



Πανεπιστήμιο Μακεδονίας
Μεταπτυχιακό Εφαρμοσμένης Πληροφορικής

Διπλωματική Εργασία

«Μελέτη αρχιτεκτονικών σχεδίασης νευρωνικών δικτύων για
ταξινόμηση φωτογραφιών πλαγκτού»

Αναστάσιος Αγακίδης

Θεσσαλονίκη, Φεβρουάριος 2022

ΜΕΛΕΤΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΣΧΕΔΙΑΣΗΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΓΙΑ
ΤΑΞΙΝΟΜΗΣΗ ΦΩΤΟΓΡΑΦΙΩΝ ΠΛΑΓΚΤΟΥ

Αναστάσιος Αγακίδης
Πτυχίο Φυσικής, ΑΠΘ, 2019

Διπλωματική Εργασία

υποβαλλόμενη για τη μερική εκπλήρωση των απαιτήσεων του

ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΤΙΤΛΟΥ ΣΠΟΥΔΩΝ ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ
ΠΛΗΡΟΦΟΡΙΚΗ

Επιβλέπων Καθηγητής
Ιωάννης Ρεφανίδης

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 23/02/2022

Ιωάννης Ρεφανίδης

Νικόλαος Σαμαράς

Δημήτριος Χρήστου-
Βαρσακέλης

.....

.....

.....

Αναστάσιος Αγακίδης

Περίληψη

Ο στόχος αυτής της διπλωματικής εργασίας είναι η ψηφιακή επιμέλεια της δημιουργίας ενός συνόλου δεδομένων πλαγκτού που συναντώνται σε ελληνικά ύδατα και η κατασκευή ενός μοντέλου ταξινόμησής τους με την χρήση βαθιών νευρωνικών δικτύων.

Τα δεδομένα των φωτογραφιών έχουν ληφθεί από οπτικό μικροσκόπιο, που σημαίνει ότι η ανάλυση των φωτογραφιών ποικίλει. Συνεπώς, για την ορθή λειτουργία του προγράμματος ταξινόμησης χρειάστηκε εκτενής προ-επεξεργασία των δεδομένων για την ομοιογενοποίησή και την κατάλληλη διαλογή τους, ώστε να μπορούν να είναι χρηστικά. Για την δημιουργία του μοντέλου ταξινόμησης, μελετήθηκαν και διερευνήθηκαν διαφορετικές προσεγγίσεις και αρχιτεκτονικές νευρωνικών δικτύων με στόχο την μέγιστη δυνατή ακρίβεια προβλέψεων.

Λέξεις Κλειδιά:

Ταξινόμηση, πλαγκτόν, βαθιά νευρωνικά δίκτυα, συνελκτικά νευρωνικά δίκτυα, επεξεργασία εικόνας

Abstract

The purpose of this thesis is the curation of a plankton image dataset, that are prominent in Greek waters and the creation of an automated classification model, using deep neural networks.

The image data were obtained using optical microscopy, thus, the quality between images varies. Therefore, emerged a need for extensive pre-processing of the images for them to be usable. Different approaches and architectures of neural networks were studied and implemented to achieve the best possible accuracy of predictions.

Keywords:

Classification, plankton, deep neural networks, convolutional neural networks, image processing

Ευχαριστίες

Με την ολοκλήρωση της μεταπτυχιακής διπλωματικής μου εργασίας, θα ήθελα να εκφράσω τις ευχαριστίες μου σε όλους όσους συνέβαλλαν στην εκπόνησή της.

Ευχαριστώ θερμά τον Επιβλέποντα Καθηγητή μου, κ. Ιωάννη Ρεφανίδη, για την εμπιστοσύνη που μου έδειξε εξ' αρχής, αναθέτοντάς μου το συγκεκριμένο θέμα, την επιστημονική του καθοδήγηση και τις υποδείξεις του.

Θα ήθελα να εκφράσω την βαθύτατη ευγνωμοσύνη μου στην Καθηγήτρια κα. Τριανταφύλλου Μαρία, για την παραχώρηση της άδειας χρήσης των φωτογραφιών πλαγκτού που χρησιμοποιήθηκαν για την εκπαίδευση των νευρωνικών δικτύων. Δίχως τη δική της συνδρομή δεν θα ήταν δυνατή η εκπόνηση της εργασίας αυτής.

Τέλος, θα ήθελα εκφράσω την ευγνωμοσύνη μου στην οικογένεια και τους φίλους μου, για τη συμπαράσταση και την κατανόησή τους, καθ' όλη τη διάρκεια των σπουδών μου.

Περιεχόμενα

1	Εισαγωγή	1
1.1	Πρόβλημα – Σημαντικότητα του θέματος	1
1.2	Σκοπός – Στόχοι	1
1.3	Διάρθρωση της μελέτης	1
2	Πλαγκτόν	3
2.1	Η αξία των παράκτιων υδάτων	3
2.2	Φυτοπλαγκτόν	4
2.3	Ομάδες Φυτοπλαγκτού	6
2.3.1	Κυανοβακτήρια	7
2.3.2	Διάτομα	7
2.3.3	Δινομαστιγωτά	9
2.4	Ζωοπλαγκτόν	10
2.5	Ομάδες ζωοπλαγκτόν	11
2.5.1	Τρηματοφόρα	12
2.5.2	Ακτινόζωα	13
3	Μηχανική μάθηση και μέθοδοι επεξεργασίας εικόνας	15
3.1	Τεχνητή Νοημοσύνη	15
3.2	Μηχανική Μάθηση	15
3.3	Τεχνητά Νευρωνικά Δίκτυα	16
3.3.1	Ταξινομητές	17
3.3.2	Συνάρτηση Ενεργοποίησης	17
3.3.3	Δομή Νευρωνικών Δικτύων	18
3.3.4	Εκπαίδευση	18
3.3.5	Συνάρτηση softmax	21
3.4	Συνελκτικά Νευρωνικά Δίκτυα (CNN)	21
3.4.1	Επίπεδο Συνέλιξης (Convolutional Layer)	22
3.4.2	Επίπεδο Συγκέντρωσης	24
3.4.3	Υπολειμματικά μπλοκ (Residual blocks)	25
3.4.4	Ανεστραμμένα υπολειμματικά μπλοκ (Inverted residual blocks)	25
3.5	Μέθοδοι επεξεργασίας εικόνας	26
3.5.1	Τελεστής Sobel-Feldman	26

3.5.2 Γκαουσιανό φίλτρο θόλωσης (Gaussian blur filter)	26
3.5.3 Αλγόριθμος εύρεσης άκρων του Canny	27
4 Μεθοδολογία	29
4.1 Σύνολο των Δεδομένων	29
4.1.1 Καθαρισμός των εικόνων	30
4.1.2 Υλοποίηση αλγορίθμου εύρεσης άκρων του Canny	31
4.1.3 Επαύξηση των δεδομένων για εμπλουτισμό του συνόλου	35
4.2 Σύγκριση αρχιτεκτονικών σχεδίασης των νευρωνικών δικτύων	36
4.2.1 Μοντέλο πρόβλεψης χωρίς τη χρήση συνελκτικών επιπέδων	37
4.2.2 Μοντέλα πρόβλεψης με 3 ή 4 επίπεδα συνέλιξης	38
4.2.3 Μοντέλο πρόβλεψης με τη χρήση του EfficientNet	40
4.2.4 Μοντέλο πρόβλεψης με τη χρήση VGG-16	42
5 Σύνοψη και Συμπεράσματα	45
5.1 Όρια και περιορισμοί της έρευνας	47
5.2 Μελλοντικές Επεκτάσεις	47

Κατάλογος Εικόνων

Εικόνα 2-1: Απλοποιημένη απεικόνιση υδάτινου τροφικού πλέγματος.....	5
Εικόνα 3-1: Βασική αρχή λειτουργίας κλασσικού προγραμματισμού και μηχανικής μάθησης.....	16
Εικόνα 3-2: Απλοϊκό παράδειγμα δομής νευρωνικού δικτύου.....	18
Εικόνα 3-3: Σχηματική αναπαράσταση της σύνδεσης μεταξύ των εισόδων και της εξόδου ενός συνελκτικού επιπέδου.....	23
Εικόνα 3-4: Παράδειγμα λειτουργίας του επιπέδου συγκέντρωσης μεγίστου.....	24
Εικόνα 4-1: Κατανομή των ακατέργαστων φωτογραφιών.....	35
Εικόνα 4-2: Κατανομή φωτογραφιών του τελικού συνόλου δεδομένων.....	36
Εικόνα 4-3: Αρχιτεκτονική απλού νευρωνικού δικτύου ταξινόμησης.....	37
Εικόνα 4-4: Αρχιτεκτονική νευρωνικών δικτύων με χρήση συνελκτικών επιπέδων. Αριστερά με τρία, ενώ δεξιά με τέσσερα επίπεδα.....	39
Εικόνα 4-5: Μετρικές μοντέλου με τρία επίπεδα συνέλιξης.....	40
Εικόνα 4-6: Μετρικές μοντέλου με τέσσερα επίπεδα συνέλιξης.....	40
Εικόνα 4-7: Δημιουργία του μοντέλου που χρησιμοποιεί EfficientNet.....	41
Εικόνα 4-8: Αρχιτεκτονική EfficientNet.....	41
Εικόνα 4-9: Μετρικές μοντέλου με χρήση EfficientNet.....	42
Εικόνα 4-10: Αρχιτεκτονική VGG-16.....	43
Εικόνα 4-11: Μετρικές μοντέλου με χρήση VGG-16.....	44

1 Εισαγωγή

1.1 Πρόβλημα – Σημαντικότητα του θέματος

Η χρήση των νευρωνικών δικτύων και της μηχανικής μάθησης για τη διευκόλυνση, ή και τη λύση, πολλών προβλημάτων της καθημερινότητας έχει ήδη ευρεία εφαρμογή. Όμως σε πολλούς τομείς των επιστημών, παρατηρείται έλλειψη στη χρήση τους. Συνεπώς, η επέκταση της χρήσης τεχνητής νοημοσύνης, ειδικά στις θετικές επιστήμες, θα μπορούσε να λύσει πληθώρα προβλημάτων ή να μειώσει τον χρόνο που απαιτούν ορισμένες διεργασίες ώστε να ολοκληρωθούν.

Το πλαγκτόν αποτελεί πολύ σημαντικό κομμάτι του παγκόσμιου οικοσυστήματος. Παράλληλα, συνιστά σημαντικό βιολογικό δείκτη για την ποιότητα των υδάτων. Συνεπώς, η υλοποίηση ενός συστήματος αυτόματης ταξινόμησής του, δύναται να είναι χρήσιμη για την εξαγωγή συμπερασμάτων για την ποιότητας των υδάτων στην περιοχή όπου έγινε η δειγματοληψία.

1.2 Σκοπός – Στόχοι

Η εργασία αυτή περιλαμβάνει την ανάλυση και επεξεργασία ενός συνόλου ακατέργαστων φωτογραφιών πλαγκτού με σκοπό τη δημιουργία ενός συνόλου δεδομένων (dataset). Λόγω της ανομοιομορφίας και σχετικά μικρής κλίμακας του συνόλου των δεδομένων, χρειάστηκε εκτενής επεξεργασία των φωτογραφιών ώστε να είναι σε μορφή που μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός νευρωνικού δικτύου αυτόματης ταξινόμησής τους.

Για την ταξινόμηση των φωτογραφιών του πλαγκτού ακολουθήθηκαν πέντε διαφορετικές αρχιτεκτονικές κατασκευής νευρωνικών δικτύων. Οι αρχιτεκτονικές συγκρίθηκαν μεταξύ τους τόσο ως προς την ακρίβεια των προβλέψεων τους, αλλά και ως προς το χρόνο εκτέλεσής τους. Ο τελικός σκοπός είναι η δημιουργία ενός ικανοποιητικού ταξινομητή για τις φωτογραφίες πλαγκτού.

1.3 Διάρθρωση της μελέτης

Στο δεύτερο Κεφάλαιο γίνεται μία θεωρητική εισαγωγή για τους πλαγκτονικούς οργανισμούς, τη σημασία τους τόσο για το περιβάλλον όσο και σαν βιολογικοί δείκτες.

Περαιτέρω, γίνεται μία συνοπτική αναφορά σε ορισμένες σημαντικές οικογένειες και είδη πλαγκτού, κάποια από τα οποία ταξινομήθηκαν στα πλαίσια της εργασίας, καθώς ήταν μέρη του συνόλου φωτογραφιών. Στο Κεφάλαιο 3, παρουσιάζονται βασικές έννοιες της μηχανικής μάθησης και περιγράφεται συνοπτικά ο τρόπος λειτουργίας των νευρωνικών δικτύων που μελετήθηκαν. Γίνεται επίσης αναφορά στις τεχνικές που χρησιμοποιήθηκαν για την επεξεργασία των φωτογραφιών. Στο Κεφάλαιο 4, αναλύεται η μεθοδολογία που ακολουθήθηκε για την επεξεργασία των φωτογραφιών. Εν συνεχεία, αναλύονται οι διαφορετικές αρχιτεκτονικές που ακολουθήθηκαν για τη δημιουργία και εκπαίδευση των μοντέλων ταξινόμησης του πλαγκτού. Επίσης, παρουσιάζονται τα αποτελέσματα τους. Τελικώς, στο Κεφάλαιο 5 συνοψίζεται το αποτέλεσμα της διπλωματικής εργασίας, εκφράζονται μερικές ιδέες για μελλοντικές επεκτάσεις καθώς και κάποιοι περιορισμοί που υπήρξαν στην υλοποίησή της.

2 Πλαγκτόν

Ο όρος πλαγκτόν, πρώτη φορά χρησιμοποιήθηκε από τον βιολόγο Hensen. Τα πλαγκτόν είναι οργανισμοί που βρίσκονται στο νερό(ή στον αέρα) και δεν μπορούν να προωθηθούν ενάντια σε ρεύμα νερού ή αέρα. Η κατηγορία τους περιλαμβάνει φυτά, ζώα, βακτήρια, μύκητες και ιούς. Στους ωκεανούς, παρέχουν μια κρίσιμη πηγή τροφής σε πολλούς μικρούς και μεγάλους υδρόβιους οργανισμούς, όπως τα δίθυρα, τα ψάρια και τις φάλαινες. Το πλαγκτόν του γλυκού νερού είναι παρόμοιο με το θαλάσσιο πλαγκτόν. Παρότι τα περισσότερα είδη πλαγκτού είναι υδρόβια, υπάρχουν και κατηγορίες πλαγκτού που ζουν μέρος της ζωής τους παρασυρόμενες στην ατμόσφαιρα. Αυτές περιλαμβάνουν γύρη, διασκορπισμένους από τον άνεμο σπόρους, καθώς και θαλάσσιους μικροοργανισμούς που παρασύρονται στον αέρα ή τα κύματα.

2.1 Η αξία των παράκτιων υδάτων

Οι παράκτιες περιοχές λόγω της υψηλής παραγωγικότητας και αξίας τους, έχουν γίνει πόλος έλξης μεγάλης μερίδας του παγκόσμιου πληθυσμού. Συγκεκριμένα στην Ευρώπη, περισσότερο από το 1/3 του πληθυσμού, συναντάται σε περιοχές που απέχουν λιγότερο από 50 χιλιόμετρα από την ακτή. Παράκτια ύδατα, ορίζονται τα επιφανειακά ύδατα που βρίσκονται στην πλευρά της ξηράς μίας γραμμής, κάθε σημείο της οποίας βρίσκεται σε απόσταση ενός ναυτικού μιλίου (περίπου 1.8 χιλιομέτρων) προς τη θάλασσα από το πλησιέστερο σημείο της γραμμής βάσης από την οποία μετράται το εύρος των χωρικών υδάτων, και τα οποία, κατά περίπτωση, εκτείνονται μέχρι του απώτερου ορίου των μεταβατικών υδάτων.

Η εκτίμηση της οικολογικής αξίας και ποιότητας των παράκτιων υδάτων πραγματοποιείται βάση υδρομορφολογικών συνθηκών, χημικών χαρακτηριστικών αλλά και βιολογικών στοιχείων. Τα βιολογικά στοιχεία που χρησιμοποιούνται για την εκτίμηση της ποιότητας των παράκτιων υδάτων είναι: η σύνθεση και αφθονία της πανίδας βενθικών μακροασπονδύλων, η σύνθεση, αφθονία και βιομάζα του φυτοπλαγκτόν και η σύνθεση και αφθονία της λοιπής υδατικής χλωρίδας.

Για την ταξινόμηση της ποιότητας των παράκτιων υδάτινων συστημάτων, σημαντικό βιολογικό δείκτη αποτελούν τα πλαγκτόν αφού έχουν την ικανότητα να αποκρίνονται άμεσα σε περιβαλλοντολογικές ή ανθρωπογενείς πιέσεις που ασκούνται στο

σύστημά τους. Ένα από τα προτερήματα της χρήσης των διαφόρων ειδών πλαγκτού ως δείκτη, είναι ο μικρός χρόνος γενεάς τους. Αυτό έχει ως αποτέλεσμα να αντιδρούν σχετικά άμεσα στις περιβαλλοντικές αλλαγές και να μπορούν να εξαχθούν ολοκληρωμένα συμπεράσματα της πίεσης σε κλίμακες αντίστοιχες του χρόνου γενεάς. Βέβαια αυτό για να συμβεί απαιτούνται ιδιαίτερα συχνές δειγματοληψίες. Οι πιο σημαντικοί παράμετροι που λαμβάνονται υπόψιν για την εκτίμηση της ποιότητας ενός υδάτινου συστήματος, με βάση το φυτοπλαγκτόν, είναι η σύνθεση των ειδών και των ομάδων του, η βιομάζα καθώς και η διάρκεια και η ένταση των ανθίσεων του φυτοπλαγκτού [1].

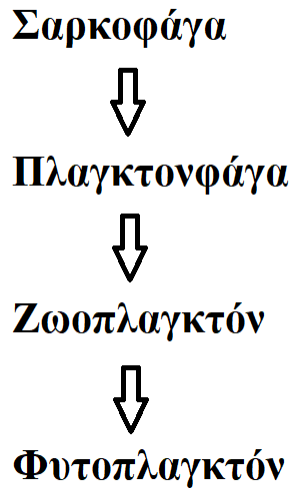
2.2 Φυτοπλαγκτόν

Το φυτοπλαγκτόν, συγκεκριμένα, αποτελείται από φωτοσυνθετικούς θαλάσσιους οργανισμούς όπως μικροφύκη και κυανοβακτήρια. Τα διαφορετικά είδη φυτοπλαγκτού έχουν προσαρμοστεί ειδικά ώστε να μπορούν να ζουν είτε σε γλυκό είτε σε θαλασσινό νερό και μετακινούνται, παθητικά, μέσω των θαλάσσιων ρευμάτων ή του ανέμου, ενώ διαθέτουν μηχανισμούς για να διατηρούνται σε αιώρηση. Αυτό τους επιτρέπει να μπορούν να επιβιώνουν σε εύρωτη ζώνη για κάποιο μέρος της ζωής τους.

Τα περισσότερα είδη φυτοπλαγκτού διαθέτουν μέγεθος όσο ενός ατόμου, δηλαδή κυμαίνονται σε ένα εύρος μεταξύ 0.2 και 200μm. Αρκετές φορές, όταν οι συγκεντρώσεις του φυτοπλαγκτού στο νερό είναι μεγάλες, δημιουργούν αποικίες οι οποίες έχουν τη δυνατότητα να ξεπεράσουν σε μέγεθος το 1mm και έτσι να γίνουν ορατές στο γυμνό μάτι. Οι φυτοπλαγκτονικοί οργανισμοί εμφανίζονται τόσο ως προκαρυωτικοί όσο και ως ευκαρυωτικοί. Παρουσιάζουν υψηλή αφθονία στα μεγέθη μεταξύ 0.6 και 2μm. Όπως αναφέρθηκε παραπάνω, αποτελούν σημαντικότερο παράγοντα της λειτουργίας των υδάτινων οικοσυστημάτων. Περαιτέρω, ο χρόνος γενεάς των φυτοπλαγκτονικών οργανισμών, συνήθως μεταξύ 24 και 48 ωρών, είναι αυτός που επηρεάζει τον ρόλο τους στα τροφικά πλέγματα αλλά και τη συνολική τους παραγωγικότητα.

