

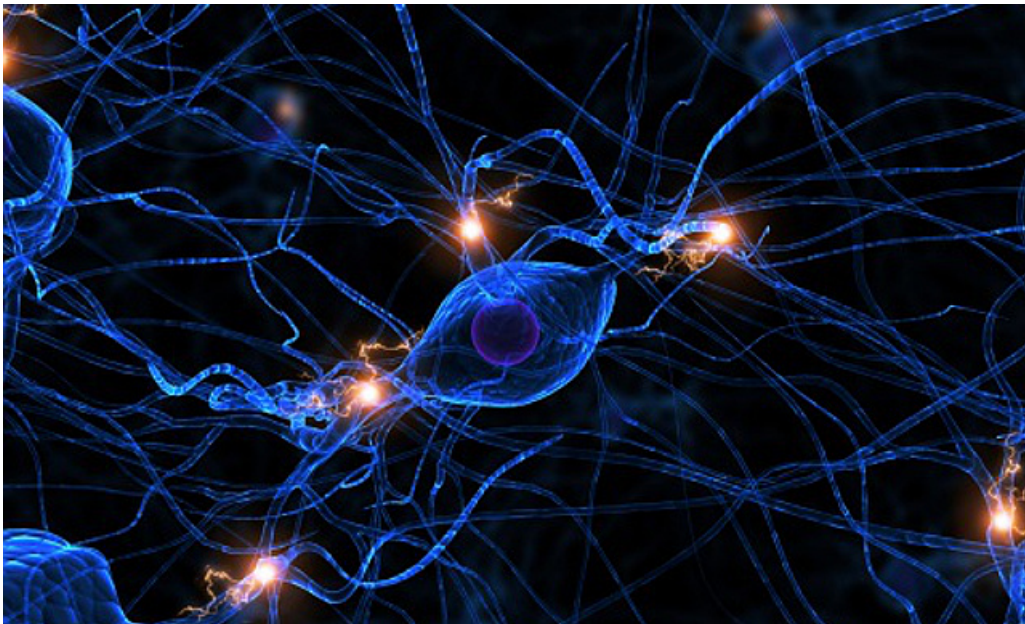


ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Χρήση τεχνικών βαθιάς μάθησης για την διάγνωση της νόσου Alzheimer μέσω μαγνητικών τομογραφιών του εγκεφάλου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ειρήνη Κληρονόμου



Επιβλέπων: Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2022



Χρήση τεχνικών βαθιάς μάθησης για την διάγνωση της νόσου Alzheimer μέσω μαγνητικών τομογραφιών του εγκεφάλου

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ειρήνη Κληρονόμου

Επιβλέπων: Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 18η Νοεμβρίου 2022.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Στέφανος Κόλλιας
Καθηγητής Ε.Μ.Π.

.....
Αθανάσιος Βουλόδημος
Επικουρος Καθηγητής Ε.Μ.Π.

.....
Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.



Copyright © - All rights reserved. Με την επιφύλαξη παντός δικαιώματος.
Ειρήνη Κληρονόμου, 2022.

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

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

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

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

(Υπογραφή)

.....
Ειρήνη Κληρονόμου

Περίληψη

Η Μηχανική Μάθηση και ειδικότερα η χρήση των τεχνητών νευρωνικών δικτύων αποτελεί ένα ραγδαία αναπτυσσόμενο κλάδο της σύγχρονης επιστήμης, με ποικίλες εφαρμογές σε πολυάριθμους επιστημονικούς τομείς, έχοντας ως κατευθυντήριο άξονα την βελτίωση της ποιότητας ζωής των ανθρώπων. Στο επίκεντρο της καλύτερης διαβίωσης του ανθρώπου τοποθετείται η επιστήμη της Ιατρικής, με το πεδίο των νευρωνικών δικτύων να έχει συνεισφέρει στη διάγνωση και την αντιμετώπιση ποικίλων ασθενειών.

Ιδιαίτερα η νόσος Alzheimer, την οποία πραγματεύεται η παρούσα εργασία, συναντάται όλο και περισσότερο στη σύγχρονη κοινωνία, χρήζοντας έτσι απαραίτητη την αποτελεσματική αντιμετώπισή της. Η διάγνωσή της αποτελεί μία από τις μεγαλύτερες προκλήσεις της ιατρικής κοινότητας, καθώς μέχρι σήμερα δεν υφίσταται κάποια συγκεκριμένη μέθοδος για την επιτέλεση του σκοπού αυτού. Τα διάφορα μοντέλα και αλγόριθμοι που έχουν αναπτυχθεί στο πλαίσιο των τεχνητών νευρωνικών δικτύων στοχεύουν σε μία ελπιδοφόρα λύση προς την κατεύθυνση αυτή.

Στην προκείμενη διπλωματική εργασία, στόχος είναι η αξιοποίηση των τεχνικών βαθιάς μάθησης στην κατηγοριοποίηση εικόνων για την έγκυρη διάγνωση της νόσου Alzheimer μέσω μαγνητικών τομογραφιών του εγκεφάλου. Συγκεκριμένα το μοντέλο που αναπτύχθηκε περιλαμβάνει την απόκτηση και την κατάλληλη προεπεξεργασία του συνόλου δεδομένων και την εκπαίδευσή του μέσω συνελκτικών και ανατροφοδοτούμενων δικτύων. Τα αποτελέσματα που παράχθηκαν και αξιολογήθηκαν με βάση κάποιες μετρικές ακρίβειας, δημιουργούν προσδοκίες για την εύρεση παρόμοιων εφαρμογών ιατρικής διάγνωσης και για άλλες νευρολογικές ασθένειες του εγκεφάλου.

Λέξεις Κλειδιά

Βαθιά Μάθηση, Συνελκτικά Νευρωνικά Δίκτυα, Ανατροφοδοτούμενα Νευρωνικά Δίκτυα, Νόσος Αλτσχάιμερ (Alzheimer), Μαγνητική Τομογραφία, Κατηγοριοποίηση Εικόνων

Abstract

Machine Learning and in particular the use of artificial neural networks is a rapidly developing sector of modern science, with diverse applications in numerous scientific fields, focusing on improving the quality of life of humans. At the centre of better human living is the science of medicine and the field of neural networks has contributed to the diagnosis and treatment of a variety of diseases.

In particular, Alzheimer disease, which is the subject of this thesis, is increasingly encountered in modern society, thus necessitating its effective treatment. Its diagnosis is one of the greatest challenges for the medical community, as to this moment there is no specific method for accomplishing this purpose. The various models and algorithms developed in the context of artificial neural networks aim at a promising solution in this direction.

In the present thesis, the aim is to utilize deep learning techniques in image classification for a valid diagnosis of Alzheimer disease through brain MRIs. Specifically, the developed model involves the acquisition and appropriate preprocessing of the dataset and its training through convolutional and recurrent networks. The results generated and evaluated based on some accuracy metrics, raise expectations for finding similar medical diagnosis applications for other neurological brain diseases.

Keywords

Deep Learning, Convolutional Neural Networks, Recurrent Neural Networks, Alzheimer's Disease, MRI, Image Classification

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ.Στέφανο Κόλλια που μου έδωσε την ευκαιρία να ασχοληθώ με ένα τόσο ενδιαφέρον και πολύπλευρο θέμα και μου εμπιστεύτηκε την εκπόνησή του.

Επιπλέον οφείλω να ευχαριστήσω την κυρία Παρασκευή Τζούβελη και τον Αναστάση Αρσένο για τη συνεχή καθοδήγηση και υποστήριξή τους κατά την σύνθεση της παρούσας εργασίας.

Τέλος θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένεια και τους φίλους μου, που ήταν δίπλα μου καθ' όλη την διάρκεια των σπουδών μου και με την στήριξη και την αγάπη τους έκαναν την κάθε στιγμή ξεχωριστή.

Αθήνα, Νοέμβριος 2022

Ειρήνη Κληρουόμου

Περιεχόμενα

Περίληψη	7
Abstract	9
Ευχαριστίες	11
1 Εισαγωγή	19
1.1 Η νόσος Alzheimer	19
1.2 Διάγνωση της νόσου Alzheimer	19
1.2.1 Απεικόνιση Μαγνητικής Τομογραφίας (Magnetic Resonance Imaging)	20
1.3 Σκοπός και δομή της εργασίας	21
I Θεωρητικό Μέρος	23
2 Θεωρητικό υπόβαθρο	25
2.1 Νευρωνικά Δίκτυα	25
2.1.1 Βιολογικά Νευρωνικά Δίκτυα	25
2.1.2 Τεχνητά Νευρωνικά Δίκτυα	26
2.1.3 Συναρτήσεις ενεργοποίησης	28
2.1.4 Αλγόριθμοι Βελτιστοποίησης	30
2.1.5 Αλγόριθμος Οπισθοδιάδοσης (Back-Propagation Algorithm)	33
2.2 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)	35
2.3 Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks)	38
2.3.1 Δομή και Μοντελοποίηση	38
2.3.2 Οπισθοδιάδοση στο Χρόνο (Back Propagation Through Time)	40
2.3.3 Είδη Ανατροφοδοτούμενων Νευρωνικών Δικτύων	41
2.4 Μεταφορά Μάθησης (Transfer Learning)	44
2.5 Πλεονεκτήματα και Μειονεκτήματα των Τεχνητών Νευρωνικών Δικτύων	45
2.5.1 Πλεονεκτήματα	46
2.5.2 Μειονεκτήματα	46
2.5.3 Εφαρμογές	46
II Πρακτικό Μέρος	49
3 Ανάλυση Μεθοδολογίας	51

3.1	Δεδομένα και Προεπεξεργασία	51
3.2	Επαύξηση Δεδομένων (Data Augmentation)	52
3.3	ResNet	52
3.3.1	Υπολειπόμενη Εκμάθηση (Residual Learning)	53
3.3.2	Χαρτογράφηση Ταυτότητας με Συντομεύσεις (Identity Mapping by Shortcuts)	53
3.3.3	Εφαρμογές	54
3.4	DenseNet	55
3.4.1	Πυκνή συνδεσιμότητα (Dense connectivity)	55
3.4.2	Σύνθετη συνάρτηση (Composite function)	56
3.4.3	Στρώματα Συσσώρευσης (Pooling layers)	56
3.4.4	Ρυθμός ανάπτυξης (Growth rate)	56
3.4.5	Στρώματα Συμφόρησης (Bottleneck layers)	56
3.4.6	Συμπίεση (Compression)	57
3.4.7	Εφαρμογές	57
3.5	EfficientNet	57
3.5.1	Mobile inverted bottleneck MBConv	59
3.5.2	Squeeze-and-excitation	60
3.5.3	Εφαρμογές	61
3.6	CNN - RNN	61
4	Πειραματική Διαδικασία	63
4.1	Μέθοδοι και Παραμετροποίηση	63
4.2	Μετρικές αξιολόγησης	64
4.3	Αποτελέσματα	64
III	Επίλογος	69
5	Επίλογος	71
5.1	Συμπεράσματα	71
5.2	Μελλοντικές Επεκτάσεις	71
	Βιβλιογραφία	79

Κατάλογος Εικόνων

1.1	Στεφανιαία τομή MRI εγκεφάλου ενός υγιούς ατόμου (αριστερά) και ενός ατόμου με νόσο Alzheimer (δεξιά) [1]	20
2.1	Βιολογικός νευρώνας (αριστερά) και τεχνητός νευρώνας (δεξιά) [2]	26
2.2	Απλή αρχιτεκτονική ενός τεχνητού νευρωνικού δικτύου [3]	27
2.3	Αρχιτεκτονική δικτύου πρόσθιας τροφοδότησης (FNN) (αριστερά) και ανατροφοδοτούμενου δικτύου (RNN) (δεξιά) [3]	27
2.4	Γραφική αναπαράσταση των συναρτήσεων ενεργοποίησης [4]	30
2.5	Διακυμάνσεις του αλγορίθμου Stochastic Gradient Descent[5]	31
2.6	Ο αλγόριθμος SGD χωρίς τη χρήση της μεθόδου Momentum (αριστερά) και με αυτήν (δεξιά) [5]	32
2.7	Η βασική αρχιτεκτονική ενός συνελκτικού νευρωνικού δικτύου [6]	36
2.8	Διαδικασία Max-Pooling [6]	37
2.9	Αναπαράσταση ενός πλήρως συνδεδεμένου στρώματος [6]	38
2.10	Οπτικοποίηση των διαφορών μεταξύ νευρωνικών δικτύων πρόσθιας διάδοσης (αριστερά) και ανατροφοδοτούμενων νευρωνικών δικτύων (δεξιά) [7]	39
2.11	Διαφορετικές διαδικασίες εκμάθησης μεταξύ (a) της κλασικής μηχανικής μάθησης και (b) της μεταφοράς μάθησης [8]	44
3.1	Αρχιτεκτονική ενός δομικού block [9]	54
3.2	Αρχιτεκτονική των μοντέλων τύπου ResNet [9]	54
3.3	Ένα dense block 5 στρωμάτων με ρυθμό ανάπτυξης $k = 4$. Κάθε στρώμα λαμβάνει ως είσοδο όλους τους προηγούμενους χάρτες χαρακτηριστικών [10]	55
3.4	Ένα δίκτυο DenseNet με τρία dense blocks. Τα στρώματα ανάμεσα σε δύο γειτονικά blocks αναφέρονται σαν στρώματα μετάβασης [10]	56
3.5	Αρχιτεκτονική των μοντέλων τύπου DenseNet [10]	57
3.6	Συμβατικές μέθοδοι κλιμάκωσης (a-d), σε αντίθεση με την μέθοδο σύνθετης κλιμάκωσης (e) [11]	58
3.7	Η βασική αρχιτεκτονική του δικτύου EfficientNet-B0. Κάθε γραμμή περιγράφει ένα στάδιο i με \hat{L}_i στρώματα, ανάλυση εισόδου $\langle \hat{H}_i, \hat{W}_i \rangle$ και έξοδο τα κανάλια \hat{C}_i [11]	59
3.8	Residual block [12]	60
3.9	Inverted residual block [12]	60
3.10	Squeeze-and-excitation block [13]	61
4.1	Αριθμός των slices ανά MRI σε καθεμία από τις κλάσεις AD, MCI, CN	63

4.2	Αποτελέσματα του μοντέλου ResNet-50 στο σύνολο εκπαίδευσης	65
4.3	Αποτελέσματα του μοντέλου ResNet-50 στο σύνολο επαλήθευσης	65
4.4	Αποτελέσματα του μοντέλου DenseNet-121 στο σύνολο εκπαίδευσης	66
4.5	Αποτελέσματα του μοντέλου DenseNet-121 στο σύνολο επαλήθευσης	66
4.6	Αποτελέσματα του μοντέλου EfficientNet-B0 στο σύνολο εκπαίδευσης	67
4.7	Αποτελέσματα του μοντέλου EfficientNet-B0 στο σύνολο επαλήθευσης	67

Κατάλογος Πινάκων

4.1	Ακρίβεια του συνόλου εκπαίδευσης για τα μοντέλα Resnet-50, DenseNet-121, EfficientNet-B0	66
4.2	Ακρίβεια του συνόλου επαλήθευσης για τα μοντέλα Resnet-50, DenseNet-121, EfficientNet-B0	67

Εισαγωγή

1.1 Η νόσος Alzheimer

Η νόσος Alzheimer (Alzheimer's disease) αποτελεί μία προοδευτική, νευρολογική διαταραχή του εγκεφάλου και είναι η κύρια αιτία πρόκλησης της άνοιας [14]. Πρόκειται για μία νόσο η οποία πιστεύεται πως μπορεί να ξεκινάει 20 ή και περισσότερα χρόνια πριν την εμφάνιση των συμπτωμάτων, πραγματοποιώντας μικρές αλλαγές στον εγκέφαλο του ατόμου, οι οποίες είναι αρχικά μη παρατηρήσιμες. Έπειτα από εκτεταμένες εγκεφαλικές μεταβολές με την πάροδο των χρόνων, το άτομο αρχίζει να εμφανίζει αισθητά συμπτώματα τα οποία έγκειται κυρίως στην αλλοίωση των γνωστικών του λειτουργιών όπως είναι η απώλεια μνήμης και γλωσσικών δεξιοτήτων και η αδυναμία εκτέλεσης καθημερινών εργασιών.

Καθώς η ασθένεια εξελίσσεται, οι νευρώνες σε άλλα σημεία του εγκεφάλου εκφυλίζονται ή και καταστρέφονται. Αυτό έχει ως αποτέλεσμα, δραστηριότητες οι οποίες αποτελούσαν αναπόσπαστο κομμάτι της ζωής του ατόμου να καθίστανται πλέον αδύνατες ως προς την εκτέλεσή τους. Τελικά, επηρεάζονται νευρώνες σε σημεία του εγκεφάλου τα οποία είναι υπεύθυνα για βασικές σωματικές λειτουργίες όπως το περπάτημα ή η κατανάλωση τροφής. Οι ασθενείς στα τελευταία στάδια της νόσου Alzheimer χρήζουν συνεχής παρακολούθησης από τρίτους, με συνήθη κατάληξη το θάνατο.

1.2 Διάγνωση της νόσου Alzheimer

Η έγκαιρη και έγκυρη διάγνωση της νόσου Alzheimer αποτελεί μία από τις μεγαλύτερες προκλήσεις της ιατρικής κοινότητας και τη δεδομένη στιγμή δεν υφίσταται κάποιο μεμονωμένο και εξειδικευμένο τεστ για το σκοπό αυτό. Μία ορθή διάγνωση θα μπορούσε να συμβάλλει στην πρόληψη ή και τον περιορισμό της ασθένειας του Alzheimer, σώζοντας τη ζωή εκατομμυρίων ανθρώπων. Η διαδικασία της διάγνωσης συγκροτείται από ένα σύνολο ποικίλων προσεγγίσεων και εργαλείων, οι οποίες τελούνται με τη συνεργασία διαφόρων ιατρικών ειδικοτήτων όπως νευρολόγοι, ψυχίατροι και ακτινολόγοι και περιλαμβάνει τα εξής:

- Λήψη ιατρικού και οικογενειακού ιστορικού από το άτομο, συμπεριλαμβανομένου του ψυχιατρικού ιστορικού και του ιστορικού γνωστικών και συμπεριφορικών αλλαγών.
- Παροχή πληροφοριών από το οικογενειακό περιβάλλον για τις μεταβολές του ατόμου που αφορούν στις γνωστικές και συμπεριφορικές του δεξιότητες.

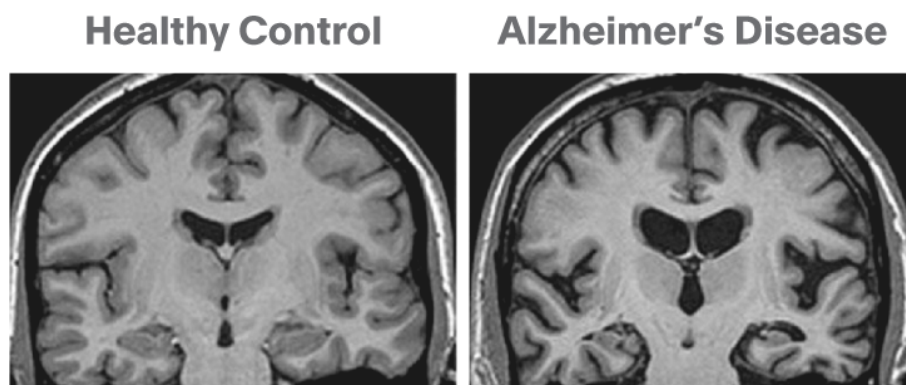
- Διενέργεια γνωστικών τεστ και σωματικών και νευρολογικών εξετάσεων.
- Υποβολή του ατόμου σε εξετάσεις αίματος και απεικόνισης εγκεφάλου για να αποκλειστούν άλλες πιθανές αιτίες συμπτωμάτων άνοιας, όπως ένας όγκος ή ορισμένες ανεπάρκειες βιταμινών.

1.2.1 Απεικόνιση Μαγνητικής Τομογραφίας (Magnetic Resonance Imaging)

Όσον αφορά στην απεικόνιση του εγκεφάλου ως μέθοδο διάγνωσης, αυτή συναντάται με ποικίλες μεθόδους όπως είναι οι ακτινογραφίες (X-rays), οι positron-emission tomography (PET) scans, οι computed tomography (CT) scans και οι magnetic resonance imaging (MRI) scans. Στην παρούσα εργασία έχει επιλεγθεί η κατηγοριοποίηση εικόνων που προήλθαν από MRIs εγκεφάλου.

Οι MRIs χρησιμοποιούνται για την απόκτηση λεπτομερών εικόνων του εγκεφάλου. Αυτό είναι σημαντικό για την έρευνα σχετικά με την άνοια, καθώς μπορεί να αποκαλύψει εάν τμήματα του εγκεφάλου μεταβάλλονται και να δείξει στους ερευνητές πώς η ασθένεια επηρεάζει τον εγκέφαλο [15]. Ένα μαγνητικό πεδίο, ραδιοσυχνότητες και ένας υπολογιστής μπορούν να δημιουργήσουν εικόνες των περισσότερων εσωτερικών τμημάτων του σώματος. Οι MRIs του εγκεφάλου συμπληρώνουν το έργο των ιατρών και τους βοηθούν να εντοπίσουν τυχόν ανωμαλίες του εγκεφάλου και να προβλέψουν ακόμη αν ελλοχεύει ο κίνδυνος ανάπτυξης της νόσου Alzheimer [16]. Οι MRIs μπορούν να γνωστοποιήσουν αν οι άνθρωποι αναπτύσσουν διαφορετικούς τύπους άνοιας και να υποδείξουν πώς ανταποκρίνονται στις θεραπείες.

Η διάγνωση της νόσου Alzheimer μέσω MRIs αποτελεί μία μέθοδο που έχει εφαρμοστεί από επιστήμονες τόσο του κλάδου της ιατρικής [17], [18] όσο κι εκείνου της μηχανικής μάθησης [19], [20], για αυτό και επιλέχθηκε ως μέθοδος απεικόνισης στην παρούσα εργασία. Συγκεκριμένα, επειδή η νόσος επηρεάζει τα σημεία εκείνα του εγκεφάλου στα οποία βρίσκονται ο ενδορινικός φλοιός και ο ιππόκαμπος, η πιο αντιπροσωπευτική τους απεικόνιση γίνεται στις στεφανιαίες τομές του εγκεφάλου οι οποίες και εξήχθησαν εδώ.



Εικόνα 1.1: Στεφανιαία τομή MRI εγκεφάλου ενός υγιούς ατόμου (αριστερά) και ενός ατόμου με νόσο Alzheimer (δεξιά) [1]

Ωστόσο, πολλές φορές ακόμη και μετά από τη διεξαγωγή των εξετάσεων που αναφέρονται δεν παρέχονται επαρκή αποτελέσματα ώστε να ταυτοποιηθεί η νόσος Alzheimer ως αιτία της

άνοιας. Η πιο ακριβής μέθοδος μέχρι και σήμερα είναι εκείνη της αυτοψίας [21], η οποία ενέχει τον κίνδυνο του ανθρώπινου λάθους.

