



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ
ΜΗΧΑΝΙΚΩΝ- ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

**«Τεχνικές Βαθιάς Μάθησης για την χαρτογράφηση
των αλλαγών χρήσεων γης με χρήση
τηλεπισκοπικών απεικονίσεων υψηλής ανάλυσης»**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ελισσάβητ Παναγοπούλου

Επιβλέπουσα καθηγήτρια: Βασιλεία Καραθανάση, Καθηγήτρια ΕΜΠ

Αθήνα, Ιούνιος 2024



**NATIONAL TECHNICAL UNIVERSITY OF
ATHENS**

**SCHOOL OF RURAL, SURVEYING AND
GEOINFORMATICS ENGINEERING**

LABORATORY OF REMOTE SENSING

**«Deep Learning Techniques for mapping land use
changes using high resolution remote sensing
imagery»**

DIPLOMA THESIS

Elissavet Panagopoulou

Supervising Professor: Vassilia Karathanassi, Professor NTUA

Athens, June 2024



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ
ΜΗΧΑΝΙΚΩΝ- ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΕΡΓΑΣΤΗΡΙΟ ΤΗΛΕΠΙΣΚΟΠΗΣΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ελισσάβητ Παναγοπούλου

Επιβλέπουσα καθηγήτρια: Βασιλεία Καραθανάση, Καθηγήτρια ΕΜΠ

Εγκρίθηκε από την τριμελή επιτροπή τον Ιούνιο του 2024

.....
Καραθανάση Βασιλεία

Καθηγήτρια ΕΜΠ

.....
Παπουτσής Ιωάννης

Επίκουρος Καθηγητής ΕΜΠ

.....
Δουλάμης Αναστάσιος

Επίκουρος Καθηγητής ΕΜΠ

Αθήνα, Ιούνιος 2024

Ελισσάβητ Παναγοπούλου

Αγρονόμος, Τοπογράφος Μηχανικός- Μηχανικός Γεωπληροφορικής, Ε.Μ.Π.

Copyright © Ελισσάβητ Παναγοπούλου, 2024.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς την συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν την συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

ΕΥΧΑΡΙΣΤΙΕΣ

Αρχικά, θα ήθελα να ευχαριστήσω την επιβλέπουσα καθηγήτρια της διπλωματικής μου εργασίας, κυρία Καραθανάση Βασιλεία, για την καθοδήγηση που μου προσέφερε.

Το μεγαλύτερο ευχαριστώ το οφείλω στην οικογένειά μου, η οποία είναι ότι πολυτιμότερο έχω σε αυτή τη ζωή. Αρχικά, να ευχαριστήσω τους γονείς μου, οι οποίοι στάθηκαν στο πλάι μου και πίστεψαν σε εμένα όταν ακόμη και εγώ με υποτίμησα. Χάρη σε αυτούς πιστεύω στην ανιδιοτελή αγάπη, την προσωπική θυσία και είμαι σίγουρη πως πάντα θα μπορώ να στηρίζομαι σε αυτούς ότι κι αν συμβεί, μιας και μου το έχουν αποδείξει. Το ευχαριστώ είναι λίγο για τη μητέρα μου Μαρία, η οποία αποτελεί πρότυπο γυναίκας για εμένα, καθώς με τον δυνατό και ακέραιο χαρακτήρα της, ξεπερνά τις δυσκολίες, και είναι πάντα δίπλα στην οικογένειά μας. Όντας η καλύτερή μου φίλη με γνωρίζει και με νιώθει καλύτερα από τον καθένα δημιουργώντας μου απόλυτη εμπιστοσύνη και ασφάλεια. Ελπίζω να γνωρίζει πόσο την αγαπώ και πόσο περήφανη είμαι για εκείνη. Επίσης, είμαι πολύ ευγνώμων και οφείλω ένα τεράστιο ευχαριστώ στον πατέρα μου Τάκη, ο οποίος πάντα πίστευε σε μένα πιο πολύ από όλους. Με έναν μαγικό τρόπο καταφέρνει να με καθησυχάζει και με ενθαρρύνει να κυνηγάω τα όνειρά μου με κάθε κόστος. Αποτελεί για εμένα πηγή έμπνευσης, μιας και σε όλη του την πορεία, αποδεικνύει τη δυναμικότητα του χαρακτήρα του και το θάρρος απέναντι στη ζωή. Πέρα από εξαιρετικός καθηγητής, είναι για εμένα πάνω από όλα πατέρας, φίλος, καλός συνομιλητής και πρότυπο ανθρώπου που αντιμετωπίζει τη ζωή με γέλιο. Έπειτα, να ευχαριστήσω τον αδερφό μου Νάσο, το αποκούμπι μου σε αυτή τη ζωή, ο οποίος αν και μικρότερος αντιμετωπίζει τις δύσκολες καταστάσεις με περισσότερη ψυχραιμία από ότι εγώ και είναι πάντα δίπλα μου με τα αστεία του και τον αισιόδοξο τρόπο σκέψης του να με ενθαρρύνει και να με στηρίζει. Του εύχομαι ολόψυχα τα καλύτερα σε όλους τους τομείς γιατί τα αξίζει. Θα είμαι δίπλα σου για πάντα και το εννοώ. Δυστυχώς, τα λόγια είναι πολύ φτωχά για να περιγράψουν τη σχέση μου και τα συναισθήματά μου για αυτούς τους ανθρώπους.

Ακόμη, να ευχαριστήσω όλους μου τους φίλους και τις φίλες, με τους οποίους δημιουργώ υπέροχες αναμνήσεις καθημερινά και τους συμφοιτητές και τις συμφοιτήτριές μου, μελλοντικούς συναδέλφους, με τους οποίους πέρασα πολύ ευχάριστα φοιτητικά χρόνια και ανυπομονώ να τους δω να εξελίσσονται επαγγελματικά.

Στη συνέχεια, θα ήθελα να ευχαριστήσω όλους τους Καθηγητές της Σχολής για όλες τις γνώσεις και τις συμβουλές που απλόχερα μου παρείχαν, συμβάλλοντας ο καθένας με τον τρόπο του στο να αγαπήσω την επιστήμη του Αγρονόμου- Τοπογράφου Μηχανικού και Μηχανικού Γεωπληροφορικής. Τέλος, ευχαριστώ όλο το ίδρυμα του Εθνικού Μετσόβιου Πολυτεχνείου για τις δυνατότητες εξέλιξης που παρέχει με το υψηλό επίπεδο σπουδών που προσφέρει.

Παναγοπούλου Ελισσάβητ

Αθήνα, Ιούνιος 2024

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια, παρατηρείται μεγάλη ροή δορυφορικών δεδομένων Παρατήρησης Γης, η οποία μπορεί να συμβάλει σημαντικά σε προβλήματα, τα οποία αφορούν σε διάφορους επιστημονικούς κλάδους των Μηχανικών. Ταυτόχρονα, η συνεχής τεχνολογική εξέλιξη στα πεδία της Τηλεπισκόπησης (remote sensing), της Φωτογραμμετρίας (photogrammetry), της Γεωπληροφορικής (geoinformatics), της Βαθιάς Μάθησης (Deep Learning) και της επεξεργασίας εικόνας (image processing) παρέχει νέα εργαλεία και αυτοματοποιημένες λύσεις για αξιοποίηση των δορυφορικών δεδομένων, σε μελέτες πολεοδομίας, χωροταξίας και κτηματολογίου.

Ένα αρκετά δύσκολο πρόβλημα προς μελέτη είναι η ανίχνευση μεταβολών. Τα τηλεπισκοπικά δεδομένα αποτελούν μία πηγή δεδομένων για τη παροχή πληροφοριών σχετικά με τα χαρακτηριστικά κάλυψης αστικής γης και τις μεταβολές τους σε διάφορες χωρικές και χρονικές κλίμακες. Επίσης, καινοτόμες τεχνικές προσφέρουν τη δυνατότητα να βελτιωθεί σημαντικά η διαθεσιμότητα των χωρικών δεδομένων, η ικανότητα για ανάλυση, η κατανόηση και η μοντελοποίηση των αστικών περιοχών.

Επομένως, στη συγκεκριμένη Διπλωματική Εργασία, πραγματοποιήθηκε έρευνα σχετικά με την ανίχνευση μεταβολών αστικού ιστού αξιοποιώντας τεχνικές βαθιάς μάθησης. Συγκεκριμένα, εφαρμόστηκαν τα νευρωνικά δίκτυα: SAM και STANet, αφού προηγουμένως έγινε η συλλογή και η κατάλληλη επεξεργασία των δεδομένων. Τέλος, ακολούθησε αξιολόγηση των αποτελεσμάτων.

Λέξεις κλειδιά: Βαθιά μάθηση, νευρωνικά δίκτυα, ανίχνευση μεταβολών αστικού ιστού, STANet, SAM, αξιολόγηση

ABSTRACT

In recent years, there has been a large influx of Earth Observation satellite data, which can make a significant contribution to problems related to various engineering disciplines. Simultaneously, the continuous technological development in the fields of remote sensing, photogrammetry, geoinformatics, deep learning and image processing provides new tools and automated solutions for the use of satellite data in urban planning, land use planning and cadastral studies.

A rather difficult problem to study is the detection of changes. Remote sensing data is a source of data to provide information on urban land cover characteristics and their changes at different spatial and temporal scales. Innovative techniques also offer the potential to significantly improve the availability of spatial data, the ability to analyse, understand and model urban areas.

Therefore, in this thesis, research on the detection of urban fabric changes using deep learning techniques was carried out. Specifically, three neural networks were applied: SAM and STANet after having collected and properly processed the data. Finally, an evaluation of the results followed.

Key words: Deep learning, neural networks, urban tissue change detection, STANet, SAM, evaluation

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ.....	12
1.1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ.....	12
1.2 ΣΚΟΠΟΣ ΚΑΙ ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΓΑΣΙΑΣ.....	12
1.3 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ.....	12
ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ.....	14
2.1 ΑΝΙΧΝΕΥΣΗ ΜΕΤΑΒΟΛΩΝ ΑΣΤΙΚΟΥ ΙΣΤΟΥ.....	14
2.2 ΒΑΘΙΑ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (DEEP LEARNING).....	15
2.3 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ (ARTIFICIAL INTELLIGENCE).....	15
2.4 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING).....	15
2.5 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	16
2.6 ΜΗΧΑΝΙΣΜΟΙ ΠΡΟΣΟΧΗΣ-ΜΕΤΑΣΧΗΜΑΤΙΣΤΕΣ.....	22
2.7 FINETUNING.....	22
2.8 ΕΚΦΡΑΣΗ ΑΚΡΙΒΕΙΑΣ ΚΑΙ ΜΕΤΡΙΚΕΣ.....	23
2.9 ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ STANet (Spatial Temporal Attention neural network) ..	25
2.10 ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ SAM (Segment Anything Model).....	28
ΚΕΦΑΛΑΙΟ 3: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΦΟΡΑ.....	30
3.1 ΣΧΕΤΙΚΑ ΕΡΕΥΝΗΤΙΚΑ ΑΡΘΡΑ.....	30
ΚΕΦΑΛΑΙΟ 4: ΜΕΘΟΔΟΛΟΓΙΑ.....	32
4.1 ΨΗΦΙΑΚΟ ΠΕΡΙΒΑΛΛΟΝ.....	32
4.2 ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ ΔΙΑΔΙΚΑΣΙΑΣ.....	32
ΚΕΦΑΛΑΙΟ 5: ΕΦΑΡΜΟΓΗ ΣΤΗΝ ΠΕΡΙΟΧΗ ΜΕΛΕΤΗΣ.....	34
5.1 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ.....	34
5.1.1 ΕΚΔΟΧΗ 1.....	34
5.1.2 ΕΚΔΟΧΗ 2.....	39
5.2 ΕΦΑΡΜΟΓΗ ΕΚΔΟΧΗΣ 1.....	40
5.2.1 ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΕΙΚΟΝΩΝ ΜΕ ΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ SAM.....	40
5.2.2 ΑΝΙΧΝΕΥΣΗ ΜΕΤΑΒΟΛΩΝ ΑΣΤΙΚΟΥ ΙΣΤΟΥ ΜΕ ΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ STANET.....	48
5.3 ΕΦΑΡΜΟΓΗ ΕΚΔΟΧΗΣ 2.....	53
5.3.1 ΑΝΙΧΝΕΥΣΗ ΜΕΤΑΒΟΛΩΝ ΑΣΤΙΚΟΥ ΙΣΤΟΥ ΜΕ ΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ STANET.....	53
ΚΕΦΑΛΑΙΟ 6: ΑΞΙΟΛΟΓΗΣΗ.....	58
6.1 ΑΞΙΟΛΟΓΗΣΗ ΕΚΔΟΧΗΣ 1.....	58
6.1.1 ΑΞΙΟΛΟΓΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ SAM.....	58
6.1.2 ΑΞΙΟΛΟΓΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ STANet.....	61

6.2 ΑΞΙΟΛΟΓΗΣΗ ΕΚΔΟΧΗΣ 2	65
6.2.1 ΑΞΙΟΛΟΓΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ STANet	65
ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ – ΠΡΟΟΠΤΙΚΕΣ.....	65
7.1 ΣΥΜΠΕΡΑΣΜΑΤΑ	67
7.2 ΠΡΟΟΠΤΙΚΕΣ	67
ΚΕΦΑΛΑΙΟ 8: ΑΝΑΦΟΡΕΣ	68

ΕΥΡΕΤΗΡΙΟ ΕΙΚΟΝΩΝ

Εικόνα 1: Σχέση τεχνητής νοημοσύνης, μηχανικής και βαθιάς μάθησης και νευρωνικών δικτύων (πηγή: https://armonitor.com/)	16
Εικόνα 2: Βιολογικός νευρώνας (πηγή: https://medium.com/@ismailghallou/build-your-perceptron-neural-net-from-scratch-e12b7be9d1ef)	16
Εικόνα 3: Μοντέλο τεχνητού νευρώνα (πηγή: https://www.v7labs.com/blog/neural-networks-activation-functions)	17
Εικόνα 4: Λειτουργία νευρωνικού δικτύου (πηγή: https://bigblue.academy/gr/neuronika-diktua)	18
Εικόνα 5: Λειτουργία νευρωνικού δικτύου. (πηγή: https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa)	19
Εικόνα 6: Βελτιστοποίηση (πηγή: https://www.analyticsvidhya.com/).....	21
Εικόνα 7: Διαγράμματα υπό-προσαρμογής, υπερ-προσαρμογής και καλής προσαρμογής (πηγή: https://opencourses.auth.gr/).....	21
Εικόνα 8: Αποτέλεσμα μικρής και μεγάλης τιμής ρυθμού εκμάθησης (https://www.javatpoint.com/).....	22
Εικόνα 9: Confusion matrix (πηγή: https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62)	23
Εικόνα 10: Intersection over Union (https://www.v7labs.com/).....	24
Εικόνα 11: a) Η αρχιτεκτονική του νευρωνικού δικτύου STANet, b) Εξαγωγέας χαρακτηριστικών (feature extractor), c) Η Basic spatial-temporal attention module (BAM), d) Η μέθοδος Pyramid spatial-temporal attention module (PAM).	26
Εικόνα 12: Η αρχιτεκτονική του νευρωνικού δικτύου SAM.	28
Εικόνα 13: Αποτέλεσμα μεθόδων MKL (αριστερά) και HM (δεξιά).....	35
Εικόνα 14: Αποτέλεσμα μεθόδων Reinhard (αριστερά) και MVGD (δεξιά).	36
Εικόνα 15: Αποτέλεσμα μεθόδων HM-MVGD-HM (αριστερά) και HM-MKL-HM (δεξιά).	36
Εικόνα 16: Η τελική εικόνα του 2007, έπειτα από την περικοπή στη περιοχή μελέτης.	37
Εικόνα 17: Η τελική εικόνα του 2023, έπειτα από την περικοπή στη περιοχή μελέτης.	37

ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ

Πίνακας 1: Ημερομηνίες λήψης και δορυφόροι εικόνων εκδοχής 2	39
Πίνακας 2: Αποτελέσματα SAM για τα patches της εικόνας 2007.....	40
Πίνακας 3: Αποτελέσματα SAM για τα patches της εικόνας 2023.....	44
Πίνακας 4: Αποτελέσματα του pre-trained δικτύου STANet (1η εκδοχή).....	49
Πίνακας 5: Αποτελέσματα της διαδικασίας finetuning δικτύου STANet (1η εκδοχή)...	50
Πίνακας 6: Αποτελέσματα της διαδικασίας εκπαίδευσης δικτύου STANet (1η εκδοχή)	52
Πίνακας 7: Αποτελέσματα του pre-trained δικτύου STANet (2η εκδοχή).....	54
Πίνακας 8: Αποτελέσματα της διαδικασίας finetuning δικτύου STANet (2η εκδοχή)...	55
Πίνακας 9: Αποτελέσματα της διαδικασίας εκπαίδευσης δικτύου STANet (2η εκδοχή)	56
Πίνακας 10: Αξιολόγηση SAM για την εικόνα του 2007	59
Πίνακας 11: Αξιολόγηση SAM για την εικόνα του 2023	60
Πίνακας 12: Αξιολόγηση STANet για την 1 ^η εκδοχή	64
Πίνακας 13: Αξιολόγηση STANet για την 2 ^η εκδοχή	65

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

1.1 ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Αντικείμενο της ανίχνευσης μεταβολών είναι ο ποσοτικός καθορισμός των αλλαγών της επιφάνειας της γης κατά την διάρκεια του χρόνου, μέσω της από κοινού ανάλυσης δύο ή περισσότερων εικόνων που απεικονίζουν την ίδια περιοχή. Οι αλλαγές αυτές μπορεί να οφείλονται για παράδειγμα σε φυσικές καταστροφές ή αστική εξάπλωση. Ακόμη, η ανίχνευση μεταβολών έχει εφαρμογές στην διαχείριση των φυσικών πόρων και στην χαρτογράφηση των χρήσεων και των καλύψεων γης (Daudt, 2020). Η αρχή του πεδίου της ανίχνευσης μεταβολών ορίστηκε από τους Weismiller et al. (1977), οι οποίοι δοκίμασαν την μέθοδο της διαφοράς μεταξύ δύο τηλεπισκοπικών εικόνων για να ανιχνεύσουν την αλλαγή μεταξύ τους. Οι μέθοδοι μπορούν να διαχωριστούν στις παραδοσιακές μεθόδους, ειδικότερα σε εκείνες που βασίζονται στις εικονοψηφίδες (pixel-based), σε εκείνες που βασίζονται στα αντικείμενα (object-based), ανάλογα με την κλίμακα των αντικειμένων, και στις μεθόδους που βασίζονται στα νευρωνικά δίκτυα. Σύμφωνα με τον Shi et al., 2020, οι εικόνες αρχικά πρέπει να υποστούν προεπεξεργασία, και ειδικότερα να γίνει αντιστοίχιση (“registration”- Vakaloroulou et al., 2015), ραδιομετρικές διορθώσεις και αφαίρεση του θορύβου. Στη συνέχεια, για την περίπτωση όπου ακολουθηθεί μια μέθοδος μηχανικής μάθησης, κατασκευάζεται το σετ με τις εικόνες εκπαίδευσης και αν χρειάζεται γίνεται επαύξηση των δεδομένων. Έπειτα, εκπαιδεύεται το μοντέλο με βάση ένα καθορισμένο κριτήριο και στη συνέχεια εφαρμόζεται στις εικόνες του σετ ελέγχου. Το τελικό αποτέλεσμα είναι ο χάρτης αλλαγών (change map).

1.2 ΣΚΟΠΟΣ ΚΑΙ ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΓΑΣΙΑΣ

Η παρούσα εργασία εστιάζει στην μελέτη της ανίχνευσης μεταβολών αστικού ιστού με χρήση νευρωνικών δικτύων. Πιο συγκεκριμένα, ανιχνεύονται μεταβολές αστικού ιστού, μεταξύ δύο εικόνων διαφορετικής ημερομηνίας λήψης. Έτσι, σε πρώτο στάδιο έγινε μελέτη ερευνητικών άρθρων, με σκοπό την επιλογή του νευρωνικού δικτύου που πετυχαίνει τη μεγαλύτερη ακρίβεια στην ανίχνευση μεταβολών αστικού ιστού. Έπειτα από δοκιμές παράχθηκαν ορισμένα αποτελέσματα, τα οποία θα μπορούσαν να χρησιμοποιηθούν σε μελέτες πολεοδομικές, ανίχνευσης αυθαίρετης δόμησης, στατιστικές μελέτες κλπ.

1.3 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ

Η εργασία αναπτύσσεται σε συνολικά 8 κεφάλαια. Στο παρόν κεφάλαιο, γίνεται μία εισαγωγή στο θέμα και στη δομή της εργασίας. Στο κεφάλαιο 2 παρουσιάζεται το θεωρητικό υπόβαθρο, στο οποίο βασίζεται η εργασία, με τις βασικές έννοιες και τις αρχιτεκτονικές των δικτύων που αξιοποιήθηκαν. Στη συνέχεια, στο κεφάλαιο 3 παρουσιάζεται μία βιβλιογραφική αναφορά (State of the art), η οποία αποτέλεσε την αφορμή για την προσωπική μου έρευνα στο συγκεκριμένο αντικείμενο. Έπειτα, στο κεφάλαιο 4, περιγράφεται συνοπτικά η διαδικασία που ακολουθήθηκε. Στο κεφάλαιο 5,

αναφέρεται η εφαρμογή της παραπάνω μεθοδολογίας στα δεδομένα της περιοχής μελέτης, εκτενέστερα. Ακολούθως, στο κεφάλαιο 6 παρουσιάζεται η αξιολόγηση των αποτελεσμάτων από τις εφαρμογές. Έπειτα, τα συμπεράσματα που προέκυψαν, καθώς και οι προτάσεις για πιθανή μελλοντική έρευνα παρουσιάζονται στο Κεφάλαιο 7. Τέλος, στο Κεφάλαιο 8, αναγράφονται οι αναφορές, οι οποίες συνέβαλαν στην επιστημονική μου κατάρτιση και στην συγγραφή της παρούσας διπλωματικής.

