



Εθνικό Μετσόβιο Πολυτεχνείο (Ε.Μ.Π)
Σχολή Μηχανολόγων Μηχανικών

ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ ΚΑΤΕΡΓΑΣΙΩΝ, ΜΑΘΗΜΑ 9^{ΟΤ} ΕΞΑΜΗΝΟΥ

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

Μηχανική Όραση

Σπουδαστές:

Λάζαρος Χριστοφορίδης (mc18012)

Μιχάλης Παπαδάκης (mc18026)

Καθηγητής:

Πανώριος Μπερνάρδος

19 Ιουλίου 2025

Περιεχόμενα

Περιεχόμενα	1
Κατάλογος Σχημάτων	2
Κατάλογος Πινάκων	2
1 Εισαγωγή	3
2 Επεξεργασία Εικόνων	4
2.1 Μορφολογική Ανάλυση	4
3 Εξαγωγή Δεδομένων	7
4 Ανάπτυξη TNΔ	8
4.1 Αρχικό Data Set	8
4.2 Αξιολόγηση αρχικού Data Set	8
4.3 Τελικό Data Set	9
4.4 Αξιολόγηση τελικού Data Set	10
5 Video	11
6 Παράρτημα	12

Κατάλογος Σχημάτων

1	Επίδραση αύξησης αντίθεσης	4
2	Συνολική επεξεργασία εικόνας εγκλείσματος	5
3	Συνολική επεξεργασία εικόνας στικτής επιφάνειας	6
4	Συνολική επεξεργασία εικόνας κηλίδας	6
5	Διαγράμματα με τις πιθανές εισόδους του TNΔ	7
6	Τεχνητό Νευρωνικό Δίκτυο κατηγοριοποίησης ελλατωμάτων με τις αρχικές εισόδους	8
7	Αξιολόγηση Αρχικού TNΔ	8
8	Επίδραση αύξησης αντίθεσης	9
9	Τελικό TNΔ	10
10	Αξιολόγηση Τελικού TNΔ	10
11	Συνολική επεξεργασία 'κακής' εικόνας κηλίδας	12
12	Συνολική επεξεργασία 'κακής' εικόνας στικτής επιφάνειας	12
13	Μορφή DataSet για την εκπαίδευση του TNΔ	12

Κατάλογος Πινάκων

1	Απόδοση εκπαιδευμένου TNΔ με διαφορετικές αρχικοποιήσεις (random states)	10
---	--	----

1. Εισαγωγή

Σκοπός της εργασίας είναι η δημιουργία μιας εφαρμογής που θα αναγνωρίζει αυτόματα τον τύπο ελαττώματος σε τεμάχια που έχουν υποστεί έλαση εν θερμώ χρησιμοποιώντας μεθόδους μηχανικής όρασης. Συνολικά εμφανίζονται τρεις κατηγορίες ατελειών: κηλίδες (spots), στικτή επιφάνεια (pitted surface) και εγκλείσματα (inclusions), οι οποίες εμφανίζουν διαφορετική μορφολογία. Τα βήματα της εργασίας είναι η αρχική επεξεργασία εικόνων των παραπάνω ελαττωμάτων, με σκοπό την βελτίωση της ποιότητας της εικόνας και τον εντοπισμό των ατελιών σε αυτές. Έστερα, εξάγονται διάφορα χαρακτηριστικά μεγέθη από την εικόνα, όπως ο αριθμός των pixels που αντιστοιχούν σε ελάττωμά κλπ. Το τελευταίο στάδιο είναι η ανάπτυξη ενός Τεχνητού Νευρωνικού Δικτύου (Τ.Ν.Δ.) κατηγοριοποίησης το οποίο δέχεται σαν είσοδο τα εξαγόμενα χαρακτηριστικά των εικόνων και εκτιμάει τον τύπο του ελαττώματος.

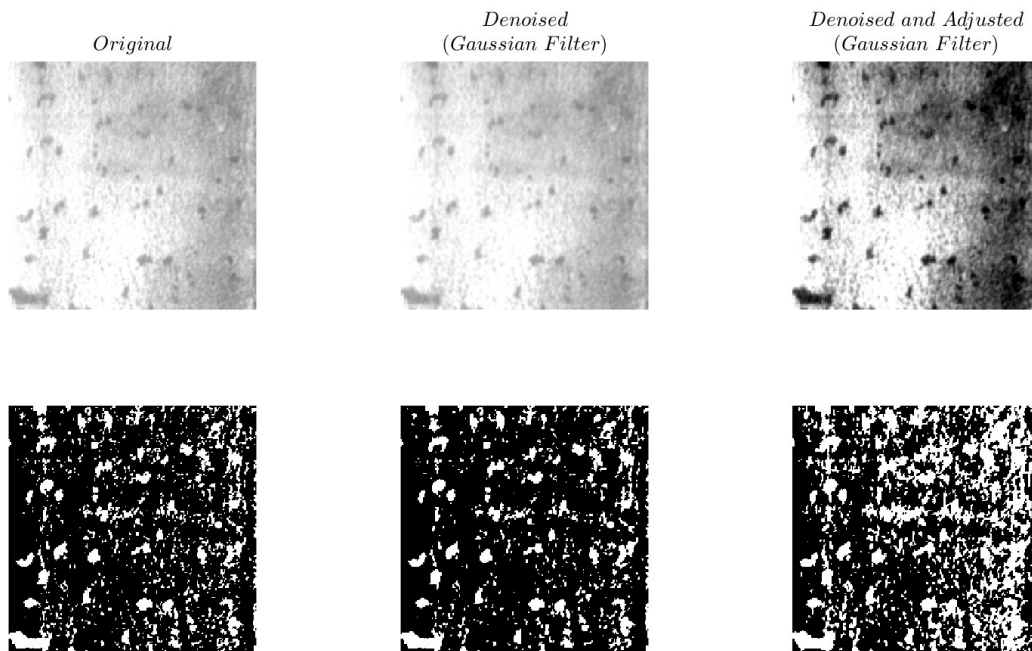
Για την ανάπτυξη της παραπάνω εφαρμογής χρησιμοποιήθηκε το *image processing toolbox*¹ της matlab, το οποίο περιέχει διάφορους αλγόριθμους για image processing, με σκοπό την αρχική επεξεργασία των εικόνων των ελαττωμάτων και την εξαγωγή των επιθυμητών χαρακτηριστικών από αυτές. Επιπλέον, αναπτύχθηκε το ΤΝΔ χρησιμοποιώντας το πακέτο *sk-learn* της python. Πιο συγκεκριμένα χρησιμοποιήθηκε το μοντέλο Multi-Layer Classifier² που είναι ένα feed-forward νευρωνικό δίκτυο κατηγοριοποίησης.

