



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

«Παραγωγή Συνθετικών Ιατρικών Δεδομένων με Παραγωγικά Ανταγωνιστικά Δίκτυα»

Εξαμηνιαία Εργασία

για το μάθημα «Τεχνολογίες Κινητής και Ηλεκτρονικής Υγείας»

των σπουδαστών

Χαραλάμπου Δάλη el17067

Αντωνίου Καραντώνη el17439

Δημητρίου Κυριακίδη el17077

Διδάσκοντες: Δ. Κουτσούρης, Κ. Νικήτα, Γ. Ματσόπουλος

Επιβλέποντες: Κ. Μήτσης, Ν. Παπαδόπουλος, Π. Καρατζά

ΠΕΡΙΛΗΨΗ

Σκοπός της εργασίας είναι η παραγωγή συνθετικών ιατρικών δεδομένων, συγκεκριμένα απεικονίσεων υπερήχων μαστού, για την χρήση τους ως δεδομένα εκπαίδευσης ή αξιολόγησης άλλων συστημάτων τεχνητής νοημοσύνης. Αυτές οι απεικονίσεις είναι μαστών με καλοήγη ή κακοήγη όγκο. Ο βασικός λόγος που τέτοιου είδους δεδομένα είναι χρήσιμα, αν όχι απαραίτητα, είναι ότι τα σύγχρονα συστήματα τεχνητής νοημοσύνης απαιτούν μεγάλο όγκο δεδομένων εκπαίδευσης, συχνά μεγαλύτερο από αυτόν που είναι διαθέσιμος στην πράξη. Προς τον σκοπό αυτόν, αξιοποιείται η τεχνολογία των Παραγωγικών Ανταγωνιστικών Δικτύων (Generative Adversarial Networks - GANs), η οποία έχει βρει πολλές εφαρμογές στο πεδίο κατασκευής τεχνητών εικόνων. Αρχικά, πραγματοποιούμε σύγκριση των εναλλακτικών αρχιτεκτονικών GAN ως προς την παραγωγή αληθοφανών συνθετικών δεδομένων. Στη συνέχεια, δοκιμάζουμε διαφορετικές υλοποιήσεις των εν λόγω δικτύων με σκοπό την παραγωγή των επιθυμητών τεχνητών εικόνων, εκπαιδεύοντας και βελτιστοποιώντας τις αρχιτεκτονικές αυτές.

«Λέξεις-κλειδιά»: μηχανική μάθηση, παραγωγικά ανταγωνιστικά δίκτυα, συνθετικά δεδομένα, υπέρηχοι μαστού, καρκίνος του μαστού.

Όλα τα αρχεία κώδικα και αποτελέσματα της εργασίας έχουν δημοσιευθεί στο εξής github repository: <https://github.com/DimK19/suns-out-gans-out>

Ο πηγαίος κώδικας του τεχνήματος:
https://github.com/antokarant/mobile_health_artifact

Αναλυτική περιγραφή αυτού στο αρχείο APP_PROTOTYPE_SRS.pdf

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ	1
1 ΠΡΟΒΛΗΜΑ ΚΑΙ ΣΤΟΧΟΙ.....	3
1-1 Η νόσος	3
1-2 Αυτοματοποιημένη διάγνωση	3
1-3 Συνθετικά δεδομένα	4
1-4 Σκοπός	4
2 ΘΕΩΡΗΤΙΚΗ ΘΕΜΕΛΙΩΣΗ.....	5
2-1 Υπόβαθρο μεθόδων Νευρωνικών Δικτύων	5
2-2 Συνελκτικά Δίκτυα - CNNs	5
2-3 Ανταγωνιστικά Παραγωγικά Δίκτυα - GANs.....	5
2-4 Εκπαίδευση GANs	7
3 ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΔΟΜΗ	7
3-1 Σύνολο δεδομένων	7
3-2 Αρχιτεκτονικές GAN	8
3-3 Εκπαίδευση	10
4 ΑΠΟΤΕΛΕΣΜΑΤΑ.....	12
4-1 Δείγματα.....	12
4-2 Αξιολόγηση και πορίσματα.....	16
4-3 Επεκτάσεις.....	18
5 ΕΠΙΛΟΓΟΣ.....	18
6 ΠΑΡΑΡΤΗΜΑ.....	19
6-1 Ευρέως χρησιμοποιούμενες συναρτήσεις ενεργοποίησης.....	19
6-2 Μετρικές απώλειας.....	20
ΒΙΒΛΙΟΓΡΑΦΙΑ	21

1 ΠΡΟΒΛΗΜΑ ΚΑΙ ΣΤΟΧΟΙ

1-1 Η νόσος

Η σημασία της παρούσας εργασίας έγκειται στην ανάγκη αντιμετώπισης του καρκίνου του μαστού. Αυτό το είδος καρκίνου είναι το δεύτερο συχνότερα εμφανιζόμενο στις γυναίκες, με εκτιμώμενη πιθανότητα εμφάνισης 1/7 ανά άτομο, και πρώτο σε θνησιμότητα μεταξύ αυτών. Ο καρκίνος του μαστού αναπτύσσεται συνήθως από τα αδενικά κύτταρα των πόρων και των λοβιδίων (δομών που περιέχουν γαλακτοπαραγωγικούς αδένες) του μαστού. Σπανιότερα εμφανίζονται σαρκώματα στον λιπώδη ή μυϊκό ιστό. Σε αντίθεση με τους καλοήθεις όγκους, οι οποίοι είναι ομαλοί και δεν κάνουν μετάσταση, οι κακοήθεις όγκοι, ο καρκίνος του μαστού δηλαδή, διηθούν τα όμορα όργανα, και μετέπειτα μπορούν να κάνουν μετάσταση σχεδόν οπουδήποτε στο σώμα. Η διάγνωση της νόσου σε πρώιμο στάδιο φαίνεται πως είναι σωτήρια, καθώς οδηγεί σε υψηλότερα ποσοστά πενταετούς επιβίωσης, αλλά και πλήρους ίασης [11], [12].

1-2 Αυτοματοποιημένη διάγνωση

Την τελευταία δεκαετία τα νευρωνικά δίκτυα έχουν βρει σημαντική εφαρμογή σε θέματα αναγνώρισης χαρακτηριστικών σε εικόνες, και συνεχίζουν να έχουν ραγδαία ανάπτυξη. Αυτό φυσικά απασχολεί την Ιατρική, καθώς μεγάλο μέρος των διαγνώσεων αφορά παρατήρηση απεικονίσεων, και ειδικότερα αφορά την εργασία, αφού τα δεδομένα επί των οποίων αυτή στηρίζεται είναι απεικονίσεις υπερήχων μαστού. Οι υπέρηχοι μαστού, όπως και οι ακτινογραφίες, αποτελούν μέθοδο πρώτης ανίχνευσης του όγκου.

Το επικρατέστερο είδος νευρωνικών δικτύων στο πεδίο της ανάλυσης εικόνων είναι τα συνελικτικά νευρωνικά δίκτυα (convolutional neural networks - CNNs). Αυτά έχουν επιδείξει ικανότητα εξαγωγής χαρακτηριστικών από εικόνες και συνεπώς χρησιμοποιούνται ευρύτατα σε προβλήματα ταξινόμησης [3].

Προκειμένου να κατασκευασθεί ένας ταξινομητής υπερήχων που θα αποφαινεται σχετικά με την παρουσία κάποιας νόσου, θα πρέπει να εκπαιδευθεί ένα CNN ώστε να εξάγει χαρακτηριστικά από τους αντίστοιχους υπερήχους (απεικονίσεις υπερήχων), και προς τον σκοπό αυτόν απαιτούνται αρκετές χιλιάδες διαφορετικές εικόνες. Επομένως, η συλλογή κατάλληλων και επαρκών εικόνων για την εκπαίδευση τέτοιων δικτύων είναι ένα εξαιρετικά καίριο ζήτημα.

1-3 Συνθετικά δεδομένα

Το πρόβλημα έλλειψης δεδομένων εκπαίδευσης είναι εξαιρετικά σύνηθες στην μηχανική μάθηση. Οι διαθέσιμες ιατρικές απεικονίσεις είναι περιορισμένες, τόσο επειδή η συγκέντρωση συναφών πληροφοριών από όλον τον κόσμο είναι δύσκολη, όσο και λόγω της νομοθεσίας προστασίας προσωπικών δεδομένων και ιατρικού απορρήτου. Ως εκ τούτου, σε πολλές περιπτώσεις τα πραγματικά δεδομένα δεν επαρκούν για τέτοιου είδους εφαρμογές. Ακόμη και όταν μεγάλο πλήθος δεδομένων είναι διαθέσιμο, αυτά μπορεί να μην έχουν την επιθυμητή μορφή: συστηματικά σφάλματα, αλλοιώσεις των απεικονίσεων, ή ανεπάρκεια δειγμάτων με συγκεκριμένους συνδυασμούς χαρακτηριστικών είναι προβλήματα που δυσχεραίνουν την εκπαίδευση. Σε αυτές τις περιπτώσεις μπορούν να εφαρμοσθούν διάφορες τεχνικές αύξησης δεδομένων.

