

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Πρόγραμμα Μεταπτυχιακών Σπουδών
Πληροφορικά Συστήματα και Υπηρεσίες
Κατεύθυνση: Μεγάλα Δεδομένα και Αναλυτική



Μεταπτυχιακή Διπλωματική Εργασία
**Συγκριτική μελέτη παραγωγικών ανταγωνιστικών δικτύων για την εξισορρόπηση
κλάσεων σε ιατρικές εικόνες**

Δημήτριος Παπατσαρουχάς

A.M. ME2032

Επιβλέπων: Ηλίας Μαγκλογιάννης, Καθηγητής

Πειραιάς Σεπτέμβριος 2023

Περίληψη

Η ανάπτυξη της τεχνολογίας και ειδικότερα της μηχανικής μάθησης φαίνεται να εξελίσσεται με γοργούς ρυθμούς τα τελευταία χρόνια, με τεχνικές βαθιάς μάθησης και υπολογιστικής όρασης, να χρησιμοποιούνται για την επίλυση διαφόρων προβλημάτων σε ποικίλους τομείς. Η συλλογή δεδομένων αποτελεί σημαντική πρόκληση σε πολλούς κλάδους, καθώς η εξασφάλιση αξιόπιστων και επαρκών δεδομένων απαιτεί προσπάθεια και χρόνο από ειδικούς του κάθε κλάδου. Ειδικά στο επιστημονικό πεδίο της ιατρικής, η συλλογή δεδομένων υγείων ατόμων μπορεί να είναι σχετικά ευκολότερη σε σύγκριση με τη συλλογή δεδομένων από ασθενείς, ιδιαίτερα όταν πρόκειται για σπάνιες ασθένειες. Αυτό δημιουργεί ένα πρόβλημα ανισορροπίας στο σύνολο δεδομένων, κάτι το οποίο κάνει δύσκολη την αξιοποίηση των τεχνικών μηχανικής μάθησης. Για την αντιμετώπιση του προβλήματος της ανισορροπίας κλάσεων έχουν προταθεί διάφορες τεχνικές όπως ο εμπλουτισμός δεδομένων με τεχνητά δείγματα. Η παρούσα διπλωματική εργασία, προτείνει την αντιμετώπιση του προαναφερθέντος προβλήματος, με τη χρήση Παραγωγικών Ανταγωνιστικών Δικτύων (GAN). Τα GANs είναι μοντέλα μηχανικής μάθησης τα οποία παράγουν ρεαλιστικές τεχνητές εικόνες. Αποτελούνται από 2 νευρωνικά δίκτυα, τον Γεννήτορα (Generator) και τον Διευκρινιστή (Discriminator), με το πρώτο να προσπαθεί να παράξει τεχνητές εικόνες όμοιες με τις πραγματικές και το δεύτερο να επιδιώκει να διαχωρίσει τις τεχνητές από τις πραγματικές εικόνες. Η εκπαίδευσή τους έχει ως στόχο τον ανταγωνισμό των δύο δικτύων με τη συνεχή βελτίωσή τους, μέχρι το σημείο όπου οι τεχνητές εικόνες να φαίνονται όμοιες με τις πραγματικές.

Η μελέτη εστιάζει στο σύνολο δεδομένων OASIS-3, το οποίο περιλαμβάνει εικόνες Μαγνητικής Απεικόνισης (MRI) εγκεφάλων από υγιή άτομα και ασθενείς με Αλτσχάιμερ. Η πειραματική διαδικασία αποτελείται αρχικά από τη λήψη και την επεξεργασία των δεδομένων. Έπειτα πραγματοποιείται η δοκιμή διαφόρων αρχιτεκτονικών GAN για τη δημιουργία τεχνητών εικόνων ασθενών, με στόχο την επιλογή των καταλληλότερων αρχιτεκτονικών. Στη συνέχεια δημιουργούνται τεχνητές εικόνες από τις επιλεγμένες αρχιτεκτονικές και προστίθενται στο αρχικό σύνολο δεδομένων, δοκιμάζοντας 5 διαφορετικές περιπτώσεις που περιλαμβάνουν διαφορετικά ποσοστά εισαγωγής τεχνητών εικόνων. Τέλος πραγματοποιείται η εκπαίδευση ενός ταξινομητή ResNet18 με

τη χρήση των διαφορετικών εμπλουτισμένων συνόλων και συγκρίνονται τα αποτελέσματα από το αρχικό σύνολο δεδομένων.

Όπως διαπιστώθηκε μετά από εκτενή μελέτη και ανάλυση των αποτελεσμάτων της έρευνας, η προσθήκη τεχνητών εικόνων φαίνεται να επηρεάζει σε μικρό βαθμό την απόδοση του ταξινομητή, παρατηρώντας ελάχιστη βελτίωση των μετρικών με την προσθήκη μικρού ποσοστού τεχνητών εικόνων (10%-20%) στην κλάση των ασθενών AD. Σε μεγαλύτερα ποσοστά τεχνητών εικόνων, παρατηρείται η αδυναμία του ταξινομητή να εντοπίσει τις περιπτώσεις της μειονοτικής κλάσης, ενώ παράλληλα οι μετρικές εμφανίζουν σταδιακή επιδείνωση.

Abstract

The development of technology, especially machine learning, seems to be advancing rapidly in recent years, with techniques such as deep learning and computer vision being used to solve various problems across diverse fields. Data collection poses a significant challenge in many sectors, as obtaining reliable and sufficient data requires effort and time from domain experts. Specifically, in the scientific field of medicine, collecting health-related data from negative samples might be relatively easier compared to collecting data from positives, especially when dealing with rare diseases. However, this creates an imbalance issue in the dataset, which complicates the utilization of machine learning techniques.

To address the problem of imbalance, various techniques like data augmentation have been proposed. This thesis suggests tackling the issue using Generative Adversarial Networks (GANs). GANs are machine learning models that generate realistic artificial images. They consist of two neural networks, the Generator, and the Discriminator. The Generator aims to produce artificial images like real ones, while the Discriminator tries to distinguish between artificial and real images. Their training involves a competition between the two networks, iteratively improving them until the artificial images closely resemble real ones.

The study focuses on the OASIS-3 dataset, which includes Magnetic Resonance Imaging (MRI) brain images from healthy individuals and Alzheimer's patients. The experimental process begins with data acquisition and processing. Various GAN architectures are then tested to create artificial patient images, with the goal of selecting the most suitable architectures. Subsequently, artificial images generated from the selected architectures are added to the original dataset, testing five different scenarios involving varying percentages of artificial image insertion. Finally, a ResNet18 classifier is trained using the enriched datasets, and the results are compared to those from the original dataset.

As observed after an extensive study and analysis of the research results, the addition of artificial images appears to have a minor impact on the classifier's performance, showing minimal improvement in metrics with the addition of a small percentage of

artificial images (10%-20%) to the class of AD patients. At higher percentages of artificial images, the classifier fails to identify instances of the minority class, while concurrently, the metrics exhibit a gradual deterioration.

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον κ. Μαγκλογιάννη Ηλία καθηγητή του τμήματος Ψηφιακών Συστημάτων του πανεπιστημίου Πειραιώς, για την εμπιστοσύνη που μου έδειξε δίνοντας μου την ευκαιρία να ασχοληθώ με ένα τόσο ενδιαφέρον και επίκαιρο θέμα. Έπειτα θα ήθελα να ευχαριστήσω το κ. Καλλιπολίτη Αθανάσιο για την πολύτιμη βοήθεια και καθοδήγηση καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής εργασίας. Τέλος, ευχαριστώ την οικογένεια μου για την στήριξή τους όλο αυτό το διάστημα.

Περιεχόμενα

Λίστα σχημάτων.....	9
Λίστα πινάκων.....	11
1. Εισαγωγή	13
2. Τεχνολογικό υπόβαθρο	17
2.1. Νευρωνικά Δίκτυα	18
2.2. Κατηγοριοποίηση (Classification)	21
2.2.1. Resnet.....	22
2.3. Παραγωγικά ανταγωνιστικά δίκτυα (Generative Adversarial Networks).....	22
2.3.1. DCGAN	25
2.3.2. LSGAN	27
2.3.3. BEGAN.....	28
2.3.4. WGAN & WGAN-GP	29
2.3.5. CGAN	31
2.3.6. CWGAN-GP	33
3. Μεθοδολογία	34
3.1. Ροή διαδικασιών	34
3.2. Παραγωγικά Ανταγωνιστικά Δίκτυα – Αρχιτεκτονικές και λεπτομέρειες υλοποίησης.....	36
3.2.1. WGAN (Gradient Penalty).....	36
3.2.2. CWGAN (Gradient Penalty).....	41
3.2.3. LSGAN	43
3.2.4. CGAN	45
3.2.5. BEGAN.....	46

4.	Πειραματικά αποτελέσματα	49
4.1.	Δεδομένα	50
4.2.	Δομή και τρόπος αξιολόγησης ταξινομητών	56
4.3.	Αποτελέσματα τεχνητών εικόνων	57
4.3.1.	WGANGP 1, με σύνολο δεδομένων ΣΔ1	58
4.3.2.	WGANGP 1, με σύνολο δεδομένων ΣΔ2	60
4.3.3.	WGANGP 1, με σύνολο δεδομένων ΣΔ3	61
4.3.4.	WGANGP 2, με σύνολο δεδομένων ΣΔ1	63
4.3.5.	WGANGP 2, με σύνολο δεδομένων ΣΔ2	64
4.3.6.	CWGANGP 1, με σύνολο δεδομένων ΣΔ3	66
4.3.7.	LSGAN, με σύνολο δεδομένων ΣΔ3	67
4.3.8.	CGAN, με σύνολο δεδομένων ΣΔ3	68
4.3.9.	BEGAN, με σύνολο δεδομένων ΣΔ1	69
4.4.	Inception Score	70
4.5.	Αποτελέσματα ταξινομητή	71
4.5.1.	Σύνολο Δεδομένων 1	74
4.5.2.	Σύνολο Δεδομένων 2	75
4.5.3.	Σύνολο Δεδομένων 3	75
5.	Συζήτηση	77
6.	Συμπεράσματα και μελλοντική εργασία	86
7.	Βιβλιογραφία	90

Λίστα σχημάτων

2-1 Residual Block [24]	22
2-2 Δομή ενός δικτύου GAN.....	24
2-3 Δομή Discriminator και Generator στο DCGAN	25
2-4 Δομή αρχιτεκτονικής BEGAN [33]	29
2-5 Δομή αρχιτεκτονικής cGAN [37]	31
3-1 Ροή εργασιών μεθοδολογίας.....	35
4-1 Δείγμα κεντρικών τομών ΣΔ1, πάνω σειρά κλάση NC κάτω σειρά κλάση AD	53
4-2 Δείγμα κεντρικών τομών του ΣΔ3, πάνω σειρά υγιής, κάτω σειρά ασθενής.....	54
4-3 Τεχνητές εικόνες WGANGP 1 στο ΣΔ1 εποχή 240.....	59
4-4 Τεχνητές εικόνες WGANGP 1 στο ΣΔ1 εποχές 10, 50 και 150	59
4-5 WGANGP, ΣΔ1 Discriminator Loss (αριστερά) και Generator Loss (δεξιά).....	60
4-6 Τεχνητές εικόνες WGANGP 1 στο ΣΔ2 εποχή 11225	60
4-7 Τεχνητές εικόνες WGANGP 1 ΣΔ2 εποχές 100, 1000 και 5000	61
4-8 Discriminator Loss WGAN 1 ΣΔ2, ως εποχή 400 (αριστερά) και ως εποχή 11225 (δεξιά).....	61
4-9 Πρώτη δοκιμή αρχιτεκτονικής WGANGP 1 στο ΣΔ3 χωρίς διαχωρισμό των τομών, εποχές 200, 1000, 1900	62
4-10 Πρώτη δοκιμή αρχιτεκτονικής WGANGP 1 στο ΣΔ3 χωρίς διαχωρισμό των τομών, απώλειες Discriminator (αριστερά) και Generator (δεξιά).....	62
4-11 Τεχνητές εικόνες από τα ξεχωριστά δίκτυα των τομών, axial, coronal και sagittal αντίστοιχα.....	63
4-12 Απώλειες Discriminator από τα ξεχωριστά δίκτυα των τομών, axial, coronal και sagittal.....	63

4-13 Τεχνητές εικόνες από την αρχιτεκτονική WGANGP 2 στο ΣΔ1, εποχές 10, 50, 170	64
4-14 Απώλειες Discriminator και Generator WGANGP 2 στο ΣΔ1	64
4-15 Τεχνητές εικόνες από την αρχιτεκτονική WGANGP 2 στο ΣΔ2, εποχές 1000, 2000, 4000	65
4-16 Τεχνητές εικόνες από την αρχιτεκτονική WGANGP 2 στο ΣΔ2, εποχές 6250 και 7660	65
4-17 Απώλεια Discriminator WGANGP 2 στο ΣΔ2.....	65
4-18 Απώλεια Generator WGANGP 2 στο ΣΔ2.....	66
4-19 Τεχνητές εικόνες από την αρχιτεκτονική CWGANGP 1 στο ΣΔ3, εποχές 100, 500, 1100	67
4-20 Απώλειες Discriminator και Generator CWGANGP 1 στο ΣΔ3	67
4-21 Τεχνητές εικόνες από την αρχιτεκτονική LSGAN στο ΣΔ3, εποχές 1000, 2500, 5000	68
4-22 Απώλειες Discriminator και Generator LSGAN στο ΣΔ3	68
4-23 Τεχνητές εικόνες από την αρχιτεκτονική cGAN στο ΣΔ3, εποχές 500, 1000, 2000	69
4-24 Απώλειες Discriminator και Generator cGAN στο ΣΔ3.....	69
4-25 Τεχνητές εικόνες από την αρχιτεκτονική BEGAN στο ΣΔ1, εποχές 500, 1000, 2000	70
4-26 Απώλειες Discriminator και Generator BEGAN στο ΣΔ1.....	70

Λίστα πινάκων

3-1 Αρχιτεκτονική WGANGP 1, Generator	37
3-2 Αρχιτεκτονική WGANGP 1, Discriminator	37
3-3 Αρχιτεκτονική WGANGP 2, Generator	39
3-4 Αρχιτεκτονική WGANGP 2, Discriminator	39
3-5 Αρχιτεκτονική CWGAN, δομή Generator.....	42
3-6 Αρχιτεκτονική CWGAN, δομή Discriminator.....	42
3-7 Αρχιτεκτονική LSGAN, δομή Generator	44
3-8 Αρχιτεκτονική LSGAN, δομή Discriminator	44
3-9 Αρχιτεκτονική CGAN, Generator.....	45
3-10 Αρχιτεκτονική CGAN, Discriminator	45
3-11 Αρχιτεκτονική BEGAN, δομή Generator	47
3-12 Αρχιτεκτονική BEGAN, δομή Discriminator	47
4-1 Επεξήγηση κλίμακας CDR.....	51
4-2 Στατιστικά του Συνόλου Δεδομένων 1 ΣΔ1 με μέθοδο MEK1.....	52
4-3 Στατιστικά του Συνόλου Δεδομένων 2 ΣΔ2 με μέθοδο MEK2.....	53
4-4 Στατιστικά του Συνόλου Δεδομένων 3 ΣΔ3 με μέθοδο MEK1.....	54
4-5 Στατιστικά διαχωρισμένου (splitted) συνόλου δεδομένων ΣΔ1.....	55
4-6 Στατιστικά διαχωρισμένου (splitted) συνόλου δεδομένων ΣΔ2.....	55
4-7 Στατιστικά διαχωρισμένου (splitted) συνόλου δεδομένων ΣΔ3.....	55
4-20 Υπερπαράμετροι πειράματος 4.4.1	59
4-21 Παράμετροι πειράματος.....	62
4-22 WGANGP 2 ΣΔ1, υπερπαράμετροι εκπαίδευσης	63
4-23 CWGANGP 1 ΣΔ3, υπερπαράμετροι εκπαίδευσης.....	66

4-24 Inception Score ανά σύνολο δεδομένων και ανά αρχιτεκτονική	71
4-25 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ1 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP1.....	74
4-26 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ1 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP2.....	75
4-27 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ2 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP1.....	75
4-28 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ2 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP2.....	75
4-29 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ3 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP1.....	75
4-30 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ3 με τεχνητές εικόνες από την αρχιτεκτονική CWGANGP1	76
4-31 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ3 με τεχνητές εικόνες από την αρχιτεκτονική LSGAN.....	76

1. Εισαγωγή

Η τεχνολογία και ειδικότερα η μηχανική μάθηση που έχει αναπτυχθεί ιδιαίτερα τα τελευταία χρόνια, δίνουν τη δυνατότητα δημιουργίας εργαλείων που επηρεάζουν θετικά πάρα πολλούς τομείς, με έναν από τους πιο σημαντικούς να είναι η ιατρική. Μερικά παραδείγματα περιλαμβάνουν προβλήματα ταξινόμησης όπως τη διάγνωση ασθενειών [1], την τμηματοποίηση εικόνων [2], την πρόβλεψη της επιβίωσης μετά από μεγάλο ρίσκου επεμβάσεις με σχήματα παλινδρόμησης [3] και την ανάλυση χρονοσειρών [4]. Ένα σημαντικό πρόβλημα στα ιατρικά δεδομένα και ειδικότερα σε ασθένειες σπάνιες, είναι η δυσκολία συλλογής δεδομένων ασθενών. Όπως είναι λογικό η συλλογή πληροφοριών από υγιή άτομα είναι πιο εύκολη, γεγονός που κατά τη δημιουργία ενός συνόλου δεδομένων, προκαλεί την ύπαρξη άνισου αριθμού δειγμάτων μεταξύ των κλάσεων (πχ ασθενείς, υγιείς). Για την επίλυση του προβλήματος αυτού έχουν προταθεί διάφορες μεθοδολογίες όπως η τυχαία υπο-δειγματοληψία (Random Under-Sampling) [5], σύμφωνα με την οποία πραγματοποιείται η τυχαία αφαίρεση δειγμάτων από την κλάση που υπερτερεί σε αριθμό δειγμάτων. Άλλος τρόπος είναι η μέθοδος συνόλων (Ensemble methods) [6] όπου χρησιμοποιούνται πολλοί ταξινομητές και εκπαιδεύονται σε διαφορετικό υποσύνολο και στο τέλος συνδυάζονται οι προβλέψεις τους. Αντίθετα από την υπο-δειγματοληψία υπάρχει και η υπερ-δειγματοληψία (Over-sampling), που περιλαμβάνει πολλές μεθοδολογίες όπως η μέθοδος SMOTE (Synthetic Minority Over-sampling Technique) και τα Παραγωγικά Ανταγωνιστικά Δίκτυα (ΠΑΝ) που πραγματοποιούν τη δημιουργία νέων συνθετικών δεδομένων.

Σκοπός της παρούσας εργασίας είναι η μελέτη της επίλυσης του προβλήματος των άνισων δειγμάτων μεταξύ κλάσεων σε ένα ιατρικό σύνολο δεδομένων (imbalance dataset). Το σύνολο που επιλέχθηκε ανήκει στον ιατρικό τομέα, συγκεκριμένα αφορά τη νόσο Αλτσχάιμερ και περιλαμβάνει αρχεία μαγνητικών τομογραφιών εγκεφάλων, υγιών αλλά και ασθενών ατόμων.

Η νόσος Alzheimer αποτελεί μια διαταραχή στον εγκέφαλο που εμφανίζεται σε ηλικιωμένους ανθρώπους. Για πρώτη φορά δόθηκε ορισμός από τον Alois Alzheimer το 1906 με κύρια συμπτώματα την απώλεια μνήμης κι αποπροσανατολισμό. Η διάγνωση της δεν είναι εύκολη. Ωστόσο, έχει γίνει προσπάθεια εύρεσης θεραπείας, ώστε να

αντιμετωπιστούν διαταραχές γλωσσικές και προβλήματα που αφορούν τη συμπεριφορά. Ο ασθενής που πάσχει από τη νόσο Alzheimer εμφανίζει κάποια συμπτώματα που οδηγούν στη διάγνωση της. Προσβάλλει τη μνήμη, τη γλώσσα, την οπτικοχωρική αντίληψη και τις ανώτερες εκτελεστικές λειτουργίες που βρίσκονται στο προμετωπιαίο φλοιό του εγκεφάλου. Ειδικότερα, διαταράσσονται η εργασιακή μνήμη/ βραχύχρονη, η προσοχή, η συγκέντρωση, η γνωστική ευελιξία, εμφανίζεται δυσκολία στις αυτόματες απαντήσεις, σκέψη, κρίση και συμπεριφορά. Πέρα από το γνωστικό επίπεδο επηρεάζεται η προσωπικότητα του πάσχοντα, παρουσιάζονται ψυχώσεις, διαταραχές στη διάθεση και ανωμαλίες στον ύπνο. Καθίσταται απαραίτητο να σημειωθεί ότι η πορεία αυτής της ασθένειας είναι αργή και μη αναστρέψιμη. Στην αρχή χάνεται σιγά- σιγά η μνήμη και ακολουθεί μια γλωσσική διαταραχή, η αφασία, ύστερα ακολουθεί, η απραξία κι η αγνωσία. Ωστόσο, κάποιοι άνθρωποι που πάσχουν από τη νόσο Alzheimer ενδέχεται στα πρώτα στάδια να παρουσιάζουν ευερεθιστότητα και αλλαγές στη συμπεριφορά. Κατά την εξέλιξη της νόσου παρουσιάζεται επίσης διαταραχή στην κίνηση.

Τα συμπτώματα αυτά αξιολογούνται μέσω μιας σειράς ενεργειών. Αρχικά, πρέπει να ληφθεί το ιστορικό του ασθενούς, συμπεριλαμβανομένων των ιατρικών, νευρολογικών, νευροψυχιατρικών και οικογενειακών δεδομένων. Στη συνέχεια, πραγματοποιούνται εργαστηριακές εξετάσεις, συμπεριλαμβανομένων αιματολογικών εξετάσεων, εξετάσεων θυροειδούς, εξέτασης επιπέδων βιταμίνης B12 και φυλλικού οξέος.

Επίσης, άλλο ένα μέσο για να διαγνωσθεί η νόσος Alzheimer είναι η μαγνητική τομογραφία (MRI). Η εξέταση αυτή κάνει χρήση ενός ισχυρού μαγνητικού πεδίου, όπου με παλμούς από ραδιοσυχνότητες και έναν υπολογιστή παρασκευάζονται εικόνες οργάνων, μαλακών ιστών, οστών και άλλων δομών εσωτερικά του σώματος. Η μαγνητική τομογραφία δύναται να προχωρήσει στην ανίχνευση ανωμαλιών του εγκεφάλου, ανωμαλίες σχετικές ακόμα και με ήπια γνωστική εξασθένηση (MCI) και δίνει τη δυνατότητα πρόβλεψης ανάπτυξης νόσου Alzheimer σε ασθενείς με γνωστική εξασθένηση (MCI). Βέβαια, κρίνεται απαραίτητο να τονισθεί ότι στα αρχικά στάδια της νόσου ενδέχεται η μαγνητική τομογραφία να φαίνεται φυσιολογική, ενώ αργότερα να φανεί η μείωση/ συρρίκνωση του μεγέθους διαφορετικών περιοχών στον εγκέφαλο, κάτι που σηματοδοτεί την ύπαρξη της νόσου.

Η μηχανική μάθηση με τους αλγορίθμους ταξινόμησης έχει αποδειχθεί σημαντικό εργαλείο σε παρόμοια προβλήματα με σύνολα εικόνων, ειδικότερα με μοντέλα βαθιάς μάθησης να μπορούν να αναγνωρίσουν μοτίβα, οδηγώντας σε πολύ υψηλά επίπεδα ακρίβειας στην ταξινόμηση. Σημαντικό ζήτημα σε τέτοιου ίδιους ιατρικά προβλήματα που αφορούν ιδιαίτερες ασθένειες, είναι η δύσκολη συλλογή δεδομένων από ασθενείς, ενώ η συλλογή δεδομένων από υγιής ανθρώπους είναι πολλές φορές πιο εύκολη. Αυτή η ανισορροπία στο δείγμα δυσκολεύει την εκπαίδευση των μοντέλων, διότι δεν είναι εύκολο να γενικευθεί ως προς την κλάση των ασθενών.

Η δομή της εργασίας αποτελείται από το τεχνολογικό υπόβαθρο, όπου περιγράφονται βασικές έννοιες των τεχνικών που θα χρησιμοποιηθούν και σχολιάζονται σχετικές μελέτες, που διερευνούν την επίλυση ιατρικών ζητημάτων με παρόμοια μεθοδολογία. Στην συνέχεια περιγράφεται αναλυτικά η μεθοδολογία που ακολουθήθηκε, καθώς και η δομή των αρχιτεκτονικών, των μοντέλων GAN που υλοποιήθηκαν και του ταξινομητή. Η επόμενη ενότητα είναι τα πειραματικά αποτελέσματα, η οποία περιλαμβάνει μια σύνοψη του συνόλου δεδομένων και την παρουσίαση των αποτελεσμάτων των πειραμάτων. Ακολούθως πραγματοποιείται η συζήτηση και σχολιασμός των αποτελεσμάτων. Τέλος παρατίθενται τα συμπεράσματα και οι προτάσεις για μελλοντική έρευνα.

Στόχος της παρούσας ερευνητικής εργασίας αποτελεί η σύγκριση διαφορετικών αρχιτεκτονικών GAN με σκοπό την επίλυση των προβλημάτων που προκύπτουν κατά την ταξινόμηση (classification) ενός ιατρικού συνόλου δεδομένων, το οποίο αποτελείται από διαφορετικό αριθμό δειγμάτων μεταξύ των κλάσεων (imbalance dataset). Η λύση που θα ακολουθηθεί κάνει χρήση των Generative Adversarial Networks, τα οποία χρησιμοποιούνται για την παραγωγή τεχνητών εικόνων από τη μειονοτική κλάση, οι οποίες αξιοποιούνται για την εξισορρόπηση του συνόλου και αξιολογούνται βάση της απόδοσής τους.

Η αξιολόγηση θα γίνει αρχικά οπτικά αλλά και με το κριτήριο inception Score. Έπειτα θα μελετηθεί η επιρροή που ασκεί η προσθήκη τεχνητών εικόνων στο σύνολο δεδομένων, στα αποτελέσματα του ταξινομητή. Η αναζήτηση της κατάλληλης αρχιτεκτονικής GAN για ένα συγκεκριμένο σύνολο δεδομένων αποτελεί μια χρονοβόρα και σύνθετη διαδικασία,

αλλά η προσπάθεια αυτή, ανεξάρτητα από το τελικό αποτέλεσμα προσφέρει σημαντικές πληροφορίες, που μπορούν να χρησιμοποιηθούν εποικοδομητικά σε μελλοντικές μελέτες.

2. Τεχνολογικό υπόβαθρο

Η ενότητα αυτή αποτελεί τη βάση για την κατανόηση της διπλωματικής εργασίας, καθώς επιδιώκει να παράσχει μια ανασκόπηση του τεχνολογικού πεδίου στο οποίο εστιάζει η παρούσα μελέτη. Στόχος είναι να παρουσιαστεί η βασική θεωρία και να αναδειχθεί η σημαντική βιβλιογραφία, καθώς με την επισκόπηση των μελετών που είναι άμεσα συσχετισμένες με το αντικείμενο της παρούσας εργασίας, δίνεται η δυνατότητα καλύτερης κατανόησης του θέματος και των σημαντικών σημείων που καθορίζουν την εξέλιξη της έρευνας. Μέσα από την ανασκόπηση προηγούμενων εργασιών στον τομέα, στοχεύεται η θεμελίωση μιας σταθερής βάσης και ο εντοπισμός σύγχρονων μεθόδων. Επιπλέον, μέσω της κριτικής αξιολόγησης των δυνατοτήτων, αδυναμιών και περιορισμών των προηγούμενων εργασιών που έχουν πραγματοποιηθεί, καθίσταται δυνατόν να εντοπίσουμε ευκαιρίες για την αντιμετώπιση των ελλείψεων και την εξερεύνηση νέων προοπτικών για μελλοντική έρευνα.

Έχουν πραγματοποιηθεί πολλές και αξιόλογες μελέτες πάνω στο ζήτημα της ανισορροπίας των κλάσεων, ειδικότερα όταν το πρόβλημα προς επίλυση ανήκει στον ιατρικό τομέα. Οι προτάσεις περιλαμβάνουν τόσο κλασσικές μεθόδους εμπλουτισμού δεδομένων όσο και νέες πρακτικές όπως η παραγωγή τεχνητών δεδομένων με τη χρήση παραγωγικών μοντέλων [7] [8]. Εξίσου αξιόλογες έρευνες έχουν πραγματοποιηθεί και για τη διάγνωση της νόσου Αλτσχάιμερ μέσω ιατρικών απεικονίσεων, όπως η μαγνητική τομογραφία, με τη χρήση μοντέλων κατηγοριοποίησης. [8] [9] [10]. Οι Yagis et al [9] διερευνούν μία λύση Συνελικτικού Νευρωνικού Δικτύου τριών διαστάσεων για τη διάγνωση Αλτσχάιμερ με χρήση μαγνητικών τομογραφιών. Έχουν πραγματοποιήσει δοκιμές σε δύο διαφορετικά σύνολα δεδομένων και καταλήγουν στο συμπέρασμα ότι οι τρισδιάστατη μέθοδος που υλοποιήθηκε έχει σημαντικά καλύτερα αποτελέσματα από τις μεθόδους δύο διαστάσεων. Οι Khagi et al [10] συγκρίνουν 3 προεκπαιδευμένα δίκτυα τα οποία έγιναν finetune σε σύνολο δεδομένων με μαγνητικές εγκεφάλων, έναντι ενός μοντέλου που εκπαιδεύτηκε για πρώτη φορά στο ίδιο σύνολο δεδομένων. Οι συγγραφείς καταλήγουν σε παρατηρήσεις όπως ότι η εξειδίκευση προεκπαιδευμένων μοντέλων σε μη ιατρικές εικόνες μπορούν να χρησιμοποιηθούν με καλά αποτελέσματα, επίσης παρατηρήθηκε ότι το μη προεκπαιδευμένο μοντέλο είχε καλύτερη απόδοση.

2.1. Νευρωνικά Δίκτυα

Ο ανθρώπινος εγκέφαλος είναι ένα βιολογικό νευρωνικό δίκτυο, το οποίο αποτελείται από 200 δισεκατομμύρια νευρώνες. Ο κάθε νευρώνας λαμβάνει εισόδους σε μορφή σημάτων, τα επεξεργάζεται και εξάγει σήματα. Ένας νευρώνας αποτελείται από δενδρίτες, το σώμα και τον άξονα. Υπεύθυνοι για τη λήψη όλων των εισερχόμενων σημάτων είναι οι δενδρίτες, το σώμα αναλαμβάνει το άθροισμα όλων των σημάτων και αν το αποτέλεσμα φτάσει το καθορισμένο κατώφλι, τότε το σήμα αποστέλλεται σε άλλους νευρώνες μέσω του άξονα. Μια άλλη δομή είναι οι συνάψεις που επιτρέπουν στους νευρώνες να περνούν τα σήματα σε άλλους γειτονικούς νευρώνες. Πολλοί διασυνδεδεμένοι νευρώνες αποτελούν ένα νευρωνικό δίκτυο.

Τα τεχνητά νευρωνικά δίκτυα προσπαθούν να προσομοιώσουν τη λειτουργία του ανθρώπινου εγκέφαλου, που περιγράφηκε παραπάνω. Ειδικότερα ένα τεχνητό νευρωνικό δίκτυο αποτελείται από πολλούς νευρώνες, οι οποίοι δέχονται τα σήματα και υπολογίζουν μία ποσότητα λαμβάνοντας υπόψιν και τις τιμές των βαρών (weights), ώστε να μεταβιβάσουν το σήμα στον επόμενο νευρώνα. Για τον υπολογισμό της ποσότητας και τη μετάδοση του σήματος χρησιμοποιείται μια συνάρτηση που ονομάζεται συνάρτηση ενεργοποίησης (activation function). Πιο συγκεκριμένα, ένα νευρωνικό δίκτυο αποτελείται από ένα επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα και ένα επίπεδο εξόδου. Ο αριθμός των κρυφών επιπέδων όπως και ο αριθμός των νευρώνων, εξαρτώνται από την πολυπλοκότητα του συστήματος και του προβλήματος προς επίλυση. Οι νευρώνες συνδέονται μεταξύ τους και οι συνδέσεις έχουν ένα αριθμητικό βάρος. Από την είσοδο της πληροφορίας ως την έξοδο, για τη μετάδοση της από το ένα επίπεδο στο άλλο πραγματοποιούνται υπολογισμοί χρησιμοποιώντας τα βάρη κάθε σύνδεσης και μεταβιβάζεται η πληροφορία στο επόμενο επίπεδο με τη χρήση μιας συνάρτησης ενεργοποίησης, δίνοντας στο τέλος μια τιμή εξόδου.

