



NATIONAL CENTRE FOR
SCIENTIFIC RESEARCH "DEMOKRITOS"



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
UNIVERSITY OF PIRAEUS



PET Attenuation Correction with GANs

Αλέξανδρος Νταγιάντας, Ηρακλής Σπύρου

Βαθιά Μάθηση

Ιούνιος 2025

Περίληψη

Στην παρούσα εργασία μελετάται η διόρθωση εξασθένισης (attenuation correction) σε τομογραφίες εκπομπής ποζιτρονίων (PET) χρησιμοποιώντας Generative Adversarial Networks (GANs). Αρχικά, μέσω Monte Carlo προσομοιώσεων στο GATE παράχθηκαν ζεύγη μηδιορθωμένων (NAC) και διορθωμένων (AC) εικόνων, τα οποία ανακατασκευάστηκαν με OSEM iterative αλγόριθμο. Ως baseline υλοποιήθηκε ένα Pix2Pix μοντέλο (Vanilla GAN) με U-Net generator και PatchGAN discriminator, ενώ στη συνέχεια εξετάστηκε η παραλλαγή Wasserstein GAN με Gradient Penalty (WGAN-GP) για βελτιωμένη σταθερότητα εκπαίδευσης. Η απόδοση των μοντέλων αξιολογήθηκε οπτικά και ποσοτικά με μετρικές PSNR και SSIM, καθώς και με έμμεση αξιολόγηση της downstream εφαρμογής object detection: οι generated AC εικόνες χρησιμοποιήθηκαν ως είσοδος σε YOLOv8n για ανίχνευση πηγών. Το WGAN-GP υπερείχε του baseline, παράγοντας εικόνες με +1.6 dB PSNR (37.1 dB vs. 35.5 dB) και ελαφρώς υψηλότερο SSIM, καθώς και πιο ομαλές καμπύλες εκπαίδευσης. Επιπλέον, η χρήση των generated εικόνων οδήγησε σε σημαντική άνοδο των mAP@90 (0.83 vs. 0.70) και IoU (0.95 vs. 0.85) στην ανίχνευση πηγών. Τα αποτελέσματα επιβεβαιώνουν ότι το WGAN-GP αποτελεί αποτελεσματική λύση για CT-free attenuation correction PET scans και ενισχύει την ακρίβεια κρίσματων downstream tasks.

Περιεχόμενα

1 Εισαγωγή	1
2 Εργαλεία & Μέθοδοι	1
2.1 Τομογραφία PET	1
2.1.1 Βασικές αρχές PET	1
2.1.2 Γραμμή Απόκρισης (Line of Response) και Sinogram	1
2.1.3 Attenuation και Scattering	2
2.2 GATE	3
2.2.1 GATE Tool	3
2.2.2 Μεθοδολογία προσομοιώσεων	4
2.3 Ανακατασκευή εικόνων	5
2.3.1 Single vs Double source scripts	6
2.4 Dataset	6
2.5 Baseline: Vanilla GAN (<i>Pix2Pix</i>)	6
2.6 Wasserstein GAN with Gradient Penalty	7
2.7 Μετρικές εικόνας (PSNR, SSIM)	8
2.8 YOLO για ανίχνευση πηγών	9
2.8.1 Data & Pre-Processing	9
2.8.2 Training	9
2.8.3 Μετρικές αξιολόγησης	10
3 Πειραματικά Αποτελέσματα	10
3.1 Ποιοτικά αποτελέσματα	10
3.1.1 Πρόοδος κατά την εκπαίδευση	10
3.1.2 Σύγκριση στο test set	14
3.2 Ποσοτικά αποτελέσματα	16
3.2.1 Καμπύλες Loss	17
3.2.2 PSNR και SSIM (validation set)	19
3.2.3 Αποτελέσματα στο test set	20
3.2.4 Αποτελέσματα detection με YOLOv8	21
4 Συμπεράσματα & Μελλοντική έρευνα	22
4.1 Κύρια Ευρήματα	22
4.2 Περιορισμοί της μελέτης	22
4.3 Προτάσεις για μετέπειτα έρευνα	23

1 Εισαγωγή

Η τομογραφία εκπομπής ποζιτρονίων (PET) αποτελεί μία από τις πιο ευαίσθητες τεχνικές ιατρικής απεικόνισης για την ανίχνευση functionalities στο ανθρώπινο σώμα. Ωστόσο, ένα από τα βασικά προβλήματα που επηρεάζουν την ποιότητα των εικόνων PET είναι η εξασθένιση (attenuation) των εκπεμπόμενων φωτονίων καθώς αυτά διαπερνούν τους ιστούς του σώματος. Η εξασθένιση οδηγεί σε υποεκτίμηση της πραγματικής συγκέντρωσης του ραδιοφαρμάκου και κατά συνέπεια σε χαμηλότερη ακρίβεια διάγνωσης. Για την αντιμετώπιση αυτού του προβλήματος, οι περισσότερες κλινικές εφαρμογές βασίζονται σε εικόνες υπολογιστικής τομογραφίας (CT), οι οποίες παρέχουν έναν χάρτη εξασθένισης ώστε να διορθωθεί η PET εικόνα. Ωστόσο, η χρήση του CT προσθέτει επιπλέον δόση ακτινοβολίας και αυξάνει το κόστος της εξέτασης. Τα τελευταία χρόνια έχουν προταθεί μέθοδοι βασισμένες στη βαθιά μάθηση (deep learning) για την παραγωγή διορθωμένων (Attenuation Corrected-AC) PET εικόνων απευθείας από τις μη διορθωμένες (Non Attenuation Corrected-NAC) εικόνες, χωρίς τη χρήση CT. Ιδιαίτερο ενδιαφέρον παρουσιάζουν τα Generative Adversarial Networks (GANs), τα οποία μπορούν να μετατρέψουν τις NAC εικόνες σε AC εικόνες μαθαίνοντας από κατάλληλα ζευγάρια δεδομένων. Στην παρούσα εργασία, διερευνούμε τη χρήση GANs για Attenuation Correction σε PET δεδομένα που έχουν παραχθεί μέσω προσομοιώσεων Monte Carlo στο GATE. Αρχικά σαν baseline, υλοποιείται ένα βασικό μοντέλο Pix2Pix, το οποίο στηρίζεται σε U-Net αρχιτεκτονική για τον generator και σε PatchGAN για τον discriminator. Στη συνέχεια, υλοποιείται ένα πιο σταθερό και θεωρητικά θεμελιωμένο μοντέλο Wasserstein GAN με Gradient Penalty (WGAN-GP), με στόχο τη βελτίωση της ποιότητας των παραγόμενων εικόνων και την σταθεροποίηση του training. Η αξιολόγηση των αποτελεσμάτων γίνεται με τρεις τρόπους: (α) μέσω μετρικών, (β) μέσω της σύγκρισης της generated εικόνας με τις αρχικές NAC αλλά και τις ground truth AC και (γ) μέσω της επίδοσης του YOLO στον εντοπισμό των ραδιοπηγών, συγκρίνοντας τις μη διορθωμένες NAC εικόνες με τις εικόνες που έχουν διορθωθεί μέσω GAN.

2 Εργαλεία & Μέθοδοι

2.1 Τομογραφία PET

2.1.1 Βασικές αρχές PET

Στην τομογραφία εκπομπής ποζιτρονίων (PET) χορηγείται σε κάποιον ασθενή ένα ραδιοφάρμακο το οποίο εκπέμπει β^+ σωμάτια (π^- , ^{18}F). Τα παραγόμενα ποζιτρόνια μετά από ένα πολύ μικρό διάστημα (0.01ns-150ns) [1] εξαϋλώνονται με ένα ηλεκτρόνιο των ιστών του ασθενούς και στη συνέχεια παράγονται δύο φωτόνια των 511 keV που κινούνται σχεδόν αντιδιαμετρικά (180°). Ο δακτύλιος ανιχνευτών του PET καταγράφει ταυτόχρονες (coincidences) αφίξεις φωτονίων και, μέσω Lines of Response, αναδομεί τη χωρική κατανομή του ραδιοφαρμάκου. Η χωρική ανάλυση και η ενασθησία εξαρτώνται από την ενεργειακή ανάλυση των ανιχνευτών, τα χρονικά παράθυρα καταγραφής και τη γωνιακή κάλυψη [2].

2.1.2 Γραμμή Απόκρισης (Line of Response) και Sinogram

Κάθε ζεύγος φωτονίων που ανιχνεύεται ταυτόχρονα (coincidence event) ορίζει μία γραμμή απόκρισης (Line of Response, LOR), η οποία ενώνει τα σημεία αλληλεπίδρασης των δύο 511 keV γ φωτονίων στον δακτύλιο των ανιχνευτών [3]. Οι μετρήσεις ανά LOR καταγράφονται ως *true*, *scattered* ή *random counts*, ανάλογα με το αν τα φωτόνια έχουν υποστεί σκεδασμό ή είναι μη συγγενή ζεύγη μέσα στο time window που ορίζουμε (συνήθως 8–16 ns) [3].

Η δομή των LORs μεταφέρεται σε έναν πίνακα *sinogram*, όπου:

- Η οριζόντια διάσταση s σχετίζεται με την απόσταση της LOR από το κέντρο του δακτυλίου (chord offset).

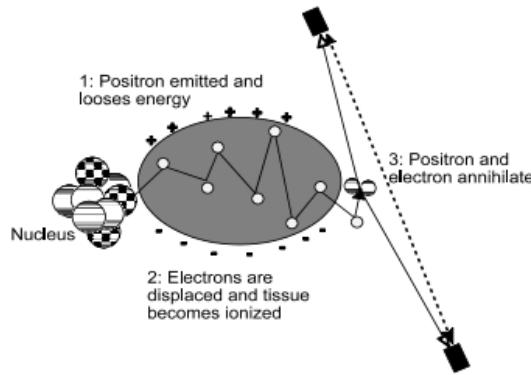


Figure 1: Βασικές αρχές λειτουργίας PET: αποβολή ενέργειας, θερμοποίηση των ποζιτρονίων, εξαύλωση με παραγωγή ζεύγους φωτονίων και LOR.

- Η κατακόρυφη διάσταση β είναι η μέση γωνία των δύο ανιχνευτών, $\beta = (P + Q)/2$, με P, Q τους δείκτες των κρυστάλλων.

Για σημειακή πηγή σε απόσταση d και γωνία α από το κέντρο, κάθε LOR ικανοποιεί τη σχέση

$$s = d \cos(\beta - \alpha)$$

η οποία απεικονίζεται στο sinogram ως καμπύλη συνημιτονοειδούς σχήματος [3].

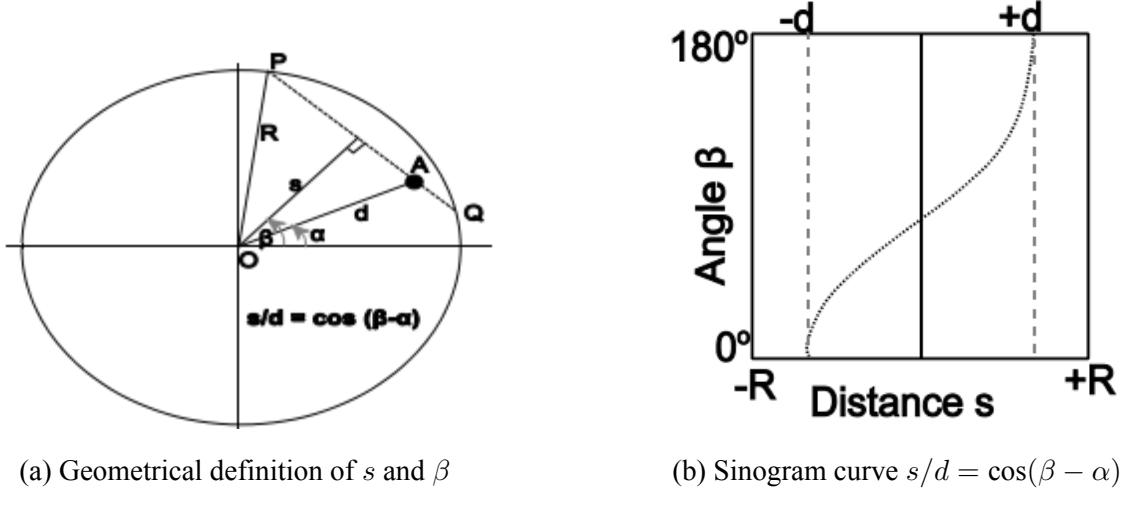


Figure 2: (a) Ορισμός των παραμέτρων s , d , R , β , α . (b) Σχηματική απεικόνιση της καμπύλης στο sinogram.

