



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΑΓΡΟΝΟΜΩΝ ΚΑΙ ΤΟΠΟΓΡΑΦΩΝ ΜΗΧΑΝΙΚΩΝ –
ΜΗΧΑΝΙΚΩΝ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗΣ
ΤΟΜΕΑΣ ΤΟΠΟΓΡΑΦΙΑΣ

**Ανίχνευση δασικών αποψιλωμένων εκτάσεων με την χρήση
τεχνικών Βαθιάς Μάθησης**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΒΑΣΙΛΙΚΗ ΚΡΗΤΙΚΟΥ

Επιβλέπων : Αναστάσιος Δουλάμης
Αναπληρωτής Καθηγητής Ε.Μ.Π - Τομέας Τοπογραφίας

Περίληψη

Τα τελευταία χρόνια έχει διαπιστωθεί ραγδαία ανάπτυξη στις τεχνικές της όρασης υπολογιστών (Computer Vision) και ιδιαίτερα σε αυτές που σχετίζονται με τη Βαθιά Μάθηση (Deep Learning). Αρκετά ενδιαφέρον πεδίο εφαρμογής των τεχνικών Βαθιάς Μάθησης αποτελεί η κατάτμηση εικόνας (Image Segmentation).

Σκοπός της παρούσας διπλωματικής εργασίας αποτέλεσε η μελέτη εφαρμογών κατάτμησης δορυφορικών εικόνων για την αυτοματοποίηση του εντοπισμού αποψιλωμένων δασικών εκτάσεων. Το σύνολο των εικόνων που χρησιμοποιήθηκε για την εκπαίδευση των νευρωνικών δικτύων που μελετήθηκαν στην συγκεκριμένη διπλωματική εργασία αντλήθηκε από την open source ιστοσελίδα GitHub και συγκεκριμένα από το μοντέλο «BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection».

Για τη μελέτη του ανωτέρω συνόλου δεδομένων χρησιμοποιήθηκαν Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks). Συγκεκριμένα, χρησιμοποιήθηκε η δημοφιλής αρχιτεκτονική UNet η οποία επιτυγχάνει πολύ καλή απόδοση σε προβλήματα κατάτμησης εικόνων. Επιπλέον, χρησιμοποιήθηκε μοντέλο με Swin Transformer βασισμένο στην αρχιτεκτονική UNet προκειμένου να πραγματοποιηθεί κατάτμηση εικόνας.

Για την αξιολόγηση των μοντέλων αυτών χρησιμοποιήθηκαν ορισμένες μετρικές, όπως ο Αρμονικός μέσος (F1 – Score), Intersection over Union (IoU), Cross - entropy loss και Accuracy (ακρίβεια). Συγκρίνοντας όλα τα μοντέλα για κάθε υπολογιζόμενη μετρική που χρησιμοποιήθηκε παρατηρείται ότι καλύτερα αποτελέσματα έχει το μοντέλο με τον Swin Transformer μετασχηματισμό. Επίσης, οι πιο αξιόπιστες μετρικές για τον σκοπό της παρούσας διπλωματικής διαπιστώθηκε ότι είναι οι μετρικές Αρμονικός μέσος (F1 – Score), Intersection over Union (IoU).

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Τεχνητή Νοημοσύνη, Βαθιά Μάθηση, Σημασιολογική Κατάτμηση (Semantic Segmentation), Αποψίλωση (Deforestation), Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks - FCNN), Μοντέλο UNet, Vision Transformer, Swin Transformer, Μετρικές.

Abstract

Over the past years, there has been a rapid development in the field of Computer Vision, especially through techniques involving Deep Learning. Image Segmentation is quite an interesting field of application of Deep Learning techniques.

The aim of this diploma is the study of satellite image segmentation applications for the automation of the detection of deforestation. The set of images used for the training of the neural networks that have been studied in this diploma was taken from the open source website GitHub and specifically from the «BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection» model.

Fully Convolutional Neural Networks were used to study the above data set. In particular, the popular UNet architecture was used, which achieves very good performance in image segmentation problems. In addition, a Swin Transformer model based on the UNet architecture was used to perform image segmentation.

Some metrics were used to evaluate these models, such as the (F1 – Score), Intersection over Union (IoU), Cross - entropy loss and Accuracy. By comparing all models for each calculated metric used, it is observed that the model with the Swin Transformer transformation has better results. Also, the most reliable metrics for the purpose of this diploma were (F1 – Score) and Intersection over Union (IoU) metrics.

KEY WORDS

Artificial Intelligence, Deep Learning, Semantic Segmentation, Deforestation, Fully Convolutional Neural Networks (FCNN), UNet Model, Vision Transformer, Swin Transformer, Metrics.

Ευχαριστίες

Η παρούσα διπλωματική εκπονήθηκε στο πλαίσιο του Προπτυχιακού Προγράμματος Σπουδών της Σχολής Αγρονόμων και Τοπογράφων Μηχανικών και Μηχανικών Γεωπληροφορικής του Εθνικού Μετσόβιου Πολυτεχνείου και σηματοδοτεί την ολοκλήρωση των σπουδών μου. Πριν όμως από οποιαδήποτε αναφορά στη διαδικασία που ακολουθήθηκε και στα αποτελέσματα που προέκυψαν, θα ήθελα να ευχαριστήσω τους ανθρώπους με τους οποίους συνεργάστηκα και συνέβαλαν στην ολοκλήρωση της εργασίας αυτής.

Κατ' αρχάς απευθύνω τις ευχαριστίες μου στον επιβλέποντα κ. Αναστάσιο Δουλάμη, Αναπληρωτή Καθηγητή Ε.Μ.Π, για την δυνατότητα που μου προσέφερε να εργαστώ σε ένα ιδιαίτερο αντικείμενο για μένα και να διευρύνω τις γνώσεις μου. Ιδιαίτερος, θα ήθελα να ευχαριστήσω την Μαρία Κασελίμη που η βοήθειά της και η στηρίξη της σε κάθε βήμα και κάθε απορία μου ήταν πολύ σημαντική για να συνεχίσω να μελετάω το συγκεκριμένο αντικείμενο που πραγματεύεται η παρούσα διπλωματική.

Τέλος, θα ήθελα να ευχαριστήσω την οικογενειά μου που με στήριξε, ώστε να ολοκληρώσω τις σπουδές μου.

Κρητικού Βασιλική, Αθήνα, Οκτώβριος 2023

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

Κεφάλαιο 1 – Εισαγωγή.....	11
1.1 Τεχνητή Νοημοσύνη.....	11
1.2 Μηχανική Μάθηση.....	11
1.2.1 Εισαγωγή.....	11
1.2.2 Είδη Μηχανικής Μάθησης.....	12
1.3 Βαθιά Μάθηση (Deep Learning).....	16
1.4 Ακραία Καιρικά Φαινόμενα - Κλιματική Αλλαγή.....	18
Κεφάλαιο 2 – Θεωρητικό Υπόβαθρο.....	20
2.1 Νευρωνικά Δίκτυα (Neural Networks).....	20
2.1.1 Εισαγωγή.....	20
2.1.2 Αρχιτεκτονική ΤΝΔ.....	20
2.1.3 Εκπαίδευση νευρωνικού δικτύου.....	21
2.2 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks) και Τηλεπισκόπηση (Remote Sensing Imagery).....	23
2.2.1 Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks).....	24
2.3 Σημασιολογική Κατάτμηση Εικόνων (Semantic Image Segmentation).....	25
2.3.1 Σημασιολογική Κατάτμηση Εικόνων (Semantic Image Segmentation) και Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks).....	26
2.4 Swin UNet - Swin Transformer.....	28
2.5 Το μοντέλο U-Net.....	31
Κεφάλαιο 3 – Φαινόμενο Αποψίλωσης και Μετρικές.....	33
3.1 Deforestation – Φαινόμενο «Αποψίλωσης» ή «Αποδάσωσης».....	33
3.1.1 Εισαγωγή.....	33
3.1.2 Ανάλυση του φαινομένου της «αποψίλωσης» δασών (deforestation).....	33
3.2 Απόδοση Νευρωνικών Δικτύων – Μετρικές (Metrics).....	34
3.2.1 Αρμονικός Μέσος - F1 – Score.....	35
3.2.2 Intersection Over Union (IOU) - Jaccard.....	37
3.2.3 Cross - entropy loss.....	39
3.2.4 Accuracy (ακρίβεια).....	40
Κεφάλαιο 4 – Δεδομένα και εργαλεία «Keras» που υλοποιήθηκαν.....	42
4.1 Περιγραφή των δεδομένων.....	42
4.2 Εργαλεία – Διεπαφή Keras.....	44
Κεφάλαιο 5 – Πειραματική Διάταξη.....	46
5.1 Μοντέλα Μηχανικής Μάθησης.....	46

5.2 Πρότυπο Μοντέλο Μηχανικής Μάθησης.....	46
5.2.1 Συναπτικά Βάρη Τεχνητού Νευρωνικού Δικτύου	47
5.2.2 Προεκπαιδευμένο Μοντέλο « <i>Model_Testing</i> » (Pre-Trained Model).....	47
Παράρτημα Κώδικα 1 – Αρμονικός Μέσος – F1 – Score:	47
Παράρτημα Κώδικα 2 - Intersection Over Union (IOU):.....	48
Παράρτημα Κώδικα 3 – Πεδίο «Plot Prediction»	49
Παράρτημα Κώδικα 4 – Υπολογισμός & Εκτύπωση F1 – Score & Intersection Over Union (IOU)	50
Παράρτημα Κώδικα 5 – Ορισμός μοντέλου (model10)	50
5.3 Μοντέλο Μηχανικής Μάθησης « <i>Swin – Unet</i> »	51
Παράρτημα κώδικα 6 – Συναρτήσεις <code>get_example</code> και <code>parse_image_function</code> για επεξεργασία του συνόλου δεδομένων.....	51
Παράρτημα κώδικα 7 – Εισαγωγή εικόνων στο μοντέλο και μείωση (downscale) αυτών σε εικόνες 128 pixels x 128 pixels με την συνάρτηση <code>block_reduce</code> της βιβλιοθήκης <code>skimage.measure</code>	52
Παράρτημα 8 – Κατηγοριοποίηση/Διαχωρισμός δεδομένων	53
Δοκιμές Εκπαίδευσης.....	53
5.4 Μοντέλο Μηχανικής Μάθησης « <i>Image Segmentation Keras: Implementation of Segnet, FCN, UNet, PSPNet</i> ».....	54
Παράρτημα κώδικα 9 – Segmentation model & Model Train	55
Κεφάλαιο 6 - Αποτελέσματα.....	57
6.1 Προεκπαιδευμένο Μοντέλο « <i>Model_Testing</i> » (Pre-Trained Model).....	57
6.2 Μοντέλο Μηχανικής Μάθησης « <i>Swin – UNet</i> »	60
Πρώτη 1 ^η Δοκιμή Εκπαίδευσης	60
Δεύτερη 2 ^η Δοκιμή Εκπαίδευσης	62
Τρίτη 3 ^η Δοκιμή Εκπαίδευσης	63
Τέταρτη 4 ^η Δοκιμή Εκπαίδευσης.....	65
Πέμπτη 5 ^η Δοκιμή Εκπαίδευσης	67
6.2.1 Πίνακες Αποτελεσμάτων μετρικών F1 – Score & Intersection over Union	69
6.2.3 Διαγράμματα « <i>Validation Performance</i> » δοκιμών εκπαίδευσης	70
6.3 Μοντέλο Μηχανικής Μάθησης « <i>Image Segmentation Keras: Implementation of Segnet, FCN, UNet, PSPNet</i> ».....	73
Πρώτη 1 ^η δοκιμή εκπαίδευσης.....	73
Δεύτερη 2 ^η δοκιμή εκπαίδευσης.....	75
Τρίτη 3 ^η δοκιμή εκπαίδευσης.....	77
6.3.1 Διαγράμματα « <i>Accuracy</i> » δοκιμών εκπαίδευσης.....	80
Κεφάλαιο 7 – Συμπεράσματα και Μελλοντική Εργασία	82

7.1 Συμπεράσματα	82
7.2 Μελλοντική Εργασία.....	83
Βιβλιογραφία.....	84

Πίνακας Περιεχομένων Εικόνων

ΕΙΚΟΝΑ 1 - ΕΠΙΠΕΔΟ ΣΥΝΕΛΙΞΗΣ	24
ΕΙΚΟΝΑ 2 – ΔΟΜΗ ΠΛΗΡΩΣ ΣΥΝΕΛΙΚΤΙΚΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ	25
ΕΙΚΟΝΑ 3 - ΠΑΡΑΔΕΙΓΜΑ ΣΗΜΑΣΙΟΛΟΓΙΚΗΣ ΚΑΤΑΤΜΗΣΗΣ ΔΟΡΥΦΟΡΙΚΗΣ ΕΙΚΟΝΑΣ	26
ΕΙΚΟΝΑ 4 - ΑΡΧΙΤΕΚΤΟΝΙΚΗ ENCODER – DECODER (DOWNSAMPLING – UPSAMPLING)	27
ΕΙΚΟΝΑ 5 - ENCODER - DECODER ΜΕ SKIP CONNECTIONS	27
ΕΙΚΟΝΑ 6 - Η ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ ΜΟΝΤΕΛΟΥ "SWIN - UNET" ΓΙΑ ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΚΑΤΑΤΜΗΣΗ, Η ΟΠΟΙΑ ΑΠΑΡΤΙΖΕΤΑΙ ΑΠΟ ENCODER - DECODER, BOTTLENECK & SKIP CONNECTIONS. ENCODER, DECODER & BOTTLENECK ΕΧΟΥΝ ΚΑΤΑΣΚΕΥΑΣΤΕΙ ΒΑΣΙΖΟΜΕΝΑ ΣΤΟ SWIN TRANSFORMER BLOCK.....	30
ΕΙΚΟΝΑ 7 - SWIN TRANSFORMER BLOCK.....	30
ΕΙΚΟΝΑ 8 - UNET ΑΡΧΙΤΕΚΤΟΝΙΚΗ (ΠΑΡΑΔΕΙΓΜΑ 32X32 PIXELS ΣΤΗ ΧΑΜΗΛΟΤΕΡΗ ΑΝΑΛΥΣΗ). ΚΑΘΕ ΜΠΛΕ ΠΛΑΙΣΙΟ ΑΝΤΙΣΤΟΙΧΕΙ ΣΕ ΜΙΑ ΑΠΕΙΚΟΝΙΣΗ ΠΟΛΛΑΠΛΩΝ ΚΛΑΣΕΩΝ. Ο ΑΡΙΘΜΟΣ ΤΩΝ ΚΛΑΣΕΩΝ ΣΗΜΕΙΩΝΕΤΑΙ ΣΤΗΝ ΚΟΡΥΦΗ ΤΟΥ ΠΛΑΙΣΙΟΥ. ΤΑ ΒΕΛΗ ΥΠΟΔΗΛΩΝΟΥΝ ΤΙΣ ΔΙΑΦΟΡΕΣ ΛΕΙΤΟΥΡΓΙΕΣ	32
ΕΙΚΟΝΑ 9 - ΜΕΤΡΙΚΕΣ PRECISION & RECALL	36
ΕΙΚΟΝΑ 10 - ΜΕΤΡΙΚΗ INTERSECTION OVER UNION (IOU).....	38
ΕΙΚΟΝΑ 11 - INTERSECTION OVER UNION	38
ΕΙΚΟΝΑ 12 - PRECISION, RECALL & IOU	38
ΕΙΚΟΝΑ 13 - ΔΙΑΓΡΑΜΜΑ ΑΠΕΙΚΟΝΙΣΗΣ ΤΗΣ ΕΚΤΑΣΗΣ ΤΩΝ ΚΛΑΣΕΩΝ FOREST, DEFOREST & OTHER ΣΤΙΣ ΕΙΚΟΝΕΣ.....	43
ΕΙΚΟΝΑ 14 - BOXPLOT OF FOREST, DEFOREST & OTHER AREAS.....	43
ΕΙΚΟΝΑ 15 - ΠΟΣΟΣΤΟ (%) ΠΛΗΡΟΦΟΡΙΑΣ FOREST, DEFOREST & OTHER ΣΤΙΣ ΕΙΚΟΝΕΣ.....	44
ΕΙΚΟΝΑ 16 - ΔΟΡΥΦΟΡΙΚΗ ΕΙΚΟΝΑ	44
ΕΙΚΟΝΑ 17 - GROUND TRUTH.....	44
ΕΙΚΟΝΑ 18 - KERAS_SEGMENTATION ΜΟΝΤΕΛΑ.....	55
ΕΙΚΟΝΑ 19 - PRE - TRAINED MODEL - TEST IMAGES, GROUND TRUTH, MODEL_PREDICTION [1/22].....	58
ΕΙΚΟΝΑ 20 - PRE - TRAINED MODEL - TEST IMAGES, GROUND TRUTH, MODEL_PREDICTION [3/22].....	58
ΕΙΚΟΝΑ 21 - PRE - TRAINED MODEL - TEST IMAGES, GROUND TRUTH, MODEL_PREDICTION [7/22].....	58
ΕΙΚΟΝΑ 22 - PRE - TRAINED MODEL - TEST IMAGES, GROUND TRUTH, MODEL_PREDICTION [10/22].....	59
ΕΙΚΟΝΑ 23 - PRE - TRAINED MODEL - TEST IMAGES, GROUND TRUTH, MODEL_PREDICTION [12/22].....	59
ΕΙΚΟΝΑ 24 - TEST IMAGE, GROUND TRUTH, PREDICTION – 1Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ	61
ΕΙΚΟΝΑ 25 - TEST IMAGE, GROUND TRUTH, PREDICTION - 2Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.	63
ΕΙΚΟΝΑ 26 - TEST IMAGE, GROUND TRUTH, PREDICTION - 3Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.	65
ΕΙΚΟΝΑ 27 - TEST IMAGE, GROUND TRUTH, PREDICTION - 4Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.	66
ΕΙΚΟΝΑ 28 - TEST IMAGE, GROUND TRUTH, PREDICTION - 5Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.	68
ΕΙΚΟΝΑ 29 - VALIDATION PERFORMANCE - 1Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ	70
ΕΙΚΟΝΑ 30 - VALIDATION PERFORMANCE - 2Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ	71
ΕΙΚΟΝΑ 31 - VALIDATION PERFORMANCE - 3Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ	71
ΕΙΚΟΝΑ 32 - VALIDATION PERFORMANCE - 4Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ	72
ΕΙΚΟΝΑ 33 - VALIDATION PERFORMANCE - 5Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.....	72

EIKONA 34 - ΜΟΝΤΕΛΟ "IMAGE SEGMENTATION KERAS" - TEST IMAGE, GROUND TRUTH, PREDICTION - 1Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.....	75
EIKONA 35 - ΜΟΝΤΕΛΟ "IMAGE SEGMENTATION KERAS" - TEST IMAGE, GROUND TRUTH, PREDICTION – 2Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.....	77
EIKONA 36 – ΜΟΝΤΕΛΟ "IMAGE SEGMENTATION KERAS" - TEST IMAGE, GROUND TRUTH, PREDICTION – 3Η ΔΟΚΙΜΗ ΕΚΠΑΙΔΕΥΣΗΣ.....	79
<i>EIKONA 37 - ΔΙΑΓΡΑΜΜΑ ΜΕΤΡΙΚΗΣ "ACCURACY" - 20 ΕΠΟΧΕΣ ΕΚΠΑΙΔΕΥΣΗΣ.....</i>	<i>80</i>
EIKONA 38 - ΔΙΑΓΡΑΜΜΑ "ACCURACY" - 40 ΕΠΟΧΕΣ ΕΚΠΑΙΔΕΥΣΗΣ.....	80
EIKONA 39 –ΔΙΑΓΡΑΜΜΑ "ACCURACY" – 20 ΕΠΟΧΕΣ ΕΚΠΑΙΔΕΥΣΗΣ – TRAINING SET 30%.....	81

Κεφάλαιο 1 – Εισαγωγή

1.1 Τεχνητή Νοημοσύνη

Η τεχνητή νοημοσύνη (artificial intelligence - AI) μπορεί να περιγραφεί ως ο κλάδος που ασχολείται με το σχεδιασμό «νοημόνων» υπολογιστών, είτε αυτοί είναι συσκευές, προγράμματα ή ρομπότ, οι οποίοι έχουν τη δυνατότητα να αντιλαμβάνονται το περιβάλλον τους και να δρουν έτσι ώστε να μεγιστοποιούν τις πιθανότητες επιτυχίας, ανάλογα με τους στόχους τους. Από τις αρχές της δεκαετίας του 1990 μέχρι και σήμερα, ο τομέας της τεχνητής νοημοσύνης έχει σταδιακά ανακάμψει και έχει πετύχει αρκετούς από τους αρχικούς στόχους που είχαν θέσει οι πρώτοι ερευνητές. Τα αποτελέσματα της έρευνας αυτών των χρόνων όλο και περισσότερο χρησιμοποιούνται σε πρακτικές εφαρμογές. Η Τεχνητή Νοημοσύνη χρησιμοποιείται στην εξόρυξη δεδομένων, την ιατρική διάγνωση και άλλους τομείς. Η επιτυχία αυτή οφείλεται στην αύξηση της υπολογιστικής ισχύος, στην επίλυση συγκεκριμένων προβλημάτων, στους νέους δεσμούς μεταξύ της Τεχνητής Νοημοσύνης και άλλων τομέων όπως στατιστική, οικονομικά και μαθηματικά και στην αφοσίωση των ερευνητών σε μαθηματικές μεθόδους και επιστημονικά πρότυπα.

Η έρευνα πάνω σε «νοήμονες» υπολογιστές επικεντρώνεται κατά κύριο λόγο σε καθαρά γνωσιακά αντικείμενα, όπως την επεξεργασία πληροφοριών, την επίλυση προβλημάτων, τη μάθηση, την αντίληψη, τη συλλογιστική και τη γλώσσα, δηλαδή σε παράγοντες που θεωρούνται απαραίτητοι για νοημοσύνη. Μάλιστα, είναι αδιαμφισβήτητο το ότι, όσον αφορά αρκετούς από αυτούς τους παράγοντες, η ικανότητα των υπολογιστών να «σκεφτούν» έχει ξεπεράσει κατά πολύ αυτή των ανθρώπων.

Πλέον διατίθενται τεράστιος όγκος δεδομένων, διότι η ανθρωπότητα λόγω της δραστηριότητας στο διαδίκτυο έχει αφήσει ένα τεράστιο ψηφιακό αποτύπωμα, το οποίο μπορούμε να εκμεταλλευτούμε για να προπονήσουμε τους αλγόριθμους τεχνητής νοημοσύνης.

1.2 Μηχανική Μάθηση

1.2.1 Εισαγωγή

Ο άνθρωπος προσπαθεί να κατανοήσει το περιβάλλον του παρατηρώντας το και δημιουργώντας μια απλοποιημένη (αφαιρετική) εκδοχή του που ονομάζεται **νοητικό μοντέλο (mental model)**. Η δημιουργία ενός τέτοιου μοντέλου, ονομάζεται **επαγωγική μάθηση (inductive learning)** ενώ η διαδικασία γενικότερα ονομάζεται **επαγωγή (induction)**. Επιπλέον, ο άνθρωπος έχει τη δυνατότητα να οργανώνει και να συσχετίζει τις εμπειρίες και τις παρατηρήσεις του δημιουργώντας νέες δομές που ονομάζονται **νοητικά πρότυπα (mental patterns)**, με αξιοποίηση και του επαγωγικού και του απαγωγικού συλλογισμού. Στη δημιουργία νέων προτύπων από παλαιά βασίζονται οι τρόποι μάθησης που εξαρτώνται σε μεγαλύτερο ή μικρότερο βαθμό από την προϋπάρχουσα γνώση για ένα πρόβλημα, όπως είναι η μάθηση από επεξηγήσεις και η μάθηση από περιπτώσεις. Με την λογική, λοιπόν, της **Μάθησης (Learning)** που αποτελεί μία από τις θεμελιώδεις ιδιότητες της νοήμονος συμπεριφοράς του ανθρώπου έχουν δημιουργηθεί υπολογιστικά συστήματα ικανά να μάθουν, να επιτύχουν, δηλαδή, τη λεγόμενη **Μηχανική Μάθηση (Machine Learning)**.

Η **Μηχανική Μάθηση** έχει ως σκοπό τη δημιουργία μηχανών ικανών να μαθαίνουν και να βελτιώνουν την απόδοσή τους σε κάποιους τομείς μέσω της αξιοποίησης προηγούμενης γνώσης και εμπειρίας. Ένας σχετικός γενικός ορισμός της Μηχανικής Μάθησης δίνεται από τον Mitchell (1997):

«Ένα πρόγραμμα υπολογιστή λέμε ότι μαθαίνει από την εμπειρία E ως προς κάποια κλάση εργασιών T και μέτρο απόδοσης P , αν η απόδοσή του σε εργασίες από το T , όπως μετριέται από το P , βελτιώνεται μέσω της εμπειρίας E .»

Συνεπώς, η ικανότητα ενός υπολογιστικού συστήματος να δημιουργεί μοντέλα ή πρότυπα από ένα σύνολο δεδομένων ονομάζεται **Μηχανική Μάθηση (Machine Learning)**. Η **Μηχανική Μάθηση** χρησιμοποιείται για να διδαχθούν οι μηχανές τον πιο αποτελεσματικό τρόπο διαχείρισης δεδομένων. Πολλές φορές η δυνατότητα διερμηνείας της εξαγόμενης πληροφορίας από το εκάστοτε σύνολο δεδομένων αποτελεί δύσκολη διαδικασία. Σε αυτή την περίπτωση χρησιμοποιείται η **Μηχανική Μάθηση (Machine Learning)** όπου σκοπός της είναι να εκπαιδευτεί και να μαθαίνει από τα δεδομένα.

Ως κλάδος της Τεχνητής Νοημοσύνης, η Μηχανική Μάθηση ασχολείται με τη μελέτη αλγορίθμων που βελτιώνουν τη συμπεριφορά τους σε κάποια εργασία που τους έχει ανατεθεί χρησιμοποιώντας την εμπειρία τους. Όσον αφορά τη σχεδίαση των συστημάτων Μηχανικής Μάθησης, για τα συστήματα που ανήκουν στη συμβολική Τεχνητή Νοημοσύνη, η δυνατότητα μάθησης προσδιορίζεται ως η ικανότητα πρόσκτησης επιπλέον γνώσης, που επιφέρει μεταβολές στην υπάρχουσα καταχωρημένη γνώση είτε αλλάζοντας χαρακτηριστικά της είτε με αυξομείωσή της. Στην περίπτωση των συστημάτων Τεχνητής Νοημοσύνης που ανήκουν στη Μη Συμβολική Τεχνητή Νοημοσύνη (όπως η περίπτωση των Τεχνητών Νευρωνικών Δικτύων), ως μάθηση προσδιορίζεται η δυνατότητα που διαθέτουν τα συστήματα στο να μετασχηματίζουν την εσωτερική τους δομή, παρά στο να μεταβάλλουν κατάλληλα τη γνώση που έχει καταχωρηθεί μέσα σε αυτά κατά το σχεδιασμό τους.

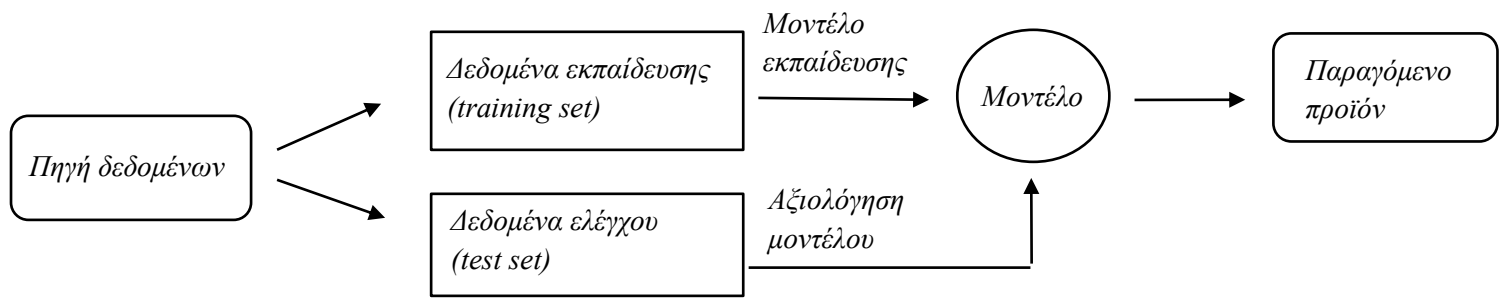
1.2.2 Είδη Μηχανικής Μάθησης

Ο τομέας της Μηχανικής Μάθησης αναπτύσσει τρεις τρόπους μάθησης, ανάλογους με τους τρόπους με τους οποίους μαθαίνει ο άνθρωπος: επιβλεπόμενη μάθηση, μη επιβλεπόμενη μάθηση και ενισχυτική μάθηση.

Πιο αναλυτικά, οι παραπάνω τεχνικές μηχανικής μάθησης οι οποίες χρησιμοποιούνται ανάλογα με τη φύση του προβλήματος:

- **Επιβλεπόμενη Μάθηση (Supervised Learning)** ή μάθηση με παραδείγματα (learning from examples) είναι η διαδικασία όπου ο αλγόριθμος κατασκευάζει μια συνάρτηση που απεικονίζει δεδομένες εισόδους (σύνολο εκπαίδευσης) σε γνωστές επιθυμητές εξόδους, με απώτερο στόχο τη γενίκευση της συνάρτησης αυτής και για εισόδους με άγνωστη έξοδο. Χρησιμοποιείται στο πρόβλημα:
 - Ταξινόμησης (Classification), αφορά στη δημιουργία μοντέλων πρόβλεψης διακριτών τάξεων (κλάσεων/κατηγοριών)

Οι αλγόριθμοι **Επιβλεπόμενης Μάθησης (Supervised Learning)** χρειάζονται εξωτερική βοήθεια. Τα εισακτέα δεδομένα διακρίνονται σε δεδομένα εκπαίδευσης (**training set**) και δεδομένα ελέγχου (**test set**). Τα δεδομένα εκπαίδευσης (**training set**) έχουν εξαγόμενες μεταβλητές οι οποίες πρέπει να ταξινομηθούν ή να προβλεφθούν. Κάθε αλγόριθμος εκπαιδευτεί από το σύνολο των δεδομένων εκπαίδευσης και δημιουργεί ορισμένα πρότυπα (patterns) τα οποία εφαρμόζει στα δεδομένα ελέγχου για ταξινόμηση ή διερμηνεία.



Σχήμα 1.1: Ροή (Workflow) Επιβλεπόμενης Μάθησης (Supervised Learning) δεδομένων

- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)** ή μάθηση από παρατήρηση (learning from observation), όπου ο αλγόριθμος κατασκευάζει ένα μοντέλο για κάποιο σύνολο εισόδων υπό μορφή παρατηρήσεων χωρίς να γνωρίζει τις επιθυμητές εξόδους. Χρησιμοποιείται σε προβλήματα:
 - Ανάλυσης Συσχετισμών (Association Analysis)
 - Ομαδοποίησης (Clustering)

Οι αλγόριθμοι **Μη Επιβλεπόμενης Μάθησης (UnSupervised Learning)**, όταν εισάγεται καινούριο σύνολο δεδομένων, χρησιμοποιούν τα χαρακτηριστικά που έχουν μάθει από την εκπαίδευση με προηγούμενα δεδομένα, ώστε να αναγνωρίζουν την κατηγορία (κλάση) των καινούριων εισακτέων δεδομένων και να τα ταξινομούν.

Στην μάθηση με επίβλεψη το σύστημα καλείται να “μάθει” επαγωγικά μια έννοια ή συνάρτηση που ονομάζεται συνάρτηση στόχος (target function) από ένα σύνολο δεδομένων, η οποία αποτελεί έκφραση του μοντέλου που περιγράφει τα δεδομένα. Αντίθετα, στη μάθηση χωρίς επίβλεψη το σύστημα πρέπει μόνο του να ανακαλύψει συσχετίσεις ή ομάδες σε ένα σύνολο δεδομένων, δημιουργώντας πρότυπα, χωρίς να είναι γνωστό αν υπάρχουν, πόσα και ποια είναι.

- **Ημι-επιβλεπόμενη μάθηση (Semi-supervised learning)**

Η μάθηση με ημι-επίβλεψη είναι μια κατηγορία τεχνικών μηχανικής μάθησης (ML) που χρησιμοποιούν τόσο επισημασμένα όσο και μη επισημασμένα δεδομένα για την εκπαίδευση – συνήθως μικρής ποσότητας επισημασμένων δεδομένων με μια μεγάλη ποσότητα μη επισημασμένων δεδομένων. Αυτή η προσέγγιση βρίσκεται μεταξύ της μάθησης χωρίς επίβλεψη (χωρίς επισημασμένα δεδομένα εκπαίδευσης) και της μάθησης με επίβλεψη (μόνο με επισημασμένα δεδομένα εκπαίδευσης). Σε πολλά σενάρια του πραγματικού κόσμου, η απόκτηση ενός ολοκληρωμένου συνόλου επισημασμένων δεδομένων μπορεί να είναι δαπανηρή ή χρονοβόρα. Τα μη επισημασμένα δεδομένα, ωστόσο, είναι συνήθως άφθονα και πιο προσιτά. Η μάθηση με ημιεπίβλεψη αξιοποιεί αυτή την αφθονία μη επισημασμένων δεδομένων για να βελτιώσει την ακρίβεια της μάθησης. Η υπόθεση στην ημι-επιβλεπόμενη μάθηση είναι ότι τα μη επισημασμένα δεδομένα, παρά το γεγονός ότι δεν έχουν ετικέτες, μπορούν να παρέχουν αρκετές πληροφορίες για να βοηθήσουν το μοντέλο να κατανοήσει την υποκείμενη δομή του συνόλου δεδομένων και να βελτιώσει την απόδοση της μάθησης.

Η μάθηση με ημιεπίβλεψη χρησιμοποιείται σε πολλές εφαρμογές όπου τα επισημασμένα δεδομένα είναι σπάνια ή είναι ακριβό να αποκτηθούν. Τα παραδείγματα περιλαμβάνουν: Επεξεργασία φυσικής γλώσσας (NLP): Εργασίες όπως η ανάλυση συναισθήματος, όπου είναι δύσκολο να βρεθούν επισημασμένα παραδείγματα, αλλά είναι διαθέσιμοι μεγάλοι όγκοι μη επισημασμένου κειμένου. Αναγνώριση εικόνας και ομιλίας: Η επισήμανση εικόνων ή δειγμάτων ομιλίας μπορεί να είναι εντάσεως εργασίας, αλλά τα μη

επισημασμένα δεδομένα είναι άφθονα. Βιοπληροφορική: Για παράδειγμα, στην ανακάλυψη φαρμάκων, όπου η επισημάνση μπορεί να είναι εξαιρετικά δαπανηρή. Με την έλευση της βαθιάς μάθησης, έχουν προταθεί νέες μέθοδοι για μάθηση με ημιεπίβλεψη με χρήση βαθιών νευρωνικών δικτύων, όπως τα ημιεπιβλεπόμενα Generative Adversarial Networks (GANs) και οι Variational Autoencoders (VAEs). Η ενεργητική μάθηση, όπου το ίδιο το μοντέλο επιλέγει τα σημεία δεδομένων που θέλει να επισημανθούν, αποτελεί επίσης έναν τομέα ενδιαφέροντος, ο οποίος συχνά συνδυάζεται με τεχνικές ημι-επιβλεπόμενης μάθησης για την ελαχιστοποίηση της ανάγκης για επισημειωμένα δεδομένα.

Η μάθηση με ημιεπίβλεψη αποτελεί μια πρακτική και αποτελεσματική λύση σε τομείς όπου τα επισημασμένα δεδομένα είναι σπάνια. Ωστόσο, η επιλογή της συγκεκριμένης μεθόδου μάθησης με ημιεπίβλεψη και η επιτυχής εφαρμογή της απαιτούν καλή κατανόηση τόσο των δεδομένων που υπάρχουν όσο και των ιδιοτήτων του προβλήματος.

- **Adversarial learning**

Adversarial learning είναι ένας τύπος μηχανικής μάθησης που περιλαμβάνει μοντέλα που μαθαίνουν να αντιμετωπίζουν adversaries. Η έννοια αυτή συνδέεται συνηθέστερα με τα Generative Adversarial Networks (GANs), αλλά έχει επίσης ευρύτερες επιπτώσεις σε τομείς όπως η ασφάλεια, η ιδιωτικότητα και η ευρωστία των μοντέλων μηχανικής μάθησης. Τα Generative Adversarial Networks (GANs) αποτελούνται από δύο νευρωνικά δίκτυα, τη γεννήτρια και τον διαχωριστή, τα οποία εκπαιδεύονται ταυτόχρονα μέσω αντιφατικών διαδικασιών. Η γεννήτρια μαθαίνει να παράγει νέα δεδομένα, ενώ ο διαχωριστής αξιολογεί τη γνησιότητά τους, δηλαδή αν τα δεδομένα είναι πραγματικά ή παραχθέντα. Η γεννήτρια προσπαθεί να παράγει δεδομένα που δεν διακρίνονται από πραγματικά δεδομένα, ενώ ο διαχωριστής προσπαθεί να αναγνωρίσει τα παραγόμενα δεδομένα. Η διαδικασία συνεχίζεται έως ότου η γεννήτρια παράγει δεδομένα αρκετά καλά ώστε να ξεγελάσει τον διαχωριστή. Τα GANs έχουν διάφορες εφαρμογές, όπως η παραγωγή εικόνων, η υπερ-ανάλυση, η μεταφορά στυλ, η μετάφραση από εικόνα σε εικόνα και άλλες.

Adversarial Robustness, πρόκειται για την εκπαίδευση των μοντέλων μηχανικής μάθησης ώστε να είναι ανθεκτικά έναντι της αντίπαλης χειραγώγησης. Ο στόχος είναι να διασφαλιστεί ότι το μοντέλο διατηρεί υψηλή απόδοση ακόμη και όταν τα δεδομένα εισόδου διαταράσσονται σκόπιμα. Οι τεχνικές για να συμβεί αυτό περιλαμβάνουν την ισχυρή βελτιστοποίηση, όπου το μοντέλο εκπαιδεύεται για την επίλυση ενός προβλήματος βελτιστοποίησης χειρότερης περίπτωσης (min-max), λαμβάνοντας υπόψη τις χειρότερες δυνατές αντίξοες αλλαγές στα δεδομένα εισόδου.

Η adversarial learning συνεχίζει να αποτελεί ενεργό τομέα έρευνας, διευρύνοντας τα όρια του εφικτού στα παραγωγικά μοντέλα, βελτιώνοντας την ευρωστία και την ασφάλεια των εφαρμογών μηχανικής μάθησης και διερευνώντας τους συμβιβασμούς μεταξύ ιδιωτικότητας και χρησιμότητας.

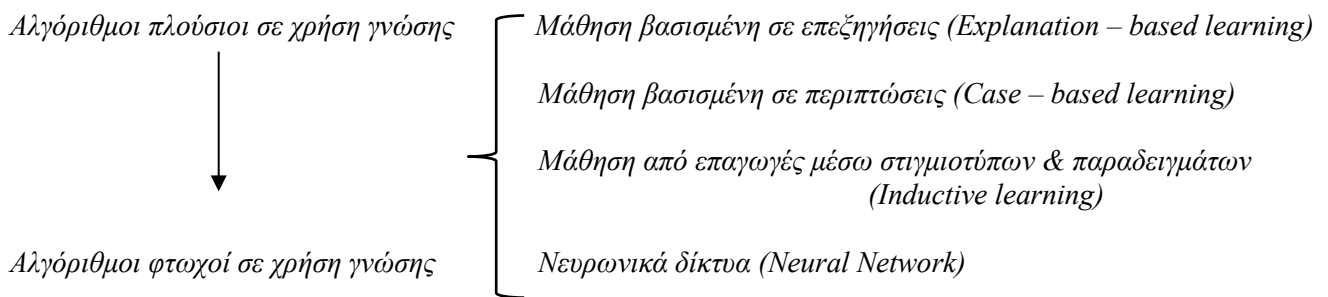
- **Αυτοεπίβλεψη (Self-supervised learning)**

Η μάθηση με αυτοεπίβλεψη είναι μια μορφή μάθησης χωρίς επίβλεψη όπου τα ίδια τα δεδομένα παρέχουν επίβλεψη. Βασίζεται στην αρχή ότι είναι δυνατόν να δημιουργηθούν αυτόματα ετικέτες από τα δεδομένα εισόδου και στη συνέχεια να χρησιμοποιηθούν αυτές οι ετικέτες για την εκπαίδευση ενός μοντέλου. Αυτή η προσέγγιση εξαλείφει ή ελαχιστοποιεί την ανάγκη για δεδομένα με ανθρώπινες ετικέτες. Η ιδέα είναι η πρόβλεψη τμημάτων των δεδομένων από άλλα τμήματα των δεδομένων. Με αυτόν τον τρόπο, το μοντέλο μαθαίνει να κατανοεί τη δομή και το πλαίσιο των δεδομένων. Συγκεκριμένα, το μοντέλο εκπαιδεύεται να προβλέπει ορισμένες πτυχές των δεδομένων εισόδου, όπως η πρόβλεψη της επόμενης λέξης σε μια πρόταση ή του επόμενου καρέ σε μια ακολουθία βίντεο. Επίσης, εκπαιδεύεται για την ανακατασκευή των αρχικών δεδομένων εισόδου από μια αλλοιωμένη έκδοση των δεδομένων

(π.χ. μια εικόνα με θόρυβο ή ένα καλυμμένο τμήμα κειμένου). Τέλος, μαθαίνει να μετασχηματίζει τα ακατέργαστα δεδομένα σε μια ουσιαστική αναπαράσταση που είναι χρήσιμη για μια μεταγενέστερη εργασία. Για παράδειγμα, ένα μοντέλο μπορεί να μάθει να μετατρέπει ακατέργαστα δεδομένα εικονοστοιχείων σε χαρακτηριστικά που είναι χρήσιμα για την αναγνώριση αντικειμένων.

Στην παρούσα διπλωματική επιλέγεται η μάθηση με επίβλεψη καθώς το πρόβλημα στο σύνολό του έχει οριστεί με τον καλύτερο δυνατό τρόπο και αναμένεται να παρθούν από το μοντέλο τα βέλτιστα αποτελέσματα καθώς η εκπαίδευση είναι συγκεκριμένη και κατευθυνόμενη.