Η εκπαίδευση των νευρωνικών δικτύων έχει σκοπό την προσαρμογή των βαρών ώστε να επιτευχθεί ο στόχος της εκπαίδευσης. Υπάρχουν τέσσερις τύποι εκπαίδευσης, η επιβλεπόμενη (supervised) και η μη επιβλεπόμενη (unsupervised), η ημι-επιβλεπόμενη (semi-supervised ή weakly-supervised) και η μάθηση με ενίσχυση (reinforcement learning). Στην περίπτωση της επιβλεπόμενης μάθησης, η αλήθεια βάσης (ground truth)

παρέχεται μαζί με την πληροφορία εισόδου, με στόχο να προσαρμοστούν τα βάρη, ώστε να ελαχιστοποιηθεί το σφάλμα των τιμών εξόδου, βάση του δοθέντος επιθυμητού αποτελέσματος. Μερικά παραδείγματα χρήσης επιβλεπόμενων νευρωνικών δικτύων είναι:

- Πρόβλεψη χρονοσειρών, όπως παρουσιάζεται από τους Sean J. Tayler et al [11], οι οποίοι υλοποίησαν μια βιβλιοθήκη με το όνομα Prophet η οποία δίνει τη δυνατότητα πρόβλεψης χρονοσειρών με ερμηνεύσιμες παραμέτρους, παρέχοντας στον αναλυτή τις απαραίτητες πληροφορίες.
- Αναγνώριση προτύπων, μια τεχνική που χρησιμοποιούν οι ερευνητές Ozaki Y. et al [12] για τον εντοπισμό 5 διαφόρων ειδών καρκίνου σε αρχικά στάδια. Προτείνουν μια λύση δύο σταδίων, στο πρώτο στάδιο ο ταξινομητής επισημαίνει σημεία ιστού ως υγιή ή ασθενή, ενώ στο δεύτερο στάδιο γίνεται η χρήση ενός SVM classifier για την επιβεβαίωση των αποτελεσμάτων.
- Πρόβλεψη ακολουθιών, μια τεχνική που χρησιμοποιείται στην μετάφραση φυσικής γλώσσας, όπως παρουσιάζουν στην έρευνα τους οι Sutskever et al [13], όπου με τη χρήση LSTM καταφέρνουν με επιτυχία την μετάφραση από Αγγλικά σε Γαλλικά.
- Ανάλυση παλινδρόμησης, όπως παρουσιάζεται και στην έρευνα [14] στην οποία οι ερευνητές δημιουργούν ένα Τεχνητό Νευρωνικό Δίκτυο και το αξιολογούν με τη χρήση του Auto MPG συνόλου δεδομένων, καταλήγοντας ότι υπερτερεί σε απόδοση έναντι άλλων μεθόδων.

Από την άλλη στην περίπτωση της μη επιβλεπόμενης μάθησης, το μοντέλο δεν λαμβάνει πληροφορίες για το επιθυμητό αποτέλεσμα, αλλά επιχειρεί να επεξεργαστεί τα δεδομένα εισόδου, με αποτέλεσμα να «κατανοήσει» τα δεδομένα και τη δομή τους. Μερικά παραδείγματα μη επιβλεπόμενης μάθησης είναι τα εξής:

- Συσταδοποίηση, οι Jiao et al [15] υλοποίησαν έναν αλγόριθμο συσταδοποίησης με τη χρήση τεχνητών νευρωνικών δικτύων, αποδεικνύοντας ότι αποδίδουν καλύτερα έναντι παραδοσιακών τεχνικών σε 4 διαφορετικά UCI σύνολα δεδομένων [16].

- Ανίχνευση ανωμαλιών, οι Schlegl et al [17] προτείνουν ένα Παραγωγικό Ανταγωνιστικό Δίκτυο για την ανίχνευση ανωμαλιών AnoGAN σε ιατρικό σύνολο δεδομένων με εικόνες. Η προτεινόμενη λύση φαίνεται να επιτυγχάνει καλά αποτελέσματα και σε νέα σύνολα δεδομένων που δεν έχει προηγηθεί εκπαίδευση.
- Παραγωγικά μοντέλα όπως τα GAN [17]

Η περίπτωση της ημι-επιβλεπόμενης μάθησης αποτελεί μια προσέγγιση που συνδυάζει τα πλεονεκτήματα των δύο προαναφερθέντων μεθόδων και χρησιμοποιείται σε περιπτώσεις που το σύνολο δεδομένων αποτελείται από μερικώς επισημασμένα (labeled) δεδομένα. Το πρώτο βήμα σε αυτό τον τύπο μάθησης είναι η εκπαίδευση ενός supervised αλγορίθμου χρησιμοποιώντας τα διαθέσιμα labeled δεδομένα και έπειτα η χρήση του εκπαιδευμένου μοντέλου για την επισήμανση των υπόλοιπων δεδομένων.

Τέλος το Reinforcement Learning RL είναι μια κατηγορία αλγορίθμων που εστιάζουν στην μάθηση από την αλληλεπίδραση με το «περιβάλλον». Στην ουσία είναι αυτοδύναμες λύσεις που εκπαιδεύονται για την επίτευξη ενός στόχου ακολουθώντας μια διαδικασία δοκιμών από τις οποίες λαμβάνουν θετικά ή αρνητικά σήματα. Σε κάθε δοκιμή πραγματοποιείται μία δράση, η οποία αν οδηγεί προς την επίτευξη του στόχου, η ανταμοιβή είναι θετική, υποδεικνύοντας στον αλγόριθμο ότι προχωράει προς την σωστή κατεύθυνση, ενώ στην περίπτωση που η δράση δεν έχει θετικό αποτέλεσμα, δεν υπάρχει ανταμοιβή αλλά τιμωρία, βελτιώνοντας έτσι τις επόμενες δράσεις. Ένα παράδειγμα εφαρμογής των αλγορίθμων είναι τα παιχνίδια, όπως παρουσιάζουν στη μελέτη τους οι Berner et al [18], τα μοντέλα RL έχουν σημαντικά υψηλότερη απόδοση από άλλες μεθόδους σε συστήματα παιχνιδιών. Επιπλέον επισημαίνουν ότι το σύστημα που αναπτύχθηκε για το παιχνίδι Dota 2, κατάφερε να κατακτήσει την πρώτη θέση σε διαγωνισμό esports. Επίσης RL μοντέλα χρησιμοποιούνται και σε εφαρμογές self-driving cars, οι Kendall et al [19] κάνουν χρήση deep RL μοντέλων για την οδήγηση αυτοκινήτου, χωρίς την ανάγκη προκαθορισμένων οδηγιών. Αναφέρουν ότι από πολύ νωρίς στην εκπαίδευση το μοντέλο μαθαίνει να ακολουθεί την λωρίδα του δρόμου, τονίζοντας ότι παρέχουν μια γενική και εύκολη λύση για την ανταμοιβή του δικτύου κατά την εκπαίδευση. Τέλος ένα άλλο παράδειγμα εφαρμογής RL είναι η επεξεργασία φυσικής γλώσσας, με

τους Li et al [20], να μελετούν μία λύση παραγωγής διαλόγου. Προτείνουν την ανταμοιβή του δικτύου βάση 3 στοιχείων και της ύπαρξής τους στις παραγόμενες ακολουθίες, πληροφοριακότητα (informativity), συνοχή (coherence) και ευκολία απάντησης (ease of answering).

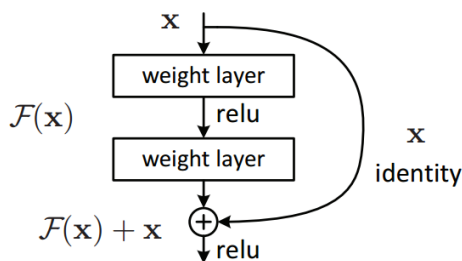
2.2. Κατηγοριοποίηση (Classification)

Η ταξινόμηση/κατηγοριοποίηση είναι μια βασική έννοια στη μηχανική μάθηση και όπως προαναφέρθηκε περιλαμβάνει την επισήμανση ενός δεδομένου βάσει συγκεκριμένων χαρακτηριστικών ή ιδιοτήτων. Όταν δημιουργείται ένα μοντέλο κατηγοριοποίησης στόχος είναι να μπορεί να αναθέτει κατάλληλες ετικέτες σε νέα δεδομένα, βασιζόμενο σε μοτίβα που έχει αναγνωρίσει και μάθει από το σύνολο εκπαίδευσης. Όπως για παράδειγμα η κατηγοριοποίηση εικόνων μαγνητικών τομογραφιών εγκεφάλου σε υγιής και πάσχοντες, με τη νόσο Alzheimer, ένα θέμα που πραγματεύεται η μελέτη των Saratxaga et al [21] όπου οι ερευνητές μελετούν τη διάγνωση της νόσου Αλτσχάιμερ με τη χρήση μαγνητικών τομογραφιών (MRI). Προτείνουν τη χρήση μιας αρχιτεκτονικής βαθιάς μάθησης (Deep Learning) για τη δημιουργία ενός ταξινομητή, που σε συνδυασμό με άλλες τεχνικές προεπεξεργασίας εικόνων επιτυγχάνει αποτελεσματικότερη διάγνωση της νόσου συγκριτικά με παρόμοιες μελέτες. Στη μελέτη όπως και στην παρούσα εργασία χρησιμοποιείται σύνολο δεδομένων που παρέχεται από τη συλλογή «Open Access Series of Imaging Studies» (OASIS). Με τη διαφορά ότι στην έρευνα [21] χρησιμοποιείται το σύνολο δεδομένων OASIS-2 [22] ενώ στην παρούσα εργασία το OASIS-3 [23]. Ομοιότητα υπάρχει επίσης και στη μεθοδολογία προεπεξεργασίας των εικόνων, καθώς οι Saratxaga et al έχουν εξάγει N αριθμό τομών (slices) από το κέντρο του εγκεφάλου, όπου προέκυψε ύστερα από πειραματική διαδικασία ότι ο βέλτιστος αριθμός τομών είναι 10. Αυτό σημαίνει ότι συλλέχθηκαν 10 εικόνες γύρω από το κέντρο του εγκεφάλου. Παρόμοια στην παρούσα εργασία έχουν πραγματοποιηθεί δοκιμές με 10, 30 και 60 εικόνες. Ύστερα από δοκιμές διαφορετικών αρχιτεκτονικών τόσο σε εικόνες δύο διαστάσεων όσο και σε ολόκληρη την MRI (3D), τα καλύτερα αποτελέσματα παρείχε το δίκτυο ResNet18 [24] με BAC (balanced accuracy) 88% για 2D εικόνες σε ταξινόμηση 3 κλάσεων (Healthy, Very mild stage, Severe stage) και 93% για εικόνες 2D σε ταξινόμηση 2 κλάσεων (Norma Cognition NC, Alzheimer Disease AD). Το συμπέρασμα ότι η απόδοση του Resnet με 18 επίπεδα

παρείχε καλύτερα αποτελέσματα λαμβάνεται υπόψιν και στη συγκεκριμένη εργασία και πραγματοποιούνται δοκιμές σύμφωνα με αυτή την αρχιτεκτονική.

2.2.1. Resnet

Το ResNet είναι μια αρχιτεκτονική βαθιάς μάθησης που έχει ιδιαίτερη απήχηση στον τομέα του computer vision, διότι δίνει λύση σε πολύπλοκα προβλήματα, αντικαθιστώντας την ανάγκη για πολύπλοκα βαθιά νευρωνικά δίκτυα και προτάθηκε από τους He et al [24]. Η βασική ιδέα των δικτύων ResNet είναι η χρήση των υπολοίπων (residuals) για την επίτευξη της μάθησης και την αποφυγή των προβλημάτων vanishing gradients degradation problem. Το degradation problem είναι ένα φαινόμενο κατά το οποίο η αύξηση του βάθους του δικτύου οδηγεί στη μείωση της απόδοσης, καθιστώντας μη εφικτή την εκπαίδευση του δικτύου. Στην υλοποίηση του ResNet η κύρια ιδέα για την αντιμετώπιση του φαινομένου αυτού, είναι να χρησιμοποιηθούν συνδέσεις απόσπασης ή συνδέσεις παράκαμψης (shortcut connections ή skip connections) για να επιτραπεί στο δίκτυο να μάθει χαρτογραφήσεις τύπου residual mapping αντί να μάθει άμεσα την υποκείμενη χαρτογράφηση (underlying mapping). Οι χαρτογραφήσεις residual mappings αντιπροσωπεύουν τη διαφορά μεταξύ της επιθυμητής εξόδου και της τρέχουσας εξόδου σε ένα συγκεκριμένο επίπεδο. Οι συνδέσεις παράκαμψης επιτρέπουν την παράκαμψη ενός ή περισσότερων επιπέδων κατά την είσοδο σε ένα μπλοκ και της σύνδεσής τους απευθείας με την έξοδο του μπλοκ. Η δομή αυτή φαίνεται στη εικόνα 2-1.



2-1 Residual Block [24]

2.3. Παραγωγικά ανταγωνιστικά δίκτυα (Generative Adversarial Networks)

Τα Παραγωγικά Ανταγωνιστικά Δίκτυα, ΠΑΔ (Generative Adversarial Networks, GANs) είναι συστήματα στα οποία εφαρμόζεται η στρατηγική ελαχιστοποίησης-μεγιστοποίησης

(ανταγωνιστικό παιχνίδι), όπου δύο αλγόριθμοι ανταγωνίζονται: ένας αλγόριθμος (γεννήτορας, Generator) δημιουργεί δεδομένα με στόχο να ξεγελάσει τον άλλο, ο οποίος επιδιώκει να διακρίνει ανάμεσα σε πλαστά και πραγματικά δεδομένα (διευκρινιστής, Discriminator). Η ιδέα των GAN εμφανίστηκε πρώτη φορά στο επιστημονικό πεδίο το 2014 από τον Ian Goodfellow [25] κι έχει αποτελέσει την πιο υποσχόμενη ιδέα στο πεδίο της μηχανικής μάθησης τα τελευταία χρόνια. Ο Generator λαμβάνει ως είσοδο τυχαίο θόρυβο $z \sim p_g(z)$ και εκπαιδεύεται με στόχο να μάθει την κατανομή των πραγματικών εικόνων $p_{data}(x)$ ώστε να παράγει ρεαλιστικές τεχνητές εικόνες μέσω της αναπαραγόμενης κατανομής $p_G(x)$. Επιπρόσθετα, εκπαιδεύουμε τον discriminator να μεγιστοποιεί την πιθανότητα σωστής ταξινόμησης των δεδομένων και τον generator να την ελαχιστοποιεί. Η διαδικασία της εκπαίδευσης και η δομή ενός δικτύου GAN απεικονίζεται στο σχήμα 2-2. Η προσπάθεια βελτιστοποίησης του μοντέλου που προαναφέρθηκε αποτυπώνεται στον τύπο 2-1:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_g(z)} [\log(1 - D(G(z)))]$$

2-1

Όπου:

G: το δίκτυο του Generator

D: το δίκτυο του Discriminator

x: τυχαίο δείγμα από την πραγματική κατανομή

z: δείγμα τυχαίου θορύβου από προκαθορισμένη κατανομή

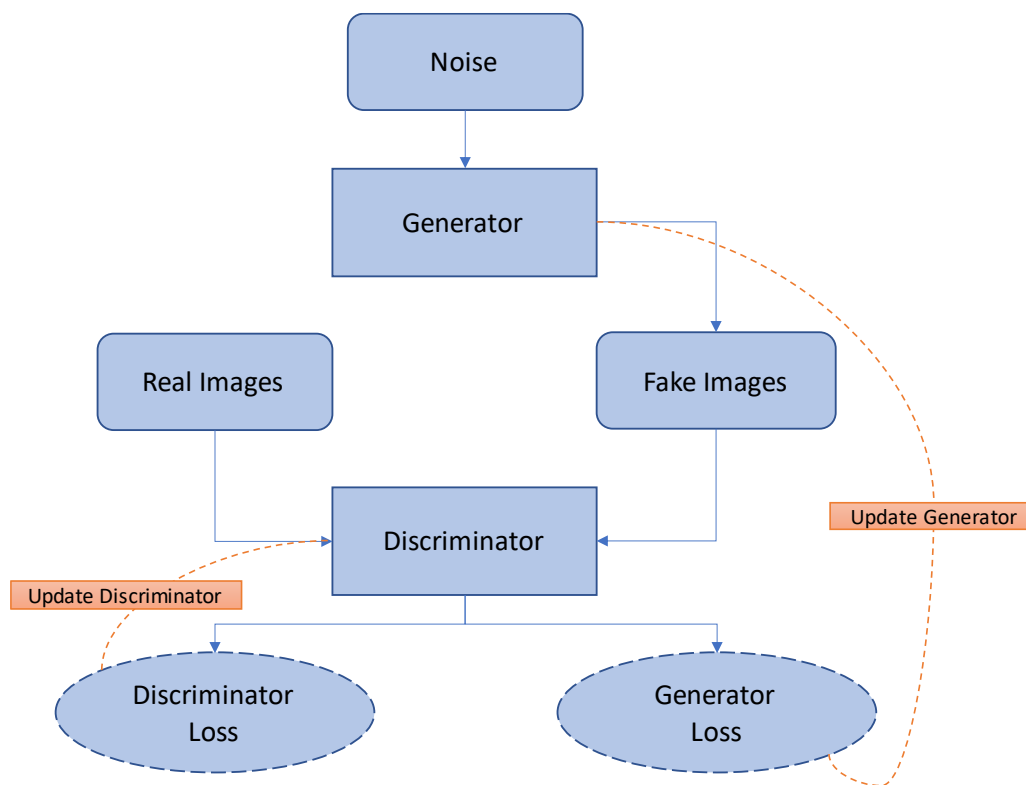
D(x): Το αποτέλεσμα του Discriminator με είσοδο πραγματικά δεδομένα x

D(G(z)): Το αποτέλεσμα του Discriminator με είσοδο δείγμα από τεχνητές εικόνες παραγόμενες από τον Generator.

Η εκπαίδευση των GAN σε όρους θεωρίας παιγνίων μπορεί να περιγραφεί ως ένα zero-sum παιχνίδι, στο οποίο 2 παίκτες ανταγωνίζονται και το κέρδος του ενός ισούται με την απώλεια του άλλου. Στα zero-sum παίγνια υπάρχει ένα σημείο όπου κανένας παίκτης

δεν μπορεί να βελτιώσει την κατάστασή του, το σημείο αυτό σύμφωνα με τη θεωρία παιγνίων ονομάζεται σημείο ισορροπίας Nash (Nash equilibrium).

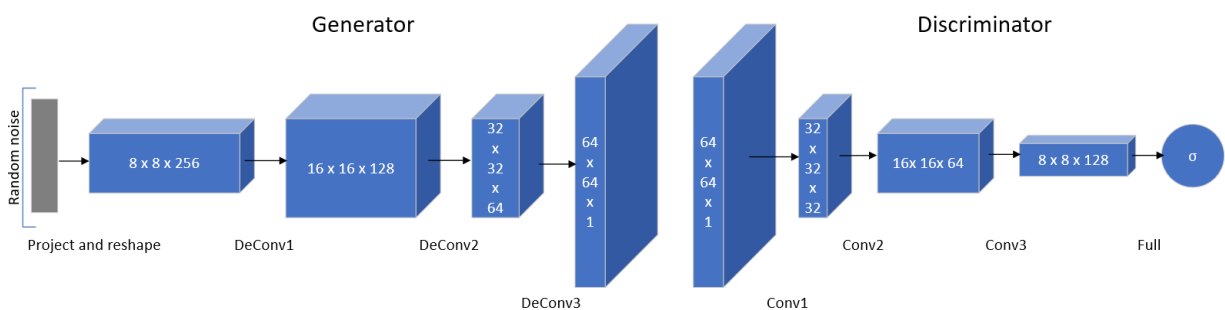
Η πρώτη δομή GAN που προτάθηκε (Vanilla GAN) αποτελείται από ένα Generator και ένα Discriminator, και τα δύο αυτά μοντέλα υλοποιούνται με πολυεπίπεδα perceptrons. Επίσης, χρησιμοποιείται η συνάρτηση απώλειας min-max που παρουσιάζεται στον τύπο 2-1. Η πρώτη εκδοχή των GAN έχει κάποια βασικά μειονεκτήματα, όπως η ασταθής εκπαίδευση, ένα πρόβλημα που καθιστά δύσκολη τη σύγκλιση του μοντέλου και την αδυναμία παραγωγής ρεαλιστικών τεχνητών εικόνων. Ένα άλλο πρόβλημα είναι το Mode Collapse, μια κατάσταση όπου ο Generator έχει μάθει να παράγει μόνο μια περιορισμένη ποικιλία αποτελεσμάτων. Μια ακόμα δυσκολία που αντιμετωπίζουν οι πρώτες αρχιτεκτονικές είναι η αδυναμία αποτύπωσης λεπτομερειών του πραγματικού συνόλου δεδομένων. Στις παρακάτω υποενότητες περιγράφονται μεταγενέστερες εκδοχές GAN με πιο εξελιγμένο τρόπο εκπαίδευσης και πρακτικές, με στόχο τη βελτίωση των αποτελεσμάτων και την αντιμετώπιση των παραπάνω προβλημάτων που υπάρχουν στην αρχική δομή. [26]



2-2 Δομή ενός δικτύου GAN

2.3.1. DCGAN

Η αρχιτεκτονική Deep Convolutional GAN, DCGAN προτάθηκε ως μια αρχιτεκτονική για την αντιμετώπιση προβλημάτων σταθερότητας στην εκπαίδευση των GAN [27], δηλαδή την αδυναμία σύγκλισης, με αποτέλεσμα ο Generator να παράγει ρεαλιστικά δείγματα. Η δομή που παρουσιάζεται εκμεταλλεύεται τις δυνατότητες των Deep Convolutional Neural Networks, τα οποία χρησιμοποιούνται και στον Discriminator και στον Generator. Συγκεκριμένα, προτάθηκε η αφαίρεση των Fully Connected layers και η αντικατάσταση των pooling operations με Convolutional Layers στον Discriminator και Transposed Convolution Layers στον Generator. Για συναρτήσεις ενεργοποίησης προτάθηκαν leaky ReLU για τον Discriminator και ReLU για όλα τα επίπεδα του Generator εκτός από το τελευταίο όπου χρησιμοποιείται tanh. Οι αλλαγές αυτές φαίνεται να βοηθούν στη σταθερότητα της εκπαίδευσης και στην παραγωγή εικόνων καλύτερης ποιότητας και μεγαλύτερης ανάλυσης [27] [28]. Ένα παράδειγμα της προτεινόμενης αρχιτεκτονικής παρουσιάζεται στο σχήμα 2-3.



2-3 Δομή Discriminator και Generator στο DCGAN

Η αρχιτεκτονική DCGAN εμφανίζεται συχνά στη βιβλιογραφία με αρκετές επιτυχημένες εφαρμογές. Στην μελέτη [7] οι ερευνητές διερεύνησαν την κατηγοριοποίηση (classification) εικόνων από συκώτι ανθρώπων με προβλήματα Cysts, metastases και Hemangiomas. Όπως αναφέρεται στην έρευνα, το πρόβλημα των περιορισμένων διαθέσιμων εικόνων σε ιατρικά προβλήματα, όπως το εξεταζόμενο, συμβάλει αρνητικά στην επιτυχή εκπαίδευση αλγορίθμων κατηγοριοποίησης, επηρεάζοντας την αποδοτικότητά τους. Παρουσιάζουν ως αποτελεσματική επίλυση του προβλήματος την παραγωγή επιπλέον δεδομένων με τη χρήση παραγωγικών ανταγωνιστικών δικτύων

GAN. Το ενδιαφέρον στη μελέτη αυτή είναι ότι συνδυάζουν κλασικές μεθόδους εμπλουτισμού του συνόλου δεδομένων με την νεότερη τεχνική των GAN. Πιο συγκεκριμένα, πραγματοποίησαν τις παρακάτω μετατροπές στις πραγματικές εικόνες για να εμπλουτίσουν το σύνολο δεδομένων που θα χρησιμοποιηθεί για την εκπαίδευση των GAN, περιστροφή (rotation, rot), αναστροφή (flipped), μεταφορά στον καρτεσιανό χώρο (translation) και αλλαγή κλίμακας του σημείου ενδιαφέροντος (scale). Συνολικά, πραγματοποιήθηκαν $N = N_{rot} \times (1 + N_{flip} + N_{trans} + N_{scale})$ αυξήσεις. Έπειτα, έγινε χρήση του εμπλουτισμένου συνόλου δεδομένων για την εκπαίδευση 3 DCGANs [27], ένα για κάθε κλάση. Παρόμοια διαδικασία ακολουθήθηκε και στην παρούσα εργασία σε ένα από τα πειράματα, όπου εκπαιδεύτηκαν 3 διαφορετικά GAN αρχιτεκτονικής WGAN-GP [29] [30] το καθένα σε διαφορετική όψη του εγκεφάλου. Τέλος, οι ερευνητές συνοψίζουν καταλήγοντας ότι η υλοποίησή τους έχει θετική επίδραση στα αποτελέσματα του classifier με την κλασική μέθοδο εμπλουτισμού να καταφέρνει 78.6% ακρίβεια, ενώ με τις συνθετικές εικόνες 85.7% ακρίβεια.

Μια ακόμα αξιολογη έρευνα είναι αυτή των Zhou et al [8], στην οποία οι μελετητές εξερευνούν τη δυνατότητα ενός τροποποιημένου GAN, να ενισχύσει την απόδοση ενός ταξινομητή (classifier) για τη νόσο Αλτσχάιμερ (AD), χρησιμοποιώντας μαγνητικές τομογραφίες (MRI) που αποκτήθηκαν σε διαφορετικές εντάσεις μαγνητικού πεδίου (1.5T και 3T). Επιλέχθηκαν σαρώσεις MRI εγκεφάλου T1-weighted από 151 συμμετέχοντες της έρευνας Alzheimer's Disease Neuroimaging Initiative (ADNI) [31], που υποβλήθηκαν ταυτόχρονα σε σαρώσεις με ένταση 1.5-Tesla (1.5T) και 3-Tesla (3T), για τη δημιουργία ενός μοντέλου GAN. Το μοντέλο κατασκευάστηκε και εκπαιδεύτηκε παράλληλα με ένα τρισδιάστατο πλήρως συνελκτικό δίκτυο (FCN), χρησιμοποιώντας τις παραγόμενες εικόνες (3T) ως είσοδο για την πρόβλεψη της κλάσης AD. Η ποιότητα των παραγόμενων εικόνων αξιολογήθηκε με τη χρήση των μετρικών signal to noise ratio (SNR), Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) and Natural Image Quality Evaluator (NIQE). Τα αποτελέσματα έδειξαν ότι ο ταξινομητής FCN βασισμένος στις MRI 3T είχε καλύτερη απόδοση από το μοντέλο FCN που εκπαιδεύτηκε με τις σαρώσεις 1,5T. Για την καλύτερη αξιολόγηση οι ερευνητές πραγματοποίησαν πειράματα με δύο ακόμα σύνολα δεδομένων με εξίσου καλά αποτελέσματα. Επιπλέον,

παρατηρήθηκε ότι η μέση ποιότητα των παραγόμενων MRI 3T ήταν συνεπώς υψηλότερη από τις εικόνες 1.5T, όπως μετρήθηκε με τα SNR, BRISQUE και NIQE στα σύνολα δεδομένων αξιολόγησης. Αυτή η μελέτη αποδεικνύει ότι οι δομές των GAN μπορούν να χρησιμοποιηθούν για να ενισχύσουν την απόδοση ταξινόμησης της νόσου Αλτσχάιμερ, παρέχοντας υψηλή ποιότητα στις τεχνητές εικόνες που παράγουν.

2.3.2. LSGAN

Μια άλλη προσπάθεια σταθερότερης εκπαίδευσης και παραγωγής πιο ποιοτικών εικόνων αποτελεί το Least Squares GAN, LSGAN [32]. Σε αντίθεση με το αρχικό GAN, όπου χρησιμοποιείται συνάρτηση κόστους Binary cross entropy, σε αυτή την εκδοχή αντικαθίσταται με τη συνάρτηση κόστους Ελαχίστων Τετραγώνων, Least Squares (απ' όπου πήρε και το όνομά της η αρχιτεκτονική).

$$V_D(D) = \frac{1}{2} E_{x \sim p_{data}(x)} [(D(x) - b)^2] + \frac{1}{2} E_{z \sim p_z(z)} [(D(G(z))) - a]^2$$

2-2

$$V_G(G) = \frac{1}{2} E_{z \sim p_z(z)} [(D(G(z))) - c]^2$$

2-3

Όπου

a: Επιθυμητή επισήμανση (label) όταν ο Discriminator δέχεται ως είσοδο τεχνητές εικόνες, στόχος είναι να επισημανθούν ως τεχνητές.

b: Επιθυμητή επισήμανση (label) όταν ο Discriminator δέχεται ως είσοδο πραγματικές εικόνες, στόχος είναι να επισημανθούν ως πραγματικές.

c: Επιθυμητή επισήμανση (label) όταν ο Discriminator δέχεται ως είσοδο τεχνητές εικόνες, στόχος είναι να επισημανθούν ως πραγματικές, καθώς πρόκειται για την συνάρτηση κόστους του Generator.

G: το δίκτυο του Generator

D: το δίκτυο του Discriminator

x: τυχαίο δείγμα από την πραγματική κατανομή

z: δείγμα τυχαίου θορύβου από προκαθορισμένη κατανομή

D(x): Το αποτέλεσμα του Discriminator με είσοδο πραγματικά δεδομένα x

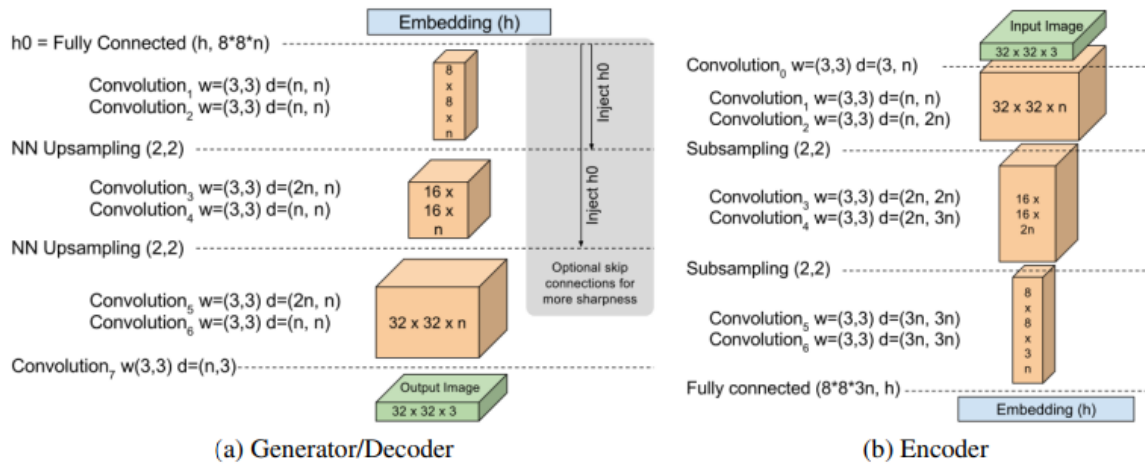
D(G(z)): Το αποτέλεσμα του Discriminator με είσοδο δείγμα από τεχνητές εικόνες παραγόμενες από τον Generator.

Οι παραπάνω συναρτήσεις αποτελούν τις συναρτήσεις κόστους του Discriminator και του Generator αντίστοιχα. Η μετατροπή αυτή βοηθάει στη μείωση του προβλήματος εξαφανιζόμενων κλίσεων (Vanishing Gradients), καθώς και στη σταθερότητα που αφορά την εκπαίδευση του δικτύου. Στην εκδοχή αυτή ο Discriminator προσπαθεί να ελαχιστοποιήσει τη συνάρτηση 2-2 και ο Generator να ελαχιστοποιήσει τη συνάρτηση 2-3.

2.3.3. BEGAN

Το Boundary Equilibrium GAN (BEGAN) είναι κι αυτό με τη σειρά του μία παραλλαγή της παραδοσιακής αρχιτεκτονικής GAN, που όπως κάθε μεταγενέστερη προσπάθεια στοχεύει στην επίτευξη μιας καλύτερης ισορροπίας μεταξύ του Discriminator και του Generator [33]. Η βασική ιδέα στην αρχιτεκτονική αυτής είναι το μέτρο της «οριακής ισορροπίας» (Boundary Equilibrium) το οποίο ποσοτικοποιεί την ισορροπία μεταξύ των δύο αντίπαλων δικτύων και η χρήση ενός auto-encoder στον Discriminator και ενός decoder στον Generator, όπως παρουσιάζεται στο σχήμα 2-4. Με τη χρήση ενός auto-encoder για Discriminator, συστήνεται η ευθυγράμμιση των κατανομών της απώλειας του τελευταίου, χρησιμοποιώντας συνάρτηση κόστους που προκύπτει από την απόσταση Wasserstein [29]. Συνάρτηση κόστους του Discriminator: $L_D = L(x) - k_t L(G(z_D))$, συνάρτηση κόστους Generator: $L_G = L(G(z_G)) = \|D(G(z)) - G(z)\|$, η παράμετρος k χρησιμοποιείται για τον έλεγχο της επιρροής του $L(G(z))$, το L εκφράζει την απώλεια από την ανακατασκευή των δοθέντων εικόνων, με τη χρήση των auto-encoders. Για τη διατήρηση της ισορροπίας στην εκπαίδευση προτάθηκε η χρήση της equilibrium loss: $E |L(G(z))| = \gamma E[L(x)]$, η παράμετρος γ χρησιμοποιείται για την εξισορρόπηση των δύο

στόχων του discriminator, του auto-encode και τις αναγνώρισεις των τεχνητών από τις πραγματικές εικόνες.



