



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΜΗΧΑΝΙΚΗΣ ΣΧΕΔΙΑΣΗΣ ΚΑΙ ΠΑΡΑΓΩΓΗΣ

ΠΡΟΓΡΑΜΜΑ ΠΡΟΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

**Εφαρμογές Τεχνητής Νοημοσύνης  
στην Προστασία του Περιβάλλοντος**

Διπλωματική Εργασία  
του  
**Σπυρίδων Α. Γεωργούλα**

Επιβλέπων: Νικολάου Γρηγόριος, Λέκτορας

Αιγάλεω, Ιούνιος 2022





UNIVERSITY OF WEST ATTICA

SCHOOL OF ENGINEERING

DEPARTMENT OF INDUSTRIAL DESIGN AND PRODUCTION ENGINEERING

UNDERGRADUATE STUDY PROGRAM

**Artificial Intelligence Applications  
in Environmental Protection**

**Spyridon A. Georgoulas's**

Diploma Thesis

Supervisor: Nikolaou Grigorios, Lecturer

Aigaleo, June 2022



**Εφαρμογές Τεχνητής Νοημοσύνης  
στην Προστασία του Περιβάλλοντος**

Διπλωματική Εργασία  
του

**Σπυρίδων Γεωργούλα**  
**A.M.: 71444271**

## Μέλη Εξεταστικής Επιτροπής

Γρηγόριος Νικολάου  
Λέκτορας

Σουλτάνα Βασιλειάδου  
Επίκουρη Καθηγήτρια

Δρόσος Χρήστος  
ΕΔΙΠ

## **Δήλωση Συγγραφέα**

Ο κάτωθι υπογεγραμμένος Σπυρίδων Γεωργούλας του Αθανασίου, με αριθμό μητρώου 71444271 φοιτητής του Πανεπιστημίου Δυτικής Αττικής της Σχολής Μηχανικών του Τμήματος Μηχανικών Βιομηχανικής Σχεδίασης και Παραγωγής, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της πτυχιακής/διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών

**Σπυρίδων Γεωργούλας**





στην Μαρία, που θα το δει και θα χαμογελάσει,

με αγάπη



## Ευχαριστίες

Με την ολοκλήρωση της παρούσας διπλωματικής εργασίας θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κ. Γρηγόρη Νικολάου που με βοήθησε να πιστέψω πως τα όνειρα μπορούν να πραγματοποιηθούν, για την εμπιστοσύνη που μου έδειξε και την καθοδήγηση που μου παρείχε.

Θα ήθελα να ευχαριστήσω πολύ τους γονείς μου, Σάκη και Νόνα, που με βοηθούν πάντα σε ό,τι χρειαστώ.

Τέλος, ευχαριστώ μέτρια τον αδερφό μου Παναή, που με άφησε να του κλέψω τον υπολογιστή για να τελειώσω την διπλωματική μου, αλλά και για όσα κάνει καθημερινά.



## Περίληψη

Η προστασία του φυσικού περιβάλλοντος είναι ένα ζήτημα που απασχολεί την ανθρωπότητα όλο και περισσότερο τα τελευταία χρόνια. Σε αυτή την διπλωματική εργασία παρουσιάζονται τρόποι με τους οποίους μπορεί η μηχανική μάθηση να βοηθήσει στην πρόβλεψη φυσικών καταστροφών και να συμβάλλει στην αποφυγή τους ή στην έγκαιρη αντιμετώπισή τους. Γίνεται εισαγωγή στην τηλεπισκόπηση και στο ραντάρ συνθετικού διαφράγματος, επεξηγούνται τα τεχνητά νευρωνικά δίκτυα και αναλύονται εφαρμογές βαθιάς μάθησης στην προστασία του φυσικού περιβάλλοντος. Παρουσιάζεται πώς μπορεί να γίνει πρόβλεψη του κινδύνου πυρκαγιάς με βαθιά μάθηση, η σημασιολογική κατάτμηση πετρελαιοκηλίδων αλλά και η ανίχνευση παγόβουνων και πλοίων με βαθιά μάθηση από δορυφορικές εικόνες SAR. Τα παγόβουνα τα οποία αποτελούν μεγάλο κίνδυνο για την ναυσιπλοΐα είναι δυνατό να ανιχνευθούν από δορυφορικές εικόνες SAR, αλλά η διάκρισή τους από τα πλοία χρειάζεται να γίνει από κάποιον άνθρωπο με ειδικές γνώσεις, κάτι το οποίο είναι πολύ χρονοβόρο. Για αυτό τον λόγο προτείνεται και παρουσιάζεται ένας αλγόριθμος βαθιάς μάθησης ο οποίος με την χρήση συνελκτικών νευρωνικών δικτύων διακρίνει αυτόματα τα παγόβουνα από τα πλοία σε εικόνες SAR.

**Λέξεις Κλειδιά:** Τεχνητή Νοημοσύνη, Μηχανική Μάθηση, Βαθιά Μάθηση, Συνελκτικά Νευρωνικά Δίκτυα, Τηλεπισκόπηση, Ραντάρ Συνθετικού Διαφράγματος, Παγόβουνο, Πλοίο, Όραση Υπολογιστών, Φυσικό Περιβάλλον

## **Abstract**

The protection of the natural environment is an issue that concerns humanity more and more in recent years. This thesis presents ways in which machine learning can help predict natural disasters and contribute to their avoidance or early response. Remote sensing and synthetic aperture radar are introduced, artificial neural networks are explained, and deep learning applications in environmental protection are discussed. It shows how fire hazard can be predicted with deep learning, the semantic segmentation of oil spills and the discrimination of icebergs and ships with deep learning in SAR satellite imagery. Icebergs that pose a great danger to navigation can be detected by SAR satellite images, but distinguishing them from ships needs to be done by someone with special knowledge, which is very time-consuming. For this reason, a deep learning algorithm is proposed and presented which, using convolutional neural networks, automatically distinguishes icebergs from ships in SAR images.

**Keywords:** Artificial Intelligence, Machine Learning, Deep Learning, Convolutional Neural Networks, Remote Sensing, Synthetic Aperture Radar, Iceberg, Ship, Computer Vision, Natural Environment

## Περιεχόμενα

1	Εισαγωγή .....	21
1.1	Παγόβουνα .....	21
1.2	Απειλή Υπεράκτιων Δραστηριοτήτων .....	22
1.3	Ανίχνευση Παγόβουνων.....	24
2	Τηλεπισκόπηση.....	27
2.1	Ο Δορυφόρος Sentinel-1 .....	27
2.2	Synthetic Aperture Radar .....	29
3	Τεχνητή Νοημοσύνη.....	34
3.1	Μηχανική Μάθηση .....	34
4	Νευρωνικά Δίκτυα .....	38
4.1.1	Βιολογικά Νευρωνικά Δίκτυα .....	38
4.1.2	Μοντέλο Τεχνητού Νευρώνα .....	39
4.1.3	Τεχνητά Νευρωνικά Δίκτυα .....	41
4.1.4	Μάθηση και Ανάκληση .....	44
4.1.5	Συνελκτικά Νευρωνικά Δίκτυα .....	48
5	Εφαρμογές Μηχανικής Μάθησης σε Φυσικές Καταστροφές.....	51
5.1	Βαθιά Μάθηση στην Καθημερινή Πρόβλεψη Κινδύνου Πυρκαγιάς.....	51
5.2	Βαθιά Μάθηση για Σηματολογική Κατάτμηση Πετρελαιοκηλίδων .....	53
6	Βαθιά Μάθηση στην Διάκριση Παγόβουνων και Πλοίων σε Εικόνες SAR .....	56
6.1	State-of-the-Art Μέθοδοι .....	57
6.2	Η Μέθοδος μας.....	57
6.2.1	Δεδομένα και Προετοιμασία.....	57
6.2.2	Το Μοντέλο .....	60
7	Αποτελέσματα και Συζήτηση .....	63
8	Συμπεράσματα .....	66

## Κατάλογος Εικόνων

Εικόνα 1.1 : Iceberg Chart για τις ακτές του Newfoundland and Labrador [2] .....	24
Εικόνα 1.2: Παγόβουνα (αριστερά) και πλοία (δεξιά) σε εικόνες SAR .....	26
Εικόνα 2.1 Ο δορυφόρος Sentinel-1 [3] .....	27
Εικόνα 2.2: Τροχιά του Sentinel-1 [3] .....	28
Εικόνα 2.3: Γεωγραφική κάλυψη του Sentinel-1 [3].....	29
Εικόνα 2.4 Γωνία πρόσπτωσης και SAR [6] .....	31
Εικόνα 2.5: Ευαισθησία μετρήσεων SAR βάσει του μήκους κύματος [7].....	32
Εικόνα 2.6: Πόλωση και Σκέδαση SAR [7] .....	33
Εικόνα 4.1: Αναπαράσταση του βιολογικού νευρώνα .....	38
Εικόνα 6.1 : Παγόβουνο σε HH (αριστερά) και HV (δεξιά) πόλωση. ....	58
Εικόνα 6.2: Πλοίο σε HH (αριστερά) και HV (δεξιά) πόλωση. ....	59



## Κατάλογος Σχημάτων

Σχήμα 2.1 : Γεωμετρία παρατήρησης για τον σχηματισμό συνθετικού διαφράγματος για στόχο P κατά $x=0$ [7] .....	30
Σχήμα 2.2 : Το ηλεκτρομαγνητικό φάσμα με ένθετες ζώνες μικροκυμάτων [7] .....	31
Σχήμα 4.1 : Μοντέλο του τεχνητού νευρώνα .....	39
Σχήμα 4.2: Βηματική συνάρτηση .....	40
Σχήμα 4.3: Συνάρτηση Προσήμου.....	41
Σχήμα 4.4: Λογιστική συνάρτηση .....	41
Σχήμα 4.5: ΤΝΔ πρόσθιας τροφοδότησης 3-4-4-2 με πλήρη διασύνδεση μεταξύ των διαδοχικών επιπέδων .....	42
Σχήμα 4.6: Γράφημα ροής του σήματος perceptron.....	43
Σχήμα 4.7 : Multilayer perceptron πολλών επιπέδων με δύο κρυφά επίπεδα .....	44
Σχήμα 4.8: Αναπαράσταση του Gradient Descent Γεωμετρικά [8] .....	46
Σχήμα 4.9: Αρχιτεκτονική CNN.....	48
Σχήμα 4.10: Συνέλιξη και μάσκα. Η εικόνα της εξόδου έχει τιμή μέγιστη εκεί που εντοπίζεται σταυρός.....	49
Σχήμα 5.1: Παραγώμενοι χάρτες διαφορετικών μοντέλων και το FWI για δυο ημερομηνίες με μέτριο (πάνω) και υψηλό (κάτω) κίνδυνο εκδήλωσης πυρκαγιάς. Ο υψηλός κίνδυνος απεικονίζεται με κόκκινο και ο χαμηλό με μπλε. [13] .....	53
Σχήμα 5.2 : Παράδειγμα 4 εικόνων ελέγχου (από πάνω προς τα κάτω): SAR εικόνες, ground truth μάσκες και μάσκες ανίχνευσης ως αποτέλεσμα του μοντέλου .....	55
Σχήμα 6.1 : Loss του Μοντέλου για εκπαίδευση για 4 εποχές.....	62
Σχήμα 6.2 : Accuracy του Μοντέλου για εκπαίδευση για 4 εποχές.....	62

## **Κατάλογος Πινάκων**

Πίνακας 1.1 : Πίνακας Μεγεθών Παγόβουνων [1].....	22
Πίνακας 3.1 : Πίνακας σύγκρισης για πρόβλημα ταξινόμησης 2 κλάσεων.....	37
Πίνακας 7.1 : Confusion Matrix για τα Δεδομένα Ελέγχου .....	63

## **Συντομογραφίες**

ΤΝΔ: Τεχνητό Νευρωνικό Δίκτυο

CNN: Convolutional Neural Network

SAR: Synthetic Aperture Radar

MSE: Mean Squared Error

EO: Earth Observation

DL: Deep Learning



# 1 Εισαγωγή

Η προστασία του φυσικού περιβάλλοντος είναι ένα ζήτημα που απασχολεί την ανθρωπότητα όλο και περισσότερο τα τελευταία χρόνια. Οι επιπτώσεις της κλιματικής αλλαγής στις ζωές μας γίνονται συνεχώς πιο έντονες και πιο συχνές. Για αυτό, επιστήμονες από όλα τα επιστημονικά πεδία μελετούν πώς με τον δικό του τρόπο ο κάθε ένας μπορεί να έχει μια θετική συμβολή στην προστασία του περιβάλλοντος και να βρουν λύσεις σε σημαντικά ζητήματα. Στην ζωή, για να μπορέσουμε να ελέγξουμε κάτι πρέπει πρώτα να μπορούμε να το παρατηρήσουμε. Ως εκ τούτου, πολλές συσκευές που έχουν την δυνατότητα να παρατηρούν την γη έχουν σταλεί σε τροχιά στο διάστημα με σκοπό την συλλογή δεδομένων τα οποία να χρησιμοποιηθούν για την δημιουργία μοντέλων ικανών να προβλέψουν κινδύνους και να βοηθήσουν στην αποφυγή τους.

## 1.1 Παγόβουνα

Παγόβουνο είναι ένας όγκος πάγου που αποκολλάται από τους πολικούς παγετώνες και επιπλέει στην θάλασσα [1]. Παγόβουνα μπορούν να παρατηρηθούν έντονα στις θάλασσες κοντά στην Ανταρκτική και την Αρκτική. Ο πάγος, στο μεγαλύτερο μέρος του είναι διάφανος, όπως ένα γυαλί, τα παγόβουνα όμως εμφανίζονται ως άσπρα διότι στο εσωτερικό τους έχουν παγιδευτεί φυσαλίδες αέρα ή είναι καλυμμένα με χιόνι. Υπάρχουν παγόβουνα τεραστίων διαστάσεων όπως μια ολόκληρη πόλη ή χώρα αλλά και παγόβουνα μικρότερων διαστάσεων. Το μεγαλύτερο παγόβουνο που έχει εντοπιστεί είχε μέγεθος περίπου όσο το Βέλγιο. Το σχήμα του μπορεί να διαμορφωθεί από το φύσημα του ανέμου και την δράση των κυμάτων. Αναλόγως το μέγεθός τους μπορούν να κατηγοριοποιηθούν όπως φαίνεται στον Πίνακα 1.1.

category	height		length	
	metres	feet	metres	feet
<b>growler</b>	< 1	< 3	< 5	< 16
<b>bergy bit</b>	1 to 4	3 to 13	5 to 14	15 to 46
<b>small berg</b>	5 to 15	14 to 50	15 to 60	47 to 200
<b>medium berg</b>	16 to 45	51 to 150	61 to 122	201 to 400
<b>large berg</b>	46 to 75	151 to 240	123 to 204	401 to 670
<b>very large berg</b>	> 75	> 240	> 204	> 670

Πίνακας 1.1 : Πίνακας Μεγεθών Παγόβουνων [1]

## 1.2 Απειλή Υπεράκτιων Δραστηριοτήτων

Τα παγόβουνα αποτελούν συχνά μεγάλο κίνδυνο για την ναυσιπλοΐα, τις πλατφόρμες πετρελαίου, τα σκάφη υποστήριξης των πλατφορμών πετρελαίου, τα αλιευτικά σκάφη και κατ' επέκταση για το φυσικό περιβάλλον. Λίγα χρόνια μετά το ναυάγιο του Τιτανικού το 1912, ιδρύθηκαν διάφοροι οργανισμοί για την επιτήρηση και επισήμανση παγόβουνων, και την ενημέρωση των ναυτιλλομένων. Επίσης, οι υδρογραφικές υπηρεσίες κάποιων χωρών όπως ο Καναδάς εκδίδουν ειδικούς χάρτες για την πλοήγηση σε περιοχές με παγόβουνα με σκοπό την ενημέρωση των πλοίων οι οποίοι ονομάζονται «iceberg-charts».

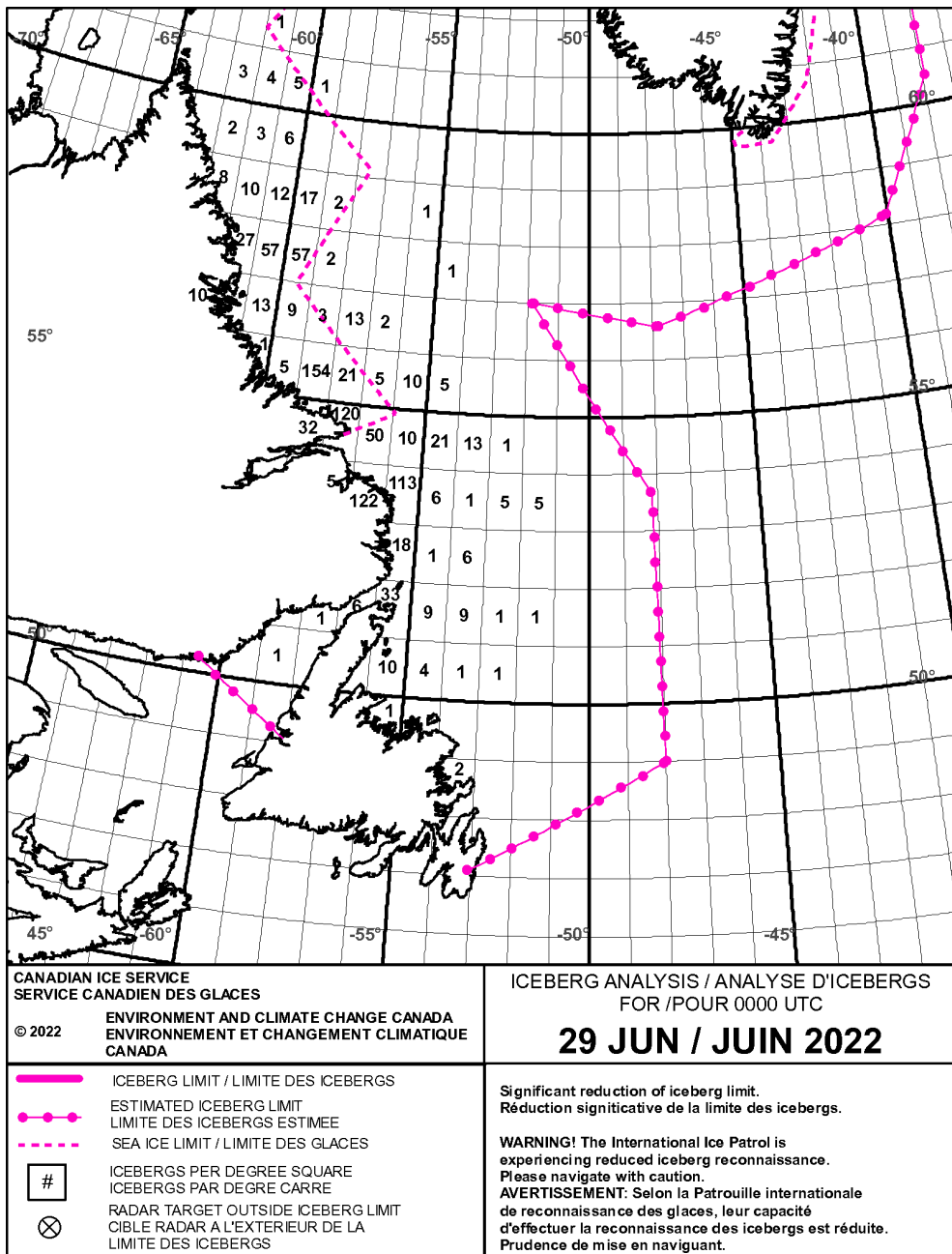
Το πρόβλημα της προστασίας της ναυτιλίας από τα παγόβουνα είναι πιο κρίσιμο σε δύο περιοχές, στα υψηλά γεωγραφικά πλάτη του Νότιου Ωκεανού και το βορειοδυτικό τμήμα του Βόρειου Ατλαντικού. Η απειλή του Νότιου Ωκεανού αυξάνεται επειδή τα μεγάλα πλοία μεταφοράς εμπορευματοκιβωτίων - εκείνα που δεν μπορούν ή δεν θέλουν να χρησιμοποιήσουν τη Διώρυγα του Παναμά - μπορούν να φτάσουν σε υψηλά νότια γεωγραφικά πλάτη κατά τη διέλευση από την Αυστραλία ή τη Νέα Ζηλανδία στο Ακρωτήριο Χορν.