Κατ' αρχάς, ας γίνει διάκριση μεταξύ των όρων «αύξηση» (που χρησιμοποιήθηκε προηγουμένως) και «επαύξηση» (augmentation) δεδομένων: με τον πρώτο εννοούμε προσθήκη νέων δεδομένων στο σύνολο εκπαίδευσης, ενώ με τον δεύτερο μετασχηματισμό των υπαρχόντων (περιστροφή ή άλλου είδους αλλοίωση). Στην εργασία αυτή ασχολούμεθα κυρίως με την πρώτη τεχνική, και ειδικότερα με την χρήση GANs [2]. Επιχειρούμε την παραγωγή των επιθυμητών τεχνητών εικόνων εκπαιδύοντας και βελτιστοποιώντας διαφορετικές υλοποιήσεις της τεχνολογίας αυτής.

1-4 Σκοπός

Το αντικείμενο της εργασίας συνδυάζει τεχνητή νοημοσύνη και ιατρική, και συγκεκριμένα τους κλάδους των νευρωνικών δικτύων και των υπερηχογραφικών καταγραφών αντιστοίχως. Στην πραγματικότητα βρίσκεται αρκετά πλησιέστερα στον πρώτο τομέα, καθώς αφορά τον σχεδιασμό εργαλείων επικουρικών στην κατασκευή συστημάτων τεχνητής νοημοσύνης, κυρίως νευρωνικών δικτύων, προοριζομένων για εφαρμογή σε θέματα ιατρικής. Τα αποτελέσματα που παράγονται είναι συνθετικές εικόνες υπερήχων μαστών, με καλοήγη ή κακοήγη όγκο. Αυτά θα μπορούσαν να χρησιμοποιηθούν για την εκπαίδευση συστημάτων αυτόματης ανίχνευσης όγκων σε πραγματικούς υπερήχους μαστών.

2 ΘΕΩΡΗΤΙΚΗ ΘΕΜΕΛΙΩΣΗ

2-1 Υπόβαθρο μεθόδων Νευρωνικών Δικτύων

Η θεμελιώδης δομή των νευρωνικών δικτύων μπορεί να περιγραφεί ως πολυεπίπεδη συστοιχία νευρώνων - perceptrons (multi-layer perceptron - MLP). Το perceptron [3] περιγράφεται από ένα διάνυσμα «βαρών» W και μία συνάρτηση «ενεργοποίησης» $f: \mathbb{R} \rightarrow D_f$. Συνήθη σύνολα τιμών είναι τα $\{0, 1\}$, $[0, 1]$ και $[-1, 1]$. Η έξοδος του perceptron για δεδομένο διάνυσμα εισόδου a είναι η $f(W \cdot a)$. Η εκπαίδευση ενός νευρωνικού δικτύου συνίσταται στην αναζήτηση των καταλλήλων διανυσμάτων βαρών για την λύση ενός προβλήματος.

Στην επιβλεπόμενη μάθηση, όπου οι σωστές αναμενόμενες έξοδοι του συστήματος για κάθε δείγμα εισόδου είναι εκ των προτέρων γνωστές, η προσέγγιση σε μία λύση περιγράφεται από την τιμή μίας συνάρτησης «απώλειας». Η συνάρτησης απώλειας, ή «κόστους», αναπαριστά την απόκλιση της εξόδου του συστήματος από την ορθή αναμενόμενη. Σε κάθε στάδιο της εκπαίδευσης ενός δικτύου («εποχή»), ακολουθείται ένας αλγόριθμος “optimizer”, που οδηγεί το σύστημα σε τοπικό ακρότατο της συνάρτησης απώλειας, τροποποιώντας τις τιμές των διανυσμάτων βαρών.

2-2 Συνελικτικά Δίκτυα - CNNs

Τα συνελικτικά δίκτυα εφαρμόζονται συνήθως σε δεδομένα δύο ή τριών διαστάσεων (όπως εικόνες ενός ή τριών χρωματικών καναλιών αντιστοίχως), και η βασική ιδέα της λειτουργίας τους είναι η εφαρμογή της γραμμικής πράξης της συνέλιξης σε διαδοχικά διαστήματα σταθερού μεγέθους των δεδομένων. Η συνέλιξη δύο συναρτήσεων f και g ορίζεται ως $(f * g)(x) = \int f(t)g(x - t) dt$, και οι στρώσεις του δικτύου στις οποίες πραγματοποιείται ονομάζονται «συνελικτικές» (convolutional layers). Συγκεκριμένα, τα βάρη κάθε νευρώνα σε αυτά τα επίπεδα αναπαρίστανται ως ένας δισδιάστατος πίνακας, ο οποίος συνδυάζεται μέσω συνέλιξης με ένα απόσπασμα ιδίων διαστάσεων της εικόνας, ώστε να προκύψει η έξοδος του νευρώνα. Οι πίνακες αυτοί ονομάζονται «φίλτρα» (filters ή convolutional layers).

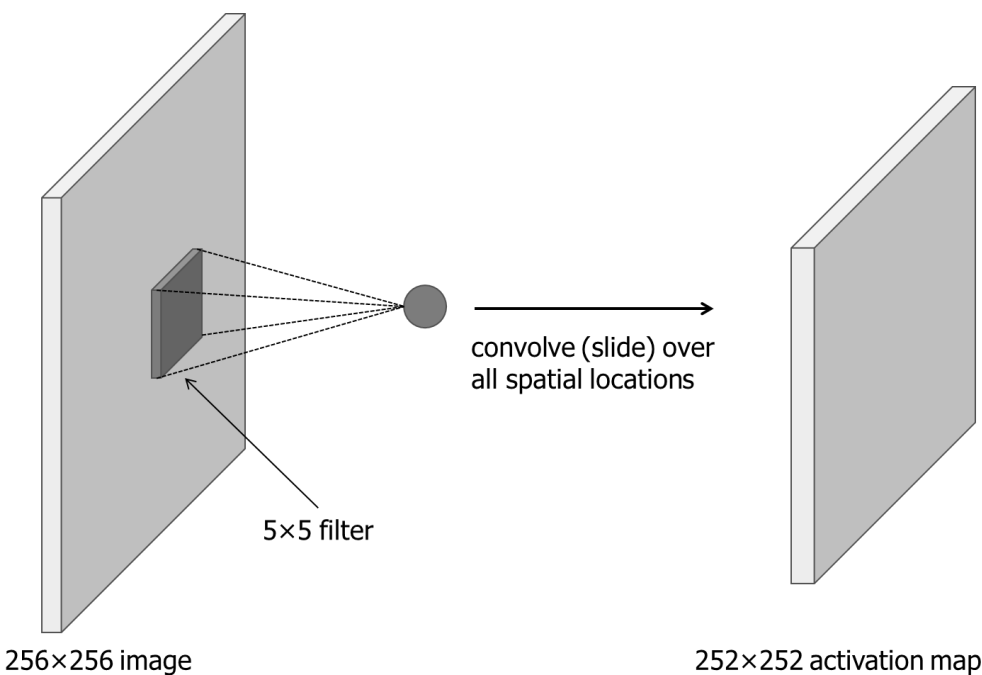
2-3 Ανταγωνιστικά Παραγωγικά Δίκτυα - GANs

Ένα GAN αποτελείται από δύο διακριτά νευρωνικά δίκτυα: τον «διευκρινιστή» (discriminator), ρόλος του οποίου είναι να αποφασίζεται αν τα δεδομένα εισόδου

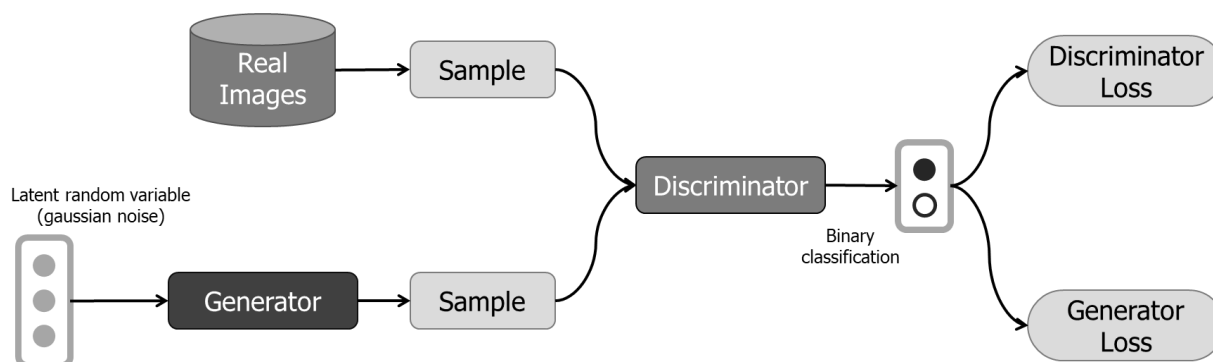
είναι αληθινά, και τον «γεννήτορα» (generator), ο οποίος «προσπαθεί» να δημιουργήσει δεδομένα που θα θεωρηθούν αληθινά από τον discriminator. Στην απλούστερη περίπτωση, ο generator δεν λαμβάνει ποτέ αληθινά δεδομένα ως είσοδο, αντ' αυτού μία τυχαία κατανομή (γκαουσιανό θόρυβο), και σταδιακά μαθαίνει την κατανομή των χαρακτηριστικών μέσω του συνεχούς ανταγωνιστικού παιχνιδιού μεταξύ των δύο δικτύων.

