

Μάθημα: Ψηφιακή επεξεργασία εικόνας και εφαρμογές

## Εργασία #2: Ανίχνευση ακμών, συμπίεση και εξαγωγή χαρακτηριστικών

Ημερομηνία παράδοσης: 20/06/2025

**Άσκηση #1, Ανίχνευση ακμών:** Θεωρήστε τις εικόνες αποχρώσεων του γκρι (grayscale image) “leaf.jpg”, “X-ray.jpeg” και “parking-lot.jpg”.

### Ζητούμενα:

- A. Σχεδιάστε, περιγράψτε και εφαρμόστε 3 διαφορετικές μεθόδους για την ανίχνευση των κύριων ακμών κάθε εικόνας (δηλαδή να δοθεί έμφαση στην ανίχνευση των ορίων και των περιγραμμάτων των εικονιζόμενων αντικειμένων).
- B. Για κάθε εικόνα, τυπώστε την αρχική εκδοχή της και τις 3 βέλτιστες δυαδικές μάσκες ανίχνευσης ακμών κάθε μεθόδου. Σχολιάστε τα αποτελέσματα.

Για όλες τις περιπτώσεις ανίχνευσης ακμών, τυπώστε μόνο τη (βέλτιστη) δυαδική μάσκα που προκύπτει από κάθε μέθοδο και εξηγήστε λεπτομερώς τη διαδικασία σχεδίασης (π.χ. επιλογή μεθόδου, καθορισμός παραμέτρων, κτλ.).

**Άσκηση #2, Κωδικοποίηση Shannon-Fano:** Θεωρήστε τις εικόνες αποχρώσεων του γκρι “airplane.jpg” και “bridge.jpg”.

### Ζητούμενα:

- A. Υλοποιήστε και εφαρμόστε τη μέθοδο συμπίεσης-κωδικοποίησης Shannon-Fano για τις εικόνες “airplane.jpg” και “bridge.jpg”. Να μη γίνει χρήση έτοιμων συναρτήσεων-υλοποιήσεων της μεθόδου.
- B. Για κάθε εικόνα παραθέστε: 1) Τις κωδικές λέξεις που προκύπτουν (αντιστοιχία τιμών φωτεινότητας με δυαδικές κωδικές λέξεις), 2) Το μέσο μήκος κωδικής λέξης, 3) Την τιμή της εντροπίας, και 4) Το λόγο συμπίεσης που επιτυγχάνεται.
- Γ. Για κάθε εικόνα, υπολογίστε τον αντίστοιχο λόγο συμπίεσης που επιτυγχάνεται με χρήση της μεθόδου συμπίεσης-κωδικοποίησης Huffman. Συγκρίνετε και σχολιάστε τα αποτελέσματα.

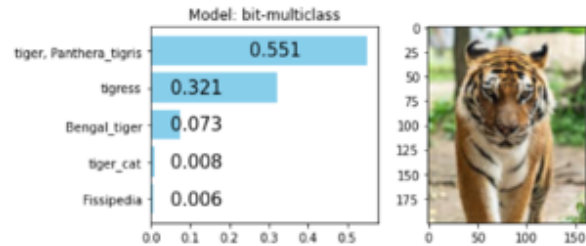
**Άσκηση #3, Εξαγωγή χαρακτηριστικών για ταξινόμηση (classification) εικόνων:** Θεωρήστε το δημόσιο σύνολο εικόνων ‘Animal image for Recognition’<sup>1</sup>, το οποίο περιέχει εικόνες από ζώα που ανήκουν στις ακόλουθες 4 κατηγορίες: α) Nilgai, β) Horse, γ) Cow, και δ) Water Buffalos.

### Ζητούμενα:

- A. Θεωρήστε τις εικόνες που βρίσκονται στα ακόλουθα υπο-σύνολα (φάκελοι): ‘chcek\chcek’ και ‘test\test’. Αφού τις μετατρέψετε σε εικόνες αποχρώσεων του γκρι, εξάγετε και αποθηκεύστε για κάθε εικόνα τα ακόλουθα καθολικά (global) χαρακτηριστικά που την περιγράφουν-αναπαριστούν στο σύνολό της: 1) Πίνακας συνεμφάνισης, 2) Χαρακτηριστικό LBP, και γ) Χαρακτηριστικό HoG.

