

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

**Μεταπτυχιακό Πρόγραμμα Σπουδών *Μεταπτυχιακό
Πρόγραμμα Σπουδών Πληροφοριακά και Επικοινωνιακά
Συστήματα***

Μεταπτυχιακή Διατριβή



**Αναγνώριση κτηρίων, πισίνων και παλιών
εγκαταλελειμμένων σπιτιών χωρίς οροφή, μέσω εικόνων
δορυφόρου, με τη χρήση Βαθιάς Μάθησης**

Μικαέλα Παπουτέ

**Επιβλέπων Καθηγητής
Αντρέας Καμηλάρης**

Μάιος 2022

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

**Μεταπτυχιακό Πρόγραμμα Σπουδών *Μεταπτυχιακό
Πρόγραμμα Σπουδών Πληροφοριακά και Επικοινωνιακά
Συστήματα***

Μεταπτυχιακή Διατριβή

**Αναγνώριση κτηρίων, πισίνων και παλιών
εγκαταλελειμμένων σπιτιών χωρίς οροφή, μέσω εικόνων
δορυφόρου, με τη χρήση Βαθιάς Μάθησης**

Μικαέλα Παπουτέ

**Επιβλέπων Καθηγητής
Αντρέας Καμηλάρης**

Η παρούσα μεταπτυχιακή διατριβή υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων για απόκτηση μεταπτυχιακού τίτλου σπουδών στα Πληροφοριακά και Επικοινωνιακά Συστήματα από τη Σχολή Θετικών και Εφαρμοσμένων Επιστημών του Ανοικτού Πανεπιστημίου Κύπρου.

Μάιος 2022

Περίληψη

Στην παρούσα μεταπτυχιακή διατριβή έγινε μια προσπάθεια εφαρμογής μεθόδων βαθιάς μάθησης για την αναγνώριση κτηρίων, πισίνων και παλιών/ εγκαταλελειμμένων σπιτιών χωρίς οροφή, με τη βοήθεια εικόνων δορυφόρου.

Πρόκειται για μια πολύπλοκη διαδικασία, η οποία εξαρτάται από την πολυπλοκότητα της δόμησης της περιοχής και της διαφορετικής δομής και κατασκευής των διάφορων κτηρίων και πισίνων.

Το πρόβλημα χρησιμοποιεί την αναγνώριση μοτίβων με σκοπό την κατηγοριοποίηση ενός συνόλου δεδομένων, για τη δημιουργία χαρτών δόμησης, οι οποίοι μπορούν να αξιοποιηθούν σε διάφορους τομείς της μηχανικής και της έρευνας.

Για τον εντοπισμό κτηρίων και πισίνων, παλαιότερα οι άνθρωποι χρησιμοποιούσαν εξοπλισμό, όπως αισθητήρες. Αυτό απαιτούσε καλή προ επεξεργασία, ώστε να είμαστε σίγουροι ότι η διαδικασία εξαγωγής γινόταν με το σωστό τρόπο κατά την αυτοματοποίηση, και είχε πολύ υψηλό κόστος.

Για την αντιμετώπιση του πιο πάνω προβλήματος, στην περίπτωση μας, χρησιμοποιήθηκε η διαδικασία βαθιάς μάθησης, η οποία μιμείται τον τρόπο που μαθαίνει ο εγκέφαλος του ανθρώπου και με τη χρήση εικόνων δορυφόρου, επιτυγχάνεται ο αυτόματος εντοπισμός σπιτιών, πισίνων και παλιών εγκαταλελειμμένων κτηρίων χωρίς οροφή.

Συγκεκριμένα, στην παρούσα μεταπτυχιακή διατριβή, χρησιμοποιήθηκε η αρχιτεκτονική του δικτύου U-Net, η οποία με την επεξεργασία των δεδομένων, αναγνωρίζει τις συγκεκριμένες κατηγορίες και κάνει χαρτογράφηση (mapping) στις υφιστάμενες εικόνες.

Αρχικά το δίκτυο αναγνωρίζει τι είναι το αντικείμενο (κτήριο, παλιό κτήριο ή πισίνα) και στη συνέχεια μαθαίνει να το χαρτογραφεί.

Για να αξιολογηθεί το δίκτυο, χρησιμοποιήθηκαν εικόνες δορυφόρου, οι οποίες συλλέχθηκαν από αξιόπιστη πηγή, η οποία είναι το Κτηματολόγιο της Κύπρου. Συγκεκριμένα χρησιμοποιήθηκαν 5131 εικόνες δορυφόρου προς εκπαίδευση από διάφορες περιοχές της Κύπρου.

Τα πειραματικά αποτελέσματα του, δείχνουν πως το μοντέλο έχει την ικανότητα να εντοπίζει σπίτια με ακρίβεια 81%, παλιά εγκαταλελειμμένα κτήρια χωρίς οροφή με ακρίβεια 94% και πισίνες με ακρίβεια 87%

Summary

In this master's thesis, deep learning methods were applied to identify buildings, swimming pools and old / abandoned houses without roof, using satellite pictures.

It is a complex process, which depends on the complexity of the construction of the area and the different structure and construction of the buildings and swimming pools.

To categorize a set of data, pattern recognition is used, to create building maps that can be used in various fields of engineering and research.

In the past, people used equipment with sensors, for the detection of buildings and swimming pools. This required a good pre-processing, to be sure that the extraction process was done correctly in the automation. Also this method has a very high cost.

The above problem was solved in our case, with the deep learning process, which mimics the way that the human brain learns, using satellite images for automatic identification of buildings, swimming pools and old / abandoned houses without roof.

Specifically, the U-Net network architecture was used, which by the data processing, identifies the specific categories and mapping the existing images.

First the network recognizes what the object is (building, old building, pool,) and then learns to map it .

To evaluate the network, satellite images were used, which were collected from a reliable source, which is the Cyprus Land Registry. Specifically, 5131 satellite images were used for training from various areas of Cyprus.

The experimental results show that the model has the ability to detect houses with 81% accuracy, old abandoned buildings without roof with 94% accuracy and swimming pools with 87% accuracy in the control set.

Ευχαριστίες

Θερμές ευχαριστίες θα ήθελα να εκφράσω στον κ. Ανδρέα Καμηλάρη, ο οποίος δέχθηκε να με αναλάβει ως επιβλέπων καθηγητής και για την ουσιαστική βοήθεια του για ολοκλήρωση της μεταπτυχιακής διατριβής μου.

Ιδιαίτερα θερμές ευχαριστίες, θα ήθελα να αποδώσω στον κ. Chirag Padubidri, για την καθοδήγηση, τη σημαντική βοήθεια του και το ενδιαφέρον του σε όλες τις φάσεις της διατριβής μου.

Τέλος θα ήθελα να ευχαριστήσω τον κ. Indrajit Kalita για την πολύτιμη βοήθεια του .

Περιεχόμενα

1	Εισαγωγή.....	1
1.1	Εισαγωγή στην Τεχνητή Νοημοσύνη	1
1.2	Η σημασία της αυτόματης εξαγωγής κτηρίων / πισίνων και παλιών εγκαταλελειμμένων σπιτιών	2
2	Ιστορική αναδρομή	4
2.1	Τεχνητή Νοημοσύνη - Το Πείραμα του Alan Turing.....	4
2.2	Οι αρχικές εφαρμογές με τη χρήση Τεχνητής.....	5
2.3	Η επανάσταση στον τομέα της βαθιάς μάθησης.....	6
2.4	Εφαρμογές Βαθιάς μάθησης σε εικόνες δορυφόρων	7
2.5	Πώς λειτουργεί η βαθιά μάθηση;	8
2.6	Συνελικτικά Νευρωνικά Δίκτυα - CNN (Convolutional Neural Networks).....	9
3	Πειραματική Διαχείριση.....	11
3.1	Συλλογή Δεδομένων	11
3.2	Ετοιμασία Δεδομένων.....	11
3.3	Επιλογή μοντέλου για βαθιά μάθηση: UNet model.....	14
3.3.1	Εκπαίδευση Αλγορίθμου.....	18
3.4	Αξιολόγηση αποτελεσμάτων - Μέτρηση απόδοσης ακρίβειας.....	19
3.4.1	Δείκτης Ακρίβειας/ precision.....	21
3.4.2	Δείκτης ανάκλησης / recall.....	21
3.4.3	Δείκτης F1 Score.....	23
3.4.4	Δείκτης Accuracy.....	23
3.4.5	Δείκτης IOU.....	24
4	Αποτελέσματα.....	25
4.1	Αποτελέσματα μοντέλου για πισίνες.....	25
4.2	Αποτελέσματα μοντέλου για Σπίτια.....	26
4.3	Αποτελέσματα μοντέλου για Παλιά εγκαταλελειμμένα κτήρια χωρίς οροφή.....	27
5.	Συζήτηση	29

5.1	Συγκρίσεις με άλλες πειραματικές μελέτες που έλυσαν το πρόβλημα για αναγνώριση αντικειμένων ίδιας κατηγορίας.....	30
5.1.1	Εντοπισμός πισίνων.....	30
5.1.2	Εντοπισμός σπιτιών.....	31
5.1.3	Εντοπισμός κατεστραμμένων σπιτιών / παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή.....	32
6.	Επίλογος	33
	Βιβλιογραφία	34

Κεφάλαιο 1

Εισαγωγή

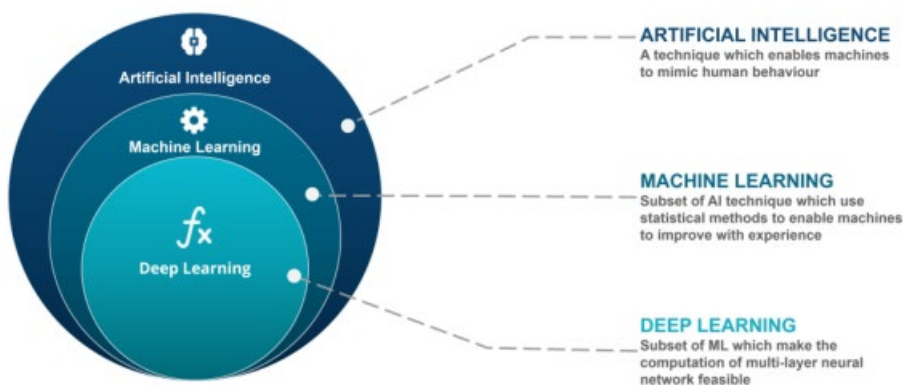
1.1 Εισαγωγή στην Τεχνητή Νοημοσύνη

Η τεχνητή νοημοσύνη[3], τα τελευταία χρόνια, και ειδικά οι πρόσφατες εξελίξεις στο χώρο της βαθιάς μάθησης, έχουν καταφέρει να μπουν δυναμικά στην καθημερινή μας ζωή, προσφέροντας μια σειρά από εφαρμογές, τις οποίες ο άνθρωπος χρησιμοποιεί στην καθημερινότητά του.

Συγκεκριμένα, η τεχνητή νοημοσύνη, αφορά το ευρύτερο σύνολο, που αντιστοιχεί σε οποιαδήποτε τεχνολογία και προσπαθεί να μιμηθεί τη νοημοσύνη ή τα πρότυπα συμπεριφοράς των ανθρώπων ή άλλων εμβρύων όντων.