Κατά τις δεκαετίες του 1950 και του 1960, έγιναν προσπάθειες από την Ακτοφυλακή των ΗΠΑ να βρουν τρόπους θραύσης των παγόβουνων που αποτελούσαν απειλή για τη ναυτιλία. Όλοι ήταν ανεπιτυχής και κυρίως οι εκρηκτικές τεχνικές, καθώς ο πάγος και το χιόνι είναι τόσο αποτελεσματικά στην απορρόφηση μηχανικών κραδασμών που δεν υπήρχε κανένα

αποτέλεσμα. Συχνά η απόδοση του κατακερματισμένου πάγου δεν ήταν μεγαλύτερη από τη μάζα του εκρηκτικού που χρησιμοποιήθηκε. Λόγω της ανάγκης υπεράσπισης των υπεράκτιων πλατφορμών γεώτρησης και παραγωγής από παγόβουνα, η βιωσιμότητα των εκρηκτικών τεχνικών έχει επανεξεταστεί πιο πρόσφατα. Διαπιστώθηκε ότι ο πολύ κρύος πάγος, όπως ο τύπος που βρίσκεται στο κάτω μέρος ενός παγόβουνου, μπορεί να κατακερματιστεί επιτυχώς με τη χρήση εκρηκτικών βραδείας καύσης όπως το Thermit. Το Thermit μπορεί να εμφυτευτεί με διάτρηση. Ωστόσο, η εμφύτευση είναι μια επικίνδυνη διαδικασία λόγω της πιθανότητας ανατροπής. Μέχρι να τελειοποιηθούν αυτές οι τεχνικές, τα παγόβουνα δεν μπορούν να καταστραφούν.

Η αδυναμία εύρεσης μεθόδων για την ασφαλή και επιτυχημένη θραύση ενός παγόβουνου με τεχνητά μέσα, ώστε να σταματήσει να αποτελεί κίνδυνο, έχει οδηγήσει τους επιστήμονες σε μία άλλη έξυπνη λύση, τον εντοπισμό τους και την παρακολούθηση της πορείας τους. Στη συνέχεια, οι τροχιές του παγόβουνου προβλέπονται από ολοένα και πιο εξελιγμένα μοντέλα υπολογιστών. Εάν είναι απαραίτητο, ορισμένα παγόβουνα ρυμουλκούνται έξω από την πορεία τους για την αποφυγή καταστροφών.

Σύμφωνα με την Καναδική Υπηρεσία Πάγου, το διάγραμμα της Εικόνας 1.1 παρουσιάζει μια εκτίμηση των συνθηκών των παγόβουνων στα ύδατα της Ανατολικής Ακτής νότια των 60°B. Οι συνθήκες βασίζονται σε οπτικές παρατηρήσεις παγόβουνων από πλοία, αεροσκάφη και περιστασιακά από δορυφορικές εικόνες. Η θέση και το μέγεθος των παγόβουνων εισάγονται σε προσαρμοσμένο λογισμικό που προβλέπει τη θέση των παγόβουνων στις 00:00 UTC. Ένα εκτιμώμενο όριο παγόβουνων παρέχεται για καθοδήγηση βόρεια των 53 °B. Αυτές οι περιοχές δεν παρακολουθούνται για παγόβουνα χρησιμοποιώντας αεροσκάφη, επομένως το όριο βασίζεται στην κλιματολογία και τις αναφορές πλοίων. Το iceberg chart περιλαμβάνει μια εκτίμηση του αριθμού των παγόβουνων ανά τετραγωνική μίρα γεωγραφικού πλάτους και μήκους, συμπεριλαμβανομένων των κομματιών του παγόβουνου, του ορίου των παγόβουνων και του ορίου του θαλάσσιου πάγου. Αυτός ο χάρτης είναι σημαντικός για όσους χρειάζονται πληροφορίες για τα παγόβουνα στα ύδατα της Ανατολικής Ακτής, όπως οι εταιρείες εκμετάλλευσης πλοίων που σχεδιάζουν διαδρομές στον Ατλαντικό.



Εικόνα 1.1 : Iceberg Chart για τις ακτές του Newfoundland and Labrador [2]

### 1.3 Ανίχνευση Παγόβουνων

Το παγόβουνο είναι ένα πολύ μεγάλο αντικείμενο που μπορεί να ανιχνευθεί στην ανοιχτή θάλασσα τόσο οπτικά όσο με ραντάρ και σόναρ. Η ανίχνευση ενός παγόβουνου εξαρτάται επίσης από την εγρήγορση των φρουρών ενός πλοίου. Ένα παγόβουνο σε αποσύνθεση ενέχει πρόσθετους κινδύνους λόγω των growlers και των bergy bits που το περιτριγυρίζουν. Αν και

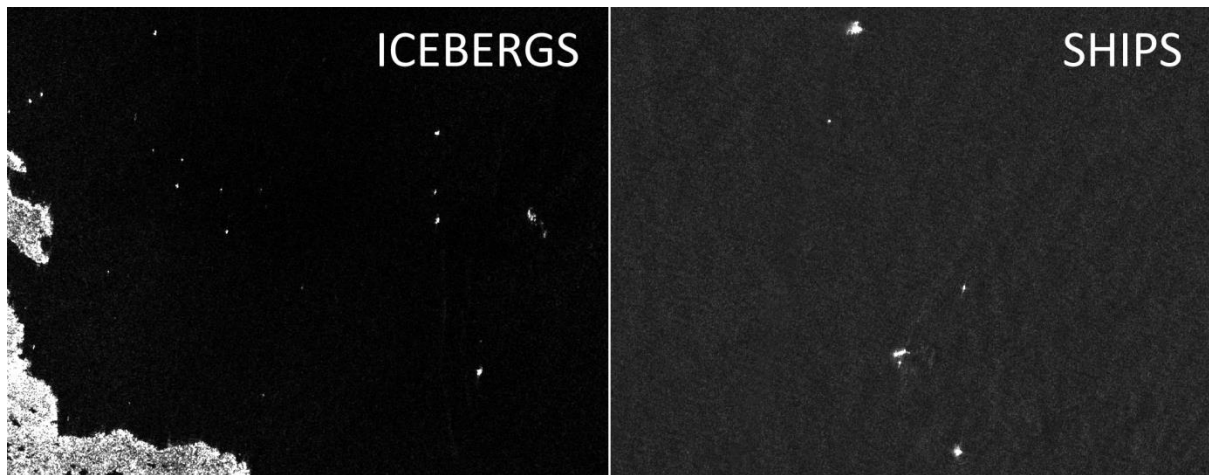


μικρά σε μέγεθος, έχουν μάζες (μέχρι 120 τόνους τα growlers, έως και 5.400 τόνους τα bergy bits) που είναι ικανές να καταστρέψουν ή να βυθίσουν πλοία. Επειδή ένα πλοίο μπορεί να κατευθυνθεί για να αποφύγει ένα μεγάλο μητρικό παγόβουνο, ενδέχεται να διατρέχει μεγαλύτερο κίνδυνο από μη ανιχνευμένους growlers και bergy bits που παρασύρονται κοντά.

Οι αναφορές των αρμόδιων οργανισμών για τις γνωστές ή εκτιμώμενες θέσεις των παγόβουνων διαβιβάζονται δύο φορές την ημέρα στα πλοία. Στο παρελθόν, οι θέσεις παγόβουνων εντοπίζονταν από πλοία ή αεροσκάφη. Ωστόσο, γίνεται όλο και πιο συνηθισμένο τα παγόβουνα να εντοπίζονται με την ερμηνεία δορυφορικών εικόνων.

Τα συστήματα τηλεπισκόπησης που χρησιμοποιούνται για την ανίχνευση παγόβουνων βρίσκονται σε δορυφόρους σε απόσταση 600 χιλιομέτρων πάνω από τη Γη. Ο δορυφορικός σχηματισμός Sentinel-1 χρησιμοποιείται για την παρακολούθηση της ξηράς και του ωκεανού. Καταγράφει εικόνες της επιφάνειας της Γης με ραντάρ C-Band το οποίο λειτουργεί με συχνότητα που «βλέπει» μέσα από το σκοτάδι, τη βροχή, το σύννεφο ακόμα και την ομίχλη. Δεδομένου ότι εκπέμπει τη δική του πηγή ενέργειας, μπορεί να τραβήξει εικόνες είτε μέρα είτε νύχτα. Το ραντάρ αυτό ονομάζεται Ραντάρ Συνθετικού Διαφράγματος (Synthetic Aperture Radar – SAR).

Ένα πρόβλημα που προκύπτει από την προσπάθεια ανίχνευσης παγόβουνων σε εικόνες SAR είναι πως όταν το ραντάρ εντοπίζει ένα αντικείμενο, δεν μπορεί να ξεχωρίσει ένα παγόβουνο από ένα πλοίο ή οποιοδήποτε άλλο στερεό αντικείμενο (Εικόνα 1.2). Το αντικείμενο πρέπει να αναλυθεί για ορισμένα χαρακτηριστικά - σχήμα, μέγεθος και φωτεινότητα - για να διαπιστωθεί τι ακριβώς είναι. Το περιβάλλον γύρω από το αντικείμενο παίζει επίσης σημαντικό ρόλο στην διαδικασία σύνθεσης της εικόνας SAR. Στους ωκεανούς, οι ισχυροί άνεμοι θα δημιουργήσουν ένα πιο φωτεινό φόντο και αντίθετα, οι χαμηλοί άνεμοι θα δημιουργήσουν ένα πιο σκούρο φόντο. Ο δορυφόρος Sentinel-1 είναι ένα ραντάρ πλάγιας όψης, που σημαίνει ότι βλέπει την περιοχή της εικόνας υπό γωνία (γωνία πρόσπτωσης). Ως εκ τούτου, το φόντο του ωκεανού θα είναι πιο σκούρο σε μεγαλύτερη γωνία πρόσπτωσης.



Εικόνα 1.2: Παγόβουνα (αριστερά) και πλοία (δεξιά) σε εικόνες SAR

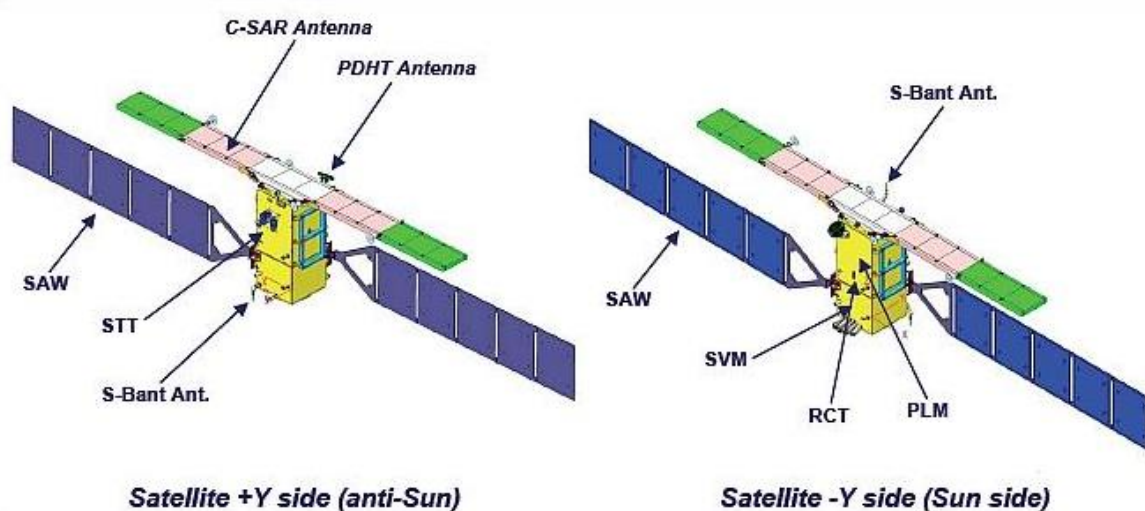
Από τα παραπάνω καταλαβαίνουμε πως η διάκριση ανάμεσα σε παγόβουνα και σε πλοία σε εικόνες SAR είναι μια χρονοβόρα διαδικασία όπου αν μη τι άλλο χρειάζεται η παρουσία ενός ανθρώπου για την ταξινόμηση των εικόνων. Για τους σκοπούς αυτής της διπλωματικής εργασίας κατασκευάστηκε ένας αλγόριθμος βαθιάς μάθησης όπου με την χρήση τεχνητών νευρωνικών δικτύων ξεχωρίζει αυτόματα τα παγόβουνα από τα πλοία σε εικόνες SAR. Απαντάει δηλαδή με επιτυχία στο ερώτημα: «Αυτό που απεικονίζεται, είναι παγόβουνο ή πλοίο;».

## 2 Τηλεπισκόπηση

Η επιστήμη της τηλεπισκόπησης (remote sensing) ασχολείται με την παρατήρηση χαρακτηριστικών και φαινομένων από απόσταση. Η τηλεπισκόπηση περιλαμβάνει ένα μεγάλο πεδίο εφαρμογών, φυσικών λειτουργιών και τεχνικών λειτουργιών. Σήμερα, όταν αναφερόμαστε στην τηλεπισκόπηση, εννοούμε την τεχνολογία και την επιστήμη μελέτης και παρατήρησης των χαρακτηριστικών της γης από απόσταση. Για τους σκοπούς της τηλεπισκόπησης, έχουν σταλεί πολλοί δορυφόροι σε τροχιά στο διάστημα, μία αποστολή από τις οποίες είναι αυτή του Sentinel-1. Το Sentinel-1 είναι η πρώτη από τις αποστολές που ανέπτυξε η European Space Agency για την πρωτοβουλία Copernicus, ενός ευρωπαϊκού προγράμματος για την μελέτη και παρακολούθηση του περιβάλλοντος της γης.

### 2.1 Ο Δορυφόρος Sentinel-1

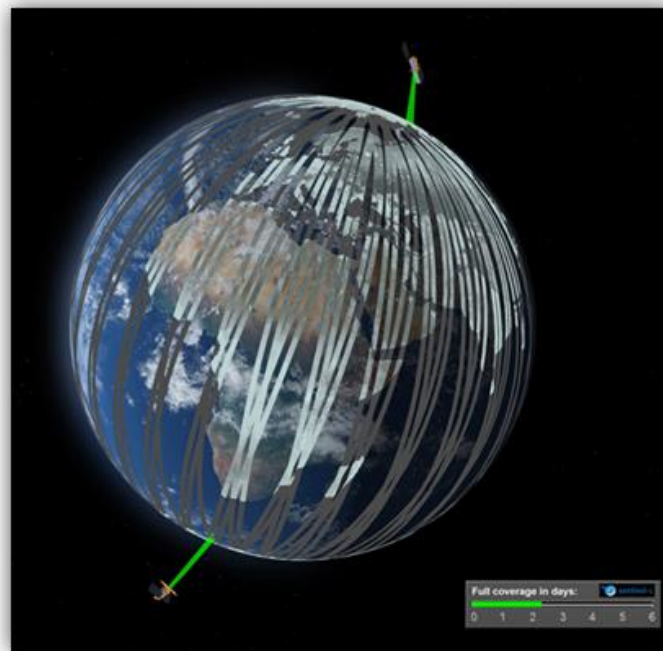
Η αποστολή Sentinel-1 αποτελείται από ένα σχηματισμό δύο δορυφόρων, τον Sentinel-1A και τον Sentinel-1B [3]. Ο δορυφόρος Sentinel-1 έχει ως κύριο στόχο την παρακολούθηση της γης και των ωκεανών. Στον δορυφόρο υπάρχει ένας αισθητήρας C-Band SAR, ο οποίος παρέχει υψηλής και μέσης ευκρίνειας απεικόνιση ανεξάρτητα από τις καιρικές συνθήκες. Το C-SAR μπορεί να κάνει λήψη εικόνων την ημέρα αλλά και την νύχτα, κάτι πολύ χρήσιμο για θαλάσσια και χερσαία παρακολούθηση [4]. Η αποστολή αυτή έχει ως στόχο την παρακολούθηση της γης, την υποστήριξη στις προσπάθειες χαρτογράφησης σε περιπτώσεις έκτακτης ανάγκης, την παρατήρηση των πάγων και των παγόβουνων, την ανίχνευση πετρελαιοκηλίδων, τον εντοπισμό σκαφών θαλάσσης κ.α.



Εικόνα 2.1 Ο δορυφόρος Sentinel-1 [3]

## Τροχιά

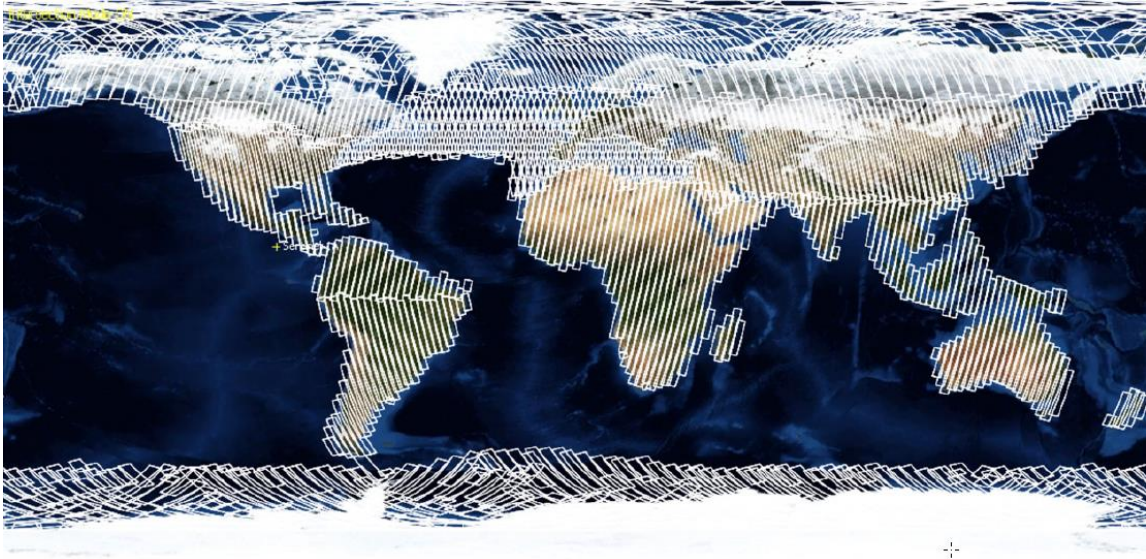
Η τροχιά του Sentinel-1 είναι σχεδόν πολική και έχει κύκλο επανάληψης 12 ημέρες και 175 τροχιές ανά κύκλο, για τον έναν δορυφόρο από τους δύο. Και ο Sentinel-1A αλλά και ο Sentinel-1B έχουν το ίδιο επίπεδο τροχιάς, όμως έχουν διαφορά φάσης τροχιάς  $180^\circ$ . Σε τροχιά οι δύο δορυφόροι διαμορφώνουν τον κύκλο επανάληψης στις έξι ημέρες.