2-4 Δομή αρχιτεκτονικής BEGAN [33]

2.3.4. WGAN & WGAN-GP

Μια ακόμα προσθήκη στην προσπάθεια σταθεροποίησης της εκπαίδευσης και σύγκλισης των παραδοσιακών GAN, αποτελεί το Wasserstein Generative Adversarial Network (WGAN) [29]. Στην αρχιτεκτονική αυτή δε χρησιμοποιείται η απόσταση Kullback-Leibler (KL) ή η απόκλιση Jensen-Shannon (JS), αλλά η Wasserstein απόσταση ή αλλιώς Earth Mover's distance. Διαισθητικά, η απόσταση Wasserstein $WD(P_r, P_s)$ μετρά την ελάχιστη προσπάθεια που απαιτείται για να μετατοπιστεί η πυκνότητα από το P_r στο P_s . Έτσι, η WD μετρά τη διαφορά μεταξύ δύο κατανομών δεδομένων πιθανότητας P_r και P_s . Η απόσταση δίνεται από τον τύπο 2-4.

$$WD(P_r, P_s) = \inf_{\Psi \in Y(P_r, P_s)} E [\|a - b\|]$$

2-4

Το inf υποδηλώνει το infimum (το μεγαλύτερο κάτω φράγμα), η κοινή κατανομή πιθανότητας των a και b είναι Ψ , ενώ οι κατανομές πιθανότητας των a και b είναι η αληθινή κατανομή πιθανότητας P_r και η συνθετική κατανομή πιθανότητας P_s αντίστοιχα. Η χρήση

της παραπάνω συνάρτησης είναι ασύμφορη λόγω του inf. Συνεπώς, με βάση την Kantorovich-Rubinstein συνάρτηση προκύπτει η επόμενη συνάρτηση 2-5.

$$WD(P_r, P_s) = \sup_C E_{a \sim P_r} [C(a)] - E_{b \sim P_s} [C(b)]$$

2-5

Σε αυτήν την προσέγγιση, αντί να υπολογίζουμε το ελάχιστο από όλες τις κοινές κατανομές $\Psi(a, b)$, πραγματοποιείται υπολογισμός του ανώτατου ορίου από όλες τις 1-Lipschitz συναρτήσεις. Μια συνάρτηση λέγεται K-Lipschitz αν υπάρχει ένας πραγματικός αριθμός K τέτοιος ώστε για κάθε ζευγάρι σημείων στο γράφημα αυτής της συνάρτησης, η απόλυτη τιμή της κλίσης της ευθείας που τα ενώνει, να μην είναι μεγαλύτερη από το K, δηλαδή αν το K είναι άνω φράγμα του μέτρου των παραγώγων της. Για το λόγο αυτό, περικόπτουμε τα βάρη του δικτύου (weight clipping) για να εξασφαλίσουμε ότι βρίσκονται εντός ενός συγκεκριμένου εύρους που καθορίζεται από μια παράμετρο. Επιπλέον, η διαφορά μεταξύ του WGAN και του τυπικού GAN, εκτός από τη συνάρτηση κόστους, είναι ότι αφαιρούμε τον τελευταίο επίπεδο (sigmoid) του Discriminator, έτσι ώστε να εξάγει μια αριθμητική τιμή αντί για μια πιθανότητα. Αυτή η νέα προσέγγιση του Discriminator ονομάζεται Critic, όπως προτείνεται από τους Arjovsky et al [29].

Τα πλεονεκτήματα του WGAN είναι πάρα πολλά κάτι το οποίο έχει αποδειχθεί σε πλήθος μελετών [34] [35] [36], αλλά η περικοπή των βαρών μπορεί να οδηγήσει σε διάφορα προβλήματα. Για την αντιμετώπιση αυτών προτάθηκε η εφαρμογή του Gradient penalty όπου βοηθάει τον Critic να καλύψει τους περιορισμούς των κριτηρίων 1-Lipschitz. Η συνάρτηση κόστους μετατρέπεται στη συνάρτηση 2-6 όπου το W^{GP} είναι η συνάρτηση 2-7.

$$L_C^{GP} = L_C^{WG} + \lambda W^{GP}$$

2-6

$$W^{GP} = E_{z \sim P_n} [\|\nabla_z C(z)\|_2 - 1]^2$$

2-7

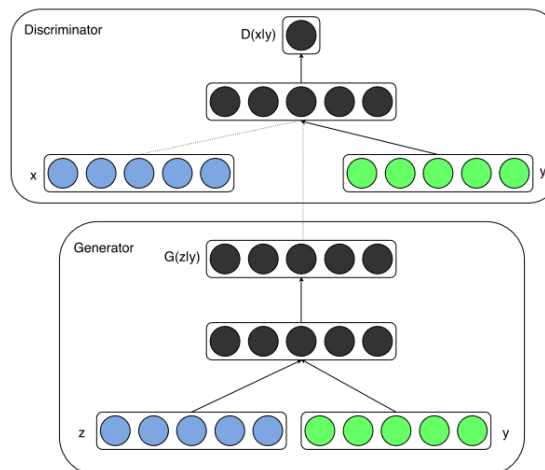
Το Gradient υπολογίζεται κάθε φορά σε δείγμα τυχαίας δειγματοληψίας $z \sim P_n$. Η παράμετρος λ είναι μια σταθερά που ορίζεται πριν από την εκπαίδευση και καθορίζει πόσο θα επηρεάζει το μέτρο της παραγώγου. Τέλος, μια άλλη διαφορά με το αρχικό WGAN, στο WGAN-GP δεν πραγματοποιείται batch normalization στον Critic.

2.3.5. CGAN

Η αρχιτεκτονική Conditional GAN (cGAN), έρχεται να λύσει ένα άλλο πρόβλημα των παραδοσιακών μοντέλων, τον διαχωρισμό των κλάσεων που υπάρχουν στις εικόνες, δηλαδή να έχει την ικανότητα το μοντέλο να καταλάβει τι απεικονίζει η τεχνητή εικόνα που παράγει. Στην έρευνα [37] πραγματοποιήθηκε αυτή η υλοποίηση όπου καθοδηγείται η διαδικασία παραγωγής τεχνητών εικόνων υπό συνθήκη κάποιον επιπλέον πληροφοριών. Η τροποποίηση γίνεται και στον Generator και στον Discriminator, ώστε και στα δύο δίκτυα μαζί με τον τυχαίο θόρυβο και τις πραγματικές/τεχνητές εικόνες αντίστοιχα, να δίνονται και οι αντίστοιχες κλάσεις σαν επιπλέον επίπεδο εισαγωγής. Η συνάρτηση βελτιστοποίησης του μοντέλου αποτελεί μετατροπή αυτής του παραδοσιακού GAN ώστε να περιλαμβάνει και τις κλάσεις (y), όπως παρουσιάζεται στην συνάρτηση 2-8. Η δομή ενός cGAN παρουσιάζεται στην εικόνα 2-5.

$$\min_G \max_D V(D, G) = E_{X \sim p_{data}(x)} [\log D(x | y)] + E_{z \sim p_g(z)} [\log(1 - D(G(z | y)))]$$

2-8



2-5 Δομή αρχιτεκτονικής cGAN [37]

Η δομή cGAN έχει αρκετές εφαρμογές, όπως η παραγωγή εικόνων βάση κάποιου κειμένου (text-to-image synthesis), η αύξηση της ανάλυσης των εικόνων (super-resolution) και το image-to-image translation όπου μια εικόνα μπορεί να μετατραπεί και σε κάτι διαφορετικό, όπως η αλλαγή ενός αλόγου σε ζέβρα.

Μια μελέτη που βασίζεται στο image-to-image translation είναι αυτή των Shin et al. [38]. Οι ερευνητές προτείνουν μια λύση για τη διάγνωση της νόσου Alzheimer (AD) που βασίζεται στην αρχιτεκτονική cGAN, για τη μετατροπή μαγνητικών τομογραφιών (MRI) σε τομογραφίες εκπομπής ποζιτρονίων (Positron Emission Tomography PET). Η τομογραφία εκπομπής ποζιτρονίων θεωρείται η πιο έγκυρη εξέταση για τη διάγνωση της AD, δυστυχώς όμως το υψηλό κόστος διεξαγωγής της και η υψηλή δόση ακτινοβολίας αποτρέπουν αρκετούς ασθενείς να την πραγματοποιήσουν. Η MRI, από την άλλη είναι πιο οικονομική και εύκολη να πραγματοποιηθεί από τον ασθενή, γεγονός που διευκολύνει τη συλλογή δεδομένων από τους ερευνητές. Ωστόσο, οι μικρές δομικές διαφορές μεταξύ υγιών ατόμων και ατόμων με AD που είναι ορατές στην MRI καθιστούν τη διάγνωση της AD μεγάλη πρόκληση. Για να αντιμετωπιστεί αυτό, οι συγγραφείς προτείνουν μια προσέγγιση βασισμένη σε cGAN που ονομάζεται GANDALF (Generative Adversarial Networks with Discriminator-Adaptive Loss Fine-tuning), η οποία συνθέτει τομογραφίες PET από MRI. Έχει χρησιμοποιηθεί το σύνολο δεδομένων ADNI [31]. Σε αντίθεση με παρόμοιες εργασίες που πραγματεύονται τη μετατροπή από MRI σε PET, οι Shin et al πραγματοποιούν της εκπαίδευση του GAN και του Classifier ταυτόχρονα. Η υλοποίηση που προτείνουν αποτελείται από ένα Generator με δομή encoder-decoder, ο οποίος λαμβάνει Real MRI (διαφορετικών κλάσεων, AD/MCI/NC κλπ) και εξάγει PET, η οποία στη συνέχεια εισάγεται σαν είσοδος στον Discriminator. Ο Discriminator έχει διπλό ρόλο, πρώτα ταξινομεί τις εισόδους σε πραγματική/τεχνητή (real/fake) κι έπειτα κατηγοριοποιεί βάσει των δοθέντων κλάσεων. Ακόμα μια τεχνική διαφορά που είναι σημαντική για την απόδοση είναι η χρήση μιας επιπλέον συνάρτησης κόστους που αφορά την ταξινόμηση AD, καθώς επίσης και η προσαρμογή ενός συντελεστή λ_G όταν οι συναρτήσεις κόστους δεν έχουν το επιθυμητό αποτέλεσμα, ενώ ταυτόχρονα επιστρέφεται το προηγούμενο σημείο αναφοράς προτού εμφανιστεί η ανεπιθύμητη τιμή στη συνάρτηση κόστους. Η προτεινόμενη προσέγγιση GANDALF πετυχαίνει υψηλή ακρίβεια με 78% σε τρεις κλάσεις

(AD/MCI/CN) και ακρίβεια σε τέσσερις κλάσεις (AD/LMCI/EMCI/CN) 37% το οποίο αν και χαμηλό αποτελεί βελτίωση έναντι παρόμοιων εργασιών.

2.3.6. CWGAN-GP

Η ιδέα του conditional Wasserstein GAN (cWGAN-GP) με gradient penalty, ξεκίνησε σαν πειραματική διαδικασία στο πλαίσιο αυτής της εργασίας τροποποιώντας τον κώδικα που υλοποιήθηκε για την αρχιτεκτονική WGAN-GP, προσθέτοντας τις κατηγορίες (labels) σαν είσοδο στον Discriminator και στον Generator. Έπειτα για να παρουσιαστεί μια ολοκληρωμένη εικόνα της βιβλιογραφίας, έγινε αναζήτηση παρόμοιων ερευνών και βρέθηκαν οι έρευνες [39] και [40] όπου έχουν πραγματοποιήσει παρόμοια υλοποίηση. Η δομή του μοντέλου είναι παρόμοια με αυτή του cGAN και φαίνεται στην εικόνα 2-5. Η συνάρτηση απώλειας του Critic παρουσιάζεται στη συνάρτηση 2-9 και του Generator στη συνάρτηση 2-10. Οι συναρτήσεις απώλειας είναι παρόμοιες με αυτές του WGAN-GP, με την μόνη διαφορά ότι δίνεται και η επιθυμητή κλάση (y).

$$L_C^{GP} = -E_{x \sim p_r} [C(x | y)] + E_{z \sim p_z(z)} [C(z | y)] + \lambda E_{z \sim p_z} [(\|\nabla_z C(z | y)\|_2 - 1)^2]$$

2-9

$$L_G^{GP} = -E_{z \sim p_z} [C(z | y)]$$

2-10

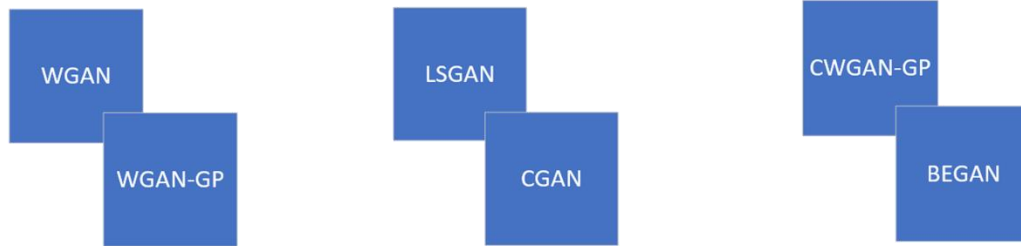
3. Μεθοδολογία

3.1. Ροή διαδικασιών

Η επίλυση ενός προβλήματος ταξινόμησης χρησιμοποιώντας ένα σύνολο δεδομένων με άνισο αριθμό δειγμάτων μεταξύ των κλάσεων, αποτελεί πρόκληση καθώς τα αποτελέσματα του ταξινομητή επηρεάζονται από την ανισορροπία. Αυτό γίνεται αντιληπτό από την αξιολόγηση των μετρικών κατά την εκπαίδευση του ταξινομητή. Ο στόχος της παρούσας εργασίας είναι η σύγκριση διαφορετικών μεθόδων για την επίλυση του προβλήματος αυτού και πιο συγκεκριμένα η βελτίωση των μετρικών του ταξινομητή. Η διαδικασία που ακολουθήθηκε αποτελείται από 4 στάδια και αποτυπώνεται στο σχήμα 3-1.

Το πρώτο βήμα είναι η διερεύνηση διαφόρων αρχιτεκτονικών με σκοπό την επιλογή των καταλληλότερων για το συγκεκριμένο σύνολο δεδομένων. Ακολουθεί η παραγωγή τεχνητών εικόνων από τα επιλεγμένα GAN. Έπειτα στο τρίτο βήμα εκπαιδεύεται ο ταξινομητής χρησιμοποιώντας το αρχικό σύνολο δεδομένων με την ανισορροπία στις κλάσεις. Τέλος, στο τέταρτο βήμα πραγματοποιείται η αξιολόγησή του κατά πόσο επηρεάζει την απόδοση του ταξινομητή η προσθήκη τεχνητών εικόνων, με τη δοκιμή 5 διαφορετικών ποσοστών 10%, 20%, 50%, 80% και 100% για κάθε σύνολο τεχνητών εικόνων που παράχθηκαν από τα επιλεγμένα GAN.

1 Διερεύνηση, αξιολόγηση και επιλογή αποδοτικότερων Αρχιτεκτονικών GAN



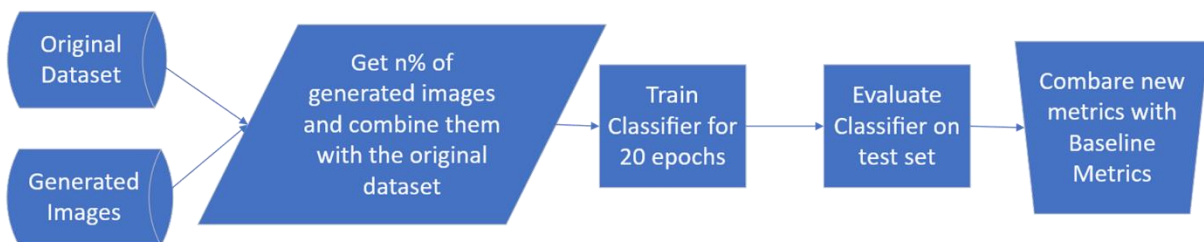
2 Παραγωγή τεχνητών εικόνων



3 Εκπαίδευση ταξινομητή στο αρχικό σύνολο δεδομένων



4 Αξιολόγηση της επιρροής των τεχνητών εικόνων στον ταξινομητή
Επανάληψη της διαδικασίας με την προσθήκη διαφορετικού ποσοστού τεχνητών εικόνων n%



3.2. Παραγωγικά Ανταγωνιστικά Δίκτυα – Αρχιτεκτονικές και λεπτομέρειες υλοποίησης

Υλοποιήθηκαν διαφορετικές αρχιτεκτονικές με σκοπό την εύρεση της καλύτερης που θα παράγει αποτελέσματα (generated images) πολύ κοντά στα πραγματικά δεδομένα. Η αρχική ιδέα ήταν η παραγωγή εικόνων διαστάσεων 64x64, καθώς οι απαιτήσεις σε επίπεδο πόρων ήταν υψηλές, κάτι το οποίο από τα αρχικά στάδια αποδείχθηκε ότι δεν είναι αποδοτικό διότι χανόταν σημαντική πληροφορία λόγω του μικρού μεγέθους. Τελικά, επιλέχθηκε η διάσταση 128x128 ως μια μέση λύση, καθώς οι πραγματικές εικόνες ήταν 256x256. Στις επόμενες υποενότητες θα παρουσιαστούν οι υλοποιήσεις που αξίζει να αναφερθούν, τόσο λόγω ποιότητας αποτελεσμάτων, όσο και αξιοσημείων παρατηρήσεων κατά τη διαδικασία.

3.2.1. WGAN (Gradient Penalty)

Μια από της σημαντικότερες προσπάθειες που έγιναν κατά τη διάρκεια της εργασίας αυτής είναι τα πειράματα με την αρχιτεκτονική Wasserstein GAN με Gradient Penalty (WGAN-GP). Επιλέχθηκε λόγω της πολύ καλής απόδοσης τους όπως έχει αποδειχθεί στη βιβλιογραφία, αλλά και για το λόγο ότι η συνάρτηση απώλειας έχει νόημα και συσχέτιση με την ποιότητα των παραγόμενων εικόνων. Το γεγονός αυτό αποτέλεσε ένα καλό κριτήριο, μαζί με την οπτική (ποιοτική) αξιολόγηση, για την αξιολόγηση της πορείας της εκπαίδευσης του μοντέλου. Κατά τη διάρκεια των πειραμάτων πραγματοποιήθηκαν δοκιμές με διάφορες παραλλαγές στα δύο δίκτυα (Discriminator και Generator), τόσο σε πολυπλοκότητα όσο και σε τιμές υπερπαραμέτρων (learning rate, lambda, n_critic_iterations κλπ.). Οι πρώτες δοκιμές έγιναν με μέγεθος παραγόμενων εικόνων σε διάσταση 64x64, κι όπως προαναφέρθηκε, δεν κατέστη δυνατή η εμφάνιση λεπτομερειών ακόμα και ύστερα από αρκετές εποχές. Έπειτα, πραγματοποιήθηκε η μετατροπή των δύο δικτύων ώστε να υποστηρίζουν τη διάσταση 128x128 προσθέτοντας επιπλέον convolutional/Transposed Convolutional επίπεδα. Η δομή χωρίς Gradient Penalty και με weight clipping δοκιμάστηκε στα πρώτα πειράματα αλλά υπερίσχυσε η μεταγενέστερη εκδοχή και δεν πραγματοποιήθηκαν περαιτέρω δοκιμές.

Η δομή της βασικής αρχιτεκτονικής που χρησιμοποιήθηκε στα πειράματα παρουσιάζεται στον πίνακα 3-1 για τον Generator και στον πίνακα 3-2 για τον Critic. Στους πίνακες παρουσιάζονται οι εκπαιδευσιμοι παράμετροι των μοντέλων. Επιπλέον επίπεδα που δεν φαίνονται θα αναφερθούν στη συνέχεια.

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	ConvTranspose2d	4,194,304
1	BatchNorm2d	2,048
2	ConvTranspose2d	8,388,608
3	BatchNorm2d	1,024
4	ConvTranspose2d	2,097,152
5	BatchNorm2d	512
6	ConvTranspose2d	524,288
7	BatchNorm2d	256
8	ConvTranspose2d	131,072
9	BatchNorm2d	128
10	ConvTranspose2d	1,025
	Σύνολο	15,340,417

3-1 Αρχιτεκτονική WGANGP 1, Generator

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Conv2d	544
1	Conv2d	32,768
2	InstanceNorm2d	128
3	Conv2d	131,072
4	InstanceNorm2d	256
5	Conv2d	524,288
6	InstanceNorm2d	512
7	Conv2d	2,097,152
8	InstanceNorm2d	1,024
9	Conv2d	8,193
	Σύνολο	2,795,937

3-2 Αρχιτεκτονική WGANGP 1, Discriminator

Ο Generator αποτελείται από 5 “blocks” τριών στοιχείων, ένα Transposed Convolution επίπεδο (ConvTranspose2d), ένα Batch Normalization επίπεδο (BatchNorm2d) και τη συνάρτηση ενεργοποίησης LeakyReLU. Στο σημείο αυτό είναι σκόπιμο να αναφερθεί ότι σε αρχικά πειράματα χρησιμοποιήθηκε η ReLU, αλλά στην πορεία παρατηρήθηκε ότι αποδιδόταν μεγαλύτερη λεπτομέρεια με την LeakyReLU. Τέλος, ακολουθεί ένα ακόμα

ConvTranspose2d με συνάρτηση ενεργοποίησης όμως την tanh. Όλα τα ConvTranspose2d επίπεδα πραγματοποιούν ανάστροφη συνέλιξη με kernel size (μέγεθος πυρήνα) 4, stride (βήμα) 2 και padding 1, εκτός από το πρώτο το οποίο έχει stride 0 και padding 1. Σε κάθε επίπεδο η εικόνα διπλασιάζει το μέγεθος της, με τη μόνη διαφορά ότι στο πρώτο επίπεδο εξάγεται εικόνα διαστάσεων 4x4 καθώς το δίκτυο έχει είσοδο 1x1 με 256 κανάλια (noise). Για το μέγεθος του θορύβου που λαμβάνει ως είσοδο ο Generator πραγματοποιήθηκαν δοκιμές με 100, 128, 256 και 512 channels. Πραγματοποιήθηκε δοκιμή αρχιτεκτονικής με διπλάσιες παραμέτρους. Δυστυχώς η εκπαίδευση δε συνεχίστηκε για μεγάλο διάστημα λόγω υψηλών υπολογιστικών απαιτήσεων και χρόνου, μέχρι το σημείο όμως που εκτελέστηκε φάνηκε να έχει παρόμοια απόδοση με το αντίστοιχο σημείο της πιο απλής αρχιτεκτονικής.

Ο Discriminator περιλαμβάνει ως πρώτο στάδιο ένα συνελικτικό επίπεδο (Convolutional Layer, Conv2d) με συνάρτηση ενεργοποίησης LeakyReLU. Έπειτα, ακολουθούν 4 «blocks», τα οποία αποτελούνται από Conv2d, InstanceNorm2d και LeakyReLU. Τελευταίο επίπεδο είναι ένα Conv2d layer. Κάθε Conv2d layer έχει kernel size 4, stride 2 και padding 1, με εξαίρεση το τελευταίο το οποίο έχει padding 0, ώστε να αποδώσει ως έξοδο ένα χαρακτηριστικό. Κάθε Conv2d layer έχει σκοπό τον υποδιπλασιασμό το μεγέθους της εικόνας.

Υλοποιήθηκε και δοκιμάστηκε μια δεύτερη αρχιτεκτονική της οποίας τα χαρακτηριστικά παρουσιάζονται στους πίνακες 3-3 και 3-4, Generator και Discriminator αντίστοιχα. Σε αυτήν την υλοποίηση πραγματοποιείται η προσθήκη τυχαίου θορύβου στα layer του Generator με σκοπό την ενίσχυση της γενίκευσης του μοντέλου και της επίτευξη σταθερότερης εκπαίδευσης [41] [42] [43].

	Όνομα Επιπέδου	Αριθμός παραμέτρων
0	Linear	131,584
1	BatchNorm1d	1,024
2	ConvTranspose2d	4,194,304
3	BatchNorm2d	1,024
4	ConvTranspose2d	2,097,152
5	BatchNorm2d	512
6	ConvTranspose2d	524,288
7	BatchNorm2d	256
8	ConvTranspose2d	131,072
9	BatchNorm2d	128
10	ConvTranspose2d	32,768
11	BatchNorm2d	64
12	ConvTranspose2d	8,192
13	BatchNorm2d	32
14	ConvTranspose2d	257
	Σύνολο	7,122,657

3-3 Αρχιτεκτονική WGANGP 2, Generator

	Όνομα Επιπέδου	Αριθμός παραμέτρων
0	Conv2d	160
1	Conv2d	8,192
2	Conv2d	32,768
3	Conv2d	131,072
4	Conv2d	524,288
5	Conv2d	2,097,152
6	Conv2d	8,193
	Σύνολο	2,801,825

3-4 Αρχιτεκτονική WGANGP 2, Discriminator

Οι διαφορές με την πρώτη αρχιτεκτονική παρατηρούνται στον Generator στα πρώτα layer όπου προστέθηκε ένα Linear layer και ένα ακόμα BatchNorm1d και τα ConvTranspose2d layers έχουν kernel size 4, stride 2 και padding 1. Ο Discriminator αποτελείται από 7 Conv2d layers. Αυτή η αρχιτεκτονική επίσης εξάγει εικόνα μεγέθους 1x128x128 (1 είναι τα κανάλια της εικόνας). Όπως μπορεί να παρατηρηθεί η δομή και των δύο δικτύων και στις δύο αρχιτεκτονικές είναι επηρεασμένη από την αρχιτεκτονική DCGAN [27].

Όπως όλα τα πειράματα στην εργασία ο κώδικας αναπτύχθηκε με τη χρήση της βιβλιοθήκης PyTorch [44]. Τα δύο δίκτυα δημιουργήθηκαν σαν δύο ξεχωριστές κλάσεις

(class) κληρονομώντας την κλάση «Module». Ορίστηκε η δομή τους με εκτέλεση κατά σειρά (Sequential) και μια forward μέθοδος, όπου στην πρώτη αρχιτεκτονική δεν πραγματοποιείται καμία αλλαγή στην είσοδο, ενώ στη δεύτερη προστίθεται τυχαίος θόρυβος στο διάνυσμα. Για την εκπαίδευση και των δύο αρχιτεκτονικών ακολουθήθηκε η ίδια διαδικασία, ορισμός παραμέτρων, αρχικοποίηση των δύο δικτύων, αρχικοποίηση συνόλου δεδομένων και εκκίνηση της εκπαίδευσης. Η αρχικοποίηση των εικόνων περιλαμβάνει μια σειρά από μετασχηματισμούς, αλλαγή μεγέθους εικόνας από 256x256 σε 128x128 με Bilinear Interpolation, μετατροπή σε tensor και τέλος κανονικοποίηση ώστε όλες οι τιμές να είναι στο διάστημα $-1, 1$. Η κάθε εποχή της εκπαίδευσης αποτελείται από N ενημερώσεις (εκπαίδευση) του Discriminator, καθώς στην αρχιτεκτονική αυτή ο Discriminator εκπαιδεύεται περισσότερες φορές σε μια εποχή από τον Generator. Σε κάθε ενημέρωση πραγματοποιείται η παραγωγή τεχνητών εικόνων από τον Generator, οι οποίες χρησιμοποιούνται μαζί με τις πραγματικές εικόνες, για τον υπολογισμό της συνάρτησης απώλειας. Η συνάρτηση απώλειας υπολογίζεται ως: $-(torch.mean(real_images) - torch.mean(fake_images)) + \lambda * gp$. Όπως παρατηρείται έχει αντιστραφεί το πρόσημο, διότι επιχειρείται η ελαχιστοποίηση αντί για μεγιστοποίηση όπως περιγράφεται στην έρευνα [30], καθώς οι μέθοδοι βελτιστοποίησης (optimizers), όπως ο Adam που χρησιμοποιείται στη συγκεκριμένη υλοποίηση πραγματοποιούν την ελαχιστοποίηση της συνάρτησης κόστους. Ύστερα από τον αριθμό ενημερώσεων που έχει τεθεί για το Discriminator, ακολουθεί η εκπαίδευση του Generator όπου υπολογίζεται η απώλεια ως: $-torch.mean(discriminator_output)$, όπου discriminator_output είναι το αποτέλεσμα του Discriminator για τις τεχνητές εικόνες που έχει δημιουργήσει ο Generator.

Με τη συγκεκριμένη αρχιτεκτονική πραγματοποιήθηκαν δοκιμές στο σύνολο ΣΔ3, το οποίο περιέχει εικόνες και από τα 3 διαφορετικά επίπεδα του εγκεφάλου (Coronal, Axial and Sagittal). Παρατηρήθηκε ότι στις παραγόμενες εικόνες σε πολλές περιπτώσεις αναμιγνύονταν τα επίπεδα, χωρίς να είναι ξεκάθαρο πιο από τα 3 απεικονίζεται. Με άλλα λόγια σε πολλές από τις τεχνητές εικόνες εμφανιζόταν μια θολή μορφή εγκεφάλου, που μοιραζόταν λεπτομέρειες από τις 3 διαφορετικές τομές, με αποτέλεσμα να μην είναι φανερό ποια τομή του εγκεφάλου απεικονίζεται. Για παράδειγμα η Coronal τομή απεικονίζει τον εγκέφαλο από μπροστά και καταλαμβάνει λιγότερο χώρο στην MRI, ενώ

η Axial τομή παρουσιάζει τον εγκέφαλο από πάνω, παρατηρήθηκαν περιπτώσεις που ο εικονιζόμενος εγκέφαλος περιείχε λεπτομέρειες στο κέντρο του από την Axial τομή, ενώ το μέγεθος και το σχήμα του στην περιφέρεια έμοιαζε περισσότερο με την Coronal τομή.

Για τον λόγο αυτό, κι εξαιτίας του ότι έδειχνε το μοντέλο να έχει προοπτική, χρησιμοποιήθηκαν τα βάρη από ένα μοντέλο που εκπαιδεύτηκε και στα 3 επίπεδα (transfer learning), για να εκπαιδευτούν 3 διαφορετικά μοντέλα, ένα για το κάθε επίπεδο.

3.2.2. CWGAN (Gradient Penalty)

Παρατηρήθηκε ότι με τη χρήση του συνόλου ΣΔ3, που περιλαμβάνει και τις 3 τομές του εγκεφάλου, τα μη εξαρτώμενα δίκτυα (non-conditional) παρήγαγαν τυχαία εικόνες από διάφορες τομές και πολλές φορές αναμειγμένες εικόνες, χωρίς δηλαδή να είναι πάντοτε ξεκάθαρο το ποια τομή απεικονίζεται. Μια προσπάθεια αντιμετώπισης του παραπάνω προβλήματος ήταν η χρήση τριών non-conditional δικτύων αρχιτεκτονικής WGAN-GP, όπως αναφέρθηκε στην προηγούμενη ενότητα. Σε αυτή την ενότητα παρουσιάζεται μια ακόμη προσπάθεια επίλυσης του παραπάνω προβλήματος, που περιλαμβάνει τον συνδυασμό της καλής απόδοσης της απόστασης Wasserstein και της χρήσης των επισημάνσεων (labels) των επιπέδων κατά την εκπαίδευση, δημιουργώντας έτσι ένα conditional GAN. Στους πίνακες 3-5 και 3-6 περιγράφεται η αρχιτεκτονική του δικτύου, με την παρουσίαση των εκπαιδευσιμων παραμέτρων τόσο του Generator όσο και του Discriminator.

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Embedding	9
1	ConvTranspose2d	8,486,912
2	BatchNorm2d	4,096
3	ConvTranspose2d	33,554,432
4	BatchNorm2d	2,048
5	ConvTranspose2d	8,388,608
6	BatchNorm2d	1,024
7	ConvTranspose2d	2,097,152
8	BatchNorm2d	512
9	ConvTranspose2d	524,288
10	BatchNorm2d	256
11	ConvTranspose2d	2,049
	Σύνολο	53,061,386

3-5 Αρχιτεκτονική CWGAN, δομή Generator

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Embedding	300
1	Linear	4,964,352
2	Conv2d	4,160
3	Conv2d	131,072
4	InstanceNorm2d	256
5	Conv2d	524,288
6	InstanceNorm2d	512
7	Conv2d	2,097,152
8	InstanceNorm2d	1,024
9	Conv2d	8,388,608
10	InstanceNorm2d	2,048
11	Conv2d	16,385
	Σύνολο	16,130,157

3-6 Αρχιτεκτονική CWGAN, δομή Discriminator

Ο Generator έχει παρόμοια δομή με τον αυτόν στο WGANGP 1 με μερικές διαφορές. Το πρώτο επίπεδο πλέον είναι ένα Embedding layer, το οποίο είναι υπεύθυνο για την εκμάθηση της κλάσης (επίπεδο του εγκεφάλου), με έξοδο 9 χαρακτηριστικών. Το επόμενο επίπεδο είναι μια ακολουθία (Sequential) με 5 «blocks» από ConvTranspose2d, BatchNorm2d και LeakyReLU layers. Το πρώτο ConvTranspose2d δέχεται ως είσοδο το άθροισμα των καναλιών του θορύβου και της εξόδου του προηγούμενου layer. Ως τελευταίο layer έχουμε μόνο ConvTranspose2d με συνάρτηση ενεργοποίησης Tanh, το

οποίο εξάγει μια εικόνα 1x128x128. Για την ένωση των Embedding και Sequential, τροποποιήθηκε η Forward μέθοδος του μοντέλου, ώστε να δέχεται ταυτόχρονα batches από τυχαίο θόρυβο και labels με τιμές από τα διάφορα επίπεδα του εγκεφάλου (κλάσεις), σύμφωνα με τις πραγματικές εικόνες, που θα χρησιμοποιηθούν για την εκπαίδευση του Discriminator.

