

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
Τμήμα Πληροφορικής



Εργασία Μαθήματος **Ανάλυση Εικόνας**

Αριθμός εργασίας – Τίτλος εργασίας	Υπολογιστική εργασία εξαμήνου
Όνομα φοιτητή	Καλλίγερος Αναστάσης
Αρ. Μητρώου	Π19253
Ημερομηνία παράδοσης	15/2/2023

Εκφώνηση εργασίας

Στόχος της συγκεκριμένης υπολογιστικής εργασίας είναι ανάπτυξη γραφοθεωρητικών αλγορίθμων για την ανάκτηση εικόνων με βάση το περιεχόμενο. Τα βασικά βήματα της προτεινόμενης αλγοριθμικής προσέγγισης έχουν ως ακολούθως:



1. Κανονικοποίηση Σειράς Κατάταξης (**Rank Normalization**)
2. Κατασκευή Υπεργράφου (**Hypergraph Construction**)
3. Υπολογισμός Ομοιότητας Υπερακμών (**Hyperedge Similarities**)
4. Υπολογισμός Καρτεσιανού Γινομένου μεταξύ των στοιχείων των Υπερακμών (**Cartesian Product of Hyperedge Elements**)
5. Υπολογισμός Ομοιότητας βάσει του κατασκευασμένου Υπεργράφου (**Hypergraph – Based Similarity**)

Λεπτομερής περιγραφή της παραπάνω αλγοριθμικής διαδικασίας μπορείτε να βρείτε στο άρθρο με τίτλο “**Multimedia Retrieval through Unsupervised Hypergraph-based Manifold Ranking**”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 28, NO. 12, DECEMBER 2019 . Το άρθρο θα αναρτηθεί στα έγγραφα του μαθήματος.

Ζητούμενα:

- i. Να παρουσιάσετε μια αναλυτική περιγραφή της υπολογιστικής διαδικασίας που παρουσιάζεται στο άρθρο ενσωματώνοντάς την στην τεκμηρίωση της εργασίας σας. **(20% του συνολικού βαθμού)**
- ii. Να αναπτύξετε κώδικα σε Matlab ή Python για την προγραμματιστική υλοποίηση των παραπάνω υπολογιστικών βημάτων. **(20% του συνολικού βαθμού)**
- iii. Το σύνολο των αντικειμενικών χαρακτηριστικών για την διανυσματική αναπαράσταση της κάθε εικόνας να εξαχθεί με την χρήση ενός προεκπαιδευμένου νευρωνικού δικτύου (**squeezenet, googlenet, resnet18, resnet50, resnet101**, κλπ.). Συγκεκριμένα, μπορείτε να χρησιμοποιήσετε ως αντικειμενικά χαρακτηριστικά της κάθε εικόνας έξοδο που αντιστοιχεί σε κάποιο από τα ενδιάμεσα κρυφά επίπεδα των προαναφερθέντων νευρωνικών δικτύων. **(20% του συνολικού βαθμού)**
- iv. Να παρουσιάσετε παραδείγματα της ορθής εκτέλεσης του κώδικά σας. Χαρακτηρίστε κάποιες από τις εικόνες της βάσης ως εικόνες στόχο (**target images**) και παρουσιάστε μια λίστα με τις συναφέστερες εικόνες της βάσης. **(20% του συνολικού βαθμού)**
- v. Προτείνετε μια συστηματική διαδικασία για την μέτρηση της ακρίβειας του συγκεκριμένου αλγορίθμου και παρουσιάστε τα αποτελέσματά της. **(20% του συνολικού βαθμού)**



ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1 Επίδειξη της λύσης.....	3
2 Βιβλιογραφικές Πηγές.....	10

1 Επίδειξη της λύσης

Η υλοποίηση της εργασίας βασίζεται στο άρθρο Multimedia Retrieval through Unsupervised Hypergraph-based Manifold Ranking. Το άρθρο αυτό αναλύει μία



