PROJECTE FINAL DE CARRERA

Reconeixement per computador de característiques artístiques en les pintures

ÀREA: VISIÓ PER COMPUTADOR

PROFESSOR: DAVID MASIP RODÓ

M.Angeles García Iranzo

17/06/2015

ÍNDEX

1. RESUMEN....3
2. ABSTRACT....4
3. INTRODUCCIÓ....5
4. PLANIFICACIÓ....6
Planificació inicial....7
Planificació real....8
5. PROCÈS DE CLASSIFICACIÓ....8
5.1. ADQUISICIÓ D'IMATGES....9
5.2. TRANSFORMACIÓ....13
5.3. SEGMENTACIÓ....14
5.4. PARAMETRITZACIÓ....14
PERCENTATGE DE PÍXELS FOSCOS....15
COEFICIENT DE GRADIENT....16
NOMBRE DE MÀXIMS LOCALS I GLOBALS EN L'HISTOGRAMA DE LUMINÀNCIA 17
PERCENTATGE DE PÍXELS EN EL MÀXIM DE L'HISTOGRAMA....18
DESVIACIÓ MITJA DEL COLOR ENTRE LA IMATGE I ELS SEUS SUBCONJUNTS...19
DESVIACIÓ DISTRIBUCIÓ COLOR DE LA DISTRIBUCIÓ GAUSS....20
5.5. CLASSIFICACIÓ I PROVES....23
SVM (Support Vector Machines)....23
K-Nearest Neighbors Classifier....26
Naïve Bayes....28
AdaBoost....31
Decission Tree Classificador....31
K-Means....32
Perceptron....33
CROSS VALIDATION....34
6. CONCLUSIONS....35
7. MILLORES PER PRÒXIMS TREBALLS....35
8. APÈNDIX....36
EINES....36
GLOSSARI....37
ÍNDEX DE FIGURES....41
ÍNDEX DE TAULES....41
BIBLIOGRAFIA....41

1. RESUMEN

Aquest treball de fi de carrera té com objectiu principal aconseguir un sistema que realitzi un entrenament supervisat amb una base de dades d'obres pictòriques suficientment amplia i determinar la precisió de les seves prediccions sobre el moviment pictòric, estil, tècnica o autor a que pertany determinada obra en virtut de diferents característiques visuals de les pintures.

Hi ha un objectiu secundari intrínsec consistent a conèixer característiques i funcionament dels diferents models o algorismes de classificació/regressió i validació.

El sistema treballa segons una seqüència de fases:

- descarrega una mostra d'obres d'art, cerca, recupera i desa les característiques principals de l'obra en una base de dades.

- extreu les dades necessàries de la imatge, calcula els paràmetres adients per realitzar els entrenaments i prediccions i els desa a la base de dades.

- realitza els entrenaments, prediccions i informa de la precisió dels càlculs.

La consideració i decisió de quines dades de la imatge eren les més adients i quins paràmetres calcular van estar condicionats pels lluïts resultats que es van obtenir a «Classification and indexing of paintings based on art movements» en el que es va portar a terme un treball de similar objectiu amb una precisió del 90% . En aquell treball, no obstant, la classificació es va fer únicament en relació a tres moviments (cubisme, classicisme i impressionisme).

2. ABSTRACT

This Bachelor's Thesis has as a main objective to achieve a system which, by doing a supervised training with a wide paintings database, it determines the accuracy of their predictions about the art movement, tendency, technique or painting's author according to different visual characteristics of the artworks.

There is an essential secondary objective: it is to know the characteristics and operating mode of different models or algorithms of classification, regression and validation.

The system works according to a phase sequence:

- it downloads an artwork's sample, searches, recovers and saves the main characteristics of the work in a database.

- it extracts the necessary data of the image, calculates the appropriate parameters for training and predictions, and it saves them in the database.

- it makes training and predictions and reports the accuracy data.

The consideration and decision of which image data were the most appropriate and which parameter to calculate have been conditioned by the clear results obtained in «Classification and indexing of paintings based on art movements». That work accomplished a similar goal and obtained an accuracy of 99%. However, in that work the classification was made just relative to three art movements (cubism, classicism, and impressionism).

3. INTRODUCCIÓ

La visió per computador es pot considerar una branca de la Intel·ligència artificial, una tecnologia horitzontal, dirigida a captar imatges i extreure informació per part de les computadores. Al llarg de les últimes dècades s'ha incrementat el nombre i el tipus de tasques que els ordinadors són capaços de fer: sempre a partir d'unes dades extretes de diferents sensors, efectuen operacions i extreuen dades. En el camp de la imatge l'extracció de dades de les imatges inclou la captació de les propietats geomètriques (dimensió, forma, ubicació...) i de les dels materials (il·luminació, textura...) no és suficient, extreure dades no significa extreure informació, sinó que cal una feina acurada d'inferir significats. Aquí entra la màquina d'aprenentatge que mitjançant nombrosos algorismes i aproximacions permet dotar les pures dades de significat adient i diferent segons el tipus de tasca al que estigui enfocada. La visió es pot dir que consta de dues parts: captació de la imatge e interpretació. La primera part des de fa ja molt de temps que no té cap dificultat, ja que tenim la tecnologia adequada per fer-ne us, però la segona part extreure informació, interpretar-la, etc. és encara objecte d'estudi de no poques investigacions. La visió per computador és una disciplina enormement complexa que es divideix en camps pel seu estudi: segmentació, agrupació, reconeixement d'objectes estàtics, reconeixement d'objectes en moviment...

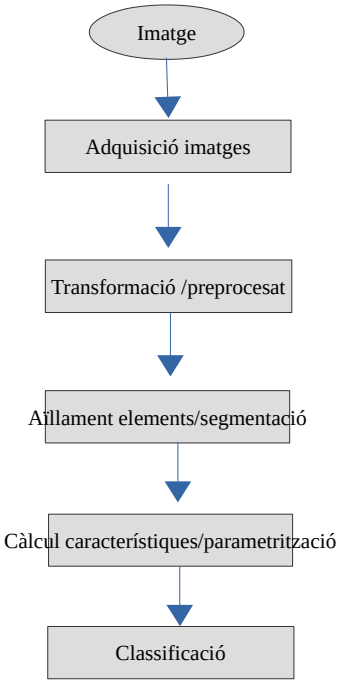
En molts dels casos és necessari que l'extracció i interpretació de la informació sigui immediata (temps de resposta limitat), que la majoria de procediments es realitzin en un espai de temps suficientment petit per poder resoldre aplicacions de la visió a sectors industrials o científics, cosa que és bastant difícil d'aconseguir si tractem amb les imatges tal com les captem. Per aconseguir tal requeriment és necessari simplificar les imatges, cal unes feines de preprocesament que es realitzen mitjançant algoritmes matemàtics (uns més adequats que altres segons la imatge, finalitat, etc.). D'altra banda s'han desenvolupat llibreries de visió per entorns estàndards.

L'aplicació de la tecnologia de la visió per computador és horitzontal i les aplicacions en producció (en sentit ampli, dins dels sectors de automòbil, metal·lúrgica, equips electrònics, i

un llarg etcètera), comunicacions, seguretat, biomedicina, xarxes socials, mobilitat, jocs, etc., són gairebé infinites.

Una bona part de la seva importància ve de la part que la industria la utilitza. Es va començar per processar imatges i reconèixer patrons aplicacions industrials.

El diagrama de flux mostra les diferents etapes de la visió artificial adaptat a la classificació i no al reconeixement d'objectes que seria motiu d'un altre treball.



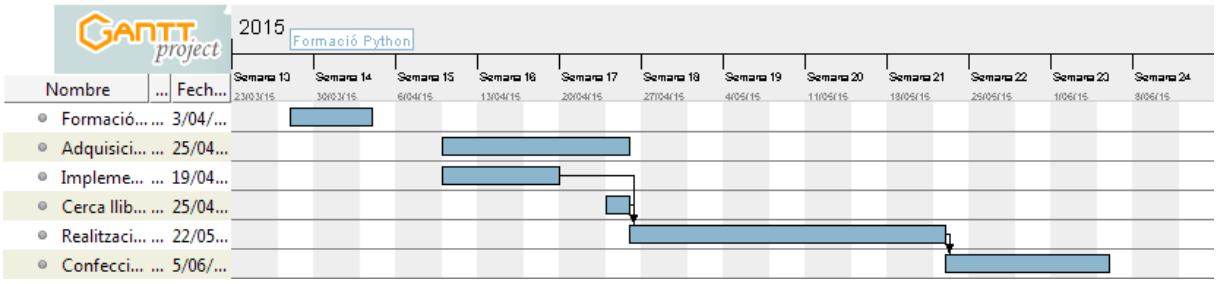
4. PLANIFICACIÓ

D'acord amb la mesura en crèdits del present projecte es van estimar que s'haurien de dedicar 225 h segons la planificació inicial que s'acompanya.

Planificació inicial:

Nom	Data Inici	Data Final	Recursos estimats
Formació Python Instal·lar Anaconda Instal·lar OpenCV Formació Python Formació llibreries	28/03/2015	03/04/2015	60 h
Adquisició base dades Decisió dades Adquisició base dades	10/04/2015	25/04/2015	20 h
Implementació codi normalització	10/04/2015	19/04/2015	60 h
Cerca llibreries entrenament/classificació Selecció Comparació Decisió	24/04/2015	25/04/2015	30 h
Realització experiments Estructurar sèries 10-Fold crossvalidation	26/04/2015	22/05/2015	30 h
Confecció memòria	23/05/2015	05/06/2015	25 h

Taula 1: Planificació inicial



Il·lustració 1: Planificació inicial, diagrama Gannt

Per raons de causa major el projecte va restar paralitzat des del dia 15/04/2015 fins el dia 14/05/2015. Les tasques es van reprendre i la finalització del projecte es va haver de postergar. Malgrat la paralització del projecte durant un mes, la quantitat d'hores totals

invertides ha estat superior a les previstes, en part per la llarga interrupció i en part pel total desconeixement dels requeriments de la base de dades.

