλη διαδικασία που ονομάζεται backpropagation χρησιμοποιεί αλγόριθμους, όπως η κάθοδος κλίσης, για να υπολογίσει σφάλματα στις προβλέψεις και στη συνέχεια προσαρμόζει τα βάρη και τις προκαταλήψεις της συνάρτησης μετακινώντας προς τα πίσω μέσα από τα στρώματα σε μια προσπάθεια να εκπαιδεύσει το μοντέλο. Μαζί, η διάδοση προς τα εμπρός και η οπισθοπορεία επιτρέπουν σε ένα νευρωνικό δίκτυο να κάνει προβλέψεις και να διορθώσει ανάλογα τυχόν σφάλματα. Με την πάροδο του χρόνου, ο αλγόριθμος γίνεται σταδιακά πιο ακριβής.
Τα παραπάνω περιγράφουν τον απλούστερο τύπο βαθιού νευρωνικού δικτύου με τους απλούστερους όρους. Ωστόσο, οι αλγόριθμοι βαθιάς μάθησης είναι απίστευτα περίπλοκοι και υπάρχουν διαφορετικοί τύποι νευρωνικών δικτύων για την αντιμετώπιση συγκεκριμένων προβλημάτων ή συνόλων δεδομένων. Για παράδειγμα

- Τα συνελικτικά νευρωνικά δίκτυα (CNN), που χρησιμοποιούνται κυρίως σε εφαρμογές υπολογιστικής όρασης και ταξινόμησης εικόνων, μπορούν να ανιχνεύσουν χαρακτηριστικά και μοτίβα μέσα σε μια εικόνα, επιτρέποντας

εργασίες, όπως ανίχνευση ή αναγνώριση αντικειμένων. Το 2015, ένα CNN κέρδισε για πρώτη φορά έναν άνθρωπο σε μια πρόκληση αναγνώρισης αντικειμένων.

- Τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) χρησιμοποιούνται συνήθως σε εφαρμογές αναγνώρισης φυσικής γλώσσας και ομιλίας, καθώς αξιοποιούν δεδομένα διαδοχικών ή χρονικών σειρών.

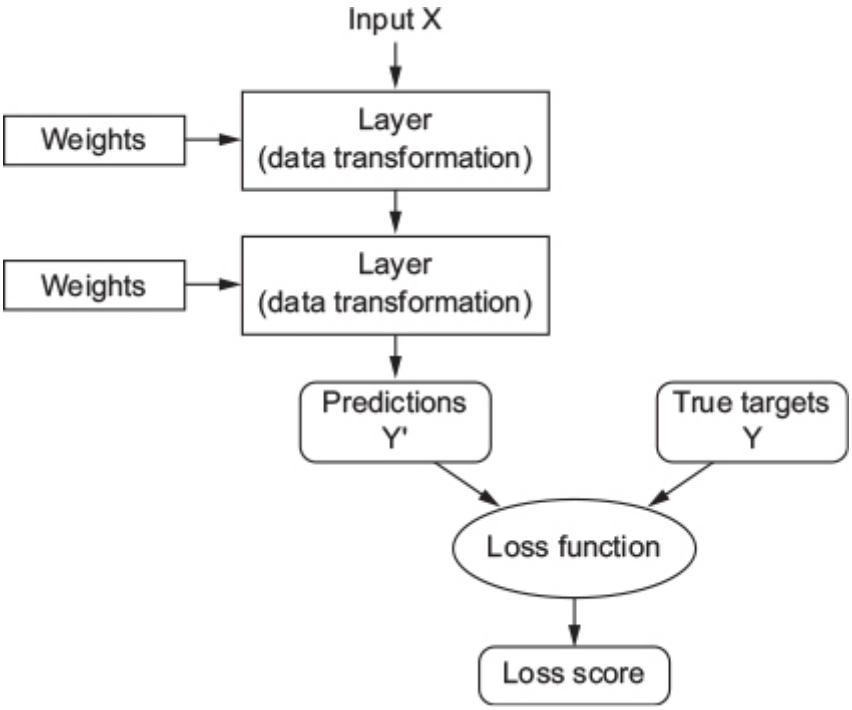
Σύμφωνα με τους Chollet και Allaire (2018), γνωρίζοντας ότι η μηχανική μάθηση αφορά τη χαρτογράφηση εισόδων (όπως εικόνες) σε στόχους (όπως η ετικέτα "γάτα"), η οποία γίνεται παρατηρώντας πολλά παραδείγματα εισόδων και στόχων γνωρίζεται επίσης ότι τα βαθιά νευρωνικά δίκτυα κάνουν αυτή τη χαρτογράφηση εισόδου-στόχου μέσω μιας βαθιάς ακολουθίας απλών μετασχηματισμών δεδομένων (επίπεδα) και ότι αυτοί οι μετασχηματισμοί δεδομένων μαθαίνονται με έκθεση σε παραδείγματα.
Ο προσδιορισμός του τι κάνει ένα στρώμα στα δεδομένα εισόδου του αποθηκεύεται στα βάρη του στρώματος, τα οποία στην ουσία είναι μια δέσμη αριθμών. Σε τεχνικούς όρους, λέγοντας ότι ο μετασχηματισμός που υλοποιείται από ένα στρώμα παραμετροποιείται από τα βάρη του (weights)(Βλ. Εικόνα 3) (Τα βάρη ονομάζονται επίσης μερικές φορές οι παράμετροι ενός στρώματος.) Σε αυτό το πλαίσιο, μάθηση σημαίνει εύρεση ενός συνόλου τιμών για τα βάρη όλων των επιπέδων σε ένα δίκτυο, έτσι ώστε το δίκτυο να χαρτογραφεί σωστά τις εισόδους παραδειγμάτων στους σχετικούς στόχους τους. Αλλά εδώ είναι το πράγμα: ένα βαθύ νευρωνικό δίκτυο μπορεί να περιέχει δεκάδες εκατομμύρια παραμέτρους. Η εύρεση των σωστών τιμών για όλα αυτά μπορεί να φαίνεται σαν μια αποθαρρυντική εργασία, ειδικά δεδομένου ότι η τροποποίηση της τιμής μιας παραμέτρου θα επηρεάσει τη συμπεριφορά όλων των άλλων.



Εικόνα 3: Ένα νευρωνικό δίκτυο παραμετροποιείται από τα βάρη του. (Chollet, 2017)

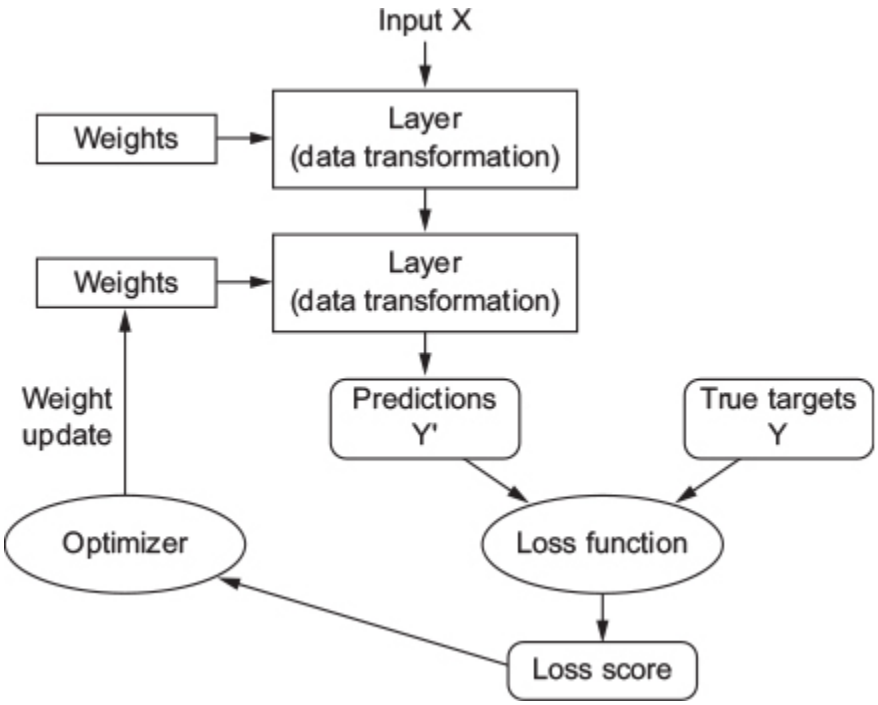
Για να ελεγχθεί κάτι, πρώτα πρέπει να είστε σε θέση να το παρατηρήσει κάποιος. Για να ελέγξει την έξοδο ενός νευρωνικού δικτύου, πρέπει να είναι σε θέση να μετρήσει πόσο μακριά είναι αυτή η έξοδος από αυτό που περιμένει. Αυτή είναι η δουλειά της λειτουργίας απώλειας του δικτύου, που ονομάζεται επίσης αντικειμενική λειτουργία. Η συνάρτηση απώλειας λαμβάνει τις προβλέψεις του δικτύου και του πραγματικού στόχου (τι ήθελε κάποιος να εξάγει το δίκτυο) και υπολογίζει μια βαθμολογία

απόστασης, καταγράφοντας πόσο καλά έχει κάνει το δίκτυο σε αυτό το συγκεκριμένο παράδειγμα (βλ. Εικόνα 4).



Εικόνα 4: Μια λειτουργία απώλειας μετρά την ποιότητα της εξόδου του δικτύου. (Chollet, 2017)

Το θεμελιώδες τέχνασμα στη βαθιά μάθηση είναι να χρησιμοποιήσει κάποιος αυτό το σκορ ως σήμα ανάδρασης για να προσαρμόσει λίγο την τιμή των βαρών, προς μια κατεύθυνση που θα μειώσει τη βαθμολογία απώλειας για το τρέχον παράδειγμα (βλ. Εικόνα 5). Αυτή η προσαρμογή είναι δουλειά του βελτιστοποιητή (optimizer), ο οποίος υλοποιεί αυτό που ονομάζεται αλγόριθμος Backpropagation: ο κεντρικός αλγόριθμος στη βαθιά μάθηση όπως αναφέρθηκε και σε προηγούμενη παράγραφο.



Εικόνα 5: Η βαθμολογία απώλειας χρησιμοποιείται ως σήμα ανάδρασης για την προσαρμογή των βαρών. (Chollet, 2017)

Αρχικά, τα βάρη του δικτύου εκχωρούνται τυχαίες τιμές, οπότε το δίκτυο απλώς υλοποιεί μια σειρά τυχαίων μετασχηματισμών. Φυσικά, η παραγωγή του απέχει πολύ από αυτό που θα έπρεπε ιδανικά να είναι και η βαθμολογία απώλειας είναι κατά συνέπεια πολύ υψηλή. Αλλά με κάθε παράδειγμα οι διαδικασίες δικτύου, τα βάρη προσαρμόζονται λίγο προς τη σωστή κατεύθυνση και η βαθμολογία απώλειας μειώνεται. Αυτη είναι η λούπα εκπαίδευσης, η οποία, επαναλαμβανόμενη αρκετές φορές (συνήθως δεκάδες επαναλήψεις σε χιλιάδες παραδείγματα), αποδίδει τιμές βάρους που ελαχιστοποιούν τη λειτουργία απώλειας. Ένα δίκτυο με ελάχιστη απώλεια είναι αυτό για το οποίο οι έξοδοι είναι όσο πιο κοντά μπορούν να είναι στους στόχους: ένα εκπαιδευμένο δίκτυο. Για άλλη μια φορά, είναι ένας απλός μηχανισμός που, μόλις κλιμακωθεί, καταλήγει να μοιάζει με μαγεία. (Chollet, 2017)

Εφαρμογές βαθιάς μάθησης για υπολογιστική όραση

Σύμφωνα με άρθρο της ESRI (χ.χ.) αναφορικά με τις εφαρμογές βαθιάς μάθησης που υπάρχουν και υποστηρίζει το λογισμικό της, χαρακτηριστικά αναφέρει ότι υπάρχουν πολλές εργασίες υπολογιστικής όρασης που μπορούν να επιτευχθούν με νευρωνικά δίκτυα βαθιάς μάθησης. Όλες αυτές οι εργασίες υπολογιστικής όρασης περιγράφονται παρακάτω, καθεμία με ένα παράδειγμα τηλεπισκόπησης και ένα γενικότερο παράδειγμα υπολογιστικής όρασης.

Σημασιολογική κατάτμηση (Semantic segmentation)

Η σημασιολογική κατάτμηση συμβαίνει όταν κάθε εικονοστοιχείο σε μια εικόνα ταξινομείται ως ανήκων σε μια κλάση. Για παράδειγμα, στην παρακάτω εικόνα, τα

εικονοστοιχεία δρόμου ταξινομούνται ξεχωριστά από τα εικονοστοιχεία εκτός δρόμου. Στα δεξιά, τα εικονοστοιχεία που αποτελούν μια γάτα σε μια φωτογραφία ταξινομούνται ως γάτα, ενώ τα άλλα εικονοστοιχεία στην εικόνα ανήκουν σε άλλες κατηγορίες. Στο GIS, αυτό αναφέρεται συχνά ως ταξινόμηση εικονοστοιχείων, κατάτμηση εικόνας ή ταξινόμηση εικόνας. Συχνά χρησιμοποιείται για τη δημιουργία χαρτών ταξινόμησης χρήσεων γης. (ESRI, χ.χ.)

Εικόνα 6: Παράδειγμα σημασιολογικής κατάτμησης. (ESRI, χ.χ.)

Κατάτμηση Στιγμιοτύπων (Instance segmentation)

Η κατάτμηση στιγμιοτύπων είναι μια πιο ακριβής μέθοδος ανίχνευσης αντικειμένων στην οποία σχεδιάζεται το όριο κάθε παρουσίας αντικειμένου. Για παράδειγμα, στην εικόνα στα αριστερά κάτω, ανιχνεύονται οι στέγες των σπιτιών, συμπεριλαμβανομένου του ακριβούς περιγράμματος του σχήματος στέγης. Στα δεξιά, εντοπίζονται αυτοκίνητα και μπορείτε να δείτε το ξεχωριστό σχήμα των αυτοκινήτων. Αυτός ο τύπος εφαρμογής βαθιάς μάθησης είναι επίσης γνωστός ως κατάτμηση αντικειμένων. (ESRI, χ.χ.)

Εικόνα 7: Παράδειγμα κατάτμησης στιγμιοτύπων. (ESRI, χ.χ.)

Πανοπτική Κατάτμηση (Panoptic segmentation)

Η πανοπτική κατάτμηση στιγμιότυπών συνδυάζει τόσο τη σημασιολογική κατάτμηση όσο και την κατάτμηση στιγμιότυπων. Για παράδειγμα, η παρακάτω εικόνα δείχνει ότι όλα τα εικονοστοιχεία είναι ταξινομημένα και κάθε μοναδικό αντικείμενο, όπως κάθε αυτοκίνητο, είναι το δικό του μοναδικό αντικείμενο. (ESRI, χ.χ.)

Εικόνα 8: Παράδειγμα πανοπτικής κατάτμησης. (ESRI, χ.χ.)

Μετάφραση Εικόνας (Image translation)

Η μετάφραση εικόνας είναι το έργο της μετάφρασης μιας εικόνας από μια πιθανή αναπαράσταση ή στυλ της σκηνής σε μια άλλη, όπως μείωση θορύβου ή υπερ-ανάλυση. Για παράδειγμα, η παρακάτω εικόνα στα αριστερά δείχνει την αρχική εικόνα χαμηλής ανάλυσης και η εικόνα στα δεξιά δείχνει το αποτέλεσμα της χρήσης ενός μοντέλου υπερ-ανάλυσης. Αυτός ο τύπος εφαρμογής βαθιάς μάθησης είναι επίσης γνωστός ως μετάφραση εικόνας σε εικόνα. (ESRI, χ.χ.)

Εικόνα 9: Παράδειγμα μετάφρασης εικόνας. (ESRI, χ.χ.)

Ανίχνευσης Αλλαγών (Change detection)

Οι εργασίες βαθιάς εκμάθησης ανίχνευσης αλλαγών μπορούν να ανιχνεύσουν αλλαγές στα χαρακτηριστικά που ενδιαφέρουν μεταξύ δύο ημερομηνιών και να δημιουργήσουν έναν λογικό χάρτη αλλαγών. Για παράδειγμα, η εικόνα στα αριστερά παρακάτω δείχνει μια οικιστική ανάπτυξη από πριν από πέντε χρόνια, η μεσαία εικόνα δείχνει την ίδια ανάπτυξη σήμερα και η εικόνα στα δεξιά δείχνει τον χάρτη λογικής αλλαγής όπου τα νέα σπίτια είναι λευκά. (ESRI, χ.χ.)

Εικόνα 10: Παράδειγμα ανίχνευσης αλλαγών. (ESRI, χ.χ.)

Ανίχνευση Αντικειμένων (Object detection)

Η ανίχνευση αντικειμένων είναι η διαδικασία εντοπισμού χαρακτηριστικών σε μια εικόνα. (Βλ. Ενότητα 1.2.1.) Για παράδειγμα, στην παρακάτω εικόνα τηλεπισκόπησης, το νευρωνικό δίκτυο βρήκε τη θέση ενός αεροπλάνου. Σε μια γενικότερη περίπτωση χρήσης υπολογιστικής όρασης, ένα μοντέλο μπορεί να είναι σε θέση να ανιχνεύσει τη θέση διαφορετικών ζώων. Αυτή η διαδικασία συνήθως περιλαμβάνει τη σχεδίαση ενός πλαισίου οριοθέτησης γύρω από τα χαρακτηριστικά που σας ενδιαφέρουν. Μπορεί να χρησιμοποιηθεί σε GIS για τον εντοπισμό συγκεκριμένων χαρακτηριστικών σε δορυφορικές, εναέριες ή drone εικόνες και για την αποτύπωση αυτών των χαρακτηριστικών σε έναν χάρτη. (ESRI, χ.χ.)

Εικόνα 11: Παράδειγμα ανίχνευσης αντικειμένων. (ESRI, χ.χ.)

Ταξινόμηση Εικόνας (Image classification)

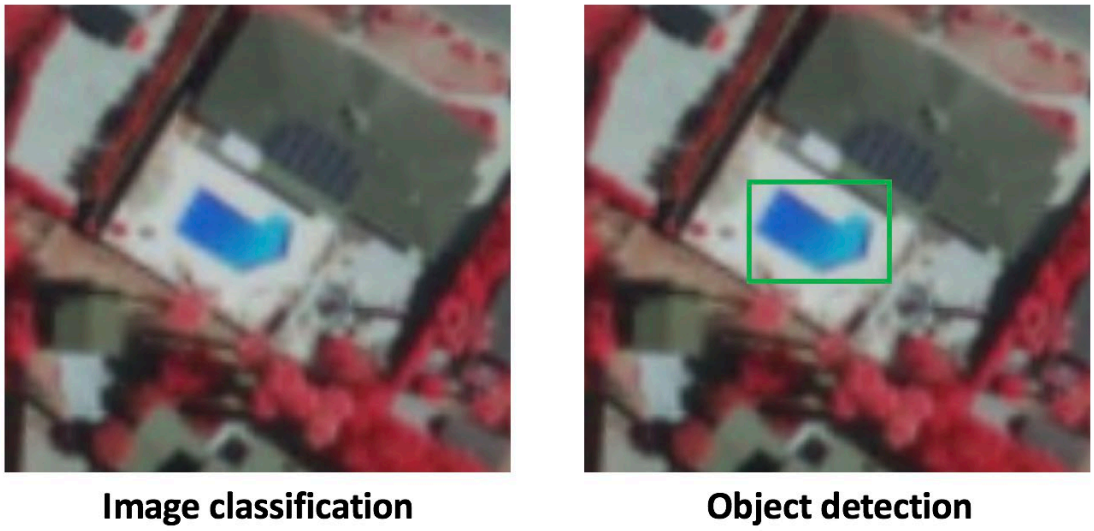
Η ταξινόμηση εικόνας περιλαμβάνει την εκχώρηση μιας ετικέτας ή κλάσης σε μια ψηφιακή εικόνα. Για παράδειγμα, η εικόνα drone στα αριστερά παρακάτω μπορεί να φέρει την ένδειξη πλήθος και η ψηφιακή φωτογραφία στα δεξιά μπορεί να φέρει την ετικέτα γάτα. Αυτός ο τύπος ταξινόμησης είναι επίσης γνωστός ως ταξινόμηση αντικειμένων ή αναγνώριση εικόνας και μπορεί να χρησιμοποιηθεί στο GIS για την κατηγοριοποίηση χαρακτηριστικών σε μια εικόνα. (ESRI, χ.χ.)

Εικόνα 12: Παράδειγμα ταξινόμησης εικόνας. (ESRI, χ.χ.)

1.2.Ανίχνευση αντικειμένων

Η ανίχνευση αντικειμένου είναι η διαδικασία προσδιορισμού του στιγμιότυπου της κλάσης στην οποία ανήκει το αντικείμενο και της εκτίμησης της θέσης του αντικειμένου εξάγοντας το πλαίσιο οριοθέτησης (ετικέτα) γύρω από το αντικείμενο. Η ανίχνευση ενός μεμονωμένου στιγμιότυπου της κλάσης από την εικόνα ονομάζεται ανίχνευση αντικειμένου μεμονωμένης κλάσης (single class object detection), ενώ η ανίχνευση των κλάσεων όλων των αντικειμένων που υπάρχουν στην εικόνα είναι γνωστή ως ανίχνευση αντικειμένου πολλαπλών κλάσεων (multi class object detection). Μερικές προκλήσεις που χρίζονται χειρισμού κατά την εκτέλεση ανίχνευσης ενός αντικειμένου είναι η μερική/πλήρης συγκάλυψη του αντικειμένου, διαφορετικές συνθήκες φωτισμού, στάσεις στην κάμερα, κλιμάκωση, κλπ. Γενικότερα, η ανίχνευση αντικειμένων είναι το πρώτο βήμα σε οποιαδήποτε δραστηριότητα οπτικής αναγνώρισης. (Pathak et al., 2018)

Η ταξινόμηση εικόνων στην υπολογιστική όραση παίρνει μια εικόνα και προβλέπει το αντικείμενο σε μια εικόνα, ενώ η ανίχνευση αντικειμένων όχι μόνο προβλέπει το αντικείμενο αλλά βρίσκει επίσης τη θέση τους όσον αφορά τα πλαίσια οριοθέτησης. Για παράδειγμα, όταν κατασκευάζουμε έναν ταξινομητή πισίνας, παίρνουμε μια εικόνα εισόδου και προβλέπουμε εάν περιέχει μια πισίνα, ενώ ένα μοντέλο ανίχνευσης αντικειμένων θα μας έλεγε επίσης τη θέση της πισίνας. (ArcGIS Developers, χ.χ.)



Εικόνα 13: Διαφορά μεταξύ ταξινόμησης και ανίχνευσης αντικειμένων. (ArcGIS Developers, χ.χ.)

Για επεξηγηματικούς σκοπούς, υποθέτοντας ότι υπάρχει το πολύ μία κλάση και ένα αντικείμενο σε μια εικόνα, η έξοδος ενός μοντέλου ανίχνευσης αντικειμένων πρέπει να περιλαμβάνει:

- Πιθανότητα ότι υπάρχει ένα αντικείμενο,
- Ύψος του κιβωτίου οριοθέτησης,
- Πλάτος του τετραγωνιδίου οριοθέτησης,
- Οριζόντια συντεταγμένη του κεντρικού σημείου του πλαισίου οριοθέτησης,
- Κατακόρυφη συντεταγμένη του κεντρικού σημείου του πλαισίου οριοθέτησης. (ArcGIS Developers, χ.χ.)

Αυτή είναι μόνο μία από τις συμβάσεις καθορισμού της εξόδου. Διαφορετικά μοντέλα και υλοποιήσεις μπορεί να έχουν διαφορετικές μορφές, αλλά η ιδέα είναι η ίδια, η οποία είναι να εξάγει την πιθανότητα και τη θέση του αντικειμένου. (ArcGIS Developers, χ.χ.)

Η ανίχνευση αντικειμένων μπορεί να πραγματοποιηθεί χρησιμοποιώντας είτε παραδοσιακές:

1. Οι τεχνικές επεξεργασίας εικόνας
2. Τα δίκτυα βαθιάς μάθησης.

Οι τεχνικές επεξεργασίας εικόνας γενικά δεν απαιτούν ιστορικά δεδομένα για την εκπαίδευση και έχουν χαρακτήρα χωρίς επίβλεψη. Το OpenCV είναι ένα δημοφιλές εργαλείο για εργασίες επεξεργασίας εικόνας. Ως εκ τούτου, αυτές οι εργασίες δεν απαιτούν σχολιασμένες εικόνες, όπου οι άνθρωποι επισήμαναν τα δεδομένα χειροκίνητα (για εποπτευόμενη εκπαίδευση). Αυτές οι τεχνικές περιορίζονται σε πολλούς παράγοντες, όπως πολύπλοκα σενάρια (χωρίς μονόχρωμο φόντο), απόφραξη (μερικώς κρυμμένα αντικείμενα), φωτισμό και σκιές και εφέ ακαταστασίας.