μέθοδο κατάταξης ομοιότητας πολυμέσων με την χρήση ενός υπεργράφου(hypergraph). Το άρθρο αυτό περιγράφει μία μέθοδο που βαθμολογεί την ομοιότητα ανάμεσα σε πολλές εικόνες ή άλλης μορφής πολυμέσα δημιουργώντας έναν γράφο ο οποίος συγκεντρώνει συνολικά τις ομοιότητες ανάμεσα στις εικόνες. Η διαδικασία αυτή ονομάζεται Log-based Hypergraph of Ranking References. Η διαδικασία αυτή βρίσκει τις k όμοιες εικόνες και βαθμολογεί την ομοιότητα των εικόνων με τιμές από το 1 έως το 0, με το 1 να είναι η πιο όμοια φωτογραφία και στο παράδειγμα μας η ίδια η εικόνα και στην συνέχεια φωτογραφίες που παρουσιάζουν ομοιότητα με την αρχική. Ο λογάριθμος είναι με βάση το k και με βάση την διάταξη της απόστασης ανάμεσα στα αντικείμενα χαρακτηριστικά με μεγαλύτερη τιμή να ανατίθεται σε αντικείμενο που παρουσιάζει μικρότερη απόσταση.

Πιο ειδικά η διαδικασία αποτελείται από 5 διαφορετικά βήματα. Αρχικά παράγουμε αντικειμενικά χαρακτηριστικά από τα πολυμέσα με κάποια μέθοδο και για κάθε αντικείμενο οι υπολογίζεται η απόσταση (ευκλίδεια) με κάθε άλλο αντικείμενο oj . Στην συνέχεια κανονικοποιούμε με αμοιβαίο τρόπο την απόσταση του στοιχείου oi με oj με την απόσταση ανάμεσα στα σημεία oj με oi . Αυτή η διαδικασία παράγει το κανονικοποιημένο πίνακα που είναι το αποτέλεσμα του πρώτου βήματος (Rank Normalization).

Στην συνέχεια με βάση τον παραπάνω πίνακα των αποστάσεων δημιουργείται ο υπεργράφος. Ο υπεργράφος αυτός είναι μία επέκταση της θεωρίας των γράφων, όπου αντί για 1 όπου υπάρχει ακμή υπάρχει μία τιμή ανάμεσα στο 0 και στο 1 που υπολογίζεται με τον τύπο: $1 - \log kx$, όπου k το πλήθος των όμοιων φωτογραφιών και το x είναι μία τιμή ανάμεσα στο 1 και το k , όπου όσο πιο μικρή η τιμή του x , τόσο μικρότερη η απόσταση ανάμεσα στις 2 εικόνες. Αυτό το βήμα παράγει τον πίνακα H ο οποίος σε κάθε γραμμή έχει μία ακμή που αντιστοιχεί σε λίστα ομοιότητας μίας εικόνας με τις κορυφές που κάθε κορυφή αντιστοιχεί μία άλλη εικόνα. Τα σημεία που έχουν τις τιμές 0 σημαίνει ότι είναι εκτός της ομάδος εικόνων που παρουσιάζουν ομοιότητα με την αρχική.

Το επόμενο βήμα υπολογίζει την ομοιότητα ανάμεσα στις υπερακμές υπολογίζοντας το γινόμενο σημείο με σημείο $Sh * Sv$. Οι πίνακες Sh και Sv είναι τα γινόμενα HHT και HTH αντίστοιχα. Ο τελικός πίνακας S την κατά στοιχείο ομοιότητα των υπερακμών.