Planificació real:

Nom	Data Inici	Data Final	Recursos reals
Formació Python Instal·lar Anaconda Instal·lar OpenCV Formació Python Formació llibreries	28/03/2015	03/04/2015	75 h
Adquisició base dades Decisió dades Adquisició base dades	10/04/2015	18/04/2015	30 h
Implementació codi	15/05/2015	25/05/2015	60 h
Cerca llibreries entrenament/classificació Selecció Comparació Decissió	15/05/2015	20/05/2015	30 h
Realització experiments Estructurar sèries 10-Fold crossvalidation	26/05/2015	04/06/2015	60 h
Confecció memòria	05/06/2015	16/06/2015	48 h

Taula 2: Planificació real

Lliurables: En finalitzar el projecte s'haurà de remetre la memòria del projecte i un fitxer comprimit contenint els arxius de codi.

5. PROCÈS DE CLASSIFICACIÓ

S'ha dedicat un capítol a cada una de les fases indicades al diagrama de flux corresponent a les etapes de la visió artificial aplicades al treball que ens interessa.
Dins de cada capítol podem trobar:

- Explicació de com s'ha executat la fase.
- Explicació de les diferents classes i mètodes utilitzats si escau.

- Explicació i raonament de les decisions preses.
- Proves realitzades si escau.
- Resultats si s'escau.

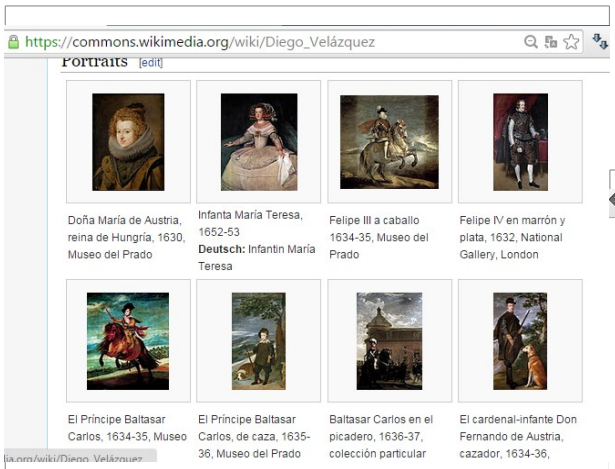
5.1. ADQUISICIÓ D'IMATGES

En realitat és la fase més important i de la que depèn l'èxit o el fracàs del projecte. En general és una etapa depenent de multitud de factors com són tècniques d'il·luminació (dins d'aquest es poden distingir nombrosos tipus com direccional, difusa, estructurada, etc.), fonts d'il·luminació, reflexió dels materials, diverses característiques de les càmeres, i molts factors més. En el cas que ens ocupa les imatges ja han estat adquirides i el nostre paper és només el de transmetre-les o descarregar-les via web.

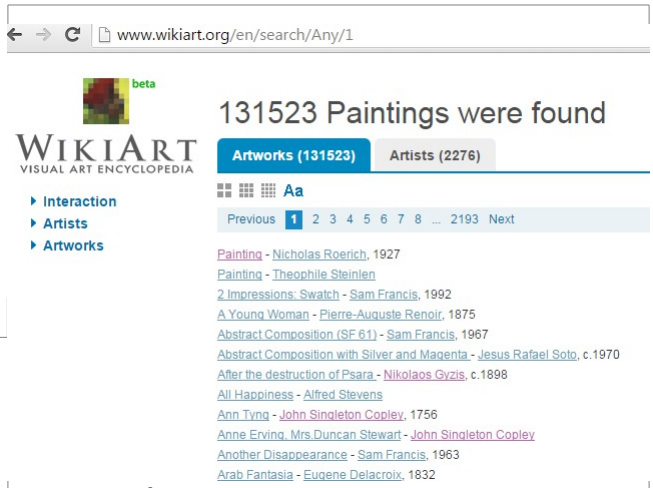
Un aspecte molt important a considerar quan es treballa amb un SVA i en concret en els processos d'adquisició i transmissió de les imatges és la problemàtica del soroll que porta a les imatges a la seva corrupció. És imprescindible en aquests casos poder eliminar aquest soroll (que pot ser de dos tipus segons en el moment en que es produeixi: gaussià en el moment d'adquisició i impulsiu en el moment de la transmissió). L'eliminació del soroll en cap cas ha d'implicar l'eliminació de textures o altres característiques. Per portar a terme aquesta correcció existeixen diferents algorismes basats en lògica difusa. En el nostre cas s'ha valorat la possible aplicació d'aquests algorismes però per simplificar i considerant la possibilitat d'aplicació en una futura ampliació del treball, s'ha vist prescindible.

En iniciar el plantejament del treball es van cercar diferents emplaçaments des d'on poder descarregar les pintures. Calia que fossin imatges amb llicència Creative Commons, que tinguessin una bona qualitat i que la seva descàrrega es pogués automatitzar de forma relativament senzilla. Les web que es van tenir en compte van ser *Google Images* (cercant per artista), *Wikimedia Commons*, *Web Gallery of art* (http://www.wga.hu/index1.html) i *Wikiart*. Es van dissenyar diferents scripts que intentaven realitzar la descàrrega però només es van

obtenir resultats satisfactoris amb *Wikimedia Commons* i amb *Wikiart* gràcies a l'estructura del site. D'aquests dos llocs es va escollir *Wikiart* ja que la seva pròpia estructura simplificava molt el codi necessari.



Il·lustració 2: Exemple Wikimedia Commons



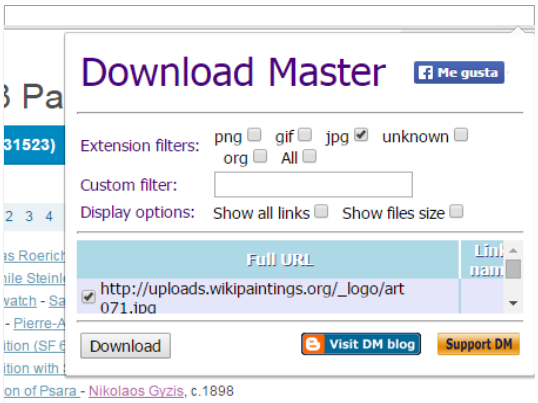
Il·lustració 3: Exemple Wikiart

Una vegada seleccionat aquest web calia decidir amb quina estructura es descarregarien i es guardarien les imatges. Es va crear en una primera fase una base de dades per autors (una carpeta per autor), seleccionant manualment els autors més rellevants dels diferents moviments. Aquesta base de dades disposava d'unes 2.400 imatges, però es va rebutjar ja que classificar amb posterioritat les obres (segons moviment, estil, gènere i tècnica) era una missió gairebé manual e impracticable.

La segona fase es va realitzar tenint ja en compte que la classificació de les obres es faria en el mateix moment de la descàrrega. El procediment seguit és :

- Creació d'una sèrie de llistes amb els enllaços que dona l'extensió de Chrome *Download Master* (també es va comprovar el funcionament d'*image collector*) a partir de la pàgina http://www.wikiart.org/en/search/Any/1 (obtenint una llista per

cada enllaç a un *artwork*). Realitzar diferents llistes, una per cada pàgina d'enllaços va ser una opció per poder controlar millor l'execució del programa.



Il·lustració 4: Extensió Download Master

- Un script llegia cada línia de llista substituint la part de l'enllaç que indicava descarregar una miniatura i realitzant altres canvis mitjançant expressions regulars.

  El guió convertia cada enllaç de la llista en un altre que enllaçava amb la pàgina on es mostren les característiques de la pintura.

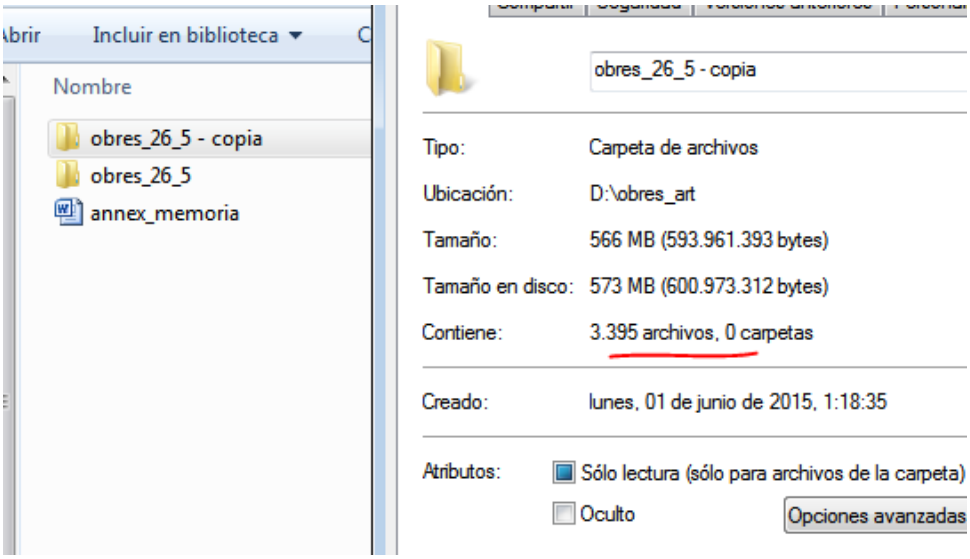
  Mitjançant el guió també es seleccionava el valor de les característiques *moviment pictòric* , *estil* , gènere, tècnica i *autor.*

  L'última funció del guió consistia en escriure en un document .csv (dades.csv) el nom de l'obra tal como es guarda a la carpeta, la tècnica, el moviment pictòric, l'estil i l'autor.