Οι μέθοδοι βαθιάς μάθησης εξαρτώνται γενικά από την εποπτευόμενη ή μη εποπτευόμενη μάθηση, με τις εποπτευόμενες μεθόδους να είναι το πρότυπο στις

εργασίες υπολογιστικής όρασης. Η απόδοση περιορίζεται από την υπολογιστική ισχύ των GPU, η οποία αυξάνεται ραγδαία χρόνο με το χρόνο.
Η ανίχνευση αντικειμένων βαθιάς μάθησης είναι σημαντικά πιο ισχυρή σε απόφραξη, πολύπλοκες σκηνές και προκλητικό φωτισμό. Απαιτείται τεράστιος όγκος δεδομένων προπόνησης. η διαδικασία σχολιασμού εικόνας είναι έντασης εργασίας και δαπανηρή. Για παράδειγμα, η επισήμανση 500.000 εικόνων για την εκπαίδευση ενός προσαρμοσμένου αλγορίθμου ανίχνευσης αντικειμένων βαθιάς μάθησης θεωρείται μικρό σύνολο δεδομένων. Ωστόσο, πολλά σύνολα δεδομένων αναφοράς (MS COCO, Caltech, KITTI, PASCAL VOC, V5) παρέχουν τη διαθεσιμότητα δεδομένων με ετικέτα. (Boesch, χ.χ.)

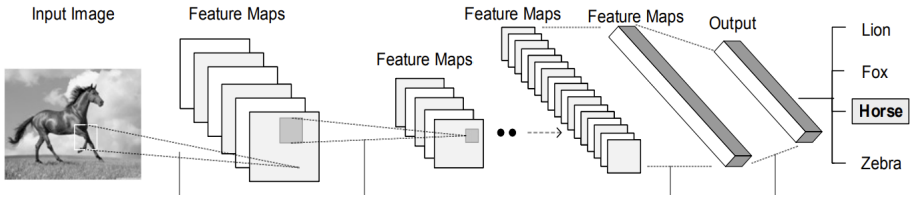
Οι κύριοι αλγόριθμοι ανίχνευσης αντικειμένων είναι δύο. Αλγόριθμοι όπως το R-CNN και το Fast(er) R-CNN χρησιμοποιούν μια προσέγγιση δύο βημάτων - πρώτα για να εντοπίσουν περιοχές όπου αναμένεται να βρεθούν αντικείμενα και στη συνέχεια να ανιχνεύσουν αντικείμενα μόνο σε αυτές τις περιοχές χρησιμοποιώντας convnet. Από την άλλη, αλγόριθμοι όπως το YOLO (You Only Look Once) (Redmon et al., 2015) και το SSD (Single-Shot Detector) (Liu et al., 2016) χρησιμοποιούν μια πλήρως συνελικτική προσέγγιση στην οποία το δίκτυο είναι σε θέση να βρει όλα τα αντικείμενα μέσα σε μια εικόνα σε ένα πέρασμα (εξ ου και «μία λήψη» ή «κοιτάξτε μία φορά») μέσω του convnet. Οι αλγόριθμοι πρότασης περιοχής έχουν συνήθως ελαφρώς καλύτερη ακρίβεια αλλά πιο αργή στην εκτέλεση, ενώ οι αλγόριθμοι μονής λήψης είναι πιο αποτελεσματικοί και έχουν εξίσου καλή ακρίβεια και σε αυτό θα επικεντρωθούμε σε αυτήν την ενότητα. (ArcGIS Developers, χ.χ.)

1.2.1. Ανίχνευση με βαθιά μάθηση

Τα τελευταία χρόνια, οι ραγδαίες εξελίξεις των τεχνικών βαθιάς μάθησης έχουν επιταχύνει σημαντικά την ορμή της ανίχνευσης αντικειμένων. Με τα δίκτυα βαθιάς μάθησης και την υπολογιστική ισχύ των GPU, η απόδοση των ανιχνευτών αντικειμένων και των ιχνηλατών έχει βελτιωθεί σημαντικά, επιτυγχάνοντας σημαντικές ανακαλύψεις στην ανίχνευση αντικειμένων. Η μηχανική μάθηση (ML) είναι ένας κλάδος της τεχνητής νοημοσύνης (AI) και ουσιαστικά περιλαμβάνει μοτίβα εκμάθησης από παραδείγματα ή δείγματα δεδομένων καθώς η μηχανή έχει πρόσβαση στα δεδομένα και έχει την ικανότητα να μαθαίνει από αυτά (εποπτευόμενη μάθηση σε σχολιασμένες εικόνες). Η βαθιά μάθηση είναι μια εξειδικευμένη μορφή μηχανικής μάθησης που περιλαμβάνει μάθηση σε διαφορετικά στάδια. (Boesch, χ.χ.)

Τα βαθιά CNNs έχουν χρησιμοποιηθεί εκτενώς για την ανίχνευση αντικειμένων. Το CNN είναι ένας τύπος τροφοδοτικού νευρωνικού δικτύου και λειτουργεί με βάση την αρχή της κατανομής βάρους. Η συνέλιξη είναι μια ολοκλήρωση που δείχνει πώς μια συνάρτηση επικαλύπτεται με μια άλλη συνάρτηση και είναι ένα μείγμα δύο συναρτήσεων που πολλαπλασιάζονται. Η Εικόνα 14 δείχνει πολυεπίπεδη αρχιτεκτονική του CNN για ανίχνευση αντικειμένων. Η εικόνα είναι περίπλοκη με τη λειτουργία ενεργοποίησης για τη λήψη χαρτών χαρακτηριστικών. Για να μειωθεί η χωρική πολυπλοκότητα του δικτύου, οι χάρτες χαρακτηριστικών αντιμετωπίζονται με

επίπεδα συγκέντρωσης για να ληφθούν αφηρημένοι χάρτες χαρακτηριστικών. Αυτή η διαδικασία επαναλαμβάνεται για τον επιθυμητό αριθμό φίλτρων και ανάλογα δημιουργούνται χάρτες χαρακτηριστικών. Τελικά, αυτοί οι χάρτες χαρακτηριστικών υποβάλλονται σε επεξεργασία με πλήρως συνδεδεμένα επίπεδα για να ληφθεί έξοδος αναγνώρισης εικόνας που δείχνει βαθμολογία εμπιστοσύνης για τις προβλεπόμενες ετικέτες κλάσης. Για τη βελτίωση της πολυπλοκότητας του δικτύου και τη μείωση του αριθμού των παραμέτρων, το CNN χρησιμοποιεί διαφορετικά είδη επιπέδων συγκέντρωσης, όπως φαίνεται στον Πίνακας 1: Συγκέντρωση επιπέδων που χρησιμοποιούνται για την ανίχνευση αντικειμένων. (Pathak et al., 2018). Τα επίπεδα συγκέντρωσης είναι αναλλοίωτα στη μετάφραση. Οι χάρτες ενεργοποίησης τροφοδοτούνται ως είσοδος στα επίπεδα συγκέντρωσης. Λειτουργούν σε κάθε έμπλαστρο στον επιλεγμένο χάρτη. (Pathak et al., 2018)



Εικόνα 14: Χρήση συνελικτικού νευρωνικού δικτύου για ανίχνευση αντικειμένων. (Pathak et al., 2018)

Pooling layer	Περιγραφή
Max pooling	Χρησιμοποιείται ευρέως στη συγκέντρωση σε CNN. Παίρνει τη μέγιστη τιμή από το επιλεγμένο patch εικόνας και τοποθετείται στη μήτρα αποθηκεύοντας τις μέγιστες τιμές από άλλα patches εικόνας.
Average pooling	Αυτή η ομαδοποίηση υπολογίζει κατά μέσο όρο τα γειτονικά εικονοστοιχεία.
Deformation pooling (Ouyang et al., 2015)	Η παραμορφώσιμη ομαδοποίηση έχει την ικανότητα να εξάγει παραμορφώσιμες ιδιότητες, γεωμετρικούς περιορισμούς των αντικειμένων..
Spatial pyramid pooling (He et al., 2015)	Αυτή η ομαδοποίηση εκτελεί δειγματοληψία προς τα κάτω της εικόνας και παράγει διάνυσμα χαρακτηριστικών με σταθερό μήκος. Αυτό το διάνυσμα χαρακτηριστικών μπορεί να χρησιμοποιηθεί για ανίχνευση αντικειμένων χωρίς να κάνει παραμορφώσεις στην αρχική εικόνα.

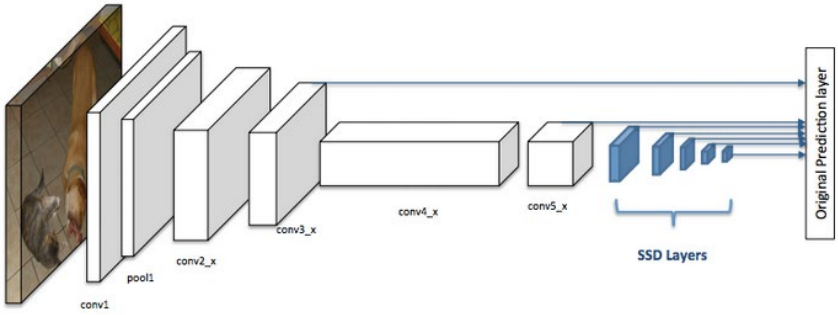
	Αυτή η συγκέντρωση είναι ανθεκτική στις παραμορφώσεις αντικειμένων.
Scale dependent pooling (Yang et al., 2015)	Αυτή η ομαδοποίηση χειρίζεται την παραλλαγή κλίμακας στην ανίχνευση αντικειμένων και βοηθά στη βελτίωση της ακρίβειας της ανίχνευσης.

Πίνακας 1: Συγκέντρωση επιπέδων που χρησιμοποιούνται για την ανίχνευση αντικειμένων. (Pathak et al., 2018)

1.2.2. Μοντέλο Single-Shot Detector

Το SSD είναι ένας δημοφιλής ανιχνευτής ενός σταδίου που μπορεί να προβλέψει πολλά. Η μέθοδος ανιχνεύει αντικείμενα σε εικόνες χρησιμοποιώντας ένα μόνο βαθύ νευρωνικό δίκτυο, διαχωρίζοντας το χώρο εξόδου των πλαισίων οριοθέτησης σε ένα σύνολο προεπιλεγμένων πλαισίων σε διαφορετικές αναλογίες διαστάσεων και κλίμακες ανά θέση χάρτη χαρακτηριστικών. Ο ανιχνευτής αντικειμένων δημιουργεί βαθμολογίες για την παρουσία κάθε κατηγορίας αντικειμένων σε κάθε προεπιλεγμένο πλαίσιο και προσαρμόζει το πλαίσιο ώστε να ταιριάζει καλύτερα στο σχήμα του αντικειμένου. Επίσης, το δίκτυο συνδυάζει προβλέψεις από πολλούς χάρτες χαρακτηριστικών με διαφορετικές αναλύσεις για το χειρισμό αντικειμένων διαφορετικών μεγεθών. Ο ανιχνευτής SSD είναι εύκολο να εκπαιδευτεί και να ενσωματωθεί σε συστήματα λογισμικού που απαιτούν ένα στοιχείο ανίχνευσης αντικειμένων. Σε σύγκριση με άλλες μεθόδους ενός σταδίου, ο SSD έχει πολύ καλύτερη ακρίβεια, ακόμη και με μικρότερα μεγέθη εικόνας εισόδου. (Liu et al., 2016) Το μοντέλο Single-Shot Detector (SSD) έχει δύο εξαρτήματα: ένα μοντέλο blackbone και μια κεφαλή SSD.

Το Backbone SSD χρησιμοποιεί ένα προ-εκπαιδευμένο δίκτυο ταξινόμησης εικόνων ως εξαγωγέα χαρακτηριστικών. Αυτό είναι συνήθως ένα δίκτυο όπως το ResNet που έχει εκπαιδευτεί στο ImageNet, από το οποίο έχουν αφαιρεθεί τα τελικά πλήρως συνδεδεμένα επίπεδα για να καταλήξουν στην προβλεπόμενη κλάση μιας εικόνας εισόδου. Έτσι, μένει ένα βαθύ νευρωνικό δίκτυο που είναι σε θέση να εξάγει σημασιολογικό νόημα από την εικόνα εισόδου διατηρώντας παράλληλα τη χωρική δομή της εικόνας αν και σε χαμηλότερη ανάλυση. Για το ResNet34, το blackbone έχει ως αποτέλεσμα 256 7x7 «χάρτες χαρακτηριστικών» ενεργοποιήσεων για κάθε εικόνα εισόδου. Κάθε ένας από αυτούς τους χάρτες 256 χαρακτηριστικών μπορεί να ερμηνευτεί ως ένα πλέγμα ενεργοποιήσεων 7x7 που ενεργοποιούνται όταν ανιχνεύεται ένα συγκεκριμένο χαρακτηριστικό στην εικόνα. Στην αρχιτεκτονική SSD, ένα ή περισσότερα συνελικτικά επίπεδα προστίθενται σε αυτό το blackbone και οι έξοδοι ερμηνεύονται ως τα πλαίσια οριοθέτησης και οι αντικειμένων στη χωρική θέση των ενεργοποιήσεων του τελικού επιπέδου. (ArcGIS Developers, χ.χ.)



Εικόνα 15: Architecture of a convolutional neural network with a SSD detector. (Liu et al., 2016)

Στην παραπάνω εικονα, τα πρώτα επίπεδα με λευκό χρώμα είναι τα blackbones ενώ εκείνα με το μπλε αποτελούν την κεφαλή SSD.

1.3. Μεταφορά γνώσης

Η μεταφορά γνώσης (transfer learning) είναι μια μέθοδος μηχανικής μάθησης όπου ένα μοντέλο που έχει αναπτυχθεί για μια εργασία επαναχρησιμοποιείται ως το σημείο εκκίνησης για ένα μοντέλο σε μια δεύτερη εργασία. (Brownlee, 2017)
Είναι μια δημοφιλής προσέγγιση στη βαθιά μάθηση όπου τα προεκπαιδευμένα μοντέλα χρησιμοποιούνται ως αφετηρία για εργασίες επεξεργασίας όρασης υπολογιστή και φυσικής γλώσσας, δεδομένου του τεράστιου υπολογισμού και των πόρων που απαιτούνται για την ανάπτυξη μοντέλων νευρωνικών δικτύων σε αυτά τα προβλήματα και από τα τεράστια άλματα δεξιοτήτων που παρέχουν σε σχετικά προβλήματα. (Brownlee, 2017)

Η εκμάθηση μεταφοράς αναφέρεται στην κατάσταση όπου αυτό που έχει μάθει σε μια ρύθμιση (π.χ. διανομή P1) αξιοποιείται για τη βελτίωση της γενίκευσης σε μια άλλη ρύθμιση (ας πούμε, διανομή P2). (Goodfellow et al., 2016)

Στη μεταφορά μάθησης, ο εκπαιδευόμενος πρέπει να εκτελέσει δύο ή περισσότερες διαφορετικές εργασίες, αλλά υποθέτουμε ότι πολλοί από τους παράγοντες που εξηγούν τις παραλλαγές στο P1 σχετίζονται με τις παραλλαγές που πρέπει να συλληφθούν για την εκμάθηση του P2. Αυτό γίνεται συνήθως κατανοητό σε ένα εποπτευόμενο μαθησιακό πλαίσιο, όπου η εισροή είναι η ίδια, αλλά ο στόχος μπορεί να είναι διαφορετικής φύσης. Για παράδειγμα, μπορεί να μάθουμε για ένα σύνολο οπτικών κατηγοριών, όπως γάτες και σκύλοι, στην πρώτη ρύθμιση και, στη συνέχεια, να μάθουμε για ένα διαφορετικό σύνολο οπτικών κατηγοριών, όπως μυρμήγκια και σφήκες, στη δεύτερη ρύθμιση. Εάν υπάρχουν σημαντικά περισσότερα δεδομένα στην πρώτη ρύθμιση (δείγμα από το P1), τότε αυτό μπορεί να βοηθήσει στην εκμάθηση αναπαραστάσεων που είναι χρήσιμες για γρήγορη γενίκευση από πολύ λίγα μόνο παραδείγματα που προέρχονται από το P2. Πολλές οπτικές κατηγορίες μοιράζονται έννοιες χαμηλού επιπέδου για τις άκρες και τα οπτικά σχήματα, τις επιπτώσεις των γεωμετρικών αλλαγών, τις αλλαγές στον φωτισμό και ούτω καθεξής. Σε γενικές γραμμές, η μάθηση μεταφοράς και η μάθηση πολλαπλών εργασιών μπορούν να επιτευχθούν μέσω της μάθησης αναπαράστασης όταν υπάρχουν χαρακτηριστικά που είναι χρήσιμα για τις διαφορετικές ρυθμίσεις ή εργασίες, που αντιστοιχούν σε

υποκείμενους παράγοντες που εμφανίζονται σε περισσότερες από μία ρυθμίσεις. (Goodfellow et al., 2016)

1.3.1. Τεχνική εκμάθηση μεταφοράς ResNet-34

Το ResNet είναι ένα σύντομο όνομα για το Residual Network, είναι ένα τεχνητό νευρωνικό δίκτυο που βοηθά στην οικοδόμηση βαθύτερου νευρωνικού δικτύου τη δυνατότητα παράβλεψης συνδέσεων ή συντομεύσεις για να μεταφερθεί πάνω από ορισμένα επίπεδα. Η παράλειψη βοηθά στη δημιουργία βαθύτερων επιπέδων δικτύου χωρίς να πέσει στο πρόβλημα της εξαφάνισης των διαβαθμίσεων. Υπάρχουν διαφορετικές εκδόσεις του ResNet, συμπεριλαμβανομένων των ResNet-18, ResNet-34, ResNet-50 και ούτω καθεξής. Οι αριθμοί υποδηλώνουν επίπεδα, αν και η αρχιτεκτονική είναι η ίδια. (Singhal, 2020)

Όπως υποδηλώνει το όνομα του δικτύου, η νέα ορολογία που εισάγει είναι η «υπολειμματική» μάθηση (residual learning). Τα βαθιά συνελικτικά δίκτυα έχουν οδηγήσει σε μία σειρά ανακαλύψεων για την ταξινόμηση των εικόνων. Έτσι, με την πάροδο του χρόνου υπάρχει μία τάση να αυξάνεται το βάθος των δικτύων έτσι ώστε να επιλύουν πιο σύνθετα προβλήματα. Όμως, όσο αυξάνεται το βάθος του δικτύου η εκπαίδευση καθίσταται πιο δύσκολη. Η «υπολειμματική» μάθηση προσπαθεί να λύσει αυτό το πρόβλημα. Σε γενικές γραμμές, σε ένα βαθύ συνελικτικό νευρωνικό δίκτυο, πολλά επίπεδα (layers) στοιβάζονται και εκπαιδεύονται. Το δίκτυο μαθαίνει χαμηλού, μεσαίου και υψηλού επιπέδου χαρακτηριστικά στο τέλος των επιπέδων του. (He et al., 2015. Θεοδωρόπουλος, 2018)

Το μοντέλο ResNet-34 είναι ένα υπολειπόμενο δίκτυο εκπαιδευμένο στο σύνολο δεδομένων ImageNet που περιέχει περισσότερες από 1 εκατομμύριο εικόνες και έχει βάθος 34 επιπέδων. (ESRI, χ.χ.)

1.4. Ακρίβεια ανίχνευσης αντικειμένων

Η ακρίβεια είναι το ποσοστό όλων των παρατηρήσεων που ταξινομούνται σωστά από το μοντέλο βαθιάς μάθησης.

Προκειμένου να εξασφαλιστεί η αποτελεσματικότητα ανίχνευσης αντικειμένου, εφαρμόζονται διάφοροι μέθοδοι.

Στην ανίχνευση και ταξινόμηση αντικειμένων, ένα μοντέλο μπορεί να προβλέψει μια θετική ή μια αρνητική κλάση και οι προβλέψεις μπορεί να είναι αληθείς ή ψευδείς. Για παράδειγμα, κατά την ανίχνευση της παρουσίας δέντρων σε μια εικόνα, η θετική κλάση μπορεί να είναι "Δέντρο", ενώ η αρνητική κλάση θα είναι "Όχι δέντρο". Μια αληθινή πρόβλεψη εμφανίζεται όταν η πρόβλεψη είναι σωστή και μια λανθασμένη πρόβλεψη εμφανίζεται όταν η πρόβλεψη είναι λανθασμένη. (ESRI, χ.χ.)

Στην παρακάτω εικόνα, τα κόκκινα πλαίσια οριοθέτησης υποδεικνύουν μια θετική πρόβλεψη, όπου το μοντέλο προέβλεψε ότι υπάρχει ένα δέντρο. Τα σκούρα μπλε πλαίσια οριοθέτησης υποδεικνύουν μια αρνητική πρόβλεψη, όπου το μοντέλο προέβλεψε ότι δεν υπάρχει δέντρο. (ESRI, χ.χ.)

Εικόνα 16: Παράδειγμα πρόβλεψης ανίχνευσης αντικειμένου. (ESRI, χ.χ.)

Η ερμηνεία κάθε πλαισίου οριοθέτησης της παραπάνω εικόνας επεξηγείται στον παρακάτω πίνακα.

Αριθμός	Περιγραφή
1	Αληθώς θετικό (True Positive) — Το μοντέλο προέβλεψε ότι υπάρχει ένα δέντρο και είναι σωστό.
2	Ψευδώς θετικό (False Positive) — Το μοντέλο προέβλεψε ότι υπάρχει ένα δέντρο και είναι λάθος.
3	Ψευδώς αρνητικό (False Negative) —Το μοντέλο προέβλεψε ότι δεν υπάρχει δέντρο και είναι λάθος.
4	Αληθές αρνητικό (True Negative) —Το μοντέλο προέβλεψε ότι δεν υπάρχει δέντρο και είναι σωστό.

Πίνακας 2: Μετρήσεις ακρίβειας πρόβλεψης της ανίχνευσης αντικειμένου. (ESRI, χ.χ.)

Οι μετρήσεις ακρίβειας περιγράφονται παρακάτω (ESRI, χ.χ.):

- **Ακρίβεια (Precisio**n)

Η ακρίβεια είναι ο λόγος του αριθμού των αληθών θετικών προς τον συνολικό αριθμό των θετικών προβλέψεων. Για παράδειγμα, εάν το μοντέλο ανίχνευσε 100 δέντρα και τα 90 ήταν σωστά, η ακρίβεια είναι 90 τοις εκατό.

Precision = (True Positive)/(True Positive + False Positive)

- **Ανάκληση (Recall)**

Η ανάκληση είναι ο λόγος του αριθμού των αληθών θετικών προς τον συνολικό αριθμό των πραγματικών (σχετικών) αντικειμένων. Για παράδειγμα, εάν το μοντέλο ανιχνεύσει σωστά 75 δέντρα σε μια εικόνα και στην πραγματικότητα υπάρχουν 100 δέντρα στην εικόνα, η ανάκληση είναι 75 τοις εκατό.

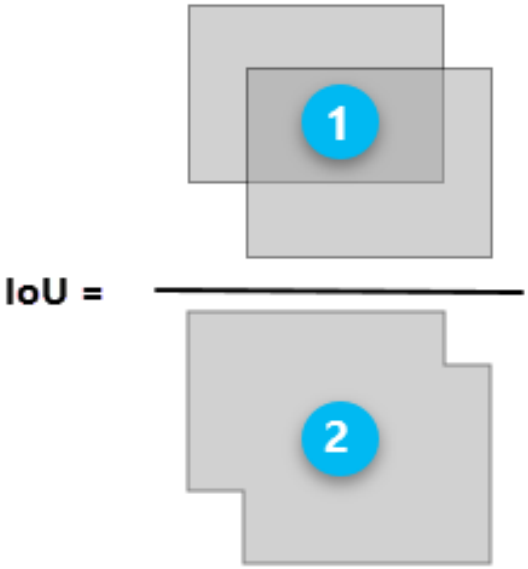
Recall = (True Positive)/(True Positive + False Negative)

- **Βαθμολογία F1 (F1 score)**

Η βαθμολογία F1 είναι ένας σταθμισμένος μέσος όρος της ακρίβειας και της ανάκλησης. Οι τιμές κυμαίνονται από 0 έως 1, όπου το 1 σημαίνει μεγαλύτερη ακρίβεια.

