

# Όραση Υπολογιστών

## 1<sup>ο</sup> Εργαστήριο

Εντοπισμός Σημείων Ενδιαφέροντος  
και  
Εξαγωγή Χαρακτηριστικών σε Εικόνες

Διαμάντης Αναστάσιος  
el15032@central.ntua.gr  
03115032

# Μέρος 1: Ανίχνευση Ακμών σε Γκρίζες Εικόνες

---

## Τεχνολογία:

Η άσκηση υλοποιήθηκε μέσω python 3.8 σε περιβάλλον anaconda με κύρια χρήση των πακέτων opencv και numpy.

## Περίληψη της άσκησης:

Στο πρώτο μέρος της άσκησης, καλούμαστε να υλοποιήσουμε μια συνάρτηση edgeDetect που ,μέσω διάφορων διεργασιών, υλοποιεί έναν αλγόριθμο ανίχνευσης ακμών , δίνοντάς μας σαν έξοδο μια δυαδική εικόνα ακμών. Η τεχνική που ακολουθούμε συνοπτικά είναι η εξής:

- Διαβάζουμε την εικόνα και της προσθέτουμε θόρυβο (λευκό Gaussian) του οποίου η απόκλιση δίνεται απο τον τύπο :  $PSNR = 20 \log_{10}(\frac{I_{max}-I_{min}}{\sigma_n})$
- Βρίσκουμε την “Laplacian of Gaussian” της εικόνας με 2 τρόπους
  1. Γραμμική Προσέγγιση : Συνελίσσοντας με ένα Gaussian φίλτρο και μετά βρίσκοντας την Λαπλασιανή της φιλτραρισμένης εικόνας :  $L_1 = \nabla^2 (G_\sigma * I)$

2. Μη γραμμική Προσέγγιση : Υπολογίζοντας μέσω μη μορφολογικών φίλτρων το :

$$L_2 = I_\sigma \oplus B + I_\sigma \ominus B - 2I_\sigma$$

- Βρίσκουμε τα zerocrossings της Laplacian με τα εξής 2 βήματα:

1. Μετατρέπουμε την εικόνα σε δυαδική ( $X = L \geq 0$ )

2. Βρίσκουμε το περίγραμμα της δυαδικής μέσω της πράξης :  $Y = X \oplus B - X \ominus B$

- Απορρίπτουμε τα zerocrossings που βρίσκονται σε περιοχές με μεγάλη ομοιογένεια , κρατώντας εκείνα που η λαπλασιανή έχει μεγάλη κλίση :

$$Y[i,j] = 1 \text{ και } ||\nabla I_\sigma[i,j]|| > \theta_{edge} * ||\nabla I_\sigma||_{max}$$

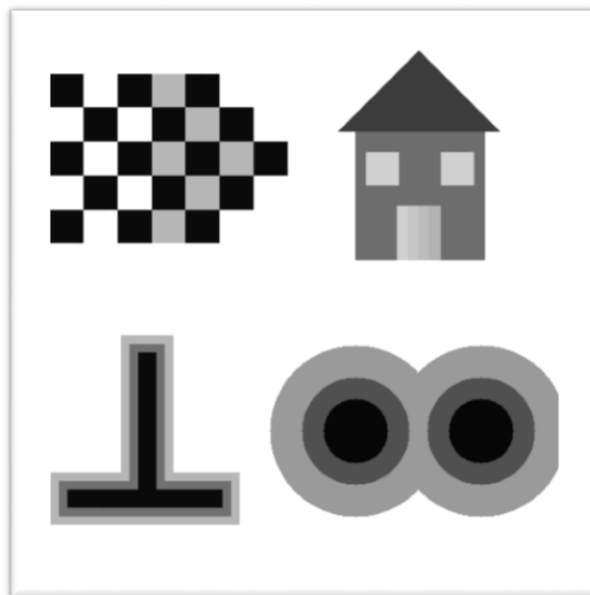
Τέλος, αξιολογούμε τα αποτελέσματα του αλγορίθμου συγκρίνοντας τα με τις «πραγματικές» ακμές με τον εξής τρόπο :

- Αληθινές ακμές:  $M = I_o \oplus B - I_o \ominus B$
- Κριτήριο ποιότητας :  $C = [\text{Pr}(D|T) + \text{Pr}(T|D)]/2$   
Όπου  $\text{Pr}(D|T)$  είναι το ποσοστό των ανιχνευθεισών ακμών που είναι αληθινές (Precision) και  $\text{Pr}(T|D)$  το ποσοστό των αληθινών ακμών που ανιχνεύθηκαν (Recall) .

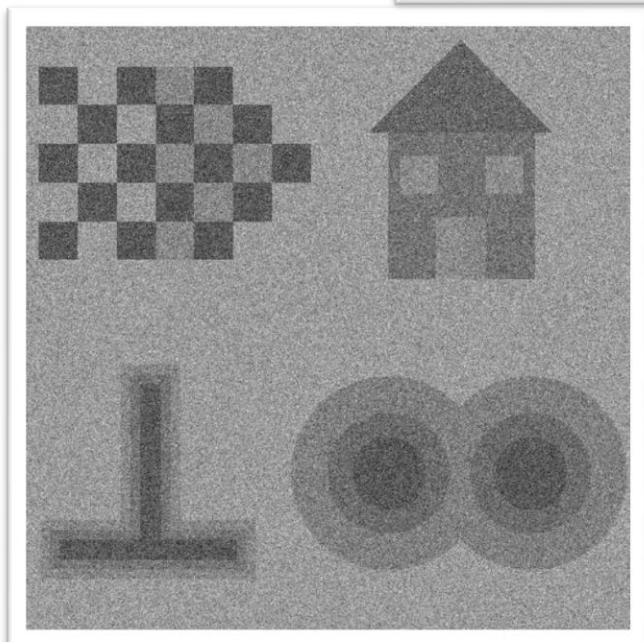
## Αναλυτική Παρουσίαση Αποτελεσμάτων:

Παρακάτω βλέπουμε την αρχική εικόνα στην οποία καλούμαστε να υλοποιήσουμε την άσκηση, καθώς και της θορυβώδεις εικόνες με  $PSNR = 10dB$  και  $PSNR = 20dB$ .

