

Πανεπιστήμιο Μακεδονίας

Νευρωνικά Δίκτυα

Style Transfer using Generative Adversarial
Networks

Γιαγκούδη Δήμητρα

Ιούλιος 2025

Περιεχόμενα

1.	Εισαγωγή	2
2.	Θεωρητικό υπόβαθρο	2
3.	Προτεινόμενη Μεθοδολογία.....	3
4.	Πειραματικά Αποτελέσματα	5
	4.1 Dataset και προεπεξεργασία	5
	4.2 Διαδικασία αξιολόγησης	5
	4.3 Loss Curves	5
	4.4 Ποσοτική Αξιολόγηση Εξόδου Μοντέλου	7
	4.5 Ποιοτική Ανάλυση Αποτελεσμάτων	7
	4.6 Συγκριτική Αξιολόγηση Εκδόσεων του CycleGAN.....	9
	4.7 Στατιστική Ανάλυση Αποτελεσμάτων	10
5.	Συμπεράσματα	10
	Βιβλιογραφία	12

Περιεχόμενα Εικόνων

Εικόνα 1: Διάγραμμα ροής (flowchart) της μεθοδολογίας που ακολουθείται για την εκπαίδευση, παραγωγή και αξιολόγηση του CycleGAN.....	4
Εικόνα 2: Αρχιτεκτονική του CycleGAN με δύο Generators και δύο Discriminators. Πηγή: [5]	5
Εικόνα 3: Καμπύλη απώλειας Generator και Discriminator για Fold 1	6
Εικόνα 4: Καμπύλη απώλειας Generator και Discriminator για Fold 2	6

Περιεχόμενα Πινάκων

Πίνακας 1: Μετρικές PSNR και SSIM για τις πέντε εικόνες δοκιμής	7
Πίνακας 2: Απεικόνιση σύγκρισης εξόδων μεταξύ Epoch 1 και Epoch 8 για τις πέντε εικόνες δοκιμής.....	8
Πίνακας 3: Μέσες τιμές PSNR και SSIM ανά epoch (1 και 8) για αξιολόγηση μοντέλων	9
Πίνακας 4: Μέσες τιμές και τυπικές αποκλίσεις PSNR και SSIM για Epoch 1 και Epoch 8	10

1. Εισαγωγή

Η παρούσα εργασία εστιάζει στην επίλυση του προβλήματος της μεταφοράς στυλ εικόνας (image style transfer), και συγκεκριμένα στην μετατροπή καλλιτεχνικών ζωγραφιών τύπου Monet σε φωτορεαλιστικές εικόνες. Το πρόβλημα αυτό ανήκει στον ευρύτερο τομέα της υπολογιστικής όρασης και της υπολογιστικής καλλιτεχνικής απεικόνισης, και παρουσιάζει ενδιαφέρον τόσο σε ερευνητικό επίπεδο όσο και σε εφαρμογές όπως η αποκατάσταση εικόνων, η καλλιτεχνική παραγωγή και η επαυξημένη πραγματικότητα.

Το βασικό πρόβλημα που αντιμετωπίζεται είναι η μετατροπή μιας καλλιτεχνικής ζωγραφιάς σε μια ρεαλιστική φωτογραφική εικόνα, χωρίς την ύπαρξη άμεσης αντιστοίχισης μεταξύ των δύο τύπων εικόνας. Με άλλα λόγια, δίνεται ως είσοδος μία εικόνα ζωγραφισμένη σε συγκεκριμένο καλλιτεχνικό στυλ (στην προκειμένη περίπτωση, έργα του Monet) και ζητείται η παραγωγή μιας φωτογραφίας που να αναπαριστά το ίδιο σκηνικό, αλλά με φυσική απόδοση φωτός, υφής και χρώματος. Το πρόβλημα αυτό εντάσσεται στην κατηγορία της μη εποπτευόμενης μετατροπής εικόνας προς εικόνα (unpaired image-to-image translation).

Για την αντιμετώπιση του προβλήματος χρησιμοποιείται μια σύγχρονη αρχιτεκτονική γεννητικών ανταγωνιστικών δικτύων (Generative Adversarial Networks – GANs), γνωστή ως CycleGAN. Το CycleGAN αποτελεί ένα ισχυρό πλαίσιο για την εκμάθηση χαρτογράφησης μεταξύ δύο κατανομών εικόνων χωρίς την ανάγκη για αντιστοιχισμένα δεδομένα. Η προσέγγιση βασίζεται στη χρήση δύο δικτύων γεννήτορα και δύο διακριτών (discriminators), τα οποία εκπαιδεύονται ανταγωνιστικά και ταυτόχρονα βελτιστοποιούν μια επιπλέον απώλεια συνέπειας κύκλου (cycle-consistency loss), ώστε να διασφαλίζεται ότι η μετατροπή από τη μία αναπαράσταση στην άλλη διατηρεί τη βασική δομή της αρχικής εικόνας. Το πρόβλημα αντιμετωπίζεται επομένως ως ένα πρόβλημα χαρτογράφησης μη εποπτευόμενης μετατροπής εικόνας προς εικόνα, χωρίς χρήση labels ή annotated δεδομένων.