¹<https://www.mathworks.com/products/image.html>

²https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html

2. Επεξεργασία Εικόνων

Η επεξεργασία των εικόνων αποσκοπεί στην βελτίωση της ποιότητάς τους για την ευκολότερη και πιο αξιόπιστη εξαγωγή χαρακτηριστικών από αυτές. Οι εικόνες του dataset ήταν εξ αρχής σε grayscale, όποτε η προ-επεξεργασία αποσκοπεί κυρίως στην αποθρομβοποίηση τους. Αρχικά, ελέγχθηκε η περίπτωση να γίνει ενίσχυση της αντίθεσης, καθώς αποτελεί ένα συνηθισμένο βήμα στην επεξεργασία, ωστόσο παρατηρήθηκε πως οδηγούσε σε ενίσχυση τόσο του θορύβου (ακόμα και μετά από την εφαρμογή φίλτρων) όσο και ανεπιθύμητων διαβαθμίσεων φωτεινότητας. Ενδεικτικά αποτελέσματα της ενίσχυσης αντίθεσης φαίνονται στην εικόνα 1, όπου παρατηρούνται όσο αναφέρθηκαν προηγουμένως, με αποτέλεσμα την έκπτωση της ποιότητας της δυαδικής εικόνας³. Για την αποθρομβοποίηση χρησιμοποιήθηκε ένα γκαουσιανό φίλτρο, το οποίο είναι ένα τυπικό φίλτρο αφαίρεσης θορύβου.



Σχήμα 1: Επίδραση αύξησης αντίθεσης

Έχοντας αφαιρέσει τον θόρυβο, γίνεται διαδικοποίηση της εικόνας. Γενικά, επιλέγεται μια τιμή έντασης (φράγμα) πάνω από την οποία όλα τα πιξελ γίνονται άσπρα (1) και κάτω από την οποία γίνονται μαύρα (0). Η matlab υλοποιεί την διαδικοποίηση μέσω της εντολής *imbinarize()*⁴ και περιέχει μια ρουτίνα που βρίσκει αυτόματα ένα βέλτιστο - προσαρμοστικό φράγμα (*adaptive threshold*), αρκεί να προσδιοριστεί αν το προσκλήνιο είναι πιο σκοτεινό ή φωτεινό από τον φόντο. Μάλιστα, η ρουτίνα αυτή αντισταθμίζει διαφορές στην φωτεινότητα που οφείλονται στις συνθήκες φωτισμού⁵.

2.1 Μορφολογική Ανάλυση

Έχοντας τη δυαδική εικόνα, γίνονται διάφορες μορφολογικές πράξεις (κλείσιμο, άνοιγμα, γέμισμα) με σκοπό να τονισθούν βασικά χαρακτηριστικά του κάθε ελαττώματος και να εξαχθούν πιο αντιπροσωπευτικά χαρακτηριστικά σε επόμενο στάδιο.

1. Αρχικά, γίνεται η πράξη του κλεισίματος, με αποτέλεσμα να κλείσουν μικρές φωτεινές τρύπες και να εξομαλυνθούν τα όριά των φωτεινών σχημάτων (Τα οποία λόγω της αντιστροφής αντιστοιχούν σε ελαττώματα και όχι στον φόντο όπως στην αρχική εικόνα). Η διαδικασία αυτή φαίνεται στις εικόνες I4 στα σχήματα 2,3 και 4.
2. Έστερα, γίνεται η διαδικασία του ανοίγματος, με συνήθως μεγαλύτερο δομικό στοιχείο. Αυτή η πράξη έχει στόχο την αφαίρεση του εναπομείνοντα θορύβου.

³ Στην συγκεκριμένη περίπτωση, έχει γίνει πρώτα αποθρομβοποίηση και ύστερα αύξηση της αντίθεσης, και πάλι γίνεται ενίσχυση του θορύβου.