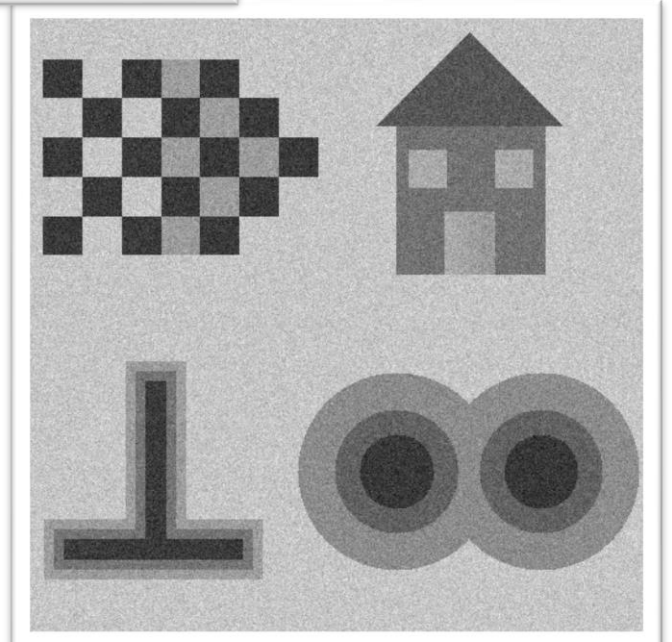
*Αρχική Εικόνα*



*PSNR = 10dB*

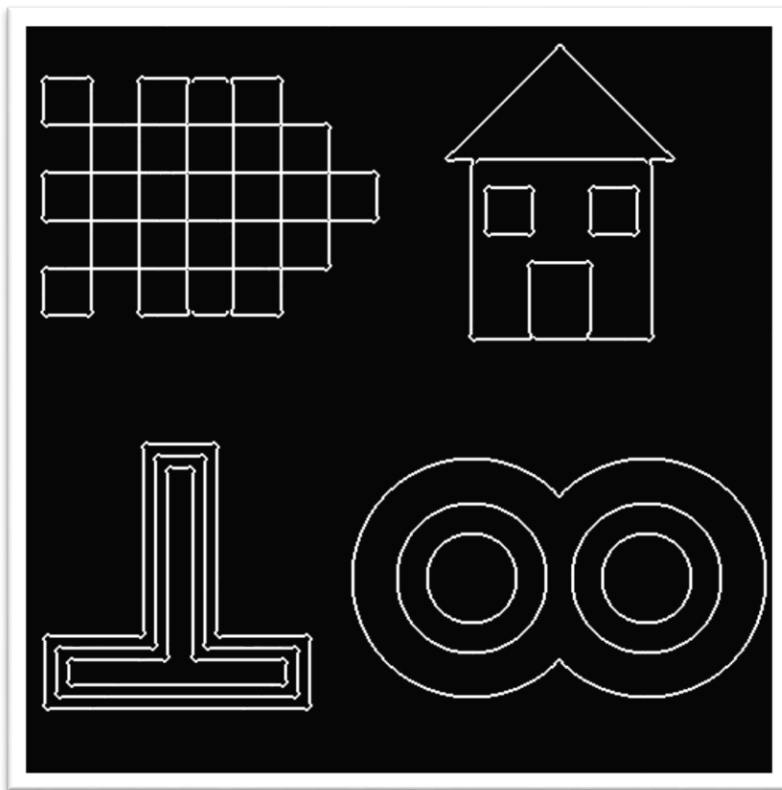


*PSNR = 20dB*



Αρχικά, πρέπει να υπολογίσουμε της αρχικής , μή  
θορυβώδης εικόνας τις αληθινές ακμές, προκειμένου να  
έχουμε ένα μέτρο σύγκρισης.

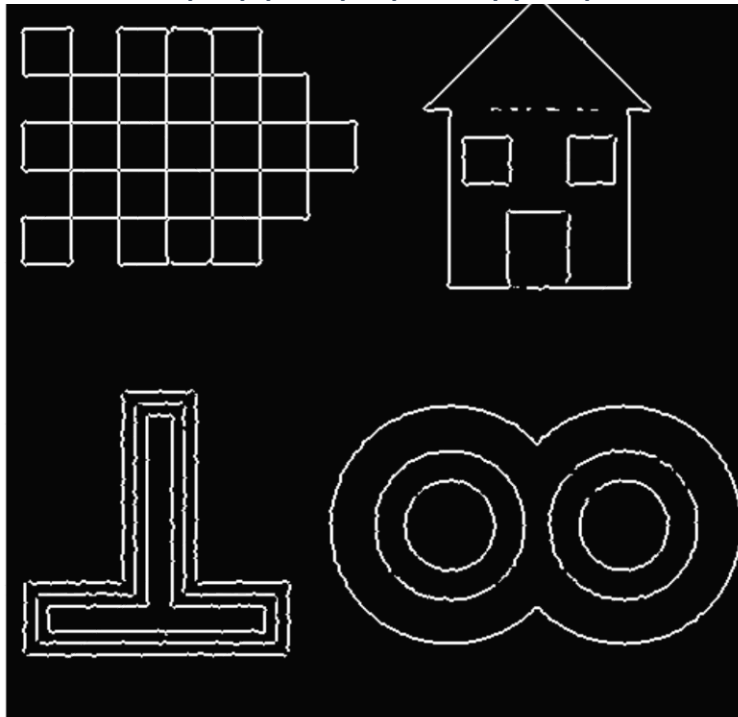
*Ακμές Αληθινής Εικόνας*



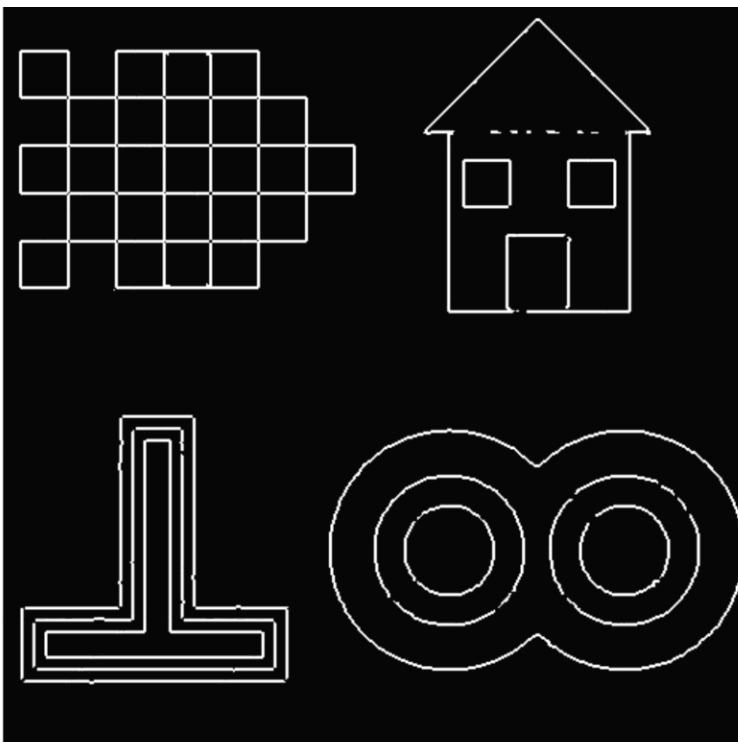
Έπειτα , προχωράμε στην εφαρμογή της `edgeDetect()` για  
τις θορυβώδεις εικόνες PSNR = 10 dB , PSNR = 20dB και για  
κάθε συνδυασμό γραμμικής , μη-γραμμικής προσέγγισης.

