

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών



Τεχνολογίες Κινητής και Ηλεκτρονικής Υγείας

Εξαμηνιαία Εργασία με θέμα
«Χρήση βαθιάς μάθησης και μηχανισμών προσοχής
για την διάγνωση της νόσου Alzheimer»

των φοιτητών:

Μυροπούλου Νεφέλη
Περίχαρος Φώτιος-Κωνσταντίνος
Πούλιος Δημήτριος
Τσούλιας Νικηφόρος

Διδάσκοντες: Δ. Κουτσούρης, Κ. Νικήτα, Γ. Ματσόπουλος, Π. Τσανάκας, Ρ. Πετροπούλου
Υπεύθυνος Άσκησης: Θεοφάνης Γανιτίδης

Περίληψη

Η παρούσα εργασία έχει ως στόχο την εφαρμογή μεθόδων βαθιάς μάθησης και μηχανισμών προσοχής για την διάγνωση της νόσου Alzheimer. Η νόσος Alzheimer αποτελεί μια εκφυλιστική νευρολογική ασθένεια η οποία επιφέρει την σταδιακή φθορά και την τελική ατροφία νευρικών κυττάρων του εγκεφάλου. Αποτελεί την πιο κοινή αιτία άνοιας, μιας διαταραχής με κύρια χαρακτηριστικά την απώλεια της μνήμης και άλλων γνωστικών ικανοτήτων του ατόμου, με σοβαρότητα ικανή να παρεμποδίσει καθημερινές δραστηριότητες του νοσούντα. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, 55 εκατομμύρια άνθρωποι πάσχουν από άνοια παγκοσμίως, με περίπου 10 εκατομμύρια νέες διαγνώσεις να γίνονται ετησίως, ενώ το 60-80% των περιπτώσεων οφείλονται στην νόσο Alzheimer. Ακόμα, η άνοια αποτελεί την έβδομη συχνότερη αιτία θανάτου και είναι ένας από τους κυριότερους παράγοντες μειωμένης λειτουργικότητας των γηραιότερων ατόμων παγκοσμίως. Με τη χρήση εγκεφαλικών MRI τομογραφιών και συνελκτικών δικτύων πολλές πρόσφατες έρευνες έχουν πετύχει υποσχόμενα αποτελέσματα στο πεδίο εντοπισμού νευρολογικών ασθενειών, όπως αυτής της νόσου Alzheimer. Αυτές οι μέθοδοι και δεδομένα θα αξιοποιηθούν και στην παρούσα εργασία, ενώ συγχρόνως θα γίνει η προσθήκη μηχανισμών προσοχής (attention mechanisms) για την βελτίωση της ερμηνευσιμότητας και της ακρίβειας του αναπτυσσόμενου μοντέλου.

Λέξεις Κλειδιά: συνελκτικά δίκτυα (CNN), νόσος Alzheimer, βαθιά μάθηση, μηχανισμοί προσοχής, Magnetic Resonance Imaging (MRI), ECA-Net, CBAM

Περιεχόμενα

1. Εισαγωγή - Θεωρητικό Υπόβαθρο	4
1.1 Νόσος Alzheimer's	4
1.1.1 Στατιστικά στοιχεία	4
1.1.2 Παθολογία της νόσου	4
1.1.3 Ιστορική αναδρομή	5
1.1.4 Αιτίες Εμφάνισης	6
1.1.5 Συμπτώματα	6
1.1.6 Θεραπεία	7
1.1.7 Διάγνωση	7
1.1.7.1 Γενικές μέθοδοι διάγνωσης	7
1.1.7.2 MRI Απεικονιστική τεχνική	7
1.2 Τεχνητά Νευρωνικά Δίκτυα - Συνελικτικά Νευρωνικά Δίκτυα	8
1.2.1 Τεχνητά Νευρωνικά Δίκτυα	8
1.2.1.1 Τεχνητός Νευρώνας - Δίκτυο Perceptron	9
1.2.1.2 Δίκτυα Multilayer Perceptron (MLP)	10
1.2.1.3 Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου	11
1.2.1.4 Μετρικές αξιολόγησης	11
1.2.2 Συνελικτικά Νευρωνικά Δίκτυα	12
1.2.2.1 Συνελικτικά Επίπεδα	12
1.2.2.2 Επίπεδα υποδειγματοληψίας (pooling layers)	14
1.2.2.3 Επίπεδα κανονικοποίησης: Batch Normalization και Dropout	14
1.3 Μηχανισμοί προσοχής	15
1.3.1 Εισαγωγή	15
1.3.2 Μηχανισμοί προσοχής σε Συνελικτικά Νευρωνικά Δίκτυα	15
1.3.2.1 Channel και Spatial Attention	15
1.3.2.2 ECA-Net	16
1.3.2.3 CBAM	17
2. Προτεινόμενη Μεθοδολογία	19
2.1 Σύνολο δεδομένων - Προεπεξεργασία	19
2.2 Σχεδιασμός και εκπαίδευση απλού συνελικτικού μοντέλου πρόβλεψης	20
2.3 Ενσωμάτωση μηχανισμών προσοχής στο υπάρχον συνελικτικό δίκτυο	23
3. Ανάλυση Αποτελεσμάτων	24
3.1 Αξιολόγηση μοντέλων με χρήση μετρικών σφάλματος	24
3.2 Αξιολόγηση μοντέλων με χρήση Gradient-weighted Class Activation Mapping	24
4. Συμπεράσματα - Συζήτηση	26
4.1 Σύνοψη	26
4.2 Μελλοντικές Προεκτάσεις	26
5. Βιβλιογραφία	27

1. Εισαγωγή - Θεωρητικό Υπόβαθρο

1.1 Νόσος Alzheimer's

1.1.1 Στατιστικά στοιχεία

Οι δημογραφικές αλλαγές του τελευταίου αιώνα στον δυτικό κόσμο, με τον όλο και γηραιότερο πληθυσμό, έχουν φέρει στο προσκήνιο την νόσο Alzheimer's ως μία νόσο με τεράστιο αντίκτυπο στην δημόσια υγεία. Ενδεικτικά στην Αμερική το 2010 είχαν υπολογιστεί περίπου 454.000 καινούριες περιπτώσεις Alzheimer's, ένα νούμερο που προβλέπεται να αυξηθεί στα 615.000 μέχρι το 2030 και να εκτιναχθεί τα 959.000 μέχρι το 2050 [1]. Παράλληλα το κόστος για την υγεία και την μακροχρόνια φροντίδα των ασθενών ανήκει στα πιο μεγάλα ανάμεσα στις χρόνιες ασθένειες, με τα συνολικά κόστη που σχετίζονται με τη νόσο να υπολογίζονται τα 236 δισεκατομμύρια δολάρια για την Αμερική το 2016 [2].

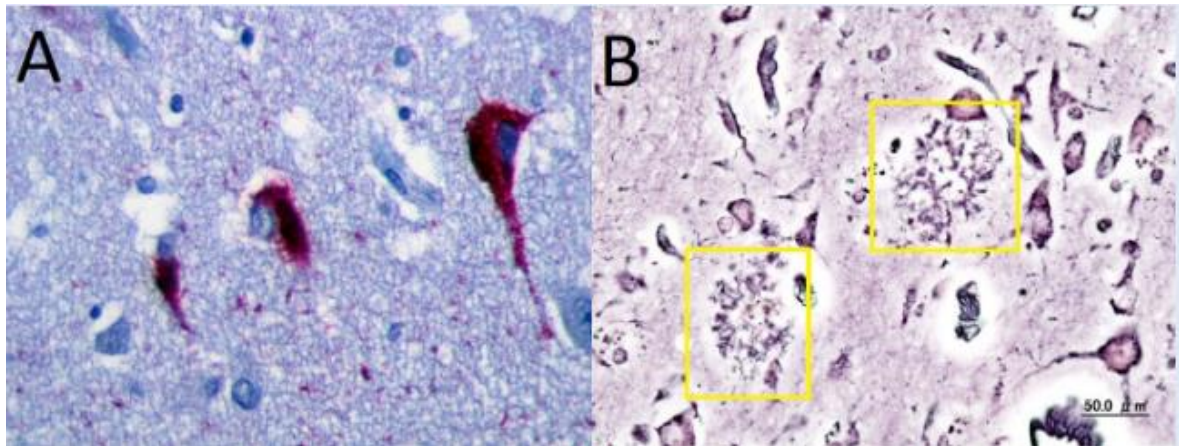
1.1.2 Παθολογία της νόσου

Η νόσος Alzheimer's αποτελεί μια εκφυλιστική νευρολογική ασθένεια, η οποία επιφέρει την σταδιακή φθορά και την τελική ατροφία νευρικών κυττάρων σε μέρη του εγκεφάλου, τα οποία συμμετέχουν στις γνωστικές λειτουργίες του ατόμου. Η φθορά και ατροφία αυτή συνεπακόλουθα επηρεάζει και άλλα μέρη του εγκεφάλου, υπεύθυνα για βασικές σωματικές λειτουργίες του ατόμου, όπως το περπάτημα ή η κατάποση. Επιπλέον, η νόσος Alzheimer's αποτελεί την πιο κοινή αιτία άνοιας, μιας διαταραχής με κύρια χαρακτηριστικά την απώλεια της μνήμης και άλλων γνωστικών ικανοτήτων του ατόμου. Η νόσος χαρακτηρίζεται θανατηφόρα, ενώ στα τελικά της στάδια, οι ασθενείς είναι κλινήρεις και χρήζοντες ολοήμερης φροντίδας. Όταν τα άτομα δυσκολεύονται να κινηθούν, είναι πιο ευάλωτα σε λοιμώξεις, συμπεριλαμβανομένης της πνευμονίας που είναι συχνά ένας παράγοντας που συμβάλλει στο θάνατο των ατόμων με νόσο του Alzheimer. [2].

Η νόσος Alzheimer's περιλαμβάνει μια τεράστια πολυπλοκότητα παθολογικών διεργασιών και αλληλεπιδράσεων πρωτεϊνών, οι οποίες έχουν ως αποτέλεσμα πολλαπλές διαταραχές της κυτταρικής λειτουργίας. Ειδικότερα, αποτελεί μία από τις 40 περίπου ταυτοποιημένες αμυλοειδώσεις, νόσους, δηλαδή, που έχουν ως κοινό παθολογικό στοιχείο την αντικανονική εναπόθεση ενδογενών και κανονικά διαλυτών πρωτεϊνών ως αμυλοειδή ινίδια σε διάφορους ιστούς. Στην περίπτωση του Alzheimer's το κύριο παθολογικό χαρακτηριστικό είναι ο σχηματισμός αμυλοειδών, αποτελούμενων κυρίως από το Amyloid-β (Aβ) πεπτιδίο, εκτός των νευρώνων, και νευροϊνικών συσσωρευμάτων (NFTs), σχηματισμένων από την φωσφορυλίωση της ταφ πρωτεΐνης, εντός των νευρώνων. Η συσσώρευση του Amyloid-β πεπτιδίου παρεμβαίνει στην επικοινωνία μεταξύ νευρώνων και συνευθύνεται για τον θάνατο των νευρικών κυττάρων. Τα NFTs μπλοκάρουν την μεταφορά θρεπτικών ουσιών και άλλων απαραίτητων μορίων προς το εσωτερικό των νευρώνων και συντελούν και αυτά στον θάνατο των νευρικών κυττάρων. Οι κακώσεις αυτές συμβαίνουν σε περιοχές του εγκεφάλου σχετιζόμενες με την μάθηση και τη μνήμη. Η όψη ενός εγκεφάλου ασθενούς σε προχωρημένο στάδιο της νόσου θα παρουσίαζε φλεγμονή, δραματική συρρίκνωση από την απώλεια κυττάρων και διάσπαρτα απομεινάρια λυμένων και λυόμενων κυττάρων [2][3][4].

Οι εγκεφαλικές αλλαγές που σχετίζονται με το Alzheimer μπορεί να ξεκινήσουν 20 ή περισσότερα χρόνια πριν εμφανιστούν τα συμπτώματα. Όταν συμβαίνουν οι αρχικές αλλαγές, ο

εγκέφαλος τις αντισταθμίζει, επιτρέποντας στα άτομα να συνεχίσουν να λειτουργούν κανονικά. Καθώς η νευρωνική βλάβη αυξάνεται, ο εγκέφαλος δεν μπορεί πλέον να αντισταθμίσει τις αλλαγές και τα άτομα ξεκινούν να παρουσιάζουν ήπια γνωστική έκπτωση. Αργότερα, η νευρωνική βλάβη είναι τόσο σημαντική που τα άτομα εμφανίζουν βαρύτερα συμπτώματα [2].



Εικόνα 1:

A: Με ερυθροϊώδες χρώμα αναδεικνύονται ενδοκυττάριας εναποθέσεις νευροϊνιδιακών σωρών (NFTs) στον ιππόκαμπο ασθενούς με νόσο Alzheimer's.

B: Απεικόνιση εξωκυττάριας πλακών αμυλοειδούς (εντός κίτρινων πλαισίων)

(Πηγή: [5])

1.1.3 Ιστορική αναδρομή

Η αρχή της αναζήτησης των νευρολογικών αιτιών της άνοιας εντοπίζεται ήδη από τον 19ο αιώνα [6]. Το 1822 ο Antoine Laurent Bayle προσδιόρισε αναγνωρίσιμες αλλαγές στην αραχνοειδή μήνιγγα του εγκεφάλου, θέτοντας έτσι τα θεμέλια για την συσχέτιση ψυχιατρικών συμπτωμάτων με παθολογικές αλλαγές στην φυσιολογία του εγκεφάλου [7]. Προς το τέλος του αιώνα ο Marce κατάφερε να συσχετίσει την φλοιική ατροφία, το “μαλάκωμα” του εγκεφαλικού ιστού και άλλα μακροσκοπικά ευρήματα με την γεροντική άνοια. Οι εξελίξεις στα μικροσκόπια τη δεκαετία του 1880 ήταν αυτές που έδωσαν στους γιατρούς την δυνατότητα να εντοπίσουν τρεις σημαντικότερες αλλαγές στον εγκέφαλο σε ασθενείς με άνοια. Αυτές ήταν η κυτταρική νέκρωση και αποσύνθεση, η πλάκα και τα νευροϊνίδια [7].