Για κάθε πρόβλημα προς επίλυση στο χώρο της Μηχανικής Μάθησης υπάρχει ένας κατάλληλος τρόπος μάθησης και για κάθε τρόπο μάθησης υπάρχει τουλάχιστον ένας κατάλληλος αλγόριθμος που μπορεί να χρησιμοποιηθεί. Όλοι οι αλγόριθμοι Μηχανικής Μάθησης διαχειρίζονται τη γνώση αναπαριστώντας την με τον καταλληλότερο τρόπο που δύναται να εκφραστεί από τον εκάστοτε αλγόριθμο. Ορισμένοι αλγόριθμοι δέχονται ως είσοδο μόνο παρατηρήσεις και άλλοι λαμβάνουν υπόψη τους λίγο ή περισσότερο την προϋπάρχουσα γνώση. Μια προσπάθεια κατάταξης των αλγορίθμων με κριτήριο τον τρόπο μάθησης βασισμένο περισσότερο ή λιγότερο στην υπάρχουσα γνώση δίνεται στο παρακάτω σχήμα:



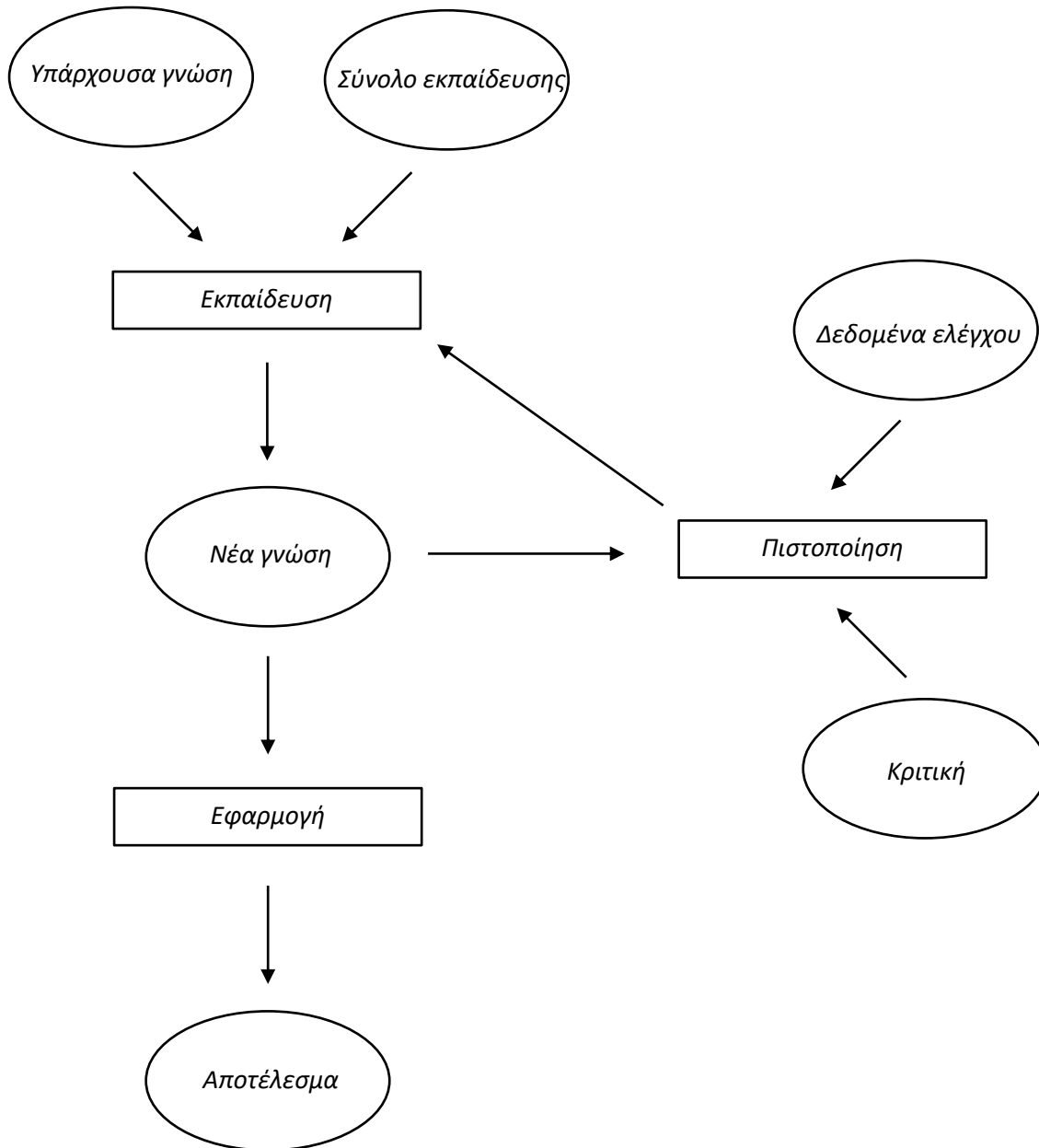
Σχήμα 1.2: Κατάταξη αλγορίθμων βάσει τρόπου μάθησης – υπάρχουσας γνώσης

Στο παρακάτω Σχήμα 1.3, αποτυπώνεται ο γενικός τρόπος λειτουργίας των αλγορίθμων Μηχανικής Μάθησης.

Αρχικά, κατηγοριοποιείται το σύνολο των δεδομένων που θα χρησιμοποιηθεί σε:

- **Δεδομένα εκπαίδευσης (training set):** αποτελεί το σύνολο των δεδομένων το οποίο θα χρησιμοποιηθεί για την εκπαίδευση του αλγορίθμου.
- **Δεδομένα αξιολόγησης (validate set):** αποτελεί το σύνολο των δεδομένων το οποίο θα χρησιμοποιηθεί για την αξιολόγηση της εκπαίδευσης του αλγορίθμου.
- **Δεδομένα ελέγχου (test set):** αποτελεί το σύνολο των δεδομένων το οποίο θα χρησιμοποιηθεί για την δοκιμή της ορθής λειτουργίας του αλγορίθμου.

Η βασικότερη φάση κάθε αλγορίθμου είναι η εκπαίδευση, όπου ο αλγόριθμος χρησιμοποιεί ως είσοδο ένα σύνολο **δεδομένων εκπαίδευσης (training set)** προς επίτευξη του σκοπού του, τη δημιουργία νέας γνώσης. Επιπλέον, μπορεί είτε να χρησιμοποιήσει λιγότερο ή περισσότερο την υπάρχουσα γνώση είτε να μην τη χρησιμοποιήσει καθόλου.



Σχήμα 1.3: Φάσεις μηχανικής μάθησης

Την εκπαίδευση ακολουθεί η φάση της πιστοποίησης της παραγόμενης νέας γνώσης. Συνήθως, η πιστοποίηση πραγματοποιείται καταρχάς από τον ίδιο τον αλγόριθμο μέσω διαδικασιών ανάκλησης (recall) με τη βοήθεια των **δεδομένων ελέγχου (test set)** και, στη συνέχεια, μέσω κριτικής που κάνει ο χρήστης βάσει των γνώσεων που διαθέτει για το πρόβλημα που επιχειρεί να λύσει ο αλγόριθμος.

1.3 Βαθιά Μάθηση (Deep Learning)

Η μηχανική μάθηση τροφοδοτεί πολλές πτυχές της σύγχρονης κοινωνίας, όπως την αναζήτηση στο διαδίκτυο έως και το φιλτράρισμα του περιεχομένου στα μέσα κοινωνικής δικτύωσης. Πλέον, η παρουσία της καθίσταται αναγκαία ακόμη και σε καθημερινά από τον άνθρωπο καταναλωτικά προϊόντα, όπως τα κινητά τηλέφωνα. Πιο συγκεκριμένα, η μηχανική μάθηση χρησιμοποιείται για την αναγνώριση αντικειμένων από εικόνες, την μετατροπή του ήχου/ομιλίας σε κείμενο, την αντιστοίχιση των ειδήσεων/αναρτήσεων ανάλογα με τα ενδιαφέροντα του χρήστη ή την δυνατότητα επιλογής σχετικών αποτελεσμάτων σε μια

αναζήτηση. Όλο και περισσότερο αυτές οι εφαρμογές χρησιμοποιούν μια κατηγορία τεχνικών που ονομάζονται Βαθιά Μάθηση (*Deep Learning*).

Η εκμάθηση μέσω της αναπαράστασης (*Representation learning*) αποτελεί σύνολο μεθόδων που επιτρέπει σε μια μηχανή να τροφοδοτείται με ακατέργαστα δεδομένα και να ανακαλύπτει αυτόματα τα δεδομένα/αναπαραστάσεις που πρέπει να ανιχνευθούν ή ταξινομηθούν. Οι μέθοδοι βαθιάς μάθησης είναι *representation learning* μέθοδοι με πολλαπλά επίπεδα αναπαράστασης που αποκτώνται με την σύνθεση απλών αλλά μη γραμμικών δομών. Κάθε μια δομή μετατρέπει την αναπαράσταση από ένα επίπεδο (ξεκινώντας με την ακατέργαστη είσοδο) σε μια αναπαράσταση υψηλότερου, ελαφρώς πιο αφηρημένου επιπέδου. Με την σύνθεση πολλαπλών τέτοιων μετασχηματισμών υπάρχει η δυνατότητα εκμάθησης σύνθετων γνώσεων.

Για παράδειγμα, μία εικόνα είναι ένας πίνακας εικονοστοιχείων. Στα πρώτα επίπεδα τα χαρακτηριστικά που «μαθαίνει» το μοντέλο βαθιάς μάθησης να αναγνωρίζει είναι την παρουσία ή απουσία ακμών σε συγκεκριμένους προσανατολισμούς και θέσεις στην εικόνα. Στα επόμενα επίπεδα το μοντέλο βαθιάς μάθησης «μαθαίνει» να αναγνωρίζει μοτίβα εντοπίζοντας συγκεκριμένες διατάξεις στις ακμές, ανεξάρτητως των μικρών αποκλίσεων στις θέσεις των ακμών. Στα τελευταία επίπεδα το μοντέλο μπορεί να συναρμολογήσει μοτίβα σε μεγαλύτερους συνδυασμούς που αντιστοιχούν σε μέρη οικείων αντικειμένων και τα επόμενα επίπεδα μπορούν να ανιχνεύσουν αντικείμενα ως συνδυασμούς αυτών των μερών. Η βασική πτυχή της βαθιάς μάθησης είναι ακριβώς αυτά τα επίπεδα αναγνώρισης ορισμένων χαρακτηριστικών τα οποία δεν έχουν σχεδιαστεί από τον άνθρωπο, αλλά μαθαίνονται από δεδομένα χρησιμοποιώντας μια διαδικασία μάθησης γενικού σκοπού.

Η βαθιά μάθηση έχει προωθήσει σημαντικά τον τομέα της όρασης υπολογιστών, οδηγώντας σε αξιοσημείωτες βελτιώσεις σε διάφορες εφαρμογές. Τα μοντέλα βαθιάς μάθησης, ιδίως τα συνελκτικά νευρωνικά δίκτυα (CNN), έχουν επιτύχει κορυφαία ακρίβεια σε εργασίες όπως η ταξινόμηση εικόνων, η ανίχνευση αντικειμένων και κατάτμησης εικόνων. Μπορούν να μαθαίνουν αυτόματα ιεραρχικές αναπαραστάσεις από τα ακατέργαστα δεδομένα, αποτυπώνοντας περίπλοκα μοτίβα που απλούστερα μοντέλα μπορεί να χάσουν. Επιπλέον, είναι ιδιαίτερα κατάλληλα ώστε να επωφεληθούν από μεγάλες ποσότητες δεδομένων. Η απόδοσή τους συνήθως βελτιώνεται με περισσότερα δεδομένα και είναι ικανά να χειρίζονται τεράστια σύνολα δεδομένων που χρησιμοποιούνται συχνά στην όραση υπολογιστών. Οι παραδοσιακές τεχνικές υπολογιστικής όρασης απαιτούν χειροκίνητη εξαγωγή χαρακτηριστικών από εικόνες. Αντίθετα, τα μοντέλα βαθιάς μάθησης μαθαίνουν αυτόματα τα χαρακτηριστικά που είναι σημαντικά για την ταξινόμηση ή άλλες εργασίες απευθείας από τα δεδομένα, εξαλείφοντας την ανάγκη για χειροκίνητη εξαγωγή χαρακτηριστικών.

Τα μοντέλα βαθιάς μάθησης έχουν την δυνατότητα να εφαρμοστούν σε ένα ευρύ φάσμα εργασιών υπολογιστικής όρασης, συμπεριλαμβανομένων, μεταξύ άλλων, της αναγνώρισης εικόνων και βίντεο, της ανίχνευσης αντικειμένων, της σύνθεσης εικόνων και της υπερ-ανάλυσης εικόνων. Η ίδια αρχιτεκτονική νευρωνικού δικτύου μπορεί συχνά να προσαρμοστεί με ελάχιστες αλλαγές σε διαφορετικές εργασίες. Με την έλευση βελτιστοποιημένου υλικού για βαθιά μάθηση (όπως οι GPU και οι TPU), πολλές εργασίες όρασης υπολογιστών που βασίζονται στη βαθιά μάθηση μπορούν πλέον να εκτελούνται σε πραγματικό χρόνο. Αυτό είναι ζωτικής σημασίας για εφαρμογές όπως τα αυτόνομα οχήματα, η ανάλυση βίντεο σε πραγματικό χρόνο και οι διαδραστικές εφαρμογές.

Τα μοντέλα βαθιάς μάθησης είναι γενικά πιο ανθεκτικά στις παραλλαγές των εικόνων, όπως οι αλλαγές στο φωτισμό, την κλίμακα ή τον προσανατολισμό. Μπορούν να γενικεύουν καλά από τα δεδομένα εκπαίδευσης και να αποδίδουν με ακρίβεια ακόμη και σε εικόνες που διαφέρουν από τα δείγματα εκπαίδευσης. Επιπλέον, έχουν την δυνατότητα να κατανοήσουν το σημασιολογικό περιεχόμενο μιας εικόνας. Για παράδειγμα, μπορούν να αναγνωρίσουν όχι μόνο τα αντικείμενα σε μια εικόνα, αλλά και το περιεχόμενό τους σε σχέση μεταξύ τους.

Γενικά μπορούν να αξιοποιηθούν προ-εκπαιδευμένα μοντέλα (που έχουν εκπαιδευτεί σε εκτεταμένα σύνολα δεδομένων εικόνων) για να επιτευχθούν υψηλές επιδόσεις με ελάχιστα δεδομένα εκπαίδευσης. Αυτή η έννοια, γνωστή ως εκμάθηση μεταφοράς, είναι ιδιαίτερα επωφελής όταν υπάρχουν περιορισμένα δεδομένα με ετικέτες για μια συγκεκριμένη εργασία.

Η βαθιά μάθηση επιτρέπει την εκμάθηση από άκρο σε άκρο, όπου τα ακατέργαστα δεδομένα μπορούν να εισαχθούν απευθείας στο μοντέλο και να εξαχθούν ακατέργαστες προβλέψεις, μειώνοντας την ανάγκη για ενδιάμεσα βήματα. Ωστόσο, είναι σημαντικό να σημειωθεί ότι τα μοντέλα βαθιάς μάθησης, ιδίως στην όραση υπολογιστών, απαιτούν σημαντικούς υπολογιστικούς πόρους και η απόδοσή τους εξαρτάται σε μεγάλο βαθμό από την ποσότητα και την ποιότητα των δεδομένων εκπαίδευσης.

Η βαθιά μάθηση έχει κάνει σημαντικές προόδους στην επίλυση προβλημάτων που χρόνια προσπαθεί να λύσει η τεχνητή νοημοσύνη. Έχει αποδειχθεί ότι η βαθιά μάθηση αποτελεί την λύση σε περιπτώσεις περίπλοκων δομών με δεδομένα υψηλών διαστάσεων και έχει εφαρμογές σε αρκετούς τομείς της επιστήμης. Είναι ευρέως γνωστό ότι η επιστήμη της βαθιάς μάθησης θα έχει πολλές περισσότερες επιτυχίες στο εγγύς μέλλον γιατί απαιτεί πολύ λιγότερη ανθρώπινη προσπάθεια και μπορεί εύκολα να επωφεληθεί από την ραγδαία αύξηση του όγκου των διαθέσιμων δεδομένων και αυτοματισμών. Νέοι αλγόριθμοι μηχανικής μάθησης και αρχιτεκτονικές που αναπτύσσονται για βαθιά νευρωνικά δίκτυα θα επιταχύνουν σημαντικά αυτή την πρόοδο.

1.4 Ακραία Καιρικά Φαινόμενα - Κλιματική Αλλαγή

Η αλλαγή του κλίματος έχει ήδη εμφανή αποτελέσματα, που εκτείνονται από την αύξηση της θερμοκρασίας έως την άνοδο της στάθμης της θάλασσας σαν αποτέλεσμα της τήξης των πολικών παγετών, καθώς και τη συχνότερη εμφάνιση καταιγίδων και πλημμυρών. Οι μεταβολές αυτές θα επιφέρουν με τη σειρά τους σοβαρές επιπτώσεις στην ακεραιότητα των οικοσυστημάτων, τους υδατικούς πόρους, τη δημόσια υγεία, την προσφορά τροφής, τη βιομηχανία, τις γεωργικές καλλιέργειες, τις μεταφορές και τις υποδομές.

Για αιώνες, οι άνθρωποι έχουν τροποποιήσει τα επίγεια και υδάτινα συστήματα για την κάλυψη βασικών ενεργειακών αναγκών καθώς και για την ικανοποίηση άλλων απαιτήσεων, που οδηγούν σε μεγάλες αλλαγές στην παγκόσμια κάλυψη γης, ατμοσφαιρικές συγκεντρώσεις αερίων θερμοκηπίου και το μέλλον της γης ικανότητα δέσμευσης άνθρακα. Μαζί, χρήση γης, η χρήση ενέργειας και η κλιματική αλλαγή αντιπροσωπεύουν τρεις κύριες, αλληλένδετες δυνάμεις της ανθρωπογενούς παγκόσμιας αλλαγής.

Οι ανθρώπινη δραστηριότητα έχει άμεσες επιπτώσεις στη γη που καταλαμβάνει και έχει τόσο άμεσες όσο και έμμεσες οικολογικές επιπτώσεις σε μεγαλύτερες κλίμακες. Για παράδειγμα, γραμμικά χαρακτηριστικά, όπως δρόμοι, σεισμικά δίκτυα εξερεύνησης πετρελαίου και φυσικού αερίου, γραμμές μεταφοράς, αποχετευτικές τάφροι, πεζοδρόμια και φράγματα, μπορούν να παρέχουν πρόσβαση σε νέες περιοχές και διαταράσσει τις οικολογικές λειτουργίες των παρακείμενων συμπεριλαμβανομένης της ιστορικής αποψίλωσης των δασών, επηρεάζουν τις παγκόσμιες συγκεντρώσεις διοξειδίου του άνθρακα (CO₂) της ατμόσφαιρας, τα αερολύματα και το άλμπεντο και συνεπώς το παγκόσμιο κλίμα. Και εξίσου σημαντικό είναι να αναφερθεί ότι οι πρακτικές διαχείρισης επηρεάζουν την τρέχουσα και τη μελλοντική ικανότητα μεταφοράς άνθρακα και η ροή των αερίων θερμοκηπίου μεταξύ γης και ατμόσφαιρας, χαρακτηρίζοντας έτσι τον ρόλο των χερσαίων οικοσυστημάτων στον μετριασμό της κλιματικής αλλαγής.

Η κλιματική αλλαγή έχει επιφέρει ακραία καιρικά φαινόμενα, όπως πλημμύρες, πυρκαγιές κλπ. τα οποία τροποποιούν σημαντικά την φυσική γήινη επιφάνεια, το ποσοστό δασικής έκτασης στη γή και το κλίμα. Είναι πολύ σημαντικό να βρεθούν αυτοματισμοί οι

οποίοι θα ελέγχουν και θα ανιχνεύουν άμεσα τις επιπτώσεις των ακραίων καιρικών φαινομένων που έχουν εμφανιστεί λόγω της κλιματικής αλλαγής. Μια σημαντική επίπτωση των φαινομένων αυτών είναι η αποψίλωση (deforestation).

Η κλιματική αλλαγή έχει αναγνωριστεί ως σημαντική απειλή για την επιμονή του είδους μας στη Γη και η πιο σημαντική πρόκληση της εποχής μας.

Η κλιματική αλλαγή είναι τόσο αιτία όσο και συνέπεια της αποψίλωσης και της υποβάθμισης των δασών. Τα ακραία φαινόμενα που προκαλεί, όπως πυρκαγιές, ξηρασίες και πλημμύρες, επηρεάζουν τα δάση. Με τη σειρά της, η απώλεια δασών είναι επιβλαβής για το κλίμα, καθώς τα δάση παίζουν σημαντικό ρόλο στην παροχή καθαρού αέρα, στη ρύθμιση του κύκλου του νερού, στη δέσμευση CO₂, στην πρόληψη της απώλειας βιοποικιλότητας και της διάβρωσης του εδάφους.

Η συγκεκριμένη εργασία αφορά άμεσα την κλιματική αλλαγή και τις επιπτώσεις της σε δασικές εκτάσεις καθώς στόχος αποτέλεσε η αυτοματοποιημένη ανίχνευση αποψιλωμένων περιοχών που έχουν προκληθεί λόγω ακραίων φαινομένων που προκαλεί η κλιματική αλλαγή όπως πυρκαγιές ή ξηρασίες. Η ανακάλυψη αυτοματοποιημένων μηχανισμών για ανίχνευση περιοχών αποψίλωσης βοηθά στην πιο εύκολη, γρήγορη και ακριβής διαδικασία εντοπισμού των συγκεκριμένων εκτάσεων με στόχο την καλύτερη αξιολόγηση του φαινομένου και την καλύτερη δυνατή γνώση των πληγέντων εκτάσεων για αντιμετώπιση και έλεγχο του φαινομένου της αποψίλωσης (*deforestation*).

Κεφάλαιο 2 – Θεωρητικό Υπόβαθρο

2.1 Νευρωνικά Δίκτυα (Neural Networks)

2.1.1 Εισαγωγή

Ο όρος **Νευρωνικά Δίκτυα (Neural Networks)** περιγράφει ένα σύνολο διαφορετικών μαθηματικών μοντέλων, εμπνευσμένων από αντίστοιχα βιολογικά μοντέλα, δηλαδή μοντέλα που προσπαθούν να μιμηθούν τη συμπεριφορά των νευρώνων του ανθρώπινου εγκεφάλου.

Τα **τεχνητά νευρωνικά δίκτυα (artificial neural networks, ANN)** αναπαριστούν πρότυπα (μοντέλα) του νευρωνικού συστήματος του ανθρώπου προκειμένου να επιτυγχάνουν αντίστοιχες διεργασίες με αυτές των βιολογικών νευρωνικών δικτύων στους ζώντες οργανισμούς μέσω περίπλοκων μαθηματικών μοντέλων. Η βασική τους διαφορά από τα βιολογικά δίκτυα είναι ότι τα νευρωνικά δίκτυα παίρνουν γνώσεις (μαθαίνουν) με την εξάσκηση και την εμπειρία, όπως ακριβώς και οι άνθρωποι, αλλά διαφέρουν στο ότι δεν ακολουθούν ορισμένους προκαθορισμένους κανόνες, που είναι χαρακτηριστικό των υπολογιστών. Έτσι στα νευρωνικά δίκτυα χρησιμοποιούνται ιδέες όπως, π.χ. ένα δίκτυο μαθαίνει και εκπαιδεύεται, θυμάται ή ξεχνά μια αριθμητική τιμή, κλπ. πράγματα που μέχρι τώρα τα αποδίδουμε μόνο στην ανθρώπινη σκέψη.

2.1.2 Αρχιτεκτονική ΤΝΔ

Τα τεχνητά νευρωνικά δίκτυα (artificial neural networks, ANN) συνήθως οργανώνονται σε επίπεδα (layers), όπως το επίπεδο εισόδου (input layer), τα ενδιάμεσα επίπεδα που καλούνται κρυμμένα επίπεδα (hidden layers) και το επίπεδο εξόδου (output layer). Το κάθε ένα από τα παραπάνω επίπεδα αποτελείται από έναν αριθμό νευρώνων (neurons) κατάλληλα συνδεδεμένων μεταξύ τους, ώστε ένας νευρώνας του πρώτου επιπέδου π.χ. επιπέδου εισόδου (input layer) να έχει συνδέσμους με πολλούς άλλους νευρώνες του επόμενου επιπέδου π.χ. ενός κρυμμένου επιπέδου (hidden layer) του τεχνητού νευρωνικού δικτύου. Οι παραπάνω σύνδεσμοι αντιπροσωπεύουν τις τιμές βάρους (w_i) οι οποίες καθορίζουν πόσο στενά είναι συνδεδεμένοι οι δύο νευρώνες που συνδέονται με ένα ορισμένο βάρος (w_i). Οι νευρώνες (neurons) είναι η πιο μικρή ανεξάρτητη μονάδα του δικτύου οι οποίοι συνεχώς και ασταμάτητα επεξεργάζονται πληροφορίες, παίρνοντας και στέλνοντας ηλεκτρικά σήματα σε άλλους νευρώνες.

Ο τρόπος σύνδεσης των νευρώνων (neurons) των παραπάνω επιπέδων (layers) καθορίζει την αρχιτεκτονική του ΤΝΔ. Υφίστανται δύο βασικές κατηγορίες ΤΝΔ:

- **πρόσθιας τροφοδότησης (feed forward)** και
- **οπίσθιας τροφοδότησης (back propagation)**

Στα **νευρωνικά δίκτυα πρόσθιας τροφοδότησης**, οι νευρώνες είναι οργανωμένοι σε διαφορετικά επίπεδα, ώστε οι νευρώνες του ενός επιπέδου να τροφοδοτούν τους νευρώνες του επόμενου επιπέδου έως ότου τροφοδοτηθούν και οι νευρώνες του τελευταίου επιπέδου, δηλαδή του επιπέδου εξόδου. Συνεπώς, δεν υπάρχει έξοδος νευρώνα ενός επιπέδου που να αποτελεί είσοδο νευρώνα του ίδιου ή προηγούμενων επιπέδων.

Στα **νευρωνικά δίκτυα οπίσθιας τροφοδότησης**, τα οποία καλούνται και ανατροφοδοτούμενα ΤΝΔ (recurrent ANN), επιτρέπεται στους νευρώνες ενός επιπέδου να τροφοδοτούν και νευρώνες του ίδιου επιπέδου ή και προηγούμενων επιπέδων. Αν η ανατροφοδότηση αφορά νευρώνες στο ίδιο επίπεδο, τότε τα δίκτυα καλούνται **αυτοσυσχετιζόμενες μνήμες (autoassociated memories)** διαφορετικά, καλούνται **ετεροσυσχετιζόμενες μνήμες (heteroassociated memories)**.

Στα ανατροφοδοτούμενα ΤΝΔ δεν υπάρχουν συνήθως άνω του ενός ενδιάμεσα (κρυμμένα) επίπεδα. Αν και τα ανατροφοδοτούμενα δίκτυα είναι πολύ χρήσιμα, τα περισσότερα των νευρωνικών δικτύων είναι πρόσθιας τροφοδότησης.

2.1.3 Εκπαίδευση νευρωνικού δικτύου

Τα μαθηματικά μοντέλα των τεχνητών νευρωνικών δικτύων, έχουν την δυνατότητα να «μαθαίνουν» ύστερα από εκπαίδευση. Η εκπαίδευση τους επιτυγχάνεται μέσω αλγορίθμων που μπορεί να υιοθετούν μοντέλα μάθησης με επίβλεψη ή χωρίς επίβλεψη. Η εκπαίδευση έχει ως βασικό στόχο να βρεθεί ένας τρόπος αλλαγής των συνδεσμικών βαρών που θα έχει ως αποτέλεσμα την αλλαγή της γενικής συμπεριφοράς του δικτύου με την αύξηση της ικανότητας του δικτύου να παρέχει στο μέλλον μία επιθυμητή έξοδο μετά από μία δεδομένη είσοδο.

- **Εκπαίδευση με επίβλεψη (Supervised Neural Network)**, όπου το νευρωνικό δίκτυο μαθαίνει να απεικονίζει δεδομένες εισόδους σε εξόδους εκ των προτέρων γνωστές (σύνολο εκπαίδευσης), με απώτερο στόχο τη γενίκευση της αναγνώρισης αυτής και για παρεμφερείς εισόδους στο μέλλον
- **Εκπαίδευση χωρίς επίβλεψη (Unsupervised Neural Network)**, όπου το νευρωνικό δίκτυο κατασκευάζει απεικονίσεις από μια αναπαράσταση σε μια άλλη.

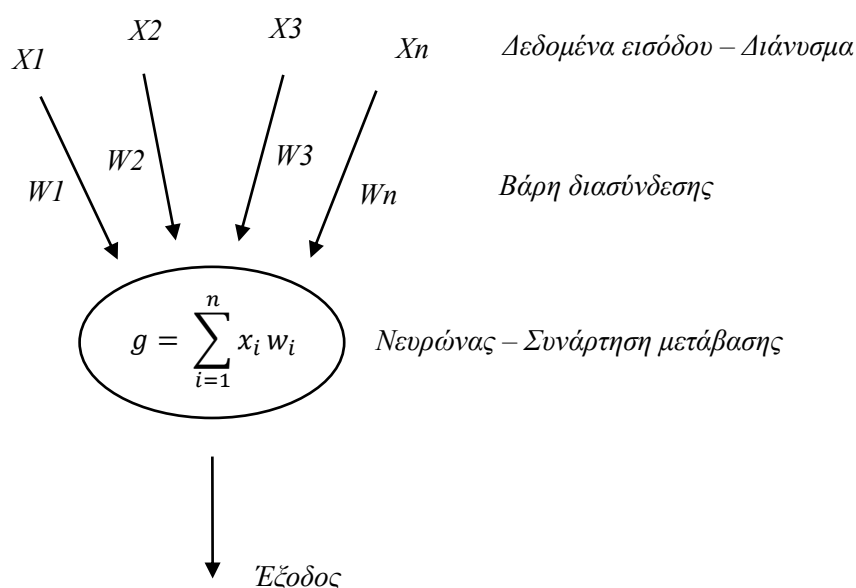
Κάθε τεχνητός νευρώνας αποτελείται από πολλές εισόδους x_i και μία μόνο έξοδο y . Κάθε είσοδος x_i «ζυγίζεται» με ένα βάρος διασύνδεσης w_i και τα αποτελέσματα αθροίζονται μέσω της συνάρτησης αθροίσματος (summation function) F :

$$F = \sum_i^n x_i w_i$$

Ο τεχνητός νευρώνας δίνει έξοδο μέσω της συνάρτησης μετάβασης (transfer function), μόνο όταν το ζυγισμένο άθροισμα των εισόδων είναι μεγαλύτερο μιας ορισμένης τιμής κατωφλίου (threshold value) θ , δηλαδή όταν:

$$\sum_i^n x_i w_i - \theta > 0$$

Στο παρακάτω Σχήμα 2.1 παρουσιάζεται το στοιχειώδες Perceptron (basic Perceptron) που αποτελεί ένα από τα απλούστερα ΤΝΔ που προσομοιώνουν τον φυσικό νευρώνα, δηλαδή ένα ΤΝΔ που αποτελείται από έναν μόνο νευρώνα. Η έξοδος του Perceptron για ένα διάνυσμα εισόδου $X=(X_1, X_2, X_3..., X_n)$ δίνεται μέσω της συνάρτησης μετάβασης g η οποία φαίνεται παρακάτω στο Σχήμα 2.1 :



Σχήμα 2.1: Στοιχειώδης τεχνικός νευρώνας (Perceptron)

Η εκπαίδευση γίνεται με την παρουσίαση μιας ομάδας από πρότυπα στο δίκτυο, αντιπροσωπευτικά ή παρόμοια με αυτά που πρέπει να μάθει το δίκτυο. Αυτό σημαίνει ότι το δίκτυο δέχεται ως εισόδους κάποια πρότυπα για τα οποία είναι γνωστό ποια πρέπει να είναι η έξοδος στο δίκτυο, είναι δηλαδή γνωστό ποιος είναι ο στόχος και τι πρέπει να δίδει το δίκτυο ως απάντηση στα πρότυπα που του έχουν παρουσιαστεί (βλ. Σχήμα 2.1). Ουσιαστικά είναι σαν να δίδεται στο δίκτυο η ερώτηση και η απάντηση που αντιστοιχεί στην εκάστοτε ερώτηση. Το δίκτυο με τα δεδομένα αυτά τροποποιεί την εσωτερική του δομή ώστε να κάνει την ίδια αντιστοιχία που του δόθηκε. Ακολουθώντας, αφού βρει την σωστή εσωτερική δομή, τότε θα μπορεί να λύνει και άλλα ανάλογα προβλήματα τα οποία δεν τα έχει δει προηγουμένως, δηλαδή δεν έχει εκπαιδευθεί στα πρότυπα των προβλημάτων αυτών. Οποσδήποτε όμως, τα προβλήματα αυτά θα πρέπει να είναι της ίδιας φύσης και των ίδιων χαρακτηριστικών όπως αυτά της εκπαίδευσης και όχι διαφορετικά. Αυτός είναι ο πιο συνηθισμένος τρόπος εκπαίδευσης.

Όταν ένας νευρώνας ενεργοποιείται, υπολογίζει μία συνάρτηση από όλα τα δεδομένα που έχει, και συγκρίνει την τιμή της συνάρτησης αυτής με μια τιμή κατωφλίου η οποία είναι χαρακτηριστική για τον νευρώνα αυτόν. Αν η τιμή της συνάρτησης είναι μεγαλύτερη από την τιμή κατωφλίου, τότε ο νευρώνας υπολογίζει την έξοδο, την οποία προωθεί ως είσοδο στον επόμενο (ή στους επόμενους) νευρώνα (νευρώνες). Κατά την διάρκεια της εκπαίδευσης το μόνο πράγμα που αλλάζει είναι η τιμές των βαρών των συνδέσεων των νευρώνων. Οι αλλαγές στις τιμές των βαρών δεν γίνεται πάντα με τον ίδιο τρόπο, αλλά εξαρτάται σημαντικά από την μέθοδο που χρησιμοποιούμε.

Μετά την εκπαίδευση (training) ακολουθεί η φάση της **ανάκλησης (recall)** δηλαδή ο έλεγχος του δικτύου με τη βοήθεια ενός συνόλου δειγμάτων τα οποία δεν είχαν πάρει μέρος στην διαδικασία της εκπαίδευσης.

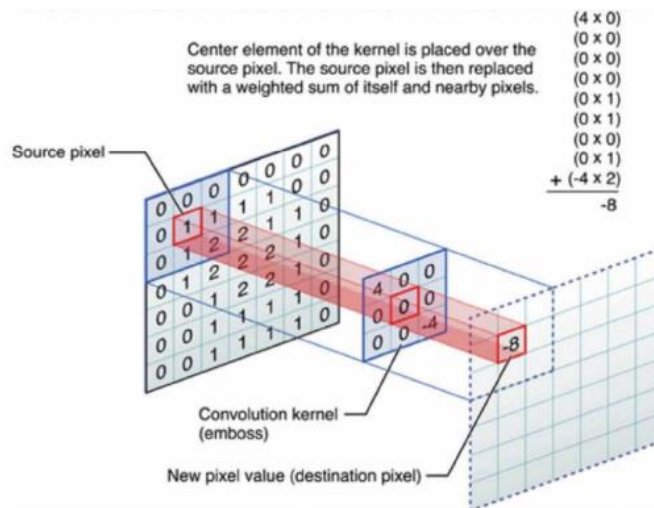
2.2 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks) και Τηλεπισκόπηση (Remote Sensing Imagery)

Η μελέτη και επεξεργασία δορυφορικών εικόνων με στόχο την συνεχή παρακολούθηση του εδάφους και των μεταβολών του καλείται Τηλεπισκόπηση (*Remote Sensing*). Η επιστήμη της τηλεπισκόπησης έχει στραφεί στο κλάδο της τεχνητής νοημοσύνης (*artificial intelligence*) και ειδικότερα της **Βαθιάς Μάθησης** (*deep learning*) λόγω της ικανότητάς της να εξάγει ζητούμενα χαρακτηριστικά από ένα σύνολο δορυφορικών δεδομένων – εικόνων και να εκτελεί σε υψηλό επίπεδο σημασιολογική κατάτμηση (*semantic segmentation*). Τα τεχνητά νευρωνικά δίκτυα τα οποία εξειδικεύονται και χρησιμοποιούνται ευρέως σε εφαρμογές Τηλεπισκόπησης, ονομάζονται Συνελκτικά Νευρωνικά Δίκτυα (**Convolutional Neural Networks (CNN)**).

Τα *Convolutional Neural Networks (CNN)* αποτελούν έναν από τους κλάδους της Βαθιάς Μάθησης (*deep learning*) και εξειδικεύονται στην αναγνώριση μοτίβων – σχημάτων (*pattern recognition*) σε ένα σύνολο εικόνων. Η αρχιτεκτονική των Συνελκτικών Νευρωνικών Δικτύων σχεδιάζεται έτσι ώστε να εκμεταλλεύεται την δισδιάστατη δομή των εικόνων εισόδου. Αυτό επιτυγχάνεται με τοπικές συνδέσεις και κατάλληλα βάρη προκειμένου να δημιουργηθούν χαρακτηριστικά ανεξαρτήτως μετατοπίσεων. Άλλο ένα πλεονέκτημα των Συνελκτικών Νευρωνικών Δικτύων αποτελεί το γεγονός ότι είναι ευκολότερα στην εκπαίδευση και έχουν πολύ λιγότερες παραμέτρους.

Η αρχιτεκτονική των **Convolutional Neural Networks (CNN)** είναι παρόμοια με την βασική δομή όλων των **τεχνητών νευρωνικών δικτύων** (*artificial neural networks, ANN*). Η διαφορά έγκειται στην ύπαρξη των λεγόμενων **φίλτρων** (*filters*) που πρέπει να ορίζονται για κάθε ένα από τα ενδιάμεσα επίπεδα, τα λεγόμενα κρυμμένα επίπεδα (*hidden layers*), του νευρωνικού δικτύου. Τα παραπάνω κρυμμένα επίπεδα (*hidden layers*) στα *Convolutional Neural Networks (CNN)*, ονομάζονται επίπεδα συνέλιξης (**convolutional layers**). Τα συνελκτικά νευρωνικά δίκτυα λοιπόν αποτελούνται από τα επίπεδα συνέλιξης (*convolutional layers*) συχνά ακολουθούμενα από ένα επίπεδο υποδειγματοληψίας ακολουθούμενο από ένα ή περισσότερα πλήρως διασυνδεδεμένα επίπεδα όπως συναντάμε και σε ένα πολυεπίπεδο Νευρωνικό Δίκτυο. Συνεπώς, για κάθε επίπεδο συνέλιξης που προστίθεται στο δίκτυο κρίνεται απαραίτητο να καθοριστεί το πλήθος των φίλτρων που θα έχει.

Η είσοδος σε ένα επίπεδο συνέλιξης είναι μια $m \times n \times c$ εικόνα όπου m και n είναι το ύψος και το πλάτος της εικόνας αντίστοιχα, ενώ το c είναι ο αριθμός των καναλιών, για παράδειγμα για έγχρωμες εικόνες RGB $c = 3$. Το επίπεδο συνέλιξης έχει k φίλτρα (*kernels*) μεγέθους $w \times w \times r$ όπου είναι μικρότερο από τη διάσταση της εικόνας και μπορεί να είναι ίδιου μεγέθους με τα κανάλια ή μικρότερου και μπορεί να ποικίλει για κάθε φίλτρο. Το μέγεθος των φίλτρων προκαλεί τοπικά συνδεδεμένη δομή όπου το καθένα συνελίσσεται με κάθε εικόνα για να παράγουν χάρτες χαρακτηριστικών μεγέθους $(m - w + 1) \times (n - w + 1)$. Κάθε χαρακτηριστικό υποδειγματοληπτείται (*downsampling*) τυπικά με κάποιο *pooling* επίπεδο σε $p \times p$ συνεχείς περιοχές όπου το p παίρνει συνήθως τιμές μεταξύ 2 και 5 αλλά για μεγάλες εικόνες εισόδου παίρνει και μεγαλύτερες τιμές. Πριν ή μετά το *pooling layer* συνήθως ακολουθεί προσθήκη μίας παραμέτρου πόλωσης (*bias*) και μια συνάρτηση ενεργοποίησης σε κάθε χάρτη χαρακτηριστικών. Η παρακάτω εικόνα 1 δείχνει παράδειγμα ενός επιπέδου συνέλιξης.



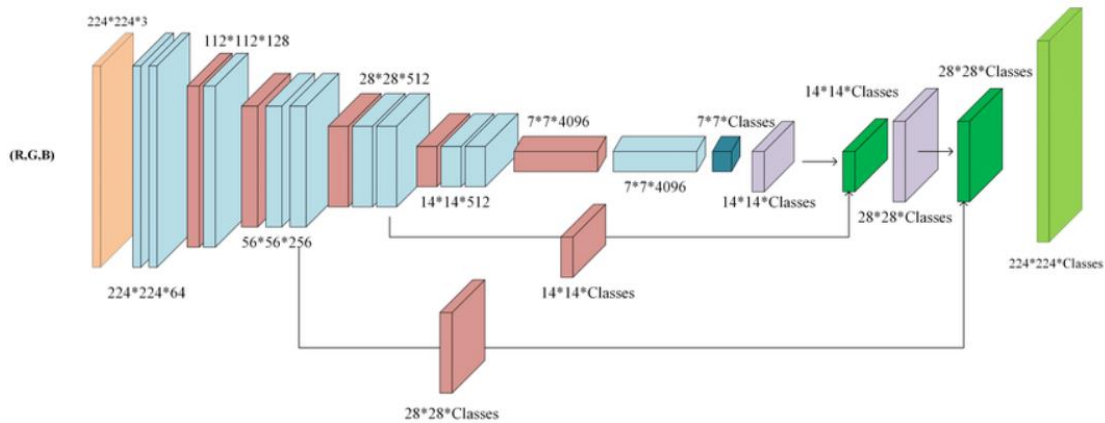
Εικόνα 1 - Επίπεδο συνέλιξης

Τα **φίλτρα**, λοιπόν, είναι πίνακες σχετικά μικρού μεγέθους π.χ 3x3 ή 4x4 ή 5x5 με τυχαίους αριθμούς. Όταν οι εικόνες εισόδου (*inputs*) εισέρχονται στο πρώτο επίπεδο συνέλιξης (*convolutional layer*), το εκάστοτε φίλτρο που ορίστηκε (έστω φίλτρο 3x3) θα εφαρμοστεί για κάθε block 3x3 εικονοστοιχείων (*pixels*) ολόκληρης της εικόνας ώστε να προκύψει νέα τιμή εικονοστοιχείου. Η διαδικασία αυτή ονομάζεται **συνέλιξη**. Οπότε, η συνέλιξη του κάθε block 3x3 της αρχικής εικόνας εισόδου (*input image*) με το εκάστοτε φίλτρο δημιουργεί μια νέα τιμή εικονοστοιχείου (*pixel value*). Η νέα εικόνα που δημιουργείται αποτελεί την καινούργια καταχώρηση (*input image*) που θα χρησιμοποιηθεί στο δεύτερο κρυμμένο επίπεδο συνέλιξης (*hidden layer*), *convolutional layer*, με σκοπό την επανάληψη της παραπάνω αναγραφόμενης διαδικασίας συνέλιξης.

Τα **convolutional layers** σε ένα **Convolutional Neural Networks (CNN)** έχουν την ικανότητα να ανιχνεύουν *patterns*. Πιο συγκεκριμένα, με κάθε ένα από τα convolutional layer που ορίζονται στο εκάστοτε convolutional neural network ορίζεται και το πλήθος των φίλτρων που πρέπει να έχει κάθε convolutional layer. Τα παραπάνω φίλτρα ανιχνεύουν και τα λεγόμενα «*patterns*». Με τον όρο «**patterns**» ορίζεται μια συγκεκριμένη γεωμετρία ή σχήμα που πρέπει να αναγνωριστεί από ένα φίλτρο όπως, για παράδειγμα γωνίες, ακμές, καμπυλότητες ή ακόμα και ολόκληρα αντικείμενα ανάλογα το φίλτρο.

2.2.1 Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks)

Μία επιπλέον κατηγορία Νευρωνικών Δικτύων αποτελούν τα Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks). Η κύρια διαφορά με τα Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks (CNN)) είναι η **απόλεια πλήρως συνδεδεμένων επιπέδων στην έξοδο** (fully-connected layers), εν αντιθέσει με τα Συνελκτικά Νευρωνικά Δίκτυα (**Convolutional Neural Networks (CNN)**) που αναλύθηκαν προηγουμένως. Δηλαδή, τα Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks) μαθαίνουν πληροφορία μόνο από φίλτρα. Τα Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks) θεωρούνται κατάλληλα για προβλήματα Σηματολογικής Κατάτμησης αντικειμένων από εικόνες και αφορούν άμεσα την παρούσα διπλωματική εργασία καθώς στόχος αποτέλεσε η πρόβλεψη σε πρόβλημα αναγνώρισης ορισμένων μοτίβων «*patterns*» σε επίπεδο εικονοστοιχείων από ένα σύνολο εικόνων μέσω της εκπαίδευσης νευρωνικού δικτύου για τον εντοπισμό αποψιλωμένων εκτάσεων.