Τα φυτοπλαγκτόν φωτοσυνθετούν τροφή για την παραγωγή ενέργειας για τους οργανισμούς ανώτερων τροφικών επιπέδων. Έτσι κατατάσσονται στους πρωτογενείς παραγωγούς του οικοσυστήματός τους. Ρυθμίζουν, δηλαδή, τη ροή της ενέργειας μέσα στο οικοσύστημα αυτό. Οι παραγωγοί, οι οποίοι είναι και αυτοί που συνθέτουν το οργανικό υλικό, αποτελούν τη βάση για τον κύκλο του άνθρακα καθώς και όλων των στοιχείων που χρησιμοποιούνται από τους μικροοργανισμούς αλλά και τους μακρο-

οργανισμούς υποστηρίζοντας έτσι όλες τις τροφικές αλυσίδες της βιόσφαιρας(όπως τους φυτοφάγους οργανισμούς, ζωοπλαγκτόν, τους πλαγκτονοφάγους οργανισμούς και τους πολύ μικρότερους πληθυσμούς των σαρκοφάγων).



Εικόνα 2-1: Απλοποιημένη απεικόνιση υδάτινου τροφικού πλέγματος

Η παρουσία τους σε ένα οικοσύστημα εξαρτάται από θεμελιώδους σημασίας αλληλοσυσχετίσεις ανάμεσα στον χρόνο γενεάς και στο βέλτιστο εύρος του φάσματος των περιβαλλοντικών παραγόντων όπως και την ικανότητα των φυτοπλαγκτονικών οργανισμών για διαχείμαση(δηλαδή τη δυνατότητα να μένουν σε ένα συγκεκριμένο μέρος κατά τους χειμερινούς μήνες). Ο χρόνος γενεάς έμμεσα αντανακλά τον ρυθμό αύξησης τους.

Τα φυτά είναι η τροφή των καταναλωτών πρώτης τάξεως. Όταν ολοκληρώσουν τον κύκλο της ζωής τους και πεθάνουν, γίνονται τροφή για τους αποικοδομητές. Οι πρωτογενείς καταναλωτές, όπως τα ζωοπλαγκτόν, είναι τροφή των καταναλωτών δεύτερης τάξεως όπως μικρά ψάρια και κάποια είδη ορνιθοπανίδας (πτηνών). Παράλληλα, παρατηρείται πως οι καταναλωτές συχνά αλλάζουν τροφικό επίπεδο στη διάρκεια του κύκλου της ζωής τους. Παραδείγματος χάριν, στα πρώτα στάδια της ζωής του ένα ψάρι μπορεί και τρέφεται με φυτοπλαγκτόν και νεκρή ή αποικοδομούμενη οργανική ύλη (σαπροφάγο). Αργότερα στον κύκλο της ζωής τους, με ζωοπλαγκτόν και στα πιο ώριμα στάδια, τρέφεται από άλλα ψάρια. Συγκεντρωτικά, τα τροφικά πλέγματα είναι συνήθως αρκετά σύνθετα και εκατοντάδες μικροοργανισμοί μετέχουν σε αυτά. Η ενέργεια που είναι διαθέσιμη από τα κατώτερα έως τα ανώτερα τροφικά επίπεδα μειώνεται με την αύξηση

επιπέδου. Έτσι πρέπει να υπάρχει μία πολύ μεγάλη βάση πρωτογενών παραγωγών ώστε να είναι ικανοί να συντηρήσουν τα υδάτινα οικοσυστήματα.

Παρατηρούμε πως επειδή η δυναμική των τροφικών αλυσίδων καθώς και οι βιογεωχημικοί κύκλοι είναι συνδεδεμένοι άρρηκτα από τη βιομάζα των πρωτογενών παραγωγών καθώς και πολλών ακόμα χαρακτηριστικών τους. Οι μικροοργανισμοί εκείνοι που επιτελούν την πρωτογενή παραγωγή είναι χρήσιμοι τόσο σε τοπική όσο και σε παγκόσμια κλίμακα καθώς σχεδόν το μισό της παγκόσμιας παραγωγής οξυγόνου συντελείται από αυτούς. Ο λόγος είναι η αφθονία του φυτοπλαγκτόν στους ωκεανούς [2].

Στα υδάτινα στρώματα που ο ρυθμός της φωτοσύνθεσης είναι πολύ υψηλός, όπως στις περιπτώσεις της άνθισης του νερού, μπορεί να εμφανιστεί υπερκορεσμός του νερού σε οξυγόνο. Το σημείο κορεσμού του νερού σε οξυγόνο αυξάνεται όσο μειώνεται η θερμοκρασία του. Σε περιόδους στρωμάτωσης της λίμνης, η μοναδική πηγή οξυγόνου για τις βαθύτερες ζώνες είναι οι φωτοσυνθετικοί οργανισμοί. Το φυτοπλαγκτόν και τα ανώτερα φυτά χρειάζονται αντίστοιχα οξυγόνο, διοξείδιο του άνθρακα και θρεπτικά συστατικά για να αναπτυχθούν και να επιζήσουν. Εκτός από κάποια ελάχιστα είδη μικροφυκών, τα υπόλοιπα δεν μπορούν να επιζήσουν σε περιβάλλον χωρίς διοξείδιο του άνθρακα. Το διοξείδιο του άνθρακα είναι πρακτικά πάντα διαθέσιμο, αφού προέρχεται από τη διάλυση των πετρωμάτων της λεκάνης απορροής, από τη λειτουργία της κυτταρικής αναπνοής των οργανισμών της λίμνης και τη διάχυση του από την ατμόσφαιρα. Τα βασικά θρεπτικά συστατικά, που είναι το άζωτο και ο φώσφορος, βρίσκονται διαλυμένα στο νερό ή στο ίζημα του πυθμένα και απορροφούνται από τους μικροφυτικούς οργανισμούς και τα ανώτερα φυτά αντίστοιχα. Η χαμηλή συγκέντρωσή τους στο νερό αποτελεί περιοριστικό παράγοντα για την ανάπτυξη των φυτών. Άλλα στοιχεία απαραίτητα για τη ζωή είναι τα κύρια ιόντα: ασβέστιο, μαγνήσιο, νάτριο και κάλιο, τα ιχνοστοιχεία σίδηρος, κοβάλτιο, μαγγάνιο, βόριο, ψευδάργυρος και είναι συνήθως παρόντα σε επαρκείς συγκεντρώσεις. Το θείο που επίσης χρειάζονται οι ζωντανοί οργανισμοί υπάρχει συνήθως στις λίμνες με τη μορφή αλάτων.

2.3 Ομάδες Φυτοπλαγκτού

Το φυτοπλαγκτόν έχει διαφορετική επίδραση στα υδάτινα οικοσυστήματα σύμφωνα με την αναλογία των παραλλαγών του στο εκάστοτε. Οι κυριότερες ομάδες του φυτοπλαγκτού διακρίνονται βάση των χρωστικών τους, αλλά και σύμφωνα με τις διαφορές

τους στη σύνθεση του κυτταρικού τοιχώματος. Οι παραλλαγές αυτές, είναι που εξηγούν και τις διαφορετικές πιέσεις που ασκεί το φυτοπλαγκτόν στο οικοσύστημα.

2.3.1 Κυανοβακτήρια

Μία από τις πολύ σημαντικές ομάδες-οικογένειες φυτοπλαγκτού είναι τα κυανοβακτήρια (ονομάζονται επίσης και γαλαζοπράσινα φύκια). Είναι μία αρχαία ομάδα φωτοσυνθετικών μικροβίων που εμφανίζονται στα περισσότερα ύδατα που βρίσκονται στη στεριά(λίμνες, ποτάμια) και μπορούν να έχουν σημαντικές επιπτώσεις στη λειτουργία του νερού και των υδάτινων οικοσυστημάτων. Το χαρακτηριστικό τους γνώρισμα είναι η ικανότητα τους να παράγουν οξυγόνο, το οποίο αποτελεί προϊόν φωτοσύνθεσης. Η συγκεκριμένη διαδικασία είναι γνωστή και ως οξυγενής φωτοσύνθεση. Τα κυανοβακτήρια έχουν μία ποικιλία τύπων κυττάρων και κυτταρικών δομών. Υπάρχουν παραπάνω από 2000 είδη χωρισμένα σε 150 γένη κυανοβακτηρίων με μεγάλη διαφορετικότητα σε σχήματα και μεγέθη. Τα χλωροξυβακτήρια, κυανόφυτα ή κυανοφύκη, τα οποία ανήκουν στην οικογένεια των κυανοβακτηρίων, συντελούν σημαντικά στη δημιουργία ευβακτηρίων. Διακρίνονται από άλλα βακτήρια από την ύπαρξη της χλωροφύλλης. Οι χλωροπλάστες των ανώτερων φυκών παρουσιάζουν την ίδια χλωροφύλλη, η οποία όμως έχει αρκετές διαφορές στη βιολογική της δομή σε σχέση με τα κυανοβακτήρια. Έχουν ιδιαίτερο ενδιαφέρον για τους διαχειριστές ποιότητας νερού, επειδή παράγουν χημικές ενώσεις γεύσης και οσμής, αρκετούς τύπους τοξινών και άλλων επιβλαβών ουσιών όπως επιφανειακά απορρίμματα. Πολλά είδη κυανοβακτηρίων, συλλέγονται ή καλλιεργούνται ως πηγές τροφής, ζωοτροφές, λιπάσματα ή χρησιμοποιούνται ακόμα και για προϊόντα υγείας [3].

2.3.2 Διάτομα

Τα διάτομα (βακιλλαριοφύκη), αποτελούν άλλη μία πολύ σημαντική ομάδα των μικροφυκών του φυτοπλαγκτού. Υπάρχουν σε αφθονία τόσο σε θαλάσσια οικοσυστήματα, στους ωκεανούς, όσο και σε εσωτερικά νερά. Τα διάτομα εμφανίστηκαν πριν 185 εκατομμύρια χρόνια, στην Ιουρασική περίοδο. Είναι μονοκύτταροι ευκαρυωτικοί οργανισμοί όπου το κυτταρικό τους τοίχωμα έχει ως βασικό στοιχείο το διοξείδιο του πυριτίου. Τα περισσότερα είδη, συναντώνται σε παράλιες ζώνες. Είναι υπεύθυνα για την παραγωγή αρκετά μεγάλου μέρους του οξυγόνου που παράγεται στον πλανήτη κάθε χρόνο

περίπου 20 με 40 τοις εκατό. Παρουσιάζουν μεγάλες ανθίσεις τόσο στους ωκεανούς αλλά και στις λίμνες και για αυτό το λόγο θεωρούνται σημαντικοί φωτοαυτότροφοι μικροοργανισμοί. Τα διάτομα είναι μονοκύτταρα και εμφανίζονται είτε ως μοναχικά κύτταρα είτε σε ομάδες (αποικίες) οι οποίες μπορούν να λάβουν διάφορα σχήματα. Τα μεμονωμένα κύτταρα κυμαίνονται σε μέγεθος από 2 έως 200 μm. Τα διάτομα έχουν δύο διακριτά σχήματα. Τα κεντρικά διάτομα είναι συμμετρικά ως προς την ακτίνα ενώ τα διάτομα με πτερύγια είναι αμφιπλεύρως συμμετρικά. Ένα μοναδικό χαρακτηριστικό της ανατομίας των διατόμων είναι ότι περιβάλλονται από ένα κυτταρικό τοίχωμα κατασκευασμένο από πυρίτιο (ενυδατωμένο διοξείδιο του πυριτίου). Τα περισσότερα διάτομα είναι πλαγκτικά, δηλαδή έχουν προσαρμοστεί ώστε να κινούνται παθητικά με αιώρηση μέσα στο νερό. Υπάρχουν όμως και κάποια είδη τα οποία μπορούν να προσκολλώνται σε σκληρές επιφάνειες όπως βράχια [4].

Μεγαλύτερες συγκεντρώσεις εμφανίζουν κατά την περίοδο της άνοιξης λόγω της υψηλής συγκέντρωσης νιτρικών και φωσφορικών (αφού τα νιτρικά αποτελούν, κατά την περίοδο της άνοιξης, κυρίαρχες μορφές του αζώτου). Τα διάτομα, έχουν τη δυνατότητα να αναπτύσσονται πιο γρήγορα, σε σχέση με άλλα μικροφύκη, στις χαμηλές θερμοκρασίες, και μπορούν έτσι να αντιμετωπίσουν τη μεγάλη διακύμανση του φωτός εκείνης της περιόδου. Παρουσία επαρκών θρεπτικών συστατικών και ηλιακού φωτός, ένα σύνολο διατόμων διπλασιάζεται περίπου κάθε μία μέρα με ασεξουαλική πολλαπλή σχάση και η μέγιστη διάρκεια ζωής των μεμονωμένων κυττάρων είναι περίπου έξι ημέρες.

Τα διάτομα χρησιμοποιούνται ως βιολογικοί δείκτες για την ποιότητα του νερού καθώς είναι ιδιαίτερα ευαίσθητα στις περιβαλλοντολογικές αλλαγές. Αποτελούν χρήσιμους δείκτες περιβαλλοντικών μεταβολών σε υδάτινους βιότοπους, παρέχοντας στοιχεία για τις μεταβολές του επιπέδου της στάθμης των θαλασσών και των λιμνών, αλλά και τον χημισμό των υδάτων. Επιπλέον, στο πλαίσιο της ανασύστασης του παλαιοπεριβάλλοντος για τον προσδιορισμό της παλαιοθερμοκρασίας χρησιμοποιούνται για την αναλογία ψυχρών/θερμών ειδών.

Πολλά διάτομα, παράγουν μία ισχυρή τοξίνη επονομαζόμενη Δομοϊκό Οξύ. Αυτή η τοξίνη είναι υπεύθυνη για την αμνησιακή δηλητηρίαση από οστρακοειδή. Η αύξηση του πληθυσμού των τοξικών διατόμων είναι ένα σημαντικό πρόβλημα καθώς απειλεί την ασφάλεια των θαλασσινών τροφίμων και συνεπώς την ανθρώπινη υγεία [5].

2.3.3 Δινομαστιγωτά

Τα δινομαστιγωτά ή αλλιώς δινοφύκη συναντώνται σε όλα τα οικοσυστήματα και περίπου το 50% των ειδών τους είναι φωτοσυνθετικά. Αρκετά είδη εμφανίζονται ως ενδοσύμβιοι αρκετών ασπόνδυλων , με πιο γνωστά τις ζωοξανθέλλες των κοραλλιών. Γενετικά, τα δινομαστιγωτά είναι αρκετά μοναδικά. Ο πυρήνας, μεγάλης πλειοψηφίας τους, είναι αρκετά διαφορετικός από άλλους ευκαρυωτικούς πυρήνες, τόσο που του έχει δοθεί ξεχωριστό όνομα, το δινοκάρυον. Τα χρωμοσώματα παραμένουν μόνιμα συμπυκνωμένα καθ' όλη τη διάρκεια του κυτταρικού κύκλου χωρίς τη βοήθεια νουκλεοσωμάτων και συνδέονται με το πυρηνικό περίβλημα. Τα γονιδιώματα των δινομαστιγωτών είναι από τα μεγαλύτερα που συναντώνται στη φύση και περιέχουν την ασυνήθιστη νουκλεοβάση υδροξυμεθυλουρακίλη [6].

Τυπικά, τα κύτταρα των δινομαστιγωτών παρουσιάζουν το διμαστίγωμα. ενώ υπάρχουν και κάποια είδη που σχηματίζουν αλυσίδες. Τα κύτταρα, με το διμαστίγωμα, έχουν δύο ανόμοια μαστίγια που βρίσκονται στα 'πλευρά' τους. Έχουν ένα εγκάρσιο μαστίγιο, σαν κορδέλα, που βρίσκεται στο 'δαχτυλίδι', που είναι ένα αυλάκι που περιβάλλει το κύτταρο και το δεύτερο είναι διαμήκες. Προσανατολίζεται περίπου ώστε να σχηματίζει ορθή γωνία με το εγκάρσιο μαστίγιο. Το εγκάρσιο μαστίγιο είναι σαν μια κυματιστή κορδέλα στην οποία μόνο το εξωτερικό άκρο κυματίζει από τη βάση μέχρι την άκρη. Η κίνηση του μπορεί και παράγει πρόωση προς τα εμπρός. Έχει και τη δυνατότητα για στροφική κίνηση. Επίσης διαθέτει τρίχες κατά μήκος του. Σε αντίθεση με το εγκάρσιο, το διαμήκη μαστίγιο έχει λίγες έως καθόλου τρίχες.

Τα κύτταρα των δινομαστιγωτών διαθέτουν ένα εξωτερικό στρώμα που ονομάζεται άμφισμα. Αυτό αποτελείται από ένα ενιαίο στρώμα κυστιδίων κάτω από την κυτταρική μεμβράνη, τις κυψελίδες ή αλλιώς αμφιεσμικά κυστίδια. Τα κύτταρα μπορούν να είναι είτε απογυμνωμένα, είτε να διαθέτουν τοίχωμα. Στο πρώτα, τα κυστίδια είναι συνήθως άδεια ή καμιά φορά περιέχουν άμορφο υλικό. Σε αυτά που διαθέτουν τοίχωμα, οι κυτταρινικές πλάκες, που εφαρμόζουν στενά, περιέχονται μέσα στις κυψελίδες, μία ανά κυστίδιο. Τα μοτίβα που σχηματίζονται είναι ουσιαστικής σημασίας για την ταξινόμηση τους.

Τα δινομαστιγωτά περιέχουν συχνά ένα ευρύ φάσμα οργανιδίων των οποίων η δουλειά είναι να εκκρίνουν υλικό προς τα έξω και πολλά από τα οποία σχετίζονται με την ακινητοποίηση του θηράματος. Αυτές περιλαμβάνουν τις τριχοκύστες, τις βλεννοκύστες και τις νηματοκύστες. Οι διαφορετικοί τύποι κυστών γενικά ορίζονται με βάση τον αριθμό

και τον τύπο των στρωμάτων στο κυτταρικό τοίχωμα και το εάν υπάρχουν μόνο βραχυπρόθεσμα ή μακροπρόθεσμα. Οι κύστες λεπτού τοιχώματος είναι γνωστές ως κύστες πελκώδους και οι κύστες διπλού τοιχώματος είναι γνωστές ως κύστες ηρεμίας. Αυτές οι κύστες δεν αποτελούν μόνο μέρος του κύκλου ζωής των δινομαστιγιωτών, αλλά είναι επίσης πολύ σημαντικές για την επιβίωση των φυκών σε δυσμενείς συνθήκες. Οι κύστες μπορούν να σχηματίσουν ‘τράπεζες σπόρων’ όταν οι συνθήκες είναι ευνοϊκές.