Υποκατηγορία αυτού του τομέα είναι η μηχανική μάθηση, η οποία περιλαμβάνει μεθόδους και προσπαθεί να αναπτύξει αλγόριθμους, οι οποίοι, μέσα από την επεξεργασία δεδομένων, προσπαθούν να μάθουν, χωρίς να χρησιμοποιούν ένα περίπλοκο σύνολο κανόνων και βρίσκονται σε θέση να κάνουν προβλέψεις.

Υποσύνολο της μηχανικής μάθησης, είναι η βαθιά μάθηση, η οποία περιλαμβάνει μεθόδους μηχανικής μάθησης, που εμπνέονται από το δίκτυο νευρώνων του εγκεφάλου και βασίζονται στα λεγόμενα βαθιά νευρωνικά δίκτυα[7].



Εικόνα : Απεικονίζει τη σχέση Τεχνητής Νοημοσύνης, Μηχανικής Μάθησης και Βαθιάς Μάθησης

1.2 Η σημασία της αυτόματης εξαγωγής κτηρίων / πισίνων και παλιών εγκαταλελειμμένων σπιτιών

Στην εποχή μας η δυνατότητα εξόρυξης δεδομένων από ψηφιακές εικόνες είναι πλέον αναγκαία, αφού οι περισσότερες εφαρμογές έχουν ψηφιακή μορφή.

Η συνεχής ανάπτυξη της τεχνολογίας σε όλους τους τομείς, προσφέρει περισσότερη ακρίβεια, λιγότερο κόστος και εξοικονομεί ανθρώπινο χρόνο.

Παλαιότερα, ο εντοπισμός και χαρτογράφηση κτηρίων και πισίνων γινόταν με τη βοήθεια της τοπογραφίας. Αυτό απαιτούσε εκτός από τη χρήση εξειδικευμένου εξοπλισμού και τη δαπάνη σημαντικού ανθρώπινου χρόνου.

Η αυτόματη χαρτογράφηση και ο εντοπισμός παλιών και νέων κτηρίων και πισίνων , αποτελεί σημαντικό προνόμιο, σε διάφορες μελέτες πριν από την ανάπτυξη κάποιου έργου, αφού βοηθά στον εντοπισμό και μελέτη της οικιστικής ζώνης στο χώρο, όπου γίνεται η μελέτη.

Αυτό μπορεί να αφορά πληθώρα εργασιών όπως χωροταξικές εργασίες, εργασίες που σχετίζονται με τη συγκοινωνία, πολεοδομικές και κτηματολογικές μελέτες, μελέτες που αφορούν το περιβάλλον, φορολογικές μελέτες, αρχαιολογικές μελέτες κ.α.

Με την ανάλυση αυτών των δεδομένων, θεσπίζονται όροι και περιορισμοί, που σκοπό έχουν τη βέλτιστη αξιοποίηση του έργου, ώστε να έχουν όσο το δυνατόν περισσότερες θετικές επιπτώσεις ως προς το περιβάλλον και τις οικιστικές ζώνες.

Στην παρούσα μεταπτυχιακή διατριβή, χρησιμοποιώντας εικόνες δορυφόρου που λήφθηκαν μέσω δορυφόρου, προτείνουμε τη χρήση υπολογιστικής όρασης, μέσω της αρχιτεκτονικής Βαθιάς μάθησης, ώστε να επιτευχθεί η καταμέτρηση σπιτιών, πισίνων, εγκαταλελειμμένων σπιτιών με όσο το δυνατόν πιο υψηλή ακρίβεια και σε σύντομο χρονικό διάστημα και με λιγότερο κόστος, σε σχέση με τις κλασσικές μεθόδους.

Η συγκεκριμένη ανάλυση είναι σημαντική και μπορεί να βρει θέση στο κτηματολόγιο του κράτους και στη διοικητική διαχείριση του κράτους, όπου η επιθεώρηση είναι δύσκολο να υποστηριχθεί.

Πιο συγκεκριμένα, η αυτόματη εξαγωγή πισίνων μέσω εικόνων δορυφόρου, μπορεί να βρει θέση σε περιπτώσεις που αφορούν την έγκριση τους, την αδειοδότηση τους καθώς και τον

εντοπισμό παράνομων πισίνων, οι οποίες δεν τηρούν τις φορολογικές υποχρεώσεις που απαιτούνται.

Επιπρόσθετα των πιο πάνω, ο εντοπισμός και η αυτόματη εξαγωγή κτηρίων μπορεί να δώσει στους υπεύθυνους μια ευρύτερη εικόνα της περιοχής. Με αυτό τον τρόπο γίνεται ο εντοπισμός επιπρόσθετων έργων που απαιτούνται, ώστε να επιτευχθεί η ομαλή λειτουργία της κάθε περιοχής και κατ' επέκταση την ομαλή λειτουργία της κάθε χώρας ξεχωριστά.

Τέλος στην παρούσα μεταπτυχιακή διατριβή έγινε ο αυτόματος εντοπισμός παλιών εγκαταλελειμμένων κτηρίων χωρίς οροφή, κάτι το οποίο εκτός των άλλων, μπορεί να βρει θέση στις διάφορες καταστροφές που μπορεί να προκληθούν, όπως για παράδειγμα πυρκαγιές, σεισμούς, πολέμους κτλ . Επιπρόσθετα, αξίζει να σημειωθεί το γεγονός ότι με τη σωστή χρήση του, μπορεί να αποτελέσει τον πυρήνα για τις διάφορες ανασκαφές εντοπισμού αρχαίων οικισμών.

Συνοψίζοντας η χρήση της βαθιάς μάθησης στον εντοπισμό διάφορων κτηρίων και πισίνων ενισχύει την καλύτερη λειτουργία της καταγραφής γεωγραφικών πληροφοριών, δικαιωμάτων και υποχρεώσεων προς αυτά και διευκολύνει την εύρυθμη διοίκηση της κάθε χώρας.

Κεφάλαιο 2

Ιστορική αναδρομή

2.1 Τεχνητή Νοημοσύνη - Το Πείραμα του Alan Turing

Από την αρχή της τεχνητής νοημοσύνης, ο Alan Turing εξέτασε το ερώτημα αν οι μηχανές μπορούν να σκεφτούν, όπως ο άνθρωπος[4].

Για να δώσει λύση στο ερώτημα αυτό, χρησιμοποίησε ένα διανοητικό πείραμα, το λεγόμενο Turing test , το οποίο αποτελεί μια πολύ σημαντική συνεισφορά στον χώρο της τεχνητής νοημοσύνης.

Συγκεκριμένα, το Turing test υπέθετε ότι ένας άνθρωπος, ανταλλάσσει γραπτά μηνύματα/ερωτήσεις, τόσο με μια μηχανή, όσο και με έναν άλλο άνθρωπο, οι οποίοι βρίσκονται πίσω από ένα χωρίς να ξέρει πού βρίσκεται η μηχανή και πού ο άνθρωπος. Μέσα από αυτή την ανταλλαγή μηνυμάτων, ο άνθρωπος είχε στόχο να προσπαθήσει να καταλάβει ποια μηνύματα προέρχονται από τη μηχανή και ποια από τον άλλο άνθρωπο. Αντίστοιχα, ο στόχος της ανάπτυξης νοημοσύνης για τη μηχανή, ήταν να μπορέσει να ξεγελάσει τον άνθρωπο ώστε αυτός να μην μπορεί να ξεχωρίσει αν τα μηνύματα προέρχονται από τη μηχανή ή από τον άλλο άνθρωπο.

Με αυτό τον τρόπο, ο Alan Turing διατύπωσε το στόχο της τεχνητής νοημοσύνης: Πώς η μηχανή μπορεί να συμπεριφέρεται σαν κάποιο άνθρωπο;

Το πιο πάνω πείραμα κατέληξε στις 2 κύριες κατευθύνσεις της τεχνητής νοημοσύνης :

1. Μεθοδολογίες, οι οποίες βασίζονται σε κανόνες
2. Μεθοδολογίες, οι οποίες βασίζονται στη μάθηση, προσπαθώντας να αναπτύξουν μεθοδολογίες, που βασίζονται στη μάθηση για να αναπτύξουν νοημοσύνη, όπως ένας άνθρωπος όταν μαθαίνει. Αυτές οι κατευθύνσεις συνεχίζουν να είναι επίκαιρες μέχρι σήμερα.[4]

2.2 Οι αρχικές εφαρμογές με τη χρήση Τεχνητής Νοημοσύνης

Το 1952, ένα πρόγραμμα του Arthur Samuel κατάφερε να παίζει ντάμα σε επίπεδο προχωρημένου ερασιτέχνη και μετά από 3 χρόνια το 1955[6], το πρόγραμμα Logic Theorist του Alan Newell και του Herbert Simon μπορούσε να αποδεικνύει μαθηματικά θεωρήματα[6].

Η τεχνητή νοημοσύνη αναλύθηκε το 1956, όπου πλήθος ερευνητών, στο Dartmouth College ορίσαν τις βάσεις της τεχνητής νοημοσύνης, θέτοντας πολύ φιλόδοξους στόχους και ευελπιστώντας ότι μέσα στα επόμενα χρόνια, θα καταφέρουν να πετύχουν πολύ σημαντικά επιτεύγματα σε αυτό τον τομέα της Τεχνητής Νοημοσύνης[5].

Έτσι λοιπόν, την εποχή εκείνη παρατηρούμε ότι υπήρξε μια μεγάλη αισιοδοξία, καθώς διατυπώνονταν από τους ερευνητές υπερ- φιλόδοξοι στόχοι, οι οποίοι δυστυχώς εκείνη τη δεκαετία δεν υλοποιήθηκαν, φέρνοντας απογοητευτικά αποτελέσματα.

Οι βασικοί λόγοι ήταν :

1. Η Περιορισμένη υπολογιστική ισχύς. Οι υπολογιστές εκείνης της εποχής δεν μπορούσαν να ανταπεξέλθουν στο χώρο που χρειαζόταν για αναζήτηση.[9]
2. Η μη αποτελεσματική διαχείριση της πολυπλοκότητας των προβλημάτων της τεχνητής νοημοσύνης. Οι προσεγγίσεις της εποχής εκείνης, απαιτούσαν χειροκίνητη ανθρώπινη εισαγωγή πληροφοριών και δεδομένων, κάτι που ήταν ανέφικτο δεδομένου της πολυπλοκότητας της τεχνητής νοημοσύνης του πραγματικού κόσμου[9].

Έτσι το τέλος της 1^{ης} περιόδου της τεχνικής νοημοσύνης ,ήρθε το 1960 και σήμανε την αρχή της νέας γενιάς της τεχνικής νοημοσύνης[9].

2.3 Η επανάσταση στον τομέα της Βαθιάς Μάθησης

Η ουσιαστική εξέλιξη στον τομέα της βαθιάς μάθησης, άρχισε το 2012, όπου εφαρμόστηκε με συνταρακτική επιτυχία η βαθιά μάθηση στην αναγνώριση των εικόνων.

Έτσι προήλθαν συστήματα τα οποία έπαιρναν σαν είσοδο μια φωτογραφία και μπορούσαν αξιόπιστα να εκτιμούν το αντικείμενο, το οποίο βρίσκεται σε αυτή τη φωτογραφία διαλέγοντας ανάμεσα σε περίπου 1000 κλάσεις. Για παράδειγμα μπορούσαν να αναγνωρίσουν αν σε μια φωτογραφία αποτυπώνεται μια γάτα, ένας σκύλος κτλ[10].