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 ΑΝΙΧΝΕΥΣΗ ΜΕΤΑΒΟΛΩΝ ΑΣΤΙΚΟΥ ΙΣΤΟΥ

Αρκετά χρόνια, γίνονται μελέτες γύρω από το θέμα της εξαγωγής δεδομένων της επιφάνειας της γης από δορυφορικές εικόνες και αεροφωτογραφίες. Τα επιστημονικά πεδία, Τηλεπισκόπηση, Φωτογραμμετρία, Συστήματα Γεωγραφικών Πληροφοριών, σε συνδυασμό με εικόνες υψηλής ανάλυσης κάνουν ευκολότερη την παρακολούθηση της γης από απόσταση. Συγκεκριμένα, η έγκαιρη και έγκυρη ανίχνευση μεταβολών κτηρίων σε αστικό και περιαστικό περιβάλλον αποτελεί μείζονος σημασίας θέμα παγκοσμίως. Η αυτόματη ανίχνευση μεταβολών προσφέρει λύση στον πυρήνα του προβλήματος, καθώς είναι μία λύση με άμεσα αποτελέσματα, τα οποία είναι αντικειμενικά, λόγω της ελάχιστης εμπλοκής του ανθρώπινου παράγοντα. Το στοιχείο αυτό καθιστά τα αποτελέσματα της μεθόδου αξιόπιστα και αδιαμφισβήτητα και άρα ιδανικά να προσφέρουν τη λύση στο πρόβλημα. Ωστόσο, παρουσιάζονται αρκετές δυσκολίες, λόγω της αυξημένης διακύμανσης και των γεωμετρικών σφαλμάτων καταγραφής που προκαλούνται από διαφορετικές κατευθύνσεις και γωνίες δορυφορικής θέασης. Τον τελευταίο καιρό, οι μέθοδοι βαθιάς μάθησης (Deep Learning: DL) που αφορούν στην αυτόματη ανίχνευση μεταβολών έχουν αποδειχθεί πολύ ελκυστικές για το πρόβλημα, λόγω της ευελιξίας τους να συνδυάζουν και να επεξεργάζονται διαφορετικούς τύπους πληροφοριών σε συνδυασμό με την αυξημένη διαθεσιμότητα συστημάτων υψηλότερης επεξεργαστικής ισχύος.

Η φυσική γήινη επιφάνεια καλύπτεται από πληθώρα φυσικών και τεχνητών αντικειμένων. Ορισμένα από αυτά είναι στατικά και δεν μεταβάλλονται καθόλου ή μεταβάλλονται σπάνια με το πέρασμα του χρόνου. Αντίθετα, κάποια άλλα χαρακτηρίζονται ως δυναμικά, διότι εμφανίζουν συχνότερες μεταβολές διαχρονικά (Jensen, 1996). Ο όρος μεταβολή αναφέρεται στην μετατροπή ενός αντικειμένου ή ενός φαινομένου σε κάτι άλλο εξαιτίας κάποιας διαδικασίας ή δράσης (Hornsby and Egenhofer, 2000). Η ανίχνευση μεταβολών είναι η διαδικασία αναγνώρισης των διαφορών σε επίπεδο αντικειμένων και φαινομένων με την παρακολούθησή τους σε διαφορετικές χρονικές στιγμές και με την επεξεργασία δεδομένων πολλαπλών χρονικών στιγμών, έτσι ώστε να αναλυθούν διαχρονικά οι επιπτώσεις τους (Singh, 1989). Σκοπός της είναι η παροχή πληροφοριών σχετικά με το είδος, την χωρική κατανομή και το ποσοστό των αλλαγών, αλλά και την ακρίβεια των αποτελεσμάτων. Η έγκαιρη και ακριβής ανίχνευση μεταβολών των χαρακτηριστικών της επιφάνειας της Γης δημιουργεί τη βάση για την καλύτερη κατανόηση των σχέσεων και αλληλεπιδράσεων μεταξύ των ανθρωπίνων και των φυσικών φαινομένων με σκοπό την καλύτερη διαχείριση και χρήση των φυσικών διαθέσιμων (Lu et al., 2004).

Η ανάγκη για ανίχνευση μεταβολών δεν αφορά μόνο σε φαινόμενα αλλά και στις χρήσεις γης ή εδαφοκαλύψεις που παρατηρούνται σε μια περιοχή. Σε αυτές τις περιπτώσεις αναλύονται οι ιδιότητες των χρήσεων/καλύψεων γης χρησιμοποιώντας γεωανηγμένα δορυφορικά δεδομένα πολλαπλών χρονικών στιγμών (Μιχαηλίδου, 2011). Τα δορυφορικά δεδομένα πρέπει να προέρχονται από κατάλληλο τηλεπισκοπικό σύστημα, ώστε να καθίσταται δυνατή η αναγνώριση των κατηγοριών κάλυψης γης που ενδιαφέρουν προς ανίχνευση (Καραγιάννης, 2011). Η ανίχνευση μεταβολών σε ανθρωπογενείς κατασκευές χρησιμοποιώντας εικόνες Τηλεπισκόπησης έχει αξία και σε πολεοδομικά θέματα, καθώς και στην ενημέρωση και αναθεώρηση χαρτών, όπου με την άμεση επαλήθευση της πληροφορίας που παρέχει ο χάρτης ή οι

βάσεις δεδομένων γεωγραφικών πληροφοριών, διευκολύνονται καθοριστικά οι χρήστες στη λήψη αποφάσεων σε μια σειρά σημαντικών θεμάτων (Καραγιάννης, 2011, Καραλής, 2008).

2.2 ΒΑΘΙΑ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (DEEP LEARNING)

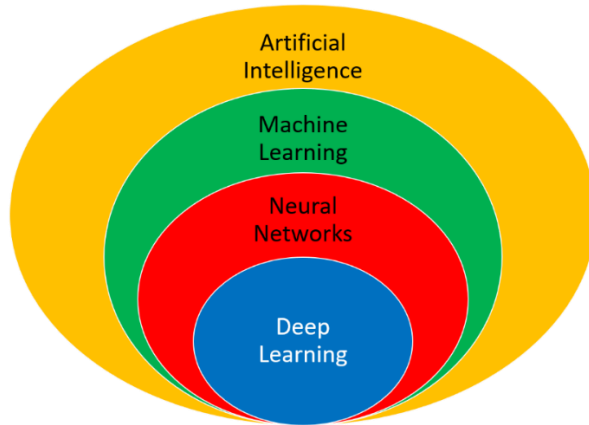
Η Βαθιά Μάθηση είναι μια εξειδικευμένη μορφή μηχανικής μάθησης, η οποία βασίζεται στα νευρωνικά δίκτυα. Αποτελεί μία τεχνική Μηχανικής Μάθησης η οποία διδάσκει στους υπολογιστές να πράττουν όπως ο ανθρώπινος εγκέφαλος: μάθηση μέσω παραδειγμάτων. Είναι μία εξ' ολοκλήρου αυτόματη διαδικασία, για την οποία απαιτείται μεγάλος όγκος δεδομένων. Ονομάζεται «βαθιά» επειδή περιλαμβάνει πολλά επίπεδα του νευρωνικού δικτύου και μαζικούς όγκους πολύπλοκων και ανόμοιων δεδομένων. Για να επιτύχει τη Βαθιά Μάθηση, το σύστημα συνδέεται με πολλαπλά επίπεδα στο δίκτυο, εξάγοντας όλο και περισσότερο εξόδους υψηλότερου επιπέδου.

2.3 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ (ARTIFICIAL INTELLIGENCE)

Η Τεχνητή Νοημοσύνη, η οποία είναι η ευρεία επιστήμη της μίμησης των ανθρώπινων ικανοτήτων, χωρίζεται σε δύο υποσύνολα: τη Μηχανική Μάθηση και τη Βαθιά Μάθηση. Σκοπός της είναι οι μηχανές να μιμηθούν την ανθρώπινη νοημοσύνη και την ανθρώπινη συμπεριφορά ως προς την εξαγωγή συμπερασμάτων. Γενικά, η μηχανική μάθηση χωρίζεται σε τέσσερις κατηγορίες: 1) επιβλεπόμενη μηχανική μάθηση (supervised machine learning), 2) μη-επιβλεπόμενη μηχανική μάθηση (unsupervised machine learning), 3) ημι-επιβλεπόμενη μάθηση (semi-supervised learning), και 4) ενισχυτική μάθηση (reinforcement learning).

2.4 ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ (MACHINE LEARNING)

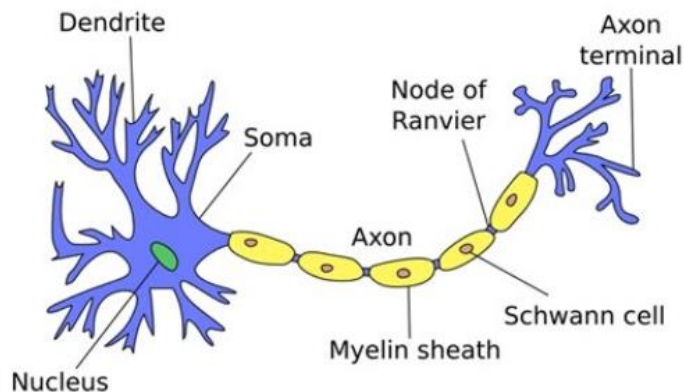
Η Μηχανική Μάθηση αποτελεί κλάδο της Τεχνητής Νοημοσύνης. Επικεντρώνεται στη διδασκαλία των υπολογιστών να μαθαίνουν από τα δεδομένα και να βελτιώνονται με την εμπειρία – αντί να είναι ρητά προγραμματισμένοι να το κάνουν. Στη Μηχανική Μάθηση, οι αλγόριθμοι εκπαιδεύονται για να βρουν μοτίβα και συσχετίσεις σε μεγάλα σύνολα δεδομένων και να λαμβάνουν τις καλύτερες αποφάσεις και προβλέψεις με βάση αυτή την ανάλυση.



Εικόνα 1: Σχέση τεχνητής νοημοσύνης, μηχανικής και βαθιάς μάθησης και νευρωνικών δικτύων (πηγή: <https://apmonitor.com/>)

2.5 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

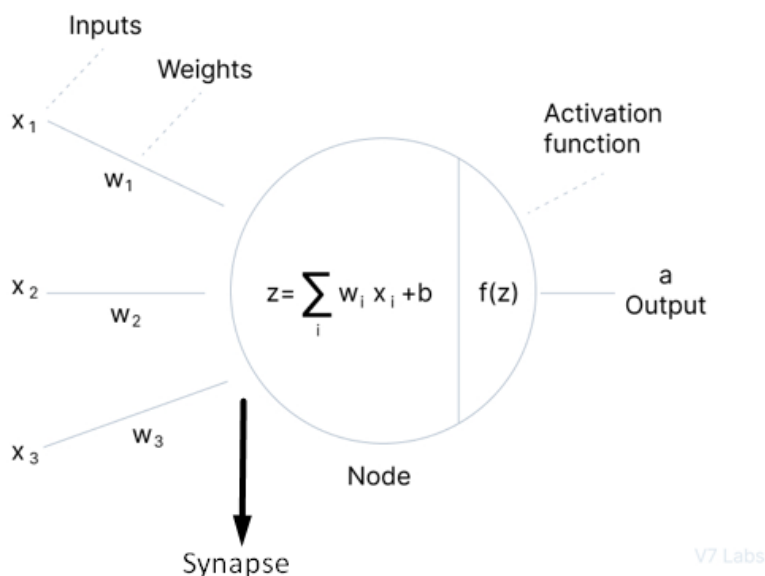
Τα νευρωνικά δίκτυα (neural networks), ή αλλιώς τεχνητά νευρωνικά δίκτυα (artificial neural networks), είναι υπολογιστικά συστήματα εμπνευσμένα από τα βιολογικά νευρικά δίκτυα που υπάρχουν στους ζωντανούς οργανισμούς. Ένας νευρώνας, αποτελείται από τους δενδρίτες, οι οποίοι δέχονται το σήμα, από τον άξονα που το μεταφέρει και από τη σύναψη, μέσα από την οποία το σήμα μεταφέρεται σε άλλο νευρώνα.



Εικόνα 2: Βιολογικός νευρώνας (πηγή: <https://medium.com/@ismailghallou/build-your-perceptron-neural-net-from-scratch-e12b7be9d1ef>)

Ένα νευρωνικό δίκτυο απαρτίζεται από συνδεδεμένους κόμβους, οι οποίοι ονομάζονται τεχνητοί νευρώνες και μοντελοποιούν τους νευρώνες ενός βιολογικού εγκεφάλου. Κάθε σύνδεση, σαν τις συνάψεις του βιολογικού εγκεφάλου, έχει την δυνατότητα να εκπέμψει ένα σήμα στους άλλους νευρώνες. Ένας τεχνητός νευρώνας λαμβάνει σήματα, τα επεξεργάζεται και στέλνει σήμα στους συνδεδεμένους σε αυτόν νευρώνες. Οι συνδέσεις ονομάζονται ακμές. Κάθε ακμή έχει ένα βάρος, η τιμή του

οποίου προσαρμόζεται κατά την εκπαίδευση. Το βάρος αυξάνει ή μειώνει την ισχύ του σήματος σε μια σύνδεση. Το σήμα σε μία σύνδεση είναι ένας πραγματικός αριθμός και η έξοδος ενός νευρώνα υπολογίζεται από μια μη γραμμική συνάρτηση του αθροίσματος της πόλωσης (bias) και των γινομένων της κάθε εισόδου του με το αντίστοιχο βάρος της σύνδεσης από την οποία προήλθε. Η έξοδος του νευρώνα συγκρίνεται με ένα κατώφλι μέσω της συνάρτησης ενεργοποίησης. Εάν η τιμή της εξόδου είναι μεγαλύτερη από το κατώφλι, ο νευρώνας ενεργοποιείται.



Εικόνα 3: Μοντέλο τεχνητού νευρώνα (πηγή: <https://www.v7labs.com/blog/neural-networks-activation-functions>)

Στην Εικόνα 3, παρουσιάζεται ένα μοντέλο τεχνητού νευρώνα. Τα inputs με τα βάρη τους δημιουργούν ένα σταθμισμένο άθροισμα, το οποίο μεταφέρεται στο νεύρο και εκεί εφαρμόζεται μια συνάρτηση, η λεγόμενη activation function, που κρίνει αν εν τέλει θα περάσει το σήμα ή όχι. Αν περάσει παράγεται το output, όπου και εδώ εφαρμόζεται μια αντίστοιχη συνάρτηση έτσι ώστε να βγει το τελικό output, το οποίο θα είναι είτε μία τιμή είτε σε binary μορφή (0,1). Πρακτικά, με μαθηματικούς όρους παράγονται δεδομένα εφαρμόζοντας μια συνάρτηση στα εισαγόμενα δεδομένα.

Κάθε τεχνητός νευρώνας αποτελείται από πολλές εισόδους (inputs) x_i και μία μόνο έξοδο (output) a . Κάθε είσοδος (input) x_i «ζυγίζεται» με ένα συναπτικό βάρος w_i (weight) και τα αποτελέσματα αθροίζονται σταθμισμένα, μέσω της συνάρτησης αθροίσματος z (summation function). Το άθροισμα αυτό μεταφέρεται στο νεύρο και εκεί ο τεχνητός νευρώνας πιθανότατα δίνει έξοδο μέσω της συνάρτησης ενεργοποίησης (activation function), η οποία κρίνει αν τελικά το σήμα θα περάσει ή όχι. Αν περάσει παράγεται το output, όπου και εδώ εφαρμόζεται μια αντίστοιχη συνάρτηση έτσι ώστε να βγει το τελικό output, το οποίο θα είναι είτε τιμή είτε σε binary μορφή (0,1).

Στα επίπεδα δεν είναι απαραίτητη η χρήση της ίδιας συνάρτησης (activation function). Ορισμένες συναρτήσεις που χρησιμοποιούνται συχνά είναι οι εξής:

- Συνάρτηση κατώφλι (Threshold):

$$f(x) = \begin{cases} 1, & \text{για } x \geq 0 \\ 0, & \text{για } x < 0 \end{cases}$$

Αυτή η συνάρτηση για τιμές μεγαλύτερες ή ίσες του 0 εξάγει την τιμή 1, ενώ για τιμές μικρότερες του 0 την τιμή 0.

- Σιγμοειδής:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Αυτή η συνάρτηση δέχεται οποιαδήποτε πραγματική τιμή ως είσοδο και εξάγει τιμές στο εύρος 0 έως 1. Όσο μεγαλύτερη είναι η είσοδος, τόσο πιο κοντά στο 1 θα είναι η τιμή εξόδου, ενώ όσο μικρότερη είναι η είσοδος (πιο αρνητική), τόσο πιο κοντά στο 0 θα είναι η έξοδος.

- Tanh:

$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$$

Αυτή η συνάρτηση είναι αρκετά παρόμοια με τη σιγμοειδή, με τη διαφορά ότι το εύρος των τιμών είναι από -1 έως 1.

- Rectified Linear Unit (ReLU):

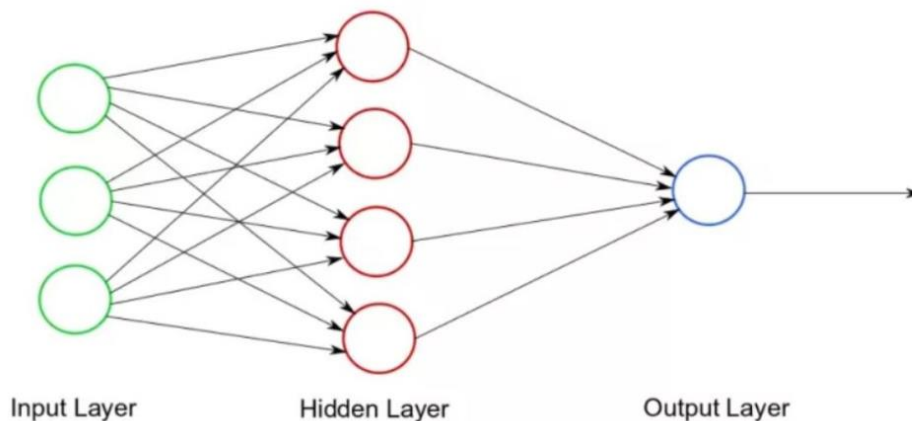
$$f(x) = \max(0, x)$$

Αυτή η συνάρτηση για είσοδο μια τιμή μικρότερη του 0 το αποτέλεσμα τείνει στο 0, ενώ για είσοδο μεγαλύτερη του 0 το αποτέλεσμα τείνει στο 1.

- LeakyReLU:

$$f(x) = \max(a * x, x)$$

Αυτή η συνάρτηση είναι παρόμοια με την ReLU, με τη διαφορά ότι για αρνητική είσοδο υπάρχει συνεχής κλίση, η οποία είναι μια παράμετρος a που βάζει ο χρήστης.

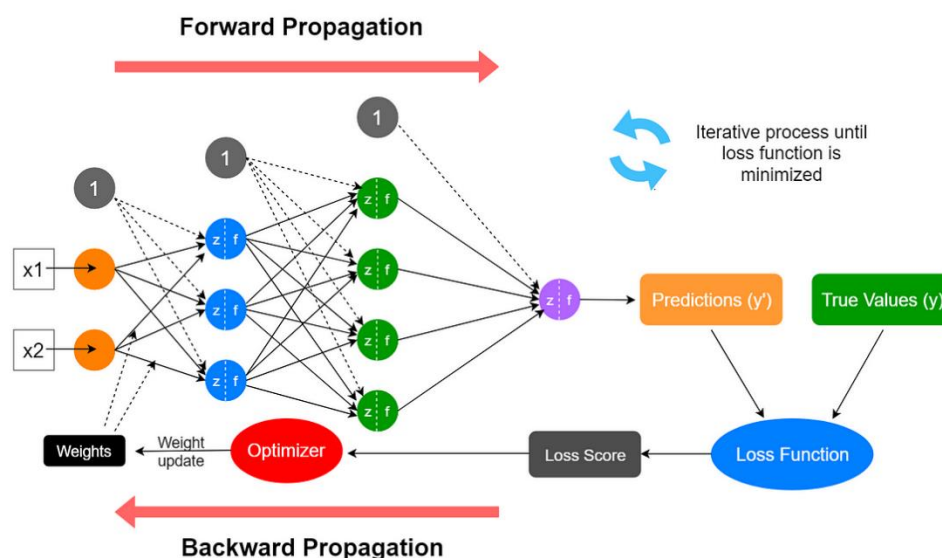


Εικόνα 4: Λειτουργία νευρωνικού δικτύου (πηγή: <https://bigblue.academy/gr/neuronika-diktua>)

Ένα βασικό νευρωνικό δίκτυο έχει διασυνδεδεμένους τεχνητούς νευρώνες σε τρία επίπεδα:

- Επίπεδο εισόδου (input layer): Πρόκειται για το πρώτο επίπεδο σε ένα νευρωνικό δίκτυο που λαμβάνει τα αρχικά δεδομένα εισόδου. Οι κόμβοι εισόδου επεξεργάζονται τα δεδομένα αυτά, τα αναλύουν και τα μεταβιβάζουν στο επόμενο επίπεδο.
- Κρυμμένο επίπεδο (hidden layer): Τα κρυμμένα στρώματα αποτελούν τα ενδιάμεσα επίπεδα ανάμεσα στο επίπεδο εισόδου και εξόδου και επιτελούν το μεγαλύτερο μέρος του υπολογισμού. Ενδέχεται να υπάρχουν πολλά κρυμμένα επίπεδα σε ένα νευρωνικό δίκτυο.
- Επίπεδο εξόδου (output layer): Το επίπεδο εξόδου είναι το τελικό επίπεδο σε ένα νευρωνικό δίκτυο και οδηγεί στην έξοδο του δικτύου. Ο αριθμός των νευρώνων στο επίπεδο αυτό εξαρτάται από το εκάστοτε πρόβλημα που λύνεται.