Η πειραματική προσέγγιση βασίστηκε σε ένα δημόσια διαθέσιμο dataset γνωστό ως monet2photo, το οποίο περιλαμβάνει δύο σύνολα εικόνων: ένα με ζωγραφίες του Claude Monet και ένα με φωτογραφίες τοπίων που αντιστοιχούν σε παρόμοιο περιεχόμενο. Τα δεδομένα αυτά είναι μορφής στατικών εικόνων RGB και διαχωρίζονται σε υποφακέλους trainA, trainB, testA και testB, χωρίς όμως να υπάρχει αντιστοιχία εικόνα προς εικόνα. Αυτό καθιστά το πρόβλημα πιο σύνθετο, αφού η μάθηση πρέπει να βασιστεί αποκλειστικά σε στατιστικά χαρακτηριστικά των δύο κατηγοριών. Όλες οι εικόνες χρησιμοποιούνται κατόπιν επεξεργασίας (resize, normalization), με στόχο να τροφοδοτηθούν στο νευρωνικό δίκτυο με συνεπή τρόπο.

2. Θεωρητικό υπόβαθρο

Η μη εποπτευόμενη μετατροπή εικόνας προς εικόνα (unpaired image-to-image translation) αποτελεί ενεργό πεδίο έρευνας στην υπολογιστική όραση, με σημαντικές εφαρμογές σε domains όπως η τέχνη, η ιατρική απεικόνιση και η χαρτογράφηση. Ο βασικός μηχανισμός που επιτρέπει αυτού του είδους τη μετατροπή βασίζεται στη χρήση Generative Adversarial Networks (GANs), ενός πλαισίου στο οποίο δύο νευρωνικά δίκτυα, ο γεννήτορας (generator) και ο διακριτής (discriminator), εκπαιδεύονται αντιπαραθετικά: το πρώτο παράγει συνθετικές εικόνες ενώ το δεύτερο επιχειρεί να τις διακρίνει από τις πραγματικές.

Η αρχιτεκτονική CycleGAN επεκτείνει το βασικό πλαίσιο των GANs ώστε να υποστηρίζει μη εποπτευόμενη εκμάθηση χαρτογράφησης μεταξύ δύο κατανομών εικόνας, χωρίς την ανάγκη για ζευγάρια εικόνων εισόδου-στόχου. Το κύριο χαρακτηριστικό της είναι η cycle-consistency loss, η οποία διασφαλίζει ότι μια εικόνα που μετατρέπεται από το domain A στο B και ξανά πίσω στο A διατηρεί τη βασική της δομή.

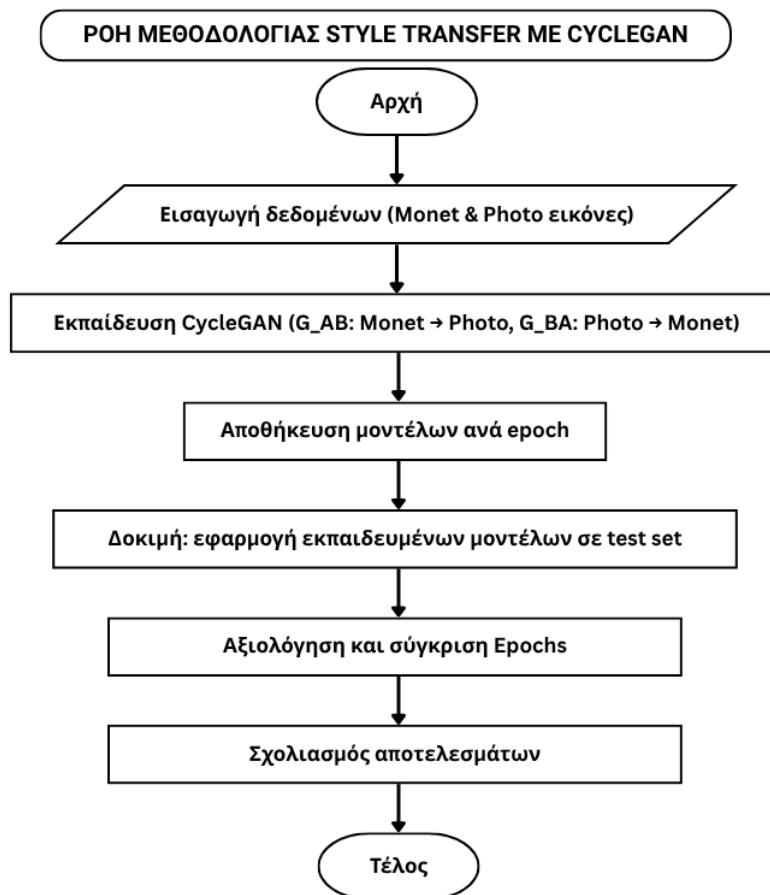
Στη σύγχρονη βιβλιογραφία, διάφορες μελέτες εστιάζουν στη χρήση των GANs για μεταφορά στυλ εικόνας και βελτιστοποίηση της απόδοσής τους σε εφαρμογές επεξεργασίας εικόνας. Σύμφωνα με τη μελέτη των Jamalpur et al. [1], η χρήση Generative Adversarial Networks (GANs) για μεταφορά καλλιτεχνικού στυλ επιτρέπει την αποδοτική μετατροπή εικόνων, διατηρώντας τη δομή του περιεχομένου και τη συνοχή της υφής. Η εργασία τους εστιάζει στη σταθερότητα της εκπαίδευσης, την αντιληπτική ποιότητα των εξόδων και τη χρήση deep learning τεχνικών για επεξεργασία υψηλής ανάλυσης. Οι Adhikari et al. [2] προτείνουν το iHQGAN, ένα υβριδικό quantum-classical GAN για unpaired image-to-image μεταφορά. Το μοντέλο τους συνδυάζει ελαφριά αρχιτεκτονική με αντιστρεψιμότητα, γεγονός που ενισχύει την αποδοτικότητα και καθιστά την τεχνική κατάλληλη για περιβάλλοντα με περιορισμένους υπολογιστικούς πόρους. Τα αποτελέσματα που παρουσιάζονται επιβεβαιώνουν τη λειτουργικότητα προσεγγίσεων τύπου CycleGAN. Στη μελέτη των Park et al. [3], παρουσιάζεται το KC-UNIT, ένα μοντέλο unpaired μεταφοράς στυλ με perceptual guidance, το οποίο εφαρμόζεται σε δεδομένα ιατρικής απεικόνισης. Οι συγγραφείς ενσωματώνουν perceptual loss στην αρχιτεκτονική του CycleGAN, με στόχο τη βελτίωση της ποιότητας και της ευκρίνειας των παραγόμενων εικόνων, επιτυγχάνοντας ικανοποιητικά αποτελέσματα σε πραγματικά δεδομένα. Τέλος, οι Karim et al. [4] συγκρίνουν πειραματικά μοντέλα paired και unpaired μεταφοράς εικόνας. Όπως καταδεικνύεται στην εργασία τους, σε περιβάλλοντα με περιορισμένη διαθεσιμότητα αντιστοιχισμένων δεδομένων, τα unpaired GANs, όπως τα CycleGAN, UNIT και CUT, παρουσιάζουν μεγαλύτερη ευελιξία και ισχυρότερη δυνατότητα γενίκευσης, καθιστώντας τα ιδανικά για πρακτικές εφαρμογές με ανεξάρτητα domains.