Επειδή το σύστημα συντίθεται από δύο δίκτυα με διαφορετικούς στόχους, πρέπει αυτά να εκπαιδευθούν χωριστά. Κάθε επανάληψη εκπαίδευσης έχει δύο διακριτές φάσεις: στην πρώτη εκπαιδεύεται ο discriminator, ως δυαδικός ταξινομητής μεταξύ αληθινών εικόνων από το σύνολο δεδομένων, και πλαστών κατασκευασμένων από τον generator. Κατά την οπισθοπαλινδρόμηση (backpropagation) βελτιστοποιούνται μόνο τα βάρη του discriminator. Στην δεύτερη φάση εκπαιδεύεται ο generator, τροφοδοτώντας τον discriminator μόνο με τεχνητές εικόνες και τροποποιώντας μόνο τα βάρη του πρώτου κατά την backpropagation.

Τα GANs δεν είναι το μοναδικό είδος παραγωγικών δικτύων: ορισμένες αρχιτεκτονικές αυτοκωδικοποιητών βρίσκουν καλή εφαρμογή στον τομέα αυτόν, όπως οι Sparse και οι Variational Autoencoders [1]. Ωστόσο, προς το παρόν τα GANs είναι η κορυφαία τεχνολογία του κλάδου ως προς την ποιότητα και ακρίβεια των αποτελεσμάτων.



Σχ. 1: Σχηματική αναπαράσταση της λειτουργίας ενός convolutional layer.



Σχ. 2: Αναπαράσταση συστήματος GAN.

2-4 Εκπαίδευση GANs

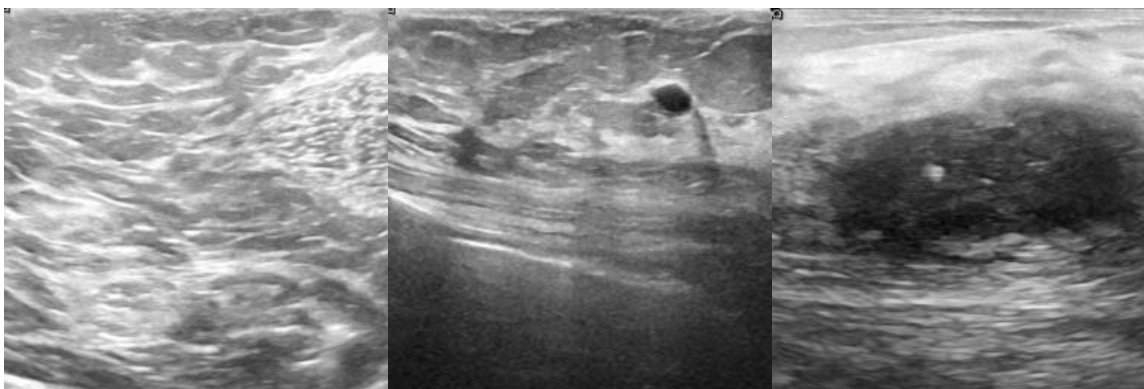
Εκτός από τα συνήθη προβλήματα που προκύπτουν κατά την εκπαίδευση ενός νευρωνικού δικτύου, όπως η διαχείριση δεδομένων μεγάλων διαστάσεων και η υπερπροσαρμογή, στα GANs εμφανίζονται ζητήματα που οφείλονται στην ανταγωνιστική φύση του συστήματος. Παρ' ότι θεωρητικά ένα GAN συγκλίνει σε μοναδική θέση ισορροπίας (Nash equilibrium) [2], όπου ο generator παράγει απολύτως αληθοφανείς εικόνες και ο discriminator επιτυγχάνει βέλτιστη επίδοση 50% με τυχαία πρόβλεψη, στην πράξη παρατηρούνται φαινόμενα υπερπροσαρμογής, όπως η «κατάρρευση επικρατούσας τιμής» (mode collapse). Σε αυτήν την περίπτωση ο generator παράγει καλύτερα δείγματα από μία κατηγορία ή υποσύνολο κατηγοριών, με αποτέλεσμα σταδιακά να παράγονται δείγματα αποκλειστικά από το υποσύνολο αυτό. Τότε η βέλτιστη απόκριση για τον discriminator είναι να απορρίπτει όλα τα δείγματα, συνεπώς ο generator μεταπηδά σε υπερπροσαρμογή άλλης κλάσης, και τελικά δεν επέρχεται ποτέ ισορροπία [1]. Ένα ακόμη πρόβλημα είναι αυτό της «κυριαρχίας», κατά το οποίο ένα από τα δύο δίκτυα κυριαρχεί επί του άλλου με αποτέλεσμα η εκπαίδευση πρακτικά να σταματά και για τα δύο.

3 ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΔΟΜΗ

3-1 Σύνολο δεδομένων

Η εκπαίδευση των δικτύων σε αυτήν την εργασία στηρίχθηκε στο ανοικτά διαθέσιμο σύνολο δεδομένων από υπέρηχους μαστών από την δημοφιλή πλατφόρμα “Kaggle”

[9]. Αυτό αποτελείται από 780 εικόνες ενός χρωματικού καναλιού, διαστάσεων 500x500, προερχόμενες από 600 διαφορετικές ασθενείς ηλικιών μεταξύ 25 και 75 ετών, που παρήχθησαν το 2018 στο νοσοκομείο Baheya του Καΐρου. Εξ αυτών, οι 437 απεικονίζουν μαστούς με καλοήγη όγκο, οι 210 με κακοήγη, και οι υπόλοιπες 133 υγιείς μαστούς [4]. Επειδή σε όλα τα πειράματα η εκπαίδευση έγινε για κάθε μία από τις δύο πρώτες κλάσεις χωριστά, δηλαδή κατασκευαζόταν ένα δίκτυο για την παραγωγή υπερηχογραφημάτων καλοηθών όγκων και ένα άλλο για κακοηθών, η ανομοιόμορφη κατανομή των δειγμάτων στις κατηγορίες-κλάσεις δεν χρειάστηκε να ληφθεί υπ' όψιν.



Δείγματα των εικόνων εκπαίδευσης. Από αριστερά προς δεξιά, υγιής μαστός, καλοήθης όγκος, κακοήθης όγκος.

Χρειάστηκε, ωστόσο, να γίνουν πειράματα με επαύξηση δεδομένων, καθώς οι διαθέσιμες εικόνες ήταν σχετικά ανεπαρκείς για τις συνήθεις απαιτήσεις των πολύπλοκων νευρωνικών δικτύων όπως τα GANs. Η επαύξηση αυτή συνεστάθη στην προσθήκη αληθινών υπερηχογραφημάτων από το υποσύνολο των υγιών μαστών, το οποίο ήταν και το μεγαλύτερο, και στην τυχαία περιστροφή και μεγέθυνση κάθε μίας αληθινής εικόνας κατά τυχαίο και μικρό παράγοντα. Η δεύτερη πραγματοποιήθηκε μέσω ειδικών μεθόδων της βιβλιοθήκης TensorFlow, που για κάλυψη των κενών στις νέες εικόνες χρησιμοποιούν διγραμμική παρεμβολή (bilinear interpolation).

3-2 Αρχιτεκτονικές GAN

Το πρώτο μοντέλο που κατασκευάστηκε τελεί ως μέτρο σύγκρισης για τα υπόλοιπα. Έχει πολύ απλή δομή, συγκεκριμένα τα δύο δίκτυα (generator και discriminator) αποτελούνται από πυκνά επίπεδα (Dense layers), για τα οποία δοκιμάστηκαν διάφορες διαστάσεις (πλήθη νευρώνων) και συναρτήσεις ενεργοποίησης (relu, selu, sigmoid).

Model: "generator"			Model: "discriminator"		
Layer (type)	Output Shape	Param #	Layer (type)	Output Shape	Param #
dense_6 (Dense)	(None, 100)	3100	flatten_1 (Flatten)	(None, 250000)	0
dense_7 (Dense)	(None, 150)	15150	dense_9 (Dense)	(None, 150)	37500150
dense_8 (Dense)	(None, 250000)	37750000	dense_10 (Dense)	(None, 100)	15100
reshape_1 (Reshape)	(None, 500, 500, 1)	0	dense_11 (Dense)	(None, 1)	101
Total params: 37,768,250			Total params: 37,515,351		
Trainable params: 37,768,250			Trainable params: 37,515,351		
Non-trainable params: 0			Non-trainable params: 0		

Σχ. 3: Σύνοψη του μοντέλου από το TensorFlow: το συγκεκριμένο είχε κατασκευασθεί για διαστάσεις εισόδου 30 στον generator, και 500x500 στον discriminator.