Ο Discriminator έχει κι αυτός παρόμοια δομή με αυτόν στο WGANP 1 με την εξής διαφορά, για τη μάθηση των κλάσεων δημιουργήθηκε άλλο ένα εσωτερικό δίκτυο με δύο επίπεδα, τα οποία εκτελούνται εντός Sequential container, ενώ το υπόλοιπο δίκτυο παραμένει ίδιο αλλά με περισσότερες παραμέτρους σε κάθε layer.

Η εκπαίδευση της αρχιτεκτονικής αυτής διαφέρει από την εκπαίδευση του WGANP στο ότι προστίθεται η διαχείριση των κλάσεων (labels). Σε κάθε εποχή μαζί με τις πραγματικές εικόνες λαμβάνεται ως είσοδος και τα labels των εικόνων αυτών. Τα labels χρησιμοποιούνται τόσο στον Generator όσο και στον Discriminator. Με τον τρόπο αυτό επιτυγχάνεται σταθερότερη εκπαίδευση με τη χρήση της συνάρτησης κόστους Wasserstein, η εκπαίδευση του μοντέλου διακρίνει τα διαφορετικά επίπεδα του εγκεφάλου και στο τέλος είναι εφικτός ο έλεγχος του τύπου επιπέδου των παραγόμενων εικόνων.

3.2.3. LSGAN

Η συγκεκριμένη αρχιτεκτονική δοκιμάστηκε λόγω των εμφανώς ποιοτικών αποτελεσμάτων που έχουν παρουσιαστεί σε διάφορες μελέτες [45] [46]. Αυτή η αρχιτεκτονική διαφέρει αρκετά από τις άλλες. Αποτελείται από Upsample layers όπου πραγματοποιούν τον διπλασιασμό των Pixels των εικόνων, από Conv2d layers (συνελκτικά), BatchNorm2d και LeakyReLU. Οι εκπαιδευσιμοι παράμετροι παρουσιάζονται στους πίνακες 3-7 και 3-8.

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Linear	13,238,272
1	Conv2d	147,584
2	BatchNorm2d	256
3	Conv2d	73,792
4	BatchNorm2d	128
5	Conv2d	577
	Σύνολο	13,460,609

3-7 Αρχιτεκτονική LSGAN, δομή Generator

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Conv2d	160
1	Conv2d	4,640
2	BatchNorm2d	64
3	Conv2d	18,496
4	BatchNorm2d	128
5	Conv2d	73,856
6	BatchNorm2d	256
7	Linear	8,193
	Σύνολο	105,793

3-8 Αρχιτεκτονική LSGAN, δομή Discriminator

Ο Generator σε αυτήν την αρχιτεκτονική αποτελείται από ένα πλήρως συνδεδεμένο layer το οποίο μετατρέπει το θόρυβο σε μια συνδεδεμένη προβολή, ώστε να χρησιμοποιηθεί σαν είσοδος στο συνελικτικό τμήμα του δικτύου. Το συνελικτικό τμήμα αποτελείται από 2 ομάδες με διαδοχικά layers Upsampling, Conv2d, BatchNorm2d και LeakyReLU. Τις 2 ομάδες ακολουθεί ένα Conv2d layer με συνάρτηση ενεργοποίησης Tanh, για την παραγωγή της τελικής εικόνας εξόδου με διαστάσεις 1x128x128.

Ο Discriminator αποτελείται επίσης από ένα συνελικτικό νευρωνικό δίκτυο. με διαδοχικές συνελίξεις στο πρώτο τμήμα (Conv2d layer), BatchNorm2d και LeakyReLU, με σκοπό τη μείωση των διαστάσεων της εικόνας εισόδου. Τέλος, την έξοδο του συνελικτικού τμήματος υποδέχεται ένα γραμμικό επίπεδο (Linear layer) που έχει ως έξοδο την πιθανότητα της εικόνας να είναι πραγματική ή τεχνητή.

3.2.4. CGAN

Η αρχιτεκτονική Conditional GAN, δημιουργήθηκε σαν προσπάθεια παραγωγής εικόνων διαφορετικού είδους τομών. Όπως αναφέρθηκε και στην ενότητα του WGANP, δημιουργήθηκαν 3 διαφορετικά μοντέλα για κάθε είδος τομής. Η αρχιτεκτονική του δικτύου αυτού είναι τελείως διαφορετική από τα προηγούμενα καθώς δεν εφαρμόζονται συνελκτικά επίπεδα. Η δομή αποτελείται από fully connected layers και παρουσιάζεται στους πίνακες 3-9 και 3-10, για τον Generator και Discriminator αντίστοιχα.

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Embedding	9
1	Linear	13,312
2	Linear	33,024
3	BatchNorm1d	512
4	Linear	131,584
5	BatchNorm1d	1,024
6	Linear	525,312
7	BatchNorm1d	2,048
8	Linear	16,793,600
	Σύνολο	17,500,425

3-9 Αρχιτεκτονική CGAN, Generator

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Embedding	9
1	Linear	8,390,656
2	Linear	262,656
3	Linear	262,656
4	Linear	513
	Σύνολο	8,916,490

3-10 Αρχιτεκτονική CGAN, Discriminator

Ο Generator όμοια με το CWGAN έχει 2 μοντέλα, ένα το οποίο διαχειρίζεται τις κλάσεις (labels) και ένα που διαχειρίζεται και μαθαίνει τις εικόνες. Αρχικά, την είσοδο τη λαμβάνει ένα Embedding layer, και ακολουθούν 5 «blocks» που περιέχουν Linear, BatchNorm1d και LeakyReLU, εκτός από το πρώτο που δεν έχει BatchNorm1d και το τελευταίο που δεν έχει ούτε BatchNorm1d ούτε LeakyReLU.

Ο Discriminator αποτελείται από 2 μοντέλα όπως και ο Generator, ένα για τα labels και ένα για την κατηγοριοποίηση των εικόνων σε πραγματικές και μη. Πρώτα εφαρμόζεται ένα

Embedding layer για τη διαχείριση των labels και ακολουθούν 4 Linear με συνάρτηση ενεργοποίησης LeakyReLU, εκτός από το τελευταίο το οποίο καταλήγει σε έξοδο ενός καναλιού όπου αναδεικνύεται αν η εικόνα είναι πραγματική ή τεχνητή. Επίσης, τα δύο ενδιάμεσα Linear layers ακολουθούνται από Dropout με πιθανότητα 0.4.

Στην αρχή κάθε εποχής δημιουργούνται 2 tensors, ένα για τις πραγματικές εικόνες με τιμές 1 και ένα για την τεχνητές εικόνες με τιμές 0, (0 ή $1 \times \text{batch_size}$). Χρησιμοποιήθηκε η συνάρτηση απώλειας Mean Cross-Entropy Loss (MCELoss). Η εκπαίδευση του δικτύου πραγματοποιείται διαδοχικά, με πρώτα να εκπαιδεύεται ο Generator, δίνοντας ως είσοδο τυχαίο θόρυβο με 100 κανάλια ($\text{batch_size} \times \text{noise_channels}$) και τα αντίστοιχα labels των επιπέδων του εγκεφάλου των πραγματικών εικόνων, που έχουν επιλεγεί για την εκάστοτε εποχή. Ο Generator παράγει εικόνες οι οποίες εισάγονται στον Discriminator, μαζί με το tensor για τις τεχνητές εικόνες και με βάση τα αποτελέσματα ενημερώνονται τα βάρη του Generator. Έπειτα, ακολουθεί η εκπαίδευση του Discriminator η οποία πραγματοποιείται σε 2 βήματα. Πρώτα λαμβάνει ως είσοδο τεχνητές εικόνες και υπολογίζεται η απώλεια. Στη συνέχεια το ίδιο γίνεται και για τις πραγματικές εικόνες. Για την ενημέρωση των βαρών του Discriminator λαμβάνεται η μέση τιμή των δύο απωλειών που υπολογίστηκαν.

3.2.5. BEGAN

Η αρχιτεκτονική Boundary equilibrium GAN (BEGAN) έχει ενδιαφέρον, διότι χρησιμοποιούνται auto-encoders και η εκπαίδευση γίνεται υπολογίζοντας τη διαφορά των ανακατασκευασμένων εικόνων, πραγματικών και τεχνητών. Οι εκπαιδευσιμοι παράμετροι της αρχιτεκτονικής παρουσιάζονται στους πίνακες 3-11 και 3-12, για τον Generator και τον Discriminator αντίστοιχα.

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Linear	3,309,568
1	BatchNorm2d	256
2	Conv2d	147,584
3	BatchNorm2d	256
4	Conv2d	73,792
5	BatchNorm2d	128
6	Conv2d	577
	Σύνολο	3,532,161

3-11 Αρχιτεκτονική BEGAN, δομή Generator

	Όνομα Επιπέδου	Αριθμός Παραμέτρων
0	Conv2d	640
1	Linear	2,097,184
2	BatchNorm1d	64
3	Linear	2,162,688
4	BatchNorm1d	131,072
5	Conv2d	577
	Σύνολο	4,392,225

3-12 Αρχιτεκτονική BEGAN, δομή Discriminator

Ο Generator δέχεται ως είσοδο ένα πίνακα τυχαίου θορύβου 100 καναλιών, το οποίο εισάγεται στο πρώτο επίπεδο που είναι ένα πλήρως συνδεδεμένο layer με σκοπό τη δημιουργία μιας αρχικής αναπαράστασης, δηλαδή ενός πίνακα 128x256 κι ακολουθεί ένα επίπεδο κανονικοποίησης (BatchNorm2d). Έπειτα, η έξοδος εισάγεται σε 2 «blocks» τα οποία περιλαμβάνουν Upsample, Conv2d, BatchNorm2d και LeakyReLU. Τέλος, ακολουθεί ένα συνελικτικό δίκτυο το οποίο εξάγει μια εικόνα με διαστάσεις 1x64x64 και έχει συνάρτηση ενεργοποίησης Tanh. Στην υλοποίηση έχει τροποποιηθεί η συνάρτηση forward, ώστε να λαμβάνει την εικόνα, να την εισάγει στο πρώτο γραμμικό επίπεδο και να μετασχηματίζει κατάλληλα το αποτέλεσμα με σκοπό να το εισάγει στα συνελικτικά επίπεδα.

Ο Discriminator δέχεται εικόνες πραγματικές και τεχνητές ως είσοδο στο πρώτο συνελικτικό επίπεδο το οποίο μειώνει τις διαστάσεις της αρχικής εικόνας εισόδου. Έπειτα, ακολουθούν 2 «blocks» που αποτελούνται από Linear, BatchNorm1d και ReLU επίπεδα. Η έξοδος αυτών εισάγεται σε ένα Upsample επίπεδο, το οποίο αναλαμβάνει την αύξηση των διαστάσεων, ώστε να συνεχίσει στο συνελικτικό επίπεδο που αναλαμβάνει τη

διαμόρφωση των τελικών διαστάσεων της εικόνας να είναι ίδια με την αρχική εικόνα εισόδου. Παρόμοια με τον Generator και στο δίκτυο του Discriminator έχει τροποποιηθεί η forward μέθοδος, ώστε να λαμβάνει την είσοδο και να την εισάγει στο πρώτο τμήμα που πραγματοποιεί τη μείωση των διαστάσεων, έπειτα πραγματοποιείται ο μετασχηματισμός της εξόδου, με σκοπό να χρησιμοποιηθεί στο τμήμα των γραμμικών επιπέδων (πλήρως συνδεδεμένων). Τέλος, η έξοδος μετασχηματίζεται στις κατάλληλες διαστάσεις με σκοπό να δοθεί στο τελευταίο τμήμα όπου θα πραγματοποιηθεί η αύξηση των διαστάσεων για να επαναφέρει την εικόνα στις αρχικές διαστάσεις.

Η εκπαίδευση πραγματοποιείται σε εποχές, όπου σε κάθε εποχή επεξεργάζεται ένα μέρος των πραγματικών εικόνων. Πρώτα, εκπαιδεύεται ο Generator, όπου παράγεται τυχαίος θόρυβος διαστάσεων $batch_size \times 100$ ο οποίος δίνεται στον Generator, για να παράξει τεχνητές εικόνες διαστάσεων $1 \times 64 \times 64$. Υπολογίζεται η απώλεια με βάση την ικανότητα να «ξεγελάσει» τον Discriminator, συγκρίνοντας τις πραγματικές με τις τεχνητές εικόνες, λαμβάνοντας υπόψη την υπερπαραμέτρο K , $loss_real - k * loss_fake$. Έπειτα, ακολουθεί η εκπαίδευση του Discriminator, ο οποίος καλείται 2 φορές, μια για τις πραγματικές και μια για τις τεχνητές εικόνες. Έχοντας τα αποτελέσματα του Discriminator υπολογίζονται οι απώλειες ως, $mean(abs(discriminator_output - images))$, όπου $discriminator_output$ είναι η ανακατασκευή των εικόνων και $images$ είναι οι εικόνες που έλαβε ως είσοδο. Στη συνέχεια υπολογίζεται η απώλεια για την ενημέρωση των βαρών του Discriminator ως: $loss_real - k * loss_fake$, όπου $loss_real$ είναι η απώλεια ανακατασκευής των πραγματικών εικόνων και $loss_fake$ είναι η απώλεια ανακατασκευής των τεχνητών εικόνων. Τέλος, πραγματοποιείται η ενημέρωση των υπερπαραμέτρων k και M ως, $k + lambda_k * diff$ και $(loss_real + abs(diff))$, με $diff = mean(gamma * loss_real - loss_fake)$, όπου $gamma$ και $lambda_k$ σταθερές υπερπαραμέτροι. Η παράμετρος M είναι μετρική σύγκλισης, δηλαδή βοηθάει στην παρατήρηση της εξέλιξης της εκπαίδευσης και στο αν η ποιότητα των παραγόμενων εικόνων βελτιώνεται, αναμένοντας μικρότερες τιμές πιο κοντά στο μηδέν, όταν οι παραγόμενες εικόνες είναι ποιοτικές, δηλαδή όμοιες με τις πραγματικές

4. Πειραματικά αποτελέσματα

Σε αυτή την ενότητα θα μελετηθούν τα δεδομένα, οι ρυθμίσεις των υπερπαραμέτρων, οι αρχιτεκτονικές GANs που υλοποιήθηκαν, τα πειραματικά αποτελέσματα τόσο των GAN αρχιτεκτονικών όσο και του ταξινομητή και των αποδόσεών του με διαφορετικά ποσοστά τεχνητών εικόνων. Τα πειράματα πραγματοποιήθηκαν στα 3 διαφορετικά σύνολα δεδομένων που αναφέρθηκαν στην αντίστοιχη ενότητα.

Σημαντικές πληροφορίες σχετικά με την ανάγνωση των αποτελεσμάτων, τα γραφήματα με τις απώλειες στον οριζόντιο άξονα περιλαμβάνουν τα βήματα και όχι τις εποχές. Μια εποχή μπορεί να περιέχει 2-4 βήματα αναλόγως το σύνολο δεδομένων και τις ανάγκες του κάθε πειράματος (π.χ. τότε θέλουμε να αποθηκεύονται τα βάρη). Για τα αποτελέσματα ταξινόμησης έχει χρησιμοποιηθεί σε όλα τα πειράματα, ένα CNN Resnet 18 επιπέδων, προεκπαιδευμένο στο σύνολο δεδομένων ImageNet [47]. Για την εξειδίκευση του ταξινομητή στο επιλεγμένο σύνολο δεδομένων, πραγματοποιήθηκε η εκπαίδευση του μοντέλου έχοντας «παγώσει» όλα τα επίπεδα εκτός του τελευταίου. Το τελευταίο επίπεδο είναι ένα Linear Layer με έξοδο 1000 κλάσεων, σύμφωνα με το σύνολο ImageNet, κάτι το οποίο τροποποιήθηκε σε 2 κλάσεις ώστε να αντιστοιχεί στον αριθμό των κλάσεων του προβλήματος που μελετάται

Θα χρησιμοποιηθούν οι εξής όροι για την παρουσίαση των αποτελεσμάτων:

- CHANNEL_IMG: Ο αριθμός των καναλιών των εικόνων. (Χρησιμοποιούνται 1 και 3 κανάλια)
- IMAGE_SIZE: Το μέγεθος της εικόνας που εισάγεται ως είσοδος κατά τη διαδικασία εκπαίδευσης, οι εικόνες είναι τετράγωνες επομένως ο αριθμός δείχνει και τις δύο διαστάσεις.
- BATCH_SIZE: Το μέγεθος της παρτίδας των εικόνων που τροφοδοτείται κάθε εποχή κατά την εκπαίδευση.
- Z_DIM: Είναι τα κανάλια του τυχαίου θορύβου που χρησιμοποιείται για την εκπαίδευση του Generator.

- CRITIC_ITERATIONS: Αφορά τις αρχιτεκτονικές GAN και είναι ο αριθμός των επαναλήψεων που εκπαιδεύεται ο Discriminator για 1 επανάληψη του Generator.
- LAMBDA_GP: Υπερπαράμετρος της αρχιτεκτονικής WGAN με gradient penalty που ελέγχει την επιρροή του gradient penalty στον υπολογισμό της Wasserstein απόστασης.

4.1. Δεδομένα

Τα δεδομένα που χρησιμοποιήθηκαν για τη διεξαγωγή των πειραμάτων έχουν συλλεχθεί από το σύνολο δεδομένων Open Access Series of Imaging Studies (OASIS) και συγκεκριμένα από το OASIS-3 [23]. Το έργο OASIS brain αποτελεί μια προσπάθεια συλλογής εικόνων MRI και PET αλλά και των σχετικών κλινικών δεδομένων από διαφορετικές έρευνες που διεξήχθησαν σε διάρκεια 15 ετών από συνολικά 1379 συμμετέχοντες, στο ερευνητικό κέντρο Knight Alzheimer Disease Research Center (Knight ADRC), του τμήματος νευρολογίας του πανεπιστημίου της Washington. Το σύνολο δεδομένων αποτελείται από 755 νοητικά φυσιολογικούς ενήλικες και 622 ασθενείς σε διάφορα στάδια γνωστικής εξασθένησης σε ηλικίες από 42 μέχρι και 95 χρόνων. Έχουν συλλεχθεί συνολικά 2842 MRI συνεδρίες, 2157 PET συνεδρίες και 1472 κλινικές δοκιμές. Οι συνεδρίες MRI περιέχουν τομές T1-weighted (T1w), T2-weighted (T2w), FLAIR, Arterial Spin Labeling (ASL), Susceptibility Weighted Imaging (SWI), καθώς και επεξεργασμένες τομές με το πρόγραμμα FreeSurfer [48].

Για τη δημιουργία των συνόλων δεδομένων χρησιμοποιήθηκε το πρόγραμμα που παρέχεται στην αποθήκη <https://github.com/NrgXnat/oasis-scripts>, αφού πρώτα επιλέχθηκαν δύο λίστες δειγμάτων (σε μορφή csv) με τις κωδικές ονομασίες των MRI συνεδριών προς λήψη. Οι λίστες αυτές δημιουργήθηκαν από ένα πρόγραμμα το οποίο έχει σαν είσοδο δύο csv αρχεία, ένα με τους κωδικούς και τις τιμές CDR (Clinical Dementia Rating) των κλινικών δοκιμών και ένα με τους κωδικούς των MRI συνεδριών για κάθε ασθενή. Για τον λόγο ότι δεν υπήρχε αντιστοιχία μεταξύ των κλινικών δοκιμών και των MRI συνεδριών, ο διαχωρισμός πραγματοποιήθηκε βάσει κάποιων κανόνων. Στο σημείο αυτό αξίζει να διευκρινιστεί η χρήση της μετρικής CDR [49], η κλινική βαθμολόγηση της

άνοιας (Clinical Dementia Rating, CDR) είναι μια παγκοσμίως χρησιμοποιούμενη κλίμακα για την αξιολόγηση του επιπέδου της άνοιας λόγω της νόσου Αλτσχάιμερ, προκύπτει ύστερα από νοητικές και συμπεριφορικές αξιολογήσεις, και η κλίμακα εξηγείται στον πίνακα 4-1.

Τιμή CDR	Επεξήγηση
0	Χωρίς άνοια / No dementia
0.5	Υπό αμφισβήτηση / Questionable dementia
1	Ήπια διαταραχή / Mild cognitive impairment
2	Σοβαρή διαταραχή / Severe cognitive impairment

4-1 Επεξήγηση κλίμακας CDR

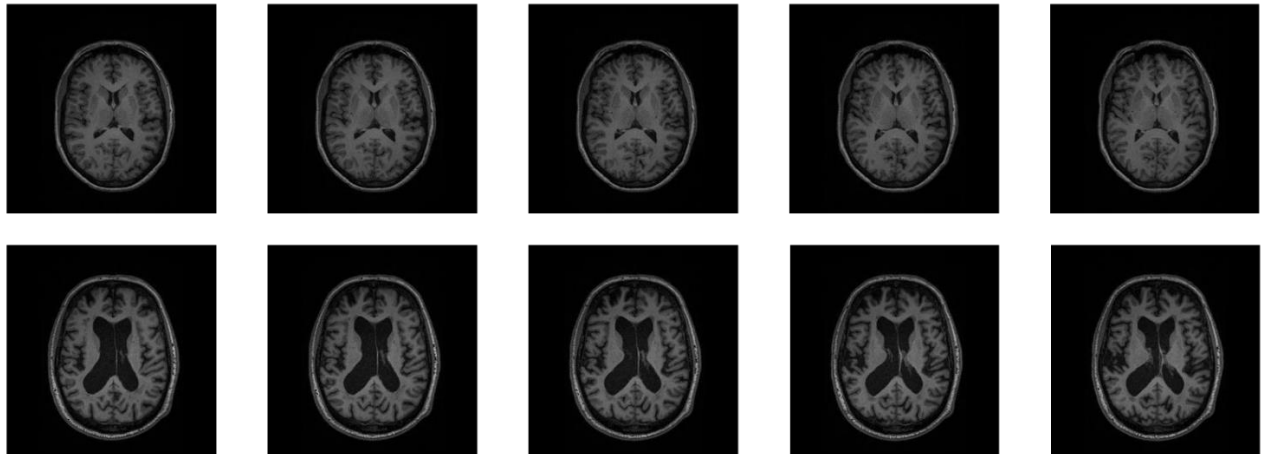
Ο διαχωρισμός έγινε σε δύο κλάσεις, σε δείγματα χωρίς άνοια (NC) και σε δείγματα με άνοια (AD). Επίσης, πραγματοποιήθηκαν δύο μεθοδολογίες για την επιλογή κλάσης. Στην πρώτη (Μέθοδος Επιλογής Κλάσης 1, **MEK1**) ταξινομούνται στην NC κλάση δείγματα που στις κλινικές δοκιμές είχαν CDR=0 και στην AD κλάση δείγματα με CDR>=0.5. Στη δεύτερη μέθοδο (Μέθοδος Επιλογής Κλάσης 2, **MEK2**) ταξινομήθηκαν δείγματα με CDR=0 και των οποίων η τιμή CDR παρέμεινε 0 καθ' όλη τη διάρκεια της έρευνας (ως το σημείο που πραγματοποιήθηκε η λήψη του συνόλου δεδομένων, καθώς η έρευνα συνεχίζει και ενημερώνεται), αντίθετα στην κλάση AD επιλέχθηκαν δείγματα με τιμή CDR>=2. Η MEK2 πραγματοποιήθηκε για να μειωθεί η υποψία «επιρροής» του ταξινομητή λόγω πιθανών ομοιοτήτων μεταξύ CDR 0, 0.5 και 1, καθώς και για τον λόγο ότι τα δείγματα που εμφάνισαν στην πορεία της έρευνας συμπτωματολογία της νόσου ενδεχομένως να υπήρχε δομική αλλοίωση και σε μαγνητικές πριν την αλλαγή της τιμής CDR. Να σημειωθεί πως το τελευταίο αποτελεί προσωπική παρατήρηση/υπόθεση βάσει των πειραμάτων που διεξήχθησαν. Όπως προαναφέρθηκε υπήρξε πρόβλημα με την αντιστοίχιση των εγγραφών των κλινικών δοκιμών και των συνεδριών MRI, δηλαδή όταν ένας ασθενής πραγματοποίησε κλινική δοκιμή δεν είναι απαραίτητο ότι ταυτόχρονα (ή σε κοντινό διάστημα) διεξήχθη και μαγνητική τομογραφία. Για την αντιστοίχιση μιας MRI με μια τιμή CDR, χρησιμοποιήθηκε η μεταβλητή d##### που υπάρχει στον κωδικό της κάθε εγγραφής, η οποία ερμηνεύεται ως η διάρκεια σε ημέρες που πραγματοποιήθηκε η συνεδρία/δοκιμή από την πρώτη εμφάνιση του δείγματος στο σύνολο δεδομένων (τιμή d0000 για την πρώτη εμφάνιση, τιμή d0100 για επανεμφάνιση μετά από 100 μέρες).

Επομένως, η συνεδρία MRI θα αντιστοιχισθεί με την τιμή CDR που είναι πιο κοντά της, χωρίζοντας τα διαστήματα στη μέση.

Όπως έγινε επιλογή δύο μεθοδολογιών για τον διαχωρισμό των δειγμάτων σε κλάσεις, έτσι επιλέχθηκαν και δύο μεθοδολογίες βάσει του τύπου των εικόνων. Το πρώτο σύνολο δεδομένων (Σύνολο Δεδομένων 1, **ΣΔ1**), δημιουργήθηκε από τις οριζόντιες τομές (axial) των T1w MRI χωρίς προεπεξεργασία. Τα αρχεία των MRI τομών ήταν σε μορφή .nii NIfTI (Neuroimaging Informatics Technology Initiative), ένα αρχείο τριών διαστάσεων που περιλάμβανε τομές απ' όλα τα επίπεδα του εγκεφάλου. Οι διαστάσεις των αρχείων ήταν 176x256x256, όπου 176 είναι ο αριθμός των τομών ανά επίπεδο (επίπεδα είναι Axial, Coronal και Sagittal) και οι δύο αριθμοί 256x256 είναι το μέγεθος της εικόνας. Επιλέχθηκαν οι 30 κεντρικές τομές δηλαδή από την τομή 61 έως την 91, με σκοπό να περιοριστεί η περιττή πληροφορία όπως άλλα χαρακτηριστικά του προσώπου, καθώς επίσης και να μεγιστοποιηθεί η πιθανότητα επιλογής τομών που υπάρχει εμφανές πρόβλημα [21]. Στον πίνακα 4-2 φαίνεται ο αριθμός των ατόμων ανά κλάση, ο αριθμός των συνεδριών, ο συνολικός αριθμός εικόνων (slices) που συλλέχθηκαν καθώς και η μέση τιμή και η τυπική απόκλιση των εικόνων. Ο διαχωρισμός των MRI σε κλάσεις πραγματοποιήθηκε σύμφωνα με τη μέθοδο **MEK1**, κι ένα δείγμα εικόνων του συνόλου παρουσιάζεται στο σχήμα 4-1.

	AD	NC
Αριθμός ατόμων	294	667
Αριθμός συνεδριών	339	1143
ΜΟ συνεδριών ανά άτομο	1.15	1.71
ΜΟ εικόνων ανά άτομο	53	78
Συνολικός αριθμός εικόνων	15570	52110
Images mean	0.0863	0.0892
Images std	0.1285	0.1333

4-2 Στατιστικά του Συνόλου Δεδομένων 1 ΣΔ1 με μέθοδο MEK1



4-1 Δείγμα κεντρικών τομών ΣΔ1, πάνω σειρά κλάση NC κάτω σειρά κλάση AD

Το δεύτερο σύνολο δεδομένων (Σύνολο Δεδομένων 2, **ΣΔ2**) ακολουθεί τους ίδιους κανόνες επιλογής εικόνων, αλλά ο διαχωρισμός των MRI σε κλάσεις έγινε σύμφωνα με τη μέθοδο **MEK2**. Τα στατιστικά χαρακτηριστικά του συνόλου παρουσιάζονται στον πίνακα 4-3, όπου είναι φανερή η μείωση του δείγματος λόγω των επιπρόσθετων περιορισμών που προαναφέρθηκαν. Τα μορφολογικά χαρακτηριστικά των εικόνων είναι παρόμοια με το δείγμα στο σχήμα 4-1.

	AD	NC
Αριθμός ατόμων	70	504
Αριθμός συνεδριών	76	911
ΜΟ συνεδριών ανά άτομο	1.10	1.81
ΜΟ εικόνων ανά άτομο	53	81
Συνολικός αριθμός εικόνων	3690	40620
Images mean	0.0850	0.0894
Images std	0.1261	0.1335

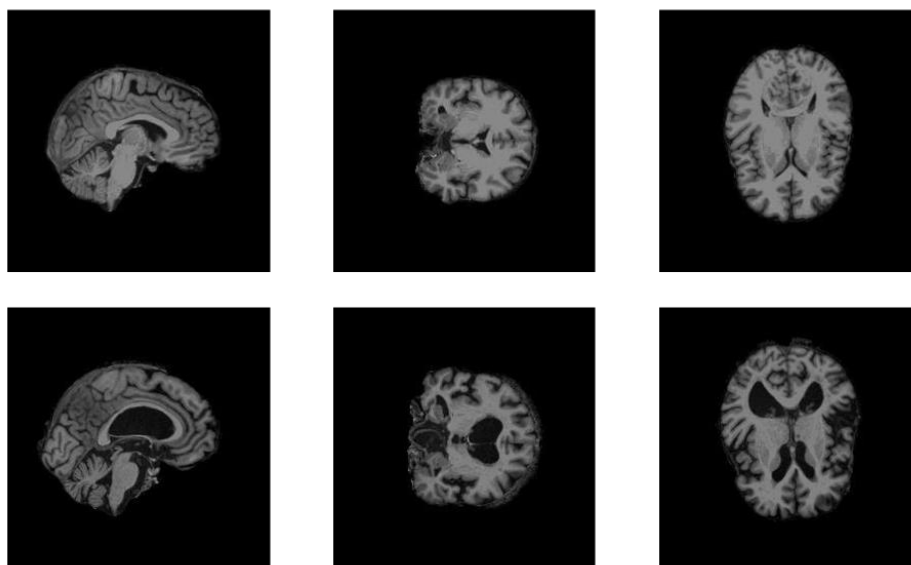
4-3 Στατιστικά του Συνόλου Δεδομένων 2 ΣΔ2 με μέθοδο MEK2

Τέλος δημιουργήθηκε ένα τρίτο σύνολο δεδομένων (Σύνολο Δεδομένων 3, **ΣΔ3**) που αποτελεί τη δεύτερη μεθοδολογία για τη δημιουργία των συνόλων δεδομένων. Δημιουργήθηκε παίρνοντας τομές (slices) και από τα τρία επίπεδα (Axial, Coronal, Sagittal). Τα αρχεία που χρησιμοποιήθηκαν σε αυτήν τη διαδικασία είναι προεπεξεργασμένες MRI με το λογισμικό FreeSurfer [48]. Μερικά στάδια της προεπεξεργασίας είναι η αφαίρεση ιστού που δεν αποτελεί μέρος του εγκεφάλου (non-brain tissue), διόρθωση κίνησης και υπολογισμό μέσου όρου των ογκομετρικών T1w εικόνων, αφαίρεση του κρανίου και χαρακτηριστικών όπως μάτια και μύτη (skull strip) κλπ

[50]. Εδώ επιλέχθηκαν 10 κεντρικές τομές από κάθε επίπεδο, δηλαδή σύνολο 30 τομές (slices). Ο διαχωρισμός των MRI σε κλάσεις πραγματοποιήθηκε με τη μέθοδο MEK1. Στατιστικά χαρακτηριστικά του συνόλου αυτού παρουσιάζονται στον πίνακα 4-4, ενώ ένα δείγμα εικόνων παρουσιάζεται στο σχήμα 4-2.

	AD	NC
Αριθμός ατόμων	343	749
Αριθμός συνεδριών	425	1838
ΜΟ συνεδριών ανά άτομο	1.24	2.45
ΜΟ εικόνων ανά άτομο	37	74
Συνολικός αριθμός εικόνων	12750	55140
Images mean	0.0746	0.0765
Images std	0.1557	0.1580

4-4 Στατιστικά του Συνόλου Δεδομένων 3 ΣΔ3 με μέθοδο MEK1



4-2 Δείγμα κεντρικών τομών του ΣΔ3, πάνω σειρά υγής, κάτω σειρά ασθενής.