Εικόνα 2.2: Τροχιά του Sentinel-1 [3]

## Γεωγραφική Κάλυψη

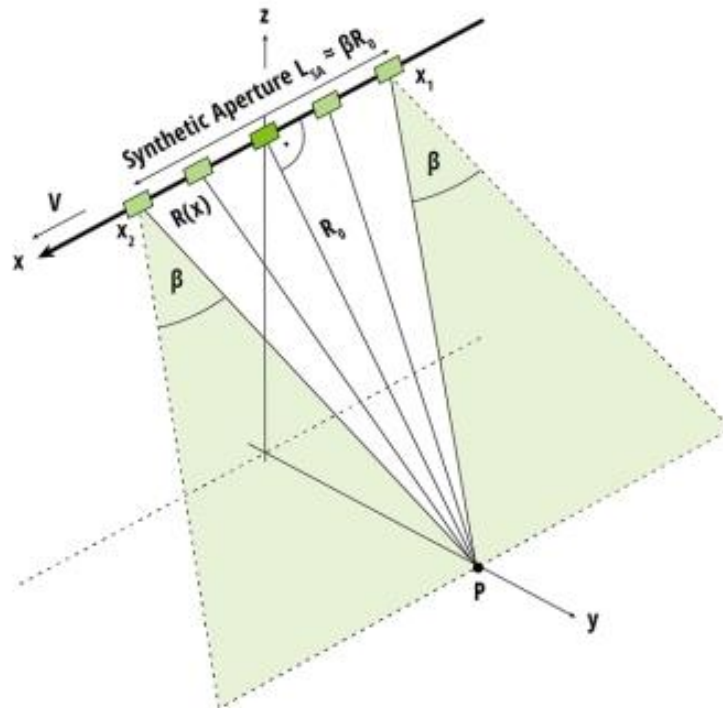
Ένας μόνο από τους δύο δορυφόρους Sentinel-1 μπορεί να χαρτογραφήσει ολόκληρη την γη σε 12 ημέρες. Ο σχηματισμός των δύο δορυφόρων προσφέρει κύκλο επανάληψης έξι ημερών. Ο σχηματισμός έχει επαναλαμβανόμενη συχνότητα (αύξουσα/φθίνουσα) 3 ημερών στον ισημερινό, λιγότερο από 1 ημέρα στην Αρκτική και παρέχει κάλυψη σε Ευρώπη, Καναδά και κύριες ακτοπλοϊκές γραμμές σε 1-3 ημέρες, ανεξάρτητα από τις καιρικές συνθήκες. Τα δεδομένα των ραντάρ παραδίδονται στις υπηρεσίες Copernicus εντός μίας ώρας από την απόκτησή τους.



Εικόνα 2.3: Γεωγραφική κάλυψη του Sentinel-1 [3]

## 2.2 Synthetic Aperture Radar

Το Synthetic Aperture Radar (SAR) ή όπως ονομάζεται στα ελληνικά «Ραντάρ συνθετικού διαφράγματος» είναι ένας τύπος για συλλογή δεδομένων όπου ένας αισθητήρας παράγει τη δική του ενέργεια και μετά καταγράφει την ποσότητα αυτής που ανακλάται πίσω μετά την αλληλεπίδραση με τη Γη, η οποία ονομάζεται “backscatter” [5]. Τα δεδομένα SAR απαιτούν διαφορετικό τρόπο σκέψης, καθώς η λειτουργία τους δεν είναι παρόμοια με αυτή μιας κλασικής φωτογραφίας. Το σήμα που παράγεται ανταποκρίνεται σε χαρακτηριστικά επιφάνειας όπως η υγρασία και η δομή [6]. Η χωρική ανάλυση των δεδομένων από το ραντάρ σχετίζεται άμεσα με την αναλογία του μήκους κύματος του αισθητήρα προς το μήκος της κεραίας του αισθητήρα. Για μια συγκεκριμένη τιμή μήκους κύματος, όσο μεγαλύτερη είναι η κεραία, τόσο μεγαλύτερη είναι η χωρική ανάλυση. Το ραντάρ αυτό έχει την δυνατότητα ανάλυσης ίση με 20 m ανά pixel. Από έναν δορυφόρο στο διάστημα που λειτουργεί σε μήκος κύματος περίπου 5 cm (ραντάρ C-band), για να έχουμε χωρική ανάλυση 10 m, θα χρειαστούμε μια κεραία ραντάρ μήκους περίπου 4.250 m, κάτι το οποίο δεν είναι πρακτικό για έναν αισθητήρα στο διάστημα. Ως εκ τούτου, επιστήμονες και μηχανικοί έχουν καταλήξει στην λύση του συνθετικού διαφράγματος. Δηλαδή μία ακολουθία λήψεων από μια μικρότερη κεραία συνδυάζονται για την προσομοίωση μιας πολύ μεγαλύτερης κεραίας, παρέχοντας έτσι δεδομένα υψηλότερης ανάλυσης (Σχήμα 2.1) [7].



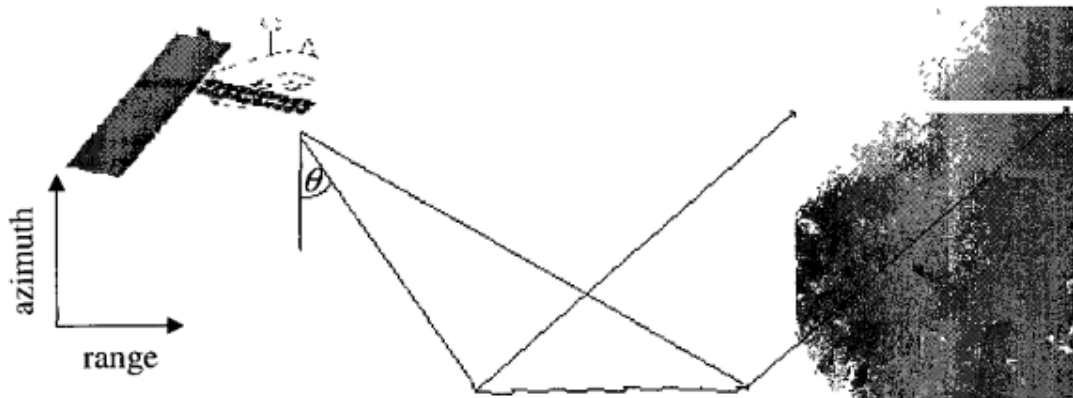
Σχήμα 2.1 : Γεωμετρία παρατήρησης για τον σχηματισμό συνθετικού διαφράγματος για στόχο P κατά  $x=0$  [7]

Μια εικόνα μπορεί να οριστεί ως μια δισδιάστατη συνάρτηση,  $f(x, y)$ , όπου  $x$  και  $y$  είναι χωρικές συντεταγμένες, και η τιμή της  $f$  σε οποιοδήποτε ζεύγος συντεταγμένων  $(x, y)$  ονομάζεται ένταση ή gray level της εικόνας σε εκείνο το σημείο. Όταν τα  $x, y$  και  $f(x, y)$  είναι όλα πεπερασμένα και διακριτά μεγέθη, η εικόνα θεωρείται ψηφιακή εικόνα. Η ψηφιακή επεξεργασία εικόνας περιλαμβάνει οποιαδήποτε τροποποίηση ή εξαγωγή πληροφοριών από μια εικόνα με υπολογιστή.

## Γωνία Πρόσπτωσης

Τα ραντάρ αυτά τα οποία είναι πλευρικής όψης και πετούν στην κατεύθυνση του αζιμουθίου εκπέμπουν και λαμβάνουν μικροκύματα στην κατεύθυνση εμβέλειας (ορθογώνια προς αζιμούθιο). Συνέπεια της χρήσης τεχνολογίας ραντάρ με πλευρική όψη είναι ότι συλλέγεται μια διαφορετική γωνία πρόσπτωσης ή στα αγγλικά “incidence angle”. Αυτή η γωνία πρόσπτωσης είναι σημαντικό να ληφθεί υπόψη, καθώς το backscatter του στόχου μπορεί να ποικίλλει ανάλογα με τις συναρτήσεις των γωνιών πρόσπτωσης [6].

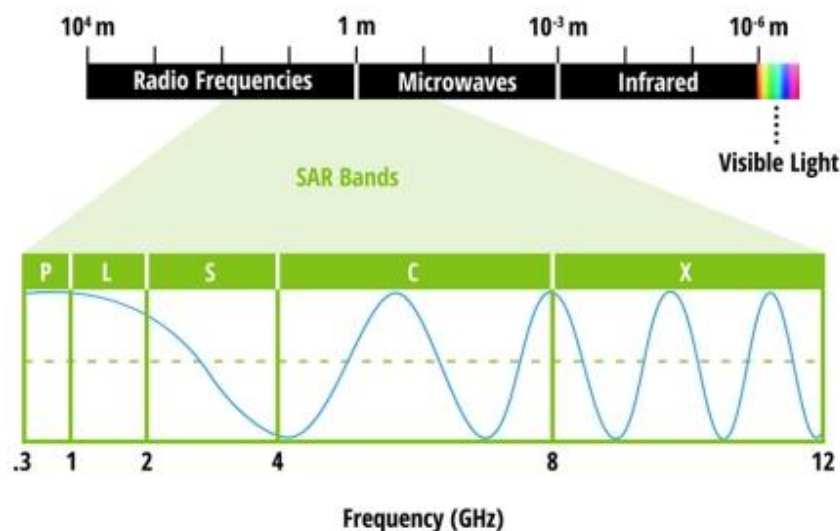




Εικόνα 2.4 Γωνία πρόσπτωσης και SAR [6]

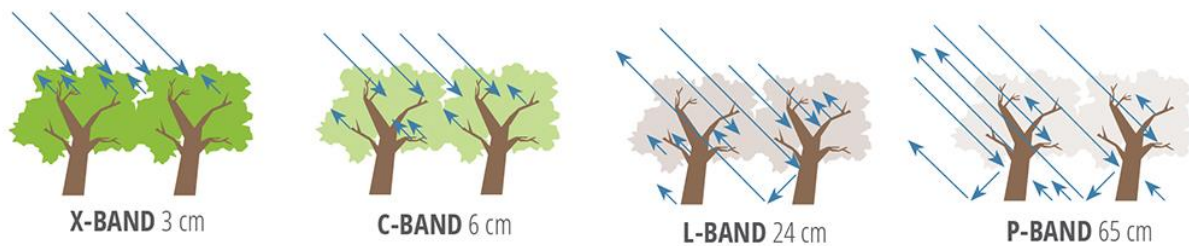
## Συχνότητα και Μήκος Κύματος

Οι αισθητήρες ραντάρ χρησιμοποιούν μήκη κύματος της κλίμακας του εκατοστού έως του μέτρου (Σχήμα 2.2), έτσι έχουν την ικανότητα να βλέπουν μέσα από σύννεφα, την ημέρα ή την νύχτα. Τα διαφορετικά μήκη κύματος του SAR αναφέρονται συχνά ως ζώνες, με χαρακτηρισμούς γραμμάτων όπως X, C, L και P. Η C-band έχει συχνότητα 4-8 GHz και μήκος κύματος 7.5-3.8 cm. Χρησιμοποιείται κυρίως για παγκόσμια χαρτογράφηση, ανίχνευση αλλαγών, παρακολούθηση περιοχών με χαμηλή έως μέτρια διείδυση, υψηλότερη συνοχή όπως στον πάγο και στην θαλάσσια πλοήγηση.



Σχήμα 2.2 : Το ηλεκτρομαγνητικό φάσμα με ένθετες ζώνες μικροκυμάτων [7]

Το μήκος κύματος είναι ένα σημαντικό χαρακτηριστικό που πρέπει να λαμβάνεται υπόψη όταν εργαζόμαστε με SAR, καθώς καθορίζει τον τρόπο με τον οποίο το σήμα του ραντάρ αλληλεπιδρά με την επιφάνεια και πόσο βαθιά μπορεί να διεισδύσει το σήμα σε ένα μέσο. Για παράδειγμα, ένα σήμα της ζώνης C διεισδύει μόνο στα ανώτερα στρώματα του θόλου ενός δάσους και επομένως θα αντιμετωπίσει ως επί το πλείστον σκέδαση τραχύτητας σε συνδυασμό με περιορισμένη ποσότητα σκέδασης όγκου. Το μήκος κύματος δεν επηρεάζει μόνο το βάθος διείσδυσης στα δάση, αλλά και σε άλλους τύπους κάλυψης γης, όπως το έδαφος και ο πάγος.



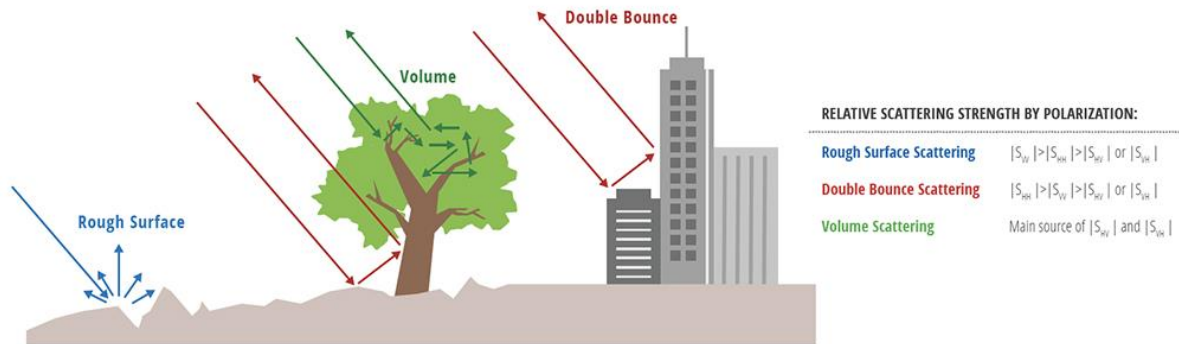
Εικόνα 2.5: Ευαισθησία μετρήσεων SAR βάσει του μήκους κύματος [7]

## Μηχανισμοί Πόλωσης και Σκέδασης

Το ραντάρ μπορεί επίσης να συλλέγει σήματα σε διαφορετικές πολώσεις, ελέγχοντας την αναλυόμενη πόλωση τόσο στη διαδρομή εκπομπής όσο και στη διαδρομή λήψης. Η πόλωση αναφέρεται στον προσανατολισμό του επιπέδου στο οποίο ταλαντώνεται το εκπεμπόμενο ηλεκτρομαγνητικό κύμα. Ενώ ο προσανατολισμός μπορεί να συμβεί σε οποιαδήποτε γωνία, οι αισθητήρες SAR συνήθως εκπέμπουν γραμμικά πολωμένα. Η οριζόντια πόλωση υποδεικνύεται με το γράμμα H και η κατακόρυφη πόλωση με το V. Το πλεονέκτημα των αισθητήρων ραντάρ είναι ότι η πόλωση του σήματος μπορεί να ελεγχθεί με ακρίβεια τόσο στη μετάδοση όσο και στη λήψη. Τα σήματα που εκπέμπονται σε κατακόρυφη (V) και λαμβάνονται σε οριζόντια (H) πόλωση θα υποδεικνύονται με ένα VH. Εναλλακτικά, ένα σήμα που εκπέμπεται σε οριζόντια (H) και λαμβάνεται σε οριζόντια (H) θα υποδεικνύεται με HH, και ούτω καθεξής. Η εξέταση της ισχύος του σήματος από αυτές τις διαφορετικές πολώσεις φέρει πληροφορίες σχετικά με τη δομή της απεικονιζόμενης επιφάνειας, με βάση τους ακόλουθους τύπους σκέδασης: τραχιά επιφάνεια, όγκο και διπλή αναπήδηση (Εικόνα 2.6). Η σκέδαση τραχιών επιφανειών, όπως αυτή που προκαλείται από γυμνό έδαφος ή νερό,



είναι πιο ευαίσθητη στη σκέδαση VV. Η διασπορά όγκου, για παράδειγμα, που προκαλείται από τα φύλλα και τα κλαδιά σε ένα δάσος, είναι πιο ευαίσθητη σε διασταυρούμενα πολωμένα δεδομένα όπως το VH ή το HV. Ο τελευταίος τύπος σκέδασης, η διπλή αναπήδηση, προκαλείται από κτίρια, κορμούς δέντρων ή πλημμυρισμένη βλάστηση και είναι πιο ευαίσθητος σε ένα πολωμένο σήμα HH.



Εικόνα 2.6: Πόλωση και Σκέδαση SAR [7]

### 3 Τεχνητή Νοημοσύνη

Οι ορισμοί της Τεχνητής Νοημοσύνης που έχουν διατυπωθεί ανά τους καιρούς είναι διάφοροι, από τους οποίους μερικοί εστιάζουν στην σκέψη και τον συλλογισμό και άλλοι στην συμπεριφορά.

Μερικοί από αυτούς είναι:

«Η προσπάθεια να κατασκευάσουμε υπολογιστές με διανοητική ικανότητα με την πλήρη και κυριολεκτική έννοια του όρου» - Haugeland

«Η μελέτη των υπολογισμών που καθιστούν εφικτή την αντίληψη, την λογική σκέψη και την αντίδραση» - Winston

Η έρευνα στην Τεχνητή Νοημοσύνη έχει προσφέρει βελτιώσεις σε πολλούς τομείς, όπως αυτοί της μηχανικής μάθησης, της μηχανικής όρασης, της φυσικής επεξεργασίας γλώσσας κ.α. [8].

#### 3.1 Μηχανική Μάθηση

Σημαντικό για να κατανοήσουμε την έννοια της Μηχανικής Μάθησης, είναι πρώτα να ορίσουμε τι είναι η «μάθηση».

*Μάθηση* είναι η διαδικασία που ακολουθούμε για βελτιωθεί η απόδοση ενός συστήματος σε μια εργασία συγκεκριμένη αφότου πρώτα έχει προηγηθεί η παρατήρηση πολλών παραδειγμάτων [8].

Υπάρχουν τρεις προϋποθέσεις για να υπάρξει μάθηση:

1. η ύπαρξη ενός περιβάλλοντος από το οποίο να προσφέρονται τα δεδομένα σε μορφή παραδειγμάτων στο σύστημα.
2. η ύπαρξη ενός κριτηρίου αξιολόγησης για την επίδοση του συστήματος.
3. η ύπαρξη μιας συγκεκριμένης λειτουργίας που το σύστημα πρέπει να εκτελέσει [9].

Κατά τον Άρθουρ Σάμουελ, μηχανική μάθηση είναι «το πεδίο μελέτης που δίνει την ικανότητα στους υπολογιστές να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί» [10].

## Μάθηση με Επίβλεψη

Όπως στον φυσικό κόσμο, έτσι και στον κόσμο των υπολογιστών, υπάρχουν διαφορετικές διαδικασίες μάθησης. Στην μηχανική μάθηση, υπάρχει η «μάθηση με επίβλεψη» και η «μάθηση χωρίς επίβλεψη». Την συγκεκριμένη στιγμή θα αναλύσουμε την μάθηση με επίβλεψη.

Στην επιβλεπόμενη μάθηση ζητείται από το σύστημα να «μάθει» μια συνάρτηση στόχο από ένα σετ δεδομένων τα οποία περιγράφουν ένα μοντέλο. Στο σύστημα δηλαδή υπάρχει κάποιος «επιβλέπων» ο οποίος στο σύστημα της σωστή τιμή εξόδου της συνάρτησης, για τα συγκεκριμένα δεδομένα που εξετάζονται εκείνη την στιγμή. Η έκφραση του μοντέλου που περιγράφει τα δεδομένα είναι η συνάρτηση στόχος. Στόχος της συνάρτησης είναι να προβλέψει την τιμή της «μεταβλητής εξόδου» βάσει των τιμών που δέχεται το μοντέλο στην είσοδό του και ονομάζονται «μεταβλητές εισόδου». «Σύνολο περιπτώσεων» ονομάζεται το σύνολο των διαφορετικών τιμών που μπορεί να δεχτεί η συνάρτηση. Ένα σύνολο «χαρακτηριστικών» δύναται να περιγράφει κάθε περίπτωση. Τέλος, «δεδομένα εκπαίδευσης» ονομάζονται τα δεδομένα - περιπτώσεις εκείνα για τα οποία γνωρίζουμε και την τιμή εισόδου και την τιμή εξόδου. «Υποθέσεις» ονομάζονται οι διαφορετικές συναρτήσεις που δοκιμάζει το σύστημα ώστε να πετύχει τελικά την καλύτερη συνάρτηση στόχο [8].

