

# Pràctica 2: Etiquetatge

## Part 3: Millores i Avaluació

### Intel·ligència Artificial

Departament de Ciències de la Computació  
Universitat Autònoma de Barcelona

## 1 Introducció

En aquesta Part 3 de la pràctica continuem treballant per resoldre el problema d'etiquetatge automàtic d'imatges, combinarem el Kmeans per l'etiquetatge del color i el KNN per l'etiquetatge de la forma, per poder obtenir resultats similars al que es veu en la figura 1.



Figure 1: Objectius de la pràctica

A més a més, en aquesta Part 3 també fareu experiments per treure mesures de l'eficàcia dels mètodes i les millores que heu implementat i que implementareu en aquesta tercera entrega. És a dir, en aquesta entrega, a part de construir l'etiquetador, haureu d'analitzar-lo, usant les mesures d'anàlisi que implementareu, i n'haureu d'exposar els resultats i conclusions.

## 2 Fitxers necessaris

Farem servir els mateixos fitxers que ja vam descarregar, però ens centrarem sobretot en el fitxer `my_labeling.py`. Com que el KNN és un algorisme d'aprenentatge supervisat

necessitarem les imatges que hi ha al directori `Images`, i els conjunts d'aprenentatge i de test que hi tenim:

1. `images`: Carpeta que conté les bases de dades amb les imatges que utilitzarem.  
Dins d'aquesta carpeta trobareu:
  - (a) `test`: Conjunt d'imatges que farem servir com a conjunt de test.
  - (b) `train`: Conjunt d'imatges que utilitzarem com a conjunt d'entrenament per a la classificació de formes.
  - (c) `gt.json`: Arxiu amb la informació del *Ground-Truth* de les imatges.
  - (d) `gt_reduced`: Arxiu amb informació complementària sobre una part de les imatges que conformen el training set.
2. `test`: Carpeta amb fitxers per a fer els tests (En aquesta part 3 ja no els farem servir).
3. `codi`: A l'arrel del fitxer comprimit trobareu tot el codi necessari i els fitxers que haureu d'omplir (`my_labeling.py`). Els fitxers necessaris per treballar són:
  - (a) `utils.py`: Conté una sèrie de funcions que us poden ser útils principalment per a la part del KMeans.
  - (b) `TestCases_kmeans.py`: Arxiu amb el qual podreu comprovar si les funcions que programeu en el fitxer `Kmeans.py` donen el resultat esperat.
  - (c) `TestCases_knn.py`: Arxiu amb el qual podreu comprovar si les funcions que programeu en el fitxer `KNN.py` donen el resultat esperat.
  - (d) `Kmeans.py`: Arxiu on ja hi haureu programat les funcions necessàries per a implementar Kmeans per classificar el COLOR.
  - (e) `KNN.py`: Arxiu on haureu de programar les funcions necessàries per a implementar KNN per classificar el NOM de la peça de roba.
  - (f) `utils_data.py`: Conté una sèrie de funcions que us poden ser útils en aquesta última part.
  - (g) `my_labeling.py`: Arxiu on combinareu els dos mètodes d'etiquetatge i les vostres millores per a obtenir el nom finals de les imatges.

### 3 Funcions útils (`utils_data.py`)

El fitxer `utils_data.py` que us donem programat, conté un seguit de funcions que us serviran per fer l'anàlisi dels vostres resultats.

`read_dataset`: Funció que carrega les imatges de `train` i `test`, i el *Ground-Truth* de color i de forma. Aquesta funció ja es crida a l'inici del fitxer `my_labeling`.

