

# Entorn de conducció autònoma: Classificació automàtica de vianants, cotxes, voreres, etc., a partir de classificar i etiquetar superpíxels utilitzant CRF

Bernat Mayolas Lacambra

**Resum**— Aquest treball consisteix en la construcció d'un classificador pel reconeixement d'imatge usant una base de dades pública. Per a la construcció d'aquest classificador, les dades d'entrada és l'energia dels nodes calculada a partir del *hog*, *sift* i *color* entre altres, extretes sobre els superpíxels. S'ha dissenyat el classificador a partir de Conditional Random Fields per a l'obtenció de millors resultats ja que aquest classificador incorpora el context, és a dir, els píxels pròxims al píxel objectiu a classificar. El classificador està basat en CRF, més concretament DenseCRF, que consisteix en la creació d'un model on tots els píxels de la imatge estan connectats entre ells. Pels resultats obtinguts amb els CRF, més concretament DenseCRF, podem concloure com ambdós mètodes són d'utilitat per a la classificació automàtica en l'entorn de conducció autònoma.

**Paraules clau**—classificador, superpíxel, *Conditional Random Fields*, DenseCRF, conducció autònoma

**Abstract**— This work involves the development of a classifier for image recognition using a public database. The input data used for the development of this classifier is the node energy, calculated through the *hog*, *sift* and *color* among others. It has been designed through Conditional Random Fields, in order to achieve better results, as they take into account context, or in other words, the pixels nearby the target pixel to be classified. The classifier is based in CRF, more specifically *DenseCRF*, which involves creating a model where all the pixels in the image are interconnected. The results obtained from CRF and more specifically the *DenseCRF*, allow us to conclude that both classifiers are useful on the automatic classification algorithms in the environment autonomous driving.

**Index Terms**—classifier, superpixel, conditional random fields, DenseCRF autonomous driving

## 1 INTRODUCCIÓ

La conducció autònoma està basada en el reconeixement d'imatges mitjançant classificadors, es genera una nova imatge cada cert temps, i en temps real, el classificador ha de ser capaç d'haver etiquetat tots els components de la imatge, aconseguint el reconeixement de cotxes, carrils, vianants etc.. per a la correcta circulació e interpretació de la imatge per la decisió de les accions que ha de prendre per seguir circulant per la via pública.

El projecte consisteix en la creació d'un classificador *Conditional Random Field*[1], que a diferència d'altres classificadors, té en compte el context alhora de classificar regions d'imatges preses per un vehicle que circuli per la via pública, permetent identificar els elements que componen la imatge. La distinció d'aquestes regions, pertanyents a elements: cotxe, carretera o vianant té una aplicabilitat immediata en el desenvolupament d'eines per la conducció autònoma.

Aquest projecte s'integrarà en un framework (anomenat FrameworkAdas), software desenvolupat pel grup ADAS

- E-mail de contacte: [bernat.mayolas@e-campus.uab.cat](mailto:bernat.mayolas@e-campus.uab.cat)
- Menció realitzada: Computació
- Treball tutoritzat per: Gemma Sánchez Albaladejo
- Curs 2014/15

del centre de visió per computador, l'objectiu del qual és poder integrar el software desenvolupat en el cotxe presentat a la figura 1, dotat de cameres i la possibilitat de controlar les seves accions.

Mitjançant el framework anteriorment esmentat, es farà un preprocessat de les imatges per extreure regions de píxels, que anomenarem superpíxels[2], aconseguint reduir així la carga computacional, i s'extrauran característiques d'aquestes regions de les imatges obtingudes.



Figura 1 Vehicle intel·ligent Grup ADAS CVC

Amb les característiques extretes d'aquests superpíxels, construirem el classificador. Aquest seleccionarà les regions i les ubicarà en diferents classes, anteriorment deter-

minades, que engloben tots els possibles elements que es poden percebre des del cotxe.

Les classes són les següents: cel, edifici, carretera, vorera, tanca, vegetació, pals (en referència a senyals de trànsit), cotxe, senyal, vianant i ciclista. A la figura 2 es pot observar un exemple d'aquest etiquetatge amb les següents classes: vermell dels edificis, groc per vegetació i miscel·lani, groc més pàlid per els pals de les senyals, rosa per la carretera, lila pel cotxe, blau fosc per la vorera i blau cel pel ciclista.



Figura 2 Etiquetatge de les classes predeterminades

El desenvolupament de l'algorisme classificador estarà basat en l'eina matemàtica anomenada *Conditional Random Fields*. Aquesta eina, per contra d'altres algorismes de classificació, empra la informació dels nodes veïns al node objectiu a classificar; sent més acurada *a priori*.

L'article queda estructurat en els següents apartats:

2. Objectius, recull dels principals objectius a complir durant el desenvolupament del treball. 3 Estat de l'art, amb una breu explicació del context actual d'altres algorismes classificadors, i explicació dels conceptes teòrics necessaris. 4 la metodologia, hi ha el procés de disseny del software que s'ha seguit, i explicació de la feina realitzada abans de la construcció del classificador principal CRF. 5 el Disseny de l'algorisme principal, estan anotades les explicacions pertinents a la creació de l'algorisme classificador CRF. 6 resultats tenim les dades obtingudes de l'entrenament i test de l'algorisme principal. 7 les conclusions extretes dels resultats i el treball.