Η επιλογή του CycleGAN στην παρούσα εργασία δικαιολογείται τόσο θεωρητικά όσο και εμπειρικά από τη σχετική βιβλιογραφία. Οι παραπάνω μελέτες επιβεβαιώνουν ότι οι τεχνικές μη εποπτευόμενης μεταφοράς στυλ που βασίζονται σε adversarial learning και cycle-consistency μπορούν να αποδώσουν εξαιρετικά αποτελέσματα ακόμα και με σύνθετα ή ασυσχέιστα δεδομένα εικόνας.

3. Προτεινόμενη Μεθοδολογία

Η παρούσα εργασία αφορά στη μετατροπή καλλιτεχνικών εικόνων τύπου Monet σε φωτορεαλιστικές απεικονίσεις, μέσω χρήσης γεννητικών ανταγωνιστικών δικτύων και ειδικότερα του CycleGAN. Η προσέγγιση ανήκει στην κατηγορία των μη εποπτευόμενων image-to-image translation τεχνικών, και αξιοποιεί ως είσοδο μη συσχετισμένες εικόνες από δύο διαφορετικά domains: ζωγραφιές και φωτογραφίες τοπίων.

Για την αντιμετώπιση του προβλήματος χρησιμοποιείται ένα μοντέλο τύπου CycleGAN, το οποίο εκπαιδεύεται από την αρχή (from scratch) σε custom dataset που βασίζεται στο Monet2Photo του Kaggle. Το σύστημα περιλαμβάνει δύο γεννήτορες και δύο διακριτές, και εκπαιδεύεται με στόχο να μάθει τις χαρτογραφήσεις Monet-to-Photo και αντίστροφα. Η παρακάτω ροή συνοψίζει τα βασικά στάδια της μεθοδολογίας:



Εικόνα 1: Διάγραμμα ροής (flowchart) της μεθοδολογίας που ακολουθείται για την εκπαίδευση, παραγωγή και αξιολόγηση του CycleGAN

Το CycleGAN αποτελείται από τέσσερα κύρια υποδίκτυα, τα οποία συνεργάζονται σε ένα πλήρως Η αρχιτεκτονική που υλοποιήθηκε βασίζεται στο μοντέλο CycleGAN όπως προτάθηκε από τους Zhu et al. (2017), προσαρμοσμένη για εικόνες μεγέθους 256×256 και υλοποιημένη με τη βιβλιοθήκη PyTorch.

Το δίκτυο αποτελείται από δύο μετατροπείς στυλ (Generators G_{AB} και G_{BA}) και δύο διακριτές μονάδες αξιολόγησης (Discriminators D_A και D_B) με δομή PatchGAN, που αξιολογούν τοπικές περιοχές αντί ολόκληρης εικόνας, επιτρέποντας καλύτερη εκμάθηση υφής.

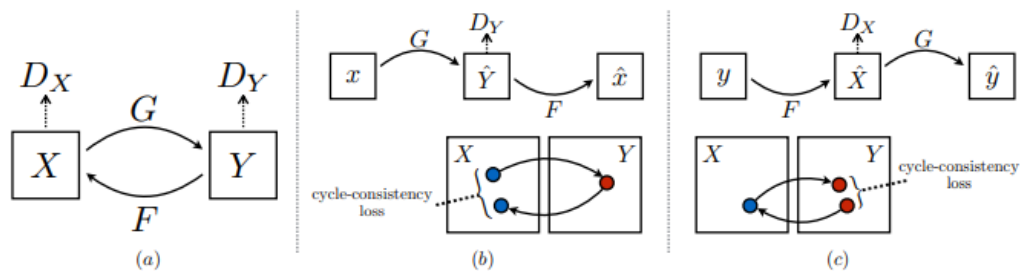
Ο Generator είναι ένα πλήρως συνελικτικό δίκτυο με την εξής δομή:

- 3 layers downsampling με Conv2D + InstanceNorm + ReLU
- 9 Residual Blocks που επιτρέπουν διατήρηση πληροφορίας και σταθερότητα κατά την εκπαίδευση
- 2 layers upsampling με ConvTranspose2D
- Ενεργοποίηση Tanh στην έξοδο για παραγωγή εικόνων κανονικοποιημένων στο διάστημα $[-1, 1]$

Το Discriminator υλοποιείται ως PatchGAN με 5 συνελικτικά επίπεδα:

- Χρήση LeakyReLU, InstanceNorm, και Conv2D
- Έξοδος ενός πίνακα (30×30) που αξιολογεί κατά τόπους την αυθεντικότητα των περιοχών της εικόνας

Η σχεδίαση αυτή επιτρέπει στο μοντέλο να μάθει τοπικά μοτίβα και να γενικεύει καλύτερα σε εικόνες υψηλής ποικιλομορφίας.