Η μάθηση με επίβλεψη στηρίζεται στην θεωρία της *επαγωγικής μάθησης*, δηλαδή μια υπόθεση που έχει δοκιμαστεί και αποδειχθεί ότι για ένα μεγάλο σύνολο παραδειγμάτων προσεγγίζει ικανοποιητικά την συνάρτηση στόχο, τότε θα την προσεγγίζει το ίδιο καλά ακόμα και για περιπτώσεις που δεν έχει εξεταστεί [8].

Κάτω από την ομπρέλα της μηχανικής μάθησης με επίβλεψη, υπάρχουν δύο κατηγορίες προβλημάτων. Τα προβλήματα *ταξινόμησης (classification)* και *παρεμβολής (regression)*. [8]

## Ταξινόμηση

Η κατηγορία της *ταξινόμησης* ασχολείται με την δημιουργία μοντέλων πρόβλεψης κλάσεων. Ένα παράδειγμα ταξινόμησης θα μπορούσε να είναι η διάκριση ζώων σε τρεις κλάσεις, σκύλος – γάτα – πρόβατο. Ή καλύτερα, όπως και στην εφαρμογή της συγκεκριμένης εργασίας, η διάκριση «παγόβουνο ή πλοίο».

## Κριτήρια Επίδοσης στην Ταξινόμηση

Αναλόγως το πρόβλημα που λύνει το κάθε μοντέλο, επιλέγεται και το ανάλογο κριτήριο αξιολόγησης της επίδοσής του. Διαφορετικά κριτήρια χρησιμοποιούνται για τα προβλήματα ταξινόμησης και διαφορετικά για τα προβλήματα παλινδρόμησης.

Σε μια εφαρμογή που στοχεύει στην ταξινόμηση δύο κλάσεων είναι πολύ σύνηθες να συναντήσουμε η μια κλάση να αντιπροσωπεύει την εκπλήρωση της εξεταζόμενης συνθήκης, δηλαδή TRUE, και η άλλη κλάση να αντιπροσωπεύει την μη εκπλήρωση της εξεταζόμενης συνθήκης, δηλαδή FALSE. Αναπαριστούμε τα TRUE με την τιμή «1» και τα FALSE με την τιμή «0». Για παράδειγμα σε μια εφαρμογή αναγνώρισης ασθενών που έχουν γρίπη, τα  $X_i$  πρότυπα περιέχουν πληροφορίες για τον ασθενή όπως, θερμοκρασία, ύπαρξη ή μη πόνου στον λαιμό, ύπαρξη ή μη βουλωμένης μύτης κ.α. Η κλάση ένα περιέχει τα πρότυπα των ασθενών με γρίπη και η κλάση 0 τα πρότυπα των υγιών ασθενών. Η συνάρτηση διαχωρισμού που δημιουργεί ο ταξινομητής είναι η  $Y_i = f(X_i, W)$ . Η συνάρτηση αυτή παραμετροποιείται από το  $W$  το οποίο κατά την εκπαίδευση μαθαίνει.

- Αν έχουμε  $Y_i = 1$ , δηλαδή το πρότυπο βγει θετικό, τότε το  $X_i$  ταξινομείται στην κλάση 1 (ασθενής που έχει γρίπη)
- Αν έχουμε  $Y_i = 0$ , δηλαδή το πρότυπο βγει αρνητικό, τότε το  $X_i$  ταξινομείται στην κλάση 0 (ασθενής υγιής)

Το ιδανικό θα ήταν για όλα τα πρότυπα  $X_i$ , το  $Y_i$  να ταυτίζεται με την την ετικέτα  $t_i$  που έχει η κλάση στην οποία ανήκει το πρότυπο. Επειδή όμως δεν συμβαίνει πάντα αυτό, υπάρχουν τέσσερις περιπτώσεις:

1. True Negatives, δηλαδή  $Y_i = 0$  και  $t_i = 0$
2. False Negatives, δηλαδή  $Y_i = 1$  και  $t_i = 0$
3. False Positives, δηλαδή  $Y_i = 0$  και  $t_i = 1$
4. True Positives, δηλαδή  $Y_i = 1$  και  $t_i = 1$

Η επίδοση ενός ταξινομητή αξιολογείται με τον λεγόμενο πίνακα σύγχυσης. Στην ιδανική περίπτωση θα θέλαμε τα False Positives και τα False Negatives να είναι ίσα με μηδέν. Από τα παραπάνω προκύπτει ένα μέτρο εκτίμησης επίδοσης το οποίο ονομάζεται «Ακρίβεια» και ορίζεται ως:

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

		ACTUAL	
		Negative	Positive
PREDICTION	Negative	TRUE NEGATIVE	FALSE NEGATIVE
	Positive	FALSE POSITIVE	TRUE POSITIVE

Πίνακας 3.1 : Πίνακας σύγκρισης για πρόβλημα ταξινόμησης 2 κλάσεων

Προκύπτει πως  $0 \leq \text{Accuracy} \leq 1$ , και η επιθυμητή τιμή θα ήταν  $\text{Accuracy}=1$ .

Ορισμένες φορές η *Ακρίβεια* λόγω της γενικότητάς της δεν είναι αντιπροσωπευτική της επιτυχίας του μοντέλου. Για αυτόν τον λόγο υπάρχουν και άλλα κριτήρια επίδοσης τα οποία ονομάζονται «*Ευστοχία*» (“*Precision*”) και «*Ανάκληση*» (“*Recall*”).

*Ευστοχία* είναι το ποσοστό εκείνο των προτύπων που το μοντέλο κατηγοριοποιεί ως Θετικά και ανήκουν όντως στην θετική κατηγορία. Δηλαδή:

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

Όσο μεγαλύτερη είναι η Ευστοχία τότε τόσο μικρότερο είναι το σύνολο των FP.

*Ανάκληση* είναι το ποσοστό εκείνο των προτύπων που ανήκουν στην Θετική κατηγορία, και το μοντέλο κατηγοριοποιεί ορθώς ως Θετικά. Δηλαδή:

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

Όσο μεγαλύτερη είναι η Ανάκληση τότε τόσο μικρότερος είναι ο αριθμός των Θετικών δειγμάτων που έχουν κατηγοριοποιηθεί λάθος.

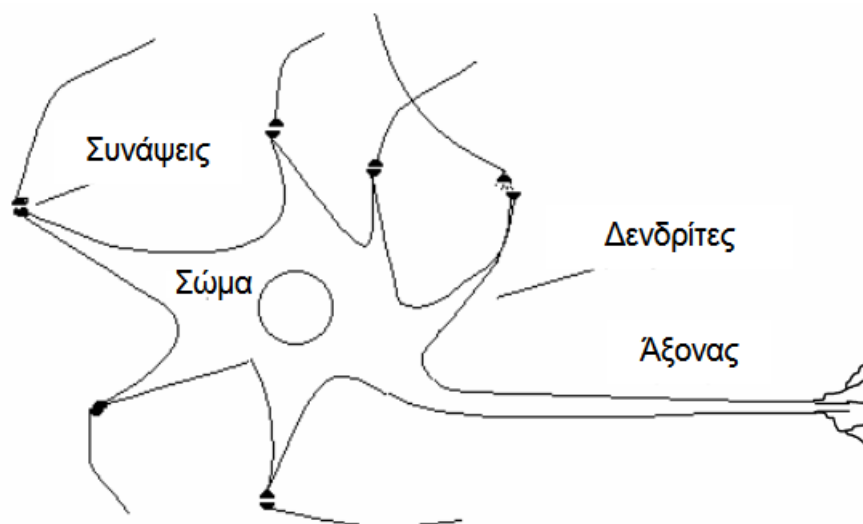
## 4 Νευρωνικά Δίκτυα

Το επιστημονικό πεδίο των Τεχνητών Νευρωνικών Δικτύων, που πολλές φορές για συντομία αναφέρονται ως «νευρωνικά δίκτυα», βασίσει το έργο του στο γεγονός πως ο ανθρώπινος εγκέφαλος εκτελεί τους απαραίτητους υπολογισμούς για την εκτέλεση κάποια λειτουργίας με πολύ διαφορετικό τρόπο από έναν τυπικό ηλεκτρονικό υπολογιστή. Ο μη γραμμικός και πολύπλοκος ανθρώπινος εγκέφαλος μπορεί να οργανώνει τους νευρώνες του οι οποίοι λειτουργούν ως δομικά στοιχεία με τέτοιο τρόπο ώστε να είναι σε θέση να φέρουν εις πέρας υπολογισμούς με πολλαπλάσια ταχύτητα από αυτή που πετυχαίνει ο καλύτερος ηλεκτρονικός υπολογιστής σήμερα. [11]

Τα νευρωνικά δίκτυα είναι μια μοναδική προσέγγιση στην προσπάθεια δημιουργίας συστημάτων που να διαθέτουν νοημοσύνη καθώς δεν αναπαριστούν ρητά την γνώση και δεν υιοθετούν αλγορίθμους ειδικά σχεδιασμένους. Για την λειτουργία τους όμως βασίζονται σε βιολογικά πρότυπα εμπνευσμένα από τον ανθρώπινο εγκέφαλο. [8]

### 4.1.1 Βιολογικά Νευρωνικά Δίκτυα

Η ανθρώπινη ικανότητα της σκέψης, της μνήμης και της επίλυσης προβλημάτων αποδίδεται στον εγκέφαλο, του οποίου η δομική μονάδα είναι ο *νευρώνας*.



Εικόνα 4.1: Αναπαράσταση του βιολογικού νευρώνα

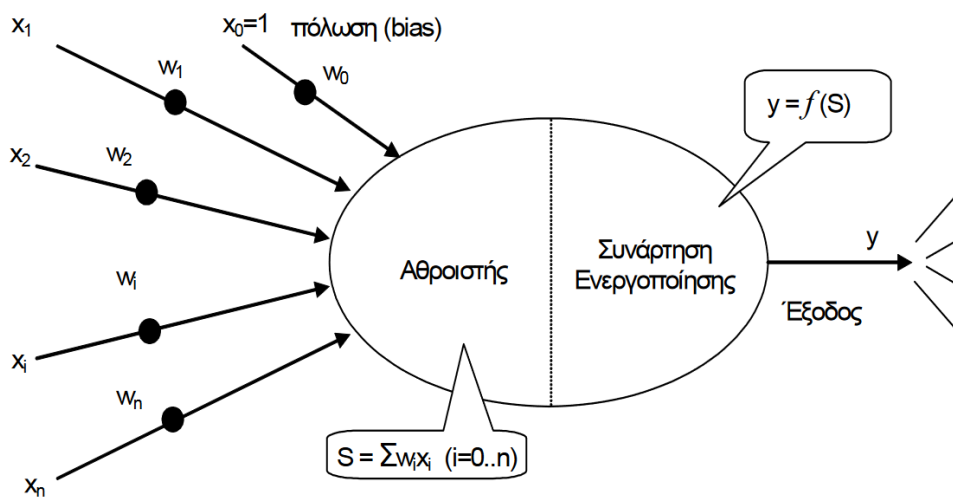
Ένας βιολογικός νευρώνας του εγκεφάλου απαρτίζεται από το σώμα το οποίο είναι ο πυρήνας του νευρώνα, τους δενδρίτες από τους οποίους δέχεται σήματα από κοντινούς

νευρώνες που λειτουργούν ως σημεία εισόδου. Αποτελείται επίσης από τον άξονα ο οποίος λειτουργεί ως έξοδος του νευρώνα και το μέσο για την σύνδεσή του με άλλους νευρώνες. Σημαντικό ρόλο για ένα νευρώνα παίζει ένα απειροελάχιστο κενό που υπάρχει σε κάθε δενδρίτη, το οποίο ονομάζεται σύναψη. Οι συνάψεις επιβραδύνουν ή επιταχύνουν την ροή ηλεκτρικών φορτίων στο σώμα του νευρώνα. Η ικανότητα των συνάψεων να μεταβάλλουν την αγωγιμότητά τους δίνει τις ικανότητες στον εγκέφαλο της μάθησης και της μνήμης. Τα ηλεκτρικά σήματα που μπαίνουν στο σώμα των νευρώνων από τους δενδρίτες συνδυάζονται και αν το αποτέλεσμα τους ξεπερνά μια τιμή τότε το σήμα μεταφέρεται μέσω του άξονα και στους υπόλοιπους νευρώνες. [8]

Ενδεικτικά αναφέρεται πως ο εγκέφαλος ενός νεογέννητου ανθρώπου περιέχει περίπου εκατό τρισεκατομμύρια συνάψεις. Τα τεχνητά μοντέλα που κατασκευάζονται σήμερα έχουν το πολύ ένα εκατομμύριο συνάψεις. [8]

#### 4.1.2 Μοντέλο Τεχνητού Νευρώνα

Ο τεχνητός νευρώνας (Σχήμα 2.2) χαρακτηρίζεται ως ένα υπολογιστικό μοντέλο το οποίο έχει αντίστοιχα μέρη με αυτά του βιολογικού νευρώνα.



Σχήμα 4.1 : Μοντέλο του τεχνητού νευρώνα

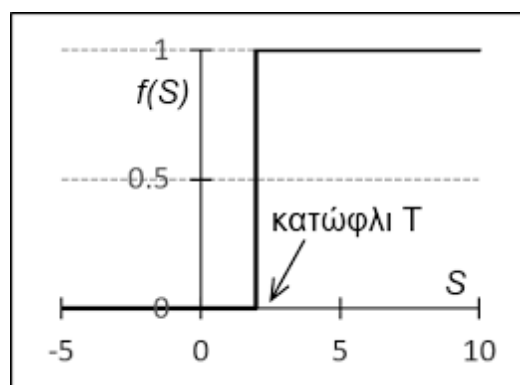
Ως είσοδο, ένας τεχνητός νευρώνας δέχεται κάποια σήματα εισόδου  $X_i$ , τα οποία σε αντίθεση με τους ηλεκτρικούς παλμούς του εγκεφάλου, αυτά είναι συνεχείς μεταβλητές. Τα σήματα εισόδου μεταβάλλονται από μια τιμή  $W_i$  η οποία ονομάζεται «βάρος». Τα βάρη σε ένα τεχνητό νευρώνα είναι οι αντίστοιχες συνάψεις του βιολογικού νευρώνα. Η τιμή του

βάρους μπορεί να πάρει και θετικές και αρνητικές τιμές, σε αναλογία με το αν η λειτουργία της σύναψης θα ήταν επιταχυντική ή επιβραδυντική. Ένας νευρώνας μπορεί να έχει κάποιο βάρος  $W_0$  το οποίο λέγεται πόλωση και λειτουργεί ως προδιάθεση του νευρώνα.

Το κυρίως σώμα του τεχνητού νευρώνα αποτελείται από δύο μέρη. Το ένα είναι ο «αθροιστής» και το άλλο είναι η «συνάρτηση ενεργοποίησης». Ο αθροιστής προσθέτει τα σήματα εισόδου. Η συνάρτηση ενεργοποίησης λειτουργεί ως φίλτρο διαμορφώνοντας την τιμή του σήματος εξόδου  $y$ . Ένας νευρώνας μπορεί να έχει περισσότερες από μια εξόδους, αλλά η τιμή σε όλες θα είναι ίδια. [8]

Υπάρχουν τρεις βασικές περιπτώσεις για την συνάρτηση ενεργοποίησης του τεχνητού νευρώνα:

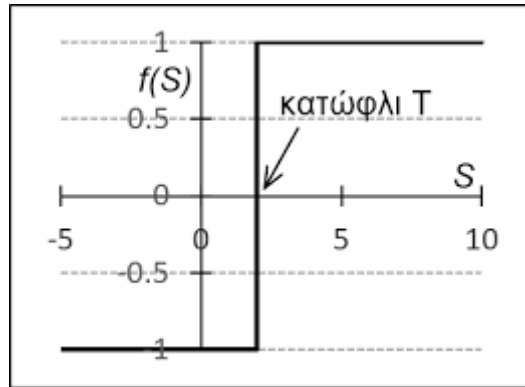
- Η *βηματική συνάρτηση* (Σχήμα 2.3), δίνει αποτέλεσμα στην έξοδο μόνο αν η τιμή που υπολογίζεται από τον αθροιστή είναι μεγαλύτερη από τιμή κατώφλιου  $T$ .



Σχήμα 4.2: Βηματική συνάρτηση

- Η *συνάρτηση προσήμου* (Σχήμα 2.4), δίνει στην έξοδο θετική ή αρνητική τιμή αναλόγως με το αν η τιμή που υπολογίζεται από τον αθροιστή είναι μικρότερη ή μεγαλύτερη από την τιμή κατώφλιου  $T$ .



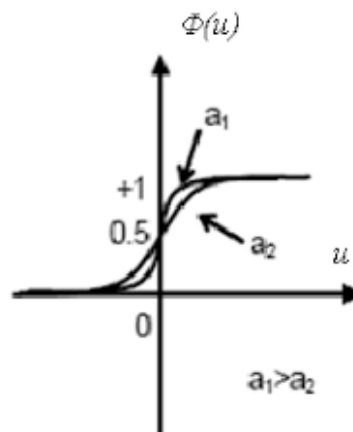


Σχήμα 4.3: Συνάρτηση Προσήμου

- Η *λογιστική συνάρτηση* (Σχήμα 2.5), η οποία εκφράζεται από την παρακάτω σχέση,

$$f(S) = \frac{1}{1 + e^{-\alpha * S}}$$

με το  $\alpha$  να είναι συντελεστής ο οποίος ρυθμίζει την ταχύτητα μεταξύ των δύο ασυμπτωματικών τιμών.