## 2 OBJECTIUS

El projecte consta de 3 objectius principals:

- I. Obtenció de les imatges d'una base de dades pública Canvid[3], més explicada en detall en l'estat de l'art, i de la indexació corresponent entre els píxels i els superpíxels als quals pertanyen, i l'energia dels nodes obtinguda mitjançant el FrameworkADAS, sobre les imatges de la base de dades anterior.
- II. Estudi i anàlisi dels classificadors *Conditional Random Fields*, decantant-nos per la crea-

ció de *DenseCRF*, tipus de CRF que té en compte tots els píxels de la imatge, amb el corresponent entrenament i emprant les dades obtingudes en (I) per a la classificació dels elements en les diferents classes predeterminades.

- III. Integració de l'algorisme desenvolupat en (II) al FrameworkADAS, adaptant tant el codi del propi classificador, com el del Framework. Això permetrà l'ús del classificador i les modificacions pertinents per qualsevol base de dades que posteriorment interressi utilitzar durant el desenvolupament d'aquest.

## 3 ESTAT DE L'ART

### 3.1 Conceptes teòrics

#### 3.1.1 Superpíxels

Les imatges de la base de dades pública han sigut processades mitjançant el FrameworkADAS, extraient-ne regions de píxels, que anomenarem superpíxels, per aconseguir una reducció en la carga computacional alhora d'assignar l'energia corresponent a cada node.

Els superpíxels són segments d'una imatge digital, més gran que un píxel, i que han sigut assignats segons la semblança d'una sèrie de característiques, el que permet ajuntar aquestes regions de píxels semblants.

Per tal de generar els superpíxels s'extreuen una sèrie de dades sobre la imatge, entre aquestes hi trobem:

- *Hog*[3]: La tècnica de descriptor HOG comptabilitza el número de vegades que apareix el gradient orientat en porcions localitzades d'una imatge.
- *Sift*[4] (Scale-Invariant Feature Transform): El detector extreu un número de regions basats en la il·luminació, punt de vista etc.. i el descriptor associa a aquestes regions una signatura.
- *Color*[5] : El color de la regió seleccionada, s'acostuma a analitzar fent un histograma dels valors RGB de tots els píxels, de l'espai de color XYZ o CIELAB .

### 3.1.2 Base de dades

Les imatges usades durant l'entrenament i test de l'algorisme han sigut extretes d'una base de dades pública, Camvid[6] dataset, a la figura 2 en podem observar un exemple abans de ser classificat i el seu corresponent Ground truth[7]. El terme ground truth fa referència a l'etiquetatge ideal, ja sigui manual o automàtic, de les classes existents a la imatge corresponent.

A la figura 3 podem observar el total d'etiquetatges de la base de dades, obtenint una referència dels elements que apareixen en les imatges preses a la via pública.

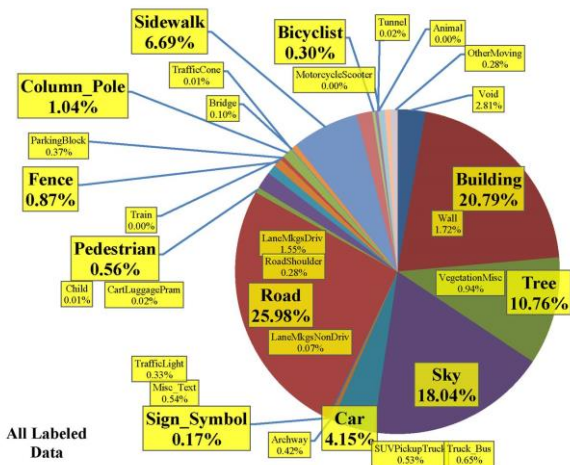


Figura 3: Representació de classes en Camvid

### 3.1.3 Conditional Random Fields

Els classificadors són un conjunt d'algorismes que donats uns paràmetres determinats són capaços de classificar una sèrie d'elements, en diversos grups, agrupant-los per tenir característiques similars o comunes.

Per tal de construir un classificador més acurat, s'ha utilitzat un algorisme que té en compte la informació del context dels elements de les imatges, els Conditional Random Fields (CRFs).

A la figura 4 es pot observar un exemple de CRF. On els cercles taronges representen les connexions entre nodes veïns, els blaus entre nodes més distants, i els quadrats son les diferents energies, ja siguin dels propis nodes, o bé de l'energia generada a partir de les connexions de 2 nodes diferents.