Εικόνα 2: Αρχιτεκτονική του CycleGAN με δύο Generators και δύο Discriminators. Πηγή: [5]

4. Πειραματικά Αποτελέσματα

Η παρούσα ενότητα παρουσιάζει τα αποτελέσματα που προέκυψαν από την εκπαίδευση του CycleGAN μοντέλου με στόχο τη μετάφραση στυλ από φωτογραφίες σε πίνακες του Monet. Η αξιολόγηση έγινε με χρήση της μεθόδου 5-fold cross-validation, από την οποία έχουν ολοκληρωθεί οι δύο πρώτοι folds, επιτρέποντας την ποσοτική και γραφική ανάλυση των trends κατά την εκπαίδευση.

4.1 Dataset και προεπεξεργασία

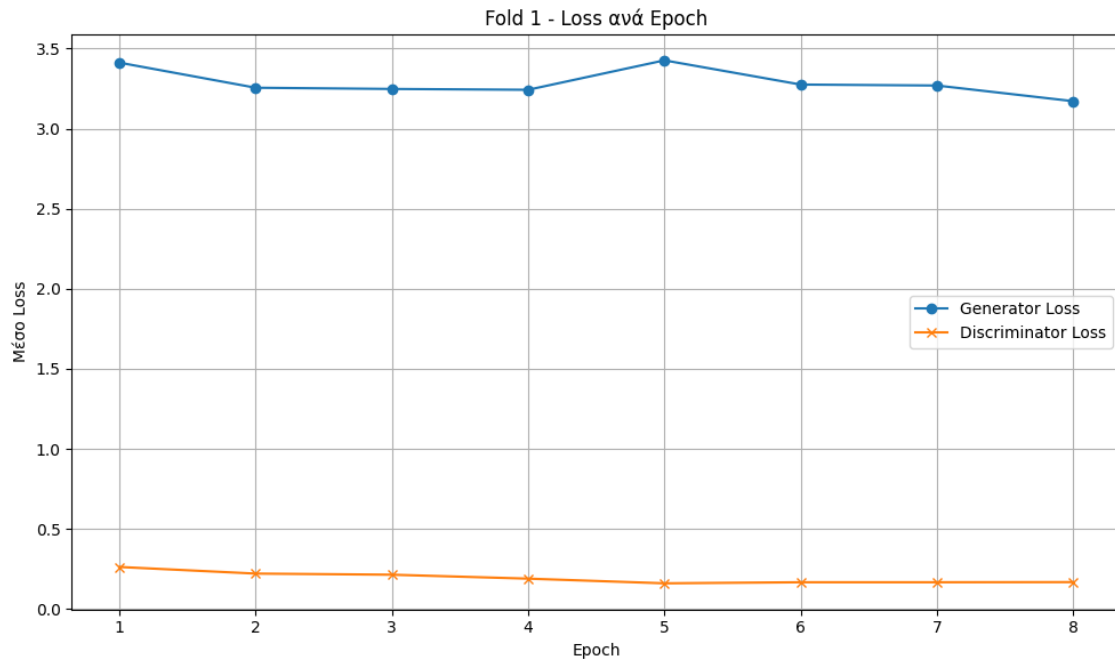
Το dataset Monet2Photo αποτελείται από δύο μη συζευγμένους χώρους εικόνων: φωτογραφίες τοπίων (domain B) και έργα ζωγραφικής του Claude Monet (domain A). Αρχικά, το σύνολο περιλάμβανε περίπου 8.000 εικόνες. Για λόγους αποδοτικότερης εκπαίδευσης και χρονικού περιορισμού, επιλέχθηκε υποδειγματοληψία: 800 εικόνες ανά domain. Οι εικόνες μετατράπηκαν σε διαστάσεις 256×256 pixels και κανονικοποιήθηκαν στο εύρος $[-1, 1]$.

4.2 Διαδικασία αξιολόγησης

Για την αξιολόγηση της επίδοσης, υιοθετήθηκε η προσέγγιση K-fold cross-validation με $K=5$. Οι δύο πρώτοι folds εκπαιδεύτηκαν για 8 εποχές έκαστος ενώ για κάθε epoch καταγράφηκε η μέση τιμή απώλειας (loss) για τα υποδίκτυα generator και discriminator.