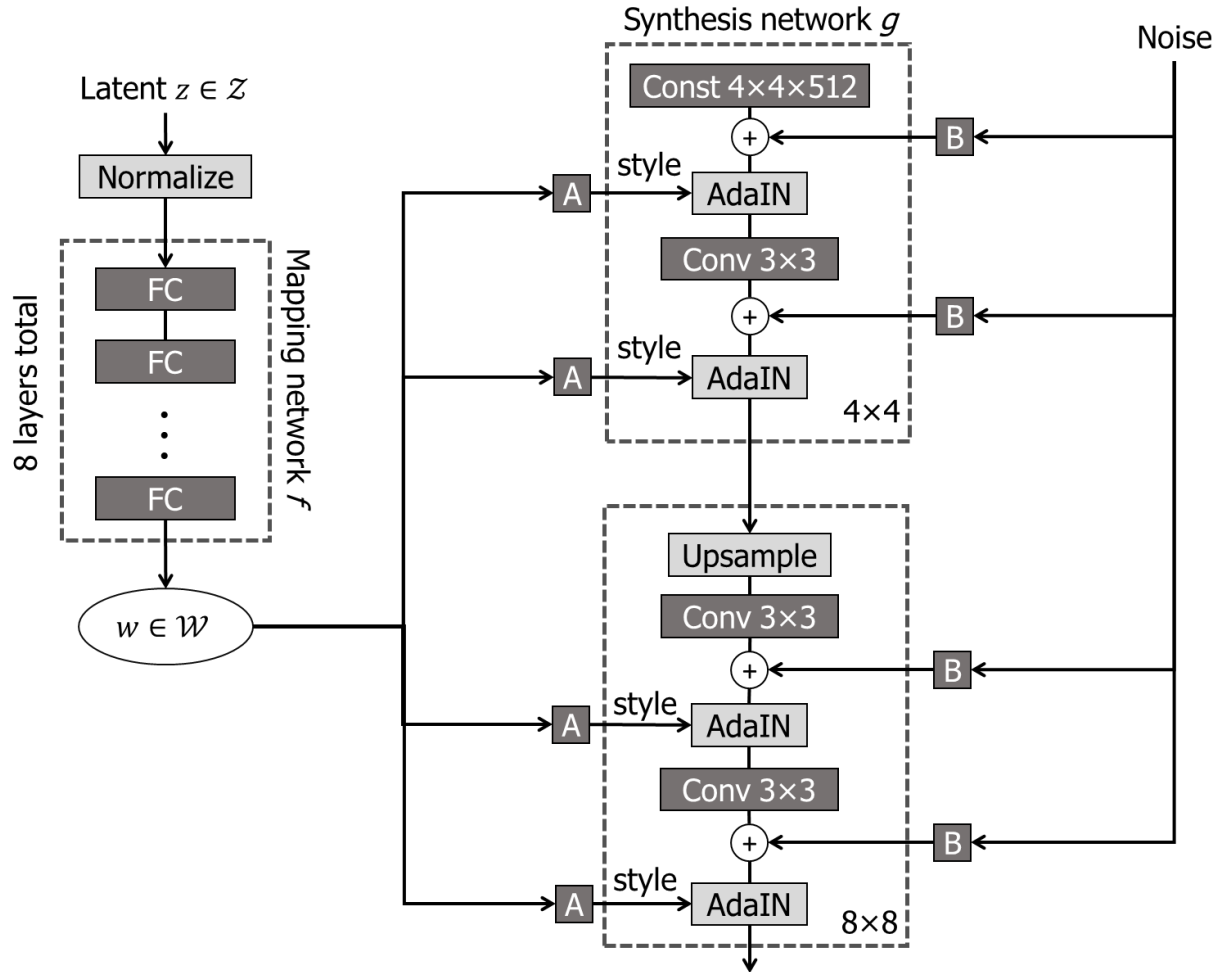
Το επόμενο μοντέλο που δοκιμάστηκε ακολουθεί την αρχιτεκτονική Deep Convolutional GAN (DCGAN) [8]. Αυτή υπαγορεύει ορισμένες αλλαγές σε σχέση με τον σχεδιασμό της αρχικής δημοσίευσης [2], όπως την χρήση συνελκτικών επιπέδων και τεχνικές batch normalization. Αν και υποδεικνύεται η χρήση συγκεκριμένων συναρτήσεων ενεργοποίησης για κάθε επίπεδο (tanh στο τελευταίο layer του generator, relu σε όλα τα υπόλοιπα, leaky relu σε όλα τα layers του discriminator), δοκιμάστηκαν και άλλες στο πλαίσιο του πειραματισμού.

Model: "generator"			Model: "discriminator"		
Layer (type)	Output Shape	Param #	Layer (type)	Output Shape	Param #
dense_12 (Dense)	(None, 2000000)	202000000	conv2d (Conv2D)	(None, 250, 250, 64)	1664
reshape_2 (Reshape)	(None, 125, 125, 128)	0	dropout (Dropout)	(None, 250, 250, 64)	0
batch_normalization (Batch Normalization)	(None, 125, 125, 128)	512	conv2d_1 (Conv2D)	(None, 125, 125, 128)	204928
conv2d_transpose (Conv2DTranspose)	(None, 250, 250, 64)	204864	dropout_1 (Dropout)	(None, 125, 125, 128)	0
batch_normalization_1 (Batch Normalization)	(None, 250, 250, 64)	256	flatten_2 (Flatten)	(None, 2000000)	0
conv2d_transpose_1 (Conv2DTranspose)	(None, 500, 500, 1)	1601	dense_13 (Dense)	(None, 1)	2000001
Total params: 202,207,233			Total params: 2,206,593		
Trainable params: 202,206,849			Trainable params: 2,206,593		
Non-trainable params: 384			Non-trainable params: 0		

Σχ. 4: Σύνοψη του μοντέλου για είσοδο μήκους 100 και έξοδο 500x500. Οι μη εκπαιδευσιμες παράμετροι στον generator οφείλονται στα επίπεδα BatchNormalization.

Το τελευταίο που υλοποιήθηκε είναι το πολύ πιο σύγχρονο (state of the art - “sota”) StyleGAN-2 [7]. Η αρχιτεκτονική αυτού είναι αρκετά πολύπλοκη. Ενδιαφέρον παρουσιάζει ο σχεδιασμός του δικτύου generator, το οποίο αποτελείται από δύο υπο-δίκτυα: “mapping”, ένα MLP αποτελούμενο από οκτώ πλήρως συνδεδεμένα επίπεδα (fully connected layers - fc), υπεύθυνο για την επεξεργασία της εισόδου,

και “synthesis”, αποτελούμενο από συνελικτικά επίπεδα. Φαίνεται πως η εισαγωγή θορύβου ενδιάμεσα στο δίκτυο synthesis, εκτός από την αρχική είσοδο, είχε σημαντική επίδραση στα αποτελέσματα [1].



Σχ. 5: Η αρχιτεκτονική του generator του StyleGAN [1], [6].

3-3 Εκπαίδευση

Για όλες τις διαφορετικές αρχιτεκτονικές δικτύων εφαρμόστηκε εν πολλοίς η ίδια μεθοδολογία στην εκπαίδευση. Αρχικά έγινε προεπεξεργασία των εικόνων εκπαίδευσης, συγκεκριμένα μετασχηματισμός αυτών ώστε να έχουν όλες τις ίδιες διαστάσεις. Για τα δύο πρώτα δίκτυα αυτές ήταν 500×500 , ενώ για το StyleGAN, σύμφωνα με τις προδιαγραφές αυτού, που απαιτούν διαστάσεις ίσες με δύναμη του 2, έγινε σμίκρυνση των δεδομένων εκπαίδευσης σε 256×256 . Δεν προτιμήθηκαν οι 512×512 για δύο λόγους: αφ' ενός πολλές εικόνες είχαν την μία διάσταση αρκετά μικρότερη από 500, οπότε θα προκαλείτο σημαντική αλλοίωση σε αυτές, και αφ'

ετέρου διατηρώντας χαμηλότερη την διαστατικότητα των δεδομένων δεν αυξήθηκε ο ήδη μεγάλος απαιτούμενος χρόνος εκπαίδευσης του περίπλοκου αυτού δικτύου.

