

Εντοπισμός Θαλάσσιων ρύπων μέσω δορυφορικών εικόνων με χρήση τεχνικών βαθιάς μηχανικής μάθησης

Μεταπτυχιακή Εργασια

του

ΧΑΤΣΙΚΙΑΝ ΑΛΕΝ

Επιβλέπων: Κωνσταντίνος Καράντζαλος Αναπληρωτής Καθηγητής ΕΜΠ

Αθήνα, Σεπτέμβριος 2022



Εντοπισμός Θαλάσσιων ρύπων μέσω δορυφορικών εικόνων με χρήση τεχνικών βαθιάς μηχανικής μάθησης

Μεταπτυχιακή Εργασια

του

ΧΑΤΣΙΚΙΑΝ ΑΛΕΝ

Επιβλέπων: Κωνσταντίνος Καράντζαλος Αναπληρωτής Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 16 Σεπτεμβρίου 2022.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

Κωνσταντίνος Καράντζαλος
Αθανάσιος Βουλόδημος
Διονύσης Ραίτσος
Αναπληρωτής Καθηγητής ΕΜΠ
Επίκουρος Καθηγητής ΕΜΠ



Copyright \bigodot – All rights reserved. Me thn epiquidaxh pantós dikaiώmatos. Xatoikián Alén, 2022.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ενυπογράφως ότι είμαι αποκλειστικός συγγραφέας της παρούσας Πτυχιακής Εργασίας, για την ολοκλήρωση της οποίας κάθε βοήθεια είναι πλήρως αναγνωρισμένη και αναφέρεται λεπτομερώς στην εργασία αυτή. Έχω αναφέρει πλήρως και με σαφείς αναφορές, όλες τις πηγές χρήσης δεδομένων, απόψεων, θέσεων και προτάσεων, ιδεών και λεκτικών αναφορών, είτε κατά κυριολεξία είτε βάσει επιστημονικής παράφρασης. Αναλαμβάνω την προσωπική και ατομική ευθύνη ότι σε περίπτωση αποτυχίας στην υλοποίηση των ανωτέρω δηλωθέντων στοιχείων, είμαι υπόλογος έναντι λογοκλοπής, γεγονός που σημαίνει αποτυχία στην Πτυχιακή μου Εργασία και κατά συνέπεια αποτυχία απόκτησης του Τίτλου Σπουδών, πέραν των λοιπών συνεπειών του νόμου περί πνευματικών δικαιωμάτων. Δηλώνω, συνεπώς, ότι αυτή η Πτυχιακή Εργασία προετοιμάστηκε και ολοκληρώθηκε από εμένα προσωπικά και αποκλειστικά και ότι, αναλαμβάνω πλήρως όλες τις συνέπειες του νόμου στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής άλλης πνευματικής ιδιοκτησίας.

(Υπογραφή)

Χατσικιάν Αλέν 1 Σεπτεμβρίου 2022

Περίληψη

Το θαλάσσιο περιβάλλον, έχει υποστεί σημαντικές πιέσεις που οφείλονται στην ανθρώπινη δρατηριότητα. Τα θαλάσσια απορρίμματα, οι πετρελαιοκηλίδες, άλλες χημικές ουσίες, ο ευτροφισμός και ο θόρυβος αποτελούν τα κύρια είδη ρύπανσης που επιβαρύνουν το θαλάσσιο οικοσύστημα. Τα τελευταία χρόνια υπάρχει ανάγκη πρόληψης και ταχύτερης αντιμετώπισης των φαινομένων ρύπανσης, με την χρήση καινοτομιών. Η τηλεπισκόπηση, που αφορά την επιστήμη παρατήρησης και μελέτης των χαρακτηριστικών της επιφάνειας της γής από απόσταση, παρέχει ψηφιακές εικόνες που έπειτα απο επεξεργασία μπορούν να τροφοδοτηθούν σε μοντέλα τεχνητής νοημοσύνης για την παρακολούθηση και τον ταχύ εντοπισμό της θαλάσσιας ρύπανσης. Το τηλεσκοπικό σύνολο δεδομένων MARIDA, αποτελεί μία σημαντική πηγή γεωδεδομένων απο τον δορυφόρο Sentinel-2, με τα οποία μπορούμε να αξιολογήσουμε και να αναπτύξουμε μοντέλα και τεχνικές αναγνώρισης για τον εντοπισμό θαλάσσιων απορριμμάτων και τον διαχωρισμό τους από τα υπόλοιπα επιπλέοντα υλικά. Στην παρούσα εργασία, αναπτύσονται και αξιολογούνται μοντέλα βαθιάς μάθησης που βασίζονται σε συνελικτικά νευρωνικά δίκτυα, τα οποία τα τελυταία χρόνια εμφανίζουν εξαιρετικά αποτελέσματα στην επίλυση προβλημάτων αναγνώρισης.

Στόχος της διπλωματικής είναι η εύρεση ενός κατάλληλου μοντέλου για την ταξινόμηση - σημασιολογική τμηματοποίηση των διαφόρων επιπλέοντων θαλάσσιων υλικών. Τα συνελικτικά μοντέλα που χρησιμοποιήθηκαν είναι τα DeepLabV3+ και U-Net, επιπροσθέτως αναπτύχθηκε ένα μοντέλο με περιορισμένο οπτικό πεδίο, σχεδιασμένο για την καλύτερη αντιμετώπιση του συγκεκριμένου προβλήματος ταξινόμησης. Για κάθε αρχιτεκτονική παραθέτουμε, πίνακες παραμέτρων που χρησιμοποιήθηκαν για την βέλτιστη αποτελεσματικότητα του εκάστοτε μοντέλου, γραφικές παραστάσεις και αποτελέσματα για τις διαδικασίες εκπαίδευσης και ελέγχου και προβλέψεις σε εικόνες των δεδομένων ελέγχου. Τέλος, για το καλύτερο μοντέλο χρησιμοποιήθηκαν τεχνικές για την ερμηνεία των αποτελεσμάτων με απεικονίσεις για τον τρόπο λειτουργίας του.

Λέξεις Κλειδιά

Θαλάσσια ρύπανση, Τηλεπισκόπηση, Συνελικτικά νευρωνικά δίκτυα (CNN), MARIDA, DeepLabV3+, U-Net

Abstract

The marine environment has suffered significant pressures due to human activity. Marine litter, oil spills, other chemicals, eutrophication and noise are the main types of pollution affecting the marine ecosystem. In recent years, there has been a need to prevent and deal more quickly with pollution phenomenon, with the appliance of innovations. Remote sensing, refers to the process of observing and studying features of the earth's surface from a long distance. Special cameras collect remotely sensed images, which are processed and fed into artificial intelligence (AI) models. The MARIDA telescope dataset is a source of geodata by Sentinel-2 satellite, wherewith we can evaluate and develop models and identification techniques to detect marine debris. Deep learning models, based on convolutional neural networks, are mainly used to solve recognition problems, since they export features based on the data we provide them.

Our goal is to find an appropriate model for classification - semantic segmentation with respect to our input data. The convolutional models we utilized are, DeepLabV3+, U-Net and a convolutional model built from scratch with restricted receptive field. For each architecture we provide tables of used parameters to optimize the effectiveness of each model, graphs and results for the training and testing processes, and predictions over images out of test data. Finally, for the best model, we used techniques to interpret the results with illustrations over the operating principle.

Keywords

Marine Pollution, Remote Sensing, Convolutional Neural Networks, MARIDA, DeepLabV3+, U-Net

Ευχαριστίες

Θα ήθελα καταρχήν να ευχαριστήσω τον καθηγητή κ.Καράντζαλο για την επίδλεψη αυτής της διπλωματικής εργασίας. Επίσης ευχαριστώ ιδιαίτερα τον Ιωάννη Κακογεωργίου, ο οποίος ήταν δίπλα μου καθ' όλη την διάρκεια εκπόνησης της παρούσας διπλωματικής με τις συμβουλές, την καθοδήγηση και τις γνώσεις του. Τέλος, θα ήθελα να ευχαριστήσω τους γονείς και τον αδερφό μου.

Αθήνα, Σεπτέμβριος 2022

Χατσικιάν Αλέν

Περιεχόμενα

Περίληψη			1	
Al	Abstract			
E۱	Ευχαριστίες			
1	Εισ	ανωγή	13	
	1.1	Θαλάσσια Ρύπανση	13	
	1.2	Αντιμετώπιση Θαλάσσιας ρύπανσης με συστήματα Τεχνητής Νοημοσύνης	14	
2	ωзΘ	ορητικό υπόβαθρο	17	
	2.1	Τεχνητή Νοημοσύνη	17	
	2.2	Μηχανική Μάθηση	18	
		2.2.1 Ταξινόμηση (Classification)	19	
		2.2.2 Σημασιολογική Τμηματοποίηση (Semantic Segmentation)	19	
	2.3	Τεχνητά Νευρωνικά Δίκτυα	19	
		2.3.1 Feed-forward Neural Networks	20	
	2.4	Convolutional Neural Networks	21	
		2.4.1 Συνελικτικό Επίπεδο (Convolutional Layer)	22	
		2.4.2 Επίπεδο Υπό-δειγματοληψίας (Pooling Layer)	23	
		2.4.3 Πλήρως Συνδεδεμένο Επίπεδο (Fully Connected Layer)	24	
	2.5	Μοντέλα Βαθιάς Μάθησης	24	
		2.5.1 Skip Connections	25	
		2.5.2 ResNet	25	
		2.5.3 DeepLabV3+	26	
		2.5.4 U-Net	28	
	2.6	Συναρτήσεις Ενεργοποίησης (Activation Functions)	29	
		2.6.1 Συνάρτηση Διορθωμένης Γραμμικής Μονάδας (ReLU)	29	
		2.6.2 Softmax	29	
	2.7	Συναρτήσεις Σφάλματος (Loss Functions)	30	
		2.7.1 Σταυροειδής Εντροπία (Cross Entropy Loss)	31	
		2.7.2 Εστιακή Απώλεια (Focal Loss)	31	
	2.8	Αλγόριθμος Βελτιστοποίησης (Optimization Algorithm)	32	
		2.8.1 Οπισθοδιάδοση (Back-propagation)	32	
		2.8.2 Adam	34	
	2.9	Μετρικές Αξιολόγησης	35	

		2.9.1 Πίνακας Σύγχυσης (Confusion Matrix)	36
		2.9.2 Accuracy	37
		2.9.3 Precision	37
		2.9.4 Recall	37
		2.9.5 F1-score	37
		2.9.6 IoU	38
	2.10	Explainable AI (XAI)	38
3	Περ	ιγραφή Δεδομένων - Μεθοδολογία	41
	3.1	Περιγραφή Δεδομένων	41
	3.2	Παραμετροποίηση U-Net - DeepLabV3+	45
	3.3	Μοντέλο με περιορισμένο οπτικό πεδίο	45
	3.4	Προεπεξεργασία Δεδομένων	46
		3.4.1 Κανονικοποίηση	46
		3.4.2 Προετοιμασία δεδομένων	46
	3.5	Ερμηνεία Αποτελεσμάτων με Explainable AI (XAI)	47
4	Πει	ραματική Διαδικασία - Αποτελέσματα	49
	4.1	DeepLabV3+	49
		4.1.1 Πειραματισμοί	49
		4.1.2 Βέλτιστο Μοντέλο	50
	4.2	Unet	52
		4.2.1 Πειραματισμοί	52
		4.2.2 Βέλτιστο Μοντέλο	54
	4.3	Μοντέλο με περιορισμένο οπτικό πεδίο	56
		4.3.1 Πειραματισμοί	56
		4.3.1.1 Linear Learning Rate Scheduler	56
		4.3.1.2 Exponential Learning Rate Scheduler	57
		4.3.2 Βέλτιστα Μοντέλα	58
		4.3.2.1 Linear Learning Rate Scheduler	58
		4.3.2.2 Exponential Learning Rate Scheduler	61
	4.4	Explainable AI (XAI)	63
	4.5	Σύγκριση Μοντέλων	67
5	Συμ	περάσματα	69
	5.1	Αξιολόγηση Αποτελεσμάτων	69
	5.2	Προβλήματα	70
	5.3	Μελλοντικές προσεγγίσεις	71

Βι6λιογραφία

75

Κατάλογος Σχημάτων

1.1	Επιπλέοντα Πλαστικά στην επιφάνεια της Θάλασσας (Πηγή)	14
2.1	Σύγκριση ενός τεχνητού με ένα πραγματικό νευρωνικό δίκτυο [1]	20
2.2	Single Layer Feed Forward Network (Πηγή)	20
2.3	Multi Layer Feed Forward Network (Πηγή)	21
2.4	Συνελικτικό Επίπεδο [2]	22
2.5	Υπό-Δειγματοληψία (Πηγή)	23
2.6	Πλήρως Συνδεδεμένο Επίπεδο (Πηγή)	24
2.7	Residual Block [3]	25
2.8	Αρχιτεκτονική VGG-16 [4]	26
2.9	Αρχιτεκτονική Resnet-18 [5]	26
2.10	Τεχνική Atrous Convolution	27
2.11	Αρχιτεκτονική DeepLabV3+ [6]	27
2.12	Αρχιτεκτονική Unet [7]	28
2.13	Γραφική Παράσταση Διορθωμένης Γραμμικής Μονάδας ReLU (Πηγή)	29
2.14	Μαθηματική Συνάρτηση της Softmax [8]	30
2.15	Γραφική Παράσταση Softmax (Πηγή)	30
2.16	Σύγκριση Cross Entropy Loss - Focal Loss [9]	32
2.17	Παράδειγμα υπολογισμου οπισθοδιάδοσης	33
2.18	Οπισθοδιάδοση Backpropagation (Πηγή)	34
2.19	Πίνακας Σύγχυσης (Confusion Matrix) (Πηγή)	36
2.20	Τύπος υπολογισμού του ΙοU [10]	38
2 1	Διαδυγασία Πορτουμασίας του δεδομόνου [11]	49
3.1		42
0.2 3 3		44
2.0		44
3.4		47
4.1	Διαγράμματα σφάλματος για τα κριτήρια σφάλματος	52
4.2	Διαγράμματα για τις διάφορες τιμές των Batch	53
4.3	Διαγράμματα Δικτύου U-Net	54
4.4	Αποτελέσματα Δικτύου Unet	56
4.5	Διαγράμματα για τις διάφορες τιμές των Batch	57
4.6	Διαγράμματα για τις διάφορες τιμές των Batch	58
4.7	Διαγράμματα Δικτύου (Linear)	59
4.8	Αποτελέσματα Μοντέλου με περιορισμένο οπτικό πεδίο (Linear)	61

4.9	Διαγράμματα Δικτύου (Εξπονεντιαλ)	62
4.10	Αποτελέσματα Μοντέλου με περιορισμένο οπτικό πεδίο (Exponential)	63
4.11	XAI για Πυκνά Sargassum	64
4.12	ΧΑΙ για Επιπλέοντα Πλαστικά (Marine Debris)	65
4.13	ΧΑΙ για πλοίο, που ταξινομήθηκε ως πλαστικό	66
4.14	XAI για Sargassum, που ταξινομήθηκε ως πλαστικό	66

Κατάλογος Πινάκων

3.1	Κανάλια Sentinel-2, που χρησιμοποιήθηκαν στη πειραματικη διαδικασία [13]	42
3.2	Πίνακας κλάσεων του MARIDA [11]	43
3.3	Πίνακας συγχώνευσης κλάσεων	44
3.4	Αρχιτεκτονική μοντέλου με περιορισμένο οπτικό πεδίο	45
4.1	Πίνακας σύγκρισης αποτελεσμάτων για τα κριτήρια σφάλματος	49
4.2	Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size	50
4.3	Συγκεντρωτικός πίνακας παραμέτρων δικτύου DeepLabV3+	50
4.4	Πίνακας Αποτελεσμάτων δικτύου DeepLabV3+	51
4.5	Πίνακας μέσου όρου μετρικών DeepLabV3+	51
4.6	Πίνακας σύγκρισης αποτελεσμάτων για τα κριτήρια σφάλματος	52
4.7	Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size	53
4.8	Συγκεντρωτικός πίνακας παραμέτρων δικτύου U-Net	54
4.9	Πίνακας Αποτελεσμάτων Unet	55
4.10	Πίνακας μέσου όρου μετρικών Unet	55
4.11	Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size	56
4.12	Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size	57
4.13	Συγκεντρωτικός πίνακας παραμέτρων δικτύου (Linear)	59
4.14	Πίνακας Αποτελεσμάτων Μοντέλου (Linear)	60
4.15	Πίνακας μέσου όρου μετρικών (Linear)	60
4.16	Συγκεντρωτικός πίνακας παραμέτρων δικτύου (Exponential)	61
4.17	Πίνακας Αποτελεσμάτων Μοντέλου (Exponential)	62
4.18	Πίνακας μέσου όρου μετρικών (Exponential)	63
4.19	Πίνακας παραμέτρων Noise Tunnel	63
4.20	Δείκτες Βλάστησης	64
4.21	Πίνακας Σύγκρισης Μοντέλων	67

Κεφάλαιο 1

Εισαγωγή

1.1 Θαλάσσια Ρύπανση

Το θαλάσσιο περιβάλλον υποβάλλεται σε μόλυνση που εμφανίζεται σε πολλές διαφορετικές μορφές, όπως τοξικές χημικές ουσίες (π.χ. οργανικές ενώσεις, μέταλλα, φαρμακευτικά προϊόντα, αέριο), στερεά απόβλητα (π.χ. πλαστικά), αυξημένα θρεπτικά συστατικά (π.χ. νιτρικά και φωσφορικά άλατα) και εισροές ιζημάτων που οφείλονται σε ανθρώπινες δραστηριότητες (π.χ. βιομηχανία, γεωργία, αποψίλωση δασών, απόρριψη λυμάτων, υδατοκαλλιέργεια), ραδιενέργεια, και πετρελαιοκηλίδες [14]. Η θαλάσσια μόλυνση αλλάζει τα φυσικά, χημικά και βιολογικά χαρακτηριστικά των ωκεανών και των παράκτιων ζωνών και δυνητικά απειλεί τους θαλάσσιους οργανισμούς, τα οικοσυστήματα και τη βιοποικιλότητα, και συνεπώς επηρεάζει την ποιότητα και την παραγωγικότητα των θαλάσσιων οικοσυστημάτων. Σε αυτό το πλαίσιο, η μόλυνση που προκαλεί ζημιά ή αρνητικές επιπτώσεις στο θαλάσσιο οικοσύστημα ονομάζεται ρύπανση. Η τελική επίδραση της ρύπανσης στους θαλάσσιους πόρους εξαρτάται από τη μορφή και την ένταση της μόλυνσης, με ορισμένα θαλάσσια περιβάλλοντα, οικοσυστήματα και είδη να είναι πιο ευαίσθητα από άλλα στη ρύπανση. Η Σύμβαση των Ηνωμένων Εθνών του 1982 για το Δίκαιο της Θάλασσας όρισε τη θαλάσσια ρύπανση ως «την εισαγωγή από τον άνθρωπο, άμεσα ή έμμεσα, ουσιών ή ενέργειας στο θαλάσσιο περιβάλλον, η οποία έχει ως αποτέλεσμα ή είναι πιθανό να οδηγήσει σε τέτοιες επιβλαβείς επιπτώσεις όπως βλάβη στους ζωντανούς πόρους και στη θαλάσσια ζωή». Οι πηγές ρύπανσης ανάλογα με τον τρόπο που διοχετεύουν τις ρυπογόνες ουσίες και τα απορρίμματα στους τελικούς αποδέκτες μπορούν να διακριθούν σε σημειακές (άμεσα) και μη-σημειακές (έμμεσα). Στις σημειακές πηγές, τα ρυπογόνα απόβλητα οδηγούνται σε συγκεκριμένες τοποθεσίες μέσω στραγγιστικών αγωγών ή υπονόμων, συνεπώς μπορούν να ανιχνεύονται και να ελέγχονται και κατά συνέπεια, να περιορίζεται σημαντικά η ρύπανση του υδάτινου οικοσυστήματος. Αντίθετα, οι μη-σημειακές πηγές, είναι διάχυτες και δύσκολα αντιμετωπίσιμες. Η απορροή αποτελεί παράδειγμα μη-σημειακής πηγής ρύπανσης, η οποία εμφανίζεται όταν η βροχή ή το χιόνι μετακινεί τους ρύπους από το έδαφος στον ωκεανό. Επίσης, τα εργοστάσια παραγωγής και τα πλοία, αποτελούν μη-σημειακές πηγές ρύπανσης. Τα πρώτα, καθώς, απελευθερώνουν τοξικά απόβλητα και λύματα στον ωκεανό και στα πλοία προκύπτουν διαρροές αργού πετρελαίου λόγω βλαβών. Εν κατακλείδι, υπάρχει συσσώρευση πολλών ρυπογόνων ουσιών στην θάλασσα, με ένα μεγάλο ποσοστό να αποτελούν τα πλαστικά απορρίμματα. Το πλαστικό

είναι απαραίτητο αγαθό στην καθημερινή μας ζωή, ωστόσο, όταν καταλήγει στους ωκεανούς, αποτελεί σοβαρή απειλή. Συνεπώς, κρίνεται απαραίτητη, η χρήση καινοτομιών για την ανίχνευση και συλλογή των επιπλεόντων πλαστικών που δεν αποσυντίθενται (σχήμα 1.1).