Για την αμερόληπτη αξιολόγηση των διαφορετικών πειραμάτων πραγματοποιήθηκε ο διαχωρισμός του κάθε συνόλου δεδομένων εκ των προτέρων και παρέμεινε σταθερό για όλα τα πειράματα. Κάθε σύνολο δεδομένων (ΣΔ1, ΣΔ2, ΣΔ3) διαχωρίστηκε σε σύνολα εκπαίδευσης (train), επαλήθευσης (validation) και δοκιμής (test) με ποσοστά 70%, 15% και 15% αντίστοιχα. Σε κάποιες περιπτώσεις θα παρατηρηθεί διαφορά στον αριθμό των εικόνων μεταξύ validation και test συνόλων, αυτό θα συμβεί διότι ο διαχωρισμός έγινε σε επίπεδο ασθενή και όχι σε επίπεδο εικόνων ή συνεδριών, καθώς όπως φαίνεται και στους πίνακες 4-2, 4-3 και 4-4, δεν έχουν όλοι οι ασθενείς ίδιο αριθμό επισκέψεων άρα και

εικόνων, με αυτόν τον τρόπο διασφαλίστηκε η παρουσία ενός ασθενή μόνο σε ένα υποσύνολο και η αποφυγή της διαρροής δεδομένων (data leakage). Σε αυτό το στάδιο δεν αφαιρέθηκαν οι επιπλέον διαθέσιμες εικόνες από την κλάση NC, αλλά όπως θα δούμε στην ενότητα 4, εκτελέστηκαν πειράματα στα οποία έγινε εξισορρόπηση των κλάσεων αφαιρώντας τυχαία τις επιπλέον εικόνες από την κλάση NC. Στους πίνακες 4-5, 4-6 και 4-7 παρουσιάζονται τα στατιστικά των διαχωρισμένων συνόλων δεδομένων για τα ΣΔ1, ΣΔ2 και ΣΔ3 αντίστοιχα.

	train		validation		test	
	AD	NC	AD	NC	AD	NC
Αριθμός ατόμων	206	467	43	99	45	101
Αριθμός συνεδριών	239	803	50	165	50	175
ΜΟ συνεδριών ανά άτομο	1.16	1.71	1.16	1.66	1.11	1.73
ΜΟ εικόνων ανά άτομο	54	78	50	76	51	81
Συνολικός αριθμός εικόνων	11160	36450	2130	750	2280	8160
Images mean	0.0867	0.0891	0.0848	0.0889	0.0857	0.0899
Images std	0.1287	0.1330	0.1264	0.1332	0.1290	0.1351

4-5 Στατιστικά διαχωρισμένου (splitted) συνόλου δεδομένων ΣΔ1

	train		validation		test	
	AD	NC	AD	NC	AD	NC
Αριθμός ατόμων	48	351	12	82	10	71
Αριθμός συνεδριών	54	631	12	144	10	136
ΜΟ συνεδριών ανά άτομο	1.13	1.80	1.00	1.76	1.00	1.92
ΜΟ εικόνων ανά άτομο	54	81	46	74	55	86
Συνολικός αριθμός εικόνων	2583	28434	553	6092	554	6094
Images mean	0.0846	0.0895	0.0868	0.0904	0.0848	0.0875
Images std	0.1262	0.1339	0.1255	0.1349	0.1262	0.1306

4-6 Στατιστικά διαχωρισμένου (splitted) συνόλου δεδομένων ΣΔ2

	train		validation		test	
	AD	NC	AD	NC	AD	NC
Αριθμός ατόμων	241	525	50	111	52	113
Αριθμός συνεδριών	306	1309	58	258	61	271
ΜΟ συνεδριών ανά άτομο	38	75	35	70	35	72
ΜΟ εικόνων ανά άτομο	1.27	2.50	1.16	2.32	1.17	2.40
Συνολικός αριθμός εικόνων	9180	39270	1740	7740	1830	8130
Images mean	0.0747	0.0763	0.0724	0.0776	0.0758	0.0762
Images std	0.1564	0.1580	0.1515	0.1596	0.1561	0.1562

4-7 Στατιστικά διαχωρισμένου (splitted) συνόλου δεδομένων ΣΔ3

4.2. Δομή και τρόπος αξιολόγησης ταξινομητών

Η επιλογή ταξινομητή και η εύρεση του κατάλληλου για το συγκεκριμένο πρόβλημα δεν αποτελεί το βασικό ζήτημα αυτής της εργασίας, αλλά το ενδιαφέρον επικεντρώνεται περισσότερο στη βελτίωση των μετρικών του ταξινομητή μέσω της προσθήκης τεχνητών εικόνων παραγόμενων από τα GAN. Για τον λόγο αυτό, ύστερα από εύρεση κάποιων αρχιτεκτονικών που αποδίδουν καλά σε εικόνες MRI εγκεφάλων [21] [10] [51], επιλέχθηκε να πραγματοποιηθούν αρχικές δοκιμές στις αρχιτεκτονικές Densenet [52] και Resnet [24]. Στις αρχικές δοκιμές, επίσης, πραγματοποιήθηκε σύγκριση προεκπαιδευμένων μοντέλων (στο σύνολο εικόνων ImageNet) και μοντέλων χωρίς χρήση έτοιμων βαρών, που εκπαιδεύτηκαν απευθείας πάνω στα δεδομένα της εργασίας. Από τις δοκιμές αυτές παρατηρήθηκε μικρή διαφορά με καλύτερη απόδοση στα προεκπαιδευμένα μοντέλα. Επομένως τα βασικά πειράματα για την αξιολόγηση των αποτελεσμάτων με τεχνητές εικόνες πραγματοποιήθηκαν με τη χρήση προεκπαιδευμένων μοντέλων. Τελικώς επιλέχθηκε η αρχιτεκτονική Resnet18, δηλαδή ένα μοντέλο Deep Residual Networks των 18 επιπέδων, διότι από τη σύγκρισή του με το Densenet παρατηρήθηκε ότι απέδιδε ελάχιστα καλύτερα. Ωστόσο η σημαντικότερη διαφορά που οδήγησε στην επιλογή του, ήταν η ταχύτερη εκπαίδευση, που επιτυγχάνεται από τον μικρό αριθμό παραμέτρων, ένα σημαντικό πλεονέκτημα δεδομένου του περιορισμού τόσο σε χρόνο όσο και σε υπολογιστικούς πόρους.

Για την αξιολόγηση των μοντέλων με τεχνητές εικόνες πραγματοποιήθηκε η σταδιακή αύξηση του ποσοστού τεχνητών εικόνων μέχρι την εξισορρόπηση των δύο κλάσεων. Τα ποσοστά που δοκιμάστηκαν είναι τα εξής, 10%, 20%, 50%, 80% και 100%. Στο ποσοστό 100% περιέχονται μόνο τεχνητές εικόνες στη μειονοτική κλάση.

Τα δίκτυα εκπαιδεύτηκαν για 20 εποχές, κάτι που ίσως περιορίζει την επιτυχία, αλλά έγινε λόγω περιορισμού σε χρόνο και υπολογιστικούς πόρους. Για την προετοιμασία των εικόνων για την εκπαίδευση πραγματοποιήθηκαν οι μετατροπές που καθορίζονταν από το προεκπαιδευμένο μοντέλο και πρέπει να ακολουθηθούν για τη σωστή εκπαίδευση. Οι μετατροπές είναι οι εξής:

- Αλλαγή μεγέθους των πραγματικών εικόνων σε 128x128, ώστε να έχουν το ίδιο μέγεθος με τις τεχνητές εικόνες και να αντιμετωπίζονται με τον ίδιο τρόπο
- Αλλαγή του μεγέθους των εικόνων σε διαστάσεις 256x256
- Κεντρική περικοπή των εικόνων σε διάσταση 224x224
- Τυχαία οριζόντια αναστροφή, μετατροπή σε tensor
- Κανονικοποίηση της εικόνας βάσει των εξής μέσων τιμών και τυπικών αποκλίσεων:
 - Οι μέσες τιμές για κάθε κανάλι ορίζονται ως [0.485, 0.456, 0.406].
 - Οι τυπικές αποκλίσεις για κάθε κανάλι ορίζονται ως [0.229, 0.224, 0.225].

Η επιλογή του καλύτερου μοντέλου έγινε βάσει της απόδοσής του στο validation σύνολο και συγκεκριμένα με τη μετρική ισορροπημένης ακρίβειας (balanced accuracy BAC). Με άλλα λόγια, η εποχή που είχε την υψηλότερη τιμή BAC, χρησιμοποιήθηκε για τη δοκιμή και αξιολόγηση του μοντέλου στο test σύνολο. Η τιμή BAC επιλέχθηκε ως κριτήριο, δεδομένου του προβλήματος της επίλυσης της ανισορροπίας των κλάσεων και συγκεκριμένα της αδυναμίας καλής απόδοσης (αναγνώρισης) σε εικόνες της μειονοτικής κλάσης (minority class). Επίσης, δόθηκε σημασία στη μετρική recall, η οποία μας δείχνει το ποσοστό που ταξινομήθηκε σωστά από την κλάση των ασθενών, δηλαδή ατόμων με Alzheimer. Επιπλέον, καταγράφηκαν οι μετρικές Accuracy και precision, αλλά δεν προσφέρουν ιδιαίτερη πληροφορία, καθώς στο validation και στο test σύνολο για την προσομοίωση πραγματικών συνθηκών διατηρήθηκε η ανισορροπία των κλάσεων.

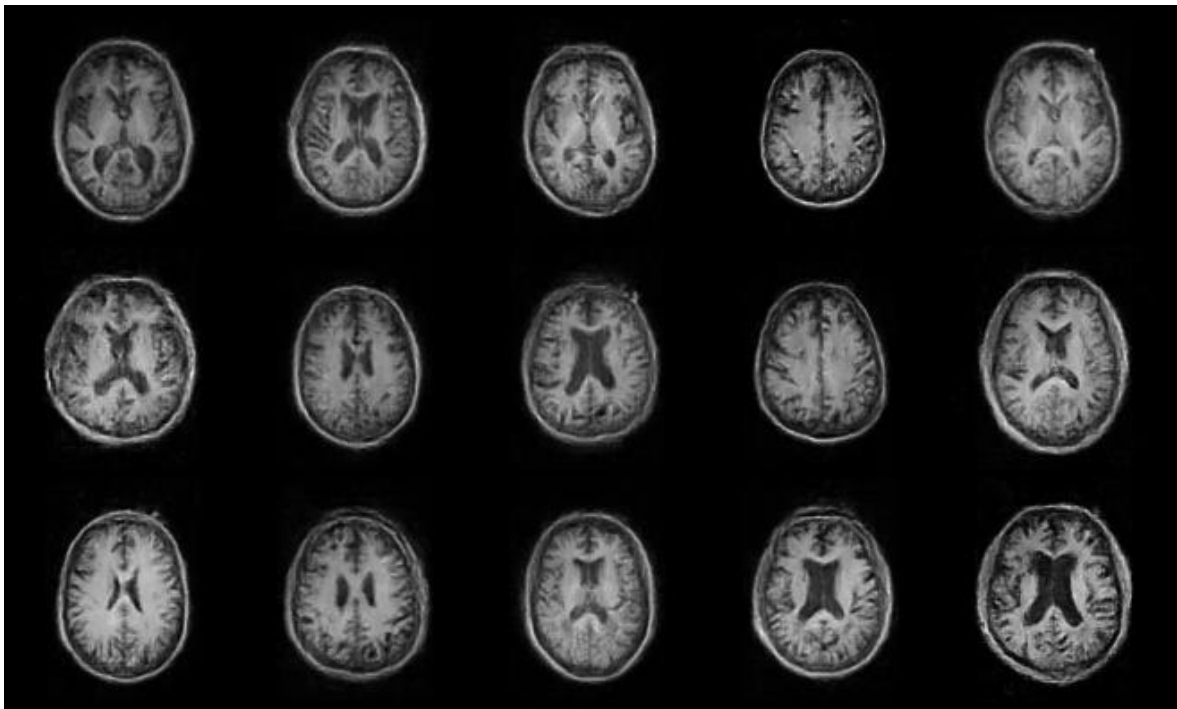
4.3. Αποτελέσματα τεχνητών εικόνων

Σε αυτή την ενότητα θα παρουσιαστούν τα αποτελέσματα των δοκιμών διαφορετικών αρχιτεκτονικών GAN, για την δημιουργία τεχνητών εικόνων. Κάθε υποενότητα αφορά ένα σύνολο δεδομένων και μια αρχιτεκτονική GAN και περιλαμβάνει εικόνες τεχνητών MRI κατά τη διάρκεια της εκπαίδευσης, τις παραμέτρους που χρησιμοποιήθηκαν και τα γραφήματα των απωλειών του Discriminator και του Generator. Ο στόχος της απώλειας είναι να προσεγγίσει το μηδέν, κάτι το οποίο είναι εμφανές και από τα αντίστοιχα

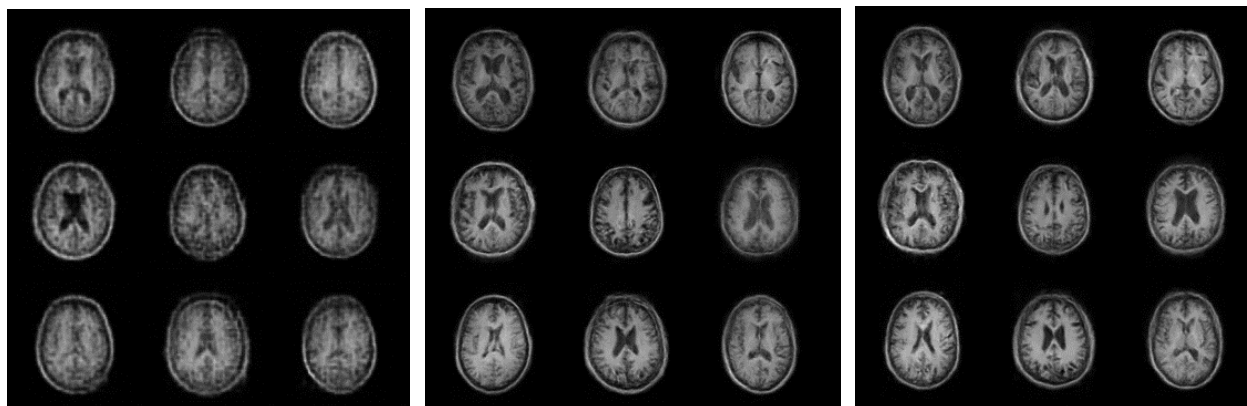
διαγράμματα. Η απώλεια στα αρχικά στάδια της εκπαίδευσης ξεκινάει από ένα τυχαίο σημείο σύμφωνα με τα τυχαία βάρη που έχουν οριστεί στο εκάστοτε δίκτυο. Στις περισσότερες περιπτώσεις παρατηρείται μεγάλη πτώση ως μια αρνητική τιμή στις πρώτες εποχές και έπειτα η σταδιακή άνοδος προς το μηδέν. Παίρνοντας ως παράδειγμα το WGAN, ο Discriminator προσπαθεί να μεγιστοποιήσει την απώλεια, δηλαδή την τιμή της Wasserstein απόστασης, ενώ αντίθετα ο Generator προσπαθεί να την ελαχιστοποιήσει. Για την υλοποίηση του κώδικα χρειάστηκε να αντιστραφεί το πρόσημο της απώλειας στον Generator, επομένως πλέον προσπαθεί να μεγιστοποιήσει την αρνητική απώλεια. Οι αρνητικές τιμές, λοιπόν, στα γραφήματα της απώλειας στα αρχικά στάδια της εκπαίδευσης εξηγούνται από την αδυναμία και των δύο δικτύων να αποδώσουν χωρίς να έχουν εκπαιδευτεί. Στα γραφήματα που παρουσιάζονται οι απώλειες των δικτύων, ο κάθετος άξονας Y δείχνει το μέγεθος της απώλειας, ενώ ο άξονας X τις επαναλήψεις (iterations) κατά τη διάρκεια της εκπαίδευσης του κάθε GAN.

4.3.1. WGANP 1, με σύνολο δεδομένων ΣΔ1

Στα σχήματα 4-3 και 4-4 παρουσιάζονται οι τεχνητές εικόνες που παράγει το μοντέλο σε διαφορετικές εποχές, με την 240 εποχή να είναι και η τελική που έφτασε η εκπαίδευση του μοντέλου με το συγκεκριμένο σύνολο δεδομένων. Στο σχήμα 4-5 φαίνεται η εξέλιξη των απωλειών του Generator και του Discriminator κατά τη διάρκεια της εκπαίδευσης. Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι υπερπαραμέτροι του πίνακα 4-8.



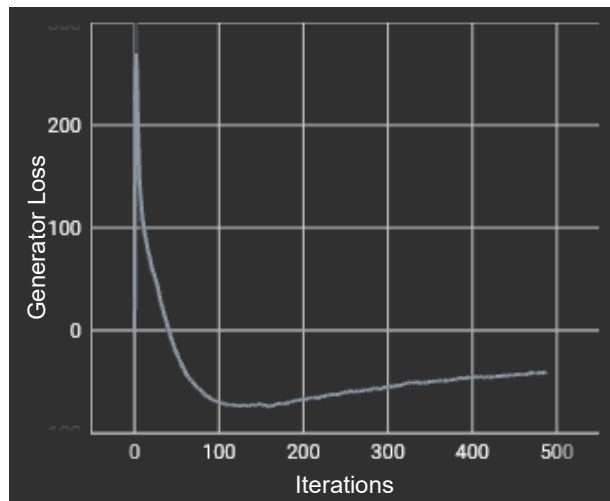
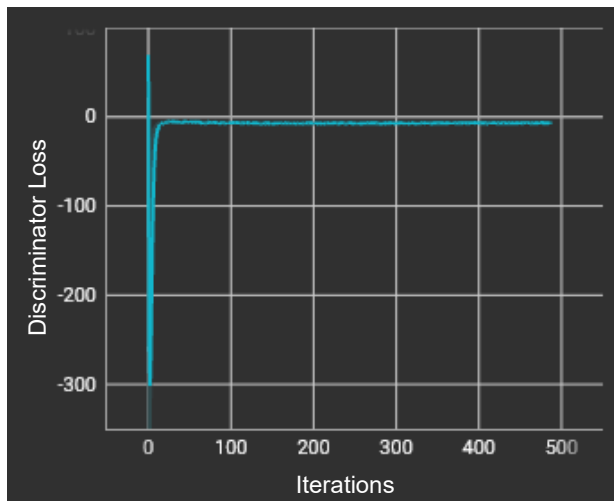
4-3 Τεχνητές εικόνες WGAN-GP 1 στο ΣΔ1 εποχή 240



4-4 Τεχνητές εικόνες WGAN-GP 1 στο ΣΔ1 εποχές 10, 50 και 150

LEARNING_RATE	0.0001
BATCH_SIZE	64
IMAGE_SIZE	128
CHANNELS_IMG	1
Z_DIM	256
CRITIC_ITERATIONS	10
LAMBDA_GP	10

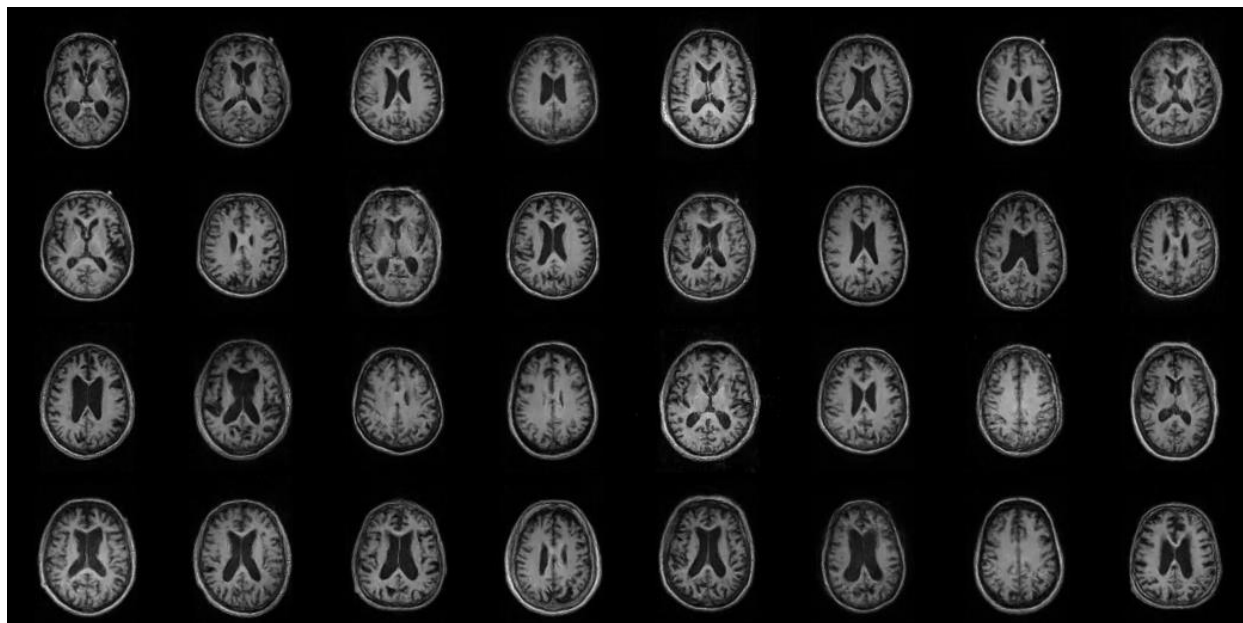
4-8 Υπερπαραμέτροι πειράματος 4.3.1



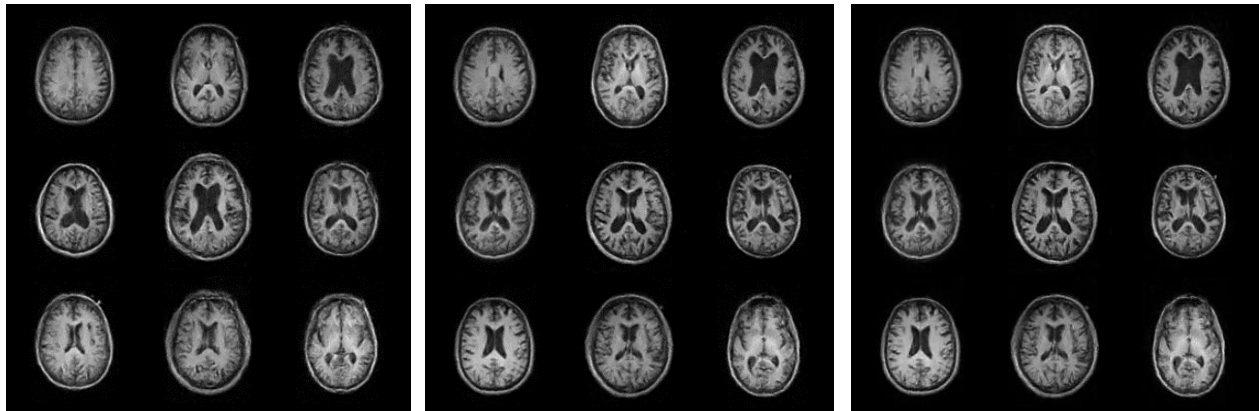
4-5 WGANGP, ΣΔ1 Discriminator Loss (αριστερά) και Generator Loss (δεξιά)

4.3.2. WGANGP 1, με σύνολο δεδομένων ΣΔ2

Στο σχήμα 4-6 παρουσιάζονται τα αποτελέσματα με τις παραγόμενες τεχνητές εικόνες από την τελευταία εποχή και στο σχήμα 4-7 αποτελέσματα παλαιότερων εποχών. Οι υπερπαραμέτροι παρέμειναν ίδιοι με το πείραμα της ενότητας 4.3.1, επομένως παρουσιάζονται στον πίνακα 4-8.

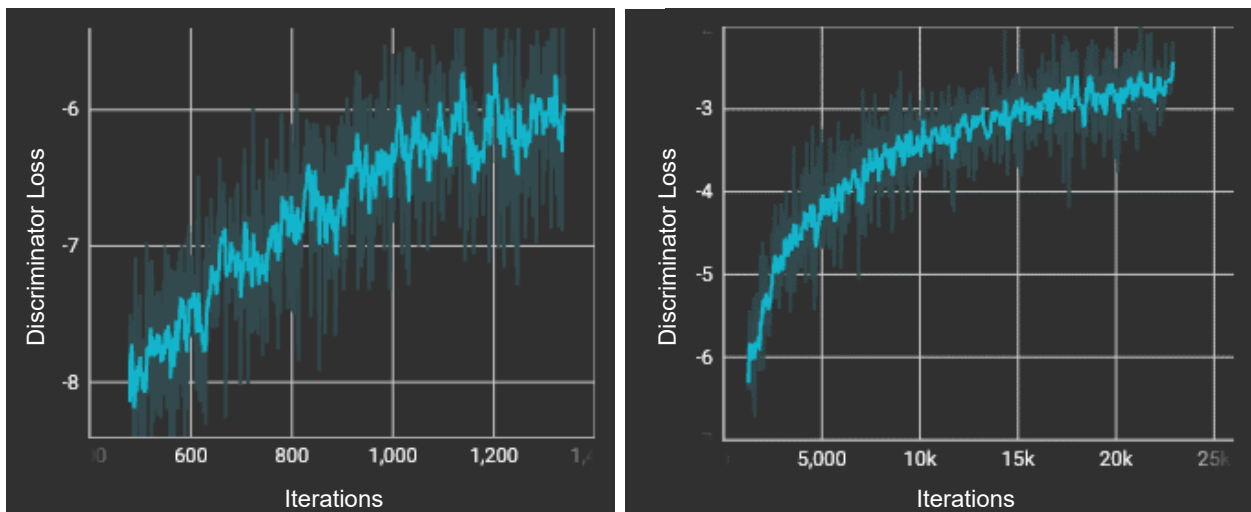


4-6 Τεχνητές εικόνες WGANGP 1 στο ΣΔ2 εποχή 11225



4-7 Τεχνητές εικόνες WGANGP 1 ΣΔ2 εποχές 100, 1000 και 5000

Στο σχήμα 4-8, παρουσιάζονται οι απώλειες του Discriminator, για τις πρώτες 400 εποχές και για τις υπόλοιπες αντίστοιχα.



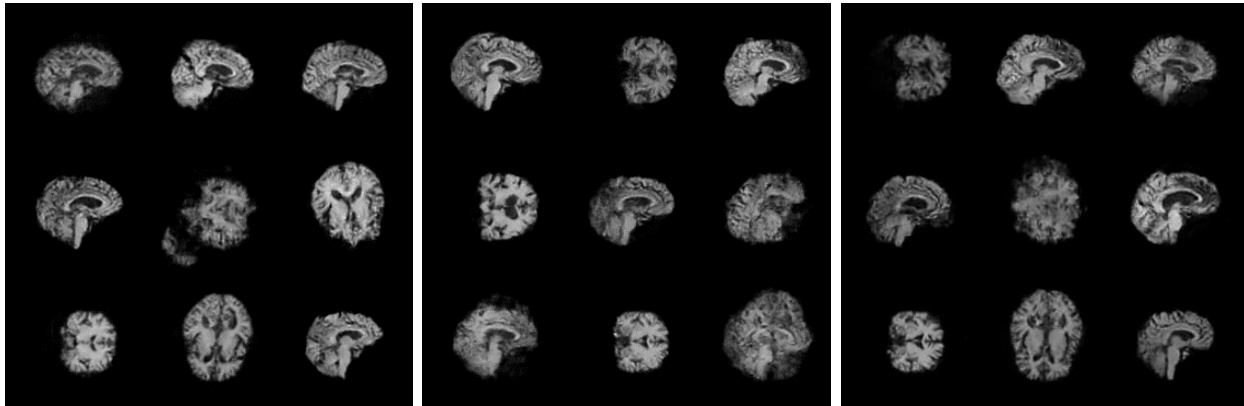
4-8 Discriminator Loss WGAN 1 ΣΔ2, ως εποχή 400 (αριστερά) και ως εποχή 11225 (δεξιά)

4.3.3. WGANGP 1, με σύνολο δεδομένων ΣΔ3

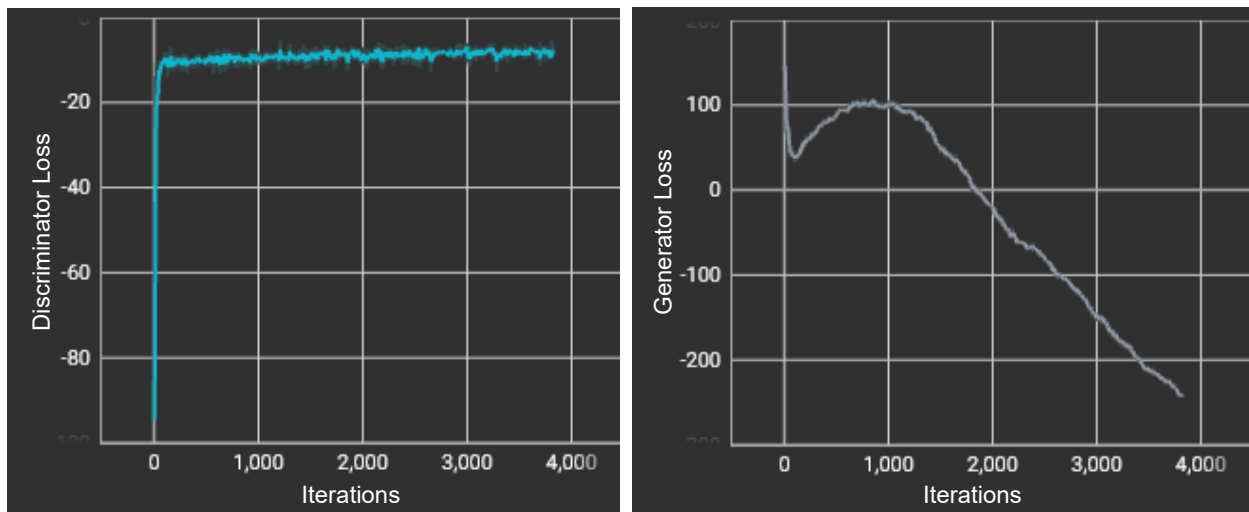
Παραδείγματα των τεχνητών εικόνων κατά τη διάρκεια της εκπαίδευσης του πειράματος φαίνονται στο σχήμα 4-9 και της εξέλιξης των απωλειών τόσο του Discriminator όσο και του Generator στο σχήμα 4-10. Οι παράμετροι που επιλέχθηκαν για την εκπαίδευση φαίνονται στον πίνακα 4-9. Να σημειωθεί ότι πραγματοποιήθηκε δοκιμή και για BATCH_SIZE=128 χωρίς διαφορά στην ποιότητα και σε απώλειες, αλλά σημαντική αύξηση του χρόνου εκπαίδευσης, επομένως η προσπάθεια ματαιώθηκε.

LEARNING_RATE	0.0001
BATCH_SIZE	64
IMAGE_SIZE	128
CHANNELS_IMG	1
Z_DIM (Κανάλια θορύβου)	256
CRITIC_ITERATIONS	10
LAMBDA_GP	10

4-9 Παράμετροι πειράματος

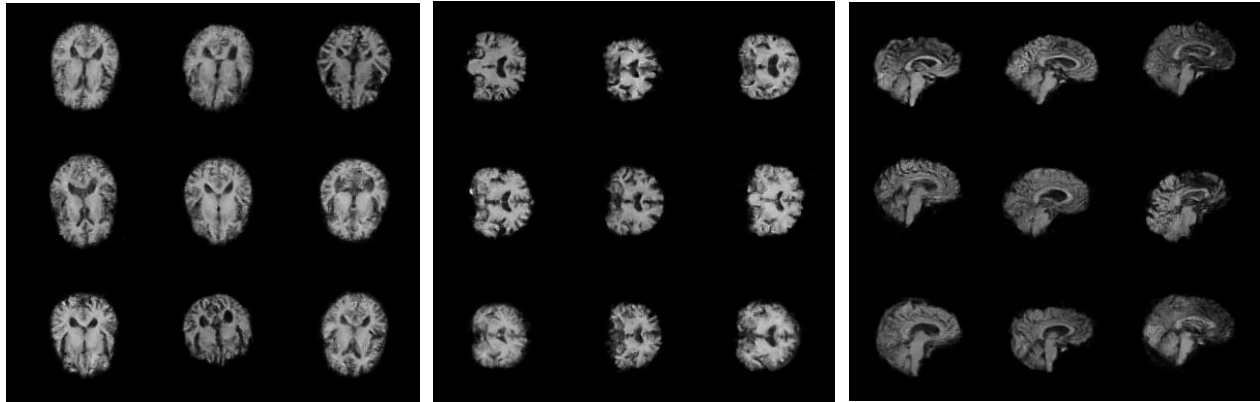


4-9 Πρώτη δοκιμή αρχιτεκτονικής WGANGP 1 στο ΣΔ3 χωρίς διαχωρισμό των τομών, εποχές 200, 1000, 1900

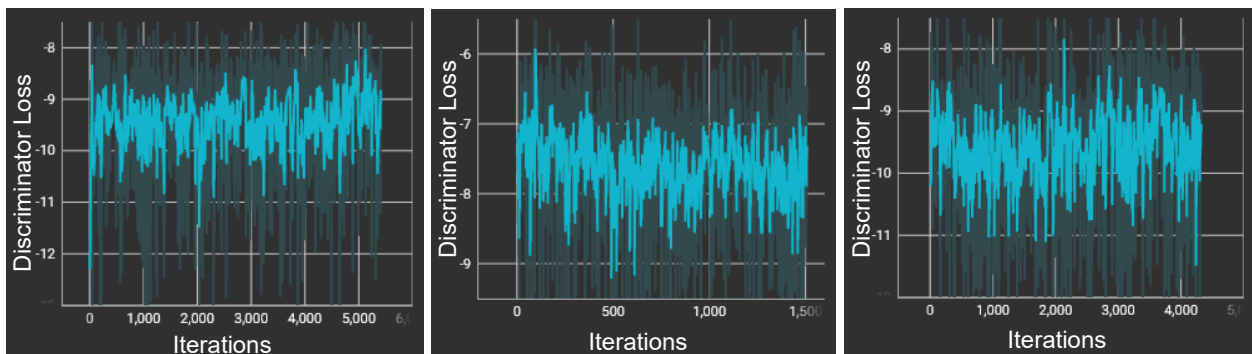


4-10 Πρώτη δοκιμή αρχιτεκτονικής WGANGP 1 στο ΣΔ3 χωρίς διαχωρισμό των τομών, απώλειες Discriminator (αριστερά) και Generator (δεξιά)

Αποτελέσματα της αρχιτεκτονικής από την υλοποίηση τριών διαφορετικών μοντέλων, ένα για κάθε τομή παρουσιάζονται στα σχήματα 4-11 και 4-12. Τα δίκτυα εκπαιδεύτηκαν με τις παραμέτρους του πίνακα 4-9.