Για PSNR = 20dB ( $\sigma_n = 1.5, \theta = 0.2$ ):

*Γραμμική Προσέγγιση*

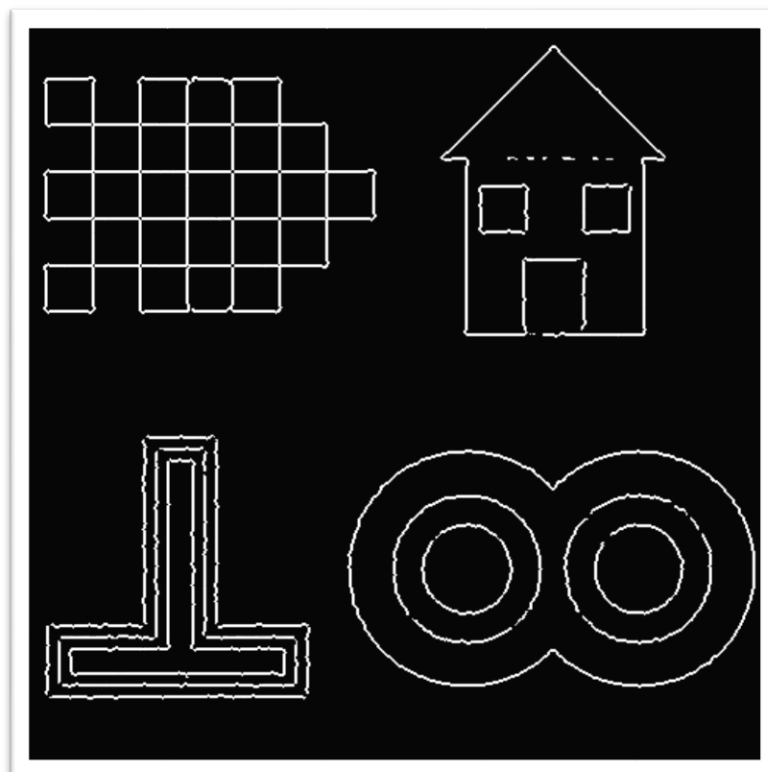
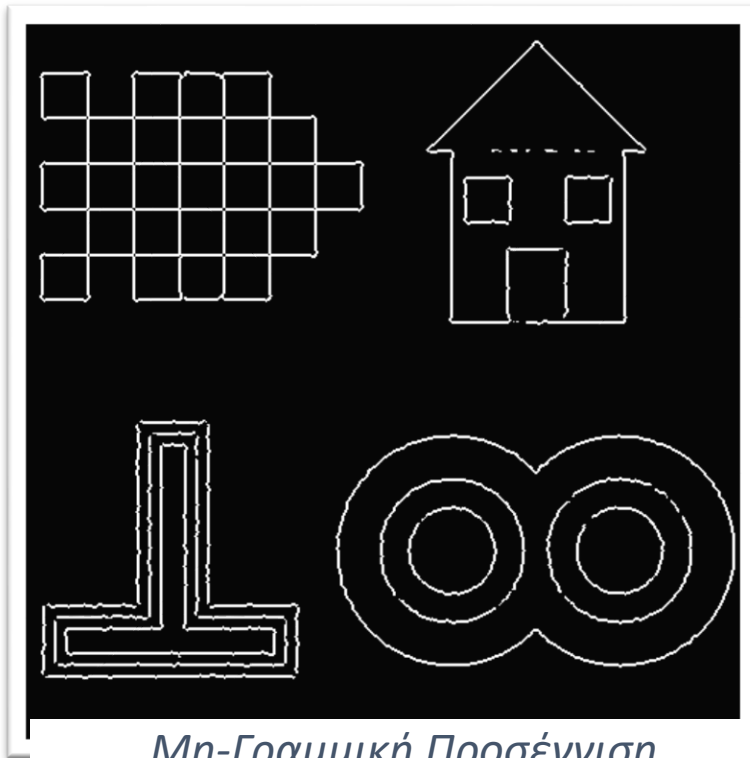


*Μη-Γραμμική Προσέγγιση*



Για PSNR = 10dB ( $\sigma_n = 1.5, \theta = 0.2$ ):

*Γραμμική Προσέγγιση*



Χρησιμοποιώντας την συνάρτηση `evalEdges()` μπορούμε να υπολογίσουμε μια ενδεικτική τιμή ποιότητας της υπολογισθέντος εικόνας .