1.3 Σκοπός και δομή της εργασίας

Όπως γίνεται ευκόλως αντιληπτό από τα ανωτέρω, κρίνεται απαραίτητη η θεμελίωση μίας άρτιας και έγκυρης μεθόδου διάγνωσης για τη νόσο Alzheimer, εγχείρημα το οποίο καλείται να εκπονήσει η παρούσα εργασία μέσω της διάγνωσης της νόσου από μαγνητικές τομογραφίες εγκεφάλου με τη χρήση τεχνικών βαθιάς μάθησης. Η διάρθρωση της εργασίας γίνεται ως εξής:

Στο κεφάλαιο 2, αναπτύσσεται το θεωρητικό υπόβαθρο που απαιτείται για την κατανόηση των διάφορων μεθόδων και εννοιών που πραγματεύεται η εργασία.

Στο κεφάλαιο 3, περιγράφεται η αρχιτεκτονική του μοντέλου που χρησιμοποιείται.

Στο κεφάλαιο 4, παρουσιάζεται η πειραματική διαδικασία καθώς και τα αποτελέσματα που εξήχθησαν.

Στο τελευταίο κεφάλαιο 5, αναφέρονται τα συμπεράσματα που προέκυψαν καθώς και οι πιθανές μελλοντικές επεκτάσεις της παρούσας έρευνας.

Μέρος I

Θεωρητικό Μέρος

Κεφάλαιο 2

Θεωρητικό υπόβαθρο

2.1 Νευρωνικά Δίκτυα

Τα Νευρωνικά Δίκτυα υπόκεινται στο πεδίο της Μηχανικής Μάθησης και της Τεχνητής Νοημοσύνης και αποτελούν μία αλγοριθμική προσέγγιση επίλυσης υπολογιστικών προβλημάτων, μιμούμενα τη λειτουργία των βιολογικών νευρωνικών δικτύων. Η διαδικασία αυτή πραγματοποιείται μέσω της εκπαίδευσης του δικτύου σε ένα σύνολο παραδειγμάτων τα οποία αφορούν στο εκάστοτε πρόβλημα προς επίλυση.

2.1.1 Βιολογικά Νευρωνικά Δίκτυα

Ο ανθρώπινος εγκέφαλος είναι μία από τις πιο περίπλοκες και ταυτοχρόνως θαυμαστές δομές στη φύση, με την κατανόηση της λειτουργίας του να αποτελεί αίνιγμα για την επιστημονική κοινότητα.

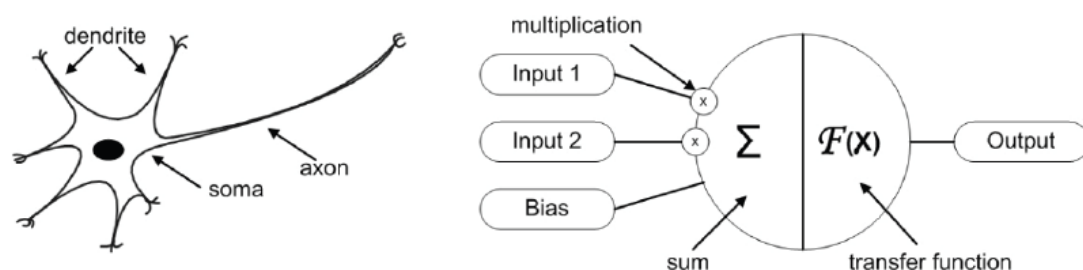
Βασική δομική μονάδα των νευρωνικών δικτύων του εγκεφάλου αποτελούν οι νευρώνες, οι οποίοι παρά τις επιμέρους μορφολογικές διαφορές που παρουσιάζουν ανάλογα με την λειτουργία τους, διαθέτουν κάποια κοινά δομικά χαρακτηριστικά που είναι τα εξής: το κυτταρικό σώμα, οι συνάψεις, οι δενδρίτες και ο άξονας. Οι 10^{11} νευρώνες που απαρτίζουν τον εγκέφαλο, σχηματίζουν ένα δίκτυο επικοινωνίας μέσω των συνάψεων, με τους δενδρίτες να λειτουργούν ως είσοδο και τον άξονα ως έξοδο της σύνδεσης αυτής. Ο κάθε νευρώνας μπορεί να επικοινωνεί με χιλιάδες άλλους νευρώνες με αποτέλεσμα το πλήθος των συνάψεων να υπερβαίνει συνολικά τις 10^{14} .

Ενώ κάθε νευρώνας μεμονωμένα διαθέτει ένα σχετικά αργό σύστημα επεξεργασίας των πληροφοριών που λαμβάνει (χρονική κλίμακα της τάξης του 1 ms), ο εκτεταμένος παραλληλισμός επεξεργασίας πληροφοριών σε πολλαπλές συνάψεις ταυτόχρονα, οδηγεί σε μία επεξεργαστική ισχύ η οποία είναι ανώτερη ακόμη και από εκείνη των σύγχρονων υπερυπολογιστών.

Όταν ένας νευρώνας πυροδοτείται, στέλνει έναν ηλεκτρικό παλμό ο οποίος μεταδίδεται στις συνάψεις και προκαλεί την απελευθέρωση χημικών νευρο-διαβιβαστών στους επόμενους νευρώνες. Το είδος της εκάστοτε σύναψης μπορεί είτε να αυξήσει είτε να μειώσει την πιθανότητα πυροδότησης ενός μεταγενέστερου νευρώνα. Αυτό συμβαίνει καθώς κάθε σύναψη είναι συσχετισμένη με ένα «βάρος», βάσει του οποίου αποφασίζεται το μέγεθος της επίδρασης του ηλεκτρικού παλμού που αποστέλλεται σε επόμενους νευρώνες. Στη συνέχεια υπολογίζεται ένα άθροισμα των βαρών που έχει λάβει ο εκάστοτε νευρώνας και αν αυτό υπερβαίνει ένα

συγκεκριμένο κατώφλι, εκείνος πυροδοτείται.

Τόσο τα βιολογικά όσο και τα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα να διαφοροποιούν τις αποκρίσεις τους συναρτήσει των εξωτερικών ερεθισμάτων που λαμβάνουν. Η ικανότητα αυτή αναφέρεται γενικά ως εκμάθηση και συμβαίνει κυρίως μέσω των αλλαγών στα βάρη των συνάψεων [2].



Εικόνα 2.1: Βιολογικός νευρώνας (αριστερά) και τεχνητός νευρώνας (δεξιά) [2]

2.1.2 Τεχνητά Νευρωνικά Δίκτυα

Όπως και στα βιολογικά, έτσι και στα τεχνητά νευρωνικά δίκτυα βασική δομική μονάδα είναι ο τεχνητός νευρώνας, ο οποίος παρουσιάζει σημαντικές ομοιότητες με τον βιολογικό ως προς την δομή και τη λειτουργία του.

Στον τεχνητό νευρώνα η πληροφορία διαβιβάζεται μέσω ενός πλήθους εισόδων οι οποίες πολλαπλασιάζονται με τα αντίστοιχα βάρη τους [3]. Στη συνέχεια, οι σταθμισμένες αυτές εισοδοί αθροίζονται μέσω του κεντρικού σώματος του νευρώνα και το άθροισμα αυτό μεταφέρεται στις εξόδους του. Η επεξεργασμένη πληροφορία δίνεται ως είσοδος σε μία συνάρτηση ενεργοποίησης, ενώ σε πολλές περιπτώσεις στο αποτέλεσμα προστίθεται και ένα μέγεθος που ονομάζεται πόλωση (bias). Η παραπάνω διαδικασία περιγράφεται στην παρακάτω μαθηματική σχέση:

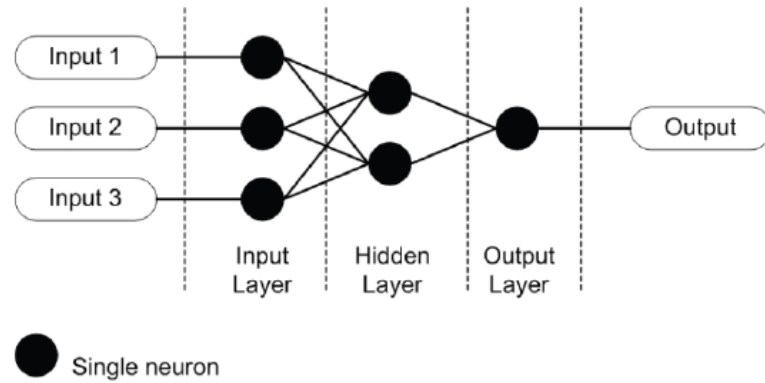
$$y(k) = F\left(\sum_{i=0}^m w_i(k)x_i(k) + b\right) \quad (2.1)$$

όπου έχουμε:

- $\mathbf{x}_i(k)$: η τιμή της εισόδου i σε μια διακριτή χρονική στιγμή k όπου το i κυμαίνεται μεταξύ του 0 και του m
- $\mathbf{w}_i(k)$: η τιμή του βάρους σε μια διακριτή χρονική στιγμή k όπου το i κυμαίνεται μεταξύ του 0 και του m
- \mathbf{b} : η τιμή της πόλωσης (bias)
- \mathbf{F} : η συνάρτηση ενεργοποίησης
- $\mathbf{y}_i(k)$: η τιμή της εξόδου σε μία διακριτή χρονική στιγμή k

Ο συνδυασμός δύο ή περισσότερων τεχνητών νευρώνων δημιουργεί ένα τεχνητό νευρωνικό δίκτυο. Τα τεχνητά νευρωνικά δίκτυα έχουν την ικανότητα να επιλύουν σύνθετα προβλήματα

του πραγματικού κόσμου, γεγονός που επιτυγχάνεται μέσω της επεξεργασίας της πληροφορίας από τα δομικά τους στοιχεία (τεχνητοί νευρώνες), με μη γραμμικό, καταναμημένο, παράλληλο και τοπικό τρόπο.

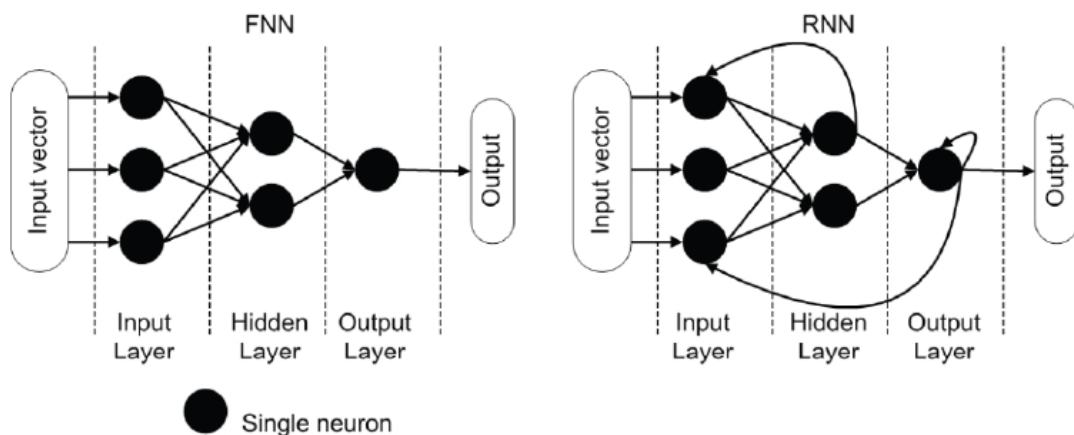


Εικόνα 2.2: Απλή αρχιτεκτονική ενός τεχνητού νευρωνικού δικτύου [3]

Ο τρόπος με τον οποίο διασυνδέονται μεταξύ τους οι τεχνητοί νευρώνες ονομάζεται τοπολογία, αρχιτεκτονική ή γράφος του τεχνητού νευρωνικού δικτύου.

Η διασύνδεση αυτή επιτυγχάνεται με πολυάριθμους τρόπους, οι οποίοι διαχωρίζονται σε δύο βασικές κατηγορίες: τα δίκτυα πρόσθιας τροφοδότησης (feed-forward networks) και τα ανατροφοδοτούμενα δίκτυα (recurrent networks). Στην τοπολογία πρόσθιας τροφοδότησης, η ροή της πληροφορίας πραγματοποιείται από το input στο output layer μόνο μέσω μίας κατεύθυνσης, ενώ στην τοπολογία ενός ανατροφοδοτούμενου δικτύου η πληροφορία μεταδίδεται στα στρώματα με περισσότερες από μία κατευθύνσεις, αφού η έξοδος ενός στρώματος μπορεί να χρησιμοποιηθεί ως είσοδος σε στρώμα ενός ίδιου ή και προηγούμενου επιπέδου.

Οι παραπάνω κατηγορίες τοπολογίας απεικονίζονται στο σχήμα 2.3 .



Εικόνα 2.3: Αρχιτεκτονική δικτύου πρόσθιας τροφοδότησης (FNN) (αριστερά) και ανατροφοδοτούμενου δικτύου (RNN) (δεξιά) [3]

2.1.3 Συναρτήσεις ενεργοποίησης

Όπως φαίνεται από την εξίσωση 2.1, η συνάρτηση ενεργοποίησης αποτελεί τον άγνωστο παράγοντα του μοντέλου του τεχνητού νευρωνικού δικτύου και είναι εκείνη που καθορίζει τις ιδιότητές του [3]. Τα τεχνητά νευρωνικά δίκτυα χωρίς την χρήση κάποιας συνάρτησης ενεργοποίησης λειτουργούν σαν μοντέλα γραμμικής παλινδρόμησης, γεγονός που τα καθιστά περιορισμένα ως προς την απόδοση και την ισχύ τους [22].

Οι βασικές συναρτήσεις ενεργοποίησης που χρησιμοποιούνται ευρέως είναι οι εξής:

1. Γραμμική συνάρτηση (Linear function)

Η γραμμική συνάρτηση ενεργοποίησης είναι ανάλογη με την τιμή της εισόδου και ορίζεται ως:

$$F(x) = ax \quad (2.2)$$

όπου η τιμή της μεταβλητής a μπορεί να είναι οποιαδήποτε σταθερά, η οποία επιλέγεται από το χρήστη.

Την παράγωγο της γραμμικής συνάρτησης ενεργοποίησης και κατά συνέπεια και την κλίση της, αποτελεί η σταθερά a . Η σταθερά αυτή είναι ανεξάρτητη της τιμής x της εισόδου, γεγονός που υποδηλώνει ότι τα βάρη και η πόλωση κατά την εφαρμογή του αλγόριθμου οπίσθιας διάδοσης θα ενημερώνονται, παρόλο που ο παράγοντας ενημέρωσης θα παραμένει αμετάβλητος. Παράλληλα, η σταθερή κλίση της συνάρτησης λειτουργεί ως τροχοπέδη στην βελτίωση της απόδοσης του δικτύου σε κάθε επανάληψη, συνεπώς και στην εκπαίδευσή του. Τέλος, το δίκτυο δεν έχει την δυνατότητα να αναγνωρίσει πιο περίπλοκα μοτίβα στα δεδομένα που δίνονται ως είσοδος.

Ωστόσο, η γραμμική συνάρτηση ενεργοποίησης αποτελεί την ιδανική επιλογή σε περιπτώσεις που απαιτείται ερμηνευσιμότητα αλλά και σε απλές εργασίες.

2. Σιγμοειδής συνάρτηση (Sigmoid function)

Η σιγμοειδής συνάρτηση είναι μία μη γραμμική, συνεχής και παραγωγίσιμη συνάρτηση με σύνολο τιμών το διάστημα $[0, 1]$. Περιγράφεται από την εξίσωση:

$$F(x) = \frac{1}{1 + e^{-x}} \quad (2.3)$$

Το γεγονός πως το σύνολο τιμών της συνάρτησης είναι το διάστημα $[0, 1]$, έχει ως αποτέλεσμα την συσσώρευση ενός μεγάλου χώρου εισόδων σε έναν πιο περιορισμένο χώρο [23]. Με αυτόν τον τρόπο, μία αισθητή μεταβολή στις τιμές της εισόδου μπορεί να σημειώσει μικρή μεταβολή στις τιμές της εξόδου καθώς τα βάρη και η πόλωση του μοντέλου δεν θα ενημερώνονται σωστά κατά την εκπαίδευση. Το πρόβλημα αυτό ονομάζεται πρόβλημα της εξαφάνισης κλίσεων (ανισοινηγ γραδιεντ προβλεμ) και γίνεται εντονότερο όσο αυξάνονται τα στρώματα του δικτύου. Επομένως, η χρήση της σιγμοειδούς συνάρτησης δεν συνιστάται για βαθιά νευρωνικά δίκτυα.

3. Συνάρτηση υπερβολικής εφαπτομένης (Hyperbolic Tangent function)

Η συνάρτηση της υπερβολικής εφαπτομένης αποτελεί μία παραλλαγή της σιγμοειδούς συνάρτησης, με σύνολο τιμών το διάστημα $[-1, 1]$ και ορίζεται ως εξής:

$$F(x) = \tanh(x) = 2 \operatorname{sigmoid}(2x) - 1 \quad (2.4)$$

όπου η συνάρτηση $\operatorname{sigmoid}$ δίνεται από την σχέση 2.3.

Όπως και η σιγμοειδής, έτσι και η συνάρτηση υπερβολικής εφαπτομένης, συσσωρεύει τις εισόδους σε ένα περιορισμένο διάστημα και συγκεκριμένα στο $[-1, 1]$. Σε αντίθεση με τη σιγμοειδή όμως, εδώ δεν υφίσταται το πρόβλημα της εξαφάνισης κλίσεων με αποτέλεσμα να υπάρχει μεγαλύτερος ρυθμός εκπαίδευσης του μοντέλου και ταχύτερη σύγκλιση.

4. Συνάρτηση ReLU (ReLU function)

Η συνάρτηση ReLU ορίζεται ως:

$$F(x) = \max(0, x) = \begin{cases} x & \text{αν } x \geq 0 \\ 0 & \text{αν } x < 0 \end{cases} \quad (2.5)$$

δηλαδή όταν δοθεί μία θετική (ή μηδενική τιμή) x , τότε η συνάρτηση ισούται με την τιμή x αυτή, ενώ αν δοθεί αρνητική τιμή x , τότε ισούται με το 0.

Η συνάρτηση αυτή διαθέτει υψηλή υπολογιστική απόδοση, καθώς δεν απαιτεί τον υπολογισμό εκθετικών συναρτήσεων όπως η σιγμοειδής, και συγκλίνει πολύ γρήγορα. Ταυτοχρόνως, λόγω της μηδενικής εξόδου στην περίπτωση αρνητικής εισόδου, επιτρέπει στο δίκτυο να αποκτήσει εύκολα αραιή αναπαράσταση (sparse representation), παράγοντας που συμβάλλει στην απλοποίηση του μοντέλου και στην επιτάχυνση της μάθησης [23], [24].

5. Συνάρτηση Softmax (Softmax function)

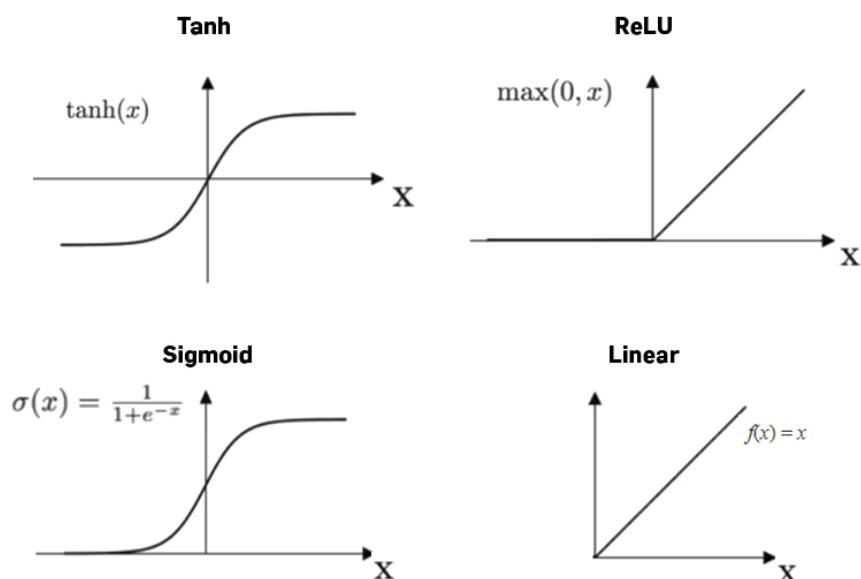
Η συνάρτηση Softmax αποτελεί συνδυασμό πολλαπλών σιγμοειδών συναρτήσεων. Ο τρόπος με τον οποίο λειτουργεί είναι μετατρέποντας ένα διάνυσμα K πραγματικών τιμών, σε ένα διάνυσμα K πραγματικών τιμών που αθροίζονται στο 1. Η μετατροπή αυτή πλαισιώνει ένα σύνολο τιμών στο διάστημα $[0, 1]$, με αποτέλεσμα οι τιμές που προκύπτουν να μπορούν να μεταφραστούν σαν πιθανότητες.

Ορίζεται ως:

$$\sigma(\vec{x})_i = \frac{e^{x_i}}{\sum_{j=1}^K e^{x_j}} \quad (2.6)$$

Σε αντίθεση με την σιγμοειδή συνάρτηση, η οποία μπορεί να χρησιμοποιηθεί μόνο για δυαδική κατηγοριοποίηση κλάσεων, η συνάρτηση Softmax χρησιμοποιείται και σε μοντέλα με πολλαπλές κλάσεις.

Στην εικόνα 2.4 παρουσιάζεται μία γραφική απεικόνιση των συναρτήσεων ενεργοποίησης που αναλύθηκαν παραπάνω [4].



Εικόνα 2.4: Γραφική αναπαράσταση των συναρτήσεων ενεργοποίησης [4]

2.1.4 Αλγόριθμοι Βελτιστοποίησης

Η χρήση των αλγόριθμων βελτιστοποίησης αποτελεί πολύτιμο εργαλείο του πεδίου των νευρωνικών δικτύων αφού συμβάλλει σημαντικά στη βελτίωση της επίδοσής τους.

Αλγόριθμος Επικλινούς Καθόδου (Gradient Descent Algorithm)

Ο αλγόριθμος επικλινούς καθόδου είναι ένας από τους πιο δημοφιλείς αλγορίθμους βελτιστοποίησης και αποτελεί τον συνηθέστερο τρόπο βελτιστοποίησης των νευρωνικών δικτύων [5].

Η επικλινής κάθοδος είναι ένας τρόπος ελαχιστοποίησης μίας συνάρτησης κόστους $L(\theta)$, όπου θ οι παράμετροι του δικτύου, ενημερώνοντας τις παραμέτρους προς την αντίθετη κατεύθυνση της κλίσης της συνάρτησης κόστους $\nabla_{\theta} \mathcal{L}(\theta)$. Ο ρυθμός εκμάθησης (learning rate) καθορίζει το μέγεθος των βημάτων που απαιτούνται ώστε να επιτευχθεί ένα (τοπικό) ελάχιστο. Ουσιαστικά, ακολουθείται η κατεύθυνση της καμπύλης της επιφάνειας που σχηματίζεται από τη συνάρτηση κόστους μέχρι μία "κοιλάδα".

Υπάρχουν τρεις παραλλαγές του αλγορίθμου, με ειδοποιό διαφορά τον όγκο των δεδομένων που χρησιμοποιούνται για τον υπολογισμό της κλίσης της συνάρτησης κόστους. Με βάση τον όγκο αυτό πραγματοποιείται μια αντιστάθμιση μεταξύ της ακρίβειας της ενημέρωσης παραμέτρων και του χρόνου που απαιτείται για την εκτέλεση μιας ενημέρωσης.