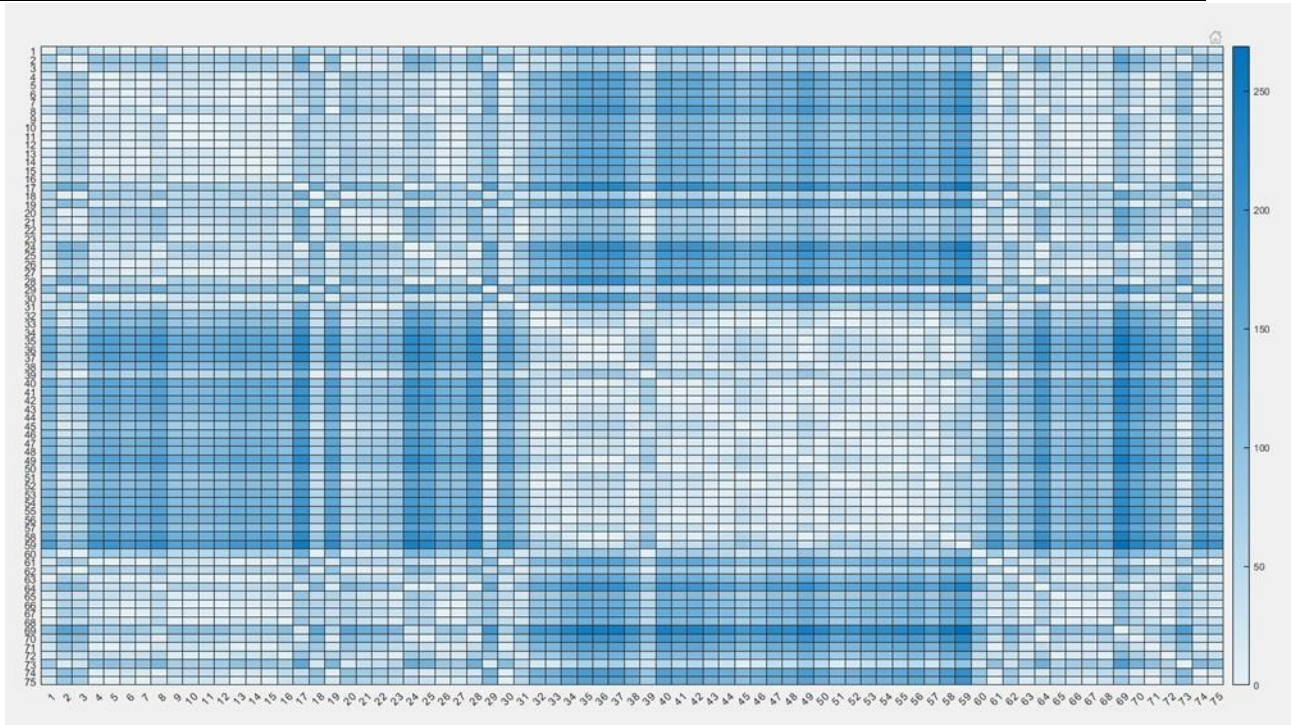


Στο προ-τελευταίο βήμα υπολογίζεται το καρτεσιανό γινόμενο των υπερακμών στον πίνακα C, όπου κάθε σημείο του πίνακα είναι το συνολικό βάρος της υπερακμής, επί την τιμή του πίνακα H σε αυτό το σημείο.

Η ομοιότητα στο τελευταίο βήμα υπολογίζεται με βάση το ανα σημείο γινόμενο $C * S$ το οποίο παράγει τον πίνακα W. Αυτός ο πίνακας είναι κατά μία έννοια ένα διάνυσμα βαρών ανα γραμμή με το μεγαλύτερο βάρος να ανατίθεται στην εικόνα με την μεγαλύτερη ομοιότητα.