Όταν υπάρχουν περισσότερα από ένα κρυμμένα επίπεδα (hidden), τότε το δίκτυο ονομάζεται βαθύ (deep).



Εικόνα 5: Λειτουργία νευρωνικού δικτύου. (πηγή: <https://medium.com/data-science-365/overview-of-a-neural-networks-learning-process-61690a502fa>)

Η διαδικασία εκμάθησης (training) ενός νευρωνικού δικτύου, όπως παρουσιάζεται στην Εικόνα 5, είναι μια επαναληπτική διαδικασία κατά την οποία οι υπολογισμοί εκτελούνται προς τα εμπρός και προς τα πίσω μέσω κάθε στρώματος του δικτύου μέχρι να ελαχιστοποιηθεί η συνάρτηση κόστους (διαδικασία gradient descent).

Ολόκληρη η διαδικασία μάθησης μπορεί να χωριστεί σε τρία κύρια μέρη:

- Προς τα εμπρός διάδοση (Forward propagation)
- Υπολογισμός της συνάρτησης απώλειας
- Προς τα πίσω διάδοση (Backpropagation)

Ένα νευρωνικό δίκτυο αποτελείται από πολλαπλούς νευρώνες (perceptrons), οι οποίοι στοιβάζονται σε επίπεδα. Οι συνδέσεις μεταξύ των στρωμάτων πραγματοποιούνται

μέσω των παραμέτρων (που αναπαρίστανται με βέλη) του δικτύου. Οι παράμετροι είναι τα βάρη (weights) και οι πολώσεις (biases). Τα βάρη ελέγχουν το επίπεδο σημαντικότητας κάθε εισόδου, ενώ οι πολώσεις (biases) καθορίζουν πόσο εύκολα πυροδοτείται ή ενεργοποιείται ένας νευρώνας. Αρχικά, τίθενται μη μηδενικές τυχαίες τιμές στα βάρη και τις πολώσεις (biases). Αυτό ονομάζεται αρχικοποίηση των παραμέτρων του δικτύου. Με βάση αυτές τις ανατεθείσες τιμές και τις τιμές εισόδου, εκτελούνται οι ακόλουθοι υπολογισμοί σε κάθε νευρώνα του δικτύου:

- Υπολογισμός της γραμμικής συνάρτησης του νευρώνα
- Υπολογισμός της συνάρτησης ενεργοποίησης του νευρώνα

Αυτοί οι υπολογισμοί πραγματοποιούνται σε ολόκληρο το δίκτυο. Μετά την ολοκλήρωση των υπολογισμών στους κόμβους του στρώματος εξόδου, λαμβάνεται η τελική έξοδος, της προς τα εμπρός διάδοσης στην πρώτη επανάληψη.

Στον εμπρόσθιο πολλαπλασιασμό, οι υπολογισμοί γίνονται από το στρώμα εισόδου στο στρώμα εξόδου (από αριστερά προς τα δεξιά) μέσω του δικτύου.

Η διαδικασία forward propagation είναι το αρχικό βήμα στην εκπαίδευση ενός νευρωνικού δικτύου, όπου τα δεδομένα εισόδου τροφοδοτούνται μέσω του δικτύου για τη δημιουργία μιας πρόβλεψης. Εισάγονται τα δεδομένα εισόδου μέσω του δικτύου και κάθε νευρώνας στα κρυφά στρώματα υπολογίζει το σταθμισμένο άθροισμα των εισόδων του και εφαρμόζει μια συνάρτηση ενεργοποίησης σε αυτό. Αυτή η έξοδος στη συνέχεια περνάει στο επόμενο στρώμα και αυτή η διαδικασία επαναλαμβάνεται μέχρι να εισαχθεί στο στρώμα εξόδου και να ληφθεί η τελική έξοδος (προβλεπόμενη τιμή).

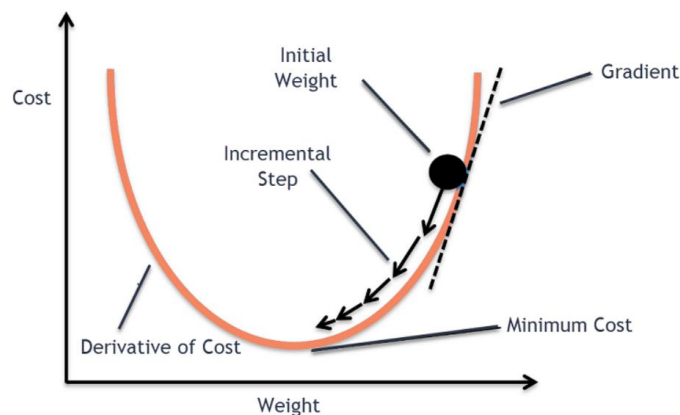
Έπειτα, ακολουθεί ο υπολογισμός του σφάλματος. Γίνεται σύγκριση της εξόδου του δικτύου με την πραγματική έξοδο που αναμένεται (το ground truth από τα δεδομένα εκπαίδευσης). Η διαφορά μεταξύ της αναμενόμενης και της πραγματικής εξόδου είναι το σφάλμα πρόβλεψης.

Στη συνέχεια, ακολουθεί η διαδικασία backward propagation. Τώρα, αυτό το σφάλμα περνάει προς τα πίσω μέσω του δικτύου. Σε αυτήν τη διαδικασία το δίκτυο εκπαιδεύεται. Χρησιμοποιώντας αυτό το σφάλμα, υπολογίζονται οι κλίσεις του σφάλματος σε σχέση με τα βάρη και τις πολώσεις. Η διαδικασία backward propagation είναι μια μέθοδος που χρησιμοποιείται για τον υπολογισμό της κλίσης της συνάρτησης απώλειας (the gradient of the loss function) σε σχέση με τα βάρη του δικτύου.

Τέλος, ακολουθεί η διαδικασία βελτιστοποίησης (gradient descent). Η συνάρτηση κόστους δίνει ένα μέτρο του πόσο απέχουν οι προβλέψεις από τις πραγματικές τιμές. Για να βελτιωθεί το μοντέλο, πρέπει να ελαχιστοποιηθεί αυτή η συνάρτηση κόστους, χρησιμοποιώντας έναν αλγόριθμο βελτιστοποίησης γνωστό ως gradient descent. Η διαδικασία gradient descent ρυθμίζει επαναληπτικά τις παραμέτρους, τείνοντάς τις προς την κατεύθυνση που μειώνει τη συνάρτηση κόστους μέχρι να εντοπίσει τις παραμέτρους με το χαμηλότερο δυνατό κόστος. Ο μαθηματικός τύπος, ο οποίος εκφράζει την τιμή μιας νέας παραμέτρου προς βελτιστοποίηση είναι ο εξής:

$$x' = x - \epsilon \nabla_x f(x),$$

όπου ϵ : ο ρυθμός εκμάθησης (τιμή η οποία ορίζει την ταχύτητα με την οποία συγκλίνει ο αλγόριθμος).

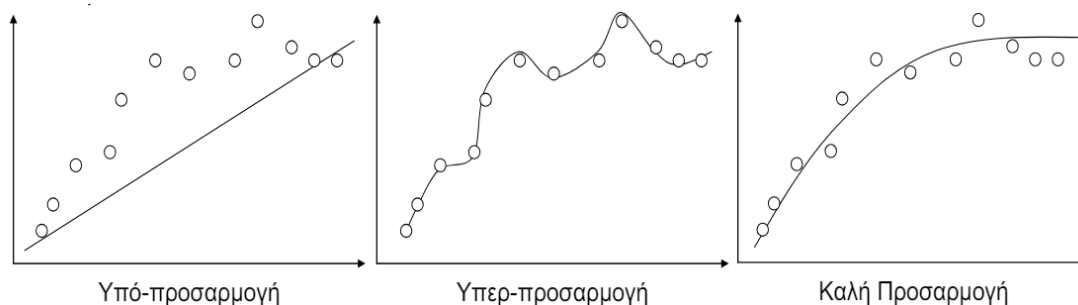


Εικόνα 6: Βελτιστοποίηση (πηγή: <https://www.analyticsvidhya.com/>)

Στην εκπαίδευση μπορεί να εμφανιστούν φαινόμενα:

A) υπό-προσαρμογής ή ατελούς μάθησης (underfitting): αποτυχία μοντελοποίησης δεδομένων εκπαίδευσης, π.χ. λόγω απλότητας του δικτύου.

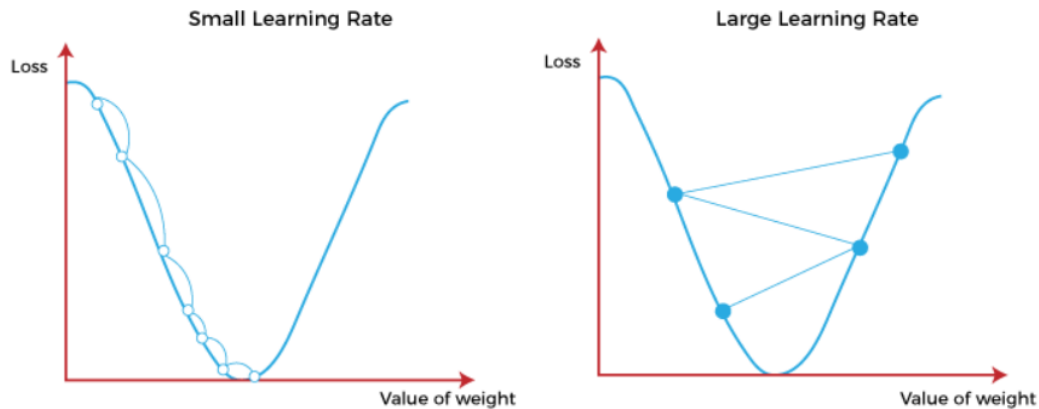
B) υπερ-προσαρμογής (overfitting): ένα πολύπλοκο δίκτυο ενδέχεται να μοντελοποιήσει υπερβολικά τα δεδομένα εκπαίδευσης, καθώς και το θόρυβο που πιθανώς υπάρχει σε αυτά. Επομένως, θα δίνει σωστή πρόβλεψη για τα δεδομένα εκπαίδευσης αλλά λανθασμένες προβλέψεις για άλλα δεδομένα εισόδου. Αυτό συμβαίνει όταν η τιμή του ρυθμού εκμάθησης είναι μικρή και άρα υπάρχει αργή σύγκλιση. Αντίθετα, μεγάλες τιμές οδηγούν σε μεγάλες τροποποιήσεις στις παραμέτρους.



Εικόνα 7: Διαγράμματα υπό-προσαρμογής, υπερ-προσαρμογής και καλής προσαρμογής (πηγή: <https://opencourses.auth.gr/>)

Αυτός ο κύκλος (forward propagation, error calculation, backpropagation, and weights και biases update) επαναλαμβάνεται για πολλές επαναλήψεις (ή εποχές) κατά τη διάρκεια της διαδικασίας εκπαίδευσης, βελτιώνοντας σταδιακά την ακρίβεια των προβλέψεων του δικτύου.

Επίσης, αξίζει να σημειωθεί ότι στη μηχανική μάθηση και στα στατιστικά στοιχεία, ο ρυθμός εκμάθησης (learning rate) είναι μια παράμετρος σε έναν αλγόριθμο βελτιστοποίησης που καθορίζει το μέγεθος του βήματος σε κάθε επανάληψη, ενώ κινείται προς την ελάχιστη συνάρτηση απώλειας. Η τιμή της παραμέτρου αυτής, τροποποιεί σημαντικά τα αποτελέσματα.



Εικόνα 8: Αποτέλεσμα μικρής και μεγάλης τιμής ρυθμού εκμάθησης (<https://www.javatpoint.com/>)

2.6 ΜΗΧΑΝΙΣΜΟΙ ΠΡΟΣΟΧΗΣ-ΜΕΤΑΣΧΗΜΑΤΙΣΤΕΣ

Οι μετασχηματιστές (transformers) αποτελούν μια σημαντική αρχιτεκτονική στον τομέα της βαθιάς μάθησης, γνωστή για τις εξαιρετικές επιδόσεις της στην επεξεργασία φυσικής γλώσσας. Το σημείο που ξεχωρίζει τους μετασχηματιστές είναι ο μηχανισμός αυτό-προσοχής που διαθέτουν, ο οποίος τους επιτρέπει να επεξεργάζονται αποτελεσματικά τις ακολουθίες εισόδου εξετάζοντας ταυτόχρονα τις σχέσεις μεταξύ όλων των στοιχείων. Αυτός ο μηχανισμός αυτό-προσοχής επιτρέπει στους μετασχηματιστές να καταγράφουν εξαρτήσεις υψηλής εμβέλειας. Ο πυρήνας της αρχιτεκτονικής ενός δικτύου transformer αποτελείται από μια δομή κωδικοποιητή (encoder) -αποκωδικοποιητή (decoder), η καθεμία από τις οποίες περιλαμβάνει πολλαπλά στρώματα αυτό-προσοχής και νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης. Ο encoder κωδικοποιεί την ακολουθία εισόδου σε μία ενδιάμεση αναπαράσταση. Στη συνέχεια, η αναπαράσταση αυτή τροφοδοτείται στον decoder, ο οποίος παράγει την ακολουθία εξόδου. Ο μηχανισμός attention επιτρέπει στο μοντέλο, σε κάθε βήμα, να επικεντρώνεται σε εκείνο το τμήμα της ακολουθίας εισόδου, το οποίο θεωρεί σημαντικό για τον υπολογισμό της εξόδου. Με άλλα λόγια, το μοντέλο μαθαίνει να διακρίνει τις σχέσεις μεταξύ των συμβόλων της ακολουθίας εισόδου.

2.7 FINETUNING

Στο στάδιο του finetuning, το γενικό προ-εκπαιδευμένο μοντέλο βελτιστοποιείται σε ένα συγκεκριμένο task. Η εκπαίδευση πραγματοποιείται για έναν μικρό αριθμό από εποχές στο τελικό σύνολο δεδομένων. Ανάλογα με το προς επίλυση πρόβλημα, στην έξοδο του μοντέλου τοποθετείται ένα δίκτυο, το οποίο παράγει τη ζητούμενη έξοδο. Συνήθως, το δίκτυο αυτό αποτελείται από ένα πλήρως συνδεδεμένο επίπεδο, το οποίο προβάλλει τις τελικές αναπαραστάσεις των συμβόλων της ακολουθίας σε έναν μικρότερο χώρο, λειτουργώντας ως ταξινομητής.

2.8 ΕΚΦΡΑΣΗ ΑΚΡΙΒΕΙΑΣ ΚΑΙ ΜΕΤΡΙΚΕΣ

Αρχικά, ένας από τους πιο συχνούς τρόπους έκφρασης της ακρίβειας της ταξινόμησης είναι η δημιουργία πινάκων σφαλμάτων ταξινόμησης, οι οποίοι ονομάζονται πίνακες σύγχυσης. Οι πίνακες αυτοί συγκρίνουν τη σχέση μεταξύ αληθών δεδομένων και των αντίστοιχων αποτελεσμάτων μιας αυτόματης διαδικασίας ταξινόμησης κατηγορία προς κατηγορία. Απεικονίζουν και συνοψίζουν την απόδοση ενός αλγορίθμου ταξινόμησης.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Εικόνα 9: Confusion matrix (πηγή: <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>)

Στον παραπάνω πίνακα αναφέρονται ως: TP (True Positive): τιμή αληθής και θετική, FP (False Positive): τιμή ψευδής και θετική, FN (False Negative): η πρόβλεψη είναι ψευδής, ενώ το αντικείμενο υφίσταται στην πραγματικότητα, τότε η τιμή είναι ψευδής αρνητική και TN (True Negative): η πρόβλεψη είναι ψευδής, και το αντικείμενο δεν υφίσταται στην πραγματικότητα, τότε η τιμή θεωρείται αληθής αρνητική.

Στη Μηχανική Μάθηση, μια μετρική είναι ένα μέτρο που χρησιμοποιείται για την αξιολόγηση της απόδοσης ενός μοντέλου. Οι μετρικές παρέχουν ένα τρόπο ποσοτικοποίησης της απόδοσης ενός μοντέλου σε μια δεδομένη εργασία.

Ορισμένες μετρικές οι οποίες θα χρησιμοποιηθούν στη συνέχεια της εργασίας είναι οι εξής:

- **Precision**

Η ακρίβεια ορίζεται ως το ποσοστό των προβλέψεων που είναι πραγματικά σωστές σε σχέση με το σύνολο των προβλέψεων που έγιναν για μια συγκεκριμένη κατηγορία. Υπολογίζεται ως η αναλογία των σωστών προβλέψεων (True Positives) προς τον συνολικό αριθμό προβλέψεων (True Positives + False Positives):

Precision: $(\text{True Positive}) / (\text{True Positive} + \text{False Positive}) = (\text{True Positive} / \text{Total Predicted Positive})$

- **Recall**

Η ανάκληση ορίζεται ως το ποσοστό των πραγματικά σωστών προβλέψεων σε σχέση με το σύνολο των πραγματικών σωστών μιας κατηγορίας. Υπολογίζεται ως η αναλογία των True Positives προς το σύνολο των πραγματικών αντικειμένων (True Positives + False Negatives), δηλαδή αντιπροσωπεύει τον αριθμό των δειγμάτων που σωστά προβλέφθηκαν ως θετικά προς το σύνολο των πραγματικών θετικών:

Recall: (True Positive) / (True Positive + False Negative) = (True Positive/Total Actual Positive)

- *F1-score*

Η ακρίβεια είναι η αναλογία των αληθινών θετικών προς το άθροισμα των αληθινών θετικών και των ψευδώς θετικών, ενώ η ανάκληση είναι η αναλογία των αληθινών θετικών προς το άθροισμα των αληθινών θετικών και των ψευδώς αρνητικών. Με άλλα λόγια, η ακρίβεια (Precision) μετρά την αναλογία των σωστά αναγνωρισμένων θετικών δειγμάτων από όλα τα δείγματα που προβλέπονται ως θετικά, ενώ η ανάκληση (Recall) μετρά την αναλογία των σωστά αναγνωρισμένων θετικών δειγμάτων από όλα τα πραγματικά θετικά δείγματα.

Η βαθμολογία F1 είναι η αρμονική μέση ακρίβεια και ανάκληση. Παρέχει μια ενιαία τιμή που συνδυάζει τόσο την ακρίβεια όσο και την ανάκληση σε ένα μόνο μέτρο απόδοσης. Ο μαθηματικός τύπος για τον υπολογισμό της βαθμολογίας F1 είναι:

$F1 = (2 * Precision * Recall) / (Precision + Recall)$, και οι τιμές που λαμβάνει κυμαίνονται από 0 έως 1.

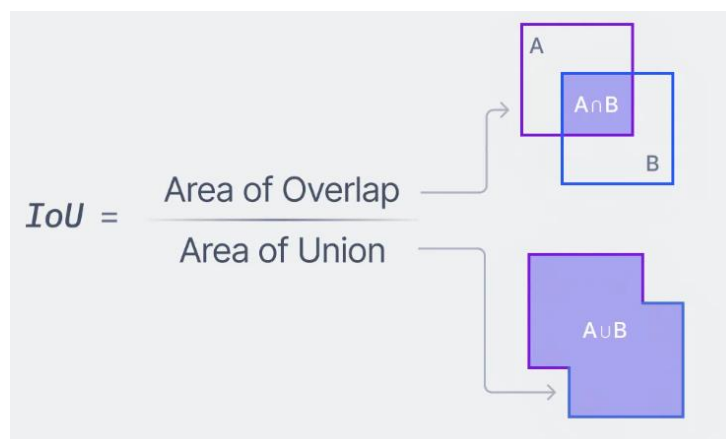
- *Overall Accuracy:*

Η συνολική ακρίβεια υπολογίζεται διαιρώντας το συνολικό αριθμό των ορθά ταξινομημένων εικονοστοιχείων με το συνολικό αριθμό των εικονοστοιχείων αναφοράς. Ο μαθηματικός τύπος για τον υπολογισμό της συνολικής ακρίβειας είναι ο εξής:

Overall Accuracy = (True Positive + True Negative) / (True Positive + False Negative + False Positive + True Negative)

- *IoU (Intersection over Union):* (Precision * Recall) / (Precision + Recall - Precision * Recall)

Η μετρική αυτή αξιολογεί το πόσο καλά εφάρμοσε το μοντέλο την αναγνώριση αντικειμένων σε εικόνες. Υπολογίζεται ως η αναλογία του μεγέθους της κοινής περιοχής (Intersection) μεταξύ της πρόβλεψης και του πραγματικού αντικειμένου, προς το μέγεθος του συνόλου της περιοχής (Union).