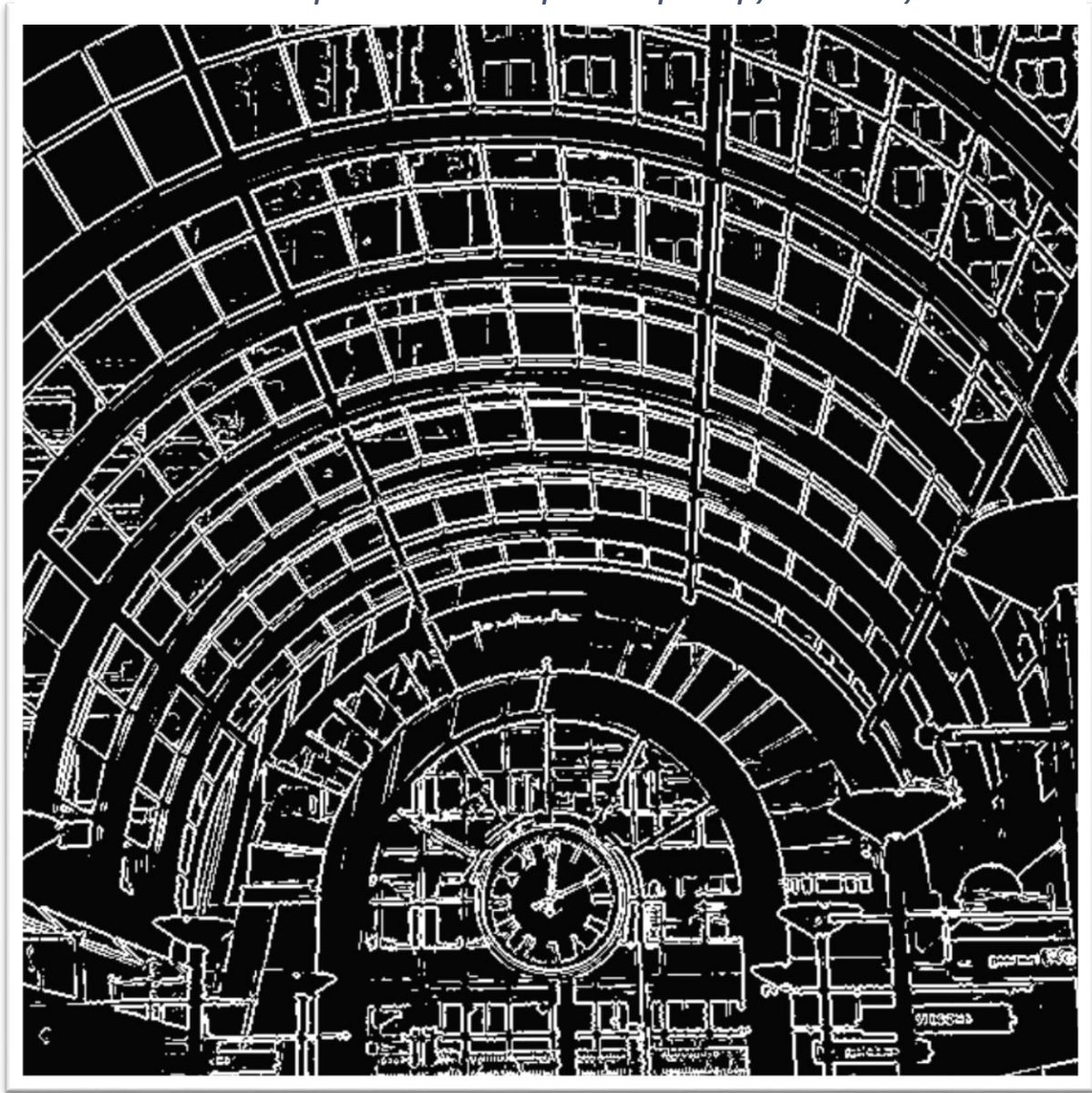
```
Evaluation for linear approach PSNR: 20dB, sn: 1.5, theta: 0.2 is: 0.9263383049696609
Evaluation for non-linear approach PSNR: 20dB, sn: 1.5, theta: 0.2 is: 0.963043403937817
Evaluation for linear approach PSNR: 10dB, sn: 1.5, theta: 0.2 is: 0.6709955385290559
Evaluation for non-linear approach PSNR: 10dB, sn: 1.5, theta: 0.2 is: 0.7540141743182192
```

Απο την παραπάνω αξιολόγηση των αποτελεσμάτων βλέπουμε έως 96% για την εικόνα με PSNR = 20dB και έως 75% για την εικόνα με PSNR = 10dB. Παρατηρούμε ότι έχουμε καλύτερα αποτελέσματα με την μη-γραμμική προσέγγιση.

Παρακάτω εκτελούμε τον ίδιο αλγόριθμο για την πραγματική εικόνα 'urban\_edges.jpg' και παρουσιάζουμε το καλύτερο αποτέλεσμα που λάβαμε πειραματίζοντας με τις τιμές  $s_n$  και  $theta$ .



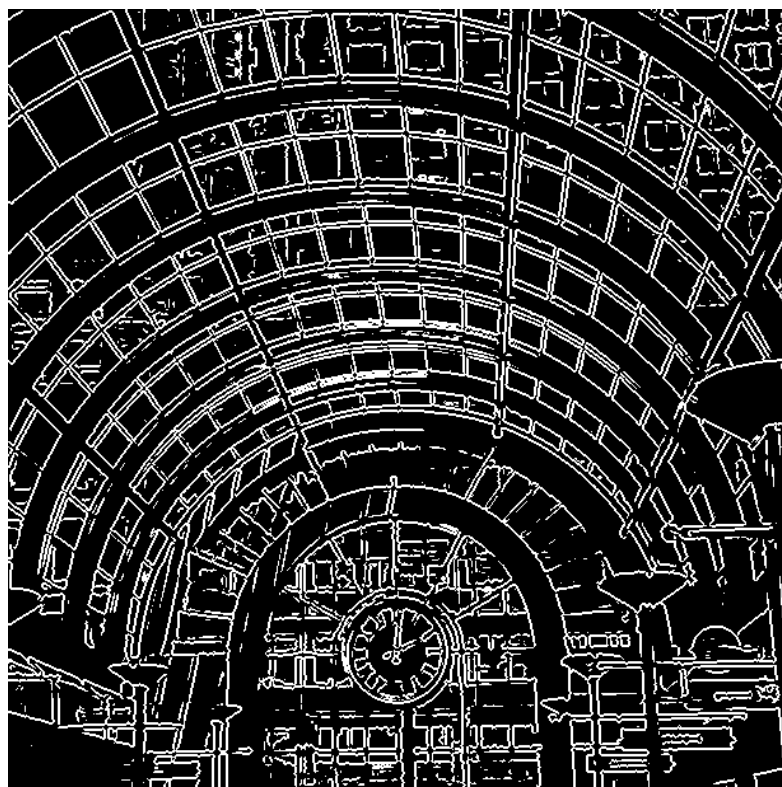
### Καλύτερο Αποτέλεσμα Αληθινής Εικόνας