2.1.3 Attenuation και Scattering

Η εξασθένιση (attenuation) συμβαίνει όταν τα 511 keV φωτόνια αλληλεπιδρούν με τους ιστούς (απορρόφηση ή σκέδαση) πριν φτάσουν στους ανιχνευτές, με αποτέλεσμα τη μείωση των *true counts* και υποεκτίμηση της πραγματικής δραστηριότητας στους αντίστοιχους LORs [4]. Αυτό οδηγεί σε σκούρες (low counts) περιοχές στην ανακατασκευασμένη εικόνα, όπου η συγκέντρωση του ραδιοφαρμάκου φαίνεται χαμηλότερη από την πραγματική και διασκορπισμένη σε μεγαλύτερη ακτίνα.

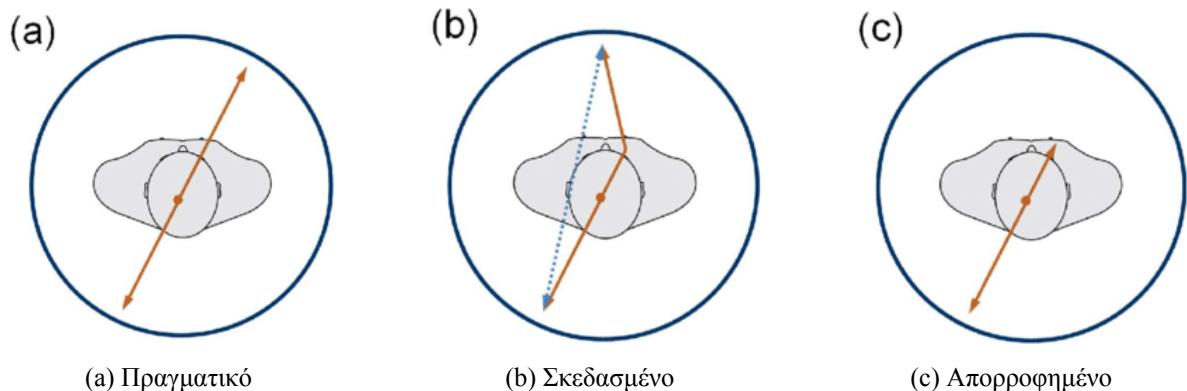


Figure 3: (a) LOR σε true coincidence, (b) LOR από σκεδασμένο φωτόνιο, (c) Απορρόφηση φωτονίου με αποτέλεσμα να μην καταγραφεί από τους κρυστάλλους.

2.2 GATE

2.2.1 GATE Tool

Το GATE (Geant4 Application for Tomographic Emission) είναι ένα open-source toolkit βασισμένο στον Geant4, σχεδιασμένο για προσομοιώσεις PET/SPECT¹. Βασίζεται στο *Monte Carlo* πλαίσιο του Geant4 και προσομοιώνει βήμα προς βήμα την αλληλεπίδραση σωματιδίων με την ύλη σε τρία κύρια στάδια:

(1) Πρωτογενή γεγονότα: για κάθε διάσπαση ραδιονουκλιδίου δειγματοληπτείται ο χρόνος αποσύνθεσης, προσομοιώνεται η εκπομπή ποζιτρονίου, γίνεται tracked μέχρι την εξαύλωση και τέλος, παράγεται ζεύγος φωτονίων 511 keV.

(2) Μεταφορά σωματιδίων: κάθε φωτόνιο προχωρά σε διαδοχικά βήματα. Το μήκος ελεύθερης διαδρομής l προκύπτει από $l = -\ln(\xi)/\mu$, όπου ξ τυχαίος αριθμός και μ ο συνολικός συντελεστής εξασθένισης. Στη συνέχεια επιλέγεται διεργασία (φωτοηλεκτρικό, Compton ή και Rayleigh), ενημερώνονται η ενέργεια και η διεύθυνση και το tracking συνεχίζεται ώσπου η ενέργεια πέσει κάτω από 1 keV ή το σωματίδιο εξέλθει της γεωμετρίας. Επαναλαμβάνοντας αυτή τη διαδικασία για εκατομμύρια γεγονότα επιτυγχάνεται στατιστική σύγκλιση.

(3) Ψηφιοποίηση και καταγραφή: τα σημεία αλληλεπίδρασης στους κρυστάλλους μετατρέπονται σε *hits*. Τα ζεύγη hits που πληρούν το κριτήριο συγχρονισμού του time window (8 ns) καταγράφονται ως λίστα *coincidences* γεγονότων για περαιτέρω επεξεργασία. Τα coincidences αποθηκεύονται σε αρχεία root.

To GATE συνοπτικά υποστηρίζει:

- Ορισμό λεπτομερούς γεωμετρίας (δακτύλιος ανιχνευτών, phantoms, attenuation maps).
 - Ενεργοποίηση και απενεργοποίηση physics lists.
 - Παραγωγή digitized events (coincidences) και εξαγωγή ROOT αρχείων.
 - Διαχείριση παραμέτρων (θερμοκρασία, energy window, time window)

¹<http://www.opengatecollaboration.org>

2.2.2 Μεθοδολογία προσομοιώσεων

Οι προσομοιώσεις έγιναν σε VirtualBox με εγκατεστημένο GATE (v9.0)². Ένα python script παράγει παραμέτρους σε .txt αρχείο (ακτίνες phantom, ακτίνες πηγών, θέσεις phantom και θέσεις πηγών), οι οποίες εισάγονται στα macro files του GATE για να παραχθούν εικόνες NAC και AC.

Macro files προσομοίωσης

Χρησιμοποιήθηκαν τέσσερα βασικά templates:

- PET_NAC_single.mac για μονή πηγή (NAC)
- PET_AC_single.mac για μονή πηγή (AC)
- PET_NAC_double.mac για διπλή πηγή (NAC)
- PET_AC_double.mac για διπλή πηγή (AC)

Κάθε macro αρχείο ορίζει:

- **Γεωμετρία ανιχνευτών** — δακτύλιος 16 BGO κρυστάλλων, εσωτερική διάμετρος 800 mm.
- **Phantoms**
 - Σφαιρικό phantom γεμισμένο με Water.
 - Ακτίνα R από 5 cm έως 200 cm.
 - Το phantom τοποθετείται με βάση τις παραμέτρους (x, y) για την θέση του ως reference για τη θέση της πηγής.
- **Πηγές (single / double)**
 - Single source: ισότροπη σημειακή πηγή ^{18}F , με activity 2000 Bq, συντεταγμένες (x, y) , ίδιες με του phantom στο οποίο ανήκουν.
 - Double source: δύο ισότροπες σημειακές πηγές, activity 2000 Bq έκαστη, θέσεις (x_1, y_1) και (x_2, y_2) ίδιες με τα phantoms που ανήκουν.
- **Physics lists**
 - NAC macros: Rayleigh, Compton, Bremsstrahlung ενεργοποιημένα.
 - AC macros: τα παραπάνω απενεργοποιούνται.
- **Digitizer** — time window για coincidence 8 ns.

Workflow

Για κάθε γραμμή του αρχείου parameters_single.txt (μονή πηγή) εκτελείται το script run_all_single.sh, το οποίο:

1. Διαβάζει παραμέτρους (R, SR, X, Y).
2. Γεννάει αυτόματα δύο macro αρχεία (run_\$count_nac.mac και run_\$count_ac.mac) με sed.
3. Εκκινεί το GATE για κάθε macro, αποθηκεύοντας τα ROOT outputs σε output/NAC και output/AC.

Αντίστοιχα, για διπλή πηγή χρησιμοποιείται το run_all_double.sh, το οποίο:

²<https://opengate.readthedocs.io/en/latest/vgate.html>

1. Διαβάζει παραμέτρους δύο πηγών από parameters_double.txt.
2. Δημιουργεί και τρέχει macro αρχεία run_\$count_nac.mac / run_\$count_ac.mac για NAC/AC.
3. Αποθηκεύει τα αποτελέσματα σε output/NAC και output/AC.

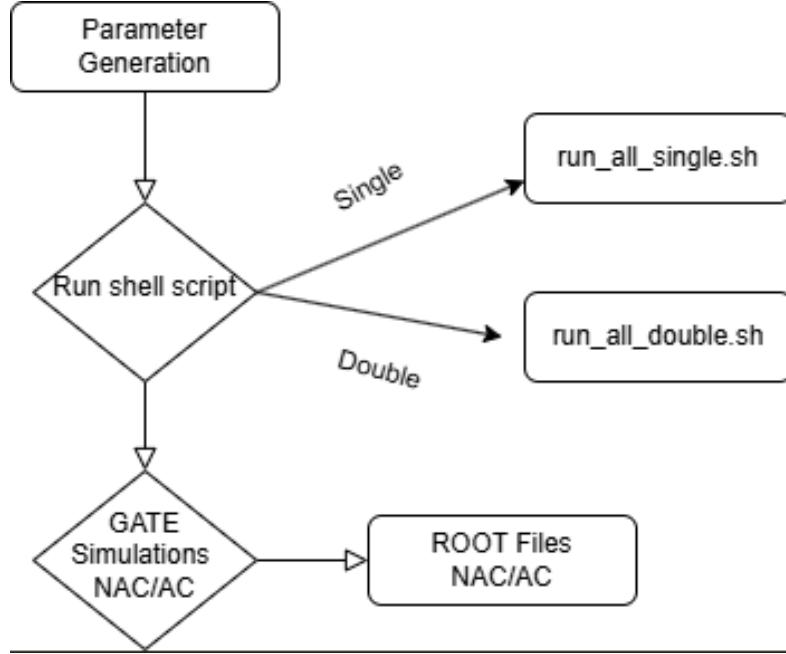


Figure 4: Workflow παραγωγής ROOT αρχείων

2.3 Ανακατασκευή εικόνων

Για την ανακατασκευή των εικόνων από τα αρχεία ROOT ακολουθήθηκαν τα παρακάτω βήματα:

1. **Φόρτωση coincidences:** Τα αρχεία .root περιέχουν τα πεδία globalPosX1, globalPosY1, globalPosX2, globalPosY2 για κάθε πηγή, τα οποία διαβάζονται με τη βιβλιοθήκη uproot μέσω των scripts iter_recon.py, iter_recon_double.py. Οι παραπάνω συντεταγμένες αντιστοιχούν στις θέσεις όπου βρέθηκαν ταυτόχρονα (coincidence) το πρώτο και το δεύτερο φωτόνιο αντίστοιχα.

2. **Σχηματισμός sinogram:** Για κάθε γεγονός υπολογίζεται

$$\theta = \arctan 2(y_2 - y_1, x_2 - x_1), \quad s = d \cos(\theta - \alpha)$$

όπου d η απόσταση από το κέντρο και α η γωνία του LOR. Τα ζευγάρια (s, θ) ιστογραμμοποιούνται σε πλέγμα 200×180 (r-θ).