Εικόνα 2 – Δομή Πλήρως Συνελκτικού Νευρωνικού Δικτύου

Μερικά θετικά στοιχεία που καταστούν τα Πλήρως Συνελκτικά Νευρωνικά Δίκτυα (Fully Convolutional Neural Networks) κατάλληλα για Σημαιολογική Κατάτμηση σε αντίθεση με άλλου είδους Νευρωνικά Δίκτυα είναι τα εξής:

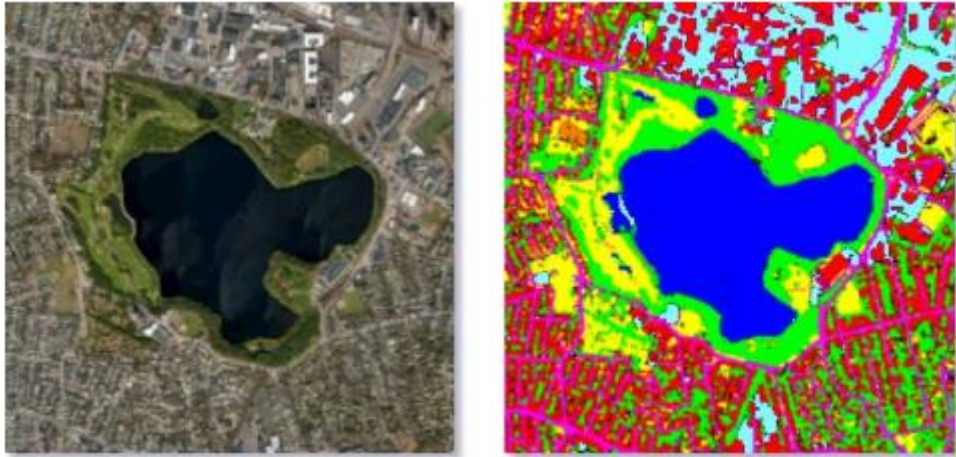
- Χρησιμοποιούν όλη την πληροφορία της εικόνας.
- Κρατάνε την χωρική πληροφορία (spatial information) από την εικόνα.
- Είναι πιο γρήγορα στην εκπαίδευση αλλά και στην συμπερασματολογία.
- Είναι αμετάβλητα ως προς το μέγεθος εισόδου της εικόνας.

2.3 Σημαιολογική Κατάτμηση Εικόνων (Semantic Image Segmentation)

Το πρόβλημα της σημαιολογικής κατάτμησης εικόνων (*semantic segmentation*) χαρακτηρίζεται ως το πρόβλημα της ταξινόμησης του συνόλου των εικονοστοιχείων (*pixels*) μιας εικόνας σε διάφορες κλάσεις με συγκεκριμένες σημαιολογικές κατηγορίες (*semantic labels*), οι οποίες ορίζονται από το περιεχόμενο των εικόνων. Δηλαδή, στο πρόβλημα της σημαιολογικής κατάτμησης αντικειμένων, σημαντική είναι η τοποθεσία ενός αντικειμένου στην εικόνα αλλά και η εύρεση των ακριβών ορίων του αντικειμένου, καθώς γίνεται η τοποθέτηση κάθε εικονοστοιχείου της εικόνας σε μια κατηγορία αντικειμένου ή ετικέτα, έτσι ώστε εικονοστοιχεία με την ίδια ετικέτα να μοιράζονται ορισμένα χαρακτηριστικά.

Η έξοδος ενός συστήματος σημαιολογικής κατάτμησης είναι μια εικόνα ίδιων διαστάσεων με την αρχική, στην οποία το κάθε εικονοστοιχείο έχει αντιστοιχηθεί σε μια κλάση (ή κατηγορία ή ετικέτα). Συνεπώς, στην διαδικασία της σημαιολογικής κατάτμησης εκτελείται η σήμανση σε επίπεδο εικονοστοιχείου με ένα σύνολο κατηγοριών (κλάσεις) (π.χ. δέντρο, νερό, έδαφος). Γίνεται, λοιπόν, αντιληπτό ότι αποτελεί γενικά ένα πιο απαιτητικό εγχείρημα από την ταξινόμηση ολόκληρης της εικόνας, η οποία προβλέπει μια ενιαία ετικέτα για ολόκληρη την εικόνα.

Η τμηματοποίηση (ή κατάτμηση) με Νευρωνικά δίκτυα βασίζεται στην επεξεργασία μικρών περιοχών μιας εικόνας. Μετά από αυτήν την επεξεργασία ο μηχανισμός λήψης αποφάσεων επισημαίνει τις περιοχές μιας εικόνας ανάλογα με την κατηγορία που αναγνωρίζει από το νευρωνικό δίκτυο. Η συγκεκριμένη διπλωματική εργασία αφορά την αυτοματοποιημένη σημαιολογική κατάτμηση δορυφορικών εικόνων για τον εντοπισμό αποψιλωμένων περιοχών (*deforestation*) μέσω συνελκτικών νευρωνικών δικτύων (*Convolutional Neural Networks*).



Εικόνα 3 - Παράδειγμα σημασιολογικής κατάτμησης δορυφορικής εικόνας

2.3.1 Σημασιολογική Κατάτμηση Εικόνων (Semantic Image Segmentation) και Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)

Συνήθως, η αρχιτεκτονική ενός μοντέλου (Συνελκτικού Νευρωνικού Δικτύου (*Convolutional Neural Network*) αποτελείται από πολλαπλά συνελκτικά επίπεδα (*convolutional layers*), μη-γραμμικές δραστηριότητες (*non-linear activations*), κανονικοποίηση παρτίδας (*batch normalization*) και συγκεντρωτικά επίπεδα (*pooling layers*). Τα αρχικά επίπεδα μαθαίνουν βασικού επιπέδου έννοιες, όπως την αναγνώριση ακμών και χρωμάτων, ενώ τα πιο βαθιά επίπεδα του μοντέλου αποκτούν υψηλού επιπέδου αντίληψη, όπως είναι η αναγνώριση διαφορετικών αντικειμένων.

Στα αρχικά επίπεδα του μοντέλου οι νευρώνες περιλαμβάνουν πληροφορία για μια μικρή περιοχή της εικόνας. Αντίθετα, στα πιο βαθιά επίπεδα του μοντέλου οι νευρώνες περιλαμβάνουν πληροφορία για ένα πολύ μεγαλύτερο τμήμα της εικόνας. Συνεπώς, όσο προστίθενται περισσότερα επίπεδα, μειώνεται όλο και περισσότερο το μέγεθος της εικόνας, ενώ τα κανάλια αντίστοιχα αυξάνονται. Η υποδειγματοληψία (*downsampling*) επιτυγχάνεται από τα επίπεδα συγκέντρωσης (*pooling layers*).

Για το πρόβλημα της σημασιολογικής κατάτμησης καθίσταται αναγκαία η διατήρηση της χωρικής πληροφορίας. Ως εκ τούτου, δεν χρησιμοποιούνται πλήρως συνδεδεμένα επίπεδα, όπως αντιθέτως συμβαίνει στην περίπτωση της ταξινόμησης (*image classification*), τα οποία “καταστρέφουν” όλη την χωρική πληροφορία. Τα Συνελκτικά Νευρωνικά Δίκτυα (*Convolutional Networks*) χρησιμοποιούνται ευρέως σε προβλήματα ταξινόμησης (*image classification*) κατά τα οποία πρέπει να εντοπιστεί ο χωρικός τανυστής (*spatial tensor*) από τα συνελκτικά επίπεδα σε ένα σταθερού μήκους διάνυσμα και προκειμένου να συμβεί αυτό πρέπει να “καταστραφεί” η χωρική πληροφορία. Ωστόσο, για το πρόβλημα της σημασιολογικής κατάτμησης ενδείκνυνται τα Πλήρως Συνελκτικά Νευρωνικά Δίκτυα.

Στην περίπτωση της σημασιολογικής κατάτμησης (*semantic image segmentation*), όπου πραγματοποιείται η παρούσα εργασία, τα συνελκτικά επίπεδα συζευγνούν με τα *downsampling* επίπεδα και παράγουν χαμηλής ανάλυσης τανυστή ο οποίος περιέχει υψηλής ανάλυσης πληροφορία. Με αυτό τον τρόπο πρέπει να παραχθούν υψηλής ανάλυσης εικόνες στις οποίες έχει εφαρμοστεί διαδικασία σημασιολογικής κατάτμησης (*segmentation outputs*). Προκειμένου να επιτευχθεί το ανωτέρω προστίθενται περισσότερα συνελκτικά επίπεδα τα οποία συζευγνούν με *downsampling* επίπεδα, ώστε να αυξηθεί το μέγεθος του χωρικού τανυστή. Καθώς αυξάνεται η ανάλυση, μειώνεται αντίστοιχα ο αριθμός των καναλιών, καθώς γίνεται επιστροφή στην πληροφορία χαμηλού επιπέδου. Η ανωτέρω διαδικασία είναι γνωστή ως δομή *encoder-decoder*, δομή κωδικοποίησης – αποκωδικοποίησης κατά την οποία τα

κατάτμησης εικόνων αποτελεί την προφανή επιλογή. Όταν χρησιμοποιούνται τα *Fully Convolutional Neural Networks* για σημασιολογική κατάτμηση εικόνων, η έξοδος είναι μια εικόνα και όχι ένα διάνυσμα σταθερού μήκους (όπως στην περίπτωση της ταξινόμησης).

Επιπλέον, αξίζει να αναφερθούν οι autoencoders. Οι autoencoders είναι ένας τύπος τεχνητού νευρωνικού δικτύου που χρησιμοποιούνται για την εκμάθηση αποτελεσματικών κωδικοποιήσεων ή αναπαραστάσεων των δεδομένων εισόδου, συνήθως με σκοπό τη μείωση της διαστατικότητας ή την εκμάθηση γεννητικών μοντέλων δεδομένων. Λειτουργούν κωδικοποιώντας μια είσοδο σε μια αναπαράσταση σταθερού μεγέθους, συνήθως μικρότερη, και στη συνέχεια αποκωδικοποιώντας την πίσω στην αρχική μορφή εισόδου. Το δίκτυο εκπαιδεύεται ώστε να αγνοεί τον θόρυβο του σήματος, με αποτέλεσμα να καταλήγει σε μια διαδικασία που συλλαμβάνει τα πιο σημαντικά χαρακτηριστικά των δεδομένων στην κωδικοποιημένη αναπαράσταση.

Encoder: Αυτό το τμήμα του δικτύου συμπιέζει την είσοδο σε μια αναπαράσταση λανθάνοντος χώρου. Συλλαμβάνει τις πληροφορίες που είναι πιο σημαντικές για την ανακατασκευή των δεδομένων εισόδου.

Decoder: Αυτό το τμήμα αποσκοπεί στην ανακατασκευή των δεδομένων εισόδου από την εσωτερική αναπαράσταση. Αντικατοπτρίζει τη δομή του κωδικοποιητή.

Bottleneck: Το στρώμα που περιέχει τη συμπιεσμένη αναπαράσταση των δεδομένων εισόδου. Η διαστατικότητα του bottleneck (δηλαδή ο αριθμός των νευρώνων) είναι μικρότερη από εκείνη των δεδομένων εισόδου, αναγκάζοντας τον αυτόματο κωδικοποιητή να μάθει μια συμπιεσμένη αναπαράσταση.

Ο συγκεκριμένος τύπος νευρωνικού δικτύου (autoencoders) ακολουθούν μια συγκεκριμένη διαδικασία εκπαίδευσης. Οι autoencoders εκπαιδεύονται με μεθόδους μάθησης χωρίς επίβλεψη, δηλαδή δεν απαιτούν επισημειωμένα δεδομένα. Στόχος είναι η ελαχιστοποίηση της διαφοράς μεταξύ της εισόδου και της εξόδου (ανακατασκευή), συνήθως χρησιμοποιώντας μια συνάρτηση απωλειών όπως το μέσο τετραγωνικό σφάλμα (MSE) για συνεχή είσοδο ή η δυαδική σταυροειδής εντροπία για δυαδική είσοδο. Για την ενημέρωση των βαρών του δικτύου χρησιμοποιείται η οπισθοδιάδοση, παρόμοια με άλλα νευρωνικά δίκτυα.

Οι autoencoders είναι ένα ισχυρό εργαλείο στο οπλοστάσιο της μηχανικής μάθησης, ικανό να μαθαίνει χωρίς επίβλεψη από την εγγενή δομή των δεδομένων και ευρέως εφαρμόσιμο σε διάφορους τομείς όπως η επεξεργασία εικόνας, η επεξεργασία φυσικής γλώσσας και άλλα.

2.4 Swin UNet - Swin Transformer

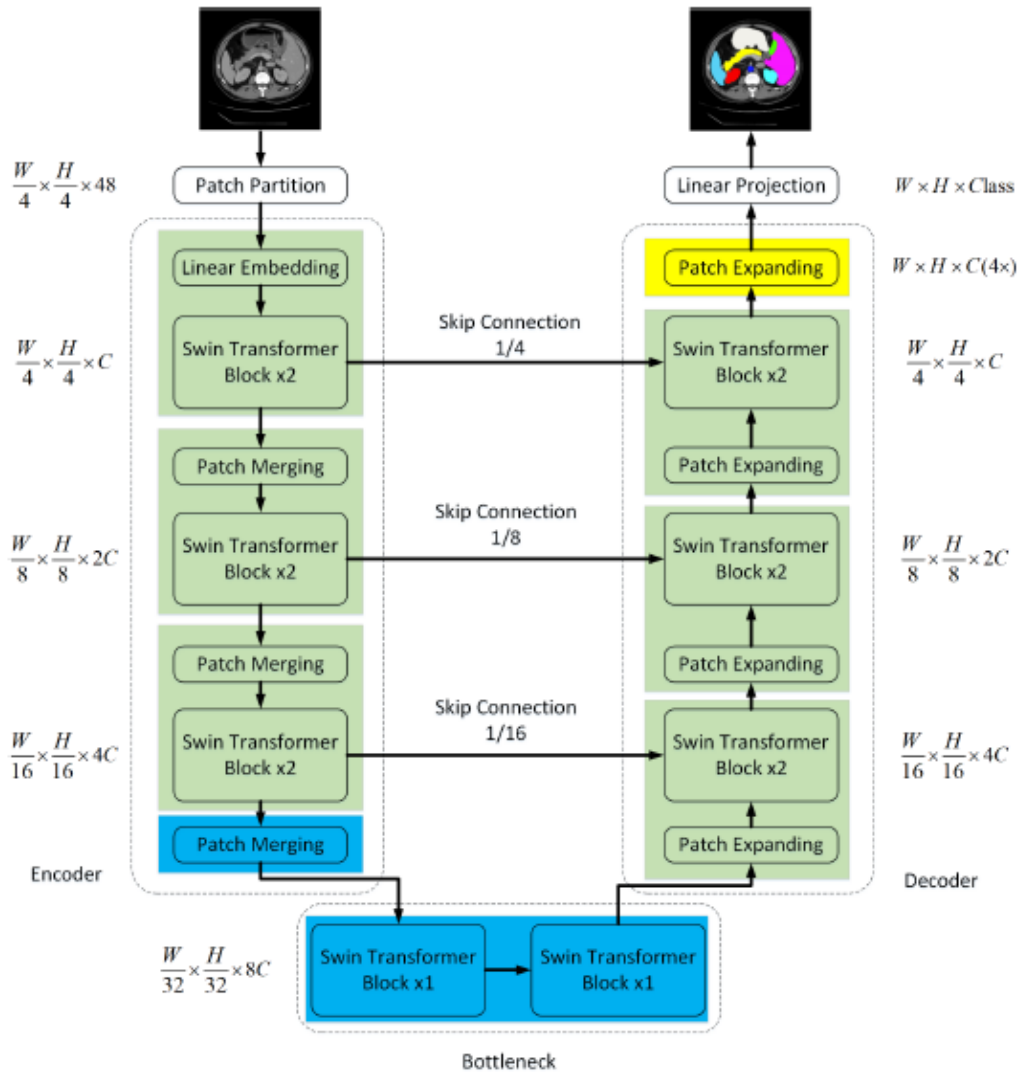
Στο πρόβλημα αναγνώρισης της πληροφορίας που περιέχει μια εικόνα οι Vision Transformers έχουν κατασκευαστεί ώστε να επιτυγχάνουν τέτοιου είδους εργασίες. Παίρνοντας 2D image patches με ενσωματωμένες θέσεις για εισαγόμενη πληροφορία και εκπαιδεύοντας το μοντέλο με μεγάλο αριθμό δεδομένων, οι Vision Transformers έχουν καταφέρει να αποδίδουν παρόμοια αποτελέσματα με τα CNN-based μοντέλα. Αντίστοιχα, έχει κατασκευαστεί ο Swin Transformer για την επίλυση του προβλήματος αναγνώρισης της πληροφορίας μιας εικόνας (vision domain). Ο Swin Transformer και γενικότερα οι Vision Transformers ενδείκνυνται για τα προβλήματα image recognition και πιο συγκεκριμένα για το πρόβλημα της κατάτμησης μιας εικόνας (segmentation).

Με αφορμή τον μετασχηματισμό Swin Transformer θα αναλυθεί το μοντέλο «Swin UNet» το οποίο αναπτύχθηκε με στόχο την κατάτμηση 2D εικόνων (image segmentation). Το μοντέλο «Swin UNet» είναι μια Transformer-based αρχιτεκτονική σε σχήμα U η οποία

αποτελείται από encoder, decoder, bottleneck και skip connections. Οι encoder, decoder και bottleneck έχουν κατασκευαστεί βασιζόμενοι στο Swin Transformer block.

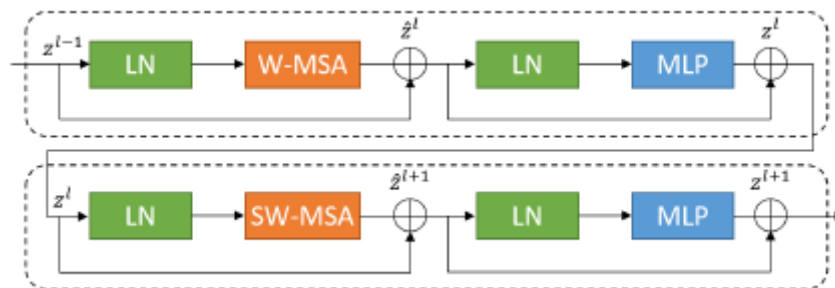
Αρχικά, οι εισαγόμενες στο ανωτέρω περιγραφόμενο μοντέλο εικόνες χωρίζονται σε μη-επικαλυπτόμενα image patches. Κάθε image patch θεωρείται από το μοντέλο ως εισαγόμενη πληροφορία για τον Transformer-based encoder με στόχο να «μάθει» να αναγνωρίζει βασικά χαρακτηριστικά και πληροφορία των απεικονίσεων – δεδομένων. Αυτά τα χαρακτηριστικά που εξάγονται από τον encoder στη συνέχεια ακολουθούν την διαδικασία του up-sampling από τον decoder μέσω των patch expanding layers και έπειτα συγχωνεύονται με τα χαρακτηριστικά που έχουν εξαχθεί από τον encoder μέσω των skip connections, ώστε να αποκατασταθεί οι χωρική πληροφορία που χάθηκε κατά την διαδικασία του down-sampling του encoder και τέλος να δημιουργηθεί η segmentation απεικόνιση, δηλαδή η απεικόνιση κατά την οποία έχει επιτευχθεί κατάτμηση.

Κατωτέρω στην *εικόνα 6* εμφανίζεται η δομή σε σχήμα U του μοντέλου Swin – UNet. Βασική μονάδα του μοντέλου Swin – UNet αποτελεί το Swin Transformer block. Στο κομμάτι του encoder προκειμένου να μετατραπούν οι εισαγόμενες εικόνες σε μια ενσωματωμένη ακολουθία, χωρίζονται σε μη επικαλυπτόμενα patches μεγέθους 4×4 . Με αυτή την προσέγγιση διαχωρισμού το κάθε patch αποκτά διάσταση $4 \times 4 \times 3 = 48$. Στη συνέχεια εφαρμόζεται μία γραμμική μέθοδος (linear embedding layer), ώστε να προβληθούν χαρακτηριστικά διαστάσεων σε μια αυθαίρετη διάσταση (αναγράφεται ως C). Τα Transformed patches εισάγονται σε πολλαπλά Swin Transformer blocks και patch merging layers, ώστε να αναγνωριστούν όλα αυτά τα χαρακτηριστικά των εικόνων τα οποία πρέπει στο τέλος της διαδικασίας να ανιχνευθούν και ταξινομηθούν με στόχο την κατάτμηση της εικόνας. Πιο συγκεκριμένα, τα patch merging layers είναι υπεύθυνα για την διαδικασία του down-sampling και τα Swin Transformer blocks αντίστοιχα είναι υπεύθυνα για την εκμάθηση των απαραίτητων χαρακτηριστικών που περιλαμβάνονται στις εικόνες. Η αρχιτεκτονική του μοντέλου «Swin – UNet» εμπνευσμένη από την συμμετρική δομή UNet αποτελείται από τον συμμετρικό Transformer-based decoder. Ο decoder ακριβώς όπως και ο encoder αποτελείται Swin Transformer blocks και patch expanding layers. Τα patch expanding layers επιτυγχάνουν την διαδικασία του up-sampling. Δηλαδή, αλλάζουν σχήμα στις απεικονίσεις με τις προσαρμοσμένες διαστάσεις σε μια μεγάλη απεικόνιση με $2 \times$ up-sampling της ανάλυσης. Στο τέλος, το τελευταίο patch expanding layer χρησιμοποιείται ώστε να επιτυγχθεί το $4 \times$ up-sampling κατά το οποίο θα αποκατασταθεί και η ανάλυση των εικόνων στην αρχική ανάλυση των εικόνων που αρχικά εισάγονται στο μοντέλο ($W \times H$) και στη συνέχεια ένας γραμμικός μετασχηματισμός εφαρμόζεται στις εικόνες ώστε να προκύψει η segmentation απεικόνιση.



Εικόνα 6 - Η αρχιτεκτονική του μοντέλου "Swin - UNet" για σημασιολογική κατάτμηση, η οποία απαρτίζεται από encoder - decoder, bottleneck & skip connections. Encoder, decoder & bottleneck έχουν κατασκευαστεί βασιζόμενα στο swin transformer block

Swin Transformer block: Το Swin Transformer block έχει κατασκευαστεί βασιζόμενο στα shifted windows. Κάθε Swin Transformer block αποτελείται από LayerNorm (LN) layer, multi-head self attention module, residual connection and 2-layer MLP with GELU non-linearity.



Εικόνα 7 - Swin Transformer block

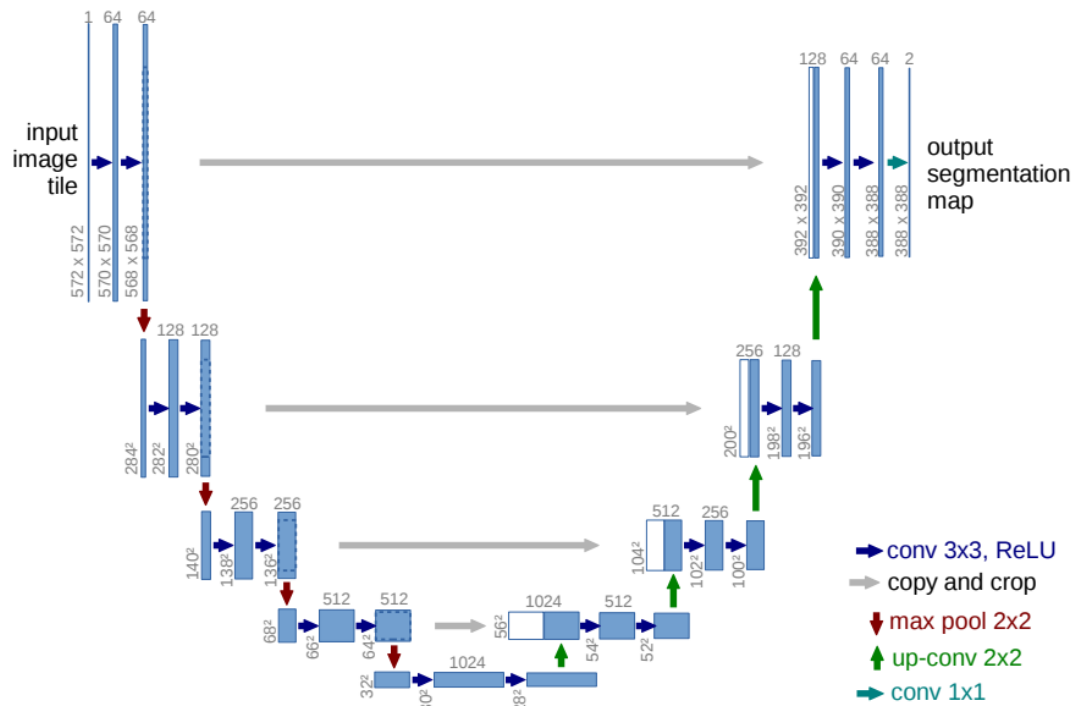
2.5 Το μοντέλο U-Net

Το μοντέλο U-Net αποτελεί αρχιτεκτονική των συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks (CNN)) και χρησιμοποιείται κυρίως σε προβλήματα κατάτμησης εικόνων. Η αρχιτεκτονική U-Net έχει αναπτυχθεί κυρίως για την αντιμετώπιση προβλημάτων σημασιολογικής κατάτμησης (semantic segmentation). Το μοντέλο U-Net αποτελεί κατά βάση ένα Πλήρως Συνελκτικό Νευρωνικό Δίκτυο (Convolutional Neural Network (CNN)) το οποίο έχει σχεδιαστεί να μαθαίνει από ένα σύνολο δεδομένων (δεδομένα εκπαίδευσης). Η κεντρική ιδέα του συγκεκριμένου δικτύου παραμένει ίδια με αυτή των (Convolutional Neural Network (CNN)), δηλαδή, η εικόνα εισόδου ακολουθεί ένα μονοπάτι συστολής (contracting path), όπου η εικόνα συμπιέζεται σε μικρότερες διαστάσεις κι ένα μονοπάτι διαστολής (expansive path), όπου η εικόνα ανοικοδομείται στις αρχικές της. Η διαφορά του δικτύου U-Net έγκειται στις συνδέσεις παράκαμψης (skip connections). Αυτές εφαρμόζονται στα σημεία όπου έχουμε ίδιες διαστάσεις κι επιτρέπουν στο κομμάτι της υπερδειγματοληψίας τη διατήρηση του μεγάλου αριθμού πληροφοριών περιβάλλοντος, ώστε το δίκτυο να μπορεί να διαδώσει της πληροφορίες περιεχομένου σε επίπεδα υψηλότερης ανάλυσης. Αυτό έχει ως απόρροια, να δημιουργείται μια συμμετρία μεταξύ των μονοπατιών συστολής και διαστολής, δημιουργώντας την αίσθηση ενός σχήματος που θυμίζει το αγγλικό γράμμα «U», εξ' ου κι η ονομασία «U-Net». Το δίκτυο δεν έχει πλήρως συνδεδεμένα επίπεδα και χρησιμοποιεί μόνο το έγκυρο κομμάτι κάθε συνέλιξης, δηλαδή η μάσκα κατάτμησης περιέχει μόνο τα εικονοστοιχεία που αντιστοιχούν στην εικόνα εισόδου.

Αυτή η διαδικασία μαζί με την πρόβλεψη εικονοστοιχείων στα σύνορα της εικόνας μέσω εξαγωγής πληροφοριών περιεχομένου μέσω κατοπτρισμού της εικόνας εισόδου, επιτρέπουν τη συνεχόμενη κατάτμηση αυθαίρετα μεγάλων εικόνων.

Πιο αναλυτικά, η αρχιτεκτονική του δικτύου αποτελείται από ένα μονοπάτι συστολής στα αριστερά κι ένα κομμάτι διαστολής στα δεξιά, όπως φαίνεται και στην παρακάτω *Εικόνα 12* το μονοπάτι της συστολής ακολουθεί την τυπική αρχιτεκτονική των συνελκτικών δικτύων. Απαρτίζεται από επαναλαμβανόμενες εφαρμογές διπλών Συνελίξεων 3×3 , τις οποίες ακολουθεί η συνάρτηση της Διορθωμένης Γραμμικής Μονάδας κι η εφαρμογή Μέγιστης Συγκέντρωσης με βήμα 2 για την επίτευξη της υποδειγματοληψίας (*downsampling*). Σε κάθε βήμα υποδειγματοληψίας διπλασιάζουμε τον αριθμό των συμμετεχόντων καναλιών. Κάθε βήμα στο μονοπάτι της διαστολής αποτελείται από την υπερδειγματοληψία των χαρτογραφημένων χαρακτηριστικών, η οποία ακολουθείται από μία Συνέλιξη 2×2 που μειώνει στο μισό τον αριθμό των συμμετεχόντων καναλιών, μια συνένωση με τα αντίστοιχα χαρτογραφημένα χαρακτηριστικά που έχουν αποκοπεί από το μονοπάτι της συστολής και δύο Συνελίξεις 3×3 , που τις ακολουθεί μία Συνάρτηση Διορθωμένης Γραμμικής Μονάδας. Στο τελικό επίπεδο, μία Συνέλιξη 1×1 χρησιμοποιείται για να αντιστοιχήσει κάθε διάνυσμα στο επιθυμητό αριθμό κλάσεων. Το δίκτυο αποτελείται από 23 συνελκτικά επίπεδα στο σύνολό του.

Τα μειονεκτήματα του δικτύου είναι ότι χάνεται μεγάλος όγκος χωρικής πληροφορίας λόγω των επιπέδων μέγιστης συγκέντρωσης και δεν έχει την ικανότητα να συσχετίσει το περιεχόμενο με τις συντεταγμένες του.



Εικόνα 8 - UNet αρχιτεκτονική (παράδειγμα 32x32 pixels στη χαμηλότερη ανάλυση). Κάθε μπλε πλαίσιο αντιστοιχεί σε μια απεικόνιση πολλαπλών κλάσεων. Ο αριθμός των κλάσεων σημειώνεται στην κορυφή του πλαισίου. Τα βέλη υποδηλώνουν τις διάφορες λειτουργίες

Κεφάλαιο 3 – Φαινόμενο Αποψίλωσης και Μετρικές

3.1 Deforestation – Φαινόμενο «Αποψίλωσης» ή «Αποδάσωσης»

3.1.1 Εισαγωγή

Στην παρούσα διπλωματική εργασία χρησιμοποιήθηκαν δορυφορικά δεδομένα με δασική και μη πληροφορία προκειμένου να μελετηθεί η εξάπλωση του φαινομένου της αποψίλωσης (deforestation) των δασών.

Η μελέτη του φαινομένου της «αποψίλωσης» (deforestation) πραγματοποιήθηκε συναρτήσει μοντέλων μηχανικής μάθησης (machine learning) που αποσκοπούν στην σημασιολογική κατάτμηση (semantic segmentation), δηλαδή στην κατηγοριοποίηση κάθε εικονοστοιχείου (pixel) μίας εικόνας. Συγκεκριμένα, στο πειραματικό μέρος της διπλωματικής εργασίας εκπαιδεύτηκαν δύο (2) μοντέλα μηχανικής μάθησης, δηλαδή νευρωνικά δίκτυα (artificial neural networks), με εικόνες που εμπεριέχουν δασική και μη πληροφορία. Στόχος αποτέλεσε ο εντοπισμός, με την βέλτιστη δυνατή ακρίβεια, εκτάσεων δασικής γης που έχουν καταστραφεί («αποψίλωση» ή «αποδάσωση») μέσω της μεθόδου της σημασιολογικής κατάτμησης δορυφορικών εικόνων.

3.1.2 Ανάλυση του φαινομένου της «αποψίλωσης» δασών (deforestation)

Τα δάση επιτελούν πολλαπλές και αλληλένδετες κοινωνικές, οικονομικές και περιβαλλοντικές λειτουργίες και συγκαταλέγεται μεταξύ των οικοσυστημάτων με τη μεγαλύτερη βιοποικιλότητα. Τα δάση είναι πολλαπλά και ανανεώσιμα και, ως εκ τούτου, αποτελούν ένα πολυτιμότερο φυσικό πόρο.

Με τον όρο αποδάσωση ή αποψίλωση αποδίδεται η, σε πολύ μεγάλη κλίμακα, καταστροφή ενός δάσους όπου τα εδάφη χρησιμοποιούνται στη συνέχεια για άλλες χρήσεις, ενώ η υποβάθμιση των δασών είναι μια πιο σταδιακή διαδικασία που σχετίζεται με την απώλεια της ικανότητας των δασών να παράγουν ουσιαστικά οφέλη, όπως το ξύλο ή η βιοποικιλότητα. Η αποψίλωση αναφέρεται στη μόνιμη απομάκρυνση των δασών και των συστάδων δέντρων για να καταστεί η γη διαθέσιμη για άλλες χρήσεις. Πρόκειται για ένα πολύπλοκο ζήτημα που περιλαμβάνει πλήθος περιβαλλοντικών, κοινωνικών και οικονομικών πτυχών.

Η αποψίλωση και η υποβάθμιση των δασών οφείλονται κυρίως σε ανθρώπινες δραστηριότητες και επηρεάζουν τους ανθρώπους σε όλο τον κόσμο. Κύριο αίτιο που οδηγεί στο παραπάνω φαινόμενο είναι η εκτεταμένη εμπορική γεωργία και, κυρίως, η παραγωγή φοινικέλαιου, σόγιας, ξυλείας και βοοειδών προϊόντων, μιας και αυτά αποτελούν τα τέσσερα εμπορεύματα που παράγονται σε μεγαλύτερη κλίμακα από την αλυσίδα παραγωγής των επιχειρήσεων σε διεθνές επίπεδο. Η μετατροπή των δασών σε γεωργική γη για τη διατροφή του αυξανόμενου αριθμού ανθρώπων παγκοσμίως είναι η κύρια αιτία της αποψίλωσης των δασών. Τα δάση αποψιλώνονται για να φιλοξενήσουν επεκτεινόμενες γεωργικές εκτάσεις, φυτείες και βοσκοτόπια. Επιπλέον, η καθαίρεση δασών αποτελεί σημαντική αιτία του φαινομένου της αποψίλωσης. Τα δέντρα υλοτομούνται για ξυλεία και χαρτοπολλά, τα οποία χρησιμοποιούνται σε διάφορες εμπορικές εφαρμογές, συμπεριλαμβανομένων των κατασκευών, της παραγωγής χαρτιού και άλλων. Φυσικά, η αστική ανάπτυξη και η ανάπτυξη των υποδομών, συμπεριλαμβανομένων των κατασκευών και της επέκτασης των δρόμων, ως μέσο προώθησης της αστικοποίησης, συμβάλει και αυτή με τη σειρά της στην όξυνση του φαινομένου. Στην Ευρώπη πρόκειται για τη κύρια αιτία αποψίλωσης των δασών.

Τέλος, άλλες επιζήμιες δραστηριότητες που σχετίζονται με ανθρώπινες δραστηριότητες και οδηγούν στο παραπάνω φαινόμενο είναι η υπερεκμετάλλευση των δασικών πόρων, μεταξύ άλλων για καύσιμα και την παράνομη ή μη βιώσιμη υλοτομία και η ανάπτυξη της

εξορυκτικής βιομηχανίας, η οποία απαιτεί την αποψίλωση μεγάλων δασικών εκτάσεων, συμβάλλει στην αποψίλωση των δασών, όπως αντίστοιχα και η κλιματική αλλαγή. Η κλιματική αλλαγή είναι τόσο αιτία όσο και συνέπεια της αποψίλωσης και της υποβάθμισης των δασών. Οι μεταβολές στα κλιματικά πρότυπα μπορούν να δημιουργήσουν συνθήκες που επιδεινώνουν την απώλεια δασών, όπως η αυξημένη συχνότητα και ένταση των πυρκαγιών. Τα ακραία φαινόμενα που προκαλεί η κλιματική αλλαγή, όπως πυρκαγιές, ξηρασίες και πλημμύρες, επηρεάζουν τα δάση. Με τη σειρά της, η απώλεια δασών είναι επιβλαβής για το κλίμα, καθώς τα δάση παίζουν σημαντικό ρόλο στην παροχή καθαρού αέρα, στη ρύθμιση του κύκλου του νερού, στη δέσμευση CO₂, στην πρόληψη της απώλειας βιοποικιλότητας και της διάβρωσης του εδάφους.

Ιδιαίτερες σημαντικές είναι οι περιβαλλοντικές περιπτώσεις που προκαλεί η επέκταση του φαινομένου της αποψίλωσης των δασών. Μερικές από αυτές παρουσιάζονται παρακάτω:

- **Απώλεια βιοποικιλότητας:** Τα δάση είναι πλούσιοι βιότοποι γεμάτοι από ποικίλα είδη. Η αποψίλωση των δασών μπορεί να οδηγήσει στην εξαφάνιση ειδών που βασίζονται σε αυτά τα ενδιαιτήματα.
- **Κλιματική αλλαγή:** Τα δέντρα διαδραματίζουν κρίσιμο ρόλο στην απορρόφηση των αερίων του θερμοκηπίου, ιδίως του διοξειδίου του άνθρακα. Όταν κόβονται, όχι μόνο χάνεται αυτή η ευεργετική λειτουργία, αλλά και ο αποθηκευμένος άνθρακας απελευθερώνεται πίσω στην ατμόσφαιρα, συμβάλλοντας στο φαινόμενο του θερμοκηπίου και στην κλιματική αλλαγή.
- **Διακοπή των κύκλων του νερού:** Τα δέντρα συμβάλλουν στον κύκλο του νερού επιστρέφοντας τους υδρατμούς πίσω στην ατμόσφαιρα. Χωρίς αυτά τα δέντρα, τα κλίματα μπορεί να γίνουν ξηρότερα.
- **Διάβρωση του εδάφους:** Οι ρίζες των δέντρων αγκυρώνουν το έδαφος. Όταν τα δέντρα κόβονται, το έδαφος είναι πιο επιρρεπές στη διάβρωση, η οποία μπορεί να οδηγήσει σε απώλεια γονιμότητας και αυξημένη ιζηματογένεση σε ποτάμια και ρυάκια.

Πρόληψη και λύσεις του φαινομένου αποψίλωσης:

- **Αειφόρος δασοκομία:** Η πρακτική αυτή περιλαμβάνει την υπεύθυνη διαχείριση των δασών ώστε να διασφαλίζεται ότι παρέχουν τους πόρους τους επ' αόριστον.
- **Αναδάσωση και αναδάσωση:** Η φύτευση δέντρων σε περιοχές όπου έχουν κοπεί (αναδάσωση) ή σε περιοχές όπου δεν υπήρχαν δέντρα πριν (αναδάσωση) μπορεί να βοηθήσει στην αποκατάσταση των οικοσυστημάτων.
- **Προστατευόμενες περιοχές:** Δημιουργία προστατευόμενων περιοχών για τη διατήρηση της βιοποικιλότητας και των δασικών τοπίων.
- **Οικονομικά κίνητρα:** Παροχή οικονομικών κινήτρων για τη διατήρηση των δασών, όπως μέσω πιστωτικών μορίων άνθρακα.
- **Νόμοι και κανονισμοί:** Εφαρμογή και επιβολή νόμων για την προστασία των δασών και τη ρύθμιση της χρήσης γης.

Η αποψίλωση των δασών είναι ένα κρίσιμο παγκόσμιο ζήτημα που απαιτεί διεθνή συνεργασία, βιώσιμες πρακτικές και προσπάθειες διατήρησης για την αποτελεσματική αντιμετώπισή του. Συνεπώς, είναι άκρως σημαντικό να κατασκευαστούν με την βοήθεια της βαθιάς μάθησης και των νευρωνικών δικτύων αυτοματισμοί ώστε εύκολα και γρήγορα να ανιχνεύονται μεγάλες εκτάσεις αποψιλωμένης περιοχής, να καταγράφεται το πρόβλημα και να ερευνώνται άμεσα λύσεις και τρόποι αντιμετώπισής του.

3.2 Απόδοση Νευρωνικών Δικτύων – Μετρικές (Metrics)

Στην παρούσα διπλωματική εργασία χρησιμοποιούνται διάφορες Μετρικές (Metrics). Μια μετρική είναι μια συνάρτηση που χρησιμοποιείται για να κρίνεται η απόδοση του μοντέλου. Με δεδομένες τις προβλεπόμενες τιμές και τις πραγματικές (ground truth) τιμές

(βλ. ενότητα 4.1) η συνάρτηση αυτή δίνει ένα βαθμωτό μέτρο της καταλληλότητας του μοντέλου, με τα υπάρχοντα δεδομένα. Τα αποτελέσματα μιας μετρικής συνάρτησης που προκύπτουν από την αξιολόγηση μιας μέτρησης δεν χρησιμοποιούνται κατά την εκπαίδευση του μοντέλου.

3.2.1 Αρμονικός Μέσος - F1 – Score

Κατά την εφαρμογή ενός μοντέλου κρίνεται απαραίτητη η αξιολόγηση της πρόβλεψης του, δηλαδή το ποσοστό ακρίβειας της πρόβλεψης που παράγει το μοντέλο σε σχέση με τα πραγματικά δεδομένα. Προκειμένου να επιτευχθεί αυτή η αξιολόγηση χρησιμοποιούνται συγκεκριμένες **μετρικές** όπως,

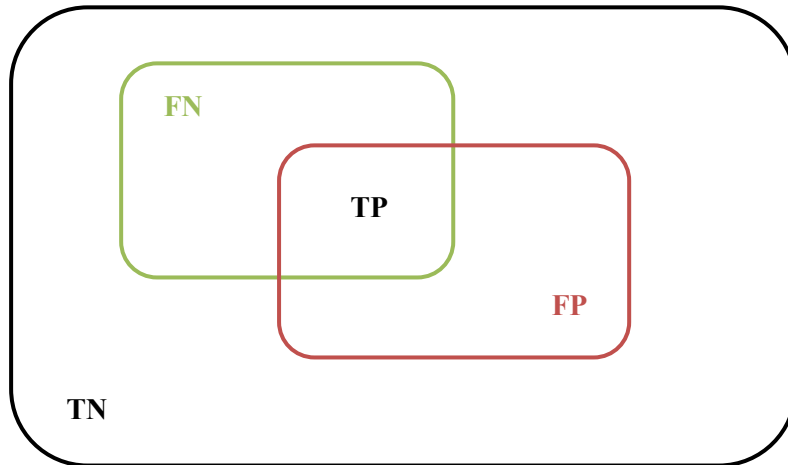
- **Precision:** Δηλώνει την ακρίβεια ή πιστότητα των αποτελεσμάτων στο ερώτημα που τέθηκε. Συγκεκριμένα, δείχνει το ποσοστό των επιτυχιών προς των συνολικών προβλέψεων μιας κλάσης από ένα μοντέλο. Το precision μπορεί να θεωρηθεί ως μέτρο ποιότητας.
- **Recall:** Δηλώνει μέτρο πληρότητας ή ποσότητας. Η ανάκληση (γνωστή και ως ευαισθησία – sensitivity) είναι μια παραδοσιακή μέτρηση αξιολόγησης, η οποία όπως όλα τα μέτρα, υποθέτουν μια έννοια βασικής αλήθειας (ground truth). Είναι το κλάσμα των επιτυχιών προβλέψεων μιας κλάσης προς τα συνολικά παραδείγματα αυτής. Η ανάκληση μπορεί να θεωρηθεί ως μέτρο πληρότητας ή ποσότητας.
- **F – Measure (ή F1 – Score):** Σταθμισμένος Αρμονικός Μέσος

Πίνακας Κατηγοριοποίησης Αποτελεσμάτων

		Correct Result / Classification	
		E1	E2
Obtained Result / Classification	E1	TP (true positive)	FP (false positive)
	E2	FN (false negative)	TN (true negative)

Επεξήγηση Πίνακα

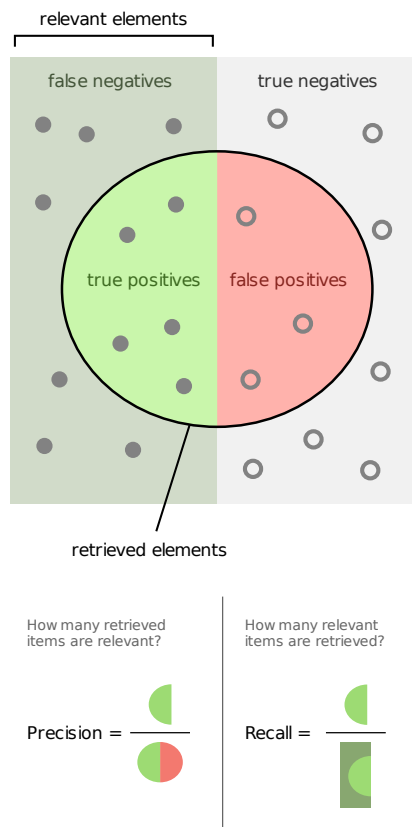
- TP (true positive): αφορά τα αποτελέσματα της πρόβλεψης που επιστράφηκαν και είναι σχετικά με μια συγκεκριμένη κλάση. Δηλαδή, αληθώς θετική ονομάζεται μια πρόβλεψη που εκτιμήθηκε ότι ανήκει σε μια συγκεκριμένη κλάση και αυτό ισχύει.
- FP (false positive): αφορά τα αποτελέσματα της πρόβλεψης που επιστράφηκαν και είναι μη σχετικά με μια συγκεκριμένη κλάση. Δηλαδή, ψευδώς θετική ονομάζεται μια πρόβλεψη που εκτιμήθηκε ότι ανήκει σε μια συγκεκριμένη κλάση και αυτό δεν ισχύει.
- FN (false negative): αφορά τα αποτελέσματα της πρόβλεψης που δεν επιστράφηκαν, είναι σχετικά με μια συγκεκριμένη κλάση και τα οποία δεν βρέθηκαν. Δηλαδή, ψευδώς αρνητική ονομάζεται μια πρόβλεψη που εκτιμήθηκε ότι *δεν* ανήκει σε μια συγκεκριμένα και αυτό *δεν* ισχύει.
- TN (true negative): αφορά τα αποτελέσματα της πρόβλεψης που δεν επιστράφηκαν και είναι σχετικά με μια συγκεκριμένη κλάση. Δηλαδή, αληθώς αρνητική ονομάζεται μια πρόβλεψη που εκτιμήθηκε ότι *δεν* ανήκει σε μια κλάση και αυτό ισχύει.