F1 score = (Precision × Recall)/[(Precision + Recall)/2]

Η ακρίβεια ενός μοντέλου ανίχνευσης αντικειμένου εξαρτάται από την ποιότητα και τον αριθμό των δειγμάτων εκπαίδευσης, τις εικόνες εισόδου, τις παραμέτρους του μοντέλου και το όριο απαίτησης για ακρίβεια. (ESRI, χ.χ.)
Ο λόγος Intersection over Union (IoU) (Βλ. Εικόνα 17) χρησιμοποιείται ως όριο για τον προσδιορισμό του εάν ένα προβλεπόμενο αποτέλεσμα είναι αληθώς θετικό ή ψευδώς θετικό. Ο λόγος IoU είναι η ποσότητα της επικάλυψης μεταξύ του πλαισίου οριοθέτησης γύρω από ένα προβλεπόμενο αντικείμενο και του πλαισίου οριοθέτησης γύρω από τα δεδομένα αναφοράς του εδάφους. (ESRI, χ.χ.)



Εικόνα 17: Ο λόγος Intersection over Union (IoU). (ESRI, χ.χ.)

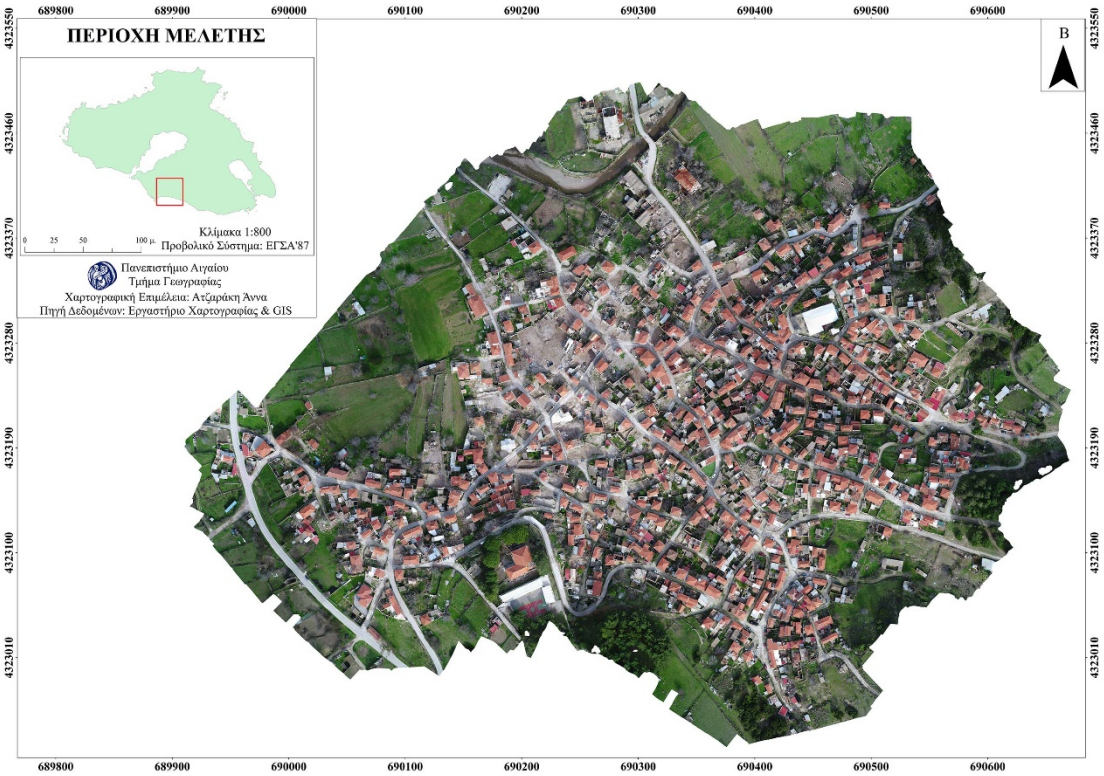
1	Η περιοχή τομής του προβλεπόμενου πλαισίου οριοθέτησης και του πλαισίου οριοθέτησης αναφοράς εδάφους
2	Η συνολική επιφάνεια του προβλεπόμενου πλαισίου οριοθέτησης και του πλαισίου οριοθέτησης αναφοράς εδάφους σε συνδυασμό.

Πίνακας 3: Χαρακτηριστικά ακρίβειας πρόβλεψης της ανίχνευσης αντικειμένου. (ESRI, χ.χ.)

2. ΠΕΡΙΟΧΗ ΜΕΛΕΤΗΣ

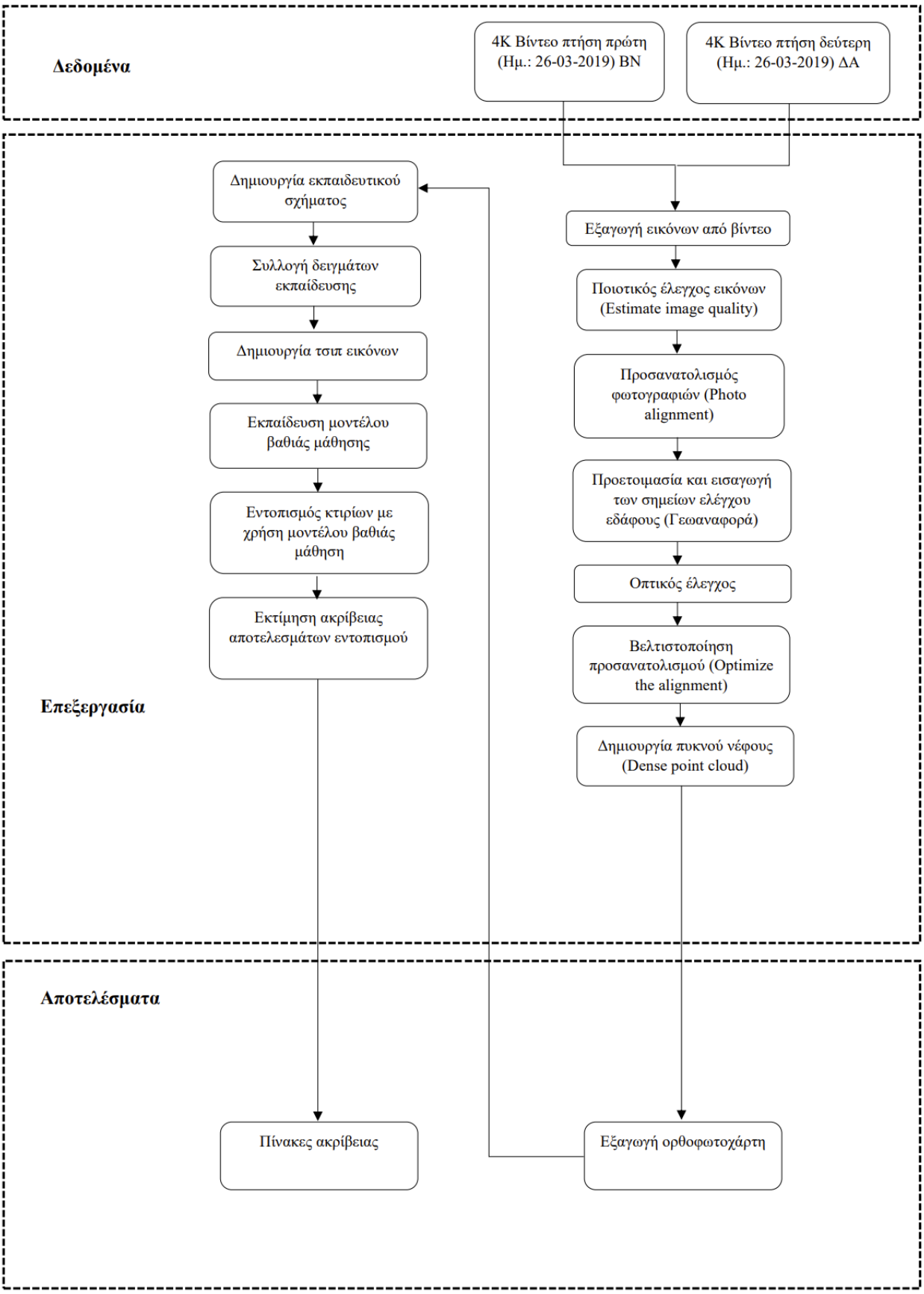
2.1. Γεωγραφικά χαρακτηριστικά Βρίσας

Η περιοχή της Βρίσας, η οποία εξετάζεται στη παρούσα εργασία, αποτελεί χωριό της νήσου Λέσβου. Διοικητικά ανήκει στον δήμο Λέσβου της Περιφερειακής Ενότητας Λέσβου ενώ αποτελεί έδρα του ομώνυμου δημοτικού διαμερίσματος του δήμου Πολιχνίτου. Βρίσκεται 50,8 χλμ. νοτιοανατολικά από τη πρωτεύουσα της Λέσβου τη Μυτιλήνη και έχει 617 κατοίκους. Έχει γεωγραφικό μήκος Ε 26° 12' 3"και γεωγραφικό πλάτος Ν 39° 2' 23", ενώ το υψόμετρο φτάνει τα 43 μέτρα από την επιφάνεια της θάλασσας. (Ατζαράκη, 2019)



Χάρτης 1: Περιοχή μελέτης Βρίσα.

3. ΔΙΑΓΡΑΜΜΑ ΡΟΗΣ



4. ΔΕΔΟΜΕΝΑ

4.1. Λήψη δεδομένων

Για την διεκπεραίωση της εργασίας συλλέχθηκαν βίντεο ποιότητας 4K που ήταν αποτέλεσμα πτήσεων drone που πραγματοποιήθηκαν στις 26/03/2019.

4.2. Βίντεο

Η πρώτη πτήση είχε τη κατεύθυνση Βορρά-Νότου ενώ η δεύτερη πτήση είχε κατεύθυνση Δύση-Ανατολή. Δηλαδή μια πτήση οριζόντιας πορείας του drone και μια κάθετης. Ακολουθεί ο πίνακας μεταδεδομένων για το κάθε βίντεο ξεχωριστά:

4.2.1. Α' Βίντεο

Τίτλος	Βίντεο Α' πτήσης
Ημερομηνία δημιουργίας	26/03/2019
Ημερομηνία τροποποίησης	27/03/2022
Περιγραφή	Βίντεο της πρώτης πτήσης drone που πραγματοποιήθηκε στις 26/03/2019 στη περιοχή της Βρίσας.
Τύπος αρχείου	MP4
Ποιότητα αρχείου	4K

Πίνακας 4: Μεταδεδομένα Α' πτήσης βίντεο.

4.2.2. Β' Βίντεο

Τίτλος	Βίντεο Β' πτήσης
Ημερομηνία δημιουργίας	26/03/2019
Ημερομηνία τροποποίησης	27/03/2022
Περιγραφή	Βίντεο της δεύτερης πτήσης drone που πραγματοποιήθηκε στις 26/03/2019 στη περιοχή της Βρίσας.
Τύπος αρχείου	MP4
Ποιότητα αρχείου	4K

Πίνακας 5: Μεταδεδομένα Β' πτήσης βίντεο.

5. ΜΕΘΟΔΟΛΟΓΙΑ

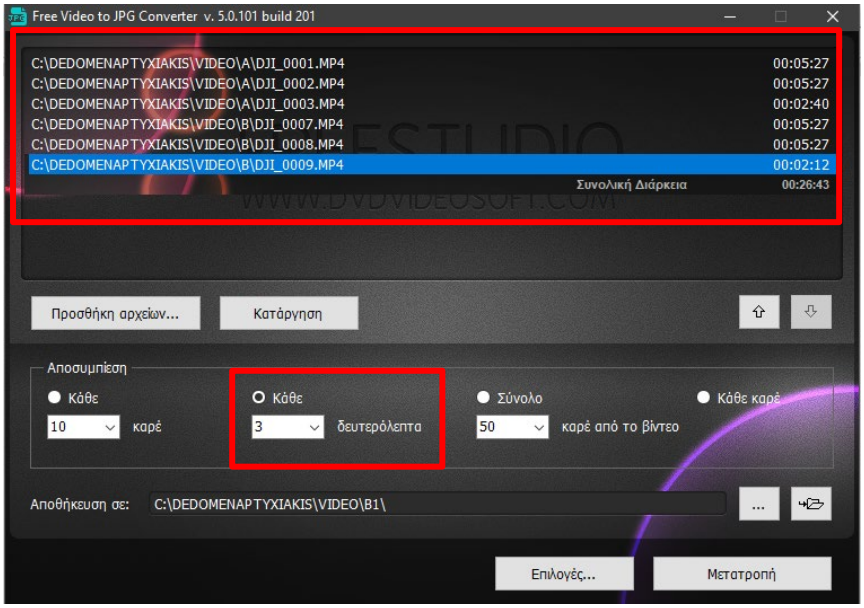
Όπως φαίνεται και στην Ενότητα 3, για τη διεκπεραίωση της εργασίας πραγματοποιήθηκαν πέντε εφαρμογές.

- Η πρώτη εφαρμογή φαίνεται στις Ενότητες 5.1 έως 5.1.7 και περιλάμβανε την προετοιμασία του συνόλου των δεδομένων που συλλέχθηκαν στις 26/03/2019, με σκοπό τη δημιουργία ορθοφωτογραφίας.
- Στη δεύτερη εφαρμογή, η οποία περιλαμβάνεται στην Ενότητα 5.2 προβάλλεται η διαμόρφωση του λειτουργικού συστήματος.
- Η τρίτη εφαρμογή φαίνεται στις Ενότητες 5.3 έως 5.3.3 και περιλάμβανε τη δημιουργία δειγμάτων εκπαίδευσης.
- Στη τέταρτη εφαρμογή, η οποία φαίνεται στις Ενότητες 5.4 έως 5.4.2, πραγματοποιήθηκε ανίχνευση κτιρίων με χρήση μοντέλου βαθιάς μάθησης.
- Η πέμπτη και τελευταία εφαρμογή στην Ενότητα 5.5 περιλαμβάνει τον υπολογισμό ακρίβειας για την ανίχνευση κτιρίου.

5.1.Προετοιμασία συνόλου δεδομένων

5.1.1. Εξαγωγή εικόνων

Στο λογισμικό Free Video to JPG Converter, προστέθηκαν τα βίντεο κάθετης και οριζόντιας κατεύθυνσης με σκοπό να μετατραπούν σε εικόνες-στιγμιότυπα και να ξεκινήσει η διαδικασία δημιουργίας του 3D μοντέλου για τις 26/03/2019 (βλ. Εικόνα 18). Έτσι, αφού εισήχθησαν τα βίντεο, απομονώθηκαν καρέ ανά 3 δευτερόλεπτα για τη βέλτιστη εξαγωγή των εικόνων (JPG). Τα βίντεο ήταν 3 από κάθε κατεύθυνση δηλαδή συνολικά 6 βίντεο ενώ οι εικόνες που εξήχθησαν ήταν περίπου 300 από κάθε πτήση (περίπου 600 εικόνες). (Ατζαράκη, 2019)



Εικόνα 18: Μετατροπή βίντεο σε εικόνες.

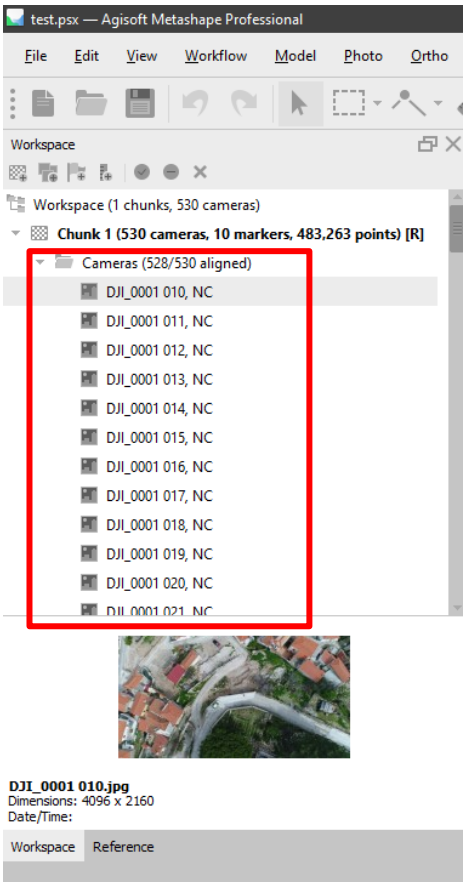
5.1.1. Οπτικός έλεγχος

Πρώτο στάδιο αφού εξήχθησαν οι εικόνες ήταν ο οπτικός έλεγχος. Η διαδικασία αυτή έγινε με σκοπό να επιλεχθούν σε πρώτη φάση οι κατάλληλες εικόνες και να απορριφθούν οι θολές και κακής ποιότητας εικόνες. Έτσι, χειροκίνητα, αφαιρέθηκαν όσες εικόνες δεν βοηθούσαν στη δημιουργία των μοντέλων και από 600 περίπου εικόνες έμειναν οι 530. (Ατζαράκη, 2019)

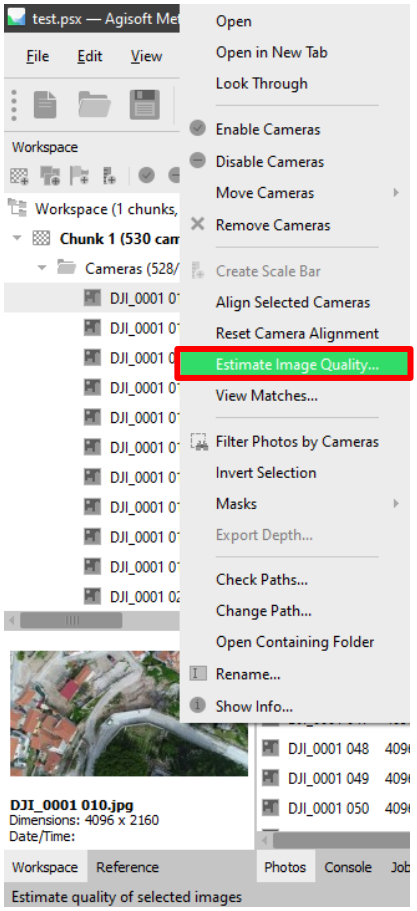
5.1.2. Ποιοτικός έλεγχος εικόνων

Έπειτα, με τη χρήση του λογισμικού Agisoft Metashape Professional πραγματοποιήθηκε ποιοτικός έλεγχος των εικόνων. Έτσι, αφού εισήχθησαν οι εικόνες στο λογισμικό (βλ.

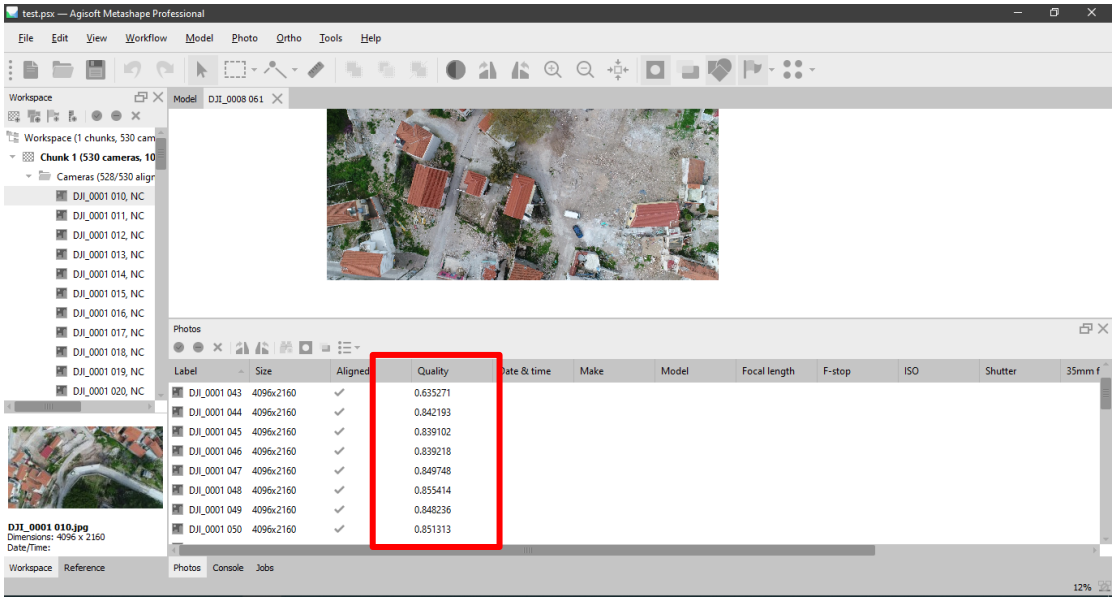
Εικόνα 19), χρησιμοποιήθηκε η διαδικασία του ποιοτικού ελέγχου (Estimate Image Quality) (βλ. Εικόνα 20), το οποίο εξίσου βοηθάει στην επιλογή των κατάλληλων φωτογραφιών καθώς υπολογίζει τη ποιότητα της κάθε εικόνας από το 0 έως το 1 (βλ. Εικόνα 21). Προτείνεται οι εικόνες με ποιότητα κάτω από 0.5, να απορρίπτονται καθώς πρόκειται για εικόνες θολές ή με μη κατάλληλες πληροφορίες εικόνες. Επιπλέον, με τη χρήση όλων των φωτογραφιών υπάρχει κίνδυνος λανθασμένων υπολογισμών και αποτελεσμάτων, ενώ ο αποκλεισμός κάποιων μπορεί να δημιουργήσει έλλειψη πληροφορίας. Με σωστή επικάλυψη, κατά τη διάρκεια λήψης των εικόνων, υπάρχουν αρκετές εικόνες, ώστε να καλύπτεται όλη η περιοχή μελέτης. Έτσι όσες εικόνες είχαν ποιότητα ίση ή κάτω από 0.5 αποκλείστηκαν (βλ. Εικόνα 22). (Ατζαράκη, 2019)



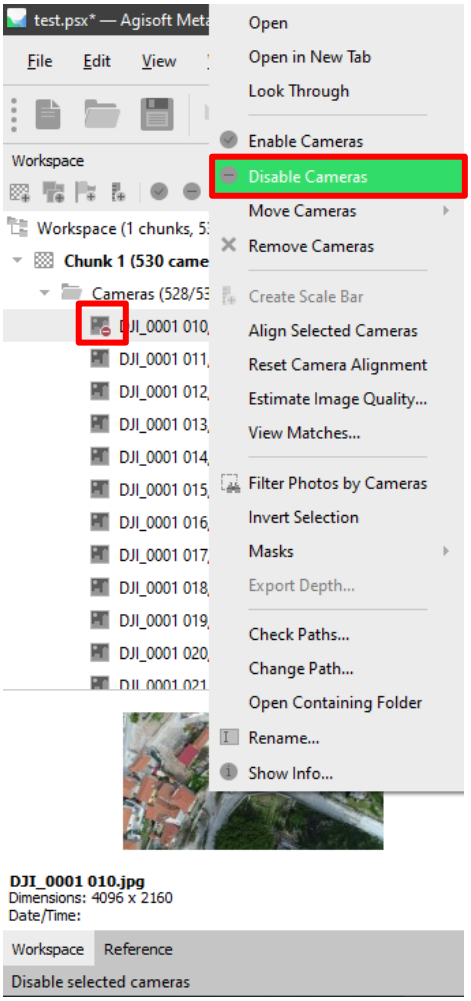
Εικόνα 19: Εισαγωγή των εικόνων στο Agistoft Metashape Professional.



Εικόνα 20: Χρήση "Estimate Image Quality".



Εικόνα 21: Υπολογισμός ποιότητας εικόνας.

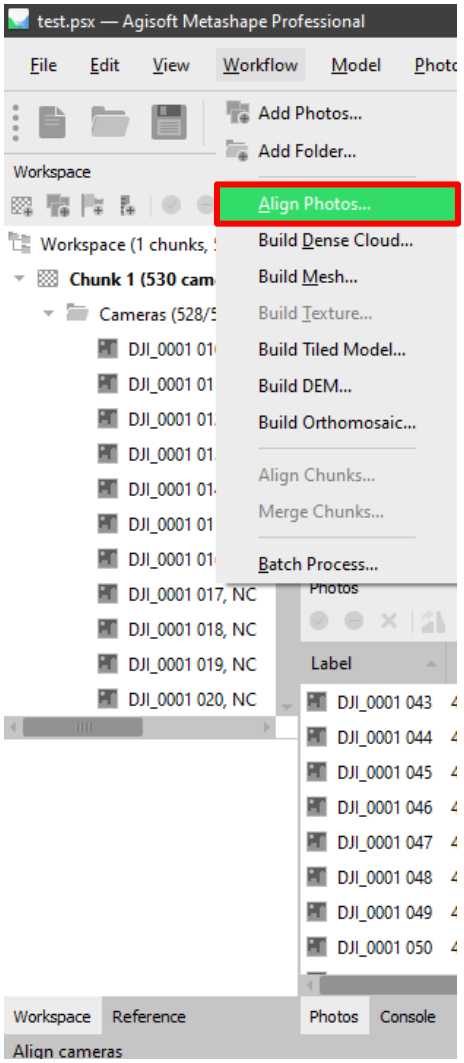


Εικόνα 22: Αποκλεισμός εικόνων με ακατάλληλη ποιότητα.