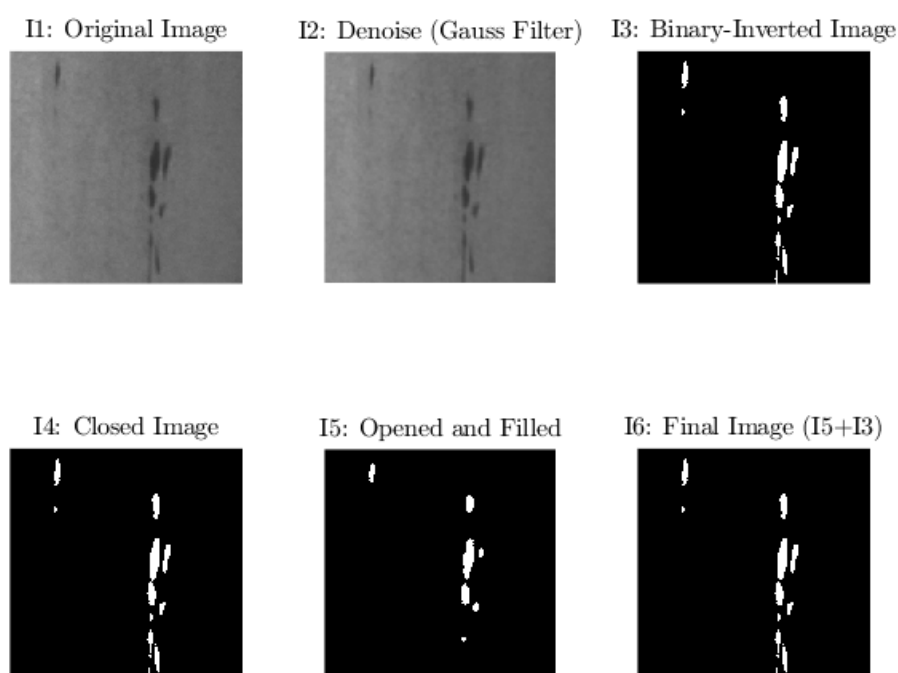
⁴ <https://www.mathworks.com/help/images/ref/imbinarize.html>

⁵ <https://www.mathworks.com/help/images/ref/adaptthresh.html>

3. Τέλος, γεμίζονται τα κενά που έχουν μείνει μέσω της εντολής *imfill*⁶. Η διαδικασία αυτή και η προηγούμενη φαίνονται στις εικόνες I5 στα σχήματα 2,3 και 4.

Η παραπάνω μορφολογική ανάλυση όπως έχει ρυθμιστεί (επιλέγοντας την αλληλουχία των πράξεων και το μέγεθος των δοκιμών στοιχείων) **αναδεικνύει τις συμπαγείς επιφάνειες**, δηλαδή τα ελαττώματα τύπου κηλίδων και τα εγκλείσματα. Λόγω της δεύτερης πράξης (άνοιγμα), πολλές φορές οι στικτές επιφάνειες χάνονται. Για την διατήρηση όλων των αντιπροσωπευτικών πληροφοριών, επιλέχθηκε η τελική εικόνα (I6 στα σχήματα 2,3 και 4) από την οποία θα εξαχθούν τα χαρακτηριστικά να είναι η άθροιση της εικόνας πριν από την μορφολογική ανάλυση (που περιέχει πληροφορίες για την στικτή επιφάνεια) και της εικόνας ύστερα από την μορφολογική ανάλυση (που αναδεικνύει συμπαγείς επιφάνειες).

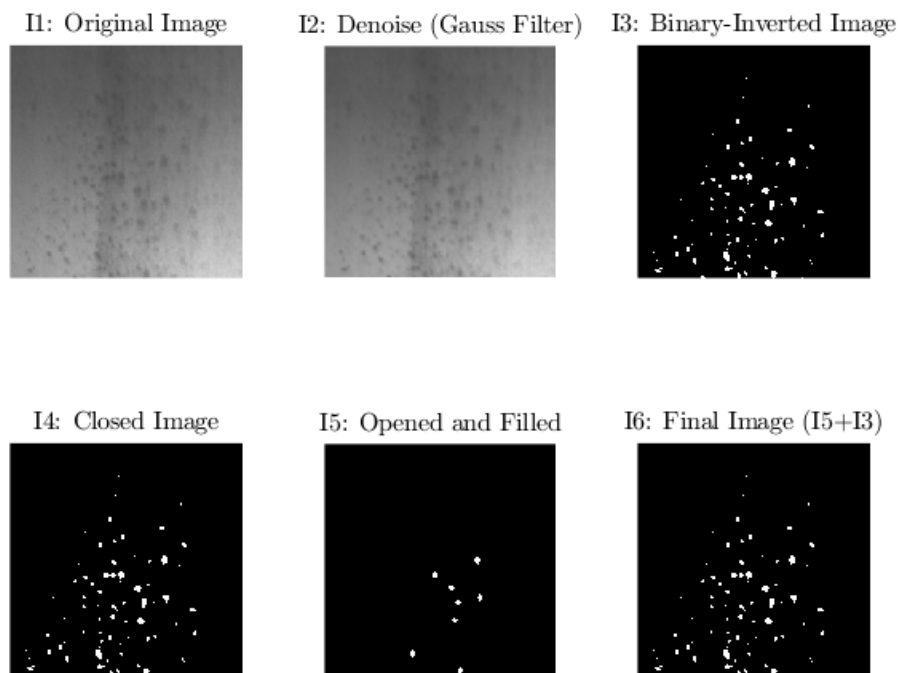
Ενδεικτικά αποτελέσματα φαίνονται στις εικόνες 2,3 και 4 που ακολουθούν. Περισσότερες εικόνες φαίνονται στο παράρτημα. Εδώ πρέπει να τονισθεί πως δεν εξετάσθηκε η βέλτιστη ρύθμιση της μορφολογικής ανάλυσης. Η τελική κατηγοριοποίηση γίνεται μέσω νευρωνικού δικτύου, το οποίο δεν είναι ευαίσθητο σε μικρές διαφορές στις εισόδους του. Οπότε, **η ανάλυση της εικόνας πρέπει να είναι τόσο ικανοποιητική, όσο τα εξαγόμενα στοιχεία να αρκούν ώστε το ΤΝΔ να κάνει υψηλής ακρίβειας classification**.⁷