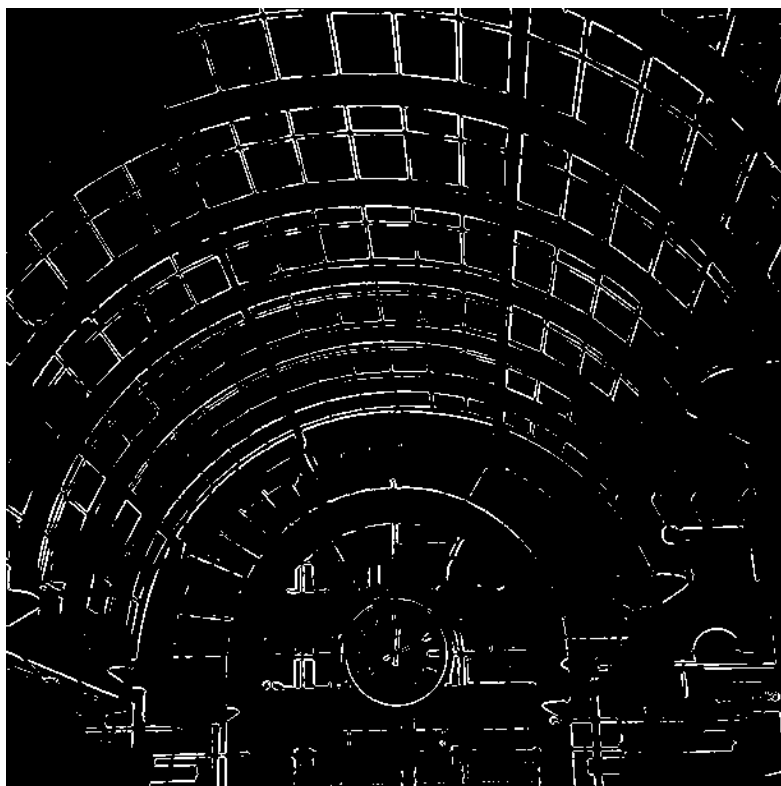
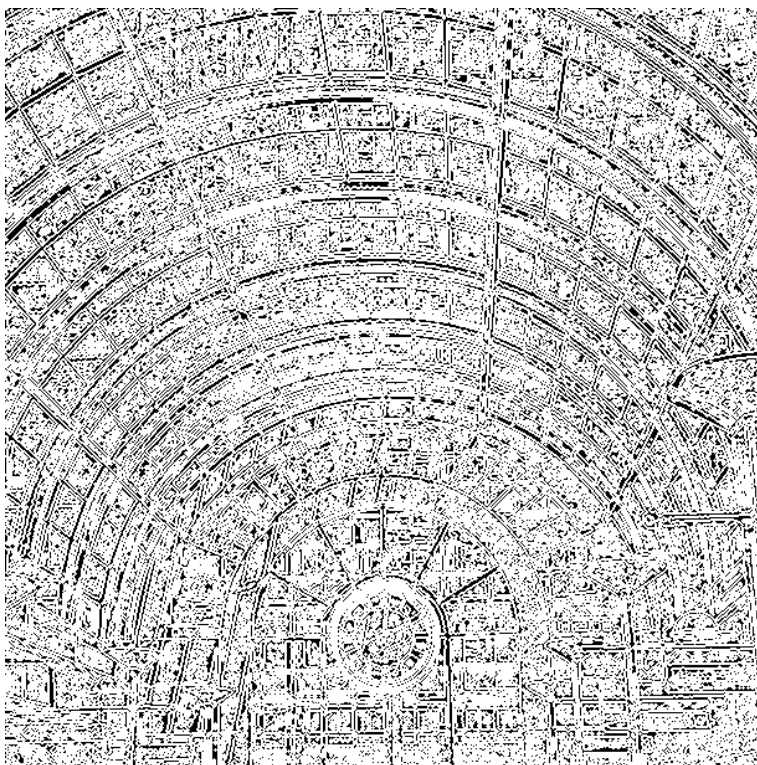
Evaluation for "urban\_edges" with sn :0.2 , theta: 0.2 is: 0.81302561894

Παρακάτω παραθέτουμε τα αποτελέσματα που λάβαμε πειραματίζοντας με τις τιμές  $s_n$  και  $\theta$  στην αληθινή εικόνα καθώς και τα συμπεράσματα που λάβαμε.

*Αυξανόμενο  $\sigma_n$ :*



*Αυξανόμενο θ:*



## **Συμπέρασματα:**

Παρατηρούμε ότι όσο αυξάνει το  $\sigma_n$  αυξάνεται το μέγεθος του Γκαουσιανού kernel με έναν παράγοντα 3 οπότε ραγδαία η αρχική εικόνα «θολώνει» και έχουμε αποτέλεσμα με όλο και λιγότερο σαφώς οριοθετημένες ακμές.

Όσον αφορά το  $\theta$  , για πολύ μικρές τιμές του , λαμβάνουμε υπερβολική πληροφορία που καθιστά την εικόνα ακμών άχρηστη , και όταν αυξάνει πολύ χάνουμε μεγάλη ποσότητα πληροφορία με αποτέλεσμα να μην διακρίνουμε τις ακμές.

## Μέρος 2: Ανίχνευση Σημείων Ενδιαφέροντος (Interest Point Detection)

---

### Ανίχνευση Γωνιών

Περίληψη της άσκησης

Στο πρώτο μέρος της άσκησης υλοποιούμε έναν ανιχνευτή γωνιών ο οποίος χρησιμοποιεί την μέθοδο Harris-Stephens. Αρχικά πρέπει να υπολογίσουμε τους δομικούς τανυστές τις εικόνες που προκύπτουν από την εξής εξίσωση:

$$J_1(x, y) = G_\rho * \left( \frac{\partial I_\sigma}{\partial x} \cdot \frac{\partial I_\sigma}{\partial x} \right) (x, y)$$

$$J_2(x, y) = G_\rho * \left( \frac{\partial I_\sigma}{\partial x} \cdot \frac{\partial I_\sigma}{\partial y} \right) (x, y)$$

$$J_3(x, y) = G_\rho * \left( \frac{\partial I_\sigma}{\partial y} \cdot \frac{\partial I_\sigma}{\partial y} \right) (x, y)$$

Οι τυπικές αποκλίσεις των κλιμάκων διαφορίσης και ολοκλήρωσης αντίστοιχα επηρεάζουν άμεσα τον δομικό τανυστή. Έπειτα υπολογίζουμε τις ιδιοτιμές του δομικού τανυστή μέσω της εξίσωσης:

$$\lambda_{\pm}(x, y) = \frac{1}{2} \left( J_1 + J_3 \pm \sqrt{(J_1 - J_3)^2 + 4J_2^2} \right)$$

Μέσω του οποίου προκύπτει το κριτήριο γωνιότητας R το οποίο εκφράζεται ως :

$$R(x, y) = \lambda_- \lambda_+ - k \cdot (\lambda_- + \lambda_+)^2$$

Με βάση το παραπάνω κριτήριο επιλέγουμε ως γωνίες τα pixel όπου :