3. **Iterative MLEM (10 iterations):** Ο αλγόριθμος (συνάρτηση mlem_reconstruction) ξεκινά με ομοιόμορφη εικόνα $f^{(0)} = 1$ και, για $k = 1 - 10$:

$$f^{(k)} = f^{(k-1)} \times R^\dagger \left(\frac{p}{R(f^{(k-1)}) + \varepsilon} \right)$$

όπου R είναι ο μετασχηματισμός Radon, R^\dagger ο αντίστροφος μετασχηματισμός (iradon χωρίς φίλτρο) και p το sinogram. Η υλοποίηση γίνεται με skimage.transform.radon / iradon.

4. **Αποθήκευση εικόνων:** Κάθε ανακατασκευασμένη εικόνα (256×256) αποθηκεύεται ως PNG, σε δύο φακέλους: NAC και AC.

2.3.1 Single vs Double source scripts

Script	Δεδομένα εισόδου	Τέξοδοι
iter_recon.py	.../NAC_single, .../AC_single	N_{single} εικόνες NAC/AC
iter_recon_double.py	.../NAC_double, .../AC_double	N_{double} εικόνες NAC/AC

Table 1: Επεξεργασία ROOT αρχείων για μία και δύο πηγές

2.4 Dataset

Το dataset που χρησιμοποιήθηκε για την εκπαίδευση και αξιολόγηση των μοντέλων δημιουργήθηκε μέσω προσομοιώσεων στο GATE, όπως περιγράφηκε στην προηγούμενη ενότητα. Κάθε ζεύγος δεδομένων αποτελείται από μία εικόνα (NAC) και την αντίστοιχη εικόνα (AC), η οποία λειτουργεί ως label (ground truth).

Η τελική κατανομή του dataset έχει ως εξής:

- **Training set:** 798 εικόνες NAC και τα αντίστοιχα labels AC.
- **Validation set:** 200 ζεύγη εικόνων (NAC, AC) για την επιλογή υπερπαραμέτρων και την αποφυγή overfitting.
- **Testing set:** 191 ζεύγη εικόνων για την τελική αξιολόγηση των μοντέλων.

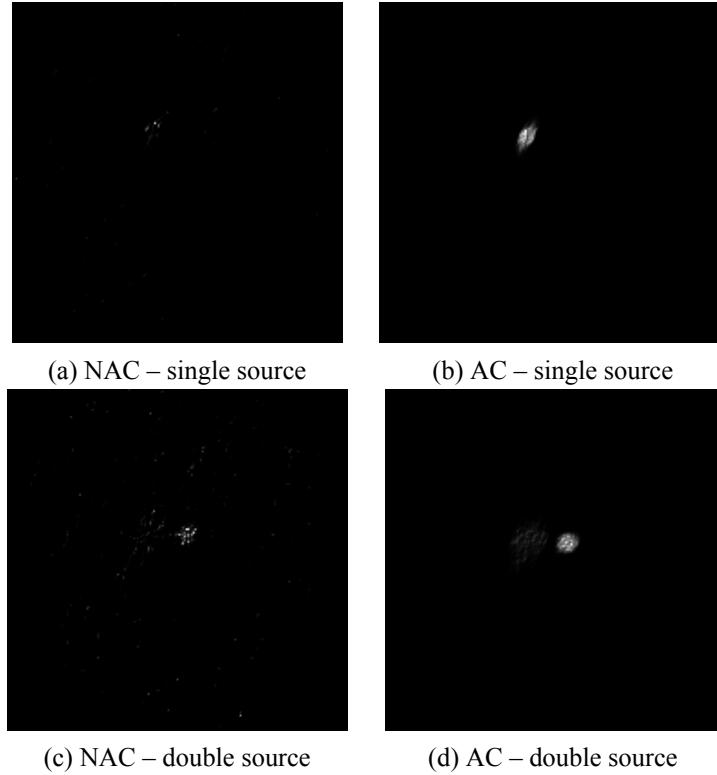


Figure 5: Ενδεικτικά παραδείγματα εικόνων NAC και AC από ανακατασκευή: (a–b) μονή πηγή, (c–d) διπλή πηγή.

2.5 Baseline: Vanilla GAN (*Pix2Pix*)

Ος baseline της παρούσας εργασίας υλοποιήθηκε ως **conditional GAN** τύπου Pix2Pix, το οποίο μαθαίνει τη χαρτογράφηση NAC \rightarrow AC σε επίπεδο εικόνας [5].

Αρχιτεκτονική

- **Generator – U-Net 8–8**

Down-path: C64 → C128 → C256 → C512 × 4

Up-path: TC512 × 4 → TC256 → TC128 → TC64 → Tanh.

Για κάθε block: Conv→InstanceNorm→LeakyReLU (ή ReLU στα up-blocks), ενώ στα τρία βαθύτερα up-blocks εφαρμόζεται Dropout 0.5.

- **Discriminator – PatchGAN 70 × 70** Συγχωνεύει είσοδο (NAC) και label (AC) κατά κανάλια. Τέσσερα Conv-blocks με φίλτρα 64→128→256→512 εξάγουν έναν πίνακα 30 × 30 πιθανοτήτων «real/fake».

Table 2: Υπερπαράμετροι εκπαίδευσης του baseline Vanilla GANN.

Batch size	16
Learning rate	2×10^{-4}
Betas (Adam)	(0.5, 0.999)
Max Epochs	500
λ_{L1}	100
Transformations	Resize 256, HFlip 0.5

Δομή training loop

1. Φόρτωση mini-batch (x, y) με εικόνες NAC και AC.

2. *Forward G*: $\hat{y} = G(x)$.

3. *Update D*:

$$\mathcal{L}_D = \frac{1}{2} [\text{BCE}(D(x, y), 1) + \text{BCE}(D(x, \hat{y}), 0)].$$

4. *Update G*:

$$\mathcal{L}_G = \text{BCE}(D(x, \hat{y}), 1) + \lambda_{L1} \cdot \|y - \hat{y}\|_1.$$

5. Αποθήκευση checkpoint κάθε 5 epochs και δειγμάτων στο φάκελο evaluation/.

Η εκπαίδευση εκτελέστηκε σε GPU A100 (Google Colab). (gen.pth.tar, disc.pth.tar) αποθηκεύονται ως baseline για τη σύγκριση στο Κεφ. 4.

2.6 Wasserstein GAN with Gradient Penalty

Με στόχο μεγαλύτερη σταθερότητα εκπαίδευσης και αποφυγή *mode-collapse*, νιοθετήσαμε την παραλλαγή **Wasserstein GAN με Gradient Penalty** (WGAN-GP). Η αρχιτεκτονική του Generator παραμένει ακριβώς η U-Net του § 2.5, ενώ ο Discriminator λειτουργεί πλέον ως *critic* (χωρίς τελική Sigmoid) και εκτιμά την Earth-Mover (Wasserstein-1) απόσταση. Η απόσταση αυτή μετρά το ελάχιστο «κόστος μεταφοράς» για να μετατρέψει τη μία κατανομή στην άλλη. Πρακτικά, πόση «μάζα» πρέπει να μετακινήσεις και πόσο μακριά. Σε αντίθεση με την κλασική KL/BCE του GAN, η Wasserstein παραμένει φραγμένη και διαφορίσιμη ακόμη κι όταν οι υποκείμενες κατανομές δεν αλληλεπικαλύπτονται, παρέχοντας σταθερότερο και πιο ενημερωτικό gradient. Για αυτό μειώνει την πιθανότητα για mode-collapse και κάνει την εκπαίδευση πιο ομαλή. Η *Wasserstein-1* απόσταση μεταξύ δύο κατανομών P_r (real) και P_g (generated) ορίζεται ως

$$W_1(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} \mathbb{E}_{(x,y) \sim \gamma} [\|x - y\|_1],$$

όπου $\Pi(P_r, P_g)$ είναι το σύνολο όλων των συζεύξεων $\gamma(x, y)$ που έχουν πρώτο περιθώριο P_r και δεύτερο P_g , ενώ το $\mathbb{E}_{(x,y) \sim \gamma}[\cdot]$ δηλώνει το μέσο όρο της συνάρτησης πάνω σε όλα τα ζεύγη (x, y) που δειγματοληπτούνται από την κατανομή γ [6].

Κύριες διαφοροποιήσεις έναντι του baseline

- Xωρίς BCE:** Η απώλεια του critic είναι $\mathcal{L}_D = \mathbb{E}[D(x, \hat{y})] - \mathbb{E}[D(x, y)] + \lambda_{GP} \mathcal{L}_{GP}$.
- Gradient Penalty** $\mathcal{L}_{GP} = (\|\nabla_{\tilde{y}} D(x, \tilde{y})\|_2 - 1)^2$, όπου $\tilde{y} = \alpha y + (1-\alpha)\hat{y}$, $\alpha \sim \mathcal{U}(0, 1)$.
- Πολλαπλά βήματα critic** 5 ενημερώσεις critic ανά 1 ενημέρωση Generator (CRITIC_ITER=5).
- Απώλεια Generator** $\mathcal{L}_G = -\mathbb{E}[D(x, \hat{y})] + \lambda_{L1} \|y - \hat{y}\|_1$.

Table 3: Υπερπαράμετροι εκπαίδευσης του WGAN-GP.

Batch size	16
Learning rate	2×10^{-4} (Adam)
Betas (Adam)	(0.5, 0.9)
Max Epochs	500
λ_{L1}	100
λ_{GP}	10
Critic updates / step	5
Augmentations	Resize 256, HFlip 0.5
Mixed precision	<code>torch.cuda.amp</code>

Δομή training loop

- Φόρτωση mini-batch** (x, y) με εικόνες NAC (είσοδος) και AC (ground-truth).
- Υπολογισμός** $\hat{y} = G(x)$ μέσω του Generator (χωρίς back-prop προς G στο βήμα του critic).
- Ενημέρωση του critic** (επαναλαμβάνεται $k = 5$ φορές)

$$\mathcal{L}_D = D(x, \hat{y}) - D(x, y) + \lambda_{GP} (\|\nabla_{\tilde{y}} D(x, \tilde{y})\|_2 - 1)^2,$$

όπου $\tilde{y} = \alpha y + (1-\alpha)\hat{y}$, $\alpha \sim \mathcal{U}(0, 1)$ και $\lambda_{GP} = 10$.

- Ενημέρωση του Generator** (μία φορά)

$$\mathcal{L}_G = -D(x, G(x)) + \lambda_{L1} \|y - G(x)\|_1, \quad \lambda_{L1} = 100.$$

- Checkpoint & δείγματα:** αποθήκευση βαρών κάθε 5 epochs και εξαγωγή παραδειγμάτων στον φάκελο `evaluation/`.

Η αντικατάσταση του BCE με τη Wasserstein απόσταση κάνει τα \mathcal{L}_D , \mathcal{L}_G μονοτονικά ως προς την ποιότητα των εικόνων, με αποτέλεσμα πιο ομαλές καμπύλες και ταχύτερη σύγκλιση

2.7 Μετρικές εικόνας (PSNR, SSIM)

Για την ποσοτική σύγκριση των παραγόμενων εικόνων (AC_{gen}) με τις αντίστοιχες ground-truth (AC_{gt}) χρησιμοποιούνται δύο συμπληρωματικές μετρικές πέρα από τα losses:

Peak Signal-to-Noise Ratio (PSNR) Μετρά την “ενεργειακή” ομοιότητα ανά pixel [7]:

$$\text{PSNR}(y, \hat{y}) = 10 \log_{10} \left(\frac{L^2}{\text{MSE}(y, \hat{y})} \right), \quad \text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2,$$

όπου $L = 1$ η μέγιστη δυνατή τιμή (οι εικόνες κανονικοποιούνται στο $[0, 1]$). Υψηλότερο PSNR αντιστοιχεί σε μικρότερο σφάλμα.