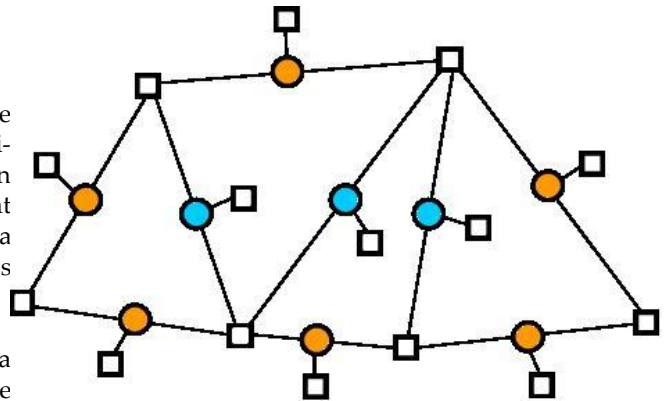


Figura 4: Representació de CRF

El classificador CRF es defineix sobre un grup de nodes, (corresponent en aquest cas al que anteriorment s'havia anomenat elements de la imatge). Aquests tenen potencials unitaris -la probabilitat de ser d'una classe o una altre- i potencials en els eixos -en les connexions entre aquests nodes- que determinen la similitud existent entre ells. Els potencials dels eixos són la característica diferencial entre un classificador normal i un CRF ja que permeten tenir en compte el context.

Concretament el classificador ha sigut construït mitjançant un DenseCRF[8], basat en l'estudi fet pel grup de recerca d'Stanford University, i on la principal diferència amb altres classificadors CRF és que aquest es un model CRF on la connexió entre nodes és completa, és a dir, tots els píxels de la imatge es relacionen amb tota la resta, aconseguint resultats molt favorables. Durant la construcció de l'algorisme s'explicarà més en detall com s'ha aconseguit aquesta connexió com s'observa a la figura 6.

### 3.1.4 Matriu de confusió

La matriu de confusió-figura 5- conté informació sobre les prediccions realitzades per un mètode o sistema de classificació, comparant pel conjunt d'individus de la taula d'aprenentatge o de testing, la predicció donada contra la classe a la que aquests realment pertanyen.

		Predicció	
		Cotxe	Cel
Valor Real	Cotxe	a	b
	Cel	c	d

Figura 5: Matriu de confusió

A l'esquerra hi ha la classe a la que realment pertanyen els elements, en aquest cas cotxe i cel, i a la dreta a quina classe han sigut classificats.

- a = True positive, cotxes que han sigut classificats correctament a la classe Cotxe
- b = False negative, predicció amb error, cotxes que han sigut classificats a la classe Cel

- $c = \text{False positive}$ , predicció amb error, regions que són de Cel que han sigut classificades com a Cotxe
- $d = \text{True negative}$ , Cel que ha sigut correctament classificat com a No-Cotxe, en aquest exemple al tindre només 2 classes, correctament classificats com a Cel.

### 3.2 Software emprat pel desenvolupament

- I. Per fer un primer apropament als algorismes CRF, ha sigut utilitzada una llibreria en C++ per models discrets de grafs anomenada openGM[9]. Aquesta permet també operacions distributives sobre els models construïts. openGM conté una sèrie de funcions integrades que faciliten la creació d'uns primers models per poder generar exemples simples de classificadors CRF.
- II. OpenCV[10]: Llibreria de programació enfocada en visió per computador en temps real, desenvolupada per Intel russia research center, facilitant la tasca de integració en el FrameworkADAs
- III. FrameworkADAs: Software desenvolupat pel grup ADAS del centre de visió per computador, i d'on s'extreuen les dades d'entrada del classificador.

## 4. METODOLOGIA

Per a la realització del projecte s'ha escollit seguir la metodologia de disseny del software SCRUM[11], ja que ens permetia desglossar el projecte en diferents subobjectius concrets, ordenats segons les necessitats de cada tasca a realitzar.

La planificació es va dur a terme mitjançant un diagrama de grantt a l'inici del projecte, fent les modificacions pertinents a l'arribar al final de cada tasca realitzada i durant les reunions de seguiment.

El projecte ha sigut desenvolupat en 3 fases principals:

### 4.1 Fase de documentació e introducció a CRF

Primerament el treball realitzat ha consistit en una profunda recerca en classificadors basats en grafs, especialment el model anteriorment explicat CRF. Llegint les publicacions sobre classificadors i treballant en la comprensió dels conceptes amb els que encara no estava familiaritzat.

Seguidament, s'ha configurat l'equip de desenvolupament del software. Així, s'ha instal·lat la llibreria openGM,

junt amb tots els components previs necessaris, com Cmake [12] (grup d'eines per a la creació i test de software). Per a l'ús d'exemples i experimentar amb la llibreria, primer es va fer una extensa consulta del tutorial openGM, on s'explica detalladament no només la instal·lació, sinó també les funcions incloses i els paràmetres modificables, així com una sèrie d'exemples.

### 4.2 Exemple de CRF mitjançant OpenGM

A continuació s'exposarà un dels exemples en que s'ha estat treballant incloent-hi les dades d'entrada i sortida i els paràmetres modificats.

Es simula un cas mol senzill, una imatge amb soroll. Així primer es genera una matriu amb números aleatoris, que posteriorment s'hauran d'ordenar.