Αξίζει να αναφερθεί ότι το 2012, το μοντέλο AlexNet[11], το οποίο βασίζονταν στη βαθιά μάθηση, κατάφερε να ξεπεράσει τις παραδοσιακές τεχνικές της όρασης υπολογιστών με μια διαφορά +10% στην αναγνώριση εικόνων. Κάτι τέτοιο, οδήγησε όλο και περισσότερους ερευνητές να χρησιμοποιούν τέτοιου είδους μεθοδολογίες και τη χρήση εξολοκλήρου μεθόδων βαθιάς μάθησης.

Αργότερα, τα δύο συστήματα βαθιάς μάθησης της Google[12] και της Microsoft[13] κατάφεραν να ξεπεράσουν ακόμα και την ανθρώπινη επίδοση, αφού μπορούσαν να αναγνωρίσουν τη κλάση στην οποία ανήκαν οι φωτογραφίες, καλύτερα και από τον οποιονδήποτε άνθρωπο προσπαθούσε να λύσει το ίδιο πρόβλημα.

Το 2016, ένα σύστημα βαθιάς μάθησης κατάφερε να κερδίσει τον παγκόσμιο πρωταθλητή του «Go», ενός παραδοσιακού κινέζικου παιχνιδιού το οποίο είναι πιο δύσκολο από το σκάκι[14].

Έτσι λοιπόν προκύπτει η ερώτηση: Γιατί η επανάσταση της βαθιάς μάθησης συμβαίνει τώρα, δεδομένου ότι οι βασικές μεθοδολογίες, τα μοντέλα και τα τεχνικά νευρωνικά δίκτυα είχαν αναπτυχθεί πριν από δεκαετίες;

Ο λόγος σχετίζεται με κυρίως 3 παράγοντες:

1. Με τις εξελίξεις στα Μεγα-δεδομένα (Big Data) . Με την πάροδο του χρόνου, προέκυψαν μεγαλύτερες βάσεις δεδομένων, οι οποίες ήταν πολύ σημαντικές για την εκπαίδευση των δικτύων [8].
2. Με την εξέλιξη στο υλικό των υπολογιστών(hardware). Συγκεκριμένα, εμφανίστηκαν και χρησιμοποιήθηκαν οι κάρτες γραφικών GPU, οι οποίες προσφέρουν μαζική παραλληλοποίηση . Αυτό ευνοεί την εκπαίδευση των δικτύων και δίνει τη δυνατότητα να εκπαιδευτούν δίκτυα πολύ μεγαλύτερης κλίμακας.[15]
3. Με τις εξελίξεις στο λογισμικό των ΗΥ. Στην πορεία του χρόνου, προέκυψαν βελτιώσεις σε τεχνικές, βελτιωμένα νέα μοντέλα και βιβλιοθήκες κώδικα, που δημιουργούσαν τεχνητά νευρωνικά δίκτυα. Αυτές ήταν δημόσια διαθέσιμες, κάτι το οποίο επέτρεψε σε ένα ευρύτερο δίκτυο επιστημόνων να χρησιμοποιήσουν αυτές τις μεθοδολογίες και να συνεισφέρουν στην ανάπτυξη της βαθιάς μάθησης.

2.4 Εφαρμογές Βαθιάς μάθησης σε εικόνες δορυφόρων

Γενικά, η εφαρμογή της βαθιάς μάθησης τα τελευταία χρόνια σε εικόνες δορυφόρου, έχει βρει θέση σε λύση προβλημάτων που σχετίζονται με την παρακολούθηση της στεριάς και της θάλασσας.

Στις μέρες μας, μέσω της επεξεργασίας των εικόνων δορυφόρου, γίνεται ανάλυση και κατανόηση των συνθηκών κάλυψης της γης και χαρτογράφηση του πυθμένα των ωκεανών.

Με τη σωστή χρήση και ανάλυση των δεδομένων που προκύπτουν, γίνεται πρόβλεψη των περιβαλλοντικών συνθηκών, ανιχνεύονται κάποιες ανωμαλίες σε στεριά και θάλασσα και γίνεται χαρτογράφηση ακόμα και σε φυσικές καταστροφές του πλανήτη μας.

Έτσι, οι ειδικοί θεσπίζουν όρους και περιορισμούς, ώστε να ευνοείται το περιβάλλον του πλανήτη μας.

Εφαρμογές χρήσης της παρούσας μεταπτυχιακής διατριβής, όπως ανέφερα και στο πιο πάνω Κεφάλαιο 1.2, μπορούν να εφαρμοστούν για παρακολούθηση της οικιστικής ζώνης στο χώρο, βοηθώντας στην εύρυθμη διοίκηση της χώρας, καταγραφή γεωγραφικών πληροφοριών, δικαιωμάτων και υποχρεώσεων των πολιτών, σε σπίτια και είδη πολυτελείας, όπως πισίνες, καθώς και για σκοπούς έγκρισης και αδειοδότησης τους.

Επιπρόσθετα, εφαρμόζονται στον εντοπισμό επιπρόσθετων έργων για τη συγκοινωνία, πολεοδομικές και κτηματολογικές μελέτες, ώστε να επιτευχθεί η ομαλή λειτουργία της κάθε περιοχής και κατ' επέκταση η ομαλή λειτουργία της κάθε χώρας με όσο το δυνατόν περισσότερες θετικές επιπτώσεις ως προς το περιβάλλον και τις οικιστικές ζώνες.

Τέλος, όπως ήδη αναφέρθηκε, εφαρμόζονται σε διάφορες καταστροφές και αποτελεί βασικό εργαλείο εύρυθμης λειτουργίας του κράτους, καθώς επίσης ο εντοπισμός τους μπορεί να αποτελέσει τον πυρήνα για τις διάφορες ανασκαφές εντοπισμού αρχαίων οικισμών σε αρχαιολογικές μελέτες.

2.5 Πώς λειτουργεί η βαθιά μάθηση;

Η βασική ιδέα της βαθιάς μάθησης, βρίσκεται στην ιδέα των Ιεραρχικών αναπαραστάσεων.

Με αυτό τον τρόπο δημιουργείται ένα pipe line επεξεργασίας, το οποίο στην αρχή θέτει και απαντά σε πολύ απλές ερωτήσεις που αφορούν τοπικές περιοχές της εικόνας.

Καθώς προχωρά η επεξεργασία, υποβάλλονται ερωτήσεις οι οποίες είναι όλο και πιο περίπλοκες και αφορούν μεγαλύτερες περιοχές της εικόνας, μέχρι να απαντήσουν στην τελική ερώτηση. Έτσι έχουμε αυτού του τύπου την ιεραρχική αναπαράσταση, από χαμηλού επιπέδου ερωτήσεις σε ψηλού επιπέδου.

Μια αντίστοιχη ιδέα, χρησιμοποιείται στα συστήματα βαθιάς μάθησης.

Η Βαθιά μάθηση αποτελείται από τεχνητούς νευρώνες (κόμβοι), οι οποίοι είναι διατεταγμένοι με τρόπο που δημιουργούν πολλαπλά επίπεδα επεξεργασίας.

Στα πρώτα επίπεδα, οι νευρώνες επιτελούν μια πολύ απλή τοπική επεξεργασία / λειτουργία.

Στη συνέχεια εξάγουν χαρακτηριστικά πιο υψηλού επιπέδου και που αφορούν ευρύτερες περιοχές της εικόνας και πιο σημασιολογική ερμηνεία της εικόνας, μέχρι να μπορούν να απαντήσουν τις ερωτήσεις και να κάνουν εκτιμήσεις όσον αφορά αν υπάρχει κτήριο, πισίνα, τί κτήριο είναι αυτό.

Ο κάθε κόμβος και ο κάθε τεχνητός νευρώνας σε αυτή τη διάταξη, δεν απαντά σε μια ερώτηση ή κάνει κάποια διεργασία για την οποία έχουμε προκαθορίσει εμείς εκ των προτέρων, αλλά κάνει κάποιους υπολογισμούς, τους οποίους τους μαθαίνει το δίκτυο από μόνο του κατά τη διάρκεια της εκπαίδευσης.

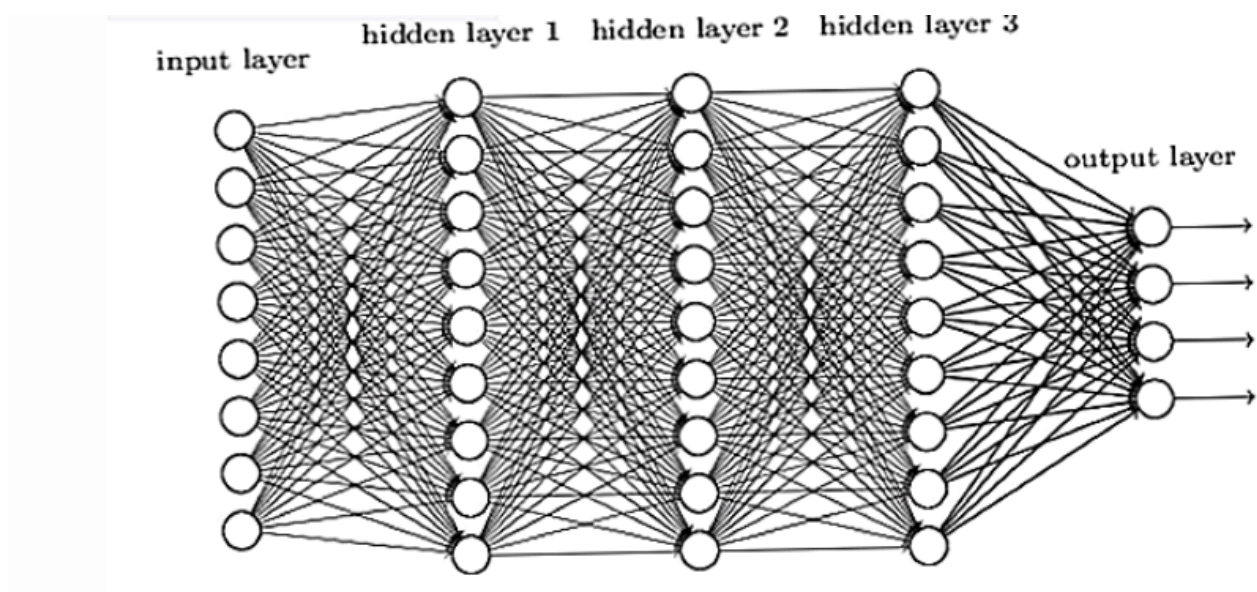


Image Credit: Νευρωνικά δίκτυα και βαθιά εκμάθηση από τον Michael A. Nielsen

Με μια πιο προσεκτική ματιά, στους νευρώνες, βλέπουμε ότι είσοδος είναι οι τεχνητοί νευρώνες του προηγούμενου επιπέδου.

Στη συνέχεια, κάνοντας απλές μαθηματικές πράξεις, εξάγει μια έξοδο, την οποία προωθεί στο επόμενο επίπεδο επεξεργασίας. Αυτό έχει κάποιες αντιστοιχίες με τους βιολογικούς νευρώνες.