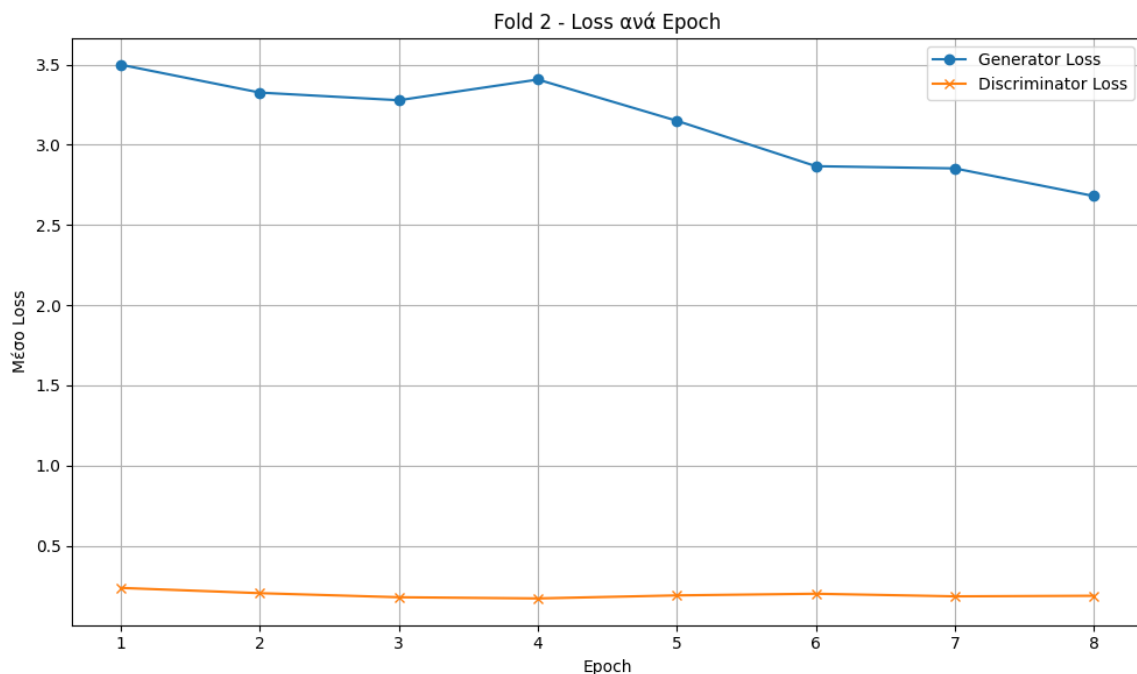
4.3 Loss Curves

Τα παρακάτω διαγράμματα (Εικόνα 3 και 4) απεικονίζουν την πορεία των τιμών loss για τα δύο πρώτα folds.



Εικόνα 3: Καμπύλη απώλειας Generator και Discriminator για Fold 1

Η Εικόνα 3 δείχνει την εξέλιξη της απώλειας κατά τη διάρκεια των 8 εποχών στο fold 1. Παρατηρείται μια μικρή αλλά σταθερή μείωση της Discriminator loss, η οποία κυμαίνεται από 0.26 στην πρώτη εποχή σε περίπου 0.16 στην όγδοη. Η Generator loss παρουσιάζει ελαφρά διακύμανση, με τιμές γύρω στο 3.2 - 3.4 χωρίς αξιοσημείωτη καθοδική τάση, γεγονός που υποδεικνύει σταθερή γεννήτρια με ήπια προσαρμογή.



Εικόνα 4: Καμπύλη απώλειας Generator και Discriminator για Fold 2

Στο δεύτερο fold (Εικόνα 4), παρατηρείται πιο καθοδική τάση στο Generator loss, το οποίο ξεκινά από 3.5 και μειώνεται σταδιακά κάτω από το 2.7 στην όγδοη εποχή. Η Discriminator loss μειώνεται επίσης ήπια (από ~0.23 σε ~0.18), με μικρή διακύμανση στα μέσα epochs. Τα

αποτελέσματα δείχνουν βελτιωμένη απόδοση στο δεύτερο fold, γεγονός που ενδέχεται να σχετίζεται με πιο αντιπροσωπευτικό διαχωρισμό δεδομένων.

4.4 Ποσοτική Αξιολόγηση Εξόδου Μοντέλου

Για την ποσοτική αποτίμηση της ποιότητας των παραγόμενων εικόνων, εφαρμόστηκαν δύο ευρέως χρησιμοποιούμενες μετρικές: Peak Signal-to-Noise Ratio (PSNR) και Structural Similarity Index (SSIM). Οι μετρήσεις πραγματοποιήθηκαν συγκρίνοντας τις εξόδους του μοντέλου από δύο χρονικά σημεία εκπαίδευσης (Epoch 1 και Epoch 8) με τις αντίστοιχες αρχικές εικόνες εισόδου από το σύνολο δοκιμής. Η μετρική PSNR αξιολογεί την ακρίβεια της παραγόμενης εικόνας σε σχέση με το ground truth σε επίπεδο pixel, με υψηλότερες τιμές να υποδηλώνουν μικρότερη παραμόρφωση. Η μετρική SSIM, από την άλλη, λαμβάνει υπόψη τη φωτεινότητα, την αντίθεση και τις δομικές ιδιότητες, παρέχοντας πιο ανθρώπινα αντιληπτή αξιολόγηση ομοιότητας.