Εικόνα 10: Intersection over Union (<https://www.v7labs.com/>)

2.9 ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ STANet (Spatial Temporal Attention neural network)

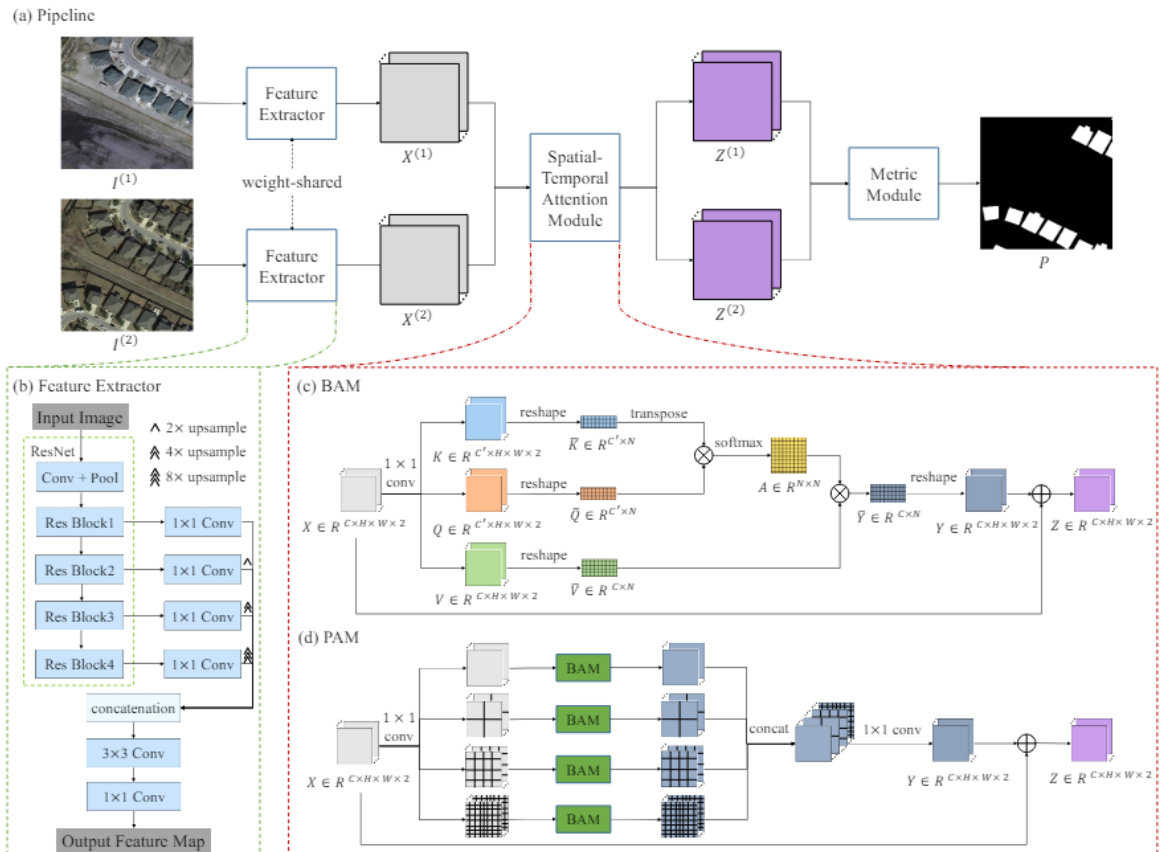
Στο επιστημονικό άρθρο Kristollari V., Karathanassi V. (2022), “Change Detection in VHR Imagery With Severe Co-Registration Errors Using Deep Learning: A Comparative Study”, αξιολογήθηκαν τέσσερις μέθοδοι αυτόματης co-registration και τέσσερις μέθοδοι αυτόματης ανίχνευσης μεταβολών, σε τέσσερις ευρωπαϊκές περιοχές με ποικίλη μορφολογία. Αποδείχθηκε ότι το νευρωνικό δίκτυο STANet εντόπισε με μεγαλύτερη επιτυχία τις αλλαγές που σχετίζονται με κτήρια. Η επιτυχία του μπορεί να αποδοθεί στον ιδιαίτερο μηχανισμό προσοχής του και στο σύνολο δεδομένων εκπαίδευσής του. Οι υπόλοιπες μέθοδοι παρουσίασαν χαμηλή απόδοση.

Το νευρωνικό δίκτυο STANet λειτουργεί ως εξής:

Δεδομένων δύο διαχρονικών εικόνων $I^{(1)}$ και $I^{(2)}$ μεγέθους $H_0 \times W_0$, στόχος της ανίχνευσης είναι η παραγωγή ενός χάρτη ετικέτας (label map) M . Συγκεκριμένα, αξιοποιείται δυαδικό σύστημα που σημαίνει ότι η ετικέτα παίρνει τιμή είτε 1 (αλλαγή), είτε 0 (καμία αλλαγή).

Το STANet αποτελείται από τρία τμήματα: i) έναν εξαγωγέα χαρακτηριστικών (feature extractor), ii) μία ενότητα προσοχής (attention module), iii) μία μετρική ενότητα (metric module).

Αρχικά, οι δύο εικόνες τροφοδοτούνται διαδοχικά στον εξαγωγέα χαρακτηριστικών (ένα πλήρως συνελκτικό δίκτυο-fully convolutional network FCN, πχ RESNet, χωρίς πλήρως συνδεδεμένα επίπεδα), για τη λήψη δύο χαρτών χαρακτηριστικών $X^{(1)}, X^{(2)} \in \mathbb{R}^{C \times H \times W}$, όπου $H \times W$ το μέγεθος του χάρτη χαρακτηριστικών και C η διάσταση του κάθε διανύσματος χαρακτηριστικών. Στη συνέχεια, αυτοί οι χάρτες χαρακτηριστικών μετατρέπονται σε $Z^{(1)}$ και $Z^{(2)}$ μέσω της ενότητας προσοχής (attention module). Αφού ενημερωθεί το μέγεθος των χαρτών χαρακτηριστικών και γίνει ίσο με αυτό των εικόνων εισόδου, η μετρική ενότητα (metric module) υπολογίζει την απόσταση μεταξύ κάθε ζεύγους εικονοστοιχείων στους δύο χάρτες χαρακτηριστικών και δημιουργεί έναν χάρτη απόστασης D . Στη φάση της εκπαίδευσης (training phase), το μοντέλο βελτιστοποιείται ελαχιστοποιώντας την απώλεια (loss) που έχει υπολογισθεί με τον χάρτη απόστασης (distance map) και με τον χάρτη ετικέτας (label map), έτσι ώστε η τιμή απόστασης του σημείου αλλαγής να είναι μεγάλη και η τιμή απόστασης του σημείου μη-αλλαγής να είναι μικρή. Στη φάση δοκιμής (testing phase), ο χάρτης πρόβλεψης P μπορεί να υπολογισθεί με απλό όριο στον χάρτη απόστασης.



Εικόνα 11: a) Η αρχιτεκτονική του νευρωνικού δικτύου STANet, b) Εξαγωγή χαρακτηριστικών (feature extractor), c) Η Basic spatial-temporal attention module (BAM), d) Η μέθοδος Pyramid spatial-temporal attention module (PAM).

Πιο συγκεκριμένα:

- Εξαγωγή χαρακτηριστικών (feature extractor)

Όπως απεικονίζεται στην Εικόνα 11, έχει σχεδιαστεί ένας εξαγωγέας χαρακτηριστικών που μοιάζει με πλήρως συνελκτικό δίκτυο- FCN. Ο συγκεκριμένος εξαγωγέας χαρακτηριστικών βασίζεται στο ResNet-18. Επειδή, το αυθεντικό ResNet είχε σχεδιαστεί για διαδικασίες ταξινόμησης εικόνων, περιλαμβάνει ένα global pooling layer και ένα πλήρως συνδεδεμένο επίπεδο του αυθεντικού ResNet. Το υπόλοιπο τμήμα αποτελείται από πέντε στάδια (καθένα από τα οποία έχει stride= 2).

Λαμβάνοντας τον χάρτη χαρακτηριστικών εξόδου του τελευταίου σταδίου και τροφοδοτώντας τον σε μια συνέλιξη (C1, 1 × 1/1) αυτός μετατρέπεται σε χάρτη διαστάσεων C1. Θα πρέπει να σημειωθεί ότι τα στρώματα batch normalization και η συνάρτηση ReLU παραλείπονται στο σχήμα, για λόγους απλότητας. Ομοίως, οι χάρτες χαρακτηριστικών εξόδου του 2ου, 3ου και 4ου σταδίου τροφοδοτούνται σε τρία διαφορετικά συνελκτικά επίπεδα και μετατρέπονται σε C1. Στη συνέχεια, το μέγεθος των μετασχηματισμένων χαρτών χαρακτηριστικών των τριών τελευταίων σταδίων μετασχηματίζεται στο 1/4 του μεγέθους των εικόνων εισόδου.

Με αυτόν τον τρόπο, προκύπτουν 4 σύνολα χαρτών χαρακτηριστικών από διαφορετικά στάδια των δικτύων. Αυτοί οι τέσσερις χάρτες χαρακτηριστικών συνυπάρχουν στη διάσταση του καναλιού (αποτέλεσμα 4 × C1) και τροφοδοτούνται σε δύο διαφορετικούς μηχανισμούς συνέλιξης (C2, 3 × 3/1 και C3, 1 × 1/1) για τη δημιουργία του τελικού χάρτη χαρακτηριστικών. Αυτά τα δύο στρώματα συνέλιξης μπορούν παράγουν πιο

συμπαγείς αναπαραστάσεις με την εκμετάλλευση των τοπικών χωρικών πληροφοριών.

- *Μηχανισμός Spatial Temporal Attention Module*

Έπειτα, ακολουθεί ο μηχανισμός Spatial Temporal Attention Module. Οι συγκεκριμένοι ερευνητές σχεδίασαν δύο μεθόδους, την Basic spatial-temporal attention module (BAM) και την Pyramid spatial-temporal attention module (PAM).

Στη μέθοδο BAM, οι δύο χάρτες χαρακτηριστικών $X^{(1)}$, $X^{(2)}$ μετασχηματίζονται σε έναν τανυστή χαρακτηριστικών $X \in \mathbb{R}^{C \times H \times W \times 2}$. Έπειτα, τροφοδοτείται στην BAM ώστε να παραχθεί ένας ενημερωμένος τανυστής χαρακτηριστικών $Z \in \mathbb{R}^{C \times H \times W \times 2}$. Τέλος, αυτός χωρίζεται σε δύο χάρτες χαρακτηριστικών $Z^{(1)}$ και $Z^{(2)}$.

Η συνάρτηση που χρησιμοποιείται για να εξαχθεί ο Z από την είσοδο X είναι: $Z = F(X) + X$, όπου $Y = F(X)$ μια υπολειπόμενη αντιστοίχιση (residual mapping) του X προς εκμάθηση.

Για να υπολογισθεί το Y παράγεται ένα σύνολο διανυσμάτων κλειδιών (keys), διανυσμάτων τιμών (values) και διανυσμάτων ερωτημάτων (queries) από τον τανυστή εισόδου, υπολογίζοντας το σταθμισμένο άθροισμα των τιμών, με σκοπό να παράγει τα διανύσματα εξόδου.

Πιο αναλυτικά, ο τανυστής εισόδου X πρώτα μετατρέπεται σε δύο τανυστές $Q, K \in \mathbb{R}^{C \times H \times W \times 2}$. Οι τανυστές Q και K λαμβάνονται από δύο διαφορετικά επίπεδα συνέλιξης ($C', 1 \times 1/1$). Αυτοί μετατρέπονται σε έναν πίνακα κλειδιών (key matrix) \bar{k} και έναν πίνακα ερωτημάτων (query matrix) $\bar{q} \in \mathbb{R}^{C' \times N}$, όπου $N = H \times W \times 2$ ο αριθμός των διανυσμάτων εισόδου. Οι πίνακες αυτοί χρησιμοποιούνται για τον υπολογισμό της προσοχής (attention). Ομοίως, το X τροφοδοτείται σε ένα άλλο επίπεδο συνέλιξης ($C, 1 \times 1/1$) για να παραχθεί ένας νέος τανυστής $V \in \mathbb{R}^{C \times H \times W \times 2}$. Αυτός μετατρέπεται σε πίνακα τιμών (value matrix) $\bar{v} \in \mathbb{R}^{C \times N}$. C' είναι η διάσταση των χαρακτηριστικών κλειδιών και των ερωτημάτων. Στο συγκεκριμένο άρθρο έχει ορισθεί στο C' η τιμή $C/8$ για μείωση της διάστασης των χαρακτηριστικών.

Έπειτα, ορίζεται ο χάρτης χωροχρονικής προσοχής $A \in \mathbb{R}^{N \times N}$, ως πίνακας ομοιότητας. Το στοιχείο $A[i, j]$ στον πίνακα ομοιότητας είναι η ομοιότητα μεταξύ του i κλειδιού και του j ερωτήματος. Στη συνέχεια, πραγματοποιείται ένας πολλαπλασιασμός του ανάστροφου πίνακα \bar{k}^T και του πίνακα \bar{q} , διαιρώντας κάθε στοιχείο με το στοιχείο $\sqrt{c'}$ και εφαρμόζοντας τη συνάρτηση softmax σε κάθε στήλη για να δημιουργηθεί ο χάρτης προσοχής A . Δηλαδή, $A = \text{softmax}(\bar{k}^T * \bar{q}) / (\sqrt{c'})$. Ο πίνακας εξόδου $\bar{y} \in \mathbb{R}^{C \times N}$ υπολογίζεται με τον πολλαπλασιασμό του πίνακα τιμών και του πίνακα ομοιότητας: $\bar{y} = \bar{v} * A$. Άρα, με αυτόν τον τρόπο το \bar{y} αλλάζει διαστάσεις σε Y .

Στη συγκεκριμένη εργασία επιλέχθηκε η μέθοδος PAM, μιας και έχει αποδειχθεί ότι παρουσιάζει καλύτερα αποτελέσματα σε σχέση με την BAM. Η συγκεκριμένη μέθοδος περιέχει τέσσερις κλάδους, καθένας από τους οποίους χωρίζει τον τανυστή χαρακτηριστικών ισότιμα σε υποπεριοχές μιας συγκεκριμένης κλίμακας. Στη μέθοδο PAM, οι δύο χάρτες χαρακτηριστικών $X^{(1)}$, $X^{(2)} \in \mathbb{R}^{C \times H \times W}$, μετασχηματίζονται σε έναν τανυστή χαρακτηριστικών $X \in \mathbb{R}^{C \times H \times W \times 2}$. Έπειτα, υπάρχουν τέσσερις παράλληλοι κλάδοι, καθένας από τους οποίους χωρίζει τον τανυστή ισομερώς σε $s \times s$ υποπεριοχές, όπου $s \in S = \{1, 2, 4, 8\}$ (το S ορίζει τέσσερις κλίμακες πυραμίδας). Στον κλάδο της κλίμακας s , κάθε περιοχή ορίζεται ως $R_{s, i, j} \in \mathbb{R}^{C \times (H/s) \times (W/s) \times 2}$, $1 \leq i, j \leq s$. Έπειτα, εφαρμόζονται τέσσερις BAMs στους τέσσερις κλάδους χωριστά. Μέσα σε κάθε κλάδο της πυραμίδας εφαρμόζεται η BAM σε όλες τις υποπεριοχές $R_{s, i, j}$ ξεχωριστά για να δημιουργηθεί ένας

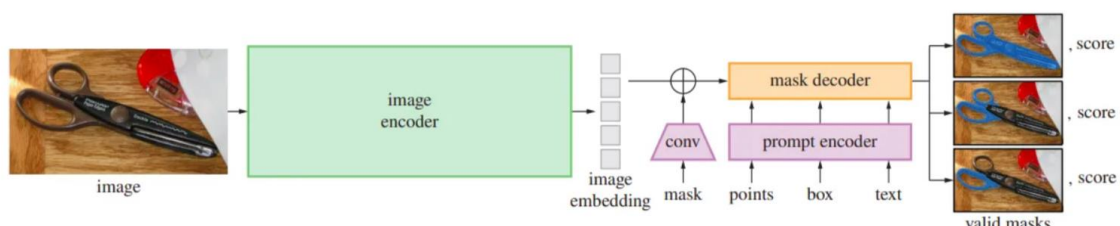
νέος ενημερωμένος residual (υπολειμματικός) τανυστής χαρακτηριστικών $Y_s \in \mathbb{R}^{C \times H \times W \times 2}$. Στη συνέχεια, συγκαταλέγονται αυτοί οι τανυστές χαρακτηριστικών Y_s , $s \in S$ και τροφοδοτούνται σε ένα επίπεδο συνέλιξης ($C, 1 \times 1 / 1$) για να δημιουργηθεί ο τελικός residual (υπολειμματικός) τανυστής χαρακτηριστικών $Y \in \mathbb{R}^{C \times H \times W \times 2}$. Τέλος, προστίθεται ο residual (υπολειμματικός) τανυστής Y και ο αρχικός τανυστής X ώστε να παραχθεί ο ενημερωμένος τανυστής $Z \in \mathbb{R}^{C \times H \times W \times 2}$.

Το συγκεκριμένο νευρωνικό δίκτυο, έχει εκπαιδευτεί στο σύνολο δεδομένων LEVIR-CD. Το LEVIR-CD είναι ένα σετ δεδομένων μεγάλης κλίμακας που αφορά το αντικείμενο της δυαδικής ανίχνευσης μεταβολών (binary change detection). Αποτελείται από 637 πολύ υψηλής ανάλυσης ζεύγη RGB εικόνων από το Google Earth μεγέθους 1024pix x 1024pix, με μέγεθος εδαφοψηφίδας 0.5m, που έχουν ληφθεί με διαφορά 5 έως 14- 25 ετών και συνοδεύονται από τον αληθή (ground truth) χάρτη μεταβολών (change map). Οι εικόνες απεικονίζουν 20 διαφορετικές περιοχές στο Τέξας των ΗΠΑ από το 2002 έως το 2018. Περιέχουν αλλαγές λόγω των διαφορετικών εποχών και λόγω της διαφορετικής φωτεινότητας. Εστιάζει στην μεταβολή διάφορων ειδών κτηρίων, όπως βίλες και αποθήκες, ψηλά κτήρια ή μικρά γκαράζ. Οι μεταβολές έχουν επισημανθεί από ειδικούς φωτερμηνευτές, χρησιμοποιώντας την τιμή 1 ως επικέτα για την αλλαγή και την τιμή 0 για την μη μεταβολή. Συνολικά το σετ δεδομένων αποτελείται από 31333 περιπτώσεις μεταβολών (<https://justchenhao.github.io/LEVIR/>).

2.10 ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ SAM (Segment Anything Model)

Το Segment Anything είναι ένα μοντέλο τμηματοποίησης εικόνων που αναπτύχθηκε από την Meta AI. Το SAM παράγει μάσκες αντικειμένων υψηλής ποιότητας από εισόδους, και μπορεί να χρησιμοποιηθεί για τη δημιουργία масκών για όλα τα αντικείμενα μιας εικόνας. Έχει εκπαιδευτεί σε ένα σύνολο δεδομένων 11 εκατομμυρίων εικόνων και 1.1 δισεκατομμυρίων масκών (είναι pre-trained).

Στις εργασίες ανίχνευσης αντικειμένων, τα αντικείμενα αναπαρίστανται με πλαίσια οριοθέτησης, τα οποία είναι σαν να σχεδιάζεται ένα ορθογώνιο γύρω από το αντικείμενο. Αυτά τα ορθογώνια δίνουν μια γενική ιδέα της θέσης του αντικειμένου, αλλά δεν δείχνουν το ακριβές σχήμα του. Μπορεί επίσης να περιλαμβάνουν τμήματα του φόντου ή άλλα αντικείμενα μέσα στο ορθογώνιο, καθιστώντας δύσκολο τον διαχωρισμό των αντικειμένων από το περιβάλλον τους. Οι μάσκες κατάτμησης, από την άλλη πλευρά, είναι σαν να σχεδιάζεται ένα λεπτομερές περίγραμμα γύρω από το αντικείμενο, ακολουθώντας το ακριβές σχήμα του. Αυτό επιτρέπει την ακριβέστερη κατανόηση του σχήματος, του μεγέθους και της θέσης του αντικειμένου.



Εικόνα 12: Η αρχιτεκτονική του νευρωνικού δικτύου SAM.

Το SAM αποτελείται από τρία στοιχεία, που απεικονίζονται στην Εικόνα 12, έναν κωδικοποιητή εικόνας (image encoder), έναν ευέλικτο κωδικοποιητή (flexible prompt encoder), και έναν αποκωδικοποιητή μάσκας (fast mask decoder).