  Amb la iteració de totes les les llistes s'aconsegueix una base de dades de pintures de més de 3000 obres.

```
def descarregar_imatges():
    f = open(obrir,'r')
    URL=f.readline()


    while URL:
        URL= URL.encode('utf-8')
        print URL
       #DESCARREGA LA IMATGE
        data = urlopen(URL).read() # descarrega i guarda la imatge en una cadena
        file = StringIO(data) # tracta la cadena com un fitxer
        img = Image.open(file)# llegeix el fitxer i retorna imatge
        mo= re.sub("http:\/\/uploads[0-9]\.wikiart\.org\/images\/.*\/",'',URL)
        mo=mo.replace("\.jpg",'')
        smo=str(mo).replace('\n','')
        smo=str(mo).replace('\t','')
        print smo
        smo=smo.rstrip("\n")
        mol=re.sub("http:\/\/uploads[0-9]\.wikiart\.org\/images\/","http://www.wikiart.org/en/",URL)
        mol=mol.replace(".jpg",'')
        print mol
        img.save(smo)
        objectiu=crea_soup(mol)
        try:
            salvar_csv(objectiu,smo)
        except:
            print 'xxxxxxxxxxx'+obrir
        URL=f.readline()
    f.close()
```

Il·lustració 5: Exemple de modificació enllaços



Il·lustració 6: Base de dades d'imatges amb 3395 fitxers

```
obra,autor,moviment,tecnica,genere,u1,u2,u3,u4,u5,u6,u6
100-000-nudes-2006.jpg,David Michael Hinnebusch,Neo-Expressionism,acrylic,nude painting (nu),33.0728915663,0.12627
1-00-detail-868-149-893-746-1967-71-1965.jpg,Roman Opalka,Conceptual Art,synthetic polymer paint on canvas,abstract,(
1024-colours-1973.jpg,Gerhard Richter,Conceptual Art,oil,abstract,15.6634343434,0.119701041667,0.0546875,0.0204363636
104-cathedral-i-1951.jpg,Friedensreich Hundertwasser,Transautomatism,watercolor,figurative,14.9935730242,0.09515887
104-perimeter-series.jpg,Robert Mangold,Minimalism,pencil,abstract,0.0,0.00719917432494,0.015625,0.246291746673,15.
12-1956.jpg,Alfred Manessier,Art Informel,oil on canvas,abstract,32.1020833333,0.025493754069,0.125,0.0271979166667,1
13-xii-1981-forever-1985.jpg,Jacek Yerka,Surrealism,acrylic,symbolic painting,10.6491582492,0.0573133417508,0.28125,0.0
17-3-69-1969.jpg,John Hoyland,Color Field Painting,acrylic,abstract,0.0,0.00762515198937,0.0234375,0.10313510917,7.1407
17-mile-drive-1918.jpg,Guy Rose,Impressionism,oil,landscape,7.72470023981,0.0623012027878,0.125,0.0148201438849,14.
1918-in-petrograd-1920.jpg,Kuzma Petrov-Vodkin,Symbolism,oil,genre painting,22.7462915601,0.0432517383312,0.21093
1919-alarm-1934.jpg,Kuzma Petrov-Vodkin,Symbolism,oil,genre painting,16.8580128205,0.0681263772035,0.21875,0.0126
1924-first-abstract-painting-chelsea-1924.jpg,Ben Nicholson,Abstract Art,oil,abstract,2.29555555556,0.0189621354167,0.2
1932-painting-1932.jpg,Ben Nicholson,Abstract Art,oil,abstract,0.0535947712418,0.0393428308824,0.2109375,0.0154771241
1934-6-painting-still-life-1936.jpg,Ben Nicholson,Cubism,oil on canvas,still life,53.8231738035,0.0205084619018,0.0039062
1936-7-no-2-1936.jpg,Clyfford Still,Expressionism,oil on canvas,abstract,85.1440030558,0.0416246299656,0.00390625,0.155
1938-n-no-1-1938.jpg,Clyfford Still,Expressionism,oil on canvas,abstract,65.6492153868,0.120551334016,0.03125,0.0407554
```

Il·lustració 7: Exemple de la base de dades en .csv

5.2. TRANSFORMACIÓ

El format gràfic d'imatges obtingudes és el mateix en que es troben al lloc web, és a dir .jpg tot i que no és el més adequat pel motiu que és un format amb compressió i pèrdua. Les imatges s'han emmagatzemat en color ja que és un possible candidat a descriptor per realitzar la parametrització de la imatge. Com es veurà més endavant, algunes de les característiques de les imatges es van obtenir mitjançant la seva transformació a escala de grisos, encara que a la base de dades es van mantenir amb el seu format de color original.

Es va considerar també la necessitat redimensionar les imatges a priori però després de valorar els factors a favor i en contra (redimensionar implica disminuir la fidelitat de les obres, encara que es guanyaria molt en temps màquina) es va estimar que es podria sacrificar temps de procés a favor de la precisió de les obres. Però una vegada iniciat l'script i després d'unes hores de funcionament es va calcular que el programa trigaria més de 120h. en processar la base de dades. Aquest temps no es considera apropiat ja que durant tot aquest temps es molt factible que pugui succeir qualsevol tipus d'incident amb la màquina que faci perdre la feina realitzada. Conseqüentment es va prendre la decisió de redimensionar totes les

imatges i es va fer indiscriminadament al 50%, fet que possiblement va repercutir negativament en els resultats finals.

Si aquesta circumstància s'hagués conegut abans s'hauria pogut replantejar el programa, però en el moment que es va descobrir ja no havia temps suficient per reescriure el programa i executar-lo convenientment.

5.3. SEGMENTACIÓ

En visió artificial, sovint és necessari el reconeixement de patrons. Seleccionar en una imatge conjunts de píxels que tinguin en comú determinada qualitat i que formin un grup amb significat físic, per exemple tenint en compte textures, línies o colors, és segmentar. Aquests objectes permetran en molts casos obviar la resta de la imatge i centrar-se en aquells per arribar al reconeixement de patrons requerit.

L'enfocament que s'ha fet del projecte que ens ocupa no està condicionat a fer cap segmentació. Un futur treball podria dirigir-se cap a la classificació o reconeixement d'obres mitjançant elements extrets en virtut de segmentació, però no és l'actual cas.

5.4. PARAMETRITZACIÓ

Arribat aquest punt cal plantejar-se diferents qüestions:

Quina és la informació rellevant que s'ha d'extreure per resoldre el nostre problema?

Com extreure la informació?

Com representar la informació obtinguda?

Basat en el paper «Classification and indexing of paintings based on art movements (Oguz Icoglu i altres, 2004)» i donats els bons resultats mencionats a l'apartat Introducció, es va

escollir extreure i utilitzar els mateixos descriptors pel present projecte.

Al mencionat paper es mostra que els les pintures dels diferents moviments d'art es poden indexar en base a sis ressenyes diferents que s'expliquen a continuació:

μ1: Percentatge de colors foscos.

μ2: Coeficient de gradient (es calcula a partir del gradient de la imatge de la pintura).

μ3: Nombre de màxims locals i globals en l'histograma de luminància.

μ4: Percentatge de píxels que s'agrupen en el màxim de l'histograma.

μ5: Desviació mitjana del nivell de gris entre determinada quantitat de blocs iguals en els que es divideix la imatge.

μ6: Desviació de la distribució de nivells de gris de la distribució de Gauss.

PERCENTATGE DE PÍXELS FOSCOS

El percentatge de colors foscos o luminància té sentit en parlar d'imatges en color RGB i es calcula si seguim la recomanació **UIT-R BT.601-7** amb la següent equació:

$$Y = 0{,}299\ R + 0{,}587\ G + 0{,}114\ B$$

Sent R, G i B els valors dels colors primaris vermell, verd i blau.

Aquests valors aniran de 0 a 255, ja que cada píxel té una composició de color de tipus RGB com s'indica a continuació:

- color vermell: (255,0,0) ,té el màxim valor pel component R i 0 tant per G com per B.

- color verd: (0,255,0) ,té el màxim valor pel component G i 0 tant per R com per B.

- color blau: (0,0,255) ,té el màxim valor pel component B i 0 tant per R com per G.

Es determinarà un llindar a partir del qual considerarem que un píxel n'és fosc: valors de Y entre 0 i 64 corresponen a píxels foscos.