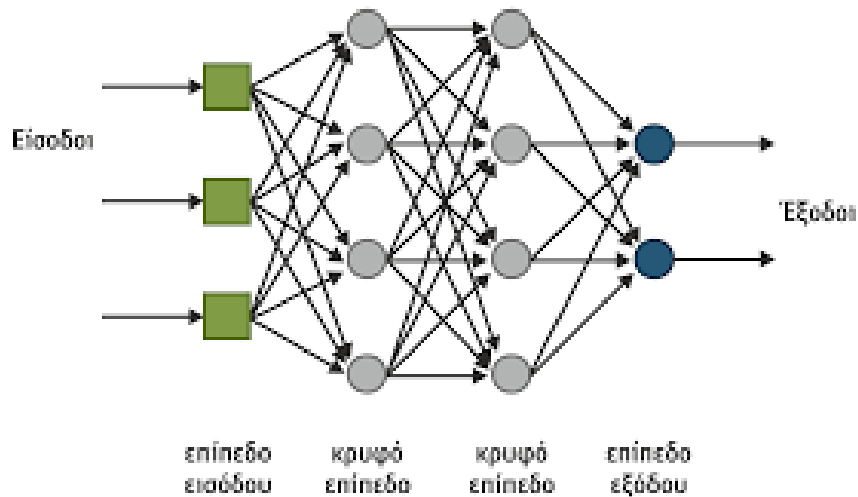
Σχήμα 4.4: Λογιστική συνάρτηση

Η μη-γραμμικότητα είναι το κοινό χαρακτηριστικό των παραπάνω συναρτήσεων. Μια συνάρτηση ενεργοποίησης γραμμική θα παρήγαγε πάντα έξοδο ανάλογη με την τιμή της εισόδου, πράγμα που θα έθετε αχρείαστους περιορισμούς. [11]

### 4.1.3 Τεχνητά Νευρωνικά Δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα είναι συστήματα τα οποία επεξεργάζονται δεδομένα. Απαρτίζονται από πληθώρα τεχνητών νευρώνων οι οποίοι είναι οργανωμένοι σε δομές οι

όμοιες με αυτές του ανθρώπινου εγκεφάλου. Οι τεχνητοί αυτοί νευρώνες είναι οργανωμένοι σε επίπεδα. Η εισαγωγή των δεδομένων στο δίκτυο γίνεται από το πρώτο επίπεδο το οποίο ονομάζεται «επίπεδο εισόδου». Στο δίκτυο, μετά το επίπεδο εισόδου, μπορεί να υπάρχουν ένα ή περισσότερα «κρυφά επίπεδα» και τέλος υπάρχει το «επίπεδο εξόδου». (Σχήμα 2.6)



Σχήμα 4.5: ΤΝΔ πρόσθιας τροφοδότησης 3-4-4-2 με πλήρη διασύνδεση μεταξύ των διαδοχικών επιπέδων

Η νευρώνες σε ένα τεχνητό νευρωνικό δίκτυο μπορούν να είναι είτε πλήρως συνδεδεμένοι μεταξύ τους είτε μερικώς. Όταν όλοι οι νευρώνες συνδέονται με όλους τους υπόλοιπους τότε λέμε πως είναι πλήρως συνδεδεμένοι ενώ σε διαφορετική περίπτωση είναι μερικώς συνδεδεμένοι. Τεχνητά νευρωνικά δίκτυα πρόσθιας τροφοδότησης χαρακτηρίζονται εκείνα όπου δεν υπάρχει σύνδεση μεταξύ των νευρώνων του ενός επιπέδου με το προηγούμενο, η πληροφορίας δηλαδή ρέει προς τα εμπρός. Σε διαφορετική περίπτωση, τα ονομάζουμε τεχνητά νευρωνικά δίκτυα με ανατροφοδότηση. [8]

## Βαθιά Μάθηση

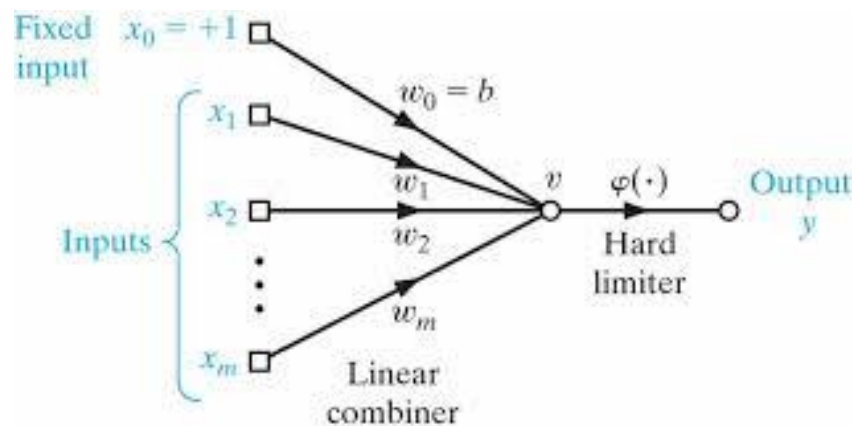
Εάν ένα μοντέλο τεχνητής νοημοσύνης έχει την δυνατότητα να αναπαρασταθεί από στρώματα νευρώνων, τότε το πλήθος αυτών των στρωμάτων το ονομάζουμε «βάθος» του μοντέλου. Ένα τεχνητό νευρωνικό δίκτυο, το οποίο ανάμεσα στην είσοδο και την έξοδό του έχει δύο στρώματα, λέμε πως έχει βάθος 2.

Ρηγά ονομάζονται τα μοντέλα με μικρό βάθος, δηλαδή από 2-4. Από την άλλη μεριά, βαθιά ονομάζονται τα μοντέλα με μεγάλο βάθος, δηλαδή μεγαλύτερο από 5. Βαθιά μοντέλα που χρησιμοποιούνται σήμερα έχουν βάθος ακόμα και μεγαλύτερο του 100.

Τα μοντέλα βαθιάς μάθησης είναι πολύ δημοφιλή τα τελευταία χρόνια καθώς πετυχαίνουν κορυφαίες επιδόσεις σε αρκετά προβλήματα.

## Multilayer Perceptron

Το perceptron είναι μια τοπολογία δικτύου πρόσθιας τροφοδότησης, απλή, χωρίς κρυφά επίπεδα, και είναι η πρώτη προσέγγιση στα τεχνητά νευρωνικά δίκτυα η οποία προτάθηκε από τον Rosenblatt το 1958. Το perceptron έχει τις βάσεις στο μοντέλο μη γραμμικού νευρώνα των McCulloch-Pitts.



Σχήμα 4.6: Γράφημα ροής του σήματος perceptron

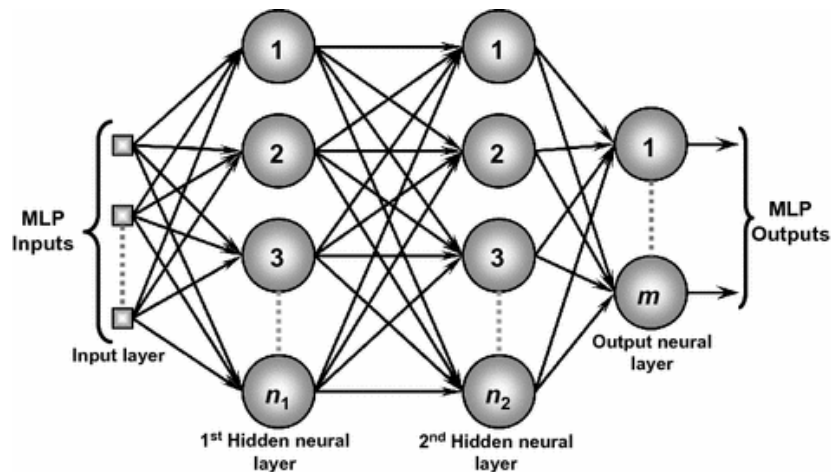
Στόχος του perceptron είναι η σωστή ταξινόμηση των εισόδων σε κλάσεις. Από το μοντέλο συμπεραίνουμε πως η είσοδος του hard limiter είναι  $u = \sum_{i=1}^m W_i X_i + b$

Επειδή οι δυνατότητες του perceptron του Rosenblatt, το οποίο στην πράξη είναι έναν νευρωνικό δίκτυο με ένα επίπεδο, περιορίζονται στην ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, θα μελετήσουμε το *perceptron πολλών επιπέδων (multilayer perceptron)*. Τα χαρακτηριστικά των multilayer perceptron είναι τα ακόλουθα:

- Μια διαφορίσιμη γραμμική συνάρτηση ενεργοποίησης περιλαμβάνεται στο μοντέλο κάθε νευρώνα.

- Το δίκτυο αποτελείται από ένα ή και περισσότερα κρυφά επίπεδα ανάμεσα στους κόμβους της εισόδου και εξόδου.
- Τα συναπτικά βάρη του δικτύου καθορίζουν την διασυνδεσιμότητα του δικτύου η οποία είναι πολύ μεγάλη.

Παρακάτω απεικονίζεται ένα δίκτυο perceptron πολλών επιπέδων, το οποίο έχει δύο κρυφά επίπεδα και ένα επίπεδο εξόδου.



Σχήμα 4.7 : Multilayer perceptron πολλών επιπέδων με δύο κρυφά επίπεδα

Τα χαρακτηριστικά αυτά είναι επίσης υπεύθυνα για την περιορισμένη γνώση μας σχετικά με την συμπεριφορά του δικτύου. Η χρήση κρυφών νευρώνων κάνει δύσκολη της πλήρη κατανόηση της διαδικασίας μάθησης.

#### 4.1.4 Μάθηση και Ανάκληση

Οι κύριες λειτουργίες που πραγματοποιούν τα τεχνητά νευρωνικά δίκτυα είναι δυο: η *μάθηση* και η *ανάκληση*. Η διαδικασία με την οποία το δίκτυο τροποποιεί τις τιμές στα βάρη του ώστε να μπορέσει να επιτύχει την επιθυμητή έξοδο βάσει των τιμών εισόδου που έχει λάβει ονομάζεται *Μάθηση*. Η μάθηση αναφέρεται συχνά και ως *εκπαίδευση* του τεχνητού νευρωνικού δικτύου. Η διαδικασία για τον υπολογισμό του διανύσματος της εξόδου για συγκεκριμένες τιμές εξόδου και βαρών ονομάζεται *Ανάκληση*. [8]

Κάποια προβλήματα που μπορεί να προκύψουν κατά την διάρκεια της εκπαίδευσης ενός μοντέλου είναι η *υποπροσαρμογή* και η *υπερπροσαρμογή*. Ένα απλό τεχνητό νευρωνικό

δίκτυο χωρίς σύνθετες παραμέτρους, μπορεί να αποτύχει στην δημιουργία ενός αντιπροσωπευτικού μοντέλου για τα δεδομένα εκπαίδευσης και έτσι να οδηγήσει σε ατελή μάθηση. Από την άλλη μεριά, ένα πολύ σύνθετο τεχνητό νευρωνικό δίκτυο μπορεί να μοντελοποιήσει σε υπερβολικό βαθμό τα δεδομένα εκπαίδευσης και το θόρυβο που ίσως βρίσκεται εκεί, και αυτό θα έχει ως αποτέλεσμα να τα απομνημονεύσει. Σε μια τέτοια περίπτωση, το δίκτυο θα μπορεί να κάνει σωστές προβλέψεις για τα δεδομένα εισόδου στα οποία έχει εκπαιδευτεί αλλά δεν θα μπορεί να κάνει σωστές προβλέψεις σε δεδομένα εισόδου που «βλέπει» για πρώτη φορά. Ένας τρόπος για την αποφυγή τέτοιων προβλημάτων είναι η χρήση μεγάλου όγκου δεδομένων εκπαίδευσης. [8]

Τα δεδομένα εκπαίδευσης χρησιμοποιούνται κυρίως σε κύκλους εκπαίδευσης οι οποίοι ονομάζονται «εποχές». «Μάθηση δέσμης» ονομάζεται η μέθοδος κατά την οποία σε κάθε εποχή, το δίκτυο δέχεται στην είσοδό του μεμονωμένα όλα τα διανύσματα εκπαίδευσης και με κατάλληλους υπολογισμούς μεταβάλλει την τιμή των βαρών. Η εκπαίδευση τερματίζεται όταν ολοκληρωθούν οι εποχές εκπαίδευσης που έχουν οριστεί ή όταν το κριτήριο ελέγχου ποιότητας που έχει οριστεί για το δίκτυο φτάσει στην επιθυμητή τιμή. Η κανονικοποίηση των δεδομένων εισόδου και εξόδου είναι ένα σημαντικό ζήτημα που επηρεάζει την ταχύτητα και την ποιότητα εκπαίδευσης του δικτύου. [8]

## Gradient Descent

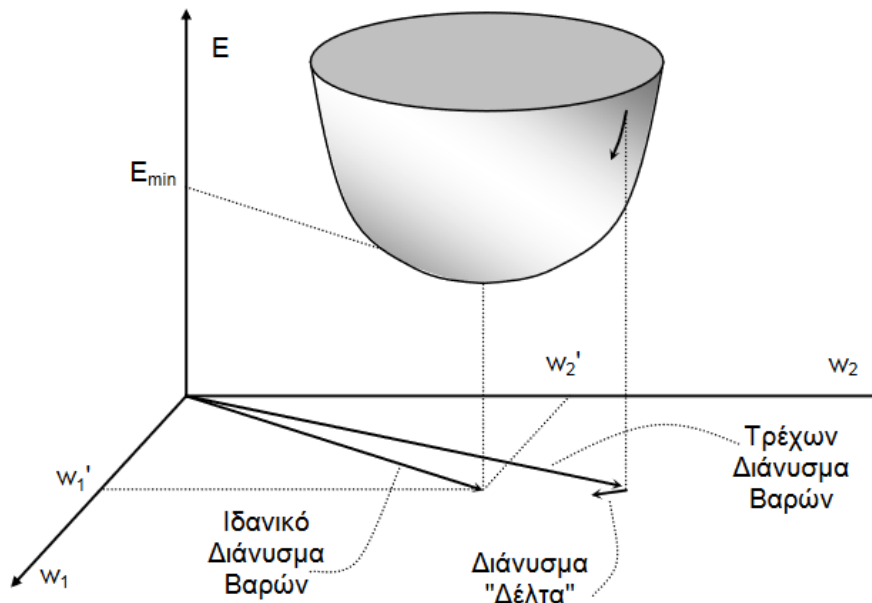
Ο κανόνας επικλινούς μεθόδου όπως ονομάζεται, αναπτύχθηκε στις αρχές της δεκαετίας του '60 και είναι ένας αλγόριθμος εκπαίδευσης όπου η μάθηση είναι καθοδηγούμενη από το σφάλμα. Ο κανόνας προκύπτει από την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (mse) των δεδομένων εκπαίδευσης και αναπαριστά την απόσταση του δικτύου από την επιθυμητή τιμή [8].

Για  $p$  διανύσματα εκπαίδευσης το μέσο τετραγωνικό σφάλμα που θα συμβολίζουμε « $E$ » περιγράφεται από την σχέση:

$$E = \frac{1}{p} \sum_{k=1}^p (t_k - input_k)^2$$

Ο δείκτης αφορά τα δεδομένα εκπαίδευσης και συμβολίζουμε το σήμα εισόδου του νευρώνα ως  $input_k$  και την επιθυμητή έξοδο του ως  $t_k$ . Το συνολικό σήμα εισόδου για κάποιο  $p$  με  $n$  να είναι ο αριθμός των σημάτων της εισόδου του νευρώνα είναι:

$$input = \sum_{i=1}^p W_i X_i$$



Σχήμα 4.8: Αναπαράσταση του Gradient Descent Γεωμετρικά [8]

Όπως φαίνεται στο Σχήμα 4.8, το τετράγωνο του σφάλματος σε συνάρτηση με τα βάρη έχει μορφή ενός παραβολοειδούς εκ περιστροφής όπου τα κοίλα είναι προς τα πάνω. Η ελάχιστη τιμή αυτού είναι το ελάχιστο μέσο σφάλμα και τα διανύσματα βαρών που αντιστοιχούν σε αυτόν διαμορφώνουν το ιδανικό διάνυσμα βαρών.

Ο κανόνας της επικλινούς μεθόδου ακολουθεί την αρνητική κλίση της επιφάνειας σφάλματος και έχει κατεύθυνση προς το ελάχιστο αυτής. Εξασφαλίζει επίσης τον καλύτερο τρόπο της μετακίνησης των βαρών για την επίτευξη του ιδανικού διανύσματος.

## Back Propagation

Ο αλγόριθμος «Back Propagation» είναι μια δημοφιλής μέθοδος εκπαίδευσης των multilayer perceptron. Ο αλγόριθμος επικλινούς καθόδου δεν μπορεί να εφαρμοστεί μόνος του σε δίκτυα με κρυφά επίπεδα διότι δεν είναι γνωστή σε κάθε περίπτωση η επιθυμητή έξοδος για

κάθε νευρώνα του δικτύου. Την επιθυμητή έξοδο την γνωρίζουμε μόνο για το επίπεδο εξόδου. Λύση σε αυτό πρόσφερε η μέθοδος ανάστροφης μετάδοσης λάθους (back propagation).

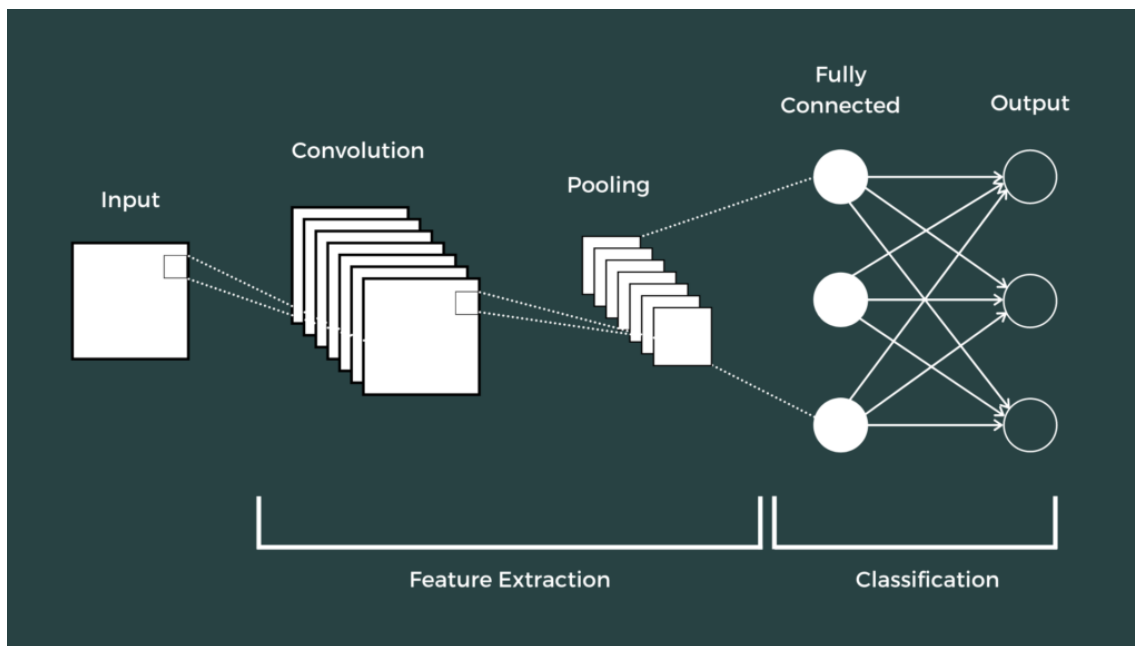
Εδώ η εκπαίδευση αποτελείται από δύο φάσεις:

1. Στην προς τα εμπρός φάση, όπου τα συναπτικά βάρη παραμένουν σταθερά και μέσω του δικτύου μεταδίδεται προς τα εμπρός το σήμα της εισόδου έως ότου φτάσει στην έξοδο. Εδώ, οι αλλαγές υφίστανται στα δυναμικά ενεργοποίησης αλλά και στις εξόδους των νευρώνων.
2. Στην προς τα πίσω φάση, συγκρίνοντας τη έξοδο του δικτύου με την επιθυμητή τιμή δημιουργείται ένα σήμα σφάλματος. Το σήμα αυτό μεταδίδεται πάλι μέσω του δικτύου αλλά αυτή την φορά υπάρχει κατεύθυνση προς τα πίσω. Στην φάση αυτή γίνονται αλλαγές στα βάρη.

Ο αλγόριθμος back propagation είναι μια διαδικασία για την βελτιστοποίηση της gradient descent η οποία ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα ανάμεσα στην έξοδο του δικτύου και της επιθυμητής εξόδου αυτού.

#### 4.1.5 Συνελκτικὰ Νευρωνικά Δίκτυα

Τα συνελκτικὰ νευρωνικά δίκτυα ή όπως ονομάζονται στα αγγλικά Convolutional Neural Networks (CNN) είναι (συνήθως βαθιά) νευρωνικά δίκτυα με πολλά στρώματα και χρησιμοποιούνται με μεγάλη επιτυχία σε προβλήματα αναγνώρισης εικόνας. Η δομή των CNN απαρτίζεται από εναλλασσόμενα μεταξύ τους στρώματα *συνέλιξης (convolution)* και *υποδειγματοληψίας (pooling)*. Μετά από αυτά ακολουθούν ένα ή και περισσότερα *πλήρως συνδεδεμένα* μεταξύ τους στρώματα τα οποία λειτουργούν ως *ταξινομητής* [12].



Σχήμα 4.9: Αρχιτεκτονική CNN

Η αλληλουχία των στρωμάτων Convolution και Pooling, δημιουργεί έναν μετασχηματισμό της εικόνας εισόδου, η οποία έχει τροφοδοτηθεί στο δίκτυο ως πίνακας, σε έναν *χάρτη χαρακτηριστικών (feature map)* στην έξοδο του πλήρως διασυνδεδεμένου στρώματος. Στην συνέχεια, αυτό ο χάρτης τροφοδοτείται ως είσοδος στον ταξινομητή [9].