2.4 Ζωοπλαγκτόν

Το ζωοπλαγκτόν είναι το ζωικό συστατικό της πλαγκτονικής κοινότητας και μοιράζεται αρκετά κοινά χαρακτηριστικά με τα φυτοπλαγκτόν όπως η αιώρηση, όπου πρακτικά παρασύρεται στους ωκεανούς τις θάλασσες και τα σώματα γλυκού νερού από τα υδάτινα ρεύματα. Είναι ετερότροφα, που σημαίνει πως δεν παράγουν τη δική τους τροφή και πρέπει να καταναλώνουν άλλα φυτά ή ζώα για να επιβιώσουν. Στην περίπτωση των ζωοπλαγκτόν, τρέφονται με φυτοπλαγκτόν όπως φαίνεται και στην Εικόνα 2-1. Τα μεμονωμένα ζωοπλαγκτόν είναι συνήθως μικροσκοπικά αλλά μερικά όπως οι μέδουσες είναι μεγαλύτερα και ορατά με γυμνό μάτι. Κατά κανόνα είναι μεγαλύτερα από τα φυτοπλαγκτόν. Πολλά πρωτόζωα, δηλαδή μονοκύτταροι οργανισμοί που ληλατούν άλλη μικροσκοπική ζωή, ανήκουν στην κατηγορία του ζωοπλαγκτού, συμπεριλαμβανομένων των ζωομαστιγιωτών, των τρηματοφόρων, των ακτινόζωων (ραδιολάρια), ορισμένων δινομαστιγιωτών και θαλάσσιων μικρόζωων. Το μακροσκοπικό ζωοπλαγκτόν περιλαμβάνει πελαγικά κνιδόζωα, μαλάκια, αρθρόποδα και χιτωνοφόρα.

Γίνεται ξεκάθαρο λοιπόν, πως το ζωοπλαγκτόν είναι μια κατηγοριοποίηση που καλύπτει μία σειρά οργανισμών με αρκετά ευρύ φάσμα μεγεθών. Περιλαμβάνει πλήρως-πλαγκτονικούς οργανισμούς, δηλαδή αυτούς που σε ολόκληρο τον κύκλο της ζωής τους είναι σε μορφή πλαγκτόν, αλλά και μερικώς-πλαγκτονικούς οργανισμούς, που περνούν μέρος της ζωής τους στο πλαγκτόν προτού μεταβληθεί η φύση τους. Αν και όπως αναφέρθηκε, το μεγαλύτερο μέρος του ζωοπλαγκτόν μεταφέρεται παθητικά, κυρίως από τα υδάτινα ρεύματα, υπάρχουν αρκετά είδη που έχουν εξελιχθεί έτσι ώστε να αποκτήσουν τη δυνατότητα κίνησης που τους είναι χρήσιμη για την αποφυγή αρπακτικών ή για την αύξηση της πιθανότητας εύρεσης θηράματος.

Μέσω της κατανάλωσης και της επεξεργασίας του φυτοπλαγκτού και άλλων πηγών τροφής, το ζωοπλαγκτόν γίνεται πόρος για τους καταναλωτές υψηλότερων

τροφικών επιπέδων (συμπεριλαμβανομένων των ψαριών). Επίσης είναι σημαντικό στη διατήρηση της ισορροπίας του υδάτινου συστήματος σε περιπτώσεις όπου υπάρχει μεγάλη άνθιση φυτοπλαγκτόν. Το ζωοπλαγκτόν αποτελεί, επίσης, βασικό συντελεστή της βιομεγέθυνσης ρύπων όπως ο υδράργυρος. (Με τον όρο βιομεγέθυνση περιγράφεται η αύξηση της συγκέντρωσης μιας ουσίας, συνήθως τοξικής, στους ιστούς των ανεκτικών οργανισμών) [7].

Το ζωοπλαγκτόν μπορεί επίσης να λειτουργήσει ως δεξαμενή ασθενειών. Το ζωοπλαγκτόν των καρκινοειδών έχει βρεθεί ότι στεγάζει το βακτήριο *Vibrio cholerae*, το οποίο προκαλεί τη χολέρα, επιτρέποντας της να προσκολληθεί πάνω του. Αυτή η συμβιωτική σχέση ενισχύει την ικανότητα του βακτηρίου να επιβιώνει σε υδάτινο περιβάλλον, καθώς παρέχει στο βακτήριο απαραίτητα συστατικά όπως ο άνθρακας και το άζωτο [8].

2.5 Ομάδες ζωοπλαγκτόν

Οι οικολογικά σημαντικές ομάδες πρωτοζωικών ζωοπλαγκτόν, περιλαμβάνουν τα τρηματοφόρα και τα ακτινόζωα. Σημαντικό μεταζωικό ζωοπλαγκτόν είναι τα κνιδόζωα όπως οι μέδουσες, τα μαλακόστρακα όπως τα κωπήποδα, τα ισόποδα, μαλάκια όπως πτερόποδα και χορδωτά όπως σάλπες και νεαρά ψάρια. Αυτό το αρκετά μεγάλο φυλογενετικό εύρος περιλαμβάνει και ένα παρομοίως ευρύ φάσμα στη διατροφική συμπεριφορά του ζωοπλαγκτόν. Η παραπάνω μπορεί να γίνει με τροφοδοσία φίλτρου, δηλαδή τη διαδικασία όπου ένας οργανισμός καταναλώνει αρκετά μικρά κομμάτια τροφής κατευθείαν, θήρευση ή συμβίωση με αυτότροφο φυτοπλαγκτόν όπως συναντάται στα κοράλλια. Το ζωοπλαγκτόν τρέφεται με βακτηριοπλαγκτόν, φυτοπλαγκτόν, άλλο ζωοπλαγκτόν (μερικές φορές και κανιβαλιστικά), ακόμα και με υπολείμματα νεκρού σωματιδιακού οργανικού υλικού.

Όπως κάθε είδος μπορεί να περιοριστεί σε μία γεωγραφική περιοχή το ίδιο συμβαίνει και με το ζωοπλαγκτόν. Ωστόσο, τα είδη του ζωοπλαγκτού δεν διασκορπίζονται ομοιόμορφα ή τυχαία σε μία περιοχή του ωκεανού. Όπως και με το φυτοπλαγκτόν, κομμάτια ειδών ζωοπλαγκτού υπάρχουν σε όλον τον ωκεανό. Συγκεκριμένα είδη ζωοπλαγκτού περιορίζονται αυστηρά από την αλατότητα και τις διαβαθμίσεις της θερμοκρασίας, ενώ άλλα είδη αντέχουν σε μεγάλες διαβαθμίσεις θερμοκρασίας και αλατότητας. Η κηλίδωση του ζωοπλαγκτού μπορεί να επηρεαστεί από βιολογικούς αλλά και από φυσικούς

παράγοντες. Οι βιολογικοί παράγοντες περιλαμβάνουν την αναπαραγωγή και τη θήρευση, τη συγκέντρωση φυτοπλαγκτού και την κατακόρυφη μετανάστευση. Ο φυσικός παράγοντας που επηρεάζει περισσότερο την κατανομή του ζωοπλαγκτού είναι η ανάμειξη της στήλης του νερού, το οποίο επηρεάζει τη διαθεσιμότητα των θρεπτικών ουσιών, αλλάζοντας έτσι την ισορροπία της, κάτι που με τη σειρά του επηρεάζει άμεσα την παραγωγή φυτοπλαγκτού.

2.5.1 Τρηματοφόρα

Τα τρηματοφόρα είναι μονοκύτταροι οργανισμοί που τα χαρακτηρίζει ένα δίκτυο κοκκιωδών ψευδοποδίων, τα οποία είναι νηματοειδείς προεκτάσεις του πρωτοπλάσματος τους. Με τη χρήση των ψευδοποδίων τα τρηματοφόρα κινούνται, συλλαμβάνουν την τροφή τους ή στερεώνονται στο υπόβαθρο. Τα τρηματοφόρα έχουν στο κέλυφός τους οπές (δηλαδή τρήματα, από εκεί οπού προέρχεται και το όνομά τους). Το ενδοπλασματικό τους κέλυφος, αποτελείται από θαλάμους που χωρίζονται μεταξύ τους με διαφράγματα. Οι πρώτοι θάλαμοι είναι οι εμβρυακοί και ακολουθούνται από περαιτέρω θαλάμους εμβρυακούς ή μετεμβρυακούς οι οποίοι συγκροτούν την εμβρυακή συσκευή. Για την κατηγοριοποίηση και την ταξινόμηση των τρηματοφόρων παίζουν πολύ σημαντικό ρόλο ο αριθμός, η διάταξη και η μορφή των μετεμβρυακών θαλάμων. Οι οπές (τρήματα) διασφαλίζουν την επικοινωνία των θαλάμων μεταξύ τους αλλά και με το περιβάλλον.

Τα τρηματοφόρα είναι μία από τις σπουδαιότερες ομάδες λόγω της μεγάλης ποικιλίας και πληθώρας των αντιπροσώπων της. Τόσο σύγχρονων όσο και παλαιότερων. Βρίσκονται σε ιζηματογενή περιβάλλοντα. Η αφθονία τους τα καθιστά μοναδικής χρησιμότητας για τη μελέτη των σύγχρονων θαλάσσιων περιβαλλόντων, καθώς και για εφαρμογές στην πετρελαϊκή έρευνα. Τα τρηματοφόρα διαθέτουν ασβεστολιθικά κελύφη, τα οποία είναι αρκετά ανθεκτικά και έτσι προσφέρονται για ισοτοπικές αναλύσεις για τον προσδιορισμό κλιματικών διακυμάνσεων στις θαλάσσιες περιοχές κατά τη διάρκεια του γεωλογικού χρόνου. Οι λόγοι αυτοί συνετέλεσαν ώστε τα τρηματοφόρα να είναι από τα πρώτα μικροαπολιθώματα που μελετήθηκαν και χρησιμοποιήθηκαν για τον προσδιορισμό της ηλικίας αλλά και του παλαιοπεριβάλλοντος απόθεσης των ιζημάτων που τα περιέχουν, εξακολουθούν δε να αποτελούν τη σημαντικότερη ομάδα μικροαπολιθωμάτων

Τα τρηματοφόρα ζουν σε ολόκληρο το υδάτινο περιβάλλον, από τις τροπικές μέχρι ακόμα και τις πολικές περιοχές. Ορισμένα είδη ζουν σε υφάλμυρα ή αλμυρά νερά, με

εξαιρεση κάποια τρηματοφόρα που ζουν σε γλυκά νερά. Υπάρχουν δύο είδη τρηματοφόρων, τα πλαγκτονικά και τα βενθονικά. Το μεγαλύτερο μέρος των τρηματοφόρων είναι βενθονικοί οργανισμοί. Τα μικρότερα βενθονικά τρηματοφόρα ζουν στο βυθό και ευδοκιμούν σε μεγάλα βάθη. Τα πλαγκτονικά (που συνήθως συμβιώνουν με φύκη) και τα μεγαλύτερα βενθονικά, ζουν σε βάθος μέχρι 150 μέτρων. Το βάθος και η περιοχή που ζουν εξαρτάται από τις περιβαλλοντικές συνθήκες των υδάτων.

2.5.2 Ακτινόζωα

Τα ακτινόζωα αποτελούν ζωοπλαγκτόν των ανοιχτών θαλασσών και των ωκεανών. Εμφανίζονται άφθονα σε μεγάλες ωκεάνιες εκτάσεις σε όλον τον κόσμο. Ορισμένα είδη περιορίζονται σε μονάχα κάποιες συγκεκριμένες περιοχές και χρησιμοποιούνται ως δείκτες ιδιοτήτων μάζας νερού, όπως η θερμοκρασία, η αλατότητα και η συνολική βιολογική παραγωγικότητα. Η αφθονία των ακτινόζων ποικίλλει ανάλογα με τη γεωγραφική περιοχή. Για παράδειγμα, οι μέγιστες πυκνότητες φτάνουν τα 10.000 ανά κυβικό μέτρο σε ορισμένες περιοχές όπως ο υποτροπικός Ειρηνικός. Αντίθετα, οι πυκνότητες τους κυμαίνονται περίπου 3-5 ακτινόζωα ανά κυβικό μέτρο στη Θάλασσα των Σαργασσών [9].

Τα ακτινόζωα ταξινομούνται σε μία μεγάλη ομάδα ευκαρυωτικών οργανισμών που περιλαμβάνει τα φύκια και τα πρωτόζωα. Τα ακτινόζωα είναι μονοκύτταρα ή αποικιακά πρωτόζωα. Τα μονοκύτταρα είδη ποικίλλουν σε μέγεθος από μικρότερα από 100 μm έως και πολύ μεγάλα είδη με διαμέτρους 1–2 mm. Τα μεγαλύτερα είναι, ταξινομικά, λιγότερο πολυάριθμα και περιλαμβάνουν κυρίως ζελατινώδη είδη που απαντώνται συνήθως στα επιφανειακά ύδατα. Τα μικρότερα είδη διαθέτουν συνήθως πυριτικούς σκελετούς εξαιρετικά πολύπλοκου σχήματος. Η σκελετική μορφολογία είναι ξεχωριστή για το κάθε είδος και χρησιμοποιείται στην ταξινόμησή τους. Τα μεγαλύτερα μη αποικιακά είδη, είναι είτε χωρίς σκελετό, δηλαδή περικλείονται μόνο από ένα ζελατινώδες κάλυμμα, είτε διαθέτουν διάσπαρτες πυριτικές πλάκες μέσα στο περιφερειακό κυτταρόπλασμα και το περιβάλλον ζελατινώδες στρώμα. Τα αποικιακά είδη περιέχουν πολυάριθμα κύτταρα που διασυνδέονται με ένα δίκτυο κυτταροπλασματικών κλώνων και περικλείονται μέσα σε ένα διαυγές, ζελατινώδες περίβλημα. Οι αποικίες ποικίλλουν σε μέγεθος από αρκετά εκατοστά έως σχεδόν ένα μέτρο σε μήκος. Το σχήμα των αποικιών είναι αρκετά μεταβλητό μεταξύ των επιμέρους ειδών. Ορισμένα είναι σφαιρικά ή ελλειψοειδή ενώ άλλα έχουν σχήμα

κορδέλας ή κυλινδρική μορφή. Τα μεγαλύτερα είδη ακτινόζων είναι αναμφισβήτητα τα πιο ξεχωριστά και μεγαλύτερα από όλα τα γνωστά πρωτόζωα. Πολλά από τα είδη που κατοικούν στην επιφάνεια συμβιώνουν με φύκη. Τα συμβιωτικά φύκη παρέχουν τροφή στον ξενιστή με έκκριση φωτοσυνθετικά παραγόμενων οργανικών προϊόντων. Οι τροφικοί πόροι απορροφούνται από το ακτινόζωο και, σε συνδυασμό με τα τρόφιμα που συλλέγονται από το περιβάλλον, χρησιμοποιούνται για την υποστήριξη του μεταβολισμού και της ανάπτυξης του.

Τα ακτινόζωα που κατοικούν στα μεγάλα βάθη, εκεί όπου το φως του ηλίου είναι αρκετά περιορισμένο ή πρακτικά ανύπαρκτο, στερούνται τη συμβίωση με φύκη. Οι πυριτικοί σκελετοί τους, εγκαθίστανται στα ιζήματα των ωκεανών όπου σχηματίζουν ένα σταθερό πλέγμα απολιθωμάτων. Αυτά τα μικροαπολιθώματα αποτελούν σημαντική πηγή δεδομένων σε βιοστρωματογραφικές και παλαιοκλιματικές μελέτες. Οι παραλλαγές στον αριθμό και το είδος των ακτινοειδών ειδών, με βάση τη σκελετική μορφή σε σχέση με το βάθος στο ιζήμα, παρέχουν πληροφορίες για τις κλιματικές και περιβαλλοντικές συνθήκες στην υπερκείμενη υδάτινη μάζα της συγκεκριμένης γεωγραφικής τοποθεσίας. Τα ακτινόζωα είναι δεύτερα μετά τα διάτομα ως κύρια πηγή βιογενούς οπαλίου, που εναποτίθεται στα ιζήματα των ωκεανών.

Ακόμη, τα ακτινόζωα χρησιμοποιούνται ευρέως στη βιοστρωματογραφική ανάλυση των πελαγικών αποθέσεων. Βρίσκονται τόσο σε ανθρακικά όσο και σε πυριτικά και κλαστικά ιζήματα. Το πυριτικό τους κέλυφος τα καθιστά πολύτιμα για τη χρονολόγηση και τη βιοζώνωση των αποθέσεων σε λεκάνες μεγάλου βάθους. Οι ραδιολαρίτες, που είναι ιζήματα που βρίσκονται στις βαθιές λεκάνες, και πήραν το όνομά τους από τα περιεχόμενα ακτινόζωα (radiolaria), αποτελούν συνήθως τα πρώτα ιζήματα που αποτίθενται πάνω στους οφιολίθους μετά από την έξοδό τους στον ωκεάνιο φλοιό. Τα περιεχόμενα ακτινόζωα επιτρέπουν τη σχετική χρονολόγηση της εξόδου του οφιολιθικού μάγματος. Μαζί με τα κωνόδοντα τα ακτινόζωα αποτελούν σημαντικά μικροαπολιθώματα για τη βιοστρωματογραφία των πελαγικών ιζημάτων του Παλαιοζωικού και Τριαδικού, καθώς και των ορίων των αντίστοιχων βαθμίδων [10].

3 Μηχανική μάθηση και μέθοδοι επεξεργασίας εικόνας

Για να ορίσουμε τα βαθιά νευρωνικά δίκτυα (deep neural networks) και να κατανοήσουμε το πως λειτουργούν, θα πρέπει να κατανοήσουμε τις διαφορές τους με τις υπόλοιπες προσεγγίσεις τεχνητής νοημοσύνης.

3.1 Τεχνητή Νοημοσύνη

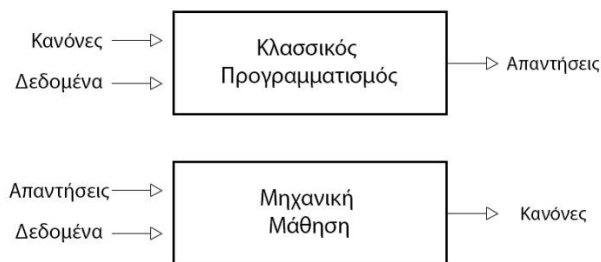
Η ιδέα της τεχνητής νοημοσύνης γεννήθηκε και ονομάστηκε πριν σχεδόν 70 χρόνια. Το 1956, ο McCarthy οργάνωσε ένα συνέδριο όπου συγκεντρώθηκε μία ομάδα επιστημόνων του εκκολαπτόμενου τομέα της επιστήμης υπολογιστών με σκοπό τη διερεύνηση των τρόπων με τους οποίους οι υπολογιστές θα μπορούσαν να κατασκευαστούν έτσι ώστε να προσομοιώσουν πτυχές της ανθρώπινης νοημοσύνης. Η βασική πρόταση προχωρούσε με την εικασία ότι κάθε πτυχή της μάθησης και γενικά της νοημοσύνης θα μπορούσε να περιγραφεί με αρκετή ακρίβεια ώστε μία μηχανή να μπορεί να την προσομοιώσει.

Συνοπτικά λοιπόν, μπορεί κάποιος να περιγράψει την τεχνητή νοημοσύνη ως την προσπάθεια αυτοματοποίησης των πνευματικών εργασιών που συνήθως εκτελούνται από ανθρώπους. Ως εκ τούτου, η τεχνητή νοημοσύνη μπορεί να οριστεί ως ένα γενικό πεδίο που περιλαμβάνει τη μηχανική μάθηση και τη βαθιά μάθηση, αλλά και τομείς και προσεγγίσεις προβλημάτων που δεν περιλαμβάνουν ‘μάθηση’ [11].

3.2 Μηχανική Μάθηση

Η μηχανική μάθηση ανέτρεψε τον τρόπο που διαχειρίζονταν τα προβλήματα μέχρι τότε ο υπολογιστής. Είναι ένας κλάδος της τεχνητής νοημοσύνης που στοχεύει στην κατασκευή προγραμμάτων ικανών να μάθουν να εκτελούν κάποιες ενέργειες μέσω μιας διαδικασίας που ονομάζεται εκπαίδευση. Ένα πρόγραμμα μηχανικής μάθησης κοιτάει τα δεδομένα εισόδου και τα αντίστοιχα δεδομένα εξόδου που αντιστοιχούν σε αυτά και προσπαθεί να βρει τους κανόνες που τα συνδέουν. Σε αντίθεση με τη συμβολική τεχνητή νοημοσύνη, ένα σύστημα μηχανικής μάθησης εκπαιδεύεται αντί για να προγραμματίζεται. Του παρουσιάζονται αρκετά δεδομένα/παραδείγματα, σχετικά με την εργασία που πρέπει να επιτελέσει, και βρίσκει μία στατιστική δομή σε αυτά τα παραδείγματα, ώστε αυτό τελικώς να του επιτρέψει να καταλήξει σε κανόνες που αυτοματοποιούν την εργασία αυτή.