Structural Similarity Index (SSIM) Αξιολογεί φωτεινότητα, αντίθεση και δομή τοπικά σε παράθυρο 11×11 [7]:

$$\text{SSIM}(y, \hat{y}) = \frac{(2\mu_y\mu_{\hat{y}} + C_1)(2\sigma_{y\hat{y}} + C_2)}{(\mu_y^2 + \mu_{\hat{y}}^2 + C_1)(\sigma_y^2 + \sigma_{\hat{y}}^2 + C_2)},$$

με μ μέση τιμή, σ^2 διασπορά και $\sigma_{y\hat{y}}$ συνδιασπορά. Κυμαίνεται στο $[0, 1]$ και όσο μεγαλύτερο, τόσο καλύτερη και η ποιότητα της εικόνας, με την τιμή 1 να αντιστοιχεί σε ταυτόσημες εικόνες.

2.8 YOLO για ανίχνευση πηγών

Για την αξιολόγηση της διόρθωσης του attenuation στο source detection πρόβλημα, χρησιμοποιήθηκε το μοντέλο **YOLOv8n** (Nano) της Ultralytics³. Εξετάστηκαν δύο κατηγορίες εισόδων: (i) οι αρχικές εικόνες χωρίς διόρθωση (**NAC**) και (ii) οι αντίστοιχες συνθετικές εικόνες (**Generated**) που παράχθηκαν από το WGAN–GP.

2.8.1 Data & Pre-Processing

Συνολικά χρησιμοποιήθηκαν **589 NAC** εικόνες, **589 Generated** εικόνες και **589** αντίστοιχα αρχεία ετικετών (.txt) με bounding boxes (x_1, y_1, x_2, y_2) που εξάχθηκαν από τα ROOT outputs του GATE. Όλες οι εικόνες αντιστοιχούν σε επροσομοιώσεις **μονής πηγής**. Για να συμβαδίσουν με το format του YOLO, οι συντεταγμένες μετατράπηκαν ως εξής:

$$x_c = \frac{x_1 + x_2}{2}, \quad y_c = \frac{y_1 + y_2}{2}, \quad w = x_2 - x_1, \quad h = y_2 - y_1,$$

και στη συνέχεια κανονικοποιήθηκαν διαιρώντας με το πλάτος/ύψος εικόνας (256×256) ώστε όλες οι τιμές να βρίσκονται στο $[0, 1]$. Δημιουργήθηκε δομή φακέλων συμβατή με το YOLOv8 format: `images/{train, val, test}`, `labels/{train, val, test}` και αρχείο `dataset.yaml` με τις διαδρομές και την κλάση source.

2.8.2 Training

To training και το validation έγιναν χρησιμοποιώντας ίδιο αριθμό εικόνων NAC και GEN.

Table 4: Παράμετροι εκπαίδευσης YOLOv8n

Παράμετρος	Περιγραφή
Μοντέλο	YOLOv8n (Nano)
Δεδομένα	80% train / 10% val / 10% test
Epochs	50
Input size	256×256 px
Batch size	16
Optimizer	AdamW
Learning rate	δυναμικό scheduler
Augmentations	RandAugment, HSV shift, Gaussian blur, CLAHE, grayscale, flip, scale
AMP	Enabled (Automatic Mixed Precision)

³<https://github.com/ultralytics/ultralytics>

2.8.3 Μετρικές αξιολόγησης

Οι παρακάτω μετρικές εφαρμόστηκαν ξεχωριστά στο test set NAC και test set GEN.

Table 5: Ορισμοί μετρικών για αξιολόγηση ανίχνευσης

Μετρική	Ορισμός
mAP@50	Μέσος όρος precision για IoU = 0.5
mAP@50–95	Μέσος όρος precision για IoU ∈ [0.5, 0.95]
mAP@90	Precision για IoU = 0.90
Precision	Ποσοστό σωστών ανιχνεύσεων επί των προβλέψεων
Recall	Ποσοστό σωστών ανιχνεύσεων επί των υπαρκτών αντικειμένων
IoU	Μέση Intersection over Union μεταξύ predicted και ground truth boxes

Η διαδικασία αυτή επέτρεψε την εκτίμηση του κατά πόσο η βελτιωμένη ποιότητα των Generated εικόνων ενισχύει την ακρίβεια εντοπισμού σε συγκριτικά δύσκολα σενάρια με θόρυβο ή ασθενές σήμα.

3 Πειραματικά Αποτελέσματα

Ο έλεγχος στο test set εκτελέστηκε με το script `test.py`, το οποίο ακολουθεί τα βήματα:

1. **Φόρτωση μοντέλου** Τα βάρη `gen_epoch_135.pth` και `disc_epoch_135.pth` φορτώνονται σε `torch.cuda`.
2. **Επεξεργασία mini-batch** (x, y_{gt}):
 - (a) $\hat{y} = G(x)$ (forward prop)
 - (b) Απενεργοποίηση $[-1, 1] \rightarrow [0, 1]$
 - (c) Υπολογισμός μετρικών: L1, SSIM, PSNR
 - (d) Υπολογισμός **generator loss**: $\mathcal{L}_G = \text{BCE}(D(x, \hat{y}), 1) + \lambda_{L1} \|\hat{y} - y_{gt}\|_1$
 - (e) Αποθήκευση της εικόνας \hat{y} σε `img_i_GEN.png`

3.1 Ποιοτικά αποτελέσματα

3.1.1 Πρόοδος κατά την εκπαίδευση

Για την ποιοτική σύγκριση της βελτίωσης των παραγόμενων εικόνων στο vanilla αλλά και στο Wasserstein GAN χρησιμοποιήθηκαν 8 ζεύγη NAC από το evaluation set και GEN εικόνων, οι οποίες μελετώνται σε διάφορες εποχές. Ως αναφορά ακολουθούν τα labels (AC εικόνες) των ζευγών που θα μελετηθούν 6. Στην πρώτη γραμμή υπάρχουν μονές πηγές ενώ στη δεύτερη διπλές.



Figure 6: Labels-AC εικόνες

Vanilla GAN

Στις παρακάτω εικόνες 7, 8, 9, 10, 11 φαίνεται η οπτική εξέλιξη του Generator κατά την εκπαίδευση. Το epoch 125 αποτέλεσε και το καλύτερο epoch για το vanilla GAN ως προς όρους validation loss.

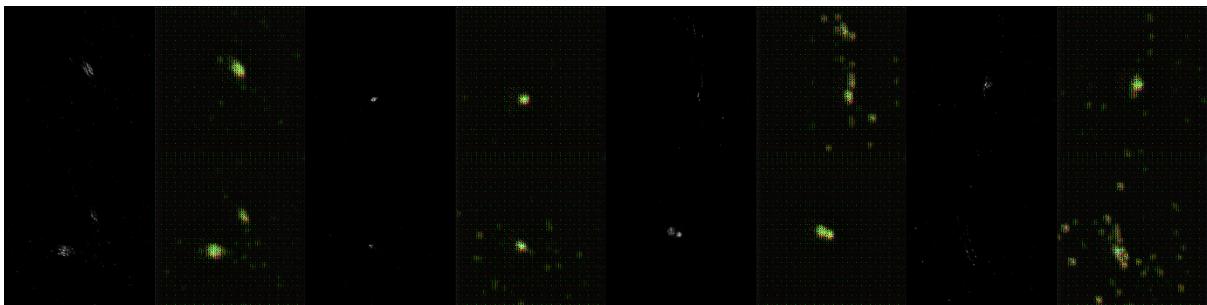


Figure 7: Vanilla GAN – αποτέλεσμα στο epoch 0.



Figure 8: Vanilla GAN – αποτέλεσμα στο epoch 10.

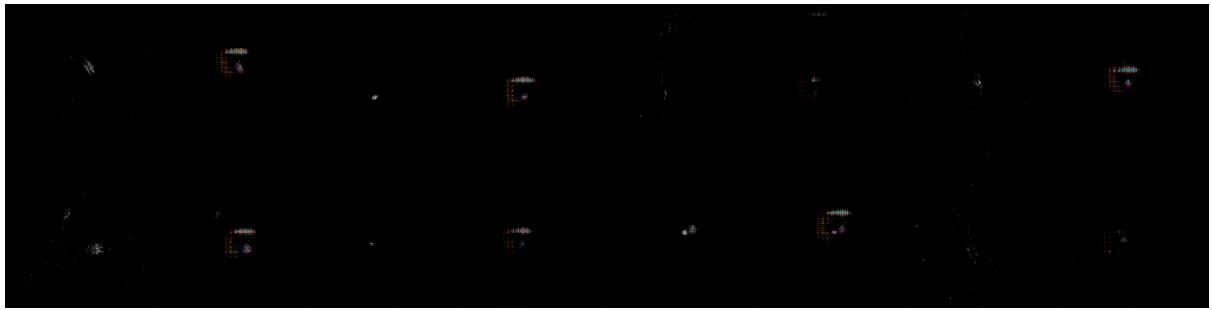


Figure 9: Vanilla GAN – αποτέλεσμα στο epoch 40.

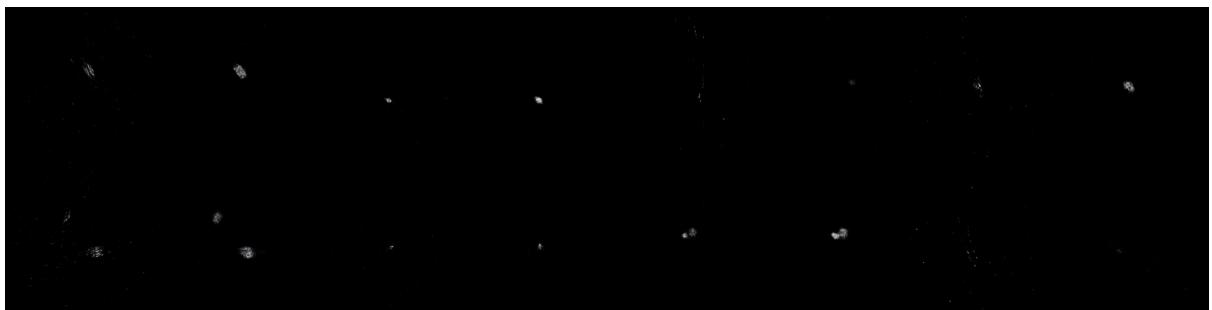


Figure 10: Vanilla GAN – αποτέλεσμα στο epoch 86.

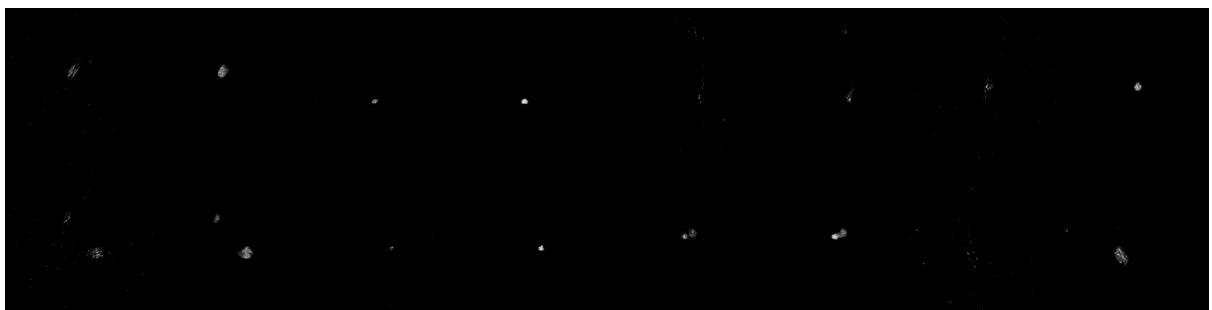


Figure 11: Vanilla GAN – αποτέλεσμα στο epoch 125.