<sup>1</sup> <https://www.kaggle.com/datasets/mdismielhossenabir/animal-image-for-recognition>

- B. Μετασχηματίστε κάθε ένα (από τα παραπάνω 3) 2Δ χαρακτηριστικό (που εξάγεται για κάθε εικόνα) σε ένα αντίστοιχο 1Δ διάνυσμα-χαρακτηριστικό, στοιβάζοντας (concatenation) διαδοχικά κάθε γραμμή του 2Δ χαρακτηριστικού στο 1Δ διάνυσμα (ΠΡΟΣΟΧΗ: το πλήθος των τιμών του χαρακτηριστικού στη 1Δ και τη 2Δ εκδοχή του πρέπει να είναι ακριβώς το ίδιο και η μετατροπή από 2Δ σε 1Δ να γίνεται με τον ίδιο ακριβώς τρόπο για όλες τις εικόνες και τα χαρακτηριστικά).
- Γ. Θεωρήστε ως αντιπροσωπευτικά δείγματα-πρότυπα για κάθε κλάση τις ακόλουθες εικόνες: 'cow\_1001.jpeg', 'horse\_1001.jpeg', 'nilgai\_1001.jpeg' και 'buffelo\_1001.jpeg'. Θεωρήστε ως μέτρο ομοιότητας (similarity metric) 2 εικόνων την Ευκλείδεια απόσταση ανάμεσα στα αντίστοιχα χαρακτηριστικά τους (του ίδιου τύπου). Μικρότερες τιμές Ευκλείδειας απόστασης υποδηλώνουν εικόνες που ομοιάζουν οπτικά περισσότερο.
- Δ. Για κάθε διαφορετικό τύπο χαρακτηριστικού, ταξινομήστε κάθε εικόνα του υπο-συνόλου 'test\test' σε μια από τις 4 υποστηριζόμενες κλάσεις, βάση της μέγιστης ομοιότητας (ελάχιστης απόστασης) από τα αντιπροσωπευτικά δείγματα-πρότυπα που έχουν οριστεί στο (Γ).
- Ε. Για κάθε τύπο χαρακτηριστικού, υπολογίστε την ακρίβεια (accuracy) ταξινόμησης των εικόνων του υπο-συνόλου 'test\test', δηλαδή το ποσοστό των εικόνων που έχουν ταξινομηθεί ορθά. Παραθέστε τις (3) υπολογιζόμενες τιμές ακρίβειας και σχολιάστε τα αποτελέσματα.
- ΣΤ. Παραθέστε ενδεικτικά αποτελέσματα ταξινόμησης (1 εικόνα από κάθε κατηγορία και τη μετρική ομοιότητάς της με κάθε υποστηριζόμενη κλάση).



A: Παράδειγμα ταξινόμησης εικόνας

Στην αναφορά να περιληφθούν απαντήσεις μόνο για τα ζητούμενα (Ε) και (ΣΤ).

**Άσκηση #4, Εξαγωγή χαρακτηριστικών για ανάκτηση (retrieval) εικόνων:** Θεωρήστε τη γενική αρχιτεκτονική ενός συστήματος ανάκτησης εικόνων με βάση το περιεχόμενο (Content-Based Image Retrieval, CBIR) που φαίνεται στο παρακάτω σχήμα. Θεωρήστε το δημόσιο σύνολο εικόνων 'Chess Pieces Detection Images Dataset'<sup>2</sup>, το οποίο περιέχει εικόνες από πιόνια σκακιού που ανήκουν στις ακόλουθες 5 κατηγορίες: 1) Queen, 2) Rook, 3) Bishop, 4) Knight, και 5) Pawn.