Σχήμα 1.1: Επιπλέοντα Πλαστικά στην επιφάνεια της Θάλασσας (Πηγή)

1.2 Αντιμετώπιση Θαλάσσιας ρύπανσης με συστήματα Τεχνητής Νοημοσύνης

Για την ανίχνευση των θαλασσίων πλαστικών απορριμμάτων έχουν αναπτυχθεί αρκετές εφαρμογές που βασίζονται σε τεχνολογίες τηλεπισκόπησης. Οι δορυφόροι, συλλέγουν οπτικά δεδομένα, τα οποία επιτρέπουν την παρατήρηση της επιφάνειας της γης, συνεπώς, ένα μεγάλο μέρος των δορυφορικών εικόνων αφορούν την επιφάνεια της θάλασσας. Τα δεδομένα παρέχονται στην συνέχεια σε εφαρμογές, που χρησιμοποιούν τεχνικές για την ανάλυση της επιφάνειας της θάλασσας. Πολλές απο αυτές τις εφαρμογές, βασίζονται στην χρήση φασματικών δεικτών, για παράδειγμα ο δείκτης FDI, που αναπτύχθηκε απο τους Biermann et al. (2020), για τον διαχωρισμό αντικειμένων που συνυπάρχουν στην επιφάνεια της θάλασσας, όπως πλαστικά, φύκια και αφρό [13]. Επίσης, οι δείκτες NDVI, NRD, χρησιμοποιήθηκαν απο τους Kikaki et al. (2022), για τον διαχωρισμό φυκιών απο επιπλέοντα πλαστικά. Εκτός απο τις εφαρμογές που βασίζονται σε φασματικούς δείκτες, υπάρχουν και αυτές που στηρίζονται σε αλγορίθμους μηχανικής μάθησης. Αυτόνομα και σε πραγματικό χρόνο συστήματα τεχνητής νοημοσύνης χρησιμοποιούνται για τη χαρτογράφηση θαλάσσιων απορριμμάτων με σκοπό την παρακολούθηση και αξιολόγηση της περιβαλλοντικής απειλής [15]. Οι αλγόριθμοι μηχανικής μάθησης, όπως τα τεχνητά νευρωνικά δίκτυα, αυτοματοποιούν τον εντοπισμό και την κατηγοριοποίηση των θαλάσσιων απορριμμάτων. Συγκεκριμένα, η Ocean Cleanup, μια ολλανδική εταιρεία, έχει αναπτύξει ένα σύστημα καθαρισμού ωκεανών χρησιμοποιώντας αλγόριθμους τεχνητής νοημοσύνης για τον εντοπισμό της βέλτιστης τοποθεσίας για την ανάπτυξη δραστηριοτήτων που συμβάλλουν στη συλλογή των σκουπιδιών. Επιπροσθέτως, ερευνητές στην Τσεχία έχουν αναπτύξει αυτοκινούμενα ρομπότ που μπορούν να αναγνωρίζουν πλαστικά, να προσκολλώνται σε αυτά και να τα διασπούν. Τέλος, η χρήση αλγορίθμων τεχνητής νοημοσύνης που παρέχουν τα εφόδια για ταχύτερη ανάλυση δεδομένων, σε συνδυασμό με τεχνικές συλλογής απορριμμάτων, μπορούν να συμβάλλουν στην μείωση των αρνητικών επιπτώσεων στο θαλάσσιο οικοσύστημα [16].

Στόχος αυτής της διπλωματικής είναι η εφαρμογή σύγχρονων τεχνικών για τον εντοπισμό επιπλεόντων πλαστικών στην επιφάνεια της θάλασσας και συγκεκριμένα η εκπαίδευση ενός συστήματος έτσι ώστε να μπορεί να εντοπίζει με την καλύτερη δυνατή ακρίβεια τα επιπλέοντα αντικείμενα μέσω δορυφορικών εικόνων. Αυτό γίνεται με την χρήση μηχανικής μάθησης και συγκεκριμένα αρχιτεκτονικών βαθιάς μηχανικής Máθησης (Deep Learning Architectures). Τα μοντέλα που χρησιμοποιήθηκαν είναι τα DeepLabV3+ / U-Net, για τα οποία έγιναν δοκιμές με διαφορετικά Loss Functions και Learning Rate Schedulers. Στην συνέχεια, αναπτύχθηκε ένα μοντέλο με περιορισμένο οπτικό πεδίο (receptive field), το οποίο θα επικεντρώνεται σε οριοθετημένα εικονοστοιχεία. Τέλος, στο μοντέλο με τις καλύτερες επιδόσεις, γίνεται χρήση των τεχνικών Explainable AI για την κατανόηση λειτουργίας του και για ερμηνεία των δεδομένων εισόδου.



Θεωρητικό υπόβαθρο

Στο κεφάλαιο αυτό, γίνεται πλήρης ανάλυση όλων των τεχνικών - μεθόδων και της δομής των μοντέλων που χρησιμοποιούνται κατά την διάρκεια της συγκεκριμένης εργασίας. Αρχικά, αναλύεται η δομή των Τεχνητών Νευρωνικών Δικτύων, ο τρόπος σύνδεσης και εκπαίδευσής τους και το πεδίο εφαρμογών τους. Στη συνέχεια, γίνεται ανάλυση των Συνελικτικών Νευρωνικών Δικτύων, όπου αναφέρεται λεπτομερώς την αρχιτεκτονική, την τροφοδότηση, τον τρόπο αξιολόγησης τους, αλλά και την διαδικασία βελτιστοποίησης.

2.1 Τεχνητή Νοημοσύνη

Η τεχνητή νοημοσύνη αναφέρεται στην προσπάθεια του ανθρώπου να δώσει σε μία μηχανή την ικανότητα να αναπαράγει τις γνωστικές λειτουργίες ενός ανθρώπου. Η ανθρώπινη νοημοσύνη μπορεί να οριστεί με τρόπο που μια μηχανή μπορεί εύκολα να τη κατανοήσει, να τη μιμηθεί και να εκτελέσει ανάλογες εργασίες. Ο υπολογιστής δέχεται δεδομένα εισόδου, τα επεξεργάζεται και στην συνέχεια καταλήγει σε κάποια απόφαση ή κάποια δράση [17]. Η τεχνητή νοημοσύνη μπορεί χωρίζεται σε δύο κατηγορίες:

- Το Weak AI, είναι ένα σύστημα τεχνητής νοημοσύνης που έχει σχεδιαστεί-εκπαιδευτεί για να ολοκληρώνει με επιτυχία μια συγκεκριμένη εργασία. Τα βιομηχανικά ρομπότ και οι εικονικοί προσωπικοί βοηθοί, όπως το Siri της Apple και την Alexa της Google.
- Το Strong AI, αποτελεί την προγραμματιστική διαδικασία που μπορεί να αναπαράγει τις γνωστικές ικανότητες του ανθρώπινου εγκεφάλου [18].

Η τεχνητή νοημοσύνη συσχετίζεται με πολλούς τεχνολογικούς κλάδους, ένας απο αυτούς είναι η Μηχανική Μάθηση.

2.2 Μηχανική Μάθηση

Μηχανική μάθηση αποτελεί υποπεδίο της επιστήμης των υπολογιστών και της τεχνητής νοημοσύνης, είναι ένα σύνολο μεθόδων που χρησιμοποιούν οι υπολογιστές για να κάνουν - βελτιώσουν προβλέψεις ή συμπεριφορές που βασίζονται σε δεδομένα. Ο αλγόριθμος καθοδηγείται από μια συνάρτηση απώλειας που κατά την διάρκεια της εκπαίδευσης ελαχιστοποιείται. Ένα πλήρως εκπαιδευμένο μοντέλο μηχανικής μάθησης μπορεί στη συνέχεια να χρησιμοποιηθεί για να κάνει προβλέψεις σε νέα δεδομένα [19]. Η μηχανική μάθηση χωρίζεται σε 3 κατηγορίες προβλημάτων:

Επιβλεπόμενη Μάθηση

Στόχος της συγκεκριμένης κατηγορίας είναι ο χαρακτηρισμός δεδομένων με βάση κάποια δεδομένα εκπαίδευσης. Τα δεδομένα εκπαίδευσης αποτελούνται από ένα σύνολο παραδειγμάτων τα οποία χρησιμοποιούνται για την εκπαίδευση του μοντέλου. Στην επιβλεπόμενη μάθηση, κάθε παράδειγμα των δεδομένων εκπαίδευσης αποτελείται από ένα διάνυσμα εισόδου και μια αντίστοιχη τιμή εξόδου (κλάση). Οι αλγόριθμοι επιβλεπόμενης μάθησης αναλύουν τα δεδομένα εκπαίδευσης, ιδανικά παράγουν μοντέλα ικανά να γενικεύουν και να καθορίζουν σωστά την κλάση καινούριων παραδειγμάτων [20]. Χρησιμοποιείται σε προβλήματα, ταξινόμησης (classification), πρόγνωσης (prediction) και διερμηνείας (interpretation).

Μη-επιβλεπόμενη Μάθηση

Αυτό που διακρίνει την μη-επιβλεπόμενη μάθηση από την επιβλεπόμενη μάθηση και την ενισχυτική (ημι-επιβλεπόμενη) μάθηση είναι ότι για τα δεδομένα εισόδου δεν έχουμε κάποιο χαρακτηρισμό (τιμή εξόδου). Στόχος της συγκεκριμένης κατηγορίας είναι ανακάλυψη πιθανής δομής που μπορεί να κρύβεται πίσω από αυτά τα δεδομένα. Το μοντέλο μας δίνει κάποιες τιμές εξόδου, τις οποίες όμως δεν μπορούμε να αξιολογήσουμε, καθώς τα δεδομένα δεν είναι επισημασμένα. Χρησιμοποιείται σε προβλήματα, ανάλυσης συσχετισμών (association analysis) και ομαδοποίησης (clustering).

Ημι-επιβλεπόμενη Μάθηση

Στην συγκεκριμένη κατηγορία ο αλγόριθμος αλληλοεπιδρά με το περιβάλλον για την εύρεση μίας σειράς στρατηγικών ενεργειών. Χρησιμοποιείται κυρίως σε προβλήματα σχεδιασμού-υλοποίησης, όπως για παράδειγμα ο έλεγχος κίνησης ρομπότ στον χώρο της παραγωγής.

Η συγκεκριμένη διπλωματική αφορά πρόβλημα επιβλεπόμενης μάθησης και συγκεκριμένα, ταξινόμησης (classification).

Η μηχανή ξεπερνά τους ανθρώπους σε πολλές εργασίες. Ακόμα κι αν το μηχάνημα δεν είναι τόσο ικανό όσο ο άνθρωπος, έχει πολλά πλεονεκτήματα όσον αφορά την ταχύτητα και την αναπαραγωγιμότητα. Επίσης, η αναπαραγωγή ενός μοντέλου μηχανικής εκμάθησης σε άλλο μηχάνημα είναι γρήγορη και φθηνή, αντίθετα η εκπαίδευση ενός ανθρώπου για μια εργασία μπορεί να διαρκέσει δεκαετίες και είναι πολύ δαπανηρή. Σημαντικό μειονέκτημα της χρήσης μηχανικής μάθησης είναι ότι οι επιμέρους γνώσεις σχετικά με τα δεδομένα και την εργασία που επιλύει το μηχάνημα είναι ιδιαίτερα περίπλοκες και πολλές φορές δυσνόητες για τον άνθρωπο. [19].

2.2.1 Ταξινόμηση (Classification)

Το πρόβλημα της ταξινόμησης αφορά προβλήματα 2 ή περισσότερων κλάσεων. Ένα μοντέλο ταξινόμησης δέχεται δεδομένα εισόδου σε συνδυασμό με την αντίστοιχη κλάση που ανήκουν. Με βάση αυτό το σύνολο δεδομένων επιχειρεί το μοντέλο να εκπαιδευτεί με σκοπό να μπορεί να αποδώσει συμπέρασμα και να γενικευτεί. Απώτερος σκοπός του μοντέλου, είναι να μπορεί να ταξινομεί νέα δεδομένα.

2.2.2 Σημασιολογική Τμηματοποίηση (Semantic Segmentation)

Η σημασιολογική τμηματοποίηση αποτελεί ένα θεμελιώδη πρόβλημα που απασχολεί τους τομείς της όρασης υπολογιστών και της μηχανικής μάθησης. Ο απλούστερος ορισμός του προβλήματος για τη σημασιολογική τμηματοποίηση είναι η διαδικασία αντιστοίχισης των κλάσεων του προβλήματος σε κάθε pixel μιας εικόνας. Αποτελεί ένα απο τα σημαντικότερα βήματα για την αποτελεσματική εξέλιξη πολύπλοκων αυτοματοποιημένων συστημάτων, όπως αυτοκίνητα και drones χωρίς την ύπαρξη οδηγού/χρήστη. Έτσι, δεν είναι περίεργο ότι εκτός από τα επιστημονικά ιδρύματα, κορυφαίες εταιρείες του κλάδου της τεχνητής νοημοσύνης επικεντρώνονται στην εξέλιξη μοντέλων για την αντιμετώπιση τέτοιων προβλημάτων [21].

2.3 Τεχνητά Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks, ANNs) όπως υποδηλώνει και το όνομα είναι βασισμένα στον τρόπο λειτουργίας ενός πραγματικού νευρικού συστήματος, όπως ο εγκέφαλος [22]. Το κύριο κομμάτι αυτής της μεθόδου είναι η πρωτοποριακή δομή του συστήματος, αποτελείται από ένα ευρύ πλήθος πλήρως συνδεδεμένων κόμβων με, που ονομάζονται νευρώνες (neurons). Οι νευρώνες είναι συνδεδεμένοι μεταξύ τους με βάρη και δουλεύουν παράλληλα για την επίλυση πολύπλοκων προβλημάτων (Σχήμα 2.1) [23]. Υπάρχουν δύο κύριες κατηγορίες αρχιτεκτονικής δικτύου ανάλογα με τον είδος των συνδέσεων μεταξύ των νευρώνων, τα feed forward και τα recurrent. Εάν δεν υπάρχει ανατροφοδότηση από τις εξόδους των νευρώνων προς τις εισόδους σε όλο το δίκτυο, τότε αναφέρεται ως feed forward, διαφορετικά αποτελεί ένα recurrent νευρωνικό δίκτυο. Στην παρούσα εργασία, χρησιμοποιούνται τα feed forward δίκτυα.



Σχήμα 2.1: Σύγκριση ενός τεχνητού με ένα πραγματικό νευρωνικό δίκτυο [1]

2.3.1 Feed-forward Neural Networks

Τα feed forward δίκτυα χωρίζονται σε δύο κατηγορίες ανάλογα με τον αριθμό των στρώσεων, ενός επιπέδου ή πολυεπίπεδα. Στο σχήμα 2.2, βλέπουμε ένα νευρωνικό δίκτυο ενός επιπέδου. Αν συμπεριλάβουμε και το επίπεδο εισόδου, βλέπουμε ότι υπάρχουν δύο στρώματα, ωστόσο το επίπεδο εισόδου δεν μετριέται καθώς δεν εκτελείται κάποιος υπολογισμός.



Σχήμα 2.2: Single Layer Feed Forward Network (Πηγή)

Στο σχήμα 2.3 βλέπουμε ότι σε αντίθεση με ένα δίκτυο ενός επιπέδου, υπάρχει τουλάχιστον ένα στρώμα «κρυμμένο» μεταξύ των επιπέδων εισόδου και εξόδου.



Σχήμα 2.3: Multi Layer Feed Forward Network (Πηγή)

Το επίπεδο εισόδου λαμβάνει το διάνυσμα εισόδου που πρόκειται να επεξεργαστεί. Η απαιτούμενη εργασία όπως η πρόβλεψη και η ταξινόμηση εκτελείται από το επίπεδο εξόδου. Ο αριθμός κρυφών επιπέδων που τοποθετούνται μεταξύ επιπέδου εισόδου και εξόδου, αποτελεί ουσιαστικά το βασικό κομμάτι του δικτύου (σχήμα 2.3). Τα δεδομένα κατευθύνονται από το επίπεδο εισόδου στο επίπεδο εξόδου.