Així, el primer paràmetre que l'anomenarem μ_1 no dependrà de la mida de la imatge i el calcularem segons la següent fórmula:

$$\mu_1 = \frac{\text{píxels foscos}}{\text{píxels totals}}$$

COEFICIENT DE GRADIENT

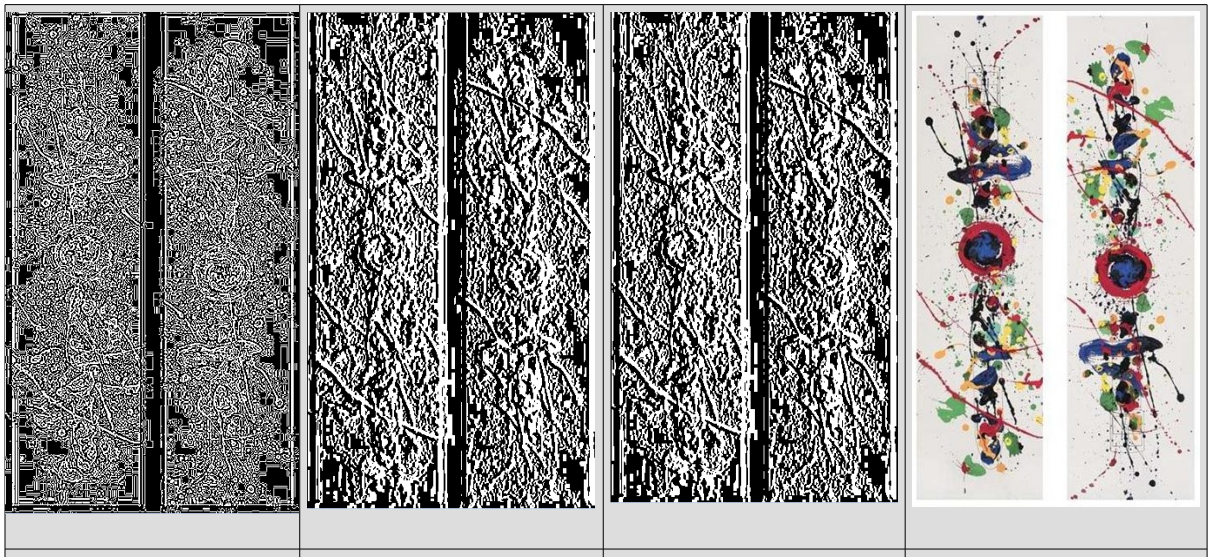
Considerarem les derivades parcials de la funció que dona el valor a cada punt, respecte de *x* i d' *y*. D'aquesta manera obtindríem el gradient, que és un vector. Si aquest vector el normalitzem resulta un escalar que en el nostre cas és el descriptor u_2 i que anomenarem coeficient de gradient.

$$\mu_2 = Normalitzat \sum_{i=0}^{r} \sum_{j=0}^{c} \sqrt{(f_{ij})_x^2 + (f_{ij})_y^2}$$

Encara que no s'ha utilitzat la llibreria cv2 per calcular el gradient, s'ha pogut constatar que el càlcul del gradient és molt senzill fent-ne ús d'aquella: l'operador *sobel* per cada punt de la imatge a processar genera el vector gradient i la norma en aquell punt. Per exemplificar el seu funcionament s'adjunta un petit script i les imatges tant del gradient com del gradient respecte d' *x* i respecte d'*y*.

```
import cv2

imatge = cv2.imread('D:/obres_art/obres_26_5/2-impressions-swatch-1992
des=cv2.Laplacian(imatge,cv2.CV_64F)

cv2.imshow("Gradient ",des)

print "premer  una tecla per continuar"
cv2.waitKey(0)

sobelx = cv2.Sobel(imatge,cv2.CV_64F,1,0,ksize=5)
cv2.imshow("Gradient a x ",sobelx)

print "premer  una tecla per continuar"
cv2.waitKey(0)
sobely = cv2.Sobel(imatge,cv2.CV_64F,0,1,ksize=5)

cv2.imshow("Gradient a y ",sobelx)

print "premer  una tecla per continuar"
cv2.waitKey(0)
```

Il·lustració 8: Exemple funcionament Sobel



Imatge gradient	Imatge gradient a x	Imatge gradient a y	Imatge original

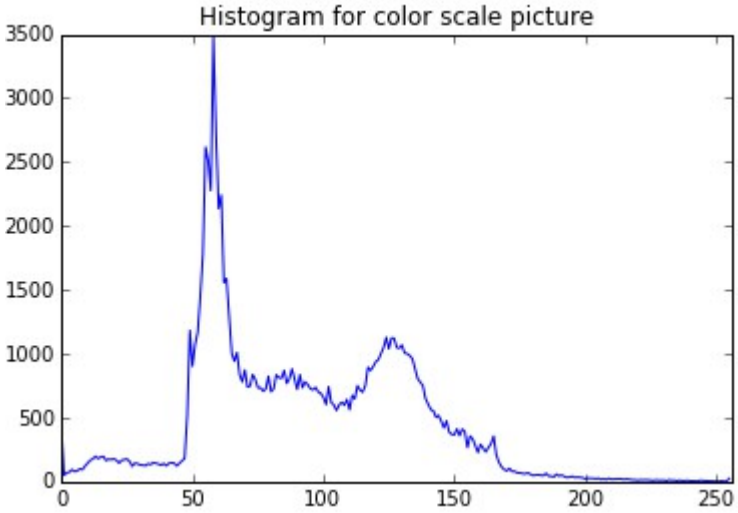
NOMBRE DE MÀXIMS LOCALS I GLOBALS EN L'HISTOGRAMA DE LUMINÀNCIA

Es calcula el nombre de màxims de la funció histograma que estan per sobre d'un llindar.

El llindar considerat és el valor resultant de dividir entre dos el valor màxim del histograma a tractar i s'ha pres de forma bastant poc científica i molt intuïtiva: hi ha moviments caracteritzats per tenir molts màxims sobre la part alta del diagrama (molts contrastos) i d'altres en que la majoria de l'histograma es troba per sota de la línia que divideix en dos parts iguals el diagrama (horitzontalment). El valor d'aquest descriptor és molt determinant d'un tipus de moviment pictòric o d'altre (segons les investigacions que es mencionen en el paper «Classification and indexing of paintings based on art movements (Oguz Icoglu i altres, 2004)».

$$\mu_3 = Nombre(|H(i) > llindar|)$$

Il·lustració 10: Histograma



A l'histograma superior que es correspon amb la fotografia de la dreta podem observar que hi ha poc contrastos: es pot veure que la major part de l'histograma es situa al quadrant inferior del diagrama.

Il·lustració 9: Imatge amb pocs contrastos

PERCENTATGE DE PÍXELS EN EL MÀXIM DE L'HISTOGRAMA

El següent descriptor compara les imatges considerant el més gran nombre de píxels que tenen el mateix color a la imatge (en percentatge). És un descriptor que encara que la informació que ofereix és diferent de l'anterior característica, està bastant relacionat.

$$\mu_4 = max_i |H(i)|$$

DESVIACIÓ MITJA DEL COLOR ENTRE LA IMATGE I ELS SEUS SUBCONJUNTS

Es calcula la mitja de color (escala grisos) de tota la imatge i després aquesta es divideix en nou blocs dels quals es calcula la mitja de color. La suma de les nou desviacions entre la mitja dels blocs i la mitja de la imatge és el nou descriptor.

$$\mu_5 = \sum_{i=1}^{9} (mitja_i - mitja) \quad \text{on mitja} \quad mitja = \frac{\sum_{i=1}^{r} \sum_{j=1}^{c} f_{ij}}{rc}$$

Es pot pensar en un valor alt de μ_5 degut a que la brillantor entre blocs és molt diferent, qüestió que podria donar-se en pintures per exemple impressionistes (i tot el contrari en cubistes).

A la llibreria cv2 existeix la possibilitat de calcular el valor de la mitja aritmètica de cada canal d'imatge (el següent codi retorna una tupla de quatre valors, dels que els tres primers són la mitja aritmètica dels tres canals B, G, R):

```
import cv2

imatge = cv2.imread('D:/p/box_in_scene.png',cv2.CV_LOAD_IMAGE_GRAYSCALE)
mitges = cv2.mean(imatge)
print mitges
```

El codi en el nostre projecte retornaria una tupla com la següent on veiem que només és significatiu el primer component, doncs treballem amb una imatge en escala de grisos:

```
(91.80233364869127, 0.0, 0.0, 0.0)
```

Fàcilment obtindríem la desviació estàndard per una imatge en escala de grisos:

```
import cv2

imatge = cv2.imread('D:/p/box_in_scene.png',cv2.CV_LOAD_IMAGE_GRAYSCALE)
mitges = cv2.mean(imatge)
mitges, desviacios = cv2.meanStdDev(imatge)
print '***MITGES***'
print mitges
print '***DESVIACIO STANDARD***'
print desviacios
```

```
***MITGES***
[[ 200.77107364]]
***DESVIACIO***
[[ 63.06591348]]
```

Il·lustració 11: Desviació estàndard amb cv2

Malgrat la facilitat d'ús d'aquest operador no es va utilitzar i els càlculs es generaren formulant mitjançant codi, ja que es va mantenir la forma com al document referenciat s'havia realitzat donats els bons resultats. Per tant, el que es va fer és realitzar una divisió de cada imatge en 9 blocs i sobre cada un d'ells es va calcular la mitja aritmètica i la desviació estàndard acumulada.

Posteriorment a la creació de la base de dades es va realitzar una prova aïllada sobre algunes obres per comprovar possibles diferències entre els càlculs fets al projecte i els que s'obtindrien amb la llibreria cv2 i efectivament s'hi van trobar discrepàncies: l'origen rau en la divisió de la imatge en els nou blocs, ja que gairebé gairebé sempre s'ignorava una petita última porció d'imatge. En una futura versió del projecte caldria incloure aquesta porció o utilitzar la citada llibreria.