Σχήμα 2.2: FP-FN-TP-TN, Η κλάση είναι το πράσινο σχήμα και το TP είναι στην τομή του πράσινου με το κόκκινο σχήμα, οι προβλέψεις που εκτιμήθηκαν ότι ανήκουν στην κλάση και ισχύει. Στο πράσινο σχήμα το FN είναι οι προβλέψεις που εκτιμήθηκαν ότι δεν ανήκουν στην κλάση και αυτό δεν ισχύει. Στο κόκκινο σχήμα το FP είναι οι προβλέψεις που εκτιμήθηκαν ότι ανήκουν στην κλάση αλλά αυτό δεν ισχύει. Τέλος, έξω από όλα τα εσωτερικά σχήματα και μέσα στο μαύρο σχήμα είναι το TN που είναι οι προβλέψεις που εκτιμήθηκαν ότι δεν ανήκουν στην κλάση και ισχύει.

Υπολογισμοί

- $Precision = TP / (TP + FP)$
- $Recall = TP / (TP + FN)$
- $F1 - Measure = 2 * [(Precision * Recall) / (Precision + Recall)]$



Εικόνα 9 - Μετρικές precision & recall

F1 – Score

Η μετρική F1 – Score χρησιμοποιείται στο κομμάτι της αξιολόγησης και συνδυάζει την ακρίβεια (Precision) και την ανάκληση (recall). Συγκεκριμένα είναι ο αρμονικός μέσος όρος της ακρίβειας και της ανάκλησης. Η F1 – Score φτάνει την καλύτερη τιμή της στο 1 (τέλεια ακρίβεια και ανάκληση) και χειρότερη στο 0.

Η μετρική F1-score χρησιμοποιείται ευρέως σε δοκιμές δυαδικής ταξινόμησης όπου η θετική κλάση παρουσιάζει ιδιαίτερο ενδιαφέρον. Ωστόσο, σε προβλήματα πολλαπλών κλάσεων, μπορεί να προσαρμοστεί μέσω μεθόδων όπως η "micro" (συνολική με άθροιση των μεμονωμένων αληθώς θετικών, ψευδώς αρνητικών κ.λπ.), η "macro" (μέσος όρος F1-score ανά κλάση) και η "σταθμισμένη" (λαμβάνοντας υπόψη την ανισορροπία στο σύνολο δεδομένων) F1-score.

Ενώ το F1-score είναι μια ισχυρή μετρική, δεν είναι χωρίς περιορισμούς. Υποθέτει ίση σημασία της ακρίβειας και της ανάκλησης, κάτι που μπορεί να μην ισχύει σε όλα τα σενάρια. Επιπλέον, σε σενάρια πολλαπλών κλάσεων, ο τρόπος υπολογισμού του ενδέχεται να αποκρύπτει προβλήματα απόδοσης σε κλάσεις με λιγότερα δείγματα (πρόβλημα σε μη ισορροπημένα σύνολα δεδομένων). Συνοψίζοντας, το F1-score είναι μια ουσιαστική μετρική για εργασίες ταξινόμησης, ειδικά σε προβλήματα όπου η απλή ακρίβεια δεν αντικατοπτρίζει πλήρως την απόδοση ενός μοντέλου λόγω ανισοβαρών κλάσεων. Παρέχει μια ισορροπία μεταξύ ακρίβειας και ανάκλησης σε μια ενιαία μετρική. Ωστόσο, όπως κάθε μετρική, θα πρέπει να χρησιμοποιείται λαμβάνοντας υπόψη το συγκεκριμένο πλαίσιο και τις απαιτήσεις της εκάστοτε εργασίας.

Συνοπτικά, η αριθμητική παράμετρος F1 – Score υπολογίζει όλα τα παραπάνω και αξιολογεί το ποσοστό αποτελεσματικότητας της πρόβλεψης του εκάστοτε μοντέλου μηχανικής μάθησης. Όσο η F1 – Score τείνει στο ένα (σε ποσοστό 100%) τόσο πιο αποτελεσματική είναι η πρόβλεψη του μοντέλου. Η μετρική F1-score είναι μια κρίσιμη μετρική στη μηχανική μάθηση που χρησιμοποιείται για τη μέτρηση της ακρίβειας σε εργασίες ταξινόμησης, εξισορροπώντας την ακρίβεια και την ανάκληση. Συνδυάζει αυτές τις δύο μετρικές μέσω του αρμονικού μέσου, με αποτέλεσμα ένα σκορ που αντικατοπτρίζει τόσο την αξιοπιστία του μοντέλου (ακρίβεια) όσο και την πληρότητά του (ανάκληση), ιδιαίτερα σημαντικό σε καταστάσεις με ανισορροπες κλάσεις ή διαφορετική σπουδαιότητα κλάσεων. Ωστόσο, υποθέτει ίση βαρύτητα για την ακρίβεια και την ανάκληση, η οποία μπορεί να μην είναι πάντα κατάλληλη, και μπορεί να αποκρύψει διαφορές επιδόσεων σε σενάρια πολλαπλών κλάσεων. Η βαθμολογία F1 είναι ιδιαίτερα πολύτιμη επειδή παρέχει μια ενιαία, εύκολη στην ερμηνεία μετρική που περιλαμβάνει την ικανότητα του μοντέλου να αποφεύγει τις λανθασμένες ταξινομήσεις, διατηρώντας παράλληλα την ευαισθησία στις πραγματικές περιπτώσεις.

3.2.2 Intersection Over Union (IOU) - Jaccard

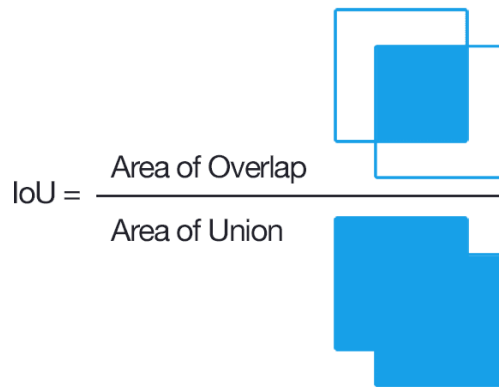
Η μετρική Intersection over Union (IoU) είναι ένα στατιστικό στοιχείο που χρησιμοποιείται για τη σύγκριση της ομοιότητας και της ποικιλομορφίας των συνόλων δειγμάτων. Ο συγκεκριμένος συντελεστής Jaccard μετρά την ομοιότητα μεταξύ συνόλων πεπερασμένων δειγμάτων και ορίζεται ως το μέγεθος της τομής διαιρούμενο με το μέγεθος της ένωσης των συνόλων δειγμάτων. Το μέγεθος της τομής που αντιστοιχεί στο True Positive ονομάζεται και intersection.

Η μετρική Intersection over Union IoU μετρά την επικάλυψη μεταξύ δύο ορίων. Για τα νευρωνικά δίκτυα, ιδίως για εκείνα που εμπλέκονται σε εργασίες όπως η ανίχνευση αντικειμένων ή η κατάμηση, είναι ζωτικής σημασίας να κατανοήσουμε πόσο καλά το μοντέλο όχι μόνο αναγνωρίζει τα αντικείμενα μέσα σε μια εικόνα, αλλά και πόσο ακριβώς

εντοπίζει τις θέσεις τους. Το IoU παρέχει μια αυστηρή μετρική για το σκοπό αυτό, λαμβάνοντας υπόψη τόσο την περιοχή επικάλυψης όσο και την περιοχή ένωσης μεταξύ του προβλεπόμενου ορίου και του πραγματικού, πραγματικού ορίου. Πρόκειται για μια πιο ισχυρή μετρική σε σύγκριση με απλούστερες μετρήσεις όπως η ακρίβεια ή η ανάκληση, καθώς λαμβάνει υπόψη τόσο το μέγεθος όσο και τη θέση της επικάλυψης.

Η διατομή πάνω από την ένωση (IoU) ή ο δείκτης Jaccard, ορίζεται ως η περιοχή τομής μεταξύ του προβλεπόμενου χάρτη τμηματοποίησης A και του χάρτη B, διαιρούμενο με την περιοχή της ένωσης μεταξύ των δύο χαρτών, και κυμαίνεται μεταξύ 0 και 1:

$$IoU = J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

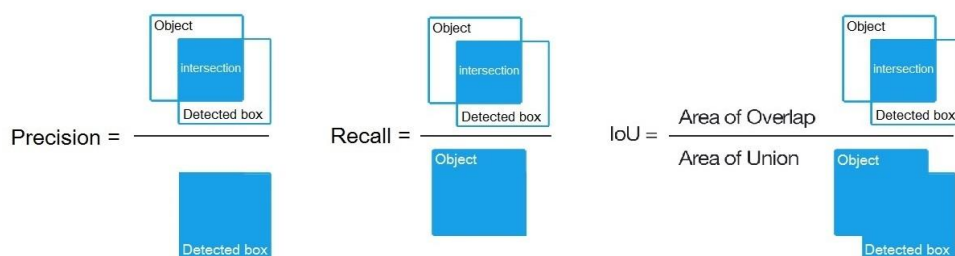


Εικόνα 10 - Μετρική Intersection over Union (IoU)



Εικόνα 11 - Intersection over Union

Από την παραπάνω εικόνα 11 προκύπτει το συμπέρασμα ότι η αριθμητική παράμετρος Intersection over Union (IoU) αποτελεί μία αναλογία. Στον αριθμητή υπολογίζεται η **area of overlap** μεταξύ της πραγματικής τιμής (Ground Truth – μάσκα) και της προβλεπόμενης από το μοντέλο τιμής (model prediction). Στον παρονομαστή υπολογίζεται η **area of union**, που αφορά το σύνολο των πραγματικών τιμών (Ground Truth) και αντίστοιχα των προβλεπόμενων από το μοντέλο τιμών (model prediction).



Εικόνα 12 - Precision, Recall & IoU

Προκειμένου, λοιπόν, να αξιολογηθεί η απόδοση του μοντέλου και η ποιότητα των αποτελεσμάτων του, δηλαδή οι πραγματικές τιμές (Ground Truth) σε σχέση με τις προβλεπόμενες από το μοντέλο τιμές (model prediction) είναι απαραίτητος ο υπολογισμός της αριθμητικής παραμέτρου Jaccard - Intersection over Union (IoU). Όσο περισσότερο τείνουν οι προβλεπόμενες τιμές στις πραγματικές η παράμετρος Intersection over Union (IoU) τείνει στο 1 (σε ποσοστό 100%) όπου είναι και το επιθυμητό.

Εφαρμογές στα νευρωνικά δίκτυα:

- Ανίχνευση αντικειμένων: Σε εργασίες όπως η ανίχνευση αντικειμένων μέσα σε εικόνες, η IoU βοηθά στην αξιολόγηση της ακρίβειας των οριοθετημένων πλαισίων που προβλέπει το μοντέλο σε σχέση με την βασική αλήθεια.
- Σημασιολογική κατάτμηση: Για εργασίες ταξινόμησης ανά εικονοστοιχείο, η IoU μπορεί να καθορίσει πόσο καλά το νευρωνικό δίκτυο ταξινομεί σωστά κάθε εικονοστοιχείο, τμηματοποιώντας έτσι το αντικείμενο από το φόντο του.
- Τμηματοποίηση στιγμιότυπων: Παρόμοια με τη σημασιολογική τμηματοποίηση, αλλά εδώ, η IoU είναι ζωτικής σημασίας για τη διάκριση και την αξιολόγηση μεμονωμένων περιπτώσεων αντικειμένων της ίδιας κλάσης.

Προκλήσεις και προβληματισμοί:

- Κατώφλι: Συχνά, τίθεται ένα κατώφλι (π.χ. IoU 0,5) για τον προσδιορισμό των θετικών ανιχνεύσεων. Ανάλογα με την εφαρμογή, αυτό το κατώφλι μπορεί να προσαρμοστεί, αλλά είναι σημαντικό να διατηρηθεί μια ισορροπία για την αποφυγή ψευδώς θετικών ή αρνητικών αποτελεσμάτων.
- Χειρισμός πολλαπλών αντικειμένων: Σε σενάρια με πολλαπλά αντικείμενα και επικαλυπτόμενα πλαίσια οριοθέτησης, ο χειρισμός των υπολογισμών IoU μπορεί να γίνει πολύπλοκος, απαιτώντας συχνά πρόσθετη λογική όπως η μη μέγιστη καταστολή (NMS) για την επίλυση συγκρούσεων και τη διατήρηση των καλύτερων προβλέψεων.

Συνοψίζοντας, η IoU (Intersection over Union) είναι μια μετρική που χρησιμοποιείται σε εργασίες όρασης υπολογιστών στο πλαίσιο νευρωνικών δικτύων για την αξιολόγηση της ακρίβειας της ανίχνευσης και της κατάτμησης αντικειμένων. Υπολογίζει τον λόγο της επικαλυπτόμενης επιφάνειας μεταξύ των προβλεπόμενων ορίων και των πραγματικών ορίων εδάφους-αλήθειας προς τη συνδυασμένη περιοχή τους. Οι βαθμολογίες IoU κυμαίνονται από 0 (καμία επικάλυψη) έως 1 (τέλεια ταύτιση), παρέχοντας ένα αυστηρό μέτρο της απόδοσης ενός μοντέλου στον ορθό εντοπισμό και την ταξινόμηση αντικειμένων ή περιοχών σε εικόνες. Είναι ιδιαίτερα πολύτιμη επειδή λαμβάνει υπόψη τόσο το μέγεθος όσο και τη θέση της επικάλυψης, καθιστώντας την πιο ισχυρή μετρική σε σύγκριση με τις βασικές μετρήσεις ακρίβειας. Ωστόσο, η εφαρμογή της σε σύνθετα σενάρια, όπως πολλαπλά επικαλυπτόμενα αντικείμενα, απαιτεί πρόσθετες τεχνικές χειρισμού, όπως η μη μέγιστη καταστολή (NMS).

Στην παρούσα εργασία χρησιμοποιήθηκαν μοντέλα (νευρωνικά δίκτυα) που εκτελούν σημασιολογική κατάτμηση εικόνων (*image segmentation*) ορισμένου συνόλου δεδομένων (dataset) με στόχο την πρόβλεψη περιοχών αποψίλωσης (deforest), δασικών περιοχών (forest) και των υπολοίπων στοιχείων του φυσικού εδάφους, όπως τα υδάτινα στοιχεία. Η απόδοση των προβλέψεων των μοντέλων κρίνεται απαραίτητο να αξιολογηθεί και αυτό γίνεται μέσω των μετρικών. Οι αριθμητικές παράμετροι που χρησιμοποιήθηκαν για την αξιολόγηση είναι οι F1 - Score και Intersection Over Union (IOU) - Jaccard όπως περιγράφονται παραπάνω.

3.2.3 Cross - entropy loss

Η μετρική Cross-entropy loss ή log loss, χρησιμοποιείται στη μηχανική μάθηση και υπολογίζει την απόδοση ενός μοντέλου ταξινόμησης και μετρείται ως ένας αριθμός μεταξύ 0 και 1, με το 0 να αποτελεί ένα τέλειο μοντέλο. Στόχος γενικά είναι η προσέγγιση του μοντέλου όσο το δυνατόν πιο κοντά στο 0. Συγκεκριμένα, η μετρική Cross-entropy loss

αυξάνεται όσο η προβλεπόμενη τιμή από το μοντέλο αποκλίνει από την πραγματική τιμή (Ground Truth) που είναι και η ζητούμενη στο τέλος της πρόβλεψης. Συνεπώς, όσο η cross-entropy loss μειώνεται τόσο καλύτερη είναι και η πρόβλεψη του εκάστοτε μοντέλου που έχει μελετηθεί.

Η cross-entropy loss ή log loss, επίσης γνωστή ως απώλεια λογαρίθμου, μετρά την απόδοση ενός μοντέλου ταξινόμησης του οποίου η έξοδος είναι μια τιμή πιθανότητας μεταξύ 0 και 1. Η cross-entropy loss ή log loss αυξάνεται καθώς η προβλεπόμενη πιθανότητα αποκλίνει από την πραγματική ετικέτα. Έτσι, η πρόβλεψη μιας πιθανότητας 0,012 όταν η πραγματική ετικέτα παρατήρησης είναι 1 θα ήταν κακή και θα οδηγούσε σε υψηλή τιμή απώλειας. Ο στόχος του μοντέλου μηχανικής μάθησης είναι η ελαχιστοποίηση της απώλειας ή της διαφοράς μεταξύ των προβλεπόμενων τιμών και των πραγματικών τιμών.

Προερχόμενη από τον τομέα της θεωρίας της πληροφορίας, η έννοια της cross-entropy loss ή log loss μετρά τη διαφορά μεταξύ δύο κατανομών πιθανότητας. Στο πλαίσιο της μηχανικής μάθησης, χρησιμοποιείται για την ποσοτικοποίηση της διαφοράς μεταξύ της πραγματικής ετικέτας και των προβλέψεων του μοντέλου. Επίσης, χρησιμοποιείται ευρέως σε προβλήματα ταξινόμησης, ιδίως στη δυαδική ταξινόμηση και στην ταξινόμηση πολλαπλών κλάσεων. Στη βαθιά μάθηση, είναι συχνά η προεπιλεγμένη συνάρτηση απωλειών για τα νευρωνικά δίκτυα λόγω της αποτελεσματικότητάς της στην ώθηση των μοντέλων προς ακριβείς προβλέψεις.

Μαθηματικός τύπος υπολογισμού της μετρικής cross-entropy loss – Τύπος δυαδικής ταξινόμησης: $-(y * \log(p) + (1 - y) * \log(1 - p))$, όπου y είναι ο δυαδικός δείκτης (0 ή 1) της πραγματικής ετικέτας της κλάσης και p είναι η προβλεπόμενη πιθανότητα του μοντέλου για την κλάση με την ετικέτα 1.

Ταξινόμηση πολλαπλών κλάσεων: Για περιπτώσεις με περισσότερες από δύο κλάσεις, ο τύπος γενικεύεται ως εξής: $-\sum [y_{o,c} * \log(p_{o,c})]$, όπου M είναι ο αριθμός των κλάσεων, $y_{\{o,c\}}$ είναι ένας δυαδικός δείκτης για το αν η κλάση c είναι η σωστή ταξινόμηση για την παρατήρηση o , και $(p_{\{o,c\}})$ είναι η προβλεπόμενη πιθανότητα η παρατήρηση (o) να ανήκει στην κλάση (c).

Η απώλεια της διασταυρούμενης εντροπίας είναι αποτελεσματική στην αντιμετώπιση σεναρίων όπου η ταξινόμηση δεν είναι βέβαιη (πιθανολογικές εξόδους) και τείνει να τιμωρεί έντονα τις σίγουρες και τις λανθασμένες προβλέψεις, γεγονός που οδηγεί το μοντέλο να κάνει πιο συντηρητικές προβλέψεις με βάση τα δεδομένα που διαθέτει. Στα νευρωνικά δίκτυα, η βελτιστοποίηση της απώλειας διασταυρούμενης εντροπίας μοιάζει με την εκτέλεση εκτίμησης μέγιστης πιθανοφάνειας στις παραμέτρους. Αυτό κάνει τις προβλέψεις του μοντέλου να είναι πιο πιθανολογικά ερμηνεύσιμες. Μπορεί να είναι ευαίσθητο σε ανισοβαρή σύνολα δεδομένων, όπου οι κλάσεις δεν αντιπροσωπεύονται εξίσου. Επίσης, όταν το μοντέλο κάνει μια πρόβλεψη με υψηλή εμπιστοσύνη που είναι λανθασμένη, υφίσταται μεγάλη ποινή στην απώλεια. Υπάρχουν επεκτάσεις της βασικής συνάρτησης απώλειας cross-entropy, όπως η σταθμισμένη cross-entropy, η εξομάλυνση ετικετών και η εστιακή απώλεια, οι οποίες αντιμετωπίζουν συγκεκριμένες προκλήσεις, όπως η ανισορροπία κλάσεων ή η υπερβολική εμπιστοσύνη του μοντέλου. Με τη συνεχή ελαχιστοποίηση της απώλειας cross-entropy κατά τη διάρκεια της φάσης εκπαίδευσης, τα μοντέλα μηχανικής μάθησης μπορούν να βελτιώσουν αποτελεσματικά την ακρίβειά τους, εξασφαλίζοντας ότι οι πιθανότητες εξόδου είναι όσο το δυνατόν πιο κοντά στην πραγματική κατανομή των δεδομένων στα οποία εκπαιδεύτηκαν.

3.2.4 Accuracy (ακρίβεια)

Η ακρίβεια (*accuracy*) είναι μια μετρική η οποία αξιολογεί όπως δηλώνει και το όνομα της, πόσο ακριβής είναι η πρόβλεψη του μοντέλου σε σχέση με τα πραγματικά δεδομένα.

Δηλαδή, η ακρίβεια είναι ο λόγος των πραγματικών αποτελεσμάτων (τόσο των αληθινών όσο και αρνητικών) μεταξύ του συνολικού αριθμού των περιπτώσεων που εξετάστηκαν.

$$\frac{TP + TN}{TP + FP + FN + TN}$$

Η ακρίβεια είναι το μέτρο του πόσο συχνά το νευρωνικό δίκτυο κάνει μια σωστή πρόβλεψη. Είναι ο λόγος του αριθμού των σωστών προβλέψεων προς τον συνολικό αριθμό των προβλέψεων (ή των εισόδων που αξιολογήθηκαν). Η ακρίβεια υπολογίζεται χρησιμοποιώντας τον ακόλουθο τύπο: (αριθμός σωστών προβλέψεων) / (συνολικός αριθμός προβλέψεων). Για παράδειγμα, εάν ένα μοντέλο κάνει 100 προβλέψεις και οι 85 από αυτές είναι σωστές, η ακρίβεια είναι 0,85 ή 85%. Χρησιμοποιείται συνήθως σε προβλήματα ταξινόμησης δυαδικών και πολλαπλών κατηγοριών όπου οι έξοδοι είναι κατηγορικές (π.χ. spam ή όχι spam, κατηγορίες εικόνων κ.λπ.). Είναι απλή και εύκολα κατανοητή, γεγονός που την καθιστά δημοφιλή μετρική ειδικά σε προκαταρκτικές αξιολογήσεις.

Η ακρίβεια μπορεί να είναι παραπλανητική όταν έχουμε να κάνουμε με ανισόρροπα σύνολα δεδομένων, δηλαδή όταν ο αριθμός των παρατηρήσεων σε διαφορετικές κλάσεις ποικίλλει σημαντικά. Για παράδειγμα, εάν ένα σύνολο δεδομένων περιέχει 95% δείγματα της κλάσης A και 5% δείγματα της κλάσης B, ένα μοντέλο που προβλέπει την κλάση A για όλες τις περιπτώσεις θα εξακολουθεί να επιτυγχάνει ακρίβεια 95% παρά το γεγονός ότι δεν έχει μάθει σημαντικά μοτίβα που σχετίζονται με την κλάση B. Επίσης, δεν λαμβάνονται υπόψη οι διαφορετικοί τύποι σφαλμάτων που μπορεί να κάνει ένα μοντέλο, όπως τα ψευδώς θετικά και τα ψευδώς αρνητικά, τα οποία μπορεί να είναι κρίσιμα σε ορισμένες.

Για τον μετριασμό των περιορισμών της, ιδίως με ανισόρροπα σύνολα δεδομένων, χρησιμοποιούνται συχνά άλλες μετρικές παράλληλα ή αντί της ακρίβειας, όπως π.χ.:

- Ακρίβεια: Ο λόγος των σωστά προβλεπόμενων θετικών παρατηρήσεων προς το σύνολο των προβλεπόμενων θετικών παρατηρήσεων.
- Ανάκληση (ευαισθησία): Ο λόγος των σωστά προβλεπόμενων θετικών παρατηρήσεων προς τις πραγματικές θετικές παρατηρήσεις.
- F1-Score: Ο σταθμισμένος μέσος όρος της ακρίβειας και της ανάκλησης.

Περιοχή κάτω από την καμπύλη λειτουργικού χαρακτηριστικού δέκτη (ROC) (AUC-ROC):
Μια μέτρηση απόδοσης για προβλήματα ταξινόμησης σε διάφορες ρυθμίσεις κατωφλίων.

Ενώ η ακρίβεια μπορεί να παρέχει μια γρήγορη εικόνα της απόδοσης του μοντέλου, είναι σημαντικό να λαμβάνονται υπόψη το πλαίσιο και οι ειδικές απαιτήσεις της εκάστοτε εργασίας. Η χρήση της σε συνδυασμό με άλλες μετρικές ή η επιλογή εναλλακτικών μετρικών που ταιριάζουν καλύτερα στο πλαίσιο του προβλήματος, μπορεί να οδηγήσει σε μια πιο ολοκληρωμένη αξιολόγηση της απόδοσης ενός νευρωνικού δικτύου.

Η ακρίβεια στα νευρωνικά δίκτυα είναι μια μετρική απόδοσης που δείχνει την αναλογία των σωστών προβλέψεων προς το σύνολο των προβλέψεων που πραγματοποιεί το μοντέλο. Αν και είναι εύκολα κατανοητή και χρησιμοποιείται συνήθως για την αξιολόγηση εργασιών ταξινόμησης, το κύριο μειονέκτημά της είναι ότι μπορεί να είναι παραπλανητική, ιδίως σε ανισόρροπα σύνολα δεδομένων όπου η κατανομή των κλάσεων είναι άνιση. Δεν λαμβάνει υπόψη το είδος των σφαλμάτων (ψευδώς θετικά, ψευδώς αρνητικά) που κάνει ένα μοντέλο, το οποίο μπορεί να είναι κρίσιμο σε ορισμένα σενάρια. Ως εκ τούτου, η ακρίβεια χρησιμοποιείται συχνά μαζί με πιο διαφοροποιημένες μετρικές όπως η ακρίβεια, η ανάκληση, το F1-Score και το AUC-ROC, ιδίως όταν πρόκειται για σύνθετες ή ευαίσθητες εργασίες ταξινόμησης.

Κεφάλαιο 4 – Δεδομένα και εργαλεία «Keras» που υλοποιήθηκαν

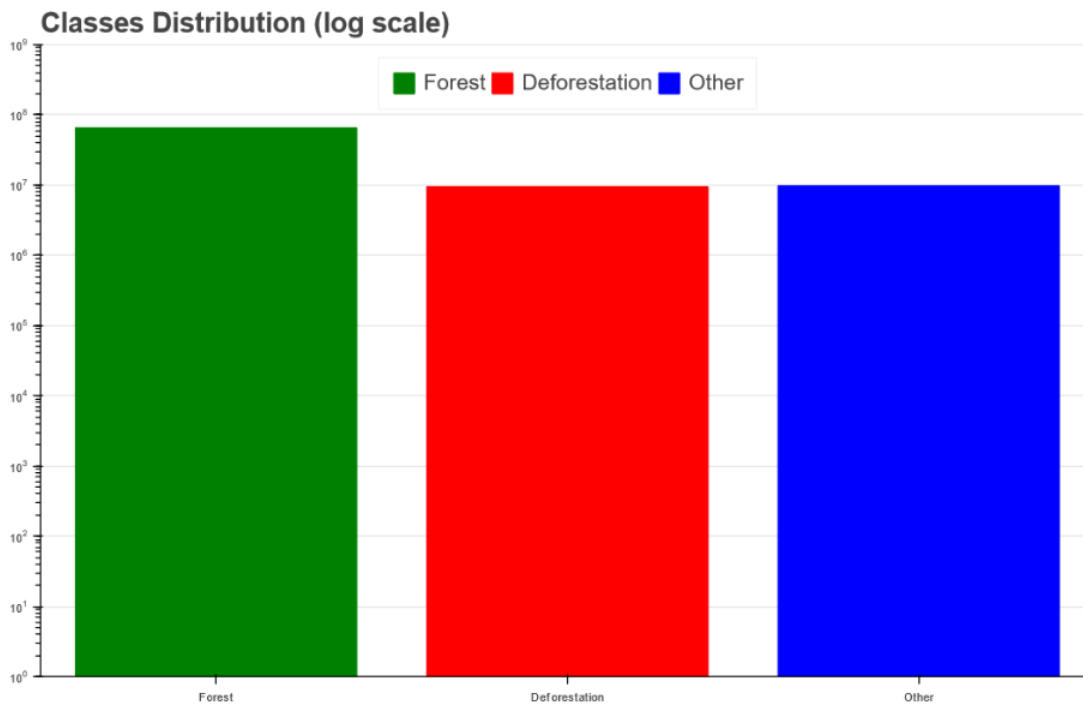
4.1 Περιγραφή των δεδομένων

Τα δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση των νευρωνικών δικτύων που παρουσιάζονται στην παρούσα διπλωματική εργασία αντλήθηκαν από την *open source* ιστοσελίδα *GitHub* και συγκεκριμένα από το προτεινόμενο μοντέλο «*BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection*».

Τα δεδομένα που χρησιμοποιήθηκαν για την εκπαίδευση των προτεινόμενων μοντέλων της συγκεκριμένης εργασίας αφορούν περιοχές ουκρανικών δασοκομιών και αντλήθηκαν μέσω του *Google Earth Pro*. Στην συνέχεια προκειμένου να προκύψουν εικόνες καθορισμένου μεγέθους και συγκεκριμένα μεγέθους 512 x 512 pixels έγινε κατάλληλη επεξεργασία με την *PyAutoGui module*. Η *PyAutoGui module* είναι βιβλιοθήκη της γλώσσας *python* η οποία χρησιμοποιείται για γραφικές διεπαφές. Η επεξεργασία των δεδομένων που αντλούνταν από τις παραπάνω πλατφόρμες απαιτούσε εικόνες υψηλής ποιότητας και η χρήση αντίστοιχων δεδομένων, όπως εικόνων *Landsat* δεν ήταν ικανοποιητική. Το πρόγραμμα έχει συγκεντρώσει 322 εικόνες από τρεις διαφορετικές περιοχές. Οι εικόνες δεν περιλαμβάνουν μόνο δασικές περιοχές ή περιοχές αποψίλωσης δασών, αλλά και περιοχές με δρόμους και οικισμούς που βρίσκονται μέσα στο δάσος, ανοιχτές πηγές νερού, όπως ποτάμια και λίμνες, ώστε το προς εκπαίδευση μοντέλο να γίνει περισσότερο εύρωστο και να αποκτήσει την μεγαλύτερη δυνατή ακρίβεια πρόβλεψης των περιοχών αποψίλωσης.

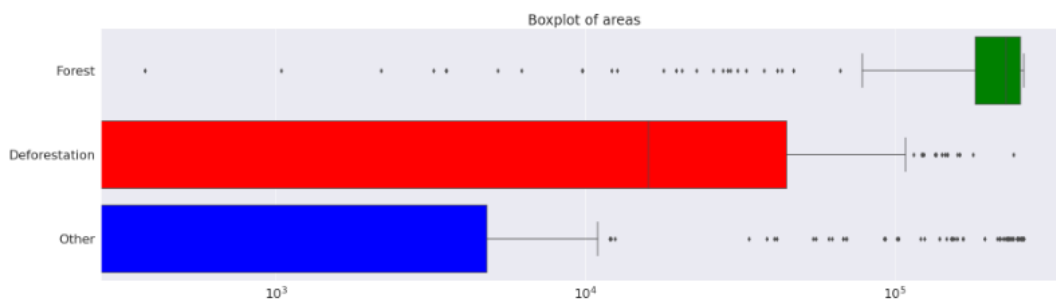
Συνεπώς, οι εικόνες διακρίνονται ουσιαστικά σε τρεις (3) βασικές κατηγορίες – κλάσεις που το μοντέλο πρέπει να ανιχνεύσει: τις δασικές περιοχές, τις περιοχές αποψίλωσης και δρόμους, οικισμούς, ποτάμια και λίμνες προκειμένου το νευρωνικό δίκτυο να εκπαιδευτεί με τον περισσότερο δυνατό βέλτιστο και ακριβή τρόπο στον εντοπισμό των περιοχών αποψίλωσης. Χωρίς την τρίτη κλάση, δηλαδή τις περιοχές με δρόμους οικισμούς, ποτάμια και λίμνες, η πρόβλεψή θα ήταν ορθή μόνο με εικόνες που εμπεριέχουν δάσος και περιοχές αποψίλωσης δασών.

Επιπλέον, επισημαίνεται ότι οι εικόνες περικόπηκαν σε μέγεθος 512 x 512 pixel με την *PyAutoGui module* και αποθηκεύτηκαν με τις αντίστοιχες μάσκες τους στο *Google Cloud Storage*. Στην εικόνα 13 παρακάτω είναι ορατό ότι η ποσότητα των εικονοστοιχείων από την κλάση «Δάσος» είναι κατά μία τάξη μεγαλύτερη από το πλήθος των εικονοστοιχείων στην κατηγορία «Αποψίλωση» και την κατηγορία «*Other*». Αυτό σημαίνει ότι το μοντέλο θα μπορούσε να προβλέψει την κατηγορία «*Forest*» με μεγαλύτερη ευκολία και υψηλότερη ακρίβεια σε σχέση με τις υπόλοιπες δύο κλάσεις «Αποψίλωση» και «*Other*» καθώς έχει την ευκαιρία να εκπαιδευτεί καλύτερα στην ανίχνευση της εφόσον υφίσταται στα δεδομένα σε αρκετά υψηλό ποσοστό.



Εικόνα 13 - Διάγραμμα απεικόνισης της έκτασης των κλάσεων forest, deforest & other στις εικόνες

Στην κατωτέρω εικόνα 14 φαίνεται ότι για καλύτερη ακρίβεια, το μοντέλο θα πρέπει να προβλέπει πολύ μικρές περιοχές αποψίλωσης των δασών στην εικόνα. Οι μικρές δασικές περιοχές στην εικόνα μπορεί να είναι ακόμα και μεμονωμένα δέντρα, κάτι που παραμένει πολύ σημαντικό να διαφοροποιηθεί από τις υπόλοιπες κατηγορίες - κλάσεις.

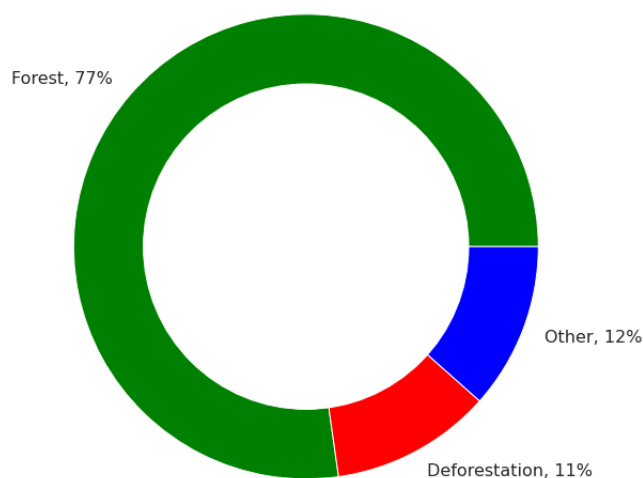


Εικόνα 14 - Boxplot of forest, deforest & other areas

Στην κατωτέρω εικόνα 15 φαίνεται ότι η δασική έκταση σε μια εικόνα καταλαμβάνει σε ποσοστό το 77% της έκτασης που απεικονίζεται σε μία εικόνα των δεδομένων, ενώ η περιοχές αποψίλωσης σε μια εικόνα καταλαμβάνουν σε ποσοστό το 11% της εικόνας. Αντίστοιχα, η κατηγορία «other» που αφορά τους δρόμους, τους οικισμούς, τα ποτάμια και τις λίμνες καταλαμβάνει σε ποσοστό το 12% της έκτασης της εικόνας.

Παρακάτω στις εικόνες 16 & 17 παρατίθενται ως παράδειγμα μία εικόνα από τα δεδομένα με την αντίστοιχη μάσκα της (Ground Truth):

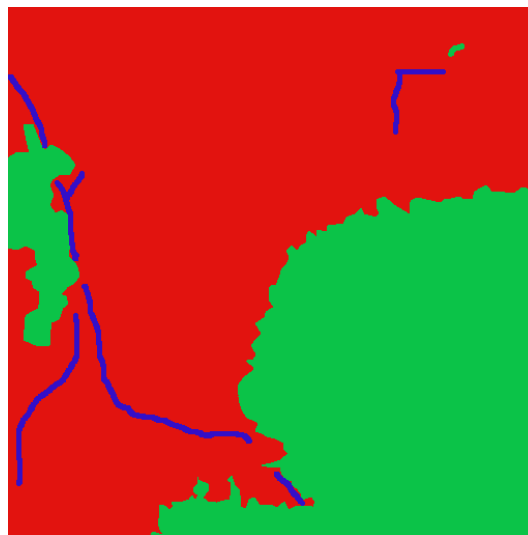
Average percentage of class area per image



Εικόνα 15 - Ποσοστό (%) πληροφορίας forest, deforest & other στις εικόνες



Εικόνα 16 - Δορυφορική εικόνα



Εικόνα 17 - Ground Truth

Τα δεδομένα λαμβάνονται σε μορφή αρχείων *.tfrec* (π.χ *tfrecords_v2_part_1.tfrec*) και συνολικά διατίθενται δέκα τέσσερα (14) αρχεία *.tfrec* τα οποία περιλαμβάνουν το κάθε ένα είκοσι δύο (22) εικόνες με τις αντίστοιχες μάσκες τους (*Ground Truth*). Συνεπώς, τα διαθέσιμα δεδομένα για την συγκεκριμένη εργασία είναι συνολικά 308 εικόνες με τις αντίστοιχες μάσκες τους. Η ανωτέρω μορφή αρχείου χρησιμοποιήθηκε και στα μοντέλα που επεξεργάστηκαν στην παρούσα διπλωματική και παρουσιάζονται στις κατωτέρω ενότητες 5.2, 5.3 και 5.4.

4.2 Εργαλεία – Διεπαφή Keras

Για την υλοποίηση της παρούσας διπλωματικής εργασίας επιλέχθηκαν μοντέλα που χρησιμοποιούν το KERAS, μια διεπαφή προγραμματισμού εφαρμογών (Application Programming Interface-API) υψηλού επιπέδου νευρωνικών δικτύων, το οποίο είναι γραμμένο

σε Python και είναι ικανό να τρέχει πάνω από το TensorFlow και άλλες βιβλιοθήκες. Το Keras περιέχει πολλές υλοποιήσεις των δομικών στοιχείων που χρησιμοποιούνται συνήθως σε νευρωνικά δίκτυα, όπως επίπεδα (layers), σφάλματα (losses), λειτουργίες ενεργοποίησης (activation functions), βελτιστοποιητές (optimizers), μετρικές (metrics). Εκτός από το KERAS χρησιμοποιήθηκαν και συναρτήσεις από το tensorflow.

Κεφάλαιο 5 – Πειραματική Διάταξη

5.1 Μοντέλα Μηχανικής Μάθησης

Στην παρούσα εργασία χρησιμοποιήθηκαν συνολικά τρία μοντέλα μηχανικής μάθησης για την αντιμετώπιση του προβλήματος εντοπισμού αποψιλωμένων περιοχών και τη σύγκριση μεταξύ τους ως προς τα αποτελέσματα που παράγει το κάθε μοντέλο. Συνεπώς, στη συγκεκριμένη ενότητα θα περιγραφεί αναλυτικά η εκπαίδευση του κάθε μοντέλου, ο στόχος που είχε η κάθε μια αναφορικά με το πρόβλημα του εντοπισμού αποψιλωμένων περιοχών και αντίστοιχα η ποιότητα των αποτελεσμάτων που παράγει η κάθε μια.

Η αναζήτηση των μοντέλων μηχανικής μάθησης για την υλοποίηση του προβλήματος που πραγματεύεται η παρούσα διπλωματική εργασία έγινε μέσω της ιστοσελίδας *GitHub* (<https://github.com/>). Η ανωτέρω ιστοσελίδα (*GitHub*) αποτελεί μία cloud-based υπηρεσία, που δημιουργήθηκε το 2008, η οποία χρησιμοποιείται για την αποθήκευση κώδικα, όσο και για την αποθήκευση αλλαγών που πραγματοποιούνται σε ένα project, καθώς και την κοινοποίηση του εκάστοτε project σε άλλα άτομα. Συνεπώς, αναζητήθηκαν μοντέλα μηχανικής μάθησης τα οποία έχουν κατασκευαστεί ώστε να υλοποιούν σημασιολογική κατάτμηση εικόνων (*image segmentation*).

5.2 Πρότυπο Μοντέλο Μηχανικής Μάθησης

Υστερα από έρευνα στην ιστοσελίδα *GitHub* για εύρεση κατάλληλων δεδομένων και μοντέλων μηχανικής μάθησης για σημασιολογική κατάτμηση εικόνων αναφορικά με το φαινόμενο της αποψίλωσης (*deforestation*), ευρέθηκε μοντέλο με τίτλο «*Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection*».

(<https://github.com/BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection>).