4-11 Τεχνητές εικόνες από τα ξεχωριστά δίκτυα των τομών, axial, coronal και sagittal αντίστοιχα



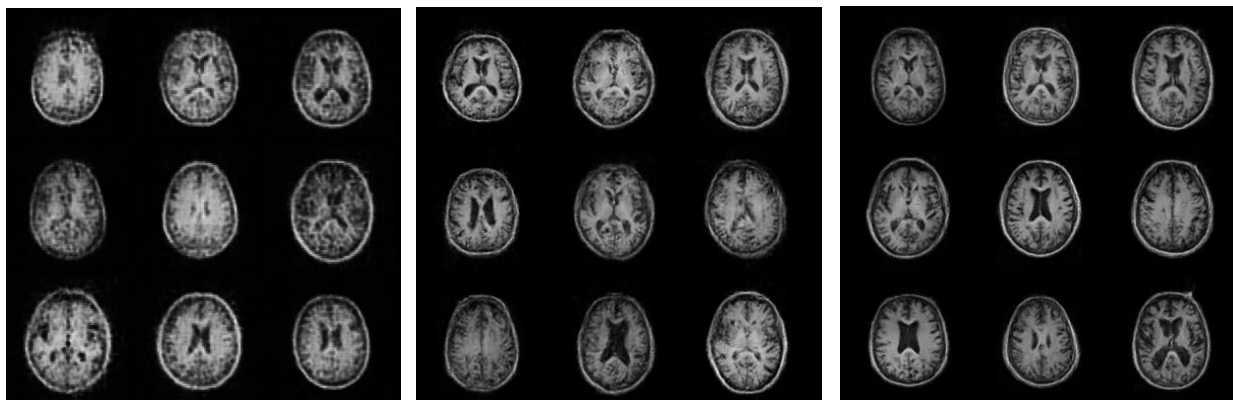
4-12 Απώλειες Discriminator από τα ξεχωριστά δίκτυα των τομών, axial, coronal και sagittal

4.3.4. WGANGP 2, με σύνολο δεδομένων ΣΔ1

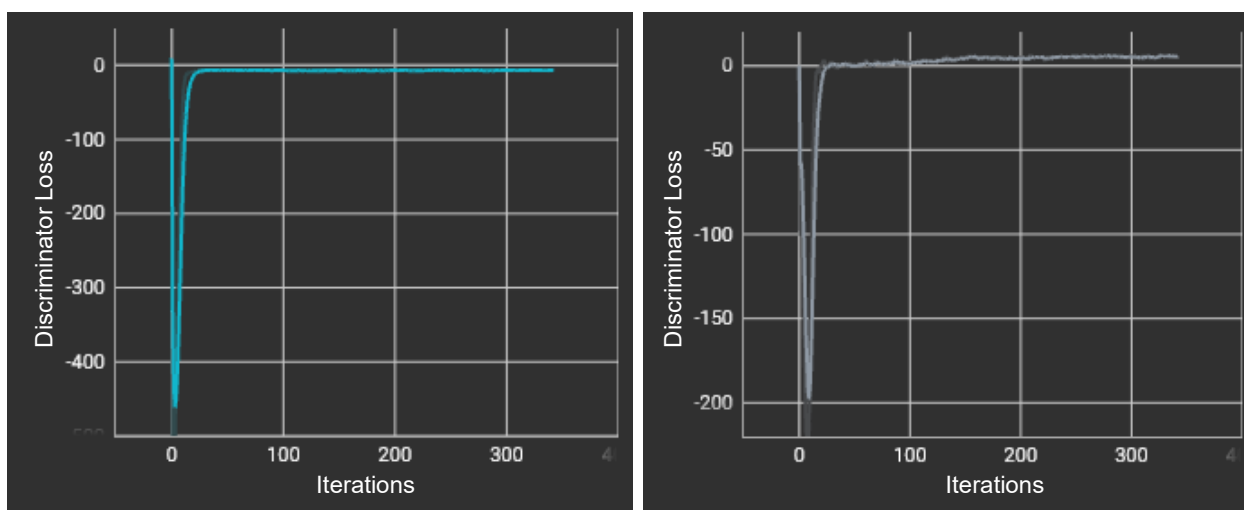
Στα σχήματα 4-13 και 4-14, παρουσιάζονται τεχνητές εικόνες από διάφορες εποχές της εκπαίδευσης και οι απώλειες του Generator και του Discriminator, από τη διεξαγωγή του πειράματος της αρχιτεκτονικής WGANGP 2 στο ΣΔ1. Ο πίνακας 4-10, περιέχει τις υπερπαραμέτρους που χρησιμοποιήθηκαν.

LEARNING_RATE	0.0001
BATCH_SIZE	64
IMAGE_SIZE	128
CHANNELS_IMG	1
Z_DIM	256
CRITIC_ITERATIONS	10
LAMBDA_GP	10

4-10 WGANGP 2 ΣΔ1, υπερπαραμέτροι εκπαίδευσης



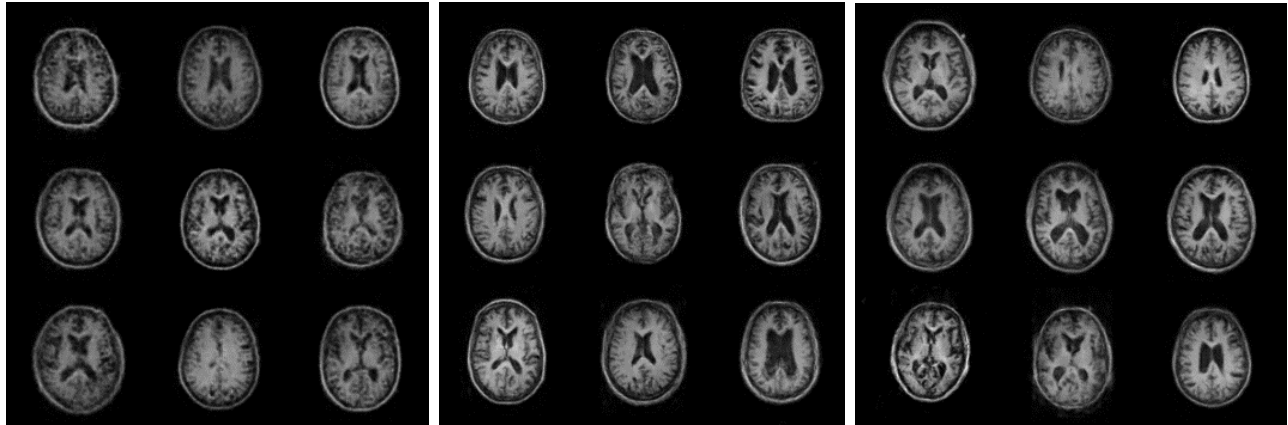
4-13 Τεχνητές εικόνες από την αρχιτεκτονική WGANGP 2 στο ΣΔ1, εποχές 10, 50, 170



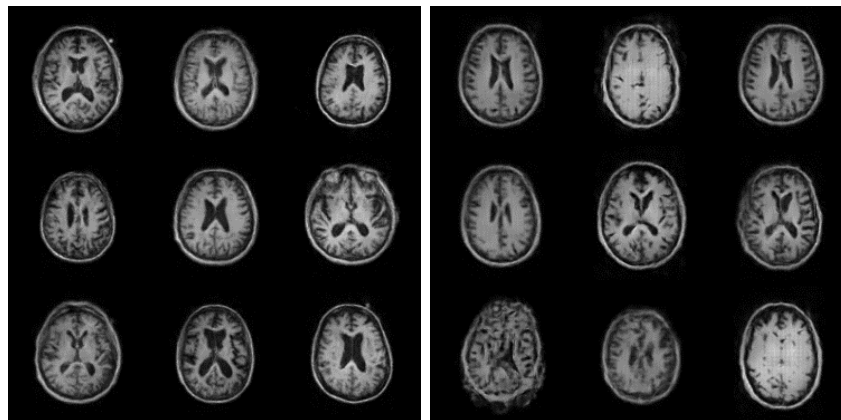
4-14 Απώλειες Discriminator και Generator WGANGP 2 στο ΣΔ1

4.3.5. WGANGP 2, με σύνολο δεδομένων ΣΔ2

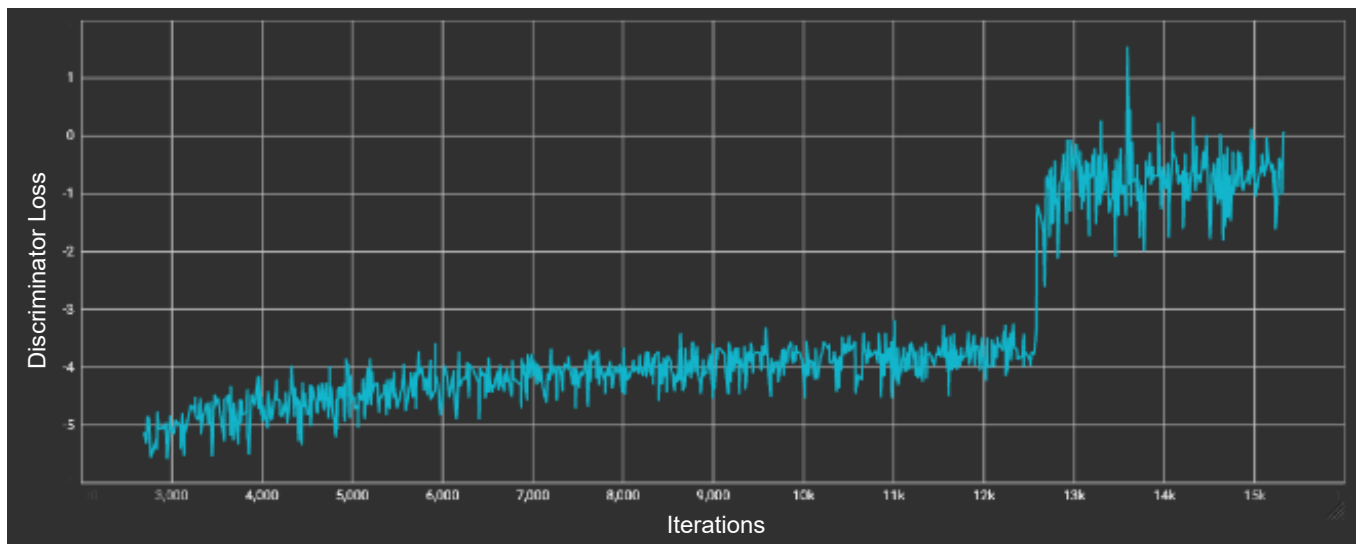
Στις εικόνες 4-15 και 4-16 παρουσιάζεται η ποιότητα των εικόνων κατά την πορεία της εκπαίδευσης και στις εικόνες 4-17 και 4-18 απεικονίζονται οι γραφικές παραστάσεις των απωλειών του Discriminator και του Generator αντίστοιχα. Οι παράμετροι που χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου παρουσιάζονται στον πίνακα 4-10.



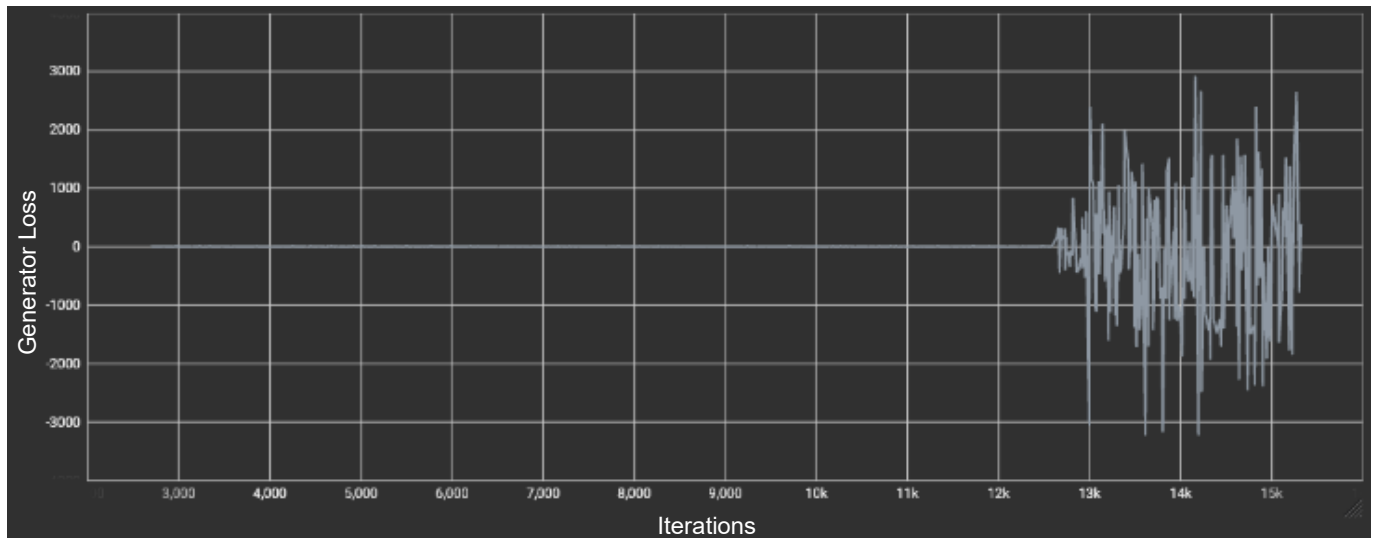
4-15 Τεχνητές εικόνες από την αρχιτεκτονική WGANGP 2 στο ΣΔ2, εποχές 1000, 2000, 4000



4-16 Τεχνητές εικόνες από την αρχιτεκτονική WGANGP 2 στο ΣΔ2, εποχές 6250 και 7660



4-17 Απώλεια Discriminator WGANGP 2 στο ΣΔ2



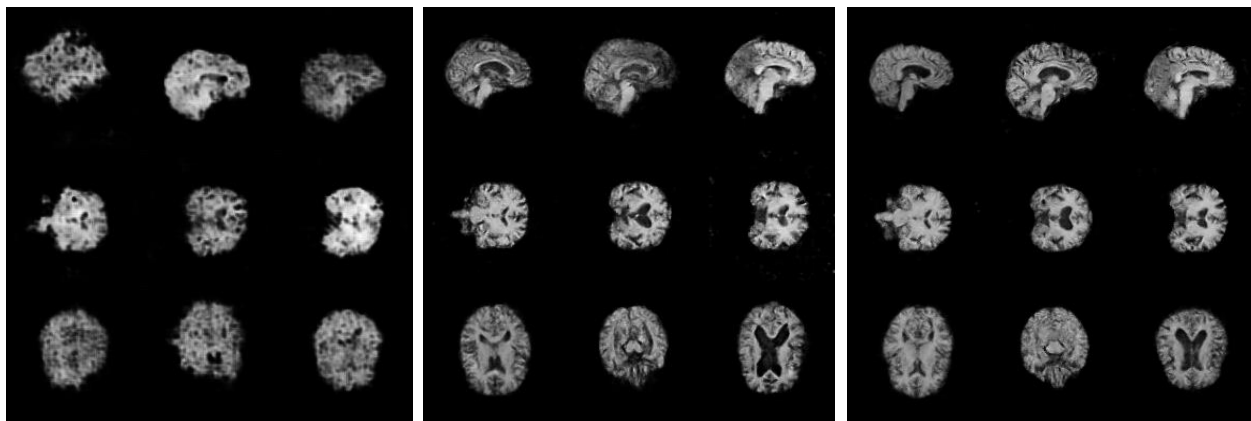
4-18 Απώλεια Generator WGANGP 2 στο ΣΔ2

4.3.6. CWGANGP 1, με σύνολο δεδομένων ΣΔ3

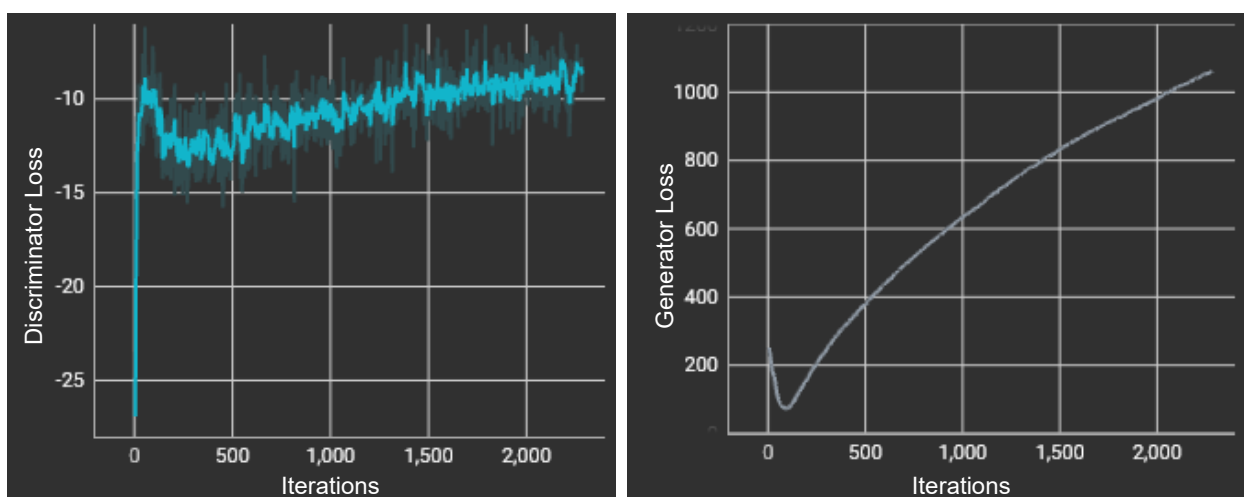
Οι παράμετροι που χρησιμοποιήθηκαν για την εκπαίδευση παρουσιάζονται στον πίνακα 4-11. Στην εικόνα 4-19 φαίνεται η πορεία των παραγόμενων εικόνων κατά τη διάρκεια της εκπαίδευσης. Τέλος, το γράφημα της απώλειας του Discriminator και του Generator βρίσκονται στην εικόνα 4-20.

LEARNING_RATE	0.0001
BATCH_SIZE	64
IMAGE_SIZE	128
CHANNELS_IMG	1
Z_DIM	256
CRITIC_ITERATIONS	5
LAMBDA_GP	5

4-11 CWGANGP 1 ΣΔ3, υπερπαράμετροι εκπαίδευσης



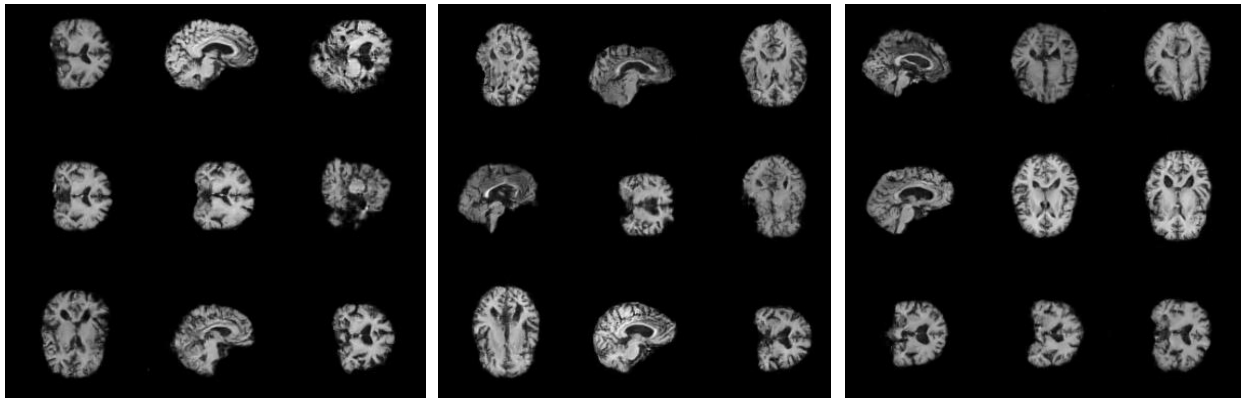
4-19 Τεχνητές εικόνες από την αρχιτεκτονική CWGAN_{GP} 1 στο ΣΔ3, εποχές 100, 500, 1100



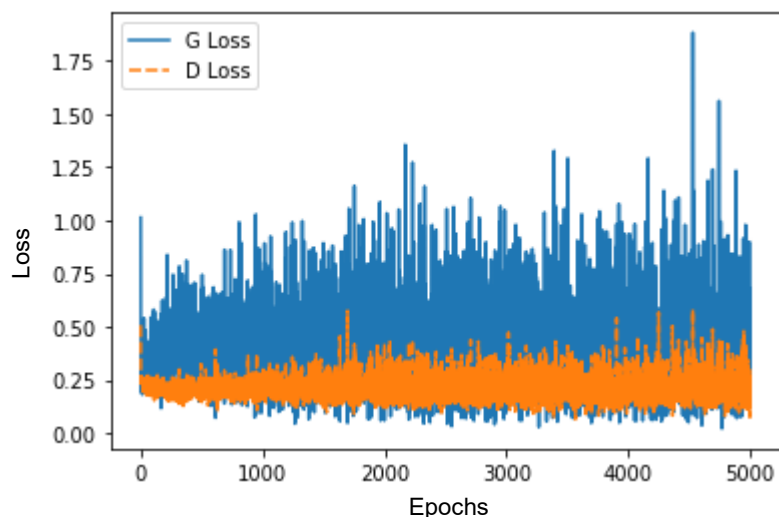
4-20 Απώλειες Discriminator και Generator CWGAN_{GP} 1 στο ΣΔ3

4.3.7. LSGAN, με σύνολο δεδομένων ΣΔ3

Στην εικόνα 4-21, παρουσιάζεται η πορεία της εκπαίδευσης του μοντέλου, μέσω του δείγματος τεχνητών εικόνων για κάποιες εποχές και στην εικόνα 4-22 υπάρχουν οι απώλειες του Discriminator και του Generator.



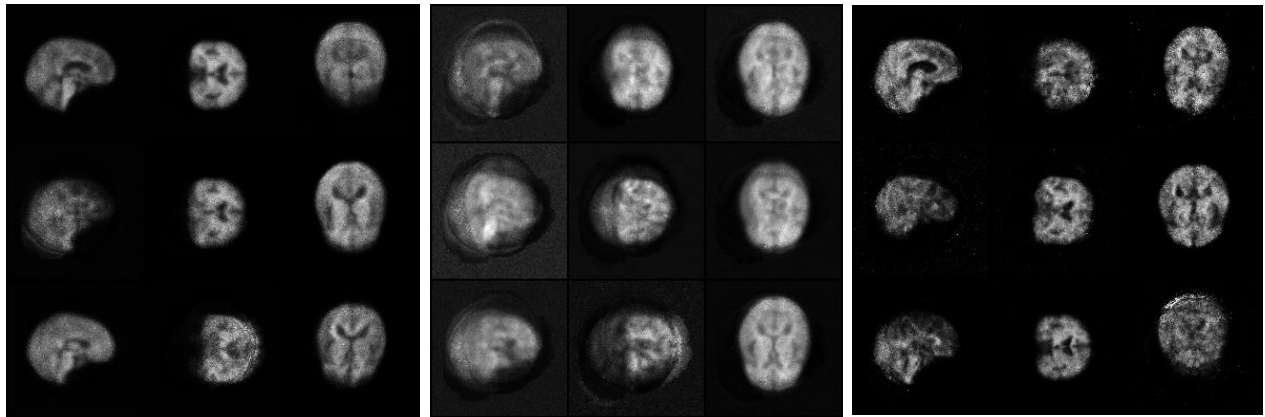
4-21 Τεχνητές εικόνες από την αρχιτεκτονική LSGAN στο ΣΔ3, εποχές 1000, 2500, 5000



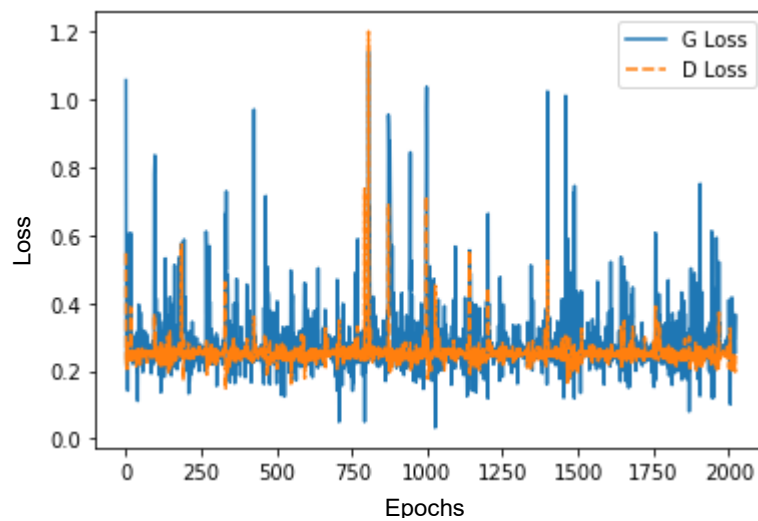
4-22 Απώλειες Discriminator και Generator LSGAN στο ΣΔ3

4.3.8. CGAN, με σύνολο δεδομένων ΣΔ3

Στην εικόνα 4-23 παρουσιάζονται τα ενδεικτικά αποτελέσματα τεχνητών εικόνων από κάποιες εποχές κατά την εκπαίδευση του μοντέλου και στην εικόνα 4-24 εμφανίζονται οι γραφικές παραστάσεις των απωλειών του Discriminator και του Generator.



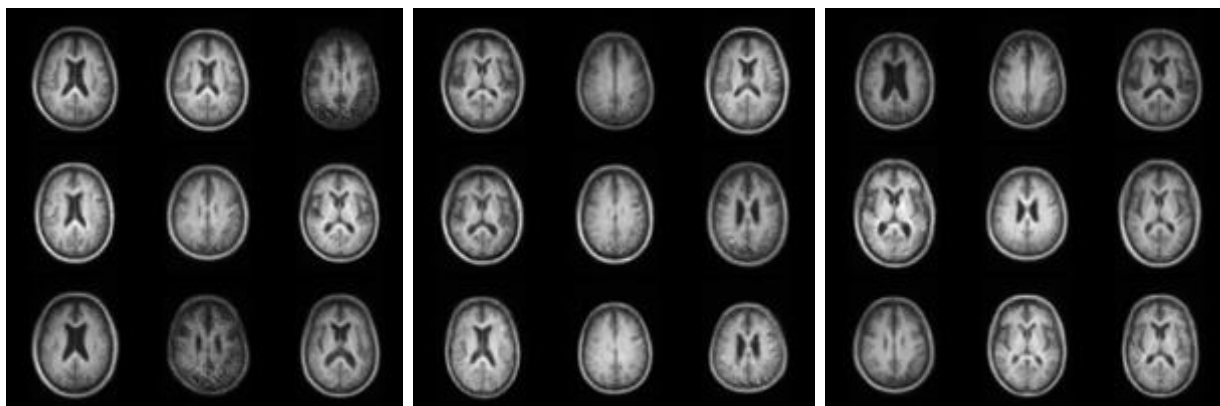
4-23 Τεχνητές εικόνες από την αρχιτεκτονική cGAN στο ΣΔ3, εποχές 500, 1000, 2000



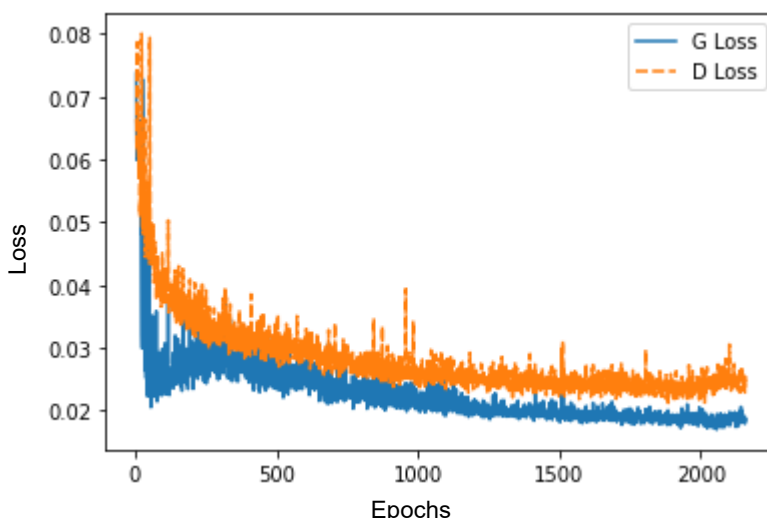
4-24 Απώλειες Discriminator και Generator cGAN στο ΣΔ3

4.3.9. BEGAN, με σύνολο δεδομένων ΣΔ1

Στις παρακάτω εικόνες προβάλλονται τα αποτελέσματα από την εκπαίδευση της αρχιτεκτονικής. Στην εικόνα 4-25 φαίνεται η πορεία της εξέλιξης των τεχνητών εικόνων κατά τη διάρκεια των εποχών και στην εικόνα 4-26 οι απώλειες του Discriminator και του Generator.



4-25 Τεχνητές εικόνες από την αρχιτεκτονική BEGAN στο ΣΔ1, εποχές 500, 1000, 2000



4-26 Απώλειες Discriminator και Generator BEGAN στο ΣΔ1

4.4. Inception Score

Για πιο ολοκληρωμένη εικόνα πραγματοποιήθηκε αξιολόγηση των τεχνητών εικόνων για κάθε αρχιτεκτονική με χρήση της μετρική Inception Score.

Το Inception Score (IS) είναι μια μετρική που ποσοτικοποιεί πόσο ρεαλιστικές είναι οι παραγόμενες τεχνητές εικόνες. Με το IS αξιολογούνται δύο παράγοντες, αν οι τεχνητές εικόνες έχουν ποικιλία, χωρίς να επαναλαμβάνουν συγκεκριμένα αποτελέσματα και αν κάθε εικόνα ξεχωριστά μοιάζει με κάποια από το πραγματικό σύνολο δεδομένων. Αν ισχύουν και τα δύο τότε το IS είναι υψηλό, κάτι που σημαίνει ότι το GAN είναι ικανό να παράξει ποικιλία τεχνητών εικόνων υψηλής ποιότητας. Για τον υπολογισμό του IS χρησιμοποιείται ο ταξινομητής Inception [53]. Εισάγοντας εικόνες στον ταξινομητή,

επιστρέφεται μια κατανομή πιθανοτήτων για κάθε εικόνα. Συνδυάζοντας αθροιστικά τις κατανομές των εικόνων υπολογίζεται η συνολική κατανομή (marginal distribution). Η κατανομή της κάθε εικόνας ιδανικά θα πρέπει να είναι κυρτή, ώστε να αναδεικνύει μια συγκεκριμένη κλάση, ενώ αντίθετα η marginal κατανομή θα πρέπει να είναι ομοιόμορφη. Τέλος χρησιμοποιείται η απόκλιση Kullback-Leibler (KL), μεταξύ των δύο κατανομών, γεγονός που καθορίζει τις υψηλές τιμές (μεγάλη απόκλιση) επιθυμητές, διότι αυτό σημαίνει πως οι τεχνητές εικόνες περιέχουν μοναδική εικόνα, ενώ ταυτόχρονα υπάρχει ποικιλία στο τεχνητό δείγμα.

Στον πίνακα 4-12 παρουσιάζονται τα αποτελέσματα για κάθε σύνολο δεδομένων (ΣΔ1, ΣΔ2 και ΣΔ3) και για κάθε αρχιτεκτονική που μελετήθηκε στις προηγούμενες ενότητες. Για τις αρχιτεκτονικές που δεν εφαρμόστηκαν σε κάποιο από τα σύνολα τα αντίστοιχα κελιά στον πίνακα παραμένουν κενά.

	ΣΔ1		ΣΔ2		ΣΔ3	
Architecture	IS AVG	IS STD	IS AVG	IS STD	IS AVG	IS STD
Real Images	1.72	0.01	1.77	0.03	2.53	0.05
WGANGP1	1.57	0.01	1.75	0.02	2.20	0.02
WGANGP2	1.65	0.01	1.74	0.02		
CWGANGP1					2.28	0.03
LSGAN					2.31	0.02
BEGAN	1.51	0.01				

4-12 Inception Score ανά σύνολο δεδομένων και ανά αρχιτεκτονική

4.5. Αποτελέσματα ταξινομητή

Σε αυτήν την ενότητα θα παρουσιαστούν τα αποτελέσματα από τις δοκιμές των ταξινομητών, πάνω στα 3 σύνολα δεδομένων, χρησιμοποιώντας τεχνητές εικόνες από τις διάφορες αρχιτεκτονικές. Οι μετρικές που εμφανίζονται είναι οι εξής:

- **Balanced Accuracy (BAC):** Η ισορροπημένη ακρίβεια είναι μία μετρική που χρησιμοποιείται για να αξιολογήσει την απόδοση ενός μοντέλου σε προβλήματα ταξινόμησης κυρίως κατά την περίπτωση όπου οι κλάσεις έχουν ανισόρροπη κατανομή. Υπολογίζεται ως η μέση τιμή ανάμεσα στην ακρίβεια (precision) και

την ανάκληση (recall). Παρέχει μια πιο ρεαλιστική αξιολόγηση σε σχέση με την απλή ακρίβεια, ιδιαίτερα σε περιπτώσεις όπου υπάρχει μεγάλη ανισότητα ανάμεσα στις κλάσεις. Στον τύπο 4-1 παρουσιάζεται ο υπολογισμός του BAC.

$$BAC = \frac{(Sensitivity + Specificity)}{2}$$

4-1

- **Specificity:** Το specificity είναι μία μετρική που υπολογίζει το ποσοστό των αρνητικών περιπτώσεων που ταξινομήθηκαν σωστά από το μοντέλο. Είναι σημαντικό να το εξετάσει κάποιος, ειδικά όταν ενδιαφέρεται για τον αντίκτυπο των ψευδώς θετικών αποτελεσμάτων. Ο υπολογισμός του Specificity αποτυπώνεται στον τύπο 4-2.

$$Specificity = \frac{True_Negative}{Negative}$$

4-2

- **Recall:** Το Recall, ή Ευαισθησία (Sensitivity) είναι μία μετρική που υπολογίζει το ποσοστό των θετικών περιπτώσεων που ταξινομήθηκαν σωστά από το μοντέλο. Είναι ιδιαίτερα σημαντικό σε περιπτώσεις όπου οι ψευδώς αρνητικοί έχουν υψηλό κόστος. Ο υπολογισμός του Recall αποτυπώνεται στον τύπο 4-3.

$$Recall = \frac{True_Positive}{Positive}$$

4-3

- **Precision:** Το Precision, είναι μία μετρική που υπολογίζει το ποσοστό των θετικών προβλέψεων που ήταν πραγματικά θετικές. Είναι ιδιαίτερα σημαντικό σε περιπτώσεις όπου οι ψευδώς θετικοί έχουν υψηλό κόστος. Ο υπολογισμός του Precision αποτυπώνεται στον τύπο 4-4.

$$Precision = \frac{True_Positive}{True_Positive + False_Positive}$$

4-4

- **F1:** Το F1 Score είναι ένας συνδυασμός της Precision και της ανάκλησης (Recall) και υπολογίζεται σαν τον αρμονικό μέσο αυτών των δύο μετρικών.

