

Detecció de les mans al volant durant la conducció

Inmaculada Navas Peña

Resum– En l'actualitat ens trobem en un estat entre la conducció tradicional i la conducció autònoma. Per a poder donar completament el salt d'una a l'altre s'ha de passar per un estat intermedi en el qual s'afegeixen elements d'automatització durant la conducció tradicional. En particular, aquest projecte detecta les mans sobre el volant en diferents entorns, el qual ajuda a aportar seguretat durant la conducció, ja que el temps de reacció d'una persona envers una situació de risc augmenta si no té les dues mans a sobre. El projecte parteix d'un projecte anteriorment realitzat per dos alumnes de la UAB l'any passat, els quals també detectaven mans però d'una manera poc automàtica, per tant, en aquest projecte es pretén millorar el sistema de detecció fent servir diferents mètodes de filtratge com background subtraction, filtratge pel color de pell, operacions morfològiques, filtratge suavitzat i filtratge per àrea. Finalment s'analitzaran les dades.

Paraules clau– Visió per computador, detecció mans, anàlisi i processament d'imatges, Matlab.

Abstract– Nowadays we are in a state between traditional driving and autonomous driving. In order to be able to pass from one to the other one it is necessary to go through an intermediate state in which automation elements are added during traditional driving. In particular, this project detects hands on the steering wheel in different environments, which helps to provide safety during driving because the reaction time of a person facing a risky situation increases if he does not have both hands on the steering wheel. This project is based on a project previously carried out by two students of the UAB last year, which also detected hands but not in a very automatic way, therefore, this project aims to improve the computer vision detection system using different filtering methods as background subtraction, skin color filter, morphological operations, blurring filtering and area filtering. Finally, the data will be analyzed.

Keywords– computer vision, hand detection, analysis and image processing, Matlab.



1 INTRODUCCIÓ

ACTUALMENT ens trobem en un moment molt específic en el que, cada vegada més, s'està intentant aconseguir automatitzar tasques que abans era impensable que les fes algú que no fos una persona. Una de les tasques que més freqüentment desenvolupa l'ésser humà és la de la conducció amb tots els riscos que aquesta pràctica comporta (l'any passat hi van haver 1.160 morts en carretera a Espanya [1]).

Si s'aconseguís automatitzar la conducció, s'eliminaria el factor humà que indueix a l'error i es podria arribar a

un estat on els cotxes poguessin circular autònomament de manera ordenada, on els accidents es poguessin evitar de forma automàtica i a temps real.

Per aconseguir arribar a aquest punt desitjat s'han de tenir en compte moltes variables dins de la conducció, entre d'altres, la que ens ocupa en aquest treball, analitzar les accions que realitza el conductor durant la conducció, que formaria part del pas intermedi entre la conducció tradicional i la conducció autònoma.

El que es pretén realitzar en aquest treball serà, a partir d'un data set d'imatges extretes d'una càmera fixa dins el cotxe enfocant al volant, tant en entorns controlats com no controlats, saber quantes mans hi ha en cada moment al volant. D'aquesta manera podrem identificar si el conductor està utilitzant una, dues o cap mà durant la conducció.

Per a poder fer-ho, s'hauran d'utilitzar diversos mètodes de filtratge d'imatges, ja que per identificar correctament les mans haurem de ser capaços de filtrar tot el fons de la

- E-mail de contacte: inmaculadaconc.navas@e-campus.uab.cat
- Menció realitzada: Computació
- Treball tutoritzat per: Aura Hernández (Ciències de la computació)
- Curs 2016/17

imatge que no correspongui a les mans com per exemple el volant, quadre de comandaments del cotxe, i qualsevol element que hi pugui aparèixer en l'escena.

Aquest treball també es podria extrapolar al camp de la seguretat viària, ja que, una vegada es sap quantes mans hi ha al volant mitjançant un sistema de visió per computador, es pot saber si el conductor està emprant les dues mans o no, ja que no utilitzar-les comporta actualment una multa de 80 euros per part de la DGT i a més és una infracció bastant comuna que derivaria en un possible accident al no poder controlar el cotxe correctament en un moment crític per no tenir les dues mans al volant.

2 ESTAT DE L'ART

Aquest projecte parteix d'un altre projecte fet l'any passat amb el nom de "Hands and foot activity analysis while driving"[3], els quals a l'hora de detectar les mans a les imatges comparaven amb una màscara del volant si hi havia sofert algun canvi amb un background subtraction. Aquest mètode era molt poc automàtic, ja que el que feien era comparar només l'àrea del volant sense mans, amb l'àrea del volant amb mans, i d'aquesta manera saber si hi havia una, dues o cap mà.