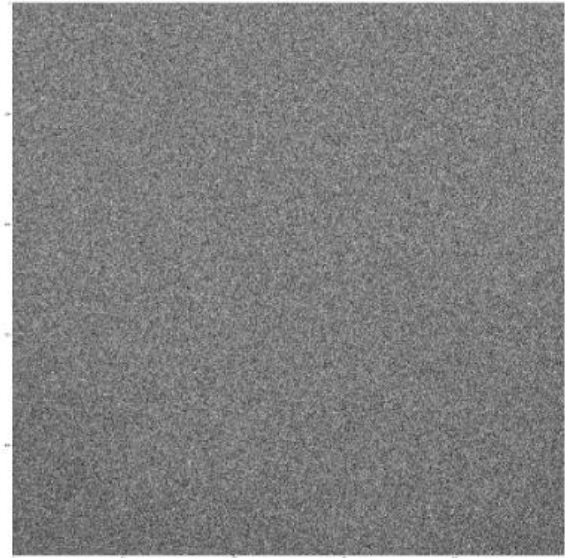
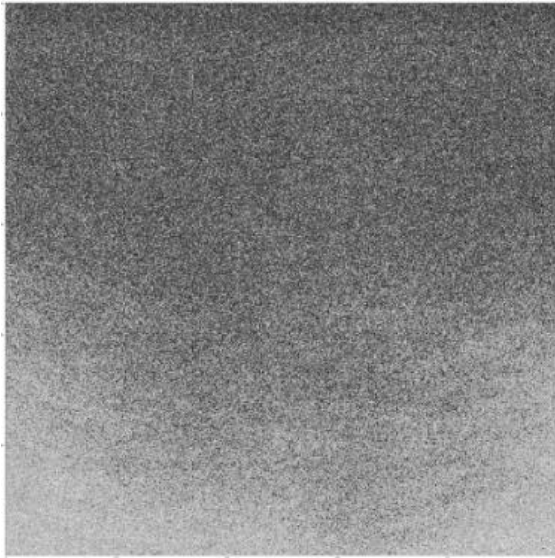
Στην συνέχεια κατασκευάστηκαν με την σειρά τα τρία συστήματα, τα δύο πρώτα χρησιμοποιώντας τις μεθόδους της βιβλιοθήκης TensorFlow της γλώσσας Python, και το τρίτο από τον επίσημο πηγαίο κώδικα της υλοποίησης σε PyTorch [7], σε περιβάλλον Google Colab. Πραγματοποιήθηκαν πειράματα με διαφορετικές συνθήκες εκπαίδευσης (πλήθος εποχών, κατανομή των training batches) αλλά και με «υπερπαραμέτρους» όπως η συνάρτηση κόστους, η συνάρτηση ενεργοποίησης, ο αλγόριθμος optimizer, το πλήθος νευρώνων στα πυκνά επίπεδα, και οι διαστάσεις και το πλήθος φίλτρων στα συνελικτικά επίπεδα.

Ως μετρική απώλειας στα δύο πρώτα δίκτυα χρησιμοποιήθηκε η “Wasserstein loss” [5], η οποία είναι η προεπιλεγμένη από την βιβλιοθήκη TF-GAN, αλλά και η binary cross-entropy. Το StyleGAN-2 εφαρμόζει την Frechet Inception Distance [7]. Αναλυτική περιγραφή όλων γίνεται στο Παράρτημα. Ως συναρτήσεις ενεργοποίησης δοκιμάστηκαν αρκετές διαφορετικές, όπως αναφέρθηκε νωρίτερα. Αναλόγως με το πεδίο τιμών της συνάρτησης ενεργοποίησης του τελικού επιπέδου σε κάθε δίκτυο, έπρεπε να γίνει κατάλληλη προσαρμογή των αποτελεσμάτων στην κλίμακα 0-255, ώστε να αναπαρασταθούν ως μονοχρωματικά 8bit pixels. Αναλυτική περιγραφή όλων των συναρτήσεων γίνεται στο Παράρτημα. Ως αλγόριθμοι βελτιστοποίησης δοκιμάστηκαν οι RMSProp και Adam (“Adaptive moment estimation”). Αυτοί είναι καταλληλότεροι από την στοχαστική κάθοδο κλίσης (stochastic gradient descent – SGD), η οποία είναι η συνηθέστερη στα νευρωνικά δίκτυα, επειδή οδηγούν σε ταχύτερη σύγκλιση σε τοπικό ακρότατο της συνάρτησης απώλειας σε μεγαλύτερα συστήματα [1]. Το μειονέκτημά τους σε σχέση με την SGD είναι ότι μπορεί να καταλήξουν σε χαμηλότερης ποιότητας λύσεις. Ο Adam είναι ο προεπιλεγμένος αλγόριθμος για το StyleGAN-2, Τέλος, τα αποτελέσματα που παρήγαγαν τα δίκτυα generator εξετάζονταν ανά εποχή εκπαίδευσης.

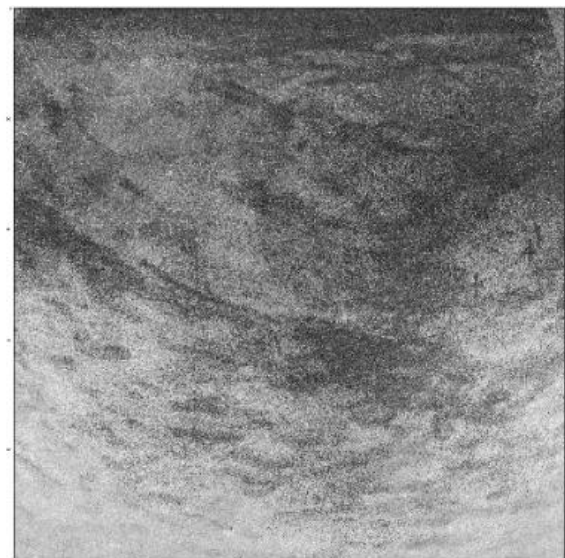
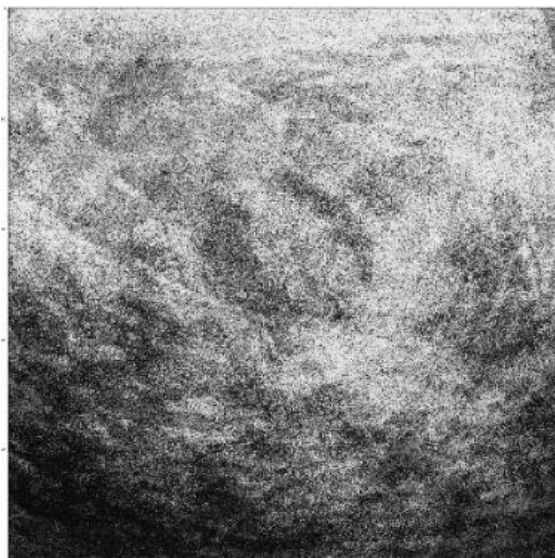
4 ΑΠΟΤΕΛΕΣΜΑΤΑ

4-1 Δείγματα

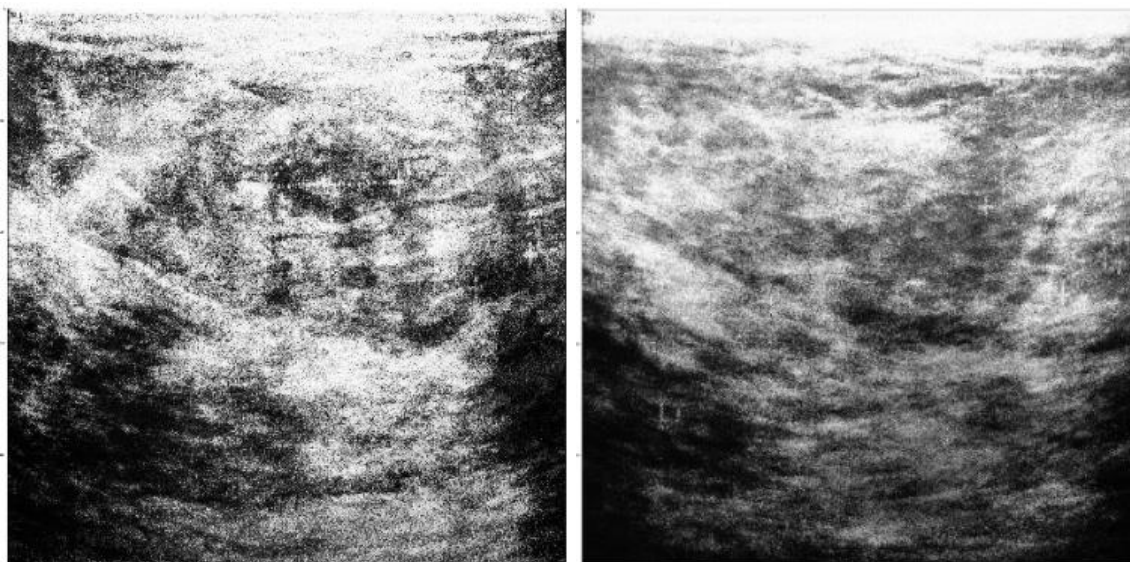
Ενδεικτικά αποτελέσματα του πρώτου συστήματος από κάθε εποχή εκπαίδευσης:



Δείγματα μετά την πρώτη εποχή εκπαίδευσης.