DESVIACIÓ DISTRIBUCIÓ COLOR DE LA DISTRIBUCIÓ GAUSS

Diferents gèneres o moviments poden respondre de forma diferent a la distribució dels valors dels seus píxels (nivell grisos) respecte d'una campana de Gauss. Precisament aquest descriptor pretén quantificar aquesta desviació.

$$\mu_6 = \frac{\sum_{i=1}^{r}\sum_{j=1}^{c}\left(\frac{f_{ij}-mitja}{\sigma}\right)^3}{rc} \quad \text{on sigma} \quad \sigma = \frac{\sqrt{\sum_{i=1}^{r}\sum_{j=1}^{c}(f_{ij}-mitja)^2}}{rc-1}$$

(sigma és la desviació estàndard d'una població)

El pas següent a la formulació dels descriptors va ser calcular els sis descriptors per cada obra.

```
def crearmatriu():

    matriu_obres=[]
    #nombre de files que tindra el csv (nombre d'obres)
    longitud
    obra = [None]*longitud#files
    #for i in range(11):
        #obra[i] = [None]*11 #columnes

    reader=csv.reader(open('D:\obres_art\obres_26_5\dades.csv','rb'))


    for index,row in enumerate(reader):
        obra[index]=[row[0],row[1],row[2],row[3],row[4],row[5],row[6],row[7],row[8],row[9],row[10]]
        #creació de la matriu np
    for i in range(longitud):
        matriu_obres.append(np.array(obra[i]))
    for i in range(longitud):
        print matriu_obres[i][0]
        percentatge=get_pfoscos('D:/obres_art/obres_26_5/'+ matriu_obres[i][0])
        matriu_obres[i][5]=percentatge
        gradient=get_gradient('D:/obres_art/obres_26_5/'+ matriu_obres[i][0])
        matriu_obres[i][6]=gradient
        extrems=get_extremshist('D:/obres_art/obres_26_5/'+ matriu_obres[i][0])
        matriu_obres[i][7]=extrems
        maxhist=get_maxhist('D:/obres_art/obres_26_5/'+ matriu_obres[i][0])
        matriu_obres[i][8]=maxhist
        desvm=get_desviaciomitja('D:/obres_art/obres_26_5/'+ matriu_obres[i][0])
        matriu_obres[i][9]=desvm
        desv=get_desviacio('D:/obres_art/obres_26_5/'+ matriu_obres[i][0])
        matriu_obres[i][10]=desv
    arxi=open('D:\obres_art\obres_26_5\dadesf.csv','a')
    for i in range(longitud):
        for j in range(10):
            arxi.write(matriu_obres[i][j])
            arxi.write(',')
        arxi.write(matriu_obres[i][10])
        arxi.write('\n')
    arxi.close()
```

Il·lustració 12: Detall càlcul descriptors

El guió realitzava el càlcul dels descriptors i escrivia el resultat al fitxer .csv obtenint un resultat com el que es visualitza a la següent imatge:

```
obra,autor,moviment,tecnica,genere,u1,u2,u3,u4,u5,u6,u6
100-000-nudes-2006.jpg,David Michael Hinnebusch,Neo-Expressionism,acrylic,nude painting (nu),33.0728915663,0.12627
1-00-detail-868-149-893-746-1967-71-1965.jpg,Roman Opalka,Conceptual Art,synthetic polymer paint on canvas,abstract,(
1024-colours-1973.jpg,Gerhard Richter,Conceptual Art,oil,abstract,15.6634343434,0.119701041667,0.0546875,0.0204363636
104-cathedral-i-1951.jpg,Friedensreich Hundertwasser,Transautomatism,watercolor,figurative,14.9935730242,0.09515887
104-perimeter-series.jpg,Robert Mangold,Minimalism,pencil,abstract,0.0,0.00719917432494,0.015625,0.246291746673,15.
12-1956.jpg,Alfred Manessier,Art Informel,oil on canvas,abstract,32.1020833333,0.025493754069,0.125,0.0271979166667,1
13-xii-1981-forever-1985.jpg,Jacek Yerka,Surrealism,acrylic,symbolic painting,10.6491582492,0.0573133417508,0.28125,0.0
17-3-69-1969.jpg,John Hoyland,Color Field Painting,acrylic,abstract,0.0,0.00762515198937,0.0234375,0.10313510917,7.1407
17-mile-drive-1918.jpg,Guy Rose,Impressionism,oil,landscape,7.72470023981,0.0623012027878,0.125,0.0148201438849,14.
1918-in-petrograd-1920.jpg,Kuzma Petrov-Vodkin,Symbolism,oil,genre painting,22.7462915601,0.0432517383312,0.21093
1919-alarm-1934.jpg,Kuzma Petrov-Vodkin,Symbolism,oil,genre painting,16.8580128205,0.0681263772035,0.21875,0.0126
1924-first-abstract-painting-chelsea-1924.jpg,Ben Nicholson,Abstract Art,oil,abstract,2.29555555556,0.0189621354167,0.2
1932-painting-1932.jpg,Ben Nicholson,Abstract Art,oil,abstract,0.0535947712418,0.0393428308824,0.2109375,0.0154771241
1934-6-painting-still-life-1936.jpg,Ben Nicholson,Cubism,oil on canvas,still life,53.8231738035,0.0205084619018,0.0039062
1936-7-no-2-1936.jpg,Clyfford Still,Expressionism,oil on canvas,abstract,85.1440030558,0.0416246299656,0.00390625,0.155
1938-n-no-1-1938.jpg,Clyfford Still,Expressionism,oil on canvas,abstract,65.6492153868,0.120551334016,0.03125,0.0407554
```

Il·lustració 13: Base de dades en .csv

En disposar de la base de dades completa, amb els descriptors calculats, es va descobrir que havia imatges que podrien desvirtuar els resultats com litografies o escultures, i per tant es va intentar esborrar-les a ma. Més endavant es va sumar un problema afegit: de vegades (molt sovint) els programes donaven errors inexplicables els quals eren conseqüència de que per alguna raó la base de dades s'havia carregat malament i hi havia camps en blanc, així com alguns registres amb els valors dels atributs intercanviats. Arribat aquest punt es van plantejar tres opcions:

- tornar a carregar des de zero la base de dades.

- intentar recalcular únicament els descriptors.

- eliminar al màxim els registres que fessin nosa i continuar amb el que es disposava tenint present que es l'objectiu es tractava més aviat d'un projecte de síntesi/aprenentatge/investigació que d'un projecte per obtenir un producte comercial.

La tercera candidata va ser l'opció seleccionada i cal dir que tot i la decepció per no obtenir uns bons resultats, resta la satisfacció d'haver aconseguit altres fites gens menyspreables i que probablement si no haguessin existit tots aquests obstacles no s'haurien abastat.

5.5. CLASSIFICACIÓ I PROVES

En el present treball s'han posat a prova els algorismes SVM (Support Vector Machines), KNN (K-Nearest Neighbors), Naïve Bayes, AdaBoost, Decission Tree Classificator i K-means, considerant algunes de les seves variants com es veurà més endavant.

SVM (Support Vector Machines)

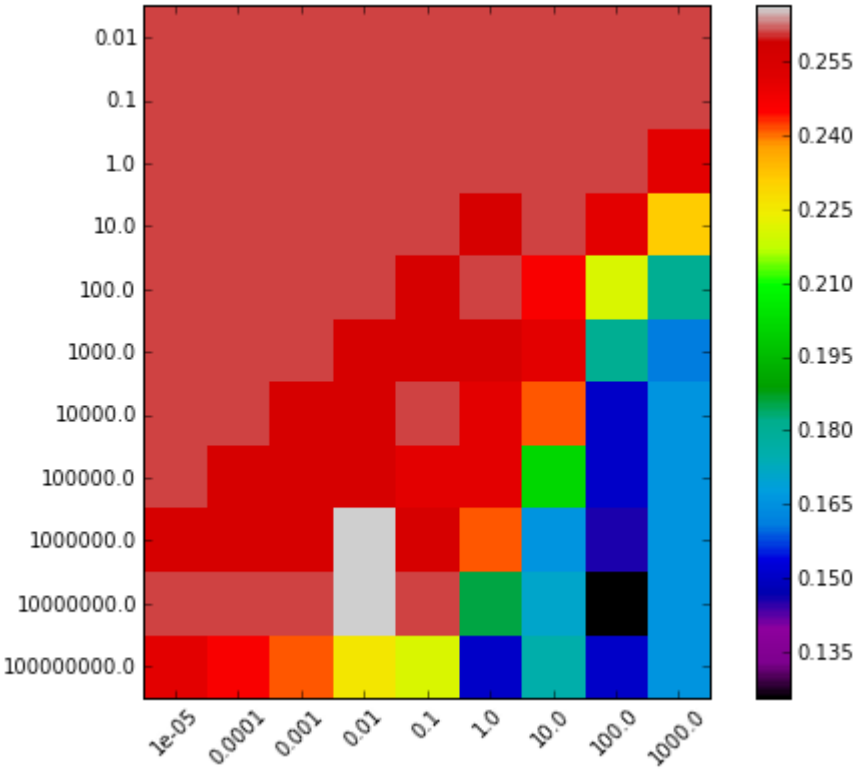
Support Vector Machines és un conjunt de mètodes d'aprenentatge supervisat per la classificació i regressió. És eficaç en espais de nombroses dimensions. Un avantatge d'aquests mètodes és que permeten especificar diferents funcions del nucli per portar a terme la decisió, així com aplicar modes personalitzats.
Encara que respon bé a casos amb un nombre de característiques més alt que el d'exemples, si la diferència és molt important els resultats no són els més adequats.
SVM calcula la puntuació d'encerts que el classificador obté tenint en compte les diferents característiques o descriptors, ja siguin preses individualment o agrupades en part o en la seva totalitat (u1,u2.....u6) en relació l'atribut que es pretén predir, i que serà un dels tres següents: *moviment, tècnica o gènere*.

Classificació

Inclosos en aquest conjunt trobem les classes SVC (Support Vector Classification), NuSVC i LinearSVC, capaces de millorar la classificació en casos de multi classes.
SVC accepta les funcions de nucli 'rbr', 'linear', 'sigmoid', 'poly', 'precomputed' així com personalitzades pel programador. En concret s'han utilitzat i s'inclouen resultats SVC amb les funcions 'rbr', 'linear' i 'sigmoid'.