Η ιστορία της νόσου Alzheimer's, ως μίας σαφώς ορισμένης νευρολογικής νόσου, ξεκινάει στις αρχές του 20ου αιώνα, όταν ο ψυχίατρος και νευροπαθολόγος Αλοΐσιος Αλτσχάιμερ δημοσίευσε την έρευνα του "Περί μιας ιδιόμορφης νόσου του εγκεφαλικού φλοιού" το έτος 1907 [6][7]. Η έρευνά του περιγράφει την περίπτωση μιας 51 ετών γυναίκας από την Φρανκφούρτη, η οποία παρουσίαζε μία εξελισσόμενη αναπηρία στις γνωσιακές της λειτουργίες, παραισθήσεις και ψυχοκοινωνικές δυσκολίες. Η μελέτη που διεξήγαγε μετά τον θάνατο της ασθενούς στον εγκέφαλο της, έδωσε τις πρώτες απαντήσεις ως προς την παθολογία της νόσου. Αξίζει να σημειωθεί πως ο Α. Αλτσχάιμερ δεν θεωρούσε απαραίτητα πως τα ευρήματά του σκιαγραφούσαν μία καινούρια και αχαρτογράφητη νόσο, αλλά πως τα ευρήματα ήταν συμπτώματα γεροντικής ψύχωσης. Ο διαχωρισμός έγινε από τον Emil Kraepelin το 1910, όταν και δόθηκε το όνομα στην καινούρια αυτή ασθένεια [7].

Η έρευνα για το Alzheimer's δεν ξεκίνησε με γοργά βήματα μετά την ανακάλυψη της. Προβλήματα στον εντοπισμό σαφών συσχετίσεων μεταξύ συμπτωμάτων και εγκεφαλικών αλλοιώσεων ώθησαν Αμερικάνους ψυχιάτρους με επικεφαλής τον David Rothschild τη δεκαετία

του 1930 να επαναπροσδιορίσουν την αντίληψη του ιατρικού κόσμου ως προς την άνοια [7][8]. Δίνοντας έμφαση σε ψυχοκοινωνικά φαινόμενα ως γενεσιουργά αίτια της άνοιας αντί για εγκεφαλο-παθολογικά κατάφεραν να ανατρέψουν την επικρατούσα αντίληψη της άνοιας ως μίας φυσιολογικής επιδείνωσης του εγκεφάλου λόγω ασθένειας ή και γήρατος. Προτάσσοντας παράγοντες όπως την κοινωνική απομόνωση του ατόμου και συναισθηματικά τραύματα ως μηχανισμούς που οδηγούν στην εμφάνιση της νόσου κατάφεραν να τραβήξουν την προσοχή της επιστημονικής κοινότητας στην μελέτη και τη θεραπεία του Alzheimer [8].

Αν και οι έρευνες που έχουν διενεργηθεί έχουν αποκαλύψει πολλά για την ασθένεια, πολλά παραμένουν ακόμη άγνωστα: οι ακριβείς βιολογικές αλλαγές που την προκαλούν, το γιατί εξελίσσεται πιο γρήγορα σε μερικούς παρά σε άλλους, καθώς και το πώς μπορεί να προληφθεί, να επιβραδυνθεί ή να σταματήσει η ασθένεια [2].

1.1.4 Αιτίες Εμφάνισης

Γενετικοί παράγοντες μπορούν και έχουν αναζητηθεί για την πρόκληση της νόσου. Μεταλλάξεις στα γονίδια της πρόδρομης πρωτεΐνης του αμυλοειδούς (APP) και της προσενιλίνης 1 και 2 (Presenilin-1, Presenilin-2) έχουν συσχετιστεί με την ανάπτυξη Alzheimer's. Σε περίπτωση μετάλλαξης τουλάχιστον ενός εκ των δύο πρώτων γονιδίων η ανάπτυξη της νόσου είναι σίγουρη, ενώ σε περίπτωση μετάλλαξης του τρίτου σε σειρά γονιδίου, η πιθανότητα ανέρχεται στο 95%. Σε αντίθεση με τη συνήθη ηλικία εμφάνισης συμπτωμάτων, που είναι τα 65 χρόνια, άτομα με τις παραπάνω μεταλλάξεις μπορούν να εμφανίσουν συμπτώματα πριν τα 65 και ήδη από την ηλικία των 30. Ένα ακόμα γονίδιο, του οποίου η παρουσία αυξάνει την πιθανότητα εμφάνισης Alzheimer's είναι το APOE 34 γονίδιο. Συγκεκριμένα, η μορφή 34 του APOE γονιδίου καθιστά τη νόσηση 3 φορές πιο πιθανή. Συγχρόνως, μεγαλύτερη πιθανότητα εμφάνισης της νόσου έχουν και όσα άτομα έχουν συγγενείς που είναι ασθενείς και ειδικότερα αν είναι πρώτου βαθμού [2]. Υπάρχουν όμως και αιτίες εμφάνισης μη γενετικές. Οι εξωγενείς παράγοντες κινδύνου για τη νόσο Alzheimer's περιλαμβάνουν την μεγάλη ηλικία, την υψηλή χοληστερόλη, τον σακχαρώδη διαβήτη, το χαμηλό μορφωτικό επίπεδο, την ύπαρξη κρανιοεγκεφαλικής κάκωσης, την κατανάλωση δίαιτας με πολλές θερμίδες, την καθιστική ζωή και τις καρδιαγγειακές παθήσεις [4].

1.1.5 Συμπτώματα

Η συμπτωματολογία της νόσου διαφοροποιείται ανάλογα τον ασθενή. Ωστόσο, το πιο κοινό πρωταρχικό σύμπτωμα είναι η δυσκολία ενθύμησης νέων πληροφοριών, που σταδιακά επιδεινώνεται. Αυτό συμβαίνει διότι τα πρώτα επηρεαζόμενα από τη νόσο νευρικά κύτταρα είναι συνήθως σε περιοχές του εγκεφάλου σχετικές με τον σχηματισμό νέων μνημών. Συγχρόνως, φθορά και καταστροφή υφίστανται και νευρικά κύτταρα σε άλλες περιοχές του εγκεφάλου, με αποτέλεσμα την δυσκολία των ασθενών και σε άλλους τομείς της ζωής τους. Συχνά συμπτώματα της νόσου Alzheimer's αποτελούν η απώλεια μνήμης (με ισχυρή επίδραση στην καθημερινότητα του ατόμου), η δυσκολία στην επίλυση προβλημάτων και στην ολοκλήρωση προηγούμενων γνωστών εργασιών, η σύγχυση όσον αφορά στον τόπο και στον χρόνο που βρίσκεται το άτομο, η δυσκολία στην κατανόηση οπτικών εικόνων και χωρικών σχέσεων και αποστάσεων, η δυσκολία στον λόγο και στο γράψιμο, η κακή κρίση και ικανότητα λήψης αποφάσεων, ο ανήσυχος ύπνος, η κοινωνική αποξένωση, το άγχος, η κατάθλιψη και ενδεχομένως και παραισθήσεις και παράνοια. Ο ρυθμός με τον οποίο τα συμπτώματα εξελίσσονται από ήπια σε μέτρια και σε σοβαρά ποικίλει από άτομο σε άτομο [2][3][4] [9][10].

1.1.6 Θεραπεία

Οι θεραπευτικές προσεγγίσεις της νόσου διακρίνονται σε φαρμακευτικές και μη φαρμακευτικές. Οι φαρμακευτικές θεραπείες δεν δύνανται να καταπολεμήσουν πλήρως τη νόσο, αλλά στοχεύουν στην επιβράδυνση ή την αναστολή της καταστροφής των νευρώνων για την αποφυγή των σοβαρών και θανατηφόρων συμπτωμάτων. Η αποτελεσματικότητά τους διαφέρει ανάλογα τον ασθενή, ενώ η οι περισσότερες θεραπείες λειτουργούν αυξάνοντας τους νευροδιαβιβαστές στον εγκέφαλο. Ειδικότερα, αναστολείς της χολινεστεράσης αυξάνουν τα επίπεδα της ακετυλοχολίνης στον εγκέφαλο, επιβροθώντας στην επικοινωνία των νευρικών κυττάρων και η ουσία μεμαντίνη μπλοκάρει την επίδραση του γλουταμινικού στον εγκέφαλο [11] .

Οι μη φαρμακολογικές θεραπείες χρησιμοποιούνται συχνά με στόχο τη διατήρηση ή τη βελτίωση της γνωστικής λειτουργίας, την ικανότητα εκτέλεσης δραστηριοτήτων της καθημερινής ζωής ή τη συνολική βελτίωση της ποιότητας ζωής. Μπορούν επίσης να χρησιμοποιηθούν με στόχο τη μείωση των συμπτωμάτων συμπεριφοράς όπως η κατάθλιψη, η απάθεια, η περιπλάνηση, οι διαταραχές ύπνου, η διέγερση και η επιθετικότητα. Παραδείγματα περιλαμβάνουν θεραπεία τέχνης, θεραπεία βάσει δραστηριότητας και εκπαίδευση μνήμης. Όπως συμβαίνει με τις τρέχουσες φαρμακολογικές θεραπείες, οι μη φαρμακολογικές θεραπείες δεν έχει αποδειχθεί ότι αλλάζουν την πορεία της νόσου Alzheimer's [2].

1.1.7 Διάγνωση

1.1.7.1 Γενικές μέθοδοι διάγνωσης

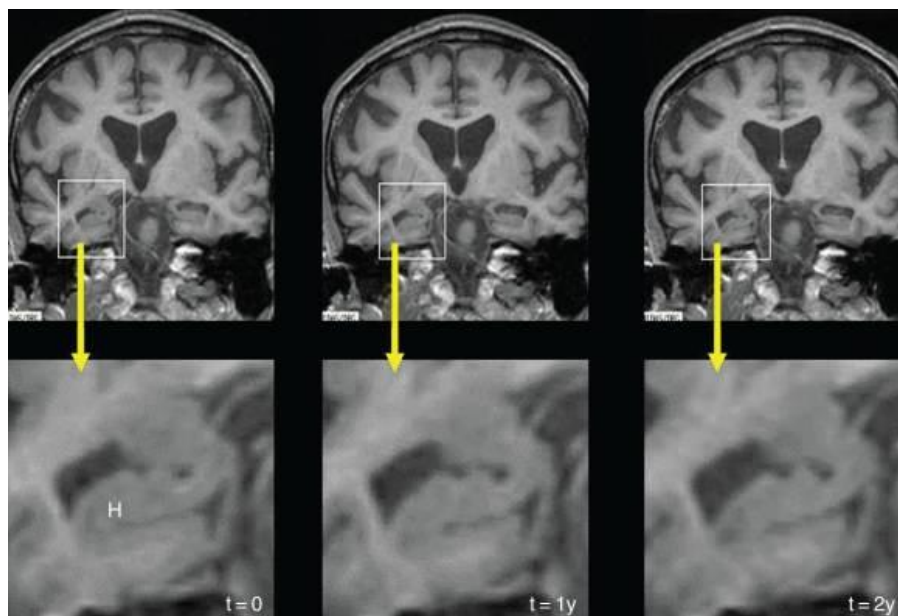
Η διάγνωση της νόσου Alzheimer's απαιτεί προσεκτική και ολοκληρωμένη ιατρική αξιολόγηση. Για τη διάγνωση της νόσου είναι αναγκαία η χρήση μιας ποικιλίας προσεγγίσεων και εργαλείων. Η λήψη ιατρικού και οικογενειακού ιστορικού από το άτομο, συμπεριλαμβανομένων του ψυχιατρικού ιστορικού και του ιστορικού γνωστικών και συμπεριφορικών αλλαγών, η διερώτηση ενός μέλους της οικογένειας ή άλλου ατόμου κοντά στον ασθενή για την παροχή πληροφοριών σχετικά με αλλαγές στη σκέψη, τις δεξιότητες ή τη συμπεριφορά, η διενέργεια γνωστικών τεστ και σωματικών και νευρολογικών εξετάσεων, η υποβολή του ατόμου σε εξετάσεις αίματος και εγκεφαλικής απεικόνισης για τον αποκλεισμό άλλων πιθανών αιτιών άνοιας, όπως ένας όγκος ή ορισμένες ανεπάρκειες βιταμινών, είναι κάποιες από αυτές. Αν και οι γιατροί μπορούν σχεδόν πάντα να προσδιορίσουν εάν ένα άτομο έχει άνοια, μπορεί να είναι δύσκολο να εντοπιστεί η ακριβής αιτία. Μπορεί να χρειαστούν αρκετές ημέρες ή εβδομάδες για να ολοκληρώσει ο ασθενής τις απαιτούμενες εξετάσεις και ο γιατρός να ερμηνεύσει τα αποτελέσματα και να κάνει μια διάγνωση. Το σημαντικότερο όπλο για την αντιμετώπιση του Αλτσχάιμερ είναι η έγκαιρη διάγνωση του κατά την φάση του MCI (Mild Cognitive Impairment), δηλαδή το στάδιο που προηγείται του συμπτώματος της άνοιας. Η διάγνωση βέβαια με MCI δεν είναι αρκετή για να ειπωθεί με ασφάλεια πως το άτομο θα νοσήσει από Alzheimer. Σε αυτό το στάδιο της νόσου οι απεικονιστικές τεχνικές είναι σε θέση να δώσουν τις απαντήσεις μέσω του εντοπισμού βιομαρτύρων (βιοχημικές αλλαγές του συστήματος), όπως του βαθμού συγκέντρωσης του Amyloid-β πεπτιδίου στον εγκέφαλο και του εντοπισμού εγκεφαλικής ατροφίας που σχετίζεται με τον μηχανισμό παθολογίας της νόσου [2]. Τον βιομάρτυρα της εγκεφαλικής ατροφίας μπορεί να παράσχει μια μαγνητική τομογραφία εγκεφάλου (MRI).