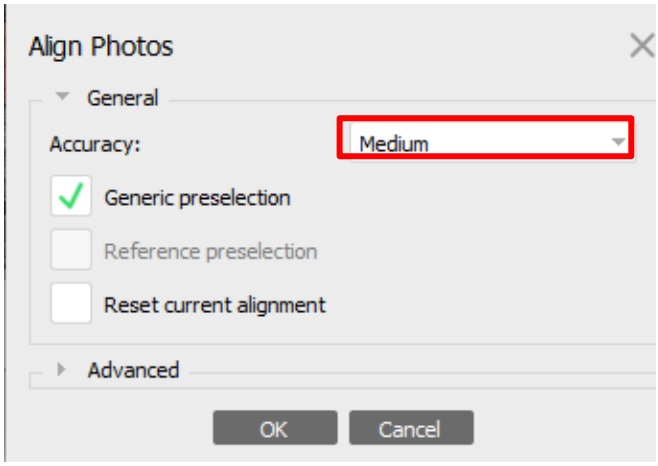
5.1.3. Προσανατολισμός φωτογραφιών

Σε αυτό το στάδιο, γίνεται ένας σχετικός προσανατολισμός όπου προσδιορίζονται τα σημεία σύνδεσης. Υπολογίζονται δηλαδή, τα κοινά σημεία μεταξύ των φωτογραφιών, ευθυγραμμίζει τις φωτογραφίες και δημιουργείται ένα αραιό νέφος σημείων (Spare Cloud). Έτσι, αφού επιλέχθηκε ο προσανατολισμός των φωτογραφιών (Align Photos) (βλ. Εικόνα 23), ορίστηκε η παράμετρο απαιτούμενης ακρίβειας του νέφους σημείων σε μέτρια (Medium) καθώς προτείνεται για τα μεγάλα σύνολα φωτογραφιών (βλ. Εικόνα 24).

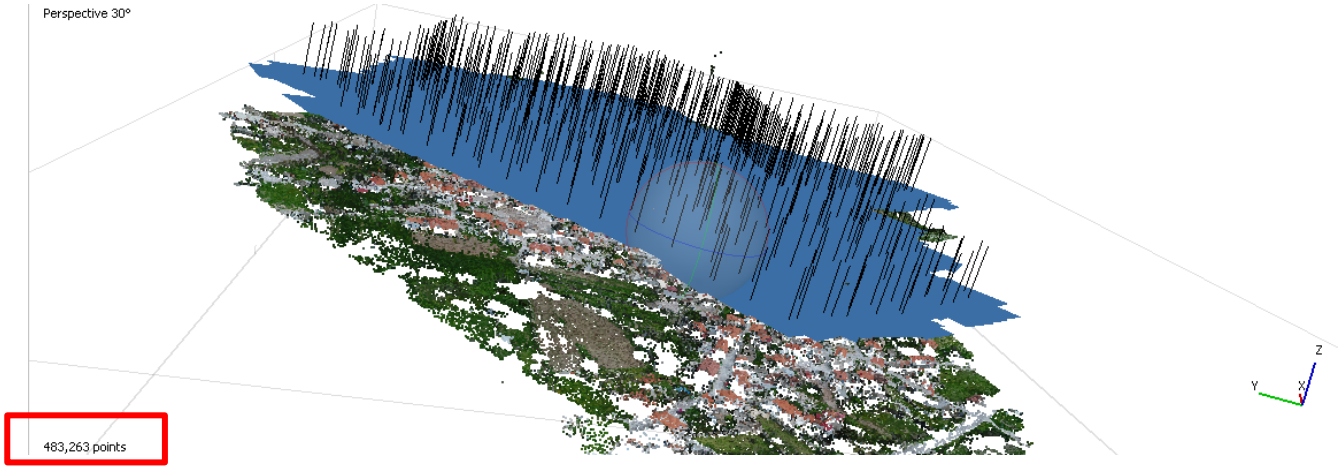
Όσο πιο μεγάλη είναι η απαιτούμενη ακρίβεια, τόσο πιο χρονοβόρα θα είναι η διαδικασία (Fayad, 2019). Το στάδιο αυτό είχε ως αποτέλεσμα ένα νέφος 483,263 σημείων (βλ. Εικόνα 25). (Ατζαράκη, 2019)



Εικόνα 23: Χρήση " Align Photos".



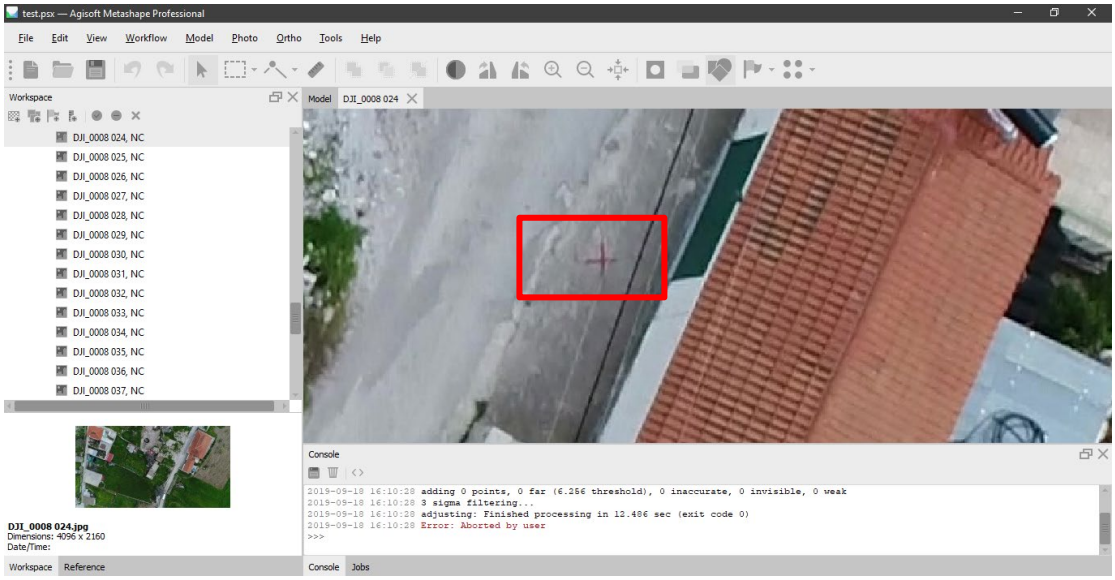
Εικόνα 24: Ορισμός της παραμέτρου ακρίβειας.



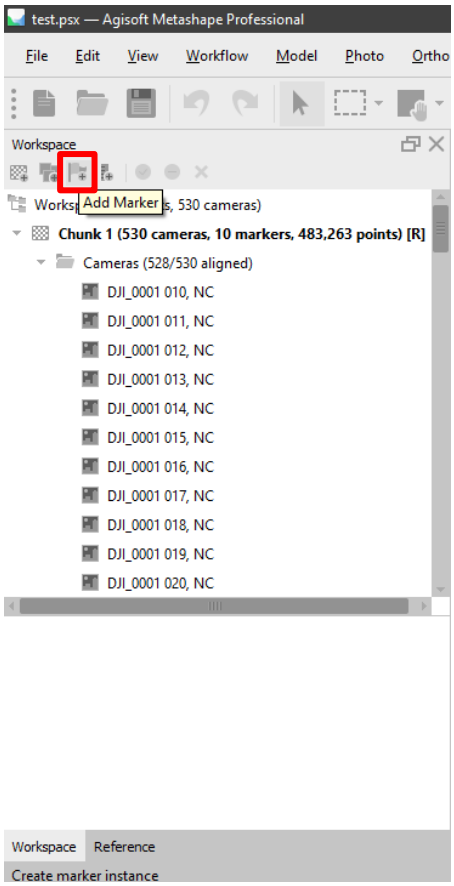
Εικόνα 25: Αποτέλεσμα προσανατολισμού φωτογραφιών.

5.1.4. Προετοιμασία & Εισαγωγή των σημείων ελέγχου εδάφους (Γεωαναφορά)

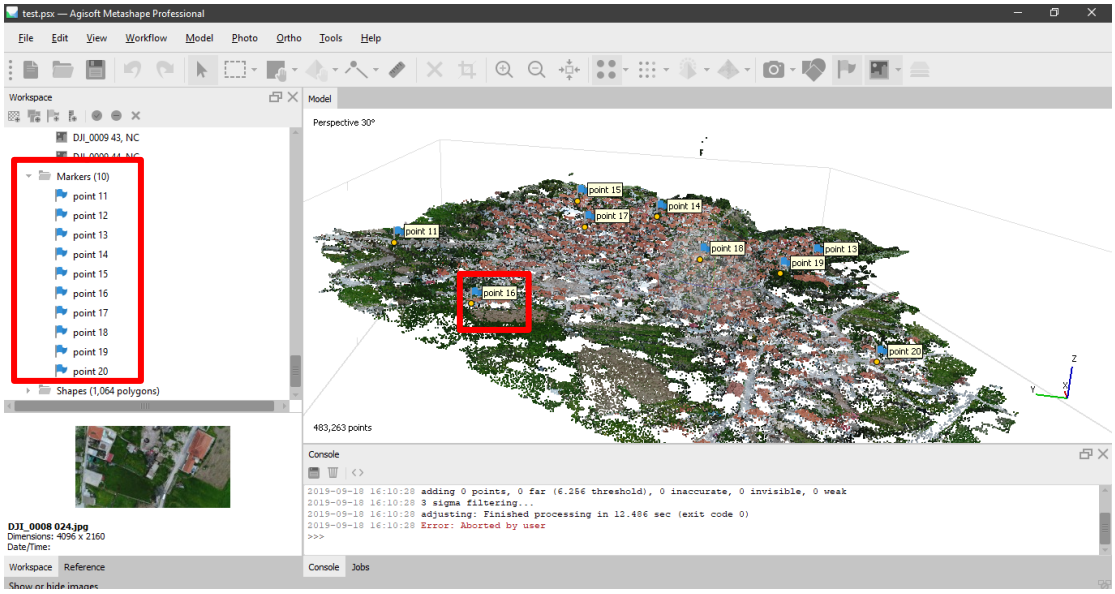
Η εισαγωγή των σημείων ελέγχου εδάφους (GSPs), είναι ένα στάδιο που απαιτείται για τη γεωαναφορά του ψηφιακού μοντέλου. Τα σημεία ελέγχου τοποθετούνται στην περιοχή μελέτης ως στόχοι με γνωστές συντεταγμένες, οι οποίες χρησιμοποιούνται για τον υπολογισμό των θέσεων των υπόλοιπων σημείων. Ανάλογα με το μέγεθος της περιοχής ορίζεται και ο αριθμός των σημείων ελέγχου που απαιτούνται ενώ ο ιδανικός αριθμός σημείων ελέγχου στα περισσότερα λογισμικά είναι πέντε. Η διάσπαρτη τοποθέτηση πολλών σημείων ελέγχου κοντά σε σημεία με απότομες υψομετρικές διαφορές δίνει ακριβέστερα αποτελέσματα. Η σήμανση τους στο έδαφος μπορεί να γίνει με διάφορους τρόπους (καρφιά κτλ.) ενώ το μέγεθος των στόχων εξαρτάται από το ύψος της κάθε πτήσης ώστε να είναι εύκολα διακριτοί στις εικόνες επεξεργασίας.
Έτσι, σε διάφορες τοποθεσίες της Βρίσας, έγινε σήμανση (με κόκκινο σπρέι) 10 σημείων ελέγχου στο έδαφος με σκοπό να γίνει η διαδικασία της γεωαναφοράς ευκολότερη και ορθότερη.
Έπειτα, αναζητήθηκε κάθε φωτογραφία που είχε σήμανση στο έδαφος (βλ. Εικόνα 26), και τοποθετήθηκε ένα σημείο ελέγχου (GSPs) (βλ. Εικόνα 27). Η διαδικασία αυτή είχε ως αποτέλεσμα να επισημανθούν 10 σημεία ελέγχου περιμετρικά και στο κέντρο της περιοχής μελέτης (βλ. Εικόνα 28). (Ατζαράκη, 2019)



Εικόνα 26: Εντοπισμός σημείων ελέγχου στο έδαφος



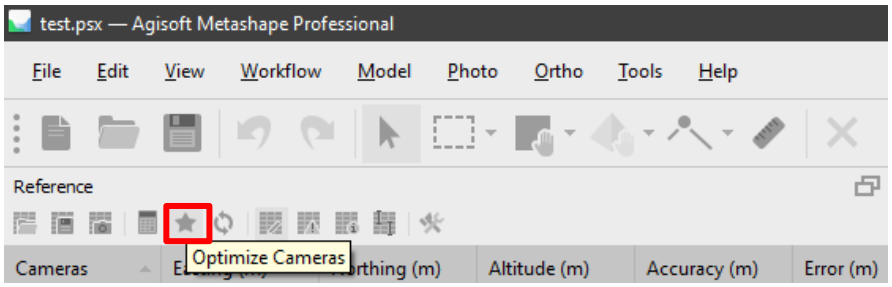
Εικόνα 27: Χρήση "Add Marker".



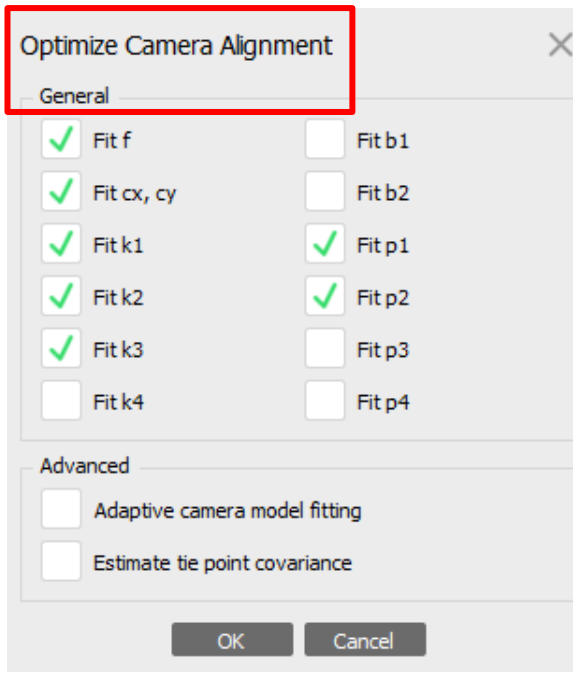
Εικόνα 28: Τοποθέτηση σημείων ελέγχου στο μοντέλο.

5.1.5. Βελτιστοποίηση προσανατολισμού

Ακολούθως, αφού τοποθετήθηκαν τα σημεία ελέγχου επόμενο στάδιο είναι η βελτιστοποίηση του προσανατολισμού (Optimize Camera ή αλλιώς Optimize the alignment) (βλ. Εικόνα 29). Έτσι, χρησιμοποιήθηκε η διαδικασία αυτή με τις προκαθορισμένες επιλογές του λογισμικού (βλ. Εικόνα 30), η οποία παράγει μια βαθμονόμηση της κάμερας, βασισμένη στις θέσεις των σημείων ελέγχου. Αυτή η διαδικασία, έχει ως αποτέλεσμα την προβολή των X, Y, Z σφαλμάτων (βλ. Εικόνα 31). Έτσι, εάν υπάρχουν σημεία με μεγαλύτερα σφάλματα από τα επιθυμητά, μπορεί να αφαιρεθεί από τη διαδικασία της γεωαναφοράς. (Ατζαράκη, 2019)



Εικόνα 29: Χρήση "Optimize Cameras".



Εικόνα 30: Προκαθορισμένες επιλογές λογισμικού.

	Easting (m)	Northing (m)	Altitude (m)	Accuracy (m)	Error (m)
1	690308.377000	4323433.023000	28.201000	0.005000	0.018075
2	689967.329000	4323160.778000	30.834000	0.005000	0.022329
3	690404.074000	4323032.069000	41.807000	0.005000	0.005109
4	690548.650000	4323202.701000	46.202000	0.005000	0.030477
5	690516.943000	4323301.311000	57.084000	0.005000	0.033986
6	690159.183000	4323327.839000	28.057000	0.005000	0.018599
7	690434.198000	4323282.739000	40.952000	0.005000	0.054343
8	690361.732000	4323160.406000	35.630000	0.005000	0.019116

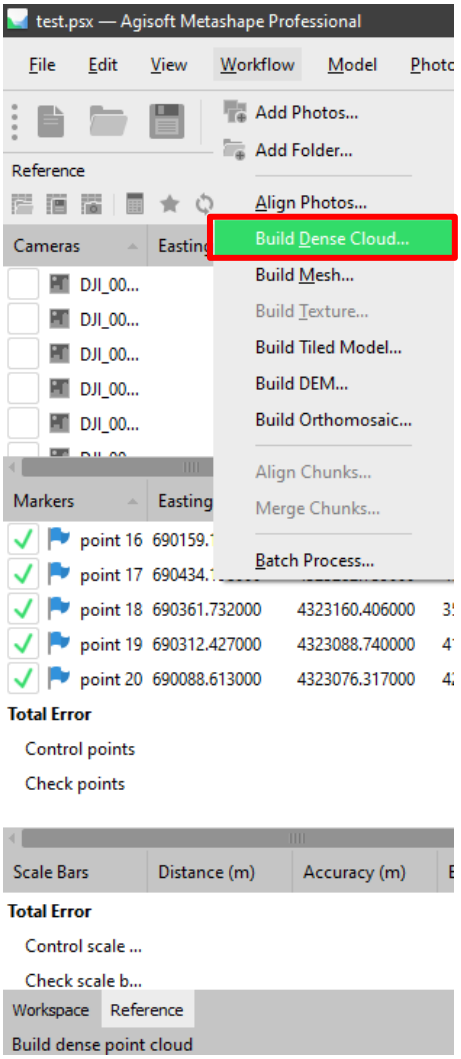
Εικόνα 31: Υπολογισμός σφάλματος.

5.1.6. Δημιουργία πυκνού νέφους σημείων

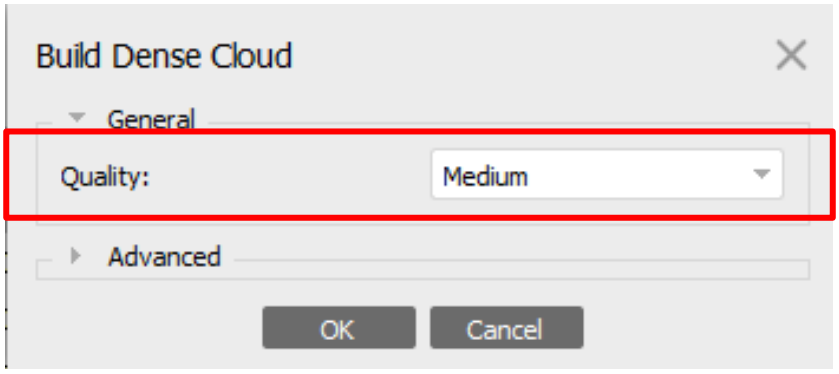
Επόμενο στάδιο είναι η δημιουργία νέφους σημείων (Dense Point Cloud) όπου είναι το ένα στάδιο πριν τη δημιουργία του Ψηφιακού Μοντέλου Επιφανείας δηλαδή της τρισδιάστατης απεικόνισης. σύννεφο σημείων είναι ένα σύνολο σημείων δεδομένων στο διάστημα .

Έτσι, αφού επιλέχθηκε η διαδικασία αυτή (βλ. Εικόνα 32), ορίστηκε η ποιότητα του νέφους σε μέτρια (Medium) καθώς προτείνεται για μεγάλα σύνολα φωτογραφιών (βλ.

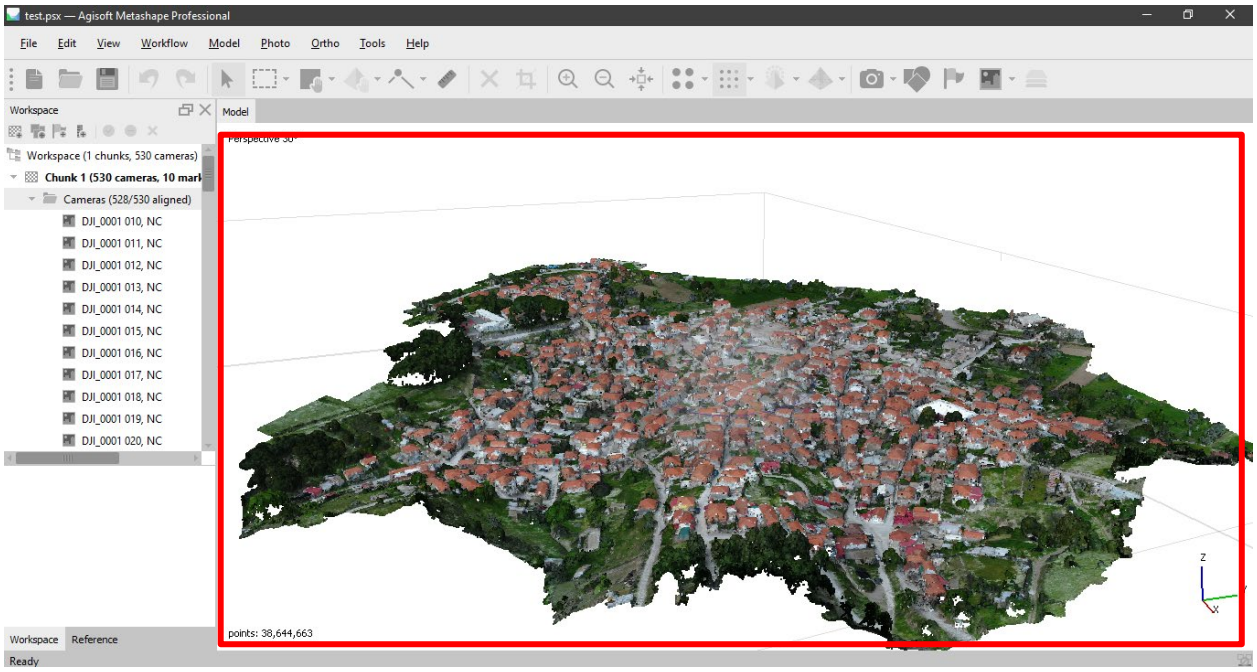
Εικόνα 33) και αυτό είχε ως αποτέλεσμα τη δημιουργία ενός πυκνού νέφους της περιοχής μελέτης (βλ. Εικόνα 34). Τα πυκνά νέφη σημείων, γενικά παράγονται με 3D σαρωτές, οι οποίοι μετρούν πολλά σημεία στις εξωτερικές επιφάνειες των αντικειμένων γύρω τους. Ως αποτέλεσμα των διεργασιών τρισδιάστατης σάρωσης, τα πυκνά νέφη σημείων χρησιμοποιούνται για πολλούς σκοπούς, συμπεριλαμβανομένης της δημιουργίας τρισδιάστατων μοντέλων και για πολλές εφαρμογές απεικόνισης, κινούμενης εικόνας, επεξεργασίας και προσαρμογής μαζικής παραγωγής. (Ατζαράκη, 2019)



Εικόνα 32: Χρήση "Build Dense Cloud".



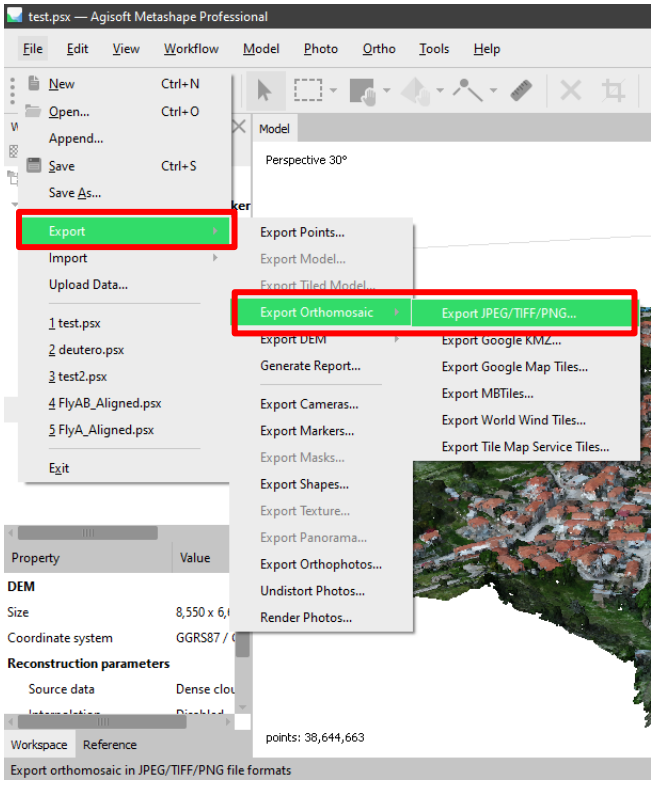
Εικόνα 33: Ορισμός της παραμέτρου ποιότητας.



Εικόνα 34: Δημιουργία πυκνού νέφους σημείων για τη Βρίσα.