Les connexions entre els nodes seran les d'un CRF d'adjacència -connexions entre nodes correlatius- com es pot observar a la figura 6.

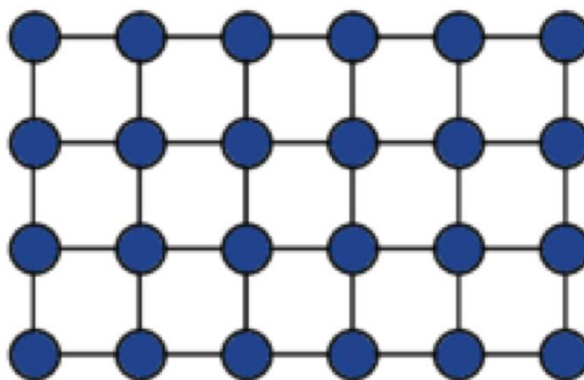


Figura 6: Connexió entre nodes CRF d'adjacència

L'entrada d'aquest exemple, espai, serà una matriu de  $30 \times 30$ , i es definirà el número de classes a 2 (0 o 1). En altres paraules, les dades d'entrada seran un espai definit de  $30 \times 30$  i dues possibles classes per node, contenint valors de 0 i 1 aleatòriament distribuïts.

La construcció del model gràfic estarà formada pels següents paràmetres:

- *Explicit Function*: Funció que defineix l'energia de cada node. Aquesta està basada en una implementació eficient de vectors multidimensionals i incorpora paràmetres que permeten la modificació del temps d'execució (número d'iteracions).
- *Potts*: Funció que defineix l'energia entre nodes (edges). Si dos nodes tenen el mateix valor majoritari, pertanyen a la mateixa classe, se'ls hi assigna un factor, *Potts*, el qual és el valor de la connexió a l'espai d'aquests dos nodes. En el cas

que els dos nodes no tinguessin el mateix valor majoritari, pertanyent a classes diferents, es penalitzarà l'energia entre nodes ja que presumptament serà on ha aparegut el soroll en la imatge.

- *Adder*: Funció que optimitza al final del procés, i que forma part de l'execució de la inferència[13]. Aquesta és el procés d'estimació de la classificació més probable utilitzant ambdós energies, la del node i la de la connexió, en aquest cas s'ha utilitzat *belief propagation*[14]

Així, com a resultat final obtindrem una matriu de 30x30 ordenada, contenint a la meitat esquerra els 0, i a la meitat dreta els 1, que seria la "simulació" d'haver aconseguit eliminar el soroll en la imatge.

### 4.3 Algorisme principal DenseCRF

Inicialment s'havia plantejat la construcció del classificador CRF mitjançant la llibreria openGM. Aquesta ha sigut de gran ajuda per l'assimilació profunda dels diversos conceptes darrera aquests tipus de classificadors i s'hi han construït diversos exemples a través seu. Tanmateix, s'ha decidit utilitzar DenseCRF per les raons que s'exposaran a continuació. Aquest algorisme, desenvolupat per 2 estudiants de la universitat d'Stanford, consisteix en un CRF on la connexió entre nodes no existeix només pels nodes veïns al node objectiu, sinó que es té en compte tots els nodes de la imatge com s'aprecia a la figura 7.

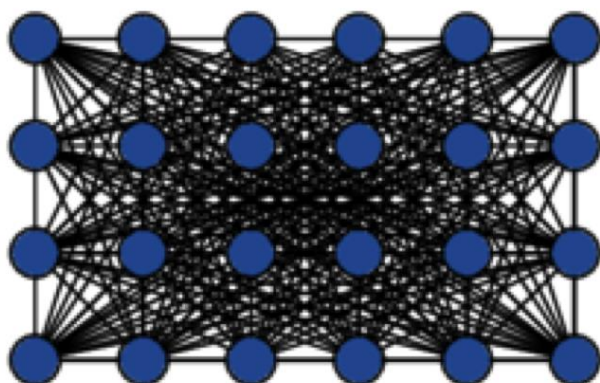


Figura 7: Connexions entre nodes en DenseCRF

DenseCRF s'ha triat per solucionar un inconvenient que presentaven els CRF per adjacència a causa que en aquests els nodes connectats són exclusivament els veïns directes.

Als CRF per adjacència el funcionament de connexions es veu limitat a les connexions locals immediates. En altres

paraules, si s'està classificant un píxel concret, només es té en compte els seus veïns directes, amb els que està directament connectat. Aquest tipus de connexió genera un suavitzat o arrodoniment de tots aquells píxels que formen part de les "fronteres" entre una classe i una altre, com es pot observar a la figura 8.