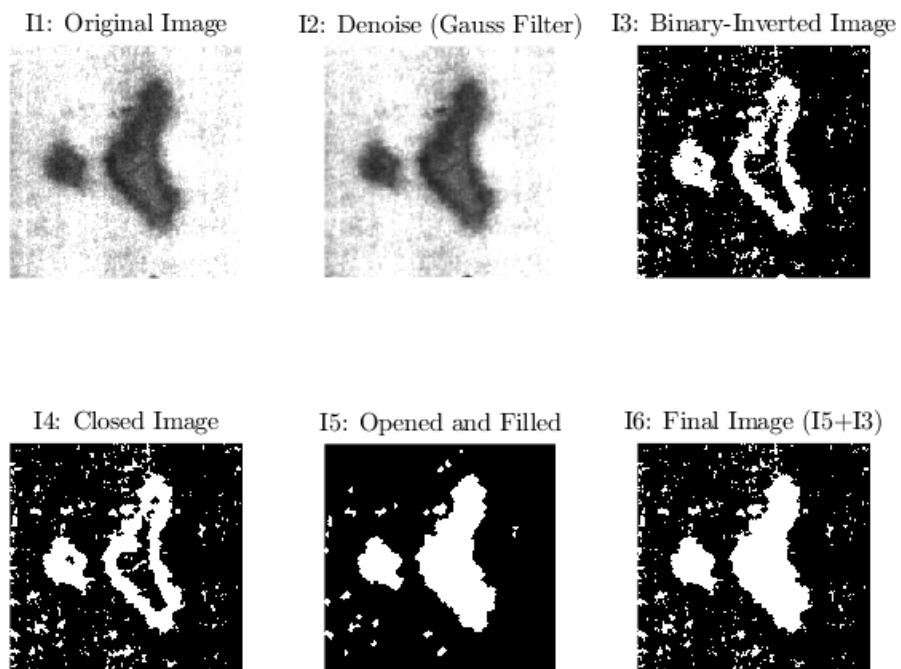
Σχήμα 2: Συνολική επεξεργασία εικόνας εγκλείσματος

⁶https://www.mathworks.com/help/images/ref/imfill.html?searchHighlight=imfill&s_tid=srchtitle_imfill_1

⁷Η τελική ρύθμιση των παραμέτρων της μορφολογικής ανάλυσης έγινε αφού είχε δημιουργηθεί ο κώδικας για την εκπαίδευση του ΤΝΔ. Έτσι, ύστερα από κάθε αλλαγή στις ρυθμίσεις της μορφολογικής ανάλυσης, γινόταν εκπαίδευση και αξιολόγηση της συμπεριφοράς του ΤΝΔ. Η τελική ρύθμιση οδήγησε σε ακριβής κατηγοριοποίηση, **οπότε κρίνεται επαρκής για τις ανάγκες της παρούσας εφαρμογής**.



Σχήμα 3: Συνολική επεξεργασία εικόνας στικτής επιφάνειας



Σχήμα 4: Συνολική επεξεργασία εικόνας κηλίδας

3. Εξαγωγή Δεδομένων

Αφού η αρχική εικόνα έχει υποστεί επεξεργασία, με τα βήματα που προαναφέρθηκαν, μπορούν να ληφθούν κατάλληλα χαρακτηριστικά. Αρχικά ως χαρακτηριστικά, επιλέχθηκαν:

- Αριθμός στοιχείων (Objects)
- Μέση Ισοδύναμη Διάμετρος
- Μέση Κυκλικότητα
- Απόλυτη Γωνία Μέγιστης Επιφάνειας
- Μέγιστη Επιφάνεια
- Μέγιστη Περίμετρος
- Αριθμός Λευκών πίξελ (σαν Εμβαδόν)

Χρησιμοποιώντας την έτοιμη εντολή της matlab (*regionprops*), εξάγονται τα παραπάνω χαρακτηριστικά, και αποθηκεύονται σε πίνακα, ώστε να βρεθούν οι μέσες και μέγιστες τιμές.

```
1 Stats=regionprops(I6,"Area","EquivDiameter","Orientation","Area","Perimeter","  
    Circularity");  
2 L=length(Stats);  
3  
4 for u=1:L  
5     Area(u)=Stats(u).Area;  
6     EquivDiameter(u)=Stats(u).EquivDiameter;  
7     ...  
8 end
```

Στο σχήμα 5 φαίνονται τα παραπάνω χαρακτηριστικά για τις ($3 \cdot 300 = 900$) φωτογραφίες.