5.1.7. Εξαγωγή ορθοφωτογραφίας

Αφού δημιουργήθηκε το πυκνό νέφος σημείων, τελευταία διαδικασία όσον αφορά τα πρωτεύοντα δεδομένα είναι η εξαγωγή ορθοφωτογραφίας (TIFF). Σε αυτό το στάδιο προσδιορίζεται μόνο η διαδικασία (βλ. Εικόνα 35). (Ατζαράκη, 2019)



Εικόνα 35: Εξαγωγή ορθοφωτογραφίας.

Εικόνα 36: Ορθοφωτοχάρτης 26-03-2019.

5.2.Διαμόρφωση λειτουργικού συστήματος

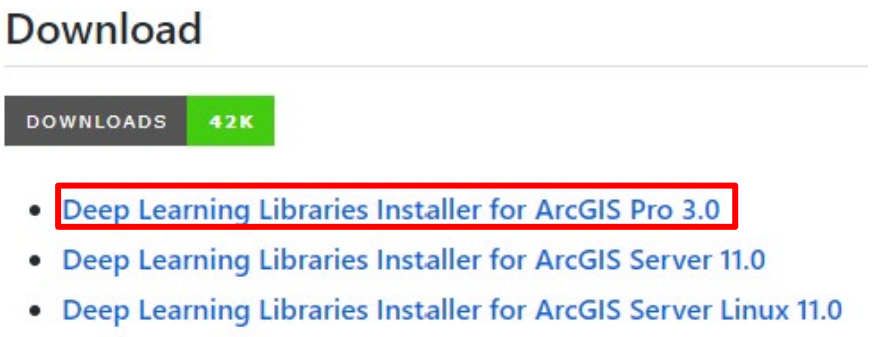
Για τις ανάγκες της εργασίας, απαραίτητη προϋπόθεση ήταν η διαμόρφωση του περιβάλλοντος επεξεργασίας στο λογισμικό δημιουργίας, διαχείρισης και επεξεργασίας ArcGIS Pro.

Με σκοπό να μπορέσει να εκτελεστεί το μοντέλο βαθιάς μάθησης, εγκαταστάθηκαν στον υπολογιστή οι απαραίτητες βιβλιοθήκες εκμάθησης για το μοντέλο που θα χρησιμοποιηθεί.

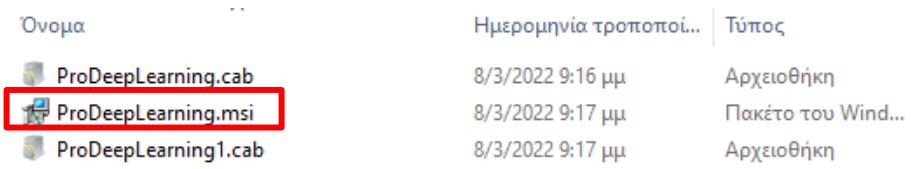
Πιο συγκεκριμένα, μέσω του ιστότοπου Github (Βλ. Ενότητα 8), έγινε λήψη και εγκατάσταση των βιβλιοθηκών βαθιάς εκμάθησης για το ArcGIS Pro. Η βαθιά εκμάθηση χρησιμοποιείται από πολλά εργαλεία λογισμικό για την επίλυση χωρικών προβλημάτων, την κατηγοριοποίηση χαρακτηριστικών και την εκτέλεση ταξινόμησης εικονοστοιχείων.

Το πρόγραμμα το οποίο εγκαταστάθηκε περιλαμβάνει ένα σύνολο στοιχείων όπως PyTorch, TensorFlow κ.ά. για την εκτέλεση εργασιών βαθιάς μάθησης και μηχανικής μάθησης, μια συνολική συλλογή 99 πακέτων. Αυτά τα πακέτα μπορούν να χρησιμοποιηθούν με τα εργαλεία Deep Learning Training, εισάγοντας τις παραμέτρους που απαιτεί ο χρήστης σύμφωνα με τις ανάγκες του.

Εικόνα 37: Βιβλιοθήκες βαθιάς μάθησης για ArcGIS. Walbridge et al., 2021



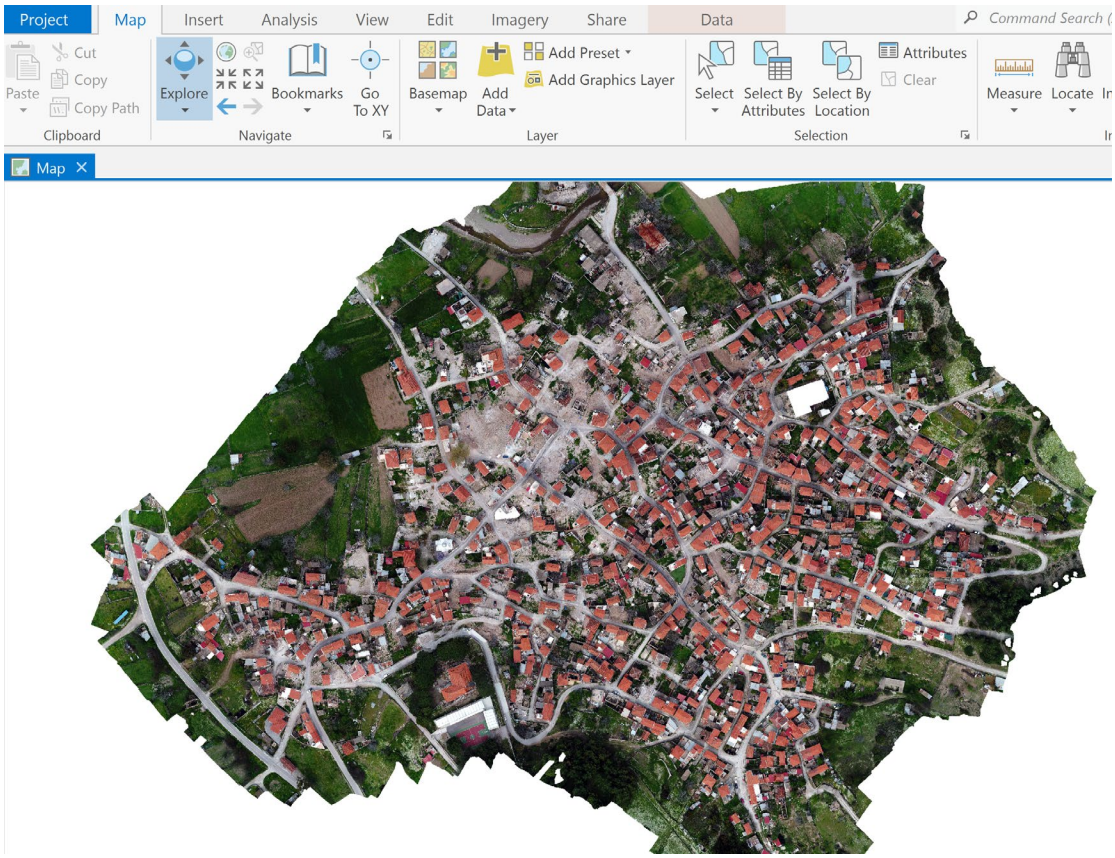
Εικόνα 38: Λήψη του πακέτου βιβλιοθηκών για το ArcGIS Pro.



Όνομα	Ημερομηνία τροποποί...	Τύπος
ProDeepLearning.cab	8/3/2022 9:16 μμ	Αρχειοθήκη
ProDeepLearning.msi	8/3/2022 9:17 μμ	Πακέτο του Wind...
ProDeepLearning1.cab	8/3/2022 9:17 μμ	Αρχειοθήκη

Εικόνα 39:Εγκατάσταση του πακέτου βιβλιοθηκών για το ArcGIS Pro.

Αφού πραγματοποιήθηκε η εγκατάσταση των βιβλιοθηκών, επόμενο βήμα ήταν η εισαγωγή της εικόνας (ορθοφωτογραφίας) στο λογισμικό επεξεργασίας.



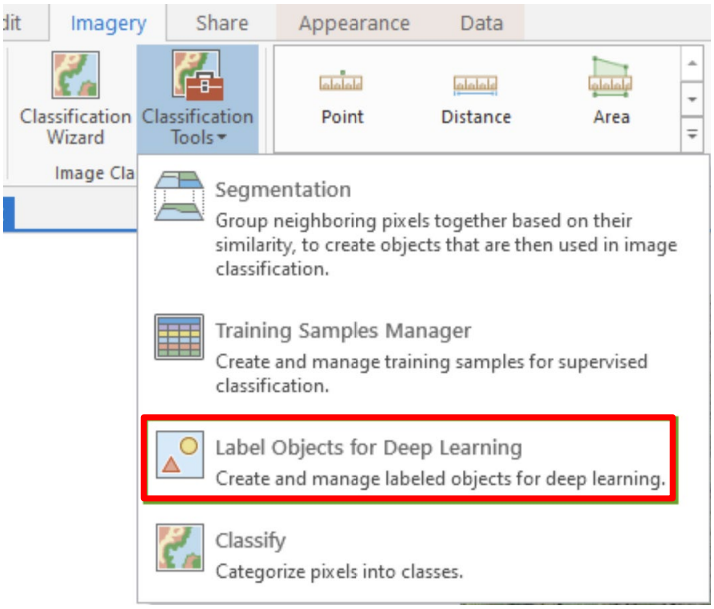
Εικόνα 40: Εισαγωγή της ορθοφωτογραφίας στο λογισμικό.

5.3.Δημιουργία δειγμάτων εκπαίδευσης

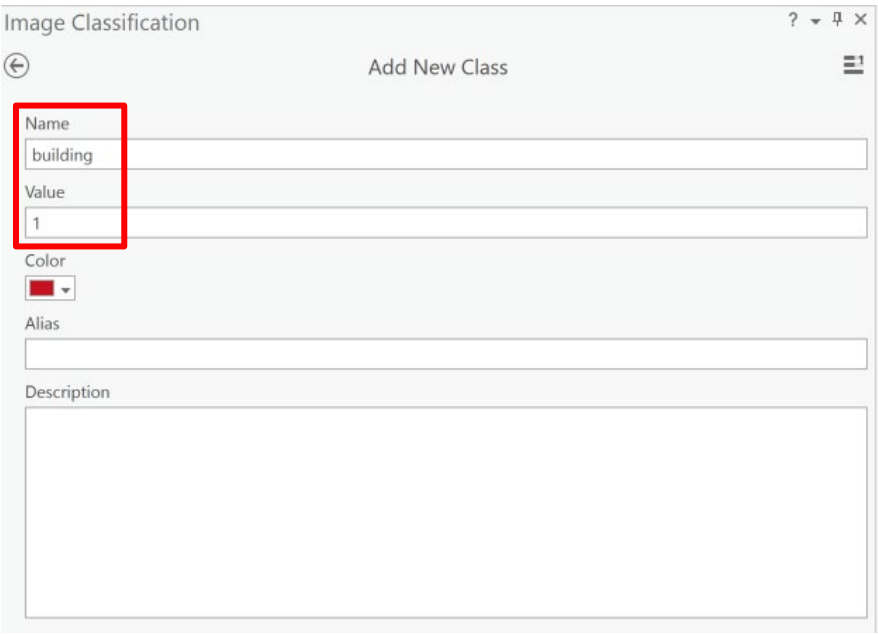
Οι επιβλεπόμενες εργασίες βαθιάς μάθησης εξαρτώνται από επισημασμένα σύνολα δεδομένων.
Αυτό σημαίνει ότι ο άνθρωπος πρέπει να εφαρμόσει τις γνώσεις του για να εκπαιδεύσει το νευρωνικό δίκτυο σε αυτό που δουλεύει για να αναγνωρίσει.
Όταν τα δεδομένα εκπαίδευσης είναι σωστά, ένα σύστημα βαθιάς μάθησης μπορεί να είναι εξαιρετικά ακριβές στην εξαγωγή χαρακτηριστικών και την ανίχνευση αντικειμένων.
Για την εκπαίδευση ενός μοντέλου βαθιά μάθησης είναι απαραίτητο να δημιουργηθούν αξιόπιστα δείγματα εκπαίδευσης. Σκοπός είναι να παρέχονται στο μοντέλο βαθιάς μάθησης οι πληροφορίες που χρειάζεται για να ανιχνεύσει όλα τα κτίρια επάνω στην εικόνα.
Επομένως, στόχος είναι μέσα από αυτή τη διαδικασία είναι να ψηφιοποιηθούν τα δείγματα εκπαίδευσης και να δημιουργηθούν τα απαραίτητα χαρακτηριστικά για ένα ορισμένο αριθμό κτιρίων και να εκπαιδευτεί το μοντέλο για το μέγεθος, το σχήμα και η φασματική υπογραφή των κτιρίων.

5.3.1. Δημιουργία εκπαιδευτικού σχήματος

Τα δείγματα αυτά, δημιουργούνται και τροποποιούνται μέσω του εργαλείου καταχώρησης ετικετών αντικειμένων για βαθιά μάθηση (Label Objects for Deep Learning).
Τα αντικείμενα με ετικέτα θα χρησιμοποιηθούν από το νευρωνικό δίκτυο για την εκπαίδευση ενός μοντέλου που μπορεί να χρησιμοποιηθεί για την εξαγωγή των συμπερασμάτων στην εικόνα.
Έτσι, από το σύνολο εργαλείων ταξινόμησης, επιλέχθηκε εκείνο για την επισήμανση των αντικειμένων.
Η επισήμανση πραγματοποιείται χειροκίνητα, ψηφιοποιώντας το αντικείμενο που είναι προς ανίχνευση δηλαδή τα κτίρια σε όλη την εικόνα.
Για τις ανάγκες δημιουργίας εκπαιδευτικού σχήματος, σχεδιάστηκε ένα σχήμα με μία μόνο κλάση καθώς η κατηγορία αντικειμένου προς ανίχνευση είναι μια, τα κτίρια. Ως τιμή για τα κτίρια δόθηκε ο αριθμός 1 ώστε να εκπαιδευτεί από τον υπολογιστή ένα μοντέλο. Η κλάση που δημιουργήθηκε είναι το κτίριο και το εκπαιδευτικό σχήμα αφορούσε τα κτίρια.



Εικόνα 41: Εργαλείο ετικετών αντικειμένων.



Εικόνα 42: Εισαγωγή κλάσης και καταχώρηση τιμής.

5.3.2. Συλλογή δειγμάτων εκπαίδευσης

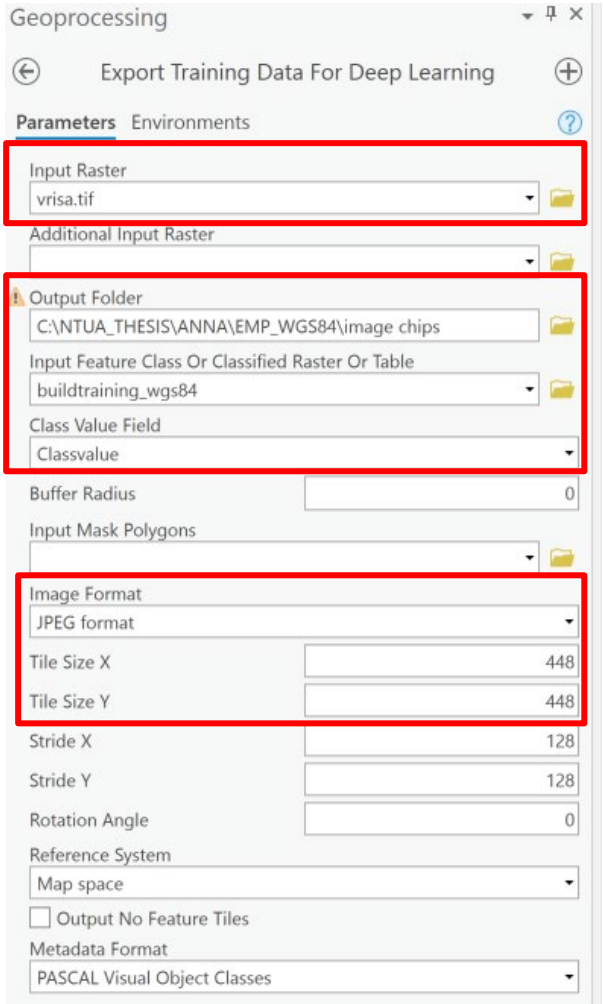
Ακολούθως, για την επισήμανση της κλάσης των κτιρίων, ορίστηκε η κατάλληλη χρωματική παλέτα για την ψηφιοποίηση τους. Τα δείγματα εκπαίδευσης που συλλέχθηκαν ήταν πολυγωνικά, διανυσματικής μορφής.
Συνολικά συγκεντρώθηκαν 280 εκπαιδευτικά δείγματα δηλαδή ψηφιοποιήθηκαν 280 κτίρια για την εκπαίδευση του νευρωνικού δικτύου.

Εικόνα 43: Τα δείγματα εκπαίδευσης που συλλέχθηκαν.

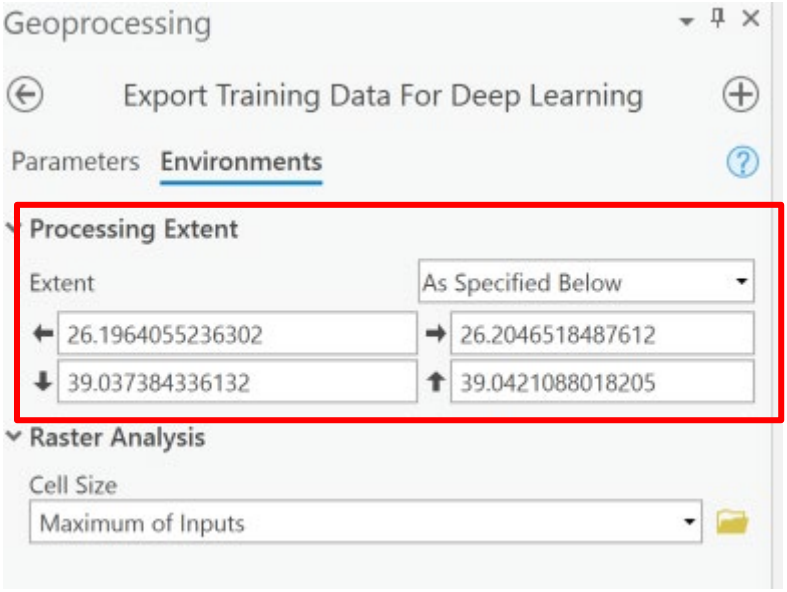
5.3.3. Δημιουργία patch εικόνων

Έπειτα αφού συλλέχθηκαν τα δείγματα εκπαίδευσης και αποθηκεύτηκαν, επόμενο στάδιο ήταν δημιουργία των patch εικόνων (Export Training Data For Deep Learning).
Τα δείγματα εκπαίδευσης κλάσεων Deep Learning βασίζονται σε μικρές δευτερεύουσες εικόνες, που ονομάζονται patch εικόνας. Αυτές οι εικόνες περιέχουν το χαρακτηριστικό ενδιαφέροντος, τα κτίρια.
Σκοπός αυτού του βήματος είναι να μετατραπούν τα δείγματα εκπαίδευσης σε σύνολα δεδομένων εκπαίδευσης βαθιάς μάθησης χρησιμοποιώντας την ορθοφωτογραφία.
Ως παράμετροι επιλέχθηκαν, η ορθοφωτογραφία, το εκπαιδευτικό σχήμα που περιέχει την τιμή κλάσης 1 (Βλ. Ενότητα 5.3.1) ως ταξινομημένο ράστερ και ως μορφή εξαγωγής JPEG. Ως μορφή μεταδεδομένων επιλέχθηκε η PASCAL Visual Object Classes.
Η εξαγωγή περιλαμβάνει ένα φάκελο με δύο φακέλους. Έναν ο οποίος περιλαμβάνει τα patch εικόνας (δηλαδή μορφής JPEG) και έναν φάκελο με αρχεία μεταδεδομένων στην μορφή XML που περιέχουν πληροφορίες σχετικά με το όνομα της εικόνας, την τιμή της κλάσης και τα πλαίσια οριοθέτησης.
Επιπλέον κατά την εξαγωγή ο φάκελος περιέχει δύο αρχεία κειμένου, ένα αρχείο .json και ένα αρχείο .emd (Βλ. Εικόνα 46).
Στη συνέχεια ορίστηκε το μέγεθος σε εικονοστοιχεία, για κάθε patch εικόνας.
Το μέγεθος του patch εικόνας καθορίζεται από το μέγεθος των χαρακτηριστικών που πρόκειται να εντοπιστούν. Εάν το χαρακτηριστικό είναι μεγαλύτερο από τις διαστάσεις x και y των πλακιδίων, το μοντέλο δεν θα παρέχει καλά αποτελέσματα. Έτσι ως μέγεθος επιλέχθηκε για x και y η τιμή 448.

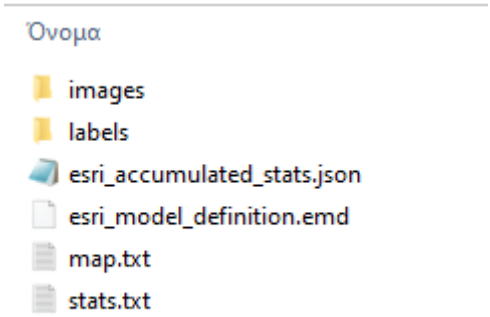
Επιπλέον παράμετρος που επιλέχθηκε πριν την ολοκλήρωση της δημιουργίας patch εικόνων είναι η ανάλυση της εικόνας ώστε τα patch εικόνας να είναι στην ίδια ανάλυση με της εικόνας.
Έτσι επιλέχθηκε ως μέγεθος εικονοστοιχείων το ίδιο με εκείνο της εικόνας και εξήχθη ο φάκελος των patch εικόνας.



Εικόνα 44: Ορισμός παραμέτρων για την δημιουργία των τσίπ.



Εικόνα 45: Ορισμός παραμέτρων για την δημιουργία των τσίπ.



Εικόνα 46: Αποτέλεσμα της εξαγωγής των patch εικόνας και τα μεταδεδομένα τους.

Εικόνα 47: Μέρος των patch εικόνας.

5.4.Ανίχνευση κτιρίων με χρήση μοντέλου βαθιάς μάθησης

Έχοντας ως στόχο την ανίχνευση κτιρίων, θα πρέπει να εκπαιδευτεί ένα μοντέλο. Για τις ανάγκες τις εργασίας, εκπαιδεύτηκαν τρία διαφορετικά μοντέλα ώστε να επιλεχθεί το βέλτιστο αποτέλεσμα.

Για την εκπαίδευση του μοντέλου βαθιάς μάθησης θα πρέπει να ληφθούν τα δεδομένα του δείγματος εκπαίδευσης και να μεταφερθούν μέσω ενός νευρωνικού δικτύου ξανά και ξανά.

Όταν δημιουργήθηκαν τα τρία διαφορετικά μοντέλα, εφαρμόστηκαν στην ορθοφωτογραφία ώστε να αναγνωριστούν τα κτίρια αυτόματα.

Η διαδικασία εντοπισμού κτιρίων πραγματοποιήθηκε τέσσερεις φορές για καθένα από τα μοντέλα εκπαίδευσης χρησιμοποιώντας διαφορετικές παραμέτρους.

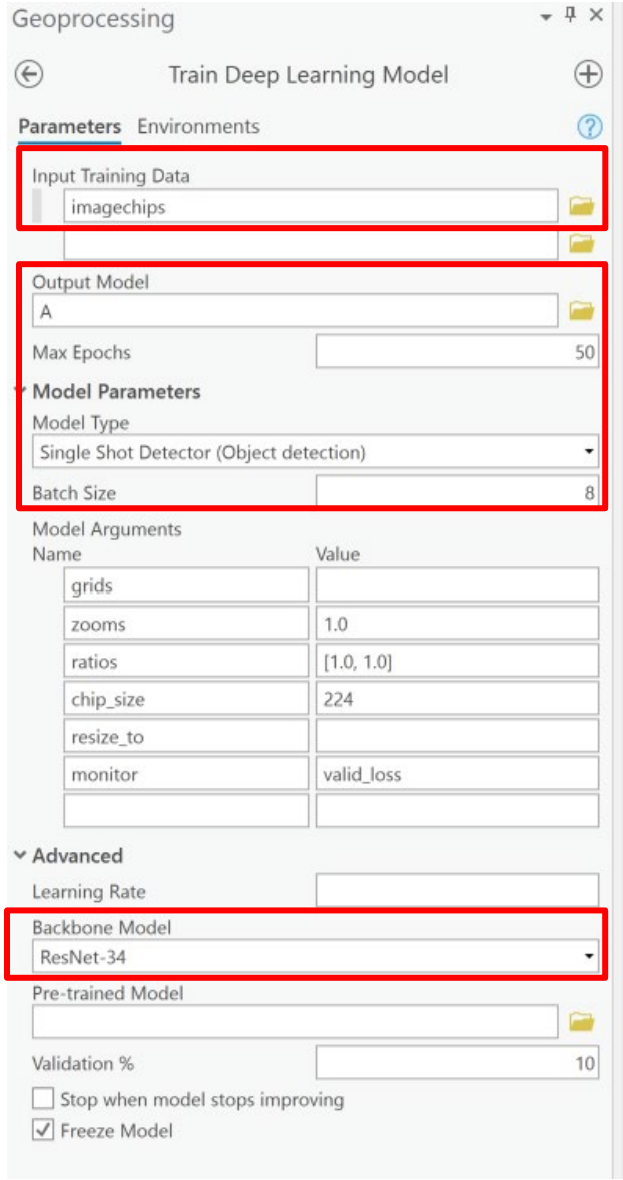
5.4.1. Εκπαίδευση μοντέλου βαθιάς μάθησης

Όπως προαναφέρθηκε, δημιουργήθηκαν 3 διαφορετικά μοντέλα, χρησιμοποιώντας διαφορετικές παραμέτρους όπως φαίνεται στον παρακάτω πίνακα.