1.1.7.2 MRI Απεικονιστική τεχνική

Η μαγνητική τομογραφία ή Magnetic Resonance Imaging (MRI) είναι μια απεικονιστική μέθοδος του εσωτερικού του ανθρώπινου σώματος που παράγει υψηλής ευκρίνειας δισδιάστατες και τρισδιάστατες εικόνες λεπτών τομών. Δεν συμπεριλαμβάνει ιονίζουσα ακτινοβολία, συνεπώς

δεν επιβαρύνει τους ασθενείς. Χρησιμοποιεί ραδιοκύματα, μαγνητικά πεδία και κλίσεις πεδίου και βασίζεται στο ότι διαφορετικοί ιστοί εκπέμπουν περισσότερο ή λιγότερο ισχυρά σήματα αναλόγως της χημικής τους δομής. Συγκεκριμένα, η μέθοδος αυτή προκύπτει από την εκμετάλλευση του φαινομένου του πυρηνικού μαγνητικού συντονισμού (ΠΜΣ) κατά τον οποίο εφαρμόζουμε μαγνητικό πεδίο σε ένα δείγμα πυρήνων που απορροφούν, κατά τη διέγερση, και εκπέμπουν, κατά την αποδιέγερση, ηλεκτρομαγνητική ακτινοβολία. Στην MRI ασχολούμαστε κυρίως με πυρήνες υδρογόνου (πρωτόνια) καθώς αυτοί βρίσκονται σε μεγάλη συγκέντρωση μέσα στο ανθρώπινο σώμα (σε όλες σχεδόν τις ενώσεις του νερού και του λίπους και σε πολλές οργανικές ενώσεις) και αποδίδουν ισχυρά σήματα. [12]

Για τη διάγνωση της νόσου Alzheimer's η Μαγνητική Τομογραφία είναι ένα πολύ χρήσιμο εργαλείο. Μέσω αυτής μπορούμε να παρατηρήσουμε τυχούσα ατροφία και διαφοροποίηση στους ιστούς του εγκεφάλου, που είναι χαρακτηριστικά του νευροεκφυλισμού και της νόσου Alzheimer's. Ειδικότερα, κάποια από αυτά τα χαρακτηριστικά είναι η ανυποχώρητη ατροφία αρχικά στο μέσο κροταφικό λοβό, αλλά και στον ιππόκαμπο, στην αμυγδαλή και στην ιπποκάμπεια έλικα [13].



Εικόνα 2: Μαγνητικές τομογραφίες εγκεφάλου του ίδιου ασθενούς με χρονική απόσταση ενός χρόνου μεταξύ τους.

Οι απεικονίσεις αυτές αναδεικνύουν την αμείωτη ατροφία του ιπποκάμπου (Πηγή: [13])

1.2 Τεχνητά Νευρωνικά Δίκτυα - Συνελκτικά Νευρωνικά Δίκτυα

1.2.1 Τεχνητά Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα -ή ΤΝΔ για συντομία- είναι εμπνευσμένα από την βιολογία του ανθρώπινου εγκεφάλου και αποσκοπούν στη προσομοίωση της λειτουργίας του. Ένα ΤΝΔ αποτελείται από ένα μεγάλο πλήθος στοιχειωδών μη γραμμικών υπολογιστικών στοιχείων - γνωστά ως τεχνητοί νευρώνες- τα οποία είναι συνδεδεμένα σε μορφή δικτύου. Η ομοιότητα με τον ανθρώπινο εγκέφαλο έγκειται στον τρόπο με τον οποίο συνδέονται και λειτουργούν αυτοί οι τεχνητοί νευρώνες. [14]

1.2.1.1 Τεχνητός Νευρώνας - Δίκτυο Perceptron

Ένας νευρώνας τεχνητού νευρωνικού δικτύου λειτουργεί ως υπολογιστική μονάδα και επιχειρεί να προσομοιάσει την λειτουργία του νευρικού κυττάρου (ή νευρώνα), του βασικού δομικού στοιχείου του εγκεφάλου. Κάθε νευρώνας υλοποιεί την παρακάτω συνάρτηση μεταφοράς:

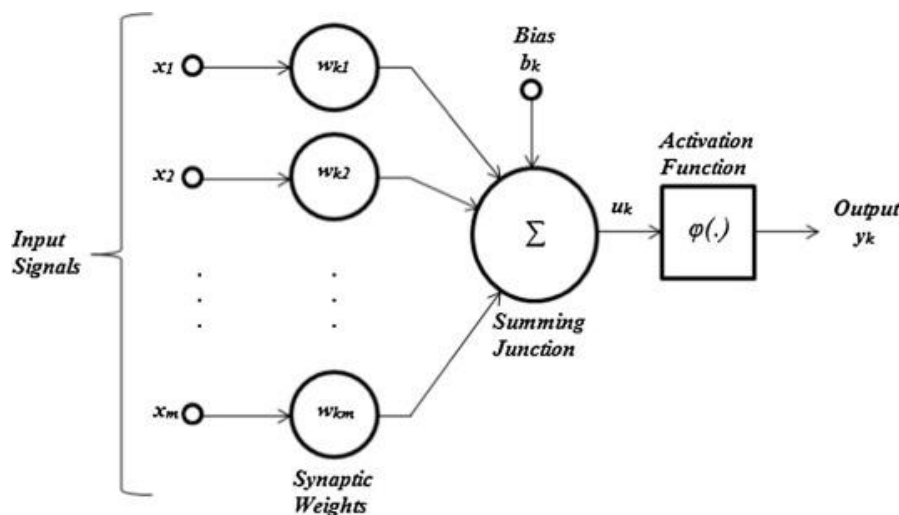
$$u = \sum_{i=1}^n w_i x_i - \theta,$$
$$y = f(u)$$

όπου x_i οι είσοδοι, w_i τα συναπτικά βάρη του νευρώνα, θ το κατώφλι ενεργοποίησης και $f(\cdot)$ η χρησιμοποιούμενη συνάρτηση ενεργοποίησης.

Η συνάρτηση ενεργοποίησης $f(\cdot)$ τροφοδοτείται από την λεγόμενη διέγερση u και δίνει την έξοδο y του νευρώνα, ενώ είναι μη γραμμική. Οι παράμετροι που ουσιαστικά καθορίζουν τη συμπεριφορά του νευρώνα είναι το διάνυσμα των συναπτικών βαρών $w = [w_1, w_2, \dots, w_n]^T$ και το κατώφλι θ .

Το πιο απλό νευρωνικό δίκτυο που μπορεί να σχεδιαστεί είναι ένα δίκτυο που αποτελείται από ένα μόνο νευρώνα. Στο δίκτυο αυτό, το οποίο ονομάζεται δίκτυο Perceptron, οι μόνες συνδέσεις που υπάρχουν είναι αυτές μεταξύ των εισόδων x_i και του νευρώνα [15], [16].

Σχηματικά η λειτουργία ενός νευρώνα φαίνεται παρακάτω:



Εικόνα 3: Λειτουργία νευρώνα, όπου με $\varphi(\cdot)$ σημειώνεται η συνάρτηση ενεργοποίησης και b_k το κατώφλι ενεργοποίησης (Πηγή: [17])

Δεδομένων κάποιων προτύπων x που ανήκουν σε δύο διαφορετικές κλάσεις και αν x_i τα χαρακτηριστικά που μοντελοποιούν το κάθε πρότυπο, σε ένα πρόβλημα ταξινόμησης αναζητείται η καμπύλη η οποία διαχωρίζει πλήρως τα πρότυπα των δύο κλάσεων στο υπερεπίπεδο των χαρακτηριστικών. Αν τα πρότυπα μπορούν να διαχωριστούν πλήρως από ένα επίπεδο, τότε οι κλάσεις ονομάζονται γραμμικά διαχωρίσιμες και υπάρχει ένα δίκτυο Perceptron που τις διαχωρίζει με βάση την σχέση:

$$\sum_{i=1}^n w_i x_i - \theta = 0$$

όπου για συνάρτηση ενεργοποίησης έχει χρησιμοποιηθεί η δυαδική συνάρτηση (1/0).

1.2.1.2 Δίκτυα Multilayer Perceptron (MLP)

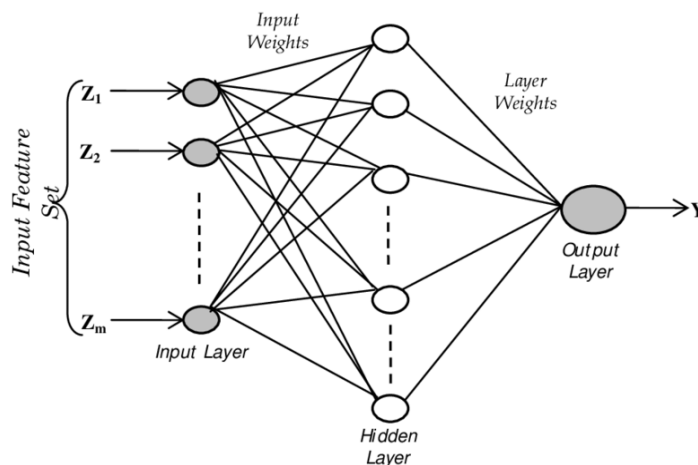
Λόγω της μορφής του, οι δυνατότητες αναπαράστασης διαχωριστικών επιφανειών είναι περιορισμένες στο δίκτυο Perceptron και περιλαμβάνουν μόνο επίπεδες επιφάνειες. Ο περιορισμός αυτός αίρεται με τη χρήση περισσότερων νευρώνων και την εισαγωγή των Perceptron Πολλών Στρωμάτων (Multilayer Perceptron ή MLP) που αποτελούν και τα τυπικά τεχνητά νευρωνικά δίκτυα. Τα δίκτυα αυτά αποτελούνται από πάνω από ένα πλήρως συνδεδεμένα (fully connected ή dense) επίπεδα. [15]

Συγκεκριμένα, οι νευρώνες μπορούν να βρεθούν σε ένα από τα τρία είδη επιπέδων:

1. Επίπεδο εισόδου: Είναι το πρώτο επίπεδο του δικτύου, το οποίο περιέχει τόσους νευρώνες όσες και οι μεταβλητές της εισόδου. Η λειτουργία των νευρώνων σε αυτό το επίπεδο είναι μονάχα να μεταβιβάζουν τα δεδομένα εισόδου στα επόμενα επίπεδα.
2. Κρυφό επίπεδο: Ένα ΤΝΔ μπορεί να περιέχει ποικίλο αριθμό κρυφών επιπέδων. Στα επίπεδα αυτά οι νευρώνες δέχονται ως είσοδο το άθροισμα των εξόδων των προηγούμενων νευρώνων, πολλαπλασιασμένων με το βάρος που αντιστοιχεί σε κάθε διασύνδεση. Στη συνέχεια η είσοδος κάθε νευρώνα περνάει από μια συνάρτηση ενεργοποίησης και το αποτέλεσμα δίνεται ως έξοδος του νευρώνα, η οποία πολλαπλασιάζεται με την σειρά της με τα αντίστοιχα βάρη, όταν διαβιβάζεται στους νευρώνες του επομένου επιπέδου. Να σημειωθεί πως για κάθε διασύνδεση ενός νευρώνα με τους νευρώνες του προηγούμενου αλλά και του επόμενου επιπέδου τα βάρη ποικίλουν και προσαρμόζονται μέσω της εκπαίδευσης..
3. Επίπεδο εξόδου: Είναι το τελικό επίπεδο του ΤΝΔ και περιέχει όσους νευρώνες όσες και οι πιθανές μεταβλητές εξόδου. Οι νευρώνες αυτοί δέχονται -με τον τρόπο που περιγράφηκε προηγουμένως- είσοδο βασισμένη στις εξόδους του τελευταίου κρυφού επιπέδου κι έχουν ως έξοδο μια (κανονικοποιημένη) τιμή (από 0 έως 1).

Να σημειωθεί ότι τα ΤΝΔ που μελετάμε δεν περιέχουν κλειστούς βρόχους και ονομάζονται επομένως δίκτυα πρόσθιας τροφοδότησης (feedforward neural networks) [14].

Σχηματικά η δομή ενός τυπικού ΤΝΔ είναι η εξής:



Εικόνα 4: Multilevel Perceptron δίκτυο με m εισόδους, ένα κρυφό επίπεδο και 1 έξοδο. (Πηγή: [18])

1.2.1.3 Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου

Η χρησιμότητα και η λειτουργικότητα ενός ΤΝΔ βασίζεται στην κατάλληλη προσαρμογή των βαρών των συνάψεων μεταξύ των νευρώνων των γειτονικών επιπέδων. Οι τιμές των συναπτικών βαρών ενός MLP είναι αυτές που καθορίζουν την τελική ταξινόμηση και την έξοδο του δικτύου, ενώ αυτές διαμορφώνονται μέσω μιας διαδικασίας που ονομάζεται εκπαίδευση του δικτύου. Επειδή οι περισσότεροι κανόνες εκπαίδευσης βασίζονται σε μεθόδους βελτιστοποίησης οι οποίες χρησιμοποιούν παραγώγους, προτιμώνται παραγωγίσιμες συναρτήσεις ενεργοποίησης.