**read\_extended\_dataset:** Funció que carrega un subconjunt d'imatges per les quals es té informació més detallada. Aquest *Ground-Truth* té com a objectiu que pugueu avaluar l'eficàcia del Kmeans d'una manera més objectiva, tenint en compte que en aquest moment encara no heu rebut formació en visió per computador, i per tant no podem demanar-vos que segmenteu la peça de roba. La informació que conté aquest és *Ground-Truth* és la següent:

1. Etiqueta del tipus de peça de roba de la imatge.
2. Coordenades d'una subfinestra de la imatge que només conté una part de la peça de roba que estem etiquetant. Intentant evitar el fons de la imatge, la pell de la persona, o altre peça de roba addicional que pot aparèixer. Per exemple, per una samarreta, la subfinestra contindrà la part central on es troben tots els colors etiquetables.
3. Etiquetes de color de la part de la peça que es mostra en la subimatge. Aquestes etiquetes han estat assignades per un humà, no per un algorisme.
4. Etiqueta de fons. Aquest camp ens diu si la subfinestra de la peça de roba que us donem conté fons (1) o no (0). En algunes peces de roba, com ara les sabates, és difícil obtenir una subfinestra que capturi tots els colors sense agafar alguna part del fons, la majoria de sabates tenen aquesta etiqueta a 1.

Aquesta funció ja es crida a l'inici del fitxer `my_labeling`.

**crop\_images:** Funció que donades unes imatges i les coordenades de les subfinestres, retalla les imatges inicials per tal que la seva nova mida sigui la definida per la subfinestra.

**visualize\_retrieval:** Funció que rep com a entrada un conjunt ordenat d'imatges, les quals han estat ordenades com a resultat d'una determinada cerca, i el nombre d'imatges que volem mostrar. La funció fa una visualització d'aquestes imatges. Com a paràmetres addicionals també pot rebre:

1. **info:** Llista de strings que es mostraran sobre cada una de les imatges donades pel primer paràmetre (per exemple: `['blue jeans', 'blue flip-flops', 'blue jeans']`).
2. **ok:** Llista d'elements de tipus `true` o `false`. Cada element fa referència a la imatge corresponent del primer paràmetre. El valor és `true` si la forma predita i el *Ground-Truth* coincideixen, retorna `false` en altre cas.
3. **title:** String que la funció posarà com a títol global a la figura.
4. **query:** Imatge respecte a la qual hem cercat les imatges més semblants (només és útil en el cas que s'hagi implementat aquesta funció).

**Plot3DCloud:** Funció que rep com a entrada la classe `Kmeans` i retorna la visualització del núvol de punts que conté la imatge dins de l'espai RGB. Cada punt apareix del color RGB corresponent al seu centroid.

**visualize\_k\_means:** Funció que rep com a entrada un objecte de la classe **Kmeans** i la mida de la imatge original, i retorna la visualització de la imatge original utilitzant només els colors dels centroides definits en l'objecte de la classe **Kmeans**.

## 4 Què s'ha de programar?

En aquesta tercera part de la pràctica programareu experiments per avaluar l'eficàcia dels vostres mètodes de classificació Kmeans i KNN, així com millores per a aquests.

Les funcions a programar es dividiran en els següents tipus:

1. **Funcions d'anàlisi qualitatiu:** Funcions que ens permeten avaluar d'una manera visual els nostres classificadors.
2. **Funcions d'anàlisi quantitatiu:** Funcions que ens permeten avaluar de manera numèrica els nostres classificadors. Com a mínim n'heu d'implementar dues.
3. **Millores als mètodes de classificació:** Canvis o noves funcionalitats en els mètodes Kmeans i KNN que heu implementat. Com a mínim haureu d'implementar tantes millores com membres té el grup.

A continuació exposem les funcions d'anàlisis que heu d'implementar, i alguns exemples de millores que podeu programar i incloure en l'anàlisi. Recordeu que podeu proposar i implementar altres millores, les quals seran valuades positivament.

### 4.1 Funcions d'anàlisi qualitatiu

L'objectiu d'aquesta anàlisi és avaluar de manera qualitativa la qualitat de les etiquetes assignades amb els nostres algorismes. Això ho farem a partir de fer cerques usant preguntes concretes, i podrem veure com varien els resultats de les cerques en funció dels paràmetres que fem servir en els nostres algorismes a l'hora d'etiquetar les imatges.

**Retrieval\_by\_color:** Funció que rep com entrada una llista d'imatges, les etiquetes que hem obtingut en aplicar l'algorisme Kmeans a aquestes imatges i la pregunta que fem per a una cerca concreta (això és, un string o una llista d'strings amb els colors que volem buscar). Retorna totes les imatges que contenen les etiquetes de la pregunta que fem. Aquesta funció pot ser millorada afegint un paràmetre d'entrada que contingui el percentatge de cada color que conté la imatge, i retorni les imatges ordenades segons el percentatge.