Τα ευρήματα των Hubel και Wiesel τα οποία σχετίζονταν με τον οπτικό φλοιό μιας γάτας αποτέλεσαν έμπνευση για αυτά τα μοντέλα. Τα ευρήματα έδειξαν πως υπάρχουν *απλοί νευρώνες* που είναι υπεύθυνοι να εξάγουν τοπικά χαρακτηριστικά από ένα σημείο μελέτης και *σύνθετοι νευρώνες* οι οποίοι συνδυάζουν εξόδους από πολλούς απλούς νευρώνες κάποια γειτονιάς και δημιουργούν χαρακτηριστικά υψηλότερου επιπέδου. Έτσι μπορούν να

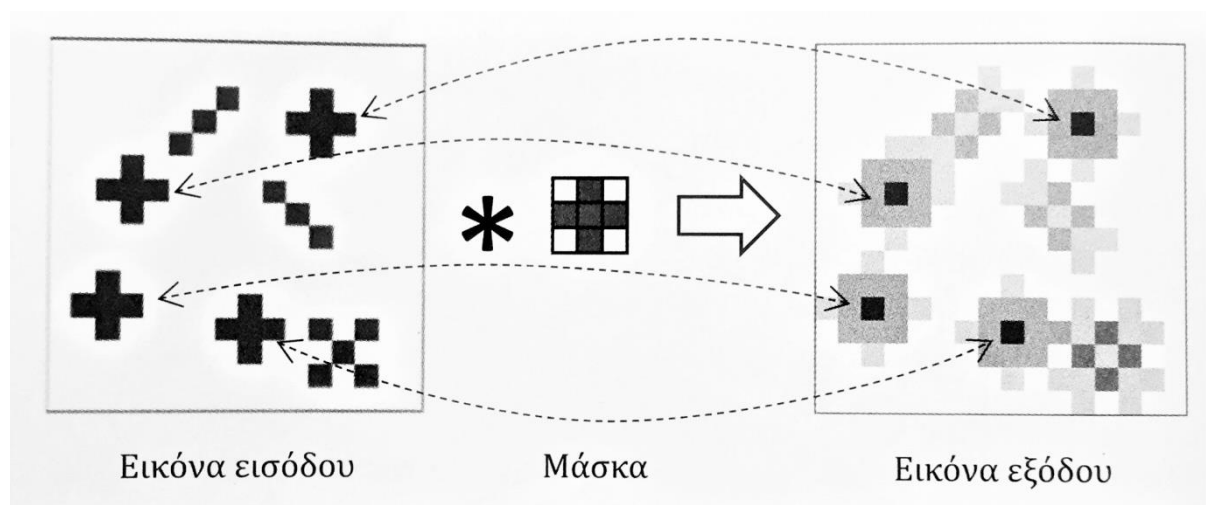


αναγνωριστούν αντικείμενα σε μία εικόνα ανεξαρτήτως από το αν μεταποτιστούν ή περιστραφούν [9].

## Convolutional Layer

Κάθε στρώμα συνέλιξης αποτελείται από ένα πλήθος feature maps (έστω  $C$ ). Κάθε feature map είναι ένα πλέγμα δύο διαστάσεων από νευρώνες, με διαστάσεις  $N \times N$  όπου ο κάθε ένας διεγείρεται από μία μικρή περιοχή του προηγούμενου στρώματος. Άρα κάθε στρώμα έχει  $N \times N \times C$  νευρώνες. Η περιοχή αυτή ονομάζεται *τοπικό υποδεκτικό πεδίο* του νευρώνα αυτού. *Μάσκα* ονομάζονται τα συναπτικά βάρη τα οποία συνδέουν έναν νευρώνα με τους αντίστοιχους νευρώνες του τοπικού υποδεκτικού του πεδίου. Το αποτέλεσμα από την συνέλιξη στην συνέχεια περνάει από συνάρτηση ενεργοποίησης νευρώνα όπως η ReLU.

Στην ουσία η συνέλιξη δουλεύει ως ένα φίλτρο το οποίο ελέγχει όλη την εικόνα εισόδου και στην συνέχεια δίνει μέγιστη απόκριση εκεί που υπάρχει κάποιο τοπικό χαρακτηριστικό παρόμοιο με το σχήμα της μάσκας.



Σχήμα 4.10: Συνέλιξη και μάσκα. Η εικόνα της εξόδου έχει τιμή μέγιστη εκεί που εντοπίζεται σταυρός.

## Pooling Layer

Συνήθως μετά από κάθε συνελκτικό στρώμα εισάγεται ένα στρώμα pooling (υποδειγματοληψίας). Κύριος στόχος αυτού του στρώματος είναι να μειώσει την ευαισθησία του στρώματος αυτού σε μικρές μετατοπίσεις των αντικειμένων που περιέχονται στην εικόνα, αλλά και να αφαιρέσει τις λεπτομέρειες που δεν μας ενδιαφέρουν για το γενικευμένο

μοντέλο που επιθυμούμε να φτιάξουμε. Αυτά βέβαια, τα καταφέρνει χωρίς υπάρχει κάποια απώλεια στην ικανότητα του δικτύου να διαχωρίζει αντικείμενα.

### **Fully-Connected Layer**

Τέλος, στα τελευταία επίπεδα του δικτύου εισάγονται κάποια πλήρως διασυνδεδεμένα στρώματα. Τα στρώματα αυτά λειτουργούν με το ίδιο τρόπο με τα κλασικά νευρωνικά δίκτυα και σκοπός τους είναι η κατάταξη της αρχικής εικόνας σε μια κλάση.

## 5 Εφαρμογές Μηχανικής Μάθησης σε Φυσικές Καταστροφές

Στο κεφάλαιο αυτό παρουσιάζονται και άλλα παραδείγματα, εκτός από αυτό της αναγνώρισης των παγόβουνων, όπου η εφαρμογή της μηχανικής μάθησης βοηθά στην διαχείριση φυσικών καταστροφών.

### 5.1 Βαθιά Μάθηση στην Καθημερινή Πρόβλεψη Κινδύνου Πυρκαγιάς

Σε παγκόσμιο επίπεδο, οι πυρκαγιές αποτελούν σημαντικό φυσικό κίνδυνο, ο οποίος διαταράσσει τις υπηρεσίες του φυσικού οικοσυστήματος, προκαλεί απώλειες ζώων, περιουσιών και υποδομών και συμβάλλει στις εκπομπές διοξειδίου του άνθρακα. Η κλιματική αλλαγή παίζει έναν αυξανόμενο ρόλο στον καθορισμό των καθεστώτων δασικών πυρκαγιών και αναμένεται να επιδεινώσει τις πυρκαγιές στα περισσότερα μέρη της Γης, και ιδιαίτερα στην ευρύτερη Μεσόγειο. Έτσι, η αξιολόγηση της πιθανότητας μεγάλων πυρκαγιών είναι υψίστης σημασίας για τις υπηρεσίες διαχείρισης πυρκαγιάς. Παραδοσιακά, ο κίνδυνος πυρκαγιάς προβλέπεται από κρατικά συστήματα αξιολόγησης κινδύνου πυρκαγιάς όπως το Ευρωπαϊκό “European Forest Fire Information System” (EFFIS), το οποίο βασίζεται σε έναν μετεωρολογικά προερχόμενο δείκτη καιρού πυρκαγιάς (FWI).

Η παρακάτω μελέτη [13] εξετάζει τα χαρακτηριστικά της βλάστησης που καταγράφονται από χρονοσειρές δεδομένων Παρατήρησης της Γης (EO) και τα επίπεδα πληροφοριών που συνδέονται με την ανθρώπινη δραστηριότητα, επιπροσθέτως με τα μετεωρολογικά δεδομένα, και μαθαίνει να προβλέπει τον κίνδυνο πυρκαγιάς με τρόπο που βασίζεται σε δεδομένα.

Δημιουργείται ένα datacube που περιέχει τόσο δεδομένα ιστορικών πυρκαγιών όσο και συμμεταβλητές που επηρεάζουν αυτές τις πυρκαγιές. Εκπαιδεύονται μοντέλα Deep Learning (DL) για την πρόβλεψη του κινδύνου πυρκαγιάς της επόμενης ημέρας και την αξιολόγηση αυτών των μοντέλων σε πραγματικές συνθήκες.

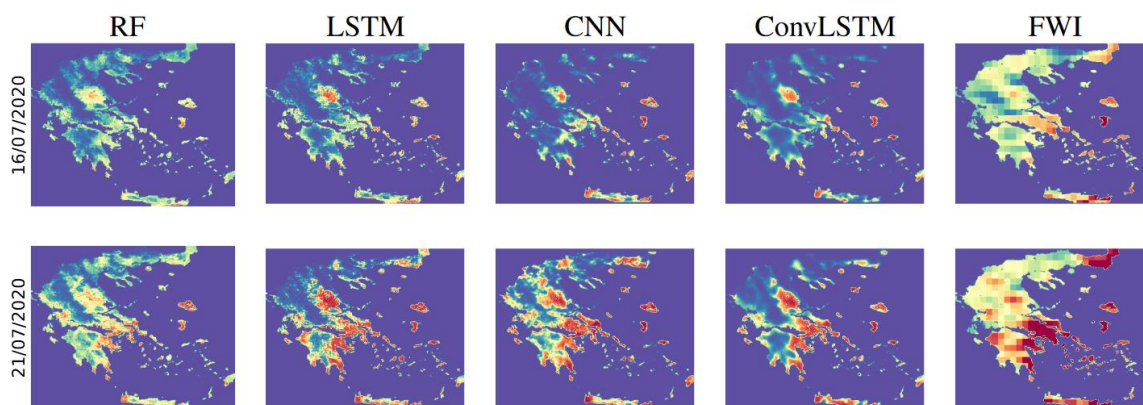
Η πρόβλεψη κινδύνου πυρκαγιάς δεν είναι μια τυπική εργασία Μηχανικής Μάθησης (ML) και θέτει τρεις μεγάλες προκλήσεις. Πρώτον, οι πυρκαγιές προκαλούνται από τις πολύπλοκες αλληλεπιδράσεις των παραγόντων που επηρεάζουν τις πυρκαγιές (καιρικές συνθήκες, χαρακτηριστικά γης και βλάστησης, ανθρώπινες δραστηριότητες). Δεύτερον, η εκδήλωση πυρκαγιών είναι εγγενώς στοχαστική. Οι ίδιες συνθήκες του περιβάλλοντος της γης μπορεί να οδηγήσουν σε μεγάλες πυρκαγιές ή μπορεί και όχι. Αυτό καθιστά τη δειγματοληψία αρνητικών παραδειγμάτων περίπλοκη, καθώς η έλλειψη εκδήλωσης πυρκαγιάς δεν σημαίνει

έλλειψη κινδύνου πυρκαγιάς. Τρίτον, οι πυρκαγιές μεγάλης έκτασης είναι σπάνια συμβάντα, τα οποία μπορεί δυνητικά να οδηγήσουν σε ένα εξαιρετικά μη ισορροπημένο σύνολο δεδομένων.

Για την συγκεκριμένη μελέτη, δημιουργείται ένα μεγάλο datacube το οποίο περιέχει μεταβλητές που συμβάλλουν στην εκδήλωση πυρκαγιάς και μεταβλητές από καμένες περιοχές για την περίοδο 2009- 2020 στην Ελλάδα, μια χώρα της Μεσογείου που είναι επιρρεπής στις πυρκαγιές. Επίσης, διαμορφώνεται η πρόβλεψη κινδύνου πυρκαγιάς ως πρόβλημα μηχανικής μάθησης. Για το σκοπό αυτό, προσδιορίζονται τα συμβάντα πυρκαγιάς που σχετίζονται με μεγάλες καμένες περιοχές, μαζί με την ημερομηνία έναρξης τους, κάτι που επιτρέπει την μοντελοποίηση της έναρξης πυρκαγιάς σε συνδυασμό με την εξέλιξη αυτής της πυρκαγιάς να είναι μεγάλη. Στη συνέχεια, αντιμετωπίζεται η πρόβλεψη πυρκαγιάς ως μια εργασία πρόβλεψης κάνοντας διάσπαση του χρόνου για την δημιουργία των σετ εκπαίδευσης, επικύρωσης και δοκιμών. Τέλος, προτείνονται και συγκρίνονται μια ποικιλία μοντέλων DL που είναι σε θέση να αποτυπώσουν χωρικό, χρονικό ή χωροχρονικό πλαίσιο. Παρουσιάζεται δηλαδή ένα επιχειρησιακό πρωτότυπο, το οποίο προβλέπει τον κίνδυνο εκδήλωσης πυρκαγιάς της επόμενης ημέρας σε εθνική κλίμακα στην Ελλάδα.

Τα δεδομένα της συγκεκριμένης μελέτης απαρτίζονται από καθημερινά μετεωρολογικά δεδομένα (θερμοκρασία, ταχύτητα του ανέμου, βροχόπτωση), δορυφορικά δεδομένα δορυφόρους της NASA (Leaf Area Index, Fraction of Photosynthetically Active Radiation), δείκτες βλάστησης, πυκνότητα πληθυσμού, τοπογραφικές μεταβλητές και ιστορικά δεδομένα καμένων περιοχών.

Τα αποτελέσματα της έρευνας αυτής, σύμφωνα με την δοκιμή στο σετ δεδομένων που περιέχει όλες τις δασικές πυρκαγιές της Ελλάδας κατά το έτος 2020 έδειξαν πως η πιο σημαντική μετρική της επίδοσης των μοντέλων που δοκιμάστηκαν είναι το AUROC, όπου το μοντέλο ConvLSTM που πρότειναν πέτυχε την καλύτερη επίδοση (0.926), το οποίο δείχνει πως η καλύτερη επίδοση μπορεί να επιτευχθεί όταν συνδυάζονται τόσο χωρικά όσο και χρονικά πλαίσια. Επιπλέον, ενώ όλα τα μοντέλα DL έχουν καλύτερη απόδοση από τα Random Forest, τα LSTM και ConvLSTM έχουν καλύτερο AUROC από το CNN, δείχνοντας ότι το χρονικό πλαίσιο είναι πιο σημαντικό από το χωρικό πλαίσιο για την πρόβλεψη κινδύνου πυρκαγιάς.



Σχήμα 5.1: Παραγώμενοι χάρτες διαφορετικών μοντέλων και το FWI για δυο ημερομηνίες με μέτριο (πάνω) και υψηλό (κάτω) κίνδυνο εκδήλωσης πυρκαγιάς. Ο υψηλός κίνδυνος απεικονίζεται με κόκκινο και ο χαμηλό με μπλε. [13]

## 5.2 Βαθιά Μάθηση για Σημασιολογική Κατάτμηση Πετρελαιοκηλίδων

Η ρύπανση από πετρελαιοκηλίδες συνδέεται στενά όχι μόνο με το οικοσύστημα των ωκεανών αλλά και με την αύξηση του θαλάσσιου εμπορίου. Δεδομένου ότι τα πρώιμα μέτρα σε τέτοιες περιπτώσεις είναι μείζονος σημασίας, έχουν παρουσιαστεί πολυάριθμοι αλγόριθμοι για την αυτόματη αναγνώριση τέτοιων σημείων ρύπανσης. Η συντριπτική πλειονότητα των σχετικών μεθόδων εκμεταλλεύεται δεδομένα που αποκτώνται από δορυφόρους που είναι εξοπλισμένοι με ραντάρ συνθετικού ανοίγματος (SAR) λόγω των πλεονεκτημάτων που εμφανίζουν. Οι εικόνες SAR όμως εισάγουν έναν τύπο αβεβαιότητας στο πρόβλημα λόγω της ύπαρξης περιοχών που μπορεί να μοιάζουν με πετρελαιοκηλίδες αλλά να μην είναι [14].

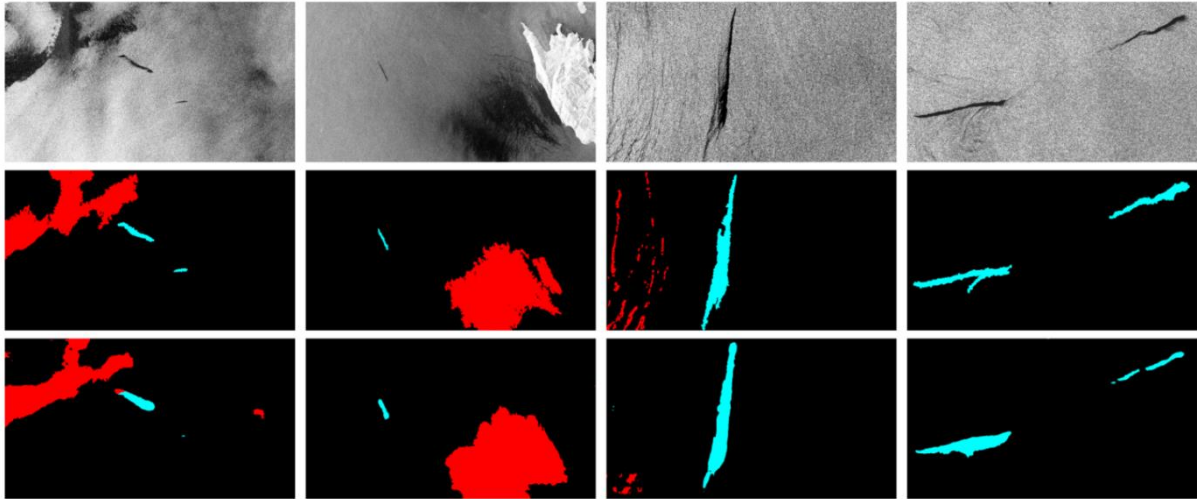
Μια τυπική διαδικασία για την ανίχνευση πετρελαιοκηλίδας μπορεί να ολοκληρωθεί σε τέσσερις διακριτές φάσεις. Το πρώτο στάδιο περιλαμβάνει την ανίχνευση των σκοτεινών σχηματισμών ενώ στη δεύτερη διαδικασία εξάγονται χαρακτηριστικά για τέτοιους σχηματισμούς. Διαδοχικά, τα εξαγόμενα χαρακτηριστικά συγκρίνονται με προκαθορισμένες τιμές και ακολουθεί ένα μοντέλο λήψης αποφάσεων για την επισημάνση κάθε σχηματισμού. Η προσέγγιση παρουσιάζει αρκετά μειονεκτήματα, κυρίως λόγω της αναγκαιότητας εξαγωγής ορισμένων χαρακτηριστικών και στη συνέχεια ομόφωνης συμφωνίας για τη φύση τους. Επιπλέον, ο περιορισμός της παροχής μιας ενιαίας ετικέτας σε καθεμία από τις εικόνες εισόδου θέτει περιορισμούς σχετικά με τα αναγνωρισμένα αντικείμενα. Οι πιο κοινές

προσεγγίσεις περιλαμβάνουν μια διαδικασία ταξινόμησης δύο τάξεων, μια κατηγορία που περιλαμβάνει μόνο πετρελαιοκηλίδες και μια δεύτερη, πιο ποικιλόμορφη κατηγορία που περιλαμβάνει οποιονδήποτε άλλο σκοτεινό σχηματισμό. αιωρούμενη βλάστηση και πάγος. Επίσης, οι πληροφορίες συμφραζομένων, οι διαδρομές πολιών κ.λπ. γύρω από τις ανιχνευμένες περιπτώσεις μπορεί να επηρεάσουν σημαντικά τον τελικό χαρακτηρισμό αυτών των "σκοτεινών σημείων". Σε αντίθεση με την παραπάνω προσέγγιση, η παρουσιαζόμενη εργασία [14] στοχεύει στην ανάπτυξη ενός βαθύ συνελκτικού νευρωνικού δικτύου (DCNN) για ανίχνευση πετρελαιοκηλίδων προκειμένου να αμβλυνθούν οι παραπάνω ελλείψεις και να μετριαστεί η αβεβαιότητα του προβλήματος. Ενώ έχουν γίνει κάποιες προσπάθειες με τη χρήση συμβατικών νευρωνικών δικτύων, αυτό απαιτεί τα χωρικά χαρακτηριστικά να υπολογιστούν στην αρχή. Επίσης, αυτοί οι αλγόριθμοι περιορίστηκαν στην ταξινόμηση εικόνων, δηλαδή στην επισήμανση ολόκληρης της εικόνας και όχι σε σημασιολογικά τμηματοποιημένες περιοχές μέσα στην εικόνα. Το πρόβλημα μπορεί να θεωρηθεί ως ένα αρκετά περίπλοκο έργο, καθώς απαιτεί εξειδικευμένη γνώση των διαφόρων φαινομένων που σχετίζονται με τα θαλάσσια περιβάλλοντα, ενώ τα οπτικά μέσα συχνά αποτυγχάνουν να δώσουν επαρκείς λύσεις. Έτσι, η τελική επιβεβαίωση της ρύπανσης περιλαμβάνει την κινητοποίηση των αρμόδιων εθνικών/περιφερειακών αρχών και τον επιτόπιο εντοπισμό. Η προτεινόμενη προσέγγιση περιλαμβάνει μια νέα εφαρμογή τέτοιων μοντέλων για την ακριβή αναγνώριση πετρελαιοκηλίδων σε εικόνες SAR χωρίς την προϋπόθεση της εξαγωγής πρόσθετων χαρακτηριστικών που οδηγούν σε ένα πλήρως αυτόματο σύστημα ανίχνευσης.