Μερικές από τις πιο συχνά χρησιμοποιούμενες συναρτήσεις ενεργοποίησης είναι οι παρακάτω:

1. Σιγμοειδής (sigmoid):

$$\varphi(x) = \frac{1}{1 + e^{-x}}$$

2. Rectified Linear Unit (ή ReLU):

$$\varphi(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

3. Υπερβολική εφαπτομένη tanh:

$$\varphi(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Κατά την δημιουργία ενός ΤΝΔ, τα βάρη αρχικοποιούνται σε τυχαίες τιμές, οι οποίες πρόκειται να αλλάξουν κατά την εκπαίδευση. Η εκπαίδευση βασίζεται στη χρήση αλγορίθμου διόρθωσης του σφάλματος, ο οποίος ονομάζεται αλγόριθμος εκπαίδευσης. Στην παρούσα εργασία περιγράφεται η εκπαίδευση με την χρήση του αλγορίθμου οπισθοδιάδοσης του σφάλματος (backpropagation error algorithm), ο οποίος απαιτεί την ύπαρξη ενός τουλάχιστον κρυφού επιπέδου κι έχει ως παραμέτρους την συνάρτηση κόστους (loss function), τον βελτιστοποιητή (optimizer), τον ρυθμό εκπαίδευσης (learning rate) και τον αριθμό των εποχών (epochs).

Για κάθε κύκλο εκπαίδευσης (εποχή), η εκπαίδευση αποτελείται από τα εξής βήματα:

1. Υπολογίζεται το σφάλμα μέσω της επιλεγμένης συνάρτησης κόστους
2. Ανανεώνονται τα βάρη σύμφωνα με τον επιλεγμένο βελτιστοποιητή και ρυθμό εκπαίδευσης. Η ανανέωση αυτή εξαρτάται από την τιμή του σφάλματος και του βήματος εκπαίδευσης, με σκοπό την ελαχιστοποίηση του καινούριου σφάλματος του δικτύου.
3. Αν δεν έχει ξεπεραστεί ο αριθμός των εποχών, επαναλαμβάνεται η διαδικασία από το βήμα 1 [15], [16].

1.2.1.4 Μετρικές αξιολόγησης

Τα μοντέλα ΤΝΔ ποικίλουν ανάλογα με την αρχιτεκτονική τους και για την αξιολόγηση της απόδοσης τους χρησιμοποιούνται διάφορες μετρικές. Μερικές από αυτές είναι:

1. Accuracy: Ποσοστό των δειγμάτων που ταξινομήθηκαν σωστά. Υπολογίζεται από την σχέση:

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

όπου:

→ TP (True Positive): σωστή πρόβλεψη (μοντέλου) θετικής κλάσης

→ TN (True Negative): σωστή πρόβλεψη αρνητικής κλάσης

- FP (False Positive): λανθασμένη πρόβλεψη θετικής κλάσης
 → FN (False Negative): λανθασμένη πρόβλεψη αρνητικής κλάσης

2. Recall: Ποσοστό αληθώς θετικών δειγμάτων προς το σύνολο των θετικών δειγμάτων. Υπολογίζεται από την σχέση:

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

3. Precision: Ποσοστό θετικών δειγμάτων προς το σύνολο των δειγμάτων που προβλέφθηκαν θετικά. Προκύπτει από την σχέση:

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

4. F1 score: Υπολογίζεται μέσω των μετρικών precision και recall (αρμονικός μέσος των δύο). Χρησιμοποιείται όταν θέλουμε να ελέγξουμε αν υπάρχει ισορροπία μεταξύ των δύο μετρικών και δίνεται από την σχέση:

$$F1_{score} = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

Οι παραπάνω μετρικές παρουσιάζονται παραπάνω για την περίπτωση της δυαδικής ταξινόμησης (binary classification), παρόλα αυτά ο ορισμός τους μπορεί να επεκταθεί για να καλύπτει και την περίπτωση ταξινόμησης με πολλαπλές κλάσεις (multi-class classification) [19]

1.2.2 Συνελικτικά Νευρωνικά Δίκτυα

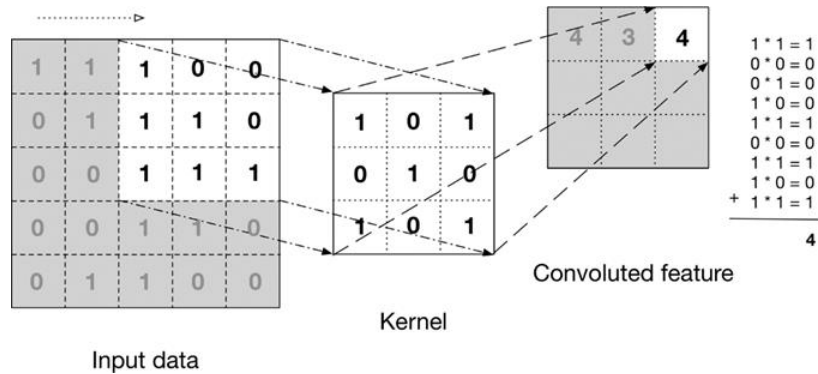
Ένας από τους πιο δημοφιλείς και εδραιωμένους τύπους Τεχνητών Νευρωνικών Δικτύων είναι τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks ή CNNs). Τα δίκτυα αυτά έχουν εφαρμοστεί εκτενώς σε προβλήματα του τομέα της όρασης υπολογιστών, όπως σε προβλήματα ταξινόμησης και κατάτμησης εικόνων, καθώς και σε προβλήματα ανίχνευσης αντικειμένων σε εικόνες [20]. Λόγω των υψηλών επιδόσεων που έχουν σημειώσει στις προαναφερθείσες περιπτώσεις, τα CNNs έχουν χρησιμοποιηθεί και σε άλλες καίριες εφαρμογές, όπως εφαρμογές επεξεργασίας φωνής και φυσικής γλώσσας [21]. Κύριο τους χαρακτηριστικό είναι η χρήση της μαθηματικής πράξης της συνέλιξης και η ενσωμάτωση συνελικτικών επιπέδων.

1.2.2.1 Συνελικτικά Επίπεδα

Τα συνελικτικά επίπεδα είναι τα θεμελιώδη συστατικά ενός CNN. Σε αυτά πραγματοποιείται δισδιάστατη συνέλιξη μεταξύ της εικόνας εισόδου και των δισδιάστατων φίλτρων (filters) ή αλλιώς πυρήνων (kernels), που περιλαμβάνονται στο επίπεδο. Ως πυρήνας ή φίλτρο μπορεί να οριστεί ένα παράθυρο συγκεκριμένης και συνήθως περιορισμένης διάστασης $n \times n$, ίδιου βάρους με την εικόνα εισόδου. Σε αναλογία με τη διακριτή συνέλιξη μονοδιάστατων σημάτων, η συνέλιξη δισδιάστατων σημάτων ορίζεται όπως φαίνεται παρακάτω:

$$(f * g)[x, y] = f[x, y] * g[x, y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1, n_2] g[x - n_1, y - n_2]$$

Κατά την συνέλιξη, το φίλτρο κυλίνεται πάνω στην εικόνα από τα αριστερά προς τα δεξιά και από πάνω προς τα κάτω, ωστόσο καλυφθούν όλες οι περιοχές της. Το «βήμα κύλισης» (stride) του φίλτρου μπορεί να οριστεί και ως διάφορο του 1. Ακόμα, επειδή μέσω της πράξης της συνέλιξης σημειώνεται μείωση των διαστάσεων της εικόνας εισόδου, πολλές φορές εφαρμόζεται «γέμισμα» των τεσσάρων πλευρών της αρχικής εικόνας με μηδενικά (zero padding), έτσι ώστε η μείωση να μην είναι τόσο σημαντική ή να διατηρηθούν οι αρχικές διαστάσεις. Τέλος, σημειώνεται ότι η έξοδος του συνελικτικού επιπέδου, ονομάζεται χάρτης ενεργοποίησης (activation map) ή χάρτης χαρακτηριστικών (feature map) [22], [23].



Εικόνα 5: Δισδιάστατη συνέλιξη εικόνας εισόδου 5x5 με φίλτρο 3x3 (Πηγή: [24])

Οι διαστάσεις του χάρτη ενεργοποίησης δίνονται από τις σχέσεις που ακολουθούν:

$$W_{out} = \frac{W_{in} - K + 2P}{S} + 1$$

$$H_{out} = \frac{H_{in} - K + 2P}{S} + 1$$

όπου $W_{in} \times H_{in}$ οι διαστάσεις της αρχικής εικόνας, $K \times K$ οι διαστάσεις του χρησιμοποιούμενου φίλτρου και P, S τα μεγέθη padding και stride που εφαρμόζονται αντίστοιχα. Ακόμα, επισημαίνουμε ότι αν D το βάθος της αρχικής εικόνας εισόδου, ο τελικός χάρτης χαρακτηριστικών θα έχει επίσης βάθος ίσο με D , καθώς το κάθε φίλτρο του επιπέδου συνελίσσεται δισδιάστατα με κάθε συνιστώσα της εικόνας εισόδου.

Οι παράμετροι των φίλτρων σε ένα συνελικτικό επίπεδο αποτελούν βάρη του δικτύου και επομένως διαμορφώνονται κατά την εκπαίδευσή του. Η κύρια λειτουργία των συνελικτικών επιπέδων είναι η εξαγωγή χρήσιμων χαρακτηριστικών από την εικόνα εισόδου (εξού και η ονομασία 'feature maps'). Καθώς είναι συνήθης η ενσωμάτωση περισσότερων του ενός συνελικτικών επιπέδων σε ένα συνελικτικό νευρωνικό δίκτυο, τα πρώτα επίπεδα τείνουν να εκπαιδεύονται έτσι ώστε να ανιχνεύουν low-level χαρακτηριστικά της εικόνας (ακμές, υφές κ.α.), ενώ τα επόμενα χρησιμοποιούν τα εξαχθέντα low-level features για να εξαγωγή πιο σύνθετα, high-level χαρακτηριστικά. Το γεγονός αυτό δίνει την δυνατότητα να χρησιμοποιηθούν αυτούσια τα ήδη εκπαιδευμένα, αρχικά επίπεδα δικτύων σε άλλα, διαφορετικά δίκτυα και να χρησιμοποιηθούν σε διαφορετικό σύνολο δεδομένων από το αρχικό στο οποίο είχαν εκπαιδευτεί. Η πρακτική αυτή ονομάζεται μεταφορά μάθησης (transfer learning) και εφαρμόζεται ευρέως για την εξοικονόμηση υπολογιστικών πόρων, αλλά και στις περιπτώσεις όπου τα διαθέσιμα δεδομένα δεν επαρκούν για την εκπαίδευση ενός μοντέλου εξ ολοκλήρου. [22],[23],[25].

Την αρχιτεκτονική τυπικών CNNs συμπληρώνουν επίπεδα υποδειγματοληψίας (pooling layers), πλήρως συνδεδεμένα, flatten, dropout αλλά και batch normalization επίπεδα τα οποία και παρουσιάζονται εκτενέστερα στην συνέχεια.

1.2.2.2 Επίπεδα υποδειγματοληψίας (pooling layers)

Τα επίπεδα υποδειγματοληψίας περιέχουν φίλτρα τα οποία εφαρμόζονται στην εικόνα εισόδου, όπως και τα συνελκτικά, με την διαφορά ότι οι τιμές των φίλτρων υποδειγματοληψίας είναι σταθερές και δεν αποτελούν βάρη του μοντέλου. Σκοπός των επιπέδων αυτών είναι η μείωση της διάστασης της εισόδου, με αποτέλεσμα τον περιορισμό του πλήθους των παραμέτρων του δικτύου. Η μείωση αυτή είναι επιθυμητή ως ένα βαθμό, για τον περιορισμό των απαιτούμενων για την εκπαίδευση υπολογιστικών πόρων, καθώς και για τον περιορισμό του ενδεχομένου υπερεκπαίδευσης (overfitting). Το είδος υποδειγματοληψίας που εφαρμόζεται κατά κύριο λόγο σε συνελκτικά δίκτυα είναι το Max Pooling, με το οποίο μόνο η μέγιστη τιμή μιας περιοχής της εικόνας διατηρείται, έπειτα από την εφαρμογή του φίλτρου στην περιοχή αυτή. Εναλλακτικά, μια άλλη επιλογή είναι η χρήση average pooling, με το οποίο διατηρείται η μέση τιμή ανά περιοχή εικονοστοιχείων. Η ενσωμάτωση επιπέδων δειγματοληψίας στην αρχιτεκτονική CNN μοντέλων έχει αποδειχθεί ότι ευνοεί σημαντικά την απόδοση τους στην πλειοψηφία των περιπτώσεων και για αυτό αποτελεί συνήθη πρακτική [22],[23].

1.2.2.3 Επίπεδα κανονικοποίησης: Batch Normalization και Dropout

Η κανονικοποίηση των δεδομένων αποτελεί συχνά χρησιμοποιούμενη μέθοδο προεπεξεργασίας, καθώς έχει φανεί ότι βελτιώνει την απόδοση του μοντέλου. Ο περιορισμός των δεδομένων εισόδου σε ένα προκαθορισμένο εύρος σταθμίζει την επιρροή που διαθέτουν τυχούσες τιμές-outliers κατά την εκπαίδευση του δικτύου, με αποτέλεσμα δεδομένα που δεν είναι χαρακτηριστικά του συνόλου εκπαίδευσης και πιθανώς να αντιστοιχούν σε λανθασμένες μετρήσεις να λαμβάνονται υπόψιν σε μικρότερο βαθμό.

Παρακάτω φαίνονται δύο βασικές σχέσεις που μπορούν να χρησιμοποιηθούν για την κανονικοποίηση των αρχικών δεδομένων:

$$x_{normalized} = \frac{x - \mu}{x_{max} - x_{min}} \text{ και } x_{normalized} = \frac{x - \mu}{\sigma},$$

όπου x_{max} , x_{min} η μέγιστη και η ελάχιστη εμφανιζόμενη τιμή αντίστοιχα και μ και σ η μέση τιμή και η μέση απόκλιση του συνόλου.