Παραδείγματος χάριν, εάν επιζητούσαμε να αυτοματοποιήσουμε μία εργασία κατά την οποία θα αναγνώριζε, ένα σύστημα μηχανικής μάθησης, πυρκαγιές σε δάση από δορυφορικές εικόνες, θα μπορούσαμε να του παρουσιάσουμε πολλά παραδείγματα εικόνων, που έχουν ήδη κατηγοριοποιηθεί από ανθρώπους, με δάση όπου σε ορισμένα θα υπάρχουν πυρκαγιές ενώ σε άλλα όχι. Το σύστημα θα προσπαθούσε να μάθει τους στατιστικούς κανόνες που συσχετίζουν τα δύο είδη εικόνων με τις αντίστοιχες ετικέτες που τα κατηγοριοποιούν.



Εικόνα 3-1: Βασική αρχή λειτουργίας κλασσικού προγραμματισμού και μηχανικής μάθησης.

Η μηχανική μάθηση σχετίζεται με τη στατιστική αλλά διαφέρει σημαντικά από αυτήν με αρκετούς τρόπους. Ασχολείται με μεγάλα, πολύπλοκα σύνολα δεδομένων όπως ένα σύνολο εκατομμυρίων εικόνων τα οποία αποτελούνται από χιλιάδες εικονοστοιχεία για τα οποία η κλασική στατιστική ανάλυση, όπως η ανάλυση Bayes, δεν θα ήταν πρακτική. Σε αντίθεση με τη στατιστική, η μηχανική μάθηση παρουσιάζει σχετικά λίγη μαθηματική θεωρία και καθοδηγείται περισσότερο από εμπειρικά ευρήματα που βασίζονται αρκετά στην πρόοδο του λογισμικού αλλά και της υπολογιστικής ισχύος.

3.3 Τεχνητά Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα είναι μία υποκατηγορία του τομέα της υπολογιστικής νοημοσύνης και της μηχανικής μάθησης. Νευρωνικό δίκτυο ονομάζεται ένα δίκτυο από απλούς υπολογιστικούς κόμβους, ονομαζόμενους και νευρώνες, διασυνδεδεμένους μεταξύ τους με τέτοιο τρόπο ώστε όλο το σύστημα να λειτουργεί ως ταξινομητής.

3.3.1 Ταξινομητές

Όταν εργαζόμαστε με στατιστική και άλλους τομείς όπου συλλέγονται και αναλύονται μεγάλοι όγκοι δεδομένων, είναι συχνά απαραίτητο να ταξινομηθούν τα δεδομένα σε μικρότερες υποομάδες. Αυτό, ειδικά όταν μιλάμε για έναν τεράστιο όγκο δεδομένων, είναι πολύ δύσκολο έργο για έναν άνθρωπο. Ο οποίος πολλές φορές δεν είναι συχνά σε θέση να αναγνωρίσει σε ποια κατηγορία ανήκουν κάποια δεδομένα. Συνεπώς, χρησιμοποιείται συχνά ένας ψηφιακός ταξινομητής. Κοινό για όλους τους ταξινομητές είναι ότι λειτουργούν με επιβλεπόμενη μάθηση (supervised learning), όπου ο ταξινομητής εκπαιδεύεται σε δεδομένα με γνωστή έξοδο και στη συνέχεια χρησιμοποιούνται δεδομένα του ίδιου είδους, που του επιτρέπουν να χρησιμοποιεί τις γνώσεις του από τα δεδομένα εκπαίδευσης για ταξινόμηση σε νέα δεδομένα. Είναι σημαντικό ο ταξινομητής να μπορεί να γενικεύει και να ταξινομεί δεδομένα που δεν έχει ξανασυναντήσει, σύμφωνα με ποια υποομάδα μοιάζουν περισσότερο [12].

3.3.2 Συνάρτηση Ενεργοποίησης

Η συνάρτηση ενεργοποίησης (activation function) είναι ένας κανόνας που χρησιμοποιείται για να δώσει την τελική τιμή που βγαίνει από έναν νευρώνα. Ένας από τους λόγους που αυτή η συνάρτηση προστίθεται σε ένα τεχνητό νευρωνικό δίκτυο είναι για να βοηθήσει το δίκτυο να μάθει πολύπλοκα μοτίβα στα δεδομένα. Αυτές οι συναρτήσεις εισάγουν μη γραμμικές ιδιότητες πραγματικού κόσμου στα τεχνητά νευρωνικά δίκτυα. Σε ένα απλό νευρωνικό δίκτυο, όπου το x ορίζεται ως είσοδος και w τα βάρη και η συνάρτηση ενεργοποίησης που δίνει την τιμή που περνά στην έξοδο του δικτύου. Αυτή θα είναι τότε η τελική έξοδος ή η είσοδος ενός άλλου στρώματος.

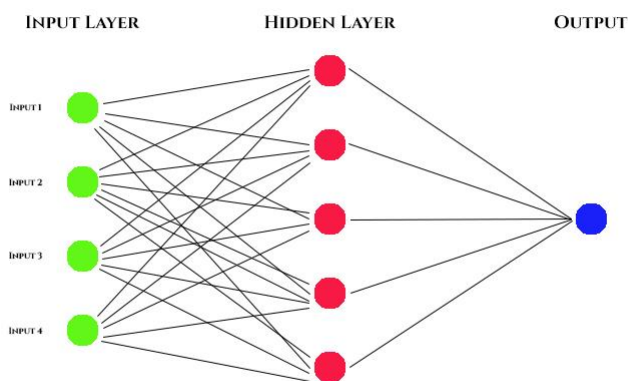
Εάν δεν εφαρμοστεί η λειτουργία ενεργοποίησης, το σήμα εξόδου γίνεται απλή γραμμική συνάρτηση. Ένα νευρωνικό δίκτυο χωρίς συνάρτηση ενεργοποίησης έχει περιορισμένη ισχύ εκμάθησης. Επίσης, πρέπει το νευρωνικό δίκτυο να μαθαίνει μη γραμμικές καταστάσεις καθώς του δίνονται σύνθετες πληροφορίες πραγματικού κόσμου όπως εικόνα, κείμενο και ήχος.

Υπάρχουν αρκετές συναρτήσεις ενεργοποίησης που χρησιμοποιούνται στα νευρωνικά δίκτυα. Οι πιο κοινές είναι οι εξής:

- Σιγμοειδής συνάρτηση (sigmoid)
- Συνάρτηση υπερβολικής εφαπτομένης (tanh)
- Συνάρτηση Relu

3.3.3 Δομή Νευρωνικών Δικτύων

Πολλαπλοί νευρώνες διατάσσονται σε στρώματα για να σχηματίσουν ένα νευρωνικό δίκτυο. Στην Εικόνα 3-6 βλέπουμε μία βασική δομή ενός νευρωνικού δικτύου



Εικόνα 3-2: Απλοϊκό παράδειγμα δομής νευρωνικού δικτύου.

Αυτή η μορφή δικτύου, χρησιμοποιεί πλήρως συνδεδεμένα επίπεδα, τα οποία αποστέλλουν την έξοδό τους στους νευρώνες που βρίσκονται παρακάτω. Κάθε εσωτερικό στρώμα αναφέρεται ως κρυφό επίπεδο. Το επίπεδο εξόδου παράγει ένα σύνολο τιμών πιθανοτήτων για ένα σύνολο αμοιβαία αλληλοαποκλειόμενων κλάσεων.

Ο όρος βαθιά νευρωνικά δίκτυα ή αλλιώς βαθιά μάθηση, αναφέρεται στην ιδέα της χρήσης πολλών ενδιάμεσων στρωμάτων μετασχηματισμένων δεδομένων για την εύρεση της εξόδου με μεγαλύτερη ακρίβεια [12].

3.3.4 Εκπαίδευση

Μέχρι στιγμής δεν έχουμε εξηγήσει πως η εκπαίδευση λειτουργεί στα τεχνητά νευρωνικά δίκτυα. Θα εξετάσουμε μονάχα την εποπτευόμενη μάθηση (supervised learning) καθώς είναι και η πιο κοινή μορφή εκπαίδευσης νευρωνικών δικτύων [13]. Στην εποπτευόμενη μάθηση, διαθέτουμε ένα εκπαιδευτικό σύνολο διανυσμάτων εισόδου \tilde{x}_n όπου $n = 1, 2, \dots, N$, με το αντίστοιχο σύνολο διανυσμάτων εξόδου \tilde{y}_n . Στόχος είναι η εκπαίδευση της συνάρτησης $f(x; \theta) = y$ του νευρωνικού δικτύου, έτσι ώστε να προσαρμόζεται για να προσεγγίζει τα δεδομένα της εκπαίδευσης. Το θ περιγράφει την παράμετρο των βαρών, τα οποία μπορούν να αλλάξουν, ή αλλιώς μαθευτούν, κατά τη διάρκεια της εκπαίδευσης. Επομένως, για να μπορεί να διαμορφωθεί, είναι απαραίτητη η

χρήση μίας συνάρτησης απώλειας. Μία απλή συνάρτηση απώλειας είναι το μέσο τετραγωνικό σφάλμα που υπολογίζει την τετραγωνική διαφορά μεταξύ της $f(x; \theta)$ και του \tilde{y} .

$$J(\theta) = \frac{1}{2N} \sum_{n=1}^N |f(\tilde{x}_n; \theta) - \tilde{y}|^2$$

Η συνάρτηση απώλειας $J(\theta)$ υπολογίζει τη μέση απώλεια σε όλο το σετ εκπαίδευσης και χρησιμοποιείται για την ελαχιστοποίηση του σφάλματος μεταξύ των δεδομένων εκπαίδευσης και του τεχνητού νευρωνικού δικτύου σε σχέση με το θ . Για την ελαχιστοποίηση της συνάρτησης απώλειας, μπορεί να χρησιμοποιηθεί η κάθοδος βασισμένη στην κλίση (gradient descent), η οποία χρειάζεται την κλίση της συνάρτησης απώλειας ως προς το θ . Η μέθοδος αυτή, είναι πρακτικά ένας επαναληπτικός αλγόριθμος βελτιστοποίησης, ο οποίος ενημερώνει τα βάρη με τη βοήθεια της κλίσης. Στην πιο απλή της μορφή, μπορεί να γραφεί μαθηματικά ως :

$$\theta \leftarrow \theta - \eta * \frac{\partial J(\theta)}{\partial \theta}$$

Η μέθοδος αυτή δεν βρίσκει απαραίτητα το ολικό ελάχιστο και μπορεί να κολλήσει σε ορισμένες περιπτώσεις σε λάθος τιμές. Η αρχικοποίηση, συνεπώς, της παραμέτρου θ σε διαφορετικές τιμές, μπορεί να οδηγήσει και σε διαφορετική λύση. Έτσι είναι σημαντική, ειδικά όταν βασίζεται στα βάρη του δικτύου.

Κάθε βάρος στο θ , μπορεί να γραφθεί ως $w_{j,k}^l$, όπου l είναι το επίπεδο, το j περιγράφει τον νευρώνα στο αντίστοιχο επίπεδο και k είναι ο νευρώνας στο προηγούμενο επίπεδο. Παρακάτω φαίνεται η εξίσωση που αρχικοποιεί τα βάρη βασισμένη στην ευρετική μέθοδο των Glorot και Bengio :

$$w_{j,k}^l = U\left(-\frac{6}{\sqrt{n_{l-1} + n_l}}, \frac{6}{\sqrt{n_{l-1} + n_l}}\right)$$

Η ευρετική αυτή χρησιμοποιεί το μέγεθος του επιπέδου n_{l-1} , του προηγούμενου επιπέδου, και το μέγεθος του επιπέδου n_l για να αρχικοποιήσει το βάρος του l . Η συνάρτηση $U[-b,b]$ περιγράφει μια ομοιόμορφη κατανομή μεταξύ $(-b,b)$ [14].

3.3.4.1 Αλγόριθμος οπισθοδιάδοσης

Ένας σημαντικός αλγόριθμος που αναφέραμε προωύτερα, είναι αυτός της οπισθοδιάδοσης (backpropagation). Πρόκειται για έναν αλγόριθμο που χρησιμοποιείται για τη διάδοση του σφάλματος, από τη συνάρτηση απώλειας, μέσω του δικτύου για τον υπολογισμό της κλίσης. Δεν είναι ένας αλγόριθμος εκμάθησης για το νευρωνικό δίκτυο, αλλά μία αποτελεσματική μέθοδος για τον υπολογισμό της κλίσης εν συναρτήσε των βαρών. Για να εξάγει τα βάρη από την κλίση χρησιμοποιεί την παρακάτω εξίσωση:

$$\frac{\partial J(\theta)}{\partial w_{j,k}^l} = \delta_j^l * y_k^{l-1}$$

Όπου δ:

$$\delta_j^l \begin{cases} \frac{\partial J(\theta)}{\partial y_j^l} * \sigma'_l(z_j^l), \text{ αν } l \text{ είναι το επίπεδο εξόδου} \\ (\sum_{i=1}^q \delta_j^{l+1} * w_{i,j}^{l+1}) \sigma'_l(z_j^l) \text{ αν } l \text{ είναι ένα κρυφό επίπεδο} \end{cases}$$

Το y_j^l περιγράφει τη έξοδο του νευρώνα j στο επίπεδο l. Το πλήθος των νευρώνων στο επίπεδο l+1 είναι το q ενώ η συνάρτηση ενεργοποίησης είναι η σ_l . Η z_j^l είναι η τιμή ενεργοποίησης του νευρώνα j στο επίπεδο l και έχει τύπο :

$$z_j^l = \sum_{i=0}^k w_{j,i}^l * y_i^{l-1}$$

Για να μπορέσει να ξεκινήσει η διαδικασία οπισθοδιάδοσης χρειάζεται πρώτα να υπολογιστούν οι τιμές ενεργοποίησης z_j^l και η έξοδος y_j^l . Αυτές θα βρεθούν με την εμπρόσθια διάδοση.

Αφού υπολογιστούν όλες οι τιμές, μπορεί να ξεκινήσει το βήμα της καθόδου βασισμένη στην κλίση, ως εξής:

$$w_{j,k}^l \leftarrow w_{j,k}^l - \eta * \frac{\partial J(\theta)}{\partial w_{j,k}^l}$$

Αυτή η διαδικασία διαρκεί μέχρι κάποιο κριτήριο να τη σταματήσει και έχει να κάνει συνήθως με συνθήκες που διαλέγει ο προγραμματιστής [14].

3.3.5 Συνάρτηση softmax

Συνήθως, το τελευταίο επίπεδο, είναι ένα επίπεδο softmax. Η softmax ή αλλιώς κανονικοποιημένη εκθετική συνάρτηση, είναι μία γενίκευση της logistic function, σε πολλαπλές διαστάσεις. Η κύρια εφαρμογή της softmax, είναι για την αναπαράσταση μιας κατανομής πιθανοτήτων σε K κλάσεις, για αυτό και χρησιμοποιείται στο τελευταίο μέρος του δικτύου. Η softmax εφαρμόζεται σε ολόκληρο το στρώμα εξόδου και χρειάζεται το γραμμικό συνδυασμό όλων των νευρώνων. Η τύπος της softmax είναι ο ακόλουθος:

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \quad \text{για } j = 1, \dots, K.$$

3.4 Συνελκτικά Νευρωνικά Δίκτυα (CNN)

Τα συνελκτικά νευρωνικά δίκτυα είναι μία κατηγορία εξειδικευμένων τεχνητών νευρωνικών δικτύων. Τα CNN είναι κατάλληλα για την επεξεργασία εικόνων, αλλά και για βίντεο ή ήχο. Στην εργασία χρησιμοποιήθηκαν τα νευρωνικά δίκτυα συνέλιξης για την επεξεργασία εικόνας, οπότε και θα αναφερθούμε σε αυτή.

Ένα CNN αποτελείται από πολλαπλά επίπεδα συνέλιξης και επίπεδα συγκέντρωσης (pooling layers). Ένα επίπεδο συγκέντρωσης εφαρμόζεται μετά από ένα ή περισσότερα επίπεδα συνέλιξης. Τα επίπεδα συνέλιξης έχουν την ιδιότητα να εξάγουν χρήσιμα χαρακτηριστικά από την είσοδο, κάτι που οδηγεί σε πολλαπλούς χάρτες χαρακτηριστικών (feature maps). Το επίπεδο συγκέντρωσης μειώνει χωρικά το μέγεθος αυτών των χαρτών.

Κατά τη διαδικασία επεξεργασίας εικόνας, εφαρμόζεται η συνέλιξη. Επεξεργάζεται έτσι την εικόνα ώστε να εξάγει κάποια χαρακτηριστικά από αυτήν, αυτό γίνεται με διάφορα φίλτρα, που έχουν διαφορετικό μέγεθος, τα οποία εφαρμόζονται σε ένα κομμάτι της εικόνας. Στην περίπτωση των CNN, αντί να γίνεται χρήση

προγραμματισμένων φίλτρων, το δίκτυο μαθαίνει αυτά τα φίλτρα από μόνο του. Για να εφαρμοστεί αυτή η διαδικασία στα νευρωνικά δίκτυα, αλλάζουν κάποια κομμάτια της δομής τους [15]. Περαιτέρω θα αναφερθούμε στα επίπεδα συνέλιξης αλλά και στα επίπεδα συγκέντρωσης

3.4.1 Επίπεδο Συνέλιξης (Convolutional Layer)

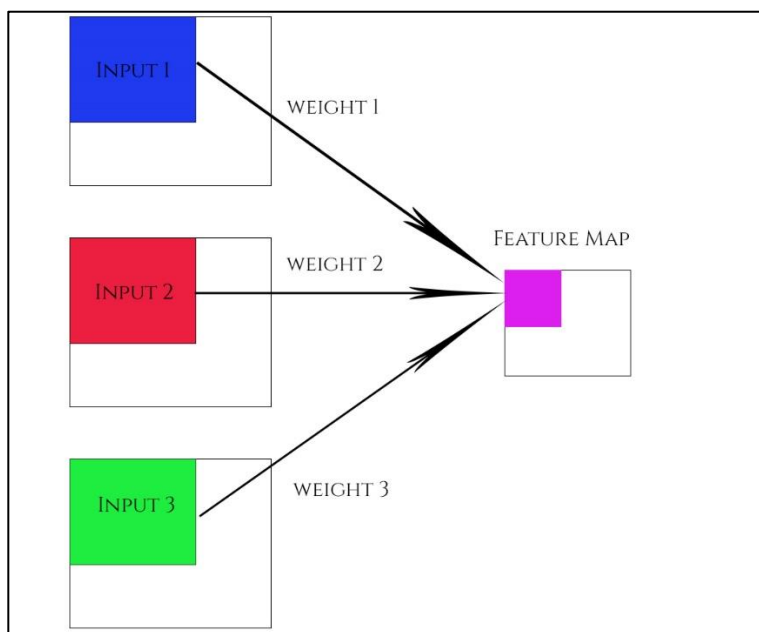
Ένα επίπεδο συνέλιξης αποτελείται από νευρώνες, όπως και ένα κανονικό νευρωνικό δίκτυο, αλλά με διαφορετική διάταξη και διαφορετική συνδεσιμότητα μεταξύ τους. Οι κύριες διαφορές με τα νευρωνικά δίκτυα είναι:

- Καταμερισμός των βαρών
- Τοπική συνδεσιμότητα
- Τρισδιάστατη διάταξη των νευρώνων

Ο λόγος που η διάταξη είναι τρισδιάστατη, έγκειται στο γεγονός ότι μελετώνται εικόνες. Οι εικόνες, όταν είναι έγχρωμες, έχουν τρία διαφορετικά κανάλια χρώματος συνήθως το κόκκινο, το πράσινο και το μπλε. Κάθε κανάλι αντιστοιχεί και σε μία διάσταση. Στην περίπτωση που οι εικόνες είναι τις γκρι κλίμακας (δηλαδή ασπρόμαυρες), θα χρησιμοποιούσαμε μία διάσταση. Ενώ, εάν γινότανε μελέτη των συνελκτικών αλγορίθμων σε βίντεο, θα έπρεπε να υπάρχει και μία επιπλέον διάσταση. Αυτή του χρόνου.