Με αυτό τον τρόπο εξασφαλίζεται ότι κάτω από συγκεκριμένες συνθήκες των εισόδων ενεργοποιείται, αλλιώς παραμένει ανενεργός και βγάζει σαν έξοδο την τιμή 0.

Κάθε ένας από τους τεχνητούς νευρώνες επιτελεί μια εξαιρετικά απλή λειτουργία, έχοντας ένα σύνολο παραμέτρων που σχετίζεται με την απλή λειτουργία που εκτελεί.

Για την εκπαίδευση των δικτύων βαθιάς μάθησης, όλες οι παράμετροι πρέπει να προσδιοριστούν. Αυτό είναι εφικτό, χρησιμοποιώντας ένα μεγάλο αριθμό δεδομένων εκπαίδευσης - Μέγα-Δεδομένα.

Η επίτευξη της τεχνητής νοημοσύνης έρχεται από το συνδυασμό των τεχνικών νευρώνων στη δομή του νευρωνικού δικτύου. Στα σύγχρονα συστήματα βαθιάς μάθησης έχουμε εκατομμύρια/ δισεκατομμύρια τέτοιου είδους νευρώνες.

2.6 Συνελικτικά Νευρωνικά Δίκτυα - CNN (Convolutional Neural Networks)

Όπως και ο ανθρώπινος εγκέφαλος, έτσι και τα συνελικτικά νευρωνικά δίκτυα[22][24], παίρνουν είσοδο, κάνουν επεξεργασία και δίνουν μια έξοδο σαν αποτέλεσμα.

Τα συνελικτικά νευρωνικά δίκτυα, ανήκουν στην κατηγορία των βαθιών νευρωνικών δικτύων και χρησιμοποιούνται κυρίως σε προβλήματα σχετικά με την αναγνώριση εικόνων, για επεξεργασία, εντοπισμό και τμηματοποίηση, ανιχνεύοντας και κατανοώντας μοτίβα με πολύ ψηλή ακρίβεια.

Ένα συνελικτικό νευρωνικό δίκτυο, χρησιμοποιεί μια τεχνική που ονομάζεται συνέλιξη για να μειώσει τις παραμέτρους και να ενισχύσει την εκπαίδευση των δεδομένων. Συγκεκριμένα κάνει διάσπαση των εικόνων σε μορφή που να είναι πιο εύκολη η επεξεργασία τους, σε πολλά κρυφά επίπεδα, διατηρώντας τα χαρακτηριστικά τους, ώστε να επιτευχθεί μια σωστή πρόβλεψη για την εξαγωγή χαρακτηριστικών.

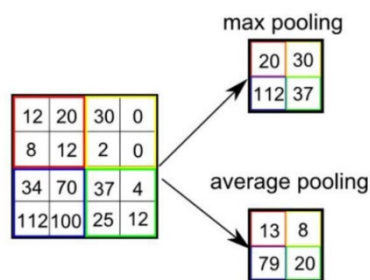
Τα δίκτυα αυτά συντίθενται από πολλαπλά επίπεδα τεχνητών νευρώνων (convolution layers), τα οποία είναι μαθηματικές συναρτήσεις που υπολογίζουν το άθροισμα των εισόδων και των εξόδων κάποιας τιμής ενεργοποίησης.

Κάθε επίπεδο δημιουργεί πολλές συναρτήσεις ενεργοποίησης που μεταβιβάζονται στο επόμενο επίπεδο.

Αρχίζοντας από το πρώτο επίπεδο, εξάγονται βασικά χαρακτηριστικά και στη συνέχεια στα επόμενα επίπεδα γίνεται ολοένα και πιο σύνθετη εξαγωγή χαρακτηριστικών, μέχρι στο τέλος να γίνει για παράδειγμα αναγνώριση αντικειμένων, όπου στην περίπτωση μας να επιτευχθεί η αναγνώριση σπιτιών και πισίνων.

Στο τελικό επίπεδο συνέλιξης εξάγεται ένα σύνολο βαθμολογιών εμπιστοσύνης (confidence scores) με τιμές 0-1, οι οποίες καθορίζουν την πιθανότητα εντοπισμού κάποιας κλάσης. Στην παρούσα μεταπτυχιακή διατριβή οι κλάσεις αυτές είναι τα σπίτια με και χωρίς οροφή και οι πισίνες.

Τέλος, το pooling layer είναι υπεύθυνο για τη μείωση της υπολογιστικής ισχύς που απαιτείται για την επεξεργασία των δεδομένων. Αυτό το επιτυγχάνει μειώνοντας τις διαστάσεις της εικόνας, διατηρώντας όμως τη χρήσιμη πληροφορία. Συνήθως αυτό επιτυγχάνεται με τη μέγιστη τιμή (ή σε κάποιες περιπτώσεις με τη μέση τιμή) σε ένα πίνακα 2 x 2 πάνω σε κάθε χάρτη χαρακτηριστικών. Ένα παράδειγμα απεικονίζεται στην πιο κάτω εικόνα:



Εικόνα: Δύο τύπο pooling layer στα CNN - Max pooling και average pooling

Κεφάλαιο 3

Πειραματική Διαχείριση

Το παρόν κεφαλαίο περιγράφει τη διαδικασία κατάτμησης των δεδομένων που συλλέχθηκαν.

Η γενική μεθοδολογία που ακολουθήθηκε στην παρούσα μεταπτυχιακή διατριβή περιγράφεται στις πιο κάτω υπό- ενότητες .

3.1 Συλλογή δεδομένων

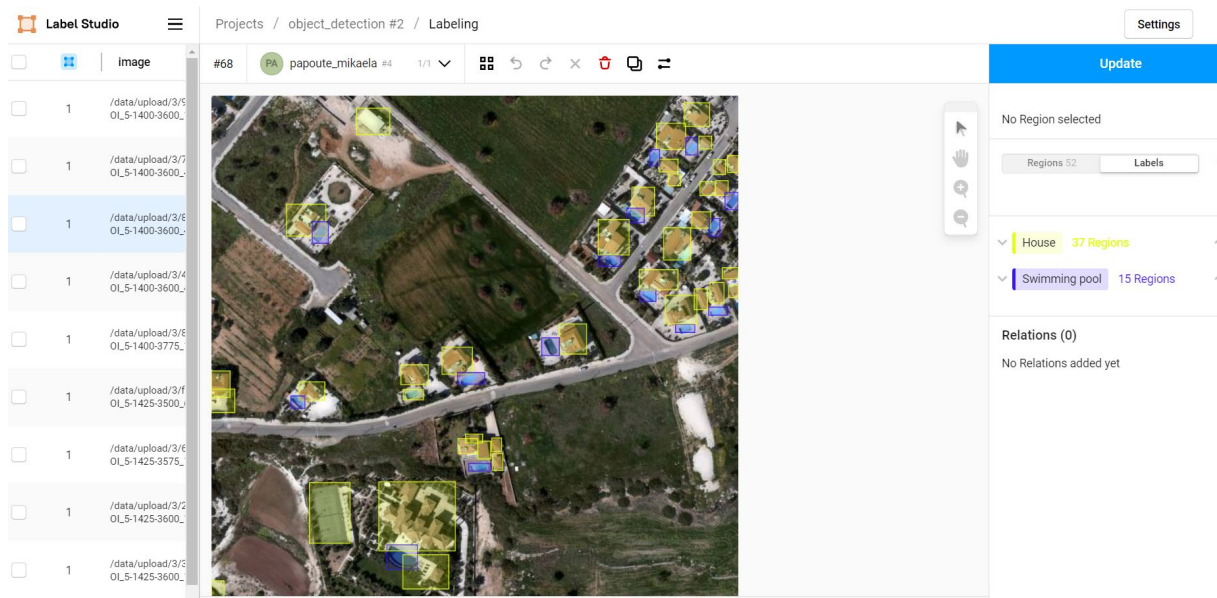
Το πρώτο στάδιο της πειραματικής διαχείρισης που ακολουθήθηκε , αφορά τη συλλογή δεδομένων. Για τους σκοπούς της παρούσας διατριβής, έχουν συλλεχθεί και αναλυθεί 5131 εικόνες δορυφόρου από το Κτηματολόγιο Κύπρου, οι οποίες είχαν ανάλυση 50cm/pixel.

Αξίζει να σημειωθεί η σημαντικότητα της συλλογής αξιόπιστων και ευδιάκριτων δεδομένων, ώστε το μοντέλο να μπορέσει να βρει τα σωστά μοτίβα και να είναι ακριβές.

3.2 Ετοιμασία Δεδομένων

Πολύ σημαντική φάση, αποτελεί η φάση της επεξεργασίας ακατέργαστων δεδομένων/ φωτογραφιών, που λήφθηκαν όπως προαναφέρθηκε στο πιο πάνω βήμα.

Για τον εντοπισμό των κατηγοριών/ κλάσεων στα δεδομένα/ εικόνες δορυφόρου που λήφθηκαν, χρησιμοποιήθηκε το εργαλείο Label-Studio .



Εικόνα: Ετοιμασία Δεδομένων με τη χρήση εργαλείου Label Studio

Για να μπορέσει να γίνει χρήση του, απαιτούσε αρχικά την εγκατάσταση του προγράμματος Anaconda Prompt και την εγκατάσταση της python version 3.7.

Στη συνέχεια έγινε η δημιουργία και η ενεργοποίηση ενός virtual environment, ώστε να μπορέσει να ανοίξει το συγκεκριμένο εργαλείο. Συγκεκριμένα χρησιμοποιήθηκαν οι εντολές:

```
conda activate DL_ENV  
label-studio
```

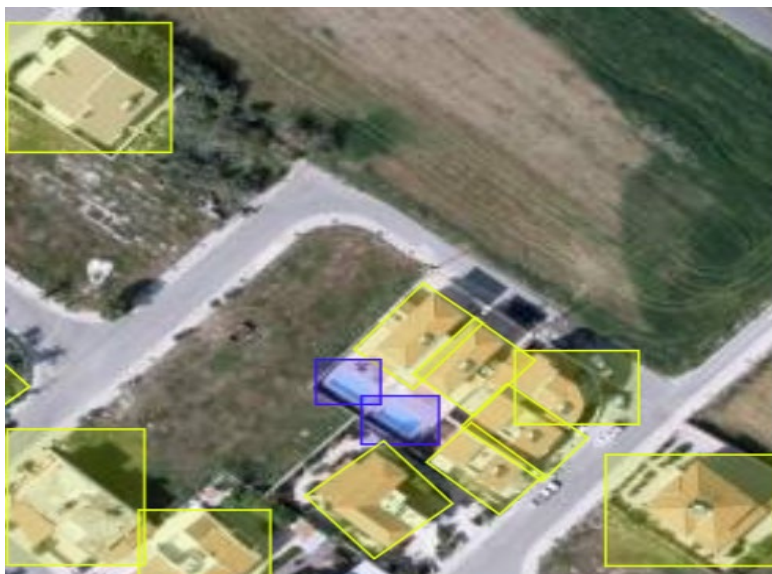
Το Label-Studio είναι ένα Open-source annotation tool, που επιτρέπει διαφορετικούς τύπους ετικετών/ annotations, σε οριοθετημένα κουτιά.

Πρόκειται για ένα φιλικό ως προς τον χρήστη web interface και οι κλάσεις/ labels που δημιουργούνται μπορούν να εξαχθούν σε διαφορετικές μορφές όπως CSV, JSON, COCO κτλ.