Καταλήγουμε στο συμπέρασμα ότι και στα δύο σχήματα 2.2 και 2.3 βλέπουμε feed forward δίκτυα, καθώς κάθε νευρώνας ενός επιπέδου είναι συνδεδεμένος με όλους τους νευρώνες του επόμενου επιπέδου, σε περίπτωση που δεν υπήρχε η πλήρη σύνδεση, θα μιλούσαμε για ένα partially connected δίκτυο [22].

2.4 Convolutional Neural Networks

Σε αυτή την ενότητα γίνεται ανάλυση των Συνελικτικών Νευρωνικών δικτύων (Convolutional Neural Networks). Τα συγκεκριμένα μοντέλα θα μας απασχολήσουν στην συνέχεια της μελέτης μας, καθώς είναι αυτά που χρησιμοποιούνται για την αναγνώριση αντικειμένων στις εικόνες εισόδου. Τα Συνελικτικά Νευρωνικά δίκτυα (CNN) αποτελούνται από πολλά επίπεδα. Κάθε επίπεδο αποτελείται από έναν αριθμό νευρώνων, οι οποίοι δέχονται μια είσοδο, πολλαπλασιάζουν την είσοδο με κάποια βάρος και το αποτέλεσμα της διαδικασίας περνά από μια συνάρτηση ενεργοποίησης. Ανάλογα με τις ενεργοποιήσεις και τα βάρη που έχουν δημιουργηθεί στο πρώτο επίπεδο προχωρά η διαδικασία στο επόμενο επίπεδο με εισόδους τις εξόδους του προηγούμενου επιπέδου. Ενδιάμεσα απο τα συνελικτικά επίπεδα, πολλές φορές παρατηρούνται και επίπεδα υπό-δειγματοληψίας (Pooling Layers). Τέλος, ακολουθεί ένα πλήρως συνδεδεμένο επίπεδο (Fully Connected Layer).

2.4.1 Συνελικτικό Επίπεδο (Convolutional Layer)

Το συνελικτικό επίπεδο (Convolutional Layer) βρίσκεται στα αρχικά επίπεδα του δικτύου και εφαρμόζεται για την εξαγωγή διαφόρων χαρακτηριστικών απο τα δεδομένα εισόδου. Σε αυτό το επίπεδο πραγματοποιείται η μαθηματική διαδικασία της συνέλιξης ανάμεσα στην εικόνα και σε ένα φίλτρο MxM. Το φίλτρο εφαρμόζεται σε όλο το μήκος της εικόνας και δίνει μια νέα εικόνα με πληροφορία για τις γωνίες, τις ακμές και άλλα χαρακτηριστικά (Σχήμα 2.4). Στην συνέχεια, αυτός ο χάρτης χαρακτηριστικών εισάγεται σε άλλα επίπεδα για την εξαγωγή νέων σημαντικών χαρακτηριστικών.

Το επίπεδο συνέλιξης έχει τέσσερις υπερπαράμετρους:

- Τα φίλτρα (Kernels) (K)
- Το μέγεθος (F) των φίλτρων. Κάθε φίλτρο έχει διαστάσεις F x F x D, όπου το D είναι ο αριθμός των καναλιών των δεδομένων εισόδου
- Το Stride (S) είναι το βήμα με το οποίο που κινείται το φίλτρο επάνω στην εικόνα εισόδου
- Η μηδενική επένδυση (Zero Padding) (P) είναι η διαδικασία πλήρωσης του περιγράμματος του δεδομένου εισόδου και είναι μια αποτελεσματική μέθοδος για να μπορεί να πραγματοποιηθεί η διαδικασία της συνέλιξης σε ολόκληρη την εικόνα.

Σε μία εικόνα μεγέθους W x H x D, το επίπεδο συνέλιξης θα επιστρέψει μία εικόνα με νέες διαστάσεις W_{new} x H_{new} x D_{new} [2], όπου:

$$W_{new} = \frac{W - F + 2P}{S} + 1$$
$$H_{new} = \frac{H - F + 2P}{S} + 1$$
$$D_{new} = K$$



Input data

Σχήμα 2.4: Συνελικτικό Επίπεδο [2]

2.4.2 Επίπεδο Υπό-δειγματοληψίας (Pooling Layer)

Το επίπεδο υπό-δειγματοληψίας (Pooling Layer) στοχεύει στη μείωση του αριθμού των παραμέτρων και της υπολογιστικής πολυπλοκότητας του μοντέλου. Αυτό πραγματοποιείται εφαρμόζοντας το σε κάθε χάρτη χαρακτηριστικών (feature map) (Σχήμα 2.5). Ανάλογα με την μέθοδο, υπάρχουν πολλοί τύποι υπό-δειγματοληψίας, θα αναφερθούμε στις δύο πιο γνωστές:

- Μέγιστη Υπό-Δειγματοληψία (Max Pooling): Κρατάμε το μεγαλύτερο στοιχείο απο τον χάρτη των χαρακτηριστικών
- Μέση Υπό-Δειγματοληψία (Average Pooling): Κρατάμε το μέσο όρο των στοιχείων απο τον χάρτη των χαρακτηριστικών [24]

Το επίπεδο Υπό-Δειγματοληψίας έχει δύο υπερπαράμετρους:

- Το μέγεθος F των κελιών, η εικόνα χωρίζεται σε τετράγωνα μεγέθους F x F
- Το Stride (S) είναι το βήμα, με το οποίο είναι χωρισμένα τα κελιά μεταξύ τους (Σχήμα 2.5)

Σε μία εικόνα μεγέθους W x H x D, το επίπεδο υπό-δειγματοληψίας θα επιστρέψει μία εικόνα με νέες διαστάσεις $W_{new} x H_{new} x D_{new}$, όπου:

$$W_{new} = \frac{W - F}{S} + 1$$
$$H_{new} = \frac{H - F}{S} + 1$$
$$D_{new} = K$$



Σχήμα 2.5: Υπό-Δειγματοβηψία (Πηγή)

2.4.3 Πλήρως Συνδεδεμένο Επίπεδο (Fully Connected Layer)

Το πλήρως συνδεδεμένο επίπεδο (Fully Connected layer) είναι παρόμοιο με τον τρόπο που οι νευρώνες διατάσσονται σε ένα νευρωνικό δίκτυο. Κάθε νευρώνας σε ένα πλήρως συνδεδεμένο επίπεδο, συνδέεται άμεσα με κάθε νευρώνα του προηγούμενου αλλά και του επόμενου επιπέδου (Σχήμα 2.6). Το κύριο μειονέκτημα ενός πλήρως συνδεδεμένου επιπέδου, είναι ότι περιλαμβάνει πολλές παραμέτρους, συνεπώς καθιστά αρκετά περίπλοκη τη διαδικασία εκπαίδευσης [25].



Σχήμα 2.6: Πλήρως Συνδεδεμένο Επίπεδο (Πηγή)

2.5 Μοντέλα Βαθιάς Μάθησης

Έχει αποδειχθεί ότι το κλειδί της επιτυχίας των νευρωνικών δικτύων είναι η ικανότητά τους να γίνονται πιο βαθιά, δηλαδή το γεγονός ότι τα δίκτυα μπορούν να αποτελούνται από ένα πολύ μεγάλο αριθμό μη γραμμικών συναρτήσεων. Διαισθητικά, πολλαπλά επίπεδα επιτρέπουν στο δίκτυο να μάθει χαρακτηριστικά στα διάφορα επίπεδα εκπαίδευσης. Ωστόσο, αυτό έχει ένα κόστος, αυτό της επεξήγησης. Δεδομένου την περίπλοκη σύνθεση μεγάλου αριθμού μη γραμμικών συναρτήσεων, την περίπλοκη σύνθεση μεγάλου αριθμού μη γραμμικών συναρτήσεων, τα μοντέλα αυτά δεν είναι εύκολα κατανοητά απο τον άνθρωπο. Έτσι, σε κρίσιμες εφαρμογές, προτιμά κανείς να χρησιμοποιεί λιγότερο ακριβή αλλά ερμηνεύσιμα από τον άνθρωπο μοντέλα, όπως π.χ η γραμμική παλινδρόμηση. Παρόλα αυτά, τα μοντέλα βαθιά μηχανικής μάθησης είναι ιδιαίτερα ικανά στην αποκάλυψη πολύτιμων γνώσεων που διαφορετικά θα ήταν δύσκολο να εξορύξουν οι άνθρωποι από εξαιρετικά μεγάλα σύνολα δεδομένων [26].

2.5.1 Skip Connections

Με τον όρο Skip Connections, μπορούμε να αντιληφθούμε ότι πρόκειται για μία διαδικασία που έχει ως αποτέλεσμα την αποφυγή κάποιων επιπέδων μέσα στο δίκτυο. Η ιδέα του Skip Connections είναι η σύνδεση της εισόδου ενός επιπέδου απευθείας στην έξοδο ενός επόμενου, παραλείποντας ενδιάμεσα επίπεδα, δημιουργώντας έτσι ένα Residual Block. Παρατηρώντας την εικόνα 2.7, το x είναι η είσοδος στο αρχικό επίπεδο και αν υποθέσουμε ότι η έξοδος από τα παραλειπόμενα επίπεδα είναι F(x). Τότε μπορούμε να πούμε ότι η έξοδος του Residual Block θα είναι F(x) + x [3].



Σχήμα 2.7: Residual Block [3]

2.5.2 ResNet

Το ResNet [3] παρουσιάζεται ως ένα μοντέλο βαθιάς μάθησης, με ιδιαίτερη αρχιτεκτονική, καθώς η απόδοση του δεν μειώνεται με την αύξηση των κρυφών επιπέδων και δεν εμφανίζεται το πρόβλημα της εξαφανιζόμενης κλίσης (Vanishing Gradient). Στα απλά δίκτυα βαθμιάς μάθησης η μετάδοση της πληροφορίας είναι ακολουθιακή (Sequential), δηλαδή η έξοδος κάθε επιπέδου συνδέεται με την είσοδο του ακριβώς επόμενου. Ένα τέτοιο απλό ακουλουθιακό μοντέλο είναι το VGG-16 (Σχήμα 2.8). Αντίθετα, στα ResNet δίκτυα η πληροφορία έχει τη δυνατότητα να παρακάμψει επίπεδα (Skip Connections). Η παράκαμψη επιπέδων, υλοποιείται με την χρήση των Residual Block, όπως αναφέραμε και παραπάνω. Η έξοδος ενός block τροφοδοτείται ως είσοδος στο επόμενο. Πολλαπλά συνδεδεμένα Residual Blocks αποτελούν ένα ResNet δίκτυο (Σχήμα 2.9).



Σχήμα 2.8: Αρχιτεκτονική VGG-16 [4]



Σχήμα 2.9: Αρχιτεκτονική Resnet-18 [5]

2.5.3 DeepLabV3+

Το DeepLabV3+ [6] είναι ένα μοντέλο σημασιολογικής τμηματοποίησης και αποτελεί μία βελτιωμένη έκδοση του μοντέλου DeepLabV3. Για την πλήρη κατανόηση του δικτύου DeepLabV3+ αρκεί να εστιάσουμε σε τρία στοιχεία, την αρχιτεκτονική του ResNet, τις τεχνικές Atrous Convolution και Atrous Spatial Pyramid Pooling (ASPP). Το δίκτυο χρησιμοποιεί ένα προ εκπαιδευμένο ResNet πάνω στο ImageNet ως κύριο δίκτυο εξαγωγής χαρακτηριστικών του και ένα Residual Block για εκμάθηση χαρακτηριστικών πολλαπλής κλίμακας. Το Atrous Convolution εισάγεται στο DeepLabV3+ ως εργαλείο προσαρμογής/ελέγχου του οπτικού πεδίου της συνέλιξης. Χρησιμοποιείται μια παράμετρος που ονομάζεται dilation rate που προσαρμόζει το οπτικό πεδίο. Είναι μια απλή αλλά ιδιαίτερα ισχυρή τεχνική, καθώς με την χρήση της διευρύνεται το οπτικό πεδίο των φίλτρων, χωρίς να επηρεάζεται ο υπολογισμός ή ο αριθμός των παραμέτρων. Είναι παρόμοιο με την κλασσική συνέλιξη, με μόνη διαφορά ότι το φίλτρο κάνει υπερ-δειγματοληψία με την εισαγωγή μηδενικών μεταξύ δύο διαδοχικών τιμών φίλτρου κατά μήκος κάθε χωρικής διάστασης, παρατηρώντας το σχήμα 2.10 γίνεται πολύ εύκολα κατανοητό. Όσον αφορά το Atrous Spatial Pyramid Pooling (ASPP), το δίκτυο προσθέτει μια σειρά απο Atrous Convolutions με διαφορετικά dilation rate για να γίνεται καταγραφή περισσότερων πληροφοριών.



Σχήμα 2.10: Τεχνική Atrous Convolution



Σχήμα 2.11: Αρχιτεκτονική DeepLabV3+ [6]

2.5.4 U-Net

Το U-Net [7] αποτελεί και αυτό ένα δίκτυο για σημασιολογική τμηματοποίηση (semantic Segmentation). Το δίκτυο χωρίζεται σε δύο διαδρομές, μια συσταλτική (contracting path) και μια επεκτατική (expansive path), όπου η εικόνα συμπιέζεται και εκτονώνεται, αντίστοιχα. Στο συγκεκριμένο δίκτυο, χρησιμοποιείται η παράκαμψη συνδέσεων (skip connections), ο λόγος για αυτές τις παρακάμψεις είναι ότι παρέχουν μια αδιάκοπη ροή κλίσης από το πρώτο ως το τελευταίο επίπεδο, αντιμετωπίζοντας το πρόβλημα του vanishing gradient. Το μονοπάτι της συστολής ακολουθεί την αρχιτεκτονική των συνελικτικών δικτύων. Αποτελείται από επαναλαμβανόμενες Συνελίξεις 3×3, τις οποίες ακολουθεί η συνάρτηση ενεργοποίησης (ReLU) κι η εφαρμογή Μέγιστης Υπό-δειγματοληψίας (Max Pooling). Τα κανάλια των χαρακτηριστικών μετά την διαδικασία της υπό-δειγματοληψίας διπλασιάζονται. Κάθε βήμα στο μονοπάτι της διαστολής περιέχει υπερ-δειγματοληψία χαρακτηριστικών, η οποία ακολουθείται από μία Συνέλιξη 2×2 που μειώνει στο μισό τον αριθμό των καναλιών, μια συνένωση με τα αντίστοιχα χαρακτηριστικά που έχουν αποκοπεί από το μονοπάτι της συστολής και δύο Συνελίξεις 3×3, που τις ακολουθεί η συνάρτηση ενεργοποίησης (ReLU). Στο τελικό επίπεδο, χρησιμοποιείται μία Συνέλιξη 1×1 για την συσχέτιση των διανυσμάτων στον επιθυμητό αριθμό κλάσεων. Το δίκτυο περιέχει 23 συνελικτικά επίπεδα (Σχήμα 2.12).



Σχήμα 2.12: Αρχιτεκτονική Unet [7]

2.6 Συναρτήσεις Ενεργοποίησης (Activation Functions)

Οι συναρτήσεις ενεργοποίησης καθορίζουν το αποτέλεσμα μίας σειράς δεδομένων εισόδου που δέχονται. Στα νευρωνικά δίκτυα τα αποτελέσματα αυτά αναπαράγουν κάποια απόφασή. Οι συναρτήσεις ενεργοποίησης που χρησιμοποιούμε σε αυτή την διπλωματική είναι:

2.6.1 Συνάρτηση Διορθωμένης Γραμμικής Μονάδας (ReLU)

Η Συνάρτηση Διορθωμένης Γραμμικής Μονάδας (ReLU) είναι μη γραμμική συνάρτηση της μορφής f(x) = max(0, x) δίνει την τιμή του x αν το x είναι θετικό αλλιώς 0. Το μεγάλο πλεονέκτημα της είναι ότι έχει μικρότερη υπολογιστική ανάγκη σε σχέση με άλλες συναρτήσεις ενεργοποίησης, καθώς έχει λιγότερες πράξεις, και επίσης λόγω της φύσης της ενεργοποιεί μόνο τους θετικούς νευρώνες, στους αρνητικούς και αυτούς που ισούνται με 0, δίνει απευθείας την τιμή 0. Εξ ορισμού, η συνάρτηση ορίζεται στο $[0, \infty]$, αυτό έχει σαν συνέπεια να σταματάει να λειτουργεί ορθά (Σχήμα 2.13) [27].



Σχήμα 2.13: Γραφική Παράσταση Διορδωμένης Γραμμικής Μονάδας ReLU (Πηγή)

2.6.2 Softmax

Η Softmax είναι μια μαθηματική συνάρτηση που μετατρέπει ένα διάνυσμα αριθμών σε ένα διάνυσμα πιθανοτήτων, αντίστοιχα για κάθε κλάση. Η πιο κοινή χρήση της συνάρτησης είναι στην εφαρμοσμένη μηχανική μάθηση, όπου χρησιμοποιείται ως συνάρτηση ενεργοποίησης σε μοντέλα νευρωνικών δικτύων. Συγκεκριμένα, το δίκτυο ρυθμίζεται να εξάγει τιμές, μία για κάθε κλάση για την διαδικασία της ταξινόμησης και η συνάρτηση Softmax χρησιμοποιείται για την μετατροπή αυτών των τιμών σε πιθανότητες που αθροίζουν στο 1 [28]. Ο τύπος της Softmax (Σχήμα 2.14):

$$\sigma(ec{z})_i \, = \, rac{e^{\, z_i}}{\sum_{j=1}^K \, e^{\, z_j}}$$

Σχήμα 2.14: Μαθηματική Συνάρτηση της Softmax [8]

Τα z_i αποτελούν τα στοιχεία απο το διάνυσμα εισόδου. Στον παρονομαστή βλέπουμε τον όρο κανονικοποίησης που διασφαλίζει ότι όλες οι τιμές εξόδου της συνάρτησης θα αθροίζονται στο 1, αποτελώντας έτσι μια έγκυρη κατανομή πιθανοτήτων.