Τα batch normalization επίπεδα εκμεταλλευόμενα τα παραπάνω οφέλη πραγματοποιούν κανονικοποίηση της εξόδου ενδιάμεσων επιπέδων κατά micro-batches, επιταχύνοντας την εκπαίδευση του δικτύου κι επιτρέποντας την χρήση υψηλότερου learning rate. Η χρήση batch normalization επιπέδων συνεισφέρει στη σταθεροποίηση της κατανομής των δεδομένων, παρά την μεταβολή που επιφέρει σε αυτήν η συνεχής αλλαγή των παραμέτρων του δικτύου κατά την διάρκεια της εκπαίδευσης (φαινόμενο 'internal covariate shift') [26].

Αντίστοιχα, τα dropout επίπεδα μηδενίζουν την συνεισφορά ορισμένων νευρώνων τού αμέσως προηγούμενου επιπέδου, με αποτέλεσμα την μείωση των βαρών του επιπέδου αυτού. Οι νευρώνες που 'νεκρώνονται' επιλέγονται κάθε φορά τυχαία, ενώ το πλήθος των νευρώνων των οποίων η επίδραση θα αγνοηθεί επιλέγεται με βάση οριζόμενη υπερπαραμέτρο του dropout επιπέδου (rate). Η χρήση dropout μειώνει τον κίνδυνο υπερεκπαίδευσης και δρα ως ένα ακόμα είδος κανονικοποίησης των δεδομένων [27].

Τέλος, τα τελευταία επίπεδα των συνελκτικών δικτύων συνήθως αποτελούνται από πλήρως συνδεδεμένα (fully connected ή dense) επίπεδα. Τα επίπεδα αυτά χρησιμοποιούνται για την τελική ταξινόμηση, η οποία πραγματοποιείται με βάση τα χαρακτηριστικά που έχουν εξαχθεί από τα προηγούμενα επίπεδα. Ανάμεσα στα πυκνά επίπεδα και τα αμέσως προηγούμενα μεσολαβεί ένα επίπεδο flatten. Το επίπεδο αυτό έχει ως μοναδική λειτουργία την μετατροπή της δισδιάστατης εισόδου του σε ένα μοναδικό μονοδιάστατο διάνυσμα χαρακτηριστικών, έτσι ώστε το διάνυσμα αυτό να λειτουργήσει ως είσοδος στα fully-connected (dense) επίπεδα που ακολουθούν [22],[23].

1.3 Μηχανισμοί προσοχής

1.3.1 Εισαγωγή

Οι μηχανισμοί προσοχής είναι μια μέθοδος που εφαρμόζεται στα νευρωνικά δίκτυα και σκοπό έχει να προσομοιώνει τον τρόπο με τον οποίο ο ανθρώπινος εγκέφαλος επεξεργάζεται την πληροφορία που λαμβάνει μέσω της όρασης. Συγκεκριμένα, τα τμήματα του πεδίου οράσεως που εστιάζει το μάτι γίνονται αντιληπτά ως “υψηλής ευκρίνειας”, ενώ τα περιβάλλοντα χαρακτηρίζονται από “μειωμένη ευκρίνεια”. Με άλλα λόγια, μεγαλύτερα βάρη αναθέτονται στα τμήματα που εστιάζει το μάτι, ενώ ταυτόχρονα ελαχιστοποιείται η πληροφορία των μη σχετιζόμενων τμημάτων, με χρήση μικρότερων βαρών. Προφανώς αυτή η ιδιότητα του εγκεφάλου του δίνει τη δυνατότητα να επεξεργαστεί αποδοτικά και με ακρίβεια τα σημαντικά σημεία του πεδίου οράσεως, αντί να χρειάζεται να επεξεργαστεί όλο τον χώρο.

Το χαρακτηριστικό αυτό της ανθρώπινης όρασης ενέπνευσε του ερευνητές να αναπτύξουν τους μηχανισμούς προσοχής. Ξεκινώντας το 2014 χρησιμοποιήθηκαν για την ανάθεση υποτίτλων στην επεξεργασία της φυσικής γλώσσας (image captioning) , και από τότε έχουν βρει χρησιμότητα σε πλήθος εφαρμογών, αλλά κυρίως σε εφαρμογές όρασης υπολογιστών [28],[29]. Η δυνατότητα των μηχανισμών προσοχής να βελτιώνουν μεθόδους που βασίζονται σε συνελκτικά νευρωνικά δίκτυα έχει μελετηθεί ευρέως, ενώ επίσης συχνά χρησιμοποιούνται σε συνδυασμό με Αναδρομικά Νευρωνικά Δίκτυα (Recurrent Neural Networks ή RNNs) ή και Νευρωνικά Δίκτυα Γράφων (Graph Neural Networks). [28], [29]

1.3.2 Μηχανισμοί προσοχής σε Συνελκτικά Νευρωνικά Δίκτυα

1.3.2.1 Channel και Spatial Attention

Σε ένα συνελκτικό νευρωνικό δίκτυο κύριες συνιστώσες αποτελούν το διάνυσμα εισόδου ($B \times D \times H \times W$) (όπου B το χρησιμοποιούμενο batch size, D το βάθος της εικόνας και $H \times W$ οι διαστάσεις της), καθώς και τα συνελκτικά επίπεδα που περιέχονται στο εξεταζόμενο CNN. Η διέλευση της εισόδου από ένα συνελκτικό επίπεδο έχει ως αποτέλεσμα την δημιουργία χαρτών χαρακτηριστικών (feature maps), των οποίων το πλήθος εξαρτάται από τον αριθμό φίλτρων που συναποτελούν το εξεταζόμενο επίπεδο. Συγκεκριμένα, κάθε συνιστώσα της εικόνας εισόδου υφίσταται δισδιάστατη συνέλιξη με κάθε πυρήνα του επιπέδου, οπότε τελικά η έξοδος αποτελείται από πολλά διαφορετικά κανάλια-συνιστώσες. Εφόσον κάθε φίλτρο του συνελκτικού επιπέδου εξάγει διαφορετικά χαρακτηριστικά από την εικόνα εισόδου, τα κανάλια της εξόδου διαθέτουν διαφορετική χρησιμότητα για την τελική πρόβλεψη. Είναι επιθυμητή επομένως η απόδοση υψηλότερης βαρύτητας (ή ισοδύναμα η απόδοση μεγαλύτερης προσοχής) στα κανάλια μεγαλύτερης αξίας, κάτι που επιχειρείται με χρήση του μηχανισμού προσοχής ‘Channel Attention’. Μία από τις πρώτες και πιο πετυχημένες υλοποιήσεις channel attention ήταν

το Squeeze-and-Excitation δίκτυο, το οποίο εισήχθη το 2018 [30]. Μέσω τριών διαδοχικών υπολογιστικών μονάδων (squeeze, excitation και scale modules) γίνεται υπολογισμός των βαρών που αντιστοιχούν στα κανάλια της εξόδου ενός συνελκτικού επιπέδου. Μετεξέλιξη του Squeeze-and-Excitation δικτύου ήταν το λεγόμενο ECA-Net, το οποίο παρουσιάζεται εκτενέστερα παρακάτω [31]. Ο μηχανισμός αυτός αποδείχτηκε υπολογιστικά οικονομικότερος, κι επομένως ήταν αυτός που χρησιμοποιήθηκε στην τρέχουσα υλοποίηση.

Ένα ακόμα είδος προσοχής που μπορεί να ωφελήσει την απόδοση ενός συνελκτικού δικτύου, είναι το λεγόμενο Spatial Attention. Το spatial attention αφορά το κάθε κανάλι μεμονωμένα και με τη χρήση του, διαφορετική βαρύτητα (προσοχή) προσδίδεται στα διαφορετικά εικονοστοιχεία του ίδιου χάρτη. Μια υλοποίηση spatial attention, σε συνδυασμό με channel attention, αποτελεί το λεγόμενο Convolution Block Attention Module (CBAM) [32], το οποίο παρουσιάζεται αναλυτικότερα παρακάτω και το οποίο εφαρμόστηκε και στα πλαίσια αυτής της εφαρμογής.

1.3.2.2 ECA-Net

Το ECA-Net είναι ένας μηχανισμός προσοχής που εμφανίστηκε στη βιβλιογραφία το 2020. Ως μηχανισμός εφαρμόζει channel attention, ενώ αποτελεί βελτίωση των Squeeze-and-Excitation δικτύων που αναφέρθηκαν παραπάνω.

Συνίσταται από 3 διαδοχικά τμήματα:

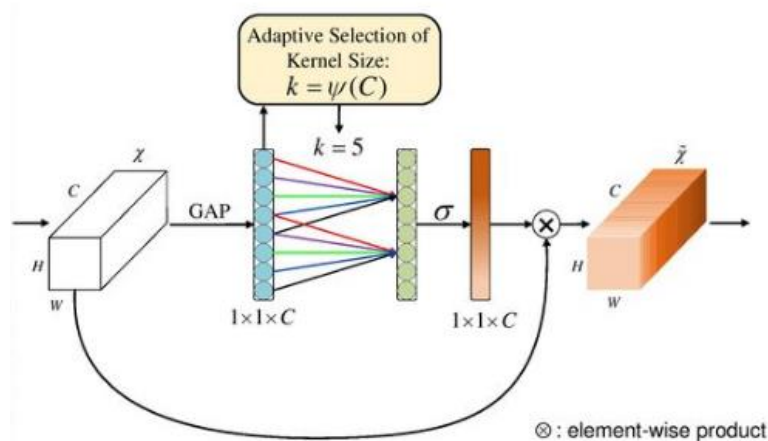
1. Global Feature Descriptor: Σε οποιοδήποτε συνελκτικό επίπεδο, η είσοδος και έξοδος έχουν διαστάσεις ($C \times H \times W$), όπου C το πλήθος των καναλιών ή αλλιώς feature maps και $H \times W$ οι διαστάσεις των επιμέρου χαρτών. Η εφαρμογή ενός Global Feature Descriptor μειώνει τη διάσταση του παραπάνω μεγέθους σε ($C \times 1 \times 1$). Ο μηχανισμός ECA-Net χρησιμοποιεί ως GFD το Global Average Pooling (GAP). Ειδικότερα, κάθε feature map αντιστοιχίζεται σε ένα μόνο pixel, του οποίου η τιμή προκύπτει ως η μέση τιμή των pixels του feature map αυτού.
2. Adaptive Neighborhood Interaction: Το στοιχείο με διαστάσεις ($C \times 1 \times 1$) στη συνέχεια υφίσταται μονοδιάστατη συνέλιξη με έναν κυλιόμενο πυρήνα του οποίου το μέγεθος k προσαρμόζεται ανάλογα με τον αριθμό των καναλιών C και προσδιορίζεται από τον ακόλουθο τύπο:

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)+b}{\gamma} \right\rfloor_{odd}$$

όπου b, γ υπερπαραμέτροι ορισμένες εκ των προτέρων σε τιμές 1 και 2 αντίστοιχα.

Μέσω αυτής της μονοδιάστατης συνέλιξης μοντελοποιείται η συσχέτιση μεταξύ των διαφορετικών καναλιών, ενώ η τιμή του k καθορίζει το πλάτος αυτής της τοπικής 'γειτονιάς' συσχετιζόμενων καναλιών.

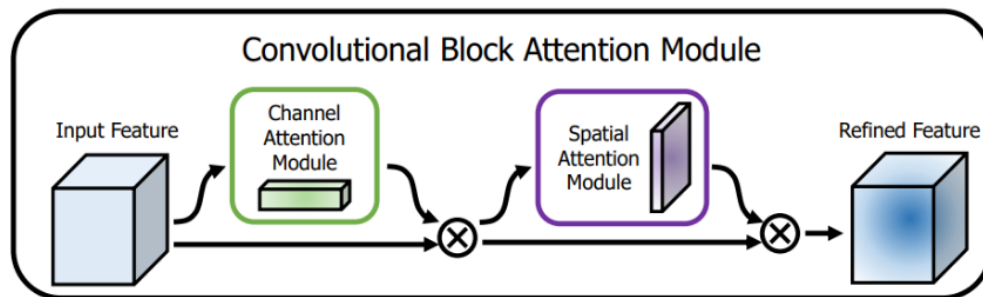
3. Broadcasted Scaling: Σε αυτό το σημείο του μηχανισμού το μέχρι τώρα επεξεργασμένο διάνυσμα με διαστάσεις ($C \times 1 \times 1$) δίνεται ως είσοδος σε μια σιγμοειδή (sigmoid) συνάρτηση, με αποτέλεσμα την κανονικοποίηση των τιμών του στο διάστημα (0,1). Εν συνεχεία, το εξαχθέν διάνυσμα γίνεται broadcast στις διαστάσεις του αρχικού στοιχείου εισόδου, δηλαδή στις διαστάσεις ($C \times H \times W$). Η μέχρι τώρα διαδικασία είχε ως αποτέλεσμα τη δημιουργία ενός πίνακα βαρών για κάθε κανάλι. Τελευταίο βήμα είναι ο πολλαπλασιασμός, κατά στοιχεία, του αρχικού διανύσματος εισόδου με τον πίνακα βαρών που υπολογίστηκε [31].



Εικόνα 6: Σχηματική αναπαράσταση ενός ECA block του μηχανισμού προσοχής ECANet (Πηγή: [33])

1.3.2.3 CBAM

Το CBAM είναι ένας μηχανισμός προσοχής που συνδυάζει Channel Attention και Spatial Attention, κι εφαρμόζεται έπειτα από κάθε συνελκτικό επίπεδο. Η χρησιμότητά του είναι ευρεία σε εφαρμογές ταξινόμησης εικόνων και αναγνώρισης αντικειμένων. Αποτελείται από δύο διαδοχικά υπομήματα, αυτό του channel (CAM - Channel Attention Module) και αυτό του spatial attention (SAM - Spatial Attention Module), με αυτή τη σειρά.



