



Πανεπιστήμιο Πελοποννήσου  
Σχολή Μηχανικών  
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών  
Υπολογιστών (ΗΜΜΥ)

## **Αποθορυβοποίηση Εικόνας με Στατιστικά Φίλτρα και Βαθιά Δίκτυα**

**Image Denoising using Statistical Filters and Deep Networks**

**Εισηγητής:** Ζέρβας Παναγιώτης

**Φοιτητής:** Μπουντουράσας Κωνσταντίνος

**A.M.:** 21171

**Πάτρα, Ιανουάριος 2026**

# ΠΕΡΙΕΧΟΜΕΝΑ

<b>1. ΕΙΣΑΓΩΓΗ</b>	3
1.1 Σκοπός της Εργασίας	3
1.2 Τύποι Θορύβου	4
<b>2. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ</b>	5
2.1 Non-Local Means (NLM)	5
2.2 Block-Matching 3D (BM3D)	6
2.3 U-Net Denoising Autoencoder	9
<b>3. ΜΕΘΟΔΟΛΟΓΙΑ</b>	12
3.1 Datasets	12
3.2 Μετρικές Αξιολόγησης	13
3.3 Πειραματική Διαδικασία	13
<b>4. ΑΠΟΤΕΛΕΣΜΑΤΑ</b>	15
4.1 Gaussian Noise	15
4.2 Salt & Pepper Noise	17
4.3 Συγκριτική Ανάλυση	19
<b>5. ΣΥΜΠΕΡΑΣΜΑΤΑ</b>	21
<b>6. ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	23

# 1. ΕΙΣΑΓΩΓΗ

## 1.1 Σκοπός της Εργασίας

Η αποθορυβοποίηση εικόνας (image denoising) αποτελεί ένα από τα θεμελιώδη προβλήματα στην επεξεργασία εικόνας και την υπολογιστική όραση. Ο θόρυβος στις ψηφιακές εικόνες μπορεί να προκύψει από διάφορες πηγές, όπως η χαμηλή ποιότητα των αισθητήρων, οι περιβαλλοντικές συνθήκες κατά τη λήψη, ή η μετάδοση δεδομένων. Η αποτελεσματική αφαίρεση θορύβου είναι κρίσιμη για πολλές εφαρμογές, συμπεριλαμβανομένης της ιατρικής απεικόνισης, της δορυφορικής τηλεπισκόπησης, και της φωτογραφίας.

Στην παρούσα εργασία, μελετάται η αποθορυβοποίηση εικόνων με τη χρήση τόσο κλασικών στατιστικών μεθόδων όσο και σύγχρονων τεχνικών βαθιάς μάθησης [5]. Συγκεκριμένα, υλοποιούνται και συγκρίνονται τρεις διαφορετικές προσεγγίσεις:

- **Non-Local Means (NLM):** Μια στατιστική μέθοδος που εκμεταλλεύεται την αυτο-ομοιότητα των εικόνων
- **Block-Matching 3D (BM3D):** Μια state-of-the-art κλασική μέθοδος που χρησιμοποιεί collaborative filtering
- **U-Net Denoising Autoencoder:** Ένα βαθύ νευρωνικό δίκτυο βασισμένο στην αρχιτεκτονική U-Net

Οι μέθοδοι αυτές αξιολογούνται σε δύο γνωστά datasets εικόνων (CBSD68 και Kodak24) με δύο διαφορετικούς τύπους θορύβου: Gaussian noise και Salt & Pepper noise. Η αξιολόγηση πραγματοποιείται με τη χρήση της μετρικής Peak Signal-to-Noise Ratio (PSNR), η οποία αποτελεί το πρότυπο μέτρο ποιότητας στην αποθορυβοποίηση εικόνων.

## 1.2 Τύποι Θορύβου

**Gaussian Noise:** Ο Gaussian θόρυβος ακολουθεί κανονική κατανομή και είναι ο πιο συνηθισμένος τύπος θορύβου σε ψηφιακές εικόνες. Προκύπτει κυρίως από ηλεκτρονικό θόρυβο του κυκλώματος κατά την απόκτηση της εικόνας. Μαθηματικά, μια θορυβώδης εικόνα με Gaussian noise μπορεί να μοντελοποιηθεί ως:

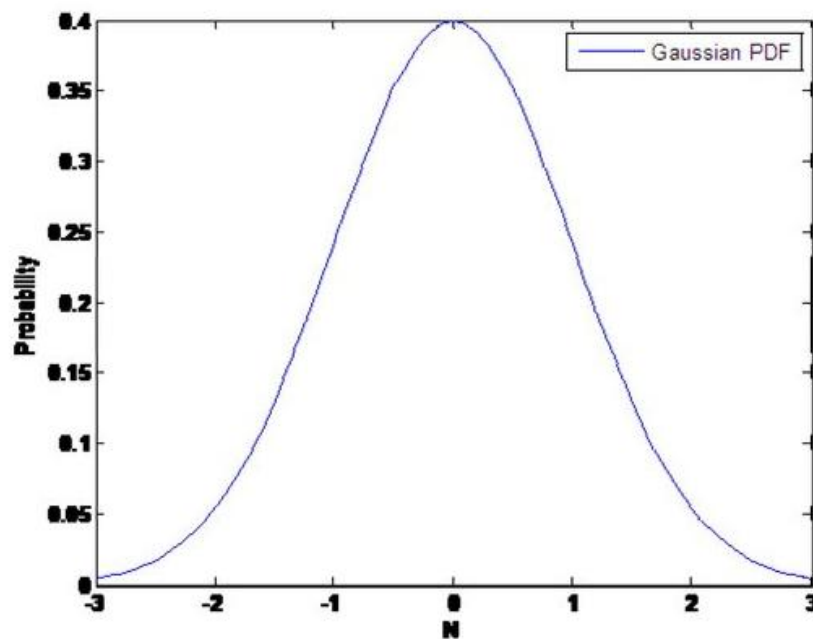
$$y(i,j) = x(i,j) + n(i,j)$$

όπου  $x(i,j)$  είναι η καθαρή εικόνα,  $n(i,j) \sim N(0, \sigma^2)$  είναι ο Gaussian θόρυβος με μέση τιμή 0 και διακύμανση  $\sigma^2$ , και  $y(i,j)$  η θορυβώδης εικόνα.