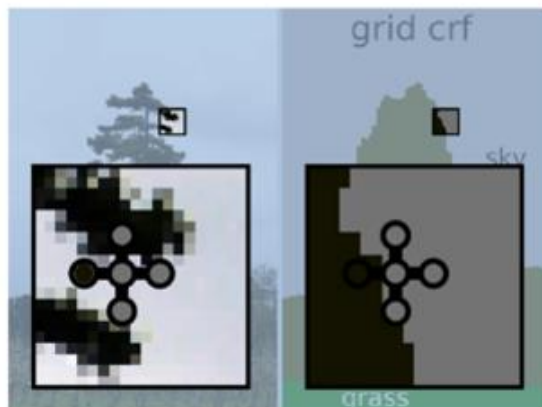


Figura 8: Observació connexió local CRF

A la creu del zoom de la figura 6 es pot com a l'esquerra el píxel negre és assignat a la classe arbre, a causa de la proximitat als seus respectius nodes veïns, mentre que la resta de píxels son assignats a la classe cel. Això fa que les branques estretes de l'arbre hagin sigut transformades per una forma més uniforme segons la proximitat de la classe majoritària al node objectiu.

Per tant, això condueix a l'arrodoniment, anomenat *smoothing*, dels píxels que formen part de les fronteres entre classes, transformant la forma original de l'arbre de la figura 9 de l'esquerra, on es pot observar cel entre les branques de l'arbre, en la forma més uniforme resultant de l'etiquetatge de les classes que observem en la part de la dreta de la imatge.

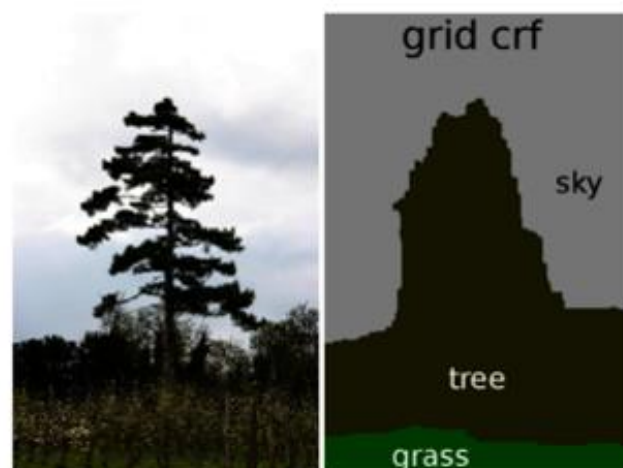


Figura 9: Smoothing a causa de la classificació CRF

Per tant, per tal d'evitar l'*smoothing* s'han fet servir els DenseCRF, permetent la connexió entre tots els nodes de la imatge assignant uns pesos diferents, segons diferents criteris que s'explicaran en detall posteriorment. Normalment el cost computacional resultant del procés de connexió amb tots els píxels de la imatge seria computacionalment inviable, per això s'ha utilitzat una combinació lineal de Gaussians[15] per a l'assignació de l'energia entre nodes que serà explicat en detall en la pròxima secció.

## 5. DISSENY DE L'ALGORISME

### 5.1 Dades d'entrada

Com s'ha explicat prèviament en l'estat de l'art, les imatges usades durant l'entrenament de l'algorisme pertanyen a Camvid dataset. Després de ser processades pel FrameworkADAS, treballarem amb 4 dades diferents per cada imatge:

- Les pròpies imatges, transformades a format ppm per a un millor processat.
- Les energies dels nodes pertanyents a cada superpíxel, en el nostre cas formen una matriu de 880 superpíxels \* 6 columnes, que són l'energia corresponent a ser classificat a cada 1 de les classes predeterminades.
- Arxiu anomenat segments, que és la corresponent indexació entre els píxels originals de la imatge i els nous superpíxels. Bàsicament ens serveix per saber quins píxels formen part de cada superpíxel, i per tant, les energies dels nodes de cada píxel són les del superpíxel del qual formen part.
- Anotacions, que és el Ground truth, la classificació ideal per cada element de la imatge.

S'ha preparat el programa en c++ perquè només sigui necessari introduir els noms de les carpetes de cada 1 dels arxius anteriorment esmentats, així com el nom de l'arxiu concret.

El primer pas del programa és el de llegir els arxius anteriorment esmentats, així com introduir els paràmetres per la creació del CRF. En el nostre cas, és la mida de les imatges en superpíxels(960-720-6), on 960x720 és el tamany de les imatges en superpíxels, i 6 el número de classes que volem classificar.

### 5.2 Processat de les dades d'entrada i energia dels nodes

El nostre algorisme, al ser classificador DenseCRF, treballa al nivell de píxel, per tant fem les transformacions necessàries entre l'arxiu contenint les energies de cada superpíxel, i l'arxiu segments, aconseguint així l'energia de cada píxel a ser classificat en una classe o altre, que a partir d'ara anomenarem *Unary energy*, per diferenciar-la més fàcilment de l'energia entre connexions.

Amb la *Unary energy*, ara ja som capaços de definir els

passos per la construcció d'energia entre connexions mitjançant la combinació lineal de Gaussians.

### 5.3 Construcció de l'energia entre connexions

El cost computacional de relacionar l'*Unary energy* de tots els píxels, així com la relació de colors i l'eliminació del soroll seria computacionalment inviable com s'ha comentat en l'explicació dels DenseCRF, per això utilitzem la combinació lineal de distribucions Gaussians.