WGAN-GP

Ακολουθούν οι αντίστοιχες εικόνες στα ίδια ζεύγη NAC και GEN 12, 13, 14, 15, 16 από το validation set για το WGAN. Το καλύτερο epoch για το WGAN ήταν το epoch 135.

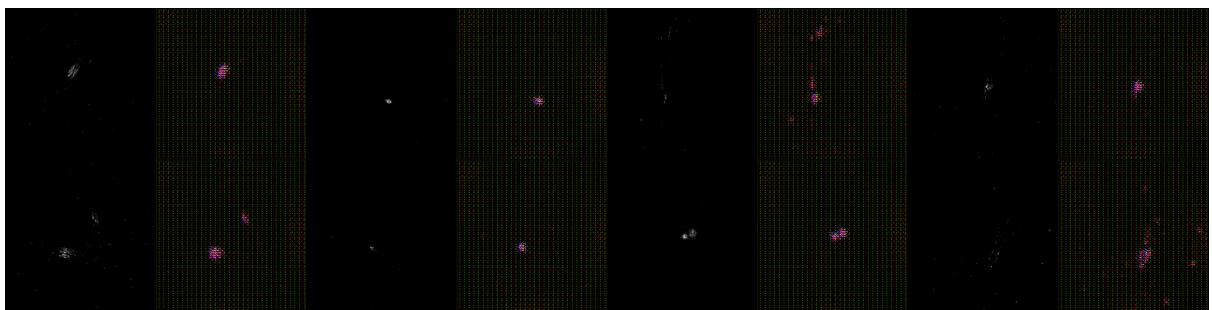


Figure 12: WGAN-GP – αποτέλεσμα στο epoch 0.

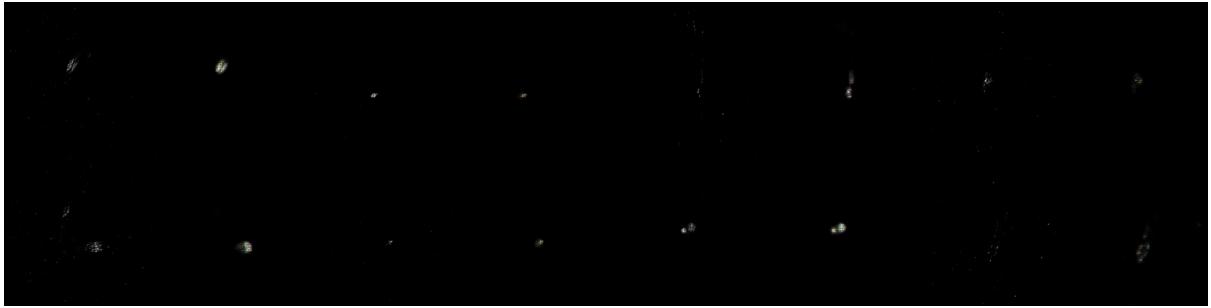


Figure 13: WGAN-GP – αποτέλεσμα στο epoch 10.



Figure 14: WGAN-GP – αποτέλεσμα στο epoch 40.

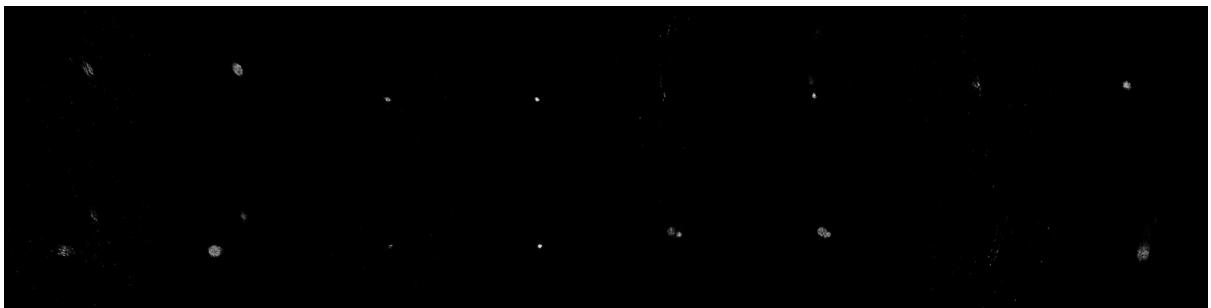


Figure 15: WGAN-GP – αποτέλεσμα στο epoch 86.

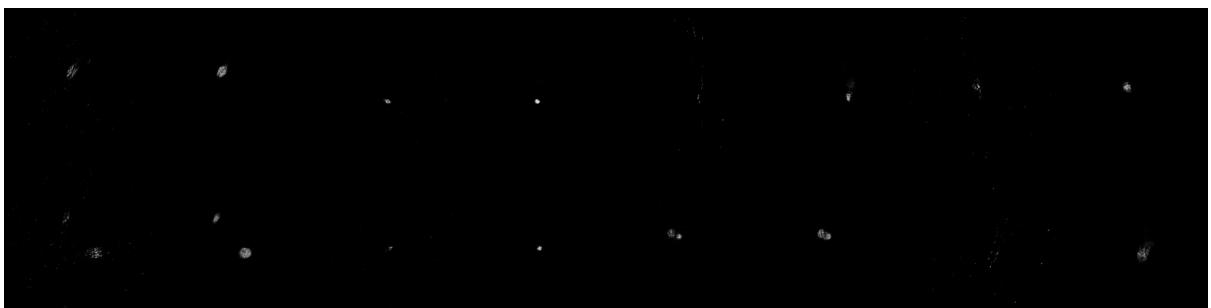


Figure 16: WGAN-GP – αποτέλεσμα στο epoch 135.

Συγκρίνοντας οπτικά τα παραγόμενα ζεύγη NAC→AC των δύο μοντέλων με το ground-truth, διακρίνεται υπεροχή του **WGAN-GP** έναντι του baseline **Vanilla GAN**:

- **Πρώιμα epochs (0–10).** Το Vanilla παρουσιάζει αποσπασματικές αναδομήσεις: μέρος του θορύβου του NAC διατηρείται και τα hot-spots εμφανίζονται θολά ή «κοκκινισμένα». Αντίθετα,

το WGAN-GP ήδη στο epoch 10 καταστέλλει τον granular θόρυβο και απομονώνει καθαρά τις πηγές.

- **Μεταγενέστερα epochs (40–125 & 135).** Στο epoch 125 το Vanilla εξακολουθεί να εμφανίζει ψευδοσημεία εκτός των πραγματικών θέσεων και ελαφρά θόλωση των spot· το WGAN-GP (epoch 135) ευθυγραμμίζεται σχεδόν πλήρως με τα labels, διατηρώντας σωστό μέγεθος και ένταση.
- **Διπλή πηγή.** Το Vanilla τείνει να συγχωνεύει ή να υποεκτιμά τη δυσδιάκριτη δεύτερη πηγή, ενώ το WGAN-GP αναπαράγει και τις δύο εστίες με σωστή σχετική φωτεινότητα (όπως με πολύ προσοχή φαίνεται για το 2ο ζεύγος της δεύτερης σειράς εικόνων).

Συνολικά, το WGAN-GP οδηγεί σε σταθερότερη εκπαίδευση και σε εικόνες AC που προσεγγίζουν πολύ πιο πιστά το ground-truth, αποφεύγοντας mode-collapses.

3.1.2 Σύγκριση στο test set

Στη φάση αυτή αντιπαραβάλλουμε άμεσα τα αποτελέσματα των δύο μοντέλων—**Vanilla GAN** (baseline) και **WGAN-GP** στο ανεξάρτητο test set (191 ζεύγη NAC/AC). Για κάθε παράδειγμα προβάλλουμε τετράδα εικόνων (i) είσοδο NAC, και (ii) ground-truth AC_{gt}, (iii) παραγόμενη AC_{gen} (Vanilla & WGAN) ώστε αξιολογήσουμε:

1. **Οπτική ποιότητα:** καθαρότητα hot-spots, καταστολή θορύβου, αλλοίωση υφής.
2. **Δομική ακρίβεια:** ευθυγράμμιση με τις πραγματικές θέσεις/σχήματα των πηγών όπως φαίνονται στη AC_{gt}.
3. **Γενίκευση:** συμπεριφορά σε δείγματα με μία ή δύο πηγές και διαφορετικές ακτίνες phantoms.

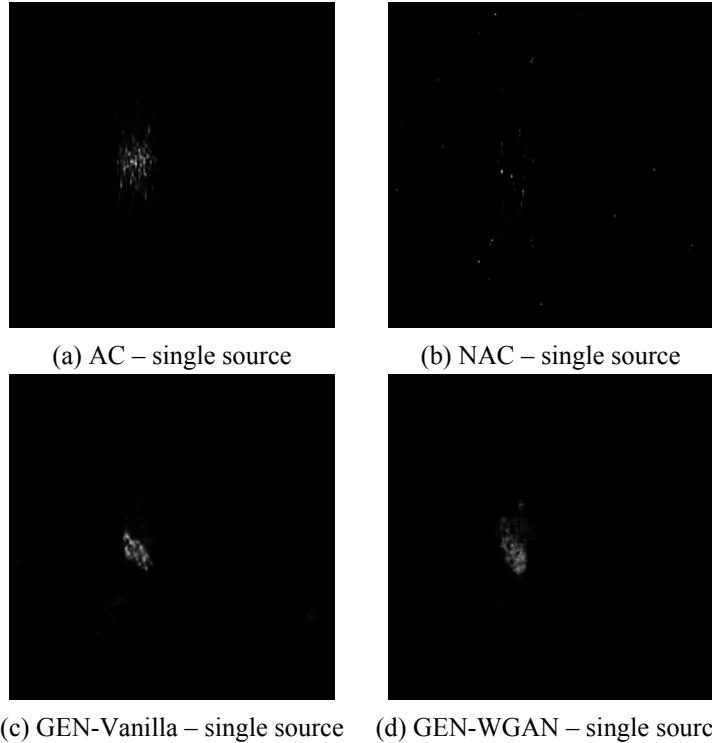


Figure 17: Παράδειγμα μονής πηγής από το test set: (a) ground-truth AC, (b) μη διορθωμένη NAC, (c) έξοδος του Vanilla GAN και (d) έξοδος του WGAN-GP. Το WGAN-GP αποδίδει πιο σωστά το μέγεθος και το σχήμα της πηγής.

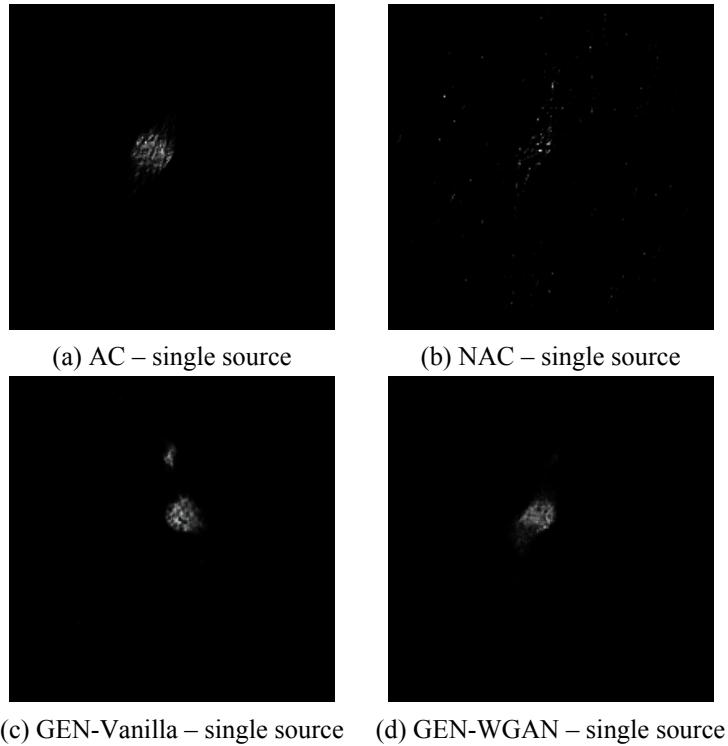


Figure 18: Παράδειγμα μονής πηγής από το test set: (a) ground-truth AC, (b) μη διορθωμένη NAC, (c) έξοδος του Vanilla GAN και (d) έξοδος του WGAN-GP. Το WGAN-GP αποδίδει καλύτερα την πηγή, ενώ το baseline δημιουργεί artefact δεύτερης πηγής.