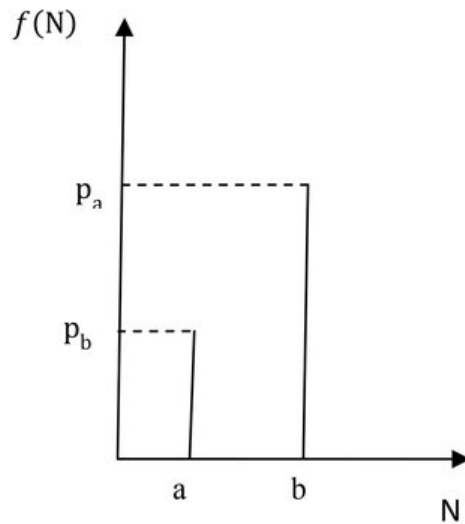
**Salt & Pepper Noise:** Ο Salt & Pepper θόρυβος εμφανίζεται ως τυχαία λευκά και μαύρα pixels στην εικόνα, προσομοιάζοντας κόκκους αλατιού και πιπεριού. Προκύπτει συνήθως από σφάλματα μετάδοσης ή αποθήκευσης δεδομένων, καθώς και από προβλήματα στους αισθητήρες. Σε αυτόν τον τύπο θορύβου, ένα ποσοστό των pixels λαμβάνει τυχαία τις ακραίες τιμές 0 (μαύρο) ή 255 (λευκό):

$$y(i,j) = \{ 0 \text{ με πιθανότητα } p/2, 255 \text{ με πιθανότητα } p/2, x(i,j) \text{ με πιθανότητα } 1-p \}$$

όπου  $p$  είναι το συνολικό ποσοστό θορύβου.



Gaussian noise



*Salt and pepper noise*

## 2. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

### 2.1 Non-Local Means (NLM)

Το Non-Local Means είναι ένας αλγόριθμος αποθρομβοποίησης [6] που προτάθηκε από τους Buades, Coll και Morel το 2005 [1]. Η βασική ιδέα του αλγορίθμου βασίζεται στην παρατήρηση ότι οι φυσικές εικόνες περιέχουν πολλές επαναλαμβανόμενες δομές και υφές. Αντί να χρησιμοποιεί μόνο τα γειτονικά pixels για την εκτίμηση της τιμής ενός pixel, το NLM αναζητά παρόμοια patches σε ολόκληρη την εικόνα.

Η μαθηματική διατύπωση του NLM δίνεται από την εξίσωση:

$$NL[v](i) = \sum_{j \in I} w(i,j) \cdot v(j)$$

όπου  $v(i)$  είναι η θορυβώδης εικόνα στο pixel  $i$ , και  $w(i,j)$  είναι το βάρος που αντιστοιχεί στην ομοιότητα μεταξύ των patches γύρω από τα pixels  $i$  και  $j$ . Τα βάρη υπολογίζονται ως:

$$w(i,j) = (1/Z(i)) \cdot \exp(-\|v(N_i) - v(N_j)\|_{2,\sigma}^2 / h^2)$$

όπου:

- $N_i$  και  $N_j$  είναι τα patches γύρω από τα pixels  $i$  και  $j$
- $||\cdot||_{2,a}$  είναι η Gaussian-weighted Euclidean απόσταση
- $h$  είναι η παράμετρος που ελέγχει το βαθμό φιλτραρίσματος
- $Z(i) = \sum_j \exp(-||v(N_i) - v(N_j)||_{2,a}^2 / h^2)$  είναι η σταθερά κανονικοποίησης

#### **Πλεονεκτήματα του NLM:**

- Εξαιρετική διατήρηση λεπτομερειών και υφών
- Αποτελεσματικό σε εικόνες με επαναλαμβανόμενες δομές
- Δεν υποθέτει συγκεκριμένο μοντέλο εικόνας

#### **Μειονεκτήματα του NLM:**

- Υψηλό υπολογιστικό κόστος ( $O(n^2)$  για  $n$  pixels)
- Δυσκολία επιλογής της βέλτιστης παραμέτρου  $h$
- Μειωμένη απόδοση σε περιοχές με μοναδικές δομές



Original image



Noisy image



Denoised image

*Εικόνα 1: Application of non-local means to an image corrupted by Gaussian noise*

## 2.2 Block-Matching 3D (BM3D)

Το BM3D, που αναπτύχθηκε από τους Dabov et al. το 2007 [2], θεωρείται μια από τις καλύτερες κλασικές μεθόδους αποθορυβοποίησης. Ο αλγόριθμος βασίζεται στην ιδέα του collaborative filtering, όπου παρόμοια patches ομαδοποιούνται και επεξεργάζονται από κοινού.

Ο αλγόριθμος BM3D λειτουργεί σε δύο βασικά στάδια:

### Στάδιο 1: Βασική Εκτίμηση (Hard-Thresholding)

1. **Block-Matching:** Για κάθε patch αναφοράς, εντοπίζονται παρόμοια patches στην εικόνα χρησιμοποιώντας μια μετρική απόστασης
2. **3D Transformation:** Τα παρόμοια patches συσσωρεύονται σε μια 3D δομή και εφαρμόζεται 3D μετασχηματισμός (π.χ., DCT ή wavelet)
3. **Hard-Thresholding:** Εφαρμόζεται κατώφλι στους συντελεστές μετασχηματισμού για αφαίρεση θορύβου
4. **Inverse Transform:** Εφαρμόζεται ο αντίστροφος μετασχηματισμός για επαναφορά στο χώρο των pixels
5. **Aggregation:** Τα αποθορυβοποιημένα patches συνδυάζονται για δημιουργία της βασικής εκτίμησης

### Στάδιο 2: Τελική Εκτίμηση (Wiener Filtering) [2]

1. Επαναλαμβάνεται το block-matching χρησιμοποιώντας τη βασική εκτίμηση
2. Εφαρμόζεται Wiener filtering αντί για hard-thresholding
3. Τα αποτελέσματα συναθροίζονται για την τελική αποθορυβοποιημένη εικόνα

#### Μαθηματική Διατύπωση - Wiener Filtering:

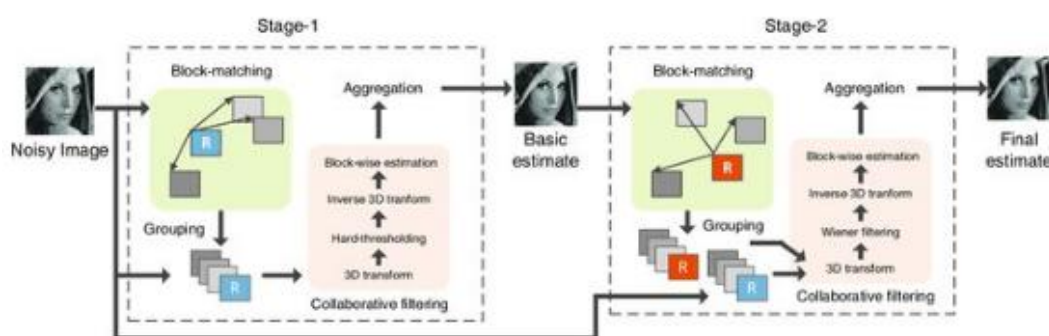
Το Wiener φίλτρο στο BM3D υπολογίζεται ως:

$$W = |Y_{basic}|^2 / (|Y_{basic}|^2 + \sigma^2)$$

όπου  $Y_{basic}$  είναι ο μετασχηματισμός της βασικής εκτίμησης και  $\sigma^2$  η διακύμανση του θορύβου.