- *Κωδικοποιητής εικόνας (image encoder):*

Ο κωδικοποιητής εικόνας αποτελεί τον πυρήνα της αρχιτεκτονικής του SAM, ένα εξελιγμένο στοιχείο που είναι υπεύθυνο για την επεξεργασία και τον μετασχηματισμό των εικόνων εισόδου σε ένα ολοκληρωμένο σύνολο χαρακτηριστικών. Χρησιμοποιώντας έναν αυτοκωδικοποιητή μάσκας, προ-εκπαιδευμένο μετασχηματιστή όρασης (ViT Vision Transformer), συμπιέζει τις εικόνες σε έναν πυκνό πίνακα χαρακτηριστικών. Μέσω αυτού του πίνακα το μοντέλο αναγνωρίζει διάφορα στοιχεία της εικόνας.

- *Κωδικοποιητής προτροπής (prompt encoder):*

Ο κωδικοποιητής προτροπής είναι μια μοναδική πτυχή του SAM που το διαφοροποιεί από τα παραδοσιακά μοντέλα τμηματοποίησης εικόνας. Ερμηνεύει διάφορες μορφές input prompts (προτροπών εισόδου), είτε αυτές βασίζονται σε κείμενο, είτε σε σημεία, είτε σε μάσκες, είτε σε συνδυασμό αυτών. Οι μάσκες ενσωματώνονται με τη χρήση συνελίξεων και αθροίζονται στοιχειομετρικά με την ενσωμάτωση εικόνας (image embedding). Αυτό επιτρέπει στο μοντέλο να εστιάζει σε συγκεκριμένες περιοχές ή αντικείμενα εντός μιας εικόνας.

- *Αποκωδικοποιητής μάσκας (fast mask decoder):*

Στον αποκωδικοποιητή μάσκας πραγματοποιείται η κατάτμηση. Συνθέτει τις πληροφορίες τόσο από τον κωδικοποιητή εικόνας όσο και από τον κωδικοποιητή προτροπής για να παράγει ακριβείς μάσκες τμηματοποίησης. Είναι υπεύθυνος για την τελική έξοδο, καθορίζοντας τα ακριβή περιγράμματα και τις περιοχές κάθε τμήματος εντός της εικόνας. Ο κωδικοποιητής εικόνας δημιουργεί πρώτα μια λεπτομερή κατανόηση ολόκληρης της εικόνας, αναλύοντάς την σε χαρακτηριστικά που μπορεί να αναλύσει η μηχανή. Στη συνέχεια, ο κωδικοποιητής προτροπής προσθέτει πλαίσιο, εστιάζοντας την προσοχή του μοντέλου με βάση την παρεχόμενη είσοδο, είτε πρόκειται για ένα απλό σημείο είτε για μια σύνθετη περιγραφή κειμένου. Τέλος, ο αποκωδικοποιητής μάσκας χρησιμοποιεί αυτές τις συνδυασμένες πληροφορίες για την ακριβή κατάτμηση της εικόνας, διασφαλίζοντας ότι η έξοδος ευθυγραμμίζεται με την πρόθεση της προτροπής εισόδου.

ΚΕΦΑΛΑΙΟ 3: ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΦΟΡΑ

3.1 ΣΧΕΤΙΚΑ ΕΡΕΥΝΗΤΙΚΑ ΑΡΘΡΑ

Αφορμή για την προσωπική μου έρευνα στο συγκεκριμένο θέμα ήταν οι πρόσφατες εξελίξεις και τα πορίσματα ερευνητών στον τομέα της ανίχνευσης μεταβολών με διάφορες τεχνικές.

Ορισμένες πρόσφατες μελέτες οι οποίες με προέτρεψαν να ασχοληθώ με το θέμα είναι οι εξής:

- *Change Detection in VHR Imagery With Severe Co-Registration Errors Using Deep Learning: A Comparative Study* (Viktoria Kristollari and Vassilia Karathanassi)

Σύμφωνα με το συγκεκριμένο άρθρο, οι προηγούμενες έρευνες έχουν αναπτύξει αρκετές αξιόλογες μεθοδολογίες Deep Learning (DL), αλλά έχουν επικεντρωθεί κυρίως σε εικόνες με μικρά σφάλματα συνταύτισης (co-registration). Με βάση αυτό, ο στόχος αυτής της μελέτης είναι να αξιολογήσει την απόδοση πέντε σύγχρονων μεθόδων DL ανίχνευσης μεταβολών, δύο μη επιβλεπόμενων και τριών επιβλεπόμενων, σε εικόνες υψηλής ανάλυσης (VHR) με μεγάλα σφάλματα συνταύτισης (co-registration). Οι μέθοδοι εφαρμόζονται σε τέσσερις αστικές ευρωπαϊκές περιοχές με διαφορετική μορφολογία. Επιπλέον, πριν από την εφαρμογή της διαδικασίας CD, αξιολογήθηκαν τέσσερις μέθοδοι αυτόματης συνταύτισης λόγω της σημασίας αυτού του βήματος προεπεξεργασίας για την επιτυχή έκβαση του προβλήματος. Αποδείχθηκε ότι η συσχέτιση φάσης που χρησιμοποιήθηκε στο μετασχηματισμό Fourier-Mellin παρήγαγε τα πιο ικανοποιητικά αποτελέσματα συνταύτισης και ότι το νευρωνικό δίκτυο STANet εντόπισε αλλαγές που σχετίζονται με κτήρια με μεγαλύτερη επιτυχία.

- *A Spatial-Temporal Attention-Based Method and a New Dataset for Remote Sensing Image Change Detection* (HaoChen and Zhenwei Shi)

Στο συγκεκριμένο άρθρο, παρουσιάζεται ένα νέο νευρωνικό δίκτυο χωροχρονικής προσοχής με βάση τη σιαμαία δομή. Σε αντίθεση με τις προηγούμενες μεθόδους που κωδικοποιούν ξεχωριστά τις εικόνες χωρίς να αναφέρονται σε οποιαδήποτε χωροχρονική εξάρτηση, αυτοί οι ερευνητές σχεδιάζουν έναν μηχανισμό αυτοπροσοχής CD για τη μοντελοποίηση των χωροχρονικών σχέσεων. Ενσωματώνουν μια νέα μονάδα αυτοπροσοχής CD στη διαδικασία εξαγωγής χαρακτηριστικών, η οποία υπολογίζει τα βάρη προσοχής μεταξύ δύο οποιωνδήποτε εικονοστοιχείων σε διαφορετικές χρονικές στιγμές και θέσεις και τα χρησιμοποιεί για τη δημιουργία πιο διακριτών χαρακτηριστικών. Λαμβάνοντας υπόψη ότι το αντικείμενο μπορεί να έχει διαφορετικές κλίμακες, το δίκτυο χωρίζει την εικόνα σε υποπεριοχές πολλαπλών κλιμάκων και εισάγουν την αυτοπροσοχή σε κάθε υποπεριοχή. Με αυτόν τον τρόπο, συλλαμβάνει τις χωροχρονικές εξαρτήσεις σε διάφορες κλίμακες, δημιουργώντας έτσι καλύτερες αναπαραστάσεις για αντικείμενα διαφόρων μεγεθών.

- *Segment Anything* (Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollar, Ross Girshick)

Στο συγκεκριμένο άρθρο, παρουσιάζεται το Segment Anything (SA): ένα νέο, μοντέλο και σύνολο δεδομένων για την τμηματοποίηση εικόνων. Χρησιμοποιώντας το αποτελεσματικό αυτό μοντέλο σε έναν βρόχο συλλογής δεδομένων, δημιουργήθηκε το μεγαλύτερο σύνολο δεδομένων τμηματοποίησης μέχρι σήμερα, με πάνω από 1 δισεκατομμύριο μάσκες, σε 11M εικόνες. Το μοντέλο σχεδιάστηκε και εκπαιδεύτηκε ώστε να μπορεί να εξάγει μάσκες τμηματοποίησης σε νέες εικόνες και εργασίες. Οι ερευνητές αξιολογούν τις δυνατότητές του σε πολυάριθμες εργασίες και διαπιστώνουν ότι οι επιδόσεις του σε είναι εντυπωσιακές - συχνά ανταγωνιστικές ή και ανώτερες από προηγούμενα αποτελέσματα με πλήρη επίβλεψη.

- *Land-Use Change Detection with Convolutional Neural Network Methods* (Cong Cao, Suzana Dragičević, and Songnian Li)

Αυτή η μελέτη στοχεύει να αξιολογήσει την απόδοση των μεθόδων CNN για ταξινόμηση και να εντοπίσει αλλαγές χρήσεων γης. Οκτώ μοντέλα CNN αξιολογήθηκαν πλήρως σε δεδομένα τηλεπισκόπησης για ταξινόμηση εικόνων με αλλαγή χρήσης γης χρησιμοποιώντας τρία προεκπαιδευμένα μοντέλα CNN, το AlexNet, το GoogLeNet και το VGGNet.

ΚΕΦΑΛΑΙΟ 4: ΜΕΘΟΔΟΛΟΓΙΑ

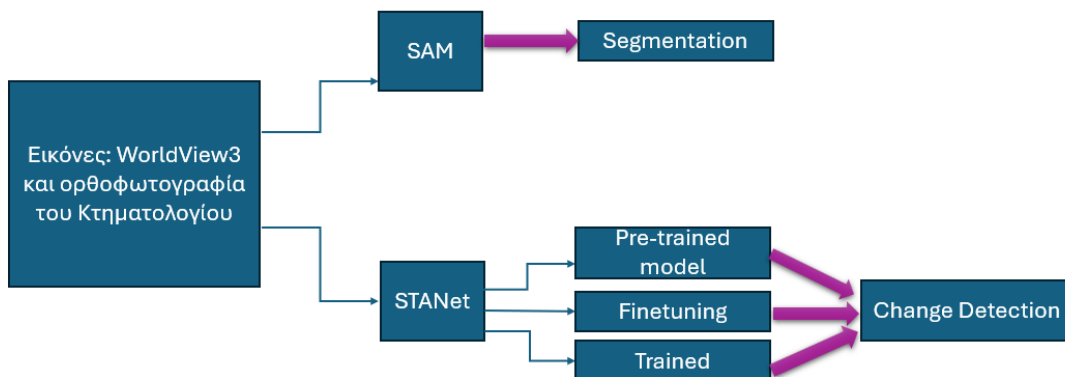
4.1 ΨΗΦΙΑΚΟ ΠΕΡΙΒΑΛΛΟΝ

Κατά την εφαρμογή χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python, η οποία διαθέτει μια σειρά από χαρακτηριστικά που σπάνια είναι διαθέσιμα σε άλλες γλώσσες προγραμματισμού. Λόγω του απλού συντακτικού, της βασικής ροής ελέγχου και της ευκολίας χρήσης βιβλιοθηκών και δομών δεδομένων, η Python είναι η καλύτερη γλώσσα για την ανάπτυξη αλγορίθμων τεχνητής νοημοσύνης. Το περιβάλλον που χρησιμοποιήθηκε είναι το Google Colab και η έκδοση της python, η οποία εγκαταστάθηκε είναι η 3.10.12. Κύριες βιβλιοθήκες που αξιοποιήθηκαν είναι: η NumPy, η Gdal, η Dominate και η PyTorch.

4.2 ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ ΔΙΑΔΙΚΑΣΙΑΣ

Στην παρούσα διπλωματική εργασία θα παρουσιαστούν δύο πειράματα-εκδοχές ανίχνευσης μεταβολών αστικού ιστού, οι οποίες παρουσιάζονται συνοπτικά στα παρακάτω διαγράμματα ροής εργασιών:

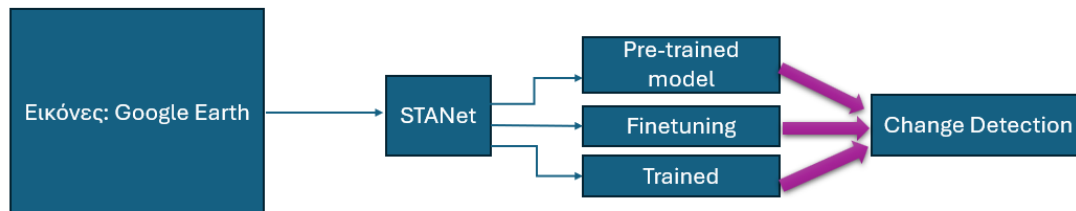
1^η εκδοχή:



Στην 1^η εκδοχή, αξιοποιήθηκαν δύο εικόνες υψηλής ανάλυσης, μια ορθοφωτογραφία του 2007, την οποία παρείχε το ΕΛΛΗΝΙΚΟ ΚΤΗΜΑΤΟΛΟΓΙΟ με pixel size 0.2m και μια δορυφορική εικόνα WorldView-3 του 2023 με pixel size 0.3m. Το πρώτο στάδιο είναι η εφαρμογή του δικτύου SAM, με σκοπό την κατάτμηση των εικόνων. Συγκεκριμένα, η κατάτμηση αυτή, δηλαδή η δημιουργία ψηφιακών εικόνων, στις οποίες θα ξεχώριζαν, σαν αντικείμενα, τα κτήρια, θα συνέβαλε στην καλύτερη αξιολόγηση της ανίχνευσης μεταβολών και ταυτόχρονα θα αποτελούσε έναν πιο αυτοματοποιημένο τρόπο παραγωγής label, τα οποία θα χρησίμευαν στην εφαρμογή του δικτύου STANet σε επόμενο βήμα. Επίσης, σκοπός ήταν η αφαίρεση των αποτελεσμάτων έπειτα από την εφαρμογή του δικτύου SAM στις δύο εικόνες, προκειμένου να προέκυπταν οι μεταβολές και να συγκρίνονταν με τα αποτελέσματα του δικτύου STANet. Όπως θα παρουσιαστούν και παρακάτω, τα αποτελέσματα δεν ήταν τα επιθυμητά. Έπειτα, ακολούθησε η εφαρμογή του δικτύου STANet, με σκοπό την ανίχνευση των μεταβολών του αστικού ιστού. Αρχικά, εφαρμόστηκε το προ-εκπαιδευμένο (pre-trained) μοντέλο,

έπειτα το μοντέλο, στο οποίο εφαρμόστηκε βελτιστοποίηση (finetuning) και τέλος έγινε εξ' αρχής εκπαίδευση του μοντέλου μόνο με τα δεδομένα της WorldView-3 και της ορθοφωτογραφίας του Κτηματολογίου. Όλες οι παραπάνω δοκιμές δεν παρουσίασαν τα επιθυμητά αποτελέσματα. Αξίζει να σημειωθεί ότι τα αποτελέσματα αξιολογήθηκαν και συγκρίθηκαν, με βάση εικόνες που προέκυψαν έπειτα από ενδελεχή φωτοερμηνεία.

2^η εκδοχή:



Στην 2^η εκδοχή, λήφθηκαν μέσω Google Earth 24 ζεύγη εικόνων, 14 από την περιοχή της Μυκόνου και 10 από την περιοχή της Πάρου, μιας και τα κτήρια ήταν αρκετά παρόμοια. Δημιουργήθηκε, έτσι ένα νέο σύνολο δεδομένων (dataset), μόνο με μεταβολές αστικού ιστού, (ενώ το σύνολο δεδομένων της 1^{ης} εκδοχής περιείχε και εικόνες που δεν παρουσίαζαν μεταβολές αστικού ιστού). Επομένως, σε αυτό το σύνολο δεδομένων εφαρμόστηκε η ίδια διαδικασία με την 1^η εκδοχή που αφορά στην εφαρμογή του δικτύου STANet, με σκοπό την ανίχνευση των μεταβολών αστικού ιστού. Αρχικά, εφαρμόστηκε το προ-εκπαιδευμένο (pre-trained) μοντέλο, έπειτα το μοντέλο, στο οποίο εφαρμόστηκε βελτιστοποίηση (finetuning) και τέλος έγινε εξ' αρχής εκπαίδευση του μοντέλου μόνο με τα δεδομένα των εικόνων Google Earth. Τα αποτελέσματα του pre-trained μοντέλου δεν ήταν αποδεκτά, ενώ τα αποτελέσματα από τη διαδικασία finetuning και της εκπαίδευσης του δικτύου με τα δεδομένα Google Earth ήταν αρκετά ικανοποιητικά. Αξίζει να σημειωθεί ότι τα αποτελέσματα αξιολογήθηκαν και συγκρίθηκαν, με βάση εικόνες που προέκυψαν έπειτα από ενδελεχή φωτοερμηνεία.

ΚΕΦΑΛΑΙΟ 5: ΕΦΑΡΜΟΓΗ ΣΤΗΝ ΠΕΡΙΟΧΗ ΜΕΛΕΤΗΣ

5.1 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

5.1.1 ΕΚΔΟΧΗ 1

Για την εισαγωγή των δεδομένων στα δίκτυα, χρειάστηκε αρχικά η επεξεργασία τους.

- Αρχικά, όσον αφορά στα δεδομένα της 1^{ης} εκδοχής ακολουθήθηκε η παρακάτω διαδικασία:

Όπως προαναφέρθηκε, τα δεδομένα της 1^{ης} εκδοχής, της παρούσας διπλωματικής, είναι δύο εικόνες, μία ορθοφωτογραφία και μία δορυφορική εικόνα, στις οποίες απεικονίζεται περιοχή της Μυκόνου. Η ορθοφωτογραφία έχει ημερομηνία λήψης το 2007 και pixel size 0.2m (την παρείχε το ΕΛΛΗΝΙΚΟ ΚΤΗΜΑΤΟΛΟΓΙΟ) και η δορυφορική εικόνα WorldView-3 έχει ημερομηνία λήψης το 2023 και pixel size 0.3m.

Σκοπός ήταν οι ημερομηνίες λήψης του ζεύγους δορυφορικών εικόνων να διαφέρουν αρκετά, προκειμένου, να υπάρχουν σημαντικές μεταβολές στον αστικό ιστό.

Αρχικά, ήταν απαραίτητη η περικοπή των δύο εικόνων στην ίδια περιοχή.

Κατά την επεξεργασία, ορίστηκε το ίδιο σύστημα αναφοράς και συγκεκριμένα το ΕΓΣΑ'87, σε κάθε ζεύγος εικόνων. Έπειτα, ακολούθησε η διαδικασία resampling, κατά την οποία η δορυφορική εικόνα με την καλύτερη ανάλυση, δηλαδή η εικόνα του 2007, υπέστη διαδικασία μείωσης ανάλυσης και από 0.2m/pixel, η ανάλυση pixel έγινε 0.3m/pixel, ίδια δηλαδή με την ανάλυση της εικόνας του 2023, προκειμένου όλες οι διαδικασίες να γίνονται σε δεδομένα με ίδια ανάλυση και τα αποτελέσματα να είναι συγκρίσιμα. Ταυτόχρονα, χρειάστηκε η μετατροπή της εικόνας του 2023 από 16-bit σε 8-bit, καθώς και η κανονικοποίηση των τιμών των εικονοστοιχείων της από 0 έως 255 και η αφαίρεση του καναλιού NIR, προκειμένου στις παρακάτω διαδικασίες οι δύο εικόνες να έχουν ίδια κανάλια.

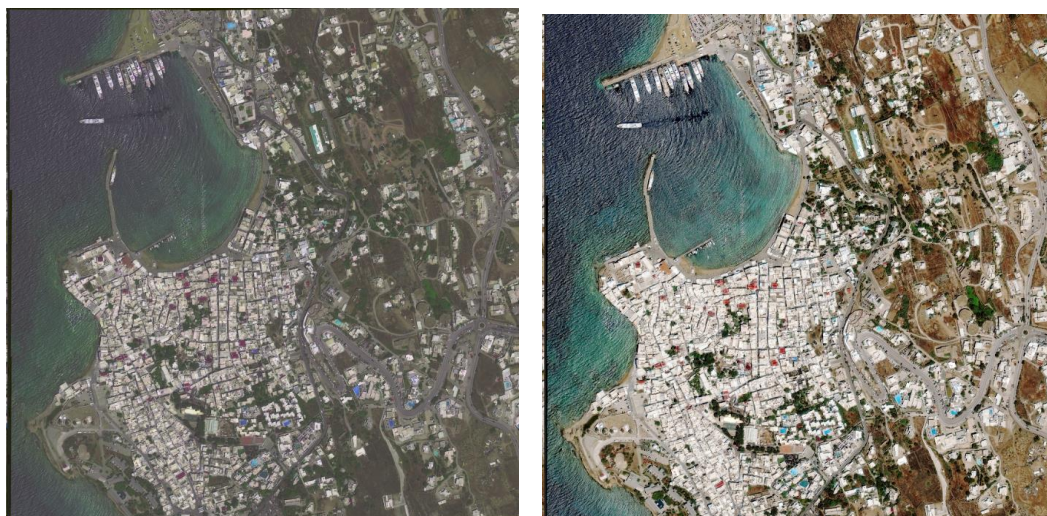
Επίσης, κρίθηκε αναγκαία η συνταύτιση των εικόνων, μιας και η εικόνα του 2007 είναι ορθοφωτογραφία, ενώ η εικόνα του 2023 δορυφορική εικόνα με παραμορφώσεις υψομέτρου, άρα η διαστρέβλωση δεν είναι ίδια σε όλη την εικόνα. Επομένως, προκειμένου να αποφευχθούν όσο το δυνατόν περισσότερο λάθη στην ανίχνευση των αλλαγών του αστικού ιστού, λόγω διαφοράς γωνίας θέασης, ακολούθησε η διαδικασία co-registration (η οποία αξιοποιεί τον μετασχηματισμό Fourier). Η διαδικασία αυτή έγινε με διάφορους τρόπους ώστε να επιτευχθεί το βέλτιστο αποτέλεσμα. Αρχικά, μέσω του plugin του QGIS, έγιναν δοκιμές με τους 4 διαφορετικούς αλγορίθμους που διαθέτει: 1) basic pixel alignment, 2) ranning pixel adjustment, 3) global και 4) local. Τα αποτελέσματα όμως δεν ήταν ικανοποιητικά. Έπειτα, ακολούθησε μία αυτοματοποιημένη διαδικασία, κατά την οποία δημιουργήθηκε κώδικας συν-ταύτισης, αλλά και πάλι τα αποτελέσματα δεν ήταν τα επιθυμητά. Επομένως, ακολούθησε μια μη- αυτοματοποιημένη διαδικασία co-registration, κατά την οποία λήφθηκαν σημεία GCP, αφού προηγουμένως κόπηκαν οι εικόνες σε εννιά τμήματα με επικάλυψη, προκειμένου να επιτευχθεί καλύτερο αποτέλεσμα. Η παραπάνω διαδικασία έγινε στο

QGIS με την εντολή Georeferencer και λήφθηκαν περίπου 25-30 σημεία σε καθένα από τα εννιά τμήματα. Επίσης, στα επικαλυπτόμενα τμήματα λήφθηκαν κοινά σημεία. Στις διαθέσιμες επιλογές μετασχηματισμών, επιλέχθηκε μετά από διάφορες δοκιμές, ο προβολικός μετασχηματισμός (projective).