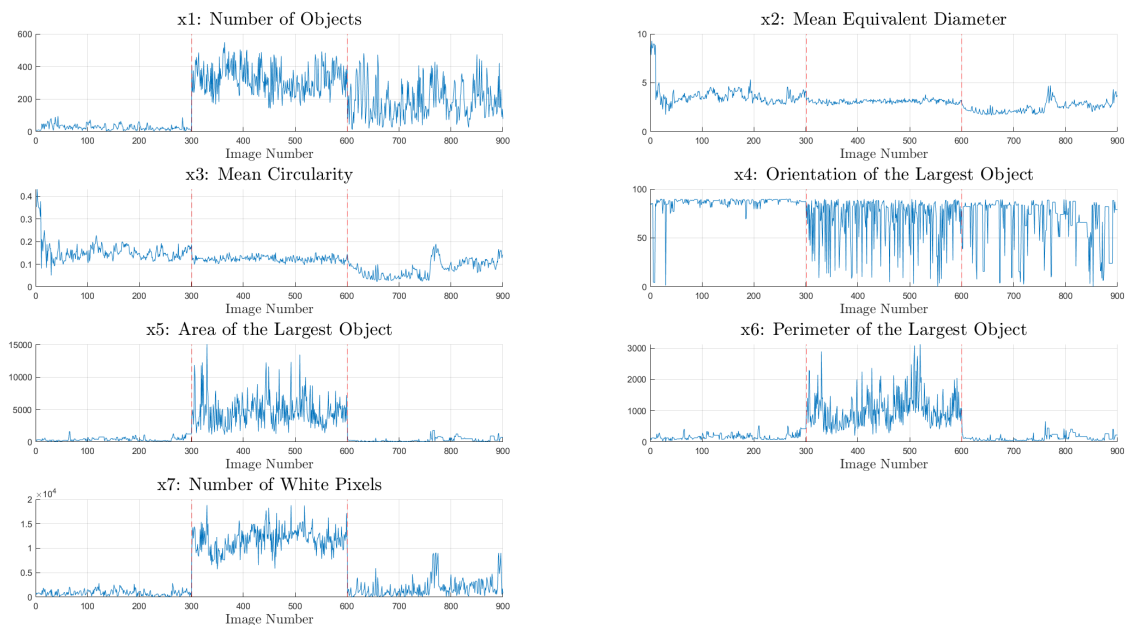


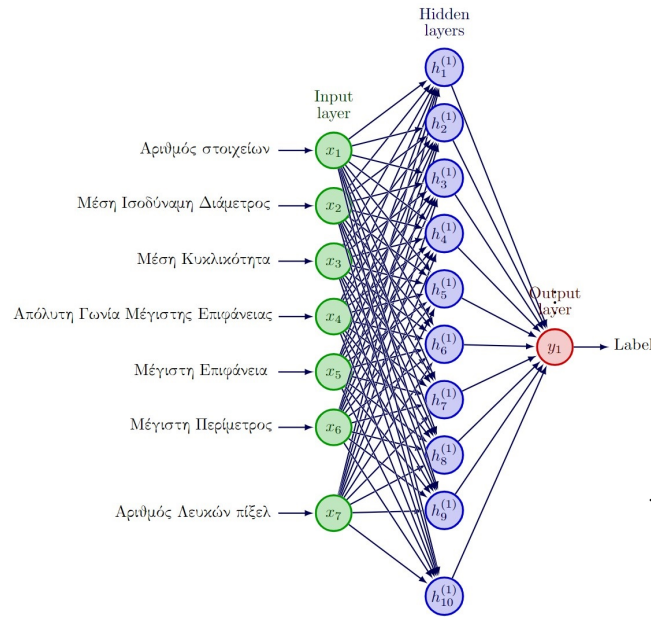
Figure 5: Διαγράμματα με τις πιθανές εισόδους του ΤΝΔ

4. Ανάπτυξη ΤΝΔ

Αρχικά αναφέρεται, πως το ποσοστό των training samples είναι 20% από Κ'ΑΘΕ κατηγορία ελαττώματος, ώστε να μην υπάρχει bias σε κάποια κατηγορία. Η κατηγορία των ελαττωμάτων είναι γνώστη και λαμβάνει τις τιμές (1=εγκλείσματα, 2=κηλίδες, 3=στικτή επιφάνεια). Τέλος σημειώνεται, πως γίνεται κανονικοποίησης των εισόδων.

4.1 Αρχικό Data Set

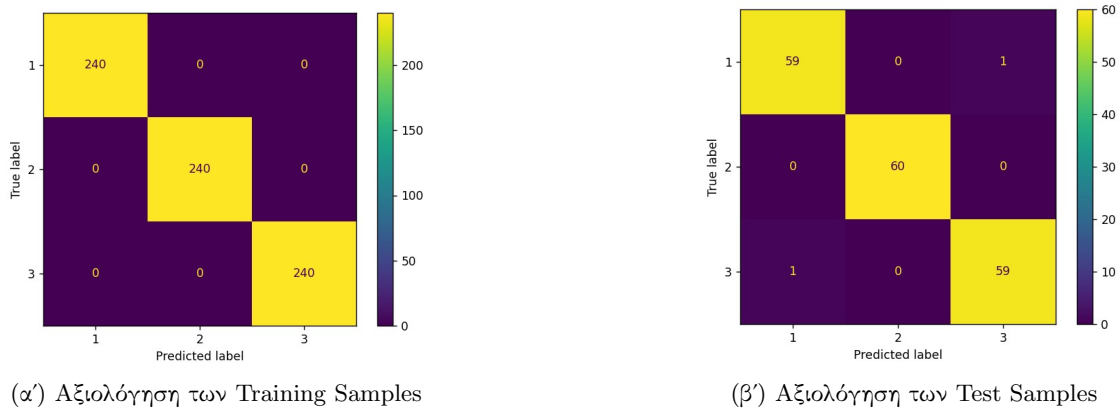
Στο σχήμα 6 παρουσιάζεται το ΤΝΔ και με τις επτά (7) εισόδους που αναφέρθηκαν.