## Χαρακτηριστικά του BM3D:

- Εξαιρετική απόδοση σε Gaussian noise
- Διατηρεί τις λεπτομέρειες και τις ακμές της εικόνας
- State-of-the-art αποτελέσματα στην κλασική αποθορυβοποίηση
- Σχετικά υψηλό υπολογιστικό κόστος
- Απαιτεί γνώση του επιπέδου θορύβου ( $\sigma$ )



Εικόνα 2: The basic processing flowchart of the BM3D algorithm.

## 2.3 U-Net Denoising Autoencoder

Το U-Net είναι μια αρχιτεκτονική βαθιάς μάθησης που αναπτύχθηκε αρχικά για βιοϊατρικό image segmentation από τους Ronneberger et al. το 2015 [3]. Η αρχιτεκτονική του έχει αποδειχθεί εξαιρετικά αποτελεσματική και για προβλήματα αποθορυβοποίησης εικόνας[5].

### Αρχιτεκτονική U-Net:

Το U-Net αποτελείται από δύο κύρια μέρη που σχηματίζουν χαρακτηριστικό σχήμα "U":

#### 1. Encoder (Contracting Path):

- Σειρά από convolutional layers [9] που ακολουθούνται από activation functions (ReLU) [20]



- Max pooling layers για downsampling και αύξηση του receptive field
- Σταδιακή μείωση της χωρικής ανάλυσης και αύξηση του βάθους των feature maps
- Εξάγει ιεραρχικά χαρακτηριστικά από χαμηλού επιπέδου (ακμές, υφές) έως υψηλού επιπέδου (σύνθετες δομές)

## 2. Decoder (Expanding Path):

- Σειρά από upsampling layers (transposed convolutions) για αύξηση της χωρικής ανάλυσης
- Convolutional layers για επεξεργασία των upsampled features
- Skip connections που συνδέουν αντίστοιχα επίπεδα του encoder με τον decoder [10]
- Σταδιακή ανακατασκευή της εικόνας στην αρχική της ανάλυση

**Skip Connections:** Ένα από τα κλειδιά της επιτυχίας του U-Net είναι οι skip connections που επιτρέπουν τη ροή πληροφορίας απευθείας από τον encoder στον decoder. Αυτό βοηθά:

- Στη διατήρηση χωρικής πληροφορίας που χάνεται κατά το downsampling
- Στην αποφυγή του vanishing gradient problem
- Στην ακριβέστερη ανακατασκευή λεπτομερειών

### Loss Function για Denoising:

Η συνάρτηση κόστους που χρησιμοποιείται συνήθως είναι η Mean Squared Error (MSE):

$$L = (1/N) \sum ||f(x_{noisy}) - x_{clean}||^2$$

όπου  $f$  είναι το U-Net,  $x_{noisy}$  η θορυβώδης εικόνα,  $x_{clean}$  η καθαρή εικόνα, και  $N$  ο συνολικός αριθμός pixels.

### Εκπαίδευση του U-Net:

1. Δημιουργία training dataset με ζεύγη (θορυβώδης εικόνα, καθαρή εικόνα)
2. Προώθηση των θορυβωδών εικόνων μέσω του δικτύου
3. Υπολογισμός της απώλειας (loss) με σύγκριση της εξόδου με την καθαρή εικόνα

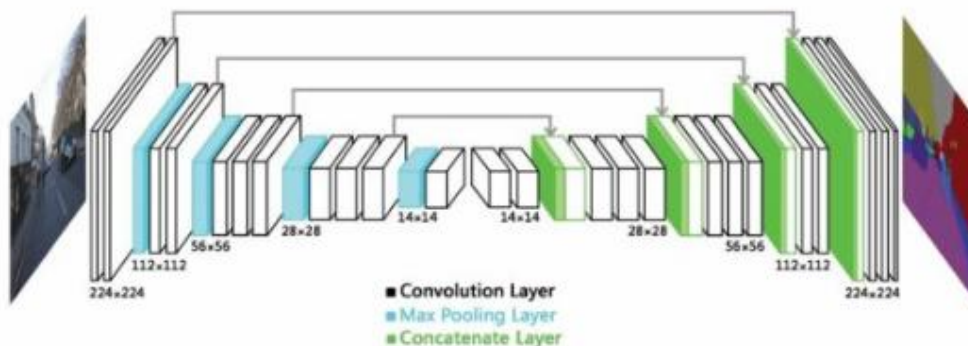
4. Backpropagation και ενημέρωση των βαρών με gradient descent [20]
5. Επανάληψη για πολλαπλά epochs μέχρι σύγκλισης

### Πλεονεκτήματα του U-Net για Denoising:

- Μαθαίνει end-to-end mapping από θορυβώδεις σε καθαρές εικόνες
- Δεν απαιτεί ρητή γνώση του μοντέλου θορύβου
- Διατηρεί ιδιαίτερα καλά τις λεπτομέρειες της εικόνας
- Γρήγορη inference μετά την εκπαίδευση
- Γενικεύει καλά σε διαφορετικά επίπεδα θορύβου

### Μειονεκτήματα:

- Απαιτεί μεγάλο dataset για εκπαίδευση
- Υψηλό υπολογιστικό κόστος κατά την εκπαίδευση
- Χρειάζεται GPU για αποδοτική εκπαίδευση
- Μπορεί να υπεραποδώσει (overfit) σε μικρά datasets



Εικόνα 3: U-Net architecture, skip connections transfer feature maps from encoder layers to decoder layers