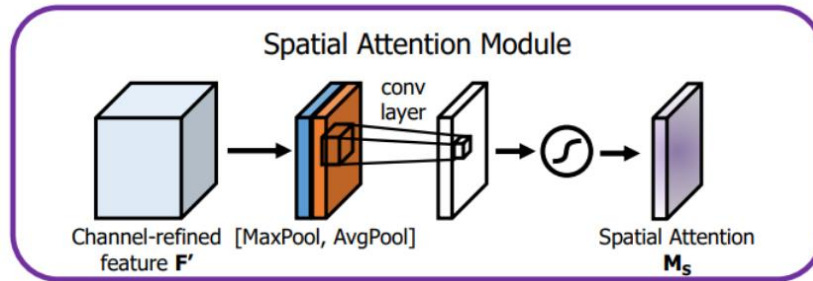
Εικόνα 7: Σχηματική αναπαράσταση ενός CBAM block του μηχανισμού προσοχής CBAM (Πηγή:[34])

Το spatial (χωρικό) attention αποτελεί μηχανισμό προσοχής στο επίπεδο μεμονωμένων feature maps. Ειδικότερα, για έναν δοθέντα χάρτη, μέσω του συγκεκριμένου είδους προσοχής, δημιουργείται μια “μάσκα”, η οποία εστιάζει στα πιο αντιπροσωπευτικά χαρακτηριστικά της εικόνας αυτής, δηλαδή στα χαρακτηριστικά στα οποία αξίζει περισσότερο να επικεντρωθεί το μοντέλο. Με αυτόν τον τρόπο βελτιώνεται η είσοδος των επιμέρους διαδοχικών συνελκτικών επιπέδων, με συνεπακόλουθη βελτίωση της απόδοσης του μοντέλου. Διαδέχεται channel attention, για το οποίο έγινε εκτενέστερη ανάλυση κατά την παρουσίαση του μηχανισμού προσοχής ECA-Net. Σημειώνεται ότι η χρήση και των δύο τεχνικών φαίνεται να αποφέρει καλύτερα αποτελέσματα από ό,τι η μεμονωμένη χρήση κάποιων από τα δύο.

Αναλυτικότερα για τα δύο υπομήματα έχουμε:

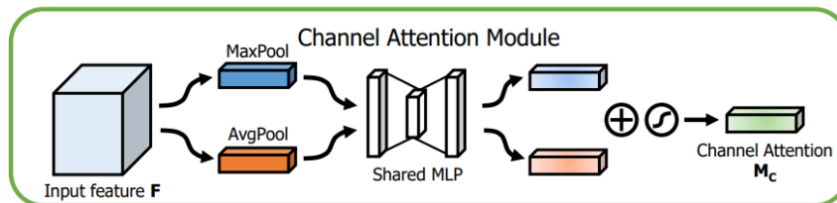
1. Spatial Attention Module (SAM): Απαρτίζεται με τη σειρά του από τρία διαδοχικά στάδια. Το αρχικό διάνυσμα μεγέθους $(C \times H \times W)$ μετατρέπεται σε $(2 \times H \times W)$, δηλαδή σε δύο

διακριτά feature maps. Ο ένας χάρτης χαρακτηριστικών είναι αποτέλεσμα εφαρμογής max pooling κατά μήκος των καναλιών, ενώ ο δεύτερος είναι αποτέλεσμα εφαρμογής αντίστοιχου average pooling. Το διάνυσμα αυτό διέρχεται μέσα από ένα συνελκτικό επίπεδο, έξοδος του οποίου αποτελεί ένα feature map, διαστάσεων $(1 \times H \times W)$. Στη συνέχεια, το αποτέλεσμα της συνέλιξης κανονικοποιείται με τη χρήση batch normalization επιπέδου. Τέλος, η έξοδος του συνελκτικού επιπέδου γίνεται τελικά είσοδος σε μια σιγμοειδή συνάρτηση και άρα κανονικοποιείται στο διάστημα $(0,1)$. Αυτή η “μάσκα” spatial attention εφαρμόζεται στη συνέχεια σε όλα τα feature maps της εισόδου, με ανά στοιχείο πολλαπλασιασμό.



Εικόνα 8: Σχηματική αναπαράσταση ενός SAM block του μηχανισμού προσοχής CBAM (Πηγή: [34])

2. Channel Attention Module (CAM): Το διάνυσμα εισόδου μετατρέπεται σε 2 υπόδιανύσματα διαστάσεων $(C \times 1 \times 1)$. Το ένα δημιουργείται με χρήση Global Max Pooling (GMP), δηλαδή εφαρμογή Max Pooling σε κάθε feature map και μετατροπή του σε ένα μόνο pixel και το δεύτερο με αντίστοιχη χρήση Global Average Pooling (GAP). Τα δύο διανύσματα γίνονται διαδοχικές είσοδοι ενός MLP δικτύου, ενώ τα διανύσματα εξόδου του δικτύου αυτού αθροίζονται κατά στοιχείο. Το MLP δίκτυο έχει τις ίδιες τιμές βαρών και για τα δύο διανύσματα εισόδου και ως συνάρτηση ενεργοποίησης χρησιμοποιεί τη συνάρτηση ReLU. Το τελικό ένα διάνυσμα διαστάσεων $(C \times 1 \times 1)$ γίνεται είσοδος μια σιγμοειδούς συνάρτησης, που παράγει τα βάρη του καναλιού. Αυτά εφαρμόζονται σε κάθε κανάλι του αρχικού διανύσματος εισόδου [32].

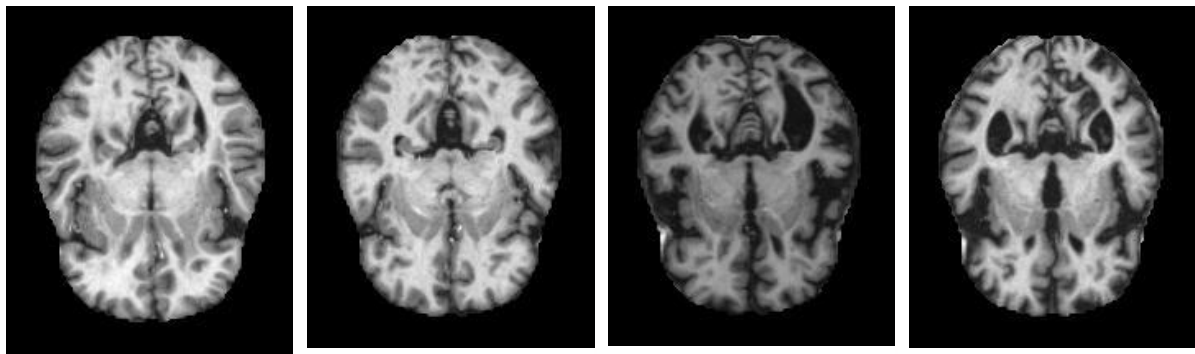


Εικόνα 9: Σχηματική αναπαράσταση ενός CAM block του μηχανισμού προσοχής CBAM (Πηγή: [34])

2. Προτεινόμενη Μεθοδολογία

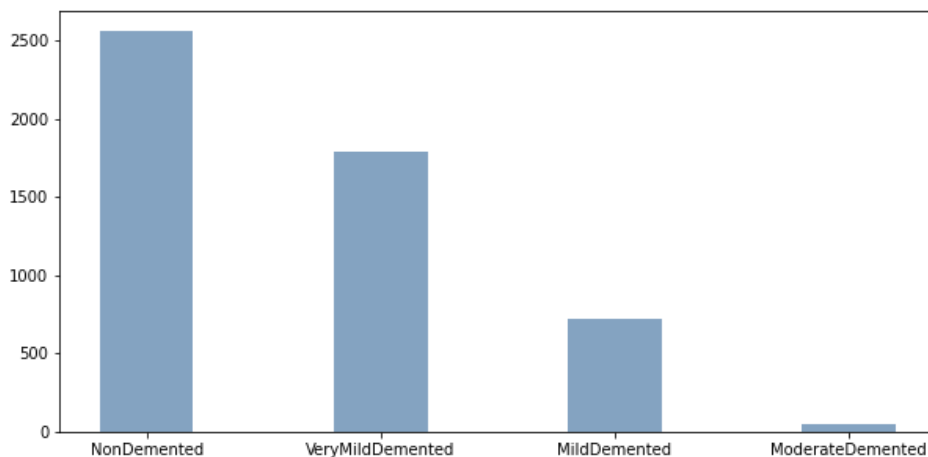
2.1 Σύνολο δεδομένων - Προεπεξεργασία

Για την εκπαίδευση του αναπτυσσόμενου μοντέλου βαθιάς μάθησης χρησιμοποιήθηκαν ως σύνολο δεδομένων 6400 εγκεφαλικές εικόνες μαγνητικής τομογραφίας από την δημοφιλή πλατφόρμα μηχανικής μάθησης Kaggle [35], [36]. Οι εικόνες αυτές μπορούν να εφαρμοστούν για την διάγνωση της νόσου Alzheimer's. Συγκεκριμένα, γίνεται διαχωρισμός των εικόνων στις εξής 4 κλάσεις: 'Non Demented' (μη περίπτωση άνοιας), 'Very Mild Demented' (πολύ ήπια περίπτωση άνοιας), 'Mild Demented' (ήπια περίπτωση άνοιας), 'Moderate Demented' (μέτριας σοβαρότητας περίπτωση άνοιας).



Εικόνα 10: Ενδεικτικές εικόνες συνόλου δεδομένων που χρησιμοποιήθηκε. (Πηγή: [36])

Κύριο χαρακτηριστικό του συνόλου δεδομένων είναι η ανισορροπία μεταξύ των τεσσάρων κλάσεων, γεγονός που απεικονίζεται και στο διάγραμμα που ακολουθεί. Ειδικότερα, η κατηγορία 'Non Demented' αποτελείται από συνολικά 3200 εικόνες, η 'Very Mild Demented' από 2240 εικόνες, η 'Mild Demented' από 896, ενώ τέλος η 'Moderate Demented' από 64 εικόνες.



Εικόνα 11: Διάγραμμα που απεικονίζει την άνιση κατανομή των εικόνων σε κατηγορίες στο σύνολο εκπαίδευσης (training set)

Σημειώνουμε ότι το σύνολο δεδομένων ήταν εξ αρχής διαχωρισμένο σε training και test υποσύνολα, με το training υποσύνολο να αποτελείται από συνολικά 5121 εικόνες και το test από 1279 εικόνες (αναλογία training-test συνόλων 80%-20%). Ακόμα, για την μετέπειτα εκπαίδευση

του δικτύου, ένα τμήμα, και ειδικότερα το 10%, του συνόλου εκπαίδευσης χρησιμοποιήθηκε ως σύνολο επικύρωσης (validation set). Το σύνολο αυτό δεν εφαρμόζεται άμεσα ως είσοδος για την εκπαίδευση του δικτύου, όπως το σύνολο εκπαίδευσης, αλλά χρησιμοποιείται για την αξιολόγηση του υπό εκπαίδευση μοντέλου με βάση τις χρησιμοποιούμενες μετρικές, έπειτα από την ολοκλήρωση της κάθε εποχής. Η απόδοση του μοντέλου στο σύνολο επικύρωσης είναι ενδεικτική του εάν έως τη στιγμή της εξέτασης η εκπαίδευση επιφέρει τα επιθυμητά αποτελέσματα. Ακόμα, η χρήση validation set προσφέρει την δυνατότητα περιορισμού του φαινομένου της υπερεκπαίδευσης (overfitting), καθώς είναι δυνατή η έγκαιρη αξιολόγηση της ικανότητας του μοντέλου να γενικεύει, δηλαδή να αποδίδει εξίσου ικανοποιητικά σε άγνωστα δεδομένα. Αντίστοιχα, το test set λειτουργεί ως μέσο αξιολόγησης του τελικού μοντέλου έπειτα από το πέρας της εκπαίδευσης του και την οριστικοποίηση των βαρών του. Συνοψίζοντας, συνολικά 4610 MRI εικόνες αποτέλεσαν το σύνολο εκπαίδευσης (training set), 511 εικόνες το σύνολο επικύρωσης (validation set), ενώ 1279 το σύνολο ελέγχου (test set).

Τέλος, πριν την είσοδο τους στο σύστημα, όλες οι εικόνες κανονικοποιήθηκαν έτσι ώστε οι τιμές των εικονοστοιχείων τους να λαμβάνουν τιμές στο διάστημα $[0, 1]$, έναντι του αρχικού διαστήματος $[0, 255]$. Ακόμα, έγινε αλλαγή της διάστασης των εικόνων από την προεπιλεγμένη 176×208 , στην συχνά χρησιμοποιούμενη κατά την εκπαίδευση συνελκτικών μοντέλων, 224×224 .

2.2 Σχεδιασμός και εκπαίδευση απλού συνελκτικού μοντέλου πρόβλεψης

Για τη διάγνωση της νόσου Alzheimer με χρήση του συνόλου MRI εικόνων που περιγράφηκε παραπάνω, αναπτύχθηκε βαθύ συνελκτικό νευρωνικό δίκτυο CNN. Τα CNNs αποτελούν μία από τις κορυφαίες και συνηθέστερες επιλογές για εφαρμογές ταξινόμησης εικόνων. Για την ανάπτυξη χρησιμοποιήθηκε η γλώσσα Python (v. 3.7.12), καθώς και οι βιβλιοθήκες Keras (v. 2.6.0) [37] και TensorFlow (v. 2.6.4) [38]. Τα πειράματα πραγματοποιήθηκαν στις πλατφόρμες Google Colab και Kaggle, οι οποίες προσφέρουν δωρεάν την δυνατότητα εκτέλεσης κώδικα σε περιβάλλον που προσομοιάζει Jupyter Notebook περιβάλλον. Η εκτέλεση πραγματοποιήθηκε με χρήση επιτάχυνσης υλικού (GPUs) με στόχο την επίτευξη ταχύτερων χρόνων εκπαίδευσης.