Η προετοιμασία των δεδομένων έγινε με τον εντοπισμό στην κάθε φωτογραφία των σπιτιών, πισίνων και εγκαταλελειμμένων σπιτιών που έπεσε η οροφή τους, με όλες τις πιθανές υποκατηγορίες/ τάξεις που μπορούσαν να παρουσιαστούν.

Συγκεκριμένα, δημιουργήθηκαν τρεις κλάσεις χρησιμοποιώντας συνολικά 5131 φωτογραφίες. Για την κατηγοριοποίηση τους έχουν επιλεγθεί 3 διαφορετικά χρώματα όπως αναφέρονται πιο κάτω:

1. Πισίνες - Μπλε χρώμα
2. Σπίτια /Κτήρια – Κίτρινο Χρώμα
3. Παλιά εγκαταλελειμμένα κτήρια χωρίς – Πράσινο χρώμα



Εικόνα: Ετοιμασία Δεδομένων για σπίτια και πισίνες- δείγμα παρούσα μεταπτυχιακή διατριβή



Εικόνα: Ετοιμασία Δεδομένων για παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή- δείγμα παρούσα μεταπτυχιακή διατριβή

Για το συγκεκριμένο πείραμα, η εξαγωγή των δεδομένων, έγινε σε μορφή . COCO.

3.3 Επιλογή μοντέλου για βαθιά μάθηση: UNet model

Η συγκεκριμένη φάση, περιλαμβάνει αλγορίθμους βαθιάς μάθησης, ώστε να επιτευχθεί η ομαδοποίηση των τριών κατηγοριών στις οποίες βασίζεται η παρούσα μεταπτυχιακή διατριβή.

Συγκεκριμένα, για το σκοπό αυτό, έχει χρησιμοποιηθεί η αρχιτεκτονική UNET[1], η οποία προσφέρει πολύ καλή και γρήγορη απόδοση στην τμηματοποίηση εικόνων.

Το μοντέλο UNET είναι ένα πλήρως συνελικτικό δίκτυο (CNN), με εξαιρετικές επιδόσεις διαμέριση εικόνων. Το μοντέλο αυτό, αρχικά αναπτύχθηκε στο τμήμα της Επιστήμης Υπολογιστών του Πανεπιστημίου του Φράιμπουργκ, και ήταν εξειδικευμένο στην τμηματοποίηση εικόνων με πολύ υψηλή ακρίβεια σε βιοϊατρικές εργασίες.

Ένα βασικό της πλεονέκτημα του είναι, ότι μπορεί να μάθει και να εκπαιδευτεί, χρησιμοποιώντας μικρό αριθμό εικόνων, δίνοντας αρκετά καλά αποτελέσματα. Επιπλέον για να εκπαιδευτεί απαιτεί λιγότερη ώρα και μικρότερη υπολογιστική ισχύ, σε αντίθεση με τις υπόλοιπες αρχιτεκτονικές.

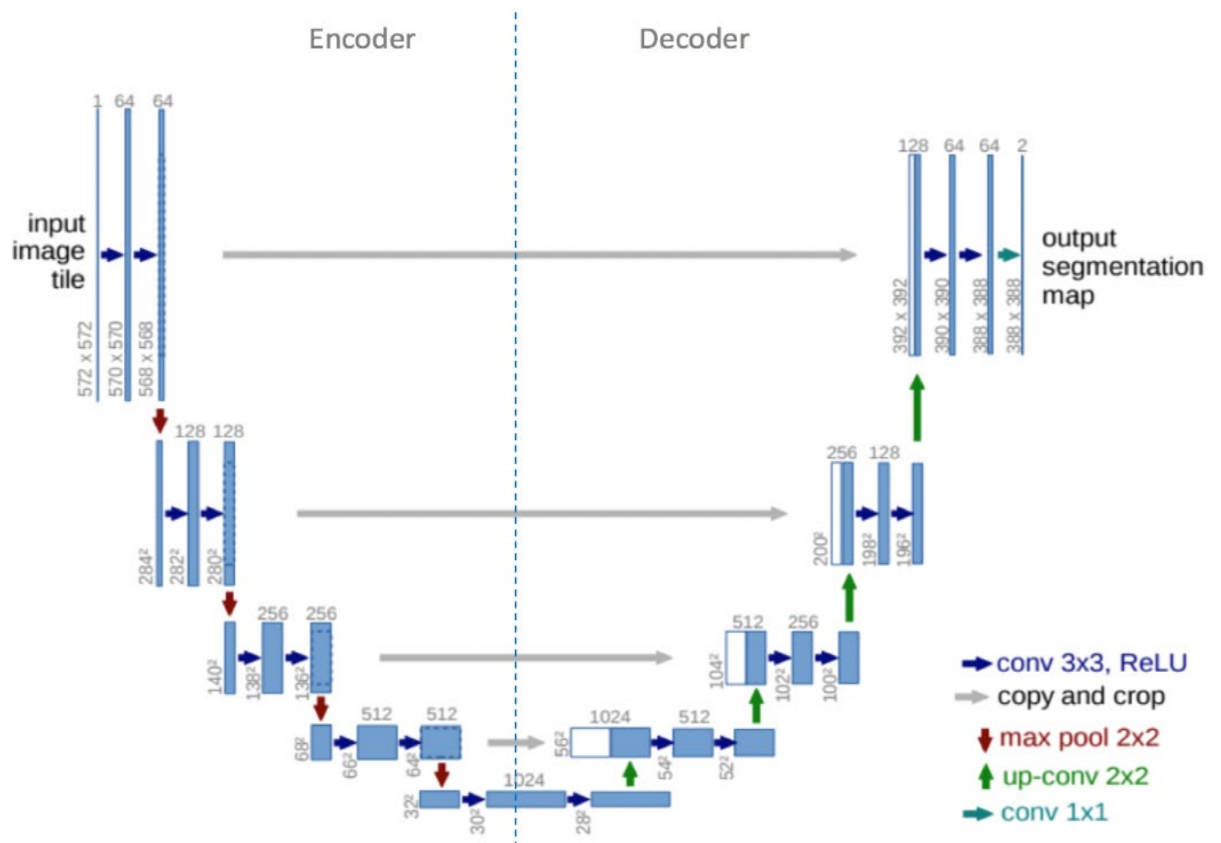
Η αρχιτεκτονική UNET, αποτελείται από 2 μέρη.

Ο πρώτο μέρος αφορά την κωδικοποίηση της εικόνας-εισόδου σε αναπαραστάσεις χαρακτηριστικών σε πολλά διαφορετικά επίπεδα- Συστελλόμενο μονοπάτι.

Το δεύτερο μέρος της αρχιτεκτονικής, αφορά την αποκωδικοποίηση-διαστελλόμενο μονοπάτι. Σε αυτό το μέρος της αρχιτεκτονικής γίνεται ο εντοπισμός και η χαρτογράφηση όλων των κλάσεων για τις οποίες έχει εκπαιδευτεί το μοντέλο.

Συγκεκριμένα η παρούσα πειραματική μελέτη, κάνει σημασιολογική κατάτμηση για τις κλάσεις των τριών αντικείμενων ενδιαφέροντος : πισίνες, σπίτια, παλιά εγκαταλελειμμένα σπίτια χωρίς οροφή.

Πέρα από την αναγνώριση των κλάσεων στις φωτογραφίες, η σημασιολογική κατάτμηση, δίνει τη δυνατότητα να δημιουργηθεί ένας ταξινομημένος χάρτης των τριών κλάσεων που αναφέρθηκαν πιο πάνω.



παράδειγμα αρχιτεκτονικής UNET

Όπως φαίνεται στην πιο πάνω εικόνα, η αρχιτεκτονική παίρνει ένα σχήμα U, στο οποίο οφείλεται και το όνομα της αρχιτεκτονικής U-NET.

Με γκριζο χρώμα απεικονίζονται οι συνδέσεις παράκαμψης, οι οποίες μεταφέρουν τις πληροφορίες από την δεξιά μεριά του κωδικοποιητή στην αριστερή μεριά του αποκωδικοποιητή.

Αρχικά οι χάρτες των χαρακτηριστικών του κωδικοποιητή, καταγράφουν λεπτομέρειες χαμηλού επιπέδου, όπως το περίγραμμα των αντικειμένων.

Προχωρώντας σε βάθος, η αρχιτεκτονική καταγράφει πληροφορίες υψηλού επιπέδου, όπως το σχήμα και τις κατηγορίες των αντικειμένων.

Για την τμηματοποίηση των αντικειμένων απαιτούνται τόσο οι πληροφορίες χαμηλού, όσο και οι πληροφορίες υψηλού επιπέδου. Έτσι χρησιμοποιούνται οι συνδέσεις

παράκαμψης, κάτι το οποίο μας δίνει τη δυνατότητα να παίρνουμε πληροφορίες ενδιάμεσων χαρακτηριστικών σε διάφορα βάθη από την δεξιά μεριά και να τις συνενώνουμε με την αριστερή, για την επεξεργασία και τη διευκόλυνση καλύτερων προβλέψεων.

Συγκεκριμένα, το δίκτυο δημιουργεί μια εικόνα, η οποία χρησιμοποιείται για την εκπαίδευση του δικτύου και η οποία μας δείχνει τη δομημένη και την αδόμητη περιοχή σε δυαδική μορφή.

Με μαύρο χρώμα, συμβολίζονται τα pixels που ανήκουν στην κλάση «έδαφος-background» και με άσπρο χρώμα, συμβολίζονται τα pixels που ανήκουν στην κλάση «κτήριο».

Ο εντοπισμός των σπιτιών σε κάθε εικόνα, όπως αναφέρθηκε πιο πάνω, ονομάζεται μάσκα. Οι μάσκες αυτές, σύμφωνα με τον αλγόριθμο στοχαστικής κλίσης, χρησιμοποιούνται για την εκπαίδευση του δικτύου.

Στις ακόλουθες εικόνες απεικονίζεται ένα παράδειγμα κατά την εξαγωγή κτηρίων που χρησιμοποιήθηκε στην παρούσα μεταπτυχιακή διατριβή.

Η πρώτη εικόνα απεικονίζει την πραγματική δορυφορική εικόνα RGB. Κάθε μια από αυτές τις εικόνες, αποτελεί μια οπτική αναπαράσταση κάποιων αντικειμένων και αποτελείται από pixels. Το κάθε χρώμα σε κάθε ένα από αυτά τα pixels χαρακτηρίζεται από ένα αριθμό RGB, που αναπαριστά τα 3 βασικά χρώματα Κόκκινο, Πράσινο, Μπλε και την ένταση του κάθε ενός από αυτά τα χρώματα που κυμαίνεται από 0-255. Κατά την ανάμιξη τους γίνεται η απεικόνιση της RGB εικόνας. Κάπως έτσι λειτουργεί και το μάτι του ανθρώπου.

Η δεύτερη εικόνα απεικονίζει τις μάσκες των σπιτιών.



Εικόνα : δορυφορική , RGB εικόνα- δείγμα παρούσα μεταπτυχιακή διατριβή



Εικόνα: μάσκα της πιο πάνω εικόνας-δείγμα παρούσα μεταπτυχιακή διατριβή

Στη συνέχεια, αφού το δίκτυο έχει εκπαιδευτεί, κατά την εισαγωγή οποιασδήποτε φωτογραφίας σε αυτό το μοντέλο, μπορεί αυτόματα να μας δώσει σαν έξοδο τις δυαδικές εικόνες που αναφέρθηκαν πιο πάνω.