Αντίστοιχα, βλέπουμε την γραφική παράσταση της Softmax (Σχήμα 2.15):



Σχήμα 2.15: Γραφική Παράσταση Softmax (Πηγή)

2.7 Συναρτήσεις Σφάλματος (Loss Functions)

Οι συναρτήσεις απώλειας (loss functions) σε ένα νευρωνικό δίκτυο ποσοτικοποιούν τη διαφορά μεταξύ του αναμενόμενου αποτελέσματος και του αποτελέσματος που παράγεται από το μοντέλο μηχανικής μάθησης. Με την βοήθεια τους, ελέγχουμε, αν βελτιώνεται η εκπαίδευση κι αν οι πραγματικές τιμές βρίσκονται κοντά με τις τιμές εξόδου. Οι συναρτήσεις σφάλματος που χρησιμοποιείται σε αυτή την διπλωματική είναι:
2.7.1 Σταυροειδής Εντροπία (Cross Entropy Loss)

Η Σταυροειδής Εντροπία (Cross Entropy Loss) χρησιμοποιείται για την προσαρμογή των βαρών του μοντέλου κατά τη διάρκεια της εκπαίδευσης. Ο στόχος είναι η ελαχιστοποίηση του σφάλματος, δηλαδή όσο μικρότερη είναι το σφάλμα τόσο καλύτερο είναι το μοντέλο. Ένα τέλειο μοντέλο έχει απώλεια σταυροειδής εντροπίας ίση με 0. Συνήθως, χρησιμεύει για ταξινομήσεις πολλαπλών κλάσεων. Για τον υπολογισμό της Σταυροειδής Εντροπίας ακολουθούμε το τύπο:

$$\sum_{k}^{K} y^{(k)} log \hat{y}^{(k)}$$

Το K αφορά το πλήθος των κλάσεων του προβλήματος ταξινόμησης. Επίσης, το $y^{(k)}$, παίρνει τιμές 0 ή 1, υποδεικνύοντας εάν η ταξινόμηση έγινε σωστά ή λάθος, αντίστοιχα. Ενώ, το $\hat{y}^{(k)}$, αφορά την προβλεπόμενη πιθανότητα του εκάστοτε δεδομένου να ανήκει στην κλάση k.

2.7.2 Εστιακή Απώλεια (Focal Loss)

Η Εστιακή Απώλεια (Focal Loss) χρησιμοποιείται για την αντιμετώπιση προβλημάτων ταξινόμησης, όπου υπάρχει ακραία ανισορροπία μεταξύ κλάσεων. Αυτή η συνάρτηση σφάλματος, αποτελεί μια απλή επέκταση της Σταυροειδής Εντροπίας (Cross Entropy Loss). Συγκεκριμένα, προστίθεται ένας νέος παράγοντας στο τύπο της Σταυροειδής Εντροπίας, με αποτέλεσμα να προκύπτει ο τύπος:

$$FL(p_t) = -(1 - p_t)^{\gamma} log(p_t)$$

Όπου, το γ είναι ο παράγοντας εστίασης και αποτελεί υπερπαράμετρο προσαρμογής, με $\gamma \in [0, 5]$ (Σχήμα 2.16). Συνεπώς, όταν το $\gamma > 0$ μειώνει τη σχετική απώλεια για καλά ταξινομημένα παραδείγματα, δηλαδή για παραδείγματα όπου $p_t > 0.5$. Έτσι, δίνεται μεγαλύτερη έμφαση σε δυσκολότερα παραδείγματα με εσφαλμένη ταξινόμηση [9].



Σχήμα 2.16: Σύγκριση Cross Entropy Loss - Focal Loss [9]

2.8 Αλγόριθμος Βελτιστοποίησης (Optimization Algorithm)

Ο Αλγόριθμος Βελτιστοποίησης (Optimization Algorithm) χρησιμοποιείται για την ελαχιστοποίηση του σφάλματος κατά της εκπαίδευση. Οι βελτιστοποιητές (Optimizers) μας είναι χρήσιμοι για την αναδιαμόρφωση του νευρωνικού δικτύου, αναθεωρούν βάρη και ρυθμό μάθηση, με σκοπό την μείωση του συνολικού σφάλματος. Στην συγκεκριμένη διπλωματική χρησιμοποιείται ο Adam.

2.8.1 Οπισθοδιάδοση (Back-propagation)

Η οπισθοδιάδοση (Backpropagation), είναι ένας αλγόριθμος για την εκπαίδευση τεχνητών νευρωνικών δικτύων με χρήση Gradient Descent. Με δεδομένο το νευρωνικό δίκτυο και μια συνάρτηση σφάλματος, η μέθοδος υπολογίζει την κλίση της συνάρτησης σφάλματος σε σχέση με τα βάρη του νευρωνικού δικτύου. Είναι στενά συνδεδεμένη με τον αλγόριθμο Gauss-Newton. Το δίκτυο ξεκινά να εκπαιδεύεται με τυχαία βάρη, εάν το αποτέλεσμα που θα λάβουμε δεν είναι το επιθυμητό τότε ανανεώνονται τα βάρη για τη βελτίωση του δικτύου έως ότου να είναι σε θέση να εκτελεί ικανοποιητικά το έργο για το οποίο εκπαιδεύεται. Η ίδια διαδικασία επαναλαμβάνεται πολλές φορές έτσι ώστε σταδιακά το λάθος να ελαττώνεται μέχρις ότου να γίνει αμελητέο (Σχήμα 2.18). Ολοκληρώνοντας αυτή την διαδικασία λέμε ότι το δίκτυο εκπαιδεύτηκε με ακρίβεια στα δεδομένα εισόδου. Θα δούμε αναλυτικά την διαδικασία ανανέωσης βαρών στο παρακάτω νευρωνικό δίκτυο (Σχήμα 2.17):



Σχήμα 2.17: Παράδειγμα υποβογισμου οπισθοδιάδοσης

Μετά το πρώτο πέρασμα του δικτύου οι τιμές των Hidden Layers και του Output, θα είναι οι εξής:

 $h_1 = w_1 * x_1 + w_2 * x_2$ $h_2 = w_3 * x_1 + w_4 * x_2$ $o_1 = w_5 * h_1 + w_6 * h_2$

Στην περίπτωση που δεν είναι επιθυμητό το αποτέλεσμα ξεκινάει η διαδικασία ανανέωσης των βαρών. Αρχικά, ορίζουμε την συνάρτηση σφάλματος, που είναι το Μέσο Τετραγωνικό Σφάλμα (MSE) και ορίζεται ως: Error(MSE) = $\frac{1}{n} \sum_{1}^{n} (y_{true} - y_{pred})^2$. Στην συνέχεια, πρέπει να υπολογιστεί το: $\frac{\partial Error}{\partial w_l}$. Παρατηρούμε την συνάρτηση σφάλματος και δεν βλέπουμε συσχέτιση με τα βάρη, συνεπώς δεν παραγωγίζεται ως προς κάποιο βάρος, έτσι, χρησιμοποιείται ο κανόνας της αλυσίδας. Για παράδειγμα, αν θέλουμε να υπολογίσουμε το w_1 γράφουμε το $\frac{\partial Error}{\partial w_l}$ ως εξής:

$$\frac{\partial Error}{\partial w_1} = \frac{\partial Error}{\partial y_{pred}} * \frac{\partial y_{pred}}{\partial h_1} * \frac{\partial h_1}{\partial w_1}$$

Με την χρήση του κανόνα της αλυσίδας, υπολογίζεται εύκολα η μερική του σφάλματος ως προς το κάθε βάρος. Τέλος, για την ανανέωση του εκάστοτε βάρους υπολογίζουμε το : $w_i^{new} = w_i - \eta * \frac{\partial Error}{\partial w_i}$, όπου το η, είναι ο ρυθμός μάθησης (learning rate). Σε περίπτωση που το $\frac{\partial Error}{\partial w_i}$ είναι θετικό, τότε το w_i μειώνεται, αντίθετα αν το $\frac{\partial Error}{\partial w_i}$ είναι αρνητικό τότε το w_i αυξάνεται [29].



Σχήμα 2.18: Οπισθοδιάδοση Backpropagation (Πηγή)

2.8.2 Adam

Ο Adam, εφευρέθηκε απο τους Kingma και Ba, είναι ένα στοχαστικό εργαλείο βελτιστοποίησης κλίσης που χρησιμοποιεί μόνο τις δύο πρώτες ροπές της κλίσης και υπολογίζει τον μέσο όρο τους. Εξομαλύνει τα βήματα του Gradient Descent έτσι ώστε να μπορεί να ακολουθήσει μια λιγότερο θορυβώδη διαδρομή και να συγκλίνει πιο γρήγορα. Ένα μεγάλο θετικό με το συγκεκριμένο αλγόριθμο είναι η μικρή απαίτηση μνήμης. Η μέθοδός έχει σχεδιαστεί για να συνδυάζει τα πλεονεκτήματα δύο Gradient Descent αλγορίθμων: του Momentum και του RMSProp [30].

Ο αλγόριθμος Momentum χρησιμοποιείται για την επιτάχυνση του αλγόριθμου gradient descent λαμβάνοντας υπόψη τον εκθετικό σταθμισμένο μέσο όρο των κλίσεων. Η χρήση μέσων όρων κάνει τον αλγόριθμο να συγκλίνει προς την ελάχιστη τιμή με ταχύτερο ρυθμό. Για τον υπολογισμό των βαρών χρησιμοποιούμαι τους τύπους:

$$w_{t+1} = w_t - a * m_t$$
$$m_t = b * m_{t-1} + (1-b) * \frac{\partial Error}{\partial w_t}$$

όπου, m_t είναι το άθροισμα των κλίσεων την χρονική στιγμή t, a_t είναι ο ρυθμός μάθησης (learning rate), b μία σταθερά στο [0, 1] και *∂Error* η μερική παράγωγος του σφάλματος.

Αντίστοιχα, ο RMSProp είναι ένας αλγόριθμος εκμάθησης που προσπαθεί να βελτιώσει το AdaGrad. Ο RMSProp χρησιμοποιεί τον εκθετικό κινούμενο μέσο όρο. Για τον υπολογισμό των βαρών χρησιμοποιούμαι τους τύπους:

$$w_{t+1} = w_t - \frac{a_t}{(u_t + \epsilon)^{1/2}} * \frac{\partial Error}{\partial w_t}$$
$$u_t = b * u_{t-1} + (1 - b) * (\frac{\partial Error}{\partial w_t})^2$$

όπου, u_t το άθροισμα των τετραγώνων περασμένων κλίσεων, δηλαδή $\frac{\partial Error}{\partial w_{t-1}}$, a_t είναι ο ρυθμός μάθησης (learning rate), b μία σταθερά στο [0, 1], $\partial Error$ η μερική παράγωγος του σφάλματος και ϵ μία μικρή θετική σταθερά.

Συνδυάζοντας τους τύπους που χρησιμοποιούνται στις δύο παραπάνω μεθόδους, παίρνουμε:

$$u_{t} = b_{1} * m_{t-1} + (1 - b_{1}) * (\frac{\partial Error}{\partial w_{t}}) * u_{t} = b_{2} * u_{t-1} + (1 - b_{2}) * (\frac{\partial Error}{\partial w_{t}})^{2}$$

Εφόσον, το m_t και το u_t αρχικοποιούνται και τα δύο ως 0, παρατηρείται ότι αποκτούν μια τάση προκατάληψης προς το 0 καθώς και τα δύο b_1 και $b_2 \approx 1$. Στην συνέχεια, ο βελτιστοποιητής υπολογίζει τα νέα m_t και u_t , απαλείφοντας την μεροληψία. Οι τύποι που χρησιμοποιούνται για τον υπολογισμό των νέων m_t και u_t είναι:

$$\hat{m}_t = \frac{m_t}{1 - b_1^t} * \hat{u}_t = \frac{u_t}{1 - b_2^t}$$

Τέλος, για τον υπολογισμό των βαρών χρησιμοποιούμαι τον ανανεωμένο τύπο:

$$w_t = w_t - \hat{m}_t * \frac{a}{\sqrt{\hat{u}_t} + \epsilon}$$

2.9 Μετρικές Αξιολόγησης

Μετά την ολοκλήρωση της εκπαίδευσης ενός μοντέλου κρίνεται απαραίτητη η γνώση την απόδοσης του. Για το υπολογισμό της απόδοσης ενός μοντέλου χρησιμοποιούμαι τις μετρικές αξιολόγησης. Η αξιολόγηση της απόδοσης δεν γίνεται αποκλειστικά απο μία μετρική. Κάθε μετρική εμβαθύνει σε κάτι διαφορετικό, με αποτέλεσμα να έχουμε μία πιο γενική εικόνα για την εκπαίδευση του μοντέλου. Είναι απαραίτητο να ορίσουμε τις παραμέτρους TP (True Positive), FN (False Negative), FP (False Positive) και TN (True Negative), καθώς είναι απαραίτητες για τον υπολογισμό των μετρικών αξιολόγησης. Στην δικιά μας περίπτωση που έχουμε περισσότερες από δύο κλάσεις εξόδου, οι παραπάνω παράμετροι υπολογίζονται χωριστά για κάθε κλάση i, με τον εξής τρόπο:

- TP_i = όσα ανήκουν στην κλάση (εξόδου) i και ταξινομήθηκαν στην i
- FN_i = όσα ανήκουν στην κλάση (εξόδου) i, αλλά δεν ταξινομήθηκαν σ' αυτήν
- FP_i = όσα δεν ανήκουν στην κλάση (εξόδου) i, αλλά ταξινομήθηκαν στην i
- TN_i = όσα δεν ανήκουν στην κλάση (εξόδου) i και δεν ταξινομήθηκαν σ' αυτήν

Στην συγκεκριμένη διπλωματική χρησιμοποιούμε τις παρακάτω μετρικές αξιολόγησής:

2.9.1 Πίνακας Σύγχυσης (Confusion Matrix)

Στον τομέα της μηχανικής μάθησης και συγκεκριμένα στα προβλήματα ταξινόμησης, ο πίνακας σύγχυσης, είναι ένας πίνακας που απεικονίζει την απόδοσης ενός αλγορίθμου επιβλεπόμενης μάθησης [31]. Κάθε γραμμή του πίνακα αντιπροσωπεύει τα αποτελέσματα πραγματικών κλάσεων, ενώ κάθε στήλη αντιπροσωπεύει τα αποτελέσματα των προβλεπόμενων ή αντίστροφα . Το όνομα προέρχεται από το γεγονός ότι καθιστά εύκολο να δει κανείς εάν το σύστημα συγχέει δύο κατηγορίες (λανθασμένα χαρακτηρίζει μία κλάση ως άλλη) [32]. Η δομή του πίνακα παρουσιάζεται στο σχήμα 2.19.



Σχήμα 2.19: Πίνακας Σύγχυσης (Confusion Matrix) (Πηγή)

2.9.2 Accuracy

Η μετρική Accuracy περιγράφει την απόδοση του μοντέλου σε όλες τις κλάσεις. Είναι ο λόγος του αριθμού των σωστών προβλέψεων προς τον συνολικό αριθμό των δειγμάτων εισόδου (2.1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
(2.1)

2.9.3 Precision

Η μετρική Precision υπολογίζει την ακρίβεια του μοντέλου στην ταξινόμηση ενός δείγματος σε μία κλάση. Είναι ο λόγος των παραδειγμάτων που ταξινομούνται σωστά σε μία κλάση προς τον συνολικό αριθμό των δειγμάτων που ταξινομούνται στην συγκεκριμένη κλάση (είτε σωστά είτε λανθασμένα) (2.2).

$$Precision = \frac{TP}{TP + FP}$$
(2.2)

2.9.4 Recall

Η μετρική Recall μετρά την ικανότητα του μοντέλου να ανιχνεύει θετικά δείγματα απο μία κλάση. Είναι ο λόγος των παραδειγμάτων που ταξινομούνται σωστά σε μία κλάση προς όλες τις σωστές προβλέψεις που θα μπορούσαν να είχαν γίνει για αυτή την κλάση (2.3). Όσο υψηλότερη είναι η ανάκληση, τόσο περισσότερα θετικά δείγματα ανιχνεύονται.

$$Recall = \frac{TP}{TP + FN}$$
(2.3)

2.9.5 F1-score

Η μετρική F1-score είναι μια μετρική που παρέχει έναν συνδυασμό τόσο της ακρίβειας (precision) όσο και της ανάκλησης (recall). Είναι ο αρμονικός μέσος όρος των δύο κλασμάτων (2.4) και είναι η πιο κοινή μετρική που χρησιμοποιείται σε μη ισορροπημένες ταξινομήσεις.

$$F1 - score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}}$$
(2.4)

2.9.6 IoU

To IoU (Intersection over Union) χρησιμοποιείται κατά τον υπολογισμό του Macro Average Precision. Είναι ένας αριθμός από το 0 έως το 1 που καθορίζει την ακρίβεια της επικάλυψης μεταξύ της προβλεπόμενης οριοθετημένης περιοχής και της περιοχής εμπειρικής αλήθειας. Ο υπολογισμός του ΙοU προκύπτει απο την διαίρεση της περιοχής επικάλυψης μεταξύ των οριοθετημένων πλαισίων με την περιοχή της ένωσης αυτών (Σχήμα 2.20) [10].



Σχήμα 2.20: Τύπος υποβογισμού του Ιου [10]

2.10 Explainable AI (XAI)

Η τεχνητή νοημοσύνη συχνά είναι δυσνόητη και δεν μπορεί να εξηγηθεί εύκολα. Το Explainable AI (XAI) προσπαθεί να αποσαφηνίσει μοντέλα και τις επιδόσεις τους έτσι ώστε να είναι κατανοητά για τον άνθρωπο. Το XAI αναφέρεται σε μεθόδους - τεχνικές που επιδιώκουν να παρέχουν πληροφορίες για το αποτέλεσμα ενός μοντέλου μηχανικής μάθησης παρουσιάζοντας ποιοτικά κατανοητούς όρους και απεικονίσεις για τον τρόπο λειτουργίας του μοντέλου [33].