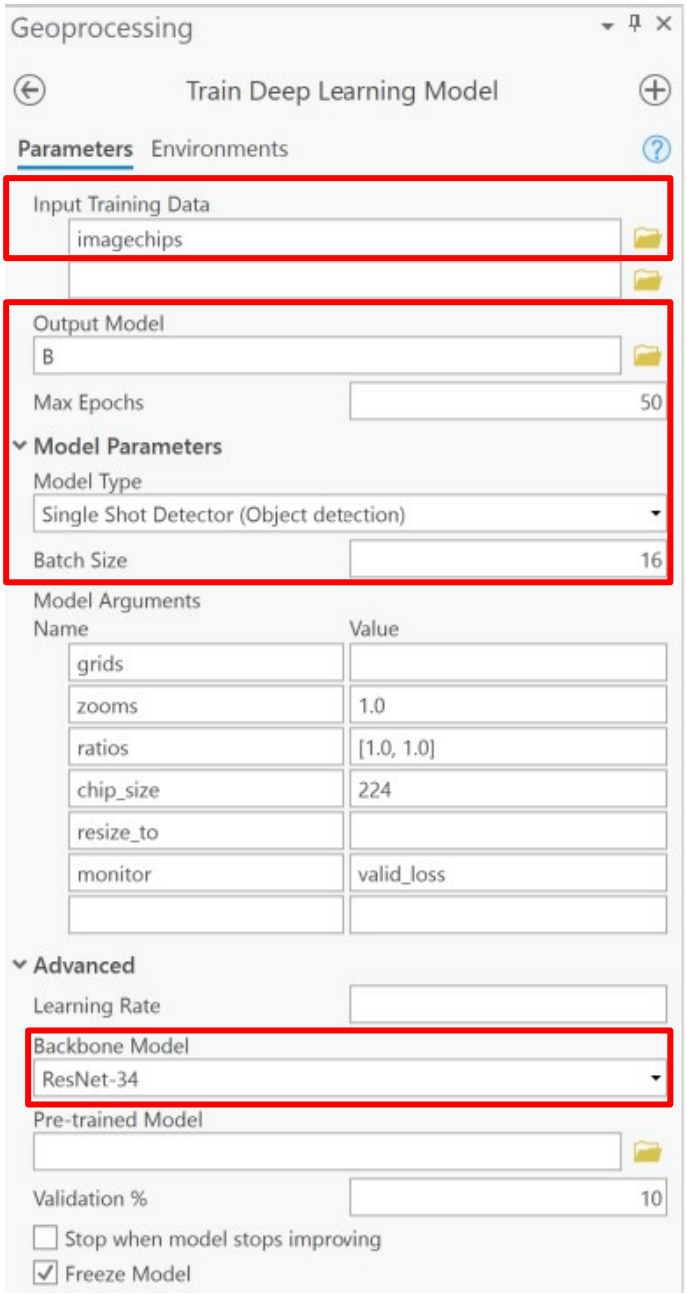
TRAINING MODEL	A	B	C
Max epochs	50	50	50
Batch	8	16	4

Πίνακας 6: Τα τρία μοντέλα εκπαίδευσης με διαφορετικές παραμέτρους.

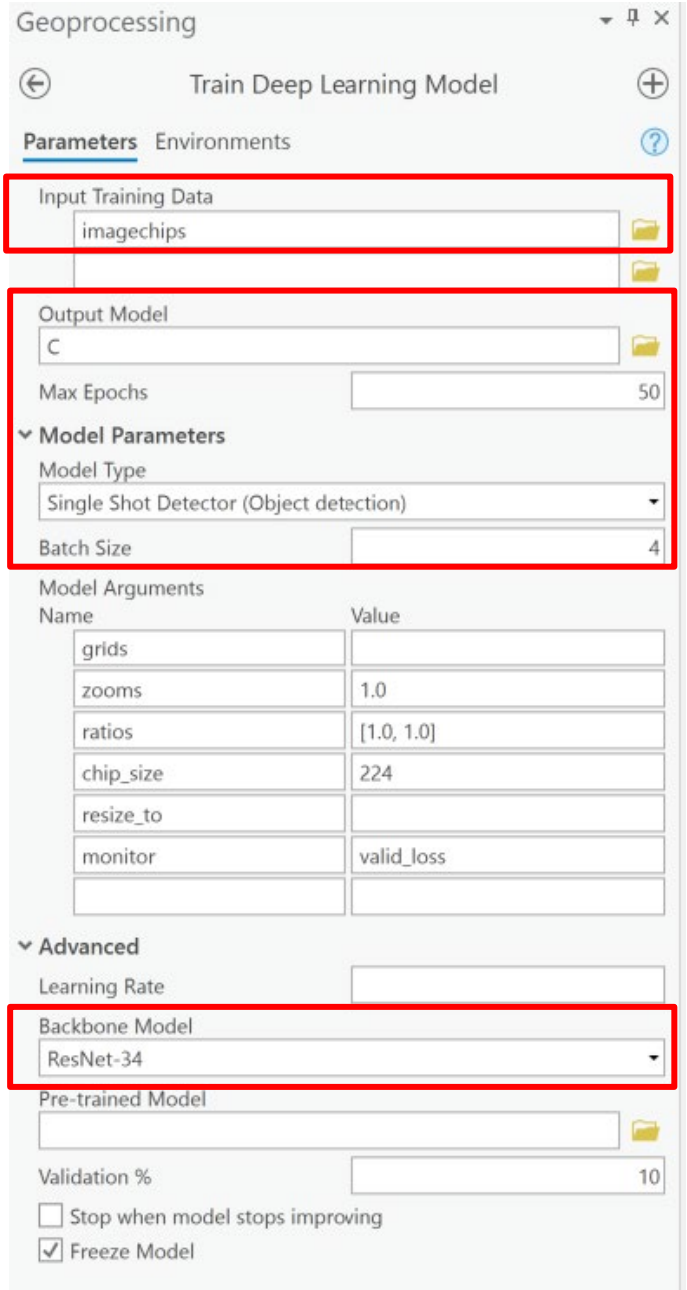
Για να εκπαιδευτεί το μοντέλο για το τι αποτελεί και τι όχι ως ένα κτίριο, η διαδικασία διεκπεραιώθηκε μέσω του εργαλείου γεωεπεξεργασίας (Train Deep Learning Model).
Το εργαλείο αυτό, χρησιμοποιεί ως δεδομένα εκπαίδευσης τα patch εικόνας που επισημάνθηκαν. Τα δεδομένα αυτά χρησιμοποιήθηκαν για να εκπαιδευτεί ένα μοντέλο βαθιάς μάθησης Single Shot Detector (SSD) και να προσδιοριστεί ποιοι συνδυασμοί εικονοστοιχείων αποτελούν ένα κτίριο.
Μια αρχιτεκτονική μοντέλου SSD που χρησιμοποιεί το Resnet-34 ως βασικό μοντέλο έχει ήδη προκαθοριστεί στο λογισμικό ArcGIS Pro, γεγονός που το καθιστά εύκολο στη χρήση.
Το προδιαμορφωμένο μοντέλο ResNet-34 είναι ένα υπολειπόμενο δίκτυο εκπαιδευμένο στο σύνολο δεδομένων ImageNet που περιέχει περισσότερες από 1 εκατομμύριο εικόνες και έχει βάθος 34 επιπέδων. (ESRI, χ.χ.)
Επομένως πρώτο στάδιο για την ολοκλήρωση της εκπαίδευσης, ήταν να εισαχθούν τα δεδομένα εκπαίδευσης δηλαδή ο φάκελος με τα patch εικόνας. Ακολούθως επόμενο στάδιο ήταν να οριστεί ο αριθμός των εποχών που θα τρέχει το μοντέλο. Και στα τρία μοντέλα ο αριθμός τον εποχών ήταν κοινός.
Μια εποχή αποτελεί ένα κύκλο μέσα από το σύνολο δεδομένων εκπαίδευσης που βρίσκονται στον φάκελο των patch εικόνας.
Το δεδομένα αυτά, κατά τη διάρκεια κάθε εποχής θα μεταφέρονται με κατεύθυνση εμπρός και πίσω μέσω του νευρωνικού δικτύου μία φορά. Έτσι και στα τρία μοντέλα (A), (B), (C) ορίστηκε η τιμή 50 ως αριθμός εποχών.
Ακόλουθο στάδιο για την εκπαίδευση του μοντέλου ήταν να οριστεί το μέγεθος της παρτίδας δηλαδή ο αριθμός των δειγμάτων που θα εκπαιδεύονταν κάθε φορά.
Η διαφοροποίηση των τριών μοντέλων βασίζεται σε αυτή τη παράμετρο καθώς οι τιμές των παρτίδων σε καθένα από αυτά ήταν διαφορετικές.
Στο πρώτο μοντέλο (A) ο αριθμός των δειγμάτων που εκπαιδεύτηκαν ήταν 8, στο δεύτερο μοντέλο (B) ο αριθμός των δειγμάτων που εκπαιδεύτηκαν ήταν 16 ενώ στο τρίτο (C), ο αριθμός των δειγμάτων ήταν 4. (Βλ. Πίνακας 6).



Εικόνα 48: Ορισμός παραμέτρων για εκπαίδευση του πρώτου (A) μοντέλου βαθιάς μάθησης.



Εικόνα 49: Ορισμός παραμέτρων για εκπαίδευση του δεύτερου (B) μοντέλου βαθιάς μάθησης.



Εικόνα 50: Ορισμός παραμέτρων για εκπαίδευση του τρίτου (C) μοντέλου βαθιάς μάθησης.

5.4.2. Εντοπισμός κτιρίων με χρήση μοντέλου βαθιάς μάθησης

Αφότου ολοκληρώθηκε το μεγαλύτερο τμήμα της εργασίας που αφορούσε την προετοιμασία των δεδομένων, τη δημιουργία των δειγμάτων εκπαίδευσης και την εκπαίδευση του μοντέλου, επόμενο στάδιο ήταν ο εντοπισμός των κτιρίων (Detect Objects Using Deep Learning).

Κάθε εκπαιδευτικό μοντέλο, χρησιμοποιήθηκε με σκοπό να ανιχνευτούν τα κτίρια στην εικόνα.

Έχοντας ως στόχο να επιτευχθεί το καλύτερο αποτέλεσμα, εφαρμόστηκαν πολλαπλές δοκιμές τροποποιώντας τις παραμέτρους. Για να εντοπιστεί ποιο μοντέλο αποδίδει καλύτερα αλλά και με ποιες παραμέτρους, οι δοκιμές διαμορφώνονται στον παρακάτω πίνακα.

DETECTION	A_1	A_2	A_3	A_4
model	A	A	A	A
padding	56	56	56	56
threshold	0.2	0.5	0.8	0.8
nms_overlap	0.1	0.1	0.1	0.5
batch_size	8	8	8	8
DETECTION	B_1	B_2	B_3	B_4
model	B	B	B	B
padding	56	56	56	56
threshold	0.2	0.5	0.8	0.8
nms_overlap	0.1	0.1	0.1	0.5
batch_size	8	8	8	8
DETECTION	C_1	C_2	C_3	C_4
model	C	C	C	C
padding	56	56	56	56
threshold	0.2	0.5	0.8	0.8
nms_overlap	0.1	0.1	0.1	0.5
batch_size	8	8	8	8

Πίνακας 7: Εφαρμογή ανίχνευσης των τριών μοντέλων με διαφορετικές παραμέτρους.

Για καθένα από τα τρία μοντέλα, πραγματοποιήθηκαν τέσσερις εφαρμογές ανίχνευσης κτιρίου διαφοροποιώντας τις παραμέτρους.

Για την εφαρμογή ανίχνευσης των κτιρίων, οι παράμετροι που συνέβαλαν στην βελτίωση των αποτελεσμάτων – της ακρίβειας ήταν το κατώφλι εμπιστοσύνης (threshold), το μέγεθος παρτίδας (batch size), περιοχή συνόρων (padding) και η επικάλυψη μη μέγιστης καταστολής (nms_overlap).

Το κατώφλι εμπιστοσύνης (threshold), αποτελεί ένα όριο το οποίο καθορίζει ο χρήστης για να επισημάνει το μέγεθος της εμπιστοσύνης που είναι αποδεκτή για να χαρακτηριστεί ένα αντικείμενο ως κτίριο.
Τα μοντέλα πρόβλεψης εξάγουν ένα επίπεδο εμπιστοσύνης για κάθε χαρακτηριστικό (δηλ. κτίριο) που ανιχνεύουν. Η παράμετρος κατωφλίου ορίζει το ελάχιστο επίπεδο εμπιστοσύνης που θα συμπεριληφθεί στην έξοδο (Alouta & Hess, 2021).
Αντίστοιχα το μέγεθος παρτίδας (batch size) είναι ένας αριθμός δειγμάτων που υποβάλλονται σε επεξεργασία πριν από την ενημέρωση του μοντέλου. Ο αριθμός των εποχών είναι ο αριθμός των ολοκληρωμένων διελεύσεων μέσω του συνόλου δεδομένων εκπαίδευσης. Το μέγεθος μιας παρτίδας πρέπει να είναι μεγαλύτερο ή ίσο με ένα και μικρότερο ή ίσο με τον αριθμό των δειγμάτων στο σύνολο δεδομένων εκπαίδευσης. (Brownlee, 2022).
Το padding είναι η περιοχή συνόρων από την οποία το μοντέλο θα απορρίψει τις ανιχνεύσεις, καθώς τείνουν να αποτελούνται από περικομμένα κτίρια που εκτείνονται σε πολλά πλακίδια κατά τη διάρκεια της εξαγωγής συμπερασμάτων (Alouta & Hess, 2021).
Κατά την εφαρμογή συνέλιξης εικόνων, στη μοντελοποίηση συνελικτικού νευρωνικού δικτύου, τα δεδομένα και τα εικονοστοιχεία «συρρικνώνονται» στην άκρη της εικόνας και χρησιμοποιούνται πολύ λιγότερο κατά την ανάλυση, σε σύγκριση με τα εσωτερικά εικονοστοιχεία. Η παράμετρος **padding** προσθέτει ένα επιπλέον όριο pixel στα εξωτερικά άκρα της εικόνας. Αυτό μειώνει την απώλεια πληροφοριών από τα έγκυρα εικονοστοιχεία άκρων και τη συρρίκνωση (Lenhardt, 2021).
Αρθρο της ESRI (χ.χ.) αναφέρει πως η επικάλυψη μη μέγιστης καταστολής (nms_overlap), είναι απαραίτητη για τη συγχώνευση επικαλυπτόμενων χαρακτηριστικών. Εάν δύο χαρακτηριστικά επικαλύπτονται περισσότερο από μια δεδομένη μέγιστη αναλογία, η δυνατότητα με τη χαμηλότερη τιμή εμπιστοσύνης θα καταργηθεί.
Έτσι, για το **πρώτο μοντέλο (Α)**, πραγματοποιήθηκαν τέσσερις εφαρμογές ανίχνευσης.

- Η πρώτη εφαρμογή, ως κατώφλι εμπιστοσύνης (threshold) είχε την τιμή 0.2, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Αντίστοιχα, η δεύτερη εφαρμογή είχε ως κατώφλι εμπιστοσύνης (threshold) την τιμή 0.5, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Η τρίτη εφαρμογή, είχε ως κατώφλι εμπιστοσύνης (threshold) την τιμή 0.8, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Η τέταρτη και τελευταία εφαρμογή ως κατώφλι εμπιστοσύνης (threshold) είχε την τιμή 0.8, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.5.

Για το **δεύτερο μοντέλο (B)**, πραγματοποιήθηκαν τέσσερις εφαρμογές ανίχνευσης.

- Η πρώτη εφαρμογή, ως κατώφλι εμπιστοσύνης (threshold) είχε την τιμή 0.2, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Αντίστοιχα, η δεύτερη εφαρμογή είχε ως κατώφλι εμπιστοσύνης (threshold) την τιμή 0.5, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Η τρίτη εφαρμογή, είχε ως κατώφλι εμπιστοσύνης (threshold) την τιμή 0.8, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Η τέταρτη και τελευταία εφαρμογή ως κατώφλι εμπιστοσύνης (threshold) είχε την τιμή 0.8, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.5.

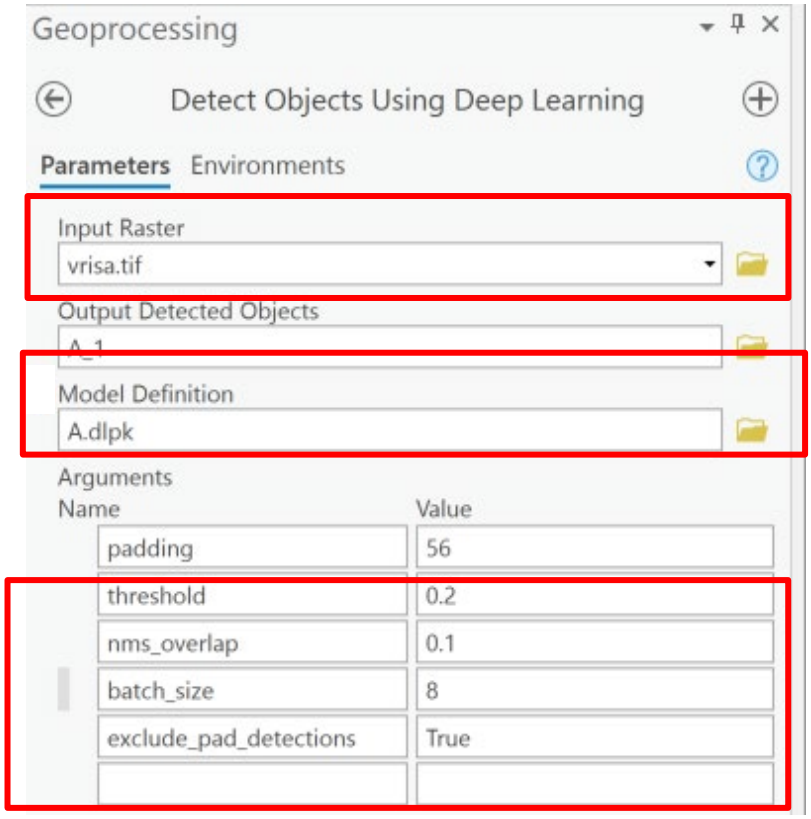
Τέλος, για το **τρίτο μοντέλο (C)**, πραγματοποιήθηκαν εξίσου τέσσερις εφαρμογές ανίχνευσης.

- Η πρώτη εφαρμογή, ως κατώφλι εμπιστοσύνης (threshold) είχε την τιμή 0.2, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Αντίστοιχα, η δεύτερη εφαρμογή είχε ως κατώφλι εμπιστοσύνης (threshold) την τιμή 0.5, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

- Η τρίτη εφαρμογή, είχε ως κατώφλι εμπιστοσύνης (threshold) την τιμή 0.8, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.1.

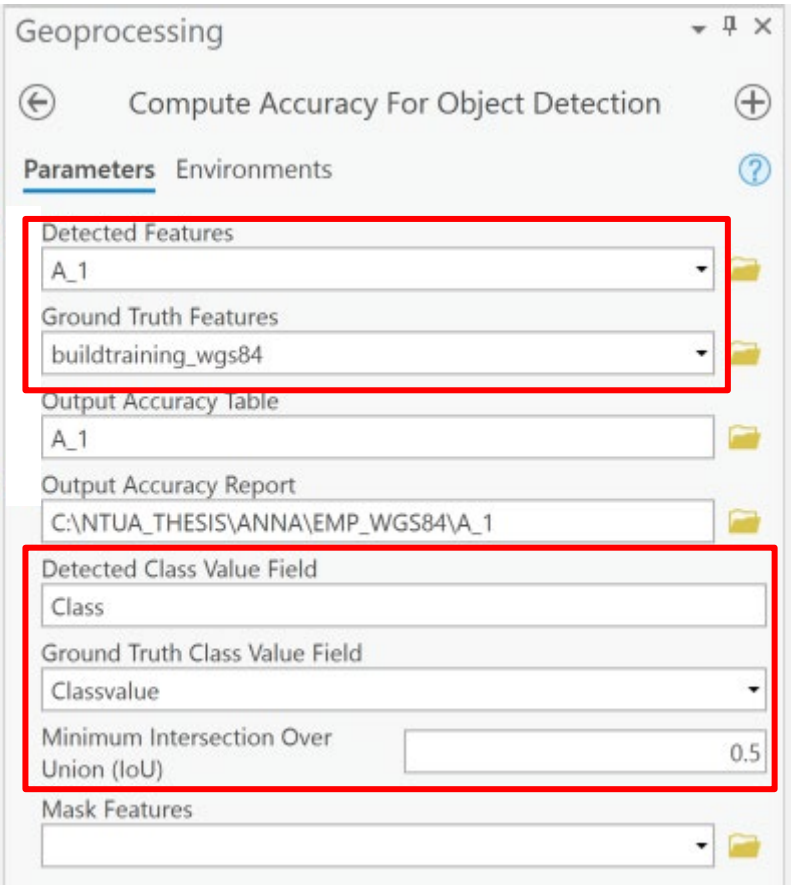
- Η τέταρτη και τελευταία εφαρμογή ως κατώφλι εμπιστοσύνης (threshold) είχε την τιμή 0.8, ως μέγεθος παρτίδας (batch size) ορίστηκε η τιμή 8, η περιοχή συνόρων (padding) 56 και η επικάλυψη μη μέγιστης καταστολής (nms_overlap) 0.5.



Εικόνα 51: Ορισμός παραμέτρων για την πρώτη εφαρμογή εντοπισμού κτιρίων του πρώτου (A) μοντέλου βαθιάς μάθησης.

5.5. Υπολογισμός ακρίβειας ανίχνευσης κτιρίων

Τέλος, με σκοπό να κατανοηθούν τα αποτελέσματα των τριών μοντέλων ανίχνευσης, υπολογίστηκε η ακρίβεια για κάθε τάξη στα δεδομένα που εντοπίστηκαν. Η ακρίβεια ενός μοντέλου βαθιάς μάθησης, υπολογίζεται συγκρίνοντας τα κτίρια που ανιχνεύτηκαν με τα δεδομένα αναφοράς . Το εργαλείο που χρησιμοποιείτε (Compute Accuracy For Object Detection) έχει ως αποτέλεσμα έναν πίνακα σχετικό με την ακρίβεια της ανίχνευσης που περιλαμβάνει μετρήσεις όπως Precision, Recall, F1_Score, True_Positive, False_Positive και False_Negative (Βλ. Ενότητα 1.4). Οι παράμετροι που λαμβάνονται θέση στον υπολογισμό της ακρίβειας είναι τα δεδομένα αναφοράς (ground truth features), το πεδίο με τη τιμή της κλάσης της εφαρμογής ανίχνευσης (detected class value field), το πεδίο με τη τιμή της κλάσης των δεδομένων αναφοράς (ground truth class value field) και η μέτρηση intersection over union (IoU). Και για τα δώδεκα πειράματα οι παράμετροι ήταν κοινοί καθώς επίσης στην μέτρηση IoU δόθηκε η τιμή 0.5 καθώς προτείνεται στην ανίχνευση αντικειμένων (Βλ. Ενότητα 1.4).



Εικόνα 52: Υπολογισμός ακρίβειας για την πρώτη εφαρμογή εντοπισμού κτιρίων του πρώτου (A) μοντέλου βαθιάς μάθησης.

Έτσι, εξήχθη ένας πίνακας ο οποίος περιέχει τα ακόλουθα πεδία: Precision, Recall, F1_Score, True_Positive, False_Positive, False_Negative.

Τα αποτελέσματα των παραπάνω εφαρμογών, προβάλλονται στην ακόλουθη ενότητα.

6. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΑΞΙΟΛΟΓΗΣΗ

6.1. Αποτελέσματα εκπαίδευσης μοντέλων βαθιάς μάθησης

6.1.1. Αποτελέσματα εκπαίδευσης πρώτου μοντέλου

Εικόνα 53: Εκπαίδευση Α μοντέλου βαθιάς μάθησης.