1. Batch gradient descent

Ο batch gradient descent υπολογίζει την κλίση της συνάρτησης κόστους ως προς τις παραμέτρους θ για ολόκληρο το σύνολο δεδομένων εκπαίδευσης:

$$\theta = \theta - \eta \cdot \nabla_{\theta} \mathcal{L}(\theta) \quad (2.7)$$

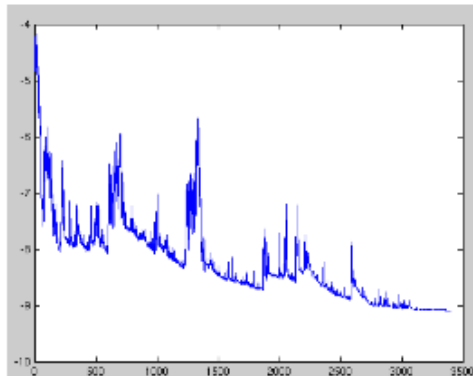
Επειδή απαιτείται ο υπολογισμός των κλίσεων ολόκληρου του συνόλου δεδομένων για μία μεμονωμένη ενημέρωση, ο αλγόριθμος batch gradient descent σημειώνει χαμηλή ταχύτητα εκτέλεσης και δεν λειτουργεί σε σύνολα δεδομένα που υπερβαίνουν την χωρητικότητα της μνήμης.

2. Stochastic gradient descent

Σε αντίθεση με τον αλγόριθμο που προαναφέρθηκε, ο stochastic gradient descent υλοποιεί την ενημέρωση των παραμέτρων για κάθε σύνολο εκπαίδευσης $x^{(i)}$ και ετικέτα $y^{(i)}$:

$$\theta = \theta - \eta \cdot \nabla_{\theta} \mathcal{L}(\theta; x^{(i)}; y^{(i)}) \quad (2.8)$$

Ο batch gradient descent πραγματοποιεί πολυάριθμους περιττούς υπολογισμούς σε μεγάλα σύνολα δεδομένων, καθώς υπολογίζει την κλίση του εκάστοτε δείγματος πριν από κάθε ενημέρωση παραμέτρων. Σε αντίθεση με αυτήν την προσέγγιση, ο stochastic gradient descent πραγματοποιεί μία ενημέρωση κάθε φορά, επιτυγχάνοντας υψηλότερη ταχύτητα και πιο σταθερή σύγκλιση αλλά δημιουργώντας μεγαλύτερο θόρυβο.



Εικόνα 2.5: Διακυμάνσεις του αλγορίθμου Stochastic Gradient Descent[5]

3. Mini-batch gradient descent

Ο mini-batch gradient descent αποτελώντας κράμα των δύο προηγούμενων αλγορίθμων, συνδυάζει τα πλεονεκτήματά τους και ενημερώνει κάθε φορά ένα υποσύνολο n παραδειγμάτων εκπαίδευσης:

$$\theta = \theta - \eta \cdot \nabla_{\theta} \mathcal{L}(\theta; x^{(i:i+n)}; y^{(i:i+n)}) \quad (2.9)$$

Με την υλοποίηση αυτή επιτυγχάνεται τόσο η μείωση της απόκλισης των ενημερώσεων των παραμέτρων και κατά συνέπεια μία σταθερότερη σύγκλιση, όσο και η υψηλή υπολογιστική απόδοση.

- **Momentum**

Ο αλγόριθμος stochastic gradient descent, παρουσιάζει πρόβλημα σε περιοχές όπου η επιφάνεια καμπυλώνει πολύ πιο απότομα στη μία διάσταση σε σχέση με την άλλη,

φαινόμενο το οποίο είναι σύνηθες γύρω από τα τοπικά βέλτιστα. Στις περιπτώσεις αυτές, ο αλγόριθμος ταλαντώνεται σε εκείνα τα σημεία, σημειώνοντας χαμηλή πρόοδο και ταχύτητα μέχρι την άφιξή του στο τοπικό βέλτιστο.

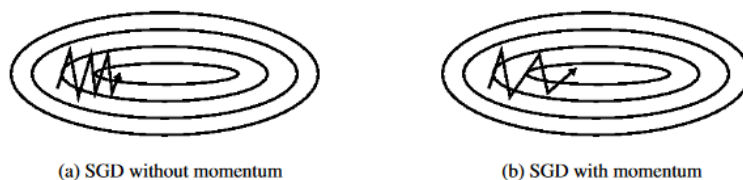
Η μέθοδος Momentum συνεπικουρεί στην επιτάχυνση του αλγορίθμου, μετριάζοντας τις ταλαντώσεις. Η βελτίωση αυτή οφείλεται στην προσθήκη του παράγοντα γ του διανύσματος ενημέρωσης του προηγούμενου βήματος, σε εκείνο του τρέχοντος:

$$v_t = \gamma v_{t-1} + \eta \nabla_{\theta} \mathcal{L}(\theta) \quad (2.10)$$

$$\theta = \theta - v_t \quad (2.11)$$

Ο παράγοντας γ συνήθως ορίζεται στην τιμή 0,9 ή σε παραπλήσιες τιμές.

Η χρήση της μεθόδου Momentum μπορεί να παραλληλιστεί με το εξής πείραμα. Έστω ότι ρίχνουμε μία μπάλα σε μία κοιλάδα. Η μπάλα όσο κατεβαίνει προς το χαμηλότερο σημείο της κοιλάδας, θα συσσωρεύει ορμή, κυλώντας όλο και ταχύτερα μέχρι να φτάσει σε μία τελική ταχύτητα. Το ίδιο παρατηρείται και στις ενημερώσεις των παραμέτρων: ο όρος momentum αυξάνεται στις διαστάσεις που έχουν σταθερή κατεύθυνση κλίσης και μειώνεται σε εκείνες των οποίων η κατεύθυνση της κλίσης είναι μεταβλητή. Με τον τρόπο αυτό μειώνεται η ταλάντωση και επιτυγχάνεται ταχύτερη σύγκλιση.



Εικόνα 2.6: Ο αλγόριθμος SGD χωρίς τη χρήση της μεθόδου Momentum (αριστερά) και με αυτήν (δεξιά) [5]

- **Nesterov accelerated gradient**

Σε συνέχεια χρήσης του προηγούμενου παραδείγματος, ιδανικά θα ήταν επιθυμητή η ύπαρξη μίας μπάλας η οποία θα έχει επίγνωση της κίνησής της και θα διαθέτει την ικανότητα να μειώσει την ταχύτητά της πριν η κλίση της κοιλάδας αρχίζει να γίνεται ξανά ανοδική.

Η μέθοδος Nesterov accelerated gradient επιτρέπει στον παράγοντα της ορμής ο οποίος προστέθηκε παραπάνω, να αποκτήσει την ικανότητα αυτής της πρόγνωσης. Για το σκοπό αυτό θα χρησιμοποιηθεί ο όρος γv_{t-1} ώστε να μετακινηθούν οι παράμετροι θ . Ο υπολογισμός του όρου $\theta - \gamma v_{t-1}$ δίνει μία προσέγγιση της επόμενης θέσης των παραμέτρων, ως προς τις οποίες θα γίνει και ο υπολογισμός της κλίσης:

$$v_t = \gamma v_{t-1} + \eta \nabla_{\theta} \mathcal{L}(\theta - \gamma v_{t-1}) \quad (2.12)$$

$$\partial = \partial - v_t \quad (2.13)$$

Ο παράγοντας γ και σε αυτήν την μέθοδο ορίζεται στην τιμή 0,9 ή σε παραπλήσιες τιμές.

Ενώ η μέθοδος Momentum υπολογίζει αρχικά την τρέχουσα κλίση και στη συνέχεια κάνει ένα μεγάλο άλμα στην κατεύθυνση της ενημερωμένης συσσωρευμένης κλίσης, η παρούσα μέθοδος κάνει πρώτα ένα μεγάλο άλμα στην κατεύθυνση της προηγούμενης συσσωρευμένης κλίσης, μετράει την κλίση και μετέπειτα πραγματοποιεί την απαιτούμενη διόρθωση. Αυτή η προκαταρκτική ενημέρωση μας εμποδίζει να προχωρήσουμε πολύ γρήγορα και έχει ως αποτέλεσμα αυξημένη απόκριση, η οποία έχει ενισχύσει σημαντικά την απόδοση των RNN σε έναν αριθμό εργασιών.

- **Adam**

Ο αλγόριθμος Adam (Adaptive Moment Estimation) αποτελεί μία μέθοδο η οποία προσαρμόζει το ρυθμό εκπαίδευσης για κάθε παράμετρο του μοντέλου. Πιο συγκεκριμένα, αποθηκεύει έναν εκθετικό κινούμενο μέσο όρο τόσο των κλίσεων όσο και των τετραγώνων τους (m_t , v_t αντίστοιχα):

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (2.14)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (2.15)$$

όπου τα m_t , v_t είναι εκτιμήσεις της ορμής πρώτης τάξης (μέση τιμή) και της ορμής δεύτερης τάξης (διασπορά) των κλίσεων αντίστοιχα. Καθώς τα μεγέθη m_t , v_t αρχικοποιούνται σε μηδενικά διανύσματα, παρατηρείται ότι παρουσιάζουν πόλωση στο 0. Προκειμένου να ισορροπηθεί η πόλωση αυτή έχουμε τις εξής εκτιμήσεις ορμής πρώτης και δεύτερης τάξης:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (2.16)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (2.17)$$

Τελικά, οι ενημερώσεις των παραμέτρων υπολογίζονται από την μαθηματική σχέση:

$$\partial_{t+1} = \partial_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t \quad (2.18)$$

2.1.5 Αλγόριθμος Οπισθοδιάδοσης (Back-Propagation Algorithm)

Ο αλγόριθμος της οπισθοδιάδοσης αποτελεί μία από τις πιο θεμελιώδεις μεθόδους στον τομέα των νευρωνικών δικτύων και η χρήση του επαφίεται στην αποτελεσματική εκπαίδευση του δικτύου.

Ο αλγόριθμος αυτός προσαρμόζει επανειλημμένα τα βάρη και τις πολώσεις των συνδέσεων στο δίκτυο με τέτοιο τρόπο ώστε να ελαχιστοποιηθεί η διαφορά μεταξύ του πραγματικού διανύσματος εξόδου με το επιθυμητό, δηλαδή η συνάρτηση κόστους. Ο βαθμός προσαρμογής εξαρτάται από τις κλίσεις της συνάρτησης κόστους στις αντίστοιχες παραμέτρους [25].

Παρακάτω περιγράφεται η κεντρική ιδέα, καθώς και το μαθηματικό υπόβαθρο που πλαισιώνει τον αλγόριθμο οπισθοδιάδοσης.

Έστω νευρωνικό δίκτυο n στρωμάτων και W^n , b^n , x^n τα βάρη, οι πολώσεις και οι εξοδοί των στρωμάτων αντίστοιχα. Αρχικά θα πραγματοποιηθεί η πρόσθια διάδοση στο δίκτυο (feed-forward). Ορίζεται διάνυσμα:

$$z^n = W^n x^{n-1} + b^n \quad (2.19)$$

Στη συνέχεια εφαρμόζεται η συνάρτηση ενεργοποίησης $F(x)$, επομένως η έξοδος του εκάστοτε στρώματος διαμορφώνεται ως εξής:

$$a^n = F(z^n) \quad (2.20)$$

Το τελικό κομμάτι του νευρωνικού δικτύου το οποίο αποτελεί το στρώμα εξόδου, θα υπολογίσει την προβλεπόμενη τιμή η οποία θα είναι:

$$s = W^n a^n \quad (2.21)$$

Το τελευταίο βήμα ώστε να ολοκληρωθεί η πρόσθια διάδοση είναι η αξιολόγηση της τιμής της προβλεπόμενης εξόδου (s) σε σύγκριση με εκείνη της επιθυμητής (y). Αυτό επιτυγχάνεται μέσω της συνάρτησης κόστους:

$$\mathcal{L} = \text{cost}(s, y) \quad (2.22)$$

Έπειτα ακολουθεί το στάδιο της οπισθοδιάδοσης. Αρχικά θα πρέπει να υπολογιστεί η παράγωγος, δηλαδή η κλίση της συνάρτησης κόστους, η οποία θα υποδείξει τις απαιτούμενες μεταβολές των παραμέτρων ώστε να ελαχιστοποιηθεί η συνάρτηση κόστους. Για τον υπολογισμό των παραγώγων αυτών θα γίνει χρήση του κανόνα της αλυσίδας. Για ένα μεμονωμένο βάρος w_{jk}^n ισχύει:

$$\frac{\partial \mathcal{L}}{\partial w_{jk}^n} = \frac{\partial \mathcal{L}}{\partial z_j^n} \frac{\partial z_j^n}{\partial w_{jk}^n} \quad (2.23)$$

λόγω του κανόνα της αλυσίδας. Επιπλέον:

$$z_j^n = \sum_{k=1}^m w_{jk}^n a_k^{n-1} + b_j^n \quad (2.24)$$

όπου m ο αριθμός των νευρώνων στο στρώμα $n-1$. Με παραγωγήση:

$$\frac{\partial z_j^n}{\partial w_{jk}^n} = a_k^{n-1} \quad (2.25)$$

και τελικά:

$$\frac{\partial C}{\partial w_{jk}^n} = \frac{\partial \mathcal{L}}{\partial z_j^n} a_k^{n-1} \quad (2.26)$$

Ομοίως για τις πολώσεις b_j^n , προκύπτουν οι εξισώσεις:

$$\frac{\partial \mathcal{L}}{\partial b_j^n} = \frac{\partial \mathcal{L}}{\partial z_j^n} \frac{\partial z_j^n}{\partial b_j^n} \quad (2.27)$$

$$\frac{\partial z_j^n}{\partial b_j^n} = 1 \quad (2.28)$$

$$\frac{\partial \mathcal{L}}{\partial b_j^n} = \frac{\partial \mathcal{L}}{\partial z_j^n} 1 \quad (2.29)$$

Μέσω του υπολογισμού των παραπάνω παραγώγων και κλίσεων περατώνεται η βελτιστοποίηση των παραμέτρων του δικτύου. Αφού ολοκληρωθούν τα βήματα αυτά, γίνεται ενημέρωση των βαρών και ο αλγόριθμος οπισθοδιάδοσης τερματίζεται όταν εκπληρωθούν οι παρακάτω προϋποθέσεις:

$$w := w - \epsilon \frac{\partial \mathcal{L}}{\partial w} \quad (2.30)$$

$$b := b - \epsilon \frac{\partial \mathcal{L}}{\partial b} \quad (2.31)$$

όπου το ϵ συμβολίζει τον ρυθμό εκμάθησης και καθορίζει το βαθμό επιρροής της κλίσης [26].

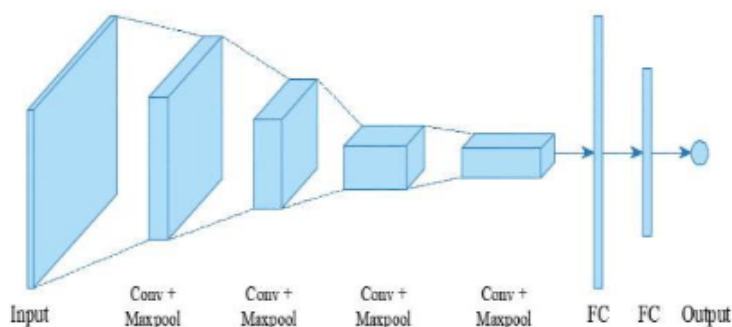
2.2 Συνελκτικκά Νευρωνικά Δίκτυα (Convolutional Neural Networks)

Τα Συνελκτικκά Νευρωνικά Δίκτυα (CNN) είναι μία υποκατηγορία των τεχνητών νευρωνικών δικτύων και αποτελούν πολύτιμο εργαλείο για την επεξεργασία και εκπαίδευση οπτικών δεδομένων όπως είναι η εικόνα και το βίντεο [6].

Τα CNN είναι μια επιβλεπόμενη μέθοδος μάθησης και η αρχιτεκτονική τους βασίζεται σε ένα στρώμα εισόδου, ένα στρώμα εξόδου και πολλαπλά κρυφά στρώματα. Τα στρώματα αυτά χωρίζονται σε τρεις κύριες κατηγορίες: το στρώμα συνέλιξης (convolution layer), το στρώμα συσσώρευσης (pooling layer) και το πλήρως συνδεδεμένο στρώμα (fully-connected layer). Συγχρόνως, θα πρέπει να υπογραμμιστεί η σημασία της εφαρμογής της συνάρτησης ενεργοποίησης στο μοντέλο μας, στοιχείο το οποίο μπορεί να θεωρηθεί ως ένα επιπλέον στρώμα.

Το γεγονός ότι τα CNN πραγματεύονται δεδομένα που περιέχουν εικόνες, τοποθετεί στο επίκεντρο της αρχιτεκτονικής τους την ανάγκη διαχείρισης τέτοιου είδους δεδομένων. Μία από τις πιο σημαντικές παραλλαγές είναι ότι τα στρώματα που απαρτίζουν τα δίκτυα αυτά περιέχουν νευρώνες που οργανώνονται σε τρεις διαστάσεις, γνωστή και ως χωρική διαστατικότητα της εισόδου. Η αρχιτεκτονική του CNN είναι μοντελοποιημένη έτσι ώστε να αξιοποιεί τα πλεονεκτήματα της δισδιάστατης δομής της εικόνας εισόδου (ή και άλλων δισδιάστατων

τύπων δεδομένων) το οποίο πραγματοποιείται μέσω γειτονικών συνδέσεων και συνδεδεμένων βαρών, ακολουθούμενα από κάποια χωρική υποδειγματοληψία. Η είσοδος και η έξοδος κάθε σταδίου του δικτύου συντελούν ένα σύνολο πινάκων το οποίο ονομάζεται χάρτης χαρακτηριστικών.



Εικόνα 2.7: Η βασική αρχιτεκτονική ενός συνελκτικού νευρωνικού δικτύου [6]

Παρακάτω γίνεται ανάλυση των επιμέρους στρωμάτων ενός CNN.

1. Στρώμα Συνέλιξης (Convolution Layer)

Το στρώμα συνέλιξης αποτελεί θεμελιώδες στοιχείο της αρχιτεκτονικής των CNN, αφού αφορά στο μεγαλύτερο μέρος των υπολογισμών. Οι παράμετροί του εστιάζουν κυρίως στη χρήση πυρήνων (ή φίλτρων) οι οποίοι μπορούν να εκπαιδευτούν. Ενώ αυτοί οι πυρήνες είναι μικροσκοπικοί ως προς την χωρική διαστατικότητα, αναπτύσσονται σε όλη τη διάσταση του βάθους της εισόδου. Μόλις η πληροφορία φτάσει σε ένα στρώμα συνέλιξης, το στρώμα συνελίσσει κάθε φίλτρο στη χωρική διαστατικότητα των δεδομένων, προκειμένου να δημιουργηθεί ένας διδιάστατος χάρτης ενεργοποίησης. Οι έξοδοι των νευρώνων οι οποίες είναι συνδεδεμένες σε τοπικές περιοχές της εισόδου, μπορούν να επαληθευτούν στο στρώμα συνέλιξης μέσω του υπολογισμού του βαθμωτού γινομένου μεταξύ των βαρών και της περιοχής που είναι συνδεδεμένη με την είσοδο. Οι νευρώνες που συντελούνται από ταυτόσημους χάρτες χαρακτηριστικών, διαμοιράζουν τα βάρη (διαμοιρασμός παραμέτρων), μειώνοντας έτσι την πολυπλοκότητα του δικτύου αφού διατηρείται χαμηλό το πλήθος των παραμέτρων.

Τα συνελκτικά στρώματα μπορούν να βελτιστοποιηθούν περαιτέρω μέσω της χρήσης τριών υπερ-παραμέτρων: του βάθους, της τιμής του βήματος (stride) και της συμπλήρωσης με μηδέν (zero-padding). Η συμπλήρωση με μηδέν αποτελεί μία αποτελεσματική μέθοδο συμπλήρωσης των οριακών τιμών της εισόδου και παρέχει επιπρόσθετη διαχείριση της διαστατικότητας της εξόδου. Για τον υπολογισμό της χωρικής διαστατικότητας των συνελκτικών στρωμάτων, χρησιμοποιείται η παρακάτω μαθηματική σχέση:

$$\frac{(V - R) + 2Z}{S + 1} \quad (2.32)$$

όπου :

- **V**: μέγεθος της εισόδου
- **R**: μέγεθος της περιοχής υποδοχής
- **Z**: μέγεθος του συνόλου συμπλήρωσης με μηδέν
- **S**: τιμή του βήματος

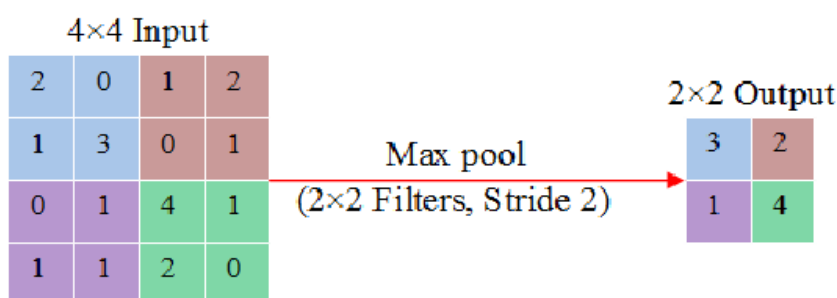
Τα CNN εκπαιδεύονται με τον αλγόριθμο οπισθοδιάδοσης, συνεπώς το βήμα της οπισθοδιάδοσης συμβάλλει περαιτέρω στη διαδικασία της συνέλιξης με χωρικά μετακινούμενα φίλτρα.

2. Στρώμα Ενεργοποίησης (Activation Layer)

Κατά το στρώμα ενεργοποίησης συντελείται η εφαρμογή των συναρτήσεων ενεργοποίησης (αναλύθηκαν στην υποενότητα 2.1.3), οι οποίες πρέπει να είναι μη γραμμικές. Με τη συμβολή του βήματος αυτού, πληθαίνουν οι μη γραμμικές ιδιότητες του δικτύου, το οποίο είναι ένα επιθυμητό χαρακτηριστικό των πολύ-στρωματικών δικτύων, ενώ συγχρόνως δεν επηρεάζει τα πεδία υποδοχής του συνελκτικού στρώματος.

3. Στρώμα Συσσώρευσης (Pooling Layer)

Τα CNN δεν περιέχουν μόνο συνελκτικά στρώματα, αλλά και στρώματα συσσώρευσης, με τα τελευταία να συναντώνται συνήθως μετά τα πρώτα. Οι μέθοδοι συσσώρευσης μειώνουν τις διαστάσεις των χαρτών χαρακτηριστικών μέσω της χρήσης συναρτήσεων που συνοψίζουν τις ιδιότητες κάποιων περιοχών, όπως η συνάρτηση της μέγιστης τιμής, της μέσης τιμής κ.α.. Τα στρώματα συσσώρευσης στοχεύουν στη σταδιακή ελάττωση της διαστατικότητας των δεδομένων, περιορίζοντας περισσότερο τον αριθμό των παραμέτρων. Το γεγονός αυτό ενισχύει την απόδοση του δικτύου, αφού ρυθμίζεται τόσο το ζήτημα της πολυπλοκότητας όσο και εκείνο της υπερ-εκπαίδευσης.