Ο Πίνακας 1 συνοψίζει τις τιμές PSNR και SSIM για πέντε δείγματα εικόνων. Παρατηρείται ότι στην πλειονότητα των περιπτώσεων, η απόδοση του μοντέλου βελτιώνεται σημαντικά μεταξύ Epoch 1 και Epoch 8, ειδικά όσον αφορά τη δομική ομοιότητα:










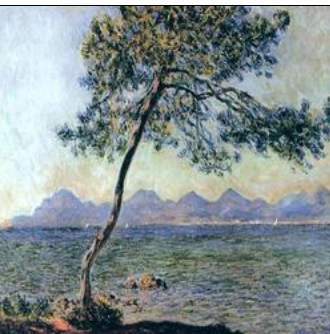





Image	PSNR_1	SSIM_1	PSNR_8	SSIM_8
Img1	11.41	0.361	16.18	0.537
Img2	12.82	0.415	11.37	0.551
Img3	13.27	0.426	14.04	0.633
Img4	13.53	0.361	15.57	0.556
Img5	15.54	0.325	14.75	0.443

Πίνακας 1: Μετρικές PSNR και SSIM για τις πέντε εικόνες δοκιμής

Στις εικόνες 1, 3 και 4, οι βελτιώσεις είναι εντονότερες και στις δύο μετρικές, γεγονός που καταδεικνύει ότι όσο προχωρά η εκπαίδευση, το μοντέλο μαθαίνει να αποδίδει καλύτερα το εικαστικό στυλ χωρίς να καταστρέφει τη δομή της αρχικής εικόνας. Η εικόνα 3 εμφανίζει την πιο σταθερή και καθαρή βελτίωση, με άνοδο της SSIM από 0.426 σε 0.633 και PSNR από 13.27 σε 14.04. Αντίθετα, στην εικόνα 2 παρατηρείται πτώση της τιμής PSNR, γεγονός που υποδηλώνει μικρή απώλεια ακρίβειας σε επίπεδο pixel, χωρίς ωστόσο να υποβαθμίζεται η δομή, αφού η SSIM αυξάνεται. Αυτό ενδέχεται να σχετίζεται με πιο επιθετική μεταφορά στυλ σε συγκεκριμένα χαρακτηριστικά της εικόνας. Συνολικά, τα αποτελέσματα ενισχύουν την υπόθεση ότι το μοντέλο βελτιώνει τη γενίκευσή του με την πρόοδο των εποχών, ιδίως ως προς τη διατήρηση της μορφολογίας και των οπτικών στοιχείων.

4.5 Ποιοτική Ανάλυση Αποτελεσμάτων

Για την αξιολόγηση της οπτικής ποιότητας των αποτελεσμάτων του μοντέλου, πραγματοποιήθηκε συγκριτική ανάλυση των παραγόμενων εικόνων στο Epoch 1 και το Epoch 8. Επιλέχθηκαν πέντε εικόνες από το test set του domain A (ζωγραφικοί πίνακες του Monet), οι οποίες μετασχηματίστηκαν μέσω του εκπαιδευμένου generator G_AB σε φωτογραφικού τύπου εικόνες του domain B. Η ανάλυση επικεντρώνεται στη μορφολογική ακρίβεια, στην αναπαραγωγή υφών, καθώς και στη φωτεινότητα και χρωματική συνέπεια των παραγόμενων εικόνων.

Image	Είσοδος	Έξοδος Epoch 1	Έξοδος Epoch 8
Img1			
Img2			
Img3			
Img4			
Img5			

Πίνακας 2: Απεικόνιση σύγκρισης εξόδων μεταξύ Epoch 1 και Epoch 8 για τις πέντε εικόνες δοκιμής

Στην πρώτη εικόνα, που απεικονίζει παραθαλάσσιο τοπίο με λόφους, το αποτέλεσμα του Epoch 1 εμφανίζει μειωμένη ευκρίνεια, θόλωση και αφύσικη αποτύπωση του ουρανού. Στο Epoch 8, η σκηνή αποδίδεται με αυξημένη ευκρίνεια και ρεαλισμό ενώ παρατηρείται καλύτερη μετάβαση χρωμάτων και βελτιωμένη φωτεινότητα. Η δεύτερη εικόνα, με θέμα ένα ορεινό τοπίο δίπλα στη θάλασσα, στο Epoch 1 παρουσιάζει έντονα παραμορφωμένα χαρακτηριστικά και μειωμένη χωρική συνέπεια. Αντίθετα, στο Epoch 8, η απόδοση των σχημάτων και των φωτεινών περιοχών είναι αισθητά βελτιωμένη, με ρεαλιστικότερη αναπαράσταση των μορφολογικών λεπτομερειών. Στην τρίτη εικόνα, που απεικονίζει ένα λιμάνι με βάρκες και δέντρα, το Epoch 1 αποτυγχάνει να αναπαράγει πειστικά τις αντανάκλασεις και τις υφές του νερού, με αποτέλεσμα η εικόνα να φαίνεται τεχνητή και ασυνεπής χρωματικά. Το αποτέλεσμα του Epoch 8 δείχνει μεγαλύτερη συνοχή και φυσικότητα, με σαφέστερες δομές και πειστικότερη σκίαση. Η τέταρτη εικόνα αφορά ένα μοναχικό δέντρο μπροστά από θάλασσα και βουνά. Στο Epoch 1, η παραμόρφωση του κορμού, η ασάφεια και ο υπερκορεσμός χρωμάτων οδηγούν σε χαμηλής ποιότητας αποτέλεσμα. Το Epoch 8 αποδίδει την εικόνα με μεγαλύτερη πιστότητα: η δομή του δέντρου διατηρείται, η θάλασσα παρουσιάζει υφή και οι φωτεινές πηγές αποδίδονται ομαλά. Τέλος, στην πέμπτη εικόνα, όπου απεικονίζεται ένα δέντρο ανάμεσα σε σπίτια, το Epoch 1 δημιουργεί αφύσικα περιγράμματα και υπερβολική φωτεινότητα σε μη ρεαλιστικά σημεία και η γενικότερη σκηνή είναι δυσανάγνωστη. Το αποτέλεσμα του Epoch 8 είναι αισθητά βελτιωμένο, με ευδιάκριτες δομές, καλύτερη διαχείριση φωτός και φυσικότερη απόδοση σκιών και υφής.