Στη συνέχεια, τα εννιά επικαλυπτόμενα τμήματα κάθε εικόνας ενώθηκαν και προέκυψαν δύο εικόνες, οι οποίες συν- ταυίζονταν σε αρκετά καλό βαθμό. Οι δύο εικόνες, όμως είχαν αρκετά μεγάλες διαφορές όσον αφορά στα χρώματα. Η εικόνα του 2007 είχε πιο κίτρινες αποχρώσεις, ενώ η εικόνα του 2023 πιο μπλε αποχρώσεις. Επομένως, για να μην επηρεαστεί το αποτέλεσμα της ανίχνευσης μεταβολών από τη διαφορά χρωμάτων, ακολούθησε η διαδικασία color-match, μέσω κώδικα στην python, κατά την οποία οι αποχρώσεις της εικόνας του 2023, προσαρμόστηκαν με βάση τις χρωματικές αποχρώσεις της εικόνας του 2007. Οι μέθοδοι πίσω από τις αντιστοιχίσεις βασίζονται στην προσέγγιση των Reinhard et al., στη γραμμικοποίηση Monge-Kantorovich (MKL) όπως προτάθηκε από τους Pitie et al. και σε μια αναλυτική λύση για μια μεταφορά πολλαπλών μεταβλητών κατανομής Gauss (MVGD) σε συνδυασμό με την κλασική αντιστοίχιση ιστογράμματος.

Έγιναν δοκιμές με όλες τις διαθέσιμες μεθόδους, την Monge-Kantorovich (MKL), την Reinhard, την Multi-Variate Gaussian Distribution (MVGD), την histogram match (HM) καθώς και με την Monge-Kantorovich και την Multi-Variate Gaussian Distribution σε συνδυασμό με αντιστοίχιση ιστογράμματος (hm-mkl-hm και την hm-mvgd-hm). Το καλύτερο αποτέλεσμα το είχε η μέθοδος Monge-Kantorovich σε συνδυασμό με αντιστοίχιση ιστογράμματος (hm-mkl-hm).

Παρακάτω παρουσιάζονται τα αποτελέσματα:



Εικόνα 13: Αποτέλεσμα μεθόδων MKL (αριστερά) και HM (δεξιά).



Εικόνα 14: Αποτέλεσμα μεθόδων Reinhard (αριστερά) και MVGD (δεξιά).

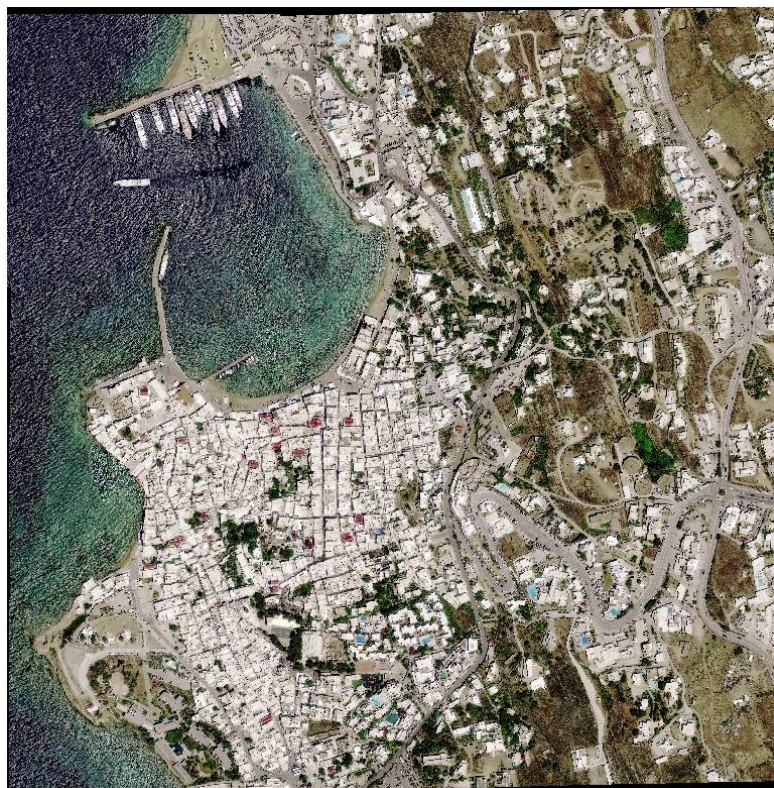


Εικόνα 15: Αποτέλεσμα μεθόδων HM-MVGD-HM (αριστερά) και HM-MKL-HM (δεξιά).

Στη συνέχεια, ακολούθησε η περικοπή των εικόνων σε συγκεκριμένη περιοχή ενδιαφέροντος, μεγέθους 1050m*1050m (3500pix*3500pix), προκειμένου να γίνεται πιο εύκολη και γρήγορη η επεξεργασία τους (οι εικόνες καταλάμβαναν πολύ μεγάλο όγκο δεδομένων). Οι τελικές εικόνες παρουσιάζονται παρακάτω:



Εικόνα 16: Η τελική εικόνα του 2007, έπειτα από την περικοπή στη περιοχή μελέτης.



Εικόνα 17: Η τελική εικόνα του 2023, έπειτα από την περικοπή στη περιοχή μελέτης.

Έπειτα, μέσω κώδικα στην python, έγινε περικοπή της κάθε εικόνας σε patches των 256pix*256pix, μιας και το νευρωνικό δίκτυο STANet λαμβάνει τέτοια δεδομένα.

Επειδή τα δεδομένα ήταν λιγοστά, για το στάδιο της εκπαίδευσης του νευρωνικού δικτύου (169 patches), ακολούθησε η διαδικασία επαύξησης των δεδομένων (augmentation), δηλαδή η δημιουργία νέων μετασχηματισμένων εκδόσεων των αρχικών εικόνων από το δεδομένο σύνολο εικόνων για αύξηση της ποικιλομορφίας του. Επομένως, μέσω κώδικα στην python έγιναν οι εξής διαδικασίες: 1) περιστροφή των αρχικών εικόνων κατά 90° , 2) περιστροφή των αρχικών εικόνων κατά 45° , 3) περιστροφή των αρχικών εικόνων κατά 180° , 4) περιστροφή των αρχικών εικόνων κατά 250° , 5) οριζόντια περιστροφή και 6) κατακόρυφη περιστροφή. Τελικά, από κάθε αρχική εικόνα προέκυψαν άλλα 6 δεδομένα augmented και έτσι το σύνολο των δεδομένων που προέκυψε είναι περίπου 1000 εικόνες, οι οποίες κρίθηκαν αρκετές για την εφαρμογή κάποιου δικτύου.

Στη συνέχεια, με φωτοερμηνεία των εικόνων ψηφιοποιήθηκαν οι μεταβολές αστικού ιστού, είτε αυτές αφορούσαν σε δημιουργία κτηρίων, δηλαδή κτηρίων που εμφανίζονται στην πιο πρόσφατη εικόνα και δεν εμφανίζονται στην παλαιότερη, είτε αφορούσαν σε κατεδάφιση κτηρίων, δηλαδή κτηρίων που δεν εμφανίζονται στην πιο πρόσφατη εικόνα και εμφανίζονται στην παλαιότερη. Τα πολύγωνα αυτά μετατράπηκαν σε binary raster αρχεία με τιμές 0 και 1. Με τιμή 1 απεικονίζονται τα pixels με τις μεταβολές που έχει εντοπίσει το δίκτυο στον αστικό ιστό, δηλαδή στα κτήρια, μεταξύ των δύο εικόνων, ενώ τα pixels όλης της υπόλοιπης μάσκας θα έχουν τιμή 0. Τα παραπάνω αποτελούν τα labels με βάση τα οποία αξιολογήθηκαν τα αποτελέσματα που προέκυψαν από το νευρωνικό δίκτυο.

5.1.2 ΕΚΔΟΧΗ 2

- Όσον αφορά στα δεδομένα της 2^{ης} εκδοχής ακολουθήθηκε η παρακάτω διαδικασία:

Τα δεδομένα, όπως προαναφέρθηκε ήταν 24 ζεύγη δορυφορικών εικόνων, τα οποία λήφθηκαν από το Google Earth, από την περιοχή της Μυκόνου και της Πάρου, σε μέγεθος 1920pix*1080pix και από ύψος περίπου 600-650m. Στον παρακάτω πίνακα, παρουσιάζονται οι ημερομηνίες λήψης των ζευγών, καθώς και ο δορυφόρος από τον οποίο λήφθηκε η κάθε εικόνα:

Πίνακας 1: Ημερομηνίες λήψης και δορυφόροι εικόνων εκδοχής 2

α/α ζεύγους	Περιοχή	Ημερομηνία λήψης εικόνας A	Δορυφόρος εικόνας A	Ημερομηνία λήψης εικόνας B	Δορυφόρος εικόνας B
1	Μύκονος	2018	WorldView	2023	Airbus
2	Μύκονος	2016	CNES/Airbus	2022	WorldView
3	Μύκονος	2018	WorldView	2023	Airbus
4	Μύκονος	2017	WorldView	2023	Airbus
5	Μύκονος	2021	WorldView	2023	Airbus
6	Μύκονος	2020	CNES/Airbus	2023	Airbus
7	Μύκονος	2018	WorldView	2023	Airbus
8	Μύκονος	2019	WorldView	2023	Airbus
9	Μύκονος	2016	CNES/Airbus	2023	Airbus
10	Μύκονος	2018	WorldView	2023	Airbus
11	Μύκονος	2018	WorldView	2023	Airbus
12	Μύκονος	2017	WorldView	2023	Airbus
13	Μύκονος	2016	CNES/Airbus	2023	Airbus
14	Μύκονος	2016	CNES/Airbus	2023	Airbus
15	Πάρος	2016	CNES/Airbus	2023	Airbus
16	Πάρος	2016	CNES/Airbus	2023	Airbus
17	Πάρος	2017	WorldView	2023	Airbus
18	Πάρος	2016	CNES/Airbus	2023	Airbus
19	Πάρος	2016	CNES/Airbus	2023	Airbus
20	Πάρος	2016	CNES/Airbus	2023	Airbus
21	Πάρος	2016	CNES/Airbus	2023	Airbus
22	Πάρος	2016	CNES/Airbus	2023	Airbus
23	Πάρος	2016	CNES/Airbus	2023	Airbus
24	Πάρος	2016	CNES/Airbus	2023	Airbus

Έπειτα, ακολούθησε η διαδικασία co-registration (η οποία αξιοποιεί τον μετασχηματισμό Fourier). Επιλέχθηκε η μη- αυτοματοποιημένη διαδικασία co-registration, κατά την οποία λήφθηκαν σημεία GCP, μιας και από την 1^η εκδοχή αποδείχθηκε ότι αυτή είχε το καλύτερο αποτέλεσμα συν-ταύτισης. Η παραπάνω διαδικασία έγινε στο QGIS με την εντολή Georeferencer και λήφθηκαν περίπου 15-20 σημεία σε κάθε ζεύγος. Στις διαθέσιμες επιλογές μετασχηματισμών, επιλέχθηκε μετά από διάφορες δοκιμές, ο προβολικός μετασχηματισμός (projective). Οι εικόνες συν-ταυτίζονταν σε αρκετά καλό βαθμό. Στη συνέχεια, οι εικόνες περικόπηκαν σε patches












των 256pix*256pix. Από αυτά απομονώθηκαν μόνο τα patches με τις μεταβολές αστικού ιστού. Συνολικά προέκυψαν περίπου 400 patches μόνο με μεταβολές αστικού ιστού.

Ακριβώς, όπως και στην εκδοχή 1 ακολούθησε η ίδια διαδικασία επαύξησης των δεδομένων (augmentation) και δημιουργίας των labels.

5.2 ΕΦΑΡΜΟΓΗ ΕΚΔΟΧΗΣ 1

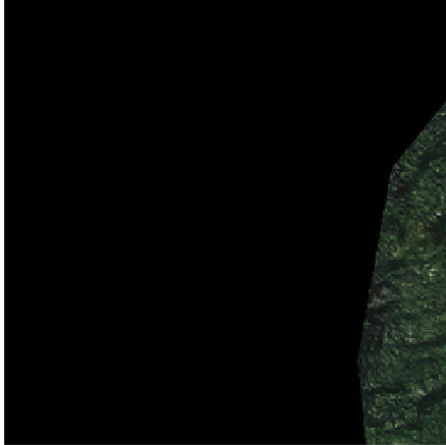

5.2.1 ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΕΙΚΟΝΩΝ ΜΕ ΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ SAM

Το δίκτυο SAM τμηματοποιεί τις εικόνες, παράγοντας μάσκες και συγκεκριμένα μία μάσκα για κάθε αντικείμενο που αναγνωρίζει. Όλες αυτές οι οντότητες που αναγνωρίστηκαν από το δίκτυο, μετατράπηκαν σε vector αρχεία και ταξινομήθηκαν σε κατηγορίες χειροκίνητα. Όμως, δεν κατάφερε το δίκτυο, ειδικά στον πυκνό αστικό ιστό, να αναγνωρίσει όλες τις οντότητες και έτσι έγινε η ταξινόμηση για μια λωρίδα της εικόνας, προκειμένου να προκύψουν ορισμένα στατιστικά, αλλά δεν ακολούθησε η παραγωγή των labels από αυτά τα αποτελέσματα. Η λωρίδα (γραμμή κατά μήκος της εικόνας) που επιλέχθηκε είναι αυτή που απεικονίζει περιοχή περίπου στο Κάστρο της Μυκόνου και περιλαμβάνει patches με αραιό, αλλά και με πυκνό αστικό ιστό. Οι κατηγορίες στις οποίες ταξινομήθηκαν οι οντότητες χειροκίνητα είναι οι εξής βασικές:

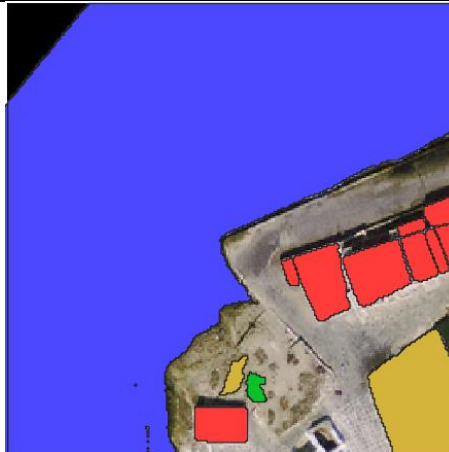
- | | |
|--|--|
| <input checked="" type="checkbox"/>  thalassa_2007 | <input checked="" type="checkbox"/>  kthrio_2023 |
| <input checked="" type="checkbox"/>  kthrio_2007 | <input checked="" type="checkbox"/>  vlastisi_2023 |
| <input checked="" type="checkbox"/>  pisina_2007 | <input checked="" type="checkbox"/>  gymno_edafos_2023 |
| <input checked="" type="checkbox"/>  dromoi_2007 | <input checked="" type="checkbox"/>  thalassa_2023 |
| <input checked="" type="checkbox"/>  gymno_edafos_2007 | <input checked="" type="checkbox"/>  dromoi_2023 |
| <input checked="" type="checkbox"/>  vlastisi_2007 | |

Τα αποτελέσματα που προέκυψαν για τα patches της εικόνας του 2007 και του 2023 παρατίθενται παρακάτω:

Πίνακας 2: Αποτελέσματα SAM για τα patches της εικόνας 2007

Αριθμός patch	Εικόνα 2007 (patch)	Ταξινόμηση αποτελεσμάτων SAM
6_0 (απουσία αστικού ιστού)		

6_1
(αραιός
αστικός
ιστός)



















6_2
(αραιός
αστικός
ιστός)



6_3
(αραιός
αστικός
ιστός)

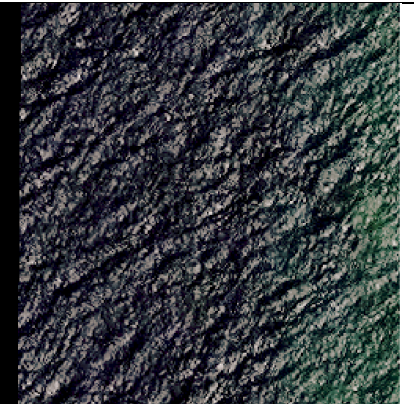
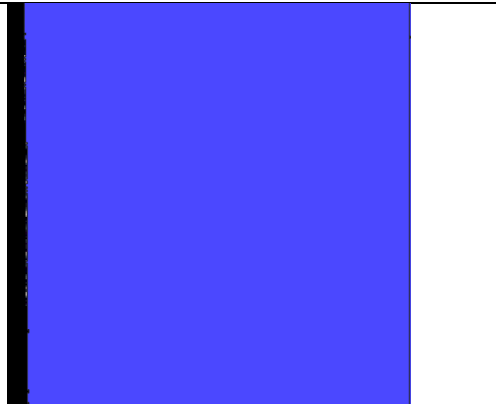









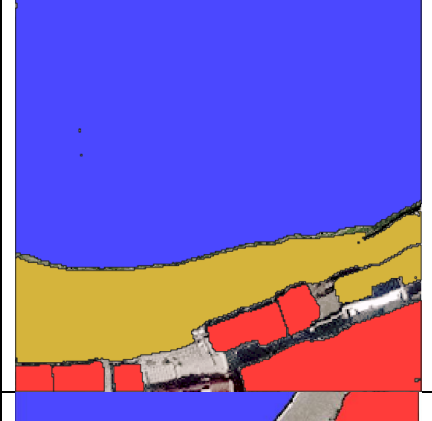


<p>6_4 (αραιός αστικός ιστός)</p>		
<p>6_5 (αραιός αστικός ιστός)</p>		
<p>6_6 (πυκνός αστικός ιστός)</p>		
<p>6_7 (πυκνός αστικός ιστός)</p>		








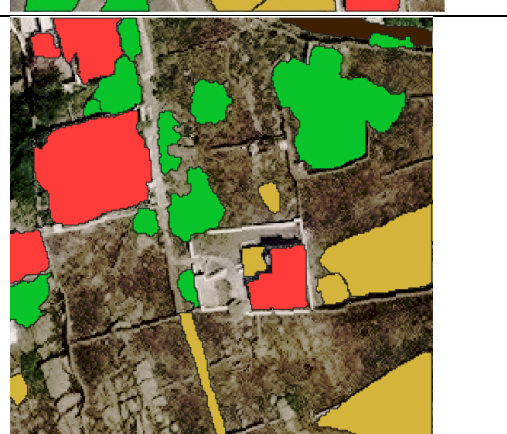
<p>6_8 (αραιός αστικός ιστός)</p>		
<p>6_9 (αραιός αστικός ιστός)</p>		
<p>6_10 (αραιός αστικός ιστός)</p>		
<p>6_11 (αραιός αστικός ιστός)</p>		




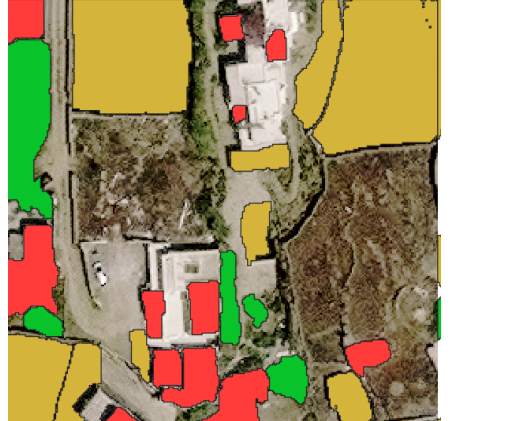




Πίνακας 3: Αποτελέσματα SAM για τα patches της εικόνας 2023

Αριθμός patch	Εικόνα 2023 (patch)	Ταξινόμηση αποτελεσμάτων SAM
6_0 (απουσία αστικού ιστού)		
6_1 (αραιός αστικός ιστός)		

<p>6_2 (αραιός αστικός ιστός)</p>		
<p>6_3 (αραιός αστικός ιστός)</p>		
<p>6_4 (αραιός αστικός ιστός)</p>		
<p>6_5 (αραιός αστικός ιστός)</p>		

<p>6_6 (πυκνός αστικός ιστός)</p>		
<p>6_7 (πυκνός αστικός ιστός)</p>		
<p>6_8 (αραιός αστικός ιστός)</p>		
<p>6_9 (αραιός αστικός ιστός)</p>		

<p>6_10 (αραιός αστικός ιστός)</p>		
<p>6_11 (αραιός αστικός ιστός)</p>		
<p>6_12 (αραιός αστικός ιστός)</p>		

5.2.2 ΑΝΙΧΝΕΥΣΗ ΜΕΤΑΒΟΛΩΝ ΑΣΤΙΚΟΥ ΙΣΤΟΥ ΜΕ ΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ STANET

Όπως προαναφέρθηκε, στην 1η εκδοχή, έπειτα από την εφαρμογή του δικτύου SAM, έγινε εφαρμογή του δικτύου STANet.