Το ανωτέρω μοντέλο μηχανικής μάθησης υλοποιεί το πρόβλημα εντοπισμού περιοχών αποψίλωσης μέσω της σημασιολογικής κατάτμησης δορυφορικών εικόνων. Αρχικά χρησιμοποιήθηκε προς επεξεργασία το προεκπαιδευμένο μοντέλο «*Model_Testing*» (https://github.com/BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection/blob/main/Model_Testing.ipynb) με έτοιμο αρχείο κατάλληλων βαρών (*U6_E_1201-F1_0.7134-IOU_0.6555.h5*) εφόσον η εκπαίδευση του συγκεκριμένου μοντέλου έχει ήδη πραγματοποιηθεί στο αντίστοιχο αναλυτικό μοντέλο «*Model_Training_TPU_Strategy*» (https://github.com/BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection/blob/main/Model_Training_TPU_Strategy.ipynb). Το προεκπαιδευμένο μοντέλο «*Model_Testing*» χρησιμοποιείται ως μοντέλο αναφοράς καθώς η εκπαίδευση του έχει πραγματοποιηθεί στο σύνολο δεδομένων (*Dataset*) που έχει ληφθεί υπόψη στην συγκεκριμένη εργασία, ο κώδικας που έχει υλοποιηθεί αφορά άμεσα το πρόβλημα εντοπισμού αποψιλωμένων περιοχών (*Deforestation*) μέσω της σημασιολογικής κατάτμησης εικόνων (*image segmentation*) και τα αποτελέσματα αναμένεται να είναι τα βέλτιστα.

Το προεκπαιδευμένο μοντέλο που χρησιμοποιήθηκε και επεξεργάστηκε αρχικά δεν θα αποτελέσει συγκριτικό στοιχείο για την αξιολόγηση των αποτελεσμάτων που προκύπτουν από τα κατωτέρω μοντέλα που περιγράφονται αναλυτικά στις ενότητες 5.3 και 5.4, καθώς αυτό έχει εκπαιδευτεί στο σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα διπλωματική εργασία.

5.2.1 Συναπτικά Βάρη Τεχνητού Νευρωνικού Δικτύου

Ένα τεχνητό δίκτυο αποτελείται από ένα σύνολο τεχνητών νευρώνων που αλληλεπιδρούν, συνδεδεμένοι μεταξύ τους με τις λεγόμενες συνάψεις (synapses). Ο βαθμός αλληλεπίδρασης είναι διαφορετικός για κάθε ζεύγος νευρώνων και καθορίζεται από τα λεγόμενα συναπτικά βάρη (synaptic weights). Συγκεκριμένα, καθώς το νευρωνικό δίκτυο αλληλεπιδρά με το περιβάλλον και μαθαίνει από αυτό, τα συναπτικά βάρη μεταβάλλονται συνεχώς, ενδυναμώνοντας ή αποδυναμώνοντας την ισχύ του κάθε δεσμού. Όλη η εμπειρική γνώση που αποκτά επομένως το νευρωνικό δίκτυο από το περιβάλλον κωδικοποιείται στα συναπτικά βάρη. Αυτά αποτελούν το χαρακτηριστικό εκείνο που δίνει στο δίκτυο την ικανότητα για εξέλιξη και προσαρμογή στο περιβάλλον.

Συνεπώς, στο προεκπαιδευμένο μοντέλο «*Model Testing*» εφαρμόζεται αρχείο κατάλληλων βαρών (*U6_E_1201-F1_0.7134-IOU_0.6555.h5*) που έχει προκύψει ύστερα από εκπαίδευση του αναλυτικού μοντέλου «*Model Training TPU Strategy*» και πιο συγκεκριμένα το παραπάνω αρχείο βαρών εφαρμόζεται στο μοντέλο «*model10*» που αναφέρεται παρακάτω στην ενότητα 4.3.1.2.

5.2.2 Προεκπαιδευμένο Μοντέλο «*Model Testing*» (Pre-Trained Model)

Στον κώδικα του προεκπαιδευμένου μοντέλου πραγματοποιήθηκαν οι κατωτέρω προσθήκες:

- Προσθήκη τμήματος κώδικα προκειμένου να οριστούν οι μετρικές (Metrics): Αρμονικός Μέσος - F1 – Score και Intersection Over Union (IOU), για την αξιολόγηση των αποτελεσμάτων του μοντέλου (βλ. Παράρτημα Κώδικα 1 - Αρμονικός Μέσος - F1 – Score & Παράρτημα Κώδικα 2 - Intersection Over Union (IOU)).
- Προσθήκη τμήματος κώδικα για την κατάλληλη εκτύπωση (plot) των αποτελεσμάτων των Μετρικών F1 – Score και Intersection Over Union (IOU) στο πεδίο «*Plot predictions*» στο οποίο γίνεται η εκτύπωση των προβλεπόμενων εικόνων από το μοντέλο (*Model Prediction*) (βλ. Παράρτημα Κώδικα 3 – Πεδίο «*Plot Prediction*» & Παράρτημα Κώδικα 4 – Υπολογισμός & Εκτύπωση F1 – Score & Intersection Over Union (IOU)).
- Αλλαγή μεγέθους εισαγόμενων εικόνων στο οριζόμενο (*model10*) από 1024 x 1024 x 3 σε 512 x 512 x 3 καθώς το σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα διπλωματική εργασία αποτελείται από εικόνες μεγέθους 512 x 512 pixels (βλ. Ενότητα 4.1 σύνολο δεδομένων) (βλ. Παράρτημα Κώδικα 5 – Ορισμός μοντέλου (*model10*)). Επιπλέον, στο συγκεκριμένο πεδίο του κώδικα γίνεται ορισμός των φίλτρων (*final_filters=2048*). Δηλαδή, γίνεται η εφαρμογή συνολικά 7 φίλτρων, από 5 έως 12, $filters = [2^5=32, 2^6=64, 2^7=128, 2^8=256, 2^9=512, 2^{10}=1024, 2^{11}=2048]$, $\log_2(2048)=11+1 = 12$.

Παράρτημα Κώδικα 1 – Αρμονικός Μέσος – F1 – Score:

```
class F1Score(tf.keras.metrics.Metric):
```

```
    def __init__(self, name, **kwargs):
```

```
        super(F1Score, self).__init__(name=name, **kwargs)
```

```
        self.f1score_forest = self.add_weight(name='f1score_forest', initializer='zeros')
```

```
        self.f1score_deforest = self.add_weight(name='f1score_deforest', initializer='zeros')
```

```
        self.f1score_other = self.add_weight(name='f1score_other', initializer='zeros')
```

```
    def update_state(self, y_true, y_pred, sample_weight=None):
```

```
        y_true = tf.reshape(y_true, (tf.shape(y_true)[0], -1, tf.shape(y_true)[-1]))
```

```

y_pred = tf.reshape(y_pred, (tf.shape(y_pred)[0], -1, tf.shape(y_pred)[-1]))
value = 0
#forest
intersection = tf.math.reduce_sum(tf.math.multiply(y_true[..., 0], y_pred[..., 0]))
union = tf.math.reduce_sum(y_true[..., 0]) + tf.math.reduce_sum(y_pred[..., 0])
value = tf.math.reduce_sum((2. * intersection + 1e-16) / (union + 1e-16))
self.f1score_forest.assign_add(value)
value=0
#deforest
intersection = tf.math.reduce_sum(tf.math.multiply(y_true[..., 1], y_pred[..., 1]))
union = tf.math.reduce_sum(y_true[..., 1]) + tf.math.reduce_sum(y_pred[..., 1])
value = tf.math.reduce_sum((2. * intersection + 1e-16) / (union + 1e-16))
self.f1score_deforest.assign_add(value)
value=0
#other
intersection = tf.math.reduce_sum(tf.math.multiply(y_true[..., 2], y_pred[..., 2]))
union = tf.math.reduce_sum(y_true[..., 2]) + tf.math.reduce_sum(y_pred[..., 2])
value = tf.math.reduce_sum((2. * intersection + 1e-16) / (union + 1e-16))
self.f1score_other.assign_add(value)
value=0

def reset_states(self):
    for s in self.variables:
        s.assign(tf.zeros(shape=s.shape))

def result(self):
    return [self.f1score_forest,self.f1score_deforest,self.f1score_other]

```

Παράρτημα Κώδικα 2 - Intersection Over Union (IOU):

```

class IOU(tf.keras.metrics.Metric):

    def __init__(self, name, **kwargs):
        super(IOU, self).__init__(name=name, **kwargs)
        self.iou_forest = self.add_weight(name='iou_forest', initializer='zeros')
        self.iou_deforest = self.add_weight(name='iou_deforest', initializer='zeros')
        self.iou_other = self.add_weight(name='iou_other', initializer='zeros')

    def update_state(self, y_true, y_pred, sample_weight=None):
        y_true = tf.reshape(y_true, (tf.shape(y_true)[0], -1, tf.shape(y_true)[-1]))
        y_pred = tf.reshape(y_pred, (tf.shape(y_pred)[0], -1, tf.shape(y_pred)[-1]))
        value = 0
        #forest
        intersection = tf.math.reduce_sum(tf.math.multiply(y_true[..., 0], y_pred[..., 0]))
        union = tf.math.reduce_sum(y_true[..., 0]) + tf.math.reduce_sum(y_pred[..., 0])
intersection
        value += tf.math.reduce_sum((intersection + 1e-16) / (union + 1e-16))
        self.iou_forest.assign_add(value)
        value=0
        #deforest
        intersection = tf.math.reduce_sum(tf.math.multiply(y_true[..., 1], y_pred[..., 1]))
        union = tf.math.reduce_sum(y_true[..., 1]) + tf.math.reduce_sum(y_pred[..., 1])
intersection
        value += tf.math.reduce_sum((intersection + 1e-16) / (union + 1e-16))

```



```

self.iou_deforest.assign_add(value)
value=0
#other
intersection = tf.math.reduce_sum(tf.math.multiply(y_true[..., 2], y_pred[..., 2]))
union = tf.math.reduce_sum(y_true[..., 2]) + tf.math.reduce_sum(y_pred[..., 2]) -
intersection
value += tf.math.reduce_sum((intersection + 1e-16) / (union + 1e-16))
self.iou_other.assign_add(value)
value=0

def reset_states(self):
    for s in self.variables:
        s.assign(tf.zeros(shape=s.shape))

def result(self):
    return [self.iou_forest,self.iou_deforest,self.iou_other]

```

Παράρτημα Κώδικα 3 – Πεδίο «Plot Prediction»

```

# Plot predictions
for i, model in enumerate(models):
    for j, image in enumerate(images):
        f1 = F1Score(name="f1-score")
        iou = IOU(name="iou")
        image = image[np.newaxis, ...]
        prediction = model.predict(image)
        f1.update_state(masks[j],prediction[0])
        iou.update_state(masks[j],prediction[0])
        f1_result = f1.result().numpy()
        iou_result = iou.result().numpy()
        if flags[i] == True:
            prediction_class1 = np.copy(prediction[..., 0]) # Forest
            prediction_class2 = np.copy(prediction[..., 1]) # Deforest
            prediction[..., 0] = prediction_class2 # RED - Deforest
            prediction[..., 1] = prediction_class1 # GREEN - Forest
            df_ax.iloc[j][i+2].imshow(prediction[0]) # i + 1 because 0 column is complete.
UPDATE: i + 2 because 1 column is gt
            if j == 0:
                df_ax.iloc[j][i+2].title.set_text(f'{model_names[i]}')
                df_ax.iloc[j][i+2].set_xticklabels([])
                df_ax.iloc[j][i+2].set_yticklabels([])
                print(f'[INFO] Image {j+1}/{len(images)}, model {i}.')
                print(f'F1_Score_forest: {f1_result[0]}", f'F1_Score_deforest: {f1_result[1]}",
f'F1_Score_other: {f1_result[2]}")
                print(f'iou_forest: {iou_result[0]}", f'iou_deforest: {iou_result[1]}", f'iou_other:
{iou_result[2]}")

            print("[INFO] Finished!")
            plt.subplots_adjust(wspace=0, hspace=0)
            plt.savefig("Predictions.png", dpi=100)
            return df_ax

```

Παρακάτω στο Παράρτημα Κώδικα 4 – Υπολογισμός & Εκτύπωση F1 – Score & Intersection Over Union (IOU) αναγράφονται οι γραμμές κώδικα που προστέθηκαν στο πεδίο

«Plot Predictions», ώστε να υπολογίζονται και να εκτυπώνονται για κάθε μια από τις εικοσιδύο (22) εικόνες ενός επιλεγμένου αρχείου «*tfrecords_v2_part_*.tfrec*» από τα συνολικά 14 .tfrec αρχεία του Dataset, οι μετρικές F1 – Score & Intersection Over Union (IOU) για κάθε μια από τις τρεις κλάσεις «Forest», «Deforest» και «Other». Όπως εμφανίζεται και παρακάτω στο Παράρτημα Κώδικα 4 – Υπολογισμός & Εκτύπωση F1 – Score & Intersection Over Union (IOU) για την καλύτερη εκτύπωση των μετρικών έχουν οριστεί οι ονομασίες «*F1_Score_forest*», «*F1_Score_deforest*» & «*F1_Score_other*» και «*iou_forest*», «*iou_deforest*» & «*iou_other*».

Παράρτημα Κώδικα 4 – Υπολογισμός & Εκτύπωση F1 – Score & Intersection Over Union (IOU)

```
f1 = F1Score(name="f1-score")
iou = IOU(name="iou")

f1.update_state(masks[j],prediction[0])
iou.update_state(masks[j],prediction[0])
f1_result = f1.result().numpy()
iou_result = iou.result().numpy()

print(f'F1_Score_forest: {f1_result[0]}", f'F1_Score_deforest: {f1_result[1]}",
f'F1_Score_other: {f1_result[2]}")
print(f'iou_forest: {iou_result[0]}", f'iou_deforest: {iou_result[1]}", f'iou_other:
{iou_result[2]}")
```

Παράρτημα Κώδικα 5 – Ορισμός μοντέλου (model10)

```
final_filters = 2048

model10 = build_unet(input_shape=(512, 512, 3),
    filters=[2 ** i for i in range(5, int(np.log2(final_filters) + 1))], # filters =
[2^5=32, 2^6=64, 128, 256, 512, 1024, 2048] log2(2048)=11+1 = 12. Συνεπώς, από 5 μέχρι
12, γίνεται εφαρμογή 7 φίλτρων. # Amount of filters in U-Net arch.
    batchnorm=False, transpose=False, dropout_flag=False)
```

Προκειμένου να επιτευχθεί ο έλεγχος για την ορθή λειτουργία του προεκπαιδευμένου μοντέλου «*Model_Testing*», χρησιμοποιήθηκε από το σύνολο των δεδομένων το αρχείο «*tfrecords_v2_part_5.tfrec*» το οποίο περιλαμβάνει συνολικά σαράντα τέσσερις εικόνες, είκοσι δύο δορυφορικές εικόνες και τις αντίστοιχες μάσκες τους, όπως αναφέρθηκε και στην ενότητα 4.1.

Η εικόνα που πρέπει να προβλέπει το μοντέλο μηχανικής μάθησης είναι μια εικόνα συμβατή με την αντίστοιχη μάσκα. Όσο μεγαλύτερη ομοιότητα υφίσταται μεταξύ μάσκας και προβλεπόμενης από το μοντέλο εικόνας τόσο ορθότερη καθίσταται η πρόβλεψη του μοντέλου. Η αξιολόγηση της πρόβλεψης του μοντέλου εμφανίζεται αριθμητικά μέσω των μετρικών: αρμονικός μέσος F1 – Score & Intersection Over Union (IOU).

Παρατίθενται link του κώδικα του προεκπαιδευμένου μοντέλου «*Model_Testing*» (Pre-Trained Model) μετά την επεξεργασία και την προσθήκη των τμημάτων κώδικα όπως περιγράφεται παραπάνω στην παρούσα ενότητα 5.2.2, <https://colab.research.google.com/drive/19uJTthEQXmPSIFC2VrJpyMW30TJTMUWj>

5.3 Μοντέλο Μηχανικής Μάθησης «Swin – Unet»

Αναφορικά με το πρόβλημα της σημασιολογικής κατάτμησης εικόνων (*image segmentation*) και συγκεκριμένα των εικόνων του συνόλου δεδομένων που χρησιμοποιείται στην παρούσα εργασία (βλ. ενότητα 4.1) αναζητήθηκαν μοντέλα μηχανικής μάθησης μέσω της ιστοσελίδας GitHub (<https://github.com/>) τα οποία έχουν κατασκευαστεί ώστε να επιτυγχάνουν την ορθή κατάτμηση εικόνων. Για τον σκοπό αυτό εκτός του μοντέλου που αναλύθηκε παραπάνω στην ενότητα 4.3.1 βρέθηκε το μοντέλο με τίτλο «*keras-vision-transformer*».

(<https://github.com/yingkaisha/keras-vision-transformer>)

Το μοντέλο «*keras-vision-transformer*» αποτελεί μια εφαρμογή του μετασχηματισμού «Swin Transformer». Ο «Swin Transformer» αποτελεί έναν μετασχηματισμό εικόνας (*vision Transformer*) και παρουσιάζει πολύ καλά αποτελέσματα σε προβλήματα σημασιολογικής κατάτμησης. Η ανάλυση για τον μετασχηματισμό «Swin Transformer» γίνεται στην ενότητα 2.4.

Στον κώδικα του μοντέλου έχουν οριστεί κάποιες βασικές παράμετροι του μετασχηματισμού «*Swin Transformer*», όπως το *patch size* και το *attention window size*. Συγκεκριμένα, το *patch_size=(4x4)* και το *attention window_size=(4,2,2,2)*.

Το συγκεκριμένο μοντέλο έχει κατασκευαστεί ώστε να λειτουργεί για ένα ορισμένο μέγεθος εικόνων *128 pixels x 128 pixels* που έχει επιλεγεί αρχικά από τον χρήστη. Για την ορθή λειτουργία του μοντέλου με το σύνολο δεδομένων που αντλήθηκε από το πρότυπο μοντέλο (βλ. ενότητα 4.3.1) και αφορά το φαινόμενο της αποψίλωσης που μελετάται στη συγκεκριμένη εργασία, πραγματοποιήθηκαν ορισμένες αλλαγές στον κώδικα του μοντέλου οι οποίες παρουσιάζονται παρακάτω στα *Παραρτήματα κώδικα 6, 7, 8*.

Παρακάτω στο *Παράρτημα 6 - Συναρτήσεις get_example και parse_image_function για επεξεργασία του συνόλου δεδομένων* αναγράφονται οι γραμμές κώδικα που προστέθηκαν στο αρχικό μοντέλο και αφορούν την επεξεργασία των εικόνων μεγέθους *512 pixels x 512 pixels* του Dataset της παρούσας διπλωματικής (*images και masks – Ground Truth*).

Παράρτημα κώδικα 6 – Συναρτήσεις get_example και parse_image_function για επεξεργασία του συνόλου δεδομένων

```
def get_example(dataset_path, image_num):
    raw_image_dataset = tf.data.TFRecordDataset(dataset_path)
    # Create a dictionary describing the features.
    image_feature_description = {
        "image": tf.io.FixedLenFeature([], tf.string),
        "mask": tf.io.FixedLenFeature([], tf.string),
    }

    def _parse_image_function(example_proto):
        # Parse the input tf.train.Example proto using the dictionary above.
        return tf.io.parse_single_example(example_proto, image_feature_description)

    parsed_image_dataset = raw_image_dataset.map(_parse_image_function)
    for image_features in parsed_image_dataset:
        if image_num != 0:
            image_raw = image_features['image'].numpy()
            mask_raw = image_features['mask'].numpy()
            image_num -= 1
        else:
```

```

break
image = tf.image.decode_png(image_raw, channels=3)
image = tf.image.resize(image, [512, 512])
image = tf.cast(image, tf.float32) / 255.0 # convert image to floats in [0, 1] range
mask = tf.io.decode_raw(mask_raw, out_type="float")
mask = tf.reshape(mask, [512, 512, 3])
mask = tf.cast(mask, tf.float32)
return image, mask

```

Παρακάτω στο *Παράρτημα 7 - Εισαγωγή εικόνων στο μοντέλο και μείωση (downscale) αυτών σε εικόνες 128 pixels x 128 pixels με την συνάρτηση block_reduce της βιβλιοθήκης skimage.measure* αναγράφονται οι γραμμές κώδικα που προστέθηκαν στο αρχικό μοντέλο και αφορούν:

- την εισαγωγή των εικόνων μεγέθους 512 x 512 pixels του συνόλου δεδομένων που αντλήθηκε από το πρότυπο μοντέλο μηχανικής μάθησης «*Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection*» όπως αυτό περιγράφεται στην ενότητα 4.3.1.
- την αλλαγή του μεγέθους των εικόνων από 512 pixels x 512 pixels σε εικόνες μεγέθους 128 pixels x 128 pixels με την συνάρτηση block_reduce και την μεταβλητή np.mean της βιβλιοθήκης skimage.measure. Προκειμένου να λειτουργεί ορθά το μοντέλο με το σύνολο των δεδομένων που μελετάται σε αυτή την εργασία, εφόσον έχει κατασκευαστεί ώστε να εισάγονται σε αυτό εικόνες μεγέθους 128 x 128 pixels, η μετατροπή αυτή ήταν απαραίτητη. Χρησιμοποιήθηκε τεχνική μείωσης μεγέθους (downscale) των εικόνων με την συνάρτηση που αναφέρεται παραπάνω (block_reduce) και την μεταβλητή (np.mean).

Παράρτημα κώδικα 7 – Εισαγωγή εικόνων στο μοντέλο και μείωση (downscale) αυτών σε εικόνες 128 pixels x 128 pixels με την συνάρτηση block_reduce της βιβλιοθήκης skimage.measure

```

images = []
masks = []

#try: we downscale each image of 512x512
#to image of 128x128
#this way we will have less data for training
from skimage.measure import block_reduce

for i in range(1,15):
    path = '/content/drive/MyDrive/Dataset/tfrecords_v2_part_'+ str(i) +'.tfrec'
    for k in range(1, 23):
        image, mask = get_example(path, k)
        resized_image = block_reduce(image,(4,4,1), np.mean)
        resized_mask = block_reduce(mask,(4,4,1), np.mean)
        images.append(resized_image)
        masks.append(resized_mask)

print("Images: ",len(images))
print("Masks: ",len(masks))
print("Image Test: ",images[0].shape)
print("Masks Test: ", masks[0].shape)

```

Μια ακόμη απαραίτητη διαδικασία που πρέπει να ακολουθηθεί, όσον αφορά τα δεδομένα, ώστε να προκύπτουν από το μοντέλο σωστές προβλέψεις είναι η κατηγοριοποίηση

τους σε πακέτο δεδομένων εκπαίδευσης (*training set*), πακέτο δεδομένων ελέγχου (*test set*) και πακέτο δεδομένων αξιολόγησης (*validation set*) (βλ. ενότητα 1.2.2). Παρακάτω στο Παράρτημα 8 – Κατηγοριοποίηση/Διαχωρισμός δεδομένων αναγράφεται το τμήμα του κώδικα κατά τον οποίο γίνεται ο τυχαίος διαχωρισμός των δεδομένων. Συγκεκριμένα, το 80% των δεδομένων χρησιμοποιείται για την εκπαίδευση και το υπόλοιπο ποσοστό 20% χρησιμοποιείται για τον έλεγχο και την αξιολόγηση του μοντέλου. Δηλαδή, 10% των δεδομένων αντιστοιχεί στον έλεγχο (*test set*) και 10% των δεδομένων στην αξιολόγηση (*validate set*). Συμπερασματικά, εάν συνολικά τα δεδομένα είναι τριακόσιες οκτώ (308) εικόνες αυτό σημαίνει ότι διακόσιες σαράντα έξι (246) εικόνες θα χρησιμοποιηθούν για την εκπαίδευση, τριάντα (30) περίπου εικόνες για τον έλεγχο και αντίστοιχα τριάντα (30) περίπου εικόνες για την αξιολόγηση.

Το πλήθος των εικόνων μετά την κατηγοριοποίηση τους όπως προκύπτει από το μοντέλο: *Training:validation:testing = 246:30:32*.

Τέλος, πρέπει να οριστούν σταθερά σύνολα εκπαίδευσης (*training set*), ελέγχου (*testing set*) και αξιολόγησης (*validation set*) για κάθε δοκιμή εκπαίδευσης προκειμένου να επιτευχθεί ορθή σύγκριση των προβλέψεων που προκύπτουν από το μοντέλο. Δηλαδή, πρέπει να χρησιμοποιηθούν για κάθε δοκιμή εκπαίδευσης οι ίδιες εικόνες για εκπαίδευση, έλεγχο και αξιολόγηση. Το ανωτέρω επιτυγχάνεται μέσω της συνάρτησης *np.random.seed(15)* όπου τυχαία επιλέγεται με τον αριθμό 15 να λαμβάνει το μοντέλο σταθερό σε κάθε δοκιμή εκπαίδευσης πακέτο εικόνων εκπαίδευσης, ελέγχου και αξιολόγησης.

Παράρτημα 8 – Κατηγοριοποίηση/Διαχωρισμός δεδομένων

```
sample_names = np.array(images)
label_names = np.array(masks)

L = len(sample_names)
np.random.seed(15)
ind_all = utils.shuffle_ind(L)

L_train = int(0.8*L); L_valid = int(0.1*L); L_test = L - L_train - L_valid
ind_train = ind_all[:L_train]; ind_valid = ind_all[L_train:L_train+L_valid]; ind_test =
ind_all[L_train+L_valid:]
print("Training:validation:testing = {}: {}: {}".format(L_train, L_valid, L_test))
```

Δοκιμές Εκπαίδευσης

Σε κάθε μία από τις κατωτέρω δοκιμές εκπαίδευσης αναγράφονται οι εποχές (*N_epoch*), τα batches per epoch (*N_batch*) και τα samples per batch (*N_sample*) που έχουν οριστεί για την εκπαίδευση του μοντέλου. Το γινόμενο των batches και samples δεν πρέπει να ξεπερνάει σε μεγάλο ποσοστό το πλήθος των δεδομένων εκπαίδευσης. Επιπρόσθετα, πρέπει να υπάρχει μια ισορροπία μεταξύ των δύο αυτών τιμών. Ύστερα από διάφορες δοκιμές βρέθηκε ότι ο καταλληλότερος συνδυασμός είναι *batches=8* και *samples=32*. Συνεπώς, με την σταθερή επιλογή των *batches=8* και *samples=32* πραγματοποιείται δοκιμή των προβλέψεων του μοντέλου για διαφορετικό πλήθος εποχών εκπαίδευσης (*epoch*). Κατά την διάρκεια της εκπαίδευσης του μοντέλου και κατά την ολοκλήρωση κάθε εποχής (*epoch*) ορίζεται η τιμή «*validation performance*» σύμφωνα με την οποία αξιολογείται η πορεία της εκπαίδευσης. Η τιμή «*validation performance*» παίρνει μια μέγιστη τιμή στην αρχή κάθε εκπαίδευσης η οποία ονομάζεται «*Initial loss*» και στην πορεία της εκπαίδευσης αναμένεται να μειωθεί έως ούτε φτάσει στην δυνατή ελάχιστη τιμή βάσει της εκάστοτε εκπαίδευσης που έχει οριστεί να πραγματοποιηθεί.

Επιπλέον, για κάθε δοκιμή εκπαίδευσης αναγράφεται η υπολογιζόμενη από το μοντέλο μετρική «*cross-entropy loss*» του testing set (*Testing set cross-entropy loss*), ώστε να

αξιολογηθεί η δοκιμή εκπαίδευσης με τις καλύτερες προβλέψεις. Η συγκεκριμένη μεταβλητή υποδηλώνει τα βέλτιστα αποτελέσματα όσο τείνει στο μηδέν.

Τέλος, για κάθε δοκιμή εκπαίδευσης αναγράφονται οι υπολογιζόμενες από το μοντέλο μετρικές «*f1-score*» & «*Intersection over Union - IoU*» ή «*Jaccard_score*» για κάθε μια από τις τρεις κατηγορίες – κλάσεις «*forest*», «*deforest*» & «*other*» ώστε να αξιολογηθεί το ποσοστό επιτυχίας εντοπισμού της κάθε κατηγορίας – κλάσης στο εκάστοτε test set που έχει ληφθεί. Οι συγκεκριμένες μεταβλητές υποδηλώνουν τα βέλτιστα αποτελέσματα όσο τείνουν στο 1 (σε ποσοστό 100%).

Στις κατωτέρω εικόνες, όπου εμφανίζονται οι προβλέψεις του μοντέλου σε κάθε μια από τις δοκιμές εκπαίδευσης, με το χρώμα πράσινο απεικονίζεται η δασική έκταση (*forest*), με το χρώμα κόκκινο απεικονίζεται η περιοχή αποψίλωσης (*deforest*) και με το χρώμα μπλε απεικονίζεται οποιοδήποτε άλλο στοιχείο της φυσικής γήινης επιφάνειας όπως υδάτινα στοιχεία ή δρόμοι (*other*).

Η εικόνα *Ground Truth* ή μάσκα του *Dataset* ορίζει την πραγματικότητα στην φυσική γήινη επιφάνεια (βλ. ενότητα 4.1). Η εικόνα *prediction* είναι η πρόβλεψη που προκύπτει από το μοντέλο βάσει της εκάστοτε δοκιμής εκπαίδευσης. Ζητούμενο αποτελεί η όσο το δυνατόν μεγαλύτερη ομοιότητα μεταξύ της προβλεπόμενης από το μοντέλο εικόνας (*prediction*) και της μάσκας (*ground Truth*). Δηλαδή, ζητούμενο είναι οι πραγματικές τιμές της μάσκας (*Ground Truth*) του *Dataset* να πλησιάζουν τις προβλεπόμενες τιμές του μοντέλου. Τα βέλτιστα αποτελέσματα υφίστανται όταν οι προβλεπόμενες τιμές του μοντέλου αντιστοιχούν στις πραγματικές τιμές της μάσκας (*Ground Truth*) του *Dataset*.

Τέλος, για σκοπούς σύγκρισης έγινε δοκιμή του μοντέλου «*Swin UNet*» με αρκετά μικρότερο training set, δηλαδή σε ποσοστό 30% που αντιστοιχεί περίπου σε ενενήντα (90) εικόνες εκπαίδευσης. Σύμφωνα με τις εικόνες που παρατίθενται ανωτέρω στην πέμπτη 5^η δοκιμή εκπαίδευσης, αλλά και τις μετρικές, προκύπτει το συμπέρασμα ότι η καλύτερη απόδοση του μοντέλου επιτυγχάνεται με μεγαλύτερο πλήθος δεδομένων εκπαίδευσης.

Παρατίθενται link με τον συνολικό κώδικα του μοντέλου όπως αυτός επεξεργάστηκε στην πλατφόρμα *Google Colab*:

https://colab.research.google.com/drive/1VWwt2UFtCjMbvXtkK3q9_riGlCqNWo3D#scrollTo=D5fpewX6cj-w

5.4 Μοντέλο Μηχανικής Μάθησης «*Image Segmentation Keras: Implementation of Segnet, FCN, UNet, PSPNet*»

Η εφαρμογή που παρουσιάζεται στην συγκεκριμένη ενότητα 4.3.3 αφορά την κατασκευή και εκπαίδευση ορισμένων *segmentation models*, με την χρήση της διεπαφής προγραμματισμού εφαρμογών, *Keras*. Δηλαδή, μοντέλων μηχανικής μάθησης κατάλληλα εκπαιδευμένων να εκτελούν σημασιολογική κατάτμηση. Η εφαρμογή έχει τίτλο «*image-segmentation-keras*». (<https://github.com/divamgupta/image-segmentation-keras/tree/master>) και θα χρησιμοποιηθεί ως κώδικας - βάση ώστε με κατάλληλη επεξεργασία στον αρχικό κώδικα να προκύπτουν οι επιθυμητές προβλέψεις.

Η βιβλιοθήκη *keras_segmentation* περιλαμβάνει ορισμένα έτοιμα προς χρήση μοντέλα (βλ. εικόνα 33) από τα οποία το μοντέλο *UNET* χρησιμοποιήθηκε στις κατωτέρω δοκιμές εκπαίδευσης και αξιολογήθηκε ως προς τα αποτελέσματα της στο *Dataset* της παρούσας εργασίας (βλ. ενότητα 4.1). Για την σωστή επιλογή του εκάστοτε «*segmentation model*» κρίνεται απαραίτητη πρώτα η κατάλληλη επιλογή του «*Base model*». Όπως φαίνεται και στην κατωτέρω εικόνα 33 τα προτεινόμενα στην παρούσα εφαρμογή *base models* είναι το *ResNet*, *VGG-16*, *MobileNet*, *Custom CNN* (π.χ *Vanilla CNN* & *Vanilla Mini CNN*), *FCN*, *SegNet*, *UNet* και *PSPNet*. Η επιλογή του κάθε μοντέλου γίνεται ανάλογα με τον στόχο της εκάστοτε εφαρμογής, όπως επίσης το πλήθος των δεδομένων εκπαίδευσης, το μέγεθος των εικόνων του συνόλου δεδομένων και την πληροφορία που περιλαμβάνουν οι εικόνες του συνόλου

δεδομένων. Συνήθως, τα μοντέλα σημασιολογικής κατάτμησης που βασίζονται σε βαθιά μάθηση κατασκευάζονται πάνω σε ένα βασικό δίκτυο Συνελκτικού Νευρωνικού Δικτύου (Convolutional Neural Networks) CNN. Συνήθως, επιλέγεται ένα τυπικό μοντέλο όπως *ResNet*, *VGG* ή *MobileNet* για το βασικό δίκτυο. Ορισμένα αρχικά επίπεδα του βασικού Μοντέλου (*base model*) χρησιμοποιούνται στον κωδικοποιητή (*encoder*) και το υπόλοιπο δίκτυο κατάτμησης είναι χτισμένο πάνω από αυτό. Για τα περισσότερα μοντέλα σημασιολογικής κατάτμησης, μπορεί να χρησιμοποιηθεί οποιοδήποτε βασικό μοντέλο.

model_name	Base Model	Segmentation Model
fcn_8	Vanilla CNN	FCN8
fcn_32	Vanilla CNN	FCN8
fcn_8_vgg	VGG 16	FCN8
fcn_32_vgg	VGG 16	FCN32
fcn_8_resnet50	Resnet-50	FCN32
fcn_32_resnet50	Resnet-50	FCN32
fcn_8_mobilenet	MobileNet	FCN32
fcn_32_mobilenet	MobileNet	FCN32
pspnet	Vanilla CNN	PSPNet
pspnet_50	Vanilla CNN	PSPNet
pspnet_101	Vanilla CNN	PSPNet
vgg_pspnet	VGG 16	PSPNet
resnet50_pspnet	Resnet-50	PSPNet
unet_mini	Vanilla Mini CNN	U-Net
unet	Vanilla CNN	U-Net
vgg_unet	VGG 16	U-Net
resnet50_unet	Resnet-50	U-Net
mobilenet_unet	MobileNet	U-Net
segnet	Vanilla CNN	Segnet
vgg_segnet	VGG 16	Segnet
resnet50_segnet	Resnet-50	Segnet
mobilenet_segnet	MobileNet	Segnet

Εικόνα 18 - Keras_segmentation μοντέλα

Στο Παράρτημα κώδικα 9 – *Segmentation model & Model Train* κατωτέρω εμφανίζεται το μοντέλο που επιλέχθηκε για δοκιμή από την βιβλιοθήκη *keras_segmentation.model* και η εντολή που χρησιμοποιήθηκε για την διαδικασία της εκπαίδευσης του μοντέλου. Επίσης, αναγράφονται οι γραμμές κώδικα που αφορούν την εκπαίδευση του μοντέλου στις οποίες το στοιχείο που αλλάζει σε κάθε δοκιμή είναι οι εποχές (*epochs*) εκπαίδευσης.

Επιπλέον, στις γραμμές κώδικα που αφορούν την κατηγοριοποίηση των δεδομένων σε δεδομένα εκπαίδευσης (*train set*) και δεδομένα ελέγχου (*test set*) προστέθηκε η εντολή *np.random.seed(15)* για την επιλογή της σταθερής τυχαιότητας των δεδομένων.

Παράρτημα κώδικα 9 – Segmentation model & Model Train

```
from keras_segmentation.models.unet import unet

model = unet(n_classes=3 , input_height=512, input_width=512)

model.train(
    train_images = "/content/train_images/",
    train_annotations = "/content/train_masks/",
    checkpoints_path = "/tmp/vgg_unet_1" , epochs=*
```

Παρατίθενται link με τον συνολικό κώδικα του μοντέλου όπως αυτός επεξεργάστηκε στην πλατφόρμα Google Colab:

<https://colab.research.google.com/drive/11HC9hfCH1yjAO3lXBa8Kmog5vQOnB4J5>

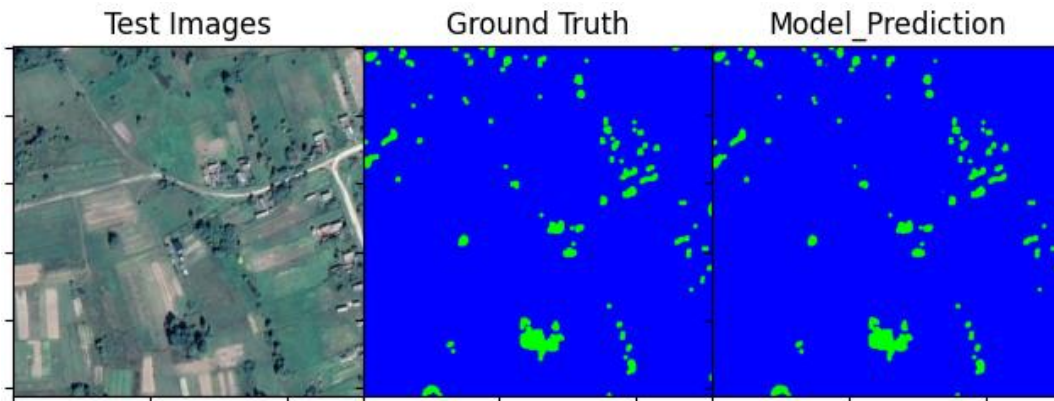
Κεφάλαιο 6 - Αποτελέσματα

6.1 Προεκπαιδευμένο Μοντέλο «Model_Testing» (Pre-Trained Model)

Για τις κατωτέρω εικόνες 19, 20, 21, 22, 23 που απεικονίζονται κατά σειρά η δορυφορική εικόνα (Test Images), η μάσκα της αντίστοιχης δορυφορικής εικόνας (Ground Truth) και η προβλεπόμενη εικόνα του μοντέλου (Model Prediction) παρατηρείται ότι η μάσκα είναι ίδια με την αντίστοιχη πρόβλεψη του μοντέλου. Αυτό συμβαίνει διότι στο προεκπαιδευμένο μοντέλο «Model_Testing» δεν είναι εφικτό να γίνει διαχωρισμός των δεδομένων σε πακέτο εικόνων εκπαίδευσης (train set), πακέτο εικόνων ελέγχου (test set) και πακέτο εικόνων αξιολόγησης (validate set) καθώς δεν πραγματοποιείται διαδικασία εκπαίδευσης στο προεκπαιδευμένο μοντέλο (Pre-Trained Model) εφόσον λαμβάνεται έτοιμο αρχείο βαρών (U6_E_1201-F1_0.7134-IOU_0.6555.h5). Επιπλέον, τα δεδομένα που χρησιμοποιήθηκαν στο συγκεκριμένο προεκπαιδευμένο μοντέλο «Model_Testing» αποτελούν τα δεδομένα που πάρθηκαν για την εκπαίδευση του όπως αυτό εμφανίζεται στο αναλυτικό μοντέλο «Model_Training_TPU_Strategy». Αναλυτικά, η διαδικασία εκπαίδευσης του μοντέλου και ο διαχωρισμός των δεδομένων όπως αναφέρθηκε ανωτέρω προκειμένου να εξεταστεί η ορθή λειτουργία του μοντέλου και το ποσοστό συμβατότητας της προβλεπόμενης από το μοντέλο εικόνας (Model Prediction) με την αντίστοιχη μάσκα (Ground Truth) πραγματοποιείται στο αντίστοιχο μοντέλο «Model_Training_TPU_Strategy».

Όσον αφορά τα αποτελέσματα των μετρικών «F1_Score_forest», «F1_Score_deforest» & «F1_Score_other» και «iou_forest», «iou_deforest» & «iou_other» όπως φαίνονται παρακάτω για κάθε μια από τις εικόνες 19, 20, 21, 22, 23 προκύπτουν τιμές οι οποίες τείνουν στο ποσοστό 100%. Ο λόγος για τον οποίο συμβαίνει αυτό εξηγήθηκε στην προηγούμενη παράγραφο και αυτές οι τιμές ήταν αναμενόμενες.

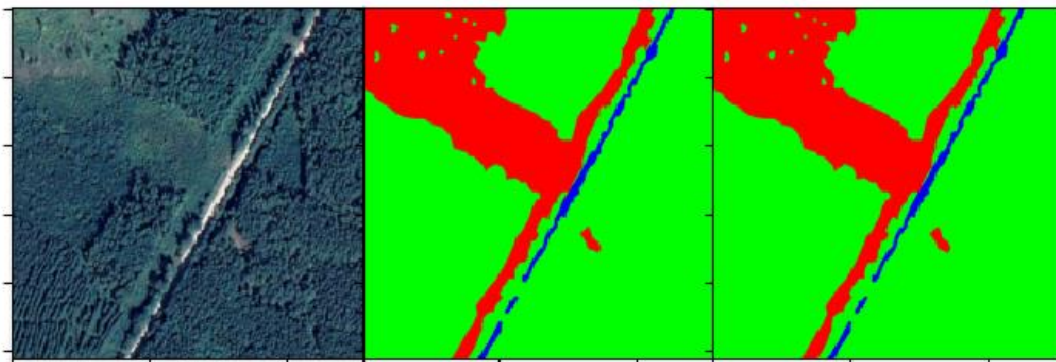
Συμπερασματικά, η επεξεργασία που πραγματοποιήθηκε στο προεκπαιδευμένο Μοντέλο «Model_Testing» (Pre-Trained Model), κυρίως στον κώδικα του μοντέλου, είχε στόχο την βαθιά κατανόηση ως προς την λειτουργία ενός νευρωνικού δικτύου, την διαφορά μεταξύ προεκπαιδευμένου μοντέλου (Model_Testing) με αναλυτικό μοντέλο (Model_Training_TPU_Strategy) και την εξοικείωση με τον κώδικα ώστε να αποτελέσει την βάση για την επεξεργασία των μοντέλων που παρουσιάζονται παρακάτω στις ενότητες 5.3 και 5.4. Τα αποτελέσματα που παρουσιάζονται παρακάτω δεν είναι σωστά είναι «πλασματικά» καθώς στο προεκπαιδευμένο μοντέλο χρησιμοποιήθηκαν τα δεδομένα στα οποία έχει εκπαιδευτεί. Ωστόσο, αναφέρονται στην παρούσα διπλωματική εργασία καθώς πραγματοποιήθηκε επεξεργασία στο κώδικα ώστε να λειτουργεί ορθά και να εμφανίζονται τα αποτελέσματα όπως εμφανίζονται στις κατωτέρω εικόνες. Στην περίπτωση βέλτιστης λειτουργίας του εκάστοτε μοντέλου τα αποτελέσματα θα έπρεπε να εμφανίζονται κατά αυτόν τον τρόπο.



Εικόνα 19 - Pre - trained Model - Test Images, Ground truth, Model_Prediction [1/22]

[INFO] Image 1/22, model0.