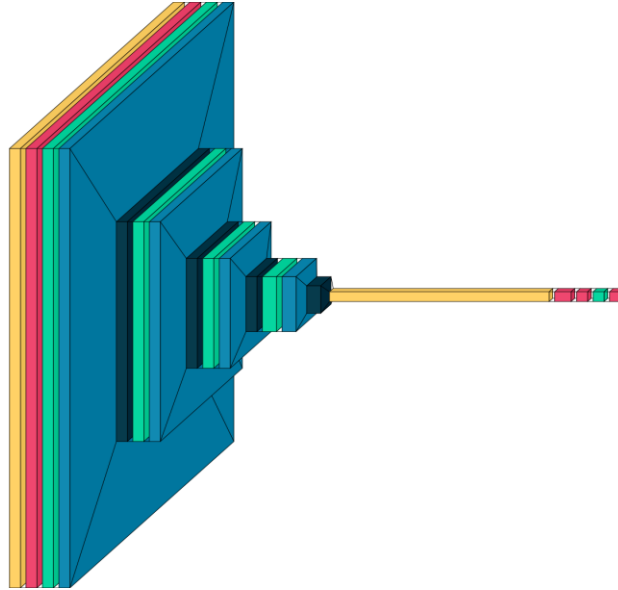
Η αρχιτεκτονική του τελικού δικτύου αποτελεί τυπικό παράδειγμα CNN ταξινομητή εικόνων και στηρίζεται στην δομή υπαρχουσών, εδραιωμένων αρχιτεκτονικών CNN, όπως το AlexNet [39] και το VGG-16. [40] Συγκεκριμένα, το αναπτυχθέν μοντέλο αποτελείται από τέσσερα συνελκτικά επίπεδα, καθένα εκ των οποίων ακολουθείται από ένα Batch Normalization και ένα Pooling επίπεδο. Όπως συνηθίζεται, το πλήθος των πυρήνων ανά συνελκτικό επίπεδο αυξάνεται καθώς το δίκτυο βαθαίνει, κι επομένως το 1ο συνελκτικό επίπεδο περιέχει 32 φίλτρα, το 2ο επίπεδο 64 φίλτρα, το 3ο 128 φίλτρα και το τελευταίο 256 φίλτρα. Οι διαστάσεις φίλτρων που χρησιμοποιούνται σε όλα τα συνελκτικά επίπεδα είναι 3×3 , ενώ επίσης χρησιμοποιείται stride ίσο με 1 και zero padding για την διατήρηση των διαστάσεων της εισόδου αμετάβλητων. Όσον αφορά τα επίπεδα υποδειγματοληψίας, χρησιμοποιούνται φίλτρα διαστάσεων 2×2 και τεχνική υποδειγματοληψίας max pooling.

Τα συνελκτικά, batch normalization και max pooling επίπεδα ακολουθούνται από τρία πλήρως συνδεδεμένα επίπεδα με 300, 150 και 4 νευρώνες αντίστοιχα (όπου 4 το πλήθος των κατηγοριών όπου γίνεται η ταξινόμηση). Επισημαίνεται ότι ανάμεσα στα αρχικά επίπεδα και τα πυκνά επίπεδα μεσολαβεί ένα flatten επίπεδο, το οποίο έχει ως στόχο την δημιουργία της μονοδιάστατης αναπαράστασης χαρακτηριστικών που θα τροφοδοτήσει τα επόμενα επίπεδα. Ακόμα, ανάμεσα στα δύο πρώτα και το τρίτο dense layer έχει τοποθετηθεί dropout επίπεδο με λόγο dropout ίσο με 0.25. Τέλος, οι έξοδοι των συνελκτικών, αλλά και των dense επιπέδων διέρχονται από την ReLU συνάρτηση ενεργοποίησης, πριν διέλθουν στο επόμενο επίπεδο. Μοναδική εξαίρεση αποτελεί το τελευταίο dense επίπεδο, το οποίο διαθέτει ως συνάρτηση ενεργοποίησης την softmax και παράγει ως έξοδο τις κανονικοποιημένες πιθανότητες που αντιστοιχούν σε κάθε κλάση για την δοθείσα εικόνα εισόδου.

Ο πίνακας που ακολουθεί συνοψίζει την παραπάνω περιγραφή:

Επίπεδο	Χαρακτηριστικά
Resizing	διαστάσεις εξόδου: (224, 224)
Rescaling	κανονικοποίηση στο διάστημα [0, 1]
Convolutional	32 φίλτρα διαστάσεων (3x3), χρήση zero padding, stride: 1, ReLU συνάρτηση ενεργοποίησης
Batch Normalization	
Max Pooling	φίλτρα 2x2
Convolutional	64 φίλτρα διαστάσεων (3x3), χρήση zero padding, stride: 1, ReLU συνάρτηση ενεργοποίησης
Batch Normalization	
Max Pooling	φίλτρα 2x2
Convolutional	128 φίλτρα διαστάσεων (3x3), χρήση zero padding, stride: 1, ReLU συνάρτηση ενεργοποίησης
Batch Normalization	
Max Pooling	φίλτρα 2x2
Convolutional	256 φίλτρα διαστάσεων (3x3), χρήση zero padding, stride: 1, ReLU συνάρτηση ενεργοποίησης
Batch Normalization	
Max Pooling	φίλτρα 2x2
Flatten	
Dense	300 νευρώνες, ReLU συνάρτηση ενεργοποίησης
Dense	150 νευρώνες, ReLU συνάρτηση ενεργοποίησης
Dropout	dropout ratio: 0.25
Dense	4 νευρώνες, softmax συνάρτηση ενεργοποίησης

Σημειώνουμε ότι η κανονικοποίηση και η αλλαγή των διαστάσεων των εικόνων που αναφέρονται παραπάνω στην υποενότητα 2.1 έχουν ενσωματωθεί στην αρχιτεκτονική του μοντέλου ως τα δύο πρώτα επίπεδα, όπως φαίνεται και από τον πίνακα. Τελικά, το μοντέλο που σχεδιάστηκε περιείχε συνολικά 15,488,230 εκπαιδευσιμες παραμέτρους.



Εικόνα 12: Η δομή του μοντέλου όπως προέκυψε με χρήση του εργαλείου οπτικοποίησης *visualkeras* [41]

Για την εκπαίδευση του δικτύου, ως μετρική σφάλματος (loss function) επιλέχθηκε η categorical crossentropy, η οποία εφαρμόζεται συχνά σε προβλήματα ταξινόμησης πολλών κλάσεων. Ως βελτιστοποιητής (optimizer) ορίστηκε ο Adam, ο οποίος τα τελευταία χρόνια έχει καταστεί πολύ δημοφιλής επιλογή έναντι του κλασικού stochastic gradient descent (SGD) βελτιστοποιητή. [42] Ακόμα, δοκιμάστηκαν δύο τιμές batch size: η τιμή 16 και η τιμή 32. Το μοντέλο φάνηκε να αποδίδει καλύτερα για μικρότερο batch size, ίσο με 16, όποτε αυτό χρησιμοποιήθηκε στη συνέχεια. Ο ρυθμός εκπαίδευσης (learning rate) ορίστηκε τελικά ίσος με 0.0001, έπειτα από δοκιμές και με υψηλότερους ρυθμούς εκπαίδευσης, ενώ 50 εποχές φάνηκαν να αρκούν για την εκπαίδευση του δικτύου, χωρίς να προκύπτει υποεκπαίδευση ή υπερεκπαίδευση αυτού.

Για την αξιολόγηση του μοντέλου χρησιμοποιήθηκαν οι μετρικές accuracy, precision, recall και AUC (Area Under the Curve). Κρίθηκε πολύ κρίσιμη η αξιολόγηση του δικτύου με πλήθος μετρικών, πέρα από την συχνά χρησιμοποιούμενη μετρική της ακρίβειας (accuracy), έτσι ώστε να δοθεί μια σφαιρική εκτίμηση της απόδοσης του, κάτι που είναι ιδιαίτερα σημαντικό, ειδικά στην περίπτωση όπου το αρχικό σύνολο δεδομένων χαρακτηρίζεται από ανισορροπία των κλάσεων. Συγκεκριμένα, στην περίπτωση προϋπάρχουσας ανισορροπίας του συνόλου δεδομένων, η χρήση της ακρίβειας ως αποκλειστικού μέσου αξιολόγησης του αναπτυσσόμενου μοντέλου μπορεί να προσφέρει παραπλανητικά υψηλά αποτελέσματα που οφείλονται σε σωστές προβλέψεις μόνο των κλάσεων πλειοψηφίας. Τέλος, σημειώνεται ότι για την αντιμετώπιση της υπάρχουσας ανισορροπίας του συνόλου εκπαίδευσης δοκιμάστηκαν ακόμη η εφαρμογή weighted learning [43], καθώς και η ενσωμάτωση επιπέδου επαύξησης δεδομένων (data augmentation layer) [44] στην αρχιτεκτονική του δικτύου. Παρόλα αυτά, επειδή οι δύο αυτές μέθοδοι δεν φάνηκαν να ευνοούν την απόδοση του μοντέλου, στην συνέχεια εγκαταλείφθηκαν.

2.3 Ενσωμάτωση μηχανισμών προσοχής στο υπάρχον συνελκτικό δίκτυο

Στην παρούσα εργασία έγινε εφαρμογή δύο διαφορετικών μηχανισμών προσοχής: του ECA-Net μηχανισμού προσοχής και του Convolution Block Attention Module (CBAM). Οι δύο αυτοί μηχανισμοί και τα δομικά τους στοιχεία παρουσιάστηκαν συνοπτικά στις υποενότητες 1.3.2.2 και 1.3.2.3 της 1ης ενότητας. Για την χρήση τους στο αναπτυχθέν μοντέλο, οι μηχανισμοί προσοχής αυτοί μοντελοποιήθηκαν ως συναρτήσεις και ενσωματώθηκαν ως νέα, custom επίπεδα στην υπάρχουσα αρχιτεκτονική. Ειδικότερα, και στις δύο περιπτώσεις, τα ECA-Net και τα CBAM blocks ενσωματώθηκαν αμέσως μετά από κάθε συνελκτικό επίπεδο, με αποτέλεσμα την απόδοση υψηλότερης βαρύτητας σε ορισμένες συνιστώσες της εξόδου, αλλά και σε ορισμένα εικονοστοιχεία εντός του κάθε καναλιού ξεχωριστά. Για την επανεκπαίδευση του μοντέλου και στις δύο περιπτώσεις χρησιμοποιήθηκε η ίδια διαρρύθμιση (χρησιμοποιούμενος βελτιστοποιητής, ρυθμός εκπαίδευσης, πλήθος εποχών εκπαίδευσης, μετρική σφάλματος κ.ο.κ.), έτσι ώστε να είναι δυνατή η σύγκριση των αποτελεσμάτων των τριών αρχιτεκτονικών επί, όσο τον δυνατόν γίνεται, ίσους όρους.

3. Ανάλυση Αποτελεσμάτων

3.1 Αξιολόγηση μοντέλων με χρήση μετρικών σφάλματος

Στην παρούσα ενότητα παρουσιάζονται τα αποτελέσματα τα οποία προέκυψαν από την εκπαίδευση των προαναφερθέντων αρχιτεκτονικών, με χρήση των ρυθμίσεων που περιγράφηκαν παραπάνω. Η απόδοση του κάθε μοντέλου αξιολογήθηκε με χρήση των τεσσάρων μετρικών accuracy, precision, recall και AUC (Area Under the Curve), ενώ τα αποτελέσματα για κάθε μοντέλο συνοψίζονται στον παρακάτω πίνακα:

	Accuracy	Precision	Recall	AUC
Απλό μοντέλο CNN	0.6724	0.6742	0.67	0.8643
Μοντέλο έπειτα από την ενσωμάτωση ECA-Net επιπέδων	0.7114	0.7112	0.7107	0.8609
Μοντέλο έπειτα από την ενσωμάτωση CBAM επιπέδων	0.7310	0.7319	0.7302	0.8769

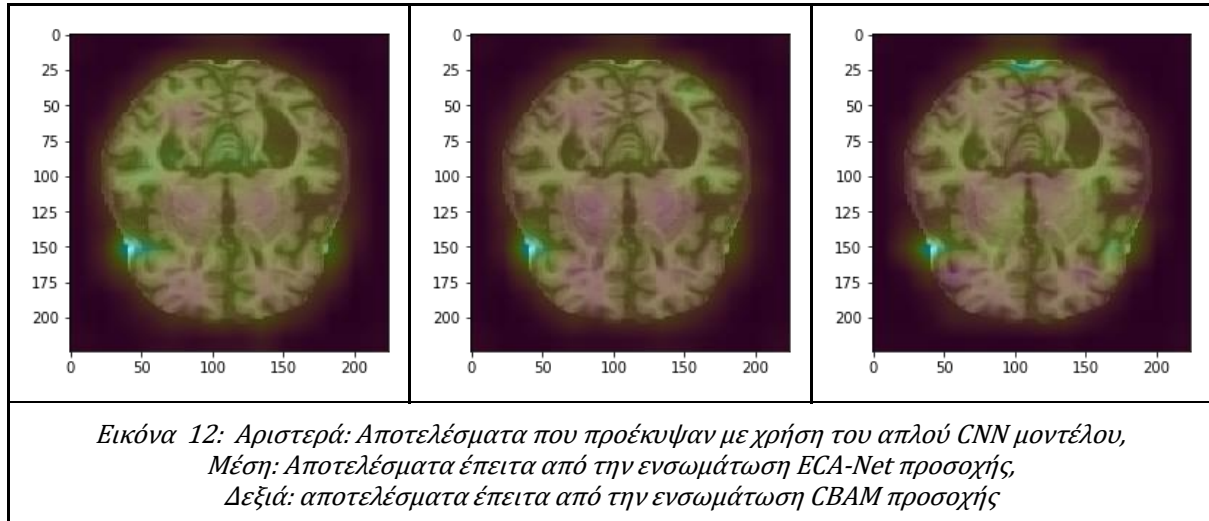
Παρατηρούμε ότι σε όλες τις περιπτώσεις η απόδοση του μοντέλου κρίνεται ικανοποιητική, λαμβάνοντας υπόψιν την απαιτητικότητα της εφαρμογής, αλλά και την αρχική, προϋπάρχουσα ανισορροπία των κλάσεων του συνόλου δεδομένων. Ειδικότερα, πραγματοποιώντας μια ανασκόπηση των υπαρχουσών υλοποιήσεων με στόχο την ταξινόμηση του συγκεκριμένου συνόλου δεδομένων στην πλατφόρμα Kaggle [45], θα παρατηρήσουμε σε αρκετές περιπτώσεις ακρίβεια μοντέλων παρόμοια με ή και μικρότερη από την δική μας, ακόμα και στις περιπτώσεις όπου χρησιμοποιούνται βαθιά, διαδεδомένα μοντέλα, όπως τα AlexNet και Xception. Ακόμα, όσον αφορά στην συγκριτική απόδοση των τριών αρχιτεκτονικών παρατηρούμε ότι η ενσωμάτωση των μηχανισμών προσοχής φαίνεται να ευνοεί την ταξινόμηση και συνεισφέρει στην καλύτερη απόδοση των μοντέλων και στις δύο περιπτώσεις.