Στην αυτή την μελέτη [14] τους, ο Ορφανίδης et. al επεξεργάστηκαν πολλές εικόνες SAR ώστε να δημιουργηθεί ένα χρήσιμο dataset με επαρκές πλήθος δεδομένων. Οι ετικέτες των εικόνων βασίστηκαν σε πληροφορίες που παρείχε η EMSA και σε ανθρώπινη χειροκίνητη εισαγωγή.

Το συνελκτικό νευρωνικό δίκτυο που χρησιμοποιήθηκε είναι εμπνευσμένο από το "DeepLab" (ένα μοντέλο σημασιολογικής κατάτμησης της Google) το οποίο έχει αποδειχθεί πως είναι πολύ αποτελεσματικό σε κατατμήσεις πολλών κλάσεων και στην συνεχεία κάνει χρήση και του ResNet-101 ενός συνελκτικού νευρωνικού δικτύου που έχει 101 στρώματα [14].

Για την συγκεκριμένη μελέτη ορίστηκαν τρεις κλάσεις ταξινόμησης, πετρελαιοκηλίδα, μοιάζει-με-πετρελαιοκηλίδα, background pixels. Η επίδοση του μοντέλου μετρήθηκε με την τακτική Pixel intersection over union (IoU).



Σχήμα 5.2 : Παράδειγμα 4 εικόνων ελέγχου (από πάνω προς τα κάτω): SAR εικόνες, ground truth μάσκες και μάσκες ανίχνευσης ως αποτέλεσμα του μοντέλου

Η καινοτομία της εφαρμογής βασίζεται στο γεγονός ότι καμία παρόμοια προσέγγιση δεν αντιμετωπίζει παρόμοια το πρόβλημα, ενώ δεν απαιτείται αρχική εξαγωγή χαρακτηριστικών.

## 6 Βαθιά Μάθηση στην Διάκριση Παγόβουνων και Πλοίων σε Εικόνες SAR

Η αλληλεπίδραση του SAR με την επιφάνεια του ωκεανού είναι εξαιρετικά πολύπλοκη. Εδώ πρέπει να λάβουμε υπόψη ότι η δυναμική της επιφάνειας της θάλασσας επηρεάζεται από τον άνεμο, τα θαλάσσια ρεύματα, τα φουσκώματα, τις ατμοσφαιρικές επιδράσεις και τη θερμοκρασία της επιφάνειας της θάλασσας. Επομένως, ο άνεμος είναι συνήθως ένας σημαντικός παράγοντας όταν εξετάζουμε το backscatter του ωκεανού. Ελλείψει ανέμου, η επιφάνεια του ωκεανού θα να είναι ομαλή και επομένως λιγότερη ενέργεια του ραντάρ θα επιστρέφεται στον αισθητήρα. Η αύξηση της ταχύτητας του ανέμου θα έχει ως αποτέλεσμα την αύξηση της τραχύτητας των ωκεανών και ως εκ τούτου αυξημένο backscatter. Παραλλαγές τυχόν χιονιού ή νερού στην επιφάνεια του παγόβουνου, που μπορεί επίσης να μειώσουν ή να αλλάξουν τη διείσδυση του σήματος. Γενικά, το φόντο του ωκεανού θα είναι πιο σκούρο σε μεγαλύτερη γωνία πρόσπτωσης και πρέπει επίσης να λαμβάνεται υπόψη η πόλωση του ραντάρ. Για HH και VV SAR, καθώς η ταχύτητα του ανέμου αυξάνεται, η πιθανότητα ανίχνευσης παγόβουνων μειώνεται. Επίσης, για HH και VV SAR, καθώς αυξάνεται η γωνία πρόσπτωσης αυξάνεται η πιθανότητα ανίχνευσης παγόβουνων. Το backscatter από ένα μεταλλικό πλοίο προκύπτει κυρίως από επιφανειακή διασπορά του σήματος. Αυτό βασίζεται στην προϋπόθεση ότι η ανακλαστικότητα ενός μεταλλικού πλοίου είναι απόλυτη. Έτσι, η ανίχνευση ενός πλοίου εξαρτάται άμεσα από το μέγεθος του πλοίου, την ταχύτητα του ανέμου και την κατεύθυνση του ανέμου. Όσο αυξάνεται η ταχύτητα του ανέμου, η πιθανότητα ανίχνευσης για τα πλοία μειώνεται. Επίσης, όσο αυξάνεται η γωνία πρόσπτωσης, αυξάνεται η πιθανότητα ανίχνευσης για πλοία. Από αυτό, οι τρόποι επιλογής για την ανίχνευση πλοίων με εικόνες SAR είναι αυτοί με μεγαλύτερες γωνίες πρόσπτωσης.

Καταλαβαίνουμε λοιπόν, πως η διάκριση ανάμεσα σε παγόβουνα και πλοία σε εικόνες SAR είναι μία χρονοβόρα διαδικασία η οποία απαιτεί την εργασία ενός ανθρώπου για την σωστή ταξινόμηση των εικόνων. Αξίζει να σημειωθεί πως μια εικόνα μπορεί να περιέχει εκατοντάδες αντικείμενα προς ταξινόμηση κάθε φορά και όχι μόνο ένα. Για αυτούς τους λόγους έχουν γίνει αρκετές προσπάθειες για την εύρεση τεχνολογίας ικανής να βοηθήσει στην επίλυση αυτού του προβλήματος.



## 6.1 State-of-the-Art Μέθοδοι

Στην ερευνά τους οι Aravaralli Naveena και J.V.D.Prasad (2020) [15] προτείνουν ένα συνελκτικό νευρωνικό δίκτυο για την επίλυση του παραπάνω προβλήματος. Με εκπαίδευση για 20 εποχές, χρήση dropout για αποφυγή overfitting και σιγμοειδή συνάρτηση ενεργοποίησης. Το μοντέλο τους πετυχαίνει Accuracy μεγαλύτερης του 90% στα δεδομένα ελέγχου.

Επίσης, ο Zhan et. al (2018) [16] προτείνουν ένα συνελκτικό νευρωνικό δίκτυο με Μεταφορά Γνώσης (Transfer Learning) το οποίο πετυχαίνει Accuracy στα δεδομένα ελέγχου μεγαλύτερη του 95%. Σημαντικό ρόλο στην επιτυχία της εφαρμογής τους έπαιξε η εμπλούτιση των διαθέσιμων δεδομένων εκπαίδευσης (data augmentation) με παραλλαγή των εικόνων με τεχνικές όπως περιστροφή, smoothing κ.α.

Νικητές τους αντίστοιχου διαγωνισμού του Kaggle ήταν οι Weimin Wang και David Austin με επίτευξη στα δεδομένα ελέγχου  $\log \text{loss}=0.08227$ .

## 6.2 Η Μέθοδος μας

Παρακάτω θα εξηγηθεί αναλυτικά η προτεινόμενη μέθοδος αυτής της διπλωματικής εργασίας για την αυτόματη διάκριση των παγόβουνων και των πλοίων σε εικόνες SAR. Η εφαρμογή υλοποιήθηκε σε Jupyter Notebook στην πλατφόρμα Google Colaboratory, σε γλώσσα προγραμματισμού Python και με χρήση βιβλιοθηκών Tensorflow.

### 6.2.1 Δεδομένα και Προετοιμασία

Τα δεδομένα που χρησιμοποιήθηκαν για την υλοποίηση της εφαρμογής είναι αυτά που δόθηκαν για τον διαγωνισμό του Kaggle “Statoil/C-CORE Iceberg Classifier Challenge” το 2018 [17].

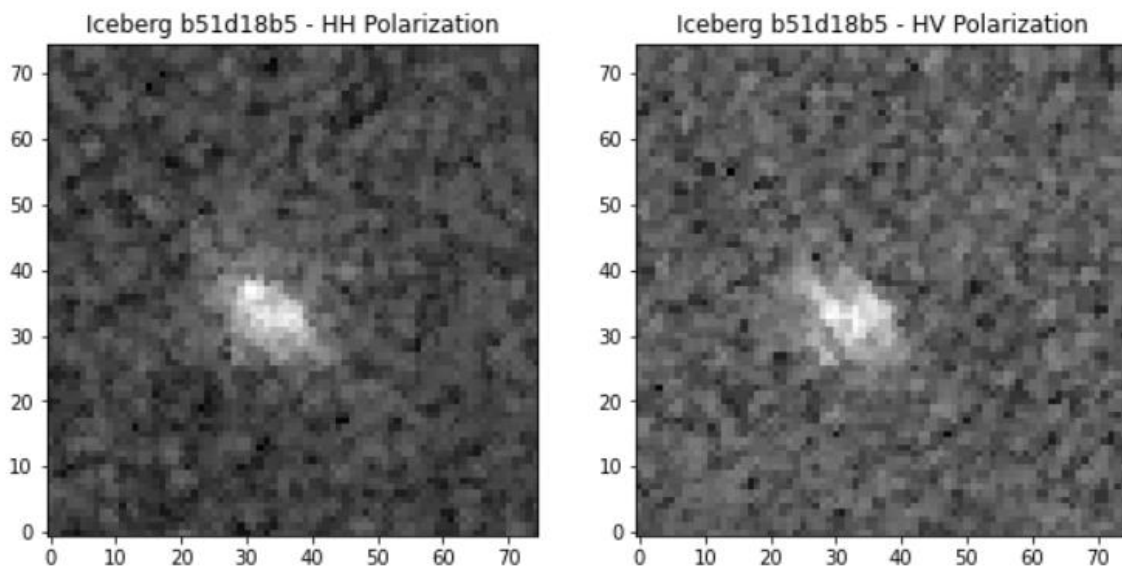
#### Dataset

Τα δεδομένα αποτελούνται από 1604 εικόνες διαστάσεων 75x75 με δύο μπάντες. Οι ετικέτες έχουν αποδοθεί από ανθρώπους ειδικούς στον χώρο και με γεωγραφικές γνώσεις της περιοχής. Τα δεδομένα προσφέρονται σε ένα αρχείο τύπου .json. Κάθε εικόνα συνοδεύεται από τα παρακάτω πεδία:

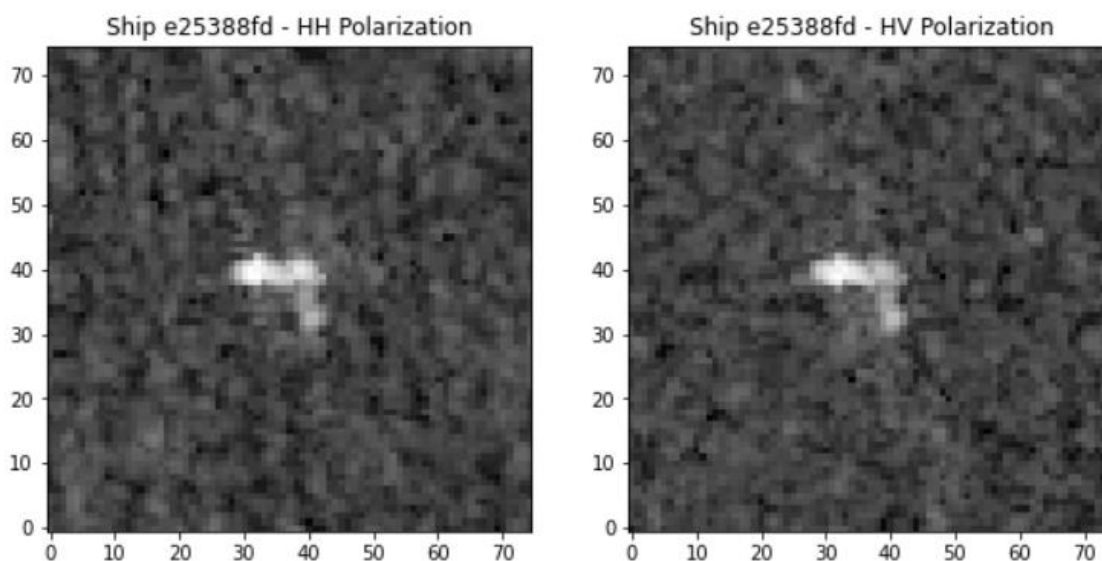
- **id:** το id της εικόνας.

- **band\_1, band\_2:** τα flattened δεδομένα της εικόνας. Κάθε εικόνα περιλαμβάνει 75x75 τιμές pixel στην λίστα, άρα 5625 στοιχεία. Αξίζει να σημειωθεί πως οι τιμές αυτές δεν είναι απλοί θετικοί ακέραιοι αριθμοί που συνηθίζουν να έχουν οι εικόνες. Οι τιμές αυτές, λόγω της μέτρησης με το ραντάρ συνθετικού διαφράγματος, είναι δεκαδικοί αριθμοί σε μονάδα μέτρησης dB και έχουν φυσική σημασία. Οι μπάντες Band 1 και Band 2 είναι σήματα τα οποία χαρακτηρίζονται από το backscatter που παράγεται από τις διαφορετικές πολώσεις του ραντάρ σε μια συγκεκριμένη γωνία πρόσπτωσης. Οι πολώσεις αντιστοιχούν σε HH (μεταδίδει και λαμβάνει οριζόντια) και HV (μεταδίδει οριζόντια και λαμβάνει κάθετα).
- **inc\_angle:** η γωνία πρόσπτωσης με την οποία τραβήχτηκε η εικόνα. Να σημειωθεί εδώ πως σε αυτό το πεδίο υπάρχουν τιμές που λείπουν όπου τα αντίστοιχα πεδία έχουν τιμή “na” (not available).
- **is\_iceberg:** η μεταβλητή στόχος. Έχει τιμή 1 εάν είναι παγόβουνο και τιμή 0 εάν είναι πλοίο.

Σε μερικές εικόνες η ταξινόμηση είναι εύκολη και σε άλλες η διάκριση ανάμεσα σε παγόβουνο και πλοίο δεν είναι τόσο ευδιάκριτη. Για καλύτερη κατανόηση των δεδομένων, παρακάτω απεικονίζονται εικόνες παγόβουνων και πλοίων σε HH και HV πόλωσεις.



Εικόνα 6.1 : Παγόβουνο σε HH (αριστερά) και HV (δεξιά) πόλωση.



Εικόνα 6.2: Πλοίο σε HH (αριστερά) και HV (δεξιά) πόλωση.

## Προετοιμασία Δεδομένων

Για να μπορέσουμε να εισάγουμε τα δεδομένα μας στο τεχνητό νευρωνικό δίκτυο, απαραίτητη προϋπόθεση είναι τα δεδομένα να μετασχηματιστούν στην κατάλληλη μορφή. Οπότε δημιουργήθηκαν οι κατάλληλοι φάκελοι και τα δεδομένα αποθηκεύτηκαν σε tensors. Ένας *Tensor* είναι ένα δοχείο για αριθμητικά δεδομένα. Είναι ο τρόπος με τον οποίο αποθηκεύουμε τις πληροφορίες που θα χρησιμοποιήσουμε στο σύστημά μας.

Στην συνέχεια, οι τιμές του πεδίου `inc_angle` που ήταν ίσες με “na” μετατράπηκαν σε 0. Αξίζει να σημειωθεί πως για την εκπαίδευση του μοντέλου δεν χρησιμοποιήθηκαν τα δεδομένα της γωνίας πρόσπτωσης. Θα ήταν ενδιαφέρον να μελετηθεί η συνεισφορά τους σε κάποια μελλοντική μελέτη. Όπως αναφέρθηκε, κάθε εικόνα έχει δύο μπάντες. Επειδή τα δεδομένα `band_1` και `band_2` ήταν ισοπεδωμένα (από εδώ και πέρα θα αναφέρεται με την αγγλική του “flattened”), μετασχηματίστηκαν σε διαστάσεις 75x75. Έτσι σχηματίστηκε ένας tensor που περιέχει τα δεδομένα όλων των εικόνων με διαστάσεις (1604,75,75,2). Όπου 1604 είναι το πλήθος των εικόνων, 75x75 είναι οι διαστάσεις των εικόνων και 2 είναι οι μπάντες.

## Feature Scaling

Τα τεχνητά νευρωνικά δίκτυα δουλεύουν καλύτερα και με γρηγορότερη απόδοση όταν οι τιμές που τους εισάγονται είναι μικρές. Επειδή τα δεδομένα από το backscatter είναι

δεκαδικοί αριθμοί χωρίς κάποιον άλλο περιορισμό, δηλαδή μπορεί να παίρνουν μεγάλες τιμές και θετικές και αρνητικές, εμείς τις μετασχηματίσαμε ώστε να είναι στο πεδίο  $[-1,1]$  με την μέθοδο *Mean Normalization* της οποίας ο τύπος φαίνεται παρακάτω:

$$x' = \frac{x - \mu}{\max(x) - \min(x)}$$

Όπου  $\mu$  είναι η μέση τιμή. Δηλαδή, υπολογίζει και αφαιρεί τον μέσο όρο για κάθε χαρακτηριστικό.

## Διαχωρισμός Δεδομένων σε Train – Validation – Test

Τα δεδομένα χωρίστηκαν σε δεδομένα εκπαίδευσης, επικύρωσης και ελέγχου όπως φαίνεται παρακάτω:

- Train = 1122 εικόνες
- Validation = 337 εικόνες
- Test = 145 εικόνες

## Data Augmentation

Επειδή η ύπαρξη περισσότερων δεδομένων εκπαίδευσης ήταν αναγκαία για την καλύτερη γενίκευση του μοντέλου μας και την αποφυγή *overfitting*, η αύξηση των δεδομένων μας επιτεύχθηκε με μία ποικιλία μετασχηματισμών στις εικόνες όπως περιστροφή και καθρέφτισμα. Όποτε τα δεδομένα εκπαίδευσής μας τώρα έχουν ως εξής:

- Train = 3366 εικόνες

Η αύξηση των δεδομένων εκπαίδευσης επιλέχθηκε εσκεμμένα να γίνει μετά τον διαχωρισμό των δεδομένων σε εκπαίδευσης – επικύρωσης – ελέγχου, και μόνο στα δεδομένα εκπαίδευσης, ώστε να αποφύγουμε τυχόν ανεπιθύμητη διαρροή πληροφοριών στο νευρωνικό δίκτυο.

### 6.2.2 Το Μοντέλο

Για την κατηγοριοποίηση των εικόνων του προβλήματος επιλέχθηκε η χρήση ενός συνελκτικού νευρωνικού δικτύου το οποίο όπως έχουμε αναφέρει προηγουμένως ανταποκρίνεται πολύ καλά σε τέτοιου είδους προβλήματα.