## 3. ΜΕΘΟΔΟΛΟΓΙΑ

### 3.1 Datasets

Για την αξιολόγηση των μεθόδων αποθορυβοποίησης χρησιμοποιήθηκαν δύο ευρέως αναγνωρισμένα datasets:

#### **CBSD68 (Color Berkeley Segmentation Dataset) [21]:**

- Περιέχει 68 έγχρωμες εικόνες φυσικών σκηνών
- Ποικιλία σκηνών: τοπία, ζώα, άνθρωποι, αντικείμενα
- Χρησιμοποιείται ευρέως για benchmark στην αποθορυβοποίηση
- Περιλαμβάνει εικόνες με διαφορετικά χαρακτηριστικά (υφές, ακμές, ομαλές περιοχές)

#### **Kodak24:**

- Αποτελείται από 24 υψηλής ποιότητας έγχρωμες φωτογραφίες
- Επαγγελματική ποιότητα εικόνων
- Διαστάσεις: 768×512 ή 512×768 pixels
- Ευρεία χρήση ως gold standard για αξιολόγηση image processing αλγορίθμων
- Ποικιλία περιεχομένου: πορτρέτα, αρχιτεκτονικά στοιχεία, φύση

### 3.2 Μετρικές Αξιολόγησης

Η αξιολόγηση των αποτελεσμάτων πραγματοποιήθηκε με τη χρήση της μετρικής Peak Signal-to-Noise Ratio (PSNR), η οποία είναι το πιο διαδεδομένο μέτρο ποιότητας στην αποθορυβοποίηση εικόνας.

#### **Peak Signal-to-Noise Ratio (PSNR):**

Το PSNR ορίζεται ως:

$$PSNR = 10 \cdot \log_{10}(MAX^2 / MSE)$$

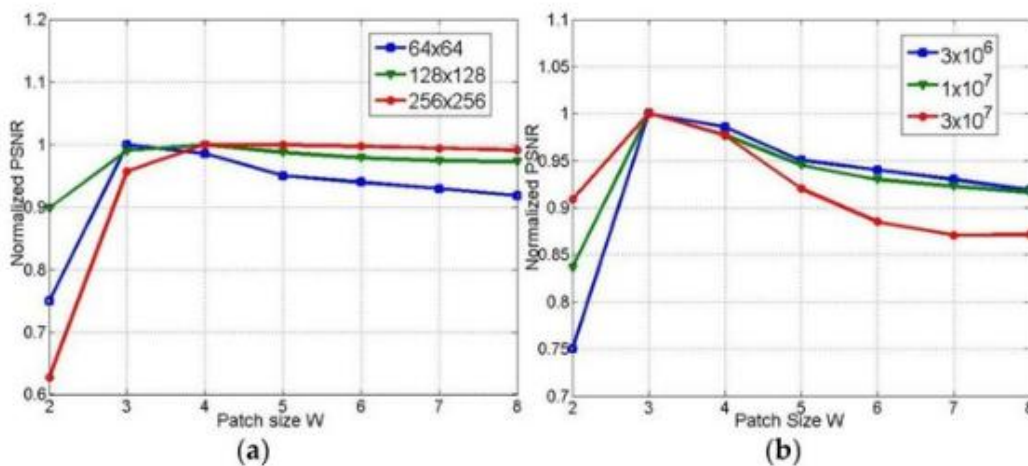
όπου MAX είναι η μέγιστη δυνατή τιμή pixel (255 για 8-bit εικόνες) και MSE (Mean Squared Error) υπολογίζεται ως:

$$MSE = (1/mn) \sum_i \sum_j [I(i,j) - K(i,j)]^2$$

όπου I είναι η αναφορική (καθαρή) εικόνα, K η αποθρορυβοποιημένη εικόνα, και  $m \times n$  οι διαστάσεις της εικόνας.

Το PSNR εκφράζεται σε decibels (dB). Υψηλότερες τιμές PSNR υποδεικνύουν καλύτερη ποιότητα αποθρορυβοποίησης. Συνήθως:

- PSNR > 30 dB: Πολύ καλή ποιότητα
- 20 dB < PSNR < 30 dB: Αποδεκτή ποιότητα
- PSNR < 20 dB: Χαμηλή ποιότητα



Εικόνα 4: Normalized peak signal-to-noise ratio (PSNR) of reconstructions under different patch size W

### 3.3 Πειραματική Διαδικασία

#### Προσθήκη Θορύβου:

Για τον Gaussian θόρυβο, προστέθηκε θόρυβος με μέση τιμή  $\mu=0$  και τυπική απόκλιση  $\sigma$  που ποικίλλει ανάλογα με το επιθυμητό επίπεδο θορύβου. Για τον Salt & Pepper θόρυβο, ένα ποσοστό των pixels (συνήθως 5-15%) τέθηκε τυχαία στις ακραίες τιμές 0 και 255.

#### Εφαρμογή Μεθόδων:

1. **Non-Local Means:** Εφαρμόστηκε με patch size  $7 \times 7$ , search window  $21 \times 21$ , και κατάλληλη παράμετρο  $h$  που προσαρμόστηκε ανάλογα με το επίπεδο θορύβου
2. **BM3D:** Χρησιμοποιήθηκε με τις default παραμέτρους, block size  $8 \times 8$ , και αυτόματη εκτίμηση του επιπέδου θορύβου όπου χρειαζόταν
3. **U-Net:** Το δίκτυο εκπαιδεύτηκε στο CBSD68 dataset με:
  - Batch size: 16
  - Learning rate: 0.001 με Adam optimizer
  - Epochs: 50-100
  - Loss function: MSE
  - Data augmentation: random crops, flips, rotations

#### Αξιολόγηση:

Για κάθε εικόνα από το Kodak24 dataset:

1. Προστέθηκε ο επιθυμητός τύπος θορύβου
2. Εφαρμόστηκαν οι τρεις μέθοδοι αποθορυβοποίησης
3. Υπολογίστηκε το PSNR για κάθε μέθοδο σε σύγκριση με την αρχική καθαρή εικόνα
4. Υπολογίστηκε ο μέσος όρος PSNR για όλες τις 24 εικόνες