3.3.1 Εκπαίδευση Αλγορίθμου

Το δίκτυο UNet για να εκπαιδευτεί σωστά υπάρχουν διάφοροι παράμετροι, οι οποίοι είναι σημαντικοί για την εκπαίδευση του.

Σε αυτό το σημείο, θα αναφερθούν κάποιες από αυτές τις παραμέτρους που χρησιμοποιήθηκαν στην παρούσα μεταπτυχιακή διατριβή.

Η μονάδα επεξεργασίας GPU ήταν το NVIDIA GeForce RTX 2060. Η μονάδα επεξεργασίας στα νευρωνικά δίκτυα είναι πολύ σημαντική. Αυτή εξοικονομεί χρόνο στην εκπαίδευση, αφού τα στοιχεία επεξεργάζονται παράλληλα σε μικρά χρονικά διαστήματα.

Οι συστάδες εικόνων (batch size) ήταν 64. Αυτός ο αριθμός δείχνει πόσες εικόνες χρησιμοποιούνται σε κάθε κύκλο κάθε φορά ώστε να επιτευχθεί η εκπαίδευση του δικτύου.

Ο ρυθμός εκμάθησης-learning Rate στην παρούσα εκπαίδευση του δικτύου, ήταν από 0.1-0.9. Ο ρυθμός είναι πολύ σημαντικός, αφού αυτός προσδιορίζει το ρυθμό μεταφοράς προς περιοχές με καλύτερες λύσεις..

Συγκεκριμένα, για μικρές τιμές του ρυθμού εκπαίδευσης, χρειάζονται περισσότερες επαναλήψεις, αφού γίνονται μικρές αλλαγές τα βάρη του δικτύου, μέχρι να επιτευχθεί καλύτερη τιμή.

Επιπλέον για πιο μεγάλες τιμές ρυθμού εκπαίδευσης, η διαδικασία είναι πιο γρήγορη, αλλά είναι πιθανόν να δημιουργηθεί ένα δίκτυο που είναι ασταθές ή και να παραβλέψουμε τη βέλτιστη λύση.

```
# compile keras model with defined optimizer, loss and metrics  
optimizer = tf.keras.optimizers.SGD(lr=0.1, momentum = 0.9)
```

Ο αλγόριθμος βελτιστοποίησης-optimizer που χρησιμοποιήθηκε για να εκπαιδευτεί το δίκτυο είναι ο αλγόριθμος οπισθοδιάδοσης- Stochastic Gradient Descent.

Συγκεκριμένα ο αλγόριθμος αυτός επιλέγει τυχαία ένα αριθμό δειγμάτων από το σύνολο δεδομένων (batches) για κάθε επανάληψη και κάνει ταξινόμηση. Με αυτό τον τρόπο, ανανεώνοντας κάθε φορά τις παραμέτρους, δε χρειάζεται να υπολογίζει συνεχώς τις κλίσεις για όλα τα δεδομένα. Αυτό κάνει το αλγόριθμο πολύ γρήγορο, σε σχέση άλλους βασικούς αλγορίθμους της ίδιας κατηγορίας. Επιπλέον με την συνεχή ανανέωση των

παραμέτρων που χρησιμοποιεί, δίνει καλύτερες τιμές στη συνάρτηση κόστους. Ο αλγόριθμος αυτός μπορεί να θεωρηθεί ως σχετικά θορυβώδης, καθώς χρειάζονται πιο πολλές επαναλήψεις, λόγω της τυχαίας επιλογής δειγμάτων στις επαναλήψεις του.

```
for i in range (m) :  
 $\theta_j = \theta_j - \alpha (\hat{y}^i - y^i) x_j^i$ 
```

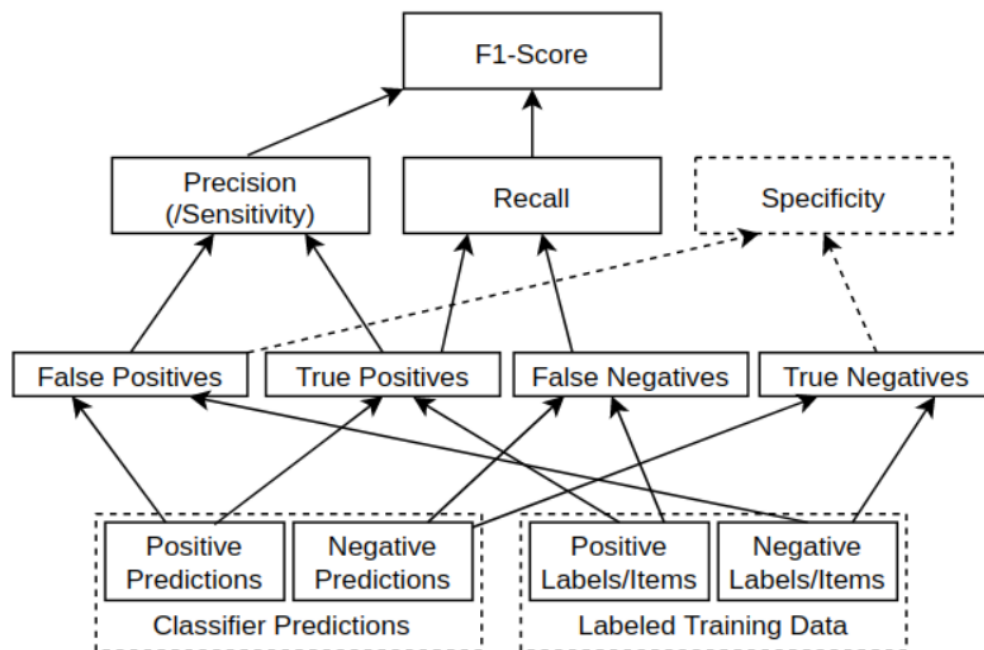
Ο χρόνος για κάθε epoch/ εποχή ήταν 18 sec και έγιναν συνολικά 207 epochs. Σε κάθε epoch/ εποχή τα δεδομένα που χρησιμοποιούνται για εκπαίδευση, ενημερώνουν τις εσωτερικές παραμέτρους του μοντέλου. Αριθμός των Epochs προσδιορίζεται ως ο αριθμός των φορών που χρειάζεται ο αλγόριθμος για να καλύψει όλα τα δεδομένα εκπαίδευσης. Αυτό μας δίνει σαν Συνολικό Χρόνο $18 \times 207 = 3726$ Sec

Τέλος, ο προγραμματισμός έγινε σε γλώσσα python, που εκτός ότι πρόκειται για μια open source γλώσσα προγραμματισμού, έχει αρκετές βιβλιοθήκες που εξυπηρετούν το πρόβλημα μας.

3.4 Αξιολόγηση αποτελεσμάτων - Μέτρηση απόδοσης ακρίβειας

Για την αξιολόγηση των ικανοτήτων του αλγορίθμου, χρησιμοποιείτε κάποιο ποσοτικό μέτρο απόδοσης. Αυτό αφορά τους δείκτες accuracy, precision, recall, F1 Score και IOU score. Αυτοί οι δείκτες μας δίνουν μια πιο αναλυτική ιδέα, για το πόσο αποδοτικός είναι ο αλγόριθμος που χρησιμοποιήθηκε.

Το πιο κάτω διάγραμμα απεικονίζει τη σχέση των δεικτών Precision, Recall και F1 Score, που θα αναλυθούν πιο κάτω[16].



Συγκεκριμένα,

TP – True Positives ορίζεται ως: Ορθά θετικά ορθός υπολογισμός- εντοπισμός πσίνας, ή εντοπισμός κτήριου και στις προβλέψεις, και στην πραγματικότητα

TN - True Negatives ορίζεται ως: Ορθά αρνητικά ορθός υπολογισμός- όχι πσίνα, όχι κτήριο ούτε στις προβλέψεις, ούτε την πραγματικότητα

FP - False Positives (ή σφάλμα TYPE I.) ορίζεται ως: Λανθασμένα Θετικά ορθός υπολογισμός - εντοπισμός πσίνας, ή εντοπισμός κτήριου στις προβλέψεις, αλλά όχι στην πραγματικότητα

FN - False Negatives (ή σφάλμα TYPE II .) ορίζεται ως: Λανθασμένα Αρνητικά ορθός υπολογισμός - όχι πσίνα, όχι κτήριο στις προβλέψεις, αλλά υπάρχουν στην πραγματικότητα

Οι πιο πάνω τιμές αποτελούν παραμέτρους ενός πίνακα σύγχυσης.

3.4.1 Δείκτης Ακριβείας/ precision

Το συγκεκριμένο μέτρο απόδοσης, δίνει τη λάθος πρόβλεψη της αρνητικής κλάσης. Προβλέπονται ως θετικά σημεία, που στην πραγματικότητα είναι αρνητικά - False Positive[16]. Το FP ονομάζεται επίσης σφάλμα TYPE I.

Ο δείκτης αυτός υπολογίζεται από το λόγο του αριθμού των σημείων που είναι σωστά προβλεπόμενα θετικά ως προς το σύνολο των θετικών σημείων που αναφέρθηκαν αρχικά. Αυτό απεικονίζεται στον πιο κάτω τύπο:

$$Precision = \frac{TP}{TP + FP}$$

Συγκεκριμένα απαντάει στην ερώτηση, «από το σύνολο όλων των σπίτια ή πισίνων που εντοπίστηκαν, πόσα από αυτά είναι πραγματικά σπίτια ή πισίνες;»

Ο δείκτης ακρίβειας, μπορεί να θεωρηθεί καλό μέτρο όταν υπάρχουν σύνολα πολύ ισορροπημένα.

Επιπλέον ο δείκτης αυτός είναι σημαντικός όταν θέλουμε να εστιάσουμε και να ελαχιστοποιήσουμε τα ψευδώς θετικά αποτελέσματα και να στοχεύσουμε στη μεγιστοποίηση της ακρίβειας.

3.4.2 Δείκτης Ανάκλησης / recall

Το μέτρο αυτό, προβλέπει λανθασμένα τη θετική κλάση.

Προβλέπονται ως αρνητικά σημεία που στην πραγματικότητα είναι θετικά. False Negative (FN)[16]. Το FN ονομάζεται επίσης σφάλμα TYPE II.

Ο δείκτης ανάκλησης υπολογίζεται από το λόγο του αριθμού των σημείων που είναι σωστά προβλεπόμενα θετικά ως προς το άθροισμα των θετικών σημείων που

αναφέρθηκαν αρχικά και των σημείων, τα οποία θα έπρεπε να αναφερθούν ως θετικά αλλά αναφέρθηκαν λανθασμένα ως αρνητικά (όλα τα πραγματικά σωστά σημεία). Αυτό απεικονίζεται στον πιο κάτω τύπο:

$$Recall = \frac{TP}{TP + FN}$$

Συγκεκριμένα απαντάει στην ερώτηση, « από το σύνολο όλων των σπιτιών ή πισίνων που υπάρχουν, πόσα από αυτά είναι εντοπίστηκαν;»

Ο δείκτης ανάκλησης, μπορεί να θεωρηθεί καλό μέτρο σε τομείς όπως η ιατρική, σε εντοπισμό διαφόρων ασθενειών, όπου η πιθανότητα απώλειας περιπτώσεων τα οποία είναι θετικά πρέπει να ελαχιστοποιηθεί .