Per defecte els paràmetres de la classe sklearn.svm.SVC són:
(*C=1.0,kernel='rbf',degree=3,gamma=0.0, coef0=0.0, shrinking=True, probability=False,tol =0.001, cache_size=200, class_weight=None, verbose=False, max_iter=-1, random_state=*

None). Calia concretar quins eren els paràmetres més adients al nostre cas, per això es va adaptar el codi trobat a la pàgina oficial de Scickit (http://scikit-learn.org/0.11/auto_examples/svm/plot_svm_parameters_selection.html) i es va obtenir com a exemple d'un cas particular, que el millor classificador s'ha de configurar de la següent forma: *(C=1000000.0, cache_size=200, class_weight=None, coef0=0.0, degree=3, gamma=0.01, kernel='rbf', max_iter=-1, probability=False, random_state=None, shrinking=True, tol=0.001, verbose=False).* El gràfic orientatiu que es va obtenir és el següent:



Il·lustració 14: Paràmetres classificador SVC

També s'han volgut comparar resultats de la classe LinearSVC de SVM.

Regressió

Com es manifesta a l'apartat «Resumen», part de l'objectiu rau en provar els diferents

algorismes. D'acord amb això calia incloure per tant, proves amb algorismes de regressió. Els resultats en aquest cas han estat excessivament inferiors als que podrien haver-se obtingut aleatòriament, qüestió del tot previsible a priori ja que per obtenir uns millors resultats les dades s'haurien d'haver preprocesat per reduir la seva dimensionalitat utilitzant per exemple LDA supervisat (Linear Discriminant Analysis).
Les proves amb la classe SVR (Support Vector Regression) igual que en altres casos de regressió exigien transformar les dades categòriques a dades quantitatives qüestió que per la manca de temps no s'ha realitzat.

A la següent taula es recullen els millors resultats obtinguts per les diferents característiques.
La característica autor no s'ha reflectit però es fa constar que de totes les combinacions efectuades, en cap moment s'ha aconseguit passar d'un 3% - 1%.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	SVC	µ1, µ2, µ3, µ4, µ5, µ6	23	20,5
		µ1 ,µ2	22,5	31,16
		µ3	28	14
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	LinearSVC	µ1, µ2, µ3, µ4, µ5, µ6	18,4	55,4
		µ1 ,µ2	73	54
		µ3	69	55
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	SVC	µ1, µ2, µ3, µ4, µ5, µ6	18,4	18
		µ1 ,µ2	20	8,5
		µ3	23	14

Taula 3: Proves amb SVC

S'observa que la característica tècnica té una bona resposta als descriptors amb que s'han realitzat les proves.

K-Nearest Neighbors Classifier o KNN

Aquest algorisme s'utilitza tant per la classificació com per la regressió supervisada, ja que necessita de casos per entrenar-se.

Classificació

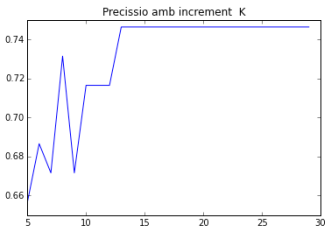
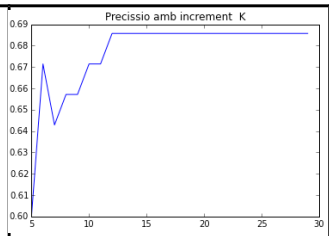
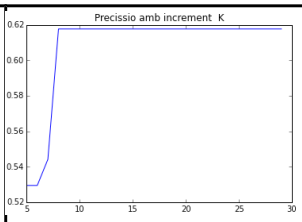
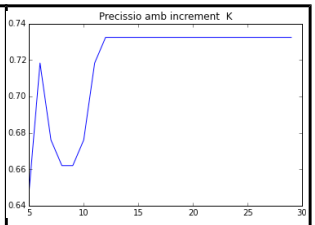
La classificació es calcula comptant els costos dels veïns més propers (majoria simple). Hi ha la possibilitat de ponderar les distàncies dels veïns d'alguna forma:

- 'uniform', assignant pesos uniformes per cada veí.
- 'distance', assigna pesos proporcionals a la inversa de la distància de la distància del punt de consulta.

S'ha de tenir en compte que els paràmetres per defecte de la classe KneighborsClassifier són els següents (n_neighbors=3,weights='uniform').

En el present treball s'han fet proves amb els dos tipus de ponderació de les distàncies sent els resultats molt similars entre si.

L'algorisme per defecte calcula tres veïns, però es va creure interessant conèixer per cada cas quin era el nombre de veïns més adequat per aconseguir la màxima precisió. A tall d'exemple es presenten quatre iteracions d'un mateix cas (considerant l'atribut 'tècnica') generats seqüencialment. Els resultats obtinguts poden ajudar a reflexionar i seleccionar el nombre més adequat de veïns: podem considerar a la vista d'aquests exemples que 8 veïns és un nombre relativament adequat pels nostres càlculs :

			
Veins: 5, Precissió: 0.656716 Veins: 6, Precissió: 0.686567 Veins: 7, Precissió: 0.671642 Veins: 8, Precissió: 0.731343 Veins: 9, Precissió: 0.671642 Veins: 10, Precissió: 0.716418 Veins: 11, Precissió: 0.716418 Veins: 12, Precissió: 0.716418 Veins: 13, Precissió: 0.746269 Veins: 14, Precissió: 0.746269 Veins: 15, Precissió: 0.746269 Veins: 16, Precissió: 0.746269 Veins: 17, Precissió: 0.746269 Veins: 18, Precissió: 0.746269 Veins: 19, Precissió: 0.746269 Veins: 20, Precissió: 0.746269 Veins: 21, Precissió: 0.746269 Veins: 22, Precissió: 0.746269 Veins: 23, Precissió: 0.746269 Veins: 24, Precissió: 0.746269 Veins: 25, Precissió: 0.746269 Veins: 26, Precissió: 0.746269 Veins: 27, Precissió: 0.746269 Veins: 28, Precissió: 0.746269 Veins: 29, Precissió: 0.746269 PRECISIÓ %: 74.6268656716	Veins: 5, Precissió: 0.600000 Veins: 6, Precissió: 0.671429 Veins: 7, Precissió: 0.642857 Veins: 8, Precissió: 0.657143 Veins: 9, Precissió: 0.657143 Veins: 10, Precissió: 0.671429 Veins: 11, Precissió: 0.671429 Veins: 12, Precissió: 0.685714 Veins: 13, Precissió: 0.685714 Veins: 14, Precissió: 0.685714 Veins: 15, Precissió: 0.685714 Veins: 16, Precissió: 0.685714 Veins: 17, Precissió: 0.685714 Veins: 18, Precissió: 0.685714 Veins: 19, Precissió: 0.685714 Veins: 20, Precissió: 0.685714 Veins: 21, Precissió: 0.685714 Veins: 22, Precissió: 0.685714 Veins: 23, Precissió: 0.685714 Veins: 24, Precissió: 0.685714 Veins: 25, Precissió: 0.685714 Veins: 26, Precissió: 0.685714 Veins: 27, Precissió: 0.685714 Veins: 28, Precissió: 0.685714 Veins: 29, Precissió: 0.685714 PRECISIÓ %: 68.5714285714	Veins: 5, Precissió: 0.529412 Veins: 6, Precissió: 0.529412 Veins: 7, Precissió: 0.544118 Veins: 8, Precissió: 0.617647 Veins: 9, Precissió: 0.617647 Veins: 10, Precissió: 0.617647 Veins: 11, Precissió: 0.617647 Veins: 12, Precissió: 0.617647 Veins: 13, Precissió: 0.617647 Veins: 14, Precissió: 0.617647 Veins: 15, Precissió: 0.617647 Veins: 16, Precissió: 0.617647 Veins: 17, Precissió: 0.617647 Veins: 18, Precissió: 0.617647 Veins: 19, Precissió: 0.617647 Veins: 20, Precissió: 0.617647 Veins: 21, Precissió: 0.617647 Veins: 22, Precissió: 0.617647 Veins: 23, Precissió: 0.617647 Veins: 24, Precissió: 0.617647 Veins: 25, Precissió: 0.617647 Veins: 26, Precissió: 0.617647 Veins: 27, Precissió: 0.617647 Veins: 28, Precissió: 0.617647 Veins: 29, Precissió: 0.617647 PRECISIÓ %: 61.7647058824	Veins: 5, Precissió: 0.647887 Veins: 6, Precissió: 0.718310 Veins: 7, Precissió: 0.676056 Veins: 8, Precissió: 0.661972 Veins: 9, Precissió: 0.661972 Veins: 10, Precissió: 0.676056 Veins: 11, Precissió: 0.718310 Veins: 12, Precissió: 0.732394 Veins: 13, Precissió: 0.732394 Veins: 14, Precissió: 0.732394 Veins: 15, Precissió: 0.732394 Veins: 16, Precissió: 0.732394 Veins: 17, Precissió: 0.732394 Veins: 18, Precissió: 0.732394 Veins: 19, Precissió: 0.732394 Veins: 20, Precissió: 0.732394 Veins: 21, Precissió: 0.732394 Veins: 22, Precissió: 0.732394 Veins: 23, Precissió: 0.732394 Veins: 24, Precissió: 0.732394 Veins: 25, Precissió: 0.732394 Veins: 26, Precissió: 0.732394 Veins: 27, Precissió: 0.732394 Veins: 28, Precissió: 0.732394 Veins: 29, Precissió: 0.732394 PRECISIÓ %: 73.2394366197

Taula 4: Precisió i veïns

Cal remarcar que no són dades extremadament significatives ja que depenen en bona part de la forma com estan repartits els casos d'entrenament i els de test (tot i que es seleccionen aleatòriament).

Els resultats per la característica autor han estat d'un màxim de 4% - 6% sense que hagi diferència notable entre els diferents algorismes.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	KNeighborsClassifier	µ1, µ2, µ3, µ4, µ5, µ6	17,7	13,4
		µ1, µ2	15	10
		µ3	15	11,5
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	KNeighborsClassifier	µ1, µ2, µ3, µ4, µ5, µ6	73,2	56,4
		µ1, µ2	70,7	56,4
		µ3	76	54
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	KNeighborsClassifier	µ1, µ2, µ3, µ4, µ5, µ6	27,6	25,1
		µ1, µ2	21,2	19,1
		µ3	20,3	18

Taula 5: Proves amb KNeighborsClassifier

Regresssió

La classe KneighborsRegressor realitza la regressió prenent els mateixos paràmetres per defecte que a la classe classificadora així com els mateixos valors.