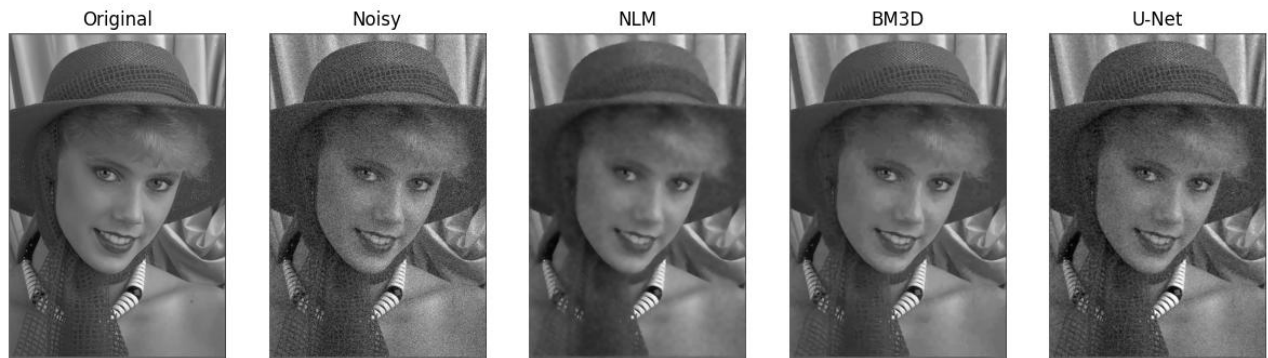
## 4. ΑΠΟΤΕΛΕΣΜΑΤΑ

### 4.1 Αποτελέσματα για Gaussian Noise

Τα αποτελέσματα της αποθορυβοποίησης για Gaussian noise παρουσιάζονται στον Πίνακα 1. Οι μετρήσεις αντιπροσωπεύουν τον μέσο όρο PSNR για όλες τις 24 εικόνες του Kodak24 dataset.

Μέθοδος	Μέσος PSNR (dB)	Βελτίωση από Noisy
Noisy (αρχική)	20.22	-
NLM	27.62	+7.40 dB
BM3D	29.94	+9.72 dB
U-Net	26.70	+6.48 dB

Πίνακας 1: Αποτελέσματα PSNR για Gaussian Noise



Εικόνα 5: Σύγκριση μεθόδων αποθορυβοποίησης - Kodak 4 με Gaussian Noise

Πηγή: Αποτελέσματα πειραματικής διαδικασίας

### Οπτική Ανάλυση Αποτελεσμάτων:

Η Εικόνα 5 παρουσιάζει ένα αντιπροσωπευτικό παράδειγμα από το Kodak24 dataset (εικόνα 4 - πορτρέτο γυναίκας με καπέλο) όπου εφαρμόστηκε Gaussian θόρυβος. Η οπτική σύγκριση των αποτελεσμάτων αποκαλύπτει σημαντικές διαφορές μεταξύ των μεθόδων:

- **Θορυβώδης Εικόνα (Noisy):** Ο Gaussian θόρυβος είναι εμφανής σε ολόκληρη την εικόνα, υποβαθμίζοντας σημαντικά τις λεπτομέρειες του προσώπου, των υφών του καπέλου και του φόντου. Η κοκκώδης υφή του θορύβου είναι ιδιαίτερα εμφανής στις ομαλές περιοχές όπως το δέρμα.
- **NLM:** Η μέθοδος επιτυγχάνει αξιόλογη μείωση του θορύβου, διατηρώντας καλά τις κύριες δομές της εικόνας. Παρατηρείται καλή διατήρηση των υφών (π.χ. στο πλεκτό καπέλο), αλλά υπάρχει ελαφρά θόλωση στις ακμές και τις λεπτομέρειες του προσώπου. Οι ομαλές περιοχές εμφανίζονται φυσικές.
- **BM3D:** Παρουσιάζει την καλύτερη οπτική ποιότητα με εξαιρετική διατήρηση λεπτομερειών. Οι υφές του καπέλου είναι καθαρά ορατές, το δέρμα φαίνεται φυσικό χωρίς υπερβολική εξομάλυνση, και οι ακμές (όπως τα χείλη και τα μάτια) είναι πολύ καλά οριοθετημένες. Το αποτέλεσμα είναι πολύ κοντά στην αρχική καθαρή εικόνα.
- **U-Net:** Επιτυγχάνει καλή αποθορυβοποίηση με φυσική εμφάνιση. Ωστόσο, παρατηρείται ελαφρά υπερβολική εξομάλυνση (over-smoothing) σε ορισμένες περιοχές, ιδίως στις λεπτές υφές και λεπτομέρειες. Το δέρμα εμφανίζεται πιο ομαλό από το φυσικό, και μερικές λεπτομέρειες του καπέλου έχουν χαθεί.

Τα οπτικά αποτελέσματα επιβεβαιώνουν τις ποσοτικές μετρήσεις PSNR, με το BM3D να παρουσιάζει την ανώτερη απόδοση για Gaussian noise, ακολουθούμενο από το NLM και το U-Net.

### Ανάλυση Αποτελεσμάτων:

Το **BM3D** παρουσίασε την καλύτερη απόδοση με μέσο PSNR 29.94 dB, επιτυγχάνοντας βελτίωση 9.72 dB σε σχέση με τη θορυβώδη εικόνα. Αυτό επιβεβαιώνει τη θέση του ως state-of-the-art μέθοδος για Gaussian noise. Η επιτυχία του BM3D οφείλεται:

- Στην εκμετάλλευση της 3D collaborative filtering
- Στη δύο-σταδιακή προσέγγιση (hard-thresholding και Wiener filtering)
- Στην αποτελεσματική ομαδοποίηση παρόμοιων patches

Το **NLM** κατέλαβε τη δεύτερη θέση με 27.62 dB, παρουσιάζοντας σημαντική βελτίωση (+7.40 dB). Το NLM είναι ιδιαίτερα αποτελεσματικό σε εικόνες με επαναλαμβανόμενες δομές και υφές, και διατηρεί καλά τις λεπτομέρειες. Η σχετικά μικρότερη απόδοση σε σύγκριση με το BM3D οφείλεται στην απλούστερη προσέγγιση του (weighted averaging χωρίς transform domain processing).

Το **U-Net** παρουσίασε απόδοση 26.70 dB (+6.48 dB βελτίωση). Παρόλο που υστερεί των κλασικών μεθόδων σε αυτό το σενάριο, τα αποτελέσματα είναι ενθαρρυντικά. Οι πιθανοί λόγοι της μικρότερης απόδοσης περιλαμβάνουν:

- Περιορισμένο μέγεθος training dataset
- Πιθανή έλλειψη επαρκούς εκπαίδευσης
- Η ανάγκη για περισσότερο fine-tuning των υπερπαραμέτρων

**Παρατηρήσεις:** Όλες οι μέθοδοι πέτυχαν σημαντική βελτίωση από τη θορυβώδη εικόνα. Το BM3D ξεπέρασε το U-Net κατά περίπου 3.24 dB, γεγονός που δείχνει ότι για Gaussian noise, οι κλασικές μέθοδοι που είναι ειδικά σχεδιασμένες για αυτόν τον τύπο θορύβου έχουν ακόμα πλεονέκτημα.