#### Ζητούμενα:

- A. Για κάθε εικόνα του συνόλου δεδομένων (αφού τη μετατρέψετε σε εικόνα αποχρώσεων του γκρι), εξάγετε και αποθηκεύστε ένα καθολικό (global) χαρακτηριστικό σάκου οπτικών λέξεων (bag of visual words), το οποίο να στηρίζεται στον υπολογισμό τοπικών (local) χαρακτηριστικών SIFT.
- B. Θεωρήστε ως εικόνες αναζήτησης τα ακόλουθα αντιπροσωπευτικά δείγματα από κάθε κλάση: "Queen-Resized\00000000\_resized.jpg", "Queen-Resized\00000001\_resized.jpg", "Rook-resize\00000001\_resized.jpg", "Rook-resize\00000002\_resized.jpg", "bishop\_resized\00000000\_resized.jpg", "bishop\_resized\ 00000002\_resized.jpg", "knight-resize\00000001\_resized.jpg",

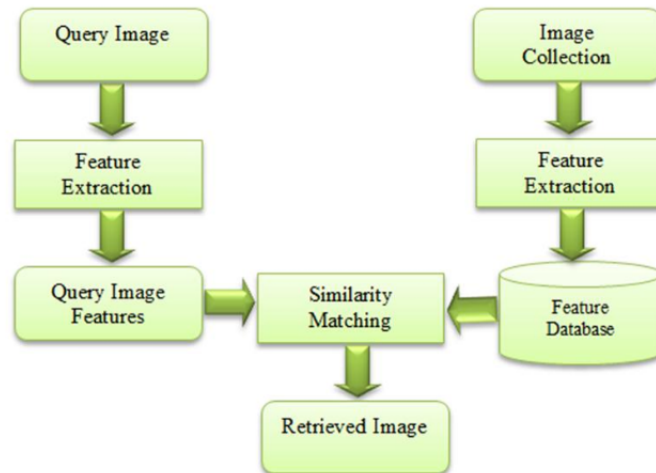
<sup>2</sup> <https://www.kaggle.com/datasets/anshulmehtakaggl/chess-pieces-detection-images-dataset>

“knight-resize\00000002\_resized.jpg”, “pawn\_resized\00000001\_resized.jpg”, και “pawn\_resized\00000002\_resized.jpg”. Θεωρήστε ως μέτρο ομοιότητας (similarity metric) 2 εικόνων την Ευκλείδεια απόσταση ανάμεσα στα αντίστοιχα χαρακτηριστικά τους. Μικρότερες τιμές Ευκλείδειας απόστασης υποδηλώνουν εικόνες που ομοιάζουν οπτικά περισσότερο.

Γ. Για τις εικόνες αναζήτησης και τη μετρική ομοιότητας που έχουν οριστεί στο (Β), υπολογίστε τις 10 πιο όμοιες εικόνες στο σύνολο δεδομένων (top-10 retrieval results).

Στα top-10 αποτελέσματα δε θα πρέπει να περιλαμβάνεται η εικόνα αναζήτησης (query image) κάθε φορά. Υπολογίστε τη συνολική μέση ακρίβεια ανάκτησης (retrieval accuracy), δηλαδή το ποσοστό των εικόνων που έχουν ανακτηθεί και ανήκουν στην ίδια κατηγορία με την εικόνα αναζήτησης κάθε φορά. Σχολιάστε τα αποτελέσματα.

Δ. Παραθέστε ενδεικτικά top-10 αποτελέσματα ανάκτησης (για 2-3 εικόνες αναζήτησης).



Στην αναφορά να περιληφθούν απαντήσεις μόνο για τα ζητούμενα (Α), (Γ) και (Δ). Για τα αποτελέσματα ανάκτησης, συμπεριλάβετε μόνο τα βέλτιστα αποτελέσματα που προκύπτουν μετά από πειραματισμό και εξηγήστε λεπτομερώς τη διαδικασία σχεδίασης (π.χ. καθορισμός παραμέτρων εξαγωγής χαρακτηριστικών σάκου οπτικών λέξεων, κτλ.).

### Σημειώσεις:

- Για την υλοποίηση των ασκήσεων μπορεί να γίνει χρήση των σχετικών βιβλιοθηκών της Python αναφορικά με image και data processing (π.χ. NumPy, scikit-image, matplotlib, κτλ.).
- Τα παραδοτέα της εργασίας είναι: 1) Μια αναφορά σε αρχείο pdf, όπου θα περιέχονται όλα τα ζητούμενα κάθε άσκησης (αποτελέσματα, σχολιασμός, περιγραφές, επεξηγήσεις, κτλ.), και 2) Αρχείο/α με τον κώδικα που έχει υλοποιηθεί για την επίλυση των ασκήσεων.