Εικόνα 2.8: Διαδικασία Max-Pooling [6]

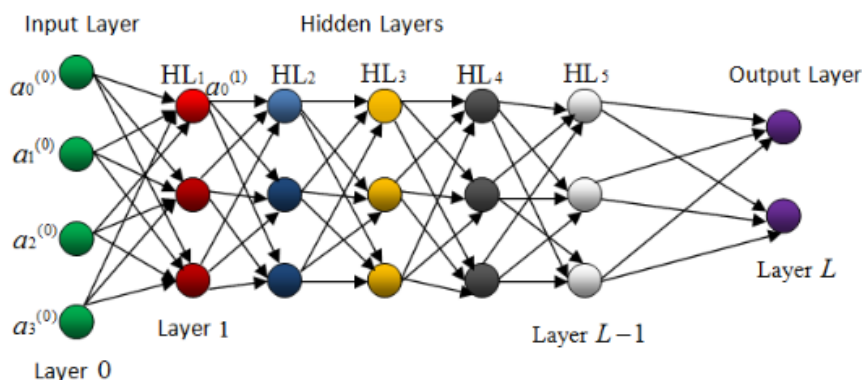
Στο σχήμα 2.8 απεικονίζεται ένα στρώμα συσσώρευσης που χρησιμοποιεί την συνάρτηση της μέγιστης τιμής (max pooling layer). Στην πλειονότητα των CNN, αυτά είναι τα διαθέσιμα μεγέθη max pooling layers με πυρήνες διάστασης 2×2 , τα οποία εφαρμόζονται με τιμή βήματος (stride) που ισούται με 2 στις χωρικές διαστάσεις της εισόδου και μειώνει τον χάρτη ενεργοποίησης στο 25 % του αρχικού του μεγέθους, διαφυλάσσοντας παράλληλα το βάθος της στο κανονικό του μέγεθος.

Υπάρχουν δύο στρατηγικές που χρησιμοποιούνται στην εφαρμογή του max pooling. Τυπικά, η τιμή του βήματος (stride) και τα φίλτρα των στρώματων συσσώρευσης ορίζονται στο 2×2 , το οποίο επιτρέπει στο στρώμα να διαδοθεί σε όλη την έκταση της χωρικής διαστατικότητας της εισόδου. Επιπροσθέτως, χρησιμοποιείται η επικαλυπτόμενη συσσώρευση (overlapping pooling), όπου το stride ισούται με 2 και ο πυρήνας είναι μεγέθους 3×3 . Ωστόσο, εξαιτίας της δομής του στρώματος συσσώρευσης, η ύπαρξη ενός πυρήνα με διάσταση μεγαλύτερη από 3 μπορεί να περιορίσει σημαντικά την απόδοση του μοντέλου.

4. Πλήρως Συνδεδεμένο Στρώμα (Fully Connected Layer)

Η υψηλού επιπέδου συλλογιστική των CNN ολοκληρώνεται με τα πλήρως συνδεδεμένα στρώματα. Σε ένα πλήρως συνδεδεμένο στρώμα, οι νευρώνες σε κάποιο στάδιο συνδέονται με όλες τις ενεργοποιήσεις του προηγούμενου στρώματος, οι οποίες υπολογίζονται με πράξεις πινάκων και αντιστάθμιση της πόλωσης.

Τα στρώματα συνέλιξης και συσσώρευσης συνεπιτελούν στην εξόρυξη υψηλού επιπέδου χαρακτηριστικών των εικόνων εισόδου, γεγονός που ενισχύουν περαιτέρω τα πλήρως συνδεδεμένα δίκτυα αφού αποτελούν μία οικονομική προσέγγιση ώστε να εξαχθούν οι μη γραμμικοί συνδυασμοί των χαρακτηριστικών αυτών. Οι νευρώνες των πλήρως συνδεδεμένων στρωμάτων δεν είναι χωρικά οργανωμένοι, καθιστώντας αδύνατο το να ακολουθείται από ένα συνελκτικό στρώμα. Το τελευταίο, προωθεί την διδιάστατη έξοδο στο στρώμα εξόδου όπου και μπορεί να εφαρμοστεί κάποια μη γραμμική συνάρτηση ενεργοποίησης. Ο στόχος των πλήρως συνδεδεμένων στρωμάτων είναι να εξομαλύνει τα χαρακτηριστικά υψηλού επιπέδου και να συνδυάσει όλα τα στοιχεία.



Εικόνα 2.9: Αναπαράσταση ενός πλήρως συνδεδεμένου στρώματος [6]

2.3 Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks)

2.3.1 Δομή και Μοντελοποίηση

Τα Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (RNN) είναι μία υποκατηγορία των τεχνητών νευρωνικών δικτύων και η χρήση τους έγκειται κυρίως στον εντοπισμό μοτίβων σε μία ακο-

λουθία δεδομένων [7]. Η κυρίαρχη διαφοροποίηση τους από τα κλασικά πολυστρωματικά δίκτυα είναι ότι η πληροφορία μεταδίδεται όχι μόνο σε στρώματα επόμενων επιπέδων αλλά και σε ίδια ή προηγούμενα επίπεδα. Η ιδιότητά τους αυτή του επιτρέπει να έχουν εφαρμογή σε περιπτώσεις όπως η μοντελοποίηση γλώσσας και παραγωγή κειμένου, η αναγνώριση ομιλίας, η παραγωγή περιγραφών στις εικόνες κ.α..

Στη συνέχεια περιγράφεται το μαθηματικό υπόβαθρο που συντελεί την ιδέα των RNN. Ορίζουμε με H ένα σύνολο από κρυφά επίπεδα (μπορεί να περιλαμβάνει ένα ή περισσότερα κρυφά επίπεδα). Συμβολίζουμε με $H_t \in \mathbb{R}^{n \times h}$ και $X_t \in \mathbb{R}^{n \times h}$ την κατάσταση των κρυμμένων δικτύων και της εισόδου αντίστοιχα, σε μία δεδομένη χρονική στιγμή t , όπου:

- n : ο αριθμός των δειγμάτων
- d : ο αριθμός των εισόδων κάθε δείγματος
- h : ο αριθμός των κρυφών στρωμάτων

Έπειτα ορίζουμε έναν πίνακα βαρών $W_{xh} \in \mathbb{R}^{d \times h}$, έναν πίνακα καταστάσεων των κρυφών επιπέδων $W_{hh} \in \mathbb{R}^{h \times h}$ και μία παράμετρο πόλωσης $b_h \in \mathbb{R}^{1 \times h}$. Στο σύνολο αυτών των πληροφοριών εφαρμόζεται μία συνάρτηση ενεργοποίησης η οποία συνήθως είναι είτε η σιγμοειδής είτε εκείνη της υπερβολικής εφαιπτομένης, ώστε να προετοιμαστούν οι κλίσεις για τον αλγόριθμο οπισθοδιάδοσης. Έτσι προκύπτουν οι εξισώσεις:

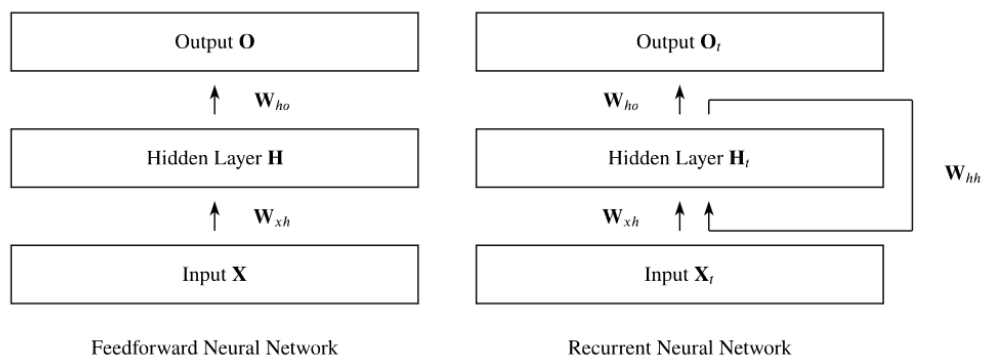
$$H_t = \phi_h (X_t W_{xh} + H_{t-1} W_{hh} + b_h) \quad (2.33)$$

που αποτελεί την μεταβλητή των κρυφών στρωμάτων.

$$O_t = \phi_o (H_t W_{ho} + b_o) \quad (2.34)$$

που αποτελεί την μεταβλητή εξόδου.

Όπως φαίνεται από την αναδρομική σχέση 2.33, το RNN περιλαμβάνει στοιχεία όλων των κρυφών σταδίων που προηγούνται του H_{t-1} , καθώς και το ίδιο το H_{t-1} .



Εικόνα 2.10: Οπτικοποίηση των διαφορών μεταξύ νευρωνικών δικτύων πρόσθιας διάδοσης (αριστερά) και ανατροφοδοτούμενων νευρωνικών δικτύων (δεξιά) [7]

2.3.2 Οπισθοδιάδοση στο Χρόνο (Back Propagation Through Time)

Ένα ερώτημα που εγείρεται με τη χρήση των RNN είναι το με ποιον τρόπο μπορεί να εφαρμοστεί σε αυτά ο αλγόριθμος οπισθοδιάδοσης, αντικείμενο που θα εξεταστεί σε αυτήν την υποενότητα. Η οπισθοδιάδοση στο χρόνο (BPTT) είναι μία προσαρμογή του αλγόριθμου οπισθοδιάδοσης για τα RNN.

Κατά την πρόσθια διάδοση της εισόδου X_t στο δίκτυο, υπολογίζουμε την κατάσταση των κρυφών στρωμάτων H_t και την τελική έξοδο O_t με ένα βήμα τη φορά. Ορίζουμε μία συνάρτηση κόστους (loss function) $\mathcal{L}(O, Y)$ προκειμένου να περιγράψουμε τη διαφορά μεταξύ των εξόδων O_t και των αναμενόμενων τιμών εξόδου Y_t . Η συνάρτηση κόστους δίνεται από τη σχέση:

$$\mathcal{L}(O, Y) = \sum_{t=1}^T \ell_t(O_t, Y_t) \quad (2.35)$$

δηλαδή αθροίζει κάθε όρο κόστους ℓ_t της κάθε ενημέρωσης μέχρι τη δεδομένη στιγμή t .

Στη συνέχεια θα πρέπει να υπολογιστεί η μερική παράγωγος κάθε πίνακα βαρών με τη χρήση του κανόνα της αλυσίδας, όπως και στον κανονικό αλγόριθμο οπισθοδιάδοσης:

$$\frac{\partial \mathcal{L}}{\partial W_{ho}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot \frac{\partial \phi_o}{W_{ho}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot H_t \quad (2.36)$$

$$\frac{\partial \mathcal{L}}{\partial W_{hh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot \frac{\partial \phi_o}{\partial H_t} \cdot \frac{\partial H_t}{\partial \phi_h} \cdot \frac{\partial \phi_h}{W_{hh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot W_{ho} \cdot \frac{\partial H_t}{\partial \phi_h} \cdot \frac{\partial \phi_h}{\partial W_{hh}} \quad (2.37)$$

$$\frac{\partial \mathcal{L}}{\partial W_{xh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot \frac{\partial \phi_o}{\partial H_t} \cdot \frac{\partial H_t}{\partial \phi_h} \cdot \frac{\partial \phi_h}{W_{xh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot W_{ho} \cdot \frac{\partial H_t}{\partial \phi_h} \cdot \frac{\partial \phi_h}{\partial W_{xh}} \quad (2.38)$$

Αφού κάθε H_t εξαρτάται από την προηγούμενη χρονική στιγμή, μπορούμε να αντικαταστήσουμε το τελευταίο μέρος των παραπάνω εξισώσεων και έτσι προκύπτουν οι σχέσεις:

$$\frac{\partial \mathcal{L}}{\partial W_{hh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot W_{ho} \sum_{k=1}^t \frac{\partial H_t}{\partial H_k} \cdot \frac{\partial H_k}{\partial W_{hh}} \quad (2.39)$$

$$\frac{\partial \mathcal{L}}{\partial W_{xh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot W_{ho} \sum_{k=1}^t \frac{\partial H_t}{\partial H_k} \cdot \frac{\partial H_k}{\partial W_{xh}} \quad (2.40)$$

Οι παραπάνω εξισώσεις μπορούν να μετασχηματιστούν περαιτέρω στις εξής:

$$\frac{\partial \mathcal{L}}{\partial W_{hh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot W_{ho} \sum_{k=1}^t (W_{hh}^\top)^{t-k} \cdot H_k \quad (2.41)$$

$$\frac{\partial \mathcal{L}}{\partial W_{xh}} = \sum_{t=1}^T \frac{\partial \ell_t}{\partial O_t} \cdot \frac{\partial O_t}{\partial \phi_o} \cdot W_{ho} \sum_{k=1}^t (W_{hh}^\top)^{t-k} \cdot X_k \quad (2.42)$$

Από τις σχέσεις που παρουσιάζονται ανωτέρω γίνεται αντιληπτό πως χρειάζεται να απο-

θηκεύουμε δυνάμεις του W_{hh}^k όσο προχωράμε σε κάθε όρο κόστους ℓ_t της συνολικής συνάρτησης κόστους \mathcal{L} που μπορεί να μεγαλώσει αρκετά. Για αυτές τις μεγάλες τιμές, η μέθοδος γίνεται αριθμητικά ασταθής, καθώς τα ιδιοδιανύσματα των οποίων η τιμή είναι μικρότερη από 1 εξαφανίζονται, ενώ εκείνα που είναι μεγαλύτερα από 1 αποκλίνουν.

Ένας τρόπος προκειμένου να επιλυθεί αυτή η δυσχέρεια είναι να περικόψουμε το άθροισμα σε ένα υπολογιστικά κατάλληλο μέγεθος. Η μέθοδος αυτή ονομάζεται Truncated BPTT και καθορίζει ένα ανώτατο όριο για τον αριθμό των χρονικών βημάτων που μπορεί να ακολουθήσει η κλίση κατά την οπισθοδιάδοση.

2.3.3 Είδη Ανατροφοδοτούμενων Νευρωνικών Δικτύων

Όπως στην πλειονότητα των τεχνητών νευρωνικών δικτύων, μείζον πρόβλημα για τα RNN αποτελούν η εκραγείσα κλίση (exploding gradient) καθώς και η εξαφάνιση κλίσης (vanishing gradient). Από τις εξισώσεις 2.39, 2.40 φαίνεται ότι ο όρος $\frac{\partial H_t}{\partial H_k}$ εισάγει στην (δυσνητικά πολύ μεγάλη) ακολουθία τον πολλαπλασιασμό πινάκων, γεγονός που συνεπάγεται πως αν υπάρχουν μικρές τιμές (< 1) στον πολλαπλασιασμό των πινάκων θα προκληθεί σταδιακή μείωση της κλίσης και τελικώς η εξαφάνισή της. Το φαινόμενο αυτό μπορεί να παρατηρηθεί και αντιστρόφως, δηλαδή αν υπάρχουν αρκετά μεγάλες τιμές (> 1) το μοντέλο θα οδηγηθεί σε έκρηξη των κλίσεων.

Τα ζητήματα αυτά, δημιούργησαν την ανάγκη ανάπτυξης κάποιων επεκτάσεων των RNN, οι οποίες θα αναλυθούν στην υποενότητα αυτή.

1. Δίκτυα Μακράς Βραχέας Μνήμης (Long Short-Term Memory Units)

Τα δίκτυα μακράς βραχέας μνήμης (LSTM) σχεδιάστηκαν προκειμένου να αντιμετωπιστεί το πρόβλημα της εξαφάνισης κλίσεων. Καθώς χρησιμοποιούν ένα πιο σταθερό σφάλμα, επιτρέπουν στα RNN να μαθαίνουν σε περισσότερα βήματα χρόνου ($\gg 1000$). Για να επιτευχθεί αυτό, τα LSTM αποθηκεύουν περισσότερη πληροφορία έξω από την κλασική ροή του νευρωνικού δικτύου, σε κάποιες δομές που ονομάζονται κελιά πύλης.

Η αρχιτεκτονική απαρτίζεται από μία πύλη εξόδου O_t για να διαβάσει τις καταχωρήσεις του κυττάρου, μία πύλη εισόδου I_t ώστε να διαβάσει τα δεδομένα μέσα στο κύτταρο και μία πύλη λήθης F_t για να ρυθμίζει το περιεχόμενο του κελιού. Οι υπολογισμοί για αυτές τις πύλες παρουσιάζονται στις παρακάτω εξισώσεις:

$$O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \quad (2.43)$$

$$I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \quad (2.44)$$

$$F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + f_o) \quad (2.45)$$

Οι ανωτέρω εξισώσεις χρησιμοποιούν ως πίνακες βαρών τους $W_{xi}, W_{xf}, W_{xo} \in \mathbb{R}^{d \times h}$ και $W_{hi}, W_{hf}, W_{ho} \in \mathbb{R}^{h \times h}$ ενώ τα $b_i, b_f, b_o \in \mathbb{R}^{1 \times h}$ συντελούν τις αντίστοιχες πολώσεις

τους. Ταυτοχρόνως, εφαρμόζεται η σιγμοειδής συνάρτηση ενεργοποίησης σ για να μετασχηματιστεί η έξοδος $\in (0, 1)$, δίνοντας ως αποτέλεσμα διανύσματα με καταχωρήσεις $\in (0, 1)$.

Στη συνέχεια κρίνεται απαραίτητη η χρήση ενός κελιού μνήμης $\tilde{C}_t \in \mathbb{R}^{n \times h}$ το οποίο προκύπτει με παρόμοιους υπολογισμούς όπως τα προηγούμενα, αλλά χρησιμοποιεί ως συνάρτηση ενεργοποίησης εκείνη της υπερβολικής εφαιπτομένης ώστε να έχει έξοδο στο διάστημα $(-1, 1)$. Το κελί αυτό διαθέτει τα βάρη $W_{xc} \in \mathbb{R}^{d \times h}$, $W_{hc} \in \mathbb{R}^{h \times h}$ και την πόλωση $b_c \in \mathbb{R}^{1 \times h}$. Η μαθηματική σχέση που συνοψίζει τη λειτουργία του κελιού μνήμης είναι ή εξής:

$$\tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \quad (2.46)$$

Με την εισαγωγή του προγενέστερου περιεχομένου μνήμης $C_{t-1} \in \mathbb{R}^{n \times h}$ σε συνάρτηση με την εισαγωγή των πυλών, ελέγχεται η ποσότητα της προγενέστερης μνήμης που θέλουμε να διατηρήσουμε ώστε να δημιουργηθεί το νέο περιεχόμενο μνήμης C_t . Δηλαδή:

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (2.47)$$

Με την εισαγωγή των κρυφών επιπέδων στη δομή που έχει δημιουργηθεί, περατώνεται η διαδικασία. Έτσι έχουμε:

$$H_t = O_t \odot \tanh(C_t) \quad (2.48)$$

όπου με τη χρήση της συνάρτησης ενεργοποίησης της υπερβολικής εφαιπτομένης, διασφαλίζεται πως κάθε στοιχείο $H_t \in (-1, 1)$

2. Βαθιά Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Deep Recurrent Neural Networks)

Τα βαθιά ανατροφοδοτούμενα νευρωνικά δίκτυα (DRNN) αποτελούν μία απλή παραλλαγή των (RNN) καθώς η αρχιτεκτονική τους απαρτίζεται από πολλαπλές στρώσεις οποιουδήποτε τύπου (RNN).

Έστω L ο αριθμός των κρυφών στρωμάτων. Κάθε κρυφό στάδιο $H_t^{(\ell)} \in \mathbb{R}^{n \times h}$ μεταβιβάζεται τόσο στο επόμενο χρονικό βήμα, του τρέχοντος στρώματος $H_{t+1}^{(\ell)}$, όσο και στο τρέχον χρονικό βήμα του επόμενου στρώματος $H_t^{(\ell)+1}$. Για το πρώτο στρώμα, υπολογίζουμε την κατάσταση του κρυφού στρώματος όπως προτάθηκε και στα προηγούμενα μοντέλα, ενώ για τα μεταγενέστερα στρώματα χρησιμοποιούμε σαν είσοδο την έξοδο των προηγούμενων στρωμάτων. Προκύπτουν οι εξισώσεις:

$$H_t^{(1)} = \phi_1(X_t, H_{t-1}^{(1)}) \quad (2.49)$$

$$H_t^{(\ell)} = \phi_\ell(H_t^{(\ell-1)}, H_{t-1}^{(\ell)}) \quad (2.50)$$

Η έξοδος $O_t \in \mathbb{R}^{n \times o}$, όπου o ο αριθμός των εξόδων απαιτεί για τον υπολογισμό του μόνο τη χρήση του κρυφού στρώματος L .

$$O_t = \phi_o \left(H_t^{(L)} W_{ho} + b_o \right) \quad (2.51)$$

3. Αμφίδρομα Ανατροφοδοτούμενα Νευρωνικά Δίκτυα (Bidirectional Recurrent Neural Networks)

Για να γίνει αντιληπτή η ανάγκη δημιουργίας των αμφίδρομων ανατροφοδοτούμενων νευρωνικών δικτύων (BRNN), θα παρουσιαστεί ένα απλό παράδειγμα μοντελοποίησης γλώσσας. Βάσει των μοντέλων που έχουν αναφερθεί ως τώρα, υπάρχει η δυνατότητα πρόβλεψης ενός στοιχείου (π.χ. μίας λέξης) λαμβάνοντας υπόψη τα προηγούμενα δεδομένα. Στην περίπτωση όμως που απαιτείται η συμπλήρωση του κενού σε μία πρόταση πρέπει να συνυπολογιστεί τόσο το μέρος της πρότασης πριν το κενό, όσο και εκείνο που ακολουθεί μετά. Η ανάγκη αυτή δημιουργίας της εξόδου όχι μόνο συναρτήσκει τον προγενέστερων αλλά και των μεταγενέστερων δεδομένων, πυροδότησε την ανάπτυξη των (BRNN).