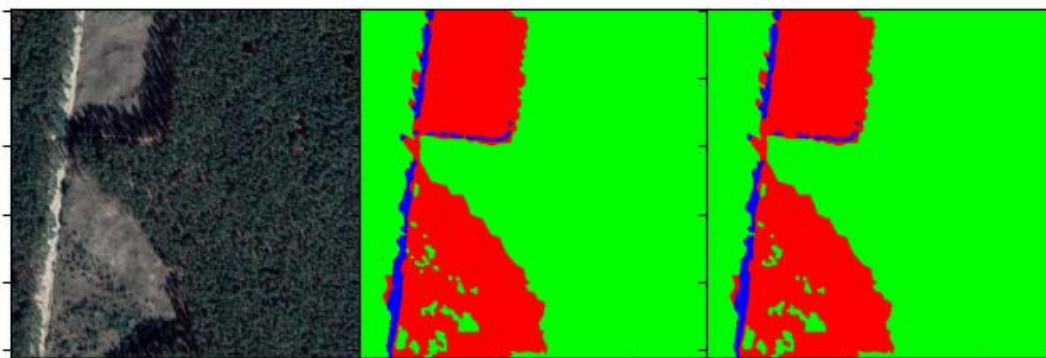
- F1_Score_forest: 0.9565525054931641
- F1_Score_deforest: 0.9999990463256836
- F1_Score_other: 0.9983355402946472
- iou_forest: 0.9167231321334839
- iou_deforest: 0.9999990463256836
- iou_other: 0.9966765642166138



Εικόνα 20 - Pre - trained Model - Test Images, Ground truth, Model_Prediction [3/22]

[INFO] Image 3/22, model0.

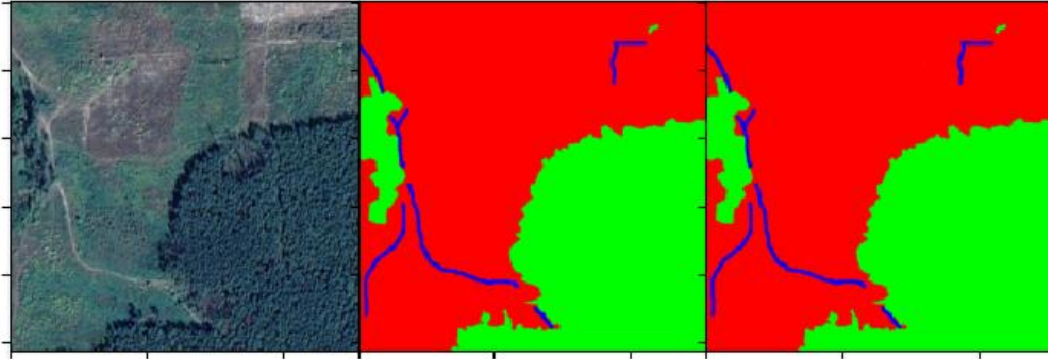
- F1_Score_forest: 0.9992837309837341
- F1_Score_deforest: 0.9977229237556458
- F1_Score_other: 0.9956244826316833
- iou_forest: 0.9985685348510742
- iou_deforest: 0.9954562187194824
- iou_other: 0.9912870526313782



Εικόνα 21 - Pre - trained Model - Test Images, Ground truth, Model_Prediction [7/22]

[INFO] Image 7/22, model0.

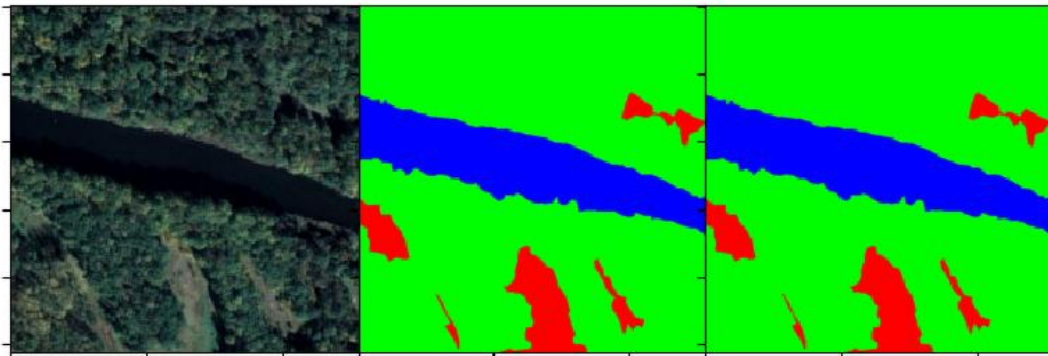
- F1_Score_forest: 0.9979519844055176
- F1_Score_deforest: 0.9986457228660583
- F1_Score_other: 0.9928860664367676
- iou_forest: 0.9959123134613037
- iou_deforest: 0.9972951412200928
- iou_other: 0.9858726263046265



Εικόνα 22 - Pre - trained Model - Test Images, Ground truth, Model_Prediction [10/22]

[INFO] Image 10/22, model0.

- F1_Score_forest: 0.9988346099853516
- F1_Score_deforest: 0.9955765008926392
- F1_Score_other: 0.9856393933296204
- iou_forest: 0.9976719617843628
- iou_deforest: 0.9911919236183167
- iou_other: 0.9716853499412537



Εικόνα 23 - Pre - trained Model - Test Images, Ground truth, Model_Prediction [12/22]

[INFO] Image 12/22, model0.

- F1_Score_forest: 0.9993254542350769
- F1_Score_deforest: 0.9979701638221741
- F1_Score_other: 0.9975786805152893
- iou_forest: 0.9986518025398254
- iou_deforest: 0.9959485530853271
- iou_other: 0.9951691031455994

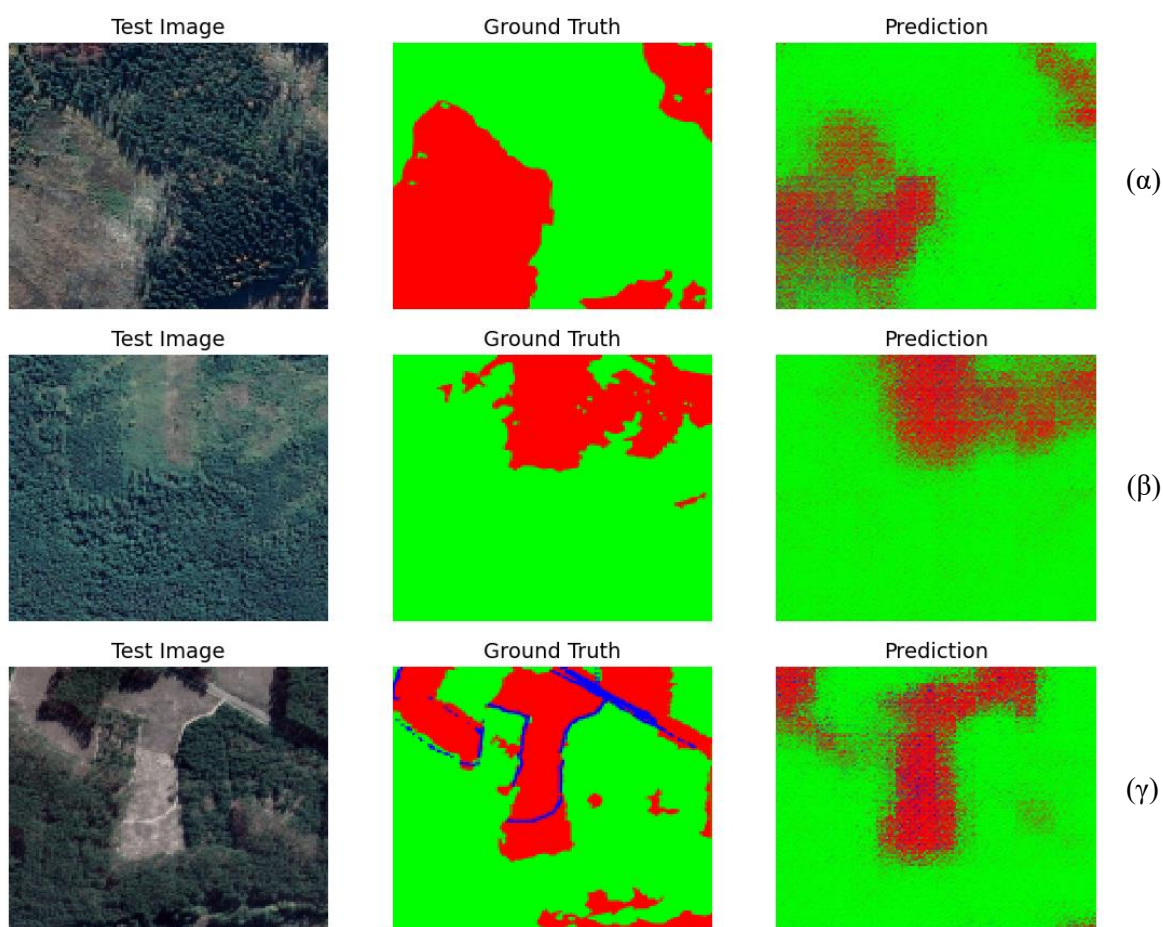
6.2 Μοντέλο Μηχανικής Μάθησης «Swin – UNet»

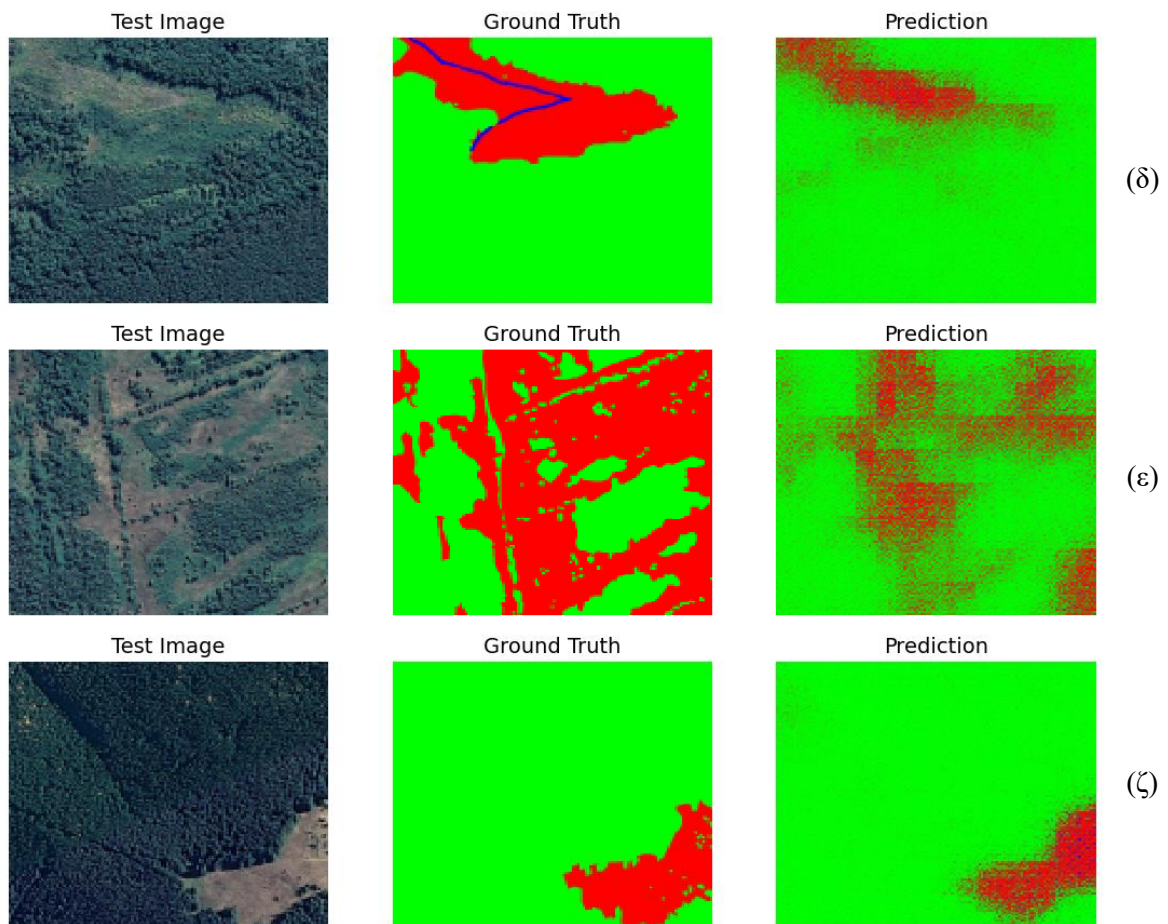
Πρώτη 1^η Δοκιμή Εκπαίδευσης

Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Δέκα εποχές ($N_{epoch}=10$)
- $Batches=8$ per epoch ($N_{batch}=8$)
- $Samples=32$ per batch ($N_{sample}=32$)

Η εκπαίδευση διήρκησε είκοσι ένα (21) λεπτά και στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (*Prediction*) σε σχέση με την μάσκα (*Ground Truth*) του *Dataset*. Κατά την διάρκεια της εκπαίδευσης το «*validation performance*» από τιμή 2,572 μειώθηκε στη τιμή 0,7136 (71,36%), ενώ αντίστοιχα προκύπτει τιμή «*cross-entropy loss*» ίση με 0,4123 (41,23%), *Testing set cross-entropy loss* = 0.4123271107673645 (41,23%).





Εικόνα 24 - Test image, Ground Truth, Prediction – 1η δοκιμή εκπαίδευσης

Στην συγκεκριμένη δοκιμή εκπαίδευσης ορίστηκαν δέκα (10) εποχές εκπαίδευσης, Batches=8 και Samples=32. Παρατηρούμε από τις ανωτέρω εικόνες ότι το μοντέλο έχει ανιχνεύσει σε κάθε μια από αυτές που βρίσκεται στην εικόνα η αποψιλωμένη έκταση. Ωστόσο, η ακρίβεια ανίχνευσης δεν είναι η βέλτιστη που αναμένεται να παραχθεί. Δηλαδή, η εικόνα Prediction να είναι όσο το δυνατόν γίνεται παρόμοια με την (Ground Truth) μάσκα. Πιο συγκεκριμένα, στην εικόνα 24 (γ) στο Prediction φαίνεται το μοντέλο να έχει ανιχνεύσει την συγκεντρωτική αποψιλωμένη έκταση, ενώ παράλληλα δεν έχει καταφέρει να εντοπίσει μικρότερη αποψιλωμένη έκταση που εμφανίζεται στην εικόνα. Αντίστοιχα, στην εικόνα 24 (ε) που η αποψιλωμένη έκταση υφίσταται σε μεγαλύτερη πολυπλοκότητα σε σχέση με τις υπόλοιπες εικόνες φαίνεται ότι το μοντέλο “δυσκολεύτηκε” να ανιχνεύσει με ακρίβεια την θέση της αποψιλωμένης έκτασης στην εικόνα. Συνεπώς, στις περιπτώσεις που η αποψιλωμένη έκταση εμφανίζεται στην εικόνα συγκεντρωτικά σε μια ορισμένη θέση το μοντέλο έχει σαφώς καλύτερα αποτελέσματα ανίχνευσης σε σχέση με την εικόνα που η αποψιλωμένη έκταση εμφανίζεται στην εικόνα με πολυπλοκότητα και “δυσκολεύει” το μοντέλο να κατανοήσει την ακριβή θέση της αποψιλωμένης έκτασης.

Συμπερασματικά, με τις παραμέτρους που επιλέχθηκαν στην πρώτη (1^η) δοκιμή εκπαίδευσης παρατηρούνται μέτρια αποτελέσματα ανίχνευσης της αποψιλωμένης έκτασης από το μοντέλο χωρίς να έχει την δυνατότητα να αναγνωρίσει με λεπτομέρεια τα ακριβή όρια της αποψιλωμένης έκτασης. Ωστόσο, παρατηρούνται καλά αποτελέσματα όσον αφορά την ανίχνευση της θέσης της αποψιλωμένης έκτασης που αυτό επιτυγχάνεται σε όλες τις εικόνες. Επιπλέον, παρατηρείται ότι το μοντέλο έχει πολύ καλά αποτελέσματα στην ανίχνευση της δασικής έκτασης και αυτό συμβαίνει διότι η δασική έκταση εμφανίζεται στα δεδομένα σε πολύ μεγαλύτερο ποσοστό από την αποψιλωμένη έκταση και την κατηγορία «Other» και το μοντέλο έχει την δυνατότητα να εκπαιδευτεί καλύτερα στην κατηγορία «forest». Τέλος, όσον

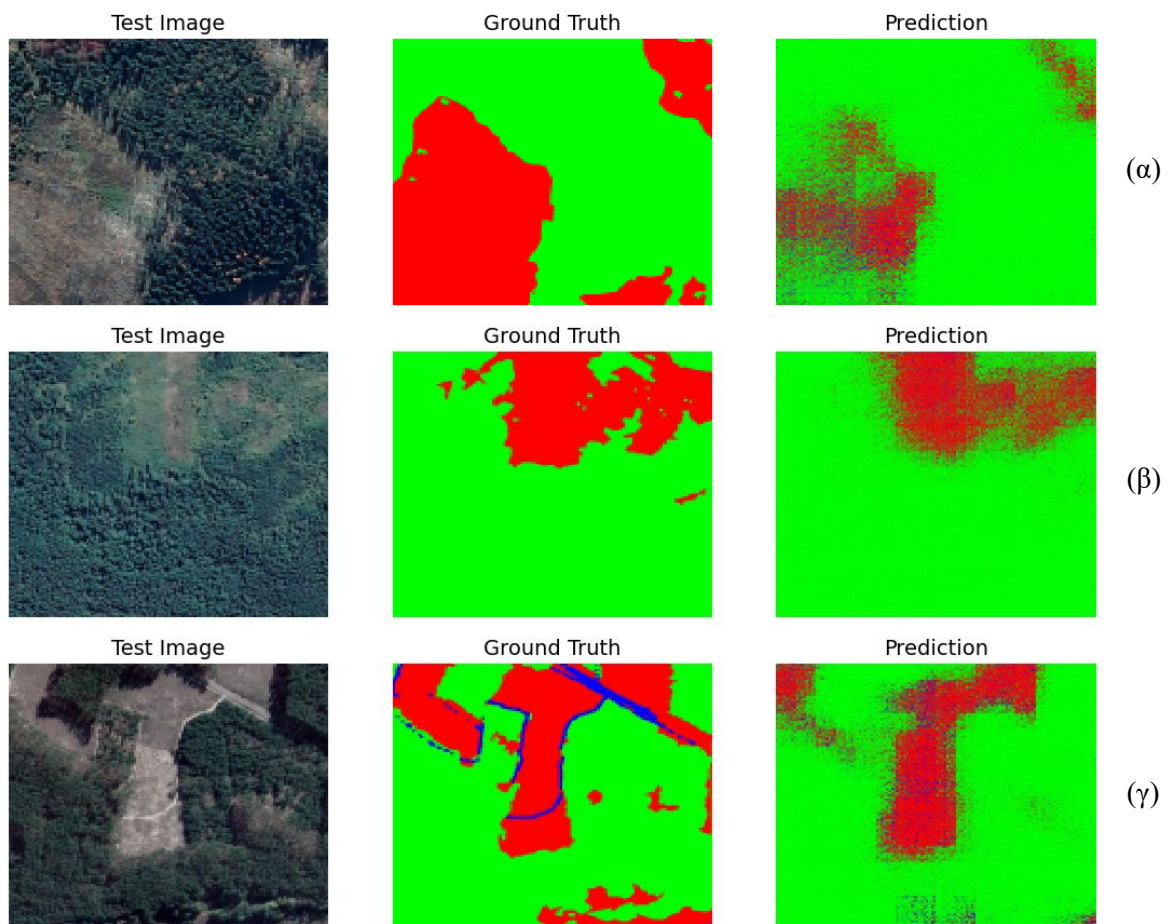
αφορά την κατηγορία *other* το μοντέλο δεν την ανιχνεύει με ακρίβεια και αποτελεί την κατηγορία που το μοντέλο δυσκολεύεται αρκετά να ανιχνεύσει.

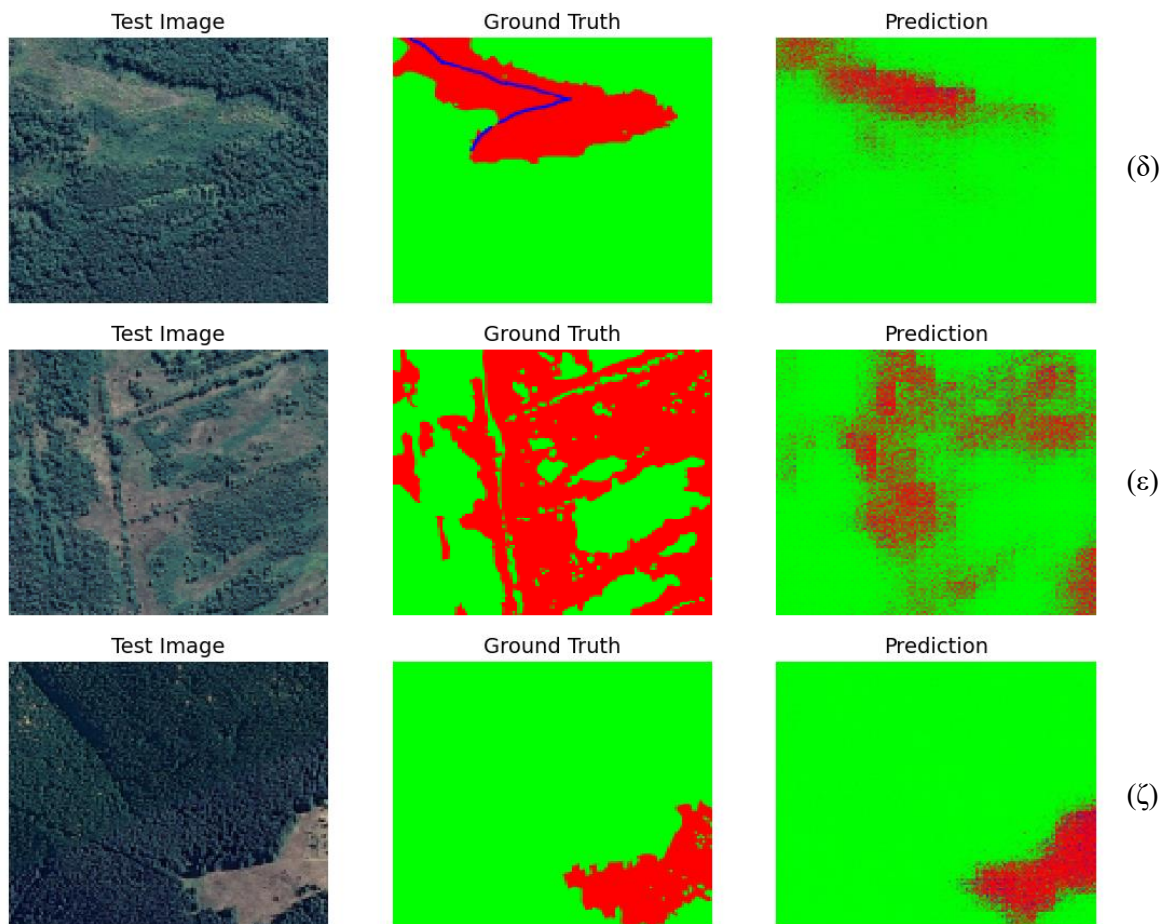
Δεύτερη 2^η Δοκιμή Εκπαίδευσης

Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Είκοσι εποχές ($N_{epoch}=20$)
- $Batches=8$ per epoch ($N_{batch}=8$)
- $Samples=32$ per batch ($N_{sample}=32$)

Η εκπαίδευση διήρκησε τριάντα 33 λεπτά με *early stopping* στις δέκα έξι εποχές προκειμένου να αποφευχθεί το *overfitting* και οι προβλέψεις του μοντέλου να μην είναι αντιπροσωπευτικές. Στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (*Prediction*) σε σχέση με την μάσκα (*Ground Truth*) του *Dataset*. Κατά την διάρκεια της εκπαίδευσης το «*validation performance*» από τιμή 2,75 μειώθηκε στη τιμή 0,5716 (57,16%), ενώ αντίστοιχα προκύπτει τιμή «*cross-entropy loss*» ίση με 0,3671 (36,71%), *Testing set cross-entropy loss* = 0.36710143089294434.





Εικόνα 25 - Test image, Ground Truth, Prediction - 2η δοκιμή εκπαίδευσης

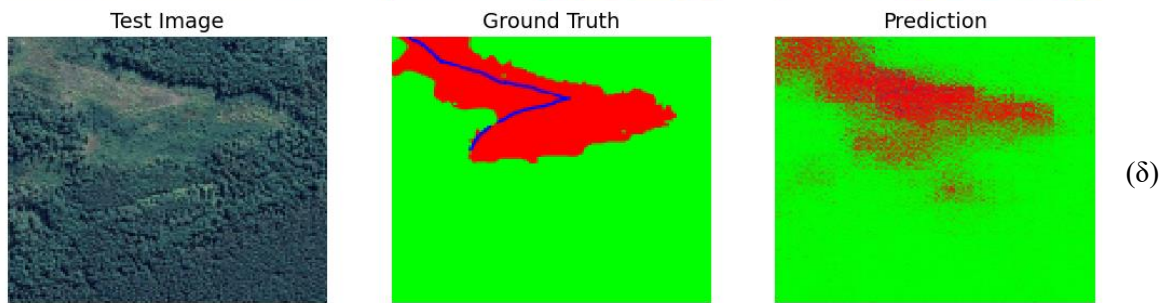
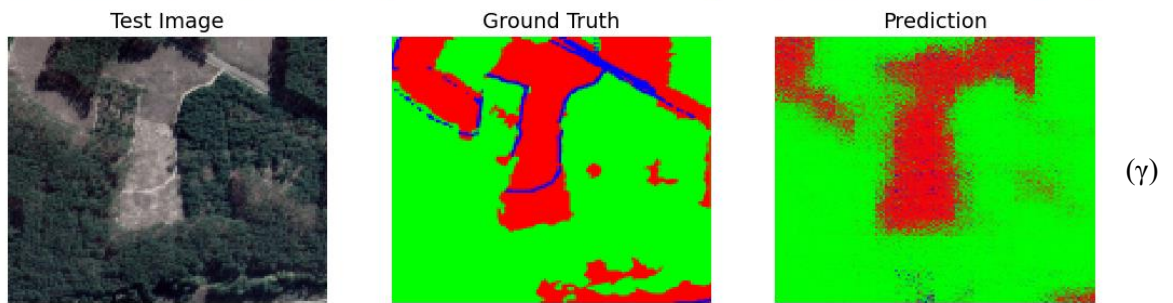
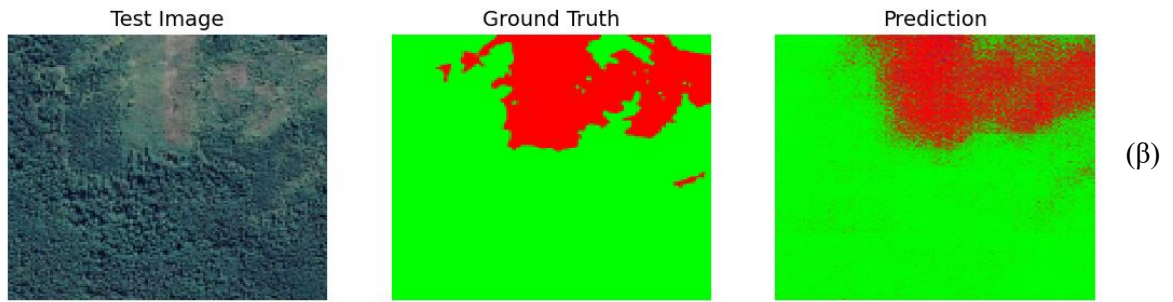
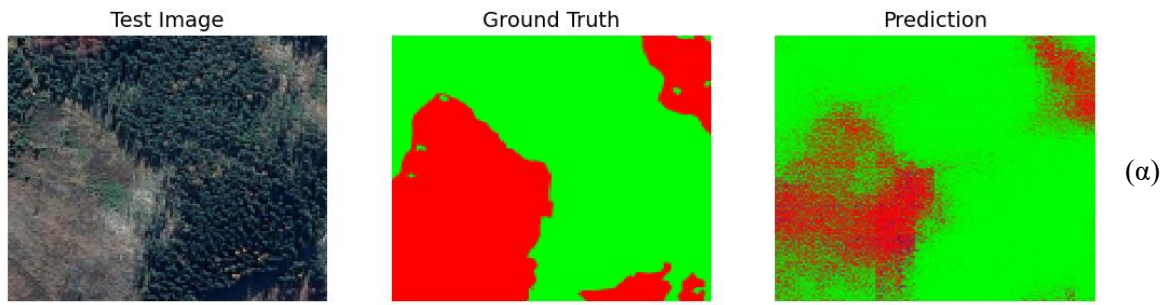
Στη συγκεκριμένη δοκιμή εκπαίδευσης ορίστηκαν είκοσι (20) εποχές εκπαίδευσης, Batches=8 και Samples=32. Σε σχέση με την πρώτη 1η δοκιμή εκπαίδευσης όπου είχαν οριστεί δέκα (10) εποχές εκπαίδευσης δεν παρατηρείται μεγάλη διαφορά στην ακρίβεια ανίχνευσης των αποψιλωμένων εκτάσεων. Τα Prediction από το μοντέλο έχουν παρουσιάσει μικρή σχετικά βελτίωση σε σχέση με τις δέκα εποχές εκπαίδευσης. Συγκεκριμένα, φαίνεται ότι το μοντέλο έχει εντοπίσει σε λίγο καλύτερο ποσοστό τις αποψιλωμένες εκτάσεις και τις έχει πυκνώσει σε καλύτερο βαθμό.

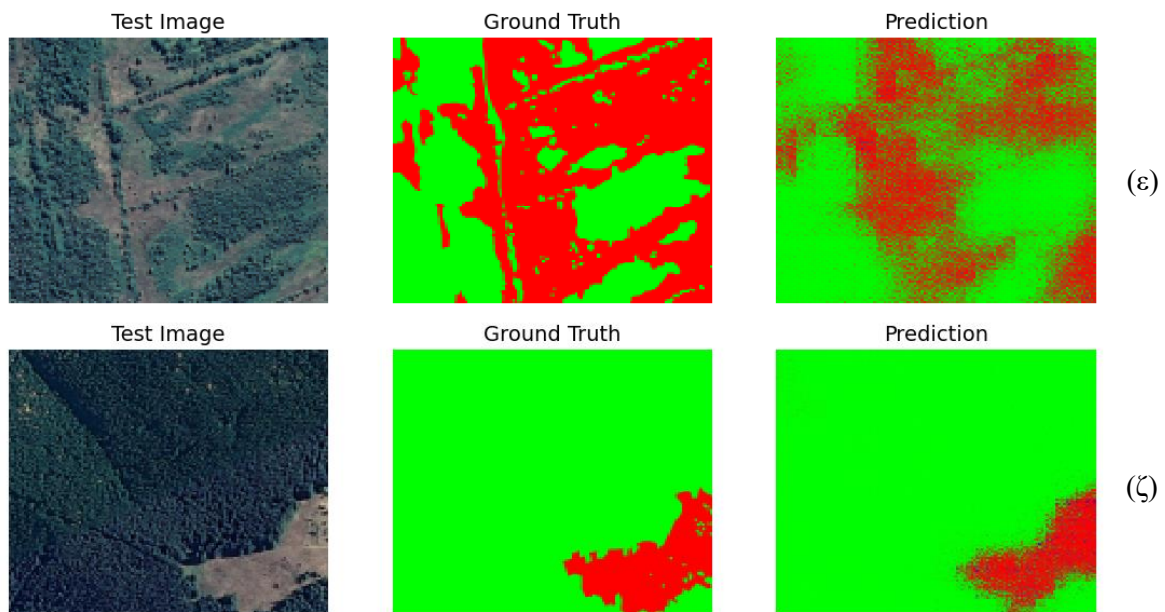
Τρίτη 3^η Δοκιμή Εκπαίδευσης

Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Τριάντα εποχές ($N_{epoch}=30$)
- Batches=10 per epoch ($N_{batch}=10$)
- Samples=32 per batch ($N_{sample}=32$)

Η εκπαίδευση διήρκεσε πενήντα (50) λεπτά με early stopping στις δέκα πέντε εποχές προκειμένου να αποφευχθεί το overfitting και οι προβλέψεις του μοντέλου να μην είναι αντιπροσωπευτικές. Στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (Prediction) σε σχέση με την μάσκα (Ground Truth) του Dataset. Κατά την διάρκεια της εκπαίδευσης το «validation performance» από τιμή 2,556 μειώθηκε στη τιμή 0,5115 (51,15%), ενώ αντίστοιχα προκύπτει τιμή «cross-entropy loss» ίση με 0,2809 (28,09%), $Testing\ set\ cross-entropy\ loss = 0.2809394299983978$.





Εικόνα 26 - Test image, Ground Truth, Prediction - 3η δοκιμή εκπαίδευσης

Στην τρίτη 3^η δοκιμή εκπαίδευσης έγινε δοκιμή αλλαγής των παραμέτρων εκπαίδευσης Batches=10 και των εποχών εκπαίδευσης σε τριάντα (30) και παρατηρήθηκε ότι το μοντέλο εκπαιδεύτηκε σε καλύτερο βαθμό και ανίχνευσε με καλύτερη ακρίβεια τις περιοχές αποψίλωσης. Συγκεκριμένα, παρατηρείται στις εικόνες 24 (γ), 24 (δ) και 24 (ε) βελτίωση στον εντοπισμό της αποψιλωμένης έκτασης, ειδικότερα στην εικόνα 24 (ε) που έχει και μεγαλύτερο ποσοστό πολυπλοκότητας φαίνεται ότι το μοντέλο έχει εκπαιδευτεί σε καλύτερο βαθμό εφόσον εντοπίζει μεγαλύτερη έκταση αποψιλωμένης περιοχής στην εικόνα.

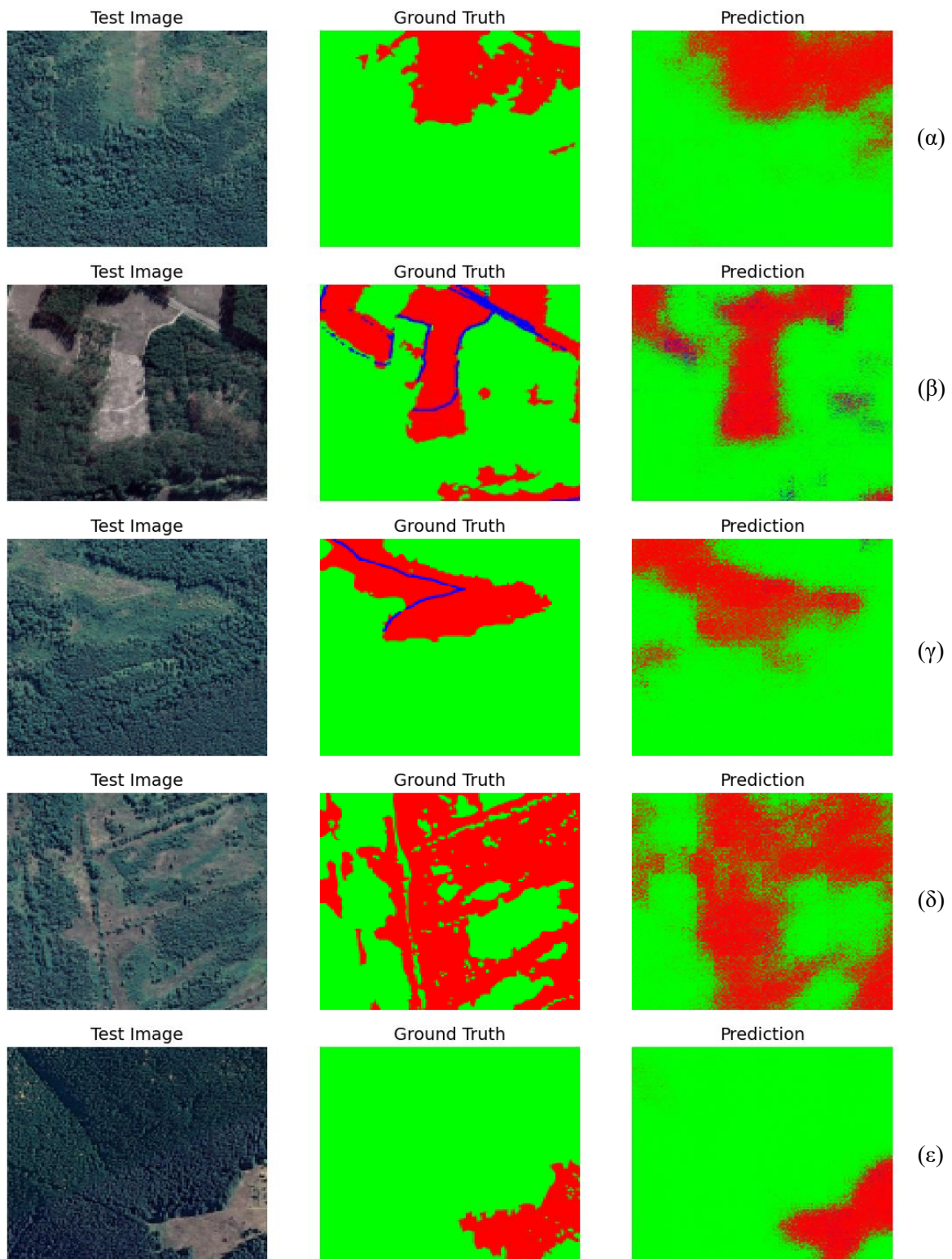
Από τα ανωτέρω σχόλια προκύπτει το συμπέρασμα ότι η καλύτερη δυνατή εκπαίδευση του μοντέλου επιτυγχάνεται όταν έχει επιλεγθεί κατάλληλος αριθμός εποχών εκπαίδευσης και παραμέτρων batches και samples. Όπως έχει ήδη αναφερθεί ανωτέρω το γινόμενο των batches και samples επιλέγεται σύμφωνα με το πλήθος των δεδομένων, δηλαδή ορίζεται από τα δεδομένα. Συγκεκριμένα, το γινόμενο των batches και samples πρέπει να είναι όσο το πλήθος των δεδομένων ή να μην το ξεπερνά σε μεγάλο ποσοστό.

Τέταρτη 4^η Δοκιμή Εκπαίδευσης

Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Δέκα οκτώ εποχές ($N_{epoch}=18$)
- Batches=12 per epoch ($N_{batch}=12$)
- Samples=42 per batch ($N_{sample}=42$)

Η εκπαίδευση διήρκησε μία ώρα και στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (Prediction) σε σχέση με την μάσκα (Ground Truth) του «Dataset». Κατά την διάρκεια της εκπαίδευσης το «validation performance» από τιμή 2,526 μειώθηκε στη τιμή 0,3646 (36,46%), ενώ αντίστοιχα προκύπτει τιμή «cross-entropy loss» ίση με 0,2387 (23,87%), $Testing\ set\ cross-entropy\ loss = 0.23873858153820038$.



Εικόνα 27 - Test image, Ground Truth, Prediction - 4η δοκιμή εκπαίδευσης

Εφόσον στην προηγούμενη δοκιμή εκπαίδευσης με την αλλαγή της παραμέτρου *batches* παρατηρήθηκαν καλύτερα αποτελέσματα, στην τέταρτη 4η δοκιμή εκπαίδευσης επιλέχθηκαν *batches=12*, *samples=42* και δέκα οκτώ εποχές (18) εποχές εκπαίδευσης. Παρατηρείται ότι το μεγαλύτερο μέρος της αποψιλωμένης έκτασης στις εικόνες έχει ανιχνευθεί και έχει οριστεί με αρκετά καλή ακρίβεια. Εμφανής είναι η διαφορά στις εικόνες 27 (γ) και 27 (δ) συγκριτικά με τις προηγούμενες εποχές εκπαίδευσης όπου οι εκτάσεις

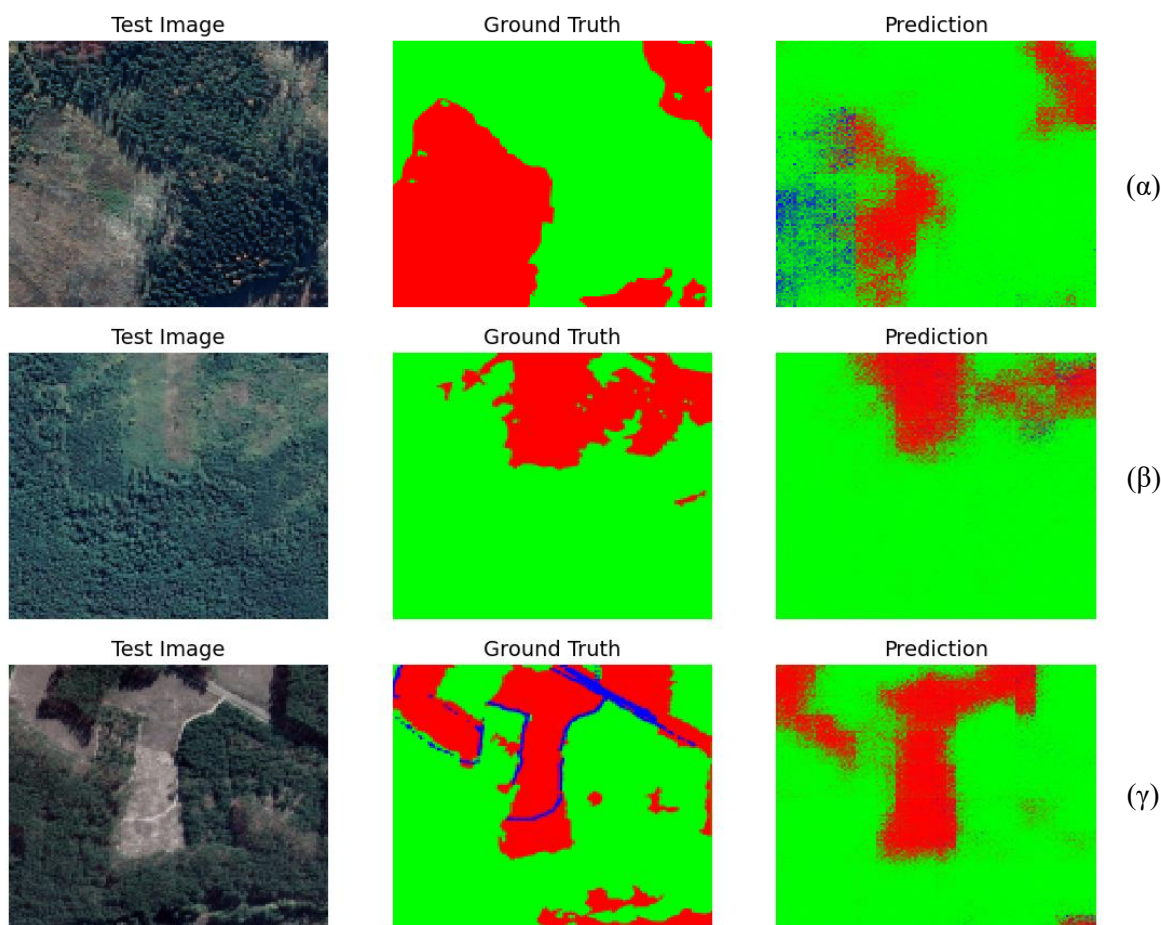
αποψίλωσης έχουν εντοπιστεί σε αρκετά καλό βαθμό και έχουν οριστεί στην εικόνα Prediction σε αρκετά καλό ποσοστό. Συνεπώς, προκύπτει το συμπέρασμα ότι οι παράμετροι εκπαίδευσης batches και samples είναι αρκετά σημαντική για την εκπαίδευση του μοντέλου και την αποτελεσματικότητα του στην ανίχνευση της ζητούμενης κατηγορίας στην εικόνα. Όσο μεγαλύτερη είναι η τιμή των παραμέτρων batches και samples αλλά και οι εποχές εκπαίδευσης τόσο καλύτερη η εκπαίδευση του μοντέλου κάτι το οποίο συνεπάγεται και καλύτερο αποτέλεσμα κατάτμησης.