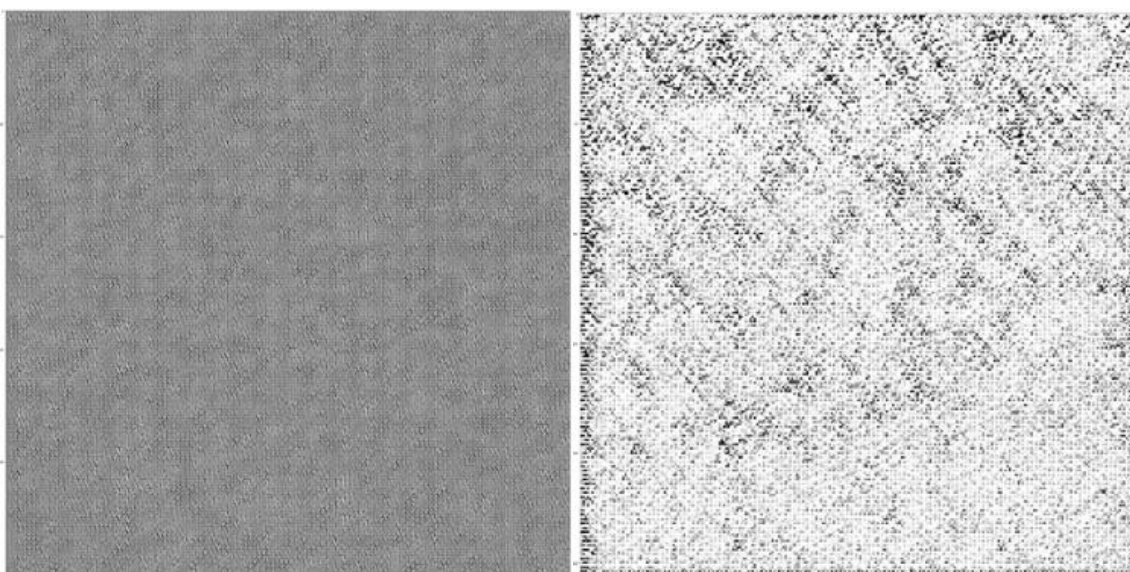
Δείγματα μετά την δεύτερη εποχή εκπαίδευσης.



Τρίτη εποχή (αριστερά) και δέκατη εποχή (δεξιά). Διακρίνονται σταυροί, οφειλόμενοι στις επισημάνσεις παρούσες σε αρκετές εικόνες εκπαίδευσης.

Στην πρώτη εικόνα φαίνεται ο τυχαίος θόρυβος που παράγεται αρχικά, ενώ σταδιακά στην δεύτερη, τρίτη, και δέκατη εποχή, τα αποτελέσματα λαμβάνουν μορφή πλησιέστερη σε αυτή των αληθινών δεδομένων.

Δείγματα αποτελεσμάτων του DCGAN, εκπαιδευμένου με διάφορες παραμέτρους:

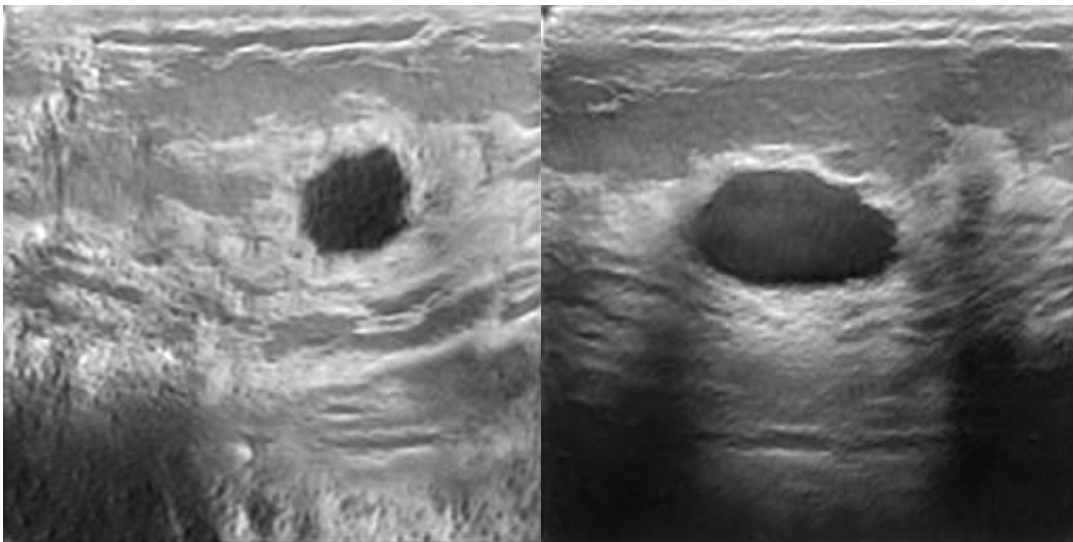


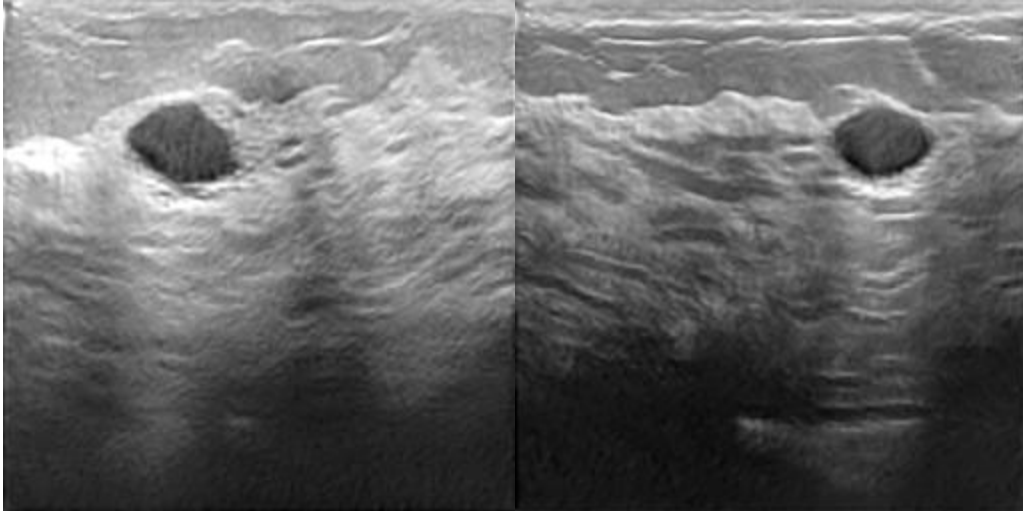
Το DCGAN δεν παρήγαγε αληθοφανή αποτελέσματα σε κανένα πείραμα.



Η μοναδική περίπτωση που συνέκλινε σε κάτι εκτός του πρακτικά τυχαίου θορύβου προέκυψε μειώνοντας τις διαστάσεις δεδομένων εισόδου. Αιτίες και πιθανές βελτιώσεις αναλύονται στην επόμενη ενότητα.

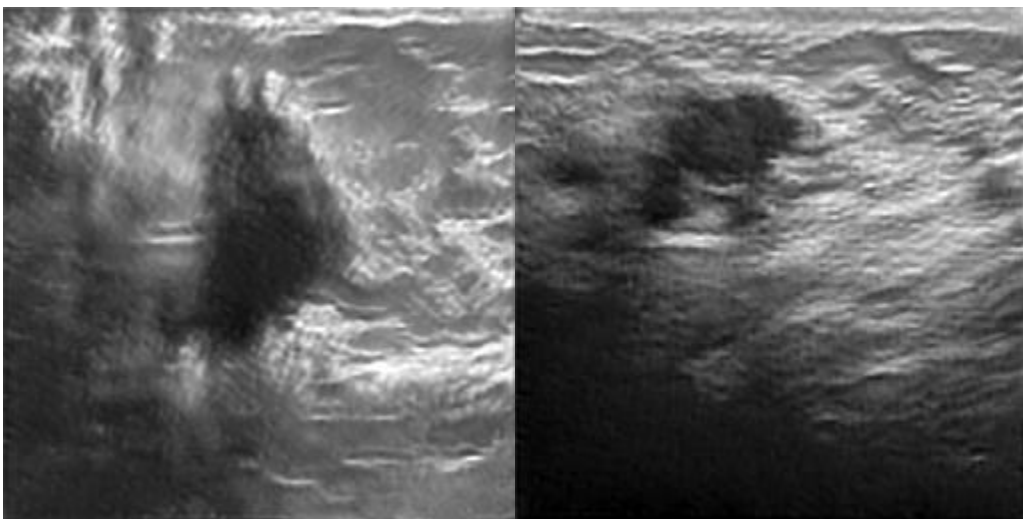
Τα αποτελέσματα του StyleGAN-2 ήταν αρκετά καλά ώστε να παρουσιασθούν αναλυτικότερα. Ορισμένα συνθετικά υπερηχογραφήματα με καλοήγη όγκο:

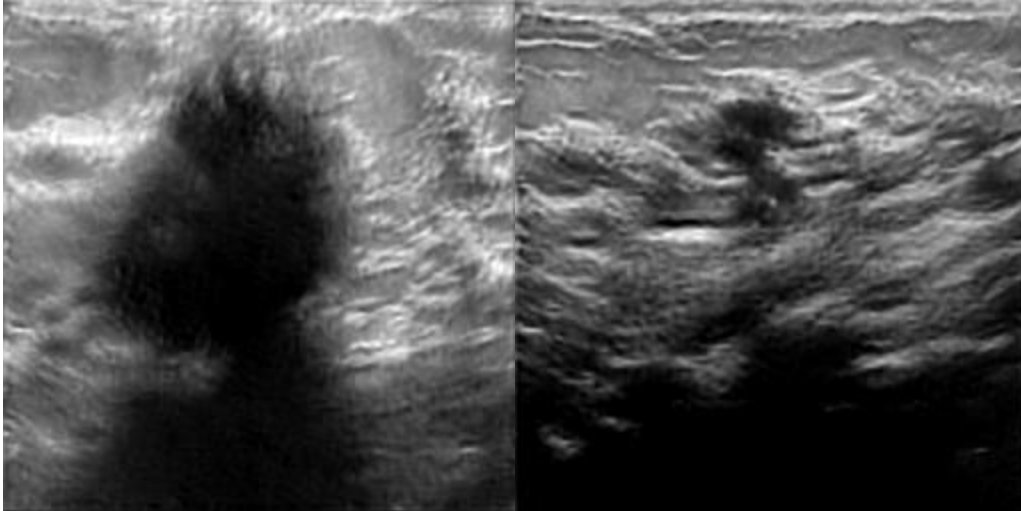




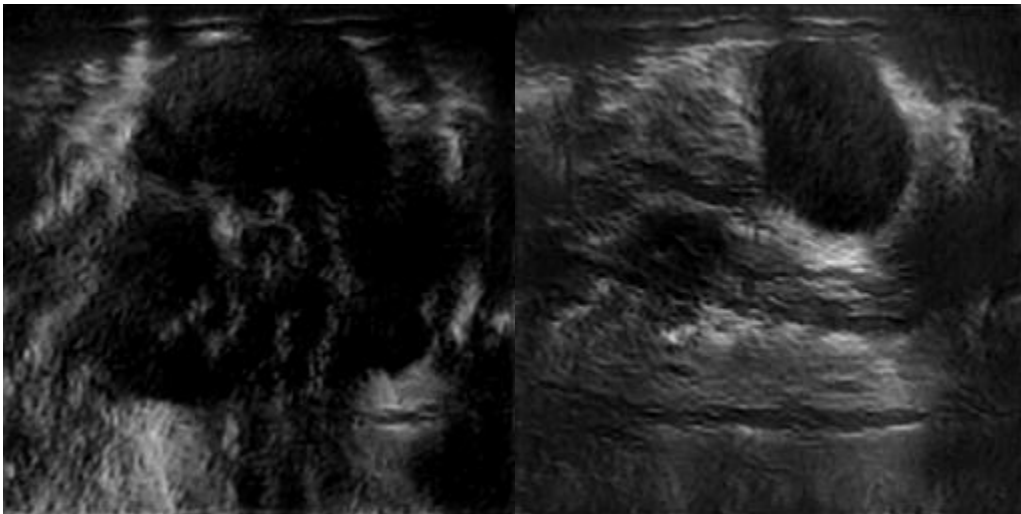
Προέκυψαν ύστερα από πέντε εποχές εκπαίδευσης.

Και με κακοήθη όγκο:





Αξίζει να παρατεθούν δείγματα κακών αποτελεσμάτων, καθώς στην πραγματικότητα τέτοια ήταν στην πλειοψηφία τους. Τα ανωτέρω εξαιρετικά δείγματα επελέγησαν μεταξύ εκατοντάδων.



4-2 Αξιολόγηση και πορίσματα

Εκ των αποτελεσμάτων γίνεται προφανές ότι το StyleGAN-2 παρήγαγε με μεγάλη διαφορά τα καλύτερα αποτελέσματα. Μάλιστα, οι δίχως ιατρική εκπαίδευση συγγραφείς αδυνατούν να τα διακρίνουν από τα αληθινά. Το μοναδικό μειονέκτημα αυτού σε σχέση με τα άλλα ήταν ο πολλαπλάσιος απαιτούμενος χρόνος εκπαίδευσης.

Το πρώτο, απλοϊκό δίκτυο, που είχε ρόλο μέτρου σύγκρισης για τα υπόλοιπα, παρήγαγε εικόνες που φέρουν ορισμένα χαρακτηριστικά των αληθινών, ωστόσο

είναι εμφανές ότι δεν είναι πραγματικά υπερηχογραφήματα. Επίσης, σε αρκετά πειράματα παρατηρήθηκε το φαινόμενο εκφυλισμού που έχει αναλυθεί σε προηγούμενη ενότητα, όπου πρακτικά το δίκτυο κατασκευάζει μόνο μία εικόνα.

Το δεύτερο δίκτυο (DCGAN) δεν μπόρεσε να συγκλίνει ποτέ σε κάποια λύση, παρά τις πολλαπλές δοκιμές, τόσο με επαύξηση δεδομένων, όσο και με αλλαγή των υπερπαραμέτρων του δικτύου. Όπως φαίνεται από τα ανωτέρω δείγματα, πρακτικά παρήγαγε τυχαίο θόρυβο. Τα αποτελέσματα εξετάζονταν σε κάθε εποχή εκπαίδευσης, με καμία ορατή αλλαγή μεταξύ αυτών. Μοναδική εξαίρεση αποτέλεσε μία υλοποίηση που χρησιμοποίησε μικρότερα δεδομένα εισόδου, συγκεκριμένα 100×100 . Σε αυτήν τα αποτελέσματα ήταν εμφανώς όχι τυχαία, αλλά πάλι δεν πλησίασαν τις αληθινές εικόνες. Πιθανές αιτίες είναι η αδυναμία του δικτύου discriminator να απορρίψει εικόνες, με αποτέλεσμα οτιδήποτε παρήγαγε ο generator να θεωρείτο αποδεκτό, ή τα συνελκτικά επίπεδα που απαρτίζουν τον generator να μην είναι κατάλληλα για την παραγωγή συνθετικών δεδομένων. Στο StyleGAN αυτά χρησιμοποιούνται σε συνδυασμό με πλήρως συνδεδεμένα επίπεδα.

	Συναρτήσεις ενεργοποίησης	Συναρτήσεις απώλειας	Optimizer	Πλήθος παραμέτρων	Χρόνος εκπαίδευσης
Simple	selu στα ενδιάμεσα επίπεδα, σιγμοειδής στα τελικά	binary cross-entropy	RMSProp	generator 37,768,250 discriminator 37,515,351	~1min/epoch
DCGAN	στον discriminator leaky relu 0.2 στα ενδιάμεσα, σιγμοειδής στο τελικό στον generator selu στα ενδιάμεσα και tanh στο τελικό	Wasserstein	RMSProp	generator 16,527,233 discriminator 16,570,658	~3min/epoch
StyleGAN-2	relu	FID	Adam	generator 23186516 discriminator 24000961	>10h

Βέλτιστες υπερπαραμέτροι και αντίστοιχοι χρόνοι εκπαίδευσης για κάθε υλοποίηση.

4-3 Επεκτάσεις

Το πρώτο βήμα για την περαιτέρω βελτίωση των αποτελεσμάτων της εργασίας θα ήταν αναζήτηση και χρήση καλύτερου dataset: αυτό που χρησιμοποιήθηκε ήταν αρκετά μικρό, και επιπλέον δεν ήταν συνεκτικό ως προς τις διαστάσεις των εικόνων, ούτε όλες οι εικόνες ήταν καλής ποιότητας. Επιπλέον, οπωσδήποτε θα έπρεπε να εξετασθούν και άλλες σύνθετες αρχιτεκτονικές, όπως το SAGAN [10], δεδομένων των εξαιρετικών αποτελεσμάτων του StyleGAN. Τέλος, πιθανότατα θα ωφελούσε να χρησιμοποιηθεί ισχυρότερο υπολογιστικό σύστημα για την εκπαίδευση των δικτύων.

5 ΕΠΙΛΟΓΟΣ

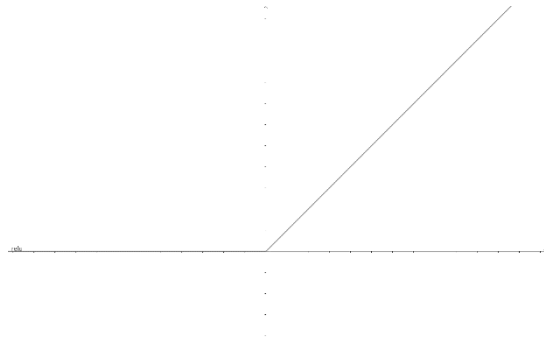
Η εργασία αυτή εξέτασε την ικανότητα των GANs να παράγουν αληθοφανή ιατρικά δεδομένα. Τα συμπεράσματα είναι ιδιαίτερος ενθαρρυντικά. Φαίνεται πως η τεχνολογία αυτή είναι ικανή να αποτελέσει χρήσιμο εργαλείο στα χέρια των ιατρών και βιοεπιστημόνων, και να επιταχύνει την ανάπτυξη λύσεων τεχνητής νοημοσύνης για αυτοματοποιημένη διάγνωση. Επομένως αξίζει να διερευνηθεί περαιτέρω και σε άλλες κατηγορίες απεικονίσεων. Παρ' ότι το αντικείμενο των GANs είναι δημοφιλέστατο και η σχετική βιβλιογραφία άφθονη, θα είναι δύσκολο να επιλεγούν οι καταλληλότερες υλοποιήσεις, ως προς την ενημερότητα και την εφαρμοσιμότητα στο συγκεκριμένο πρόβλημα. Ένα επιπλέον ζήτημα θα είναι ο βαθμός εξηγησιμότητας των παραγόμενων συστημάτων.