La construcció dels *edge potentials*-energia de la connexió entre nodes- mitjançant la distribució Gaussiana és construïda de la següent manera com es pot observar en la fórmula de la figura 10:

$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^K w^{(m)} k^{(m)}(\mathbf{f}_i, \mathbf{f}_j)$$

Figura 10: Fórmula per la construcció edge potentials

- Funció de compatibilitat entre classes: Donats 2 nodes, mira l'*Unary energy* de cada un d'ells, construint una matriu de compatibilitat de 6x6 - les 6 classes de cada un dels nodes-, per assignar-ne l'energia de compatibilitat entre classes.
- El terme anteriorment obtingut, és multiplicat per la combinació lineal dels següents kernels Gaussians i es defineix sobre l'espai generat de les connexions entre píxels.
  - I. Kernel de similitud: Utilitzant la imatge d'entrada, força els píxels propers amb el mateix color a ser assignats a la mateixa classe.
  - II. Kernel Gaussià de suavitzat: Aquí s'aplica la distribució Gaussiana, utilitzada per reduir el soroll de la imatge, afavorint l'assignació dels píxels propers a la mateixa classe.

### 5.4 Inferència i aprenentatge

Generades les energies anteriorment explicades, procedim a la inferència. Aquesta consisteix en un procés d'optimització on es minimitza l'energia i l'assignació dels píxels a cada una de les classes esmentades.

La última de les dades d'entrada per utilitzar és la de les anotacions, és a dir, el *Ground Truth*. Gràcies a les quals l'algorisme pot "apreciar" com ha sigut la seva classificació envers la òptima. Això ajuda a variar els termes que s'han d'aprendre per millorar l'algorisme classificador, fent les variacions pertinents al pes o importància que aquestes tenen. Els termes a variar són:

- Paràmetres Unary: Variacions pertinents en les energies dels nodes.
- Paràmetres en les energies entre connexions.
- Paràmetres en els Kernels anteriorment esmentats.

### 5.6 Integració en FrameworkADAs

El disseny de l'algorisme classificador ha sigut creat de forma modular, preparant les funcions que depenguin de les dades d'entrada amb les mateixes dependències que al Framework al que està destinada la integració, i de forma que s'utilitzin les mateixes dades per a tot el framework, no havent d'especificar cap variable d'entrada exclusivament pel funcionament del classificador CRF.

Per aconseguir la correcta integració, les classes creades en l'algorisme *DenseCRF*, originalment creades en c++, han sigut transformades a mòduls de python utilitzant la llibreria *boost*[17], aconseguint així una fàcil integració en el Framework.

## 6. RESULTATS

### 6.1 Test

En el test, agafarem els paràmetres o pesos apresos durant tota la fase de *training* i els posarem a prova amb una execució d'una nova imatge, no pertanyent a les del conjunt d'imatges de train.

Aquest test ens donarà una classificació que podrem comparar amb la del *Ground Truth*. A partir d'aquesta comparació, podrem generar una matriu de confusió[16] que ens permetrà analitzar els resultats obtinguts de cada classe.

### 6.2 Matriu de confusió

Tant l'entrenament com el test de l'algorisme classificador s'ha executat sobre les imatges de la base de dades pública anteriorment esmentada, que ens servirà per contrastar les imatges creades a partir del nostre classificador, amb la imatge corresponent al seu *Ground Truth*.

L'entrenament de l'algorisme ha sigut executat sobre 15 imatges, i el test sobre 650 imatges, per tindre una idea general de la precisió de classificació en diferents moments de l'entorn de la conducció. La matriu de confusió resultant del test sobre les imatges es troba a la figura 11.

A l'Annex està inclosa amb un tamany més gros per apreciar-ne millor els detalls.

```
In [6]: display_conf(confusion_matrix, label_list, acc_gb_acc, avg_acc)
```

Label names	Sky	Building	Road	Sidewalk	Vertical	Dynamic	Void	Accuracies
Sky-GT	68275256.0	1098836.0	623.0	0.0	1666208.0	8188.0	0.0	96.1770827078
Building-GT	689578.0	97272489.0	411810.0	82461.0	1355230.0	824453.0	0.0	96.7346241753
Road-GT	254.0	2295766.0	138591868.0	266196.0	50889.0	424853.0	0.0	97.7271719394
Sidewalk-GT	18.0	3531166.0	18884661.0	4669642.0	68109.0	118287.0	0.0	17.1225597238
Vertical-GT	2797663.0	41265481.0	1875522.0	183359.0	23568840.0	111560.0	0.0	33.2844663927
Dynamic-GT	8927.0	11820966.0	4144239.0	54309.0	282115.0	9875995.0	0.0	37.7116652067
Void-GT	829959.0	18523917.0	3439479.0	134560.0	1460777.0	1376591.0	0.0	0.0

Global Accuracy : 74.7424209493  
Average Accuracy : 63.1262501989

Figura 11: Matriu de confusió DenseCRF

Com es pot observar, les 3 classes principals, que agrupen els millors resultats són cel, carretera i edificis. Això es a causa de que són els conjunts més grans de la imatge i apareixen de forma molt uniforme, facilitant-ne la classificació. Per exemple un vehicle es pot percebre de dife-

rents formes, de cara, de costat, amb diferents colors, mentre que el cel, la carretera i els edificis tenen unes característiques semblants a totes les imatges.