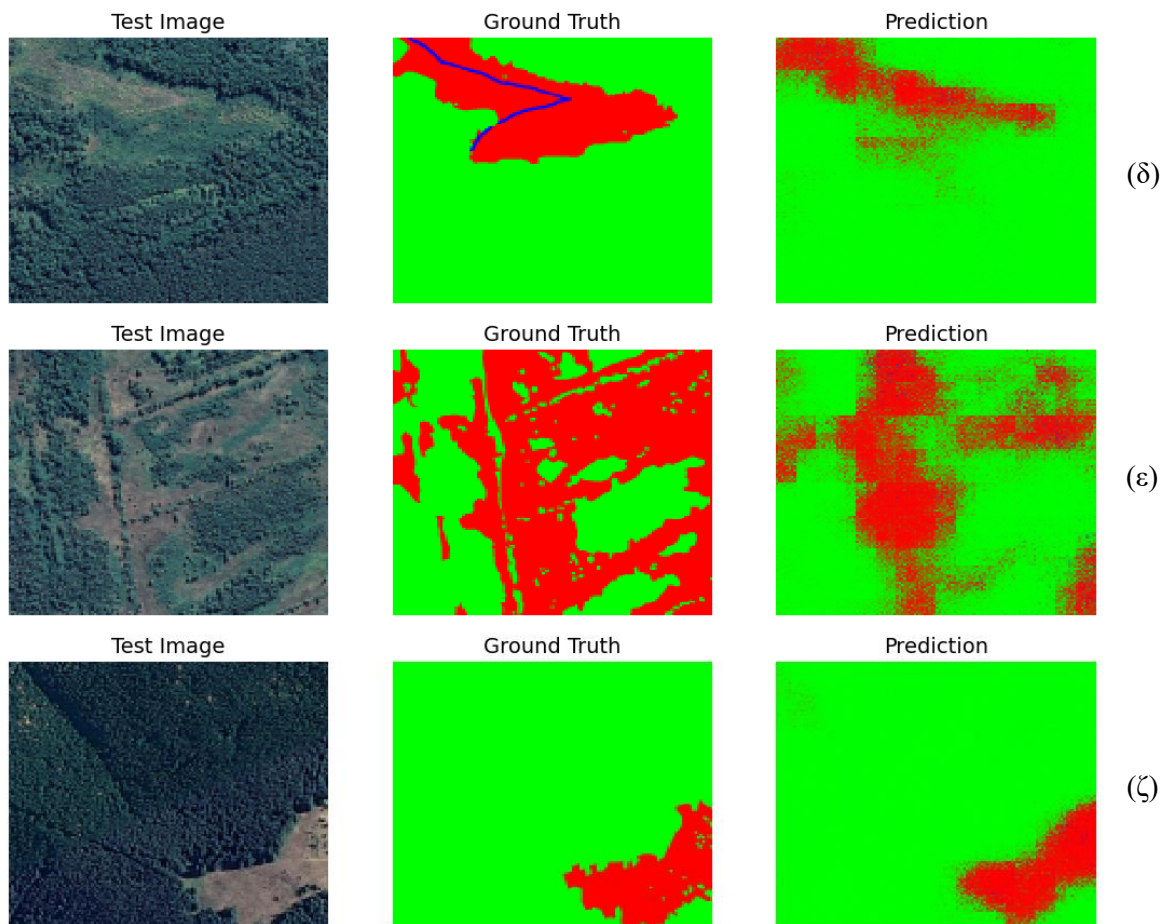
Πέμπτη 5^η Δοκιμή Εκπαίδευσης

Για σκοπούς σύγκρισης έγινε δοκιμή του συγκεκριμένου μοντέλου με αρκετά μικρότερο training set, δηλαδή σε ποσοστό 30% που αντιστοιχεί περίπου σε ενενήντα (90) εικόνες εκπαίδευσης. Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Δέκα εποχές ($N_{epoch}=10$)
- Batches=12 per epoch ($N_{batch}=12$)
- Samples=42 per batch ($N_{sample}=42$)

Η εκπαίδευση διήρκησε σαράντα (43) λεπτά και στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (Prediction) σε σχέση με την μάσκα (Ground Truth) του «Dataset». Κατά την διάρκεια της εκπαίδευσης το «validation performance» από τιμή 2,29 μειώθηκε στη τιμή 0,4214 (42,14%), ενώ αντίστοιχα προκύπτει τιμή «cross-entropy loss» ίση με (42,17%), Testing set cross-entropy loss = 0.42168164253234863.





Εικόνα 28 - Test image, Ground Truth, Prediction - 5η δοκιμή εκπαίδευσης

Η συγκεκριμένη δοκιμή εκπαίδευσης πραγματοποιήθηκε για σκοπούς σύγκρισης της λειτουργικότητας του μοντέλου για πολύ μικρότερο αριθμό δεδομένων εκπαίδευσης (training set). Πάρθηκε training set περί το 30% των δεδομένων ενώ τα υπόλοιπα δεδομένα χρησιμοποιήθηκαν ως δεδομένα ελέγχου και αξιολόγησης. Παρατηρείται ότι τα αποτελέσματα δεν είναι ικανοποιητικά και ότι το μοντέλο έχει δυσκολευτεί να ανιχνεύσει σε καλό ποσοστό τις αποψιλωμένες εκτάσεις ακόμα και με τον καλύτερο συνδυασμό batches=12 και samples=42 που έχει χρησιμοποιηθεί στις ανωτέρω δοκιμές εκπαίδευσης. Ειδικότερα αυτό εμφανίζεται στις εικόνες 28 (α) και 28 (ε). Συνεπώς, προκύπτει το συμπέρασμα ότι όσο περισσότερα είναι τα δεδομένα εκπαίδευσης δίνεται η ευκαιρία στο μοντέλο να εκπαιδευτεί καλύτερα στις κατηγορίες που πρέπει να ανιχνευθούν και να προκύψει μια αξιόπιστη από το μοντέλο πρόβλεψη.

6.2.1 Πίνακες Αποτελεσμάτων μετρικών F1 – Score & Intersection over Union

Training set: 80%

	Ten (10) epochs Batches=8, Samples=32	Twenty (20) epochs Batches=8, Samples=32	Thirty (30) epochs Batches=10, Samples=32	Eighteen (18) epochs Batches=12, Samples=42
F1-score for Deforest	0.418319466	0.418018765	0.595695451	0.671358039
F1-score for Forest	0.931481919	0.936404374	0.948610636	0.960351047
F1-score for Other	0.162964056	0.536877516	0.684280007	0.842845784

Training set: 30%

	Ten (10) epochs Batches=12, Samples=42	Ten (10) epochs Batches=12, Samples=42
F1-score for Deforest	0.558717322	IOU-score for Deforest 0.387652839
F1-score for Forest	0.938030962	IOU-score for Forest 0.883294078
F1-score for Other	0.689384461	IOU-score for Other 0.526000524

Training set: 80%

	Ten (10) epochs Batches=8, Samples=32	Twenty (20) epochs Batches=8, Samples=32	Thirty (30) epochs Batches=10, Samples=32	Eighteen (18) epochs Batches=12, Samples=42
IOU-score for Deforest	0.264477849	0.264237499	0.424192495	0.505296429
IOU-score for Forest	0.8717512	0.8804139	0.902244847	0.923726267
IOU-score for Other	0.088710325	0.366939557	0.520080268	0.728378095

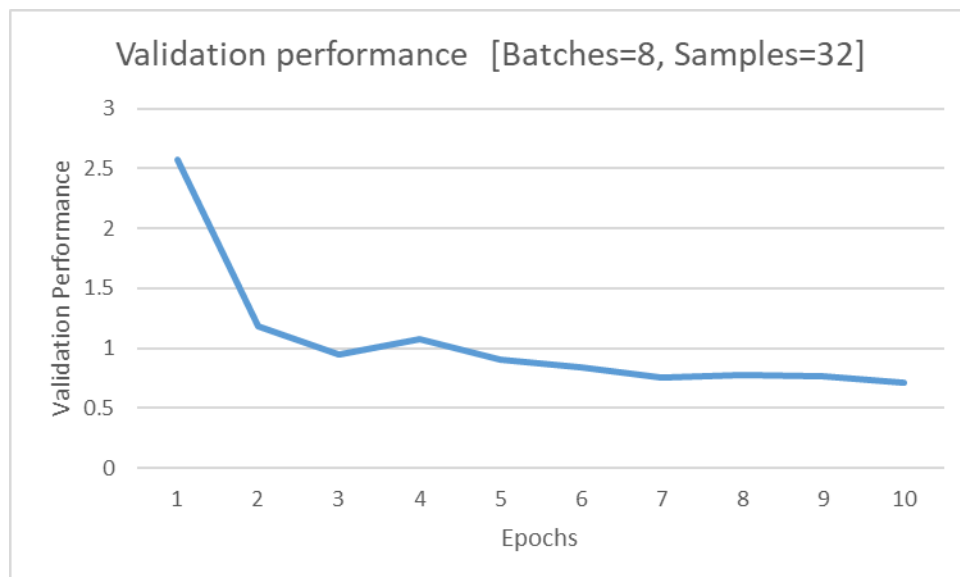
Παρατηρώντας τις παραπάνω εικόνες με τις προβλέψεις του μοντέλου σε κάθε μια από τις τέσσερις δοκιμές εκπαίδευσης προκύπτει το συμπέρασμα ότι ο κατάλληλος συνδυασμός των εποχών εκπαίδευσης με το πλήθος των batches και samples οδηγούν σε περισσότερο αξιόπιστα αποτελέσματα. Η καλή εκπαίδευση του μοντέλου οδηγεί και σε ακριβέστερη πρόβλεψη. Συγκρίνοντας τόσο τις μετρικές F1-score & IoU αλλά και την cross-entropy loss του testing set της κάθε δοκιμής προκύπτει το συμπέρασμα ότι η εκπαίδευση με τις δεκαοκτώ (18) εποχές και μεγαλύτερο πλήθος batches (12) και samples (42) είχε τις καλύτερες προβλέψεις, ενώ η εκπαίδευση με τις λιγότερες εποχές (10) και το μικρότερο πλήθος batches (8) και samples (32), δηλαδή η πρώτη 1^η δοκιμή εκπαίδευσης, είχε τις λιγότερο ακριβείς προβλέψεις. Η πρώτη 1^η δοκιμή με τις δέκα εποχές εκπαίδευσης είχε τα εξής αποτελέσματα: *Testing set cross-entropy loss* = 0.4123271107673645 (41,23%) και *F1-score for Deforest*: 0.41831946594427244 (41,83%), *IOU-score for Deforest*: 0.26447784931108836 (26,45%), ενώ η τέταρτη 4^η δοκιμή εκπαίδευσης με τις δεκαοκτώ εποχές είχε τα εξής αποτελέσματα: *Testing set cross-entropy loss* = 0.23873858153820038 (23,87%) και *F1-score for Deforest*: 0.6713580387038193 (67,14%), *IOU-score for Deforest*: 0.5052964291816162 (50,53%). Η μετρική cross-entropy loss όσο τείνει στο 0 δηλώνει ότι οι προβλέψεις του μοντέλου είναι και οι ιδανικές. Το ακριβώς αντίθετο ισχύει για τις μετρικές f1-score & IoU. Δηλαδή, όσο τείνουν στο 1 τόσο καλύτερες είναι και οι προβλέψεις του μοντέλου.

Επιπλέον, παρατηρείται ότι το μοντέλο προβλέπει με αρκετά καλή ακρίβεια τις δασικές περιοχές (*forest*) σε σχέση με τις περιοχές αποψίλωσης (*deforest*). Αυτό εμφανίζεται από τα αποτελέσματα των μετρικών *f1-score* & *Intersection over Union (IoU)*. Για παράδειγμα, στην τέταρτη 4^η δοκιμή εκπαίδευσης προκύπτει: *F1-score for Forest*: 0.960351046870264

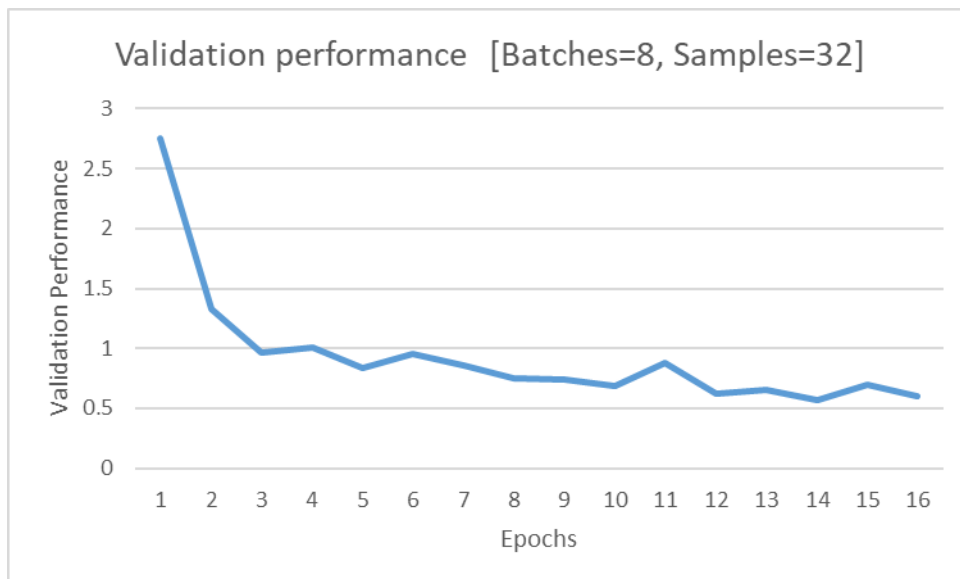
(96%), *IOU-score for Forest*: 0.923726267389819 (92,37%), ενώ *F1-score for Deforest*: 0.6713580387038193 (67,14%), *IOU-score for Deforest*: 0.5052964291816162 (50,53%). Επίσης, σύμφωνα με τα αποτελέσματα των μετρικών *f1-Score* και *Intersection over Union (IoU)* ερμηνεύεται ότι εντοπίζει με ικανοποιητική ακρίβεια τις υπόλοιπες περιοχές της φυσικής γήινης επιφάνειας (*other*) που δεν αποτελούν δασική έκταση ή αποψιλωμένη έκταση. Για παράδειγμα, στην τέταρτη 4η δοκιμή εκπαίδευσης προκύπτει: *F1-score for Other*: 0.8428457837495451 (84,28%), *IOU-score for Other*: 0.7283780950828073 (72,84%). Ωστόσο, παρατηρώντας τις προβλέψεις του μοντέλου φαίνεται ότι παρουσιάζει δυσλειτουργία ως προς τον εντοπισμό της κατηγορίας *other* κάτι το οποίο είναι λογικό καθώς η συγκεκριμένη κατηγορία εντοπίζεται στις εικόνες σε πολύ μικρότερο ποσοστό από τις υπόλοιπες δύο κατηγορίες – *forest* & *deforest*. Αυτό έχει ως αποτέλεσμα το μοντέλο να μην εκπαιδεύεται κατάλληλα στο εντοπισμό της συγκεκριμένης κατηγορίας *other*. Πιθανότατα μεγαλύτερο πλήθος δεδομένων – εικόνων που να συμπεριλαμβάνουν σε μεγαλύτερο ποσοστό την κατηγορία *other* θα έδινε την ευκαιρία στο μοντέλο να εκπαιδευτεί καλύτερα στον εντοπισμό της συγκεκριμένης κατηγορίας.

6.2.3 Διαγράμματα «Validation Performance» δοκιμών εκπαίδευσης

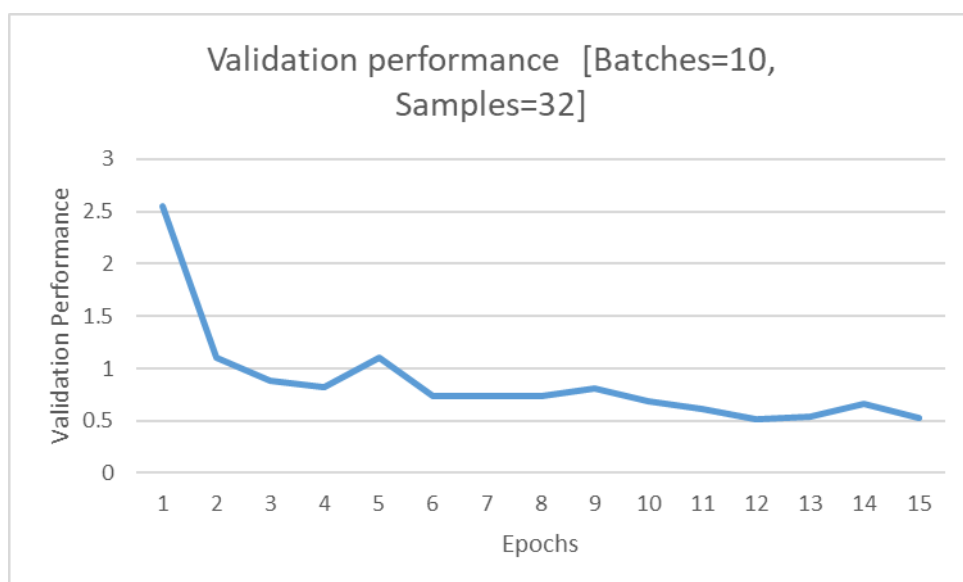
Για training set 80%:



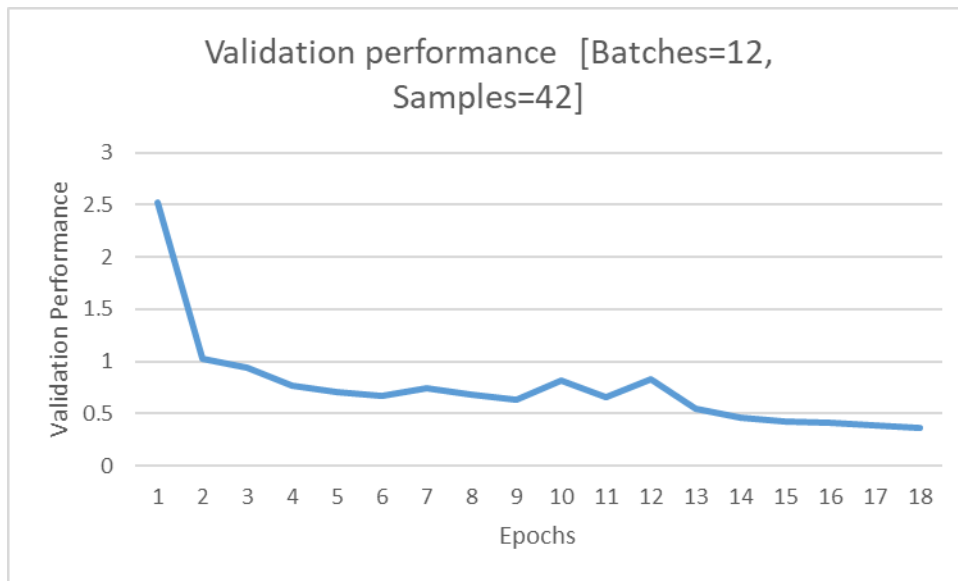
Εικόνα 29 -Validation Performance - 1η δοκιμή εκπαίδευσης



Εικόνα 30 - Validation Performance - 2η δοκιμή εκπαίδευσης

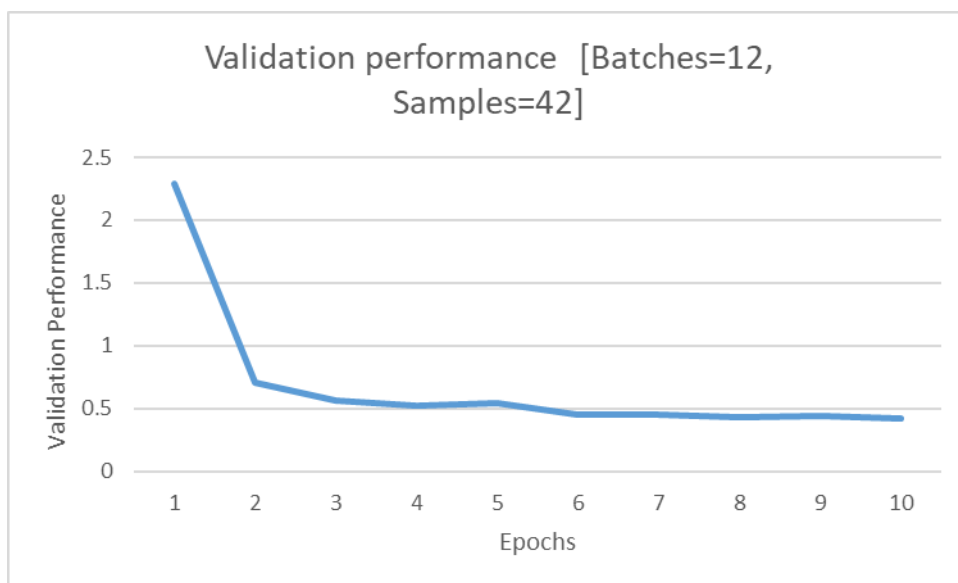


Εικόνα 31 - Validation Performance - 3η δοκιμή εκπαίδευσης



Εικόνα 32 - Validation Performance - 4η δοκιμή εκπαίδευσης

Για training set 30%:



Εικόνα 33 - Validation Performance - 5η δοκιμή εκπαίδευσης

Η τιμή Validation performance προκύπτει ύστερα από την ολοκλήρωση της κάθε εποχής. Η τιμή αρχικά ορίζεται σε αρκετά υψηλή τιμή, όπως φαίνεται και από τα ανωτέρω διαγράμματα η τιμή αρχικά κυμαίνεται στο εύρος 2 – 2.5. Κατά την διάρκεια της εκπαίδευσης του μοντέλου και όσο προχωράει η διαδικασία εκμάθησης των εικόνων από το μοντέλο και έχει αναγνωρίσει την χωρική πληροφορία που υπάρχει στις εικόνες η τιμή αυτή αναμένεται να μειωθεί και να τείνει στο 0%. Γενικά, όσο μικρότερη η τιμή validation performance, δηλαδή όσο αυτή η τιμή τείνει στο 0, τόσο καλύτερο είναι το αποτέλεσμα της εκπαίδευσης του μοντέλου.

Παρατηρείται ότι στις τρεις πρώτες δοκιμές εκπαίδευσης η ελάχιστη τιμή Validation performance κυμαίνεται στο εύρος τιμών 0,5 – 1, ενώ στην τέταρτη (4η) δοκιμή εκπαίδευσης όπου από αυτήν λαμβάνονται και τα καλύτερα δυνατά αποτελέσματα η τιμή validation performance είναι μικρότερη από 0,5, περίπου 0,3. Όσον αφορά την ελάχιστη τιμή της

Validation performance στην πέμπτη (5η) δοκιμή εκπαίδευσης με το 30% του dataset έχει λάβει τιμή περί τα ~ 0,4.

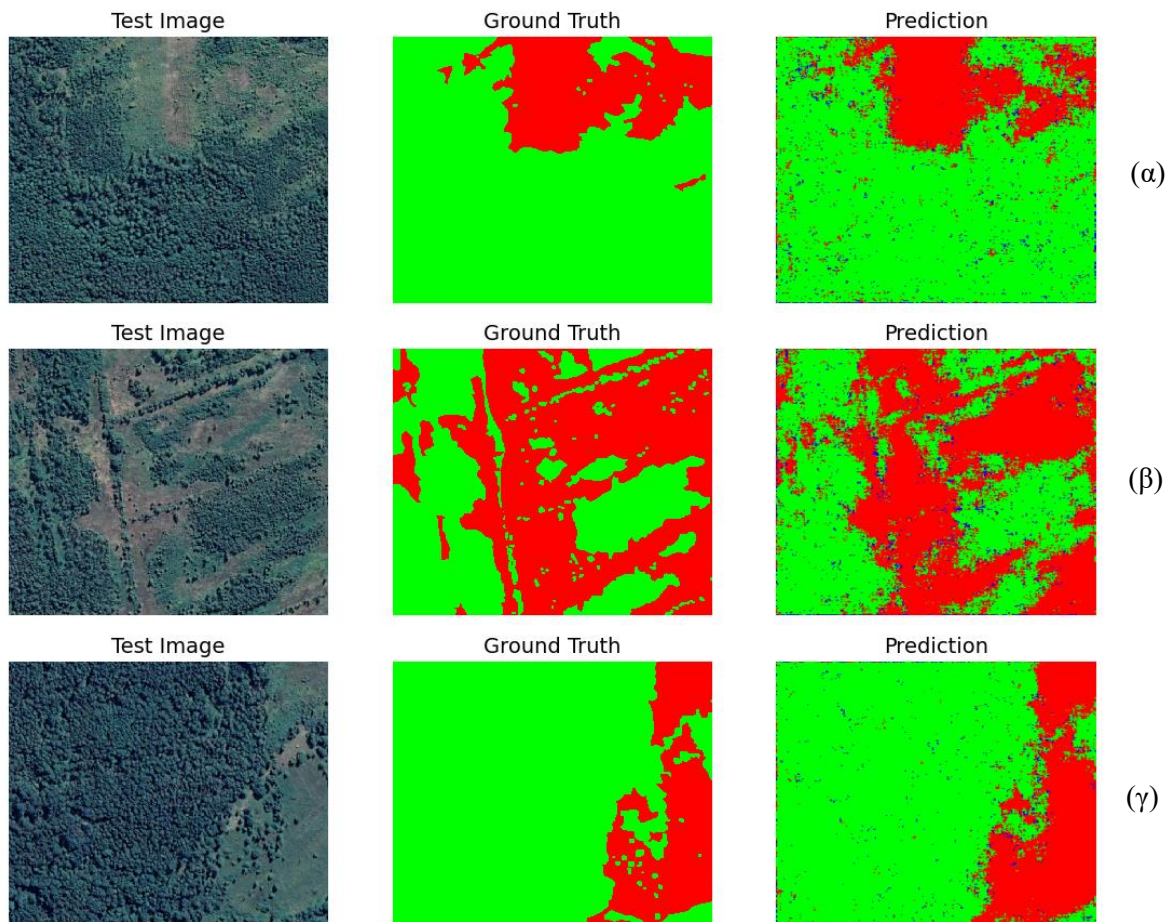
6.3 Μοντέλο Μηχανικής Μάθησης «Image Segmentation Keras: Implementation of Segnet, FCN, UNet, PSPNet»

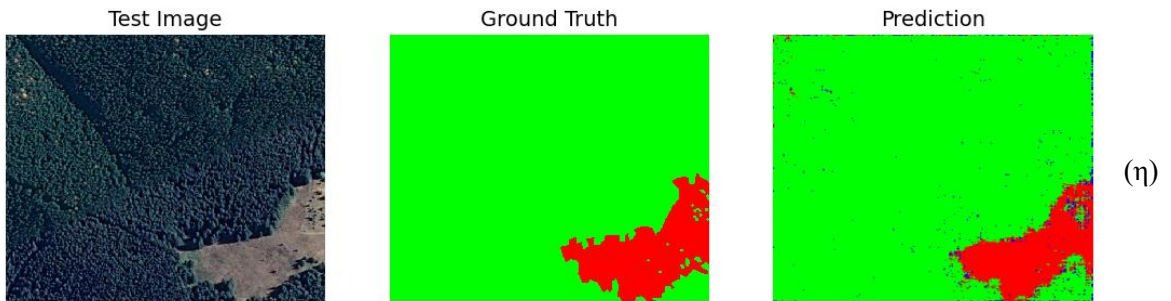
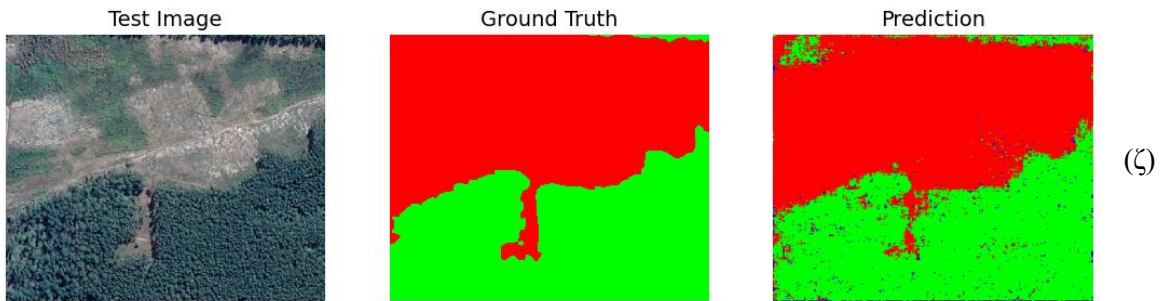
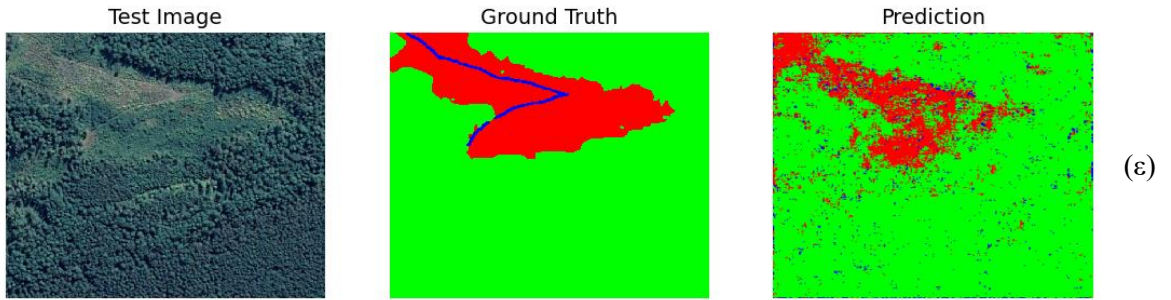
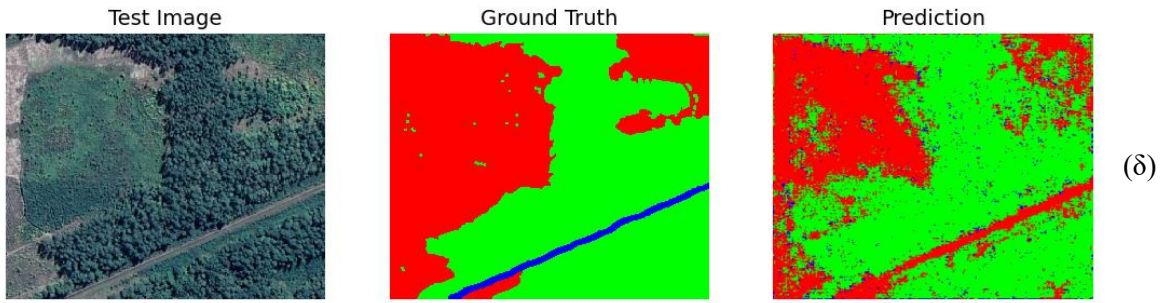
Πρώτη 1^η δοκιμή εκπαίδευσης

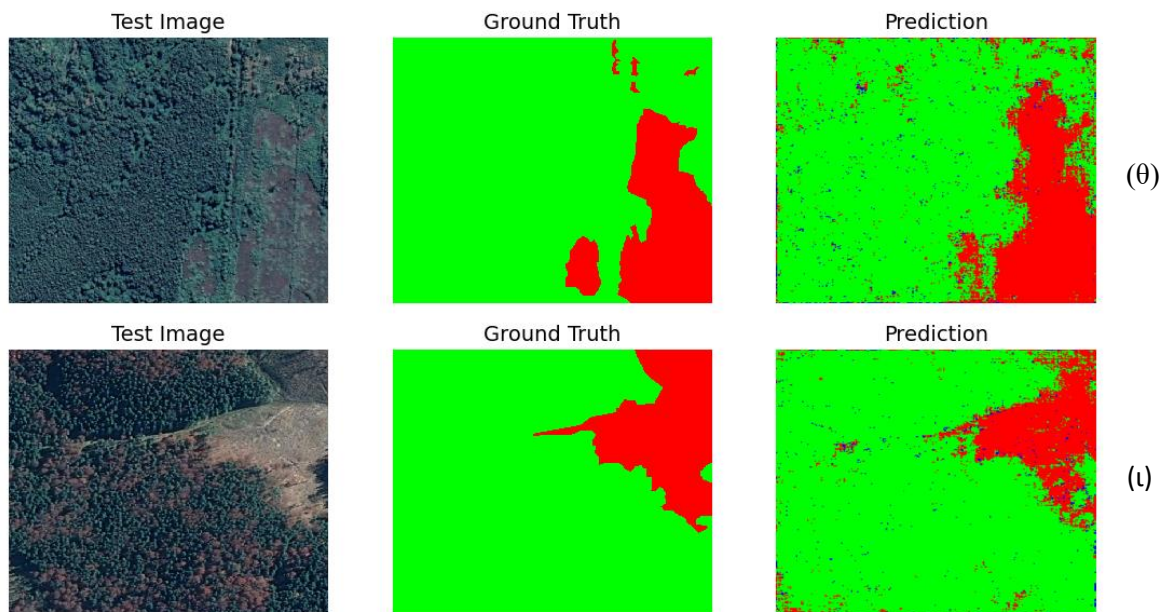
Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Είκοσι εποχές (*epochs=20*)
- *Segmentation model=unet*
- *Seed=15*

Η εκπαίδευση διήρκησε είκοσι (21) λεπτά και στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (*Prediction*) σε σχέση με την μάσκα (*Ground Truth*) του «Dataset». Η υπολογιζόμενη από το μοντέλο μετρική «ακρίβεια» (*accuracy*) ξεκίνησε με ελάχιστη τιμή 0.3973 (39,73%) και έφτασε μέγιστη τιμή 0.7392 (73,92%).







Εικόνα 34 - Μοντέλο "image segmentation keras" - Test Image, Ground Truth, Prediction - 1η δοκιμή εκπαίδευσης

Στην συγκεκριμένη δοκιμή εκπαίδευσης επιλέχθηκαν είκοσι (20) εποχές εκπαίδευσης. Από τις ανωτέρω εικόνες προκύπτει το συμπέρασμα ότι το μοντέλο ανιχνεύει σε ικανοποιητικό βαθμό την θέση της αποψιλωμένης έκτασης στην εικόνα, όπως βλέπουμε αντίστοιχα στις εικόνες 34 (ζ) και 34 (η), ενώ σε λιγότερο ικανοποιητικό βαθμό φαίνεται το μοντέλο να ανιχνεύει την αποψιλωμένη έκταση στην εικόνα 34 (ε). Ωστόσο, σε όλες τις προβλεπόμενες από το μοντέλο εικόνες (Prediction) φαίνεται να είναι αυξημένος ο θόρυβος στην εικόνα, δηλαδή το μοντέλο εντοπίζει εντός της δασικής έκτασης πάρα πολύ μικρές περιοχές αποψύλωσης ή περιοχές της κατηγορίας «other» κάτι το οποίο δεν ισχύει και συμβαίνει διότι το μοντέλο δεν έχει εκπαιδευτεί στο βέλτιστο δυνατό βαθμό και δεν έχει μάθει να ξεχωρίζει ικανοποιητικά τις τρεις κατηγορίες που πρέπει να ανιχνεύσει και να ταξινομήσει στην εικόνα.

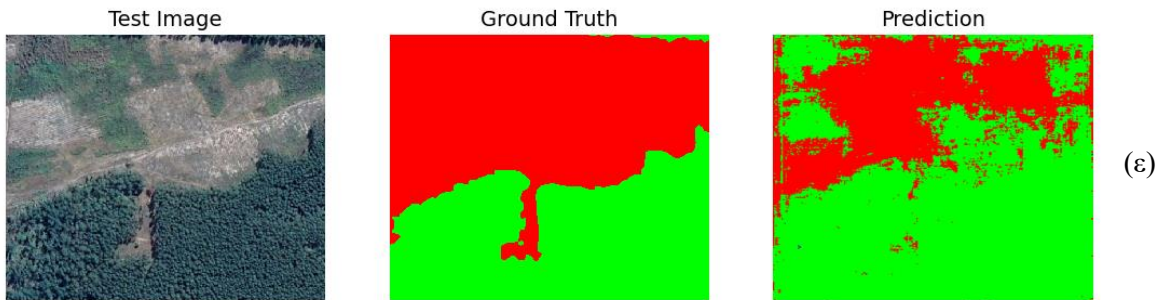
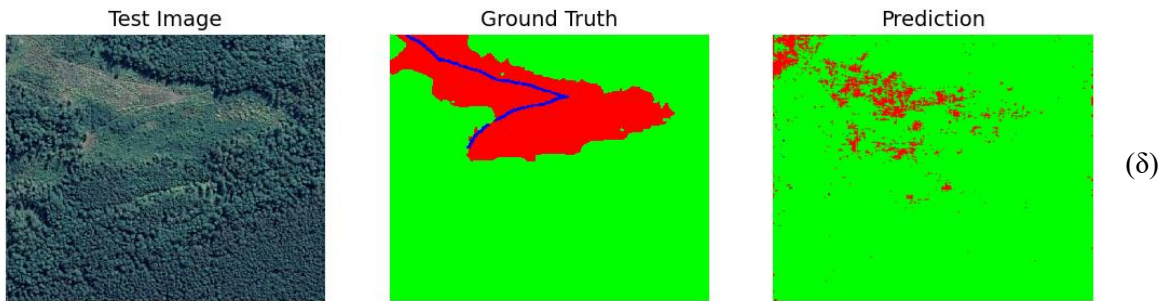
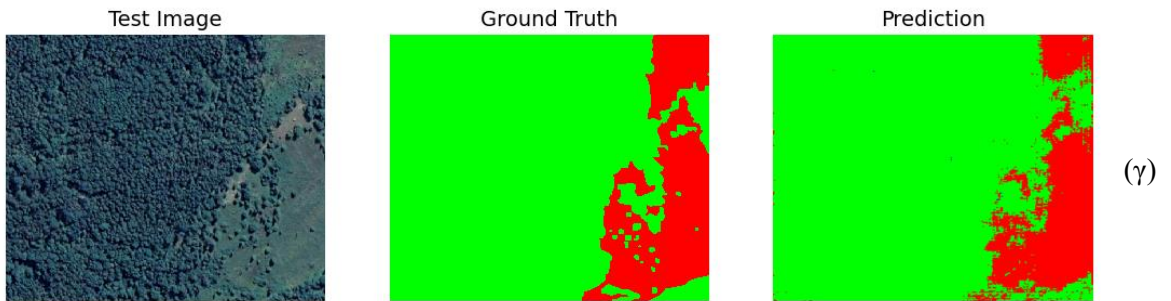
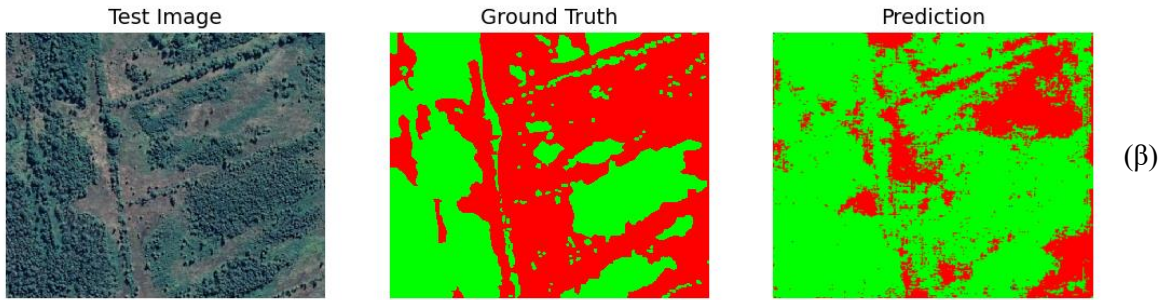
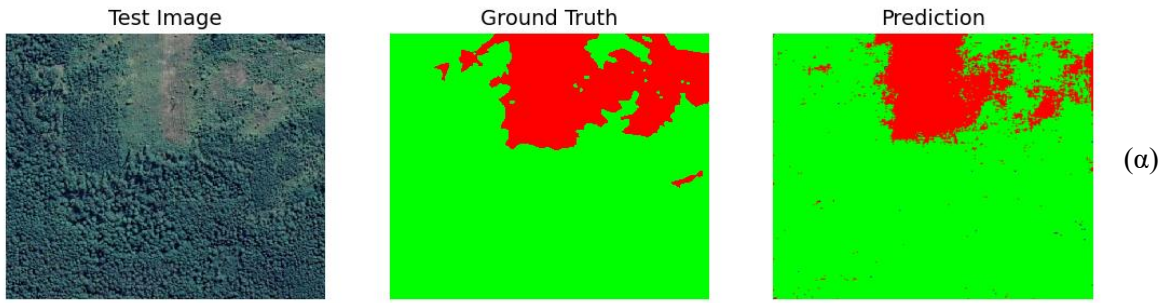
Για τον λόγο αυτό πραγματοποιήθηκε μία ακόμη δοκιμή εκπαίδευσης με τις διπλάσιες εποχές εκπαίδευσης προκειμένου να διερευνηθεί αν θα βελτιωθούν οι προβλέψεις του μοντέλου.

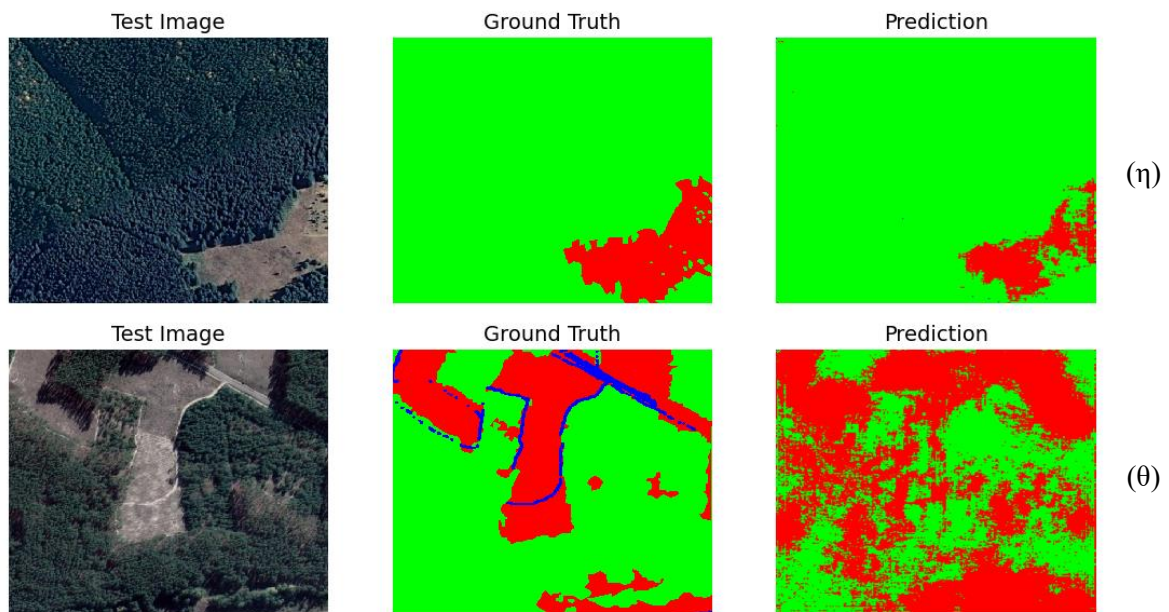
Δεύτερη 2^η δοκιμή εκπαίδευσης

Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Σαράντα εποχές (*epochs=40*)
- *Segmentation model=unet*
- *Seed=15*

Η εκπαίδευση διήρκησε σαράντα τέσσερα (44) λεπτά και στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (*Prediction*) σε σχέση με την μάσκα (*Ground Truth*) του «Dataset». Η υπολογιζόμενη από το μοντέλο μετρική «ακρίβεια» (*accuracy*) ξεκίνησε με ελάχιστη τιμή (37,73%) και έφτασε μέγιστη τιμή (94,41%).