Ο καταμερισμός των βαρών σημαίνει ότι τα ίδια βάρη χρησιμοποιούνται για πολλούς νευρώνες εξόδου. Με τον όρο τοπική συνδεσιμότητα εννοούμε ότι δεν συνδέονται όλοι οι νευρώνες της εισόδου με τους νευρώνες της εξόδου.

Η τοπική συνδεσιμότητα περιγράφεται με το μέγεθος του πυρήνα (kernel size). Όταν ασχολούμαστε με εισόδους πολλαπλών διαστάσεων, όπως είναι οι εικόνες, δεν είναι πρακτικό να συνδεθούν οι νευρώνες με όλους τους νευρώνες του προηγούμενου επιπέδου. Μια τέτοια αρχιτεκτονική δικτύου δεν λαμβάνει υπόψη τη χωρική δομή των δεδομένων. Τα συνελκτικά δίκτυα επιβάλλουν ένα αραιό μοτίβο τοπικής συνδεσιμότητας μεταξύ νευρώνων γειτονικών στρωμάτων. Σε αυτήν την περίπτωση, η σύνδεση των νευρώνων γίνεται μονάχα σε μία μικρή περιοχή του όγκου της εισόδου. Η έξοδος του νευρωνικού δικτύου είναι και πάλι ένας πολυδιάστατος πίνακας, με χάρτες χαρακτηριστικών δύο διαστάσεων για κάθε φίλτρο που υπάρχει στο επίπεδο. Πρακτικά, κάθε φίλτρο παράγει και έναν χάρτη χαρακτηριστικών [12].



Εικόνα 3-3: Σχηματική αναπαράσταση της σύνδεσης μεταξύ των εισόδων και της εξόδου ενός συνελκτικού επιπέδου.

Στην Εικόνα 3-3, η είσοδος εμφανίζεται με τρία κανάλια. Αυτά μπορεί να είναι τα χρωματικά κανάλια μιας εικόνας ή οι χάρτες χαρακτηριστικών ενός προηγούμενου επιπέδου. Η είσοδος περιγράφεται από έναν τρισδιάστατο πίνακα. Ως έξοδος απεικονίζεται ένας χάρτης χαρακτηριστικών. Για περισσότερους χάρτες χαρακτηριστικών η συμπεριφορά είναι παρόμοια. Το επίπεδο συνέλιξης έχει μόνο έναν χάρτη χαρακτηριστικών για την έξοδο. Η συνάρτηση ενεργοποίησης παραλήφθηκε. Το μέγεθος του χάρτη χαρακτηριστικών είναι μικρότερο από το μέγεθος των καναλιών εισόδου. Ο αριθμός των νευρώνων της εξόδου, εξαρτάται από το μέγεθος της εισόδου και τον αριθμό των φίλτρων.

Ο αριθμός των φίλτρων περιγράφει πόσοι χάρτες χαρακτηριστικών θα πρέπει να δημιουργηθούν. Αλλά το χωρικό μέγεθος του χάρτη χαρακτηριστικών και επομένως ο αριθμός των νευρώνων, εξαρτάται από το χωρικό μέγεθος εισόδου.

Λόγω του καταμερισμού των βαρών και της τοπικής συνδεσιμότητας, ο αριθμός των βαρών μειώνεται εκθετικά σε σύγκριση με ένα πλήρως συνδεδεμένο επίπεδο με παρόμοιο αριθμό μονάδων. Παραδείγματος χάριν, αν N το πλάτος και το ύψος ενός πυρήνα, Σ ο αριθμός των καναλιών εισόδου και B ο αριθμός των φίλτρων, που ισοδυναμεί με τον αριθμό των χαρτών χαρακτηριστικών. Τότε το επίπεδο έχει αριθμό βαρών

$N^2 * \Sigma * B$ μονάχα, πολύ λιγότερο από ένα αντίστοιχο δίκτυο πλήρως συνδεδεμένο δίκτυο [15].

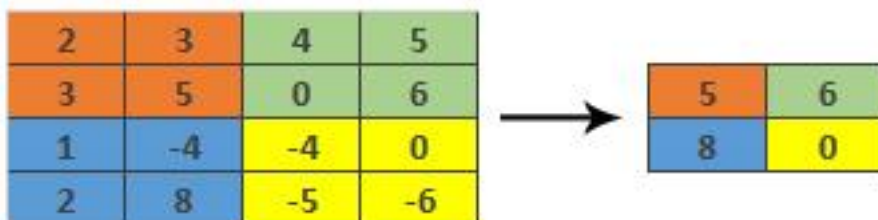
Η συνέλιξη εφαρμόζεται μεταξύ της εισόδου και των βαρών. Έστω ότι η είσοδος $\text{Input}(:, :, i) = V_i$, όπου V_i είναι ένας διδιάστατος πίνακας που περιγράφει το i -κανάλι της εισόδου. Το $W_{i,o}$ περιγράφει έναν πυρήνα με μέγεθος $N * N$ και Z_o είναι επίσης ακόμα ένας πίνακας διδιάστατος που περιγράφει το o -κανάλι της εξόδου. Τότε, ο o -χάρτης χαρακτηριστικών υπολογίζεται ως:

$$Z_o = \sigma \sum_{i=0} V_i * W_{i,o} + b * \mathbf{1}$$

Όπου $\mathbf{1}$ είναι ο μοναδιαίος πίνακας.

3.4.2 Επίπεδο Συγκέντρωσης

Το επίπεδο συγκέντρωσης (pooling layer) εφαρμόζεται συνήθως μετά από ορισμένα επίπεδα συνέλιξεων. Εν συντομία, πρόκειται για μία διαδικασία η οποία όπως η συνέλιξη, είναι υπεύθυνη για τη μείωση του μεγέθους του αντίστοιχου χάρτη χαρακτηριστικών. Μειώνει τον όγκο των παραμέτρων, επειδή το επόμενο επίπεδο θα λάβει μια μικρότερη έκδοση των χαρτών χαρακτηριστικών. Αυτό συμβαίνει για τη μείωση της υπολογιστικής ισχύος που απαιτείται για την επεξεργασία των δεδομένων, μειώνοντας τη διαστατικότητα (dimensionality). Επιπλέον, είναι χρήσιμο για την εξαγωγή των κυρίαρχων χαρακτηριστικών του κάθε χάρτη χαρακτηριστικών.



Εικόνα 3-4: Παράδειγμα λειτουργίας του επιπέδου συγκέντρωσης μεγίστου.

Τα δύο πιο κοινά επίπεδα συγκέντρωσης είναι αυτά της συγκέντρωσης μεγίστου και τη συγκέντρωσης του μέσου. Και τα δύο στρώματα συγκέντρωσης λαμβάνουν ως

παραμέτρους το μέγεθος του φίλτρου και το βήμα. Το πρώτο υπολογίζει το μέγιστο σε κάθε αντίστοιχο κομμάτι, ενώ το δεύτερο τον μέσο όρο. Ένα παράδειγμα λειτουργίας του επιπέδου συγκέντρωσης μεγίστων (max pooling layer) φαίνεται στην Εικόνα 3-4.

Σε κάποιες περιπτώσεις για να μην χαθούν δεδομένα όταν το βήμα δεν διαιρεί χωρίς υπόλοιπο τον αντίστοιχο χάρτη χαρακτηριστικών, χρησιμοποιούμε γέμισμα (padding), ώστε να μην χαθεί κάποια στήλη ή γραμμή.

3.4.3 Υπολειμματικά μπλοκ (Residual blocks)

Καθώς αυξάνεται το βάθος των νευρωνικών δικτύων, δηλαδή όσο περισσότερα επίπεδα έχουν, συχνά εμφανίζεται ένα πολύ σημαντικό πρόβλημα. Αυτό της εξαφάνισης της παραγωγού, το οποίο οδηγεί στην υποβάθμιση της απόδοσης του δικτύου. Όταν συμβαίνει αυτό, η ακρίβεια του δικτύου φτάνει σε έναν κορεσμό, και εν συνεχεία υποβαθμίζεται αρκετά γρήγορα. Για να αντιμετωπιστεί αυτό το πρόβλημα, δημιουργήθηκε μία αρχιτεκτονική που χρησιμοποιεί υπολειμματικά μπλοκ [16]. Ένα υπολειμματικό μπλοκ αποτελείται από ένα σύνολο επιπέδων που έχουν οριστεί με τέτοιο τρόπο ώστε η έξοδος ενός επιπέδου να προστίθεται σε ένα άλλο επίπεδο βαθύτερα στο μπλοκ. Η μη γραμμικότητα εφαρμόζεται όταν προστεθεί το παραπάνω, με το αποτέλεσμα στο τέλος του αντίστοιχου επιπέδου εκτός του μπλοκ. Πρακτικά, με αυτό τον τρόπο παρακάμπτονται κάποια επίπεδα σε ορισμένες επαναλήψεις και τιμές. Ένα τέτοιο δίκτυο δημιουργείται με τη στοίβαξη πολλών τέτοιων μπλοκ και ονομάζεται ResNet.

3.4.4 Ανεστραμμένα υπολειμματικά μπλοκ (Inverted residual blocks)

Ένα ανεστραμμένο υπολειμματικό μπλοκ, ονομάζεται και MBConv, είναι ένας τύπος υπολειμματικού μπλοκ που χρησιμοποιείται σε μοντέλα αναγνώρισης εικόνας, και έχει ανεστραμμένης δομής για λόγους αποδοτικότητας. Η πρώτη αρχιτεκτονική που προτάθηκε με τη συγκεκριμένη δομή ήταν το MobileNetV2, το οποίο είναι ένα συνελκτικό νευρωνικό δίκτυο. Έκτοτε έχει χρησιμοποιηθεί σε πολλά δίκτυα όπως το EfficientNet που χρησιμοποιήθηκε και στα πλαίσια της εργασίας.

Ένα παραδοσιακό υπολειμματικό μπλοκ έχει δομή που από ευρεία γίνεται στενή και μετά ξανά ευρεία, βασισμένη στον αριθμό των καναλιών. Η είσοδος του, έχει μεγάλο αριθμό καναλιών τα οποία συμπιέζονται με μία συνέλιξη 1x1 και στη συνέχεια ο αριθμός

των καναλιών αυξάνεται ξανά ώστε να μπορεί να είναι εφικτή η πρόσθεση μεταξύ εισόδου και εξόδου.

Στην περίπτωση των ανεστραμμένων υπολειμματικών μπλοκ ακολουθείται μία δομή που είναι αρχικά στενή, μετά ευρεία και τελικά ξανά στενή. Αρχικά διευρύνεται με μία συνέλιξη 1x1 η είσοδος, αφετέρου χρησιμοποιείται μία συνέλιξη 3x3 για τη μείωση των παραμέτρων, και στη συνέχεια με ακόμα μία συνέλιξη 1x1 μειώνονται τα κανάλια ώστε να μπορεί να προστεθεί η είσοδος με την έξοδο [17].

3.5 Μέθοδοι επεξεργασίας εικόνας

3.5.1 Τελεστής Sobel-Feldman

Ο τελεστής Sobel-Feldman, γνωστός και ως φίλτρο Sobel, είναι ένα εργαλείο που χρησιμοποιείται σε εφαρμογές επεξεργασίας εικόνας και υπολογιστικής όρασης. Η εφαρμογή του είναι ιδιαίτερα χρήσιμη σε περιπτώσεις που είναι αναγκαία η εύρεση των άκρων. Πήρε το όνομά του από τον Irwin Sobel και τον Gary Feldman του εργαστηρίου τεχνητής νοημοσύνης του Στάνφορντ το 1968. Πρότειναν την ιδέα ενός 3x3 ιστροπικού τελεστή εικόνας ο οποίος είναι ένας διακριτός τελεστής διαφοροποίησης ο οποίος υπολογίζει τη διαβάθμιση της συνάρτησης της έντασης της εικόνας, επιστρέφοντας το σχετικό διάνυσμα της κλίσης της. Λειτουργεί ασκώντας συνέλιξη στην εικόνα με ένα διακριτό σύνολο σημείων [18]. Περισσότερα για τον τρόπο λειτουργίας του αναφέρονται στο επόμενο κεφάλαιο όπου και εφαρμόστηκε.

3.5.2 Γκαουσιανό φίλτρο θόλωσης (Gaussian blur filter)

Κατά την επεξεργασία μίας εικόνας, συχνά υπάρχει ενδιαφέρον στην αναγνώριση των αντικειμένων που αναπαρίστανται σε αυτήν. Η εύρεση των άκρων είναι πολύ σημαντική για το παραπάνω. Η άκρες πρακτικά αντιπροσωπεύουν τη μετάβαση από μία ομάδα με παρόμοια εικονοστοιχεία σε μία άλλη. Η διαδικασία θόλωσης μίας εικόνας καταφέρνει να κάνει τη χρωματική μετάβαση από τη μία άκρη στην άλλη ομαλότερη. Επίσης μπορεί να χρησιμοποιηθεί για την ελαχιστοποίηση του θορύβου σε μία εικόνα.

Όταν εφαρμόζεται ένα φίλτρο, θεωρούμε ένα ορθογώνιο σύνολο εικονοστοιχείων που βρίσκονται γύρω από το κάθε εικονοστοιχείο. Ο πυρήνας του φίλτρου, είναι ένα ξεχωριστό σύνολο από εικονοστοιχεία που έχουν διαστάσεις ίδιες με το πρώτο σύνολο που αναφέραμε. Μία σημαντική ιδιότητα του πυρήνα του φίλτρου είναι ότι πρέπει να έχει

πάντα μονούς αριθμούς στις διαστάσεις του (ύψος και το πλάτος), έτσι ώστε το εικονοστοιχείο στο οποίο εφαρμόζεται το φίλτρο να είναι πάντα στο κέντρο. Για να εφαρμοστεί ένας πυρήνας σε κάποιο εικονοστοιχείο, υπολογίζεται ο μέσος όρος των χρωματικών τιμών που αντιστοιχούν στο περιβάλλον του εικονοστοιχείου, ο οποίος είναι σταθμισμένος με τις τιμές του πυρήνα. Στην περίπτωση του Γκαουσιανού φίλτρου θόλωσης, τα εικονοστοιχεία που βρίσκονται στο κέντρο του πυρήνα έχουν και το μεγαλύτερο βάρος, το οποίο είναι αντιστρόφως ανάλογο με την απόσταση από τον πυρήνα. Αφού σταθμιστεί ο μέσος όρος ανά κανάλι, οι μέσες τιμές το καναλιού γίνονται η νέα τιμή για το εικονοστοιχείο στη φιλτραρισμένη εικόνα. Συνεπώς, όσο μεγαλύτερος είναι ο πυρήνας, τόσο περισσότερο θα θολώνει την εικόνα σε σχέση με έναν μικρότερο πυρήνα, καθώς οι μεγαλύτεροι πυρήνες έχουν περισσότερες τιμές συνυπολογισμένες στον μέσο όρο.

3.5.3 Αλγόριθμος εύρεσης άκρων του Canny

Η ανίχνευση των άκρων είναι μία τεχνική που μπορεί να χρησιμοποιηθεί για την εξαγωγή χρήσιμων πληροφοριών από δεδομένα εικόνας. Η εφαρμογή της τεχνικής αυτής μπορεί να μειώσει, δραματικά, τον όγκο των δεδομένων προς επεξεργασία.

Ο John F. Canny δημιούργησε έναν αλγόριθμο ο οποίος μετά από κάποια συγκεκριμένα στάδια, μπορεί να εξάγει τις άκρες μίας τροφοδοτούμενης, στον αλγόριθμο, εικόνας.

Ένας αλγόριθμος εύρεσης άκρων, για να λειτουργεί ορθά θα πρέπει να έχει χαμηλό ποσοστό σφάλματος, να μην σχηματίζονται στην τελική εικόνα άκρες από στοιχεία θορύβου της εικόνας και να εντοπίζεται με ακρίβεια το κέντρο της κάθε άκρης ώστε να μην υπάρχει αλλοίωση της τελικής εικόνας.

Για να ικανοποιηθούν οι παραπάνω απαιτήσεις τα βήματα που επιτάσσει ο αλγόριθμος του Canny είναι τα εξής:

1. εφαρμογή ενός Γκαουσιανού φίλτρου θόλωσης για εξομάλυνση του θορύβου
2. εύρεση της διαβάθμισης στην έντασης της εικόνας(π.χ. με την χρήση του τελεστή Sobel-Feldman)
3. εφαρμογή μιας μεθόδου καταστολής των ψευδών άκρων
4. εφαρμογή διπλού κατωφλιού για προσδιορισμό πιθανών άκρων

5. ολοκλήρωση της ανίχνευσης άκρων καταστέλλοντας όλες τις άκρες που είναι αδύναμες και δεν συνδέονται με ισχυρές άκρες (υστέρηση)

Οι εικόνες για να χρησιμοποιηθούν από τον αλγόριθμο πρέπει να είναι σε γκρι κλίμακα (δηλαδή ασπρόμαυρες). Μετά την σωστή χρήση των παραπάνω βημάτων οι εικόνα που θα εξαχθεί από τον αλγόριθμο θα είναι ασπρόμαυρη. Μόνο οι σημαντικές άκρες θα πρέπει να φαίνονται και επιπροσθέτως, θα πρέπει να έχουν όλα τα εικονοστοιχεία της εικόνας την ίδια ένταση.

4 Μεθοδολογία

Το συγκεκριμένο κεφάλαιο ασχολείται με το κύριο κομμάτι της εργασίας. Δηλαδή την μεθοδολογία που ακολουθήθηκε για την επεξεργασία του συνόλου των δεδομένων και την εκπαίδευση του νευρωνικού δικτύου βαθιάς μάθησης, για την ταξινόμηση των οικογενειών του πλαγκτού. Χωρίζεται σε τρία μέρη, στο πρώτο γίνεται αναφορά στο σύνολο των δεδομένων, δηλαδή των εικόνων που λήφθηκαν από πολωτικό μικροσκόπιο. Στο δεύτερο μέρος αναλύεται η μεθοδολογία που χρησιμοποιήθηκε για την προεπεξεργασία των ακατέργαστων εικόνων ώστε να μπορούν να χρησιμοποιηθούν καταλλήλως από το νευρωνικό δίκτυο. Ενώ στο τελευταίο κομμάτι του κεφαλαίου αποτυπώνονται τα στάδια δημιουργίας των νευρωνικών δικτύων που υλοποιήθηκαν, γίνεται μία ανάλυση των επιμέρους κομματιών τους, σύγκριση μεταξύ τους και τελικά, εκφράζονται τα αποτελέσματα τους.

4.1 Σύνολο των Δεδομένων

Τα αρχεία των φωτογραφιών που χρησιμοποιήθηκαν για τη δημιουργία του συνόλου των δεδομένων (dataset), αποτελούνται από μία σειρά φωτογραφιών που λήφθηκαν με πολωτικό μικροσκόπιο στο εργαστήριο μικροπαλαιοντολογίας του τομέα Ιστορικής Γεωλογίας και Παλαιοντολογίας στο Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών. Υλικό ιζηματοπαγίδων από το Βόρειο Αιγαίο, παραχώρηση από Μ. Τριανταφύλλου. Προέρχεται από τη δημοσιευμένη εργασία Skampa, E., Triantaphyllou, M. V., Dimiza, M. D., Gogou, A., Malinverno, E., Stavrakakis, S., Panagiotopoulos, I.P., Parinos, C. and Baumann, K. H. (2019). Coupling plankton-sediment trap-surface sediment coccolithophore regime in the North Aegean Sea (NE Mediterranean). *Marine Micropaleontology*, 152, 101729 [19].