## Συνελικτικό Νευρωνικό Δίκτυο

Το δίκτυο αποτελείται από:

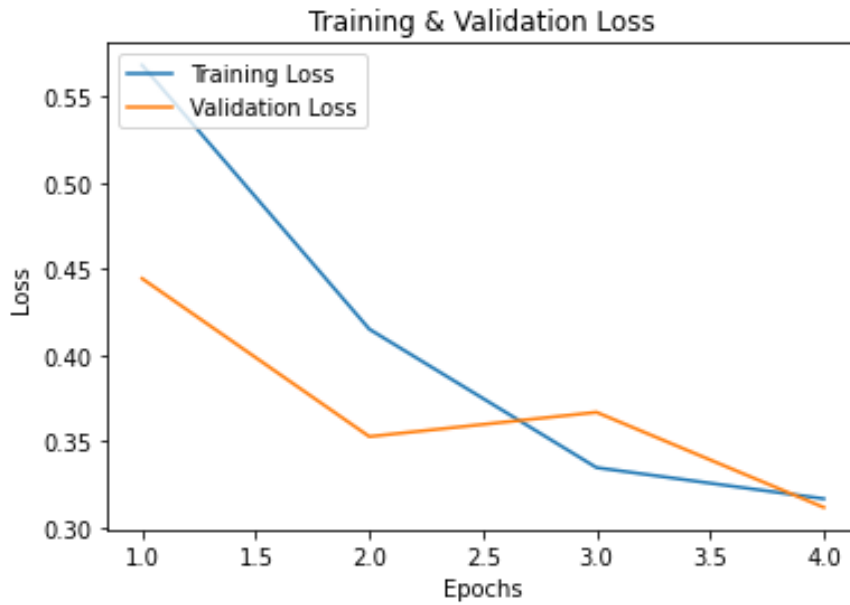
- Τέσσερα συνελικτικά στρώματα ακολουθούμενο το κάθε ένα από ένα στρώμα Max Pooling και ένα στρώμα Dropout 20% για την αποφυγή overfitting.
- Ένα στρώμα Flatten για την προετοιμασία των δεδομένων για εισαγωγή στα πλήρως διασυνδεδεμένα στρώματα.
- Δύο στρώματα Dense με συνάρτηση ενεργοποίησης ReLU και Dropout 20%.
- Ένα στρώμα Dense με σιγμοειδή συνάρτηση ενεργοποίησης.

Ως Optimization Algorithm επιλέχθηκε ο αλγόριθμος “Adam” και για συνάρτηση λάθους (Loss Function) η Binary Crossentropy. Ως μετρική ελέγχου επιλέχθηκε το Accuracy. Οι όροι αναφέρονται με την αγγλική τους ονομασία για καλύτερη κατανόηση.

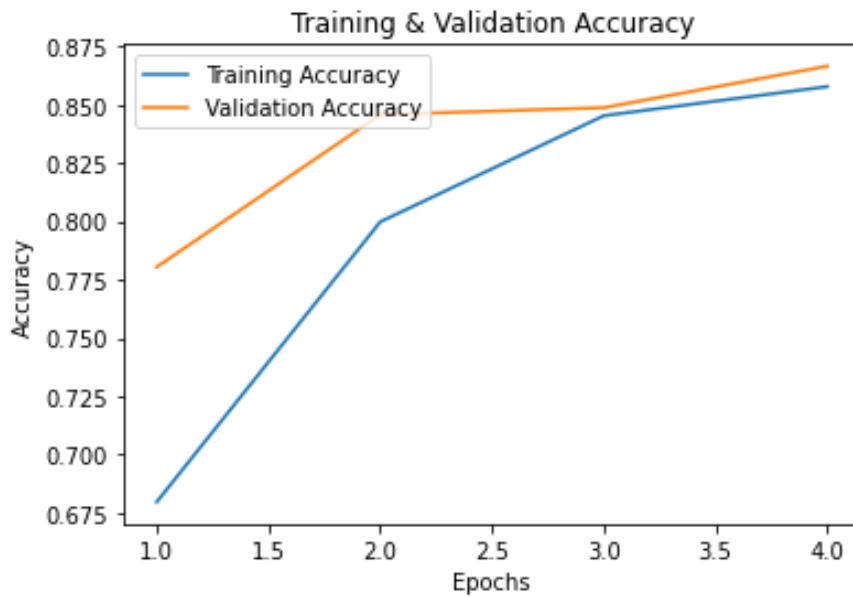
### Εκπαίδευση

Το συνελικτικό νευρωνικό δίκτυο εκπαιδεύτηκε για 4 εποχές με batch size=24. Ο χρόνος που διήρκησε η εκπαίδευση του μοντέλου είναι 30 sec.

Για να μπορέσουμε να αποφασίσουμε πόσες εποχές εκπαίδευσης είναι κατάλληλες για την επίλυση του προβλήματος μας, ικανές να δημιουργήσουν ένα γενικευμένο μοντέλο και να αποφύγουμε την υπερπροσαρμογή στα δεδομένα εκπαίδευσης, παρατηρήσαμε πως εξελίσσονται το accuracy και το loss στα δεδομένα εκπαίδευσης και επικύρωσης κατά την διάρκεια των εποχών εκπαίδευσης. Αποφασίστηκε να σταματήσει η εκπαίδευση στις 4 εποχές καθώς σε περισσότερες εποχές παρατηρήθηκε ότι το training loss συνέχισε να πέφτει αλλά το validation loss άρχισε να ανεβαίνει, κατάσταση που σηματοδοτεί overfitting.



Σχήμα 6.1 : Loss του Μοντέλου για εκπαίδευση για 4 εποχές.



Σχήμα 6.2 : Accuracy του Μοντέλου για εκπαίδευση για 4 εποχές.

Στο κεφάλαιο 6 παρουσιάστηκε η μεθοδολογία για την επίλυση του προβλήματος της διάκρισης των παγόβουνων και των πλοίων από εικόνες SAR. Παρουσιάστηκαν τα δεδομένα και οι απαραίτητες διαδικασίες επεξεργασίας τους για την εισαγωγή τους στο νευρωνικό δίκτυο και εξηγήθηκε το μοντέλο που χρησιμοποιήθηκε. Στο κεφάλαιο 7 παρουσιάζονται τα αποτελέσματα του μοντέλου μας και στην συνέχεια στο κεφάλαιο 8 παρουσιάζονται τα συμπεράσματα.

## 7 Αποτελέσματα και Συζήτηση

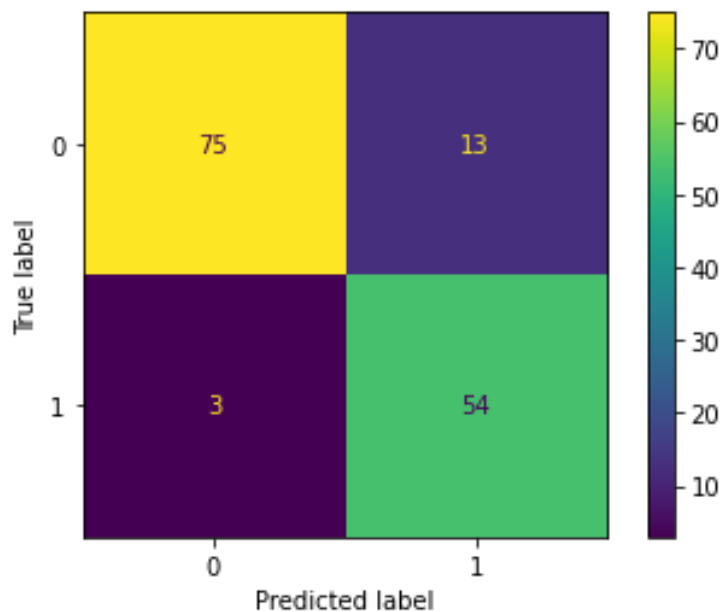
Μετά την ολοκλήρωση της εκπαίδευσης του μοντέλου, το επόμενο βήμα είναι να δοκιμαστεί η επίδοση του σε ένα σύνολο δεδομένων το οποίο να του είναι τελείως άγνωστο, σε δεδομένα που «βλέπει» για πρώτη φορά. Έτσι θα μπορέσουμε να καταλάβουμε αν έχουμε πετύχει ικανοποιητική γενίκευση του ώστε μετά να μπορεί να χρησιμοποιηθεί αποτελεσματικά σε μελλοντικές προσπάθειες ανίχνευσης παγόβουνων. Το μοντέλο ελέγχθηκε στα δεδομένα ελέγχου που είχαμε δημιουργήσει στα προηγούμενα βήματα.

### Μετρικές Ελέγχου

Ως μετρικές ελέγχου χρησιμοποιήθηκαν η Ακρίβεια, ο Πίνακας Σύγκυσης, η Ευστοχία και η Ανάκληση.

Το μοντέλο μας πέτυχε Ακρίβεια=88% στα δεδομένα ελέγχου.

Για να μπορέσουμε να κατανοήσουμε ακόμα καλύτερα την επίδοση του μοντέλου μας, είναι πολύ χρήσιμο να δημιουργήσουμε έναν πίνακα σύγκυσης των αποτελεσμάτων του. Για τα δεδομένα ελέγχου έχουμε:



Πίνακας 7.1 : Confusion Matrix για τα Δεδομένα Ελέγχου

Το πρώτο συμπέρασμα που βγάζουμε από τον πίνακα σύγκυσης είναι σχετικά με την ισορροπία των δεδομένων ελέγχου μας, βλέπουμε πως στις θέσεις TP και TN το σύνολο των

δειγμάτων είναι περίπου ίδιο, δηλαδή υπάρχει περίπου ίδιος αριθμός δειγμάτων για κάθε κλάση. Αυτό σηματοδοτεί πως έχουμε ένα ισορροπημένο σετ δεδομένων ελέγχου, άρα και πως το μοντέλο μας είναι εξίσου καλό στο να αναγνωρίζει επιτυχώς και τις δύο κλάσεις.

Στην συνέχεια του ελέγχου, επιτεύχθηκε Ευστοχία=80% και Ανάκληση=94%. Όταν σε μια εφαρμογή υπάρχει υψηλό κόστος συσχετιζόμενο με την σωστή πρόβλεψη των FN, τότε μας ενδιαφέρει το μοντέλο να πετυχαίνει υψηλό ποσοστό ανάκλησης. Όσο μεγαλύτερη είναι η ανάκληση τόσο μικρότερος είναι ο αριθμός των θετικών δειγμάτων που έχουν κατηγοριοποιηθεί λάθος. Και συγκεκριμένα για την δική μας εφαρμογή έχουμε πως FN=3, δηλαδή 3 παγόβουνα κατηγοριοποιήθηκαν λανθασμένα ως πλοία.

Για καλύτερη κατανόηση θα αναφερθεί ένα παράδειγμα. Στην περίπτωση των FP, όπου είναι ίσα με 13 σημαίνει πως 13 πλοία ταξινομήθηκαν λανθασμένα ως παγόβουνα. Σε ένα πραγματικό τέτοιο περιστατικό, ας υποθέσουμε πως τελευταία στιγμή ο καπετάνιος ενός πλοίου λαμβάνει ενημέρωση ότι στην πορεία του θα συναντήσει ένα παγόβουνο, οπότε έκανε όλες τις απαραίτητες μανούβρες για να το αποφύγει. Όταν φτάνει στην τοποθεσία του υποτιθέμενου παγόβουνου διαπιστώνει πως η ενημέρωση που είχε λάβει ήταν λάθος γιατί τελικά αυτό που συνάντησε ήταν πλοίο. Οπότε πρακτικά δεν βρέθηκε αντιμέτωπος με κάποιο μεγάλο κίνδυνο, ήταν ένας λάθος συναγερμός και το πλοίο παραμένει ασφαλές. Στην άλλη περίπτωση όμως, εκεί όπου έχουμε FN ίσα με 3 σημαίνει πως 3 παγόβουνα ταξινομήθηκαν λανθασμένα ως πλοία. Στο αντίστοιχο παράδειγμα με τον καπετάνιο, αυτό αποτελεί πολύ μεγάλο κίνδυνο καθώς λαμβάνει ενημέρωση πως στην πορεία του υπάρχει ένα πλοίο, γεγονός που δεν θέτει το πλήρωμα σε κατάσταση συναγερμού και έτσι συνεχίζουν την πορεία τους με πολύ μεγάλο κίνδυνο όταν φτάσουν κοντά στο παγόβουνο και διαπιστώσουν πως δεν πρόκειται για πλοίο, να είναι πολύ αργά και να κινδυνεύσουν να συγκρουστούν. Οπότε το γεγονός, πως ναι μεν ο αλγόριθμος μας ταξινομεί κάποια δείγματα λάθος αλλά ο μεγαλύτερος αριθμός αυτών είναι τα FP μας δείχνει πως το μοντέλο μας μπορεί να ανταποκριθεί καλά στις ανάγκες του πραγματικού κόσμου.

Σε σύγκριση με τις κορυφαίες μεθόδους που έχουν δοκιμαστεί το μοντέλο που προτείνεται σε αυτή την εργασία φτάνει αρκετά κοντά στα αποτελέσματα εκείνων (περίπου 7% χαμηλότερο accuracy). Για την περαιτέρω βελτίωση της μεθόδου θα πρέπει να εξεταστεί η χρήση των γωνιών πρόσπτωσης ως δεδομένων εκπαίδευσης, η καταλληλότητα της αρχιτεκτονικής του δικτύου και το πλήθος των δεδομένων. Ενδιαφέρον θα ήταν επίσης, να



δοκιμαστεί η μεταφορά γνώσης από κάποιο άλλο προ-εκπαιδευμένο νευρωνικό δίκτυο για την επίτευξη μεγαλύτερης ακρίβειας και γενίκευσης του μοντέλου.

Το κεφάλαιο αυτό παρέχει ποιοτικές αποδείξεις πως το μοντέλο μας ανταποκρίνεται πολύ καλά στην αντιμετώπιση του προβλήματος που είχε τεθεί. Απαντάει δηλαδή επιτυχημένα και με αυτοπεποίθηση στο ερώτημα «Αυτό που απεικονίζεται, είναι παγόβουνο ή πλοίο;».

## 8 Συμπεράσματα

Σε αυτή τη διπλωματική εργασία παρουσιάζονται τρόποι με τους οποίους η τεχνητή νοημοσύνη και συγκεκριμένα η βαθιά μάθηση και τα τεχνητα νευρωνικά δίκτυα μπορούν να βοηθήσουν στην προστασία του φυσικού περιβάλλοντος, στην δημιουργία μοντέλων ικανών να προβλέψουν φυσικές καταστροφές και να βοηθήσουν στην αποφυγή ή την γρήγορη αντιμετώπισή τους.

Συγκεκριμένα παρουσιάστηκαν εφαρμογές βαθιάς μάθησης για την καθημερινή πρόβλεψη κινδύνου πυρκαγιάς, την σημασιολογική κατάτμηση πετρελαιοκηλίδων και την διάκριση παγόβουνων και πλοίων από δορυφορικές εικόνες.

Η διάκριση παγόβουνων και πλοίων από εικόνες SAR αποτελεί μια χρονοβόρα διαδικασία η οποία χρειάζεται κάποιον άνθρωπο με ειδικές γεωγραφικές γνώσεις για την ταξινόμηση των εικόνων. Στην εργασία αυτή προτάθηκε ένα συνελκτικό νευρωνικό δίκτυο για την αυτόματη ταξινόμηση των εικόνων, ένα μοντέλο δηλαδή που απαντάει με επιτυχία στο ερώτημα «Αυτό που απεικονίζεται, είναι παγόβουνο ή πλοίο;». Τα δεδομένα που χρησιμοποιήθηκαν για την υλοποίηση της εφαρμογής είναι αυτά που δόθηκαν για τον διαγωνισμό του Kaggle “Statoil/C-CORE Iceberg Classifier Challenge”.

Η ακρίβεια (accuracy) του μοντέλου ανέρχεται στο 88% και η ανάκληση (recall) στο 94%. Αυτό σημαίνει πως είναι ένα μοντέλο το οποίο μπορεί να ανταπεξέλθει επιτυχώς στις απαιτήσεις του πραγματικού κόσμου και να προσφέρει πολύτιμη βοήθεια σε ναυτιλομένους και όχι μόνο, να συμβάλλει στην ασφάλεια ανθρώπινων ζώων και στην προστασία των ωκεανών.

Τέλος, η χρήση τεχνητών νευρωνικών δικτύων σε ζητήματα προστασίας του φυσικού περιβάλλοντος ανοίγει νέους ορίζοντες στις διαθέσιμες λύσεις τέτοιων προβλημάτων που επηρεάζουν το περιβάλλον γύρω μας και κατ' επέκταση την ποιότητα της ζωής.

## Βιβλιογραφία

- [1] P. Wadhams, "Britannica," Encyclopedia Britannica, Inc, [Online]. Available: <https://www.britannica.com/science/iceberg>. [Accessed 2022 06 2022].
- [2] Canadian Ice Service, "Government of Canada," 2022. [Online]. Available: <https://ice-glaces.ec.gc.ca/cgi-bin/getprod.pl?prodid=WIS65&wrap=1&lang=en>. [Accessed 28 06 2022].
- [3] European Space Agency, "European Space Agency," [Online]. Available: <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1>. [Accessed 2022].
- [4] Κ. Γ. Περάκης, Ι. Ν. Φαρασλής and Α. Κ. Μουσιάδης, Η Τηλεπισκόπηση σε 13 Ενότητες, Αθήνα: Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών, 2015.
- [5] K. Herndon, F. Meyer, A. Flores, E. Cherrington and L. Kucera, "Earth Data Nasa," NASA, [Online]. Available: <https://www.earthdata.nasa.gov/learn/backgrounders/what-is-sar>. [Accessed 28 06 2022].
- [6] C. Howell, Iceberg and ship detection and classification in single, dual and quad polarized synthetic aperture radar, Canada: Memorial University of Newfoundland, 2008.
- [7] NASA, The SAR Handbook - Comprehensive Methodologies for Forest Monitoring and Biomass Estimation, 2019.
- [8] Ι. Βλαχαβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κοκκορας and Η. Σακελλαρίου, Τεχνητή Νοημοσύνη, Εκδόσεις Πανεπιστημίου Μακεδονίας, Γ' Έκδοση.
- [9] Κ. Διαμαντάρας and Δ. Μπότσης, Μηχανική Μάθηση, Εκδόσεις Κλειδάριθμος, 2019.
- [10] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of research and development*, 1959.

- [11] S. Haykin, *Νευρωνικά Δίκτυα και Μηχανική Μάθηση*, Εκδόσεις Παπασωτηρίου, 2010.
- [12] F. Chollet, *Deep Learning with Python*, Manning, 2021.
- [13] I. Prapas, S. Kondylatos, I. Papoutsis, G. Camps-Valls, M. Ronco, M. Á. F. Torres, M. P. Guillem and N. Carvalhais, "Deep Learning Methods for Daily Wildfire Danger Forecasting," 2021.
- [14] G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis and I. Kompatsiaris, "A DEEP NEURAL NETWORK FOR OIL SPILL SEMANTIC SEGMENTATION IN SAR IMAGES," *IEEE*, 2018.
- [15] A. Naveena and J. V. Prasad, "Iceberg Detection in Satellite Images using Deep Learning Techniques," *International Journal of Recent Technology and Engineering (IJRTE)*, 2020.
- [16] L. Zhang, C. Zhan, Z. Zhong, S. Didi-Ooi, L. Youzuo, Y. Zhang, S. Huang and C. Wang, "Deep Learning Approach in Automatic Iceberg - Ship Detection with SARRemote Sensing Data," in *International Geophysical Conference*, Beijing, China, 2018.
- [17] Statoil/C-CORE, "Kaggle," [Online]. Available: <https://www.kaggle.com/competitions/statoil-iceberg-classifier-challenge/data>. [Accessed 28 06 2022].