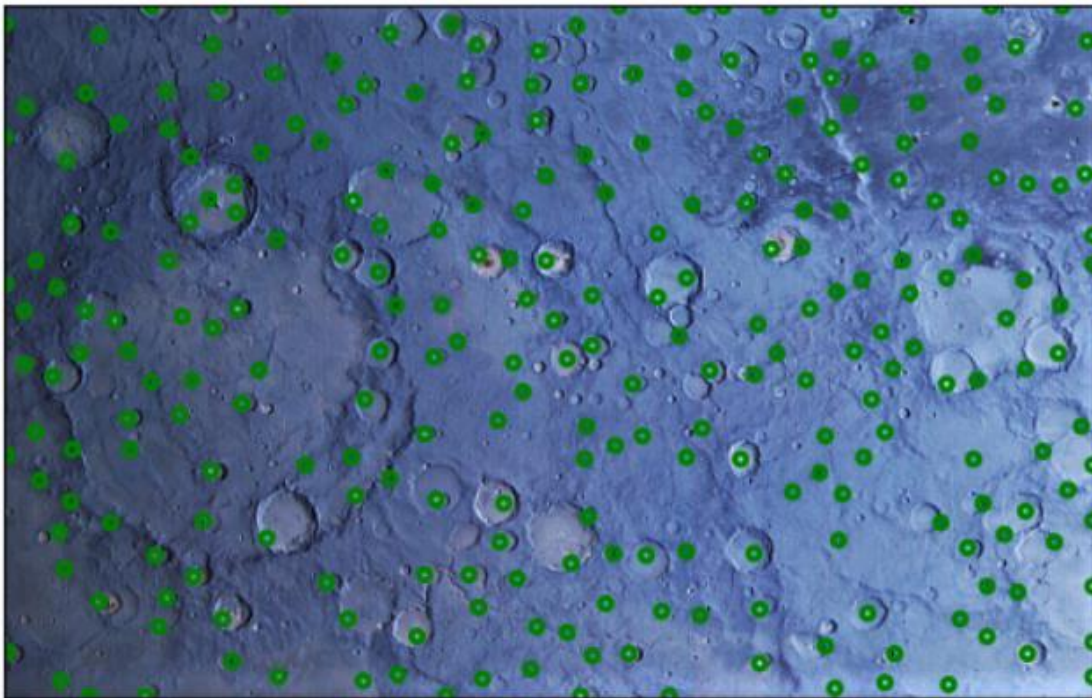
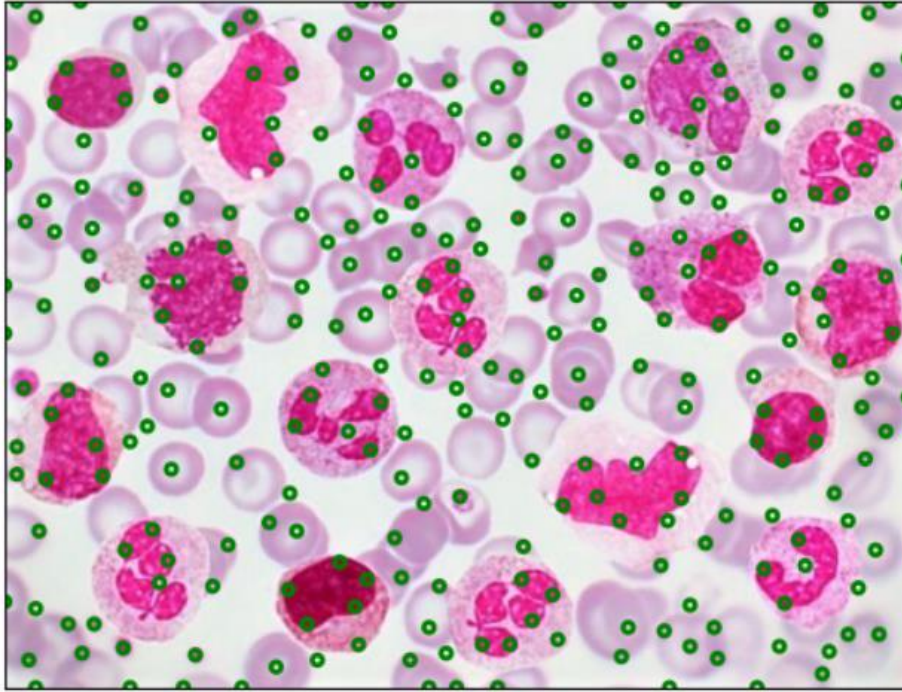
- Είναι μέγιστα του R εντός τετραγωνικών παραθύρων
- Αντιστοιχούν σε τιμή του R μεγαλύτερη απο ένα ποσοστό του συνολικού μέγιστου , δηλαδή :

$$R(x, y) > \theta_{corn} \cdot R_{max}$$

Αναλυτική Παρουσίαση Αποτελεσμάτων

Παρακάτω παραθέτουμε τα αποτελέσματα που λάβαμε για τις 2 διαφορετικές εικόνες:





Προφανώς τα αποτελέσματα δεν είναι βέλτιστα αφού επηρεάζονται απο τις επιλεγμένες τιμές των παραμέτρων.

Πολυκλιμακωτή Ανίχνευση Γωνιών

Σε αυτό εκτελούμε τον αλγόριθμο Harris-Laplacian ο οποίος πρόκειται για παραλλαγή του Harris-Stephens στην οποία έχουμε προσθέσει 2 επιπλέον στάδια:

Υπολογίζουμε διάφορες κλίμακες διαφόρισης και ολοκλήρωσης:

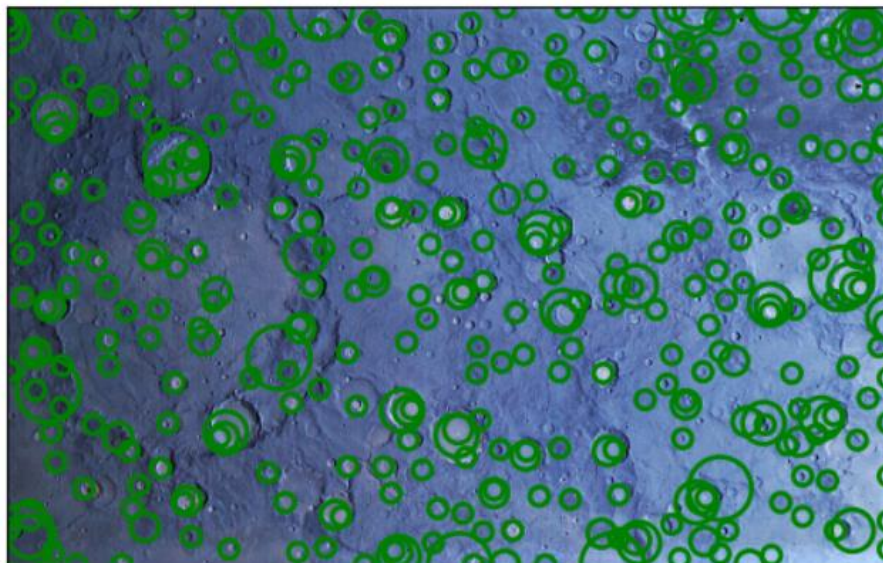
$$\begin{aligned}\sigma_0, \sigma_1, \dots, \sigma_{N-1} &= s^0 \sigma_0, s^1 \sigma_0, \dots, s^{N-1} \sigma_0 \\ \rho_0, \rho_1, \dots, \rho_{N-1} &= s^0 \rho_0, s^1 \rho_0, \dots, s^{N-1} \rho_0\end{aligned}$$