Οι φωτογραφίες που λήφθηκαν από την κάθε οικογένεια πλαγκτού ήταν ήδη ταξινομημένες, κάτι που είναι αναγκαίο για την εκπαίδευση των νευρωνικών δικτύων.

Για τη δημιουργία του τελικού συνόλου δεδομένων ακολουθήθηκαν οι εξής διαδικασίες:

1. Καθαρισμός των εικόνων και ομοιογενοποίησή τους
2. Χρήση του αλγορίθμου εύρεσης άκρων του Canny για περαιτέρω ανάλυση των δεδομένων
3. Επαύξηση των δεδομένων για τον εμπλουτισμό του συνόλου.

Στα επόμενα υποκεφάλαια γίνεται αναλυτική επισκόπηση της μεθοδολογίας που ακολουθήθηκε.

4.1.1 Καθαρισμός των εικόνων

Ένα πρόγραμμα μηχανικής μάθησης δεν μπορεί να λειτουργήσει εάν δεν υπάρχουν ικανοποιητικά δεδομένα για να εκπαιδευτεί. Στο συγκεκριμένο πρόβλημα της ταξινόμησης, ο αλγόριθμος θα πρέπει να μπορεί να ξεχωρίσει τις διαφορές μεταξύ των εικόνων του δείγματος για να εξάγει τα σημαντικότερα στοιχεία τους.

Τα δεδομένα που χρησιμοποιήθηκαν περιείχαν αρκετό θόρυβο. Ο θόρυβος μπορεί να έχει αρκετές μορφές. Σε κάποιες από τις φωτογραφίες υπάρχει αντανάκλαση του φωτός από τη λήψη της φωτογραφίας. Επίσης, πέρα από το πλαγκτόν, υπάρχουν αρκετά θορυβώδη στοιχεία του περιβάλλοντος στη φωτογραφία που θα μπορούσαν να μεπεδέψουν τον αλγόριθμο μειώνοντας την επιτυχία του. Η εξαγωγή των χαρακτηριστικών στοιχείων, που περιγράφει τις σχετικές πληροφορίες που περιέχονται σε ένα μοτίβο, είναι πολύ σημαντική και για τη μείωση της διαστατικότητας του προβλήματος. Όσο μικρότερη είναι, τόσο λιγότερους υπολογιστικούς πόρους χρειάζεται για να λειτουργήσει. Παρότι το σύνολο των δεδομένων που χρησιμοποιήθηκε στα πλαίσια αυτής της εργασίας είναι σχετικά μικρό, σε περίπτωση που η κλίμακα των δεδομένων ήταν μεγαλύτερη, η μείωση της διαστατικότητας θα ήταν πολύ σημαντική, οπότε και έγινε προσπάθεια να συμπεριληφθεί.

Επίσης, ένα άλλο πρόβλημα που αντιμετωπίστηκε ήταν η ύπαρξη πολλών διαφορετικών πλαγκτονικών οργανισμών σε αρκετές από τις εικόνες, ενώ σε άλλες ο θόρυβος ήταν πιο έντονος από την εικόνα του οργανισμού οπότε και έπρεπε να αφαιρεθούν από το σύνολο των δεδομένων.

Λόγω του ανομοιόμορφου χρωματικού προφίλ ανάμεσα στις διαφορετικές οικογένειες πλαγκτού, το πρόγραμμα έκανε εσφαλμένη ταξινόμηση στις φωτογραφίες.

Αυτό δεν συμβαίνει σε όλες της εικόνες τους, αλλά σε έναν μεγάλο αριθμό, ενώ οι εικόνες των διατόμων είναι σχεδόν όλες ασπρόμαυρες.

Για να ελεγχθεί κατά πόσο αυτό αποτελεί πρόβλημα, δημιουργήθηκε ένα απλό νευρωνικό δίκτυο ταξινόμησης και εκπαιδεύτηκε μονάχα με δύο ομάδες πλαγκτού, που παρουσίαζαν διαφορετικό χρωματικό προφίλ στις φωτογραφίες τους. Το νευρωνικό παρουσίασε μεγάλη ακρίβεια στα δεδομένα εκπαίδευσης. Για να ελεγχθεί το κατά πόσο θα επηρέαζε η διαφορά του χρωματικού προφίλ την εγκυρότητα του εκπαιδευμένου νευρωνικού, με χρήση προγράμματος επεξεργασίας εικόνας, αλλάχθηκε το χρωματικό προφίλ μερικών εκ των φωτογραφιών της μίας ομάδας ώστε να μοιάζει με αυτό της άλλης και χρησιμοποιήθηκαν αυτές για την δοκιμή του προγράμματος.

Το αποτέλεσμα ήταν να εμφανίσει το μοντέλο πολύ χαμηλή ακρίβεια στις προβλέψεις του καθώς, όπως ήταν αναμενόμενο, το χρώμα των εικόνων επηρέασε την εκπαίδευση και τον τρόπο ταξινόμησης των δειγμάτων. Συνεπώς, όλες οι φωτογραφίες θα πρέπει να μετασχηματιστούν ώστε να έχουν το ίδιο χρωματικό προφίλ. Για να καλυφθεί αυτή η ανάγκη, δημιουργήθηκε ένα πρόγραμμα στην Python που χρησιμοποιεί τη βιβλιοθήκη Scikit για τη μετατροπή των έγχρωμων φωτογραφιών σε ασπρόμαυρες.

Συγκεκριμένα έγινε χρήση της συνάρτησης `rgb2gray` με τις παρακάτω παραμέτρους:

```
def rgb2gray(original_img):  
    r, g, b = original_img[:, :, 0], original_img[:, :, 1], original_img  
   [:, :, 2]  
    grayscale_img = 0.2959 * r + 0.5570 * g + 0.1440 * b  
    return grayscale_img
```

Όλες οι φωτογραφίες εισήχθησαν στο παραπάνω πρόγραμμα και έγινε η μετατροπή τους σε ασπρόμαυρες.

4.1.2 Υλοποίηση αλγορίθμου εύρεσης άκρων του Canny

Αναφέρθηκε πριν, ότι η μείωση της διαστατικότητας είναι χρήσιμη, ειδικά σε σύνολα δεδομένων που έχουν μεγάλο όγκο. Για τη μείωση της διαστατικότητας έγινε επέκταση του κώδικα του παραπάνω προγράμματος για να ανιχνεύει τα άκρα και τις ακμές της κάθε φωτογραφίας. Η ιδέα είναι να δημιουργηθεί ένας αλγόριθμος που μπορεί να σκιαγραφήσει τις άκρες οποιουδήποτε αντικειμένου που υπάρχει στην εικόνα. Είναι

σημαντικό, γιατί κάθε εικόνα μεταφέρει μεγάλο πλήθος άχρηστων δεδομένων, έτσι μπορεί να μειωθεί αρκετά ο χρόνος επεξεργασίας του μοντέλου. Έγινε χρήση του αλγορίθμου εύρεσης άκρων που δημιούργησε ο John F. Canny.

Ο αλγόριθμος του Canny έχει 3 στάδια υλοποίησης:

1. Μείωση θορύβου
2. Υπολογισμός κλίσης – αρχική εύρεση άκρων
3. Καθαρισμός ατελειών – τελική εύρεση άκρων

Η συγκεκριμένη μέθοδος χρησιμοποιεί παραγώγους για να υπολογίζει την κλίση και είναι αρκετά ευαίσθητη στον θόρυβο της εικόνας. Επίσης απαιτεί οι εικόνες να είναι ασπρόμαυρες. Για να μειωθεί ο θόρυβος πρέπει να εφαρμοστεί κάποιο φίλτρο. Ένας τρόπος που μπορεί αυτό να επιτευχθεί είναι με την εφαρμογή ενός Γκαουσιανού φίλτρου θόλωσης (Gaussian blur).

Αυτό μπορεί να γίνει με τη χρήση μίας τεχνικής συνέλιξης της εικόνας. Ο πυρήνας της συνέλιξης καθορίζει το ποσό της θόλωσης που θα προκύψει. Όσο μεγαλύτερος είναι ο πυρήνας, τόσο πιο θολό θα είναι το αποτέλεσμα.

Ένας πυρήνας με Γκαουσιανό φίλτρο και μέγεθος $L = (2 * k + 1) \times (2 * k + 1)$ έχει τον ακόλουθο τύπο:

$$K_{ij} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(i - (k + 1) + (j - (k + 1)))^2}{2\sigma^2}\right) \text{ όπου } 1 \leq i, j \leq L$$

Αφού υλοποιήθηκε η παραπάνω σχέση στην Python και εισήχθησαν τα δεδομένα παρατηρήθηκε ότι η χρήση του φίλτρου μείωσε αρκετά την οξύτητα των στοιχείων θορύβου της εικόνας ενώ δεν αλλοίωσε σημαντικά το σημείο ενδιαφέροντος, που είναι ο πλαγκτονικός οργανισμός.

Για να βρεθούν τα άκρα της εικόνας, θα πρέπει να εφαρμόσουμε το επόμενο βήμα που είναι ο υπολογισμός της κλίσης. Θα γίνει ανίχνευση των ακμών χρησιμοποιώντας έναν τελεστή. Οι άκρες υποδηλώνουν πρακτικά την αλλαγή της έντασης των εικονοστοιχείων. Θα εφαρμοστεί ένα φίλτρο το οποίο θα βρίσκει αυτήν την αλλαγή έντασης σε δύο διαστάσεις (οριζόντια και κάθετα). Αυτό μπορεί να γίνει με τη χρήση του τελεστή Sobel–Feldman ο οποίος χρησιμοποιείται σε τέτοιου είδους

προβλήματα υπολογιστικής όρασης. Ο τελεστής χρησιμοποιεί δύο 3×3 πίνακες οι οποίοι είναι σε συνέλιξη με την αρχική εικόνα. Αν θεωρήσουμε A την αρχική εικόνα και G_x, G_y δύο εικόνες που περιέχουν σε κάθε σημείο τις κατακόρυφες και οριζόντιες προσεγγίσεις των παραγώγων, ο τύπος αποκτά αυτήν τη μορφή:

$$G_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} * A, \quad G_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} * A$$

Όπου το $*$ συμβολίζει τη συνέλιξη μεταξύ του εκάστοτε πίνακα με την αρχική εικόνα A . Η συντεταγμένη x ορίζεται ως αυξανόμενη προς τη δεξιά κατεύθυνση ενώ η y προς τα κάτω. Ο συνδυασμός των δύο προσεγγίσεων της κλίσης μπορεί να δώσει το μέγεθος της κλίσης με τον τύπο:

$$G = \sqrt{G_x^2 + G_y^2}$$

Και για να υπολογιστεί η κατεύθυνση του διανύσματος της κλίσης, χρησιμοποιείται ο παρακάτω τύπος:

$$\theta = \text{atan2}(G_y, G_x), \quad \text{όπου } -\pi < \theta \leq \pi$$

Η υλοποίηση των παραπάνω οδήγησε σε εύρεση των άκρων της εικόνας που όμως είχαν διαφορετικό πάχος, επίσης τα εικονοστοιχεία είχαν ανόμοια ένταση.

Τα επόμενα βήματα του αλγορίθμου του Canny, είναι αυτά της χρήσης της μη-μέγιστης καταστολής (για να λεπτύνουν οι άκρες) και του διπλού κατωφλιού (που στοχεύει στον εντοπισμό των ισχυρών, αδύναμων και μη σχετικών άκρων) για το τελικό στάδιο της εύρεσης άκρων και ο η διαδικασία της υστέρησης (hysteresis).

Στο βήμα της μη-μέγιστης καταστολής, ο αλγόριθμος περνά από όλα τα σημεία του πίνακα της έντασης της κλίσης και βρίσκει τη μέγιστη τιμή στις κατευθύνσεις των άκρων. Ο σκοπός του είναι να ελέγξει εάν τα εικονοστοιχεία στην ίδια κατεύθυνση είναι λιγότερο ή περισσότερο έντονα από αυτά που υποβάλλονται προς επεξεργασία. Παραδείγματος χάριν, έστω ένα εικονοστοιχείο που ελέγχεται και άλλα δύο που βρίσκονται πλησίον του, προς την ίδια κατεύθυνση. Εάν ένα από αυτά τα δύο εικονοστοιχεία έχει μεγαλύτερη τιμή, δηλαδή είναι πιο έντονο, από το αρχικό, τότε διατηρείται μόνο το πιο έντονο εικονοστοιχείο, ειδάλλως η τιμή του γίνεται μηδενική.

Για την υλοποίηση αυτού του βήματος, χρειάστηκε να δημιουργηθεί ένας μηδενικός πίνακας που έχει το ίδιο μέγεθος με τον αρχικό. Να προσδιοριστούν οι κατευθύνσεις των άκρων με βάση της τιμής της γωνίας. Έπειτα να γίνει έλεγχος για το εάν τα εικονοστοιχεία πλησίον του προς επεξεργασία εικονοστοιχείου έχουν μεγαλύτερη τιμή και τέλος να επιστραφεί η επεξεργασμένη εικόνα αφού γίνουν οι αλλαγές.

Αφού ολοκληρώθηκε και αυτό το βήμα παρατηρήθηκε πως οι άκρες ήταν πολύ πιο ομοιογενείς. Παρόλα αυτά, κάποιες ήταν πιο φωτεινές από τις άλλες.

Στο επόμενο βήμα του αλγορίθμου, δηλαδή αυτό της εφαρμογής του διπλού κατωφλιού, τα ισχυρά εικονοστοιχεία, δηλαδή αυτά που έχουν πολύ υψηλή ένταση, διατηρούνται καθώς θεωρείται ότι συμβάλλουν στην ύπαρξη άκρου. Τα εικονοστοιχεία που έχουν τιμή μικρότερη από το χαμηλό κατώφλι, δεν θεωρούνται σημαντικά και παραλείπονται. Τελικώς, τα εικονοστοιχεία που βρίσκονται ανάμεσα στα δύο κατώφλια, θεωρούνται αδύναμα και στο επόμενο βήμα του αλγορίθμου θα εξεταστεί εάν μπορούν να θεωρηθούν ισχυρά ή μη σχετικά.

Με βάση τα αποτελέσματα που εξάχθηκαν από το προηγούμενο βήμα, αυτό της εφαρμογής του διπλού κατωφλιού, ακολουθήθηκε η διαδικασία της υστέρησης. Η οποία μετατρέπει τα αδύναμα εικονοστοιχεία σε ισχυρά, μονάχα εάν γύρω τους υπάρχει κάποιο ισχυρό εικονοστοιχείο.

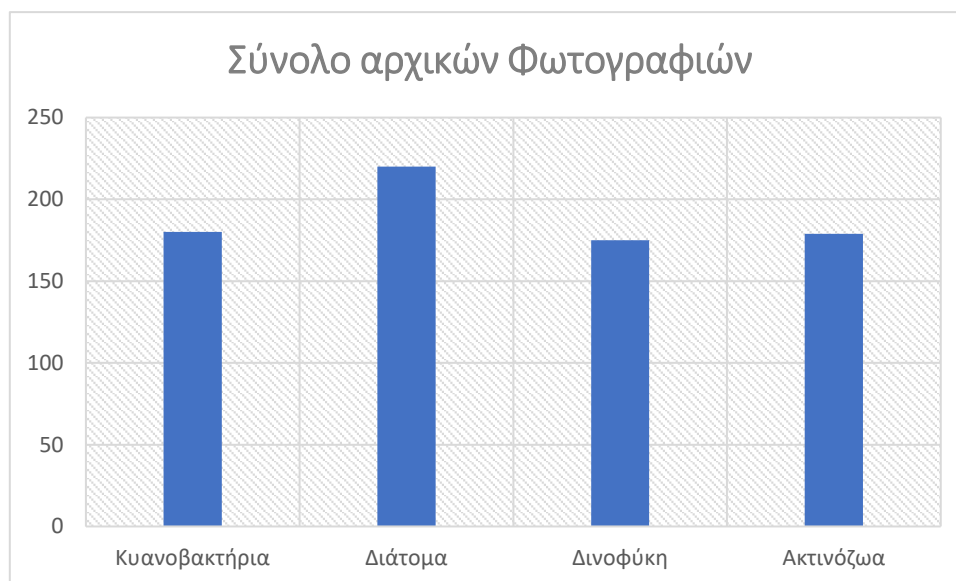
Τα παραπάνω βήματα υλοποιήθηκαν, όμως μετά από τη δοκιμή του προγράμματος στις φωτογραφίες παρατηρήθηκε ένα σημαντικό πρόβλημα. Αρκετές φωτογραφίες διαθέτουν τόσο θόρυβο ώστε να μη είναι δυνατή η χρήση αυτής της μεθόδου στα πλαίσια του συγκεκριμένου συνόλου δεδομένων. Καθώς ο θόρυβος σε πολλές περιπτώσεις υπερκαλύπτει τον ίδιο τον μικροοργανισμό.

Παρατηρείται πως ο αλγόριθμος του Canny λειτουργεί ικανοποιητικά. Δεν παρατηρούνται διακυμάνσεις της έντασης των γραμμών και όλα τα εικονοστοιχεία έχουν την ίδια φωτεινότητα.

Η τεχνική αυτή μοιάζει κατά κάποιον τρόπο με τη λειτουργία των συνελκτικών νευρωνικών δικτύων και θα μπορούσε να χρησιμοποιηθεί για τη διερεύνηση προβληματικών φωτογραφιών σε ένα σύνολο δεδομένων μεγάλου όγκου. Έτσι χρησιμοποιήθηκε για να αναγνωρισθούν ποιες εικόνες είναι προβληματικές και να αφαιρεθούν από το τελικό σύνολο δεδομένων.

Το πρώτο κομμάτι αυτής της διεργασίας, για τη μείωση του θορύβου, χρησιμοποιήθηκε σε όσες φωτογραφίες κρίθηκε απαραίτητο. Τα συνολικά δείγματα των

φωτογραφιών ήταν 754 (Εικόνα 4-1), εκ των οποίων αρκετές μετά από το φιλτράρισμα με τον αλγόριθμο του Canny, αφαιρέθηκαν από το σύνολο των εικόνων.



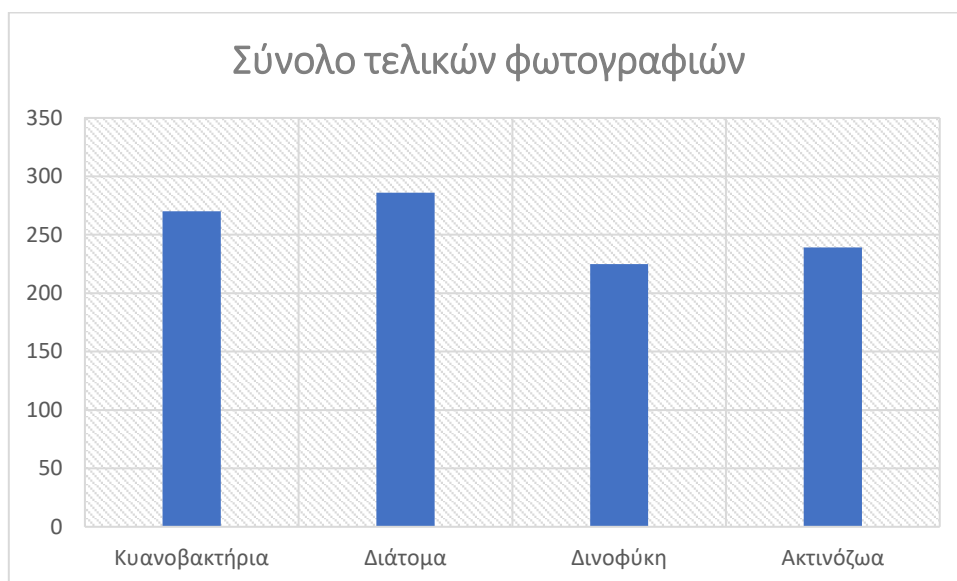
Εικόνα 4-1: Κατανομή των ακατέργαστων φωτογραφιών.