La precisió global de l'algorisme és d'un 74,74%, en gran part a causa del 96,1%, 96,7% i 97,72% de les precisions obtingudes pel cel, edificis i carretera respectivament, que contenen la majoria dels píxels que formen la imatge, de mitjana la precisió és del 63% a causa de les classes vorera vertical i dinàmic com s'observa a la figura 12.

Nom de classe	Precisió
Cel	96.177
Edificis	96.734
Carretera	97.727
Vorera	17.122
Vertical	33.284
Dinàmic(cotxes)	37.71
<b>Global accuracy</b>	<b>74,74%</b>
<b>Average accuracy</b>	<b>63,14%</b>

Figura 12: Matriu de confusió

### 6.3 Imatges generades

A continuació es presenten les imatges generades a partir de la classificació del *DenseCRF* junt amb el *Ground Truth* corresponent a l'etiquetatge de la base de dades Camvid.

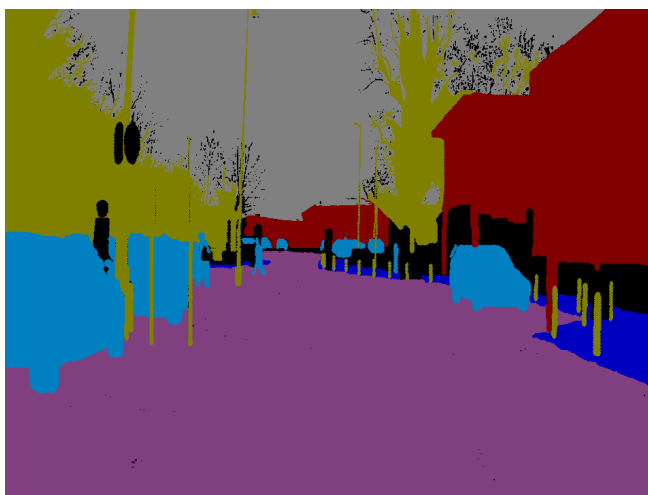


Figura 13a: Imatge generada per la classificació DenseCRF

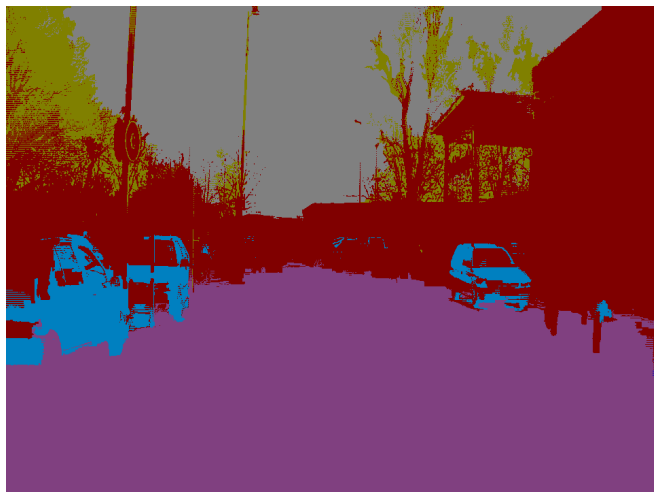


Figura 13b: Imatge Ground Truth

La primera imatge és la representació de la classificació obtinguda per l'algorisme, amb un acceptable encert per determinar el terra, el cel i els edificis, encara que en el cas dels edificis s'ha barrejat en alguns punts amb la classificació vertical (vegetació). En el cas dels cotxes/vianants, els que estan pròxims a la càmera els distingeix bé, encara que hi ha problemes per etiquetar les parts on canvien els colors (com per exemple els vidres, o la part de les rodes). A l'annex s'inclou una altra mostra de classificació de l'algorisme.

## 7. CONCLUSIÓ

### 7.1 Resultats obtinguts

El classificador DenseCRF ha classificat amb molt bons resultats les 3 classes amb més aparició per imatge-cel, carretera i edificis- però la precisió classificant les classes formades pels elements de la imatge més petits no ha sigut tan bona. En especial la classe referent a la vorera, que apareix en les imatges en menor mesura, i en unes formes molt estretes, dificultant així l'entrenament referent a aquesta classe.

Un 74,74% de precisió global, i un 63,14% de mitjana entre les classificacions de les diferents classes, són bons resultats, però ampliant el conjunt d'entrenament a una majoria d'imatges, reservant un marge acceptable pel conjunt de test, els resultats en les classificacions dels vehicles, voreres i vertical (vegetacions i senyals) millorarien considerablement. El cost computacional de l'entrenament d'un conjunt d'imatges usant l'algorisme DenseCRF és elevat i per això s'ha hagut de limitar a un petit conjunt d'imatges.