Naïve Bayes

Mètode d'aprenentatge supervisat que permet aplicar l'estimador segons diferents models de distribució estadística:

- **BernoulliNB**, l'estimador aplica una funció matemàtica d'acord amb la distribució de Bernoulli. Permet considerar múltiples variables i cada una es tracta com un valor boolean que pren un valor si pertany o no al grup.
- **MultinomialNB** a diferència de l'anterior mesura el conjunt de probabilitats de que cada característica pertanyi al grup. Addicionalment, també a diferència del model

BernoulliNB, les característiques no presents es consideren i es penalitzen reduint el valor de la freqüència o valor final de la funció.

- **GaussiaNB**, l'estimador calcula la possibilitat de tenir una característica d'acord a la distribució per campana de Gauss.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	MultinomialNB	µ1, µ2, µ3, µ4, µ5, µ6	-----	-------
		µ1, µ2	20	16,38 (amb 3 casos) 5,48 (amb 10 casos)
		µ3	16,17	16,43 (amb 3 casos) 2,8 (amb 10 casos)
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	MultinomialNB	µ1, µ2, µ3, µ4, µ5, µ6		------
		µ1, µ2	73,27	73,86 (amb 3 casos) 54,11 (amb 10 casos)
		µ3	66,17	66,92 (amb 3 casos) 54,11 (amb 10 casos)
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	MultinomialNB	µ1, µ2, µ3, µ4, µ5, µ6		------
		µ1, µ2		----
		µ3	---	---

Taula 6: Proves amb MultinomialNB

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	BernoulliNB	µ1, µ2, µ3, µ4, µ5, µ6	16,83	17,20 (amb 3 casos) 2,8 (amb 10 casos)
		µ1, µ2	16,9	17,27 (amb 3 casos) 2,8 (amb 10 casos)
		µ3		14,6 (amb 3 casos) 14,8 (amb 10 casos
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	BernoulliNB	µ1, µ2, µ3, µ4, µ5, µ6	76,16	76,5(amb 3 casos 56,14 (amb 10 casos
		µ1, µ2	74,6	75,18 (amb 3 casos 54,11 (amb 10 casos
		µ3	75	71,4 (amb 3 casos) 54,11 (amb 10 casos)
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	BernoulliNB	µ1, µ2, µ3, µ4, µ5, µ6	21,2	21 (amb 3 casos) 17,6 (amb 10 casos)
		µ1, µ2	22,2	21,9 (amb 3 casos) 18,6 (amb 10 casos)
		µ3	----	----

Taula 7: Proves amb BernoulliNB

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	GaussianNB	µ1, µ2, µ3, µ4, µ5, µ6	12,12	12,46 (amb 3 casos) 11,43 (amb 10 casos)
		µ1, µ2	18,6	18,9 (amb 3 casos) 11,29 (amb 10 casos)
		µ3	12,34	17,23 (amb 3 casos) 7,7 (amb 10 casos)
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	GaussianNB	µ1, µ2, µ3, µ4, µ5, µ6	44,95	25,20 (amb 3 casos) 61,10 (amb 10 casos)
		µ1, µ2	71,96	71,75 (amb 3 casos) 54,72 (amb 10 casos)
		µ3	62,5	65,25 (amb 3 casos) 54,62 (amb 10 casos)
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	GaussianNB	µ1, µ2, µ3, µ4, µ5, µ6		-----
		µ1, µ2		----
		µ3		-----

Taula 8: Proves amb GausianNB

En el cas de la característica *autor* no hi ha diferències significatives entre els algorismes, i en cap cas es supera el 7% d'encerts.

AdaBoost

Aquest classificador permet crear una seqüència de classificadors dèbils en que cada nova seqüència ofereix una exactitud superior a l'anterior. El punt d'eficiència inicial o de partida és del 50% (eficiència de selecció aleatòria) i si arriba un moment en què l'eficiència calculada a la nova seqüència no és superior a la prèvia, el classificador s'atura.
El valor per defecte de nombre de seqüències és de 50. En triar un nombre excessivament gran existeix el risc d'overfitting o sobreentrenament que resulta ser el fet de memoritzar dades i no aprendre-les que és el desitjable.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	AdaBoostClassifier	µ1, µ2, µ3, µ4, µ5, µ6	20,31	6,2
		µ1, µ2	16,51	8,4
		µ3	21,53	7
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	AdaBoostClassifier	µ1, µ2, µ3, µ4, µ5, µ6	75,46	54,18
		µ1, µ2	70	54,95
		µ3	72,10	54,11
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	AdaBoostClassifier	µ1, µ2, µ3, µ4, µ5, µ6	29,39	----
		µ1, µ2	3	----
		µ3	27,89	----

Taula 9: Proves amb AdaBoostClassifier

Decission Tree Classificador

És un classificador, senzill, no parametritzable de cost logarítmic que accepta dades categòriques i numèriques,.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	DecissionTreeClassifier	µ1, µ2, µ3, µ4, µ5, µ6	16,03	15,94
		µ1, µ2	11,38	13,77
		µ3	10,28	10,68
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	DecissionTreeClassifier	µ1, µ2, µ3, µ4, µ5, µ6	57,91	61,78
		µ1, µ2	54,96	59,84
		µ3	69,87	53,69
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	DecissionTreeClassifier	µ1, µ2, µ3, µ4, µ5, µ6	19,37	----
		µ1, µ2	21,67	----
		µ3	19,68	----

Taula 10: Proves amb DecissionTreeclassifer

K-Means

Un altre algorisme d'agrupament que a través d'un conjunt d'operacions matemàtiques aconseguix després de diverses iteracions arribar a crear uns centroids recalculats a cada interacció fins que considera que ha convergit.

El model per defecte genera vuit centroids o clústers i el màxim d'iteracions és de 300. El més adient seria establir com nombre de centroids la quantitat de variacions de l'atribut que es consideri.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	Kmeans	µ1, µ2, µ3, µ4, µ5, µ6	-	-
		µ1, µ2	-	-
		µ3	20,8	3,7
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	Kmeans	µ1, µ2, µ3, µ4, µ5, µ6	-	-
	Kmeans	µ1, µ2	-	-
		µ3	21	3,8
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	Kmeans	µ1, µ2, µ3, µ4, µ5, µ6	-	-
		µ1, µ2	-	-
		µ3	17,58	3,8

Taula 11: Proves amb Kmeans

Perceptron

Es tracta d'un cas determinat del SGD Classificator (Stochastic Gradient Descent), en concret els paràmetres per defecte que utilitza són (loss='perceptron', eta0=1, learning_rate='constant', penalty=None). És un algorisme que previsiblement no havia de donar resultats acceptables, ja que és un dels models lineals en que existeix una variable independent i d'altres variables dependents d'ella. Les característiques de les pintures presumiblement no segueixen un model lineal.

Característica	Algorisme	Descriptors	% encert	% encert cross validation
moviment	SGDClassifier	µ1, µ2, µ3, µ4, µ5, µ6	15,5	3
		µ1, µ2	3,8	0,9
		µ3	7	0,5
Característica	Algorisme	Descriptors	% encert	% encert cross validation
tecnica	SGDClassifier	µ1, µ2, µ3, µ4, µ5, µ6	61,37	61,74
		µ1, µ2	70,10	50,74
		µ3	74,24	47,33
Característica	Algorisme	Descriptors	% encert	% encert cross validation
genere	SGDClassifier	µ1, µ2, µ3, µ4, µ5, µ6	---	---
		µ1, µ2	----	----
		µ3	----	----

Taula 12: Proves amb perceptron

CROSS VALIDATION

Aquest algorisme s'utilitza en conjunt i amb posterioritat a qualsevol dels models de classificació explicats fins ara: la finalitat d'aquest algorisme és que els resultats de predicció que s'obtenen en els diferents models no depenguin de la forma com es fa la divisió entre les dades d'entrenament i de prova. Així, les dades de la mostra es divideixen en una sèrie de subconjunts. Si posem com exemple *10-fold cross-validation* les dades es dividiran en 10 subgrups, dels que un d'ells s'utilitza com grup de prova i els altres nou com grups d'entrenament. L'algorisme utilitzarà cada vegada un subconjunt d'entrenament diferent contra el de prova. Aquest procediment es repetirà fins que tots els subconjunts hagin estat considerats com subconjunts de prova. Amb les dades adquirides es realitzarà una mitja aritmètica que serà el valor de la validació creuada.

6. CONCLUSIONS

Els resultats obtinguts no són suficientment satisfactoris en general, exceptuant a la característica *tècnica*, amb gairebé tots els classificadors. La meva opinió és que una part important dels errors i falles passades i presents són conseqüència d'una banda de com es va generar la base de dades ja que es pot dir que es va formar pel mètode de prova error, sense conèixer gairebé fins el final quins són els requeriments més adequats. D'altra banda, la base de dades ha estat molt manipulada com s'ha comentat en anteriors apartats, i sembla ser que per aquest motiu no ha funcionat correctament en efectuar les proves per documentar la memòria (sobretot els càlculs relatius a la característica *gènere).*

En anteriors apartats s'ha fet menció a les dificultats patides en relació al projecte, i atès que cal relatar les conclusions, s'ha de tornar a comentar que gràcies als inconvenients el resultat que s'obté del projecte (sense tenir en compte els resultats directes de les proves que no han estat tant correctes o encertats com es podia esperar):

- Aprenentatge sobre els diferents descriptors, cv2, python, etc.
- Aprofundiment en el funcionament tant de la visió per computador com de la màquina d'aprenentatge.
- Assolir la capacitat de poder dissenyar un/s descriptors adients al tema.