Υπολογίζουμε την κανονικοποιημένη Λαπλασιανή της Γκαουσιανής και απορρίπτουμε τα σημεία για τα οποία η

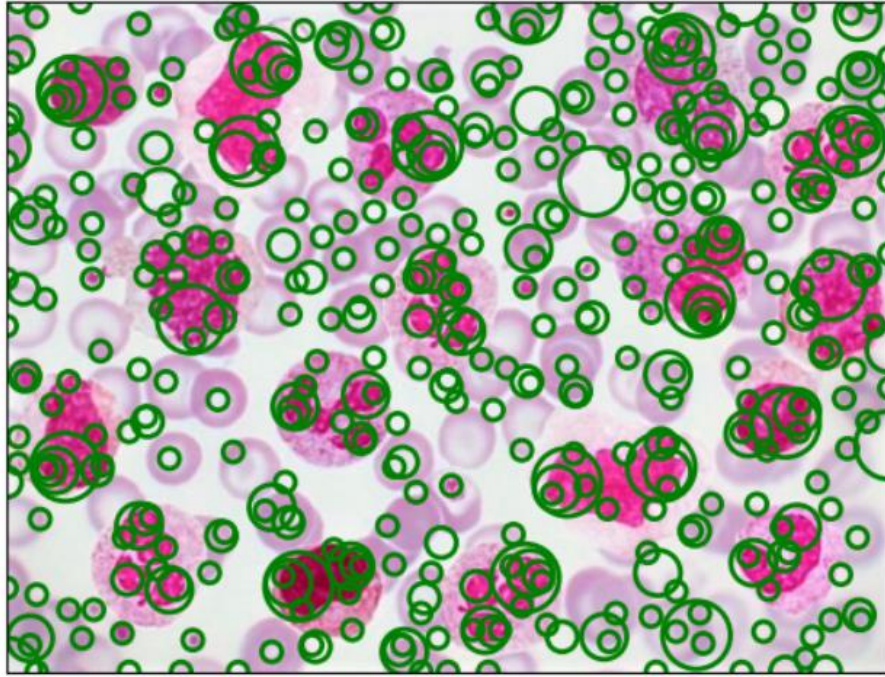
$$|LoG(\mathbf{x}, \sigma_i)| = \sigma_i^2 |L_{xx}(\mathbf{x}, \sigma_i) + L_{yy}(\mathbf{x}, \sigma_i)|, \quad i = 0, \dots, N - 1$$

κλίμακα που ανιχνεύθηκαν δεν μεγιστοποιεί την LoG μετρική σε μια γειτονιά 2 διαδοχικών κλιμάκων.

Το αποτέλεσμα που λαμβάνουμε για  $N = 4$  διαφορετικές κλίμακες διαφόρισης και ολοκλήρωσης είναι τα εξής:







## Ανίχνευση Blobs

Μια από τις σημαντικότερες κατηγορίες σημείων ενδιαφέροντος βασίζονται στην ανίχνευση 'blobs', που ορίζονται ως περιοχές με κάποια ομοιογένεια που διαφέρουν σημαντικά από την γειτονιά τους.

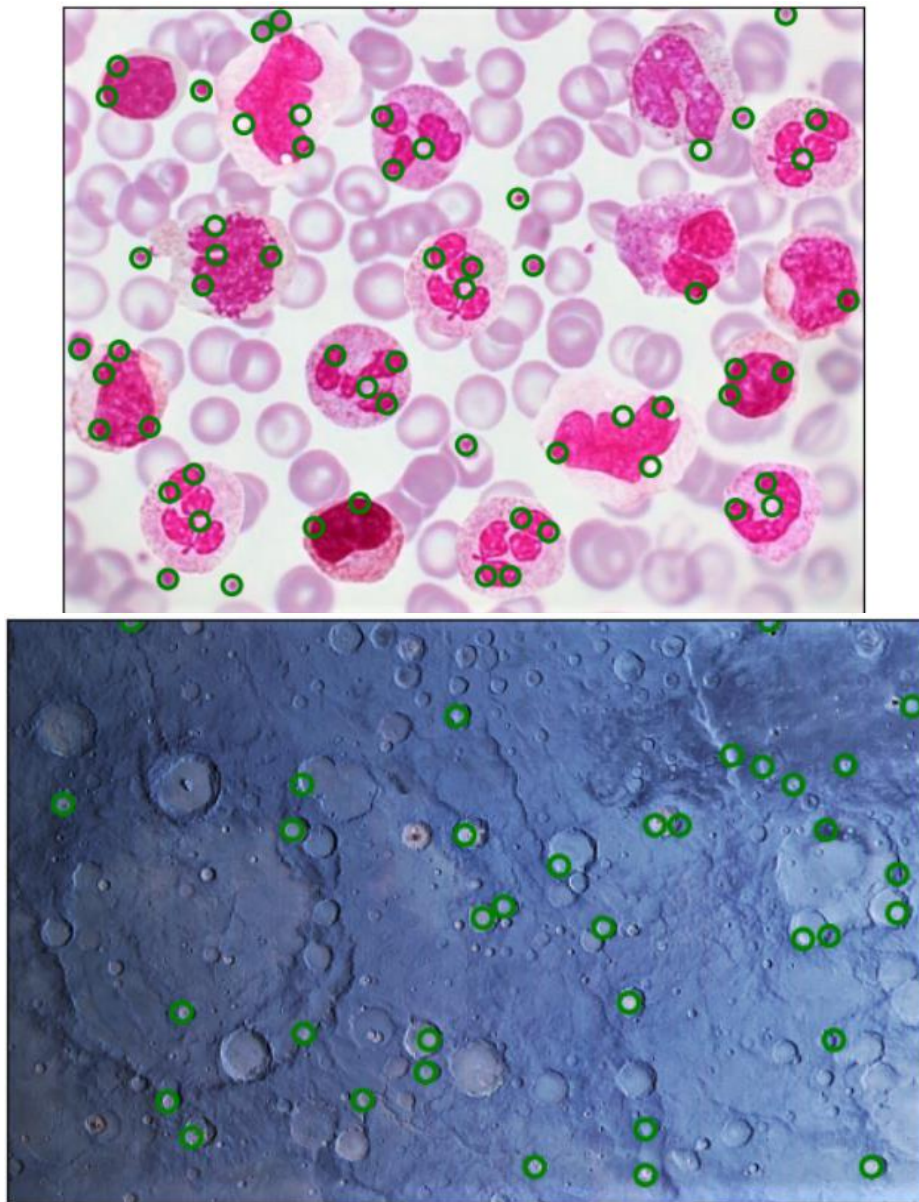
Τα Blobs τα βρίσκουμε με την βοήθεια της Hessian :