Το Captum είναι μια ανοιχτού κώδικα βιβλιοθήκη για την ερμηνεία μοντέλων που βασίζονται στο PyTorch. Το Captum παρέχει αλγόριθμους, συμπεριλαμβανομένων του Integrated Gradients, για να παρέχει στους προγραμματιστές έναν εύκολο τρόπο να κατανοήσουν ποια χαρακτηριστικά συμβάλλουν στην έξοδο του μοντέλου. Η Integrated Gradients, ορίζει μία συνάρτηση $F : \mathbb{R}^n \to [0, 1]$, που αποτελεί το δίκτυο. Έχουμε τα διανύσματα $x \in \mathbb{R}^n$ και $x' \in \mathbb{R}^n$, που αποτελούν το δεδομένο εισόδου και το επίπεδο αναφοράς. Στην περίπτωση μας, που το δεδομένο εισόδου είναι εικόνα, το επίπεδο αναφοράς, μπορεί να είναι μια μαύρη εικόνα [34]. Για τον υπολογισμό του Integrated Gradients για την *i*-οστή διάσταση χρησιμοποιούμε τον τύπο (2.5):

IntegratedGrads_i(x) =
$$(x_i - x'_i) * \int_{a=0}^{1} \frac{\partial F(x' + a * (x - x'))}{\partial x_i} da$$
 (2.5)

Μια άλλη μέθοδος είναι, η SmoothGrad, η οποία όπως υποδηλώνει και το όνομα της, αφορά την διαδικασία υπολογισμού του μέσου όρου για να εξομαλύνει την επεξήγηση, ουσιαστικά συμβάλλει στην μείωση του σφάλματος της κλίσης. Με τύπο (2.6):

$$\nabla_{smooth} F(x) = E_{\epsilon \sim \mathcal{N}(0, \sigma^2)} \nabla F(x + \epsilon)$$
(2.6)

Χρησιμοποιούμε τις δύο αυτές τεχνικές συγχρόνως, διατηρούνται τα πλεονεκτήματα του Integrated Gradients και μειώνεται ο θόρυβος, με την πρόσθεση θορύβου [35].

Κεφάλαιο 3

Περιγραφή Δεδομένων - Μεθοδολογία

Σε αυτό το κεφάλαιο γίνεται αναλυτική περιγραφή του σύνολού των δεδομένων εισόδου και τη δομή του μοντέλου με το περιορισμένο οπτικό πεδίο που χρησιμοποιείται στο κεφάλαιο της πειραματικής διαδικασίας. Επίσης, διευκρινίζεται η υποχρεωτική παραμετροποίηση των μοντέλων βαθιάς μάθησης (U-Net - DeepLabV3+) για να είναι εφικτή η εκπαίδευση. Τέλος, γίνεται εκτενή αναφορά στην διαδικασία του Explainable AI (XAI).

3.1 Περιγραφή Δεδομένων

Το MARIDA [11] είναι ένα σύνολο ελεύθερων δεδομένων που αποτελείται από επισημασμένα εικονοστοιχεία σε δορυφορικές εικόνες Sentinel-2. Ο δορυφόρος Sentinel-2 προσφέρει παγκόσμια κάλυψη σε παράκτια ύδατα, παρέχει πολυφασματικά δεδομένα και με χρόνο επανεπισκόπησης 2-5 ημερών. Τα δεδομένα για τα επιπλέοντα θαλάσσια πλαστικά απορρίμματα (marine debris), συλλέγονται μέσω αναφερόμενων συμβάντων από επιστήμονες και μέσα κοινωνικής δικτύωσης σε παράκτιες περιοχές και εκβολές ποταμών. Το τρέχον σύνολο δεδομένων αναφοράς στοχεύει στην υποστήριξη πραγματικών επιστημονικών ζητημάτων που μπορούν τελικά όχι μόνο να διευκολύνουν τις ερευνητικές προσπάθειες αλλά και να προσφέρουν λύσεις παρακολούθησης. Έτσι, το MARIDA αποτελείται από ρεαλιστικά, μη ιδανικά δορυφορικά δεδομένα (π.χ. με τον όρο μη ιδανικό, αναφερόμαστε εικόνες που περιέχουν σύννεφα και η θάλασσα δεν βρίσκεται μόνιμα σε κατάσταση ηρεμίας). Το περιεχόμενο του MARIDA, δεν αποτελείται αποκλειστικά και μόνο απο θαλάσσια απορρίμματα στον πίνακα 3.2 παρουσιάζονται οι κλάσεις για τις οποίες παρέχεται επισήμανση (annotation) και το πλήθος των εικόνων που εμφανίζεται. Κατά την διαδικασία της ψηφιοποίησης, υπήρξαν εικονοστοιχεία για τα οποία δεν υπήρξε απόλυτη σιγουριά στην ταξινόμηση τους σε κάποια κλάση για αυτό τον λόγο θεσπίστηκε ένα επίπεδο εμπιστοσύνης. Το κάθε ψηφιοποιημένο εικονοστοιχείο συνοδεύεται με ένα βαθμό εμπιστοσύνης, που καθορίζει την βεβαιότητα με την οποία ανήκει στην αναγραφόμενη κλάση.



Σχήμα 3.1: Διαδικασία Προετοιμασίας των δεδομένων [11]

Οι δορυφόροι Sentinel-2 διαθέτουν ένα όργανο πολλαπλών φασμάτων με 13 φασματικές ζώνες που κυμαίνονται από το ορατό εύρος έως το υπέρυθρο βραχέων κυμάτων (SWIR). Οι ζώνες έχουν μεταβλητή ανάλυση από 10 έως 60 μέτρα και το μήκος κύματος τους καθορίζεται με βάση συγκεκριμένους σκοπούς. Στην συγκεκριμένη εργασία, χρησιμοποιούμε 11 φασματικές ζώνες 3.1.

Κανάλια	Κεντρικ Κύματο	τό Μήκος ς (nm)	Χωρική Ανάλυση (m)
	Α	В	
Band 1 – Coastal aerosol	442.7	442.3	60
Band 2 – Blue	492.4	492.1	10
Band 3 – Green	559.8	559.0	10
Band 4 - Red	664.6	665.0	10
Band 5 – Vegetation Red Edge	704.1	703.8	20
Band 6 – Vegetation Red Edge	740.5	739.1	20
Band 7 – Vegetation Red Edge	782.8	779.7	20
Band 8 – NIR	832.8	833.0	10
Band 8A – Narrow NIR	864.7	864.0	20
Band 11 – SWIR	1613.7	1610.4	20
Band 12 – SWIR	2202.4	2185.7	20

Πίνακας 3.1: Κανάβια Sentinel-2, που χρησιμοποιήθηκαν στη πεφαματικη διαδικασία [13]

Στον πίνακα 3.2, βλέπουμε τις κλάσεις για τις οποίες παρέχεται επισήμανση (annotation) και το πλήθος των εικόνων. Η κλάση Marine Water, είναι εκείνη που εμφανίζεται στο μεγαλύτερο πλήθος των δορυφορικών εικόνων, ενώ οι υπόλοιπες κλάσεις νερού έχουν λιγότερες καταγραφές. Οι επόμενες κλάσεις με τις περισσότερες καταγραφές είναι οι Marine Debris και Turbid Water που εμφανίζονται σε 373 και 220 εικόνες, αντίστοιχα. Οι κλάσεις Sparse Sargassum, Ship, Wakes, Clouds και Mixed Water παρατηρούνται σε περισσότερες απο 100 εικόνες. Τέλος, οι υπόλοιπες κλάσεις διαθέτουν καταγραφές σε πλήθος που δεν ξεπερνάει τις 100 εικόνες.

Κλάση	Περιγραφή Κλάσης	Πλήθος Εικόνων (Patches)
ЕЮПА (Marine Debris)	Επιπλέοντα πλαστικά	373
	Ανθρωπογενή απορρίμματα	
Πυκνά Sargassum	Πυκνές συγκεντρώσεις	49
(Dense Sargassum)	επιπλέοντων Sargassum	
Αραιά Sargassum	Αραιές συγκεντρώσεις	106
(Sparse Sargassum)	επιπλέοντων Sargassum	100
Φυσικά Οργανικά Υλικά	Βλάστραρ και Ξύλο	71
(Natural Organic Material)	Bruellol kut Boko	/1
Πλοία		189
(Ship)	Ζκαφή πλεουμένα ή αγκοροσολήμενα	102
Σύννεφα		101
(Clouds)		101
Θαλασσινό Νερό	Καθαρά Θαλασσιμά Νορά	870
(Marine Water)	Κάθαρο Θάλαθοινο Νέρο	870
Νερό με ιζήματα		51
(Sediment-Laden Water)	Νερό με υψηλές συγκεντρωσεις ιζηματών	51
Αφρός	Αφρός εκβολών ποταμών	50
(Foam)	και παράκτιων περιοχών	59
Θολά Νερά		220
(Turbid Water)		220
Αβαθή Νερά	Παράκτια νερά	64
(Shallow Water)	Βυθισμένη βλάστηση	04
Κύματα	Watana and	E 4
(Waves)	κυματα	54
Σκιές Σύννεφων	Service minute and	71
(Clouds Shadows)	Ζκιές ουννέφων	/1
	Κύματα και αναταράξεις της	
Απονερά Πλοιών	θάλασσας εξαιτίας της	106
(wakes)	κίνησης σκαφών	
Μεικτό Νερό		140
(Mixed Water)	νερο οπιλά σε επιπλεοντά αντικειμένα	140
Σύνολο		1381

Πίνακας 3.2: Πίνακας κλάσεων του MARIDA [11]

Τα δεδομένα, όπως αναφέραμε και νωρίτερα, έχουν μεγάλη ανισορροπία μεταξύ των κλάσεων και σε συνδυασμό με το πλήθος τους, καθιστούν την εκπαίδευση του δικτύου ιδιαίτερα δύσκολη.

Παρατηρώντας τον πίνακα 3.2, αντιλαμβανόμαστε ότι υπάρχουν κλάσεις που παρουσιάζονται ξεχωριστά ωστόσο, μπορούμε να τις δούμε και σαν μία, καθώς είναι μέρος του θαλάσσιου νερού. Συνεπώς, αποφασίζουμαι να συγχνεύσουμε αυτές τις κλάσεις. Έτσι, απο 15 κλάσεις καταλήγουμε σε 11. Συγκεκριμένα, στον πίνακα 3.3 βλέπουμε τις κλάσεις που συγχωνεύτηκαν.



Πίνακας 3.3: Πίνακας συγχώνευσης κλάσεων

Στο παρακάτω σχήμα 3.2 μπορούμε να δούμε μερικές τοποθεσίες (κόκκινες κουκκίδες) όπου υπήρξε εκτενή αναφορά επιπλεόντων πλαστικών στο παγκόσμιο χάρτη.



Σχήμα 3.2: Χάρτης εύρεσης επιπλέοντων πλαστικών [12]

Επιπρόσθετα, στο 3.3 παρουσιάζονται κάποιες RGB εικόνες, για την καλύτερη κατανόηση των δεδομένων εισόδου.



Σχήμα 3.3: RGB εικόνες των δεδομένων εισόδου

3.2 Παραμετροποίηση U-Net - DeepLabV3+

Οι παραμετροποιήσεις στα μοντέλα U-Net - DeepLabV3+, καθιστούν δυνατή την εκπαίδευση των μοντέλων στα δεδομένα εισόδου. Συγκεκριμένα, τροποποιούμαι το πρώτο επίπεδο εισόδου των μοντέλων απο εικόνες 3 καναλιών (RGB) στα 11 κανάλια του δορυφόρου Sentinel-2 και αντίστοιχα, το τελικό επίπεδο ταξινόμησης άλλαξε και αυτό απο 3 σε 11 κανάλια για να εξάγει τις κλάσεις του MARIDA.

3.3 Μοντέλο με περιορισμένο οπτικό πεδίο

Το μοντέλο με το περιορισμένο οπτικό πεδίο το χρησιμοποιούμαι για το κύριο τμήμα της διπλωματικής αποτελείται απο 3 συνελικτικά και 2 fully connected επίπεδα, ενδιάμεσα απο κάθε επίπεδο υπάρχει η συνάρτηση ενεργοποίησης ReLU. Στο δίκτυο δοκιμάστηκαν και άλλες τεχνικές όπως Batch Normalization κτλ, ωστόσο δεν επέφεραν τα επιθυμητά αποτελέσματα και συνεπώς, δεν συμπεριλήφθηκαν στο τελικό μοντέλο. Παρουσιάζουμε το δίκτυο αναλυτικά στον πίνακα 3.4.

Επίπεδα		Πληροφορίες								
Input		Shape: 5x5x11								
	Filters	FiltersKernel SizeStridePadding								
Convolutional 1	32	(3,3)	(1,1)	(1,1)	5x5x32					
ReLU										
Convolutional 2	64	(3,3)	(1,1)	(1,1)	5x5x64					
ReLU										
Convolutional 3	128	(5,5)	(1,1)	(0,0)	1x1x128					
ReLU										
Fully Connected 1					256					
ReLU										
Fully Connected 2					11					

Πίνακας 3.4: Αρχιτεκτονική μουτέβου με περιορισμένο οπτικό πεδίο

Ουσιαστικά, το μοντέλο δέχεται μία εικόνα 5x5, ωστόσο ταξινομεί μόνο το κεντρικό pixel. Η διαδικασία αυτή έχει ως αποτέλεσμα την δημιουργία ενός οπτικού πεδίου (receptive field), η συγκεκριμένη τεχνική είναι μια από τις πιο σημαντικές έννοιες στα Συνελικτικά Νευρωνικά Δίκτυα (CNN) για την αναγνώριση αντικειμένων. Μέσα στο οπτικό πεδίο, όσο πιο κοντά είναι ένα pixel στο κέντρο του περιορισμένου παράθυρου, τόσο περισσότερο επηρεάζει τον υπολογισμό του χαρακτηριστικού εξόδου, χωρίς αυτό να σημαίνει ότι τα γύρω pixel, δεν επηρεάζουν.

Το μοντέλο θέλουμε να έχει πιο απλή δομή σε σχέση με τα μοντέλα DeepLabV3+ και

Unet, για να έχουμε εύκολες υλοποιήσεις στην διαδικασία του Explainable Artificial Intelligence (XAI).

3.4 Προεπεξεργασία Δεδομένων

3.4.1 Κανονικοποίηση

Η εκπαίδευση νευρωνικών δικτύων γίνεται πιο αποτελεσματική με εκτέλεση ορισμένων βημάτων προ- επεξεργασίας στις εισόδους-εξόδους του δικτύου. Η διαδικασία κανονικοποίησης στα δεδομένα εισόδου έχει μεγάλη επίδραση στην προετοιμασία των δεδομένων ώστε να είναι κατάλληλα για την εκπαίδευση. Χρησιμοποιείται για την κλιμάκωση των δεδομένων στο ίδιο εύρος τιμών για κάθε ένα χαρακτηριστικό εισόδου προκειμένου να ελαχιστοποιηθεί η προκατάληψη εντός του νευρικού δίκτυο. Η μη χρήση της τεχνικής της κανονικοποίησης, έχει ως αποτέλεσμα, η εκπαίδευση των νευρωνικών δικτύων να είναι πολύ αργή. Υπάρχουν πολλοί τύποι κανονικοποίησης δεδομένων. Συγκεκριμένα, χρησιμοποιούμαι την κανονικοποίηση μέσου όρου και τυπικής απόκλισης. Αυτή η συνάρτηση κανονικοποιεί τις εισόδους και τις εξόδους έτσι ώστε να έχουν μηδενική μέση τιμή και τυπική απόκλιση ίση με μονάδα. Όταν αυτή η μέθοδος χρησιμοποιείται για την προ-επεξεργασία του σετ εκπαίδευσης, τότε θα προκύψουν νέα δεδομένα εισόδου τα οποία στην συνέχεια θα αποτελέσουν το νέο σετ εκπαίδευσης για το μοντέλο [36].

3.4.2 Προετοιμασία δεδομένων

Οι εικόνες του συνόλου δεδομένων MARIDA είναι διαστάσεων 256x256. Το δίκτυο κατασκευάστηκε να δέχεται είσοδο μικρές περιοχές/ εικόνες διαστάσεων 5x5. Η διαδικασία που ακολουθούμε για κάθε εικόνα είναι να κάνουμε Reflection Padding, με αποτέλεσμα οι διαστάσεις τις εικόνας να μετατραπούν σε 260x260. Το Reflection Padding χρησιμοποιεί ουσιαστικά το περιεχόμενο της κάθε γραμμής του πίνακα για την προσθήκη τιμών ακριβώς δίπλα απο το τελευταίο στοιχείο του αρχικού. Η επιλογή αυτού του τύπου padding δεν έγινε τυχαία, παρατηρήθηκε ότι βελτιώνει την απόδοση του μοντέλου μας, επειδή τα καινούρια δεδομένα θα εξακολουθούν να μοιάζουν με τα αντίστοιχα, αρχικά. Στην συνέχεια, για κάθε pixel των padded εικόνων θα δημιουργήσουμε νέες 5x5 εικόνες, όπου το κεντρικό pixel θα είναι αντίστοιχα ένα pixel απο την αρχική μας εικόνα (Σχήμα 3.4). Σημαντικό, για την ομαλή εκπαίδευση του δικτύου είναι να γνωρίζουμε για κάθε pixel και την κλάση του, για αυτό τον λόγο για την διαδικασία της εκπαίδευσης θα χρησιμοποιήσουμε μόνο τα επισημασμένα (annotated) εικονοστοιχεία, καθώς για αυτά και μόνο γνωρίζουμε με ακρίβεια την κλάση τους. Έτσι, τα μη επισημασμένα εικονοστοιχεία δεν τα λαμβάνουμε υπόψιν. Τέλος, αποθηκεύουμε όλες τις 5x5 εικόνες και ξεκινάμε την διαδικασία εκπαίδευση. Η χρήση αυτής της τεχνικής και γενικότερα του συγκεκριμένου δικτύου, έχει ως αποτέλεσμα την μετατροπή του



προβλήματος απο Semantic Segmentation σε Multi Class Classification.

Σχήμα 3.4: Διαδικασία Προετοιμασίας Εικόνων

3.5 Ερμηνεία Αποτελεσμάτων με Explainable AI (XAI)

Για την διαδικασία του ΧΑΙ, χρησιμοποιούμε την βιβλιοθήκη Captum, όπως αναφέραμε και στο Κεφάλαιο 2.10. Συγκεκριμένα, εξετάζουμε την συνεισφορά των χαρακτηριστικών του κεντρικού εικονοστοιχείου του κάθε παράθυρου εισόδου (Attribution) (σχήμα 3.4). Αντίστοιχα, εξετάζουμε και το μέσο όρο της συνεισφοράς των χαρακτηριστικών σε ολόκληρο το παράθυρο εισόδου (Mean Attribution). Επίσης, χρησιμοποιούμε την μέθοδο Noise Tunnel, η οποία προσθέτει θόρυβο σε κάθε είσοδο και στη συνέχεια εφαρμόζει τον αλγόριθμο ερμηνείας σε καθένα από τα δείγματα, στην συγκεκριμένη περίπτωση ο αλγόριθμος απόδοσης είναι ο Integrated Gradients (2.5). Οι αποδόσεις των δειγμάτων συνδυάζονται με βάση το είδος (nt_type) εξομάλυνσης του Attribution, εμείς χρησιμοποιούμε την SmoothGrad (2.6), που επιστρέφει το μέσο όρο των αποδόσεων του δείγματος. Η μέθοδος Noise Tunnel, πέρα απο το είδος της εξομάλυνσης, έχει κάποιες υπερπαραμέτρους, όπως το πλήθος των παραδειγμάτων (nt_samples), που δημιουργούνται τυχαία ανά δείγμα εισόδου και την τυπική απόκλιση (stdevs) του θορύβου με μηδενικό μέσο όρο που προστίθεται σε κάθε είσοδο.