Εικόνα 35 - Μοντέλο "image segmentation keras" - Test image, Ground Truth, Prediction – 2η δοκιμή εκπαίδευσης

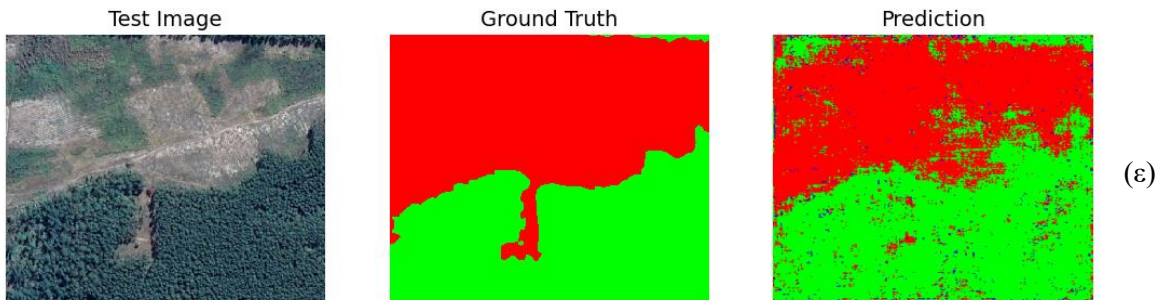
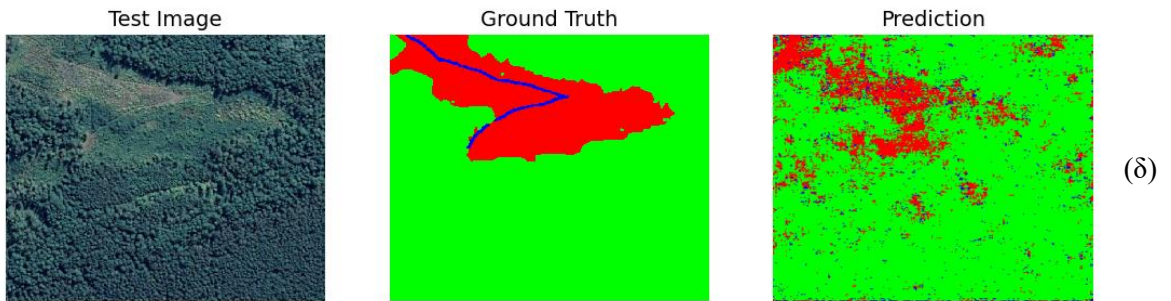
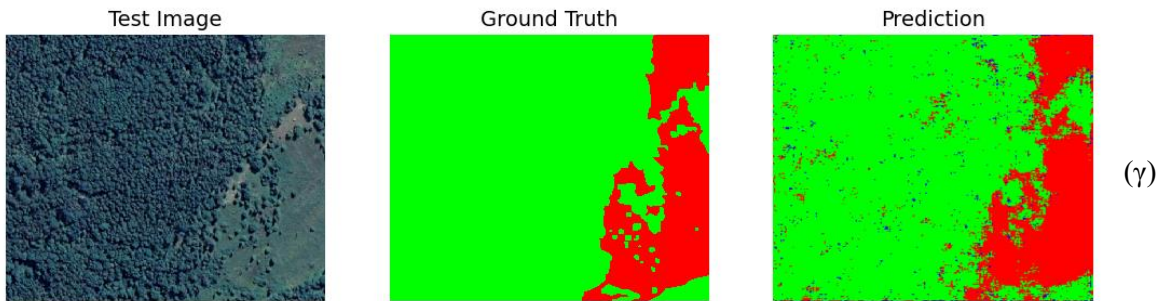
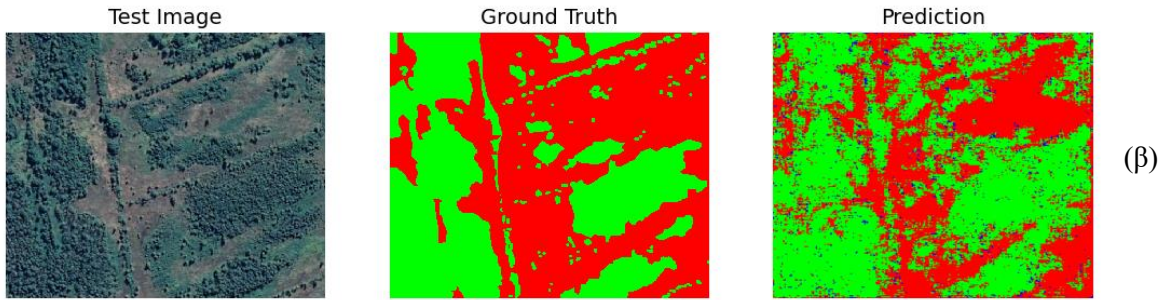
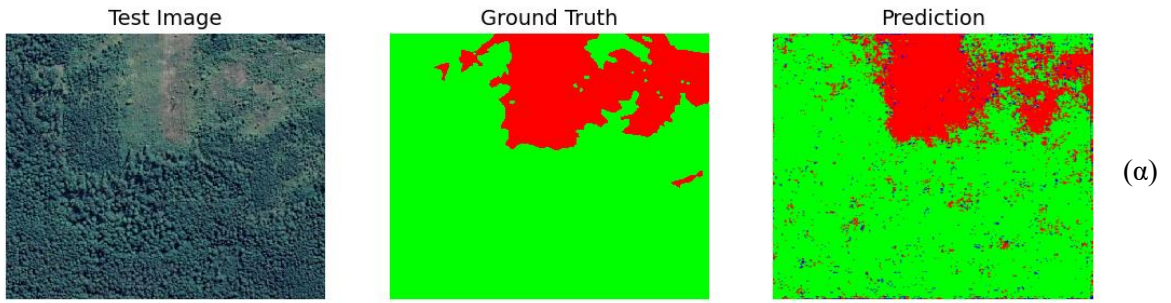
Στην συγκεκριμένη δοκιμή εκπαίδευσης χρησιμοποιήθηκαν οι διπλάσιες εποχές εκπαίδευσης σε σχέση με την πρώτη 1η δοκιμή εκπαίδευσης και παρατηρήθηκε ότι σε όλες τις εικόνες ο θόρυβος στην εικόνα έχει μειωθεί αισθητά. Παρόλα αυτά ένω το μοντέλο έχει ανιχνεύσει την θέση της περιοχής αποψίλωσης στην κάθε εικόνα φαίνεται ότι δεν υπάρχει η ίδια πληρότητα σε σχέση με την πρώτη δοκιμή εκπαίδευσης. Πιο συγκεκριμένα, παρατηρείται στην εικόνα 34 (ε) και στην εικόνα 35 (δ) εμφανής διαφορά στην ανίχνευση της αποψιλωμένης έκτασης. Επιπλέον, παρατηρείται στην εικόνα 35 (η) η οποία αποτελεί εικόνα με μεγαλύτερη πολυπλοκότητα σε σχέση με τις υπόλοιπες, δηλαδή η αποψιλωμένη έκταση δεν εμφανίζεται συγκεντρωμένη σε μια θέση της εικόνας, όπως αντίστοιχα συμβαίνει στην εικόνα 35 (ζ) ή στην εικόνα 35 (α), αλλά εντοπίζεται διάσπαρτα σε διάφορες θέσεις στην εικόνα, ότι το μοντέλο δυσκολεύεται σε αρκετά μεγάλο ποσοστό να εντοπίσει τις θέσεις των αποψιλωμένων εκτάσεων και για αυτό το λόγω προκύπτει μια λανθασμένη από το μοντέλο πρόβλεψη.

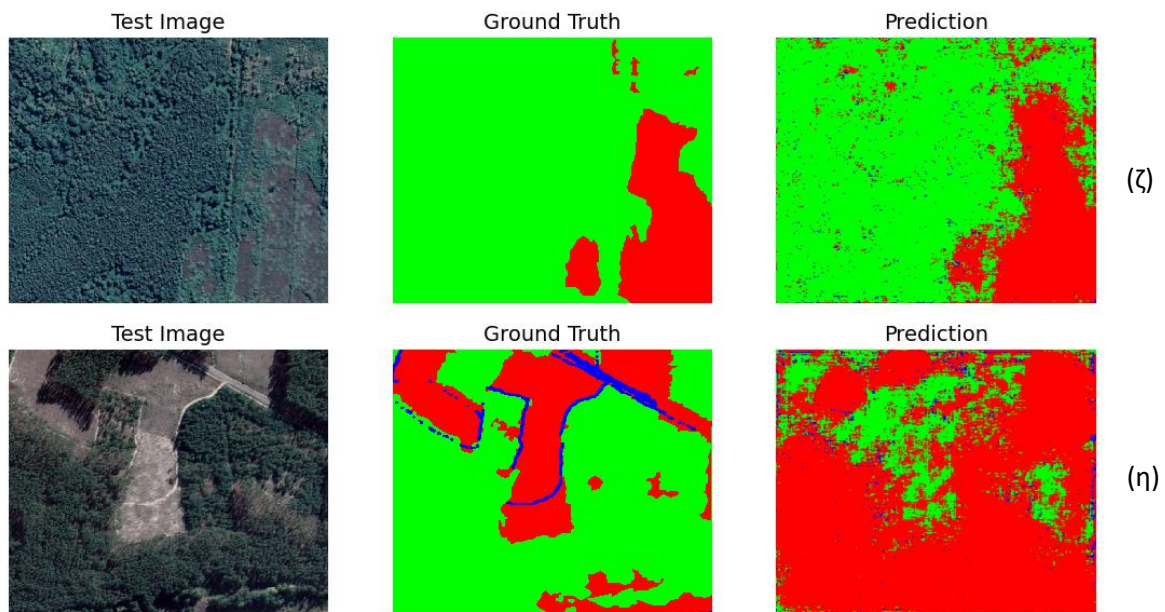
Τρίτη 3^η δοκιμή εκπαίδευσης

Για σκοπούς σύγκρισης έγινε δοκιμή του συγκεκριμένου μοντέλου με αρκετά μικρότερο training set, δηλαδή σε ποσοστό 30% που αντιστοιχεί περίπου σε ενενήντα (90) εικόνες εκπαίδευσης. Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι εξής παράμετροι:

- Σαράντα εποχές (epochs=20)
- Segmentation model=unet
- Seed=15

Η εκπαίδευση διήρκησε είκοσι δύο (22) λεπτά και στις παρακάτω εικόνες παρατίθενται οι προβλέψεις του μοντέλου (Prediction) σε σχέση με την μάσκα (Ground Truth) του «Dataset». Η υπολογιζόμενη από το μοντέλο μετρική «ακρίβεια» (accuracy) ξεκίνησε με ελάχιστη τιμή (37,14%) και έφτασε μέγιστη τιμή (80,88%).



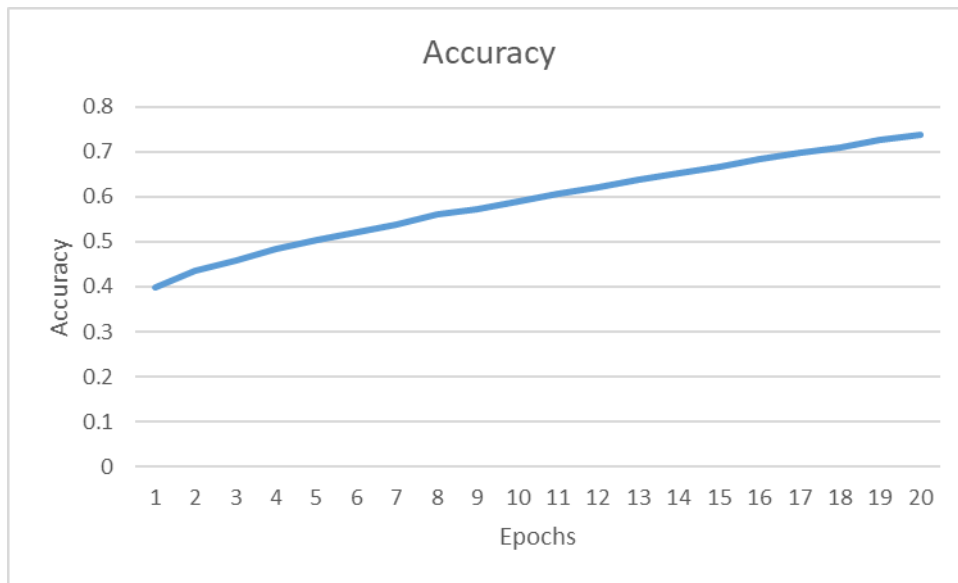


Εικόνα 36 – Μοντέλο "image segmentation keras" - Test image, Ground Truth, Prediction – 3η δοκιμή εκπαίδευσης

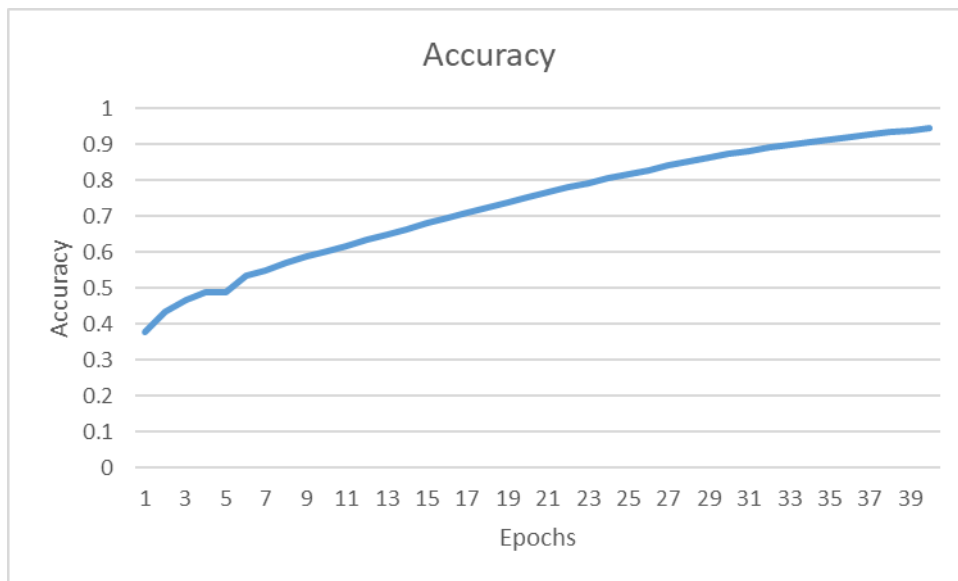
Η συγκεκριμένη δοκιμή εκπαίδευσης πραγματοποιήθηκε για σκοπούς σύγκρισης της λειτουργικότητας του μοντέλου για πολύ μικρότερο αριθμό δεδομένων εκπαίδευσης (training set). Πάρθηκε training set περί το 30% των δεδομένων ενώ τα υπόλοιπα δεδομένα χρησιμοποιήθηκαν ως δεδομένα ελέγχου. Παρατηρείται ότι τα αποτελέσματα δεν είναι ικανοποιητικά και ότι το μοντέλο έχει δυσκολευτεί να ανιχνεύσει σε καλό ποσοστό τις αποψιλωμένες εκτάσεις όπως αντίστοιχα και ο θόρυβος στις εικόνες είναι αρκετά αυξημένος. Ειδικότερα αυτό εμφανίζεται στις εικόνες 36 (α) και 36 (β) και 36 (δ). Επιπλέον, στην εικόνα (η) που έχει την μεγαλύτερη πολυπλοκότητα σε σχέση με τις υπόλοιπες εικόνες η πρόβλεψη του μοντέλου είναι λανθασμένη. Συνεπώς, προκύπτει το συμπέρασμα ότι όσο περισσότερα είναι τα δεδομένα εκπαίδευσης δίνεται η ευκαιρία στο μοντέλο να εκπαιδευτεί καλύτερα στις κατηγορίες που πρέπει να ανιχνευθούν και να προκύψει μια αξιόπιστη από το μοντέλο πρόβλεψη.

Συμπερασματικά, παρατηρώντας τις παραπάνω εικόνες με τις προβλέψεις του μοντέλου σε κάθε μια από τις δύο δοκιμές εκπαίδευσης προκύπτει το συμπέρασμα ότι στις σαράντα (40) εποχές εκπαίδευσης μειώθηκε ο «θόρυβος» συγκριτικά με τις είκοσι (20) εποχές εκπαίδευσης αλλά μειώθηκε κατά μικρό ποσοστό η ακρίβεια του εντοπισμού των τριών κατηγοριών *forest*, *deforest* & *other*. Δηλαδή, από τις εικόνες που παρατίθενται ανωτέρω *prediction* & *Ground Truth* προκύπτει το συμπέρασμα ότι στις 20 εποχές παρατηρείται πιο ακριβή πρόβλεψη των τριών κατηγοριών αλλά περισσότερος θόρυβος στην εικόνα. Ενώ, στις 40 εποχές εκπαίδευσης παρατηρείται λιγότερο ακριβή πρόβλεψη των τριών κατηγοριών αλλά αισθητά μειωμένος θόρυβος στις προβλεπόμενες εικόνες.

6.3.1 Διαγράμματα «Accuracy» δοκιμών εκπαίδευσης

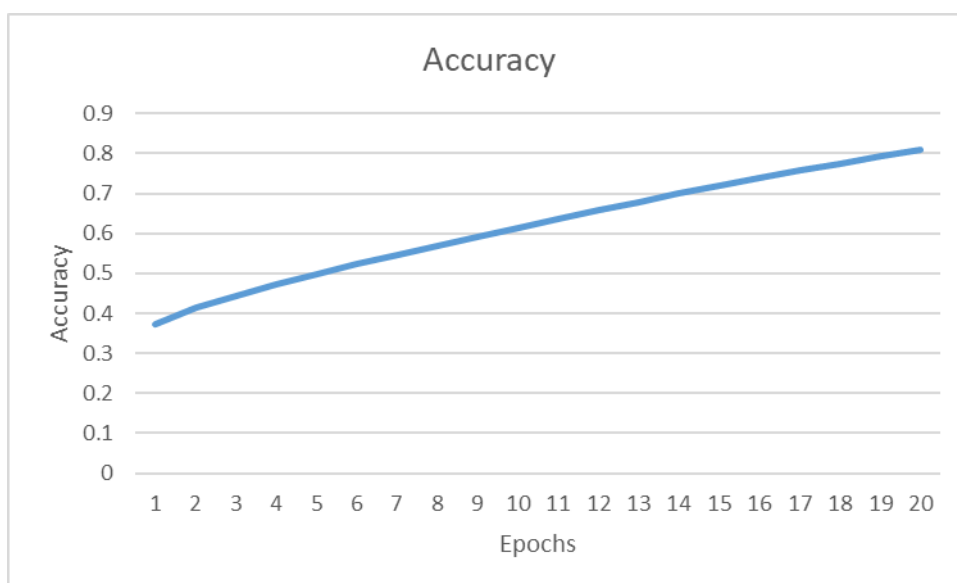


Εικόνα 37 - Διάγραμμα μετρικής "Accuracy" - 20 εποχές εκπαίδευσης



Εικόνα 38 - Διάγραμμα "Accuracy" - 40 εποχές εκπαίδευσης

Για training set 30%:



Εικόνα 39 – Διάγραμμα "Accuracy" – 20 εποχές εκπαίδευσης – training set 30%

Στα ανωτέρω διαγράμματα εμφανίζεται η πορεία της τιμής της μετρικής accuracy η οποία αναμένεται να αυξάνεται και να τείνει στην τιμή ένα (1), ποσοστό 100%, καθώς προχωράει η εκπαίδευση του μοντέλου. Η ανωτέρω πορεία της τιμής υποδηλώνει την αποτελεσματικότητα της εκπαίδευσης του μοντέλου και το πόσο καλά έχει «μάθει» το μοντέλο την χωρική πληροφορία της εικόνας. Στις δοκιμές εκπαίδευσης που πραγματοποιήθηκαν στο συγκεκριμένο μοντέλο η τιμή accuracy ξεκίνησε με ένα εύρος τιμής μεταξύ 0,3 – 0,4 και κατά την ολοκλήρωση της εκπαίδευσης του μοντέλου η τιμή ήταν μεταξύ των τιμών 0,8 – 0,9. Στην δεύτερη 2η δοκιμή εκπαίδευσης μόνο το μοντέλο έχει λάβει ανώτερη τιμή η οποία είναι μεγαλύτερη της τιμής 0,9, δηλαδή η τιμή accuracy τείνει στο 1.

Κεφάλαιο 7 – Συμπεράσματα και Μελλοντική Εργασία

7.1 Συμπεράσματα

Η συγκεκριμένη διπλωματική εργασία πραγματεύεται εφαρμογές σημασιολογικής κατάτμησης δορυφορικών εικόνων για τον εντοπισμό αποψιλωμένων περιοχών. Οι εικόνες που χρησιμοποιήθηκαν αντλήθηκαν από την open source ιστοσελίδα GitHub και συγκεκριμένα από το μοντέλο αναφοράς «*BioWar/Satellite-Image-Segmentation-using-Deep-Learning-for-Deforestation-Detection*».

Ύστερα από επεξεργασία και χρήση μοντέλων *UNet* προκύπτει η κατάτμηση των δορυφορικών εικόνων και η ανίχνευση των περιοχών αποψίλωσης (*deforestation*). Μεταξύ των δύο μοντέλων που μελετήθηκαν και παρουσιάζονται στις ενότητες 5.3 και 5.4 διαπιστώθηκε ότι καλύτερο μοντέλο καθ' όλη την διάρκεια της πειραματικής διαδικασίας είναι το «*Swin UNet*», καθώς οι προβλέψεις του ανιχνεύουν σε αρκετά καλό ποσοστό περιοχές αποψίλωσης. Επιπλέον, για τον σκοπό της παρούσας διπλωματικής, οι περισσότερο αντικειμενικές μετρικές είναι η *F1 – Score* (Αρμονικός μέσος) και *Intersection over Union (IoU)*, καθώς έχουν προστεθεί κατάλληλα στο μοντέλο «*Swin UNet*», ώστε να υπολογίζεται το ποσοστό ακρίβειας ανίχνευσης της κάθε μια κλάσης (*forest, deforest & other*). Πιο συγκεκριμένα, οι προσανξήσεις ή απομειώσεις στις συγκεκριμένες μετρικές ανάλογα με το πλήθος των εποχών εκπαίδευσης και την κατάλληλη επιλογή των παραμέτρων *batches* και *samples* υποδηλώνουν την αξιοπιστία των αποτελεσμάτων. Δηλαδή, στο μοντέλο «*Swin UNet*» παρατηρήθηκε ότι για τις δεκαοκτώ εποχές εκπαίδευσης $N_{epoch}=18$ που έγινε δοκιμή προέκυψε *F1-score for Deforest: ~ 67%* και αντίστοιχα *IOU-score for Deforest: ~ 51%*. Αντίθετα, για τις λιγότερες εποχές εκπαίδευσης $N_{epoch}=10$ που έγινε αντίστοιχα δοκιμή για την ποιότητα των προβλέψεων προέκυψε *F1-score for Deforest: ~ 42%* και *IOU-score for Deforest: ~ 26%*. Στις συγκεκριμένες μετρικές όσο η τιμή τείνει στο 1 (σε ποσοστό 100%) τόσο καλύτερη και περισσότερο αξιόπιστη είναι η πρόβλεψη του μοντέλου και αυτό σημαίνει ότι λειτουργεί με τον βέλτιστο δυνατό τρόπο. Σύμφωνα με τις ανωτέρω τιμές των μετρικών *F1 – Score* (Αρμονικός μέσος) και *Intersection over Union (IoU)* και λαμβάνοντας υπόψη τις προβλέψεις – εικόνες του μοντέλου της κάθε δοκιμής εκπαίδευσης, συμπεραίνεται ότι από το μοντέλο «*Swin UNet*» προκύπτουν αξιόπιστες προβλέψεις και αυτό επιβεβαιώνεται βάσει των αποτελεσμάτων των παραπάνω μετρικών. Επιπλέον, παρατηρείται ότι η μετρική *F1 – Score* είναι περισσότερο αξιόπιστη συγκριτικά με την *Intersection over Union (IoU)*.

Το μοντέλο «*image segmentation keras*» του οποίου τα αποτελέσματα αποτέλεσαν συγκριτικό στοιχείο διαπιστώθηκε ότι έχει μέτρια αποτελέσματα σε σχέση με το μοντέλο «*Swin – UNet*» με μικρότερο ποσοστό ακρίβειας εντοπισμού περιοχών αποψίλωσης. Συνεπώς, για τους σκοπούς της σημασιολογικής κατάτμησης προτείνεται ο *Swin – Transformer* καθώς τα αποτελέσματα όπως μελετήθηκε και ανωτέρω είναι περισσότερο αξιόπιστα.

Αυτή η εργασία έχει πολύ ενδιαφέρον για μελλοντικές μελέτες καθώς μπορεί να αποτελέσει σημαντικό εργαλείο για εφαρμογές Τηλεπισκόπησης. Με την αυτόματη κατάτμηση εικόνας και πρόβλεψη αποψιλωμένων δασικών εκτάσεων θα μετατραπεί η διαδικασία ανίχνευσης των συγκεκριμένων περιοχών σε πιο εύκολη και κυρίως πιο ακριβής διαδικασία. Η αποψίλωση των δασών είναι ένα φαινόμενο που αφορά άμεσα την επιστημονική κοινότητα και κύρια αιτία είναι η κλιματική αλλαγή με βασική συνέπεια την υποβάθμιση των δασών. Τα ακραία φαινόμενα που προκαλεί, όπως πυρκαγιές, ξηρασίες και πλημμύρες, επηρεάζουν τα δάση. Ιδιαίτερα η Ελλάδα που πλήγεται από πυρκαγιές οι περιοχές αποψίλωσης πληθαίνουν κάθε χρόνο. Το γεγονός ότι υφίστανται μοντέλα που με κατάλληλα δεδομένα μπορούν να ανιχνεύσουν τέτοιες εκτάσεις άμεσα και αξιόπιστα

αποτελεί σημαντική πρόοδο για την κατανόηση του ποσοστού της αποψιλωμένης έκτασης και στην συνέχεια την αντιμετώπιση του φαινομένου (*deforestation*).

7.2 Μελλοντική Εργασία

Υπάρχουν πολλά ακόμα μοντέλα και βελτιώσεις πάνω σε αυτά που μπορούν να γίνουν, ώστε η ανίχνευση των εκτάσεων αποψίλωσης με χρήση τεχνικών βαθιάς μάθησης να δίνει ακόμα καλύτερα αποτελέσματα. Κατ' επέκταση, μπορούν να ερευνηθούν τα μοντέλα που αναλύθηκαν και επεξεργάστηκαν στην συγκεκριμένη εργασία, ώστε με κατάλληλες βελτιώσεις να παράγουν τις βέλτιστες δυνατές προβλέψεις.

Βιβλιογραφία

- [1] Aurelien Geron. Introducing capsule networks, <https://www.oreilly.com/content/introducing-capsule-networks/>, February 6, 2018
- [2] Zewen Li, Fan Liu , Member, IEEE, Wenjie Yang, Shouheng Peng, and Jun Zhou , Senior Member, IEEE. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects, 10 June 2021
- [3] Shervin Minaee, Member, IEEE, Yuri Boykov, Member, IEEE, Fatih Porikli, Fellow, IEEE, Antonio Plaza, Fellow, IEEE, Nasser Kehtarnavaz, Fellow, IEEE, and Demetri Terzopoulos, Fellow, IEEE. Image Segmentation Using Deep Learning: A Survey, 17 February 2021
- [4] International Journal of Science and Research (IJSR). Machine Learning Algorithms - A Review, January 2020
- [5] Oludare Isaac Abiodun a,b, Aman Jantan a, Abiodun Esther Omolara c, Kemi Victoria Dada d, Nachaat AbdElatif Mohamed e, Humaira Arshad. State-of-the-art in artificial neural network applications: A survey, November 2018
- [6] Maria Kaselimi, Athanasios Voulodimos, Member, IEEE, Ioannis Daskalopoulos, Nikolaos Doulamis, Member, IEEE, and Anastasios Doulamis, Member IEEE. A vision Transformer Model for Convolution-Free Multilabel Classification of Satellite Imagery in Deforestation Monitoring, 02 February 2022
- [7] Hu Cao^{1y}, Yueyue Wang^{2y}, Joy Chen¹, Dongsheng Jiang³, Xiaopeng Zhang³, Qi Tian³, and Manning Wang². Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation, 12 May 2021
- [8] Liu Yutong Lin Yue Cao Han Hu Yixuan Wei Zheng Zhang Stephen Lin Baining Guo Microsoft Research Asia. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows, 17 Aug 2021
- [9] Divam Gupta. A Beginner's guide to Deep Learning based Semantic Segmentation using Keras, <https://divamgupta.com/image-segmentation/2019/06/06/deep-learning-semantic-segmentation-keras.html>, 06 Jun 2019
- [10] Κάλλιπος. Τεχνητή Νοημοσύνη, Μια εισαγωγική προσέγγιση, ΚΕΦΑΛΑΙΟ 4 - Μηχανική Μάθηση, Κατερίνα Γεωργούλη, http://repfiles.kallipos.gr/html_books/93/04a-main.html, 2015
- [11] Yann LeCun, Yoshua Bengio & Geoffrey Hinton. Deep learning, 2015
- [12] S Sykiotis, C Menos-Aikateriniadis, A Doulamis, N Doulamis, PS Georgilakis. A self-sustained EV charging framework with N-step deep reinforcement learning, Sustainable Energy, Grids and Networks 35, 101124, 2023

- [13] I Kavouras, R Ioannis, N Doulamis, A Doulamis. Dynamically tangible cultural heritage monitoring from web video sources, Proceedings of the 16th International Conference on Pervasive Technologies, 2023
- [14] I Georgoulas, E Protopapadakis, K Makantasis, A Doulamis. Tensor-based embedding for graph-based semi-supervised approaches, Proceedings of the 16th International Conference on Pervasive Technologies, 2023
- [15] G Kopsiaftis, M Kaselimi, E Protopapadakis, A Voulodimos, A Doulamis, N Doulamis, A Mantoglou. Performance comparison of physics-based and machine learning assisted multi-fidelity methods for the management of coastal aquifer systems, *Frontiers in Water* 5, 1195029, 2023
- [16] A Temenos, N Temenos, M Kaselimi, A Doulamis, N Doulamis. Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP, *IEEE Geoscience and Remote Sensing Letters* 20, 1-5, 2023
- [17] S Sykiotis, S Athanasoulas, M Kaselimi, A Doulamis, N Doulamis, L Stankovic, V Stankovic. Performance-aware NILM model optimization for edge deployment, *IEEE Transactions on Green Communications and Networking*, 2023
- [18] Kavouras, E Sardis, E Protopapadakis, I Rallis, A Doulamis, N Doulamis. A Low-Cost Gamified Urban Planning Methodology Enhanced with Co-Creation and Participatory Approaches, *Sustainability* 15 (3), 2297, 2023
- [19] K Labropoulos, C Zafeiropoulos, I Rallis, A Doulamis, N Doulamis, A Moropoulou. Improved Resilience and Sustainable Reconstruction of Cultural Heritage Areas to cope with Climate Change and Other Hazards based on Innovative Algorithms and Modelling Tools, *Technical Annals* 1 (1), 66-77, 2022
- [20] JN Tzortzi, A Doulamis, I Rallis, MS Lux, G Barbotti, I Tzortzis. HARMONIA: strategy of an integrated resilience assessment platform (IRAP) with available tools and geospatial services, *IOP Conference Series: Earth and Environmental Science* 1122 (1), 012020, 2022
- [21] M Kaselimi, E Protopapadakis, A Doulamis, N Doulamis. A review of non-invasive sensors and artificial intelligence models for diabetic foot monitoring, *Frontiers in Physiology* 13, 924546, 2022
- [22] S Sykiotis, C Menos-Aikateriniadis, A Doulamis, N Doulamis, PS Georgilakis. Solar power driven EV charging optimization with deep reinforcement learning, 2022 2nd International Conference on Energy Transition in the Mediterranean, 2022
- [23] IN Tzortzis, I Rallis, K Makantasis, A Doulamis, N Doulamis, A Voulodimos. Automatic inspection of cultural monuments using deep and tensor-based learning on hyperspectral imagery, 2022 IEEE International Conference on Image Processing (ICIP), 3136-3140, 2022,
- [24] IN Tzortzis, A Davradou, I Rallis, M Kaselimi, K Makantasis, A Doulamis, N Doulamis. Tensor-Based Learning for Detecting Abnormalities on Digital Mammograms, *Diagnostics* 12 (10), 2389, 2022

- [25] I Katsamenis, EE Karolou, A Davradou, E Protopapadakis, A Doulamis, N Doulamis, D Kalogeras. TraCon: A novel dataset for real-time traffic cones detection using deep learning, *Novel & Intelligent Digital Systems Conferences*, 382-391, 2022
- [26] S Sykiotis, M Kaselimi, A Doulamis, N Doulamis. An Efficient Deep Bidirectional Transformer Model for Energy Disaggregation, 2022 30th European Signal Processing Conference (EUSIPCO), 1536-1540, 2022
- [27] M Kaselimi, E Protopapadakis, A Voulodimos, N Doulamis, A Doulamis. Towards trustworthy energy disaggregation: A review of challenges, methods, and perspectives for non-intrusive load monitoring, *Sensors* 22 (15), 5872, 2022
- [28] A Davradou, E Protopapadakis, M Kaselimi, A Doulamis, N Doulamis. Diabetic foot ulcers monitoring by employing super resolution and noise reduction deep learning techniques, *Proceedings of the 15th International Conference on Pervasive Technologies*, 2022
- [29] C Zafeiropoulos, E Protopapadakis, A Chatzidaki, A Doulamis, D Vamvatsikos, N Zotos, G Bogdos, A Kostaridis. A holistic monitoring scheme for road infrastructures, *Proceedings of the 15th International Conference on Pervasive Technologies*, 2022
- [30] I Katsamenis, M Bimpas, E Protopapadakis, C Zafeiropoulos, D Kalogeras, A Doulamis, N Doulamis. Robotic maintenance of road infrastructures: The heron project, *Proceedings of the 15th International Conference on Pervasive Technologies*, 2022
- [31] I Kavouras, M Kaselimi, E Protopapadakis, N Bakalos, N Doulamis, A Doulamis. COVID-19 spatio-temporal evolution using deep learning at a European level, *Sensors* 22 (10), 3658, 2022
- [32] S Sykiotis, M Kaselimi, A Doulamis, N Doulamis. Electricity: An efficient transformer for non-intrusive load monitoring, *Sensors* 22 (8), 2926, 2022
- [33] I Katsamenis, S Sykiotis, N Dikaios, GA Kastis, SN Chatziioannou, M Metaxas, N Doulamis, A Doulamis. A few-shot U-Net deep learning model for lung cancer lesion segmentation via PET/CT imagin, *Biomedical Physics & Engineering Express* 8 (2), 025019, 2022
- [34] M Kaselimi, A Voulodimos, I Daskalopoulos, N Doulamis, A Doulamis. A vision transformer model for convolution-free multilabel classification of satellite imagery in deforestation monitoring, *IEEE Transactions on Neural Networks and Learning Systems*, 2022
- [35] M KASELIMI, A TEMENOS, I RALLIS, A DOULAMIS, N DOULAMIS, M BIMPAS, A AGGELI. Using mHealth Technologies to Promote Public Health and Well-Being in Urban Areas with Blue-Green Solutions
- [36] M KASELIMI, I RALLIS, A DOULAMIS, N DOULAMIS, M BIMPAS, A AGGELI, E PROTOPAPADAKIS, E SARDIS. Pervasive Monitoring of Public Health and Well-Being in Urban Areas with Blue-Green Solutions

- [37] M Kaselimi, A Voulodimos, N Doulamis, A Doulamis, D Delikaraoglou. Deep recurrent neural networks for ionospheric variations estimation using gnss measurements, *IEEE Transactions on Geoscience and Remote Sensing* 60, 1-15, 2021
- [38] A Voulodimos, E Protopapadakis, I Katsamenis, A Doulamis, N Doulamis. Deep learning models for COVID-19 infected area segmentation in CT images, the 14th PErvasive technologies related to assistive environments conference, 2021
- [39] K Makantasis, A Georgogiannis, A Voulodimos, I Georgoulas, A Doulamis, N Doulamis. Rank-r fnn: A tensor-based learning model for high-order data classification, *IEEE Access* 9, 58609-58620, 2021
- [40] A Doulamis, N Doulamis, A Angeli, A Lazaris, S Luthman, M Jayapala, G Silbernagel, A Napp, I Lazarou... A non-invasive photonics-based device for monitoring of diabetic foot ulcers: architectural/sensorial components & technical specifications, *Inventions* 6 (2), 27, 2021
- [41] E Protopapadakis, A Doulamis, N Doulamis, E Maltezos. Stacked autoencoders driven by semi-supervised learning for building extraction from near infrared remote sensing imagery, *Remote Sensing* 13 (3), 371, 2021
- [42] K Makantasis, A Voulodimos, A Doulamis, N Bakalos, N Doulamis. Space-time domain tensor neural networks: An application on human pose classification, 2020 25th International Conference on Pattern Recognition (ICPR), 4688-4695, 2021
- [43] N Bakalos, A Voulodimos, N Doulamis, A Doulamis, K Papatirio, M Bimpas. Fusing RGB and thermal imagery with channel state information for abnormal activity detection using multimodal bidirectional LSTM, *Cyber-Physical Security for Critical Infrastructures Protection*, 2021
- [44] M Kaselimi, N Doulamis, A Voulodimos, A Doulamis, E Protopapadakis. EnerGAN++: A generative adversarial gated recurrent network for robust energy disaggregation, *IEEE Open Journal of Signal Processing* 2, 1-16, 2020
- [45] I Katsamenis, E Protopapadakis, A Voulodimos, A Doulamis, N Doulamis. Transfer learning for COVID-19 pneumonia detection and classification in chest X-ray images, *Proceedings of the 24th Pan-Hellenic Conference on Informatics*, 170-174, 2020
- [46] E Protopapadakis, A Garbi, A Malamou, M Kaselimi, Z Pontikas, A Doulamis, N Doulamis, K Vasilakis. Creating a Baseline for Robust Energy Savings Estimation in Households, *Proceedings of the 24th Pan-Hellenic Conference on Informatics*, 131-134, 2020
- [47] I Katsamenis, E Protopapadakis, A Doulamis, N Doulamis, A Voulodimos. Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation, *International Symposium on Visual Computing*, 160-169, 2020
- [48] M Kaselimi, N Doulamis, A Doulamis, D Delikaraoglou. A sequence-to-sequence temporal convolutional neural network for ionosphere prediction using GNSS observations, *The International Archives of the Photogrammetry, Remote Sensing and Spatial*, 2020

- [49] M Kaselimi, A Voulodimos, E Protopapadakis, N Doulamis, A Doulamis. Energan: A generative adversarial network for energy disaggregation, ICASSP 2020-2020 IEEE International Conference on Acoustics, 2020
- [50] E Protopapadakis, A Doulamis, N Doulamis, E Maltezos. Semi-supervised fine-tuning for deep learning models in remote sensing applications. Advances in Visual Computing: 15th International Symposium, ISVC 2020, 2020.
- [51] N Bakalos, N Doulamis, A Doulamis. Multispectral Monitoring of Microclimate Conditions for Non-destructive Preservation of Cultural Heritage Assets. Strategic Innovative Marketing and Tourism: 8th ICSIMAT, Northern Aegean, 2020.
- [52] M Kaselimi, E Protopapadakis, A Voulodimos, N Doulamis, A Doulamis. Multi-channel recurrent convolutional neural networks for energy disaggregation. IEEE Access 7, 81047-81056, 2019.
- [53] M Kaselimi, N Doulamis, A Doulamis, A Voulodimos, E Protopapadakis. Bayesian-optimized bidirectional LSTM regression model for non-intrusive load monitoring. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech, 2019.
- [54] K Makantasis, A Doulamis, N Doulamis, A Voulodimos. Common mode patterns for supervised tensor subspace learning. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech, 2019.
- [55] M Hajdin, I Kico, M Dolezal, J Chmelik, A Doulamis, F Liarokapis. Digitization and visualization of movements of slovak folk dances. The Challenges of the Digital Transformation in Education: Proceedings of, 2019.
- [56] I Kavouras, E Protopapadakis, A Doulamis, N Doulamis. Skeleton extraction of dance sequences from 3D points using convolutional neural networks based on a new developed C3D visualization interface. The Challenges of the Digital Transformation in Education: Proceedings of, 2019.
- [57] N Doulamis, A Voulodimos, A Doulamis, M Bimpas, A Angeli, N Bakalos, A Giusti, P Philimis, A Varriale. Waterspy: A high sensitivity, portable photonic device for pervasive water quality analysis. Sensors 19 (1), 33, 2018.
- [58] I Sevilla, P Chrobocinski, F Barmpas, F Schmidt, N Kerle, A Kostaridis, A Doulamis, R Russotto. Improving Resilience of Transport Infrastructure to Climate Change and other natural and Manmade events based on the combined use of Terrestrial and Airbone Sensors and Advanced Modelling Tools. CONAMA2018, 11 p., 2018.
- [59] A Doulamis. Automatic 3D Reconstruction From Unstructured Videos Combining Video Summarization and Structure From Motion. Frontiers in ICT 5, 29, 2018.
- [60] V Moulos, G Chatzikyriakos, V Kassouras, A Doulamis, N Doulamis, G Leventakis, T Florakis, T Varvarigou. A robust information life cycle management framework for securing and governing critical infrastructure systems. Inventions 3 (4), 71, 2018.

- [61] E Protopapadakis, A Voulodimos, A Doulamis, S Camarinopoulos, N Doulamis, G Miaoulis. Dance pose identification from motion capture data: a comparison of classifiers. *Technologies* 6 (1), 31, 2018.
- [62] E Maltezos, E Protopapadakis, N Doulamis, A Doulamis, C Ioannidis. Understanding historical cityscapes from aerial imagery through machine learning. *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation ...*, 2018
- [63] E Protopapadakis, A Doulamis, C Ioannidis, S Soile, S Camarinopoulos, N Doulamis. Analyzing Folklore Dance Similarities via Spatial and Temporal Embedding and Signal Analysis. *FIG Commission 3, Spatial Information Management,: Volunteered Geographic ...*, 2017.
- [64] E Protopapadakis, A Voulodimos, A Doulamis, N Doulamis, D Dres, M Bimpas. Stacked autoencoders for outlier detection in over-the-horizon radar signals. *Computational intelligence and neuroscience 2017*, 2017.
- [65] E Protopapadakis, A Voulodimos, A Doulamis, S Camarinopoulos. A Study on the Use of Kinect Sensor in Traditional Folk Dances Recognition via Posture Analysis. *Proceedings of the 10th International Conference on PErvasive Technologies*, 2017
- [66] K Makantasis, A Nikitakis, AD Doulamis, ND Doulamis, I Papaefstathiou. Data-driven background subtraction algorithm for in-camera acceleration in thermal imagery. *IEEE Transactions on Circuits and Systems for Video Technology* 28 (9), 2090-2104, 2017.
- [67] ND Doulamis, AD Doulamis, E Varvarigos. Virtual associations of prosumers for smart energy networks under a renewable split market. *IEEE Transactions on Smart Grid* 9 (6), 6069-6083, 2017.
- [68] AD Doulamis, A Voulodimos, ND Doulamis, S Soile, A Lampropoulos. Transforming Intangible Folkloric Performing Arts into Tangible Choreographic Digital Objects: The Terpsichore Approach. *VISIGRAPP (5: VISAPP)*, 451-460, 2017
- [69] E Protopapadakis, A Grammatikopoulou, A Doulamis, N Grammalidis. Folk dance pattern recognition over depth images acquired via kinect sensor. *The International Archives of the Photogrammetry, Remote Sensing and Spatial*, 2017.
- [70] D Preka, A Doulamis. 3D building modeling in LoD2 using the CityGML standard. *The International Archives of the Photogrammetry, Remote Sensing and Spatial*, 2016.
- [71] K Makantasis, A Doulamis, N Doulamis, K Psychas. Deep learning based human behavior recognition in industrial workflows. *2016 IEEE International conference on image processing (ICIP)*, 1609-1613, 2016
- [72] V Botsis, N Doulamis, A Doulamis, P Makris, E Varvarigos. Efficient clustering of DERs in a virtual association for profit optimization. *2015 Euromicro Conference on Digital System Design*, 494-501, 2015