4.1.3 Επαύξηση των δεδομένων για εμπλουτισμό του συνόλου

Η επαύξηση των δεδομένων είναι χρήσιμη για την καλύτερη εκπαίδευση του νευρωνικού δικτύου και μπορεί να πραγματοποιηθεί είτε στο τελικό πρόγραμμα όπου δημιουργείται και το μοντέλο, είτε να γίνει στο στάδιο της προ-επεξεργασίας. Επιλέχθηκε να γίνει και στα δύο ώστε τα δεδομένα να τροφοδοτηθούν ήδη επαυξημένα, ως έναν βαθμό, στο τελικό πρόγραμμα.

Για την επαύξηση των δεδομένων δημιουργήθηκε ένα απλό πρόγραμμα το οποίο λαμβάνει τις εικόνες, και τυχαία επιλέγει μερικές ώστε να τις περιστρέψει κατά 90, 180 ή 270 μοίρες. Έτσι εμπλουτίζεται το σύνολο των διαθέσιμων φωτογραφιών κατά αρκετές εικόνες. Στο τελικό σύνολο των δεδομένων υπάρχουν 1020 εικόνες για τις 4 κατηγορίες πλαγκτού που μελετάει το πρόγραμμα.

Στην παρακάτω εικόνα φαίνεται η κατανομή των δεδομένων του τελικού συνόλου δεδομένων, που χρησιμοποιήθηκε στην εκπαίδευση του νευρωνικού δικτύου.



Εικόνα 4-2: Κατανομή φωτογραφιών του τελικού συνόλου δεδομένων.

4.2 Σύγκριση αρχιτεκτονικών σχεδίασης των νευρωνικών δικτύων

Χρησιμοποιήθηκε το Jupyter Notebook του Google Colab για την υλοποίηση των μοντέλων, καθώς εκμεταλλεύεται τον παράλληλο υπολογισμό μέσω απομακρυσμένης χρήσης κάρτας γραφικών (GPU) και επιτρέπει τη γραφή κώδικα στη γλώσσα Python μέσα σε πρόγραμμα περιήγησης. Επίσης διαθέτει προ-εγκατεστημένες αρκετές χρήσιμες βιβλιοθήκες που χρησιμοποιούνται στη δημιουργία προγραμμάτων μηχανικής μάθησης.

Για την υλοποίηση των μοντέλων, έγινε χρήση του Keras API, το οποίο είναι ένα υψηλού επιπέδου API για νευρωνικά δίκτυα σχεδιασμένο για ευκολία στη χρήση αλλά και απλοποιημένη δημιουργία πρωτοτύπων. Λειτουργεί με frameworks για μηχανική μάθηση όπως το TensorFlow, το οποίο και χρησιμοποιήθηκε στη συγκεκριμένη περίπτωση.

Το σύνολο δεδομένων μεταφορτώθηκε στο περιβάλλον του Google Drive και από εκεί φορτώθηκε στο Google Colab. Τα δεδομένα χωρίστηκαν σε 3 κατηγορίες: Σε δεδομένα εκπαίδευσης, δοκιμών και επαλήθευσης.

Επειδή το Keras χρησιμοποιεί δέσμες δεδομένων (batches), όταν γίνεται εισαγωγή δεδομένων εικόνων, δηλαδή δύο διαστάσεων, τα δεδομένα έχουν την παρακάτω μορφή (samples,size_1,size_2,channels). Όμως μία εικόνα έχει τη μορφή (size_1,size_2,channels) άρα για να δημιουργηθούν δέσμες δεδομένων από αυτά, πρέπει να γίνει τροποποίηση των εικόνων με την εισαγωγή μίας νέας διάστασης στον πίνακα. Για

αυτό, γίνεται η χρήση της συνάρτησης *preprocess_input*, μέσα από το MobilenetV2, η οποία κανονικοποιεί και προσαρμόζει τις εικόνες στη μορφή που απαιτεί το μοντέλο. Συνήθως τα περισσότερα μοντέλα που χρησιμοποιούν εικόνες κανονικοποιούν τις τιμές από 0 σε 1, ή από -1 σε 1. Το MobilenetV2 κανονικοποιεί τις τιμές από -1 σε 1 [17].

Αφού χωριστούν και κανονικοποιηθούν τα δεδομένα, το επόμενο βήμα είναι ο καθορισμός των τελευταίων λεπτομερειών, όπως η χρωματική κλίμακα και το μέγεθος των δεσμών δεδομένων (batches).

4.2.1 Μοντέλο πρόβλεψης χωρίς τη χρήση συνελκτικών επιπέδων

Αφού τα δεδομένα και οι παράμετροι του μοντέλου ρυθμίστηκαν, το επόμενο στάδιο είναι η δημιουργία του νευρωνικού. Η πρώτη αρχιτεκτονική μοντέλου που δοκιμάστηκε είναι αρκετά απλή. Αποτελείται από τρία πλήρως συνδεδεμένα επίπεδα (dense), τα πρώτα δύο χρησιμοποιούν τη συνάρτηση ενεργοποίησης Relu, ενώ το τελευταίο τη softmax.

Ανάμεσα από αυτά υπάρχουν επίπεδα dropout. Γίνεται η χρήση τους για την αποφυγή της υπερπροσαρμογής, καθώς η λειτουργία τους είναι να αγνοούν τυχαία ορισμένους αριθμούς εισόδων. Οι εισοδοί που δεν έχουν οριστεί ίσοι με το μηδέν, κλιμακώνονται με τέτοιο τρόπο ώστε το συνολικό άθροισμα εισόδων να παραμένει αμετάβλητο.

Layer (type)	Output Shape	Param #
flatten_3 (Flatten)	(None, 59536)	0
dense_9 (Dense)	(None, 128)	7620736
dropout_6 (Dropout)	(None, 128)	0
dense_10 (Dense)	(None, 128)	16512
dropout_7 (Dropout)	(None, 128)	0
dense_11 (Dense)	(None, 4)	516

Εικόνα 4-3: Αρχιτεκτονική απλού νευρωνικού δικτύου ταξινόμησης.

Δεδομένου του μικρού δείγματος των φωτογραφιών, αλλά και της χαμηλής τους ανάλυσης, συνθήκες που αυξάνουν τη δυσκολία του προβλήματος της ταξινόμησής τους, το μοντέλο δεν είχε τα επιθυμητά αποτελέσματα. Κατά τη διάρκεια της εκπαίδευσης του,

η ακρίβεια του δεν έφτασε ποτέ παραπάνω από το 32% και στην εφαρμογή του στα δοκιμαστικά δεδομένα, η ακρίβεια του μοντέλου ήταν 29%.

Έτσι κρίθηκε αναγκαία η δημιουργία μοντέλων ταξινόμησης χρησιμοποιώντας πιο περίπλοκες τεχνικές.

4.2.2 Μοντέλα πρόβλεψης με 3 ή 4 επίπεδα συνέλιξης

Τα ακόλουθα μοντέλα που δοκιμάστηκαν αποτελούνται από μία σειρά επιπέδων συνέλιξης και επιπέδων συγκέντρωσης μεγίστου. Τα επίπεδα αυτά εναλλάσσονται διαδοχικά. Στα τελευταία επίπεδα του νευρωνικού, υπάρχουν δύο πλήρως συνδεδεμένα επίπεδα. Η συνάρτηση ενεργοποίησης των επιπέδων συνέλιξης και του ενός εκ των πλήρως συνδεδεμένων επιπέδων, είναι η Relu. Ενώ στο επίπεδο εξόδου, χρησιμοποιείται η συνάρτηση ενεργοποίησης softmax.

Η συνάρτηση σφάλματος που χρησιμοποιήθηκε ήταν η κατηγορική διασταυρούμενη εντροπία. Είναι μία αρκετά συνηθισμένη συνάρτηση σφάλματος που χρησιμοποιείται σε μοντέλα ταξινόμησης. Σε αυτά τα μοντέλα, ένα αποτέλεσμα μπορεί να ανήκει σε μονάχα μία κατηγορία έτσι το μοντέλο πρέπει να αποφασίσει βασισμένο σε πιθανότητες σε ποια κατηγορία να κατατάζει το κάθε αποτέλεσμα. Η συνάρτηση της κατηγορικής διασταυρούμενης εντροπίας έχει σχεδιαστεί για την ποσοτικοποίηση της διαφοράς μεταξύ κατανομών πιθανοτήτων.

Η εκπαίδευση του μοντέλου έγινε σε 10 εποχές μέγιστο. Αυτό σημαίνει ότι θα γίνουν 10 επαναλήψεις όπου θα μεταλλάσσονται τα βάρη με τη βοήθεια του βελτιστοποιητή, για να μάθει το εκάστοτε μοντέλο να αναγνωρίζει και να κατηγοριοποιεί σωστά τις εικόνες του πλαγκτού στην κατάλληλη κατηγορία. Για την αποφυγή αχρείαστης εκπαίδευσης, χρησιμοποιήθηκε η επανάκληση (callback). Αυτό σημαίνει ότι η εκπαίδευση του μοντέλου διακόπτεται σε περίπτωση που κατά τη διάρκεια της εκπαίδευσης δεν βελτιώνεται η ακρίβεια για περισσότερο από έναν ορισμένο αριθμό εποχών.

Layer (type)	Output Shape	Param #	Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 242, 242, 32)	320	conv2d (Conv2D)	(None, 242, 242, 32)	896
max_pooling2d_4 (MaxPooling2D)	(None, 121, 121, 32)	0	max_pooling2d (MaxPooling2D)	(None, 121, 121, 32)	0
conv2d_5 (Conv2D)	(None, 119, 119, 128)	36992	conv2d_1 (Conv2D)	(None, 119, 119, 64)	18496
max_pooling2d_5 (MaxPooling2D)	(None, 59, 59, 128)	0	max_pooling2d_1 (MaxPooling2D)	(None, 59, 59, 64)	0
conv2d_6 (Conv2D)	(None, 57, 57, 128)	147584	conv2d_2 (Conv2D)	(None, 57, 57, 128)	73856
max_pooling2d_6 (MaxPooling2D)	(None, 28, 28, 128)	0	max_pooling2d_2 (MaxPooling2D)	(None, 28, 28, 128)	0
flatten_5 (Flatten)	(None, 100352)	0	conv2d_3 (Conv2D)	(None, 26, 26, 128)	147584
dropout_9 (Dropout)	(None, 100352)	0	max_pooling2d_3 (MaxPooling2D)	(None, 13, 13, 128)	0
dense_14 (Dense)	(None, 64)	6422592	flatten_4 (Flatten)	(None, 21632)	0
dense_15 (Dense)	(None, 4)	260	dropout_8 (Dropout)	(None, 21632)	0
			dense_12 (Dense)	(None, 64)	1384512
			dense_13 (Dense)	(None, 4)	260

Εικόνα 4-4: Αρχιτεκτονική νευρωνικών δικτύων με χρήση συνελκτικών επιπέδων.

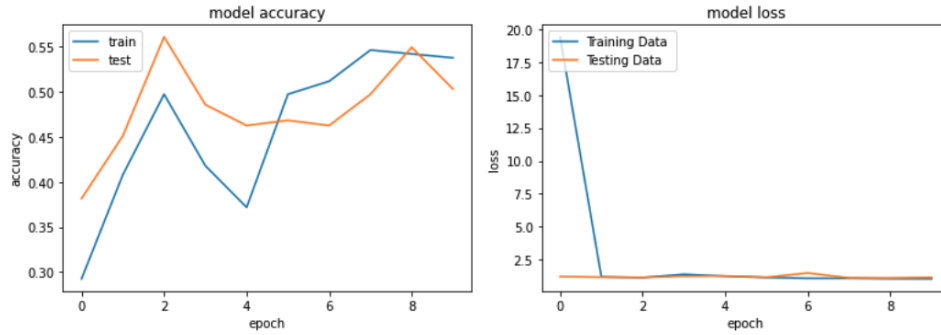
Αριστερά με τρία, ενώ δεξιά με τέσσερα επίπεδα.

Η κύρια διαφορά μεταξύ των δύο αρχιτεκτονικών, αν εξαιρεθεί το ένα επιπλέον επίπεδο συνέλιξης αλλά και το επίπεδο συγκέντρωσης μεγίστου, βρίσκεται στον αριθμό των παραμέτρων που θα τεθούν προς εκπαίδευση. Στην περίπτωση με τα τρία επίπεδα συνέλιξης, οι παράμετροι είναι 6,608,324 ενώ στη δεύτερη περίπτωση 1,625,604. Όπως είχε αναφερθεί στο προηγούμενο κεφάλαιο, τα επίπεδα συνέλιξης και συγκέντρωσης μεγίστου, μειώνουν τη διαστατικότητα του προβλήματος της ταξινόμησης. Παρατηρούμε παραπάνω από 4 φορές μείωση στις παραμέτρους εκπαίδευσης.

Στην πρώτη περίπτωση, η εκπαίδευση έφτασε να έχει ακρίβεια πλησίον του 53%, ενώ στα δοκιμαστικά δεδομένα η ακρίβεια της πρόβλεψης ήταν περίπου 50%. Είναι μία σαφής βελτίωση από το προηγούμενη, και πιο απλοϊκή, αρχιτεκτονική, αλλά ακόμη δεν λειτουργεί ικανοποιητικά.

Κατά την εκπαίδευση του νευρωνικού με τα τέσσερα επίπεδα συνέλιξης, η ακρίβεια του εκπαιδευτικού συνόλου έφτασε το 84% αλλά και η ακρίβεια στα δοκιμαστικά δεδομένα βελτιώθηκε σημαντικά. Έχοντας επιτυχία στις προβλέψεις της τάξης του 65%.

Παρακάτω φαίνονται οι πίνακες των μετρικών των δύο μοντέλων.



Εικόνα 4-5: Μετρικές μοντέλου με τρία επίπεδα συνέλιξης



Εικόνα 4-6: Μετρικές μοντέλου με τέσσερα επίπεδα συνέλιξης

Από τα αποτελέσματα του μοντέλου, αλλά και τις παραπάνω εικόνες, παρατηρείται μία αδυναμία ορθής αντιμετώπισης του προβλήματος της ταξινόμησης του συγκεκριμένου συνόλου δεδομένων, παρότι χρησιμοποιήθηκαν αρκετές μέθοδοι για να αποφευχθεί η υπερπροσαρμογή κατά τη διάρκεια της εκπαίδευσης. Συνεπώς, θα πρέπει να διερευνηθούν πιο περίπλοκοι τρόποι για να αντιμετωπιστεί το πρόβλημα.

4.2.3 Μοντέλο πρόβλεψης με τη χρήση του *EfficientNet*

Το *EfficientNet* είναι μια αρχιτεκτονική που χρησιμοποιεί επίπεδα συνέλιξης, και μία μέθοδος κλιμάκωσης νευρωνικών δικτύων, που κλιμακώνει ομοιόμορφα όλες τις διαστάσεις του βάθους και πλάτους χρησιμοποιώντας έναν σύνθετο συντελεστή. Σε αντίθεση με τις συμβατικές αρχιτεκτονικές, που η κλιμάκωση αυτών των παραγόντων γίνεται αυθαίρετα, η μέθοδος που χρησιμοποιεί το δίκτυο *EfficientNet*, τις κλιμακώνει ομοιόμορφα με ένα σύνολο σταθερών συντελεστών κλιμάκωσης. Η μέθοδος αυτή βασίζεται στο σκεπτικό ότι εάν η εικόνα εισόδου είναι μεγαλύτερη, τότε το δίκτυο χρειάζεται περισσότερα στρώματα για να αυξήσει το πεδίο λήψης και περισσότερα

κανάλια για να καταγράψει πιο λεπτά μοτίβα στη μεγαλύτερη εικόνα. Το EfficientNet παρέχει καλύτερη ακρίβεια συγκριτικά με λοιπές μεθόδους, αλλά και βελτιώνει την απόδοση των μοντέλων μειώνοντας τις παραμέτρους και τα FLOPS (floating point operations per second) [20].

Δημιουργήθηκε συνεπώς, ένα μοντέλο που αξιοποιεί το EfficientNet. Η εκπαίδευση έγινε και πάλι σε 10 εποχές, όπως και στα προηγούμενα μοντέλα.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 224, 224, 3)]	0
efficientnetb1 (Functional)	(None, 7, 7, 1280)	6575239
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dense (Dense)	(None, 56)	71736
dropout (Dropout)	(None, 56)	0
dense_1 (Dense)	(None, 4)	228

Εικόνα 4-7: Δημιουργία του μοντέλου που χρησιμοποιεί EfficientNet.

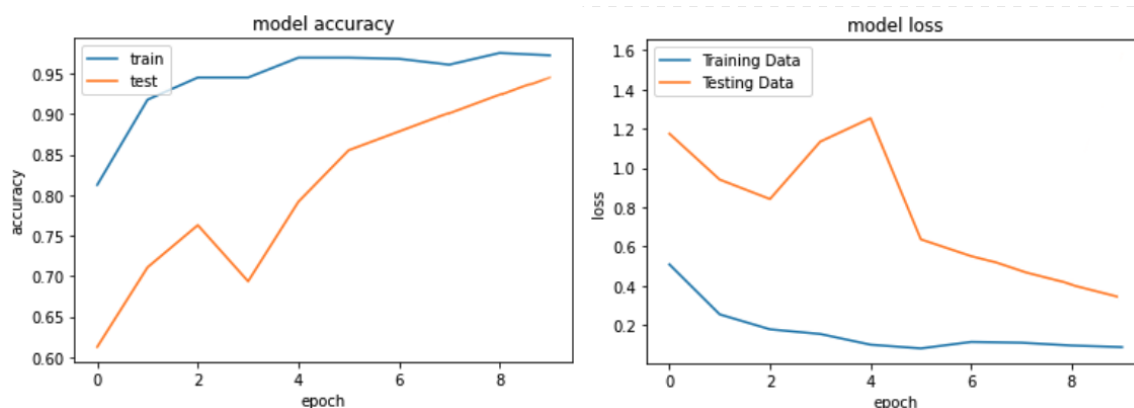
Η αρχιτεκτονική του EfficientNet είναι η ακόλουθη:

Στάδια υλοποίησης	Επίπεδο (Τελεστής)	Ανάλυση εισόδου	Αριθμός καναλιών	Αριθμός επιπέδων
1	Conv 3x3	224 x 224	32	1
2	MBConv1, 3x3	112 x 112	16	1
3	MBConv6, 3x3	112 x 112	24	2
4	MBConv6, 5x5	56 x 56	40	2
5	MBConv6, 3x3	28 x 28	80	3
6	MBConv6, 5x5	14 x 14	112	3
7	MBConv6, 5x5	14 x 14	192	4
8	MBConv6, 3x3	7 x 7	320	1
9	Conv 1x1, Pooling, FC	7 x 7	1280	1

Εικόνα 4-8: Αρχιτεκτονική EfficientNet.

Αποτελείται από ένα συνελκτικό επίπεδο 3x3 στην αρχή και ένα 1x1 πριν την εφαρμογή του πλήρως συνδεδεμένων επιπέδων, στο τέλος. Ανάμεσα τους υπάρχουν ανεστραμμένα υπολειμματικά μπλοκ, όπως αυτά που αναφέρθηκαν το προηγούμενο κεφάλαιο.