Κεφάλαιο 4

Πειραματική Διαδικασία - Αποτελέσματα

Σε αυτό το κεφάλαιο παρατίθενται όλοι οι πειραματισμοί που γίναν για όλα τα μοντέλα και τα αναλυτικά αποτελέσματα των βέλτιστων μοντέλων, καθώς και αντίστοιχες εικόνες για την καλύτερη κατανόηση της αποτελεσματικότητας του κάθε μοντέλου. Επίσης, εφαρμόζονται τεχνικές ΧΑΙ που εξηγούμαι τον λόγο για τον οποίο το μοντέλο μας ταξινομεί κάποιο εικονοστοιχείο, σε κάποια συγκεκριμένη κλάση.

4.1 DeepLabV3+

4.1.1 Πειραματισμοί

Σε αυτή την ενότητα αναφέρονται οι πειραματισμοί που εκτελούμαι για την εύρεση ενός ικανοποιητικού μοντέλου DeepLabV3+. Για την διαδικασία αυτή, γίναν δοκιμές στις υπερπαραμέτρους του Learning Rate, Loss Criterion και Batch Size. Συγκεκριμένα, για το Learning Rate μετά απο αρκετές δοκιμές καταλήγουμε στην αρχική τιμή : 2e - 4. Στον παρακάτω πίνακα 4.1 παρουσιάζουμε τα αποτελέσματα του Focal Loss και του Cross Entropy Loss για τις δοκιμές που πραγματοποιούμαι. Στις συγκεκριμένες δοκιμές το Batch size ισούται με 10.

	Loss Criterion						
	Focal Loss Cross Entropy Loss						
Macro IoU	0.44	0.39					
Macro Precision	0.64	0.57					
Macro F1-score	0.53	0.49					

Πίνακας 4.1: Πίνακας σύγκρισης αποτεβεσμάτων για τα κριτήρια σφάβματος

Απο τα παραπάνω αποτελέσματα του πίνακα 4.1, βλέπουμε ότι η συνάρτηση σφάλματος Focal Loss είναι πιο αποτελεσματική απο τη συνάρτηση Cross Entropy Loss. Το Focal Loss είναι μια βελτιωμένη έκδοση του Cross Entropy Loss που προσπαθεί να χειριστεί το πρόβλημα της ανισορροπίας μεταξύ κλάσεων, εκχωρώντας περισσότερα βάρη σε δύσκολα ταξινομημένα παραδείγματα. Συνεπώς, το Focal Loss αφοσιώνεται στη διόρθωση παραδειγμάτων που δεν έχουν ταξινομηθεί σωστά.

Μετά την επιλογή του κριτηρίου σφάλματος (Loss Criterion), θέλαμε να πειραματιστούμε με την υπερπαράμετρο του Batch Size. Στον παρακάτω πίνακα 4.2 παρουσιάζονται οι τιμές που δοκιμάστηκαν μαζί με τα αποτελέσματα τους.

	Batch Size					
	5 10 15					
Macro IoU	0.36	0.44	0.43			
Macro Precision	0.45	0.64	0.52			
Macro F1-score	0.45	0.53	0.53			

Πίνακας 4.2: Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size

Απο τα παραπάνω αποτελέσματα βλέπουμε ότι για Batch Size = 10, έχουμε την βέλτιστη απόδοση του μοντέλου DeepLabV3+ στα δεδομένα εισόδου.

4.1.2 Βέλτιστο Μοντέλο

Η εκπαίδευση του δικτύου γίνεται με batches 10 εικόνων για 30 εποχές και χρησιμοποιείται ο αλγόριθμος βελτιστοποίησης Adam. Σχετικά με το Learning Rate Scheduler χρησιμοποιούμε τον Εκθετικό, καθώς παρατηρήθηκε οτι αποδίδει τα καλύτερα αποτελέσματα. Πολύ σημαντική παράμετρος για την εκπαίδευση του δικτύου, αποτελεί η συνάρτηση απώλειας, που επικεντρώνεται στην αξιολόγηση της ορθότητας των υπολογιζόμενων πιθανοτήτων. Σε αυτό το δίκτυο επιλέχτηκε η συνάρτηση Focal, η οποία αποδίδει καλύτερα απο την Cross Entropy . Για την αποφυγή της υπερπροσαρμογής (overfitting) χρησιμοποιούμε την τεχνική του Early Stopping με υπομονή 5 εποχών, στο τέλος κάθε εποχής γίνεται σύγκριση μεταξύ της τιμής σφάλματος στα δεδομένα επικύρωση (validation), αν υπάρξει μείωση τότε κρατάμε το καλύτερο μοντέλο. Η διαδικασία σταματάει αν εκτελεστούν όλες οι εποχές ή αν εκπληρωθούμε οι εποχές υπομονής. Στη συγκεκριμένη περίπτωση, το δίκτυο διακόπτει την εκπαίδευση στην εποχή 13. Αναλυτικά ο πίνακας παραμέτρων του δικτύου (Πίνακας 4.3).

Parameters						
Batch Size	10					
Epochs	13/30					
Loss Criterion	Focal Loss					
Optimizer	Adam					
Learning Rate Scheduler	Exponential LR					
LR Scheduler Gamma	0.1					
LR Starting Value	2e-4					
Early Stopping	Enabled					
Early Stopping Patience	5 Epochs					

Πίνακας 4.3: Συγκεντρωτικός πίνακας παραμέτρων δικτύου DeepLabV3+

Στον παρακάτω πίνακα 4.4 βλέπουμε αναλυτικά αποτελέσματα για κάθε κλάση για το συγκεκριμένο δίκτυο, κάθε γραμμή του πίνακα αντιπροσωπεύει τα πραγματικά αποτελέσματα για κάθε κλάση, ενώ κάθε στήλη αντιπροσωπεύει τα αποτελέσματα της ταξινόμησης του μοντέλου:

Labels	MD	DenS	SpS	NatOMat	Ship	Clouds	MWater	SWater	Foam	TWater	ShWater
MD	81	0	0	0	0	0	296	0	0	4	0
DenS	96	607	57	0	0	0	0	0	0	0	0
SpS	594	58	140	0	0	2	87	0	0	0	0
NatOMat	31	0	0	0	14	1	9	0	0	16	0
Ship	324	0	0	0	710	21	60	12	0	45	2
Clouds	89	0	0	0	337	21159	10246	0	0	1	1011
MWater	1175	0	3	0	239	112	27469	14	0	1568	39
SWater	0	20	0	0	0	0	0	92315	0	702	0
Foam	90	0	0	48	66	37	61	0	75	10	0
TWater	23	0	0	1	50	0	389	27	0	30021	1715
ShWater	5	0	0	0	6	36	906	0	0	232	1321
IoU	0.03	0.72	0.15	0.0	0.38	0.64	0.64	0.99	0.19	0.86	0.25
Precision	0.03	0.89	0.7	0.0	0.5	0.99	0.69	1.0	1.0	0.92	0.32
F1-score	0.06	0.84	0.26	0.0	0.55	0.78	0.78	1.0	0.32	0.93	0.4

Πίνακας 4.4: Πίνακας Αποτεβεσμάτων δικτύου DeepLabV3+

Macro IoU	Macro Precision	Macro F1-score
0.44	0.64	0.53

Πίνακας 4.5: Πίνακας μέσου όρου μετρικών DeepLabV3+

Τα αποτελέσματα μας βλέπουμε πως ποικίλλουν, υπάρχουν κλάσεις που έχουν αρκετά υψηλές μετρικές, πράγμα που σημαίνει ότι το δίκτυο μπορεί να εκπαιδευτεί αρκετά καλά στην πρόβλεψη τους, αντίθετα υπάρχουν κλάσεις που τα αποτελέσματα των μετρικών τους οριακά αγγίζουν το μηδέν. Παρατηρώντας, με λεπτομέρεια τα αποτελέσματα, γίνεται αντιληπτό ότι κλάσεις που κατέχουν μεγάλη κλίμακα (π.χ. σύννεφα, θαλάσσιο νερό), αντικείμενα μεγάλου όγκου, έχουν αρκετά καλά αποτελέσματα. Απεναντίας, κλάσεις που σαν αντικείμενα είναι μικρότερα (π.χ. πλοία), βλέπουμε ότι το μοντέλο δυσκολεύεται να τα αναγνωρίσει λόγο του μικρού όγκου που κατέχουν σε μία εικόνα μεγέθους 256x256. Όσον αφορά, τα επιπλέοντα θαλάσσια απορρίμματα (marine debris) αντιλαμβανόμαστε ότι το δίκτυο δυσκολεύεται σε μεγάλο βαθμό για την αναγνώριση τους, καθώς μιλάμε για μικροσκοπικά αντικείμενα με άμορφο σχηματισμό.

4.2 Unet

4.2.1 Πειραματισμοί

Σε αυτή την ενότητα αναφέρονται οι πειραματισμοί που εκτελούμαι για την εύρεση ενός ικανοποιητικού μοντέλου U-Net. Για την διαδικασία αυτή, γίναν δοκιμές στις υπερπαραμέτρους του Learning Rate, Loss Criterion και Batch Size. Συγκεκριμένα, για το Learning Rate μετά απο αρκετές δοκιμές καταλήξαμε στην αρχική τιμή: 2*e* – 3.

Στον παρακάτω πίνακα 4.6 παρουσιάζονται τα αποτελέσματα του Focal Loss και του Cross Entropy για τις δοκιμές που κάναμε. Στις συγκεκριμένες δοκιμές το Batch size ισούται με 5.

	Loss Criterion						
	Focal Loss Cross Entropy Loss						
Macro IoU	0.12	0.55					
Macro Precision	0.45	0.80					
Macro F1-score	0.17	0.67					

Πίνακας 4.6: Πίνακας σύγκρισης αποτεβεσμάτων για τα κριτήρια σφάβματος



Σχήμα 4.1: Διαγράμματα σφάβματος για τα κριτήρια σφάβματος

Απο τα παραπάνω αποτελέσματα του πίνακα 4.6, βλέπουμε την ανωτερότητα του Cross Entropy Loss. Παρατηρώντας, τα διαγράμματα 4.1, βλέπουμε ότι στην περίπτωση του Focal, τα σφάλματα εκπαίδευσης και ελέγχου έχουν πολύ μεγάλες τιμές και οι καμπύλες τους δεν είναι ιδιαίτερα ομαλές, σε αντίθεση με το διάγραμμα του Cross Entropy, που παρατηρούμε μία ομαλή μείωση του σφάλματος εκπαίδευσης και ελέγχου.

Μετά την επιλογή του κριτηρίου σφάλματος (Loss Criterion), θέλαμε να προσπαθήσουμε να προσαρμόσουμε και την υπερπαράμετρο του Batch Size. Στον παρακάτω πίνακα 4.7 παρουσιάζονται οι τιμές που δοκιμάστηκαν μαζί με τα αποτελέσματα τους.

	Batch Size				
	5 10 15				
Macro IoU	0.55	0.46	0.53		
Macro Precision	0.80	0.78	0.75		
Macro F1-score	0.67	0.56	0.63		

Πίνακας 4.7: Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size



Σχήμα 4.2: Διαγράμματα για τις διάφορες τιμές των Batch

Απο τα παραπάνω αποτελέσματα του πίνακα 4.7, παρατηρούμε ότι για Batch = 5, προκύπτουν τα καλύτερα αποτελέσματα του μοντέλου. Στα διαγράμματα 4.2 βλέπουμε τις συναρτήσεις των σφαλμάτων που δείχνουν πως και για τις 3 τιμές του Batch το μοντέλο εκπαιδεύεται άρτια, με μοναδική διαφορά το μοντέλο με το μικρότερο Batch να έχει τις περισσότερες απότομες μεταβολές, ωστόσο επιφέρει τα καλύτερα αποτελέσματα. Επίσης, τα διαγράμματα που αφορούν την μετρική ΙοU, δείχνουν ότι τα μοντέλα με μεγαλύτερο Batch Size, έχουν μεγαλύτερη σταθερότητα, διότι δεν παρατηρούνται απότομες μεταβολές. Τα αποτελέσματα είναι λογικά, καθώς γνωρίζουμε ότι μικρότερα μεγέθη Batch έχουν ως αποτέλεσμα την ταχεία μάθηση, ωστόσο η διαδικασία μάθησης είναι άστατη και με υψηλή διακύμανση και μεγαλύτερα μεγέθη Batch επιβραδύνουν τη διαδικασία εκμάθησης, αλλά έχουν μεγαλύτερη σταθερότητα και χαμηλή διακύμανση.

4.2.2 Βέλτιστο Μοντέλο

Η εκπαίδευση του δικτύου γίνεται με batches 5 εικόνων για 30 εποχές. Ο Adam αποτελεί τον αλγόριθμο βελτιστοποίησης, για Learning Rate Scheduler χρησιμοποιούμαι τον Εκθετικό και για συνάρτηση απώλειας Cross Entropy καθώς διαπιστώνουμε καλύτερη απόδοση απο τη Focal. Δεν παρατηρούμε φαινόμενα υπερπροσαρμογής (overfitting), με αποτέλεσμα να μην χρησιμοποιούμαι την τεχνική του Early Stopping. Αναλυτικά ο πίνακας παραμέτρων του δικτύου (Πίνακας 4.8).

Parameters						
Batch Size	5					
Epochs	30					
Loss Criterion	Cross Entropy					
Optimizer	Adam					
Learning Rate Scheduler	Exponential LR					
LR Scheduler Gamma	0.01					
LR Starting Value	2e-3					
Early Stopping	Disabled					

Πίνακας 4.8: Συγκευτρωτικός πίνακας παραμέτρων δικτύου U-Net

Παρουσιάζονται τα διαγράμματα της εκπαίδευσης του δικτύου:



Σχήμα 4.3: Διαγράμματα Δικτύου U-Net

Labels	MD	DenS	SpS	NatOMat	Ship	Clouds	MWater	SWater	Foam	TWater	ShWater
MD	254	0	4	0	0	0	122	0	0	1	0
DenS	4	578	178	0	0	0	0	0	0	0	0
SpS	8	106	737	0	0	0	30	0	0	0	0
NatOMat	31	1	7	0	0	0	10	0	0	4	0
Ship	92	0	0	0	941	0	124	6	3	8	0
Clouds	30	0	3	0	31	22083	10616	6	0	59	15
MWater	153	0	5	0	217	25	28769	0	14	1436	0
SWater	0	0	81	0	0	0	17	92939	0	0	0
Foam	49	0	0	0	59	39	27	36	137	40	0
TWater	0	0	0	0	0	0	903	1253	0	29471	599
ShWater	16	0	0	0	0	0	645	10	0	916	919
IoU	0.33	0.67	0.64	0.0	0.64	0.67	0.67	0.99	0.34	0.85	0.29
Precision	0.4	0.84	0.73	1.0	0.75	1.0	0.7	0.99	0.89	0.92	0.6
F1-score	0.5	0.8	0.78	0.0	0.78	0.8	0.8	0.99	0.51	0.92	0.46

Στον παρακάτω πίνακα 4.9 βλέπουμε αναλυτικά αποτελέσματα για κάθε κλάση για το συγκεκριμένο δίκτυο:

Πίνακας 4.9: Πίνακας Αποτεβεσμάτων Unet

Macro IoU	Macro Precision	Macro F1-score		
0.55	0.80	0.67		

Πίνακας 4.10: Πίνακας μέσου όρου μετρικών Unet

Τα αποτελέσματα δείχνουν ότι το δίκτυο Unet, δίνει καλύτερα αποτελέσματα απο το αντίστοιχο δίκτυο του DeepLabV3+. Υπάρχουν αυξήσεις στις μετρικές όλων των κλάσεων. Συγκεκριμένα, παρατηρούμε βελτίωση στα επιπλέοντα θαλάσσια πλαστικά και στα πλοία, αυτό σημαίνει ότι το U-Net, τα πηγαίνει καλύτερα στην πρόβλεψη αντικειμένων μικρότερης κλίμακας. Ωστόσο, για τα επιπλεοντα πλαστικά, τον αφρό και τα ρηχά νερά, η μετρική ΙοU έχει τιμή μικρότερη του 0.5, συνεπώς το μοντέλο δεν προβλέπει ικανοποιητικά. Για την κλάση Natural Organic Material, το δίκτυο αδυνατεί να αναγνωρίσει με επιτυχία έστω και ένα εικονοστοιχείο, ωστόσο βλέπουμε ότι δεν έχει ταξινομηθεί κανένα εικονοστοιχείο λανθασμένα ως φυσικό οργανικό υλικό, καθώς το precision ισούται με 1. Τέλος, οι μέσοι όροι των μετρικών δείχνουν οτι το μοντέλο αποδίδει ικανοποιητικά, παρόλα αυτά υπάρχουν κλάσεις για τις οποίες παρουσιάζονται αδυναμίες.



Παρακάτω βλέπουμε τις προβλέψεις σε 4 εικόνες των δεδομένων ελέγχου:

Σχήμα 4.4: Αποτεβέσματα Δικτύου Unet

4.3 Μοντέλο με περιορισμένο οπτικό πεδίο

4.3.1 Πειραματισμοί

Σε αυτή την ενότητα αναφέρονται όλες οι ενέργειες που εκτελούμαι για την εύρεση ενός ικανοποιητικού μοντέλου. Για την διαδικασία αυτή, γίναν δοκιμές στις υπερπαραμέτρους του Learning Rate Scheduler και Batch Size. Αρχικά, ξεκινήσαμε με την δοκιμή στους Schedulers, όπου δοκιμάστηκαν ο Γραμμικός (Linear) και ο Εκθετικός (Exponential).

4.3.1.1 Linear Learning Rate Scheduler

Έχοντας επιλέξει τον Γραμμικό Scheduler, προσπαθούμε να προσαρμόσουμε την υπερπαράμετρο του Batch Size. Στον παρακάτω πίνακα 4.11 παρουσιάζονται οι τιμές που δοκιμάστηκαν μαζί με τα αποτελέσματα τους.