Χρησιμοποιείται για να αξιολογήσει το μοντέλο όταν χρειάζεται μία ισορροπία μεταξύ ακρίβειας και ανάκλησης. Ο υπολογισμός του F1 Score αποτυπώνεται στον τύπο 4-5.

$$F1 = \frac{2True_Positive}{2True_Positive + False_Positive + False_Negative}$$

4-5

- **Loss:** Η απώλεια (loss) είναι μία μετρική που αξιολογεί το συνολικό κόστος της απόκλισης μεταξύ των προβλέψεων του μοντέλου και των πραγματικών ετικετών (των αποκρίσεων). Σκοπός της εκπαίδευσης είναι η ελαχιστοποίηση της ζημίας (loss).
- **Accuracy:** Η ακρίβεια (Accuracy) είναι μία μετρική που μετράει το ποσοστό των σωστών προβλέψεων του μοντέλου σε σχέση με το σύνολο των προβλέψεων. Είναι συχνά χρησιμοποιούμενο για προβλήματα ταξινόμησης, αλλά πρέπει να εξετάζεται με προσοχή, καθώς μπορεί να είναι παραπλανητική σε περιπτώσεις μη ισορροπημένου αριθμού δειγμάτων μεταξύ των κλάσεων, όπως στην περίπτωση που εξετάζει η εργασία αυτή. Ο υπολογισμός του Accuracy αποτυπώνεται στον τύπο 4-6.

$$Accuracy = \frac{True_Positive + True_Negative}{Positive + Negative}$$

4-6

Στις επόμενες υποενότητες παρουσιάζονται οι πίνακες των αποτελεσμάτων ανά σύνολο δεδομένων. Κάθε πίνακας περιλαμβάνει τα αποτελέσματα από την εκτέλεση του εκπαιδευμένου ταξινομητή στο test σύνολο, χρησιμοποιώντας τεχνητές εικόνες από μια αρχιτεκτονική. Οι γραμμές του πίνακα περιλαμβάνουν τις τιμές των μετρικών που προαναφέρθηκαν. Οι στήλες είναι χωρισμένες σε δύο ομάδες, την Imbalanced και την Balanced, με την κάθε ομάδα να αποτελείται από 6 στήλες με τα ποσοστά των τεχνητών εικόνων που περιλαμβάνονταν στο σύνολο εκπαίδευσης (train set). Η διαφορά μεταξύ των ομάδων έγκειται στην δομή του συνόλου εκπαίδευσης.

Στην ομάδα Imbalanced το σύνολο εκπαίδευσης προσομοιώνει την πραγματικότητα, καθώς έχει διατηρηθεί η ανισορροπία μεταξύ των κλάσεων, δηλαδή διατηρήθηκαν όλα τα

δείγματα από την κλάση των υγείων. Το ποσοστό τεχνητών εικόνων εφαρμόζεται στον αριθμό των διαθέσιμων πραγματικών εικόνων από την κλάση των ασθενών. Αυτό σημαίνει ότι για κάθε ποσοστό προστίθενται τόσες εικόνες ώστε το δείγμα των ασθενών να περιέχει το αντίστοιχο ποσοστό τεχνητών εικόνων. Με τον τρόπο αυτό διατηρείται η ανισορροπία του αριθμού δειγμάτων μεταξύ των κλάσεων.

Αντίθετα, στην ομάδα *Balanced* πραγματοποιήθηκε εξισορρόπηση του συνόλου εκπαίδευσης, επομένως ο αριθμός δειγμάτων NC είναι ίδιος με τον αριθμό δειγμάτων AD. Αυτό πραγματοποιήθηκε αφαιρώντας τα επιπλέον δείγματα από την κλάση NC. Το ποσοστό τεχνητών εικόνων εφαρμόζεται και σε αυτή τη περίπτωση στον αριθμό πραγματικών δειγμάτων της κλάσης AD, δίνοντας κάθε φορά νέο αριθμό δειγμάτων για την κλάση AD. Για κάθε ποσοστό αφαιρείται αντίστοιχος αριθμός δειγμάτων από την κλάση NC, ώστε να διατηρηθεί ίδιος αριθμός με την AD.

Οι τιμές με πράσινο χρώμα και έντονη γραμματοσειρά αναδεικνύουν τις καλύτερες τιμές των μετρικών, ανά γραμμή και ανά ομάδα. Συγκεκριμένα, για τις μετρικές *BAC*, *Specificity*, *Recall*, *Precision*, *F1* και *Accuracy*, καλύτερη τιμή θεωρείται η μεγαλύτερη, ενώ αντίθετα για την μετρική *Loss* καλύτερη τιμή θεωρείται η μικρότερη.

4.5.1. Σύνολο Δεδομένων 1

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	61%	63%	61%	61%	55%	50%	64%	65%	63%	61%	56%	50%
Specificity	84%	73%	84%	78%	92%	100%	61%	62%	70%	80%	91%	100%
Recall	37%	54%	38%	45%	19%	0%	67%	68%	56%	43%	21%	0%
Precision	40%	36%	40%	36%	40%	0%	33%	33%	34%	38%	38%	0%
F1	39%	43%	39%	40%	26%	0%	44%	44%	42%	40%	27%	0%
Loss	0.53	0.61	0.53	0.58	0.53	3.41	0.69	0.69	0.63	0.56	0.54	3.56
Accuracy	74%	69%	74%	71%	76%	78%	63%	63%	67%	72%	75%	78%

4-13 Αποτελέσματα ταξινόμητη στο σύνολο ΣΔ1 με τεχνητές εικόνες από την αρχιτεκτονική *WGANGP1*

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	61%	63%	61%	62%	57%	50%	64%	65%	63%	61%	55%	50%
Specificity	84%	73%	84%	77%	89%	100%	61%	60%	69%	81%	92%	100%
Recall	37%	54%	38%	47%	26%	0%	67%	69%	58%	41%	19%	0%
Precision	40%	36%	40%	37%	40%	0%	33%	33%	34%	38%	39%	0%
F1	39%	43%	39%	41%	31%	0%	44%	44%	43%	40%	26%	0%
Loss	0.53	0.60	0.53	0.58	0.54	4.28	0.69	0.70	0.63	0.55	0.53	3.74
Accuracy	74%	69%	74%	71%	75%	78%	63%	62%	67%	72%	76%	78%

4-14 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ1 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP2

4.5.2. Σύνολο Δεδομένων 2

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	72%	74%	73%	68%	75%	56%	78%	79%	79%	79%	76%	61%
Specificity	92%	93%	93%	94%	89%	99%	74%	72%	75%	75%	74%	98%
Recall	52%	55%	53%	43%	60%	12%	82%	85%	83%	84%	78%	23%
Precision	36%	41%	40%	41%	34%	65%	22%	22%	24%	23%	22%	54%
F1	43%	47%	45%	42%	43%	20%	35%	35%	37%	36%	34%	32%
Loss	0.27	0.27	0.26	0.25	0.32	0.39	0.52	0.54	0.50	0.50	0.53	0.35
Accuracy	88%	90%	89%	90%	87%	92%	75%	73%	76%	75%	74%	92%

4-15 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ2 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP1

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	72%	74%	73%	69%	77%	56%	78%	79%	80%	79%	76%	61%
Specificity	92%	93%	92%	93%	86%	99%	74%	71%	75%	75%	75%	98%
Recall	52%	54%	53%	45%	67%	12%	82%	86%	85%	83%	77%	23%
Precision	36%	41%	39%	41%	34%	65%	22%	22%	24%	23%	23%	54%
F1	43%	46%	45%	43%	49%	20%	35%	35%	37%	36%	33%	32%
Loss	0.27	0.26	0.26	0.25	0.32	1.26	0.52	0.55	0.49	0.50	0.52	0.53
Accuracy	88%	90%	89%	90%	86%	92%	75%	72%	76%	76%	75%	92%

4-16 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ2 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP2

4.5.3. Σύνολο Δεδομένων 3

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	60%	59%	59%	58%	58%	50%	62%	60%	61%	60%	58%	50%
Specificity	88%	88%	86%	86%	83%	100%	64%	53%	58%	69%	84%	100%
Recall	32%	30%	32%	29%	32%	0%	59%	68%	63%	52%	32%	0%
Precision	37%	35%	34%	32%	30%	0%	27%	24%	25%	27%	31%	33%
F1	34%	32%	33%	31%	31%	0%	37%	36%	36%	35%	31%	0%
Loss	0.50	0.50	0.52	0.52	0.55	3.19	0.65	0.74	0.71	0.63	0.54	1.93
Accuracy	77%	77%	76%	76%	74%	82%	63%	55%	59%	65%	75%	82%

4-17 Αποτελέσματα ταξινομητή στο σύνολο ΣΔ3 με τεχνητές εικόνες από την αρχιτεκτονική WGANGP1

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	60%	59%	58%	58%	57%	50%	62%	60%	60%	60%	58%	50%
Specificity	88%	87%	85%	86%	84%	100%	60%	51%	56%	67%	84%	100%
Recall	32%	30%	32%	29%	30%	0%	63%	70%	65%	53%	31%	0%
Precision	37%	77%	33%	32%	30%	26%	26%	24%	25%	26%	31%	0%
F1	34%	32%	32%	31%	30%	1%	37%	36%	36%	35%	31%	0%
Loss	0.50	0.51	0.52	0.52	0.54	2.47	0.68	0.75	0.72	0.65	0.54	2.79
Accuracy	77%	77%	75%	76%	74%	81%	61%	54%	57%	64%	75%	81%

4-18 Αποτελέσματα ταξινόμητη στο σύνολο ΣΔ3 με τεχνητές εικόνες από την αρχιτεκτονική CWGANP1

Generated images %	Imbalanced						Balanced					
	0%	10%	20%	50%	80%	100%	0%	10%	20%	50%	80%	100%
BAC	60%	60%	59%	58%	57%	50%	62%	61%	61%	60%	57%	50%
Specificity	88%	84%	84%	83%	83%	99%	61%	61%	61%	66%	81%	100%
Recall	32%	35%	34%	33%	32%	0%	62%	61%	61%	53%	34%	1%
Precision	37%	33%	32%	31%	30%	14%	27%	26%	26%	26%	29%	22%
F1	34%	34%	33%	32%	31%	1%	37%	37%	37%	35%	31%	1%
Loss	0.50	0.53	0.53	0.54	0.55	1.80	0.68	0.68	0.68	0.65	0.56	1.84
Accuracy	77%	75%	75%	74%	74%	81%	61%	61%	61%	64%	72%	81%

4-19 Αποτελέσματα ταξινόμητη στο σύνολο ΣΔ3 με τεχνητές εικόνες από την αρχιτεκτονική LSGAN

5. Συζήτηση

Στην ενότητα αυτή θα συζητηθούν και θα αναλυθούν τα αποτελέσματα που παρουσιάστηκαν στη ενότητα 4, με στόχο τη σύγκριση της ποιότητας των τεχνητών εικόνων που παράχθηκαν από τις διάφορες αρχιτεκτονικές GAN, αλλά και των αποτελεσμάτων του ταξινομητή.

Το μοντέλο WGANP1 χρησιμοποιήθηκε για την παραγωγή τεχνητών εικόνων και στα 3 σύνολα δεδομένων. Για το σύνολο ΣΔ1 το μοντέλο εκπαιδεύτηκε για 240 εποχές και παρατηρείται ότι από πολύ νωρίς είχε καταφέρει να πετύχει καλή ποιότητα στις τεχνητές εικόνες, κάτι το οποίο γίνεται κατανοητό από τη γρήγορη ελαχιστοποίηση της απώλειας του Discriminator, όπως καταδεικνύεται στο σχήμα 4-5. Φαίνεται να έχει καταφέρει να αποτυπώσει τις προβληματικές περιοχές του εγκεφάλου που διαφοροποιούν τις δύο κλάσεις μεταξύ τους. Το αρνητικό στη συγκεκριμένη υλοποίηση είναι πως δείχνει να έχει σχεδόν σταθεροποιηθεί η απώλεια, μην έχοντας καταφέρει τη βέλτιστη ποιότητα μέσω οπτικής αξιολόγησης. Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκαν οι υπερπαραμέτροι του πίνακα 4-8.

Το σύνολο ΣΔ2 περιέχει τις ίδιες εικόνες με το ΣΔ1 αλλά με πιο αυστηρούς κανόνες επιλογής ασθενών. Αυτό περιορίζει τον αριθμό των εικόνων στο δείγμα, όπως φαίνεται και στον πίνακα 4-6, κάτι που δυσκολεύει την εκπαίδευση ενός GAN μοντέλου. Για την αντιμετώπιση της δυσκολίας που προκύπτει από τον μικρό αριθμό εικόνων, χρησιμοποιήθηκαν τα βάρη από παρόμοιο μοντέλο για την εκπαίδευση πάνω στο πιο εξειδικευμένο σύνολο δεδομένων ΣΔ2. Συγκεκριμένα, χρησιμοποιήθηκαν τα βάρη από το πείραμα της ενότητας 4.3.1, δηλαδή της αρχιτεκτονικής WGANP 1 πάνω στο σύνολο ΣΔ1. Με αυτόν τον τρόπο εξελίχθηκε η εκπαίδευση πιο γρήγορα αλλά και γενικότερα έγινε εφικτή η δημιουργία ενός μοντέλου ικανού να παράγει ικανοποιητικές τεχνητές εικόνες, από ένα τόσο μικρό σύνολο δεδομένων. Στο σχήμα 4-6 παρουσιάζονται τα αποτελέσματα με τις παραγόμενες τεχνητές εικόνες από την τελευταία εποχή και στο σχήμα 4-7 αποτελέσματα παλαιότερων εποχών. Οι υπερπαραμέτροι εκπαίδευσης παρέμειναν ίδιοι, επομένως παρουσιάζονται στον πίνακα 4-8. Λόγω τεχνικών προβλημάτων η εκπαίδευση διακόπηκε στην εποχή 400. Για παρόμοιες περιπτώσεις πραγματοποιούνταν η τακτική αποθήκευση των βαρών, υπήρχε επομένως τρόπος επαναφοράς της εκπαίδευσης από

την εποχή 401. Στο σχήμα 4-8, προβάλλονται οι απώλειες του Discriminator, για τις πρώτες 400 εποχές και για τις υπόλοιπες αντίστοιχα. Το παρόν πείραμα είναι το μεγαλύτερης διάρκειας από όσα πραγματοποιήθηκαν και ο λόγος είναι η αργή και σταθερή βελτίωση της απώλειας του Discriminator, όπως φαίνεται στο σχήμα 4-8. Η βελτίωση στην εμφάνιση των εικόνων (ποιοτική αξιολόγηση) δεν είναι ιδιαίτερα αισθητή. Η διακοπή της εκπαίδευσης ήταν αναγκαία για τη συνέχεια των υπόλοιπων πειραμάτων και για λόγους χρόνου και υπολογιστικών πόρων.

Τα αποτελέσματα της εκπαίδευσης του WGANGP 1 στο σύνολο ΣΔ3 παρουσιάζονται στα σχήματα 4-9 και 4-10, όπου παρατηρείται μικρή διαφορά στις εικόνες όσο περνάνε οι εποχές, αλλά εντοπίζεται επίσης η προσπάθεια απεικόνισης των διαφορετικών τομών. Σχετικά με τις απώλειες, φαίνεται να εμφανίζει πολύ μικρή βελτίωση η απώλεια του Discriminator κατευθυνόμενη προς το 0, έχοντας όμως, ακόμα, μεγάλη απόσταση να καλύψει. Η εκπαίδευση διακόπηκε, καθώς η εξέλιξη ήταν πολύ αργή και αποφασίστηκε να εκμεταλλευτούν οι πόροι στο επόμενο πείραμα που αποτελεί τη συνέχεια αυτού.

Έχοντας παρατηρήσει την ικανοποιητική απόδοση της αρχιτεκτονικής σε αυτό το σύνολο, θεωρήθηκε ότι η δυσκολία βελτίωσης οφείλεται στις διαφορετικές τομές που υπάρχουν στο σύνολο δεδομένων. Επίσης, λαμβάνοντας υπόψιν ότι μετέπειτα θα πρέπει να παραχθεί ίσος αριθμός εικόνων για κάθε κλάση (τομή εγκεφάλου), κάτι το οποίο θα ήταν αρκετά δύσκολο δεδομένου ότι το δίκτυο παράγει εικόνες τυχαία από κάθε τομή, δοκιμάστηκε, λοιπόν, η ιδέα εξειδίκευσης του δικτύου σε κάθε τμήμα ξεχωριστά. Για την επίτευξη του στόχου αυτού, αποθηκεύτηκαν τα βάρη από την τελευταία εποχή της πρώτης δοκιμής, ώστε να χρησιμοποιηθούν για την εκπαίδευση 3 ξεχωριστών μοντέλων, ένα για κάθε τομή.

Οι τεχνητές εικόνες των τελευταίων εποχών καθώς και οι απώλειες των Discriminators των ξεχωριστών μοντέλων παρουσιάζονται στα σχήματα 4-11 και 4-12, αντίστοιχα. Τα δίκτυα εκπαιδεύτηκαν με τις παραμέτρους του πίνακα 4-9. Από τις εικόνες και τα γραφήματα θα χαρακτηριζόταν ως θετική προσπάθεια, καθώς επιτεύχθηκε η εξειδίκευση του αρχικού μοντέλου σε συγκεκριμένη κλάση. Παρατηρείται μικρή βελτίωση της απώλειας του Discriminator, πράγμα που μπορεί να διαπιστωθεί από την οπτική

αξιολόγηση των εικόνων κατά τη διάρκεια της εκπαίδευσης, όπου δεν παρατηρήθηκε έντονη μεταβολή από εποχή σε εποχή.

Το μοντέλο WGANGP 2 αποτελεί παραλλαγή του WGANGP 1 και δοκιμάστηκε στα σύνολα ΣΔ1 και ΣΔ2. Για το σύνολο ΣΔ1, το μοντέλο εκπαιδεύτηκε για 170 εποχές. Από τις τεχνητές εικόνες 4-13 παρατηρείται βελτίωση με την πάροδο των εποχών. Το αξιοσημείωτο στην περίπτωση αυτή είναι ότι η πολύ μικρή διακύμανση της απώλειας του Discriminator, που δείχνει να έχει σταθεροποιηθεί από τα πρώτα μόλις στάδια της εκπαίδευσης, γεγονός που συμβαδίζει με τα αποτελέσματα των τεχνητών εικόνων, όπου η ποιότητα φαίνεται να είναι σε πολύ καλό επίπεδο από πολύ νωρίς.

Ενδιαφέρον αποτελέσματα εμφανίζει το πείραμα του WGANGP 2 με το σύνολο ΣΔ2, διότι παρατηρείται από τα σχήματα 4-15, 4-16, 4-17 και 4-18 ότι κατά τη διάρκεια της εκπαίδευσης υπάρχει εμφανής βελτίωση της ποιότητας των εικόνων, καθώς και αργή και σταθερή βελτίωση της απώλειας Wasserstein, κάτι το οποίο σημαίνει ότι ο Generator παράγει όλο και καλύτερες τεχνητές εικόνες, καταφέρνοντας να ξεγελάσει περισσότερο τον Discriminator ότι είναι πραγματικές.

Επίσης, τόσο από τα δείγματα των τεχνητών εικόνων όσο και από τις γραφικές παραστάσεις των δύο δικτύων, είναι φανερό ότι από ένα σημείο και έπειτα το δίκτυο καταρρέει. Συγκεκριμένα, μετά την εποχή 6251 οι απώλειες παρουσίασαν μια ιδιαίτερη συμπεριφορά. Ο Discriminator φαίνεται να αντιμετωπίζει δυσκολία στο να εντοπίσει τις τεχνητές εικόνες, κάτι το οποίο είναι αισθητό από την απότομη αύξηση της απώλειας και σταθεροποίησή της πολύ κοντά στο μηδέν. Σαν γεγονός το τελευταίο, θα ήταν ιδανική περίπτωση, διότι θα σήμαινε πως ο Generator παράγει πάρα πολύ καλές τεχνητές εικόνες, δηλαδή πολύ όμοιες με τις πραγματικές. Αντίθετα, όμως, ο Generator, φαίνεται να παράγει εικόνες οι οποίες γίνονται όλο και χειρότερες με το πέρασμα των εποχών, ενώ ταυτόχρονα παρατηρείται περίεργη συμπεριφορά στη απώλεια, έχοντας απότομες μεταβολές κατά τη διάρκεια της εκπαίδευσης. Γίνεται, λοιπόν, αντιληπτό ότι ο Generator βρήκε τρόπο να παράγει εικόνες με θόρυβο, οι οποίες ξεγελάνε τον Discriminator ότι είναι πραγματικές, ενώ στην πραγματικότητα δεν απεικονίζουν πλέον τόσο καλά τις λεπτομέρειες του ιστού του εγκεφάλου και σε αρκετές περιπτώσεις δεν δίνουν την εντύπωση ότι γίνεται προσπάθεια απεικόνισης ανθρώπινου εγκεφάλου.

Τέλος, σε αυτό το πείραμα, παρατηρώντας τις παραγόμενες τεχνητές εικόνες έγινε αντιληπτή η ικανότητα του μοντέλου να απεικονίσει λεπτομέρειες χαρακτηριστικών του προσώπου, όπως τα μάτια, κάτι που φαίνεται στην εικόνα 4-16.

Το πείραμα 4.3.3 έδωσε την αφορμή για τη δημιουργία της υπό συνθήκης αρχιτεκτονικής CWGAN με Gradient Penalty. Η αρχιτεκτονική CWGAN-GP που παρουσιάζεται δεν είναι η πρώτη που δημιουργήθηκε, αλλά είναι η πιο επιτυχημένη προσπάθεια και γι' αυτό παρουσιάζεται. Αποτελεί μέρος της προσπάθειας συνδυασμού της επιτυχημένης απόστασης Wasserstein και των υπό συνθήκη GAN. Η αρχιτεκτονική αυτή προσαρμόστηκε, ώστε να παράγει εικόνες 128x128 και αυξήθηκε η πολυπλοκότητα. Οι παράμετροι που χρησιμοποιήθηκαν για την εκπαίδευση παρουσιάζονται στον πίνακα 4-11. Στην εικόνα 4-19 φαίνεται η πορεία των παραγόμενων εικόνων κατά τη διάρκεια της εκπαίδευσης. Τέλος το γράφημα της απώλειας του Discriminator και του Generator παρουσιάζονται στην εικόνα 4-20. Όπως φαίνεται στις παραμέτρους, επιλέχθηκε να πραγματοποιηθούν 5 επαναλήψεις του Discriminator για 1 επανάληψη του Generator αντί για 10 κάτι το οποίο κάνει πιο «αδύναμο» τον Discriminator αλλά βοηθάει στην ταχύτητα της εκπαίδευσης. Για να εξισορροπηθεί η δύναμη του Discriminator, επιλέχθηκε η τιμή 5 για την παράμετρο λ (LAMBDA_GP) μειώνοντας την «ποινή» στην απώλεια του Discriminator, με συνέπεια να γίνει πιο «δυνατός». Από τα αποτελέσματα στην ενότητα 4.3.6 καθίσταται φανερό, τουλάχιστον από την ποιοτική αξιολόγηση, ότι υπάρχει βελτίωση στην ποιότητα των εικόνων, η οποία όμως είναι όλο και λιγότερο εμφανής όσο προχωράει η εκπαίδευση. Παρόμοια εικόνα αποτυπώνεται και στην απώλεια του Discriminator, καθώς παρατηρείται μικρή και σταθερή βελτίωση. Το σημαντικό σε αυτό το πείραμα είναι ότι έχει λειτουργήσει η αρχιτεκτονική υπό συνθήκη (Conditional), κάτι το οποίο γίνεται αντιληπτό στην εικόνα 4-19, όπου αποτυπώνονται τριάδες από τις 3 κλάσεις (τομές) του συνόλου ΣΔ3 (axial, coronal, sagittal). Με άλλα λόγια το μοντέλο έχει καταφέρει να διαχωρίσει τις κλάσεις και να αποτυπώσει τα χαρακτηριστικά της κάθε μιας, σε αρκετά καλό βαθμό, έχοντας κατορθώσει να αποτυπώσει λεπτομέρειες του ιστού και των χαρακτηριστικών της κάθε τομής. Όπως και στα προηγούμενα πειράματα, η απώλεια του Generator δεν δίνει κάποια σημαντική πληροφορία, αλλά συμπεριλαμβάνεται για να δοθεί μια ολοκληρωμένη εικόνα της συμπεριφοράς των μοντέλων.

Μια ακόμα επιτυχημένη αρχιτεκτονική είναι το LSGAN, το οποίο, αν και περιλαμβάνει «αδύναμο» από άποψη αριθμό παραμέτρων, Discriminator, όπως θα αποδειχθεί στη συνέχεια, έχει πετύχει την παραγωγή ρεαλιστικών τεχνητών εικόνων. Στην εικόνα 4-21, παρουσιάζεται η πορεία της εκπαίδευσης του μοντέλου, μέσω του δείγματος τεχνητών εικόνων για κάποιες εποχές και στην εικόνα 4-22 προβάλλονται οι απώλειες του Discriminator και του Generator. Από τις εικόνες που παρουσιάζονται στο σχήμα 4-21 είναι φανερό ότι το δίκτυο έχει καταφέρει από νωρίς να διατυπώσει τις διαφορετικές κλάσεις και να εμφανίσει εικόνες πολύ κοντά στο σχήμα του εγκεφάλου, χάνοντας βέβαια κάποιες λεπτομέρειες. Ακόμα σε κάποιες περιπτώσεις συγχέονται οι διαφορετικές τομές, με αποτέλεσμα να μην φαίνεται μόνο μια καθαρή τομή. Μέχρι την εποχή 5000 το δίκτυο έχει καταφέρει να «μάθει» την κατανομή των πραγματικών εικόνων και να δημιουργεί τεχνητές εικόνες όμοιες με τις πραγματικές. Όπως και οι υπόλοιπες αρχιτεκτονικές στο ΣΔ3 υπάρχει δυσκολία δημιουργίας της τομής sagittal, κάτι που ίσως οφείλεται στο ότι μοιάζει με την τομή axial αλλά οριζόντια. Τέλος παρατηρείται στο σχήμα 4-22 ότι οι απώλειες, σε αντίθεση με τη Wasserstein απόσταση που χρησιμοποιείται στα περισσότερα από τα υπόλοιπα πειράματα, στην αρχιτεκτονική αυτή, χρησιμοποιήθηκε η συνάρτηση απώλειας ελαχίστων τετραγώνων η οποία υπολογίζει την απόσταση μεταξύ των πραγματικών και των παραγόμενων δεδομένων και στοχεύει στην ελαχιστοποίηση του τετραγωνικού σφάλματος.

Ακολούθως η δομή του cGAN που υλοποιήθηκε είναι διαφορετική από τα πειράματα που παρουσιάστηκαν ως τώρα, καθώς δεν αποτελείται από συνελικτικά επίπεδα, αλλά κυρίως από γραμμικά (πλήρως συνδεδεμένα). Το μοντέλο εκπαιδεύτηκε για 2000 εποχές, όπου διακόπηκε η διαδικασία καθώς δεν παρατηρούνταν βελτίωση των αποτελεσμάτων κάτι που φαίνεται στο σχήμα 4-23. Πιθανές αιτίες αποτυχίας μπορεί να είναι οι ακόλουθες, ενδεχομένως να χρειαζόταν περισσότερος χρόνος εκπαίδευσης, οι παράμετροι ήταν λίγοι και δεν ήταν εφικτό να μπορέσει το μοντέλο να αντιγράψει τις λεπτομέρειες των πραγματικών εικόνων, τέλος ίσως τα συνελικτικά δίκτυα να είναι καταλληλότερα για την επίλυση του παρόντος προβλήματος, με το συγκεκριμένο σύνολο δεδομένων.

Παρόμοια με την αρχιτεκτονική cGAN στο ΣΔ3, αποτυχία αποτελεί και η εκπαίδευση του BEGAN στο σύνολο ΣΔ1. Αν και αποτελείται από συνελικτικά επίπεδα όπως και οι

περισσότερες αρχιτεκτονικές στην παρούσα εργασία, έχει μια σημαντική διαφορά: η δομή υλοποιήθηκε για την παραγωγή τεχνητών εικόνων διαστάσεων 64x64. Έτσι όπως είναι φανερό στην ενότητα 3.2.5, τόσο ο Generator όσο και ο Discriminator δεν συνιστούν περίπλοκες αρχιτεκτονικές καθώς αποτελούνται από λίγες εκπαιδύσιμες παραμέτρους. Στις εικόνες 4-25 και 4-26 παρουσιάζεται δείγμα τεχνητών εικόνων και οι απώλειες από την εκπαίδευση αντίστοιχα. Το μοντέλο εκπαιδεύτηκε για 2000 εποχές και διακόπηκε η εκπαίδευση καθώς δεν υπήρχαν σημάδια βελτίωσης των αποτελεσμάτων.

Όπως έχει προαναφερθεί κύριες μέθοδοι αξιολόγησης είναι η ποιοτική (οπτικά) και η απώλεια του Discriminator στις περιπτώσεις που έχει νόημα, όπως με τη χρήση της Wasserstein απόστασης. Για πιο ολοκληρωμένη εικόνα πραγματοποιήθηκε αξιολόγηση των τεχνητών εικόνων για κάθε αρχιτεκτονική με χρήση της μετρική Inception Score. Η υλοποίηση έγινε σε ρυθμό με τη χρήση της βιβλιοθήκης TensorFlow [54]. Ο υπολογισμός του inception score έγινε μόνο στην τελευταία εποχή κάθε μοντέλου καθώς η εκτέλεση του κώδικα για τον υπολογισμό απαιτεί αρκετό χρόνο και αν πραγματοποιούταν σε κάθε εποχή κατά την εκπαίδευση για κάθε μοντέλο, θα απαιτούσε πολύ περισσότερο χρόνο και υπολογιστικούς πόρους. Στον πίνακα 4-12 παρουσιάζονται τα αποτελέσματα για κάθε σύνολο δεδομένων (ΣΔ1, ΣΔ2 και ΣΔ3) και για κάθε αρχιτεκτονική που παρουσιάστηκε στις προηγούμενες ενότητες.

Για τον υπολογισμό του Inception Score, χρησιμοποιείται το μοντέλο Inception v3 προεκπαιδευμένο στο σύνολο δεδομένων ImageNet [47]. Το σύνολο αυτό περιλαμβάνει εικόνες από 1000 διαφορετικές κλάσεις, αλλά δεν περιλαμβάνει εικόνες με μαγνητικές τομογραφίες εγκεφάλων. Αυτό αποτελεί πρόβλημα όπως έχει διερευνηθεί στη μελέτη [55], καθώς τα αποτελέσματα μπορεί να είναι παραπλανητικά, διότι μπορεί να δοθεί υψηλή τιμή για εικόνες όχι πολύ καλής ποιότητας.

Για να ελεγχθεί αν έχει νόημα η υλοποίηση του ελέγχου πρώτα πραγματοποιήθηκαν δοκιμές στις πραγματικές εικόνες και των τριών συνόλων δεδομένων, ώστε να χρησιμοποιηθούν σαν σημείο αναφοράς, οι τιμές αυτές παρουσιάζονται στον πίνακα 4-12 στη γραμμή με όνομα «Real Images». Επίσης για να ελεγχθεί αν η τιμή επηρεάζεται από την ποιότητα των εικόνων πραγματοποιήθηκαν δοκιμές σε αρχικές εποχές των ίδιων δικτύων όπου οι παραγόμενες τεχνητές εικόνες ήταν αρκετά θολές και δεν είχε προλάβει

το δίκτυο να αποτυπώσει το πλήρες σχήμα του εγκεφάλου. Οι δοκιμές αυτές ήταν επιτυχημένες καθώς όπως ήταν αναμενόμενο οι τιμές του Inception Score ήταν πολύ χαμηλές πολύ κοντά στο 1.02, κάτι το οποίο σημαίνει ότι η ποιότητα επηρεάζει την τιμή.