Per últim vull fer constar que la idea personal en començar el treball va ser arribar a programar una aplicació que permetés predir a quin moviment, quina tècnica o a quin gènere pertany una pintura interaccionant entre l'usuari i una interfície. No ha estat possible però, a nivell de programació ha estat molt senzill incloure una línia per realitzar les prediccions.

7. MILLORES PER PRÒXIMS TREBALLS

Els histogrames ofereixen molta més informació de la que aquí s'ha utilitzat, com és el cas de determinar si en una obra predominen els tons clars o foscos (si el pes de l'histograma està cap

a la dreta o cap a l'esquerra), si predominen els tons mitjans (contrast en clau alta o baixa) o els «molt foscos» i «molt clars», etc. Caldria realitzar un estudi en profunditat de tot el que un histograma pot comunicar. Amb aquestes possibilitats al cap caldria realitzar una selecció de descriptors que ja existeixin com a tals o combinacions que puguin ser interessants per detectar diverses característiques visuals. Seria adequat com s'acaba de suggerir ampliar amb variacions dels descriptors o diferents enfocaments d'aplicació, és a dir, podria ser convenient estudiar com es comporten petites àrees dels quadres (a petita escala, potser s'obtindrien dades de la pinzellada...). Paral·lelament seria necessari conèixer a nivell avançat què és el que determina cada una de les característiques que volem entrenar (foscor, cares llargues, elements geomètrics, etc.) i provar els diferents descriptors sistemàticament sobre reductes controlats de la base de dades. Segurament amb aquestes proves es podria inferir amb molta més exactitud com entrenar el classificador.

Per acabar de millorar la feina, seria imprescindible crear una interfície gràfica que permeti interactuar amplia i fàcilment amb l'aplicació sense necessitat de modificar els diferents paràmetres directament en el codi.

8. APÈNDIX

EINES

Es va considerar la possibilitat d'utilitzar Java com llenguatge de programació per realitzar les aplicacions necessàries per ser un dels llenguatges més coneguts i utilitzats per l'autora del treball. Es va valorar la possibilitat de programar en Python per recomanació del professor tutor i finalment veient la facilitat d'aprenentatge d'aquest llenguatge i la quantitat de documentació i exemples disponibles a la Xarxa es va optar per seleccionar Python. L'entorn de desenvolupament escollit per treballar ha estat Anaconda sobre Windows 7 de 64 bits. Anaconda és un entorn lliure per Python que disposa de gran quantitat de paquets necessaris pel processament de dades, anàlisis predictiu i computació científica. Alguns dels paquets inclosos i utilitzats són Beautiful-soup, Matplotlib, NumPy, Pandas, Scickit-learn, SciPy així

com un de molt important que va ser necessari instal·lar posteriorment ja que no està inclòs d'origen: Opencv en la seva versió cv2, llibreria de programació open source enfocada a la visió per computador i tractament d'imatges.

GLOSSARI

- ALGORISME

Conjunt ordenat i finit d'operacions que permet trobar la solució d'un problema.

- APRENENTATGE SUPERVISAT

Tècnica per deduir una funció a partir de dades d'entrenament. L'objectiu de l'aprenentatge supervisat és el de crear una funció capaç de predir el valor corresponent a qualsevol objecte d'entrada vàlida després d'haver vist una sèrie d'exemples, les dades d'entrenament.

- BASE de DADES

Bancs de dades organitzades per temàtiques i categories, que comparteixen vincles o relació entre si.

- CAMPANA DE GAUSS

És una representació gràfica de la distribució normal d'un grup de dades. Aquests es reparteixen en valors baixos, mitjans i alts, creant un gràfic de forma acampanada i simètrica respecte a un determinat paràmetre. El punt màxim de la corba correspon a la mitjana, i té dos punts d'inflexió a banda i banda.

- CENTROID

És un punt que defineix el centre geomètric de diferents punts geomètricament distants. La seva localització es pot determinar a partir de formules semblants a les utilitzades per determinar el centre de gravetat o el centre de massa dels punts.

- DADES CATEGÒRIQUES

Són dades que només poden prendre un nombre finit no numerable de possibles valors.

- DESCRIPTOR

Paraula clau que defineix el contingut d'un objecte i ens ajuda a localitzar-lo en un arxiu manual o informatitzat.

-DESVIACIÓ ESTÀNDARD

Mesura de dispersió respecte a la mitja aritmètica.

- DISTRIBUCIÓ DE BERNOULLI

és una distribució de probabilitat discreta, que pren valor 1 per a la probabilitat d'èxit (p) i valor 0 per a la probabilitat de fracàs (q = 1-p).

- HISTOGRAMA

És una manera de representar gràficament dades quantitatives procedents d'una imatge, s'acostuma a utilitzar per representar el nombre de píxels amb un determinat valor de color comú.

- IL·LUMINACIÓ DIFUSA

La il·luminació difusa és una llum tènue sense la intensitat ni la resplendor de la llum directa. Aquesta disseminada i ve des de totes les direccions.

- IL·LUMINACIÓ ESTRUCTURADA

Tipus d'il·luminació utilitzada normalment per aconseguir l'efecte de 3D i per aconseguir que els dispositius de visió artificial puguin obtenir la distància de l'objecte enfocat.

- IL·LUMINACIÓ DIRECCIONAL

És generada per llums directes, que il·luminen àrees relativament petites, té un feix de llum molt marcat que produeix ombres denses i ben definides.

- INTEL·LIGÈNCIA ARTIFICIAL:

Sistemes digitals que pensen de manera racional (es caracteritzen per imitar el pensament lògic dels humans, és a dir, en una qüestió concreta raonen com a experts).

- LÓGICA DIFUSA

La Lògica Difusa utilitza expressions que no són totalment certes ni completament falses, és a dir, que poden prendre un valor qualsevol de veracitat.

- LLIBRERIA

Conjunt d'implementacions funcionals, codificades en un llenguatge de programació, que ofereix una interfície ben definida per a la funcionalitat que s'invoca.

- LLICÈNCIES CREATIVE COMMONS

Diverses llicències de drets d'autor/a que permeten exercir el dret que la llei li confereix en exclusiva per decidir certes condicions d'ús i distribució de la seva obra .

-MITJA ARITMÈTICA

Paràmetre estadístic resultant de sumar determinats valors i dividir entre el total dels elements que hi han participat a la suma.

- PARAMETRITZACIÓ

Determinar, fixar determinats valors que modelen algunes funcions matemàtiques per adaptar-les al tractament matemàtic necessari.

- PÍXELS

Cadascuna de les Mínimes fraccions en què és dividida una imatge en en sotmetre-la a un escombratge electrònic.

- SOROLL GAUSSIÀ

El soroll Gaussià es troba associat amb la radiació electromagnètica ja que no podem tenir comunicació elèctrica sense electrons i per tant és impossible evitar el soroll.

- SOROLL IMPULSIU

La intensitat d'aquest soroll augmenta bruscament durant un impuls. Incideix fonamentalment en la transmissió de les dades, es deu bàsicament a fortes induccions conseqüència de commutacions electromagnètiques.

-TUPLA

Llista ordenada d'elements.

- UIT R BT.601-7

La Recomanació UIT-R BT.601-7, antigament denominada CCIR 601, és la primera recomanació sobre televisió digital, realitzada per la Unió Internacional de Telecomunicacions (UIT) que es refereix al mostreig dels senyals de vídeo i àudio.

- VALOR booleà

Referencia a una expressió o una variable que pot tenir només un valor vertader o fals.

ÍNDEX DE FIGURES

Planificació inicial, diagrama Gannt....7
Exemple Wikimedia Commons....10
Exemple Wikiart....10
Extensió Download Master....11
Exemple de modificació enllaços....12
Base de dades d'imatges amb 3395 fitxers....12
Exemple de la base de dades en .csv....13
Exemple funcionament Sobel....16
Imatge amb pocs contrastos....18
Histograma....18
Desviació estàndard amb cv2....19
Detall càlcul descriptors....21
Base de dades en .csv....22
Paràmetres classificador SVC....24

ÍNDEX DE TAULES

Planificació inicial....7
Planificació real....8
Proves amb SVC....25
Precisió i veïns....27
Proves amb KNeighborsClassifier....28
Proves amb MultinomialNB....29
Proves amb BernoulliNB....30
Proves amb GausianNB....30
Proves amb AdaBoostClassifier....31
Proves amb DecissionTreeclassifer....32
Proves amb Kmeans....33
Proves amb perceptron....34

BIBLIOGRAFIA

ERIC SOLEM, J. (2012) *Programming Computer Vision with Python.* O'Reilly

GRADSKI, G. i KAEHLER A. (2008) *Learning OpenCV ComputerVision with the OpenCV Library.* O'Reilly

Pàgines web

Curs Machine Learning Udacity. ***https://www.udacity.com/course/machine-learning-supervised-learning--ud675*** [Última visita 16/06/2015]

Portal OpenCV. ***http://docs.opencv.org/master/index.html*** [Última visita 15/06/2015]

Scikit learn ***http://scikit-learn.org/stable/*** [Última visita 15/06/2015]

Springer Link ***http://link.springer.com/article/10.1007/s00138-014-0621-6*** [Visita 24/05/2015]

University of Nevada ***http://www.cse.unr.edu/~bebis/CS791E/*** [Visita 24/04/2015]

Oguz Icoglu, Bilge gunsel, Sanem Sariel
http://web.itu.edu.tr/sariel/publications/eusipco_04_sariel.pdf [Última visita 15/06/2015]