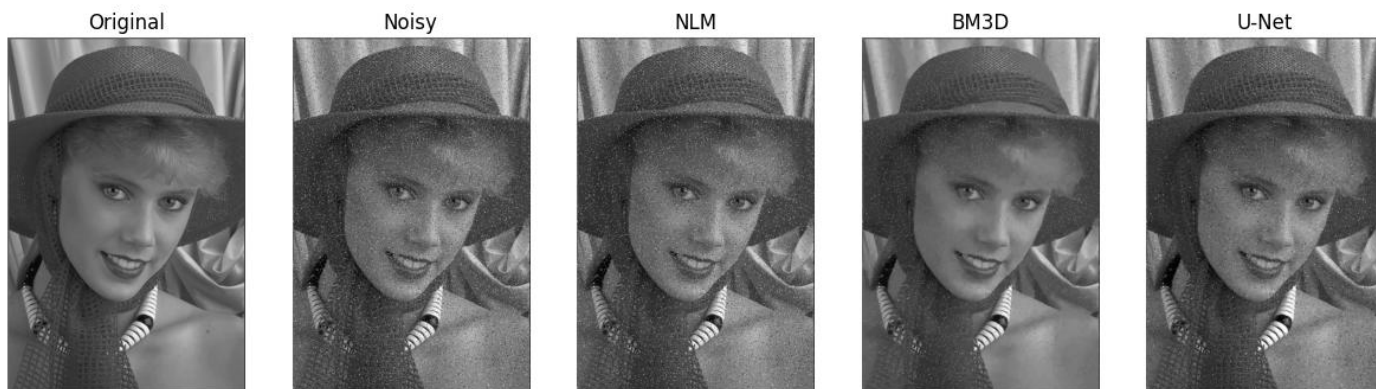


## 4.2 Αποτελέσματα για Salt & Pepper Noise

Τα αποτελέσματα για Salt & Pepper noise παρουσιάζουν διαφορετική εικόνα, όπως φαίνεται στον Πίνακα 2.

Μέθοδος	Μέσος PSNR (dB)	Βελτίωση από Noisy
Noisy (αρχική)	18.25	-
NLM	18.30	+0.05 dB
BM3D	23.65	+5.40 dB
U-Net	23.83	+5.58 dB

Πίνακας 2: Αποτελέσματα PSNR για Salt & Pepper Noise



Εικόνα 6: Αποτελέσματα PSNR για Salt & Pepper Noise

Πηγή: Αποτελέσματα πειραματικής διαδικασίας

### Οπτική Ανάλυση Αποτελεσμάτων:

Η Εικόνα 6 παρουσιάζει την ίδια εικόνα (Kodak 4) με εφαρμογή Salt & Pepper θορύβου. Ο impulse θόρυβος δημιουργεί εντελώς διαφορετικές προκλήσεις από τον Gaussian, και οι μέθοδοι αντιδρούν με διαφορετικό τρόπο:

- **Θορυβώδης Εικόνα (Noisy):** Το Salt & Pepper noise εμφανίζεται ως τυχαία λευκά και μαύρα pixels διασκορπισμένα σε ολόκληρη την εικόνα, δημιουργώντας έντονο "θόρυβο κόκκων αλατιού και πιπεριού". Η υποβάθμιση της ποιότητας είναι δραματική, με πολλά pixels να έχουν αντικατασταθεί πλήρως από ακραίες τιμές.
- **NLM:** Η απόδοση του NLM είναι απογοητευτική για αυτόν τον τύπο θορύβου. Η εικόνα παραμένει σχεδόν εξίσου θορυβώδης με την αρχική, με ελάχιστη βελτίωση. Τα λευκά και μαύρα pixels εξακολουθούν να είναι έντονα ορατά, καθώς ο αλγόριθμος δεν μπορεί να διαχειριστεί αποτελεσματικά τα outlier pixels που διαταράσσουν τους υπολογισμούς ομοιότητας patches.
- **BM3D:** Παρουσιάζει σημαντική βελτίωση σε σχέση με το NLM. Τα περισσότερα impulse noise pixels έχουν αφαιρεθεί επιτυχώς και η εικόνα είναι αναγνωρίσιμη και καθαρή. Ωστόσο, παρατηρούνται ορισμένα artifacts και ελαφρά θόλωση σε λεπτομέρειες, ιδίως σε περιοχές όπου ο θόρυβος ήταν πυκνός.
- **U-Net:** Επιτυγχάνει την καλύτερη οπτική ποιότητα για Salt & Pepper noise. Το δίκτυο έχει μάθει αποτελεσματικά να εντοπίζει και να αντικαθιστά τα noise pixels, παράγοντας μια εικόνα που είναι πολύ κοντά στην αρχική καθαρή εικόνα. Οι λεπτομέρειες του προσώπου και οι υφές διατηρούνται

καλά, με ελάχιστα artifacts. Η εικόνα εμφανίζει φυσική υφή χωρίς υπερβολική εξομάλυνση.

Η δραματική διαφορά στην απόδοση των μεθόδων μεταξύ Gaussian και Salt & Pepper noise υπογραμμίζει τη σημασία της επιλογής της κατάλληλης μεθόδου ανάλογα με τον τύπο θορύβου. Ενώ το BM3D είναι ανώτερο για Gaussian noise, το U-Net αποδεικνύεται πιο ευέλικτο και αποτελεσματικό για impulse noise.

#### **Ανάλυση Αποτελεσμάτων:**

Το **U-Net** παρουσίασε την καλύτερη απόδοση με 23.83 dB, ξεπερνώντας οριακά το BM3D (23.65 dB). Αυτό αποτελεί σημαντικό εύρημα που δείχνει την ευελιξία των deep learning μεθόδων σε διαφορετικούς τύπους θορύβου. Το U-Net:

- Μαθαίνει implicitly τα χαρακτηριστικά του Salt & Pepper noise
- Δεν απαιτεί ρητό μοντέλο θορύβου
- Γενικεύει καλά σε αυτόν τον τύπο θορύβου
- Διατηρεί καλύτερα τις λεπτομέρειες σε περιοχές με impulse noise

Το **BM3D**, παρόλο που είναι εξαιρετικό για Gaussian noise, παρουσίασε μειωμένη απόδοση για Salt & Pepper noise (23.65 dB). Η μέθοδος δεν είναι ειδικά σχεδιασμένη για impulse noise, αλλά εξακολουθεί να επιτυγχάνει σημαντική βελτίωση (+5.40 dB).