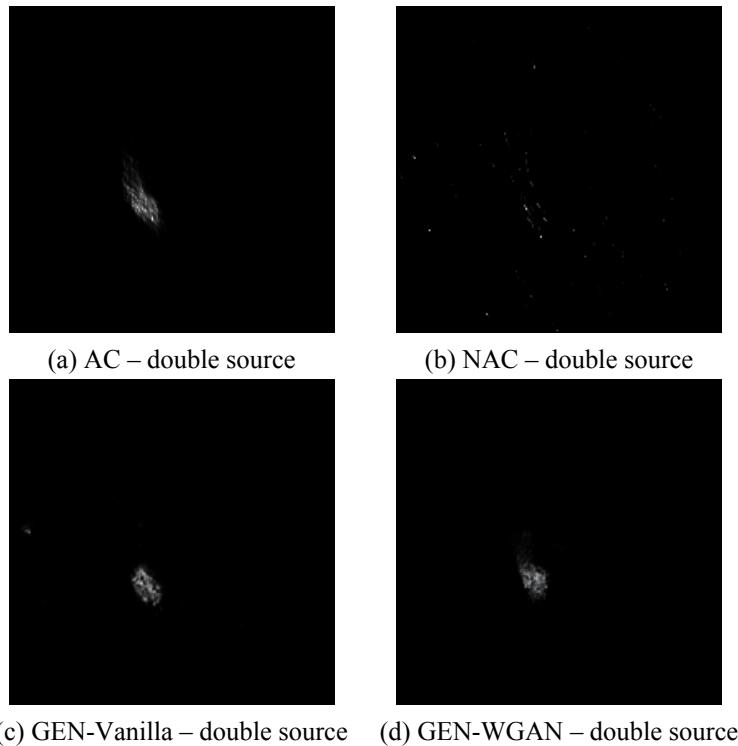


Figure 19: Παράδειγμα διπλής (επικαλυπτόμενης) πηγής από το test set: (a) ground-truth AC, (b) μη διορθωμένη NAC, (c) έξοδος του Vanilla GAN και (d) έξοδος του WGAN-GP. Το WGAN-GP αποδίδει καθαρότερα τις δύο πηγές και μειώνει τον υπόλοιπο θόρυβο σε σύγκριση με το baseline.

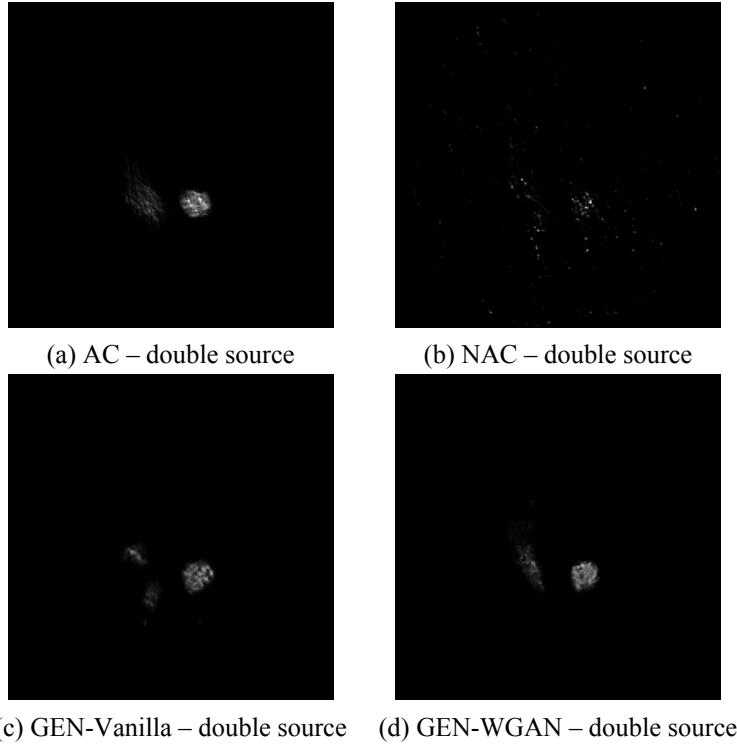


Figure 20: Παράδειγμα διπλής πηγής από το test set: (a) ground-truth AC, (b) μη διορθωμένη NAC, (c) έξοδος του Vanilla GAN και (d) έξοδος του WGAN-GP. Το WGAN-GP αποδίδει καθαρότερα τις δύο πηγές, ενώ το baseline φαίνεται να δημιουργεί artefact τρίτης πηγής

Συνοψίζοντας τις οπτικές συγκρίσεις των Σχ. 17–20, το **WGAN-GP** αποδίδει συστηματικά εικόνες πιο κοντά στο ground-truth AC σε σχέση με το **Vanilla Pix2Pix**:

- *Μονή πηγή* (Σχ. 17, 18): το WGAN-GP αναπαράγει πιστά το μέγεθος και το σχήμα της πηγής, καταστέλλοντας σχεδόν πλήρως τον υποκείμενο θόρυβο. Το baseline αφήνει υπολειμματικές κηλίδες, ενώ φαίνεται να δημιουργεί και artifacts.
- *Διπλή πηγή* (Σχ. 19, 20): το WGAN-GP διαχωρίζει καθαρά τις δύο εστίες και διατηρεί μηδενικό background, ενώ ο Vanilla είτε δημιουργεί μία μόνο πηγή είτε δημιουργεί ψευδή τρίτη.

Τα παραδείγματα επιβεβαιώνουν ότι το WGAN-GP παρέχει καλύτερη μείωση θορύβου από τις NAC και μειώνει αισθητά τα artifacts σε σύγκριση με το κλασικό adversarial loss του Vanilla GAN.

3.2 Ποσοτικά αποτελέσματα

Προχωρώντας στα ποσοτικά αποτελέσματα, θα παρουσιάσουμε για το baseline και για το WGAN αντίστοιχα τις καμπύλες loss για discriminantor και generator (train & val), καθώς και τις μετρικές test loss, PSNR, SSIM για το test set. Τα train loss curves απεικονίζουν το loss/batch, ενώ τα val loss curves απεικονίζουν το loss/epoch.

3.2.1 Καμπύλες Loss

Vanilla GAN

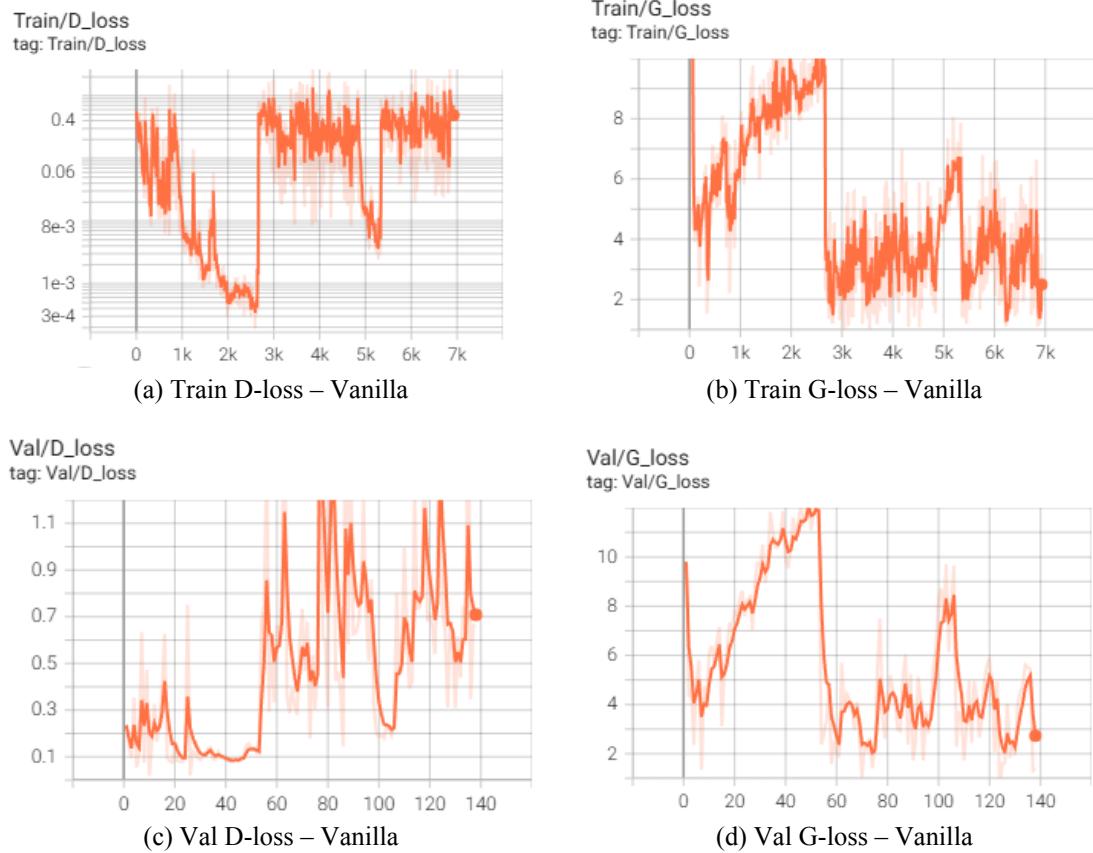


Figure 21: Εξέλιξη των loss curves για το baseline.

Οι καμπύλες Train/Val του Vanilla Pix2Pix (Fig. 21) αποτυπώνουν την αναμενόμενη διαδικασία εκπαίδευσης BCE-GAN:

- **Train-D:** απότομη πτώση στα πρώτα 2 k batches, έπειτα έντονες διακυμάνσεις και τρία μεγάλα spikes (≈ 3 k, 4.5 k, 6.5 k), ένδειξη ότι ο Discriminator γίνεται κατά διαστήματα υπερβολικά ισχυρός.
- **Train-G:** συμμετρική συμπεριφορά, όσο ο D κερδίζει (spike προς τα κάτω), ο G χάνει (spike προς τα πάνω).
- **Val-D / Val-G:** σταθερή μείωση έως epoch 40 και στη συνέχεια ανεβοκατεβάσματα, που υποδεικνύουν overfitting του D, κάτι το οποίο μπορεί να δικαιολογεί στο ότι το Vanilla GAN δημιουργούσε μία πηγή αντί για δύο στο κεφάλαιο §3.1. Ελάχιστο val loss για τον generator παρατηρείται στο epoch 125.