6 ΠΑΡΑΡΤΗΜΑ

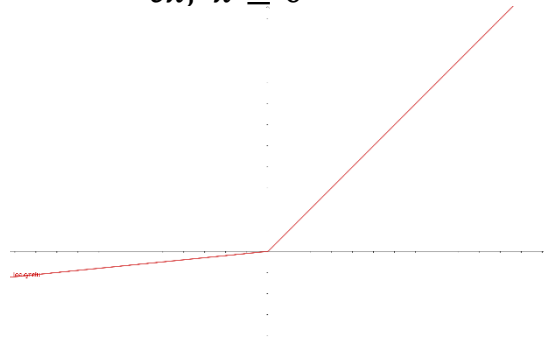
6-1 Ευρέως χρησιμοποιούμενες συναρτήσεις ενεργοποίησης

Rectified Linear Unit Function (ReLU) “Leaky” ReLU

$$f(x) = \max\{0, x\}$$



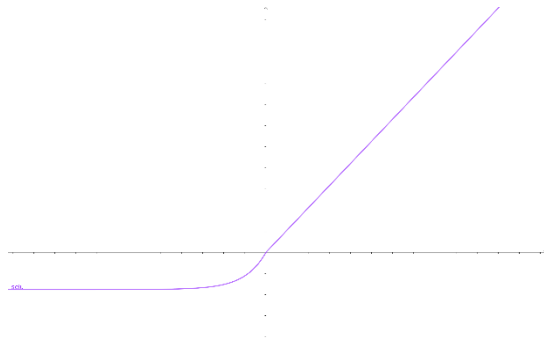
$$f(x) = \begin{cases} x, & x > 0 \\ cx, & x \leq 0 \end{cases}, \quad 0 < c < 1$$



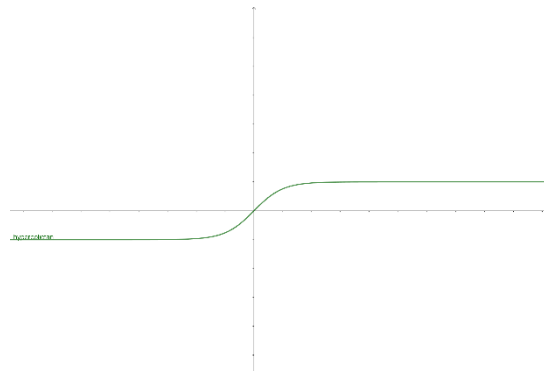
Scaled Exponential Linear Unit (SELU) Υπερβολική εφαπτομένη

$$f(x) = \begin{cases} \lambda x, & x \geq 0 \\ \lambda \alpha (e^x - 1), & x < 0 \end{cases}$$

$$\lambda \simeq 1.0507, \alpha \simeq 1.6733$$

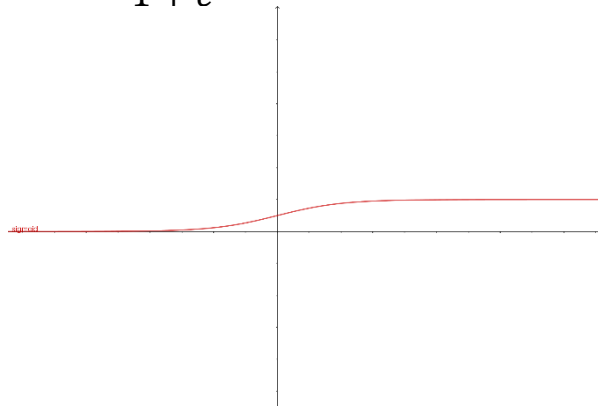


$$\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$$



Σιγμοειδής

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



6-2 Μετρικές απώλειας

Minimax loss (η συνάρτηση που χρησιμοποιείται από την πρώτη δημοσίευση [2])

$$\mathbb{E}_x[\log D(x)] + \mathbb{E}_z[\log(1 - D(G(z)))], \text{ όπου}$$

$D(x)$ είναι η εκτίμηση του discriminator για την πιθανότητα τα αληθινά δεδομένα x να είναι αληθινά,

\mathbb{E}_x είναι η αναμενόμενη τιμή για όλα τα δείγματα αληθινών δεδομένων x ,

$G(z)$ είναι η έξοδος του generator δεδομένου θορύβου z ,

$D(G(z))$ είναι η εκτίμηση του discriminator για την πιθανότητα συνθετικά δεδομένα αυτά να είναι αληθινά, και

\mathbb{E}_z είναι η αναμενόμενη τιμή για όλες τις τυχαίες εισόδους στον generator.

Επειδή το δίκτυο generator δεν μπορεί να επηρεάσει ευθέως την ποσότητα $\log D(x)$, για αυτό η ελαχιστοποίηση της συνάρτησης απώλειας ανάγεται σε ελαχιστοποίηση του $\log(1 - D(G(z)))$.

Frechet Inception Distance

Αν τα χαρακτηριστικά των εικόνων που παράγονται από το GAN ακολουθούν κανονική κατανομή $\mathcal{N}(\mu, \sigma)$, και τα ίδια χαρακτηριστικά των αληθινών εικόνων εκπαίδευσης ακολουθούν κατανομή $\mathcal{N}(\mu_w, \sigma_w)$, τότε η συνάρτηση αυτή ορίζεται ως

$$\|\mu - \mu_w\|_2^2 + \text{tr}\left(\sigma + \sigma_w - 2(\sigma^{1/2}\sigma_w\sigma^{1/2})^{1/2}\right)$$

όπου $\text{tr}(A)$ είναι το ίχνος του τετραγωνικού πίνακα A .

Wasserstein loss

Απώλεια discriminator: $D(x) - D(G(z))$

Απώλεια generator: $D(G(z))$

Σε αυτήν την προσέγγιση, ο discriminator δεν επιτελεί ακριβώς δυαδική ταξινόμηση, αλλά έχει ρόλο κριτή, και εκπαιδεύεται με στόχο την μεγιστοποίηση της ανωτέρω συνάρτησης. Προς την ίδια κατεύθυνση εκπαιδεύεται και ο generator.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Έντυπη

- [1] Geron, Aurelien. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol: O'Reilly, 2019. Print.
- [2] Goodfellow, Ian, et al. “Generative Adversarial Nets”. *Proceedings of the 27th International Conference on Neural Information Processing Systems 2*, 2014, p. 2672-2680.
- [3] Goodfellow, Ian, et al. *Deep Learning*. Cambridge: MIT Press, 2017. Print.

Ηλεκτρονική

- [4] Al-Dhabyani, Walid, et al. “Dataset of Breast Ultrasound Images”. *Data in Brief*, vol. 28, 2020, n.p. <<https://doi.org/10.1016/j.dib.2019.104863>>.
- [5] Arjovsky, Martin, et al. “Wasserstein GAN”. *arXiv.org*, Cornell University, 6 Dec. 2017, <<https://arxiv.org/abs/1701.07875>>.
- [6] Karras, Tero, et al. “A Style-Based Generator Architecture for Generative Adversarial Networks”. *arXiv.org*, Cornell University, 29 Mar. 2019, <<https://arxiv.org/abs/1812.04948>>.
- [7] Karras, Tero, et al. “Analyzing and Improving the Image Quality of StyleGAN”. *arXiv.org*, Cornell University, 23 Mar. 2020, <<https://arxiv.org/abs/1912.04958>>.
- [8] Radford, Alec, et al. “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks”. *arXiv.org*, Cornell University, 7 Jan. 2016, <<https://arxiv.org/abs/1511.06434>>.
- [9] Shah, Arya. “Breast Ultrasound Images Dataset”. *Kaggle*. Kaggle Inc., 2021. Web. 01 Apr. 2022. <<https://www.kaggle.com/datasets/aryashah2k/>>

breast-ultrasound-images-dataset>.

- [10] Zhang, Han, et al. “Self-Attention Generative Adversarial Networks”. *arXiv.org*, Cornell University, 14 June 2019, <<https://arxiv.org/abs/1805.08318>>.

- [11] «Καρκίνος μαστού». *Άλμα Ζωής Πανελλήνιος Σύλλογος Γυναικών με Καρκίνο Μαστού*. Άλμα Ζωής, n.d., <<https://almazois.gr/καρκίνος-μαστού>>.

- [12] «Οκτώβριος: Μήνας πρόληψης και ενημέρωσης για τον καρκίνο του μαστού». *Υπουργείο Υγείας*. Υπουργείο Υγείας, 19 Οκτωβρίου 2019, <<https://www.moh.gov.gr/articles/news/1436-oktwbrios-mhnas-prolhps-shs-kai-enhmerwshs-gia-ton-karkino-toy-mastoy>>.