Επιπλέον ο δείκτης αυτός είναι σημαντικός όταν θέλουμε να εστιάσουμε και να ελαχιστοποιήσουμε τα ψευδώς θετικά αποτελέσματα και να στοχεύσουμε στη μεγιστοποίηση της ακρίβειας.

Με βάση τα πιο πάνω σχετικά με την Ακρίβεια και την Ανάκληση (precision and recall), βλέπουμε ότι από όλα τα θετικά αποτελέσματα, το μοντέλο βρίσκει το πραγματικά θετικό ποσοστό.

Οι τιμές και του precision και του recall κυμαίνονται μεταξύ 0 και 1.

3.4.3 Δείκτης F1 Score

Το μέτρο αυτό, είναι μια ενδιάμεση μέτρηση της ακρίβειας και της ανάκλησης.

Ο δείκτης F1 Score, είναι πολύ σημαντικός, αφού μας δείχνει τη γενική εικόνα και όχι μεμονωμένα την *ακρίβεια* και την *ανάκληση*. Λαμβάνει υπόψη τόσο τα ψευδώς θετικά όσο και τα ψευδώς αρνητικά[16].

Αυτός απεικονίζεται στον πιο κάτω τύπο:

$$F1\ score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2 * (Precision * Recall)}{(Precision + Recall)}$$

Η καλύτερη τιμή είναι 1 και η χειρότερη 0

Θετικά στοιχεία του Δείκτη F1 Score είναι ότι έχει καλή απόδοση σε κάποιο σύολο που δεν είναι ισορροπημένο.

3.4.4 Δείκτης Accuracy

Το μέτρο αυτό μας δίνει πόσο κοντά είμαστε στην ακριβή τιμή.

Ο δείκτης Accuracy, υπολογίζεται από όλες τις τιμές TP, TN, FP, FN που απεικονίζονται στον πίνακα σύγχυσης, ο οποίος παρουσιάστηκε πιο πάνω.

Συγκεκριμένα, υπολογίζει την προβλεπόμενη εικόνα του κάθε μοντέλου και τη θεωρητικά σωστή εικόνα.

Ο Δείκτης ακρίβειας, υπολογίζεται με τον πιο κάτω τύπο:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

Accuracy = Number of correctly detected pixels / Total Number of actual pixels

3.4.5 Δείκτης IOU

Το μέτρο αυτό δίνει το λόγο της τομής του συνόλου των εικόνων που πραγματικά απεικονίζουν τη κλάση που θέλουμε (X) και του συνόλου των εικόνων με τα προβλεπόμενα δυτικά (Y), ως προς την ένωση των δύο αυτών συνόλων[23].

Ο Δείκτης IOU, υπολογίζεται με τον πιο κάτω τύπο:

$$IoU = \frac{\text{Intersection } X \cap Y}{\text{Union } X \cup Y}$$

Κεφάλαιο 4

Αποτελέσματα

4.1 Αποτελέσματα μοντέλου για πισίνες:

Το μοντέλο εκπαιδεύτηκε χρησιμοποιώντας εικόνες για να προβλέψει black και white pixel.

Συνολικά έγιναν δύο εκπαιδεύσεις στο συγκεκριμένο μοντέλο.

Ο αριθμός δειγμάτων που χρησιμοποιήθηκαν για το Μοντέλο 1 -πρώτο training- ήταν συνολικά 1863 και για το Μοντέλο 2 -δεύτερο training- ήταν συνολικά 1923 δείγματα.

Το άσπρο(white) και το μαύρο (black) pixel αναφέρεται στα annotation για τις πισίνες.

Συγκεκριμένα,

White	αναφέρεται ως η περιοχή όπου βρίσκονται πισίνες (W-swimming pool)
Black	αναφέρεται ως η περιοχή που δεν βρίσκονται πισίνες (B-background)

Πιο αναλυτικά, για το Μοντέλο 1 και Μοντέλο 2, χρησιμοποιήθηκαν:

Για εκπαίδευση	80% των συνολικών δεδομένων
Για testing	10% των συνολικών δεδομένων
Για επικύρωση- Validation	10% των συνολικών δεδομένων

Ο πιο κάτω πίνακας απεικονίζει τις τιμές των precision “P” και recall “R” που υπολογίστηκαν για τα 2 μοντέλα που χρησιμοποιήθηκαν:

Class	Model1		Model2	
	P	R	P	R
Black	1.00	0.99	1.00	1.00
White	0.64	0.86	0.72	0.77
Average	0.82	0.93	0.86	0.88

Εικόνα: Πίνακας που απεικονίζει τις τιμές των precision “P” και recall “R” για τα 2 Μοντέλα.

Βάση των πιο πάνω τιμών υπολογίζουμε τον δείκτη F1 Score όπως πιο κάτω:

F1 Score		
Μοντέλο 1	1,5252/1,75	0.872
Μοντέλο 2	1,5136/1,75	0.870

4.2 Αποτελέσματα μοντέλου για Σπίτια:

Συνολικά έγιναν δύο εκπαιδεύσεις στο συγκεκριμένο μοντέλο για αναγνώριση σπιτιών.

Ο αριθμός δειγμάτων που χρησιμοποιήθηκαν για το Μοντέλο 1 - πρώτο training- ήταν περίπου 30000 και για το Μοντέλο 2- δεύτερο training- ήταν συνολικά 80100 δείγματα (801 εικόνες*100 δείγματα η κάθε εικόνα).

Η Class 1 (Κλάση 1) και η Class 2 (Κλάση 2) αναφέρονται στα δείγματα για τα σπίτια. Συγκεκριμένα,

Class 1	αναφέρεται σε περιοχή όπου δε βρίσκονται σπίτια (background)
Class 2	αναφέρεται σε περιοχή όπου βρίσκονται σπίτια (House)

Πιο αναλυτικά, για το Μοντέλο 1 και Μοντέλο 2, χρησιμοποιήθηκαν:

Για εκπαίδευση	70% των συνολικών δεδομένων
Για testing	10% των συνολικών δεδομένων
Για επικύρωση- Validation	20% των συνολικών δεδομένων

Αξιολόγηση μοντέλου για σπίτια έγινε με τη χρήση του δείκτη IOU:

IOU	
Μοντέλο 1	0.62
Μοντέλο 2	0.68

Επιπλέον ο δείκτης Precision για τα 2 μοντέλα απεικονίζονται πιο κάτω:

Precision	
Μοντέλο 1	0,72
Μοντέλο 2	0,81

4.3 Αποτελέσματα μοντέλου για Παλιά εγκαταλειμμένα κτήρια χωρίς οροφή:

Το μοντέλο εκπαιδεύτηκε χρησιμοποιώντας 280 δείγματα/ annotations.

Η Class 1 (Κλάση 1) και η Class 2 (Κλάση 2) αναφέρονται στα annotation για τις παλιά εγκαταλειμμένα κτήρια χωρίς οροφή.

Συγκεκριμένα,

Class 1	αναφέρεται σε περιοχή όπου δε βρίσκονται παλιά εγκαταλελειμμένα κτήρια χωρίς οροφή (background)
Class 2	αναφέρεται σε περιοχή όπου βρίσκονται παλιά εγκαταλελειμμένα κτήρια χωρίς οροφή (House)

Αξιολόγηση μοντέλου για παλιά εγκαταλελειμμένα σπίτια χωρίς οροφή έγινε με τη χρήση του δείκτη IOU:

IOU
0.94

Επιπλέον ο δείκτης Presicion του μοντέλου για εντοπισμό παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή είναι:

Presicion
0.94



Κεφάλαιο 5

Συζήτηση

Παρατηρώντας σε διάφορες φάσεις τα αποτελέσματα του αλγορίθμου, συμπεραίνουμε ότι ο αλγόριθμος εντοπίζει σε αρκετά καλό βαθμό τις πισίνες, σπίτια και παλιά κτήρια χωρίς οροφή και δεν επηρεάζεται σε ψηλό βαθμό από τη δόμηση που μπορεί να έχει κάποια περιοχή.

Συγκεκριμένα, για τις πισίνες, το μοντέλο εκπαιδεύτηκε με ακρίβεια 0,87% με τη χρήση 1923 annotations.

Κατά την αξιολόγηση βλέπουμε ότι υπήρχαν περιπτώσεις όπου εντοπίστηκαν πισίνες, χωρίς να υπάρχουν στην πραγματικότητα και αντίστοιχα ενώ στην πραγματικότητα υπήρχαν πισίνες δεν έγινε ο εντοπισμός τους.

Αυτό οφείλεται σε διάφορες περιπτώσεις, κάποιες από τις οποίες είναι τα σχήματα των πισινών, τα οποία σε αρκετές περιπτώσεις είναι ακανόνιστα, κάποιες από τις πισίνες επίσης είναι καλυμμένες, σε κάποιες φωτογραφίες λόγω της ύπαρξης ψηλών δένδρων μπορεί να μην είναι ευδιάκριτες, καθώς επίσης υπήρχαν περιπτώσεις όπου οι πισίνες ήταν πολύ μικρές και όχι πολύ ευδιάκριτες.

Αντίστοιχα, για τον εντοπισμό σπιτιών, το μοντέλο εκπαιδεύτηκε με ακρίβεια 0,81 % με τη χρήση 80100 annotations.

Κατά την αξιολόγηση βλέπουμε ότι υπήρχαν περιπτώσεις όπου εντοπίστηκαν σπίτια, χωρίς να υπάρχουν στην πραγματικότητα και αντίστοιχα ενώ στην πραγματικότητα υπήρχαν σπίτια, δεν έγινε ο εντοπισμός τους.

Αυτό οφείλεται σε περιπτώσεις όπου υπάρχουν θερμοκήπια ή φάρμες ή στέγαστρα, των οποίων η οροφή τους μοιάζει με αυτή των σπιτιών. Ακόμα, εντοπίστηκε ότι το μοντέλο δυσχεραίνεται στο να κάνει τη σωστή πρόβλεψη σε περιπτώσεις με συγκροτήματα σπιτιών ή σπίτια μεζονέτες, τα οποία εφάπτονται το ένα πάνω στο άλλο. Τέλος σε

κάποιες εικόνες δορυφόρου, ο εντοπισμός τους επηρεάζεται λόγω της ύπαρξης ψηλών δένδρων κάτι τι οποίο δεν τα κάνει ευδιάκριτα .

Τέλος, για εντοπισμό παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή, το μοντέλο εκπαιδεύτηκε με εξαιρετική ακρίβεια 94%.

5.1 Συγκρίσεις με άλλες πειραματικές μελέτες που έλυσαν το πρόβλημα για αναγνώριση αντικειμένων ίδιας κατηγορίας

5.1.1 Εντοπισμός πισίνων:

Στην μελέτη [19] και [20], βλέπουμε τα πιο κάτω πειραματικά αποτελέσματα σε σχέση με τα πειραματικά αποτελέσματα που προέκυψαν από την παρούσα μεταπτυχιακή διατριβή:

Μελέτη για αναγνώριση πισίνων στην μελέτη [19] και μελέτη [20]	Παρούσα μελέτη για αναγνώριση σπιτιών
Precision : 0.98	Precision: 0.86

Βάση του πιο πάνω πίνακα βλέπουμε ότι το μοντέλο που χρησιμοποιήθηκε στην παρούσα μεταπτυχιακή διατριβή έχει μικρότερες τιμές συγκρίνοντας τις τιμές του Precision παρουσιάζονται.