WGAN

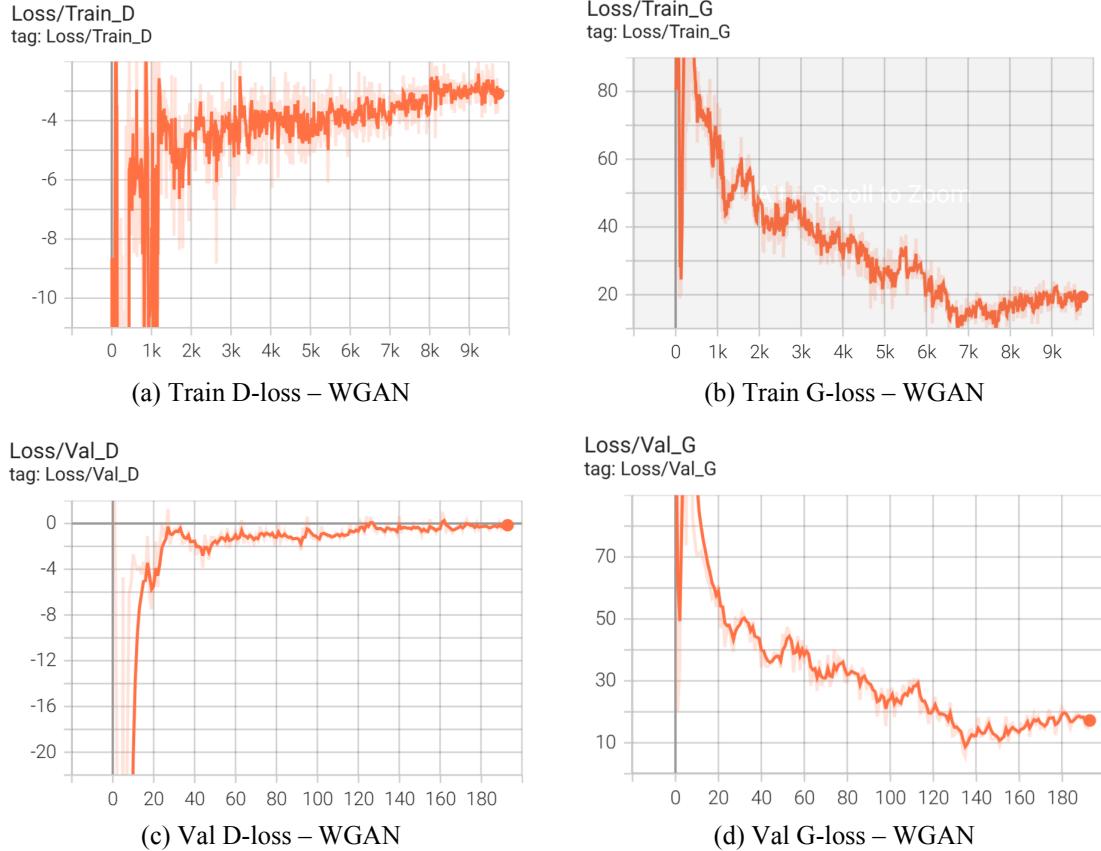


Figure 22: Εξέλιξη των loss curves για το WGAN.

Στο WGAN (Fig. 22) παρατηρούμε διαφορετική συμπεριφορά:

- **Train–Critic (D):** γρήγορη μετατόπιση προς τιμές κοντά στο 0 και σταδιακή εξομάλυνση· η μικρή αρνητική τάση είναι φυσιολογική, καθώς ο Critic μαθαίνει την ierárχηση real–fake.
- **Train–G:** μονοτονικά πτωτική, χωρίς «εκρήξεις»· δηλώνει σταθερή βελτίωση του Generator.
- **Val–Critic / Val–G:** ομαλές και σταθερές καμπύλες γεγονός που μαρτυρά απουσία overfitting και ισορροπημένη αντιπαλότητα. Ελάχιστο val loss για τον generator παρατηρείται στο epoch 135, στη συνέχεια φαίνεται να υπάρχει μία μικρή, αλλά, ανοδική τάση.

Συγκριτικά σχόλια: Σε αντίθεση με το baseline, όπου τα spikes υπονομεύουν τη σύγκλιση, το WGAN-GP εμφανίζει μονότονα και ερμηνεύσιμα loss–curves. Η ομαλότητα του critic μεταφράζεται σε σταθερή μείωση του Generator loss και τελικά σε υψηλότερη ποιότητα εικόνας.

3.2.2 PSNR και SSIM (validation set)

Vanilla GAN

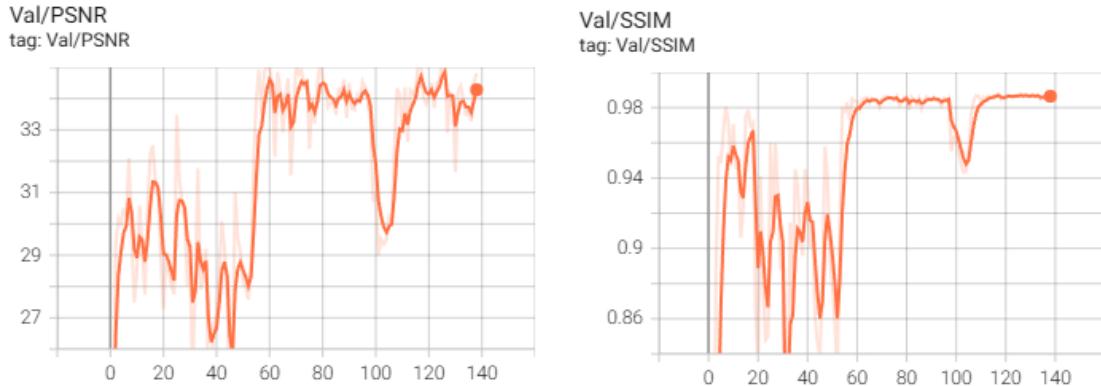


Figure 23: PSNR και SSIM ανά epoch – Vanilla GAN.

Για τα metrics PSNR και SSIM του Vanilla GAN παρατηρούμε τα εξής:

- PSNR:** έντονες διακυμάνσεις μέχρι το epoch 60. Βελτιώνεται από 28–29 dB σε 33–34 dB αλλά εμφανίζει απότομη πτώση γύρω στο epoch 95 προτού ανακάμψει.
- SSIM:** τιμές 0.86–0.90 στα πρώτα epochs με αυξημένη αστάθεια. Σταθεροποιείται κοντά στο 0.98 μετά το epoch 70, ωστόσο παρατηρείται πτώση (-3 %) στην ίδια περιοχή (epoch 95) όπου πέφτει και το PSNR.

WGAN

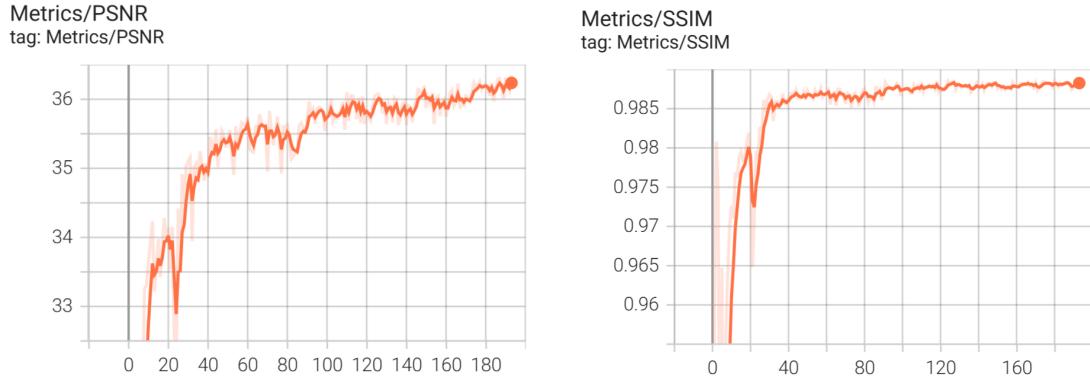


Figure 24: PSNR και SSIM ανά epoch – WGAN.

Για τα metrics PSNR και SSIM του WGAN παρατηρούμε τα εξής:

- PSNR:** μονότονη άνοδος από 33 dB (στο epoch 0) σε 36–36.5 dB, χωρίς σημαντικές ταλαντώσεις.
- SSIM:** εκτινάσσεται στο 0.985 μέσα στα πρώτα 30 epochs και παραμένει σχεδόν αμετάβλητο. Η διακύμανση δεν υπερβαίνει το 0.002.

Συγκριτικά σχόλια: Σε σχέση με το Vanilla Pix2Pix, το WGAN-GP παρουσιάζει +2-3 dB υψηλότερο τελικό PSNR και +0.005–0.007 μεγαλύτερο SSIM, ενώ οι καμπύλες του είναι σχεδόν ομαλές. Οι απότομες πτώσεις του Vanilla γύρω στο epoch 95 επιβεβαιώνουν την αστάθεια που χαρακτηρίζει την BCE-GAN εκπαίδευση, ενώ ο Wasserstein critic διατηρεί συνεχή βελτίωση χωρίς ταλαντώσεις, παρέχοντας όχι μόνο καλύτερες αλλά και πιο προβλέψιμες επιδόσεις.

3.2.3 Αποτελέσματα στο test set

Στη συνέχεια θα παρουσιάσουμε τις μετρικές και τα γραφήματα PSNR και SSIM καθώς και τα losses των generators στο test set.

Vanilla GAN

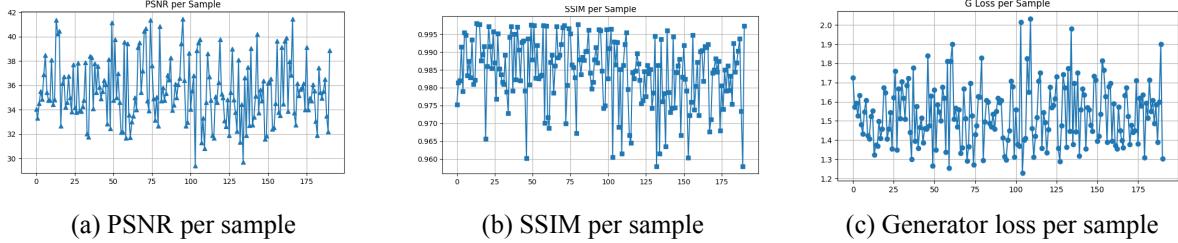


Figure 25: Κατανομή PSNR, SSIM και \mathcal{L}_G στο test set για το Vanilla GAN.

- PSNR.** Η πλειονότητα των δειγμάτων κυμαίνεται στα 34–38 dB, με αρκετές αιχμές έως 46 dB αλλά και «ουρές» κάτω από τα 32 dB. Η διασπορά υποδηλώνει ανομοιόμορφη ποιότητα από slice σε slice.
- SSIM.** Τιμές 0.96–0.995, ωστόσο 15 % των δειγμάτων πέφτουν κάτω από 0.97, γεγονός που συνδέεται με τις χαμηλότερες τιμές PSNR.
- Loss \mathcal{L}_G .** Συγκλίνει γύρω στο 1.5–1.8, αλλά παραμένουν αρκετά δείγματα με τιμές > 2.0 .

WGAN

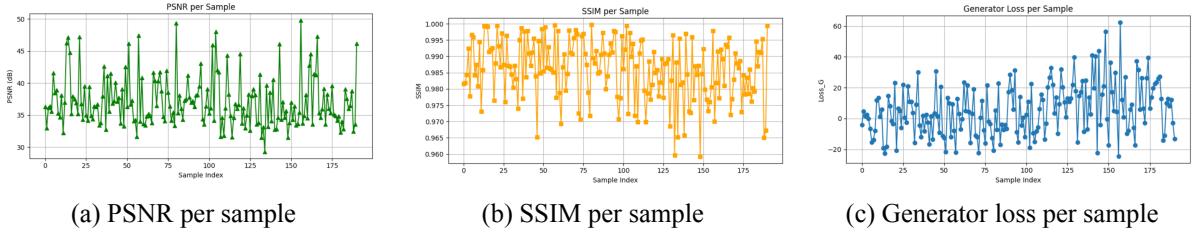


Figure 26: Κατανομή PSNR, SSIM και \mathcal{L}_G στο test set για το WGAN-GP.