Σχήμα 6: Τεχνητό Νευρωνικό Δίκτυο κατηγοριοποίησης ελαττωμάτων με τις αρχικές εισόδους

4.2 Αξιολόγηση αρχικού Data Set

Ως κριτήριο αξιολόγησης της απόδοσης του ΤΝΔ επιλέχθηκε ο πίνακας σύγχυσης, ώστε να παρατηρείται και πια ήδη ελαττωμάτων συγγέει ο ταξινομητής, ώστε να τροποποιηθεί η ανάλυση εικόνας και ο τύπος των εισόδων με τρόπο που να αναδεικνύουν τις διαφορές αυτών των τύπων των ελαττωμάτων. Ο πίνακας σύγχυσης φαίνεται για το συγκεκριμένο ΤΝΔ, και για ένα συγκεκριμένο δείγμα training samples, φαίνεται στην εικόνα 7.



Σχήμα 7: Αξιολόγηση Αρχικού ΤΝΔ

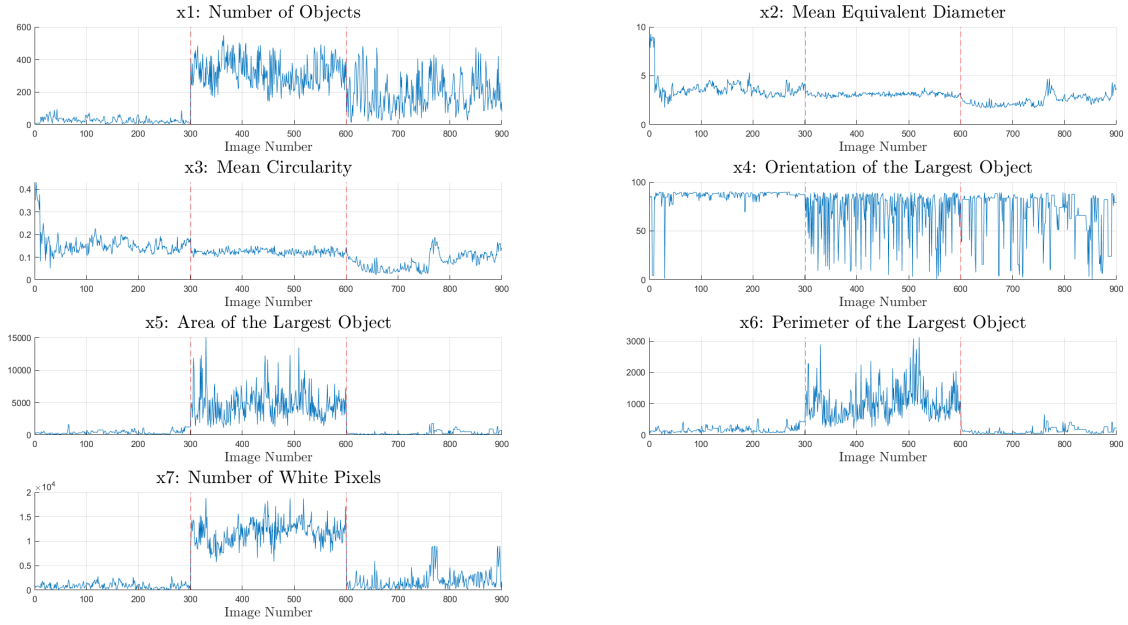
Ένα metric που συνοψίζει την συνολική απόδοση του ΤΝΔ είναι η ακρίβεια του μοντέλου (accuracy), η οποία ορίζεται από την παρακάτω σχέση 1. Εκπαιδεύτηκαν είκοσι (20) διαφορετικά ΤΝΔ με κάθε φορά διαφορετικά training samples ώστε να εξεταστεί και το robustness.

$$acc = \frac{\text{number of correct predictions}}{\text{number of predictions}}, \quad \bar{acc} = \frac{\sum_{i=1}^{i=20} acc_i}{20} \quad (1)$$

Αξιολογώντας τα training και test samples, προκύπτει $\bar{acc}_{train} = 0.999$, $\bar{acc}_{test} = 0.983$. Τα αποτελέσματα είναι σχεδόν άριστα, αλλά κρίνεται ότι μπορούν να αφαιρεθούν εισοδοί με πάλι όμοια απόδοση, ώστε να γίνει αποφυγή overfitting.

4.3 Τελικό Data Set

Ξαναπαρουσιάζεται το Σχήμα των χαρακτηριστικών της κάθε φωτογραφίας, ώστε να γίνει μια σημαντική παρατήρηση.