Η εκπαίδευση του μοντέλου έδωσε πολύ καλύτερα αποτελέσματα σε σχέση με τα προηγούμενα μοντέλα. Συγκεκριμένα, κατά την εκπαίδευση η ακρίβεια έφτασε το 97% ενώ στην πρόβλεψη των δοκιμαστικών δεδομένων υπήρξε ακρίβεια περίπου 93%. Στην Εικόνα 4-9 φαίνονται και οι αντίστοιχες μετρικές του μοντέλου.



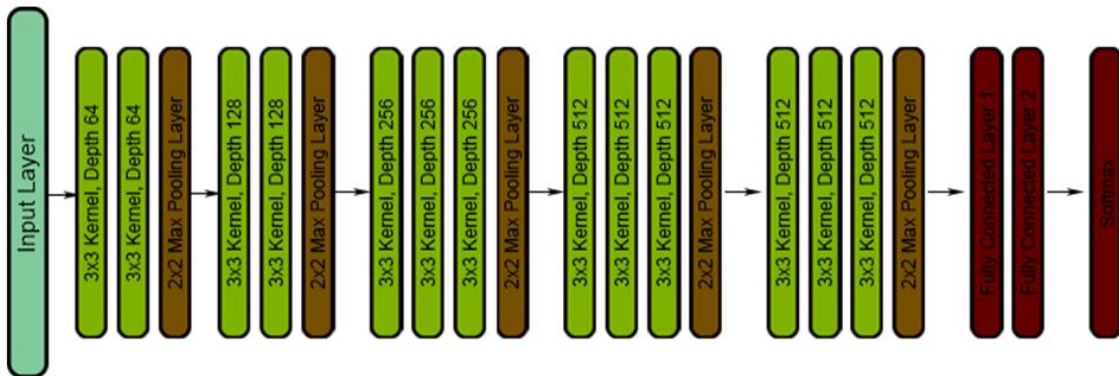
Εικόνα 4-9: Μετρικές μοντέλου με χρήση EfficientNet

Η συγκεκριμένη υλοποίηση φαίνεται να είναι ικανοποιητικά ακριβής εν συναρτήσει με το σύνολο δεδομένων στο οποίο εκπαιδεύτηκε. Περαιτέρω θα μελετήσουμε μία ακόμη αρχιτεκτονική, που χρησιμοποιεί ένα ήδη εκπαιδευμένο δίκτυο και θα γίνει χρήση της τεχνικής μεταφοράς γνώσης μεταξύ δικτύων.

4.2.4 Μοντέλο πρόβλεψης με τη χρήση VGG-16

Το VGG αναπτύχθηκε από τους Zisserman, Simoyan από το πανεπιστήμιο της Οξφόρδης. Είναι μία αρχιτεκτονική βαθιάς μάθησης που χρησιμοποιεί ένα συνελκτικό νευρωνικό δίκτυο με πολλαπλά επίπεδα. Ο αριθμός 16, αναφέρεται στον αριθμό των στρωμάτων που αποτελείται το δίκτυο. Δύο από τα πιο γνωστά είναι το VGG-16 και το VGG-19. Παραμένει ακόμα μία από τις πιο δημοφιλείς αρχιτεκτονικές αναγνώρισης εικόνων [21]. Το δίκτυο του VGG, είναι κατασκευασμένο από πολύ μικρά συνελκτικά φίλτρα. Το VGG-16 αποτελείται από 13 επίπεδα συνέλιξης, 2 πλήρως συνδεδεμένα επίπεδα και ένα επίπεδο ταξινόμησης, με τη συνάρτηση ενεργοποίησης softmax.

Τα κρυφά επίπεδα, χρησιμοποιούν τη συνάρτηση ενεργοποίησης Relu. Συμπερασματικά, είναι ένα αρκετά μεγάλο δίκτυο όπως φαίνεται και στην παραπάνω εικόνα, αλλά όχι τόσο μεγάλο όσο το δίκτυο του EfficientNet.

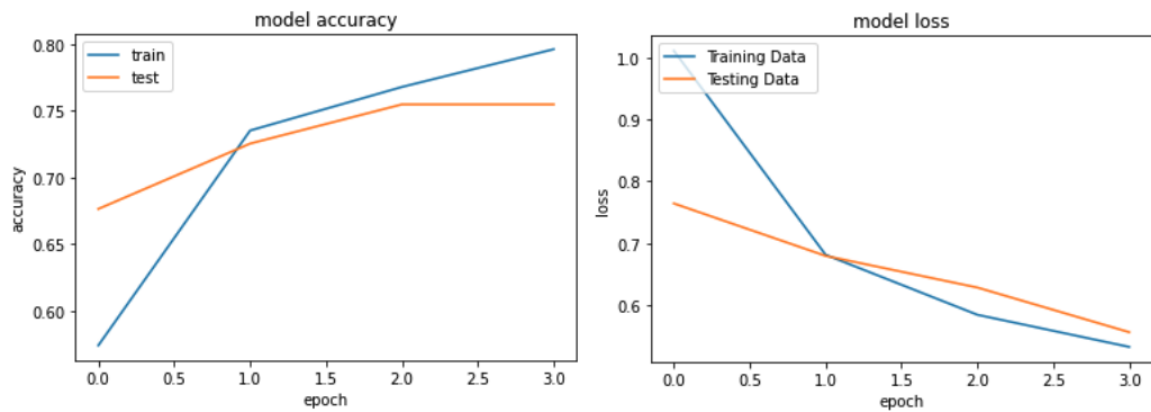


Εικόνα 4-10: Αρχιτεκτονική VGG-16

Το πολύ σημαντικό πλεονέκτημα του δικτύου αυτού, είναι το γεγονός ότι έχει εκπαιδευτεί στο σύνολο δεδομένων του ImageNet. Το ImageNet, είναι ένα ερευνητικό έργο για την ανάπτυξη μίας μεγάλης βάσης δεδομένων ταξινομημένων εικόνων. Οι εικόνες είναι εκατομμύρια, και οι κατηγορίες πολλές χιλιάδες. Καθώς το μοντέλο έχει ήδη εκπαιδευτεί σε ένα τόσο μεγάλο σύνολο δεδομένων, έχει μάθει να αναγνωρίζει χαρακτηριστικά χαμηλού επιπέδου όπως άκρες και σχήματα. Αυτή του η ιδιότητα μπορεί να χρησιμοποιηθεί για να επιτρέψει τη μεταφορά γνώσης και να λειτουργήσει ως εξαγωγέας χαρακτηριστικών για νέες εικόνες σε διάφορα προβλήματα υπολογιστικής όρασης, όπως αυτό της ταξινόμησης. Παρότι οι νέες εικόνες μπορεί να είναι εντελώς διαφορετικές από αυτές που έχει ήδη εκπαιδευτεί το δίκτυο, το προεκπαιδευμένο μοντέλο μπορεί να εξάγει τα χαρακτηριστικά από αυτές, βάση της προηγούμενης του εκπαίδευσης [22].

Αφού φορτώσουμε το προεκπαιδευμένο μοντέλο του VGG στο πρόγραμμα, πρέπει να συνδυάσουμε τα δύο μοντέλα. Το τελευταίο επίπεδο ενός συνελκτικού δικτύου είναι συνήθως ένα πλήρως συνδεδεμένο επίπεδο που έχει ίδιο αριθμό νευρώνων όσο και το πλήθος των προς ταξινόμηση κλάσεων δεδομένων. Στη συγκεκριμένη περίπτωση τέσσερις. Έτσι πρέπει να διαμορφωθεί το τελευταίο επίπεδο ώστε να έχει ως είσοδο την έξοδο του προ-εκπαιδευμένου μοντέλου ενώ σαν έξοδο πρέπει να έχει ίδιο αριθμό νευρώνων όσες και οι κλάσεις.

Υλοποιώντας και την παραπάνω λειτουργία, μπορεί να ξεκινήσει η εκπαίδευση του μοντέλου. Η εκπαίδευση επιλέχθηκε να γίνει σε 4 εποχές. Ο χρόνος εκτέλεσης κάθε εποχής ήταν σημαντικά μεγαλύτερος από τον αντίστοιχο χρόνο των προηγούμενων μοντέλων που δημιουργήθηκαν. Η ακρίβεια κατά τη διάρκεια της εκπαίδευσης πλησίασε το 80%, ενώ στα δοκιμαστικά δεδομένα η ακρίβεια έφτασε το 75%. Παρακάτω φαίνονται και οι μετρικές του μοντέλου.



Εικόνα 4-11: Μετρικές μοντέλου με χρήση VGG-16

5 Σύνοψη και Συμπεράσματα

Για την εκπόνηση της διπλωματικής εργασίας χρειάστηκαν αρκετά δεδομένα φωτογραφιών πλαγκτού. Οι φωτογραφίες ήταν σε ακατέργαστη μορφή, οπότε και υπήρξε η ανάγκη της επεξεργασίας και επιμέλειά τους ώστε να μπορούν να είναι χρηστικές. Τα δεδομένα επεξεργαστήκαν και χωριστήκαν σε υποφακέλους ταξινομημένους ανά οικογένεια πλαγκτού.

Χρησιμοποιήθηκε φίλτρο αλλαγής της χρωματικής κλίμακας των εικόνων (από έγχρωμες σε ασπρόμαυρες) γιατί το διαφορετικό χρωματικό προφίλ μεταξύ των φωτογραφιών ήταν σημαντικός παράγοντας εσφαλμένης εκπαίδευσης του μοντέλου. Περαιτέρω, εφαρμόστηκαν φίλτρα εξομάλυνσης της εικόνας σε αρκετά δείγματα. Συγκεκριμένα έγινε η χρήση του Γκαουσιανού φίλτρου θόλωσης.

Για τη διαλογή των καλύτερων δειγμάτων, έγινε χρήση της μεθοδολογίας του αλγορίθμου εύρεσης άκρων του Canny. Ο αλγόριθμος λειτούργησε σα δείκτης ώστε να αφαιρεθούν αρκετές εικόνες του συνόλου φωτογραφιών, που περιέχουν αρκετό θόρυβο ώστε να μην μπορεί να γίνει σωστή επεξεργασία τους.

Προτού τα δεδομένα τροφοδοτηθούν, προς εκπαίδευση, στο εκάστοτε δίκτυο, πραγματοποιήθηκε περαιτέρω επεξεργασία και ενίσχυση των φωτογραφιών. Κυρίως χρησιμοποιώντας τυχαίες περιστροφές και μεγέθυνση των εικόνων. Ο σκοπός αυτής της διεργασίας ήταν η αποφυγή του προβλήματος της υπερπροσαρμογής, που συνιστά σημαντικό παράγοντα εσφαλμένης εκπαίδευσης νευρωνικών δικτύων.

Στο τελικό στάδιο της εργασίας μελετήθηκαν πέντε διαφορετικές αρχιτεκτονικές σχεδίασης μοντέλων ταξινόμησης με τη χρήση νευρωνικών δικτύων.

Αφού μελετήθηκε η αντίστοιχη βιβλιογραφία, έγινε επιλογή των αρχιτεκτονικών σχεδίασης νευρωνικών δικτύων. Ο στόχος ήταν η μέγιστη ακρίβεια πρόβλεψης με το μικρότερο κόστος σε υπολογιστικούς πόρους.

Παρουσιάστηκαν αρκετά σφάλματα και αποτυχίες σε αυτό το στάδιο. Χρειάστηκε να αλλάξουν αρκετές παράμετροι του προγράμματος και της κάθε αρχιτεκτονικής.

Μελετήθηκαν πέντε διαφορετικές προσεγγίσεις για το πρόβλημα της ταξινόμησης. Η πρώτη, ήταν ένα απλοϊκό μοντέλο που χρησιμοποιεί πλήρως συνδεδεμένα επίπεδα. Το μοντέλο είχε σημαντικά χαμηλή ακρίβεια. Δεδομένου του γεγονότος ότι η πιθανότητα τυχαία η επιλογή κάποιας οικογένειας πλαγκτού να είναι σωστή προσεγγίζει το 25%, το

πρώτο μοντέλο απέτυχε πλήρως στο πρόβλημα ταξινόμησης, έχοντας ακρίβεια λίγο κάτω από 30%.

Οι επόμενες δύο περιπτώσεις που μελετήθηκαν είχαν να κάνουν με τη χρήση απλών αρχιτεκτονικών χρησιμοποιώντας συνελκτικά επίπεδα. Τα συνελκτικά νευρωνικά δίκτυα είναι πολύ ικανά στην αντιμετώπιση προβλημάτων ταξινόμησης εικόνων [12], για αυτό και επιλέχθηκαν οι συγκεκριμένες αρχιτεκτονικές.

Η ακρίβεια των δικτύων που χρησιμοποιούν τρία και τέσσερα επίπεδα συνέλιξης αντίστοιχα, ήταν αυξημένη αρκετά σε σχέση με το απλό δίκτυο που έκανε χρήση μόνο πλήρως συνδεδεμένων επιπέδων. Το δίκτυο που χρησιμοποιούσε τρία επίπεδα συνέλιξης είχε ακρίβεια κοντά στο 50%, ενώ το επίπεδο με τα τέσσερα επίπεδα πλησίον του 65%. Είναι μία σαφής βελτίωση της ποιότητας του ταξινομητή αλλά τα αποτελέσματα και πάλι δεν είναι αρκετά ακριβή.

Η επόμενη αρχιτεκτονική που μελετήθηκε ήταν αυτή του EfficientNet το οποίο χρησιμοποιεί συνελκτικά επίπεδα με πρωτότυπες τεχνικές σύνδεσης και κλιμάκωσής τους [17],[20]. Το συγκεκριμένο δίκτυο εμφάνισε τη μεγαλύτερη ακρίβεια συγκριτικά με τα υπόλοιπα που μελετήθηκαν στα πλαίσια της εργασίας. Η ακρίβεια των προβλέψεών του άγγιξε το 93%. Πολύ ικανοποιητικός αριθμός για μοντέλα ταξινόμησης.

Το τελευταίο δίκτυο που χρησιμοποιήθηκε ήταν το VGG-16. Σε αυτήν την περίπτωση, έγινε χρήση της τεχνικής μεταφοράς γνώσης από το ήδη εκπαιδευμένο δίκτυο του VGG-16. Είναι μία συνηθισμένη πρακτική που ακολουθείται σε προβλήματα ταξινόμησης καθώς παρουσιάζει ικανοποιητικά αποτελέσματα εκπαίδευσης και προβλέψεων [22]. Το VGG-16, εμφάνισε ακρίβεια κοντά στο 75%. Περισσότερο από τα απλά συνελκτικά δίκτυα αλλά λιγότερο από αυτό EfficientNet. Συγκρίνοντας τους χρόνους εκπαίδευσης των δύο δικτύων με τη μεγαλύτερη ακρίβεια, το VGG-16 ήταν αρκετά πιο αργό από τους αντίστοιχους χρόνους του EfficientNet.

Όσο εξελίσσεται η επιστήμη της πληροφορικής, ολοένα και περισσότερες μέθοδοι αναπτύσσονται πάνω στον τομέα της μηχανικής μάθησης που κάνουν ευκολότερη την επίλυση αρκετών προβλημάτων. Αυτό φαίνεται ξεκάθαρα και από τη μεγάλη ακρίβεια που επέδειξε το πιο πρόσφατο δίκτυο (EfficientNet) σε σύγκριση με τα υπόλοιπα δίκτυα. Ακόμα και χρησιμοποιώντας ένα πολύ μικρό και θορυβώδες σύνολο δεδομένων, το πρόβλημα της ταξινόμησης εικόνων αντιμετωπίστηκε με αρκετή επιτυχία.

5.1 Όρια και περιορισμοί της έρευνας

Τα δεδομένα που χρησιμοποιήθηκαν ήταν περιορισμένα σε πλήθος το οποίο, όσον αφορά μοντέλα τεχνητής νοημοσύνης που ο σκοπός τους είναι η ταξινόμηση φωτογραφιών, είναι πολύ σημαντικό πρόβλημα. Στην περίπτωση που υπήρχε μεγαλύτερος όγκος δεδομένων, θα μπορούσε το πρόγραμμα να έχει περισσότερη ακρίβεια, καθώς επίσης και θα ήταν πιο εύκολη η εκκαθάριση του συνόλου από αρκετές πιθανώς προβληματικές φωτογραφίες που αναγκαστικά συμπεριελήφθησαν.

5.2 Μελλοντικές Επεκτάσεις

Μία πιθανή μελλοντική επέκταση θα ήταν στην περίπτωση που συγκεντρωθούν αρκετά δεδομένα/φωτογραφίες και από άλλες οικογένειες πλαγκτού, να εκπαιδευτεί ένα δίκτυο που θα μπορεί να τις ταξινομήσει. Ίσως όχι μόνο βάση οικογένειας αλλά και είδους εφόσον υπάρξει ταξινόμηση και σε αυτό το επίπεδο. Εάν επιτευχθεί αυτό, θα μπορούσε στα πλαίσια μίας πιο ολοκληρωμένης εφαρμογής να συνεισφέρει στην εξαγωγή σημαντικών συμπερασμάτων ως προς τη βιοποικιλότητα και την υγεία των θαλάσσιων υδάτων και των λιμνών [23].

Παρόμοια τεχνική με αυτήν που χρησιμοποιήθηκε κατά την ταξινόμηση των δεδομένων πλαγκτού, θα μπορούσε να είναι χρήσιμη και σε άλλους τομείς, τόσο στο πεδίο της γεωλογίας όσο και στις υπόλοιπες θετικές επιστήμες.

Βιβλιογραφία

1. Colin S. Reynolds (2006). *The Ecology of Phytoplankton*.
2. David L. Kirchman (2009). *Microbial Ecology of the Oceans*.
3. W.F. Vincent (2009). *Encyclopedia of Inland Waters*.
4. Nakov, T., Theriot, E.C. and Alverson, A.J. (2014). Using phylogeny to model cell size evolution in marine and freshwater diatoms.
5. Steven L. Percival and David W. Williams (2014). *Microbiology of Waterborne Diseases (Second Edition)*.
6. Jennifer H. Wisecaver and Jeremiah D. Hackett (2011). Dinoflagellate genome evolution.
7. C. Lalli and R. Parson (1993). *An Introduction to Biological Oceanography*.
8. Jude, B.A.; Kirn, T.J.; Taylor R.K. (2005). A colonization factor links *Vibrio cholerae* environmental survival and human infection.
9. O. Roger Anderson (1983). *Radiolaria*.
10. Xia Wenchen, Zhang Ning, Wang Guoqing, and Youshitaka Kakuwa (2004). Pelagic radiolarian and conodont biozonation in the Permo-Triassic boundary interval and correlation to the Meishan GSSP.
11. J. McCarthy, Marvin L. Minsky, Nathaniel Rochester, and Claude E. Shannon (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence.
12. François Chollet (2017). *Deep Learning with Python*.
13. Yann Lecun, Yoshua Bengio and Geoffrey Hinton (2015). *Deep learning*.
14. Xavier Glorot and Yoshua Bengio (2010). Understanding the difficulty of training deep feedforward neural networks.
15. Ian Goodfellow, Yoshua Bengio and Aaron Courville (2016). *MIT Deep Learning Book*.
16. He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition.
17. Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov and Liang-Chieh Chen (2019). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*
18. Irwin Sobel (2014). *History and Definition of the Sobel Operator*.

19. Skampa, E., Triantaphyllou, M. V., Dimiza, M. D., Gogou, A., Malinverno, E., Stavrakakis, S., Panagiotopoulos, I.P., Parinos, C. and Baumann, K. H. (2019). Coupling plankton- sediment trap-surface sediment coccolithophore regime in the North Aegean Sea (NE Mediterranean). *Marine Micropaleontology*, 152, 101729.
20. Mingxing Tan and Quoc V. Le (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.
21. Karen Simonyan and Andrew Zisserman (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.
22. Srikanth Tammina (2019). Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images.
23. Donald M. Anderson, Allan D. Cembella and Gustaaf M. Hallegraeff (2017). Progress in understanding harmful algal blooms (HABs): Paradigm shifts and new technologies for research, monitoring and management.