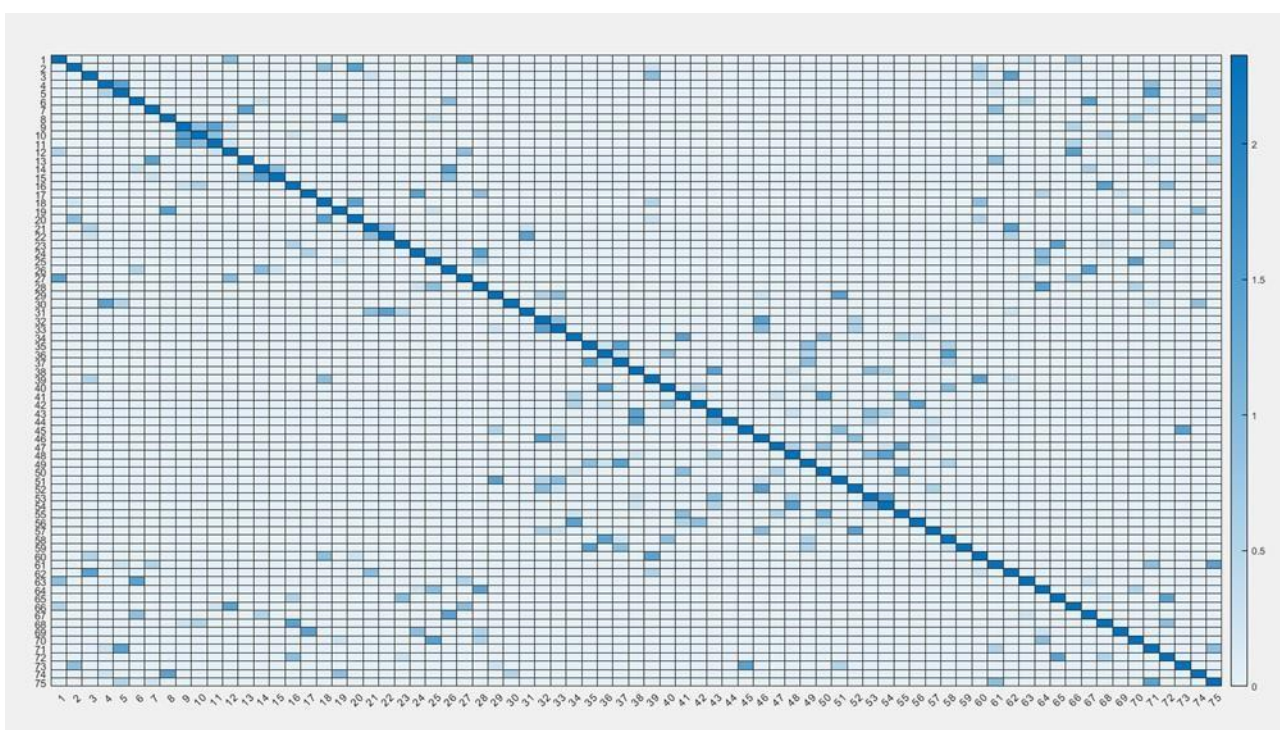
Ο κώδικας που βρίσκεται στο αρχείο `image_analysis.m` είναι κώδικας matlab και υλοποιεί το σύνολο της διαδικασίας. Ως είσοδο χρησιμοποιούμε το ενσωματωμένο Dataset MerchData του matlab που αποτελείται από 75 φωτογραφίες κάποιων αντικειμένων χωρισμένα σε 5 κατηγορίες αντικειμένων. Για κάθε μία από αυτές τις εικόνες υπολογίζουμε με την χρήση του `resnet18` το διάνυσμα αντικειμενικών χαρακτηριστικών το οποίο έχει μέγεθος 1×512 . Στην συνέχεια υπολογίζουμε τους πίνακες που παρουσιάσαμε στο παραπάνω βήμα, με βάση την διαδικασία που περιγράφηκε. Τέλος παρουσιάζουμε τις 4 πιο όποιες εικόνες για 4 τυχαίες εικόνες του Dataset.

Παρακάτω θα παρατεθούν εικόνες της μορφής heatmap για τους πίνακες του βήματος 1 και 5 τα οποία οπτικοποιούν τις σχέσεις ανάμεσα στις εικόνες.



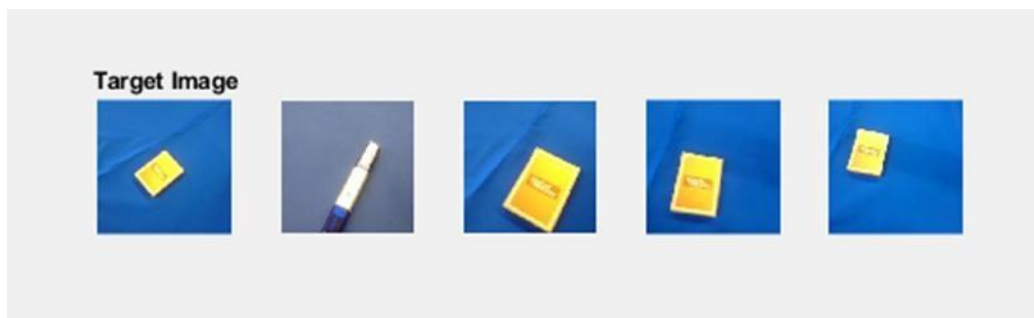
Heatmap του πίνακα που παρουσιάζει το Rank Normalization