Η αρχιτεκτονική των (BRNN) βασίζεται στην προσθήκη ενός επιπλέον κρυφού στρώματος το οποίο εκτελεί ανάποδα την ακολουθία, ξεκινώντας από το τελευταίο στοιχείο. Αρχικά εισάγουμε ένα πρόσθιο και ένα οπίσθιο κρυφό στρώμα που ορίζονται ως $\vec{H}_t \in \mathbb{R}^{n \times h}$ και $\overleftarrow{H}_t \in \mathbb{R}^{n \times h}$ αντίστοιχα. Οι μαθηματικές σχέσεις παρουσιάζονται παρακάτω:

$$\vec{H}_t = \phi \left(X_t W_{xh}^{(f)} + \vec{H}_{t-1} W_{hh}^{(f)} + b_h^{(f)} \right) \quad (2.52)$$

$$\overleftarrow{H}_t = \phi \left(X_t W_{xh}^{(b)} + \overleftarrow{H}_{t+1} W_{hh}^{(b)} + b_h^{(b)} \right) \quad (2.53)$$

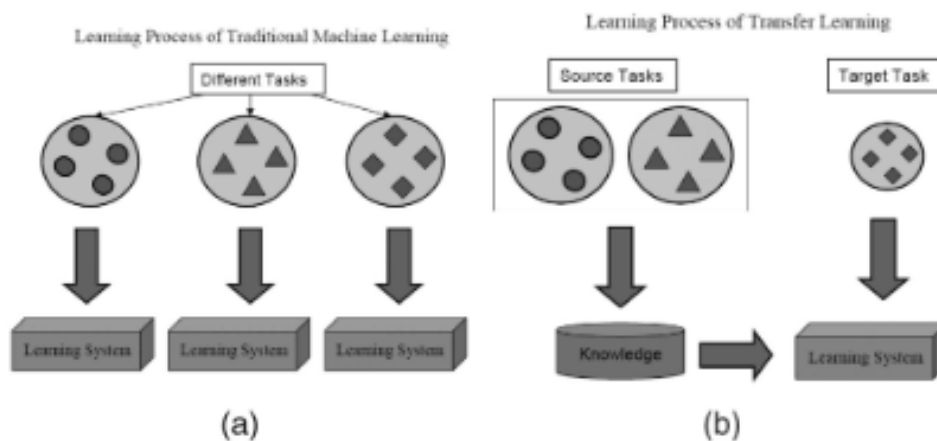
Οι πίνακες βαρών που προκύπτουν είναι παρόμοιοι με εκείνους που ορίστηκαν παραπάνω, αλλά σε αυτήν την περίπτωση χωρίζονται σε δύο σύνολα. Το ένα σύνολο απαρτίζεται από τις πρόσθιες κρυφές καταστάσεις $W_{xh}^{(f)} \in \mathbb{R}^{d \times h}$ και $W_{hh}^{(f)} \in \mathbb{R}^{h \times h}$, ενώ το άλλο από τις οπίσθιες κρυφές καταστάσεις $W_{xh}^{(b)} \in \mathbb{R}^{d \times h}$ και $W_{hh}^{(b)} \in \mathbb{R}^{h \times h}$. Οι αντίστοιχες πολώσεις τους είναι $b_h^{(f)} \in \mathbb{R}^{1 \times h}$ και $b_h^{(b)} \in \mathbb{R}^{1 \times h}$. Με αυτά τα δεδομένα υπολογίζουμε την έξοδο ως εξής:

$$O_t = \phi \left(\left[\vec{H}_t \overleftarrow{H}_t \right] W_{ho} + b_o \right) \quad (2.54)$$

όπου o είναι ο αριθμός των εξόδων και το \frown συμβολίζει την συγκέντρωση των δύο πινάκων στον άξονα 0. Ταυτόχρονα έχουμε τους πίνακες βαρών $W_{ho} \in \mathbb{R}^{2h \times o}$ και τις παραμέτρους πόλωσης $b_o \in \mathbb{R}^{1 \times o}$. Τέλος, οι δύο κατευθύνσεις μπορούν να έχουν διαφορετικό αριθμό κρυφών δομών.

2.4 Μεταφορά Μάθησης (Transfer Learning)

Όπως είναι αντιληπτό από τα παραπάνω, το πεδίο της μηχανικής μάθησης έχει σημειώσει θαυμαστή πρόοδο τα τελευταία χρόνια, αναπτύσσοντας εξελιγμένες μεθόδους και τεχνικές ώστε να βελτιστοποιήσει τα αποτελέσματα που παρέχει [8]. Ωστόσο, πολλές μέθοδοι μηχανικής εκμάθησης λειτουργούν καλά μόνο υπό την προϋπόθεση ότι τα δεδομένα εκπαίδευσης και επαλήθευσης αντλούνται από τον ίδιο χώρο χαρακτηριστικών και έχουν την ίδια κατανομή. Η μεταβολή της κατανομής επιβάλλει στα δεδομένα να δημιουργηθούν ξανά από το μηδέν (from scratch), χρησιμοποιώντας πρόσφατα δεδομένα εκπαίδευσης. Σε πολλές εφαρμογές του πραγματικού κόσμου, η διαδικασία αυτή απαιτεί υψηλό υπολογιστικό κόστος και σε κάποιες περιπτώσεις καθίσταται αδύνατη η ανάκτηση των δεδομένων. Η χρήση της μεταφοράς μάθησης (transfer learning) αποτελεί τη λύση σε αυτό το τροχοπέδη.



Εικόνα 2.11: Διαφορετικές διαδικασίες εκμάθησης μεταξύ (α) της κλασικής μηχανικής μάθησης και (β) της μεταφοράς μάθησης [8]

Στην αρχή του κεφαλαίου αναφέρθηκε και αναλύθηκε πώς τα τεχνητά νευρωνικά δίκτυα μιμούνται και αντλούν την έμπνευσή τους από τα βιολογικά νευρωνικά δίκτυα του εγκεφάλου. Η περίπτωση της μεταφοράς μάθησης δεν αποτελεί εξαίρεση καθώς και πάλι πρότυπο είναι ο ανθρώπινος οργανισμός και πιο συγκεκριμένα, η ικανότητά του να προσαρμόσει και να εφαρμόσει με ευφυή τρόπο την πρότερη γνώση που έχει αποκτήσει για την ταχύτερη και καλύτερη διεκπεραίωση των νέων προβλημάτων που καλείται να αντιμετωπίσει. Για παράδειγμα, η γνώση ενός μουσικού οργάνου διευκολύνει την εκμάθηση ενός νέου μουσικού οργάνου λόγω του μουσικού υπόβαθρου που έχει καλλιεργηθεί προγενέστερα. Η λογική αυτή ακολουθείται και στην μεταφορά μάθησης.

Παρακάτω θα αναλυθεί το μαθηματικό πλαίσιο που θεμελιώνει την ιδέα της μεταφοράς μάθησης.

Αρχικά, ορίζουμε ως "τομέα" (domain) ένα μέγεθος που αποτελείται από δύο συστατικά: ένα χώρο χαρακτηριστικών \mathcal{X} και μία οριακή κατανομή πιθανοτήτων $P(X)$, όπου το $X = \{x_1, \dots, x_n\} \in \mathcal{X}$. Αν δύο τομείς είναι διαφορετικοί τότε θα έχουν και διαφορετικούς χώρους χαρακτηριστικών ή διαφορετικές οριακές κατανομές πιθανοτήτων.

Με δεδομένο έναν τομέα $\mathcal{D} = \{\mathcal{X}, P(X)\}$, μία "εργασία" (task) αποτελείται επίσης από δύο

συστατικά: ένα χώρο ετικετών \mathcal{Y} και μία συνάρτηση πρόβλεψης $f(\cdot)$, και συμβολίζεται με $\mathcal{T} = \{\mathcal{Y}, f(\cdot)\}$. Η συνάρτηση πρόβλεψης μπορεί να εξαχθεί από τα δεδομένα εκπαίδευσης τα οποία απαρτίζονται από ζεύγη $\{x_i, y_i\}$ όπου $x_i \in X$ και $y_i \in \mathcal{Y}$ και χρησιμοποιείται για την πρόβλεψη των αντίστοιχων ετικετών $f(x)$ ενός δείγματος x .

Για λόγους απλότητας, υποθέτουμε ότι έχουμε έναν μόνο τομέα - πηγή \mathcal{D}_S και έναν τομέα - στόχο \mathcal{D}_T . Ειδικότερα:

$$\mathcal{D}_S = \{(x_{S_1}, y_{S_1}), \dots, (x_{S_{n_S}}, y_{S_{n_S}})\} \quad (2.55)$$

όπου $x_{S_i} \in \mathcal{X}_S$ είναι το σύνολο δεδομένων και $y_{S_i} \in \mathcal{Y}_S$ οι αντίστοιχες ετικέτες κλάσης. Ομοίως για τον τομέα - στόχο:

$$\mathcal{D}_T = \{(x_{T_1}, y_{T_1}), \dots, (x_{T_{n_T}}, y_{T_{n_T}})\} \quad (2.56)$$

όπου $x_{T_i} \in \mathcal{X}_T$ είναι η είσοδος και $y_{T_i} \in \mathcal{Y}_T$ η αντίστοιχη έξοδος.

Έχοντας επεξηγήσει τα παραπάνω μεγέθη, μπορεί να δοθεί ο ορισμός για την μεταφορά μάθησης. Με δεδομένα έναν τομέα - πηγή \mathcal{D}_S , έναν τομέα - στόχο \mathcal{D}_T και τις αντίστοιχες εκπαιδευόμενες εργασίες τους \mathcal{T}_S και \mathcal{T}_T , η μεταφορά μάθησης εστιάζει στην βελτίωση της εκμάθησης της συνάρτησης πρόβλεψης του στόχου $f(\cdot)_T$ στο \mathcal{D}_T , εκμεταλλευόμενη την γνώση στα \mathcal{D}_S και \mathcal{T}_S , με $\mathcal{D}_S \neq \mathcal{D}_T$ και $\mathcal{T}_S \neq \mathcal{T}_T$.

Η μέθοδος της μεταφοράς μάθησης χωρίζεται σε ποικίλες υποκατηγορίες βάσει του τρόπου υλοποίησής της και των παραδοχών που λαμβάνονται ανάλογα με το εκάστοτε πρόβλημα. Σύμφωνα με αξιολογήσεις και πειράματα που έχουν διεξαχθεί ώστε να γίνει σύγκριση της απόδοσης των μεθόδων μεταφοράς μάθησης, έναντι των κλασικών μεθόδων μηχανικής μάθησης έχει αποδειχθεί ότι σε πολλαπλές περιπτώσεις η μεταφορά μάθησης ήταν εκείνη που παρείχε τα καλύτερα αποτελέσματα [27],[28],[29]. Οι εφαρμογές της επαφίενται σε πληθώρα προβλημάτων του πραγματικού κόσμου όπως είναι εκείνα της κατηγοριοποίησης (εγγράφων, εικόνων, συναισθημάτων) αλλά και η επεξεργασία φυσικής γλώσσας.

Ωστόσο, όπως είναι φυσικό, συναντώνται και κάποιοι περιορισμοί οι οποίοι αποτελούν πρόκληση για την ερευνητική κοινότητα του πεδίου των νευρωνικών δικτύων. Σε ορισμένες περιπτώσεις, συνήθως αν η αντιστοίχιση των δεδομένων δεν είναι ορθή, υφίσταται η έννοια της αρνητικής μεταφοράς μάθησης, παρέχοντας χειρότερα αποτελέσματα από εκείνα που θα λαμβάναμε χωρίς τη χρήση της. Παράλληλα, εμφανίζεται και το πρόβλημα της ετερογενούς μεταφοράς μάθησης, το οποίο σημειώνεται όταν οι εκάστοτε χώροι χαρακτηριστικών του τομέα-πηγή και τους τομέα-στόχου είναι διαφορετικοί.

2.5 Πλεονεκτήματα και Μειονεκτήματα των Τεχνητών Νευρωνικών Δικτύων

Στο κεφάλαιο αυτό εξετάστηκε το θεωρητικό υπόβαθρο στο οποίο στηρίχτηκε η δημιουργία των νευρωνικών δικτύων, καθώς επίσης και οι διάφορες μέθοδοι και παραλλαγές που αναπτύχθηκαν ώστε να προσαρμοστούν στην εκάστοτε κατηγορία προβλήματος. Το κείμενο ερώτημα που αξίζει να θέσουμε όμως είναι, ποια είναι τελικά τα οφέλη που μας παρέχουν τα

νευρωνικά δίκτυα και ποιες οι προκλήσεις που καλούνται να αντιμετωπίσουν οι επιστήμονες του κλάδου ώστε να βελτιστοποιήσουν περαιτέρω·

2.5.1 Πλεονεκτήματα

Όσον αφορά στα πλεονεκτήματα της χρήσης των τεχνητών νευρωνικών δικτύων, αυτά είναι πολυάριθμα. Αρχικά, το κύριο και υψίστης σημασίας γνώρισμά τους είναι ότι έχουν την ικανότητα να μαθαίνουν και να εκπαιδεύονται σε ένα σύνολο δεδομένων, εφαρμόζοντας την γνώση που απέκτησαν σε διαφορετικά δεδομένα με παρόμοια χαρακτηριστικά. Ταυτόχρονα, τόσο οι ανεπαρκείς πληροφορίες όσο και η καταστροφή ενός μέρους των δεδομένων δεν αποτελεί εμπόδιο στην εκτέλεσή τους και στην παραγωγή εξόδου, με την ορθότητα της εξόδου να εξαρτάται από την σημαντικότητα των πληροφοριών που λείπουν. Ένα άλλο πολύ σημαντικό πλεονέκτημα αποτελεί το γεγονός ότι διαθέτουν κατανεμημένη μνήμη. Προκειμένου το δίκτυο να εκπαιδευτεί ώστε να παραχθεί η επιθυμητή έξοδος, την προσδιορίζει πλήρως με παραδείγματα. Η επιτυχία του δικτύου είναι ευθέως ανάλογη με τις επιλεγμένες περιπτώσεις παραδειγμάτων οι οποίες διανέμονται σε όλο το δίκτυο ώστε να αποφεύγεται η λανθασμένη έξοδος. Τέλος τα τεχνητά νευρωνικά δίκτυα διαθέτουν την δυνατότητα της παράλληλης επεξεργασίας καθώς έχουν την υπολογιστική ισχύ να εκτελέσουν περισσότερες από μία διεργασίες συγχρόνως.

2.5.2 Μειονεκτήματα

Πέρα όμως από τα σημαντικά πλεονεκτήματα που αναφέρθηκαν υπάρχουν και κάποια μειονεκτήματα. Αξίζει όμως να σημειωθεί ότι με την πάροδο του χρόνου και την πρόοδο της επιστήμης σε αυτόν τον τομέα αίρονται ένα προς ένα τα μειονεκτήματα, καθιστώντας τη χρήση νευρωνικών δικτύων καίρια για την αντιμετώπιση σημαντικών προβλημάτων.

Αρχικά, η δυνατότητα παράλληλης επεξεργασίας των δεδομένων που επισημάνθηκε μπορεί να ισοδυναμεί με την βελτίωση της απόδοσης του δικτύου, αλλά συνεπάγεται ταυτοχρόνως την εξάρτησή του από το hardware του υπολογιστή το οποίο πρέπει να πληροί συγκεκριμένες προδιαγραφές. Επιπροσθέτως, δεν υφίσταται κάποιος συγκεκριμένος κανόνας για τον προσδιορισμό της δομής των τεχνητών νευρωνικών δικτύων. Η κατάλληλη δομή δικτύου επιτυγχάνεται μέσω εμπειρίας και δοκιμής. Κάτι άλλο που οφείλει να υπογραμμιστεί είναι η δυσκολία της παρουσίασης των δεδομένων του εκάστοτε προβλήματος στο δίκτυο. Επειδή τα νευρωνικά δίκτυα λειτουργούν με πληροφορίες σε αριθμητική μορφή, όλα τα προβλήματα θα πρέπει να μεταφραστούν σε αριθμητικές τιμές ώστε να εισαχθούν στο δίκτυο. Ο μηχανισμός που θα καθοριστεί εδώ επηρεάζει άμεσα την απόδοση του δικτύου και εξαρτάται από ικανότητα του χρήστη. Τέλος, το πιο σημαντικό μειονέκτημα αποτελεί η ανεξήγητη συμπεριφορά που παρατηρείται πολλές φορές στο δίκτυο κατά την παραγωγή ενός αποτελέσματος, γεγονός που μειώνει την αξιοπιστία του [30].

2.5.3 Εφαρμογές

Εξετάστηκαν τα πλεονεκτήματα των τεχνητών νευρωνικών δικτύων και τα προβλήματα που προέκυψαν κατά τη διάρκεια της χρήσης τους. Όπως προαναφέρθηκε τα μειονεκτήματα

τους εξαλείφονται βαθμιαία ενώ τα πλεονεκτήματά τους αυξάνονται συνεχώς. Τα τεχνητά νευρωνικά δίκτυα είναι ένας αναπτυσσόμενος κλάδος της επιστήμης που σταδιακά θα γίνει αναπόσπαστο κομμάτι της ζωής του σύγχρονου ανθρώπου.

Αρχιτεκτονικές βαθιών νευρωνικών δικτύων έχουν υλοποιηθεί και χρησιμοποιηθεί σε διάφορες εφαρμογές από μέλη του Εργαστηρίου Συστημάτων Τεχνητής Νοημοσύνης και Μάθησης του ΕΜΠ. Ειδικότερα τεχνικές CNN και CNN-RNN έχουν εφαρμοστεί για κατηγοριοποίηση αντικειμένων, στην ιατρική διάγνωση νευροεκφυλιστικών ασθενειών, όπως της νόσου του Πάρκινσον [31], [32], [33], [34], [35] ή της Covid-19 [36], [37], [38], βασισμένες σε δισδιάστατες ή τρισδιάστατες εικόνες. Έμφαση έχει δοθεί στην διαφάνεια και στην προσαρμογή των μοντέλων [39], [40], [41] αλλά και στην ανάπτυξη πλέον σύνθετων αρχιτεκτονικών, μπαϋεσιανών, με κάψουλες και αβεβαιότητα [42], [43], [44], [45]. Βαθιές τρισδιάστατες νευρωνικές αρχιτεκτονικές έχουν εφαρμοστεί στην ανίχνευση βλαβών σε πυρηνικούς αντιδραστήρες [46], [47], στην πρόβλεψη της παραγωγής στον αγροτικό τομέα [48], [49] και στην αναγνώριση και σύνθεση συναισθήματος [50], [51], [52], [53], ενώ άλλες εφαρμόζονται σε προβλήματα ανάλυσης εικόνων και αλληλεπίδρασης ανθρώπου-υπολογιστή [54], [55], [56].

Μέρος 

Πρακτικό Μέρος

Ανάλυση Μεθοδολογίας

Η μεθοδολογία που εφαρμόζεται στην παρούσα εργασία και επεξηγείται αναλυτικά σε αυτό το κεφάλαιο περιλαμβάνει τα εξής βήματα :

- Την απόκτηση και την κατάλληλη **προεπεξεργασία του συνόλου δεδομένων** ώστε να μπορεί να εισαχθεί και να αναπτυχθεί από το μοντέλο.
- Την εφαρμογή της μεθόδου της **επαύξησης των δεδομένων** ώστε να αντιμετωπιστεί το πρόβλημα της υπερεκπαίδευσης.
- Την εισαγωγή του συνόλου δεδομένων αρχικά σε ένα **συνελικτικό δίκτυο** και συγκεκριμένα στα δίκτυα τύπου **ResNet50, DenseNet121, EfficientNet-B0** για την εξαγωγή των χωρικών του χαρακτηριστικών και στη συνέχεια σε ένα **ανατροφοδοτούμενο δίκτυο** για την εξαγωγή των χρονικών του χαρακτηριστικών. Η διαδικασία ολοκληρώνεται με την ύπαρξη των στρωμάτων `global average pooling`, `flatten` και `dense`.

3.1 Δεδομένα και Προεπεξεργασία

Όπως έχει ήδη εξεταστεί στο εισαγωγικό κεφάλαιο, οι μαγνητικές τομογραφίες του εγκεφάλου αποτελούν έναν ενδεδειγμένο και αποτελεσματικό τρόπο για τη διάγνωση της νόσου Alzheimer. Το σύνολο δεδομένων που χρησιμοποιήθηκε για την αποπεράτωση της εργασίας αυτής αποκτήθηκε από το Alzheimer’s Disease Neuroimaging Initiative (ADNI). Το ADNI αποτελεί μία πρωτοβουλία η οποία παρέχει στους ερευνητές δεδομένα μελέτης ώστε να ορίσουν την εξέλιξη της νόσου Alzheimer. Οι ερευνητές του ADNI συλλέγουν, επικυρώνουν και χρησιμοποιούν δεδομένα, συμπεριλαμβανομένων εικόνων MRI και PET, γενετικής, γνωστικών δοκιμασιών, βιοδεικτών του CSF και του αίματος ως προγνωστικών της νόσου. Οι πόροι και τα δεδομένα της μελέτης διατίθενται διαδικτυακά και περιλαμβάνουν ασθενείς με νόσο Alzheimer, άτομα με ήπια γνωστική εξασθένηση και υγιή άτομα [57]. Το ADNI προτείνει τη χρήση ενός τυποποιημένου συνόλου δεδομένων, το οποίο έχει δημιουργηθεί για λόγους συνοχής στις διάφορες ερευνητικές εργασίες που το χρησιμοποιούν, και παρέχει πολύ ικανοποιητικά αποτελέσματα σε εφαρμογές συνελικτικών δικτύων. Αυτός είναι και ο λόγος που επιλέχθηκε το συγκεκριμένο σύνολο δεδομένων.

Οι συνολικά 3523 μαγνητικές τομογραφίες που αποκτήθηκαν ήταν αρχικά σε μορφή NIfTI, το οποίο αποτελεί έναν τύπο αρχείου που χρησιμοποιείται στην απεικόνιση ιατρικών

εικόνων που αφορούν στην νευροεπιστήμη. Για λόγους ευκολίας και απλοποίησης της υλοποίησης, έγινε τροποποίηση της μορφής NIfTI σε PNG. Επειδή η NIfTI μορφή είναι μία τρισδιάστατη μέθοδος απεικόνισης, για την μετατροπή σε PNG, η οποία είναι μία δισδιάστατη απεικόνιση, έπρεπε να επιλεγεί ο κατάλληλος άξονας κατά μήκος του οποίου θα γινόταν ο “τεμαχισμός” της τρισδιάστατης απεικόνισης. Όπως αναφέρθηκε και στο κεφάλαιο της εισαγωγής, η πιο αντιπροσωπευτική απεικόνιση για τη νόσο Alzheimer γίνεται στις στεφανιαίες τομές της μαγνητικής τομογραφίας του εγκεφάλου, επομένως ο “τεμαχισμός” επιλέχθηκε να γίνει κατά μήκος του άξονα y . Έτσι για καθεμία από τις αρχικές τομογραφίες προέκυψαν έως και 200 (ο αριθμός διαφέρει ανά τομογραφία) slices και χωρίστηκαν σε τρεις κλάσεις κατά αυτόν τον τρόπο:

- **AD (Alzheimer’s Disease):** MRIs ατόμων που νοσούν από τη νόσο Alzheimer (698 MRIs)
- **MCI (Mild Cognitive Impairment):** MRIs ατόμων με ήπια γνωστική εξασθένηση (1673 MRIs)
- **CN (Cognitively Normal):** MRIs υγιών ατόμων (1152 MRIs)

3.2 Επαύξηση Δεδομένων (Data Augmentation)

Η επαύξηση δεδομένων ακολουθεί την προσέγγιση της δημιουργίας πρόσθετων δεδομένων εκπαίδευσης από τα υπάρχοντα παραδείγματα με την επαύξησή τους, χρησιμοποιώντας τυχαίους μετασχηματισμούς που παράγουν αληθοφανείς εικόνες. Αυτό βοηθάει στην έκθεση του μοντέλου σε περισσότερες πτυχές των δεδομένων και στην καλύτερη γενίκευσή του.

Η επαύξηση δεδομένων αποτελεί ένα τρόπο αντιμετώπισης της υπερ-εκπαίδευσης, η οποία συμβαίνει είτε όταν το σύνολο δεδομένων δεν είναι επαρκές είτε όταν εκείνο απαρτίζεται από όμοια δεδομένα εκπαίδευσης, γεγονός που συμβαίνει στο σύνολο δεδομένων της παρούσας εργασίας. Οι τροποποιήσεις που πραγματοποιούνται κατά την εφαρμογή της μεθόδου έγκειται σε διάφορα χαρακτηριστικά των εικόνων όπως η κλίση, η μεγέθυνση, ο προσανατολισμός κ.α., συμβάλλοντας έτσι στην αρτιότερη εκπαίδευση του μοντέλου [58].