$$H(x, y) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix}$$

Στην συνέχεια για κάθε pixel της εικόνας κατασκευάζουμε το κριτήριο  $R(x, y) = \det(H(x, y))$  και επιλέγουμε σαν σημεία ενδιαφέροντος εκείνα τα οποία είναι τοπικά μέγιστα και έχουν τιμή μεγαλύτερη απο ένα κατάλληλα ορισμένο

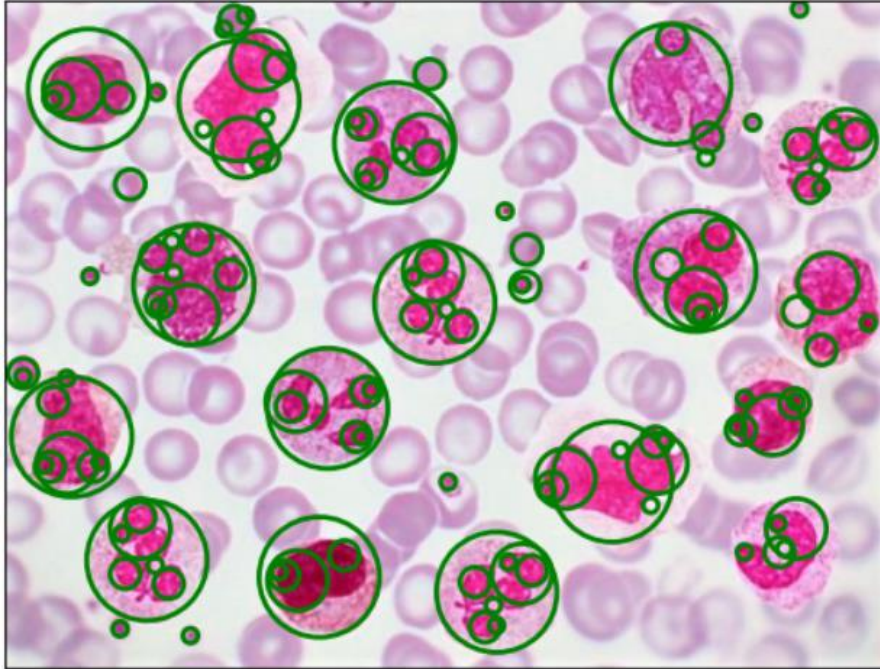
κατώφλι .

Παρακάτω παραθέτουμε τα αποτελέσματα που λάβαμε:

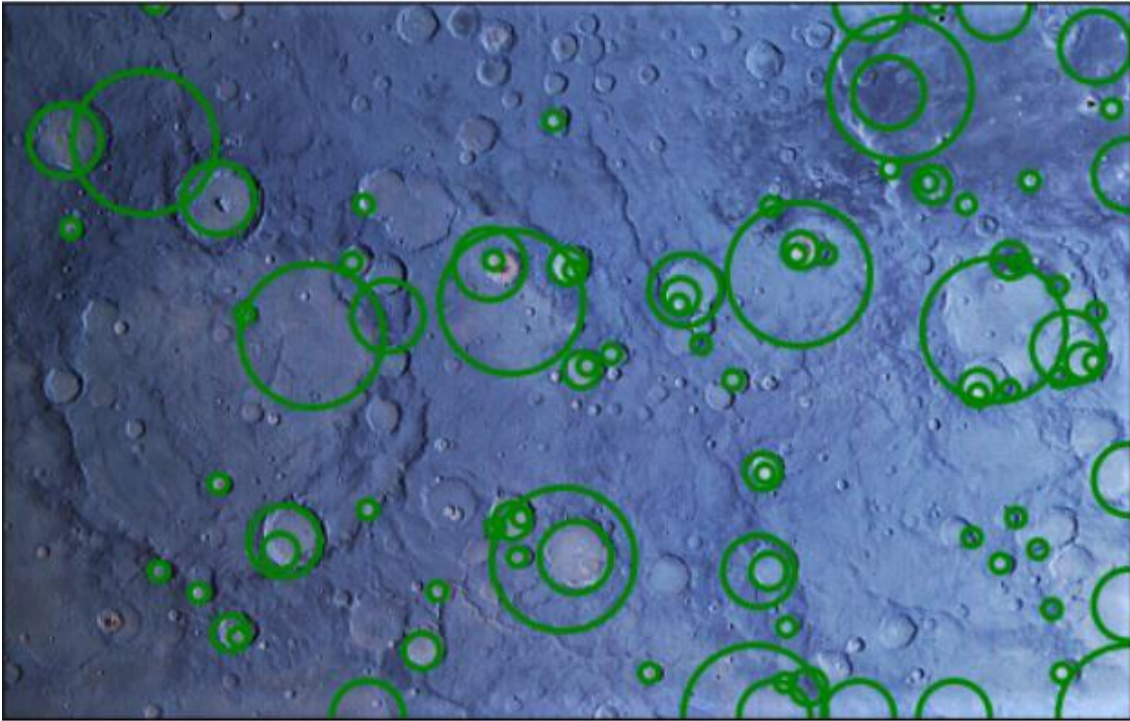


Πολυκλιμακωτή Ανίχνευση Blobs

Σε αυτό το τμήμα επαναλαμβάνουμε την ίδια διαδικασία ακριβώς με την πολυκλιμακωτή ανίχνευση γωνιών. Παραθέτουμε παρακάτω τα αποτελέσματα για  $N = 4$  κλίμακες.







Εδώ παρατηρούμε την σημαντικότητα των blobs καθώς με την πολυκλιμακωτή ανίχνευσή τους εντοπίζουμε στην πρώτη εικόνα τις περιοχές με ομοιογένεια (υποπτεύομαι ότι είναι τα ερυθρά αιμοσφαίρια) και στην δεύτερη τους κρατήρες στην επιφάνειες του Άρη.