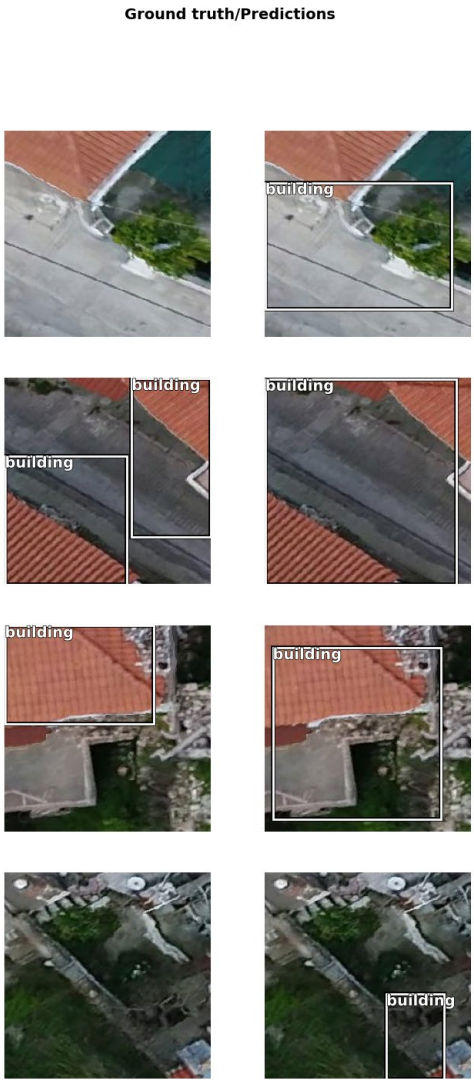
Στην παραπάνω εικόνα, φαίνονται τα αποτελέσματα της εκπαίδευσης του πρώτου (Α) μοντέλου βαθιάς μάθησης. Αριστερά της εικόνας φαίνονται τα σημεία στην εικόνα που έχουν οριστεί ως δείγματα εκπαίδευσης. Δεξιά της εικόνας φαίνονται τα σημεία που πρόβλεψε το μοντέλο. Στο πρώτο δείγμα, δεν δηλώθηκε κτίριο (αριστερά) και αντίστοιχα το μοντέλο δεν πρόβλεψε κτίριο (δεξιά). Στο δεύτερο δείγμα μαζί με το κτίριο δηλώθηκε και ένα τμήμα του δρόμου (αριστερά) με αποτέλεσμα το μοντέλο να προβλέψει λανθασμένα ολόκληρο το τμήμα του δρόμου ως κτίριο (δεξιά). Στο τρίτο δείγμα, δηλώθηκε το κτίριο ως δείγμα εκπαίδευσης (αριστερά) και το μοντέλο ορθά πρόβλεψε το κτίριο (δεξιά). Σημαντικό είναι να αναφερθεί πως σε αυτό το σημείο το μοντέλο πρόβλεψε καλύτερα από το χρήστη το κτίριο καθώς αριστερά φαίνεται ένα τμήμα εκτός σκεπής του κτιρίου και δεξιά αυτό το τμήμα έχει παραληφθεί.
Στο τέταρτο δείγμα, αριστερά φαίνεται ότι δεν έχει δηλωθεί κτίριο ενώ δεξιά το μοντέλο πρόβλεψε ένα κτίριο. Στο πέμπτο και τελευταίο δείγμα φαίνεται ένα τμήμα που έχει δηλωθεί ως κτίριο (αριστερά) και το μοντέλο πρόβλεψε και τα γειτονικά κτίρια μαζί με εκείνο (δεξιά).

6.1.2. Αποτελέσματα εκπαίδευσης δεύτερου μοντέλου

Εικόνα 54: Εκπαίδευση Β μοντέλου βαθιάς μάθησης.

Στην παραπάνω εικόνα, φαίνονται τα αποτελέσματα της εκπαίδευσης του δεύτερου (B) μοντέλου βαθιάς μάθησης. Αριστερά της εικόνας φαίνονται τα σημεία στην εικόνα που έχουν οριστεί ως δείγματα εκπαίδευσης. Δεξιά της εικόνας φαίνονται τα σημεία που πρόβλεψε το μοντέλο. Στο πρώτο δείγμα, δεν δηλώθηκε κτίριο (αριστερά) και αντίστοιχα το μοντέλο δεν πρόβλεψε κτίριο (δεξιά). Στο δεύτερο δείγμα μαζί με το κτίριο δηλώθηκε και ένα τμήμα του δρόμου (αριστερά) με αποτέλεσμα το μοντέλο να προβλέψει λανθασμένα ολόκληρο το τμήμα του δρόμου ως κτίριο (δεξιά). Στο τρίτο δείγμα, δηλώθηκε το κτίριο ως δείγμα εκπαίδευσης (αριστερά) και το μοντέλο ορθά πρόβλεψε το κτίριο (δεξιά). Σημαντικό είναι να αναφερθεί πως σε αυτό το σημείο το μοντέλο πρόβλεψε καλύτερα από το χρήστη το κτίριο καθώς αριστερά φαίνεται ένα τμήμα εκτός σκεπής του κτιρίου και δεξιά αυτό το τμήμα έχει παραληφθεί.
Στο τέταρτο δείγμα, αριστερά φαίνεται ότι δεν έχει δηλωθεί κτίριο ενώ δεξιά το μοντέλο πρόβλεψε ένα κτίριο. Στο πέμπτο και τελευταίο δείγμα φαίνεται ένα τμήμα που έχει δηλωθεί ως κτίριο (αριστερά) και το μοντέλο πρόβλεψε και τα γειτονικά κτίρια μαζί με εκείνο (δεξιά). Παρατηρούμε ότι οι προβλέψεις του μοντέλου είναι ίδιες με εκείνες του προηγούμενου μοντέλου.

6.1.3. Αποτελέσματα εκπαίδευσης τρίτου μοντέλου



Εικόνα 55: Εκπαίδευση C μοντέλου βαθιάς μάθησης.

Στην παραπάνω εικόνα, φαίνονται τα αποτελέσματα της εκπαίδευσης του τρίτου (C) μοντέλου βαθιάς μάθησης. Αριστερά της εικόνας φαίνονται τα σημεία στην εικόνα που έχουν οριστεί ως δείγματα εκπαίδευσης. Δεξιά της εικόνας φαίνονται τα σημεία που πρόβλεψε το μοντέλο. Στο πρώτο δείγμα, δεν δηλώθηκε κτίριο (αριστερά) και αντίστοιχα το μοντέλο πρόβλεψε λανθασμένα κτίριο (δεξιά). Στο δεύτερο δείγμα μαζί με το κτίριο δηλώθηκε και ένα τμήμα του δρόμου (αριστερά) με αποτέλεσμα το μοντέλο να προβλέψει λανθασμένα ολόκληρο το τμήμα του δρόμου ως κτίριο (δεξιά). Στο τρίτο δείγμα, δηλώθηκε το κτίριο ως δείγμα εκπαίδευσης (αριστερά) και το μοντέλο ορθά πρόβλεψε το κτίριο (δεξιά). Σημαντικό είναι να αναφερθεί πως σε αυτό το σημείο το μοντέλο πρόβλεψε χειρότερα από το χρήστη το κτίριο καθώς αριστερά φαίνεται ένα τμήμα εκτός σκεπής του κτιρίου και δεξιά αυτό το τμήμα έχει επεκταθεί αρκετά εκτός του κτιρίου.
Στο τέταρτο δείγμα και τελευταίο δείγμα, αριστερά φαίνεται ότι δεν έχει δηλωθεί κτίριο ενώ δεξιά το μοντέλο πρόβλεψε ένα κτίριο λανθασμένα.

6.2. Αποτελέσματα ανίχνευσης κτιρίων

6.2.1. Αποτελέσματα ανίχνευσης κτιρίων πρώτου μοντέλου

Εικόνα 56: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_1.

Στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα A_1. Όπως φαίνεται, η ανίχνευση κτιρίων δεν είχε βέλτιστα αποτελέσματα καθώς πολλά κτίρια εντοπίζονται σε περιοχές με πράσινο όπου δεν υφίστανται καθώς και σε περιοχές με έδαφος. Επίσης γύρω από τα όρια του ορθοφωτοχάρτη παρατηρούνται περισσότερα κτίρια απ' όσα υφίστανται.

Εικόνα 57: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_2.

Ακολούθως, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα A_2. Όπως φαίνεται, η ανίχνευση κτιρίων δεν είχε βέλτιστα αποτελέσματα και σε αυτό το πείραμα καθώς πολλά κτίρια εντοπίζονται σε περιοχές όπου δεν υφίστανται καθώς στα όρια της ορθοφωτογραφίας εμφανίζονται περισσότερα κτίρια.

Εικόνα 58: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_3.

Στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα A_3. Όπως φαίνεται οπτικά, η ανίχνευση κτιρίων καλύτερα αποτελέσματα συγκριτικά με τα προηγούμενα μοντέλα όμως εξακολουθούν να ανιχνεύονται κτίρια σε περιοχές που δεν υφίστανται. Στα όρια της της ορθοφωτογραφίας, εντοπίζονται πολλά λιγότερα ψευδή κτίρια αλλά εντοπίζονται κτίρια ψευδή εκτός της ορθοφωτογραφίας.

Εικόνα 59: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης A_4.

Στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα A_4. Η ανίχνευση κτιρίων δεν είχε καλά αποτελέσματα καθώς πολλά κτίρια εντοπίζονται σε περιοχές που δεν υφίστανται καθώς ενώ παρατηρείται ότι δεν υπάρχουν ψευδή κτίρια στα όρια και εκτός των ορίων της ορθοφωτογραφίας.

6.2.2. Αποτελέσματα ανίχνευσης κτιρίων δεύτερου μοντέλου

Εικόνα 60: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_1.

Ακολούθως, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα B_1. Τα αποτελέσματα σε αυτό το πείραμα προβάλλουν πολλά κτίρια σε περιοχές που υφίστανται καθώς και στα όρια της ορθοφωτογραφίας δεν υπάρχουν ψευδή κτίρια. Σε αυτό το πείραμα φαίνεται μια βελτίωση ως προς την ανίχνευση των κτιρίων.

Εικόνα 61: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_2.

Αντίστοιχα, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα B_2. Η ανίχνευση κτιρίων είχε εξίσου καλά αποτελέσματα σε αυτό το πείραμα καθώς πολλά κτίρια εντοπίζονται σε περιοχές όπου υφίστανται αλλά συγκριτικά με το προηγούμενο πείραμα σε περιοχές με πράσινο παρατηρούνται περισσότερα ψευδή κτίρια.

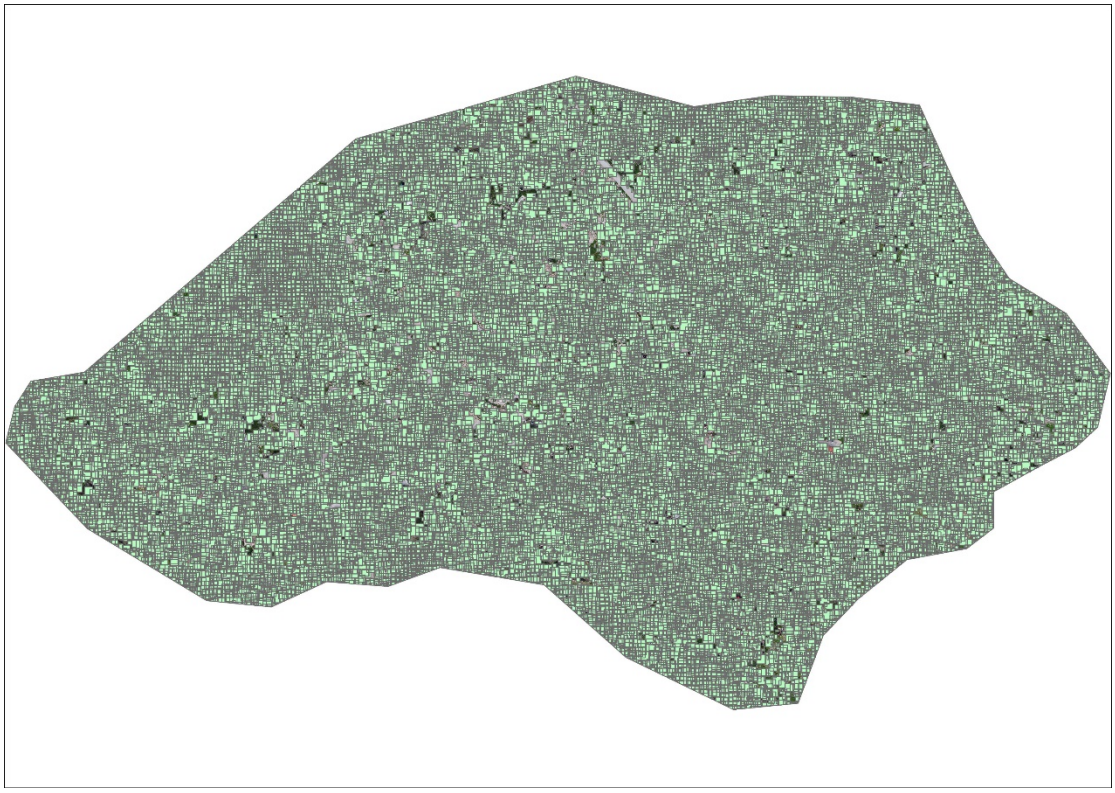
Εικόνα 62: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_3.

Στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα B_3. Σε αυτό το πείραμα ανιχνεύτηκαν λιγότερα ψευδή κτίρια ενώ σε περιοχές με πράσινο δεν παρατηρούνται καθόλου ψευδή κτίρια. Στα όρια της ορθοφωτογραφίας φαίνονται ελάχιστα ψευδή κτίρια ενώ αξίζει να σημειωθεί πως σε αυτό το πείραμα τα αποτελέσματα είναι επαρκή και λιγότερο λανθασμένα.

Εικόνα 63: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης B_4.

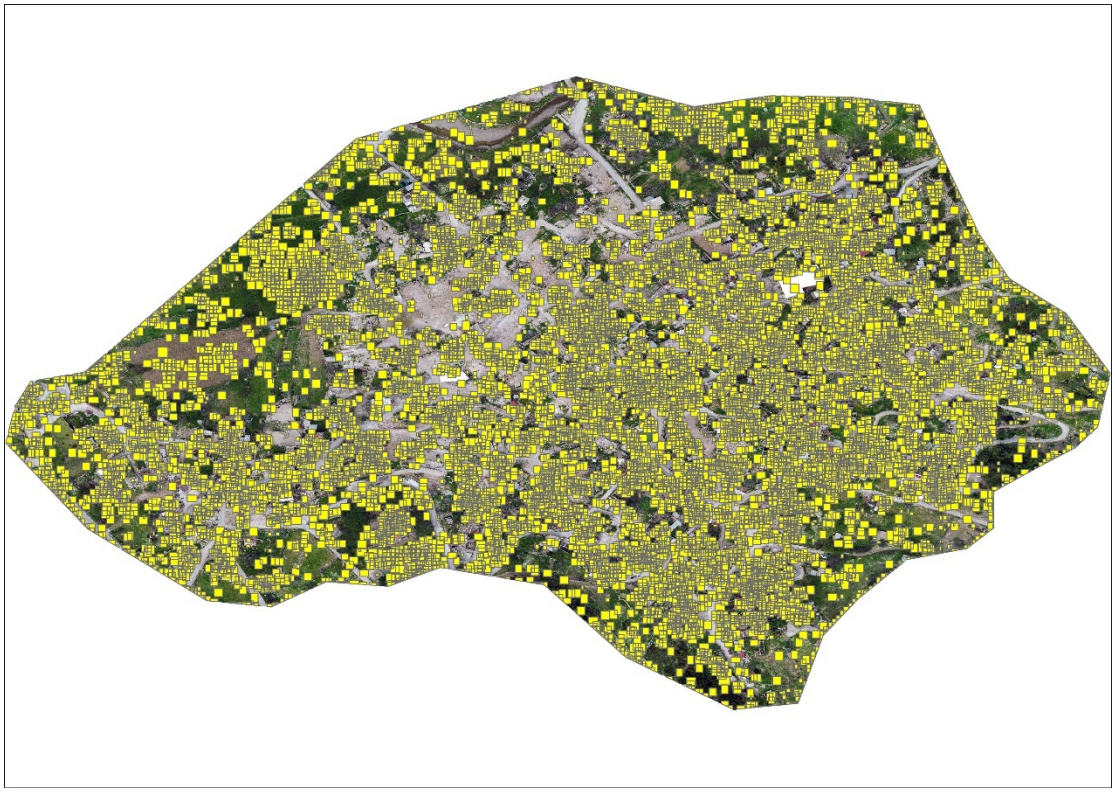
Έπειτα, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα B_4. Τα αποτελέσματα είναι ακόμα καλύτερα συγκριτικά με το προηγούμενο πείραμα. Φαίνεται να ανιχνεύτηκαν σχεδόν όλα τα κτίρια ενώ στα όρια και σε περιοχές με πράσινο δεν υπάρχει κανένα ψευδές κτίριο. Τα αποτελέσματα σε αυτό το πείραμα είναι τα καλύτερα καθώς ανιχνεύτηκαν όλα τα κτίρια και δεν υπήρχαν σχεδόν καθόλου ψευδή κτίρια.

6.2.3. Αποτελέσματα ανίχνευσης κτιρίων τρίτου μοντέλου



Εικόνα 64: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_1.

Στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα C_1. Όπως φαίνεται, η ανίχνευση κτιρίων είναι εσφαλμένη καθώς ανιχνεύονται σε ολόκληρο τον ορθοφωτοχάρτη κτίρια. Η εικόνα είναι σχεδόν όλη επικαλυπτόμενη με ψευδή κτίρια επομένως ο εντοπισμός είναι λανθασμένος.



Εικόνα 65: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_2.

Ακολούθως, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα C_2. Όπως φαίνεται οπτικά, η ανίχνευση κτιρίων είχε κακά αποτελέσματα και σε αυτό το πείραμα καθώς πολλά κτίρια εντοπίζονται σε περιοχές με πράσινο όπου δεν υφίστανται καθώς και σε περιοχές με έδαφος αλλά και στα όρια της ορθοφωτογραφίας. Τα κτίρια είναι ψευδή σχεδόν σε όλη την έκταση της ορθοφωτογραφίας.

Εικόνα 66: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_3.

Ακολούθως, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα C_3. Αντίστοιχα, η ανίχνευση κτιρίων είχε λίγο καλύτερα από το προηγούμενο πείραμα αλλά εξακολουθεί να είναι εσφαλμένη. Σε αυτό το πείραμα εντοπίζονται κτίρια σε περιοχές με πράσινο καθώς επίσης στα όρια της ορθοφωτογραφίας. Όλη η έκταση της περιοχής και σε αυτό το πείραμα είναι επικαλυπτόμενη με ψευδή κτίρια.

Εικόνα 67: Αποτέλεσμα εντοπισμού κτιρίων με χρήση πειράματος βαθιάς μάθησης C_4.

Τέλος, στην παραπάνω εικόνα φαίνονται τα αποτελέσματα ανίχνευσης κτιρίων για το πείραμα C_4. Όπως φαίνεται, η ανίχνευση κτιρίων δεν έχει έντονες διαφορές με τα αποτελέσματα του προηγούμενου πειράματος. Ψευδή κτίρια ανιχνεύτηκαν στα όρια και κυρίως εκτός ορίων της ορθοφωτογραφίας. Σε όλη την έκταση των ορίων της περιοχής ανιχνεύονται ψευδή κτίρια.

6.3. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων

6.3.1. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων πρώτου μοντέλου

A_1 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.4768**
Recall	**0.4913**
F1 Score	**0.4839**
True Positive	**113**
False Positive	**124**
False Negative	**117**

Πίνακας 8: Ακρίβεια ανίχνευσης κτιρίων πειράματος A_1.

Ο παραπάνω πίνακας, προβάλλει την αξιολόγηση του πειράματος A_1. Όπως φαίνεται, η ακρίβεια (Precision) είναι μέτρια καθώς φτάνει περίπου το 0.4. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν περίπου το 0.4. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 113 και 117. Τα ψευδώς θετικά (False Positive) είναι 124.

A_2 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.4379**
Recall	**0.4786**
F1 Score	**0.4573**
True Positive	**134**
False Positive	**172**
False Negative	**146**

Πίνακας 9: Ακρίβεια ανίχνευσης κτιρίων πειράματος A_2.

Ακολούθως, ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος A_2. Όπως φαίνεται, η ακρίβεια (Precision) φτάνει περίπου το 0.4 και σε αυτό το πείραμα. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν το 0.4 εξίσου. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 134 και 146. Τα λάθος θετικά (False Positive) είναι 172.

A_3 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.6353**
Recall	**0.6036**
F1 Score	**0.6190**
True Positive	**169**
False Positive	**97**
False Negative	**111**

Πίνακας 10: Ακρίβεια ανίχνευσης κτιρίων πειράματος A_3.

Για το πείραμα A_3, η ακρίβεια (Precision) είναι καλύτερη καθώς φτάνει περίπου το 0.6. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν περίπου το 0.6. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 169 και 111. Τα ψευδώς θετικά (False Positive) είναι 97.

A_4 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.6996**
Recall	**0.6808**
F1 Score	**0.6901**
True Positive	**177**
False Positive	**76**
False Negative	**83**

Πίνακας 11: Ακρίβεια ανίχνευσης κτιρίων πειράματος A_4.

Ακολούθως, ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος A_4. Όπως φαίνεται, η ακρίβεια (Precision) είναι εξίσου αποδοτική καθώς φτάνει περίπου το 0.6. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν περίπου το 0.6 εξίσου. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 177 και 83. Τα λάθος θετικά (False Positive) είναι 76.

6.3.2. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων δεύτερου μοντέλου

B_1 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.7331**
Recall	**0.7750**
F1 Score	**0.7535**
True Positive	**217**
False Positive	**79**
False Negative	**63**

Πίνακας 12: Ακρίβεια ανίχνευσης κτιρίων πειράματος B_1.

Η αξιολόγηση του πειράματος B_1, φαίνεται ότι έχει σημαντικές αλλαγές καθώς η ακρίβεια (Precision) φτάνει περίπου το 0.7. Η ανάκληση είναι και το F1 σκορ (F1-score) φτάνουν εξίσου το 0.7. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 217 και 63. Τα ψευδώς θετικά (False Positive) είναι 79.

B_2 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.6649**
Recall	**0.5913**
F1 Score	**0.6260**
True Positive	**123**
False Positive	**62**
False Negative	**85**

Πίνακας 13: Ακρίβεια ανίχνευσης κτιρίων πειράματος B_2.

Η αξιολόγηση του πειράματος B_2 στον παραπάνω πίνακα είναι εξίσου ικανοποιητική. Όπως φαίνεται, η ακρίβεια (Precision) είναι φτάνει περίπου το 0.6. Η ανάκληση είναι 0.5 ενώ το F1 σκορ (F1- score) 0.6. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 123 και 85. Τα λάθος θετικά (False Positive) είναι 62.

B_3 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.7793**
Recall	**0.7167**
F1 Score	**0.7483**
True Positive	**226**
False Positive	**88**
False Negative	**64**

Πίνακας 14: Ακρίβεια ανίχνευσης κτιρίων πειράματος B_3.

Για το πείραμα B_3, η ακρίβεια (Precision) είναι ακόμα καλύτερη καθώς φτάνει το 0.7. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν εξίσου το 0.7. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 226 και 64. Τα ψευδώς θετικά (False Positive) είναι 88.

B_4 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.8231**
Recall	**0.8643**
F1 Score	**0.8432**
True Positive	**242**
False Positive	**52**
False Negative	**38**

Πίνακας 15: Ακρίβεια ανίχνευσης κτιρίων πειράματος B_4.

Ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος B_4. Όπως φαίνεται, η ακρίβεια (Precision) είναι υψηλότερη από κάθε πείραμα καθώς φτάνει περίπου το 0.8. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν το 0.8 αντίστοιχα. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 242 και 38. Τα λάθος θετικά (False Positive) είναι 52.

6.3.3. Αποτελέσματα ακρίβειας ανίχνευσης κτιρίων τρίτου μοντέλου

C_1 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.0001**
Recall	**0.0108**
F1 Score	**0.0002**
True Positive	**3**
False Positive	**27397**
False Negative	**275**

Πίνακας 16: Ακρίβεια ανίχνευσης κτιρίων πειράματος C_1.

Ακολούθως, ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος C_1. Όπως φαίνεται, η ακρίβεια (Precision) είναι αρκετά χαμηλή καθώς φτάνει περίπου το 0. Η ανάκληση και το F1 σκορ (F1- score) είναι αντίστοιχα 0. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 3 και 275. Τα λάθος θετικά (False Positive) είναι 27397.