**Retrieval\_by\_shape:** Funció que rep com entrada una llista d'imatges, les etiquetes que hem obtingut en aplicar l'algorisme KNN a aquestes imatges i la pregunta que fem per a una cerca concreta (això és, un string definint la forma de roba que volem buscar). Retorna totes les imatges que contenen l'etiqueta de la pregunta que fem. Aquesta funció pot ser millorada afegint un paràmetre d'entrada que contingui el percentatge de K-neighbors amb l'etiqueta que busquem i retorni les imatges ordenades segons el percentatge.

**Retrieval\_combined:** Funció que rep com a entrada una llista d'imatges, les etiquetes de forma i les de color, una pregunta de forma i una pregunta de color. Retorna les imatges que coincideixen amb les dues preguntes, per exemple: Red Flip Flops. Com en les funcions anteriors, aquesta funció pot ser millorada introduint les dades de percentatge de color i forma de les etiquetes.

## 4.2 Funcions d'anàlisi quantitatiu

**Kmean\_statistics:** Funció que rep com a entrada la classe `Kmeans` amb un conjunt d'imatges i un valor, `Kmax`, que representa la màxima `K` que volem analitzar. Per cada valor des de `K=2` fins a `K=Kmax` executarà la funció `fit` i calcularà la `WCD`, el nombre d'iteracions i el temps que ha necessitat per convergir, etc. Finalment, farà una visualització amb aquestes dades.

**Get\_shape\_accuracy:** Funció que rep com a entrada les etiquetes que hem obtingut en aplicar el KNN i el *Ground-Truth* d'aquestes. Retorna el percentatge d'etiquetes correctes.

**Get\_color\_accuracy:** Funció que rep com a entrada les etiquetes que hem obtingut en aplicar el kmeans i el *Ground-Truth* d'aquestes. Retorna el percentatge d'etiquetes correctes. Cal tenir en compte que per a cada imatge podem tenir més d'una etiqueta, per tant, heu de pensar com puntuareu si la predicció i el *Ground-Truth* coincideixen parcialment. A la classe de teoria us varen donar algunes idees per mesurar la similitud entre aquests conjunts.

## 4.3 Millores als mètodes de classificació

**Inicializations de Kmeans:** Dins de la funció `init_centroids` de la classe `Kmeans`, heu d'implementar com a mínim dos mètodes d'inicialització a més a més del "first".

**Diferents heurístiques per BestK:** Durant la Part 1 d'aquesta pràctica vam programar l'heurística `WithinClassDistance` per a mesurar quina era la millor `K`, però hi ha altres heurístiques com per exemple, la *inter-class distance* o el coeficient de Fisher, els quals podeu implementar. Heu d'implementar com a mínim dues heurístiques diferents a més a més de la `WithinClassDistance`.

**Find\_BestK:** Durant la Part 1 d'aquesta pràctica vam decidir que per a trobar la millor `K` utilitzaríem la fórmula següent, posant un llindar al 20%:

$$\%DEC_k = 100 \frac{WCD_k}{WCD_{k-1}} \quad (1)$$

En aquesta Part 3 us demanem que proveu d'obtenir millors resultats jugant amb el valor del llindar, o utilitzant altres equacions per determinar quina és la millor `K`.

**Features for KNN:** Durant la Part 2 d'aquesta pràctica per determinar els veïns més propers vam utilitzar com a característiques la intensitat de cada píxel en la imatge transformada a escala de grisos, i com a distància, la distància euclídia. En aquesta Part 3 us demanem que proveu d'obtenir millors resultats provant diferents distàncies o bé modificant l'espai de característiques usat en la Part 2. Podeu reduir la mida de la imatge, cosa que farà que es redueixi el nombre de característiques usades, o podeu crear els vostres propis espais de característiques (per exemple: valor mitjà dels píxels, variància dels píxels, valor dels píxels superiors vs valor dels píxels inferiors, etc.).