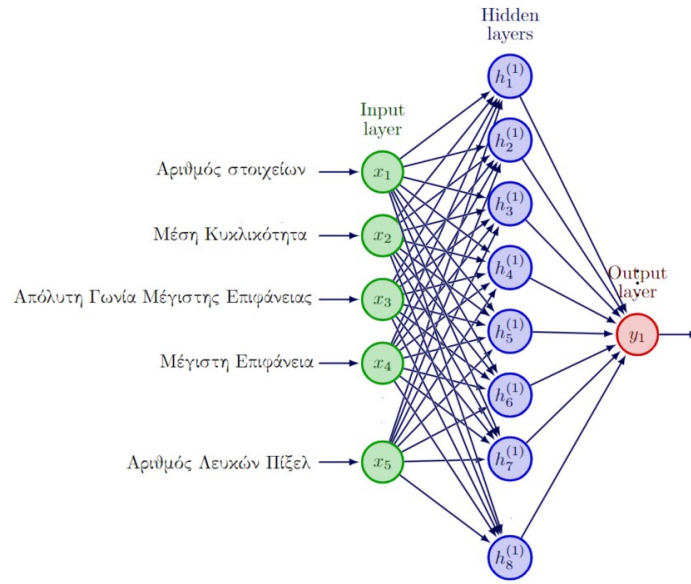


Σχήμα 8: Επίδραση αύξησης αντίθεσης

Μετά από δοκιμές, παρατηρήθηκε πως οι μεταβλητές x_2, x_6 δεν προσφέρουν αύξηση στην απόδοση του ΤΝΔ, επομένως και αφαιρούνται. Βλέποντας το παραπάνω σχήμα, μπορεί να ειπωθεί, πως η μεταβλητή x_2 δεν έχει κάποια σημασία στην τιμή της κατηγοριοποίησης. Επιπλέον η τιμή x_6 μπορεί να θεωρηθεί πως είναι όμοια με την τιμή x_5 . Επομένως στο τελικό data set υπάρχουν οι παρακάτω εισοδοί.

- Αριθμός στοιχείων (Objects)
- Μέση Κυκλικότητα
- Απόλυτη Γωνία Μέγιστης Επιφάνειας
- Μέγιστη Επιφάνεια
- Αριθμός Λευκών πίξελ (σαν Εμβαδόν)

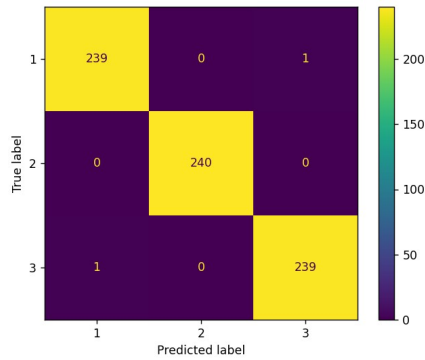
Το τελικό ΤΝΔ φαίνεται παρακάτω.



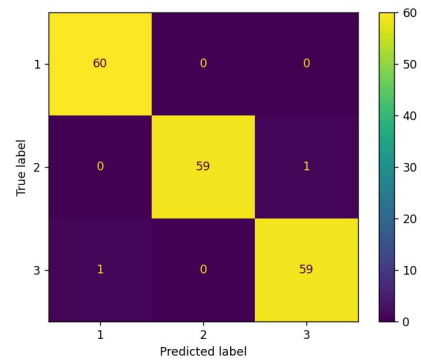
Σχήμα 9: Τελικό ΤΝΔ

4.4 Αξιολόγηση τελικού Data Set

Αξιολογώντας τα training και test samples, προκύπτει $\bar{acc}_{train} = 0.999$, $\bar{acc}_{test} = 0.985$. Τα αποτελέσματα όπως φαίνεται είναι πάλι άριστα και ίδια με τα προηγούμενα, αλλά με λιγότερες εισόδους



(α') Αξιολόγηση των Training Samples



(β') Αξιολόγηση των Test Samples

Σχήμα 10: Αξιολόγηση Τελικού ΤΝΔ

Παρουσιάζονται τα αποτελέσματα για πέντε (5) από τα είκοσι (20) ΤΝΔ που εκπαιδεύτηκαν.

Δοκιμές	1η	2η	3η	4η	5η
acc_{train}	1.0	1.0	1.0	1.0	1.0
acc_{test}	0.983	0.977	0.988	0.988	0.994

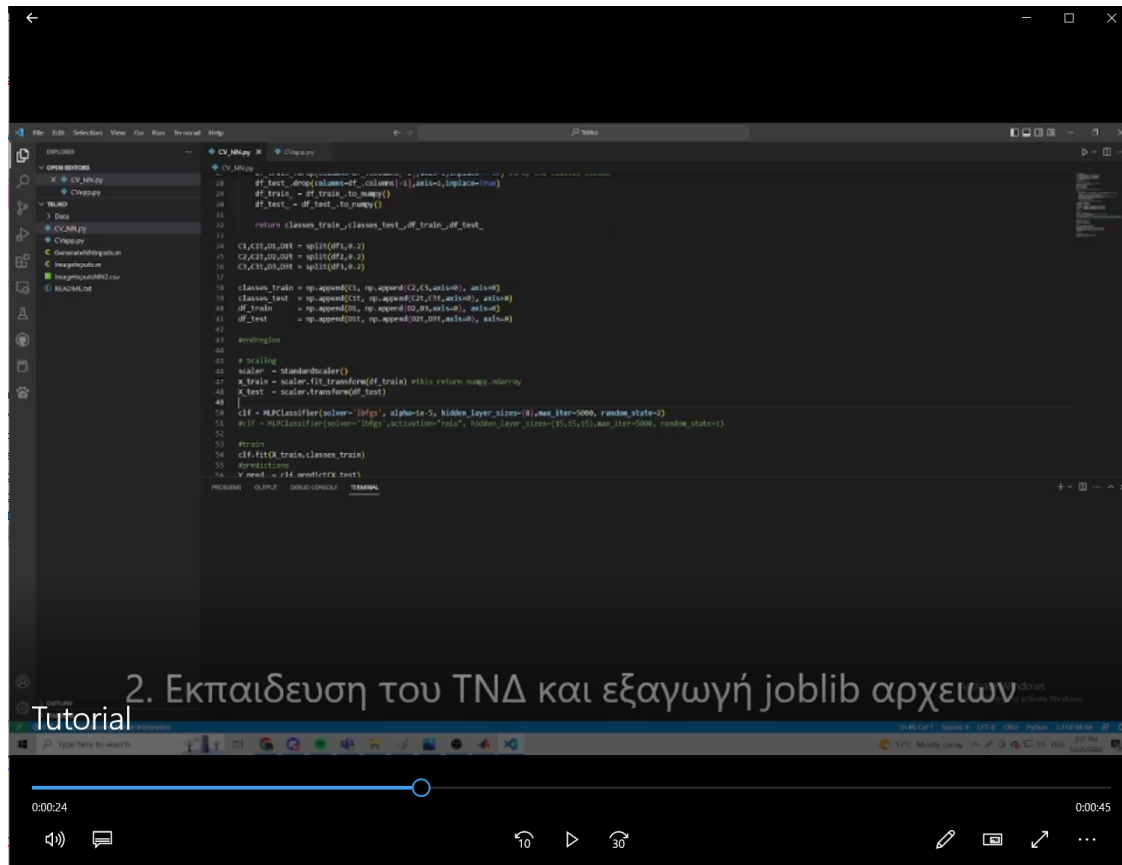
Πίνακας 1: Απόδοση εκπαιδευμένου ΤΝΔ με διαφορετικές αρχικοποιήσεις (random states)