Αυτό μπορεί να οφείλεται σε διάφορους παράγοντες που σχετίζονται με την ανάλυση των εικόνων δορυφόρου που χρησιμοποιήθηκαν.

Συγκεκριμένα, στις μελέτες [19] και [20] βλέπουμε ότι οι πισίνες στις εικόνες δορυφόρου είναι πολύ πιο ευδιάκριτες, αφού οι εικόνες είχαν καλύτερη ευκρίνεια από αυτές που μας δόθηκαν από το Κτηματολόγιο Κύπρου και οι οποίες χρησιμοποιήθηκαν στην παρούσα πειραματική μελέτη.

Επίσης κατά την προεργασία δεδομένων στην παρούσα μεταπτυχιακή διατριβή εντοπίστηκαν παράλληλα στην κάθε εικόνα δορυφόρου 3 διαφορετικές κατηγορίες αντικειμένων(σπίτια και πισίνες), που στις πιο πολλές περιπτώσεις ήταν το ένα μέσα στο άλλο, όπως για παράδειγμα πισίνες μέσα σε σπίτια ή ξενοδοχεία. Έτσι έκανε τον εντοπισμό τους μη ευδιάκριτο και το «τετράγωνο» εντοπισμού κατά την προεργασία, να μην εφάπτεται ακριβώς στο αντικείμενο «πισίνα». Σε αντίθεση για τις μελέτες [19] και [20] στην κάθε εικόνα έγινε εντοπισμός μόνο πισινών και κατά συνέπεια το «τετράγωνο» εντοπισμού κατά την προεργασία τους είχε μεγαλύτερη εμβέλεια.

Αυτό φαίνεται στις πιο κάτω εικόνες:



Εικόνες: Αριστερά προεργασία δεδομένων από μελέτη [20] και Δεξιά προεργασία δεδομένων από παρούσα μεταπτυχιακή διατριβή

Τέλος μπορεί να οφείλεται και στον αριθμό δεδομένων τα οποία για την κατηγορία πισίνες ήταν λιγότερα σε σχέση με αυτά των άλλων 2 μελετών που συγκρίθηκαν.

5.1.2 Εντοπισμός σπιτιών:

Στην μελέτη [17], στην οποία χρησιμοποιήθηκε επίσης το μοντέλο UNET βλέπουμε τα πιο κάτω πειραματικά αποτελέσματα σε σχέση με τα πειραματικά αποτελέσματα που προέκυψαν από την παρούσα μεταπτυχιακή διατριβή:

Μελέτη για αναγνώριση σπιτιών με τη χρήση μοντέλου UNET[17]	Παρούσα μελέτη για αναγνώριση σπιτιών με τη χρήση μοντέλου UNET
Presicion: 0.81	Presicion: 0.81

Βάση του πιο πάνω πίνακα βλέπουμε ότι το μοντέλο που χρησιμοποιήθηκε στην παρούσα μεταπτυχιακή διατριβή έχει ίδιες τιμές συγκρίνοντας τις τιμές Precision.

5.1.3 Εντοπισμός κατεστραμμένων σπιτιών / παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή

Στην μελέτη [21], βλέπουμε τα πιο κάτω πειραματικά αποτελέσματα σε σχέση με τα πειραματικά αποτελέσματα που προέκυψαν από την παρούσα μεταπτυχιακή διατριβή:

Table 1. The comparison for the performances of the four models based on the metrics Precision (%), Recall (%), F1 (%), AP (%), FPS (f/s), Training Time (h), Parameter Size (MB).

Model	Precision (%)	Recall (%)	F1 (%)	AP (%)	FPS (f/s)	Training Time (h)	Parameter Size (MB)
YOLOv3	-	-	-	81.63	27	5.6	235
YOLOv5s	84.97	84.56	84.76	89.17	100	1.417	14.4
YOLOv5s-ViT	88.46	87.34	87.90	90.40	100	1.467	14.4
YOLOv5s-ViT-BiFPN	89.01	89.37	89.19	90.94	80	1.203	16.5

Μελέτη για αναγνώριση / παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή[21]	Παρούσα μελέτη για αναγνώριση σπιτιών
Precision: 0.89	Precision: 0.94

Εδώ βλέπουμε ότι το μοντέλο που χρησιμοποιήθηκε στην παρούσα μεταπτυχιακή διατριβή έχει καλύτερες τιμές .

Κεφάλαιο 6

Επίλογος

Η παρούσα μεταπτυχιακή διατριβή, επίλυσε το πρόβλημα της αυτόματης ανίχνευσης και εξαγωγής κτηρίων, πισίνων, παλιών εγκαταλελειμμένων σπιτιών χωρίς οροφή, μέσα από δορυφορικές εικόνες/ εικόνες δορυφόρου.

Συγκεκριμένα, αποδείχτηκε ότι με λίγη προ-επεξεργασία, χωρίς τη χρήση αισθητήρων και κατ'έκταση με μικρό οικονομικό κόστος και με χαμηλό κόστος υπολογιστικής μονάδας, εξάγονται αυτόματα σημαντικές πληροφορίες, όπως τη δόμηση κάποιας περιοχής σε αδόμητη και δομημένη με τη χρήση εικόνων δορυφόρου.

Ο αλγόριθμος U-Net , που χρησιμοποιήθηκε στην παρούσα μεταπτυχιακή διατριβή , ο οποίος ανήκει στα πλήρως συνελκτικά νευρωνικά δίκτυα, έχει αποδώσει αρκετά καλά στο πρόβλημα, έχοντας ένα σχετικά χαμηλό αριθμό παραμέτρων για εκπαίδευση καθώς επίσης απαιτεί μικρή υπολογιστική ισχύ λεπτά για να εκπαιδευτεί.

Πρόκειται για ένα καινοτόμο αλγόριθμο, ο οποίος μέσω της σημασιολογικής κατάτμησης, μας έχει αποδείξει τα πιο πάνω πλεονεκτήματα που αναφέρθηκαν, αφού με 208 και 1923 δείγματα, μπορεί να εκπαιδευτεί, πετυχαίνοντας σχετικά ψηλές αποδόσεις στο σετ καθώς στο σετ ελέγχου .

Βιβλιογραφία

[18] Wierzbicki D., Matuk O., and Bielecka E. (2021) , “Polish Cadastre Modernization with Remotely Extracted Buildings from High-Resolution Aerial Orthoimagery and Airborne LiDAR”

B Lima, L Ferreira, JM Moura – (2021) Helping to detect legal swimming pools with Deep Learning and Data Visualization

E Fernandes, P Wildemberg (2020)Water Tanks and Swimming Pools Detection in Satellite Images: Exploiting Shallow and Deep-Based Strategies

[19]C Coelho, MFP Costa, LJL Ferrás, AJ Soares (2021)Development of a machine learning model and a user interface to detect illegal swimming pools

[21]Yafei Jing, Yuhuan Ren Yalan Liu, Dacheng Wang, Linjun Yu (2022) ,Automatic Extraction of Damaged Houses by Earthquake Based on Improved YOLOv5: A Case Study in Yangbi

David Garcia-Retuerta, Roberto Casado-Vara, Jose L. Calvo-Rolle, Héctor Quintián, Javier Prieto (2020)Deep Learning for House Categorisation, a Proposal Towards Automation in Land Registry

[17] Α Τέμενος (2021)Αυτόματος εντοπισμός κτηρίων απο RGB δορυφορικές εικόνες μέσω αλγορίθμου βαθιάς μάθησης

Mohammed Shuaib , Salwani Mohd Daud , Shadab Alam , Wazir Zada Khan4 (2020) Blockchain-based framework for secure and reliable land registry system

Aneta Jadwiga Florczyk; Stefano Ferri; Vasileios Syrris; Thomas Kemper; Matina Halkia; Pierre Soille; Martino Pesaresi (2015)A new European settlement map from optical remotely sensed data

Maria Vakalopoulou; Norbert Bus; Konstantinos Karantzas; Nikos Paragios (2017) Integrating edge/boundary priors with classification scores for building detection in very high resolution data

S Wang, X Hou, X Zhao (2020) Automatic building extraction from high-resolution aerial imagery via fully convolutional encoder-decoder network with non-local block

[4] Andrew Hodges (2007) Alan Turing and the Turing Test

[6] AL Samuel (1967) Some studies in machine learning using the game of checkers. II—Recent progress

[7] Y Guo, Y Liu, A Oerlemans, S Lao, S Wu, MS Lew (2016) Deep learning for visual understanding: A review

Object recognition using deep convolutional features transformed by a recursive network structure (2016) HM Bui, M Lech, E Cheng, K Neville, IS Burnett

MM Najafabadi, F Villanustre, TM Khoshgoftaar (2015) Deep learning applications and challenges in big data analytics

C Goutte, E Gaussier (2005) A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation

[1] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

[2] Α Μπερδελής (2019) Εφαρμογή τεχνικών μηχανικής μάθησης για την πρόβλεψη της κατανομής των επισκεπτών σε μεγάλες εκδηλώσεις, κάνοντας χρήση δεδομένων τοποθεσίας

[3] Ε Αναστασοπούλου (2020) Η Τεχνητή Νοημοσύνη και οι εφαρμογές της

[5] Dick, S. (2019). Artificial Intelligence. Harvard Data Science Review, 1(1).
<https://doi.org/10.1162/99608f92.92fe150c>

[6] L Gugerty (2006) Newell and Simon's logic theorist: Historical background and impact on cognitive modeling

[8] Big Data and Artificial Intelligence: How They Work Together
<https://online.maryville.edu/blog/big-data-is-too-big-without-ai/>

[9] Rockwell Anyoha, The History of Artificial Intelligence
<https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>

[10] L Deng, 2012 Three Classes of Deep Learning Architectures and Their Applications: A Tutorial Survey

[11] Md Zahangir Alom, Tarek M. Taha, Christopher Yakopcic, Stefan Westberg, Paheding Sidike, Mst Shamima Nasrin, Brian C Van Esesn, Abdul A S. Awwal, Vijayan K. Asari (2018) The history began from alexnet: A comprehensive survey on deep learning approaches

[12] History of A.I.: Artificial Intelligence (Infographic) (2014)
<https://www.livescience.com/47544-history-of-a-i-artificial-intelligence-infographic.html>

[13] Allison Linn (2015) Microsoft researchers win ImageNet computer vision challenge

[14] E Aminanto, K Kim (2016) Deep learning in intrusion detection system: An overview

[15] Making the Most of GPUs for Your Deep Learning Project
<https://www.run.ai/guides/gpu-deep-learning>

[16] Teemu Kanstrén (2020) A Look at Precision, Recall, and F1-Score

[23] Adrian Rosebrock , 2016 Intersection over Union (IoU) for object detection

[20] Divyansh Jha and Rohit Singh, (2018) Swimming pool detection and classification using deep learning

[22] Manav Mandal (2021) Introduction to Convolutional Neural Networks (CNN)

[24] Debasish Kalita (2022) Basics of CNN in Deep Learning