	Batch Size					
	500 800 1000					
Macro IoU	0.68	0.68	0.71			
Macro Precision	0.84	0.83	0.85			
Macro F1-score	0.78	0.79	0.81			

Πίνακας 4.11: Πίνακας σύγκρισης αποτεβεσμάτων για τα Batch Size



Σχήμα 4.5: Διαγράμματα για τις διάφορες τιμές των Batch

Απο τα παραπάνω αποτελέσματα του πίνακα 4.11, παρατηρούμε ότι για Batch = 1000, προκύπτουν τα καλύτερα αποτελέσματα του μοντέλου. Στα διαγράμματα 4.5 παρατηρούμε ότι με το Batch Size = 500, δείχνει να είναι πιο αποτελεσματικό το μοντέλο σε σχέση με το Batch = 800, τόσο στα σφάλματα αλλά και στην ακρίβεια πρόβλεψης. Τέλος, βλέπουμε την ανωτερότητα του μοντέλου με το υψηλότερο Batch, καθώς κατέχει το χαμηλότερο σφάλμα ελέγχου αλλά και την μεγαλύτερη ακρίβεια.

4.3.1.2 Exponential Learning Rate Scheduler

Αντίστοιχα, έχοντας επιλέξει τον Εκθετικό Scheduler, προσπαθούμε να προσαρμόσουμε την υπερπαράμετρο του Batch Size. Στον παρακάτω πίνακα 4.12 παρουσιάζονται οι τιμές που δοκιμάστηκαν μαζί με τα αποτελέσματα τους.

	Batch Size			
	200	400	1000	
Macro IoU	0.70	0.71	0.69	
Macro Precision	0.81	0.82	0.82	
Macro F1-score	0.69	0.79	0.79	

Πίνακας 4.12: Πίνακας σύγκρισης αποτελεσμάτων για τα Batch Size



Σχήμα 4.6: Διαγράμματα για τις διάφορες τιμές των Batch

Απο τα παραπάνω αποτελέσματα του πίνακα 4.12, παρατηρούμε ότι για Batch = 400, προκύπτουν τα καλύτερα αποτελέσματα του μοντέλου. Στα διαγράμματα 4.6 παρατηρούμε το σφάλμα ανά εποχή και την ακρίβεια για τις 3 τιμές του Batch Size.

4.3.2 Βέλτιστα Μοντέλα

Σε αυτό το σημείο θα παρουσιάσουμε τα βέλτιστα μοντέλα για κάθε Learning Rate Scheduler.

4.3.2.1 Linear Learning Rate Scheduler

Η εκπαίδευση του δικτύου γίνεται με batches 1000 εικόνων 5x5 για 20 εποχές. Ο Adam αποτελεί τον αλγόριθμο βελτιστοποίησης, για Learning Rate Scheduler χρησιμοποιούμαι τον Γραμμικό και για συνάρτηση απώλειας Focal, διαπιστώνουμε καλύτερη απόδοση απο την Cross Entropy. Δεν παρατηρούμε φαινόμενα υπερπροσαρμογής (overfitting), με αποτέλεσμα να μην χρησιμοποιούμαι την τεχνική του Early Stopping. Αναλυτικά ο πίνακας παραμέτρων του δικτύου (πίνακας 4.13).

Parameters				
Batch Size	1000			
Epochs	20			
Loss Criterion	Focal Loss			
Optimizer	Adam			
Learning Rate Scheduler	Linear LR			
LR Scheduler Gamma	0.5			
LR Starting Value	2e-3			
Early Stopping	Disabled			

Πίνακας 4.13: Συγκευτρωτικός πίνακας παραμέτρων δικτύου (Linear)



Παρουσιάζονται τα διαγράμματα της εκπαίδευσης του δικτύου:

Σχήμα 4.7: Διαγράμματα Δικτύου (Linear)

Παρατηρώντας τα διαγράμματα βλέπουμε ότι όλες οι μετρικές αξιολόγησης βρίσκονται σε υψηλά επίπεδα. Ο ρυθμός μάθησης (learning rate), μειώνεται συνεχώς μέχρι την 14η εποχή, όπου η τιμή του είναι τόσο μικρή που θεωρείται αμελητέα. Αν παρατηρήσουμε το διάγραμμα του σφάλματος, βλέπουμε ότι μετά την 14η εποχή, οι καμπύλες της εκπαίδευσης (μπλε) αλλά και της επαλήθευσης (πορτοκαλί), παραμένουν σταθερές.

Labels	MD	DenS	SpS	NatOMat	Ship	Clouds	MWater	SWater	Foam	TWater	ShWater
MD	358	0	0	0	6	3	13	0	1	0	0
DenS	0	725	30	5	0	0	0	0	0	0	0
SpS	0	74	792	15	0	0	0	0	0	0	0
NatOMat	22	6	0	19	0	0	2	0	0	0	0
Ship	20	0	0	0	1055	28	47	5	13	0	6
Clouds	32	0	0	0	7	28344	4407	0	1	27	25
MWater	59	0	94	1	377	2572	25759	2	4	1592	159
SWater	15	14	0	0	0	7	0	93001	0	0	0
Foam	3	0	0	0	11	0	9	3	355	6	0
TWater	0	0	0	0	0	24	801	73	0	28977	2351
ShWater	3	0	5	4	24	15	252	0	0	327	1877
IoU	0.67	0.85	0.78	0.26	0.66	0.8	0.71	1.0	0.87	0.85	0.37
Precision	0.7	0.89	0.86	0.43	0.71	0.91	0.82	1.0	0.95	0.94	0.42
F1-score	0.8	0.92	0.88	0.41	0.8	0.89	0.83	1.0	0.93	0.92	0.54

Παρακάτω βλέπουμε και τον Confusion Matrix (πίνακα σύγχυσης):

Πίνακας 4.14: Πίνακας Αποτεβεσμάτων Μοντέβου (Linear)

Macro IoU	Macro Precision	Macro F1-score
0.71	0.85	0.81

Πίνακας 4.15: Πίνακας μέσου όρου μετρικών (Linear)

Τα αποτελέσματα δείχνουν ότι το μοντέλο με το γραμμικό scheduler, δίνει πολύ καλύτερα αποτελέσματα απο το δίκτυο U-Net. Υπάρχουν αυξήσεις στις μετρικές όλων των κλάσεων. Πιο συγκεκριμένα, οι μόνες κλάσεις που δεν προβλέπουν ικανοποιητικά, σύμφωνα με την μετρική IoU είναι τα φυσικά οργανικά υλικά και τα ρηχά νερά. Τα επιπλέοντα θαλάσσια πλαστικά και τα πλοία έχουν υψηλές μετρικές, συνεπώς το μοντέλο διαχειρίζεται αντικείμενα μικρότερης κλίμακας. Τέλος, οι μέσοι όροι των μετρικών αποδεικνύουν την ικανότητα του μοντέλου στα δεδομένα εισόδου.

Παρακάτω βλέπουμε τις προβλέψεις σε 4 εικόνες των δεδομένων ελέγχου:



Σχήμα 4.8: Αποτεβέσματα Μοντέβου με περιορισμένο οπτικό πεδίο (Linear)

4.3.2.2 Exponential Learning Rate Scheduler

Η εκπαίδευση του δικτύου γίνεται με batches 400 εικόνων 5x5 για 10 εποχές. Ο Adam αποτελεί τον αλγόριθμο βελτιστοποίησης, για Learning Rate Scheduler χρησιμοποιούμαι τον Γραμμικό και για συνάρτηση απώλειας Focal, διαπιστώνουμε καλύτερη απόδοση απο την Cross Entropy. Αναλυτικά ο πίνακας παραμέτρων του δικτύου (πίνακας 4.16).

Parameters				
Batch Size	400			
Epochs	10			
Loss Criterion	Focal Loss			
Optimizer	Adam			
Learning Rate Scheduler	Exponential LR			
LR Scheduler Gamma	0.5			
LR Starting Value	2e-3			
Early Stopping	Disabled			

Πίνακας 4.16: Συγκευτρωτικός πίνακας παραμέτρων δικτύου (Exponential)

Παρουσιάζονται τα διαγράμματα της εκπαίδευσης του δικτύου:



Σχήμα 4.9: Διαγράμματα Δικτύου (Εξπονευτιαλ)

Παρακάτω	βλέπουμε	και τον	Confusion	Matrix	(πίνακα	σύγχυσης):
----------	----------	---------	-----------	--------	---------	------------

Labels	MD	DenS	SpS	NatOMat	Ship	Clouds	MWater	SWater	Foam	TWater	ShWater
MD	343	0	0	2	6	0	18	0	12	0	0
DenS	0	700	57	3	0	0	0	0	0	0	0
SpS	1	41	828	10	0	0	1	0	0	0	0
NatOMat	16	12	0	15	1	0	4	0	1	0	0
Ship	11	0	0	0	1055	29	58	2	13	1	5
Clouds	34	0	0	0	3	2431	26048	0	2	6	219
MWater	31	0	9	1	311	2431	26048	0	2	1703	83
SWater	0	0	0	0	0	8	0	93029	0	0	0
Foam	2	0	0	0	31	2	16	11	323	0	0
TWater	0	0	0	0	0	23	885	85	0	29039	2194
ShWater	0	0	0	1	10	64	116	0	0	483	1832
IoU	0.72	0.86	0.87	0.23	0.69	0.69	0.66	1.0	0.78	0.84	0.37
Precision	0.78	0.86	0.87	0.23	0.69	0.69	0.66	1.0	0.78	0.84	0.37
F1-score	0.84	0.93	0.93	0.37	0.81	0.82	0.79	1.0	0.88	0.92	0.54

Πίνακας 4.17: Πίνακας Αποτεβεσμάτων Μοντέβου (Exponential)

Macro IoU	Macro Precision	Macro F1-score
0.7	0.82	0.8

Πίνακας 4.18: Πίνακας μέσου όρου μετρικών (Exponential)

Με την χρήση του εκθετικού scheduler, βλέπουμε αύξηση σε όλες τις μετρικές των επιπλεόντων πλαστικών και αντίστοιχα σε κάποιες μετρικές άλλων κλάσεων, όπως Dense, Sparse Sargassum. Οι μέσοι όροι των μετρικών μειώθηκαν, συνεπώς το μοντέλο με τον γραμμικό scheduler, αποτελεί ένα πιο αποτελεσματικό μοντέλο, ωστόσο σε κάποιες επιμέρους κλάσεις η συγκεκριμένη διαφοροποίηση του scheduler προσφέρει καλύτερες προβλέψεις.



Παρακάτω βλέπουμε τις προβλέψεις σε 4 εικόνες των δεδομένων ελέγχου:

Σχήμα 4.10: Αποτελέσματα Μοντέλου με περιορισμένο οπτικό πεδίο (Exponential)

4.4 Explainable AI (XAI)

Σε αυτή την ενότητα γίνεται χρήση εργαλείων ΧΑΙ για την κατανόηση και ερμηνεία των αποφάσεων του μοντέλου. Την συγκεκριμένη διαδικασία την εφαρμόζουμε στο μοντέλο με το Exponential Learning Rate, καθώς είναι το μοντέλο με τα καλύτερα αποτελέσματα όσον αφορά τα επιπλέοντα θαλάσσια πλαστικά. Παραθέτουμε, τον πίνακα παραμέτρων 4.19 του Noise Tunnel, που χρησιμοποιούμε για το Explainable AI.

nt_type	nt_samples	stdevs	
SmoothGrad	10	0.3	

Πίνακας 4.19: Πίνακας παραμέτρων Noise Tunnel

NDVI Normalized Difference Vegetation Index	$NDVI = \frac{NIR - Red}{NIR + Red}$
NRD Near Infrared Red Difference	NRD = NIR – Red

Πίνακας 4.20: Δείκτες Βλάστησης

Αρχικά, θα χρησιμοποιήσουμε την τεχνική του ΧΑΙ για να μπορέσουμε να εξηγήσουμε τον λόγο για τον οποίο το μοντέλο μας χαρακτήρισε σωστά δύο εικονοστοιχεία ως πυκνά sargassum. Συγκεκριμένα, στην εικόνα S2_14_9_18_16PCC_14 των δεδομένων ελέγχου υπάρχει μεγάλη συγκέντρωση πυκνών sargassum. Στο σχήμα 4.11 μπορούμε να δούμε αναλυτικά τις επεξηγηματικές γραφικές παραστάσεις. Αν παρατηρήσουμε το διάγραμμα του Attribution και για τα δύο εικονοστοιχεία, βλέπουμε, πως το μοντέλο δίνει μεγάλη έμφαση στα κανάλια 842 (NIR) και σε μικρότερο βαθμό στα κανάλια 783 (Vegetation Red Edge) και 560 (Green). Αντίστοιχα, για το διάγραμμα του Mean Attribution, παρατηρείται η έμφαση στο κανάλι 842 (NIR). Η διαφορά ανακλαστικότητας μεταξύ του NIR και Red, δηλαδή ο δείκτης NRD (σχήμα 4.20), μπορεί να χρησιμοποιηθεί για τη διάκριση των επιπλεόντων πλαστικών από τα φύκια [37], καθώς οι ερμηνείες που παρέχονται από την μεθοδολογία ΧΑΙ συμβαδίζουν με τον φασματικό δείκτη. Όμοια, ο δείκτης NDVI, βασίζεται στο γεγονός ότι η χλωροφύλλη απορροφά το κόκκινο ενώ η δομή του μεσόφυλλου εκπέμπει το υπέρυθρο. Οι τιμές επομένως κυμαίνονται από -1 έως 1, όπου οι αρνητικές τιμές αντιστοιχούν σε απουσία βλάστησης. Λαμβάνοντας υπόψιν τους δείκτες βλάστησης NDVI και NRD (σχήμα 4.20), παρατηρούμε έντονη βλάστηση στα εικονοστοιχεία και συμβαδίζουν με τις ερμηνείες ΧΑΙ.



Σχήμα 4.11: ΧΑΙ για Πυκνά Sargassum

Αντίστοιχα, θέλουμε να εξηγήσουμε το ίδιο και για τα επιπλέοντα πλαστικά. Στην εικόνα S2_22_12_20_18QYF_0 παρατηρούμε συγκέντρωση αρκετών Marine Debris, τα οποία ταξινομήθηκαν σωστά απο το μοντέλο. Στο σχήμα 4.12 μπορούμε να δούμε αναλυτικά τις επεξηγηματικές γραφικές παραστάσεις. Παρατηρώντας το διάγραμμα του Attribution, βλέπουμε, πως το μοντέλο δίνει μεγάλη έμφαση στα κανάλια 842 (NIR) και 665 (Red). Αντίστοιχα, το διάγραμμα του Mean Attribution βλέπουμε ότι δίνει μεγαλύτερη έμφαση στο κανάλι 665 (Red), σε αντίθεση με το XAI με τα Dense Sargassum. Συνεπώς, οι δείκτες NDVI και NRD, μπορούν να διαχωρίσουν σωστά τα επιπλέοντα πλαστικά.



Σχήμα 4.12: ΧΑΙ για Επιπλέοντα Πλαστικά (Marine Debris)

Εκτός απο τα σωστά ταξινομημένα εικονοστοιχεία, υπάρχουν και αρκετά τα οποία δεν μπόρεσε να ταξινομήσει σωστά το μοντέλο. Κάνουμε την ίδια διαδικασία για να προσπαθήσουμε να κατανοήσουμε τις αιτίες που οδήγησαν το δίκτυο στην λάθος ταξινόμηση τους. Στην εικόνα S2_15_11_20_16PCC_13 παρατηρείται ένα εικονοστοιχείο, το οποίο είναι πλοίο ωστόσο, ταξινομείται σαν πλαστικό. Στο σχήμα 4.13 μπορούμε να δούμε αναλυτικά τις επεξηγηματικές γραφικές παραστάσεις. Παρατηρώντας, την φασματική υπογραφή (Signature), βλέπουμε μία φθίνουσα πορεία, διότι, το εικονοστοιχείο δεν περιέχει αποκλειστικά και μόνο το πλοίο, αλλά και αρκετό θαλάσσιο νερό. Τα επιπλέοντα πλαστικά και τα πλοία παρουσιάζουν παρόμοιες φασματικές ιδιότητες λόγω της πολυμερούς σύνθεσης τους. Για παράδειγμα, μικρά σκάφη και επιπλέοντα πλαστικά μπορούν απεικονίζονται ως μεμονωμένα εικονοστοιχεία, οδηγώντας σε μια δύσκολη διαδικασία ταξινόμησης. Τα παραπάνω έχουν ως αποτέλεσμα να περιπλέκεται το μοντέλο και να μην μπορεί να ταξινομήσει σωστά το συγκεκριμένο εικονοστοιχείο.



Σχήμα 4.13: ΧΑΙ για πλοίο, που ταξινομήθηκε ως πλαστικό

Ομοίως, στην εικόνα S2_14_9_18_16PCC_14 παρατηρούνται δύο εικονοστοιχεία τα οποία δεν μπόρεσε το μοντέλο να ταξινομήσει σωστά. Τα συγκεκριμένα εικονοστοιχεία, αποτελούν πυκνά Sargassum, ωστόσο το μοντέλο τα χαρακτήρισε σαν πλαστικά. Παρατηρώντας, τις φασματικές υπογραφές (Signature), βλέπουμε μία φθίνουσα πορεία, διότι, τα εικονοστοιχεία δεν περιέχουν μόνο Sargassum, αλλά και αρκετό νερό. Επίσης, βλέπουμε τα διαγράμματα των Mean Attribution, που υποδηλώνουν υψηλές τιμές στα κανάλια 665 (Red), 842 (NIR), 560 (Γρεεν). Οι δείκτες NDVI και NRD, σύμφωνα με τα διαγράμματα δεν θα έχουν μεγάλες τιμές, άρα δεν παρατηρούμε έντονη βλάστηση σε αυτές τις περιοχές. Αυτό έχεις ως αποτέλεσμα να επηρεάζει το μοντέλο αρνητικά.



Σχήμα 4.14: XAI για Sargassum, που ταξινομήθηκε ως πβαστικό
4.5 Σύγκριση Μοντέλων

	DeepLabV3+			U-Net			CNN-Linear			CNN-Exponential		
Class	IoU	Prec	F1	IoU	Prec	F1	IoU	Prec	F1	IoU	Prec	F1
MD	0.03	0.03	0.06	0.33	0.40	0.50	0.67	0.70	0.80	0.72	0.78	0.84
DenS	0.72	0.89	0.84	0.67	0.84	0.80	0.85	0.89	0.92	0.86	0.86	0.93
SpS	0.15	0.70	0.26	0.64	0.73	0.78	0.78	0.86	0.88	0.87	0.87	0.93
NatM	0	0	0	0	1	0	0.26	0.43	0.41	0.23	0.23	0.37
Ship	0.38	0.50	0.55	0.64	0.75	0.78	0.66	0.71	0.80	0.69	0.69	0.81
Clouds	0.64	0.69	0.78	0.67	1	0.80	0.80	0.91	0.89	0.69	0.69	0.82
MWater	0.64	0.69	0.78	0.67	0.70	0.80	0.71	0.82	0.83	0.66	0.66	0.79
SLWater	0.99	1	1	0.99	0.99	0.99	1	1	1	1	1	1
Foam	0.19	1	0.32	0.32	0.89	0.51	0.87	0.95	0.93	0.78	0.78	0.88
TWater	0.86	0.92	0.93	0.85	0.92	0.92	0.85	0.94	0.92	0.84	0.84	0.92
SWater	0.25	0.32	0.40	0.29	0.60	0.46	0.37	0.42	0.54	0.37	0.37	0.54
Average	0.44	0.64	0.53	0.55	0.80	0.67	0.71	0.85	0.81	0.70	0.82	0.80

Στον πίνακα 4.21 παραθέτουμε τα αποτελέσματα όλων των μοντέλων αντίστοιχα για κάθε κλάση και την μέση τιμή των μετρικών.