Επομένως επαληθεύεται και η ανεξαρτησία την επιλογής των training samples. Γενικά σημειώνεται πως ο αριθμός και τα επίπεδα των νευρώνων δεν έχουν σημαντική σημασία στην απόδοση του ΤΝΔ. Αυτός είναι ο λόγος που δεν γίνεται βελτιστοποίηση.

5. Video

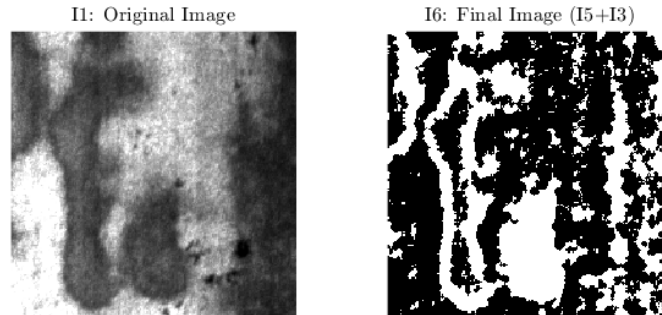
Στον τελικό φάκελο παράδοσης, υπάρχει βίντεο ενός λεπτού, στο οποίο φαίνονται οι διαδικασίες

- Παραγωγή σετ δεδομένων εκπαίδευσης (matlab)
- Εκπαίδευση νευρωνικού δικτύου και αξιολόγηση (python)
- Παραγωγή σετ δεδομένων για αξιολόγηση, συγκεκριμένης φωτογραφίας (matlab)
- Αξιολόγηση συγκεκριμένης φωτογραφίας (python)

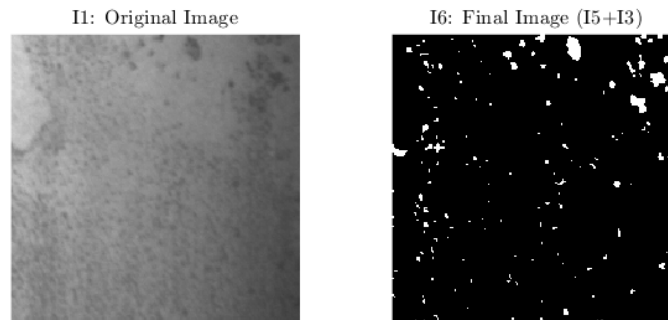


6. Παράρτημα

Παρουσιάζονται δύο 'κακές' επεξεργασμένες εικόνες. Όπως είδαμε, το ΤΝΔ έχει άριστη ακρίβεια, παρά το γεγονός ότι χάνεται πληροφορία μετά την επεξεργασία των εικόνων. Ο λόγος είναι πως τα κύρια χαρακτηριστικά (Μέγιστη Επιφάνεια, και αριθμός στοιχείων) δεν αλλιώνονται σημαντικά.



Σχήμα 11: Συνολική επεξεργασία 'κακής' εικόνας κηλίδας



Σχήμα 12: Συνολική επεξεργασία 'κακής' εικόνας στικτής επιφάνειας

Εδώ παρουσιάζεται η μορφή των σετ δεδομένων, που λαμβάνονται για την εκπαίδευση και αξιολόγηση των ΤΝΔ. Σημειώνεται πως χρησιμοποιούνται οι υπογραμμισμένες στήλες, καθώς καταλήξαμε σε μείωση των inputs.

#Objects	Mean EqDiam	Mean Circularity	Orientation of Largest	Largest Area	Largest Perim	#White Pixels	Label
11	8.903505332	0.352793071	85.21889489	353	127.997	901	1
7	7.848529737	0.361515587	84.08084333	201	66.5	517	1
9	9.285402281	0.43253219	85.21889489	353	127.997	800	1
9	8.617122831	0.381056889	85.19286152	352	127.352	705	1
406	3.181257413	0.15217291	70.09780333	3966	986.277	9719	2
254	3.393587076	0.118286005	83.13841086	4450	974.709	13121	2
242	3.342700192	0.12513099	64.15698643	5002	1406.711	14288	2
375	3.439594013	0.126904099	84.95345373	2699	604.429	12027	2
79	2.823734959	0.135155387	38.99586433	373	230.717	807	3
175	2.680212162	0.111145059	38.99586433	373	230.717	1437	3
222	2.633673654	0.111707847	82.18727327	227	117.604	1642	3
45	2.452905174	0.099488798	82.18727327	227	117.604	279	3

Σχήμα 13: Μορφή DataSet για την εκπαίδευση του ΤΝΔ