Αρχικά, ταξινομήθηκαν τα δεδομένα σε φακέλους με κατάλληλη δομή, έτσι όπως ακριβώς απαιτείται με βάση το άρθρο HaoChen and Zhenwei Shi (2020), “A Spatial-Temporal Attention-Based Method and a NewDataset for Remote Sensing Image Change Detection”. Η δομή είναι η παρακάτω:

Dataset:



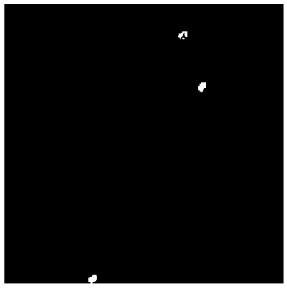




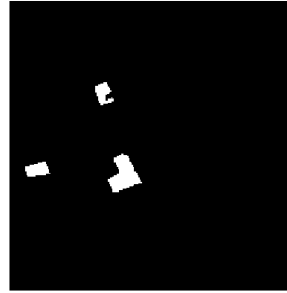
```
├─train
  │
  │   ├──A
  │   ├──B
  │   └─label
  └─val
    │
    │   ├──A
    │   ├──B
    │   └─label
  └─test
    │
    │   ├──A
    │   ├──B
    └─label
```

όπου train ο φάκελος με τα δεδομένα εκπαίδευσης, val ο φάκελος με τα δεδομένα επικύρωσης και test ο φάκελος με τα δεδομένα ελέγχου. Επίσης, σε κάθε φάκελο A ορίζονται τα patches της παλαιότερης εικόνας, σε κάθε φάκελο B ορίζονται τα patches της πιο πρόσφατης εικόνας και σε κάθε φάκελο label ορίζονται οι μάσκες ανίχνευσης αλλαγών αστικού ιστού. Τα δεδομένα χωρίστηκαν στους φακέλους, με ποσοστά: περίπου 60% του συνόλου δεδομένων στο train, περίπου 20% του συνόλου δεδομένων στο val και περίπου 20% του συνόλου δεδομένων στο test.

- Pre-trained

Αρχικά, εφαρμόστηκε το pre-trained μοντέλο, με τα βάρη της μεθόδου PAM, του ήδη εκπαιδευμένου δικτύου STANet στο σύνολο δεδομένων LEVIR-CD. Τα αποτελέσματα δεν ήταν ικανοποιητικά. Ορισμένα ενδεικτικά αποτελέσματα είναι τα εξής:

Πίνακας 4: Αποτελέσματα του pre-trained δικτύου STANet (1η εκδοχή)

α/α	A	B	Label	Αποτέλεσμα pre-trained μοντέλου
1				
2				
3				

Στα αποτελέσματα α/α 1, ενώ δεν παρατηρείται κάποια μεταβολή κτηρίων, και το label, το οποίο έχει δημιουργηθεί, έχει τιμή 0 σε όλα τα rixel, το pre-trained μοντέλο έχει εντοπίσει κάποιες μικρές μεταβολές, οι οποίες δεν αφορούν σε κτήρια, αλλά οφείλονται σε διαφορετική γωνία λήψης των εικόνων.

Στα αποτελέσματα α/α 2, ενώ παρατηρείται μεταβολή και συγκεκριμένα εξάπλωση των υπαρχόντων κτηρίων στην B εικόνα, το δίκτυο έχει ανιχνεύσει κάποιες μεταβολές, οι οποίες όμως δεν αφορούν στον αστικό ιστό. Οι μεταβολές αυτές αφορούν σε ορισμένα αυτοκίνητα που υπάρχουν στο επάνω μέρος της B εικόνας, ενώ δεν υπάρχουν στην A εικόνα.

Στα αποτελέσματα α/α 3, ενώ παρατηρείται μεταβολή και συγκεκριμένα δημιουργία κτηρίων στην εικόνα B, το δίκτυο έχει ανιχνεύσει μεταβολές οι οποίες δεν ισχύουν. Η μία εξ αυτών αφορά στη δημιουργία πισίνας στην εικόνα B.



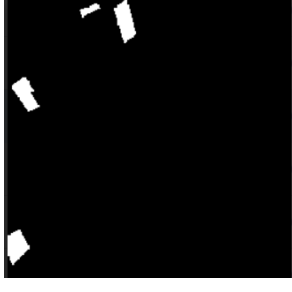
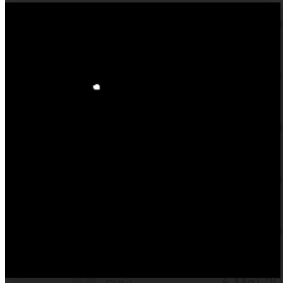


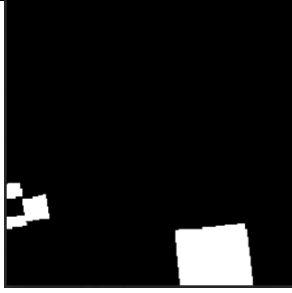



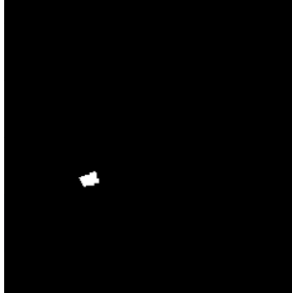
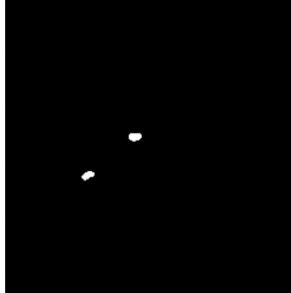
Πιθανότατα, το pre-trained μοντέλο απέτυχε στον εντοπισμό μεταβολών αστικού ιστού, μιας και τα δεδομένα στα οποία έχει εκπαιδευτεί το δίκτυο, δηλαδή στο σύνολο

δεδομένων LEVIR-CD, αποτελούνται από μεταβολές κτηρίων, τα οποία έχουν σκεπές και όχι τσιμεντένιες οροφές όπως τα δεδομένα της Μυκόνου.

- Finetuning

Έπειτα, ακολούθησε η διαδικασία finetuning, κατά την οποία η εκπαίδευση του δικτύου με τα δεδομένα LEVIR-CD, συνεχίστηκε με τα δεδομένα της εκδοχής 1 και με learning rate ίσο με 0.0002. Τα αποτελέσματα δεν ήταν ικανοποιητικά. Ορισμένα ενδεικτικά αποτελέσματα είναι τα εξής:

Πίνακας 5: Αποτελέσματα της διαδικασίας finetuning δικτύου STANet (1η εκδοχή)

α/α	A	B	Label	Αποτέλεσμα finetuning
1				
2				
3				

Στα αποτελέσματα α/α 1, ενώ παρατηρούνται μεταβολές κτηρίων, και συγκεκριμένα δημιουργία κτηρίων στην Β εικόνα, το δίκτυο έχει εντοπίσει μία πολύ μικρή μεταβολή, η οποία δεν αφορά σε μεταβολή κτηρίου, αλλά βλάστησης.

Στα αποτελέσματα α/α 2, ενώ παρατηρούνται μεταβολές κτηρίων, και συγκεκριμένα δημιουργία κτηρίων στην Β εικόνα, το δίκτυο έχει ανιχνεύσει κάποιες μεταβολές. Οι μεταβολές αυτές αφορούν σε ορισμένα αυτοκίνητα που υπάρχουν στη Β εικόνα, ενώ δεν υπάρχουν στην Α εικόνα, καθώς και σε κάποιες μικρές μεταβολές βλάστησης.



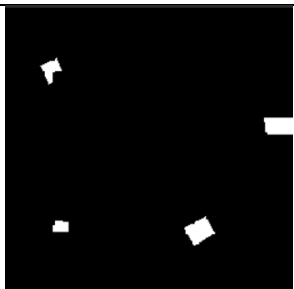







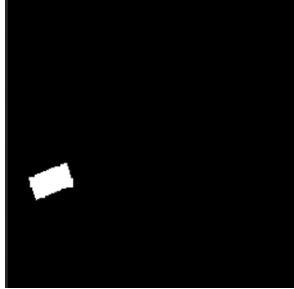
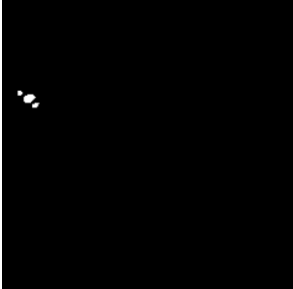
Στα αποτελέσματα α/α 3, παρατηρείται μεταβολή και συγκεκριμένα δημιουργία ενός μικρού κτηρίου στην εικόνα Β. Το δίκτυο έχει ανιχνεύσει τη μεταβολή αυτή, αλλά ταυτόχρονα έχει ανιχνεύσει και μία μεταβολή η οποία αφορά σε ένα αυτοκίνητο, το οποίο υπάρχει στην Β εικόνα, ενώ στην Α δεν υπάρχει.

Πιθανότατα, το finetuning απέτυχε στον εντοπισμό μεταβολών αστικού ιστού, μιας και τα δεδομένα στα οποία έχει εκπαιδευτεί το δίκτυο, δηλαδή στο σύνολο δεδομένων LEVIR-CD, είναι πολύ περισσότερα, σε σχέση με τα επιπλέον δεδομένα της εκδοχής 1, τα οποία προστέθηκαν για να συνεχιστεί η εκπαίδευση.

- Εκπαίδευση

Έπειτα, ακολούθησε η εκπαίδευση του δικτύου μόνο με τα δεδομένα της εκδοχής 1. Τα αποτελέσματα δεν ήταν καλά. Ορισμένα ενδεικτικά αποτελέσματα είναι τα εξής:

Πίνακας 6: Αποτελέσματα της διαδικασίας εκπαίδευσης δικτύου STANet (1η εκδοχή)

α/α	A	B	Label	Αποτέλεσμα εκπαίδευσης
1				
2				
3				

Στα αποτελέσματα α/α 1, ενώ παρατηρούνται μεταβολές κτηρίων, το δίκτυο έχει εντοπίσει ορθά μόνο τη μία μεταβολή, η οποία αφορά στην κατεδάφιση ενός κτηρίου, δηλαδή ενός κτηρίου το οποίο εμφανίζεται στην εικόνα A και στη δημιουργία πισίνας στην θέση του, στην εικόνα B. Όλες οι υπόλοιπες μεταβολές που εντόπισε, οφείλονται είτε σε μεταβολές στη βλάστηση, είτε στη διαφορετική γωνία λήψης των εικόνων. Σε κάποιες ακόμα, δεν υπάρχει καμία μεταβολή.

Στα αποτελέσματα α/α 2, ενώ παρατηρείται μεταβολή και συγκεκριμένα δημιουργία κτηρίων στην B εικόνα, το δίκτυο έχει ανιχνεύσει κάποιες μεταβολές, οι οποίες όμως δεν αφορούν στον αστικό ιστό. Οι μεταβολές αυτές αφορούν σε ορισμένα αυτοκίνητα που υπάρχουν δεξιά της B εικόνας, ενώ δεν υπάρχουν στην A εικόνα.

Στα αποτελέσματα α/α 3, παρατηρείται μεταβολή και συγκεκριμένα κατεδάφιση ενός κτηρίου στην εικόνα A, το οποίο στην εικόνα B εμφανίζεται ως γυμνό έδαφος. Το δίκτυο

δεν έχει ανιχνεύσει κάποια αλλαγή, παρά μόνο ορισμένα μικρά τμήματα στη θάλασσα, ίσως λόγω κυματισμού.

Έπειτα, από όλες αυτές τις δοκιμές στο συγκεκριμένο σύνολο δεδομένων και ειδικά έπειτα από την αξιολόγηση των δεδομένων της εκπαίδευσης, καταλήγουμε στο ότι πιθανότατα το σύνολο δεδομένων αυτό, επειδή περιέχει εικόνες με αλλαγές, κατά βάση, σε αρκετά πυκνό αστικό ιστό, δεν είναι το καταλληλότερο για την εκπαίδευση του συγκεκριμένου δικτύου. Το δίκτυο δυσκολεύεται να εκπαιδευτεί και τα αποτελέσματα είναι πολύ κατώτερα από τα αναμενόμενα με βάση τη βιβλιογραφία.

5.3 ΕΦΑΡΜΟΓΗ ΕΚΔΟΧΗΣ 2

5.3.1 ΑΝΙΧΝΕΥΣΗ ΜΕΤΑΒΟΛΩΝ ΑΣΤΙΚΟΥ ΙΣΤΟΥ ΜΕ ΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ STANET

Όπως προαναφέρθηκε, στην 2η εκδοχή, έγινε κατευθείαν εφαρμογή του δικτύου STANet. Τα δεδομένα ταξινομήθηκαν με τη δομή που απαιτείται με βάση το ερευνητικό άρθρο, όπως ακριβώς και στην εκδοχή 1.

- Pre-trained

Αρχικά, εφαρμόστηκε το pre-trained μοντέλο, με τα βάρη της μεθόδου PAM, του ήδη εκπαιδευμένου δικτύου STANet στο σύνολο δεδομένων LEVIR-CD. Τα αποτελέσματα δεν ήταν ιδιαίτερα ικανοποιητικά.

Ορισμένα ενδεικτικά αποτελέσματα είναι τα εξής:

Πίνακας 7: Αποτελέσματα του pre-trained δικτύου STANet (2η εκδοχή)

α/α	A	B	Label	Αποτέλεσμα pre-trained
1				
2				
3				

Στα αποτελέσματα α/α 1, παρουσιάζεται δημιουργία ενός κτηρίου στην Β εικόνα. Το δίκτυο δεν έχει εντοπίσει καμία μεταβολή.

Στα αποτελέσματα α/α 2, παρουσιάζεται δημιουργία κτηρίων στην Β εικόνα. Το δίκτυο έχει εντοπίσει μόνο τη μία μεταβολή.

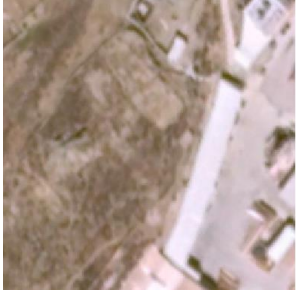





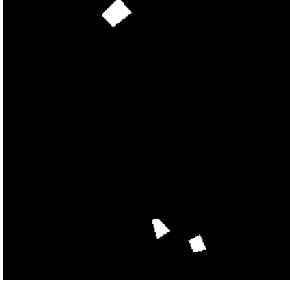



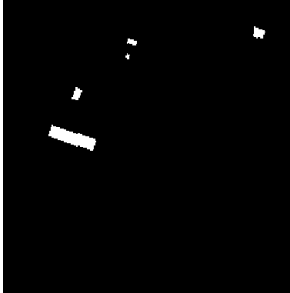

Στα αποτελέσματα α/α 3, παρουσιάζεται δημιουργία κτηρίων στην Β εικόνα. Το δίκτυο έχει εντοπίσει μόνο τις δύο μεταβολές.

Πιθανότατα, το pre-trained μοντέλο, και στην εκδοχή 2, απέτυχε στον εντοπισμό μεταβολών αστικού ιστού, μιας και τα δεδομένα στα οποία έχει εκπαιδευτεί το δίκτυο, δηλαδή στο σύνολο δεδομένων LEVIR-CD, αποτελούνται από μεταβολές κτηρίων, τα οποία έχουν σκεπές και όχι τσιμεντένιες οροφές όπως τα δεδομένα της Μυκόνου και της Πάρου.

- Finetuning

Έπειτα, ακολούθησε η διαδικασία finetuning, κατά την οποία η εκπαίδευση του δικτύου με τα δεδομένα LEVIR-CD, συνεχίστηκε με τα δεδομένα της εκδοχής 2 και με learning rate ίσο με 0.0002. Τα αποτελέσματα ήταν ικανοποιητικά. Ορισμένα ενδεικτικά αποτελέσματα είναι τα εξής:

Πίνακας 8: Αποτελέσματα της διαδικασίας finetuning δικτύου STANet (2η εκδοχή)

α/α	A	B	Label	Αποτέλεσμα finetuning
1				
2				
3				

Στα αποτελέσματα α/α 1, παρουσιάζεται δημιουργία ενός κτηρίου στην Β εικόνα. Το δίκτυο έχει εντοπίσει σε αρκετά καλό βαθμό την συγκεκριμένη μεταβολή.

Στα αποτελέσματα α/α 2, παρουσιάζεται δημιουργία κτηρίων στην Β εικόνα. Το δίκτυο έχει εντοπίσει πολύ καλά όλες τις μεταβολές.







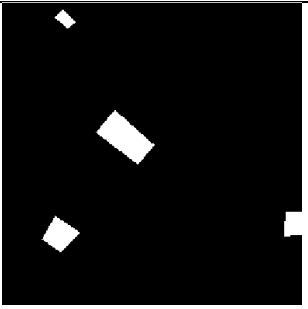





Στα αποτελέσματα α/α 3, παρουσιάζεται δημιουργία κτηρίων στην Β εικόνα. Το δίκτυο έχει εντοπίσει αρκετά καλά τις μεταβολές.

Το δίκτυο, έπειτα από την αξιολόγηση του finetuning, παρουσιάζει πολύ καλύτερα αποτελέσματα σε σχέση με όλες τις προηγούμενες διαδικασίες. Το δίκτυο έχει αρχίσει να εκπαιδεύεται στα δεδομένα. Το dataset 2 πιθανότατα παρουσιάζει καλύτερα αποτελέσματα λόγω του ότι εμφανίζει πιο ξεκάθαρες αλλαγές.

- Εκπαίδευση

Έπειτα, ακολούθησε η διαδικασία εκπαίδευσης, η οποία πραγματοποιήθηκε μόνο με τα δεδομένα της εκδοχής 2. Τα αποτελέσματα ήταν ικανοποιητικά. Ορισμένα ενδεικτικά αποτελέσματα είναι τα εξής:

Πίνακας 9: Αποτελέσματα της διαδικασίας εκπαίδευσης δικτύου STANet (2η εκδοχή)

α/α	A	B	Label	Αποτέλεσμα εκπαίδευσης
1				
2				
3				

Στα αποτελέσματα α/α 1, παρουσιάζεται δημιουργία ενός κτηρίου στην Β εικόνα. Το δίκτυο έχει εντοπίσει σε αρκετά καλό βαθμό την συγκεκριμένη μεταβολή (παρατηρείται ότι έχει ξεχωρίσει και την πισίνα και δεν την έχει συμπεριλάβει στη μεταβολή).

Στα αποτελέσματα α/α 2, παρουσιάζεται δημιουργία κτηρίων στην Β εικόνα. Το δίκτυο έχει εντοπίσει καλά τις μεταβολές αστικού ιστού, όμως ταυτόχρονα έχει εντοπίσει και ορισμένες μεταβολές, οι οποίες δεν αφορούν σε μεταβολές αστικού ιστού, αλλά σε διαφοροποίηση του εδάφους.

Στα αποτελέσματα α/α 3, παρουσιάζεται δημιουργία κτηρίου στην Β εικόνα. Το δίκτυο έχει εντοπίσει πολύ καλά τη μεταβολή.

Το δίκτυο, έπειτα από την αξιολόγηση της εκπαίδευσης με τα δεδομένα της εκδοχής 2, παρουσιάζει πολύ καλύτερα αποτελέσματα σε σχέση με όλες τις προηγούμενες διαδικασίες. Το δίκτυο εκπαιδεύεται στα δεδομένα. Το dataset 2 πιθανότατα παρουσιάζει καλύτερα αποτελέσματα λόγω του ότι εμφανίζει πιο ξεκάθαρες αλλαγές.

ΚΕΦΑΛΑΙΟ 6: ΑΞΙΟΛΟΓΗΣΗ

6.1 ΑΞΙΟΛΟΓΗΣΗ ΕΚΔΟΧΗΣ 1

6.1.1 ΑΞΙΟΛΟΓΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ SAM

Όπως προαναφέρθηκε, στην εκδοχή 1 πραγματοποιήθηκε τμηματοποίηση των εικόνων με το δίκτυο SAM και χειροκίνητη ταξινόμηση των αποτελεσμάτων αυτού. Δεν κατάφερε να αναγνωρίσει όλες τις οντότητες και έτσι δεν χρησιμοποιήθηκαν τα αποτελέσματα ως labels. Ορισμένα στατιστικά, τα οποία προέκυψαν παρουσιάζονται παρακάτω, τόσο σε επίπεδο αναγνώρισης αντικειμένων όσο και σε επίπεδο αναγνώρισης pixel.