3.3 ResNet

Τα βαθιά συνελκτικά δίκτυα έχουν οδηγήσει σε μια σειρά από ανακαλύψεις όσον αφορά στην ταξινόμηση των εικόνων. Ποικίλες μελέτες έχουν αναδείξει τον παράγοντα του βάθους ενός δικτύου ως υψίστης σημασίας για την απόδοση του μοντέλου [59], [60]. Ενώ πολλαπλές εργασίες οπτικής αναγνώρισης έχουν επωφεληθεί από μοντέλα με μεγάλο βάθος, εκείνο ελλοχεύει τον κίνδυνο της εξαφάνισης ή έκρηξης της κλίσης (vanishing, exploding gradient), δυσχεραίνοντας την διαδικασία της εκπαίδευσης του μοντέλου. Το πρόβλημα αυτό έχει αντιμετωπιστεί σε μεγάλο βαθμό με τη συνδρομή της κανονικοποιημένης αρχικοποίησης και των ενδιάμεσων στρωμάτων κανονικοποίησης, τα οποία επιτρέπουν σε δίκτυα με δεκάδες

στρώματα να αρχίσουν να συγκλίνουν στον stochastic gradient descent με οπισθοδιάδοση [61].

Βέβαια, οι τεχνικές που προαναφέρθηκαν δεν επαρκούν πάντα ώστε να βελτιωθεί η επίδοση του δικτύου. Όταν τα βαθύτερα δίκτυα είναι σε θέση να αρχίσουν να συγκλίνουν, δημιουργείται ένα πρόβλημα υποβάθμισης (degradation problem): με την αύξηση του βάθους του δικτύου, η ακρίβεια μπορεί να κορεστεί και στη συνέχεια υποβαθμίζεται ραγδαία.

Οι αρχιτεκτονικές τύπου ResNet, εισάγοντας την έννοια της Υπολειπόμενης Εκμάθησης (Residual Learning), επιτυγχάνουν υψηλού επιπέδου αποτελέσματα, διατηρώντας μεγάλο βάθος στα δίκτυα.

3.3.1 Υπολειπόμενη Εκμάθηση (Residual Learning)

Θεωρούμε $\mathcal{H}(x)$ μία υποκείμενη χαρτογράφηση που πρέπει να ενσωματωθεί σε κάποια στοιβαγμένα στρώματα (όχι απαραίτητα σε όλο το δίκτυο), με το x να συμβολίζει τις εισόδους στο πρώτο από αυτά τα στρώματα. Υποθέτοντας πως πολλαπλά μη γραμμικά στρώματα μπορούν να προσεγγίσουν ασυμπτωτικά πολύπλοκες συναρτήσεις, τότε μπορούν να πράξουν αναλόγως και για υπολειπόμενες συναρτήσεις όπως είναι η $\mathcal{H}(x) - x$. Επομένως, αντί τα στοιβαγμένα στρώματα να προσεγγίζουν την $\mathcal{H}(x)$, εκείνα προσεγγίζουν την $\mathcal{F}(x) := \mathcal{H}(x) - x$. Με τον νέο αυτό ορισμό η αρχική συνάρτηση ισοδυναμεί με $\mathcal{F}(x) + x$.

Η ανάγκη για αυτόν τον ανασχηματισμό γεννήθηκε λόγω του προβλήματος υποβάθμισης που αναφέρθηκε ανωτέρω. Αν κάποια επιπλέον επίπεδα κατασκευάζονται ως χαρτογραφήσεις ταυτότητας (identity mappings), ένα βαθύτερο μοντέλο δεν θα πρέπει να έχει σφάλμα εκμάθησης μεγαλύτερο από εκείνο που εμφανίζεται σε στρώματα ανώτερων επιπέδων. Το πρόβλημα υποβάθμισης δυσκολεύει την προσέγγιση χαρτογραφήσεων ταυτότητας από πολλαπλά μη γραμμικά στρώματα. Με την υπολειπόμενη εκμάθηση, αν η χαρτογραφήσεις ταυτότητας είναι βέλτιστη, τα βάρη των μη γραμμικών αυτών στρωμάτων μπορούν να αναχθούν στο μηδέν ώστε να την πλησιάσουν.

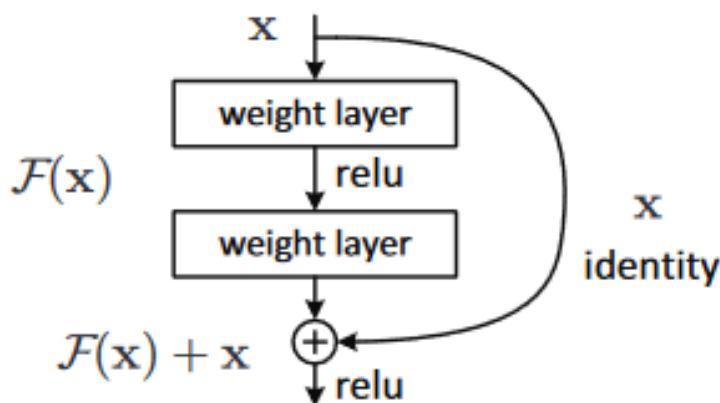
3.3.2 Χαρτογράφηση Ταυτότητας με Συντομεύσεις (Identity Mapping by Shortcuts)

Ορίζουμε ένα δομικό block ως εξής:

$$y = \mathcal{F}(x, \{W_i\}) + x \quad (3.1)$$

Τα x, y συμβολίζουν τα διανύσματα εισόδου και εξόδου των στρωμάτων, ενώ η συνάρτηση $\mathcal{F}(x, \{W_i\})$ την υπολειπόμενη χαρτογράφηση στην οποία πρέπει να γίνει εκμάθηση. Στο παρακάτω σχήμα απεικονίζεται η αρχιτεκτονική ενός δομικού block.

Βάσει του παραπάνω σχήματος όπου το block διαθέτει δύο στρώματα, η συνάρτηση διαμορφώνεται ως $\mathcal{F} = W_2 \sigma(W_1 x)$ με το σ να συμβολίζει την συνάρτηση ενεργοποίησης ReLU (οι πλώσεις παραλείπονται για απλοποίηση). Η πράξη $\mathcal{F} + x$ πραγματοποιείται μέσω μίας σύνδεσης συντόμευσης και πρόσθεσης ανά στοιχείο. Οι συνδέσεις συντόμευσης από την εξίσωση 3.1 δεν επιβαρύνουν το μοντέλο ούτε με επιπλέον παραμέτρους αλλά ούτε και με υπολογιστική πολυπλοκότητα.



Εικόνα 3.1: Αρχιτεκτονική ενός δομικού block [9]

Οι διαστάσεις του x και της \mathcal{F} θα πρέπει να ταυτίζονται, ειδάλλως συντελείται γραμμική προβολή W_s από τις συνδέσεις συντόμευσης ώστε να ταιριάζουν οι διαστάσεις:

$$y = \mathcal{F}(x, \{W_i\}) + W_s x \quad (3.2)$$

Παράλληλα μπορεί να χρησιμοποιηθεί ένας τετραγωνικός πίνακας W_s στην εξίσωση 3.1, όμως βάσει πειραμάτων που έχουν εξαχθεί, αποδεικνύεται ότι η χαρτογράφηση ταυτότητας αποτελεί μία επαρκή και οικονομική μέθοδο για την αντιμετώπιση του προβλήματος υποτίμησης.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Εικόνα 3.2: Αρχιτεκτονική των μοντέλων τύπου ResNet [9]

3.3.3 Εφαρμογές

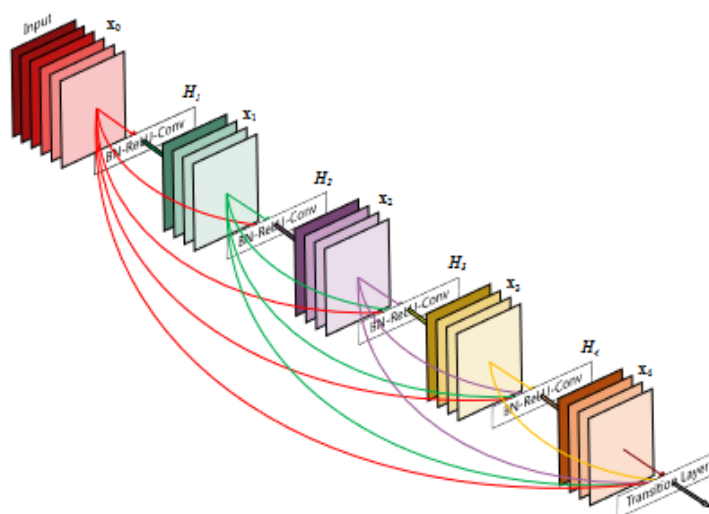
Τα δίκτυα τύπου Resnet, όπως το ResNet50 που έχει χρησιμοποιηθεί στην παρούσα εργασία, έχουν πολυάριθμες εφαρμογές κυρίως σε προβλήματα κατηγοριοποίησης εικόνων. Πιο συγκεκριμένα η χρήση τους έγκειται συχνά σε ιατρικές απεικονίσεις όπως είναι ο εντοπισμός και η κατηγοριοποίηση όγκων του εγκεφάλου [62], [63] αλλά και η διάγνωση της

νόσου Alzheimer μέσω εγκεφαλικών μαγνητικών τομογραφιών [64].

3.4 DenseNet

Τα δίκτυα τύπου DenseNet καλούνται να επιλύσουν το ίδιο πρόβλημα με παραπάνω, δηλαδή εκείνο της εξαφάνισης ή έκρηξης κλίσης που σημειώνεται στα βαθιά συνελκτικά δίκτυα. Τόσο η αρχιτεκτονική ResNet όσο και άλλες που έχουν προταθεί για την αντιμετώπιση αυτού του ζητήματος, δημιουργούν σύντομα μονοπάτια από τα ανώτερα στα κατώτερα στρώματα.

Τα δίκτυα DenseNet προτείνουν μία αρχιτεκτονική που βασίζεται σε ένα απλό μοτίβο συνδεσιμότητας: για να εξασφαλιστεί η μέγιστη δυνατή ροή πληροφοριών μεταξύ των επιπέδων του δικτύου, όλα τα επίπεδα (με αντίστοιχα μεγέθη χαρτών χαρακτηριστικών) συνδέονται απευθείας μεταξύ τους. Για να διατηρηθεί ο χαρακτήρας της πρόσθιας διάδοσης, κάθε στρώμα λαμβάνει πρόσθετες εισόδους από όλα τα προηγούμενα στρώματα και διαβιβάζει τους δικούς του χάρτες χαρακτηριστικών σε όλα τα επόμενα στρώματα. Η εικόνα 3.3 απεικονίζει σχηματικά αυτή τη διάταξη. Είναι κεφαλαιώδους σημασίας ότι, σε αντίθεση με τα ResNets, δεν συνδυάζονται ποτέ χαρακτηριστικά μέσω της άθροισης προτού περάσουν σε ένα επίπεδο. Αντίθετα, συνδυάζονται με τη συνένωσή τους [10].



Εικόνα 3.3: Ένα dense block 5 στρωμάτων με ρυθμό ανάπτυξης $\kappa = 4$. Κάθε στρώμα λαμβάνει ως είσοδο όλους τους προηγούμενους χάρτες χαρακτηριστικών [10]

3.4.1 Πυκνή συνδεσιμότητα (Dense connectivity)

Όπως αναφέρθηκε και παραπάνω, στο παρόν μοντέλο προτείνεται η απευθείας σύνδεση ενός στρώματος με όλα τα ακόλουθά του. Έτσι το ℓ -οστό στρώμα θα λαμβάνει τους χάρτες χαρακτηριστικών όλων των προγενέστερων στρωμάτων $x_0, \dots, x_{\ell-1}$, σαν είσοδο:

$$x_\ell = H_\ell([x_0, x_1, \dots, x_{\ell-1}]) \quad (3.3)$$

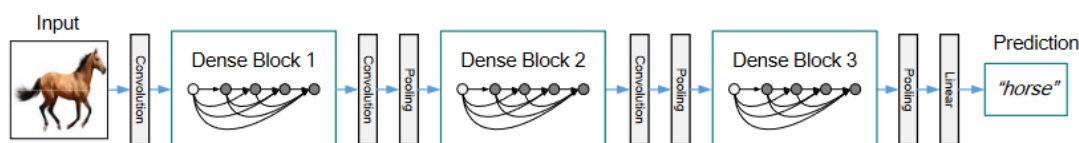
όπου τα $[x_0, \dots, x_{\ell-1}]$ αφορούν στη συνένωση των χαρτών χαρακτηριστικών που παράγονται στα στρώματα $0, \dots, \ell - 1$. Για λόγους ευκολίας στην υλοποίηση, συνενώνουμε τις πολλαπλές εισόδους της $H_\ell(\cdot)$ σε έναν μεμονωμένο τανυστή.

3.4.2 Σύνθετη συνάρτηση (Composite function)

Σύμφωνα με όσα αναλύονται στο [65], η συνάρτηση $H_\ell(\cdot)$ ορίζεται σαν μία σύνθετη συνάρτηση η οποία απαρτίζεται από τρεις διαδοχικές συναρτήσεις: την κανονικοποίηση παρτίδας (batch normalization), την συνάρτηση ενεργοποίησης ReLU και μία 3×3 συνέλιξη.

3.4.3 Στρώματα Συσσώρευσης (Pooling layers)

Η συνάρτηση συνένωσης της σχέσης 3.3 δεν είναι βιώσιμη όταν το μέγεθος των χαρτών χαρακτηριστικών μεταβάλλεται. Για το λόγο αυτό, η υποδειγματοληψία, η οποία περιγράφηκε και στο προηγούμενο κεφάλαιο, αποτελεί αναπόσπαστο κομμάτι των συνεκτικών δικτύων. Για να διευκολυνθεί η διαδικασία αυτή, το δίκτυο χωρίζεται σε πολυάριθμα dense blocks, στα οποία συναντάται πυκνή διασύνδεση. Η δομή ενός δικτύου τύπου DenseNet με την ύπαρξη των dense blocks παρουσιάζεται στην εικόνα 3.4. Τα στρώματα ανάμεσα σε δύο γειτονικά blocks αναφέρονται σαν στρώματα μετάβασης (transition layers) και μεταβάλλουν τον χάρτη χαρακτηριστικών μέσω της συνέλιξης και της συσσώρευσης.



Εικόνα 3.4: Ένα δίκτυο DenseNet με τρία dense blocks. Τα στρώματα ανάμεσα σε δύο γειτονικά blocks αναφέρονται σαν στρώματα μετάβασης [10]

3.4.4 Ρυθμός ανάπτυξης (Growth rate)

Αν κάθε συνάρτηση $H_\ell(\cdot)$ παράγει k χάρτες χαρακτηριστικών, τότε το ℓ -οστό στρώμα θα έχει $k_0 + k \times (\ell - 1)$ χάρτες χαρακτηριστικών, όπου το k_0 θα είναι ο αριθμός των καναλιών στο στρώμα εισόδου. Η υπερ-παράμετρος k αναφέρεται ως *ρυθμός ανάπτυξης (growth rate)* του δικτύου. Με έναν σχετικά μικρό ρυθμό ανάπτυξης μπορεί να επιτευχθεί ένα άρτιο αποτέλεσμα, γεγονός που οφείλεται στο ότι κάθε στρώμα έχει πρόσβαση σε όλους τους προηγούμενους χάρτες χαρακτηριστικών και των block τους, και κατέπεκταση στην "συνολική γνώση" του δικτύου. Αν οι χάρτες χαρακτηριστικών θεωρηθούν ως η συνολική κατάσταση του δικτύου κάθε στρώμα της προσθέτει k χάρτες χαρακτηριστικών του, ο ρυθμός ανάπτυξης ρυθμίζει τη συμβολή κάθε στρώματος στην ποσότητα καινούριας πληροφορίας στην συνολική αυτή κατάσταση.

3.4.5 Στρώματα Συμφόρησης (Bottleneck layers)

Παρόλο που το κάθε στρώμα έχει ως έξοδο k χάρτες χαρακτηριστικών, έχει πολύ περισσότερες εισόδους. Στις βιβλιογραφικές πηγές [66], [67] αναφέρεται ότι μία 1×1 συνέλιξη

μπορεί να θεωρηθεί σαν στρώμα συμφόρησης (bottleneck layer) πριν κάθε 3×3 συνέλιξη περιορίσει τον αριθμό των χαρτών χαρακτηριστικών της εισόδου, βελτιώνοντας έτσι την υπολογιστική απόδοση του μοντέλου.

3.4.6 Συμπίεση (Compression)

Για να ενισχυθεί περαιτέρω η ύπαρξη ενός συμπαγούς δικτύου, ο αριθμός των χαρτών χαρακτηριστικών μπορεί να μειωθεί στα στρώματα μετάβασης. Αν ένα dense block περιέχει m χάρτες χαρακτηριστικών, επιτρέπουμε στο επόμενο στρώμα μετάβασης να δημιουργήσει $\lfloor \theta m \rfloor$ χάρτες χαρακτηριστικών ως έξοδο, όπου το $0 < \theta \leq 1$ συμβολίζει τον παράγοντα συμπίεσης. Όταν το $\theta = 1$, ο αριθμός των χαρτών στα στρώματα μετάβασης παραμένει αμετάβλητος.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Εικόνα 3.5: Αρχιτεκτονική των μοντέλων τύπου DenseNet [10]

3.4.7 Εφαρμογές

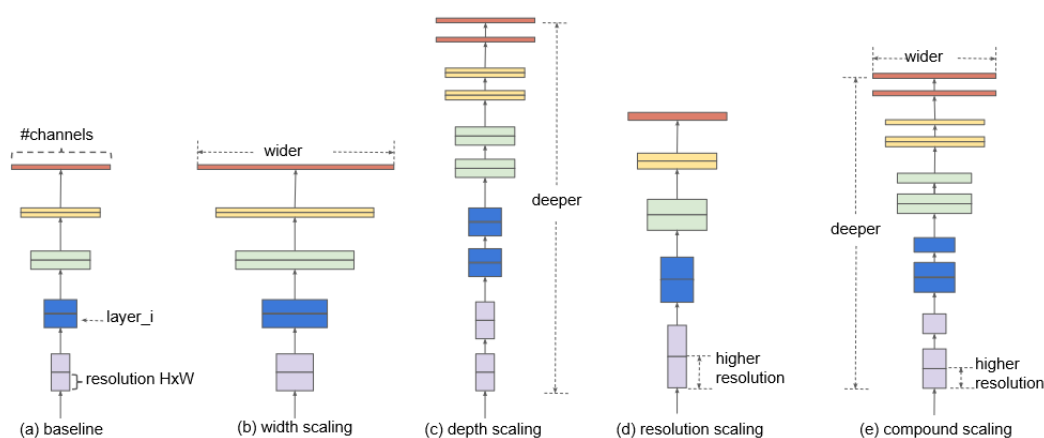
Στην επιλογή των δικτύων τύπου DenseNet, και ειδικότερα του DenseNet-121, συνέβαλε το γεγονός ότι συναντάται σε ποικίλες εφαρμογές κατηγοριοποίησης εικόνων και εντοπισμού εγκεφαλικών αλλοιώσεων μέσω μαγνητικών απεικονίσεων [68], [69].

3.5 EfficientNet

Το EfficientNet αποτελεί μία αρχιτεκτονική δικτύου που προτείνεται στο [11]. Τα συνελκτικά νευρωνικά δίκτυα συνήθως αναπτύσσονται με σταθερό προϋπολογισμό πόρων, και στη συνέχεια αναβαθμίζονται για μεγαλύτερη ακρίβεια αν περισσότεροι πόροι είναι διαθέσιμοι. Έπειτα από συστηματική έρευνα που έχει διεξαχθεί για την κλιμάκωση των μοντέλων, έχει αποδειχθεί πως η προσεκτική εξισορρόπηση του βάθους, του πλάτους και της ανάλυσης του δικτύου μπορεί να οδηγήσει σε καλύτερες επιδόσεις. Την ιδιότητα αυτή εκμεταλλεύεται η αρχιτεκτονική των δικτύων τύπου EfficientNet, παρουσιάζοντας μία νέα μέθοδο κλιμάκωσης

που κλιμακώνει ομοιόμορφα όλες τις διαστάσεις του βάθους/πλάτους/ανάλυσης και εφαρμόζοντάς την με αποτελεσματικότητα σε δίκτυα τύπου MobileNet [70] και ResNet. Αξίζει να αναφερθεί ότι τα δίκτυα EfficientNet όχι μόνο επιτυγχάνουν εξαιρετική ακρίβεια αλλά συγχρόνως διαθέτουν μικρότερο μέγεθος και μεγαλύτερη ταχύτητα.

Όπως λοιπόν προαναφέρθηκε, για την ανάπτυξη του EfficientNet, προτάθηκε μία απλή αλλά αποτελεσματική μέθοδος σύνθετης κλιμάκωσης (*compound scaling method*), σύμφωνα με την οποία τα μεγέθη του βάθους, του πλάτους και της ανάλυσης κλιμακώνονται ομοιόμορφα με ένα σύνολο σταθερών συντελεστών κλιμάκωσης. Για παράδειγμα, αν χρησιμοποιηθούν 2^N φορές περισσότερους υπολογιστικούς πόρους, τότε απλώς θα αυξηθεί το βάθος κατά a^N , το πλάτος κατά β^N και το μέγεθος της εικόνας κατά γ^N , όπου a, β, γ είναι οι σταθεροί συντελεστές που έχουν οριστεί βάσει ενός grid search.



Εικόνα 3.6: Συμβατικές μέθοδοι κλιμάκωσης (a-d), σε αντίθεση με την μέθοδο σύνθετης κλιμάκωσης (e) [11]

Στην μέθοδο σύνθετης κλιμάκωσης χρησιμοποιείται ένας σύνθετος συντελεστής ϕ που κλιμακώνει ομοιόμορφα τις διαστάσεις σύμφωνα με τον παρακάτω τρόπο:

- Βάθος: $d = a^\phi$
- Πλάτος: $w = \beta^\phi$
- Ανάλυση: $r = \gamma^\phi$

Διασηθητικά, ο καθορισμός του παράγοντα ϕ έγκειται στη διακριτική ευχέρεια του χρήστη και ορίζει πόσοι πόροι είναι διαθέσιμοι για την κλιμάκωση του μοντέλου, ενώ τα a, β, γ αποφασίζουν πώς θα ανατεθούν αυτοί οι επιπλέον πόροι στις διαστάσεις του δικτύου.

Τα FLOPS (Floating-point Operations per Second) μίας συμβατικής συνέλιξης είναι ανάλογα των d, w^2, r^2 , ενώ με τη μέθοδο σύνθετης κλιμάκωσης ο συνολικός αριθμός των FLOPS αυξάνεται κατά $(a \cdot \beta^2 \cdot \gamma^2)^\phi$. Βάσει της προτεινόμενης μεθόδου ο παράγοντας $(a \cdot \beta^2 \cdot \gamma^2)$ περιορίζεται στο $a \cdot \beta^2 \cdot \gamma^2 \approx 2$ έτσι ώστε για κάθε νέο ϕ , ο αριθμός των των FLOPS να αυξάνεται περίπου κατά 2^ϕ .