Πίνακας 4.21: Πίνακας Σύγκρισης Μουτέλων

Λαμβάνοντας υπόψη τα παραπάνω αποτελέσματα 4.21 το μοντέλο που επιλέχθηκε ως το πιο κατάλληλο για την κατάτμηση των εικόνων είναι το δίκτυο με τον Γραμμικό Scheduler. Οι επιλογή αυτή έγινε, καθώς το συγκεκριμένο μοντέλο παρουσίασε τον καλύτερο μέσο όρο στις μετρικές IoU, Precision και F1. Επίσης, τα ποιοτικά αποτελέσματα του μοντέλου εμφανίζουν την καλύτερη δυνατή σημασιολογική τμηματοποίηση των εικόνων σε σχέση με τα υπόλοιπα μοντέλα. Τέλος, να αναφέρουμε πως για την κλάση των επιπλεόντων πλαστικών το δίκτυο με τον Εκθετικό Scheduler, εμφάνισε τα καλύτερα αποτελέσματα.

Κεφάλαιο 5

Συμπεράσματα

Σε αυτό το κεφάλαιο γίνεται αξιολόγηση των τελικών αποτελεσμάτων - ερμηνείας του Explainable AI (XAI), αναφορά των προβλημάτων και μελλοντικών προσεγγίσεων.

5.1 Αξιολόγηση Αποτελεσμάτων

Στην παρούσα διπλωματική εργασία αντιμετωπίζεται το πρόβλημα του εντοπισμού θαλάσσιων ρύπων μέσω δορυφορικών εικόνων με χρήση τεχνικών βαθιάς μηχανικής μάθησης. Στην εργασία αποδεικνύεται πως είναι εφικτή η αναγνώριση αντικειμένων στην επιφάνεια της θάλασσας με την χρήση βαθιάς μηχανικής μάθησης και συγκεκριμένα των συνελικτικών νευρωνικών δικτύων. Για την διαδικασία της εκπαίδευσης χρησιμοποιούμαι 3 μοντέλα με διαφορετική αρχιτεκτονική, με κοινό χαρακτηριστικό τους είναι η ύπαρξη συνελικτικών επιπέδων.

Αρχικά, δοκιμάζουμε το μοντέλο DeepLabV3+, το οποίο δεν ανταποκρίνεται σε επιθυμητό βαθμό. Απο τα αποτελέσματα φαίνεται πως το μοντέλο ταξινομεί σε ικανοποιητικό βαθμό τα αντικείμενα μεγαλύτερης κλίμακας, ενώ αντίθετα σε αντικείμενα μικρότερης κλίμακας και με άμορφο σχηματισμό, όπως τα επιπλέοντα πλαστικά, δεν εκπαιδεύεται με επιτυχία. Αντίστοιχα, το U-Net, επιφέρει πολύ καλύτερα αποτελέσματα και με την Cross Entropy ως συνάρτηση απώλειας, παρά το γεγονός ότι η Focal θεωρείται ιδανικότερη συνάρτηση σφάλματος στην περίπτωση που υπάρχουν αντικείμενα μικρής κλίμακας και ανισορροπία στο δείγμα εκπαίδευσης. Τα μοντέλα DeepLabV3+ και U-Net λόγω των πολλαπλών επιπέδων συγκέντρωσης, επηρεάζουν τα μεμονωμένα χαρακτηριστικά των εικόνων εισόδου με αποτέλεσμα να εξαφανίζονται και το δίκτυο να μην μπορεί να τα ταξινομεί σωστά. Στο μοντέλο με το περιορισμένο οπτικό πεδίο, σε αντίθεση με τα άλλα 2 μοντέλα δεν τροφοδοτούμαι τις εικόνες ατόφιες αλλά κατόπιν επεξεργασίας. Συγκεκριμένα, τα δεδομένα εισόδου αποτελούνται απο 5x5 παράθυρα , όπου σκοπός μας είναι η ταξινόμηση του κεντρικού εικονοστοιχείου. Με την συγκεκριμένη τεχνική, παρατηρούμαι τα καλύτερα αποτελέσματα, καθώς η αρχιτεκτονική του μοντέλου σε συνδυασμό με την εικόνα εισόδου, επιτρέπουν την ανάκτηση χαρακτηριστικών και συνεπώς την καλύτερη εκπαίδευση του μοντέλου.

Τέλος, με την χρήση του Explainable AI (XAI) και τους δείκτες NDVI και NRD, αιτιολογούμαι την ταξινόμηση μερικών εικονοστοιχείων, ένας μέρος των οποίων έχουν ταξινομηθεί σωστά, ενώ τα υπόλοιπα εσφαλμένα. Συγκεκριμένα, είδαμε ότι για τα επιπλέοντα πλαστικά τα διαγράμματα των Mean Attribution, παρουσιάζουν μεγάλες τιμές στο κανάλι 665 (Red), συνεπώς οι δείκτες βλάστησης, έχουν τιμές στο σύνολο τιμών [–1, 0]. Αντιθέτως, για τα Sargassum, τα διαγράμματα των Mean Attribution, παρουσιάζουν μεγάλες τιμές στο κανάλι 842 (NIR), συνεπώς, οι δείκτες βλάστησης έχουν τιμές στο σύνολο τιμών [0, 1]. Όσον αφορά τις εσφαλμένες ταξινομήσεις, πολλές φορές τα εικονοστοιχεία δεν περιέχουν αποκλειστικά μία μοναδική κλάση με αποτέλεσμα να αντιμετωπίζει δυσκολίες το μοντέλο ταξινόμησης. Για παράδειγμα, τα πλοία έχουν πανομοιότυπη σύνθεση με τα πλαστικά, στις περισσότερες των περιπτώσεων τα εικονοστοιχεία δεν περιέχουν αποκλειστικά και μόνο μέρη του πλοίου αλλά και θαλάσσιο νερό με αποτέλεσμα να οδηγεί το μοντέλο στην λανθασμένη ταξινόμηση τους.

5.2 Προβλήματα

Το κυρίαρχο πρόβλημα κατά την διάρκεια της διπλωματικής αποτελεί, η ανισορροπία μεταξύ των κλάσεων. Υπάρχουν κλάσεις που τα δεδομένα τους δεν είναι επαρκή για την έγκυρη εκπαίδευση του μοντέλου. Η συλλογή δεδομένων, αποτελεί μία αρκετά επίπονη διαδικασία, με συνέπεια να υπάρχει ανάγκη χρήσης εικονοστοιχείων με χαμηλό επίπεδο εμπιστοσύνης.

Τα χαρακτηριστικά των εικόνων επηρεάζουν και αυτά την ακρίβεια της ταξινόμησης. Συγκεκριμένα, η αποτελεσματική επισήμανση των εικονοστοιχείων, απαιτεί την ύπαρξη δεδομένων με πολύ υψηλή διακριτική ικανότητα. Επίσης, υπάρχουν πολλές περιπτώσεις, όπου συνυπάρχουν 2 ή περισσότερα αντικείμενα των κλάσεων σε ένα εικονοστοιχείο, με αποτέλεσμα να μην είναι εύκολη η επισήμανση του. Τέτοιου είδους περιπτώσεις προκαλούν θόρυβο στο σύνολο δεδομένων. Επιπλέον, η ύπαρξη εκτεταμένης νεφοκάλυψης, καταστρέφει ένα μεγάλο ποσοστό δεδομένων.

Στην παρούσα μελέτη εκπαιδεύουμε μοντέλα επιβλεπόμενης μάθησης, συνεπώς υπάρχει απαίτηση επισημασμένων δεδομένων. Ένα μεγάλο μέρος των ήδη υπάρχων δορυφορικών εικόνων του συνόλου δεδομένων MARIDA, αποτελείται απο μη επισημασμένα δεδομένα που μένουν ανεκμετάλλευτα και δεν συμβάλλουν στην καλύτερη εκπαίδευση των δικτύων.

Ένα ακόμη σημαντικό πρόβλημα στην παρούσα μελέτη, είναι ο χρονικός περιορισμός. Λόγω του χρονικού περιορισμού δεν είναι εφικτή η ακριβής προσαρμογή πολλών υπερπαραμέτρων και γενικά, η διαδικασία βελτιστοποίησης για όλες τις υπερπαραμέτρους που υπάρχουν στα μοντέλα που δοκιμάζουμε.

5.3 Μελλοντικές προσεγγίσεις

Όσον αφορά τις μελλοντικές προσεγγίσεις, ένα σημαντικό βήμα είναι η εκπαίδευση ημιεπιβλεπόμενων μοντέλων (Semi-Supervised Models) που εκμεταλλεύονται πλήρως όλα τα διαθέσιμα δεδομένα. Η χρήση ενός ημι-επιβλεπόμενου μοντέλου μειώνει το πλήθος των επισημασμένων δεδομένων που χρησιμοποιούνται, καθώς έχουν την δυνατότητα να χρησιμοποιούν και τα μη επισημασμένα για την εκπαίδευση τους.

Μια εξελισσόμενη τεχνική μηχανικής μάθησης, που μπορεί να βοηθήσει στην εκπαίδευση μοντέλων για την αντιμετώπιση αντίστοιχων προβλημάτων είναι η αυτό-εποπτευόμενη μάθηση (Self-Supervised Learning). Η αυτό-εποπτευόμενη μάθηση, στοχεύει στην επίλυση της πρόκλησης που θέτει η υπερβολική εξάρτηση των επισημασμένων δεδομένων. Η κατασκευή έξυπνων συστημάτων με χρήση μεθόδων μηχανικής μάθησης εξαρτάται σε μεγάλο βαθμό από επισημασμένα δεδομένα καλής ποιότητας. Κατά συνέπεια, το κόστος των υψηλής ποιότητας επισημασμένων δεδομένων αποτελεί σημαντικό εμπόδιο στη συνολική εκπαιδευτική διαδικασία. Η ανάπτυξη μηχανισμών μάθησης με μη επισημασμένα δεδομένα μπορούν να κλιμακώσουν την έρευνα και την ανάπτυξη γενικών συστημάτων τεχνητής νοημοσύνης με χαμηλό κόστος. Πρακτικά, είναι αδύνατο να συλλέξουμε και να επισημάνουμε κάθε είδους ποικίλα δεδομένα. Σε αυτή τη διαδικασία, το πρόβλημα χωρίς επίβλεψη μετατρέπεται σε εποπτευόμενο πρόβλημα με την αυτόματη δημιουργία ετικετών. Για να γίνει χρήση της τεράστιας ποσότητας δεδομένων χωρίς ετικέτα, είναι σημαντικό να τεθούν οι σωστοί μαθησιακοί στόχοι για την επίβλεψη από τα ίδια τα δεδομένα. Η διαδικασία της αυτό-εποπτευόμενης σκοπεύει στην προ-εκπαίδευση ενός μοντέλου στα δεδομένα εισόδου με σκοπό την κατανόηση των δεδομένων χωρίς την ανάγκη ύπαρξης μεγάλου πλήθους επισημασμένων δεδομένων, όπου στη συνέχεια, μπορεί να επανεκπαιδευτεί σαν ένα πρόβλημα επιβλεπόμενης μάθησης.

Για την βελτίωση της απόδοσης των μοντέλων θα μπορούσαμε να χρησιμοποιήσουμε τεχνικές για την αναγνώριση συγκεκριμένων ετικετών, όπως για παράδειγμα την αναγνώριση σύννεφων και τον αντίστοιχων, σκιών τους. Συγκεκριμένα, μπορούμε να χρησιμοποιήσουμε μάσκες (Fmask), που βασίζονται σε κανόνες και στατιστικά στοιχεία που προέρχονται από τα φυσικά χαρακτηριστικά τους. Τα σύννεφα αναγνωρίζονται με βάση τα χαρακτηριστικά τους, «λευκό», «φωτεινό», «κρύο» και «υψηλό» και οι σκιές νεφών προσδιορίζονται λαμβάνοντας υπόψη το «σκοτεινό» χαρακτηριστικό τους και με βάση την υπόθεση ότι μοιράζονται παρόμοιο σχήμα με τα αντίστοιχα σύννεφα [38].

Επίσης, όσο εξελίσσεται η τεχνολογία, οι δορυφόροι αποκτούν όλο και καλύτερη φασματική και χωρική ανάλυση αλλά και νέες τεχνολογίες καταγραφής των δεδομένων, όπως για παράδειγμα συσκευές RADAR του τύπου SAR (Synthetic Aperture Radar), που θα βελτίωσαν σε πολύ μεγάλο βαθμό την ανάλυση και την ποιότητα των δορυφορικών εικόνων, με συνέπεια να καταλήγουμε σε καλύτερα τελικά αποτελέσματα.

Τέλος, απο όλα τα παραπάνω γίνεται αντιληπτό πως η αντιμετώπιση ενός τέτοιου προβλήματος περιέχει πολλά εμπόδια, ωστόσο με κατάλληλα εργαλεία και τεχνικές μπορούν να αποδοθούν ικανοποιητικά αποτελέσματα.

Βιβλιογραφία

- [1] Giorgio Roffo. Ranking to Learn and Learning to Rank: On the Role of Ranking in *Pattern Recognition Applications*.
- [2] Li Yin. A Summary of Neural Network Layers.
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun. *Deep Residual Learning for Image Recognition.*
- [4] Mo Deng, Alexandre Goy, Shuai Li, Kwabena Arthur και George Barbastathis. Probing shallower: perceptual loss trained Phase Extraction Neural Network (PLT-PhENN) for artifact-free reconstruction at low photon budget.
- [5] Farheen Ramzan, Muhammad Usman Ghani Khan, Asim Rehmat, Sajid Iqbal, Tanzila Saba, Amjad Rehman και Zahid Mehmood. A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer's Disease Stages Using Resting-State fMRI and Residual Neural Networks.
- [6] Liang Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff και Hartwig Adam. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.
- [7] Olaf Ronneberger, Philipp Fischer кал Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation.
- [8] Kunal Banerjee, Vishak Prasad C, Rishi Raj Gupta, Karthik Vyas, Anushree H ка Biswajit Mishra. *Exploring Alternatives to Softmax Function*.
- [9] Tsung Yi Lin, Priya Goyal, Ross Girshick, Kaiming He και Piotr Dollár. Focal Loss for Dense Object Detection.
- [10] Eric Hofesmann. IoU a better detection evaluation metric.
- [11] Katerina Kikaki, Ioannis Kakogeorgiou, Paraskevi Mikeli, Dionysios E. Raitsos και Konstantinos Karantzalos. MARIDA: A benchmark for Marine Debris detection from Sentinel-2 remote sensing data.
- [12] P. Mikeli, K. Kikaki, I. Kakogeorgiou και K. Karantzalos. HOW CHALLENGING IS THE DISCRIMINATION OF FLOATING MATERIALS ON THE SEA SURFACE USING HIGH RESOLUTION MULTISPECTRAL SATELLITE DATA?

- [13] Lauren Biermann, Daniel Clewley, Victor Martinez-Vicente και Konstantinos Topouzelis. Finding Plastic Patches in Coastal Waters using Optical Satellite Data.
- [14] Edward D. Goldberg. *Emerging problems in the coastal zone for the twenty-first century.*
- [15] Odei Garcia-Garin, Toni Monleón-Getino, Pere López-Brosa, Asunción Borrell, Alex Aguilar, Ricardo Borja-Robalino, Luis Cardona και Morgana Vighi. Automatic detection and quantification of floating marine macro-litter in aerial images: Introducing a novel deep learning approach connected to a web application in R.
- [16] Nitin Agarwala. Monitoring the Ocean Environment Using Robotic Systems: Advancements, Trends, and Challenges.
- [17] Mr Manoj Kumar. Emerging Trends in Big Data, IoT and Cyber Security.
- [18] Johnathan Charles Flowers. Strong and Weak AI: Deweyan Considerations.
- [19] Christoph Molnar. Interpretable Machine Learning.
- [20] Mehryar Mohri, Afshin Rostamizadeh και Ameet Talwalkar. *Foundations of machine learning*. Adaptive computation and machine learning series.
- [21] Huo Yingge, Imran Ali Kai Kang Yoon Lee. Deep Neural Networks on Chip A Survey.
- [22] Murat Sazli. A brief review of feed-forward neural networks.
- [23] Igor Aleksander και Helen Morton. *An Introduction to Neural Computing*. Chapman and Hall.
- [24] Hossein Gholamalinezhad και Hossein Khosravi. *Pooling Methods in Deep Neural Networks, a Review.*
- [25] Sakshi Indolia, Anil Kumar Goswami, S.P. Mishra και Pooja Asopa. *Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach.*
- [26] Oana Maria Camburu. Explaining Deep Neural Networks.
- [27] Jason Brownlee. A Gentle Introduction to the Rectified Linear Unit (ReLU).
- [28] Jason Brownlee. Softmax Activation Function with Python.
- [29] Daniel Svozil, Vladimír Kvasnicka και Jirí Pospichal. *Introduction to multi-layer feedforward neural networks*.
- [30] Sebastian Bock, Josef Goppold και Martin Weiß. *An improvement of the convergence proof of the ADAM-Optimizer.*
- [31] David Powers. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation.

- [32] Stephen V. Stehman. Selecting and interpreting measures of thematic classification accuracy.
- [33] Julie Gerlings, Arisa Shollo και Ioanna Constantiou. *Reviewing the Need for Explainable Artificial Intelligence (xAI)*.
- [34] Mukund Sundararajan, Ankur Taly και Qiqi Yan. Axiomatic Attribution for Deep Networks.
- [35] Wojciech Samek, Grégoire Montavon, Sebastian Lapuschkin, Christopher J. Anders kai Klaus Robert Müller. Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications.
- [36] Santhakumaran A. Statistical Normalization and Back Propagation for Clasification.
- [37] Chuanmin Hu. Remote detection of marine debris using satellite observations in the visible and near infrared spectral range: Challenges and potentials.
- [38] Shi Qiu, Zhe Zhu και Binbin He. *Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 imagery.*