## 5 Informe i presentació

Per a l'avaluació d'aquesta última part de la pràctica, haureu d'elaborar un informe i fer una presentació explicant els resultats que heu posat en aquest informe.

L'informe haurà de contenir les següents parts:

1. **Portada:** Ha de contenir el títol, l'assignatura, els noms i els NIUs dels participants.
2. **Introducció a la pràctica:** Explicació breu sobre el problema que la pràctica plantejava, l'algorisme Kmeans, i el KNN.
3. **Mètodes d'anàlisi implementats:** Explicació dels mètodes d'anàlisi que heu implementat per valorar com funciona el vostre Kmeans i KNN. Heu d'explicar com funcionen els mètodes qualitius i quantitatius que heu programat. Després heu de mostrar els resultats que obteniu sobre el programa base de Kmeans i KNN. Podeu posar totes les imatges, gràfiques o taules que facin falta per explicar tots els experiments ben justificats que heu fet.
4. **Millors sobre Kmeans i KNN:** Per a cada una de les millors o paràmetres que heu provat de modificar:
  - (a) Explicació de quin paràmetre esteu modificant.
  - (b) Resultats que obtenim quan modifiqueu aquest paràmetre o quan utilitzeu la millora proposada. (Utilitzant els mètodes d'anàlisi)
  - (c) Conclusió sobre els resultats. Comparació amb els resultats obtinguts sense millora i explicació de per què hem obtingut un resultat millor/pitjor o semblant.
5. **Conclusió global:** Conclusions globals de l'anàlisi elaborada, dels resultats que obté l'etiquetador implementat i possibles aplicacions del que heu desenvolupat.

La presentació seguirà la mateixa estructura que el vostre informe i durarà 15 minuts (10 minuts de presentació i 5 minuts de preguntes).

## 6 Entrega de la Part 3

Per a l'avaluació d'aquesta última part, s'hauran de fer dues entregues:

- **Un esborrany de l'anàlisi i de les millores** que penseu fer. S'ha d'entregar **abans del 12 de Maig a les 23:55h**. Abans d'aquest dia heu de **seleccionar una hora per a la Presentació Oral**. Si no agafeu hora, aleshores assumirem que no voleu entregar el project 2 i per tant deixeu el curs. Obrirem l'agenda al cv.uab.cat el dia **3 de Maig a les 20:00h**.
- **Entrega final** inclou codi, informe i presentació. S'ha d'entregar **2 dies abans de la Presentació Oral que heu seleccionat**, hi haurà 3 tasques d'entrega al cv.uab.cat pels següents dies:

abans del 20 de Maig a les 23:55h	si la teva Presentació Oral és dia 22 de Maig
abans del 21 de Maig a les 23:55h	si la teva Presentació Oral és dia 23 de Maig
abans del 25 de Maig a les 23:55h	si la teva Presentació Oral és dia 27 de Maig

Tot seguit expliquem com s'han d'organitzar aquestes entregues:

**Entrega Esborrany (12 de Maig)**. L'objectiu d'aquesta entrega és poder validar, abans de l'entrega final, si les avaluacions dels vostres etiquetadors que heu dissenyat van pel bon camí. Heu d'entregar al cv.uab.cat un fitxer PDF (2 pàgines com a màxim) que expliqui:

1. Les funcions d'anàlisi qualitatiu i quantitatiu que heu decidit implementar, tant les que ja tingueu implementades com les que no.
2. Una descripció dels experiments que heu dissenyat per a mostrar els resultats i les conclusions amb les eines d'anàlisi implementades.
3. Les millores o canvis que heu decidit implementar als algorismes del k-means i del kNN.
4. Una descripció de la visualització amb la qual teniu pensat presentar els resultats.
5. Un resum dels resultats preliminars que heu obtingut fins al moment.

**Entrega Final (20, 21 o 25 de Maig)**. El codi que hàgiu implementat i l'informe final. Heu de pujar al cv.uab.cat un fitxer zip amb els següents fitxers:

1. Codi (Kmeans.py, KNN.py i my\_labeling.py).
2. Informe en format PDF.
3. Presentació en format PDF o ppt.