- PSNR.** Η πλειονότητα μετατοπίζεται σε 35–40 dB, με κορυφές έως 50 dB και αισθητά λιγότερες τιμές < 32 dB· υποδηλώνει βελτίωση περίπου +2 dB στον μέσο όρο.
- SSIM.** Συγκεντρώνεται στενά στο 0.98–0.995, με ελάχιστα outliers—καλύτερη συνοχή σε σχέση με το Vanilla.
- Loss \mathcal{L}_G .** Μεγαλύτερη διασπορά του loss αλλά ο μέσος όρος είναι κοντά στο 5.

Συγκριτικά σχόλια Το WGAN-GP βελτιώνει το PSNR κατά περίπου +2 dB και συμπιέζει τη διασπορά SSIM, πετυχαίνοντας πιο ομοιόμορφη ποιότητα σε ολόκληρο το test set. Το διάγραμμα loss_G ανά δείγμα (Σχ. 26c) παρουσιάζει μεγαλύτερη διασπορά και μέσο όρο ~ 5 σε σχέση με το αντίστοιχο Vanilla (~ 1.5). Αυτό είναι αναμενόμενο, διότι το Wasserstein based generator loss

$$\mathcal{L}_G^{\text{WGAN}} = -D(x, \hat{y}) + \lambda_{L1} \|y - \hat{y}\|_1$$

δεν είναι άμεσα συγκρίσιμο με το BCE + λ_{L1} του Vanilla Pix2Pix. Ο όρος $-D(x, \hat{y})$ μπορεί να λάβει θετικές ή αρνητικές τιμές (βλ. τιμές έως 60 και κάτω από -20 στο plot), μεγαλώνοντας τη διακύμανση. Παρά την υψηλότερη αριθμητική τιμή, το σημαντικό είναι ότι οι τιμές συγκλίνουν γύρω από μια σταθερή

τάση χωρίς εκρήξεις, ενώ η ευρύτερη διασπορά αντανακλά την παραδοχή του Wasserstein critic να αποδίδει διαφορετικό «κόστος μεταφοράς» σε κάθε δείγμα. Συνεπώς, δεν συγκρίνουμε τα απόλυτα νούμερα των \mathcal{L}_G μεταξύ των δύο μοντέλων, αλλά τη συνοχή τους και την αντιστοίχισή τους με τις μετρικές ποιότητας (PSNR, SSIM), όπου το WGAN-GP παραμένει ανώτερο.

Table 6: Μετρικές στο test set και βέλτιστο epoch.

Model	Best epoch	GEN Loss ↓	PSNR (dB) ↑	SSIM ↑
Vanilla GAN	125	1.5330	35.54	0.9842
WGAN-GP	135	5.6188	37.10	0.9859

3.2.4 Αποτελέσματα detection με YOLOv8

Στον Πίνακα 7 συνοψίζονται οι μετρικές ανίχνευσης στα test set για τις NAC και Generated (μόνο από το WGAN-GP) εικόνες. Όπως φαίνεται, το YOLOv8n αποδίδει καλύτερα όταν τροφοδοτείται με τις generated AC εικόνες.

Table 7: Μετρικές ανίχνευσης στο test set (NAC vs Generated).

Metric	NAC	Generated
mAP@50	0.932	0.991
mAP@50–95	0.622	0.747
mAP@90	0.70	0.83
Precision	0.914	0.963
Recall	0.900	0.966
IoU	0.85	0.95

Οπτικοποίησεις Για να δημιουργηθεί οπτική σύγκριση της απόδοσης, παρατίθενται στα Σχ. 27a–27d τα bar-plots των Precision vs Recall, mAP@50–95, mAP@90 και IoU:

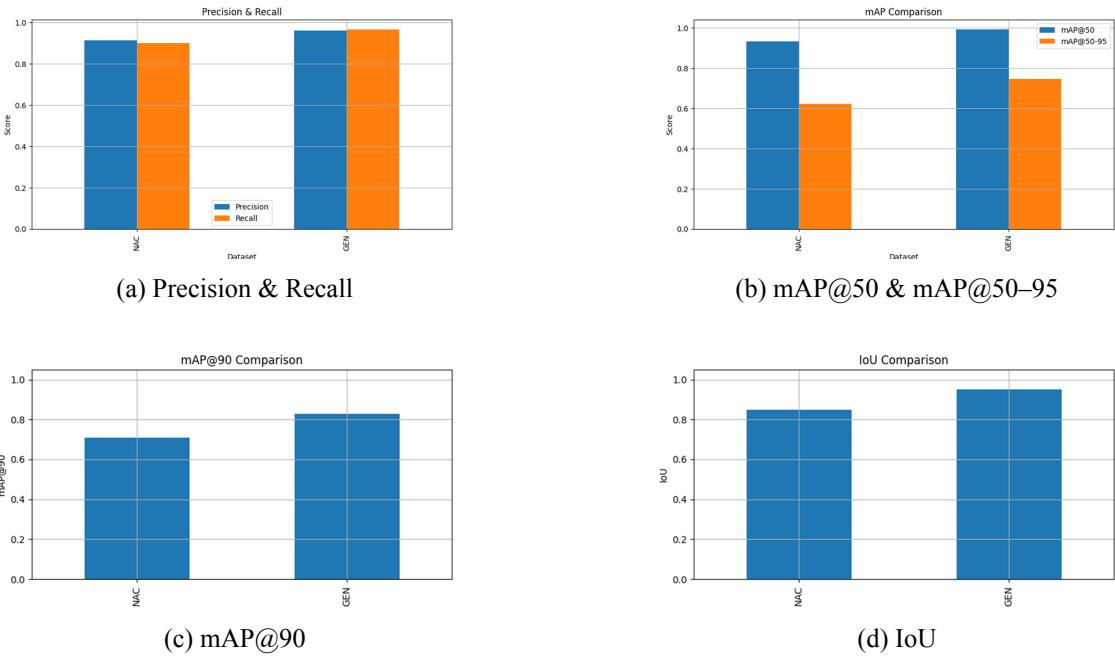


Figure 27: Σύγκριση μετρικών ανίχνευσης YOLOv8n για NAC vs Generated.

Παρατηρούμε ότι:

- Σε όλες τις μετρικές η απόδοση στο Generated σύνολο είναι υψηλότερη.
- Το mAP@50–95 και το mAP@90 αυξάνονται σημαντικά (+0.127 και +0.13), δείχνοντας ότι το μοντέλο προβλέπει πιο οξιόπιστα bounding boxes όταν η εικόνα έχει προέλθει από attenuation correction generation images.
- Η βελτίωση στο IoU (+0.10) υποδηλώνει πιο ευθυγραμμισμένες προβλέψεις με τα ground-truth boxes.

4 Συμπεράσματα & Μελλοντική έρευνα

4.1 Κύρια Ευρήματα

Τα πειραματικά αποτελέσματα επιβεβαιώνουν ότι η Wasserstein–GP εκδοχή των GANs υπερέχει σημαντικά σε σχέση με το παραδοσιακό Vanilla Pix2Pix, τόσο ως προς την ποιότητα των παραγόμενων εικόνων όσο και ως προς την υποστήριξη downstream εφαρμογών. Συγκεκριμένα, το WGAN–GP παράγει AC εικόνες με αισθητά καθαρότερο φόντο και πιο πιστή απεικόνιση των hot-spots σε μονοπηγές και διπλοπηγές σενάρια. Ποσοτικά, στο test set επιτεύχθηκε μέσος PSNR **37.1 dB** έναντι 35.5 dB του baseline και SSIM **0.990** έναντι 0.984. Οι loss-curves του WGAN–GP ήταν ομαλές, χωρίς τα μεγάλα «spikes» που εμφανίστηκαν κοντά στο 95ο epoch στο Vanilla GAN, υποδεικνύοντας πιο σταθερή εκπαίδευση. Παράλληλα, η χρήση των βελτιωμένων εικόνων ως είσοδο στο YOLOv8n οδήγησε σε σημαντική άνοδο της απόδοσης ανίχνευσης πηγών: το mAP@50 αυξήθηκε από 0.932 σε 0.991, το mAP@50–95 από 0.622 σε 0.747 και το mAP@90 από 0.70 σε 0.83, ενώ Precision και Recall βελτιώθηκαν αντίστοιχα από 0.914 → 0.963 και 0.900 → 0.966, με τον μέσο IoU να ανέρχεται στο 0.95 έναντι 0.85 στις πρωτότυπες NAC εικόνες. Αυτό καταδεικνύει ότι η attenuation correction μέσω WGAN–GP δεν βελτιώνει μόνο την οπτική ποιότητα αλλά ενισχύει και την ακρίβεια κρίσιμων downstream tasks.

4.2 Περιορισμοί της μελέτης

Παρά τη σαφή βελτίωση που καταγράφηκε, η παρούσα εργασία ενσωματώνει ορισμένες παραδοχές (biases) οι οποίες περιορίζουν τη γενίκευση των αποτελεσμάτων:

- **Απλοϊκή γεωμετρία phantoms.** Χρησιμοποιήθηκαν αποκλειστικά σφαιρικά phantoms ομοιογενώς γεμισμένα με *Water*. Στην κλινική εξέταση, ο ανθρώπινος κορμός παρουσιάζει σύνθετους ιστούς διαφορετικής πυκνότητας, γεγονός που θα επηρεάσει την κατανομή εξασθένισης.
- **Εκπαίδευση YOLO μόνο σε εικόνες με μία πηγή.** Για την αξιολόγηση της βελτίωσης στην ανίχνευση πηγών μετά τη διόρθωση χρησιμοποιήθηκαν αποκλειστικά εικόνες με μόνο μία πηγή. Η συμπεριφορά του συστήματος σε διπλές ή πολλαπλές πηγές δεν εξετάστηκε.
- **Αποκλειστικά συνθετικά δεδομένα.** Όλα τα slices προήλθαν από Monte-Carlo προσομοιώσεις απουσιάζουν συστημικοί θόρυβοι πραγματικών τομογράφων, καθώς και κίνηση ασθενούς.
- **Απουσία κλινικής επαλήθευσης.** Δεν συμπεριλήφθηκαν πραγματικά CT-based AC volumes, άρα η κλινική μεταφορά των μοντέλων απαιτεί πρόσθετη έρευνα.

4.3 Προτάσεις για μετέπειτα έρευνα

Με βάση τα παραπάνω ευρήματα, οι κατευθύνσεις για επόμενη έρευνα είναι τρεις. Πρώτον, θα πρέπει να αντικατασταθούν τα ομοιογενή υδάτινα phantoms με ρεαλιστικές τομές που περιλαμβάνουν οστά και μαλακούς ιστούς. Δεύτερον, η μελέτη της ανίχνευσης των πηγών θα πρέπει να εξετασθεί και σε εικόνες με πολλαπλές πηγές. Τρίτον, απαιτείται κλινική επαλήθευση με πραγματικά ζεύγη PET/CT προκειμένου να εκτιμηθεί η γενίκευση εκτός συνθετικού θορύβου.

References

- [1] Bangyan Huang et al. “Fast high-resolution lifetime image reconstruction for positron lifetime tomography”. In: *Communications Physics* 8.1 (2025), p. 181.
- [2] Vibhu Kapoor, Barry M McCook, and Frank S Torok. “An introduction to PET-CT imaging”. In: *Radiographics* 24.2 (2004), pp. 523–543.
- [3] Christopher J. Thompson. “PET Imaging Physics and Instrumentation”. In: 2002. url: <https://api.semanticscholar.org/CorpusID:15329565>.
- [4] Arman Rahmim and Habib Zaidi. “PET versus SPECT: strengths, limitations and challenges”. In: *Nuclear medicine communications* 29.3 (2008), pp. 193–207.
- [5] Phillip Isola et al. “Image-to-image translation with conditional adversarial networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.
- [6] Ishaan Gulrajani et al. “Improved training of wasserstein gans”. In: *Advances in neural information processing systems* 30 (2017).
- [7] Zhou Wang et al. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE transactions on image processing* 13.4 (2004), pp. 600–612.