Αναλύοντας τα αποτελέσματα του Inception Score παρατηρείται ότι για το σύνολο ΣΔ1, για το οποίο δημιουργήθηκαν εικόνες με τις αρχιτεκτονικές WGANGP 1 και 2, το IS της τελευταίας εποχής είναι 1.57 και 1.65 αντίστοιχα. Το IS στις δύο περιπτώσεις αυτές έχει μικρή διαφορά και είναι αρκετά κοντά σε αυτό των πραγματικών δεδομένων 1.72, κάτι που επιτρέπει να υποθέσουμε ότι η εικόνες έχουν πλησιάσει αρκετά τις πραγματικές. Στη συνέχεια στο σύνολο ΣΔ2 δοκιμάστηκαν οι αρχιτεκτονικές WGANGP1 και WGANGP 2, με το IS score να λαμβάνει τιμές 1.75 και 1.74 αντίστοιχα. Παρόμοια και σε αυτήν την περίπτωση το IS είναι πολύ κοντά για τις δύο αρχιτεκτονικές και πολύ κοντά στο IS των πραγματικών εικόνων. Τέλος για το σύνολο ΣΔ3 πραγματοποιήθηκε παραγωγή τεχνητών εικόνων με χρήση των WGANGP 1, CWGANGP και LSGAN, με τιμές IS 2.20, 2.28 και 2.31. Οι τιμές της μετρικής IS, αντιστοιχούν και με τα οπτικά αποτελέσματα, καθώς οι τεχνητές εικόνες από το LSGAN δείχνουν πιο «καθαρές» και με καλύτερη λεπτομέρεια

Το τελευταίο στάδιο της πειραματικής διαδικασίας είναι η χρήση των τεχνητών εικόνων που δημιουργήθηκαν από τα επιλεγμένα GAN, στην εκπαίδευση του ταξινομητή Resnet18, δοκιμάζοντας 5 διαφορετικά ποσοστά τεχνητών εικόνων στο σύνολο εκπαίδευσης, 10%, 20%, 50%, 80% και 100%. Επίσης πραγματοποιήθηκε η εκπαίδευση του ταξινομητή με τη χρήση του αρχικού συνόλου χωρίς τεχνητές εικόνες και με διατήρηση της ανισορροπίας, ώστε να χρησιμοποιηθούν οι μετρικές για τη σύγκριση με τα υπόλοιπα πειράματα. Επιπλέον, για τα ίδια ποσοστά έγιναν πειράματα έχοντας εξισορροπήσει το δείγμα, δηλαδή αφαιρέθηκαν τυχαία εικόνες από την πλειονοτική κλάση ώστε ο αριθμός του δείγματος να είναι ίδιος με αυτόν της μειονοτικής κλάσης. Η σύγκριση των αποτελεσμάτων θα γίνει ανά σύνολο δεδομένων. Στους πίνακες που παρουσιάζονται στην ενότητα 4.5, έχουν συλλεχθεί οι μετρικές από την εκτέλεση του ταξινομητή στο test σύνολο δεδομένων και κάθε πίνακας αφορά μια αρχιτεκτονική και ένα σύνολο δεδομένων, με αποτελέσματα από την εκτέλεση σε εξισορροπημένο σύνολο δεδομένων και σε μη εξισορροπημένο. Οι μετρικές που θα δοθεί μεγαλύτερη σημασία είναι η Recall και το

Balanced Accuracy, διότι στο σύνολο δοκιμών (test set) έχει διατηρηθεί η ανισότητα των κλάσεων για τη δοκιμή σε πραγματικές συνθήκες.

Στο σύνολο ΣΔ1 εκπαιδεύτηκαν αρκετές αρχιτεκτονικές με τις WGANGP 1 και 2 να έχουν καταφέρει τα καλύτερα αποτελέσματα σε ποιότητα τεχνητών εικόνων, ώστε να κριθεί ωφέλιμο να χρησιμοποιηθούν οι τεχνητές εικόνες για την εκπαίδευση του ταξινομητή. Σαν βάση σύγκρισης, όπως και στα επόμενα αποτελέσματα, θα λαμβάνεται η απόδοση του ταξινομητή στις δοκιμές χωρίς προσθήκη τεχνητών εικόνων (0% στους πίνακες). Από τους πίνακες 4-13 και 4-14, παρατηρείται ότι στο ΣΔ1 και οι δύο αρχιτεκτονικές έχουν παρόμοια απόδοση. Αρχικά, είναι φανερό ότι η αύξηση του ποσοστού των τεχνητών εικόνων πάνω από 10% μειώνει το Recall και το BAC. Επίσης, και στους δύο πίνακες τόσο σε ισορροπημένο όσο και σε μη ισορροπημένο σύνολο, το ποσοστό με τις καλύτερες τιμές στις μετρικές που παρατηρούνται είναι το 10%, δηλαδή το σύνολο της κλάσης AD (ασθενείς) έχει εμπλουτιστεί και περιέχει 10% τεχνητές εικόνες. Τέλος παρατηρείται ότι και στα δύο πειράματα, δηλαδή με τεχνητές εικόνες από 2 διαφορετικά μοντέλα τα αποτελέσματα είναι αρκετά κοντά.

Το σύνολο δεδομένων ΣΔ2, αποτελείται από εικόνες όμοιες με αυτές του ΣΔ1, αλλά όπως προαναφέρθηκε στην περιγραφή των δεδομένων έχει μικρότερο αριθμό δειγμάτων διότι η επιλογή ασθενών και υγείων ατόμων είναι αυστηρότερη. Παρόμοια με τα αποτελέσματα στην ενότητα 4.5.1, φαίνεται και σε αυτή τη περίπτωση οι δύο αρχιτεκτονικές να έχουν δημιουργήσει τεχνητές εικόνες, οι οποίες επηρεάζουν με παρόμοιο τρόπο τον ταξινομητή. Λαμβάνοντας υπόψιν τις μετρικές Recall και BAC, είναι φανερό ότι για τις περιπτώσεις που διατηρούμε το σύνολο χωρίς εξισορρόπηση, καλύτερα αποτελέσματα εμφανίζονται στην περίπτωση όπου οι τεχνητές εικόνες αποτελούν το 80% του δείγματος των ασθενών. Στην αρχιτεκτονική WGANGP 2 παρατηρείται ότι τα ποσοστά που μας ενδιαφέρουν είναι ελάχιστα καλύτερα, έχοντας όμως επηρεάσει τα αποτελέσματα της κλάσης των υγείων. Όσον αφορά το εξισορροπημένο σύνολο, παρατηρείται πάλι η συμπεριφορά της μείωσης του Recall σε μεγαλύτερα ποσοστά τεχνητών εικόνων. Σύμφωνα με τις υπογραμμισμένες τιμές καλύτερη τιμή Recall εμφανίζεται στο 10% και στις δυο αρχιτεκτονικές. Παρατηρώντας όμως την τιμή BAC, παίρνει καλύτερη τιμή στο 20% με ελάχιστα μικρότερο Recall, το οποίο όμως είναι μεγαλύτερο από το σημείο

αναφοράς. Άρα θεωρείται καλύτερο αποτέλεσμα το 20%, έχοντας πετύχει αύξηση του Recall κατά 3% του BAC κατά 2% του F1 score κατά 2%, μείωση της απώλειας και αύξηση της ακρίβειας κατά 1%.

Στην περίπτωση του συνόλου ΣΔ3, πραγματοποιήθηκαν 3 πειράματα με διαφορετικές αρχιτεκτονικές. Αρχικά, οι τεχνητές εικόνες και από τις 3 δοκιμές παρατηρείται ότι έχουν παρόμοια επιρροή στον ταξινομητή, καθώς φαίνεται τα ποσοστά των μετρικών να έχουν μικρή διαφορά. Και στις τρεις περιπτώσεις δε φαίνεται να βελτιώνεται η απόδοση στις μετρικές αυξάνοντας το ποσοστό τεχνητών εικόνων, με τις δύο πρώτες να έχουν επιτύχει ελάχιστα καλύτερο Recall αλλά ταυτόχρονα μικρότερο BAC. Από τις 3 προσπάθειες φαίνεται ότι οι τεχνητές εικόνες από την αρχιτεκτονική LSGAN, έχουν καλύτερη επιρροή στον ταξινομητή σε σχέση με τις άλλες δύο.

6. Συμπεράσματα και μελλοντική εργασία

Οι δυνατότητες των GAN είναι φανερό πως ποικίλουν και δύναται να εφαρμοστούν σε πολλούς διαφορετικούς τομείς, συνεισφέροντας σημαντικά στην επίλυση προβλημάτων όπως αυτό που πραγματεύεται αυτή η εργασία, της ανισότητας των κλάσεων ιατρικών δεδομένων. Το ερευνητικό ενδιαφέρον φαίνεται να είναι μεγάλο, κάτι πολύ ελπιδοφόρο για μελλοντικές έρευνες, τόσο για την ανάπτυξη του ίδιου του κλάδου των παραγωγικών μοντέλων όσο και για τη συνεισφορά σε άλλους τομείς.

Στην παρούσα διπλωματική εργασία αναπτύχθηκαν 6 κύριες διαφορετικές αρχιτεκτονικές με σκοπό την επίλυση του προβλήματος της ανισότητας των κλάσεων στο σύνολο δεδομένων OASIS-3, όπου ο αριθμός των υγείων υπερτερεί του αριθμού των ασθενών ατόμων. Δημιουργήθηκαν 3 διαφορετικά σύνολα δεδομένων (ΣΔ1, ΣΔ2 και ΣΔ3) αποτελούμενα από εικόνες δύο διαστάσεων με τομές από διαφορετικά επίπεδα του εγκεφάλου. Η δημιουργία τους ήταν εφικτή καθώς το αρχικό σύνολο αποτελούταν από αρχεία μαγνητικών τομογραφιών τριών διαστάσεων, κάτι που επέτρεψε την εξαγωγή συγκεκριμένων τομών (slices). Η μέθοδος επίλυσης που ακολουθείται είναι η αύξηση του αριθμού των εικόνων στην κλάση των ασθενών, με χρήση τεχνητών εικόνων που παράχθηκαν από τα επιλεγμένα μοντέλα GAN.

Η δημιουργία και η εκπαίδευση των GAN είναι μια περίπλοκη και χρονοβόρα διαδικασία. Αυτό οφείλεται στο γεγονός ότι για την βελτιστοποίηση μιας αρχιτεκτονικής απαιτείται η δοκιμή πολλών υπερπαραμέτρων κάτι που σε συνδυασμό με την αργή εκπαίδευση, απαιτεί σημαντικούς υπολογιστικούς πόρους και αρκετό χρόνο. Οι αρχιτεκτονικές επιλέχθηκαν βάση αποτελεσματικότητας, κάτι που αξιολογήθηκε βάση προηγούμενων μελετών και δοκιμάστηκαν στα 3 σύνολα δεδομένων, όπου αυτό ήταν εφικτό. Κατά τη διάρκεια των δοκιμών παρατηρήθηκε ότι η χρήση μιας αρχιτεκτονικής σε διαφορετικά σύνολα δεδομένων, διατηρώντας τις ίδιες υπερπαραμέτρους, είχε παρόμοια απόδοση, όπως για παράδειγμα η αρχιτεκτονική WGANP1 που χρησιμοποιήθηκε και στα 3 σύνολα. Όσον αφορά τις υπερπαραμέτρους και την επιρροή που έχει η αλλαγή των τιμών τους στις τεχνητές εικόνες, μεγαλύτερη επίδραση παρατηρήθηκε με τη μεταβολή του αριθμού των κρυφών επιπέδων (hidden layers) στα συνελκτικά μοντέλα, με καλύτερα αποτελέσματα να επιτυγχάνονται σε μεγαλύτερο αριθμό κρυφών επιπέδων. Επίσης, η

αλλαγή του learning rate, φαίνεται να επιδρά στην εκπαίδευση, με χαμηλότερες τιμές να βοηθούν στην ομαλή εκπαίδευση, αλλά ταυτόχρονα να απαιτούν περισσότερο χρόνο, ενώ αντίθετα υψηλές τιμές να παρέχουν γρηγορότερη εκπαίδευση αλλά με κίνδυνο κατάρρευσης του δικτύου, κάτι που αντιμετωπίστηκε αρκετά συχνά κατά τη διάρκεια των πειραμάτων. Το πρόβλημα της κατάρρευσης όπως περιγράφεται και στην προηγούμενη ενότητα, παρουσιάζει ξαφνική αλλοίωση των τεχνητών εικόνων ενώ το δίκτυο αρχικά είχε καταφέρει να παράξει καλής ποιότητας εικόνες. Η αλλαγή των καναλιών του θορύβου που εισάγεται στον Generator φαίνεται να επιδρά στην ποιότητα των αποτελεσμάτων, με βέλτιστη τιμή τα 256 κανάλια. Η επιλογή μικρότερου αριθμού επηρεάζει αρνητικά, ενώ μεγαλύτερες τιμές αυξάνουν εκθετικά τον χρόνο εκπαίδευσης και δεν παρουσιάζουν σημαντική βελτίωση στα αποτελέσματα. Το Batch Size επηρεάζει την ταχύτητα της εκπαίδευσης και η βέλτιστη τιμή βάση των υπολογιστικών πόρων είναι 64. Για τις αρχιτεκτονικές που βασίζονται στην απόσταση Wasserstein σημαντικό ρόλο παίζουν και οι τιμές Critic Iterations και Lambda, οι οποίες καθορίζουν το πόσο ισχυρός είναι ο Critic, με βέλτιστη τιμή 10 και για τις δύο παραμέτρους. Τέλος, μια ακόμα παράμετρος που επηρεάζει σημαντικά τα αποτελέσματα είναι το μέγεθος των πραγματικών εικόνων που χρησιμοποιούνται για την εκπαίδευση των GAN. Στα αρχικά πειράματα δοκιμάστηκε μέγεθος εικόνων 64x64, κάτι που οδήγησε σε αποτελέσματα χωρίς πολλές λεπτομέρειες. Έπειτα πραγματοποιήθηκαν δοκιμές με μεγέθη 128x128 και 224x224, με τα δύο αυτά μεγέθη να αποδίδουν καλύτερα αποτελέσματα. Τελικά χρησιμοποιήθηκε το μέγεθος 128x128 καθώς οι τεχνητές εικόνες ήταν ποιοτικές και με λεπτομέρειες. Φαίνεται πως η χρήση εικόνων μεγαλύτερων διαστάσεων βελτιώνει τα αποτελέσματα αλλά ταυτόχρονα απαιτεί περισσότερο χρόνο για την εκπαίδευση των GAN, κάτι το οποίο αποτελεί το κύριο λόγο που επιλέχθηκε το μέγεθος 128x128. Δεδομένου ότι οι αρχικές διαστάσεις των εικόνων είναι 256x256, η χρήση εικόνων μεγαλύτερων διαστάσεων ώστε να εκμεταλλευτεί όλη η διαθέσιμη πληροφορία είναι πιθανό να αποφέρει καλύτερα αποτελέσματα και θα μπορούσε να μελετηθεί σε μελλοντικές εργασίες.

Η αξιολόγηση των τεχνητών εικόνων που παράχθηκαν από τα επιλεγμένα GAN πραγματοποιήθηκε με κριτήριο την επίδρασή τους στις μετρικές του ταξινομητή, από την τιμή του Inception Score (IS), αλλά και με οπτική αξιολόγηση. Μελετώντας τα αποτελέσματα ανά σύνολο δεδομένων, σύμφωνα με το Inception Score, για το ΣΔ1,

καλύτερα αποτελέσματα τεχνητών εικόνων παρήγαγε η αρχιτεκτονική WGANGP2 με τιμή 1.62. Για το ΣΔ2 η WGANGP1 απέδωσε καλύτερα με τιμή 1.75, ενώ για το ΣΔ3 η LSGAN με τιμή 1.31. Η οπτική αξιολόγηση συμβαδίζει με τις τιμές του IS, καθώς σε μεγαλύτερες τιμές παρατηρήθηκαν καλύτερες τεχνητές εικόνες.

Οι παραγόμενες τεχνητές εικόνες χρησιμοποιήθηκαν για τον εμπλουτισμό των συνόλων δεδομένων σε 5 διαφορετικά ποσοστά, 10% 20%, 50%, 80% και 100%. Επίσης πραγματοποιήθηκε και η δημιουργία εξισορροπημένων συνόλων για κάθε εμπλουτισμένο σύνολο, αφαιρώντας τα επιπλέον δείγματα από την κλάση των υγιών. Εκπαιδεύθηκαν ταξινομητές resnet18 για κάθε νέο εμπλουτισμένο σύνολο για 20 εποχές και έπειτα δοκιμάστηκαν στο ίδιο δοκιμαστικό σύνολο δεδομένων (test set). Τα αποτελέσματα από το test set συγκρίθηκαν με αυτά του ταξινομητή που εκπαιδεύτηκε στα αρχικά σύνολα δεδομένων. Μελετώντας τα αποτελέσματα ανά σύνολο δεδομένων, για το ΣΔ1 όπου χρησιμοποιήθηκαν τεχνητές εικόνες από τις αρχιτεκτονικές WGANGP1 και WGANGP2, φαίνεται να είναι παρόμοια. Κρίνοντας βάση του BAC βέλτιστες τιμές εμφανίστηκαν στο εμπλουτισμένο σύνολο με 10% τεχνητές εικόνες, με τιμή 63% έναντι του μη εμπλουτισμένου που πέτυχε 61%. Υψηλότερη απόδοση φαίνεται να έχουν τα εξισορροπημένα σύνολα, πάλι με το ποσοστό 10% τεχνητών εικόνων να έχει καλύτερη τιμή BAC 65%, έναντι του αρχικού συνόλου 64%. Στο σύνολο ΣΔ2 κρίνοντας πάλι μόνο βάση του BAC, καλύτερη απόδοση είχε ο ταξινομητής που εκπαιδεύτηκε στο εμπλουτισμένο σύνολο με 80% τεχνητές εικόνες από την αρχιτεκτονική WGANGP2, με τιμή BAC 77%, ενώ ο ταξινομητής που εκπαιδεύτηκε στο αρχικό σύνολο δεδομένων πέτυχε 72%. Καλύτερη απόδοση έχει το εξισορροπημένο σύνολο με 20% τεχνητές εικόνες από την αρχιτεκτονική WGANGP2, με BAC 80%, έναντι του 78% που κατάφερε το μη εμπλουτισμένο εξισορροπημένο σύνολο. Τέλος στο σύνολο ΣΔ3 παρατηρείται ότι οι τεχνητές εικόνες και από τις 3 αρχιτεκτονικές που δοκιμάστηκαν δεν βελτίωσαν τα αποτελέσματα, με μόνη εξαίρεση την περίπτωση του εμπλουτισμένου συνόλου με 10% τεχνητές εικόνες από το LSGAN, όπου παρατηρήθηκε τιμή BAC 60.4%, έναντι του 60% που παρατηρήθηκε στο μη εμπλουτισμένο σύνολο.

Η προσθήκη τεχνητών εικόνων από τα GAN στο σύνολο εκπαίδευσης του ταξινομητή, επιδρά θετικά στην απόδοσή του καθώς φαίνεται να βελτιώνει την ικανότητά του να

αναγνωρίζει περισσότερες περιπτώσεις ασθενών, σε σύγκριση με την απόδοσή του έχοντας εκπαιδευτεί στο μη εμπλουτισμένο σύνολο δεδομένων.

Μελλοντικές κατευθύνσεις ερευνητικών εφαρμογών, περιλαμβάνουν τη χρήση των μοντέλων αυτών σε διαφορετικά σύνολα δεδομένων μαγνητικών τομογραφιών, ώστε να εκμεταλλευθούν τα βάρη από την υπάρχουσα εκπαίδευση. Ακόμα προτείνεται η δοκιμή των τεχνητών εικόνων που δημιουργήθηκαν σε διαφορετικούς ταξινομητές ώστε να διερευνηθεί και η επιρροή των εικόνων σε άλλες αρχιτεκτονικές ταξινόμησης. Σαν μελλοντική βελτίωση επίσης θα μπορούσε να πραγματοποιηθεί η επιλογή των καλύτερων τεχνητών εικόνων με χρήση διαφόρων τεχνικών αξιολόγησης, ώστε να αποφευχθούν χαμηλής ποιότητας εικόνες. Τέλος θα μπορούσε να διερευνηθεί η επιρροή τόσο των προαναφερόμενων μοντέλων GAN όσο και της απόδοσης του ταξινομητή από τη χρήση κλασικών τεχνικών επαύξησης δεδομένων (Data Augmentation).

7. Βιβλιογραφία

- [1] N. Dhungel, G. Carneiro και A. P. Bradley, «Automated Mass Detection in Mammograms Using Cascaded Deep Learning and Random Forests,» σε *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 2015.
- [2] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin και H. Larochelle, «Brain tumor segmentation with Deep Neural Networks,» *Medical Image Analysis*, τόμ. 35, pp. 18-31, 2017.
- [3] M. R. Cooperberg, S. J. Freedland, D. J. Pasta, E. P. Elkin, J. C. Presti, C. L. Amling, M. K. Terris, W. J. Aronson, C. J. Kane και P. R. Carroll, «Multiinstitutional validation of the UCSF cancer of the prostate risk assessment for prediction of recurrence after radical prostatectomy,» *Cancer*, τόμ. 107, αρ. 10, pp. 2384-2391, 2006.
- [4] Q. Sun, M. V. Jankovic, L. Bally και S. G. Mougiakakou, «Predicting Blood Glucose with an LSTM and Bi-LSTM Based Deep Neural Network,» σε *2018 14th Symposium on Neural Networks and Applications (NEUREL)*, 2018.
- [5] M. Kubat, *An introduction to machine learning*, Springer, 2015.
- [6] Y. Freund και R. E. Schapire, «A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,» *Journal of Computer and System Sciences*, τόμ. 55, αρ. 1, pp. 119-139, 1997.
- [7] F.-A. Maayan, K. Eyal, A. Michal, G. Jacob και G. Hayit, «Synthetic data augmentation using GAN for improved liver lesion classification,» *IEEE Journal of Biomedical and Health Informatics*, pp. 289-293, 2018.

- [8] X. Zhou, S. Qiu, P. S. Joshi, C. Xue, R. J. Killiany, A. Z. Mian, S. P. Chin, R. Au και V. B. Kolachalama, «Enhancing magnetic resonance imaging-driven Alzheimer's disease classification performance using generative adversarial learning,» *Alz Res Therapy*, 2021.
- [9] E. Yagis, L. Citi, S. Diciotti, C. Marzi, S. Workalemahu Atnafu και A. G. Seco De Herrera, «3D Convolutional Neural Networks for Diagnosis of Alzheimer's Disease via Structural MRi,» *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 65-70, 2020.
- [10] B. Khagi, B. Lee, J.-Y. Pyun, G.-R. Kwon και G.-R. Kwon, «CNN Models Performance Analysis on MRI images of OASIS dataset for distinction between Healthy and Alzheimer's patient,» σε *2019 International Conference on Electronics, Information, and Communication (ICEIC)*, 2019.
- [11] S. J. Taylor και B. Letham, «Forecasting at Scale,» *The American Statistician*, τόμ. 72, αρ. 1, pp. 37-45, 2018.
- [12] Y. Ozaki, H. Yamada, H. Kikuchi, A. Hirotsu, T. Murakami, T. Matsumoto, T. Kawabata, Y. Hiramatsu, K. Kamiya, T. Yamauchi, K. Goto, Y. Ueda, S. Okazaki, M. Kitagawa, H. Takeuchi και H. Konno, «Label-free classification of cells based on supervised machine learning of subcellular structures,» *PLOS ONE*, τόμ. 14, p. e0211347, {2019},.
- [13] I. Sutskever, O. Vinyals και Q. V. Le, Sequence to Sequence Learning with Neural Networks, arXiv:1409.3215, 2014.
- [14] R. Mushfiqur και A. Md, Implementation of Artificial Neural Network on Regression Analysis, IEEE, 2021.
- [15] G. Jiao και W. Li, «Neural Network Data Mining Clustering Optimization Algorithm,» *IETE Journal of Research*, τόμ. 0, pp. 1-11, 2021.

- [16] M. Kelly, R. Longjohn και K. Nottingham, «The UCI Machine Learning Repository,» [Ηλεκτρονικό]. Available: <https://archive.ics.uci.edu>.
- [17] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth και G. Langs, Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery, arXiv:1703.05921v1, 2017.
- [18] C. Berner, G. Brockman, B. Chan και V. Cheung, Dota 2 with Large Scale Deep Reinforcement Learning, arXiv:1912.06680, 2019.
- [19] A. Kendall, J. Hawke, D. Janz, P. Mazur, D. Reda, J.-M. Allen, V.-D. Lam, A. Bewley και A. Shah, Learning to Drive in a Day, arXiv:1807.00412, 2018.
- [20] J. Li, W. Monroe, A. Ritter, M. Galley, J. Gao και D. Jurafsky, Deep Reinforcement Learning for Dialogue Generation, arXiv:1606.01541, 2016.
- [21] C. L. Saratxaga, I. Moya, A. Picón, M. Acosta, A. Moreno-Fernandez-de-Leceta, E. Garrote and A. Bereciartua-Perez, "MRI Deep Learning-Based Solution for Alzheimer's," *J. Pers. Med.*, vol. 11, no. 9, p. 902, 2021.
- [22] D. S. Marcus, A. F. Fotenos, J. G. Csernansky, J. C. Morris και R. L. Buckner, «Open Access Series of Imaging Studies: Longitudinal MRI Data in Nondemented and Demented Older Adults,» *J Cogn Neurosci*, τόμ. 22, αρ. 12, p. 2677–2684, 2010.
- [23] P. J. LaMontagne, T. L. Benzinger, J. C. Morris, S. Keefe, R. Hornbeck, C. Xiong, E. Grant, J. Hassenstab, K. Moulder, A. G. Vlassenko, M. E. Raichle, C. Cruchaga και D. Marcus, «OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease,» *medRxiv* , 2019.
- [24] H. Kaiming, Z. Xiangyu, R. Shaoqing και S. Jian, Deep Residual Learning for Image Recognition, arXiv:1512.03385, 2015.

- [25] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville και Y. Bengio, «Generative Adversarial Networks,» arXiv:1406.2661, 2014.
- [26] J. Langr και V. Bok, GANs in Action, Manning Publications Co., September 2019.
- [27] A. Radford, L. Metz και S. Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, arXiv:1511.06434, 2016.
- [28] L. Bingqi , L. Jiwei , F. Xinyue , L. Jie και Z. Tianyi , «Application of an Improved DCGAN for Image Generation,» τόμ. 2022, 2022.
- [29] M. Arjovsky, S. Chintala και L. Bottou, Wasserstein GAN, arXiv:1701.07875, 2017.
- [30] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin και A. Courville, Improved Training of Wasserstein GANs, arXiv:1704.00028 [cs.LG], 2017.
- [31] C. R. Jack Jr, M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, A. M. Dale, J. P. Felmlee, J. L. Gunter, D. L. G. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. DeCarli, K. Gunnar, H. A. Ward, G. J. Metzger, K. T. Scott, R. Mallozzi, D. Blezek, J. Levy, J. P. Debbins, A. S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler και M. W. Weiner, «The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods,» *J. Magn. Reson. Imaging*, τόμ. 27, αρ. 4, pp. 685-691, 2008.
- [32] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang και S. P. Smolley, Least Squares Generative Adversarial Networks, arXiv:1611.04076, 2017.

- [33] D. Berthelot, T. Schumm και L. Metz, BEGAN: Boundary Equilibrium Generative Adversarial Networks, arXiv:1703.10717, 2017.
- [34] K. Gihyun , H. Chihye και K. Dae-shik, Generation of 3D Brain MRI Using Auto-Encoding Generative Adversarial Networks, arXiv:1908.02498, 2019.
- [35] H. Ilyass, G. Jayant Sen και V. Feuillard, Anomaly detection with Wasserstein GAN, arXiv:1812.02463, 2018.
- [36] H. B. Hussain, A. Ifrah , S. Mohammad , A. M. Joshi και N. Kanwal, «Wasserstein GAN based Chest X-Ray Dataset Augmentation for Deep Learning Models: COVID-19 Detection Use-Case,» σε *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2022, pp. 2058-2061.
- [37] M. Mehdi και O. Simon, «Conditional Generative Adversarial Nets,» arXiv:1411.1784 [cs, stat], 2014.
- [38] H.-C. Shin, A. Ihsani, Z. Xu, S. Mandava, S. T. Sreenivas, C. Forster και J. Cha, GANDALF: Generative Adversarial Networks with Discriminator-Adaptive Loss Fine-tuning for Alzheimer's Disease Diagnosis from MRI, arXiv:2008.04396, 2020.
- [39] M. Zheng, T. Li, R. Zhu, Y. Tang, M. Tang, L. Lin και Z. Ma, «Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification,» *Information Sciences*, τόμ. 512, pp. 1009-1023, 2020.
- [40] C. Fabbri, Conditional Wasserstein Generative Adversarial Networks.
- [41] R. Feng, D. Zhao και Z.-J. Zha, «On Noise Injection in Generative Adversarial Networks,» 2021.

- [42] S. Jenni και P. Favaro, «On Stabilizing Generative Adversarial Training with Noise,» *CoRR*, τόμ. abs/1906.04612, 2019.
- [43] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever και R. Salakhutdinov, «Improving neural networks by preventing co-adaptation of feature detectors,» *CoRR*, τόμ. abs/1207.0580, 2012.
- [44] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, Kopf, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai και S. Chintala, «PyTorch: An Imperative Style, High-Performance Deep Learning Library,» σε *Advances in Neural Information Processing Systems 32*, Curran Associates, Inc., 2019, pp. 8024--8035.
- [45] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang και S. P. Smolley, «Least Squares Generative Adversarial Networks,» σε *2017 IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 2813-2821.
- [46] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang και S. P. Smolley, «On the Effectiveness of Least Squares Generative Adversarial Networks,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, τόμ. 41, αρ. 12, pp. 2947-2960, 2019.
- [47] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li και L. Fei-Fei, «Imagenet: A large-scale hierarchical image database,» σε *2009 IEEE conference on computer vision and pattern recognition*, IEEE, 2009, pp. 248--255.
- [48] *FreeSurfer software suite*.
- [49] M. Mendez, «Chapter 16 - General Mental Status Scales, Rating Instruments, and Behavior Inventories,» σε *The Mental Status Examination Handbook*, Philadelphia, Elsevier, 2022, pp. 181-199.

- [50] A. M. Dale, B. Fischl και M. I. Sereno, «Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction,» *NeuroImage*, τόμ. 9, αρ. 2, pp. 179-194, 1999.
- [51] F. Li και M. Liu, «Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks,» *Comput Med Imaging Graph*, pp. 101-110, 2018.
- [52] G. Huang, Z. Liu, L. van der Maaten και K. Q. Weinberger, Densely Connected Convolutional Networks, arXiv:1608.06993 [cs.CV], 2018.
- [53] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens και Z. Wojna, Rethinking the Inception Architecture for Computer Vision, arXiv:1512.00567, 2015.
- [54] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz, Y. Jia, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, M. Schuster, R. Monga, S. Moore, D. Murray, C. Olah, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu και X. Zheng, {TensorFlow}: Large-Scale Machine Learning on Heterogeneous Systems, Software available from tensorflow.org, 2015.
- [55] S. Barratt και R. Sharma, «A Note on the Inception Score,» τόμ. arXiv:1801.01973, arXiv:1801.01973, 2018.
- [56] C. C. Aggarwal, *Neural Networks and Deep Learning*, Springer Cham, 2018.
- [57] I. H. Witten, F. Eibe, M. A. Hall και C. . J. Pal, *Data Mining (Fourth Edition)*, Morgan Kaufmann, 2017.
- [58] A. S. Schachter και K. L. Davis, «Alzheimer's disease,» *Dialogues in clinical neuroscience*, τόμ. 2, αρ. 2, pp. 91-100, 2000.

- [59] X. Zhou, S. Qiu, . P. S. Joshi, C. Xue, R. J. Killiany, A. Z. Mian, S. P. Chin, R. Au και V. B. Kolachalama, *Alz Res Therapy*, 2021.
- [60] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris και R. L. Buckner, «Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults,» *J Cogn Neurosci*, τόμ. 19, αρ. 9, p. 1498–1507, 2007.
- [61] A. Krizhevsky, I. Sutskever και G. E. Hinton, «ImageNet Classification with Deep Convolutional Neural Networks,» σε *Advances in Neural Information Processing Systems*, τόμ. 25, Curran Associates, Inc., 2012.
- [62] L. Khedher, J. Ramírez, J. M. Górriz, A. Brahim και F. Segovia, «Early diagnosis of Alzheimer's disease based on partial least squares, principal component analysis and support vector machine using segmented MRI images,» *Neurocomputing*, τόμ. 151, pp. 139-150, 2015.
- [63] S. Wang, D. Jia και X. Weng, Deep Reinforcement Learning for Autonomous Driving, arXiv:1811.11329, 2019.
- [64] P. Zhong, D. Wang και C. Miao, An Affect-Rich Neural Conversational Model with Biased Attention and Weighted Cross-Entropy Loss, arXiv:1811.07078, 2018.
- [65] Y. Djenouri, A. Belhadi και L. Jerry Chun Wei, «Recurrent neural network with density-based clustering for group pattern detection in energy systems,» *Sustainable Energy Technologies and Assessments*, τόμ. 52, αρ. Part D, p. 102308, 2022.