3.2 Αξιολόγηση μοντέλων με χρήση Gradient-weighted Class Activation Mapping (Grad-CAM)

Τέλος, οι τρεις αρχιτεκτονικές μελετήθηκαν και με χρήση της τεχνικής Gradient-weighted Class Activation Mapping, ή απλούστερα Grad-CAM. [46] Η μέθοδος αυτή μπορεί να χρησιμοποιηθεί για την οπτικοποίηση των περιοχών μιας εικόνας που κρίνονται από ένα μοντέλο ως οι πιο σημαντικές για την κατηγοριοποίησή της σε μία κλάση. Η χρήση της Grad-CAM τεχνικής ενισχύει την ερμηνευσιμότητα της υπό εξέταση αρχιτεκτονικής και μπορεί ακόμα να εξυπηρετήσει ως μέσο αξιολόγησης της ορθής λειτουργίας του μοντέλου. Συγκεκριμένα, είναι δυνατή η διαπίστωση του εάν το μοντέλο εξετάζει περιοχές της εικόνας σχετιζόμενες με την ταξινόμηση ή όχι, καθώς και το εάν χρησιμοποιεί τις περιοχές αυτές για την διεκπεραίωση της τελικής

ταξινόμησης. Για την ανίχνευση των σημαντικότερων περιοχών χρησιμοποιούνται τα βάρη του τελευταίου συνελκτικού επιπέδου του εξεταζόμενου μοντέλου.

Παρακάτω φαίνονται τα αποτελέσματα της εφαρμογής της Grad-CAM τεχνικής στην ίδια εικόνα του συνόλου δεδομένων. Επισημασμένες με πιο ανοιχτά, κίτρινα και πράσινα χρώματα παρουσιάζονται οι περιοχές στις οποίες βασίστηκε περισσότερο το αντίστοιχο μοντέλο για την τελική του απόφαση. Παρατηρούμε ότι με χρήση των μηχανισμών προσοχής το μοντέλο εστιάζει περισσότερο σε συγκεκριμένες περιοχές της εικόνας για την διεκπεραίωση ταξινόμησης.



4. Συμπεράσματα - Συζήτηση

4.1 Σύνοψη

Συνοψίζοντας, στην παρούσα εργασία καταπιαστήκαμε με την πρόκληση της κατηγοριοποίησης εικόνων MRI υλοποιώντας το δικό μας συνελκτικό νευρωνικό δίκτυο. Για να έχουμε μια εποπτεία η οποία ανταποκρίνεται στην πραγματικότητα, όσον αφορά την απόδοση του μοντέλου, χρησιμοποιήσαμε μια πλειάδα μετρικών: Accuracy, Precision, Recall, AUC. Από τα αποτελέσματα αυτών αποφαινόμαστε ότι το μοντέλο λειτουργεί σε ικανοποιητικό βαθμό, βασιζόμενο στην αρχιτεκτονική του χωρίς την ενσωμάτωση κάποιου μηχανισμού προσοχής. Στη συνέχεια, υλοποιήθηκαν δύο νέα συγγενικά μοντέλα, έχοντας ως βάση το αρχικό, τα οποία διαφοροποιούνται λόγω της ενσωμάτωσης ECA-Net και CBAM μηχανισμών επιπέδων, αντίστοιχα. Και τα δύο αυτά μοντέλα παρουσίασαν γενική βελτίωση της απόδοσης, όπως και αναμενόταν.

4.2 Μελλοντικές Προεκτάσεις

Αν και η απόδοση του αρχικού μοντέλου κρίνεται ήδη σχετικά ικανοποιητική, παρατηρούμε ότι θα μπορούσε να βελτιωθεί περαιτέρω με την κατάλληλη προεπεξεργασία των δεδομένων. Πιο συγκεκριμένα, οπτικοποιώντας το dataset παρουσιάστηκε το πρόβλημα της ανισορροπίας μεταξύ των κλάσεων, όπως φαίνεται στην Εικόνα 11. Ουσιαστικά, παρουσιάζεται αυξανόμενη διαφορά στον αριθμό δειγμάτων όσο η κλάση αναφέρεται σε πιο βαρύ στάδιο άνοιας. Αυτή η ανισορροπία ενδεχομένως να προέρχεται από το γεγονός ότι σε ένα δείγμα ανθρώπων είναι πιο πιθανό κάποιος να είναι υγιής παρά να νοσεί από κάποια μορφή άνοιας. Μελλοντικά, θα μπορούσαμε να μελετήσουμε και να υλοποιήσουμε και άλλες τεχνικές αντιμετώπισης αυτής της ανισορροπίας των κλάσεων, έτσι ώστε να βελτιστοποιηθεί η διαδικασία της εκπαίδευσης του μοντέλου και συνεπώς να αυξηθεί η απόδοση του. Επίσης, επιθυμητή θα ήταν η περαιτέρω βελτιστοποίηση της αρχιτεκτονικής του CNN μοντέλου που υλοποιήσαμε, μέσω μιας πιο ενδεδειγμένης διαδικασίας fine-tuning των υπερπαραμέτρων του δικτύου. Όσον αφορά τους μηχανισμούς προσοχής, ως μελλοντικός στόχος μπορεί να τεθεί η βαθύτερη ανάλυση και επεξήγηση των αποτελεσμάτων των δύο μοντέλων με βάση και τα αποτελέσματα της Grad-CAM μεθόδου, καθώς και η εφαρμογή και άλλων μηχανισμών προσοχής που συναντώνται στη βιβλιογραφία.

5. Βιβλιογραφία

- [1] Sloane, P. D., Zimmerman, S., Suchindran, C., Reed, P., Wang, L., Boustani, M., & Sudha, S. (2002). The public health impact of Alzheimer's disease, 2000-2050: Potential implication of treatment advances. In *Annual Review of Public Health* (Vol. 23). <https://doi.org/10.1146/annurev.publhealth.23.100901.140525>
- [2] Gaugler, J., James, B., Johnson, T., Scholz, K., & Weuve, J. (2016). 2016 Alzheimer's disease facts and figures. *Alzheimer's and Dementia*, 12(4). <https://doi.org/10.1016/j.jalz.2016.03.001>
- [3] Duyckaerts, C., Delatour, B., & Potier, M. C. (2009). Classification and basic pathology of Alzheimer disease. In *Acta Neuropathologica* (Vol. 118, Issue 1). <https://doi.org/10.1007/s00401-009-0532-1>
- [4] Finder, V. H. (2010). Alzheimer's disease: A general introduction and pathomechanism. In *Journal of Alzheimer's Disease* (Vol. 22, Issue SUPPL. 3). <https://doi.org/10.3233/JAD-2010-100975>
- [5] <https://www.neyrologos.gr/alzheimer-symptomata-stadia-aitia-therapeia/>
- [6] Goedert, M., & Spillantini, M. G. (2006). A century of Alzheimer's disease. In *Science* (Vol. 314, Issue 5800). <https://doi.org/10.1126/science.1132814>
- [7] Berrios, G. E. (1990). Alzheimer's disease: A conceptual history. *International Journal of Geriatric Psychiatry*, 5(6). <https://doi.org/10.1002/gps.930050603>
- [8] Ballenger, J. F. (2006). Progress in the history of Alzheimer's disease: The importance of context. In *Journal of Alzheimer's Disease* (Vol. 9, Issue SUPPL. 3). <https://doi.org/10.3233/jad-2006-9s302>
- [9] Merriam, A. E., Aronson, M. K., Gaston, P., Wey, S. -L, & Katz, I. (1988). The Psychiatric Symptoms of Alzheimer's Disease. *Journal of the American Geriatrics Society*, 36(1). <https://doi.org/10.1111/j.1532-5415.1988.tb03427.x>
- [10] Lyketsos, C. G., Carrillo, M. C., Ryan, J. M., Khachaturian, A. S., Trzepacz, P., Amatniek, J., Cedarbaum, J., Brashear, R., & Miller, D. S. (2011). Neuropsychiatric symptoms in Alzheimer's disease. In *Alzheimer's and Dementia* (Vol. 7, Issue 5). <https://doi.org/10.1016/j.jalz.2011.05.2410>
- [11] <https://www.nhs.uk/conditions/alzheimers-disease/treatment/>
- [12] “Καρκίνος του Πνεύμονα | LungCancer.gr - Απεικονιστικές Μέθοδοι.” http://www.lungcancer.gr/portal/content/karkinos/oz_20071017287.php3
- [13] Johnson KA, Fox NC, Sperling RA, Klunk WE. Brain imaging in Alzheimer disease. *Cold Spring Harb Perspect Med*. 2012;2(4):a006213. doi:10.1101/cshperspect.a006213
- [14] Gonzalez R.C & Woods R.E, (2018), Ψηφιακή Επεξεργασία Εικόνων, 4η Έκδοση, Εκδόσεις Τζιόλα
- [15] Διαμαντάρας Κ., (2007), Τεχνητά Νευρωνικά Δίκτυα, Εκδόσεις Κλειδάριθμος
- [16] Haykin S. O., (2009), Neural Networks and Learning Machines, 3η Έκδοση, Εκδόσεις McMaster University, Canada
- [17] Erdi Tosun, Kadir Aydin, Mehmet Bilgili, Comparison of linear regression and artificial neural network model of a diesel engine fueled with biodiesel-alcohol mixtures, *Alexandria Engineering Journal*, Volume 55, Issue 4, 2016, Pages 3081-3089, ISSN 1110-0168, <https://doi.org/10.1016/j.aej.2016.08.011>.
- [18] Swarup, Shanti. (2009). STUDY OF NEURAL NETWORK MODELS FOR SECURITY ASSESSMENT IN POWER SYSTEMS. *International Journal of Research and Reviews in Applied Sciences*. 1.
- [19] Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for Multi-Class Classification: an Overview. *ArXiv*, abs/2008.05756.
- [20] Sultana, F., Sufian, A., & Dutta, P. (2018). Advancements in Image Classification using Convolutional Neural Network. 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 122-129.
- [21] Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *EMNLP*.
- [22] O'Shea, K., & Nash, R. (2015). An Introduction to Convolutional Neural Networks. *ArXiv*, abs/1511.08458.
- [23] Wu, J. (2017). Introduction to Convolutional Neural Networks.
- [24] Swarup, Shanti. (2009). STUDY OF NEURAL NETWORK MODELS FOR SECURITY ASSESSMENT IN POWER SYSTEMS. *International Journal of Research and Reviews in Applied Sciences*. 1.
- [25] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, 109, 43-76.
- [26] Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *ArXiv*, abs/1502.03167.

- [27] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1 (January 2014), 1929–1958.
- [28] Ghaffarian, S.; Valente, J.; van der Voort, M.; Tekinerdogan, B. Effect of Attention Mechanism in Deep Learning-Based Remote Sensing Image Processing: A Systematic Literature Review. *Remote Sens.* 2021, 13, 2965. <https://doi.org/10.3390/rs13152965>
- [29] Zhaoyang Niu, Guoqiang Zhong, Hui Yu, A review on the attention mechanism of deep learning, *Neurocomputing*, Volume 452, 2021, Pages 48-62, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2021.03.091>.
- [30] Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-Excitation Networks. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7132-7141.
- [31] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11531-11539.
- [32] Woo, S., Park, J., Lee, J.Y., Kweon, I.S. (2018). CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) *Computer Vision – ECCV 2018*. ECCV 2018. Lecture Notes in Computer Science(), vol 11211. Springer, Cham. https://doi.org/10.1007/978-3-030-01234-2_1
- [33] <https://blog.paperspace.com/attention-mechanisms-in-computer-vision-ecanet/>
- [34] <https://blog.paperspace.com/attention-mechanisms-in-computer-vision-cbam/>
- [35] <https://www.kaggle.com/>
- [36] <https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images>
- [37] <https://keras.io/>
- [38] <https://www.tensorflow.org/>
- [39] Krizhevsky, Alex & Sutskever, Ilya & Hinton, Geoffrey. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*. 25. 10.1145/3065386.
- [40] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.
- [41] <https://github.com/paulgavrikov/visualkeras/>
- [42] Kingma, D.P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980.
- [43] He, J., & Cheng, M. X. (2021). Weighting Methods for Rare Event Identification From Imbalanced Datasets. *Frontiers in big data*, 4, 715320. <https://doi.org/10.3389/fdata.2021.715320>
- [44] Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. *ArXiv*, abs/1712.04621.
- [45] <https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images/code>
- [46] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.nced