Η ποιοτική αξιολόγηση καταδεικνύει πως η αύξηση των εποχών εκπαίδευσης βελτιώνει σημαντικά την ικανότητα του μοντέλου να αναπαράγει φωτορεαλιστικά στοιχεία, μειώνοντας τα αρχικά visual artifacts και αποδίδοντας τις σκηνές με υψηλότερη πιστότητα ως προς τη μορφή και το ύφος του φωτογραφικού domain.

4.6 Συγκριτική Αξιολόγηση Εκδόσεων του CycleGAN

Για να αξιολογηθεί περαιτέρω η αποτελεσματικότητα της προτεινόμενης προσέγγισης, πραγματοποιήθηκε σύγκριση μεταξύ δύο εκδόσεων του CycleGAN μοντέλου: μίας πρώιμης μορφής (Model A), η οποία εκπαιδεύτηκε μόνο για 1 epoch, και της πλήρως εκπαιδευμένης έκδοσης (Model B) μετά από 8 epochs εκπαίδευσης. Αν και πρόκειται για το ίδιο μοντέλο, η αξιολόγηση αυτών των δύο σταδίων παρέχει χρήσιμα συμπεράσματα ως προς την επίδραση του αριθμού εποχών στην ποιότητα των παραγόμενων εικόνων. Η σύγκριση πραγματοποιήθηκε με βάση τις μετρικές PSNR και SSIM για πέντε ενδεικτικές εικόνες από το σύνολο δοκιμής.

Μοντέλο	Μέση τιμή PSNR	Μέση τιμή SSIM
Model A: epoch 1	13.31 dB	0.377
Model B: epoch 8	14.78 dB	0.544

Πίνακας 3: Μέσες τιμές PSNR και SSIM ανά epoch (1 και 8) για αξιολόγηση μοντέλων

Όπως παρατηρείται, η μέση τιμή του PSNR αυξήθηκε κατά +1.47 dB ενώ η SSIM παρουσίασε ακόμη πιο σημαντική βελτίωση (+0.167 μονάδες), γεγονός που υποδηλώνει αισθητά καλύτερη δομική και οπτική ομοιότητα των εξόδων του μοντέλου B σε σχέση με τις αντίστοιχες εικόνες εισόδου. Η ποιοτική αξιολόγηση επιβεβαίωσε τις παραπάνω παρατηρήσεις, με τις εξόδους του πλήρως εκπαιδευμένου μοντέλου να εμφανίζουν μεγαλύτερη καθαρότητα, φυσικότερες υφές και πιο συνεπές στυλ Monet. Αντίθετα, οι έξοδοι του Model A παρουσίαζαν υπερβολική παραμόρφωση και απώλεια σημαντικών χαρακτηριστικών σκηνής. Η σύγκριση αυτή τεκμηριώνει τη σημασία της επαρκούς εκπαίδευσης του μοντέλου και καταδεικνύει ότι ακόμη και απλές παραλλαγές του ίδιου αλγορίθμου μπορούν να οδηγήσουν σε σημαντικές διαφορές απόδοσης.

4.7 Στατιστική Ανάλυση Αποτελεσμάτων

Προκειμένου να αξιολογηθεί αν η βελτίωση των ποσοτικών μετρικών μεταξύ του CycleGAN στο Epoch 1 και στο Epoch 8 είναι στατιστικά σημαντική, εφαρμόστηκε ο paired t-test για τις μετρικές PSNR και SSIM, χρησιμοποιώντας τις τιμές που υπολογίστηκαν για τις πέντε ενδεικτικές εικόνες του συνόλου δοκιμής.

Μετρική	Epoch 1	Epoch 8
PSNR	13.31 ± 1.33 dB	14.38 ± 1.67 dB
SSIM	0.378 ± 0.038	0.544 ± 0.061

Πίνακας 4: Μέσες τιμές και τυπικές αποκλίσεις PSNR και SSIM για Epoch 1 και Epoch 8

Ο στατιστικός έλεγχος επέστρεψε τα εξής αποτελέσματα:

- Για τη μετρική PSNR: $t = 0.964$, $p = 0.3895$
- Για τη μετρική SSIM: $t = 9.750$, $p = 0.0006$

Με βάση τα παραπάνω η διαφορά στην PSNR μεταξύ των δύο μοντέλων δεν είναι στατιστικά σημαντική ($p > 0.05$), γεγονός που υποδηλώνει ότι η συνολική ποιότητα σήματος σε επίπεδο pixel ενδέχεται να παρουσιάζει μικρές μόνο βελτιώσεις, πιθανώς λόγω του χαρακτήρα της style transfer διαδικασίας, η οποία δεν στοχεύει άμεσα στη μέγιστη πιστότητα pixel προς pixel. Αντίθετα, η βελτίωση στη SSIM είναι στατιστικά σημαντική ($p < 0.001$), γεγονός που καταδεικνύει σαφή βελτίωση στην δομική ομοιότητα και στην οπτική συνοχή των παραγόμενων εικόνων. Το αποτέλεσμα αυτό επιβεβαιώνει ότι το μοντέλο μαθαίνει αποτελεσματικά να διατηρεί τη μορφή και τα χαρακτηριστικά της σκηνής καθώς αυξάνεται ο αριθμός εποχών εκπαίδευσης.