Els classificadors amb *Conditional Random Fields*, i concretament els DenseCRF anteriorment explicats són una

bona eina en el reconeixement d'imatge. Podent ser una de les branques a tenir en compte en la construcció de d'algorismes classificadors pel desenvolupament de la conducció autònoma.

### 7.2 Treball futur

Ampliant el conjunt d'entrenament a uns marges raonables envers el conjunt de test es traurien resultats interessants, presumptament amb una precisió més elevada, així com fer el training amb una altra base de dades pública i contrastada, obtenint coneixements que permetin analitzar possibles aspectes a millorar en l'algorisme DenseCRF.

També s'afegiria la classe void (es podria entendre com la resta, lo que no pertanyi a les classes a classificar), així com més classes per aconseguir un classificador més complet.

## AGRAÏMENTS

Agrair a la meua tutora Gemma Sánchez per la guia oferida durant tot el projecte, especialment al principi, sent un projecte amb tants aspectes teòrics a assolir.

Un especial agraïment també per Prassanna Ganesh Ravishankar, per tota l'ajuda amb els detalls tècnics sobre la implementació del classificador, així com la preparació de l'entorn de desenvolupament.

## BIBLIOGRAFIA

- [1] <http://www.nowozin.net/sebastian/papers/nowozin2011structured-tutorial.pdf>
- [2] [http://scikitimage.org/docs/dev/auto\\_examples/plot\\_segmentations.html](http://scikitimage.org/docs/dev/auto_examples/plot_segmentations.html)
- [3] <http://www.colormatters.com/color-and-design/color-systems-rgb-and-cmyk>
- [4] <http://www.vlfeat.org/overview/hog.html>
- [5] <http://www.vlfeat.org/overview/sift.html>
- [6] [http://www.embedded-visibility.com/sites/default/files/apress/computervisionmetrics/chapter7/9781430259299\\_Ch07.pdf](http://www.embedded-visibility.com/sites/default/files/apress/computervisionmetrics/chapter7/9781430259299_Ch07.pdf)
- [7] [http://resources.mpi-inf.mpg.de/conferences/up2013/up2013\\_files/up2013-abstracts/kondermann/daniel-kondermann.pdf](http://resources.mpi-inf.mpg.de/conferences/up2013/up2013_files/up2013-abstracts/kondermann/daniel-kondermann.pdf)
- [8] Krähenbühl, P., & Koltun, V. (2012). Efficient inference in fully connected crfs with gaussian edge potentials. *arXiv preprint arXiv:1210.5644*.
- [9] <http://hci.iwr.uni-heidelberg.de/opengm2/>
- [10] <http://docs.opencv.org/opencv2refman.pdf>
- [11] <http://www.softeng.es/es-es/empresa/metodologias-de-trabajo/metodologia-scrum.html>
- [12] <http://www.cmake.org/documentation/>
- [13] [http://www.cse.psu.edu/~rtc12/CSE586/lectures/06\\_Learning\\_And\\_Inference\\_BobEdits.pdf](http://www.cse.psu.edu/~rtc12/CSE586/lectures/06_Learning_And_Inference_BobEdits.pdf)
- [14] <http://www.merl.com/publications/docs/TR2001-22.pdf>



- [15] <http://mathworld.wolfram.com/NormalDistribution.html>  
 [16] <http://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/>  
 [17] <http://www.boost.org/>

## APÈNDIX

### A1. MATRIU DE CONFUSIÓ DENSECRF

```
In [15]: acc,gb_acc,avg_acc = find_statistics(confusion_matrix)
```

```
In [6]: display_conf(confusion_matrix, label_list, acc,gb_acc,avg_acc)
```

Label names	Sky	Building	Road	Sidewalk	Vertical	Dynamic	Void	Accuracies
Sky-GT	68275256.0	1098836.0	623.0	0.0	1606208.0	8188.0	0.0	96.1770827078
Building-GT	609578.0	97272489.0	411810.0	82461.0	1355230.0	824453.0	0.0	96.7346241753
Road-GT	254.0	2295766.0	130591868.0	266196.0	50089.0	424853.0	0.0	97.7271719394
Sidewalk-GT	18.0	3531166.0	18884661.0	4669642.0	68109.0	118287.0	0.0	17.1225507238
Vertical-GT	2797663.0	41265481.0	1875522.0	185359.0	23568040.0	1115860.0	0.0	33.2844663927
Dynamic-GT	9927.0	11820966.0	4144239.0	54309.0	282115.0	9875595.0	0.0	37.7116052067
Void-GT	829959.0	10523917.0	3439479.0	134560.0	1460777.0	1376591.0	0.0	0.0

Global Accuracy : 74.7424209493  
 Average Accuracy : 63.1262501909

### A2. IMATGE DENSECRF I GROUND TRUTH

A l'esquerra la imatge generada amb DenseCRF  
 A la Dreta el Ground Truth de la imatge