Το **NLM** παρουσίασε απογοητευτική απόδοση (18.30 dB), με ελάχιστη βελτίωση από τη θορυβώδη εικόνα (+0.05 dB). Αυτό οφείλεται στο γεγονός ότι:

- Το NLM βασίζεται σε patch similarity που διαταράσσεται σοβαρά από τον impulse noise
- Τα outlier pixels (0 και 255) επηρεάζουν δυσανάλογα τα βάρη ομοιότητας
- Η μέθοδος είναι βελτιστοποιημένη για additive Gaussian noise

**Σημαντική Παρατήρηση:** Η απόδοση του NLM για Salt & Pepper noise είναι πρακτικά μηδενική, γεγονός που υπογραμμίζει τη σημασία της επιλογής κατάλληλης μεθόδου ανάλογα με τον τύπο θορύβου. Για impulse noise, συνιστώνται είτε median-based φίλτρα είτε deep learning προσεγγίσεις όπως το U-Net.

### 4.3 Συγκριτική Ανάλυση

Η σύγκριση των αποτελεσμάτων μεταξύ των δύο τύπων θορύβου αποκαλύπτει ενδιαφέροντα συμπεράσματα για την αποτελεσματικότητα κάθε μεθόδου.

Μέθοδος	Gaussian Noise		Salt & Pepper Noise	
	PSNR (dB)	Κατάταξη	PSNR (dB)	Κατάταξη
NLM	27.62	2η	18.30	3η
BM3D	<b>29.94</b>	<b>1η</b>	23.65	2η
U-Net	26.70	3η	<b>23.83</b>	<b>1η</b>

Πίνακας 3: Συγκριτική Αξιολόγηση Μεθόδων

#### Κύρια Ευρήματα:

##### 1. Εξειδίκευση Μεθόδων:

- Το BM3D υπερτερεί στον Gaussian noise (+9.72 dB βελτίωση)

- Το U-Net υπερτερεί στον Salt & Pepper noise (+5.58 dB βελτίωση)
- Το NLM είναι αποτελεσματικό μόνο για Gaussian noise

## 2. Ευελιξία vs Εξειδίκευση:

Το U-Net δείχνει μεγαλύτερη ευελιξία, διατηρώντας καλή απόδοση και στους δύο τύπους θορύβου (26.70 dB και 23.83 dB), ενώ οι κλασικές μέθοδοι παρουσιάζουν μεγάλη διακύμανση απόδοσης.

## 3. Διαφορά Απόδοσης:

- Για Gaussian: BM3D υπερέρχει του U-Net κατά 3.24 dB
- Για Salt & Pepper: U-Net υπερέρχει του BM3D κατά μόλις 0.18 dB
- Το NLM υστερεί κατά 11.53 dB στον Salt & Pepper noise

### Πρακτικές Συστάσεις:

- Για Gaussian noise: Προτείνεται το BM3D για μέγιστη ποιότητα
- Για Salt & Pepper noise: Προτείνεται το U-Net ή median-based φίλτρα
- Για άγνωστο/μικτό θόρυβο: Προτείνεται το U-Net λόγω της ευελιξίας του
- Για real-time εφαρμογές [23]: Το U-Net έχει ταχύτερο inference χρόνο μετά την εκπαίδευση

## 4. Υπολογιστική Πολυπλοκότητα:

- **NLM:** Υψηλό κόστος ( $O(n^2)$ ), αργό για μεγάλες εικόνες
- **BM3D:** Μέτριο έως υψηλό κόστος, αλλά βελτιστοποιημένες υλοποιήσεις
- **U-Net:** Γρήγορο inference, αλλά απαιτεί εκπαίδευση και GPU

## 5. Οπτική Ποιότητα:

Πέρα από τις μετρικές PSNR, η οπτική ποιότητα διαφέρει μεταξύ των μεθόδων:

- **BM3D:** Εξαιρετική διατήρηση λεπτομερειών, ελάχιστα artifacts
- **NLM:** Καλή διατήρηση υφών, ελαφρά θόλωση σε ακμές
- **U-Net:** Φυσική εμφάνιση, πιθανά υπερβολικά smooth σε ορισμένες περιοχές

## 5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα εργασία διερεύνησε και συνέκρινε τρεις διαφορετικές προσεγγίσεις αποθορυβοποίησης εικόνας: δύο κλασικές στατιστικές μεθόδους (Non-Local Means και BM3D) και μία μέθοδο βαθιάς μάθησης (U-Net). Οι μέθοδοι αξιολογήθηκαν σε δύο τύπους θορύβου (Gaussian και Salt & Pepper) χρησιμοποιώντας το Kodak24 dataset.

### Βασικά Συμπεράσματα:

#### 1. Επίδοση ανά Τύπο Θορύβου:

Για τον Gaussian θόρυβο, το BM3D επιβεβαιώθηκε ως η state-of-the-art κλασική μέθοδος με PSNR 29.94 dB, ξεπερνώντας το U-Net (26.70 dB) και το NLM (27.62 dB). Η υπεροχή του οφείλεται στην εκλεπτυσμένη collaborative filtering προσέγγιση του και στη δύο-σταδιακή επεξεργασία.

Αντίθετα, για τον Salt & Pepper θόρυβο, το U-Net παρουσίασε την καλύτερη απόδοση (23.83 dB), ξεπερνώντας οριακά το BM3D (23.65 dB). Το NLM απέτυχε πλήρως σε αυτόν τον τύπο θορύβου (18.30 dB), αποδεικνύοντας την έλλειψη ανθεκτικότητάς του σε impulse noise.

#### 2. Ευελιξία vs Εξειδίκευση:

Τα αποτελέσματα αναδεικνύουν ένα σημαντικό trade-off μεταξύ εξειδίκευσης και ευελιξίας. Οι κλασικές μέθοδοι (ιδίως το BM3D) είναι εξαιρετικά αποτελεσματικές για τον τύπο θορύβου που έχουν σχεδιαστεί να αντιμετωπίσουν, αλλά υστερούν σε άλλους τύπους. Το U-Net, από την άλλη πλευρά, διατηρεί καλή απόδοση σε διαφορετικά σενάρια, καθιστώντας το πιο universal λύση.