Με τη συνδρομή του [71], αναπτύχθηκε το βασικό δίκτυο, αξιοποιώντας μία αναζήτηση στο δίκτυο που βελτιστοποιεί τόσο την ακρίβεια όσο και τα FLOPS. Ο στόχος βελτιστοποίησης ορίζεται ως:

$$ACC(m) \times [FLOPS(m)/T]^w \quad (3.4)$$

όπου:

- m : το μοντέλο
- $ACC(m)$: η ακρίβεια του μοντέλου m
- $FLOPS(m)$: τα FLOPS του μοντέλου m
- T : τα FLOPS - στόχοι
- w : υπερ-παράμετρος που ελέγχει την αντιστάθμιση μεταξύ της ακρίβειας και των FLOPS και ισούται με -0.07

Με αυτόν τον τρόπο, παράγεται ένα αποτελεσματικό δίκτυο το οποίο ονομάστηκε EfficientNet-B0 και αποτελεί τη βάση των δικτύων τύπου EfficientNet. Τα βασικά δομικά του blocks αποτελούν τα mobile inverted bottleneck MBConv [12], στα οποία προστίθεται η βελτιστοποίηση squeeze-and-excitation [13].

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels \hat{C}_i	#Layers \hat{L}_i
1	Conv3x3	224×224	32	1
2	MBConv1, k3x3	112×112	16	1
3	MBConv6, k3x3	112×112	24	2
4	MBConv6, k5x5	56×56	40	2
5	MBConv6, k3x3	28×28	80	3
6	MBConv6, k5x5	14×14	112	3
7	MBConv6, k5x5	14×14	192	4
8	MBConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

Εικόνα 3.7: Η βασική αρχιτεκτονική του δικτύου EfficientNet-B0. Κάθε γραμμή περιγράφει ένα στάδιο i με \hat{L}_i στρώματα, ανάλυση εισόδου $\langle \hat{H}_i, \hat{W}_i \rangle$ και έξοδο τα κανάλια \hat{C}_i [11]

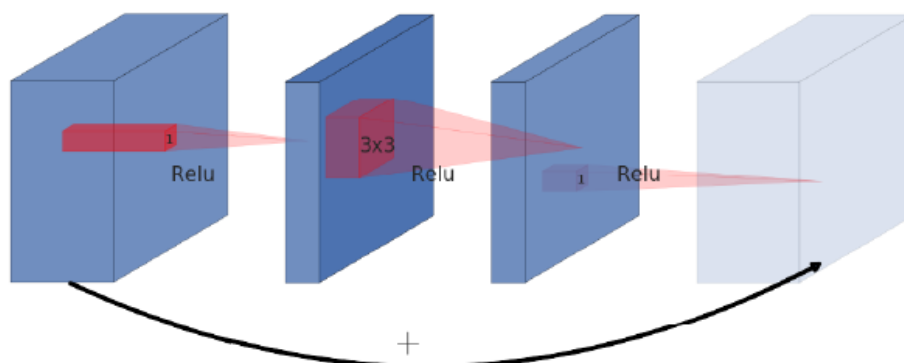
3.5.1 Mobile inverted bottleneck MBConv

Ένα inverted residual block, το οποίο συχνά αναφέρεται ως MBConv block, είναι ένας τύπος residual block, ο οποίος εφαρμόζεται σε μοντέλα εικόνας που χρησιμοποιούν μια ανεστραμμένη δομή για λόγους αποδοτικότητας. Προτάθηκε αρχικά για την αρχιτεκτονική MobileNetV2 και έκτοτε επαναχρησιμοποιήθηκε για ποικίλες βελτιστοποιήσεις συνελκτικών μοντέλων όπως είναι το EfficientNet.

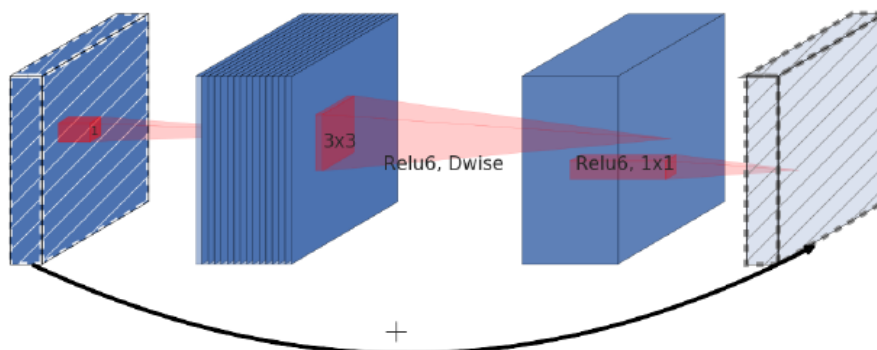
Ένα παραδοσιακό residual block έχει μια ευρεία \rightarrow στενή \rightarrow ευρεία δομή με τον αριθμό των καναλιών. Η είσοδος έχει μεγάλο αριθμό καναλιών, τα οποία συμπιέζονται με μια

συνέλιξη 1×1 . Ο αριθμός των καναλιών στη συνέχεια αυξάνεται ξανά με μια 1×1 συνέλιξη, ώστε να μπορούν να προστεθούν η είσοδος και η έξοδος.

Αντίθετα, στο inverted residual block ακολουθείται μια στενή \rightarrow ευρεία \rightarrow στενή προσέγγιση, για αυτό λαμβάνει χώρα και η αντιστροφή. Πρώτα διευρύνεται με μια 1×1 συνέλιξη, στη συνέχεια χρησιμοποιείται μια 3×3 συνέλιξη βάθους (η οποία μειώνει σημαντικά τον αριθμό των παραμέτρων) και τέλος εφαρμόζεται μια 1×1 συνέλιξη για να μειωθεί ο αριθμός των καναλιών ώστε να μπορούν να προστεθούν η είσοδος και η έξοδος [72].



Εικόνα 3.8: *Residual block* [12]



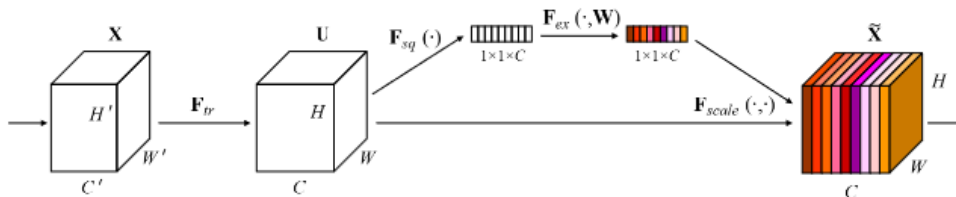
Εικόνα 3.9: *Inverted residual block* [12]

3.5.2 Squeeze-and-excitation

Το squeeze-and-excitation block είναι μια αρχιτεκτονική μονάδα που έχει σχεδιαστεί για να βελτιώσει την αντιπροσωπευτική ισχύ ενός δικτύου επιτρέποντάς του να εκτελεί δυναμική αναβαθμονόμηση χαρακτηριστικών ανά κανάλι. Η διαδικασία που ακολουθείται από τη μέθοδο είναι η εξής [73]:

- Το block έχει ως είσοδο ένα συνελικτικό block.
- Κάθε κανάλι συμπιέζεται σε μια ενιαία αριθμητική τιμή με τη χρήση της συσσώρευσης με βάση τον μέσο όρο (average pooling).

- Ένα πυκνό στρώμα (dense layer) ακολουθούμενο από μία συνάρτηση ενεργοποίησης ReLU προσθέτει μη γραμμικότητα και η πολυπλοκότητα των καναλιών εξόδου μειώνεται κατά μία αναλογία.
- Ένα άλλο πυκνό στρώμα ακολουθούμενο από μία σιγμοειδή συνάρτηση ενεργοποίησης δίνει σε κάθε κανάλι μια ομαλή συνάρτηση πύλης.
- Τέλος, κάθε χάρτης χαρακτηριστικών του συνελκτικού block σταθμίζεται συναρτήσει του πλευρικού δικτύου.

Εικόνα 3.10: *Squeeze-and-excitation block* [13]

3.5.3 Εφαρμογές

Όπως και τα προηγούμενα μοντέλα, έτσι και τα δίκτυα τύπου EfficientNet, συμπεριλαμβανομένου του EfficientNet-B0 που χρησιμοποιήθηκε στην προκείμενη περίπτωση, έχουν σημειώσει υψηλές επιδόσεις (υψηλότερες μάλιστα από τα προγενέστερα μοντέλα) στις διάφορες εφαρμογές επεξεργασίας και ανάλυσης μαγνητικών τομογραφιών του εγκεφάλου [74], [75].

3.6 CNN - RNN

Ο συνδυασμός CNN και RNN αρχιτεκτονικών για την πραγματοποίηση μίας εργασίας κατηγοριοποίησης εικόνων, αποτελεί ένα ιδιαίτερα αποτελεσματικό υβριδικό μοντέλο που χρησιμοποιείται ευρέως [76], [77], [78]. Αφενός τα CNN αξιοποιούν τη χωρική διάσταση της εισόδου, εξαγάγοντας τα χαρακτηριστικά των δεδομένων και αφετέρου τα RNN εκμεταλλεύονται τη χρονική τους διάσταση όταν πρόκειται για ακολουθιακά δεδομένα, όπως συμβαίνει στην περίπτωση της παρούσας εργασίας.

Στο προκείμενο μοντέλο, το σύνολο δεδομένων δόθηκε ως είσοδος σε ένα συνελκτικό δίκτυο ώστε να εντοπιστούν και να εξαχθούν τα χωρικά χαρακτηριστικά των εικόνων των τομογραφιών. Όπως έχει ήδη επισημανθεί, τα συνελκτικά δίκτυα που χρησιμοποιήθηκαν για να εξυπηρετήσουν αυτό το σκοπό είναι τα ResNet50, DenseNet121 και EfficientNet-B0. Στη συνέχεια προστέθηκε επιπλέον ένα στρώμα global average pooling.

Η έξοδος που παράχθηκε από την ανωτέρω διαδικασία τροφοδοτήθηκε σε ένα time distributed στρώμα το οποίο αποτελεί μία εφαρμογή των LSTM ανατροφοδοτούμενων δικτύων. Το στρώμα αυτό εφαρμόζεται σε κάθε ένα από τα χρονικά ακολουθιακά δεδομένα και συμβάλει στην περάτωση της διαδικασίας της συνέλιξης των δεδομένων, διατηρώντας την χρονική

τους ακεραιότητα [79], [80]. Έπειτα από πειράματα που έγιναν, ο αριθμός των GRU(Gated Recurrent Unit) ορίστηκε στο 8 καθώς μας έδινε τα καλύτερα αποτελέσματα.

Η διαδικασία αυτή ακολουθήθηκε από την ύπαρξη ενός στρώματος flatten ώστε να μειωθεί η διάσταση των διανυσμάτων σε 1, βήμα το οποίο απαιτείται για την τελική εισαγωγή σε ένα dense στρώμα που θα παράγει το τελικό αποτέλεσμα.

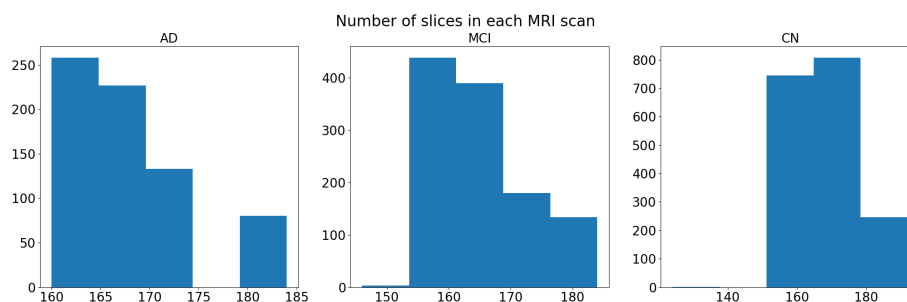
Πειραματική Διαδικασία

4.1 Μέθοδοι και Παραμετροποίηση

Για την διεκπεραίωση της μεθοδολογίας που αναλύθηκε στο προηγούμενο κεφάλαιο, κεφαλαιώδους σημασίας είναι η ορθή παραμετροποίηση για την εξαγωγή των βέλτιστων δυνατών αποτελεσμάτων. Οι παράμετροι που ορίστηκαν στα πειράματα της παρούσας εργασίας προέκυψαν έπειτα από εμπειρικές δοκιμές.

Αρχικά το μέγεθος των slices που αντιστοιχούν σε κάθε μαγνητική τομογραφία ορίστηκε σε 224×224 και στη συνέχεια δημιουργούνται δύο αντίγραφα κάθε slice ώστε να αποκτήσουν τελική διάσταση $224 \times 224 \times 3$. Αυτό συμβαίνει καθώς τα συνελκτικά δίκτυα στα οποία εισάγονται εν συνεχεία τα slices λαμβάνουν ως είσοδο εικόνες με 3 κανάλια.

Επιπλέον δημιουργήθηκε μία μάσκα ώστε να γίνεται padding στον αριθμό των slices και η παράμετρος max pad len ρυθμίστηκε στο 200, καθώς όπως φαίνεται και από την εικόνα 4.1, ο αριθμός των slices ανά MRI σε κάθε κλάση δεν υπερβαίνει τον αριθμό αυτόν, επομένως δεν υπάρχει απώλεια πληροφορίας.



Εικόνα 4.1: Αριθμός των slices ανά MRI σε καθεμία από τις κλάσεις AD, MCI, CN

Ο διαχωρισμός του συνόλου δεδομένων σε σύνολο εκπαίδευσης (training set) και σύνολο επαλήθευσης (validation set) έγινε ως εξής: 70% (training set) και 30% (validation set). Το σύνολο εκπαίδευσης υφίσταται ώστε να μπορέσει να εκπαιδευτεί το μοντέλο σε ένα σύνολο δεδομένων, ενώ το σύνολο επαλήθευσης ελέγχει την ικανότητα του μοντέλου να πραγματοποιεί ορθή κατηγοριοποίηση σε δεδομένα που δεν έχει επεξεργαστεί ξανά.

Παράλληλα, όσον αφορά στην επαύξηση δεδομένων στην οποία έχει γίνει αναφορά παραπάνω, επιλέχθηκε rotation (περιστροφή) κατά 28 μοίρες, αφού το επιθυμητό είναι να υπάρχει επαρκής διαφοροποίηση της μετασχηματισμένης εικόνας από την αρχική ώστε να

μην λαμβάνονται ως ταυτόσημες από το δίκτυο, αλλά όχι πολύ μεγάλη ώστε να αλλοιώνεται το περιεχόμενό της. Επιπλέον επιλέχθηκε horizontal flip.

Παρόλο που τα συνελκτικά μοντέλα που χρησιμοποιήθηκαν είναι προεκπαιδευμένα στο σύνολο ImageNet, επειδή αυτό διαθέτει αρκετά διαφορετικά δεδομένα από αυτά που έχει το παρόν σύνολο δεδομένων (δεν διαθέτει ιατρικά δεδομένα), ορίστηκαν 40 trainable στρώματα εκ νέου.

Ο αλγόριθμος βελτιστοποίησης που επιλέχθηκε για την εκπαίδευση όλων των μοντέλων είναι ο Adam, ενώ η συνάρτηση κόστους η sparse categorical crossentropy. Προκειμένου να ρυθμιστεί ο ρυθμός εκπαίδευσης, χρησιμοποιήθηκε η συνάρτηση tf.keras.optimizers.schedules.ExponentialDecay, με παραμέτρους: initial learning rate, decay steps=10000, decay rate=0.96, staircase=True.

Τέλος ο αριθμός των εποχών για τα μοντέλα ResNet-50 και DenseNet-121 ορίστηκε στις 30, ενώ για το μοντέλο EfficientNet-B0 στις 20. Ο λόγος που στο τελευταίο οι εποχές ήταν λιγότερες είναι επειδή έπειτα από πειράματα αποκαλύφθηκε πως για μεγαλύτερο αριθμό εποχών γίνεται υπερ-εκπαίδευση αλλοιώνοντας τα αποτελέσματα.

Το μεγαλύτερο μέρος της εκπαίδευσης του μοντέλου πραγματοποιήθηκε με τη χρήση της βιβλιοθήκης Tensorflow της γλώσσας προγραμματισμού python.

4.2 Μετρικές αξιολόγησης

Δεδομένου ότι η ακρίβεια χρησιμοποιείται σαν κύρια μετρική αξιολόγησης, η παράμετρος "metrics" ορίστηκε ως "accuracy". Στο Keras, η ακρίβεια δίνεται από τη σχέση:

$$\text{Accuracy} = \frac{\sum_{i=1}^n B^i}{n} \quad (4.1)$$

όπου n είναι ο αριθμός των δειγμάτων, και η συνάρτηση B^i είναι μία boolean συνάρτηση που ορίζεται ως εξής:

$$B^i = \begin{cases} 0, & y_{\text{true}}^i \neq y_{\text{predicted}}^i \\ 1, & y_{\text{true}}^i = y_{\text{predicted}}^i \end{cases} \quad (4.2)$$

δηλαδή παίρνει την τιμή 0 όταν γίνει λανθασμένη πρόβλεψη της ετικέτας ενός δείγματος, και την τιμή 1 όταν η πρόβλεψη είναι σωστή [20].

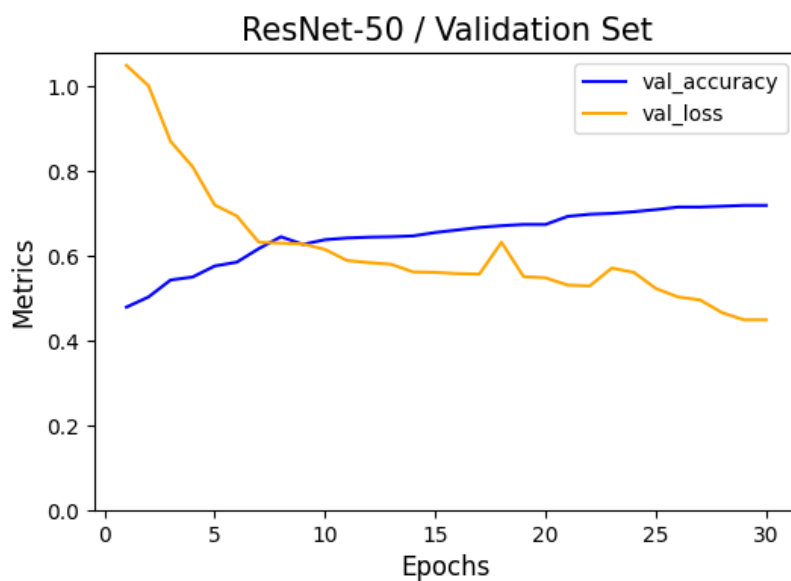
4.3 Αποτελέσματα

Παρακάτω παρατίθενται διαγραμματικά τα αποτελέσματα της κατηγοριοποίησης στις κλάσεις AD, MCI, CN των συνόλων εκπαίδευσης και επαλήθευσης με την εφαρμογή των μοντέλων Resnet-50, DenseNet-121, EfficientNet-B0.

Η ακρίβεια εκείνη που αποτελεί τον πιο αντιπροσωπευτικό παράγοντα αξιολόγησης των αποτελεσμάτων είναι η ακρίβεια του συνόλου επαλήθευσης (val accuracy, καθώς καταδεικνύει την ικανότητα του μοντέλου να προβλέψει σωστά την ετικέτα ενός δείγματος στο οποίο δεν έχει εκπαιδευτεί και αντιμετωπίζει πρώτη φορά. Λαμβάνοντας υπόψη αυτό, τα αποτελέσματα που εξήχθησαν είναι ικανοποιητικά. Ειδικότερα παρατηρούμε ότι το μοντέλο EfficientNet-B0,

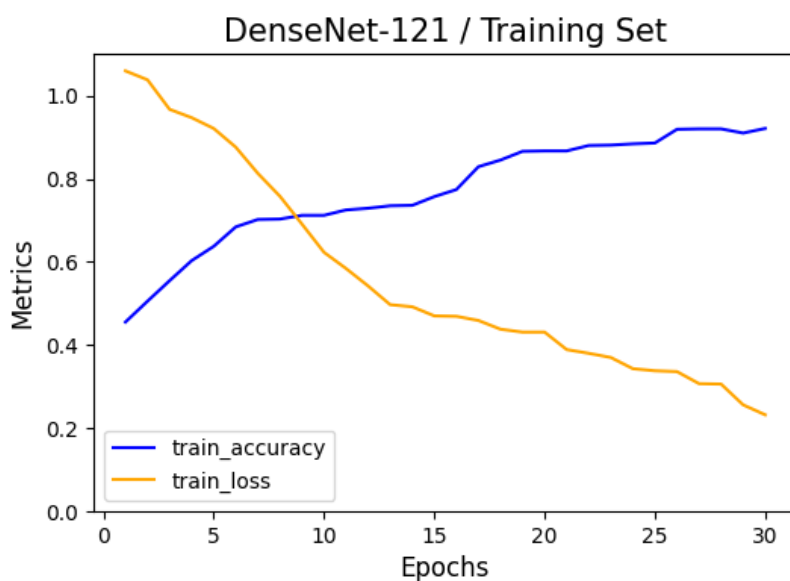


Εικόνα 4.2: Αποτελέσματα του μοντέλου ResNet-50 στο σύνολο εκπαίδευσης

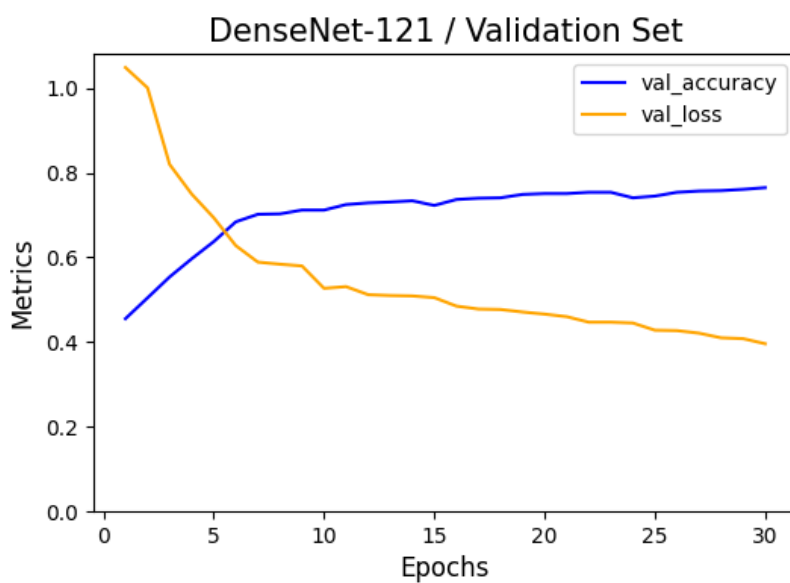


Εικόνα 4.3: Αποτελέσματα του μοντέλου ResNet-50 στο σύνολο επαλήθευσης

παρουσίασε εμφανώς καλύτερη απόδοση σε σχέση με τα άλλα δύο, ενώ το μοντέλο ResNet-50 ήταν εκείνο που σημείωσε τη χαμηλότερη ακρίβεια επαλήθευσης, χωρίς όμως αυτή να είναι απογοητευτική.