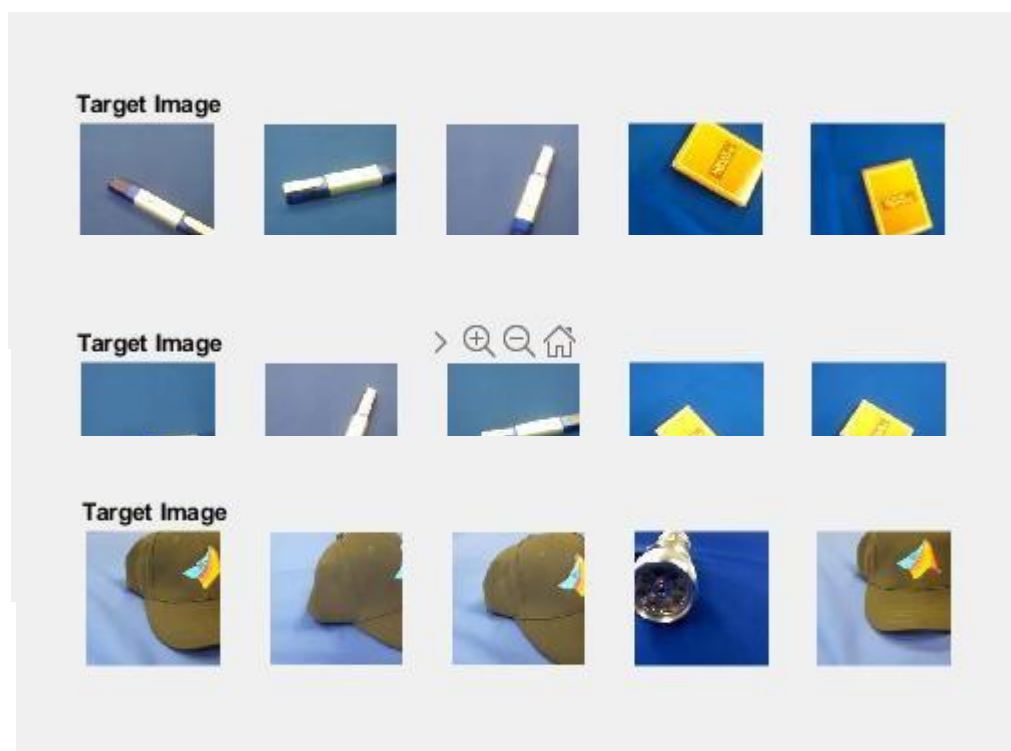
Το παραπάνω Heatmap οπτικοποιεί την απόσταση ανάμεσα στις εικόνες. Στο πιο σκούρο χρώμα αντιστοιχεί μεγαλύτερη απόσταση και παρατηρούμε ότι τα σημεία όπου $i=j$ η απόσταση είναι 0 και αυτό συμβαίνει καθώς εξετάζεται η απόσταση ενός διανύσματος με τον εαυτό του που προφανώς αντιστοιχεί στην τιμή 0.



Το heatmap του πίνακα W

Το παραπάνω Heatmap αντιστοιχεί στον πίνακα της τελικής ομοιότητας W το οποίο αντιστοιχεί μεγαλύτερη ομοιότητα σε μεγαλύτερη τιμή. Όπως και παραπάνω μπορεί να παρατηρηθεί ότι τα στοιχεία που $i=j$ έχουν την ίδια τιμή, αυτή την φορά η τιμή είναι η μέγιστη της κάθε γραμμής και αυτό συμβαίνει επειδή το αντικείμενο i συγκρίνεται με το αντικείμενο i άρα παρουσιάζει την μέγιστη ομοιότητα. Με βάση τον πίνακα W παρατίθενται οι 5 πιο όμοιες εικόνες με την αρχική για 4 τυχαίες κάθε φορά εικόνες. Παρακάτω θα παρατεθούν κάποια τέτοια γραφήματα.







2 Βιβλιογραφικές Πηγές

- **“Multimedia Retrieval through Unsupervised Hypergraph-based Manifold Ranking”**, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 28, NO. 12, DECEMBER 2019