Επομένως, παρότι η βελτίωση στην PSNR είναι περιορισμένη, η στατιστικά τεκμηριωμένη αύξηση της SSIM υποδηλώνει ότι η μεταφορά στυλ γίνεται με πιο συνεπή και δομικά αξιόπιστο τρόπο, ενισχύοντας τη χρησιμότητα της μεθόδου για εφαρμογές όπου η αντιληπτική ποιότητα υπερισχύει της ακρίβειας ανά pixel.

5. Συμπεράσματα

Η παρούσα εργασία εστίασε στη μελέτη και εφαρμογή των Generative Adversarial Networks (GANs) για την επίλυση του προβλήματος μεταφοράς καλλιτεχνικού στυλ μεταξύ δύο διαφορετικών domains εικόνων: φωτογραφιών και πινάκων ζωγραφικής του Claude Monet. Η υλοποίηση βασίστηκε στο μοντέλο CycleGAN, το οποίο επιτρέπει τη μάθηση χαρτογραφήσεων ανάμεσα σε μη αντιστοιχισμένα σύνολα δεδομένων, αξιοποιώντας την αρχή της κυκλικής συνέπειας.

Η συνολική προσέγγιση ακολούθησε βέλτιστες πρακτικές πειραματισμού: πραγματοποιήθηκε προεπεξεργασία και υποδειγματοληψία του dataset Monet2Photo, εφαρμόστηκε k-fold cross-validation, καταγράφηκαν καμπύλες απωλειών (loss curves) και υλοποιήθηκαν ποσοτικές μετρήσεις ποιότητας με βάση τις μετρικές PSNR και SSIM. Η δοκιμαστική σύγκριση μεταξύ πρώιμου και ώριμου μοντέλου CycleGAN (Epoch 1 vs Epoch8) ανέδειξε τη σημαντική επίδραση της εκπαίδευσης στην ποιότητα των εξόδων.

Η ανάλυση των πειραματικών αποτελεσμάτων έδειξε ότι οι εικόνες εξόδου από το πλήρως εκπαιδευμένο μοντέλο παρουσίασαν βελτιωμένη οπτική συνέπεια, ρεαλιστικότερες υφές, και μεγαλύτερη δομική ομοιότητα σε σχέση με την είσοδο, όπως αποτυπώθηκε στις αυξημένες τιμές SSIM και στις ποιοτικά ανώτερες εικόνες. Παράλληλα, η αξιολόγηση μέσω στατιστικού ελέγχου (paired t-test) επιβεβαίωσε ότι η βελτίωση στη SSIM ήταν στατιστικά σημαντική ($p = 0.0006$),

ενισχύοντας την αξιοπιστία των ευρημάτων. Αν και η διαφορά στη μετρική PSNR μεταξύ των δύο εκδόσεων του μοντέλου δεν κρίθηκε στατιστικά σημαντική ($p = 0.3895$), αυτό ερμηνεύεται εύλογα από τον χαρακτήρα της μεταφοράς στυλ, η οποία δεν στοχεύει στην πιστή αναπαραγωγή pixel αλλά στη διατήρηση της αισθητικής και της δομής της σκηνής.

Συνεπώς, η εφαρμογή του CycleGAN αποδείχθηκε ιδιαίτερα αποτελεσματική για το συγκεκριμένο πρόβλημα, καθιστώντας το μοντέλο κατάλληλο για εφαρμογές όπως καλλιτεχνική επεξεργασία, ψηφιακή αναπαλαίωση έργων τέχνης, και δημιουργία περιεχομένου βασισμένου σε ύφος.

Για περαιτέρω βελτίωση των αποτελεσμάτων, μελλοντική εργασία μπορεί να εστιάσει στη χρήση περισσότερων εποχών εκπαίδευσης, την ενσωμάτωση perceptual loss μέσω προεκπαιδευμένων μοντέλων (π.χ. VGG), ή τη δοκιμή εναλλακτικών αρχιτεκτονικών όπως CUT ή StyleGAN2. Παράλληλα, η προσθήκη attention μηχανισμών, πλουσιότερου augmentation ή και χρήση μετα-μάθησης θα μπορούσαν να ενισχύσουν περαιτέρω την ικανότητα γενίκευσης του μοντέλου. Τέλος, θα ήταν ενδιαφέρον να υλοποιηθεί υποκειμενική αξιολόγηση (user study) των αποτελεσμάτων, ώστε να συμπληρωθεί η ποσοτική ανάλυση με πραγματικές αντιλήψεις ανθρώπινων χρηστών για την ποιότητα της μεταφοράς στυλ.

Βιβλιογραφία

- [1] B. Jamalpur et al., "Generative Adversarial Networks (GANs) for Artistic Style Transfer in Images," 2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC), Bengaluru, India, 2024, pp. 1–6, doi: 10.1109/ICDSCNC62492.2024.10941273.
- [2] B. Adhikari et al., "iHQGAN: A Lightweight Invertible Hybrid Quantum-Classical Generative Adversarial Network for Unsupervised Image-to-Image Translation," IEEE Transactions on Quantum Engineering, 2024.
- [3] S. Park et al., "KC-UNIT: Multi-kernel Conversion using Unpaired Image-to-Image Translation with Perceptual Guidance in Chest CT Imaging," Computerized Medical Imaging and Graphics, vol. 105, 2023.
- [4] M. Karim et al., "Paired and Unpaired Image to Image Translation using Generative Adversarial Networks," arXiv preprint arXiv:2505.16310, 2025.
- [5] Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In ICCV 2017, doi: 10.1109/ICCV.2017.244