C_2 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.0016**
Recall	**0.0679**
F1 Score	**0.0031**
True Positive	**19**
False Positive	**12025**
False Negative	**261**

Πίνακας 17: Ακρίβεια ανίχνευσης κτιρίων πειράματος C_2.

Ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος C_2. Όπως φαίνεται, η ακρίβεια (Precision) είναι χαμηλή καθώς φτάνει περίπου το 0. Η ανάκληση είναι 0 ενώ το F1 σκορ (F1- score) 0. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 19 και 261. Τα ψευδώς θετικά (False Positive) είναι 12025.

C_3 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.1429**
Recall	**0.1857**
F1 Score	**0.1615**
True Positive	**52**
False Positive	**312**
False Negative	**228**

Πίνακας 18: Ακρίβεια ανίχνευσης κτιρίων μοντέλου C_3.

Ακολούθως, ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος C_3. Όπως φαίνεται, η ακρίβεια (Precision) είναι μη αποδοτική καθώς φτάνει περίπου το 0.1. Η ανάκληση όπως και το F1 σκορ (F1- score) φτάνουν και εκείνα το 0.1. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 52 και 228. Τα ψευδώς θετικά (False Positive) είναι 312.

C_4 Experiment	
Building	**Min_IoU_is_50_percent**
Precision	**0.2525**
Recall	**0.2714**
F1 Score	**0.2616**
True Positive	**76**
False Positive	**225**
False Negative	**204**

Πίνακας 19: Ακρίβεια ανίχνευσης κτιρίων μοντέλου C_4.

Ακολούθως, ο παραπάνω πίνακας προβάλλει την αξιολόγηση του πειράματος C_4. Όπως φαίνεται, η ακρίβεια (Precision) φτάνει περίπου το 0.2. Η ανάκληση και το F1 σκορ (F1- score) φτάνουν εξίσου το 0.2. Τα αληθώς θετικά (True Positive) και τα ψευδώς αρνητικά (False Negative) είναι αντίστοιχα 76 και 204. Τα ψευδώς θετικά (False Positive) είναι 225.

7. ΣΥΖΗΤΗΣΗ – ΣΥΜΠΕΡΑΣΜΑΤΑ

Ο παρακάτω πίνακας, προβάλει συνολικά τα αποτελέσματα των πειραμάτων που πραγματοποιήθηκαν με τα τρία μοντέλα βαθιάς μάθησης.

Min_IoU_is_50_percent	Precision	Recall	F1 Score	True Positive	False Positive	False Negative
A_1 Experiment	**0.4768**	**0.4913**	**0.4839**	**113**	**124**	**117**
A_2 Experiment	**0.4379**	**0.4786**	**0.4573**	**134**	**172**	**146**
A_3 Experiment	**0.6353**	**0.6036**	**0.6190**	**169**	**97**	**111**
A_4 Experiment	**0.6996**	**0.6808**	**0.6901**	**177**	**76**	**83**
B_1 Experiment	**0.7331**	**0.7750**	**0.7535**	**217**	**79**	**63**
B_2 Experiment	**0.6649**	**0.5913**	**0.6260**	**123**	**62**	**85**
B_3 Experiment	**0.7793**	**0.7167**	**0.7483**	**226**	**88**	**64**
B_4 Experiment	**0.8231**	**0.8643**	**0.8432**	**242**	**52**	**38**
C_1 Experiment	**0.0001**	**0.0108**	**0.0002**	**3**	**27397**	**275**
C_2 Experiment	**0.0016**	**0.0679**	**0.0031**	**19**	**12025**	**261**
C_3 Experiment	**0.1429**	**0.1857**	**0.1615**	**52**	**312**	**228**
C_4 Experiment	**0.2525**	**0.2714**	**0.2616**	**76**	**225**	**204**

Πίνακας 20: Συνολικά αποτελέσματα εκπαίδευσης.

Τα πειράματα A_1 έως A_4 έχουν μέτρια αποτελέσματα ανίχνευσης των κτιρίων αφού τα πειράματα A_1 και A_2 λαμβάνουν ακρίβεια (Precision) και ανάκληση (Recall) μικρότερη του 50%. Αυτό σημαίνει πως στο A_1 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν 113 κτίρια και είναι σωστό ενώ πρόβλεψε πως 124 δεν αποτελούν κτίρια είναι λάθος. Επιπλέον πρόβλεψε πως για 117 δεν υπάρχει κτίριο και είναι λάθος. Αντίστοιχα για το A_2 πείραμα το μοντέλο πρόβλεψε πως υπάρχουν 134 κτίρια και είναι σωστό ενώ πρόβλεψε πως 172 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 146 δεν υπάρχει κτίριο και είναι λάθος.

Τα πειράματα A_3 και A_4, έχουν καλύτερα αποτελέσματα καθώς η ακρίβεια (Precision) και η ανάκληση (Recall) είναι πάνω από 50%.

Πιο συγκεκριμένα, στο A_3 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν 169 κτίρια και είναι σωστό ενώ πρόβλεψε πως 97 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 111 δεν υπάρχει κτίριο και είναι λάθος. Αντίστοιχα για το A_4 πείραμα το μοντέλο πρόβλεψε πως υπάρχουν 177 κτίρια και είναι σωστό ενώ πρόβλεψε πως 76 δεν αποτελούν κτίρια είναι λάθος. Επιπλέον πρόβλεψε πως για 83 δεν υπάρχει κτίριο και είναι λάθος.

Τα πειράματα B_1 έως B_4 φαίνεται να είχαν τα καλύτερα αποτελέσματα καθώς η ακρίβεια (Precision) και η ανάκληση (Recall) τους είναι μεγαλύτερα από 60%.

Στο B_1 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν 217 κτίρια και είναι σωστό ενώ πρόβλεψε πως 79 δεν αποτελούν κτίρια είναι λάθος. Επιπλέον πρόβλεψε πως για 63 δεν υπάρχει κτίριο και είναι λάθος. Αντίστοιχα για το B_2 πείραμα το μοντέλο πρόβλεψε πως υπάρχουν 123 κτίρια και είναι σωστό ενώ πρόβλεψε πως 62 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 85 δεν υπάρχει κτίριο και είναι λάθος. Στο B_3 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν 226 κτίρια και είναι σωστό ενώ πρόβλεψε πως 88 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 64 δεν υπάρχει κτίριο και είναι λάθος.

Όπως φαίνεται, το πιο βέλτιστο αποτέλεσμα είναι το **B_4**, καθώς έχει ακρίβεια και ανάκληση που λαμβάνει περίπου 80%. Από το σύνολο των 280 δειγμάτων, φαίνεται ότι τα 242 είναι θετικά αληθή ενώ τα 38 είναι αρνητικά ψευδή. Αυτό σημαίνει ότι το μοντέλο από τα 280 πρόβλεψε πως υπάρχουν 242 κτίρια και είναι σωστό ενώ πρόβλεψε πως 52 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 38 δεν υπάρχει κτίριο και είναι λάθος.

Τα πειράματα C_1 έως C_4 αποδείχτηκαν αποτυχημένα καθώς τα ποσοστά είναι μηδενικά με τα μοντέλα C3 και C4 να λαμβάνουν περίπου 1% ποσοστό ακρίβειας και ανάκλησης.

Στο C_1 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν μόνο 3 κτίρια και είναι σωστό ενώ πρόβλεψε πως 27397 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 275 δεν υπάρχει κτίριο και είναι λάθος. Αντίστοιχα Στο C_2 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν μόνο 19 κτίρια και είναι σωστό ενώ πρόβλεψε πως 12025 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 261 δεν υπάρχει κτίριο και είναι λάθος. Στο C_3 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν μόνο 52 κτίρια και είναι σωστό ενώ πρόβλεψε πως 312 δεν αποτελούν κτίρια και είναι λάθος. Επιπλέον πρόβλεψε πως για 228 δεν υπάρχει κτίριο και είναι λάθος.

Στο C_4 πείραμα, από τα 280 κτίρια, το μοντέλο πρόβλεψε πως υπάρχουν μόνο 76 κτίρια και είναι σωστό ενώ πρόβλεψε πως 225 δεν αποτελούν κτίρια είναι λάθος. Επιπλέον πρόβλεψε πως για 204 δεν υπάρχει κτίριο και είναι λάθος.

Εν κατακλείδι, στη παρούσα εργασία μελετήθηκαν μοντέλα βαθιάς μάθησης με στόχο τον εντοπισμό κτιρίων. Σε όλα τα μοντέλα χρησιμοποιήθηκε η τεχνική της μεταφοράς γνώσης, όπου επιλέχθηκε προεκπαιδευμένο βαθύ νευρωνικό δίκτυο ResNet-34.

Συνολικά, συλλέχθηκαν 280 δείγματα εκπαίδευσης, εκπαιδεύτηκαν τρία μοντέλα βαθιάς μάθησης ενώ πραγματοποιήθηκαν δώδεκα πειράματα ανίχνευσης κτιρίων.

Από τις τρεις εκπαιδεύσεις μοντέλων βαθιάς μάθησης που πραγματοποιήθηκαν τα πειράματα του πρώτου μοντέλου (A1 έως A_4) αποδείχτηκαν ικανοποιητικά σε μέτριο βαθμό όσον αφορά την ακρίβεια αξιολόγησης και με βάση τη φωτοερμηνεία. Τα πειράματα του δευτέρου μοντέλου (B_1 έως B_4) ήταν εκείνα με τη μεγαλύτερη ακρίβεια ενώ τα πειράματα του τρίτου μοντέλου (C_1 έως C_4) αποδείχθηκαν ανακριβή και αναξιόπιστα.

Συγκρίνοντας τα μοντέλα μεταξύ τους είναι σημαντικό να αναφερθεί πως το μοντέλο με το μικρότερο μέγεθος παρτίδας (Μοντέλο C με batch size 4) είναι εκείνο με τη μικρότερη αξιολόγηση πράγμα που σημαίνει ότι η απόδοση του είναι σχεδόν ανεπαρκής.

Αντίστοιχα το μοντέλο με το μεγαλύτερο μέγεθος παρτίδας (Μοντέλο B με batch size 16) είναι εκείνο με τη μεγαλύτερη αξιολόγηση επομένως η απόδοση του μοντέλου είναι αρκετά υψηλή.

Επιπλέον, το μοντέλο με μεσαίο μέγεθος παρτίδας (Μοντέλο A με batch size 8) είχε καλή αξιολόγηση με μια μέτρια απόδοση.

Αξιοσημείωτο είναι ότι διπλασιάζοντας το μέγεθος της παρτίδας το μοντέλο αποδίδει καλύτερα. Το μέγεθος της παρτίδας εξετάζει την ακρίβεια της εκτίμησης της διαβάθμισης σφάλματος κατά την εκπαίδευση του νευρωνικού δικτύου.

Οι παράμετροι που συνέβαλαν στα πειράματα των μοντέλων δεν παρουσιάζουν έντονες μεταβολές στα αποτελέσματα.

Συμπερασματικά η ανίχνευση αντικειμένων με βαθιά μάθηση παρέχει στο χρήστη ένα γρήγορο και ακριβές μέσο για την πρόβλεψη της θέσης ενός αντικειμένου σε μια εικόνα. Η τεχνική μηχανικής εκμάθησης, μαθαίνει αυτόματα τα χαρακτηριστικά της εικόνας με αποτέλεσμα τον εντοπισμό του αντικειμένου.

Η μέθοδος που χρησιμοποιείται στο λογισμικό ArcGIS Pro, δίνει τη δυνατότητα στον χρήστη να πειραματιστεί χωρίς πολύπλοκες απαιτήσεις έχοντας ως αποτέλεσμα την εύχρηστη εφαρμογή βαθιάς μάθησης. Η διαδικασία για την επίτευξη των στόχων ανίχνευσης αντικειμένων, λαμβάνει θέση με ένα σταθερό μοτίβο έως ότου ο χρήστης να φτάσει στο επιθυμητό αποτέλεσμα.

Σε μελλοντική εργασία τα αποτελέσματα ανίχνευσης κτιρίων θα μπορούσαν να λάβουν ένα μεγαλύτερο εύρος αξιολόγησης χρησιμοποιώντας και συγκρίνοντας διαφορετικά μοντέλα ανίχνευσης (πχ. RetinaNet, YOLOv5 κλπ.). Διατηρώντας το μοντέλο Single Shot Detector η συλλογή μεγαλύτερου μεγέθους δειγμάτων εκπαίδευσης ενδεχομένως θα οδηγούσε σε πιο αποδοτικά μοντέλα. Επιπλέον, θα ήταν εφικτό να εφαρμοστούν διαφορετικά πειράματα και να συμβάλουν περισσότεροι παράμετροι στη βελτίωση των αποτελεσμάτων εκτός από το κατώφλι εμπιστοσύνης (threshold), το μέγεθος παρτίδας (batch size), τη περιοχή συνόρων (padding) και την επικάλυψη μη μέγιστης καταστολής (nms_overlap).

8. ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνόγλωσση

Ατζαράκη, Α. 2019. *3D Χαρτογράφηση μεταβολών οικισμού Βρίσας με μεθόδους γεωπληροφορικής*. Πτυχιακή εργασία. Μυτιλήνη: Πανεπιστήμιο Αιγαίου.

Ευρωπαϊκό Κοινοβούλιο, 2020. *Τι είναι η τεχνητή νοημοσύνη και πώς χρησιμοποιείται;* [online] Διαθέσιμο από: *https://www.europarl.europa.eu/news/el/headlines/society/20200827STO85804/ti-einai-i-techniti-noimosuni-kai-pos-chrisimopoieitai*

[Ημερομηνία ανάκτησης: 20 Ιουνίου 2022]

Θεοδωρόπουλος, Χ., 2018. *Χρήση τεχνικών βαθιάς μάθησης για την αναγνώριση συναισθημάτων μέσα από εκφράσεις του προσώπου.* Διπλωματική εργασία. [Online]. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο. Διαθέσιμο από: http://artemis.cslab.ece.ntua.gr:8080/jspui/bitstream/123456789/17103/1/Thesis_Christos_Theodoropoulos.pdf

[Ημερομηνία ανάκτησης: 20 Φεβρουαρίου 2022]

Χρονοπούλου, Α., 2019. *Τεχνικές Μεταφοράς Μάθησης σε Βαθιά Νευρωνικά Δίκτυα για Ανάλυση Συναισθήματος και Σημασιολογική Μοντελοποίηση.* Διπλωματική εργασία. [Online]. Αθήνα: Εθνικό Μετσόβιο Πολυτεχνείο. Διαθέσιμο από: http://artemis.cslab.ece.ntua.gr:8080/jspui/handle/123456789/17239

[Ημερομηνία ανάκτησης: 20 Ιουλίου 2022]

Ξενόγλωσση

Alouta, R., Hess K., 2021. *Deep Learning with ArcGIS Pro Tips & Tricks: Part 2*. [Online]. Διαθέσιμο από: https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/deep-learning-with-arcgis-pro-tips-tricks-part-2/

[Ημερομηνία ανάκτησης: 20 Φεβρουαρίου 2022]

ArcGIS Developer, χ.χ. *Train Deep Learning Model*. [Online]. Διαθέσιμο από: https://developers.arcgis.com/rest/services-reference/enterprise/train-deep-learning-model.htm

[Ημερομηνία ανάκτησης: 20 Ιουνίου 2022]

ArcGIS Developers, χ.χ. How single-shot detector (SSD) works? [Online]. Διαθέσιμο από: https://developers.arcgis.com/python/guide/how-ssd-works/

[Ημερομηνία ανάκτησης: 9 Απριλίου 2022]

Bellman, R., E., 1978. *An introduction to Artificial Intelligence: Can Computers Think?* Boyd & Fraser Publishing Company, San Francisco. [Online]. Διαθέσιμο από: https://www.researchgate.net/publication/256121747_Artificial_intelligence_The_very_idea_J_Haugeland_MIT_Press_Cambridge_MA_1985_287_pp

[Ημερομηνία ανάκτησης: 3 Φεβρουαρίου 2022]

Burns, E., Brush, K., 2021. Deep learning. [Online]. Διαθέσιμο από: https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network

[Ημερομηνία ανάκτησης: 20 Φεβρουαρίου 2022]

Boesch, G., χ.χ., *Object Detection in 2022: The Definitive Guide*. [Online]. Διαθέσιμο από:
https://viso.ai/deep-learning/object-detection/

[Ημερομηνία ανάκτησης: 6 Ιούνιου 2022]

Brownlee, J., 2017. *A Gentle Introduction to Transfer Learning for Deep Learning.* [Online]. Διαθέσιμο από:
https://machinelearningmastery.com/transfer-learning-for-deep-learning/

[Ημερομηνία ανάκτησης: 20 Ιουνίου 2022]

Brownlee, J., 2022. *Difference Between a Batch and an Epoch in a Neural Network* [Online]. Διαθέσιμο από:
https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/

[Ημερομηνία ανάκτησης: 20 Ιουνίου 2022]

Charniak, E. και McDermott, D., 1985. *Introduction to Artificial Intelligence*. Addison-Wesley, Reading, Massachusetts. . [Online]. Διαθέσιμο από:
https://link.springer.com/chapter/10.1007/978-3-7091-2784-1_2

[Ημερομηνία ανάκτησης: 3 Φεβρουαρίου 2022]

Chollet F., και Allaire J.J., 2018. *Deep Learning with R.* [Online]. Διαθέσιμο από:

https://livebook.manning.com/book/deep-learning-with-r/chapter-1/14

[Ημερομηνία ανάκτησης: 9 Μαρτίου 2022]

Chollet F., 2017. *Deep Learning with Python, Second Edition.* [Online]. Διαθέσιμο από:
https://www.manning.com/books/deep-learning-with-python

[Ημερομηνία ανάκτησης: 7 Φεβρουαρίου 2022]

ESRI, χ.χ. *Non Maximum Suppression (Image Analyst).* [Online]. Διαθέσιμο από:
https://pro.arcgis.com/en/pro-app/2.8/tool-reference/image-analyst/non-maximum-suppression.htm

[Ημερομηνία ανάκτησης: 20 Αυγούστου 2022]

ESRI, χ.χ. *Compute Accuracy For Object Detection (Image Analyst)*. [Online]. Διαθέσιμο από:

https://pro.arcgis.com/en/pro-app/2.9/tool-reference/image-analyst/compute-accuracy-for-object-detection.htm

[Ημερομηνία ανάκτησης: 7 Σεπτεμβρίου 2022]

ESRI, χ.χ. *How the Compute Accuracy For Object Detection tool works*. [Online]. Διαθέσιμο από:

https://pro.arcgis.com/en/pro-app/2.9/tool-reference/image-analyst/how-compute-accuracy-for-object-detection-works.htm

[Ημερομηνία ανάκτησης: 7 Σεπτεμβρίου 2022]

Goodfellow. I., Bengio, Y., Courville, A., 2016. *Deep Learning (Adaptive Computation and Machine Learning series).* [Online]. Διαθέσιμο από:
https://www.academia.edu/38223830/Adaptive_Computation_and_Machine_Learning_series_Deep_learning_The_MIT_Press_2016_pdf

[Ημερομηνία ανάκτησης: 16 Μαρτίου 2022]

He, K., Zhang, X., Ren, S. και Sun, J., 2015. *Deep Residual Learning for Image Recognition.* [Online]. Διαθέσιμο από:
https://arxiv.org/pdf/1512.03385.pdf

[Ημερομηνία ανάκτησης: 12 Φεβρουαρίου 2022]

He, K., Zhang, X., Ren, S. και Sun, J., 2015. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. [Online]. Διαθέσιμο από: https://arxiv.org/pdf/1406.4729.pdf)%C3%AC%20%CB%9C

[Ημερομηνία ανάκτησης: 9 Απριλίου 2022]

Haugeland, J., 1985. *Artificial Intelligence: The Very Idea*. MIT Press, Cambridge, Massachusetts. [Online]. Διαθέσιμο από: https://www.researchgate.net/publication/256121747_Artificial_intelligence_The_very_idea_J_Haugeland_MIT_Press_Cambridge_MA_1985_287_pp

[Ημερομηνία ανάκτησης: 3 Φεβρουαρίου 2022]

IBM Cloud, 2020. *Deep Learning.* [Online]. Διαθέσιμο από:

https://www.ibm.com/cloud/learn/deep-learning

[Ημερομηνία ανάκτησης: 9 Μαρτίου 2022]

Knight, K. 1999. *A statical mt tutorial workbook*. Prepared in connection with the Johns Hopkins University summer workshop. [Online]. Διαθέσιμο από: https://kevincrawfordknight.github.io/papers/wkbk.pdf

[Ημερομηνία ανάκτησης: 19 Φεβρουαρίου 2022]

Kurzweil, R., 1990. *The Age of Intelligent Machines*. MIT Press, Cambridge, Massachusetts. [Online]. Διαθέσιμο από: https://sicm.mitpress.mit.edu/books/age-intelligent-machines

[Ημερομηνία ανάκτησης: 19 Φεβρουαρίου 2022]

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu., C., 2016. *SSD: Single Shot MultiBox Detector.* [Online]. Διαθέσιμο από: https://arxiv.org/abs/1512.02325

[Ημερομηνία ανάκτησης: 9 Απριλίου 2022]

Mor-Yosef, S., Samueloff, A., Modan, B., Navot, D., και Schenker, J., G., 1990. *Ranking the risk factors for cesarean: logistic regression analysis of a nationwide study.* Obstet Gynecol.

Nilsson, N. J. 1998. *Artificial Intelligence : A New Synthesis*. Morgan Kaufmann, San Mateo, California. [Online]. Διαθέσιμο από: https://core.ac.uk/download/pdf/82021614.pdf

[Ημερομηνία ανάκτησης: 7 Φεβρουαρίου 2022]
Ouyang, W., Wang, X., Zeng, X., Qiu, S., Luo, P., Tian, Y., Li, H., Yang, S., Wang, Z., Loy, C., 2015. *DeepID-Net: Deformable Deep Convolutional Neural Networks for Object Detection.* . [Online]. Διαθέσιμο από: https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Ouyang_DeepID-Net_Deformable_Deep_2015_CVPR_paper.pdf

[Ημερομηνία ανάκτησης: 9 Απριλίου 2022]

Ovid and Martin, C., 2004. *Metamorphoses*. W.W. Norton.

Poole, D., Mackworth, A. K., και Goebel, R. 1998. *Computational intelligence: A logical approach*. Oxford University Press, Oxford, UK. [Online]. Διαθέσιμο από: https://www.cs.ubc.ca/~poole/ci/front.pdf

[Ημερομηνία ανάκτησης: 7 Φεβρουαρίου 2022]

Pathaka A., R., Pandey, M. και Rautaray S., 2018. *Application of Deep Learning for Object Detection.* [Online]. Διαθέσιμο από:

https://www.semanticscholar.org/paper/Application-of-Deep-Learning-for-Object-Detection-Pathak-Pandey/9a51e8cf0bd510c286bebe70c4f1a57ac8085af1

[Ημερομηνία ανάκτησης: 9 Μαρτίου 2022]

Redmon, J., Divvala, S., Girshick, R. και Farhadi, A., 2015. *You Only Look Once: Unified, Real-Time Object Detection*. [Online]. Διαθέσιμο από: https://arxiv.org/abs/1506.02640

[Ημερομηνία ανάκτησης: 9 Απριλίου 2022]

Sparkes, B., 1996. *The Red and the Black: Studies in Greek Pottery*. Routledge.

Singhal, G., 2020. *Transfer Learning with ResNet in PyTorch.* [Online]. Διαθέσιμο από: https://www.pluralsight.com/guides/introduction-to-resnet

[Ημερομηνία ανάκτησης: 4 Ιουνίου 2022]

Tandy, D., W., 1997. *Works and Days: A Translation and Commentary for the Social Sciences*. University of California Press.

Winston, P., H., 1992. *Artificial Intelligence (Third (Ed.))*. Addison-Wesley, Reading, Massachusetts. [Online]. Διαθέσιμο από: https://courses.csail.mit.edu/6.034f/ai3/rest.pdf

[Ημερομηνία ανάκτησης: 19 Φεβρουαρίου 2022]

Walbridge, S., Clee, B., Guillochon, J., Kumar, S., Lenhardt, J., 2021. *Deep Learning Libraries Installers for ArcGIS.* [Online]. Διαθέσιμο από: https://github.com/Esri/deep-learning-frameworks/blob/master/README.md?rmedium=links_esri_com_b_d&rsource=https%3A%2F%2Flinks.esri.com%2Fdeep-learning-framework-install

[Ημερομηνία ανάκτησης: 20 Φεβρουαρίου 2022]

Yang, F., Wongun, C., and Yuanqing, L., 2016. *Exploit All the Layers: Fast and Accurate Cnn Object Detector with Scale Dependent Pooling and Cascaded Rejection Classifiers*. [Online]. Διαθέσιμο από: https://openaccess.thecvf.com/content_cvpr_2016/html/Yang_Exploit_All_the_CVPR_2016_paper.html

[Ημερομηνία ανάκτησης: 6 Ιούνιου 2022]