Πίνακας 10: Αξιολόγηση SAM για την εικόνα του 2007

2007								
Patch	Type of urban area	Recognized Objects SAM	Total objects	Percentage (%)	Recognized Pixels	Unrecognized Pixels	Total Pixels	Percentage (%)
6_0	-	-	-	-	-	-	-	-
6_1	sparse	6	6	100.00	3528	-	3528	100.00
6_2	sparse	18	24	75.00	7709	4084	11793	65.37
6_3	sparse	8	23	34.78	3429	6771	10200	33.62
6_4	sparse	10	12	83.33	5736	596	6332	90.59
6_5	sparse	16	30	53.33	12629	11301	23930	52.77
6_6	dense	58	86	67.44	28720	9645	38365	74.86
6_7	dense	26	40	65.00	13524	5397	18921	71.48
6_8	sparse	5	5	100.00	7038	-	7038	100.00
6_9	sparse	4	5	80.00	6222	276	6498	95.75
6_10	sparse	5	8	62.50	2864	4113	6977	41.05
6_11	sparse	9	13	69.23	4641	1299	5940	78.13
6_12	-	-	-	-	-	-	-	-

Συνολικά:

Recognized Objects SAM	Total objects
165	252
	65,5%
Recognized Pixels	Total Pixels
96040	139522
	68,8%

Πίνακας 11: Αξιολόγηση SAM για την εικόνα του 2023

2023								
Patch	Type of urban area	Recognized Objects SAM	Total objects	Percentage (%)	Recognized Pixels	Unrecognized Pixels	Total Pixels	Percentage (%)
6_0	-	-	-	-	-	-	-	-
6_1	sparse	8	8	100.00	7945	-	7945	100.00
6_2	sparse	23	28	82.14	19742	3865	23607	83.63
6_3	sparse	15	18	83.33	8015	4416	12431	64.48
6_4	sparse	6	9	66.67	6826	858	7684	88.83
6_5	sparse	15	21	71.43	20445	5722	26167	78.13
6_6	dense	35	45	77.78	29523	7913	37436	78.86
6_7	dense	22	26	84.62	21514	3049	24563	87.59
6_8	sparse	5	5	100.00	5026	-	-	100.00
6_9	sparse	5	5	100.00	6653	-	-	100.00
6_10	sparse	10	11	90.91	8406	424	8830	95.20
6_11	sparse	12	21	57.14	5755	2893	8648	66.55
6_12	sparse	1	1	100.00	557	-	-	100.00

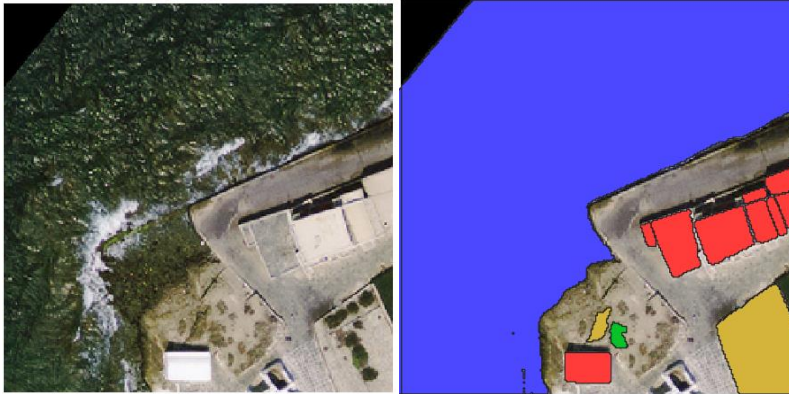
Συνολικά:

Recognized Objects SAM	Total objects
157	197
	79,7%
Recognized Pixels	Total Pixels
139850	157311
	88,9%

Με βάση τον προηγούμενο πίνακα παρατηρείται:

Στην εικόνα του 2007, μεγαλύτερο ποσοστό αναγνώρισης (ποσοστό 100%), τόσο σε επίπεδο αναγνώρισης αντικειμένων όσο και σε επίπεδο αναγνώρισης pixel εμφάνισαν τα patches 6_1 και 6_8, τα οποία απεικονίζουν τμήματα αραιού αστικού ιστού:

6_1 (αραιός αστικός ιστός)



6_8 (αραιός αστικός ιστός)



Επίσης, στην εικόνα του 2007, μικρότερο ποσοστό αναγνώρισης παρουσίασε το patch 6_3, με ποσοστό αναγνώρισης αντικειμένων 34.78% και ποσοστό αναγνώρισης pixel 33.62%, το οποίο απεικονίζει τμήμα αραιού αστικού ιστού:

6_3 (αραιός αστικός ιστός)



Στην εικόνα του 2023, μεγαλύτερο ποσοστό αναγνώρισης (ποσοστό 100%), τόσο σε επίπεδο αναγνώρισης αντικειμένων όσο και σε επίπεδο αναγνώρισης pixel εμφάνισαν τα patches 6_1, 6_8, 6_9 και 6_12 τα οποία απεικονίζουν τμήματα αραιού αστικού ιστού:

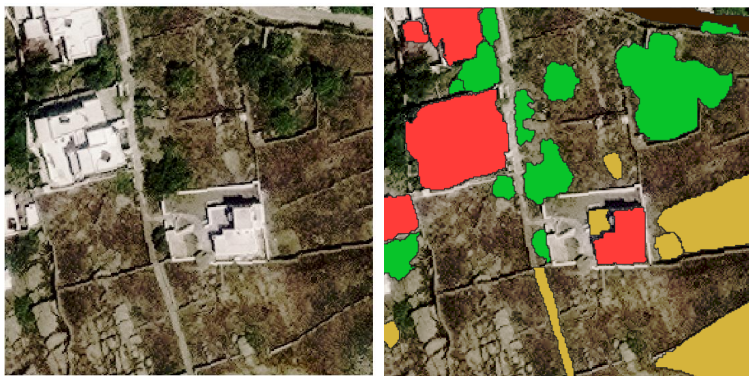
6_1 (αραιός αστικός ιστός)



6_8 (αραιός αστικός ιστός)



6_9 (αραιός αστικός ιστός)



6_12 (αραιός αστικός ιστός)



Επίσης, στην εικόνα του 2023, μικρότερο ποσοστό αναγνώρισης παρουσίασε το patch 6_11, με ποσοστό αναγνώρισης αντικειμένων 57.14% και ποσοστό αναγνώρισης pixel 66.55%, το οποίο απεικονίζει τμήμα αραιού αστικού ιστού:

6_11 (αραιός αστικός ιστός)



6.1.2 ΑΞΙΟΛΟΓΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ STANet

Στη συνέχεια, παρουσιάζονται τα αποτελέσματα που προέκυψαν έπειτα από την εφαρμογή του pre-trained μοντέλου, του μοντέλου βελτιστοποίησης (finetuning) και της εκπαίδευσης του δικτύου STANet για την 1η εκδοχή.

Αυτό που παρατηρήθηκε, στα αποτελέσματα που εμφανίζονταν καθώς έτρεχαν οι τρεις παρακάτω διαδικασίες είναι το γεγονός ότι το validation loss αυξανόταν πολύ αργά σε σύγκριση με το train loss και σε ορισμένες εποχές μειωνόταν. Αυτό σημαίνει ότι το δίκτυο εκπαιδεύονταν καλά στα δεδομένα train και δυσκολευόταν να γενικεύσει στο validation (overfitting). Ενώ έγιναν δοκιμές με μικρότερο learning rate και με άλλες loss functions το δίκτυο δεν κατάφερε να εμφανίσει τα επιθυμητά αποτελέσματα.

Πίνακας 12: Αξιολόγηση STANet για την 1^η εκδοχή

VERSION 1	Overall Accuracy	Mean_IoU	Precision	Recall	F1
PRE-TRAINED	0.985	0.020	0.101	0.028	0.085
FINETUNING	0.974	0.031	0.156	0.037	0.059
TRAINED (με δορυφορική εικόνα WorldView-3 και ορθοφωτογραφία Κτηματολογίου)	0.975	0.040	0.240	0.046	0.078

Παρακάτω, παρατηρούνται ορισμένα αποτελέσματα, όπως παρουσιάστηκαν και στο υποκεφάλαιο 5.2.2. Όπως προαναφέρθηκε, τα αποτελέσματα αυτά δεν ήταν ικανοποιητικά και το δίκτυο δεν κατάφερε να εντοπίσει τις μεταβολές αστικού ιστού. Τα αποτελέσματα f1 score ήταν πολύ χαμηλά περίπου 0.06-0.08.

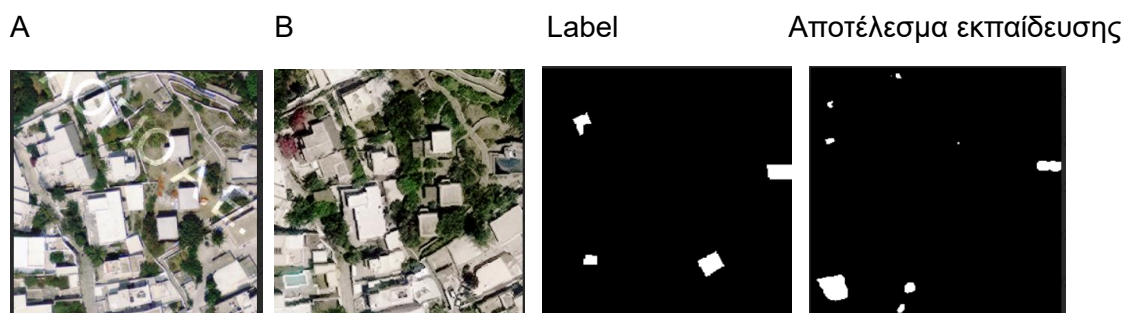
Pre-trained



Finetuning



Εκπαίδευση



6.2 ΑΞΙΟΛΟΓΗΣΗ ΕΚΔΟΧΗΣ 2

6.2.1 ΑΞΙΟΛΟΓΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ STANet

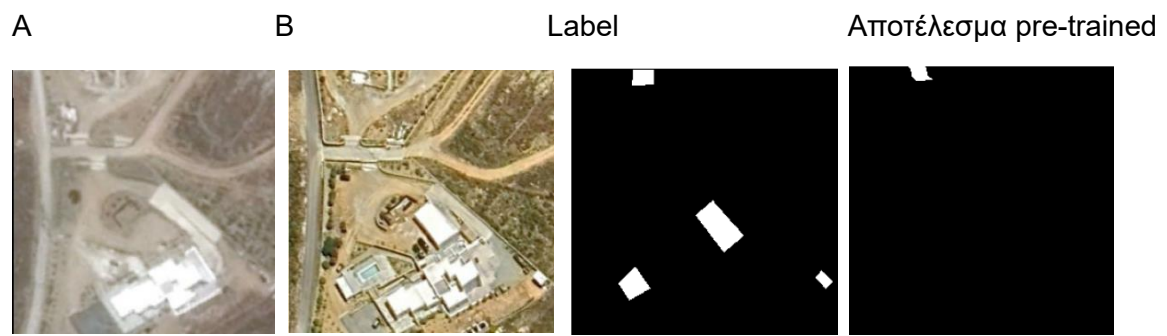
Στη συνέχεια, παρουσιάζονται τα αποτελέσματα που προέκυψαν έπειτα από την εφαρμογή του pre-trained μοντέλου, του μοντέλου βελτιστοποίησης (finetuning) και της εκπαίδευσης του δικτύου STANet για την 2η εκδοχή. Παρατηρήθηκε μόνο κατά την εφαρμογή του pre-trained μοντέλου, overfitting, όπως και στις διαδικασίες της 1ης εκδοχής, το validation loss αυξανόταν πολύ αργά σε σύγκριση με το train loss και σε ορισμένες εποχές μειωνόταν. Στις διαδικασίες finetuning και εκπαίδευσης με τα δεδομένα της εκδοχής 2, δεν παρατηρήθηκαν τέτοια φαινόμενα και όπως φαίνεται και από τον παρακάτω πίνακα τα αποτελέσματα είναι ικανοποιητικά.

Πίνακας 13: Αξιολόγηση STANet για την 2^η εκδοχή

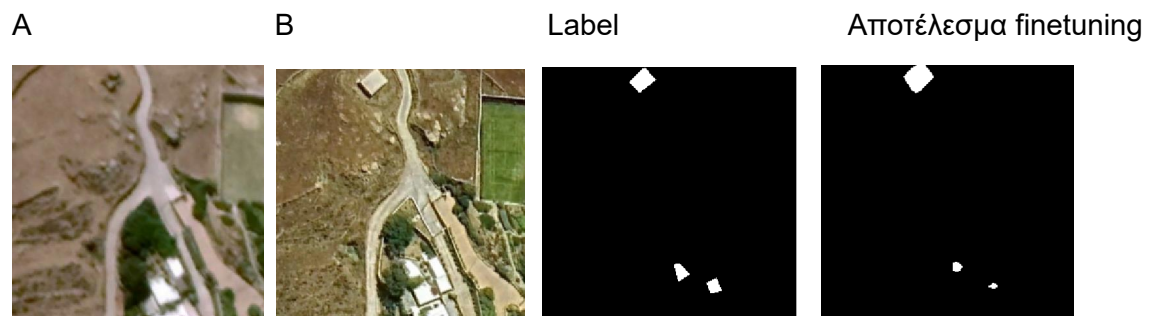
VERSION 2	Overall Accuracy	Mean_IoU	Precision	Recall	F1
PRE-TRAINED	0.943	0.485	0.158	0.030	0.051
FINETUNING	0.975	0.805	0.707	0.865	0.778
TRAINED (με εικόνες Google Earth)	0.978	0.816	0.755	0.833	0.792

Παρακάτω, παρατηρούνται ορισμένα αποτελέσματα, όπως παρουσιάστηκαν και στο υποκεφάλαιο 5.3.1. Όπως προαναφέρθηκε, τα αποτελέσματα του pre-trained μοντέλου δεν ήταν ικανοποιητικά και το δίκτυο δεν κατάφερε να εντοπίσει τις μεταβολές αστικού ιστού (ποσοστό f1 score 0.05), ενώ πολύ καλά αποτελέσματα παρουσίασε η διαδικασία finetuning και η εκπαίδευση (ποσοστά f1 score περίπου 0.80).

Pre-trained



Finetuning



Εκπαίδευση



ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ – ΠΡΟΟΠΤΙΚΕΣ

7.1 ΣΥΜΠΕΡΑΣΜΑΤΑ

Έπειτα από την ολοκλήρωση των παραπάνω διαδικασιών, γίνεται αντιληπτό ότι το ζήτημα ανίχνευσης μεταβολών αστικού ιστού με χρήση νευρωνικών δικτύων είναι αρκετά περίπλοκο. Με βάση τα πειράματα της παρούσας διπλωματικής, συμπεραίνονται τελικά τα εξής:

Η εκδοχή 1, με τις εικόνες WorldView-3 και Κτηματολογίου, δεν παρουσίασε τα επιθυμητά αποτελέσματα. Συγκεκριμένα:

- Το δίκτυο SAM δεν είχε τα επιθυμητά αποτελέσματα, αλλά πιθανότατα να λειτουργούσε καλύτερα σε εικόνες με ακόμα πιο αραιό αστικό ιστό.
- Το δίκτυο STANet στο συγκεκριμένο dataset δεν είχε επιτυχία (αποτελέσματα f1 score 0.06-0.09), τόσο στο pre-trained μοντέλο και στη διαδικασία finetuning όσο και στη διαδικασία εκπαίδευσης με δεδομένα της εκδοχής 1. Πιθανότατα, οι συγκεκριμένες εικόνες δεν ήταν οι πιο κατάλληλες, μιας και έγιναν δοκιμές με διάφορες παραμέτρους και όλα τα αποτελέσματα έδειξαν ότι το δίκτυο δυσκολευόταν πολύ να εκπαιδευτεί.
- Αρκετά από τα patches δεν είχαν μεταβολές και επομένως αυτό έκανε το δίκτυο να εκπαιδεύεται καλά στη μη-αλλαγή και όχι στην αλλαγή.

Όσον αφορά στην εκδοχή 2, με τις εικόνες που λήφθηκαν από το Google Earth:

- Τα αποτελέσματα του pre-trained μοντέλου STANet δεν ήταν αποδεκτά, με αποτέλεσμα f1 score 0.05.
- Τα αποτελέσματα έπειτα από τις διαδικασίες finetuning και training, του δικτύου STANet ήταν ικανοποιητικά με f1 score περίπου 0.80.
- Το συγκεκριμένο dataset αποτελούνταν μόνο από patches με μεταβολές αστικού ιστού και επομένως συνέβαλε στο να εκπαιδευτεί το δίκτυο καλά και στην αλλαγή.

7.2 ΠΡΟΟΠΤΙΚΕΣ

Με την ολοκλήρωση της εργασίας αναδείχθηκαν αρκετά θέματα, τα οποία χρήζουν περαιτέρω διερεύνησης. Συγκεκριμένα, υπάρχει η δυνατότητα δοκιμής άλλων νευρωνικών δικτύων, όπως Convolutional Neural Networks (CNNs), Auto-Encoders (AEs), Deep Belief Networks (DBNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), AlexNet, GoogLeNet και το VGGNet, τα οποία με βάση τη βιβλιογραφία παρουσιάζουν καλά αποτελέσματα, με καλύτερες ακρίβειες και καλύτερα οπτικά αποτελέσματα στα σχήματα και είναι πιο εξελιγμένα από το STANet. Επίσης, μελέτες μπορούν να επικεντρωθούν στην εφαρμογή των μοντέλων σε άλλες περιοχές, σε διαφορετικού τύπου δεδομένα και ταυτόχρονα βελτίωσης των παραμέτρων. Επόμενες έρευνες θα μπορούσαν να έχουν ως αντικείμενο την εφαρμογή των παραπάνω δικτύων και σε άλλες χώρες, να γίνουν δοκιμές σε περιοχές με διαφορετικό ανάγλυφο, με κτήρια άλλης μορφής και διαφορετικής πυκνότητας αστικού ιστού. Ταυτόχρονα, η βελτίωση των παραμέτρων του αυτόματου co-registration των εικόνων, μέσω νευρωνικών δικτύων είναι ένα ενδιαφέρον ερευνητικό αντικείμενο. Απλώς για το συγκεκριμένο θέμα ανίχνευσης μεταβολών αστικού ιστού πιθανότατα να είναι πολύ δύσκολο να αποφευχθεί η λήψη σημείων GCP χειροκίνητα.

ΚΕΦΑΛΑΙΟ 8: ΑΝΑΦΟΡΕΣ

- Viktoria Kristollari and Vassilia Karathanassi (2022), “*Change Detection in VHR Imagery With Severe Co-Registration Errors Using Deep Learning: A Comparative Study*” (https://www.researchgate.net/publication/359448426_Change_Detection_in_VHR_Imagery_With_Severe_Co-Registration_Errors_Using_Deep_Learning_A_Comparative_Study)
- HaoChen and Zhenwei Shi (2020), “*A Spatial-Temporal Attention-Based Method and a NewDataset for Remote Sensing Image Change Detection*” (<https://www.mdpi.com/2072-4292/12/10/1662-https://github.com/justchenhao/STANet>)
- Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollar, Ross Girshick (2023), “*Segment Anything*” (<https://arxiv-org-https://github.com/facebookresearch/segment-anything>)
- Haiping Yang ,Yuanyuan Chen, Wei Wu, Shiliang Pu, Xiaoyang Wu, Qiming Wan and Wen Dong (2023), “*A Lightweight Siamese Neural Network for Building Change Detection Using Remote Sensing Images*” (<https://www.mdpi.com/2072-4292/15/4/928-https://github.com/yanghplab/LightCDNet>)
- Cong Cao,Suzana Dragičević, and Songnian Li (2019), “*Land-Use Change Detection with Convolutional Neural Network Methods*” (<https://www.mdpi.com/2076-3298/6/2/25>)
- Maria Vakalopoulou (2017), Doctoral Dissertation: “*Mapping Changes in Very High Resolution Satellite Data: Coupling Registration, Semantic Segmentation and Change Detection*” (<https://freader.ekt.gr/eadd/index.php?doc=41274#p=1>)
- Maria Papadomanolaki (2021), Doctoral Dissertation: “*Registration, Semantic Segmentation and Change Detection with Deep Learning in Urban Environments*” (<https://freader.ekt.gr/eadd/index.php?doc=51911&lang=el#p=1>)
- D. Gupta and M. K. Patil (2014), “*A review on image registration*” (<https://www.ijert.org/a-review-on-image-registration>)
- <http://dspace.lib.ntua.gr>
- www.gdal.org
- <https://github.com/>
- <https://www.sciencedirect.com/journal/neural-networks>
- <https://pytorch.org/>
- <https://medium.com/>
- <https://scikit-learn.org/>
- <http://www.learnartificialneuralnetworks.com/#Intro>
- <https://apmonitor.com/>
- <https://www.v7labs.com>

- <https://www.analyticsvidhya.com/>
- <https://opencourses.auth.gr/>
- <https://www.javatpoint.com/>
- <https://towardsdatascience.com/>