Εικόνα 4.4: Αποτελέσματα του μοντέλου DenseNet-121 στο σύνολο εκπαίδευσης

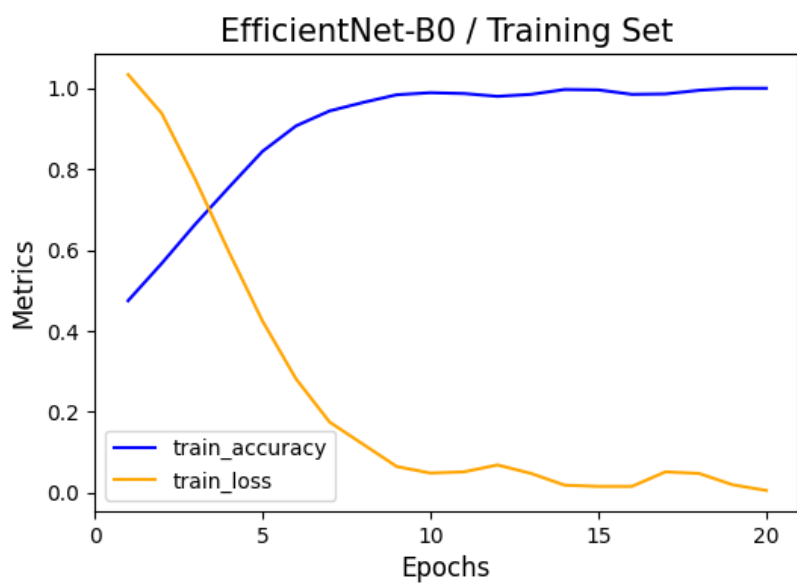


Εικόνα 4.5: Αποτελέσματα του μοντέλου DenseNet-121 στο σύνολο επαλήθευσης

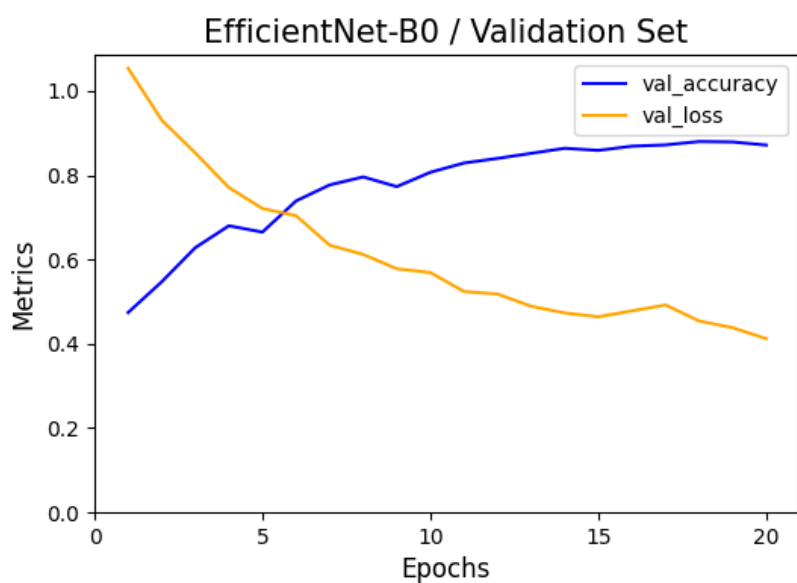
Σύνολο εκπαίδευσης

	ResNet-50	DenseNet-121	EfficientNet-B0
Accuracy	0.891	0.921	1.0

Πίνακας 4.1: Ακρίβεια του συνόλου εκπαίδευσης για τα μοντέλα Resnet-50, DenseNet-121, EfficientNet-B0



Εικόνα 4.6: Αποτελέσματα του μοντέλου *EfficientNet-B0* στο σύνολο εκπαίδευσης



Εικόνα 4.7: Αποτελέσματα του μοντέλου *EfficientNet-B0* στο σύνολο επαλήθευσης

Σύνολο επαλήθευσης

	ResNet-50	DenseNet-121	EfficientNet-B0
Accuracy	0.719	0.765	0.872

Πίνακας 4.2: Ακρίβεια του συνόλου επαλήθευσης για τα μοντέλα *Resnet-50*, *DenseNet-121*, *EfficientNet-B0*

Μέρος 

Επίλογος

Επίλογος

5.1 Συμπεράσματα

Στην παρούσα διπλωματική εργασία αναπτύχθηκε ένα αξιόπιστο μοντέλο κατηγοριοποίησης με στόχο τη διάγνωση της νόσου Alzheimer μέσω μαγνητικών τομογραφιών του εγκεφάλου. Πολύτιμους αρωγούς στο εγχείρημα αυτό αποτέλεσαν τόσο η εφαρμογή των προ-εκπαιδευμένων συνελκτικών δικτύων Resnet-50, DenseNet-121, EfficientNet-B0, και ιδιαίτερα του τελευταίου που παρουσίασε αξιοσημείωτη απόδοση, όσο και η ενσωμάτωση των ανατροφοδοτούμενων νευρωνικών δικτύων. Την χρήση των μοντέλων αυτών πλαισίωσε η απαραίτητη προ-επεξεργασία που έγινε στα δεδομένα, η οποία έγκειται στην επιλογή του κατάλληλου άξονα τομής της μαγνητικής τομογραφίας, ώστε να αξιοποιηθούν οι σημαντικότερες πληροφορίες, καθώς και στην επαύξησή τους ώστε να αποφευχθεί η υπερ-εκπαίδευση του μοντέλου. Οι παράγοντες αυτοί συντέλεσαν στην δημιουργία ενός δικτύου το οποίο σημείωσε αποτελέσματα υψηλής ακρίβειας.

5.2 Μελλοντικές Επεκτάσεις

Τα τεχνητά νευρωνικά δίκτυα αποτελούν έναν ραγδαία εξελισσόμενο κλάδο και οι εφαρμογές τους στις διάφορες πτυχές της ανθρώπινης ζωής, έχουν βελτιώσει σημαντικά την ποιότητά της. Ιδιαίτερα η ενσωμάτωσή τους στην ιατρική και ειδικότερα σε ασθένειες νευρολογικής φύσεως, όπως είναι το Alzheimer, μπορεί να αποκαλύψει νέους ορίζοντες στην αντιμετώπισή τους.

Τα αποτελέσματα της παρούσας εργασίας μπορούν να αποτελέσουν τον θεμέλιο λίθο για διάφορες μελλοντικές επεκτάσεις οι οποίες θα φέρουν την επιστήμη πιο κοντά στην πρόληψη και την θεραπεία αυτής της νόσου. Αρχικά καίριας σημασίας θα ήταν η υλοποίηση διαφόρων προσεγγίσεων του ίδιου προβλήματος με αυτό που αναπτύχθηκε εδώ. Ο ορισμός διαφορετικών κλάσεων, η χρήση διαφορετικών συνελκτικών μοντέλων καθώς και η απεικόνιση διαφορετικών τομών του εγκεφάλου αποτελούν κάποιες από τις παραμέτρους με τις οποίες μπορεί να γίνει μελλοντικός πειραματισμός και σύγκριση με την προκείμενη εργασία. Παράλληλα, μία προσπάθεια που θα μπορούσε να γίνει είναι η παρακολούθηση των ασθενών που κατηγοριοποιήθηκαν με ήπιες γνωστικές δυσλειτουργίες (MCI), ώστε να διαπιστωθεί αν μετέπειτα ανέπτυξαν ή όχι τη νόσο Alzheimer, γεγονός που θα συνέβαλλε στην πρόληψη της νόσου για άτομα με παρόμοιες εγκεφαλικές αλλοιώσεις. Μία άλλη ενδιαφέρουσα προ-

σέγγιση, θα ήταν η προσπάθεια αποκάλυψης των μοτίβων εκείνων του εγκεφάλου τα οποία οδηγούν βαθμιαία στην ανάπτυξη του Alzheimer μέσω των τεχνικών βαθιάς μάθησης. Αν επιτευχθεί αυτό, θα έχει πραγματοποιηθεί ένα σημαντικό βήμα στην προσπάθεια ανάσχεσης των μοτίβων αυτών μέσω του σχεδιασμού της κατάλληλης θεραπείας, προλαμβάνοντας έτσι την εξέλιξη της νόσου.

Βιβλιογραφία

- [1] Omega PDS MRI& CT Scan. *MRI can help diagnose alzheimer's disease*, 2019.
- [2] Chris M. Bishop. *Neural networks and their applications*, 1994.
- [3] Kenji Suzuki. *Artificial Neural Networks: Methodological advances and biomedical applications*. InTech, 2011.
- [4] *Activation function*.
- [5] Sebastian Ruder. *An overview of gradient descent optimization algorithms*, 2017.
- [6] Shadman Sakib, Nazib Ahmed, Ahmed Jawad Kabir και Hridon Ahmed. *An overview of convolutional neural network: Its architecture and applications*.
- [7] Robin M. Schmidt. *Recurrent Neural Networks (RNNs): A gentle Introduction and Overview*, 2019.
- [8] Yang Q. Pan, S. J. *A survey on Transfer Learning*.
- [9] Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun. *Deep residual learning for image recognition*, 2015.
- [10] Gao Huang, Zhuang Liu, Laurens Van Der Maaten και Kilian Q Weinberger. *Densely connected convolutional networks*. *Proceedings of the IEEE conference on computer vision and pattern recognition*, σελίδες 4700–4708, 2017.
- [11] Mingxing Tan και Quoc Le. *Efficientnet: Rethinking model scaling for convolutional neural networks*. *International conference on machine learning*, σελίδες 6105–6114. PMLR, 2019.
- [12] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov και Liang Chieh Chen. *Mobilenetv2: Inverted residuals and linear bottlenecks*. *Proceedings of the IEEE conference on computer vision and pattern recognition*, σελίδες 4510–4520, 2018.
- [13] Jie Hu, Li Shen και Gang Sun. *Squeeze-and-excitation networks*. *Proceedings of the IEEE conference on computer vision and pattern recognition*, σελίδες 7132–7141, 2018.
- [14] Alzheimer's Association. *2019 Alzheimer's disease facts and figures*. *Alzheimer's & Dementia*, 15(3):321–387.
- [15] *Having a MRI scan*, 2022.

- [16] Developer. *Role of mris in detecting alzheimer's*, 2022.
- [17] TR Stoub, M Bulgakova, S Leurgans, DA Bennett, D Fleischman, DA Turner και LdeToledo Morrell. *MRI predictors of risk of incident Alzheimer disease: a longitudinal study*. *Neurology*, 64(9):1520–1524, 2005.
- [18] R Duara, DA Loewenstein, E Potter, J Appel, MT Greig, R Urs, Q Shen, A Raj, B Small, W Barker και others. *Medial temporal lobe atrophy on MRI scans and the diagnosis of Alzheimer disease*. *Neurology*, 71(24):1986–1992, 2008.
- [19] Duaa AlSaeed και Samar Fouad Omar. *Brain MRI Analysis for Alzheimer's Disease Diagnosis Using CNN-Based Feature Extraction and Machine Learning*. *Sensors*, 22(8), 2022.
- [20] Rachna Jain, Nikita Jain, Akshay Aggarwal και D Jude Hemanth. *Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images*. *Cognitive Systems Research*, 57:147–159, 2019.
- [21] Jason Weller και Andrew Budson. *Current understanding of alzheimer's disease diagnosis and treatment*, 2018.
- [22] Siddharth Sharma, S Sharma και Anidhya Athaiya. *ACTIVATION FUNCTIONS IN NEURAL NETWORKS*. 2020.
- [23] Bin Ding, Huimin Qian και Jun Zhou. *Activation functions and their characteristics in deep neural networks*. *2018 Chinese Control And Decision Conference (CCDC)*, σελίδες 1836–1841, 2018.
- [24] Jason Brownlee. *A gentle introduction to the rectified linear unit (ReLU)*, 2020.
- [25] ROBERT HECHT-NIELSEN. *III.3 - Theory of the Backpropagation Neural Network**Based on "nonindent" by Robert Hecht-Nielsen, which appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593-611, June 1989. © 1989 IEEE. Neural Networks for Perception* Harry Wechsler, επιμελητής, σελίδες 65–93. Academic Press, 1992.
- [26] *Understanding Backpropagation Algorithm*.
- [27] Xiao Ling, Wenyuan Dai, Gui Rong Xue, Qiang Yang και Yong Yu. *Spectral Domain-Transfer Learning*. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08*, σελίδα 488–496, New York, NY, USA, 2008. Association for Computing Machinery.
- [28] Jing Gao, Wei Fan, Jing Jiang και Jiawei Han. *Knowledge Transfer via Multiple Model Local Structure Mapping*. *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '08*, σελίδα 283–291, New York, NY, USA, 2008. Association for Computing Machinery.

- [29] Xiaoxiao Shi, Wei Fan και Jiangtao Ren. *Actively Transfer Domain Knowledge. Machine Learning and Knowledge Discovery in Databases* Walter Daelemans, Bart Goethals και Katharina Morik, επιμελητές, σελίδες 342–357, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.
- [30] Maad M. Mijwel. *Artificial Neural Networks Advantages and Disadvantages*.
- [31] Dimitrios Kollias, Athanasios Tagaris, Andreas Stafylopatis, Stefanos Kollias και Georgios Tagaris. *Deep neural architectures for prediction in healthcare. Complex & Intelligent Systems*, 4(2):119–131, 2018.
- [32] Athanasios Tagaris, Dimitrios Kollias και Andreas Stafylopatis. *Assessment of Parkinson’s disease based on deep neural networks. International Conference on Engineering Applications of Neural Networks*, σελίδες 391–403. Springer, 2017.
- [33] Athanasios Tagaris, Dimitrios Kollias, Andreas Stafylopatis, Georgios Tagaris και Stefanos Kollias. *Machine learning for neurodegenerative disorder diagnosis—survey of practices and launch of benchmark dataset. International Journal on Artificial Intelligence Tools*, 27(03):1850011, 2018.
- [34] Ilianna Kollia, Andreas Georgios Stafylopatis και Stefanos Kollias. *Predicting Parkinson’s disease using latent information extracted from deep neural networks. 2019 International Joint Conference on Neural Networks (IJCNN)*, σελίδες 1–8. IEEE, 2019.
- [35] James Wingate, Ilianna Kollia, Luc Bidaut και Stefanos Kollias. *Unified deep learning approach for prediction of Parkinson’s disease. IET Image Processing*, 14(10):1980–1989, 2020.
- [36] Dimitrios Kollias, Anastasios Arsenos, Levon Soukissian και Stefanos Kollias. *Mia-cov19d: Covid-19 detection through 3-d chest ct image analysis. Proceedings of the IEEE/CVF International Conference on Computer Vision*, σελίδες 537–544, 2021.
- [37] Dimitrios Kollias, Anastasios Arsenos και Stefanos Kollias. *Ai-mia: Covid-19 detection & severity analysis through medical imaging. arXiv preprint arXiv:2206.04732*, 2022.
- [38] Anastasios Arsenos, Dimitrios Kollias και Stefanos Kollias. *A Large Imaging Database and Novel Deep Neural Architecture for Covid-19 Diagnosis. 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, σελίδες 1–5. IEEE, 2022.
- [39] Dimitrios Kollias, Miao Yu, Athanasios Tagaris, Georgios Leontidis, Andreas Stafylopatis και Stefanos Kollias. *Adaptation and contextualization of deep neural network models. 2017 IEEE symposium series on computational intelligence (SSCI)*, σελίδες 1–8. IEEE.
- [40] D Kollias, N Bouas, Y Vlaxos, V Brillakis, M Seferis, I Kollia, L Sukissian, J Wingate και S Kollias. *Deep Transparent Prediction through Latent Representation Analysis. arXiv preprint arXiv:2009.07044*, 2020.

- [41] Dimitris Kollias, Y Vlaxos, M Seferis, Ilianna Kollia, Levon Sukissian, James Wingate και S Kollias. *Transparent adaptation in deep medical image diagnosis. International Workshop on the Foundations of Trustworthy AI Integrating Learning, Optimization and Reasoning*, σελίδες 251–267. Springer, 2020.
- [42] Fabio De Sousa Ribeiro, Francesco Caliva, Mark Swainson, Kjartan Gudmundsson, Georgios Leontidis και Stefanos Kollias. *Deep bayesian self-training. Neural Computing and Applications*, 32(9):4275–4291, 2020.
- [43] Fabio De Sousa Ribeiro, Georgios Leontidis και Stefanos Kollias. *Capsule routing via variational bayes. Proceedings of the AAAI Conference on Artificial Intelligence*, τόμος 34, σελίδες 3749–3756, 2020.
- [44] Fabio De Sousa Ribeiro, Georgios Leontidis και Stefanos Kollias. *Introducing routing uncertainty in capsule networks. Advances in Neural Information Processing Systems*, 33:6490–6502, 2020.
- [45] Nikolaos Simou και Stefanos Kollias. *Fire: A fuzzy reasoning engine for imprecise knowledge*. Citeseer.
- [46] Francesco Caliva, Fabio Sousa De Ribeiro, Antonios Mylonakis, Christophe Demazière, Paolo Vinai, Georgios Leontidis και Stefanos Kollias. *A deep learning approach to anomaly detection in nuclear reactors. 2018 International joint conference on neural networks (IJCNN)*, σελίδες 1–8. IEEE, 2018.
- [47] Bashar Alhnaity, Stefanos Kollias, Georgios Leontidis, Shouyong Jiang, Bert Schamp και Simon Pearson. *An autoencoder wavelet based deep neural network with attention mechanism for multi-step prediction of plant growth. Information Sciences*, 560:35–50, 2021.
- [48] Bashar Alhnaity, Simon Pearson, Georgios Leontidis και Stefanos Kollias. *Using deep learning to predict plant growth and yield in greenhouse environments. International Symposium on Advanced Technologies and Management for Innovative Greenhouses: GreenSys2019 1296*, σελίδες 425–432, 2019.
- [49] Stefanos Kollias, Miao Yu, James Wingate, Aiden Durrant, Georgios Leontidis, Georgios Alexandridis, Andreas Stafylopatis, Antonios Mylonakis, Paolo Vinai και Christophe Demaziere. *Machine learning for analysis of real nuclear plant data in the frequency domain. Annals of Nuclear Energy*, 177:109293, 2022.
- [50] Andreas Psaroudakis και Dimitrios Kollias. *MixAugment & Mixup: Augmentation Methods for Facial Expression Recognition. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, σελίδες 2367–2375, 2022.
- [51] Dimitrios Kollias και Stefanos Zafeiriou. *Training deep neural networks with different datasets in-the-wild: The emotion recognition paradigm. 2018 International Joint Conference on Neural Networks (IJCNN)*, σελίδες 1–8. IEEE, 2018.

- [52] Dimitrios Kollias και Stefanos Zafeiriou. *Va-stargan: Continuous affect generation. International Conference on Advanced Concepts for Intelligent Vision Systems*, σελίδες 227–238. Springer, 2020.
- [53] G Caridakis, A Raouzaïou, K Karpouzis και S Kollias. *Synthesizing Gesture Expressivity Based on Real Sequences. Workshop Programme*, τόμος 10, σελίδα 19.
- [54] Phivos Mylonas, Evaggelos Spyrou, Yannis Avrithis και Stefanos Kollias. *Using visual context and region semantics for high-level concept detection. IEEE Transactions on Multimedia*, 11(2):229–243, 2009.
- [55] Stefanos Kollias και Dimitris Anastassiou. *A unified neural network approach to digital image halftoning. IEEE Transactions on signal processing*, 39(4):980–984, 1991.
- [56] Paraskevi Tzouveli, Andreas Schmidt, Michael Schneider, Antonis Symvonis και Stefanos Kollias. *Adaptive reading assistance for the inclusion of students with dyslexia: The AGENT-DYSL approach. 2008 Eighth IEEE International Conference on Advanced Learning Technologies*, σελίδες 167–171. IEEE, 2008.
- [57] *Alzheimer’s disease neuroimaging initiative.*
- [58] *Image classification nbsp;: nbsp; Tensorflow Core.*
- [59] Karen Simonyan και Andrew Zisserman. *Very deep convolutional networks for large-scale image recognition*, 2015.
- [60] Rupesh Kumar Srivastava, Klaus Greff και Jürgen Schmidhuber. *Highway Networks*, 2015.
- [61] Yann Lecun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard και L.D. Jackel. *Backpropagation applied to handwritten zip code recognition*, 1989.
- [62] Hassan Ali Khan, Wu Jue, Muhammad Mushtaq και Muhammad Umer Mushtaq. *Brain tumor classification in MRI image using convolutional neural network. Math. Biosci. Eng.*, 17(5):6203–6216, 2020.
- [63] Ahmet Çinar και Muhammed Yildirim. *Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture. Medical hypotheses*, 139:109684, 2020.
- [64] Lawrence V Fulton, Diane Dolezel, Jordan Harrop, Yan Yan και Christopher P Fulton. *Classification of Alzheimer’s disease with and without imagery using gradient boosted machines and ResNet-50. Brain sciences*, 9(9):212, 2019.
- [65] Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun. *Identity mappings in deep residual networks. European conference on computer vision*, σελίδες 630–645. Springer, 2016.
- [66] Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun. *Deep residual learning for image recognition. CVPR. 2016. arXiv preprint arXiv:1512.03385*, 2016.

- [67] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens και Zbigniew Wojna. *Rethinking the inception architecture for computer vision. Proceedings of the IEEE conference on computer vision and pattern recognition*, σελίδες 2818–2826, 2016.
- [68] Ram Deepak Gottapu και Cihan H Dagli. *DenseNet for anatomical brain segmentation. Procedia Computer Science*, 140:179–185, 2018.
- [69] Juan Ruiz, Mufti Mahmud, Md Modasshir, M Shamim Kaiser, for the Alzheimer’s Disease Neuroimaging Initiative και others. *3D DenseNet ensemble in 4-way classification of Alzheimer’s disease. International Conference on Brain Informatics*, σελίδες 85–96. Springer, 2020.
- [70] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto και Hartwig Adam. *Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861*, 2017.
- [71] Mingxing Tan, Bo Chen, Ruoming Pang, Vijay Vasudevan, Mark Sandler, Andrew Howard και Quoc V Le. *Mnasnet: Platform-aware neural architecture search for mobile. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, σελίδες 2820–2828, 2019.
- [72] *Papers with code - inverted residual block explained.*
- [73] *Papers with code - squeeze-and-excitation block explained.*
- [74] Dillip Ranjan Nayak, Neelamadhab Padhy, Pradeep Kumar Mallick, Mikhail Zymbler και Sachin Kumar. *Brain Tumor Classification Using Dense Efficient-Net. Axioms*, 11(1):34, 2022.
- [75] Serkan Savaş. *Detecting the stages of Alzheimer’s disease with pre-trained deep learning architectures. Arabian Journal for Science and Engineering*, 47(2):2201–2218, 2022.
- [76] Jiang Wang, Yi Yang, Junhua Mao, Zhiheng Huang, Chang Huang και Wei Xu. *Cnn-rnn: A unified framework for multi-label image classification. Proceedings of the IEEE conference on computer vision and pattern recognition*, σελίδες 2285–2294, 2016.
- [77] Md Zabirul Islam, Md Milon Islam και Amanullah Asraf. *A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. Informatics in medicine unlocked*, 20:100412, 2020.
- [78] Mohit Dua, Drishti Makhija, PYL Manasa και Prashant Mishra. *A CNN-RNN-LSTM based amalgamation for Alzheimer’s disease detection. Journal of Medical and Biological Engineering*, 40(5):688–706, 2020.
- [79] Ronald Mutegeki και Dong Seog Han. *A CNN-LSTM approach to human activity recognition. 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIC)*, σελίδες 362–366. IEEE, 2020.

[80] Riya Thakore. *Time distributed layer in Keras with example in Python*, 2020.