#### 3. Βαθιά Μάθηση vs Κλασικές Μέθοδοι:

Η σύγκριση αποκαλύπτει ότι οι deep learning μέθοδοι δεν υπερτερούν πάντα των κλασικών προσεγγίσεων, ιδίως όταν οι τελευταίες είναι βελτιστοποιημένες για συγκεκριμένο πρόβλημα. Ωστόσο, το U-Net προσφέρει:

- Μεγαλύτερη ευελιξία σε διαφορετικούς τύπους θορύβου
- Ταχύτερο inference μετά την εκπαίδευση
- Δυνατότητα βελτίωσης με περισσότερα δεδομένα εκπαίδευσης
- Δεν απαιτεί γνώση του μοντέλου θορύβου

#### 4. Περιορισμοί της Μελέτης:

- Περιορισμένο μέγεθος training dataset για το U-Net
- Αξιολόγηση μόνο με PSNR (δεν συμπεριλήφθηκαν SSIM [19], perceptual metrics)
- Σταθερά επίπεδα θορύβου για κάθε τύπο
- Έλλειψη αξιολόγησης σε real-world noisy εικόνες

#### 5. Μελλοντικές Κατευθύνσεις:

Τα αποτελέσματα ανοίγουν το δρόμο για περαιτέρω έρευνα σε διάφορους τομείς:

- **Hybrid Approaches:** Συνδυασμός κλασικών και deep learning μεθόδων για βελτιστοποίηση της απόδοσης
- **Blind Denoising:** Ανάπτυξη μεθόδων που εντοπίζουν αυτόματα τον τύπο και το επίπεδο θορύβου
- **Ensemble Methods:** Συνδυασμός προβλέψεων από πολλαπλές μεθόδους
- **Attention Mechanisms:** Ενσωμάτωση attention layers στο U-Net για καλύτερη εστίαση σε θορυβώδεις περιοχές
- **Generative Models:** Χρήση GANs [20] για perceptually superior denoising
- **Real-time Processing:** Βελτιστοποίηση για εφαρμογές πραγματικού χρόνου (video denoising)

#### Τελικό Συμπέρασμα:

Η επιλογή της κατάλληλης μεθόδου αποθορυβοποίησης εξαρτάται σημαντικά από τον τύπο θορύβου, τις απαιτήσεις της εφαρμογής, και τους διαθέσιμους υπολογιστικούς πόρους. Για εφαρμογές όπου ο τύπος θορύβου είναι γνωστός και σταθερός (π.χ. ιατρική απεικόνιση με Gaussian noise), οι εξειδικευμένες κλασικές μέθοδοι όπως το BM3D παραμένουν η βέλτιστη επιλογή. Για πιο γενικές εφαρμογές με ποικίλους τύπους θορύβου, οι deep learning προσεγγίσεις όπως το U-Net προσφέρουν καλύτερη ευελιξία και adaptability.

Η συνεχής εξέλιξη των deep learning αρχιτεκτονικών και η αυξανόμενη διαθεσιμότητα υπολογιστικής ισχύος υποδηλώνουν ότι το μέλλον της αποθορυβοποίησης εικόνας βρίσκεται στην κατεύθυνση των νευρωνικών δικτύων, τα οποία θα συνεχίσουν να βελτιώνονται με μεγαλύτερα datasets και πιο εξελιγμένες αρχιτεκτονικές.

## 6. ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] Buades, A., Coll, B., & Morel, J. M. (2005). A non-local algorithm for image denoising. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2, 60-65.
- [2] Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8), 2080-2095.
- [3] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, LNCS, 9351, 234-241.
- [4] Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017). Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, 26(7), 3142-3155.
- [5] Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., & Lin, C. W. (2020). Deep learning on image denoising: An overview. *Neural Networks*, 131, 251-275.
- [6] Lebrun, M., Buades, A., & Morel, J. M. (2013). A nonlocal Bayesian image denoising algorithm. *SIAM Journal on Imaging Sciences*, 6(3), 1665-1688.
- [7] Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images. *IEEE Transactions on Image Processing*, 16(5), 1395-1411.
- [8] Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008). Extracting and composing robust features with denoising autoencoders. *Proceedings of the 25th International Conference on Machine Learning*, 1096-1103.
- [9] Jain, V., & Seung, S. (2009). Natural image denoising with convolutional networks. *Advances in Neural Information Processing Systems*, 21, 769-776.
- [10] Mao, X., Shen, C., & Yang, Y. B. (2016). Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. *Advances in Neural Information Processing Systems*, 29, 2802-2810.
- [11] Zhang, K., Zuo, W., & Zhang, L. (2018). FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing*, 27(9), 4608-4622.
- [12] Dong, W., Zhang, L., Shi, G., & Li, X. (2013). Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4), 1620-1630.
- [13] Elad, M., & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12), 3736-3745.
- [14] Mairal, J., Bach, F., Ponce, J., Sapiro, G., & Zisserman, A. (2009). Non-local sparse models for image restoration. *IEEE 12th International Conference on Computer Vision*, 2272-2279.
- [15] Levin, A., & Nadler, B. (2011). Natural image denoising: Optimality and inherent bounds. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2833-2840.



- [16] Singh, N. P., Kumar, R., & Srivastava, R. (2019). An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. *Proceedings of 2nd International Conference on Communication, Computing and Networking*, Springer, Singapore, 409-419.
- [17] Gu, S., Zhang, L., Zuo, W., & Feng, X. (2014). Weighted nuclear norm minimization with application to image denoising. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2862-2869.
- [18] Burger, H. C., Schuler, C. J., & Harmeling, S. (2012). Image denoising: Can plain neural networks compete with BM3D? *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2392-2399.
- [19] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600-612.
- [20] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [21] Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *Proceedings 8th IEEE International Conference on Computer Vision (ICCV)*, 2, 416-423.
- [22] Wikipedia contributors. (2024). Non-local means. *Wikipedia, The Free Encyclopedia*. Retrieved from [https://en.wikipedia.org/wiki/Non-local\\_means](https://en.wikipedia.org/wiki/Non-local_means)
- [23] Bahrami, A., & Kamal, M. (2020). An FPGA-based Hardware Accelerator for Real-Time Block-Matching and 3D Filtering. *International Journal of Advanced Computer Science and Applications*, 11(6), 451-458.

**ΤΕΛΟΣ ΑΝΑΦΟΡΑΣ**

Πανεπιστήμιο Πελοποννήσου - Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών  
Ιανουάριος 2026