Per aquest motiu s'ha volgut millorar aquest sistema, per a fer-lo més automàtic i no haver de dependre d'una màscara del volant.

Existeixen sistemes relacionats amb aquest projecte tal com el que es pot llegir a l'article LEE: A Photorealistic Virtual Environment for Assessing Driver-Vehicle Interactions in Self-Driving Mode [5], el qual consisteix en un entorn de conducció virtual per a simular sistemes automatitzats de conducció de manera segura.

Donat que els simuladors comercials que existeixen avui dia són molt costosos, com per exemple el simulador de Toyota [2], és molt interessant desenvolupar aquesta tasca amb mitjans molt més assequibles i de baix cost.

Aquest treball podria formar part d'un sistema d'aquest caire, ja que es fan experiments en un entorn controlat amb càmeres de baix cost i a més s'estan supervisant les mans sobre el volant com a mètode de seguretat del conductor.

3 OBJECTIUS

Aquest projecte pretén ser una peça clau més per arribar a assolir un objectiu molt més gran, que seria el pas intermedi entre la conducció tradicional tal com la coneixem i la conducció autònoma. Un altre punt clau d'aquest projecte és que està realitzat amb components assequibles, el que comportaria que el cost de fabricació d'un cotxe d'aquestes característiques fos el més econòmic possible.

Per això, aquest projecte es centrarà en la detecció de les mans durant la conducció, és a dir, partint d'imatges aconseguides mitjançant càmeres convencionals enfocades cap al volant, aconseguir identificar-les en diferents entorns.

En un primer moment el que es va pretendre va ser millorar els resultats que es van obtenir al treball de final de grau de l'any passat, els quals treballaven amb un data set tret per ells mateixos en un ambient controlat com és el laboratori, on a les imatges només apareixia el volant negre i les mans del conductor.

A més, a part de les imatges enfocades al volant, en aquest treball també es van extreure imatges dels peus, i en un primer moment també es va voler millorar els resultats de la identificació dels peus als pedals, però per falta de temps i donat que el resultat dels peus era millor que els de les mans, es va voler atacar la part més crítica que eren les mans.

Per tant, com la part crítica eren les mans, es va voler treballar més a fons aquesta part afegint un data set més complex que pertany al Laboratory for Intelligent & Safe Automobiles (LISA) [4], el qual representava també un desafiament, ja que a diferència amb el data set inicial, aquestes imatges no estaven preses en un entorn controlat, sinó en un entorn real amb canvis d'il·luminació constants i en un cotxe real, el que comporta que hi apareguin elements més complexos, com són el quadre de comandaments, un volant amb diferents tonalitats i qualsevol element més que hi pugui aparèixer. Per tant, per a la realització del sistema d'anàlisi de les mans al volant, es compta amb dos data sets d'imatges diferents, el qual també ens aportarà robustesa a l'hora d'implementar aquest sistema en altres entorns diferents.

L'objectiu principal del projecte, per tant, és:

- Millorar els resultats de la identificació de les mans del projecte anterior.

Com a conseqüència d'això, els sub objectius a assolir són:

- Aconseguir un bon pretractament de les imatges perquè els mètodes a utilitzar tinguin un resultat exitós.
- Filtrar les mans en imatges capturades en un entorn controlat, com són les imatges del data set original.
- Filtrar les mans en imatges en un entorn no controlat i amb un fons complex, com són les imatges del data set LISA.
- Saber en cada moment quantes mans hi ha sobre el volant.
- Analitzar els resultats.

En cas que els resultats amb el data set LISA siguin prou bons, hi ha la possibilitat d'enviar els resultats a la web des d'on es van extreure les imatges perquè aparegui el projecte amb els resultats obtinguts.

4 METODOLOGIA

Per a la realització d'aquest projecte es va escollir Matlab com a llenguatge de programació, ja que ofereix moltes possibilitats per al que es vol assolir.

El projecte es pot desglossar en tots els mètodes realitzats per tal de filtrar les mans de les imatges el màxim possible. Els mètodes emprats han sigut:

1. Pretractament de la imatge.
2. Background subtraction amb threshold
3. Filtratge per color de la pell.
4. Operacions morfològiques.

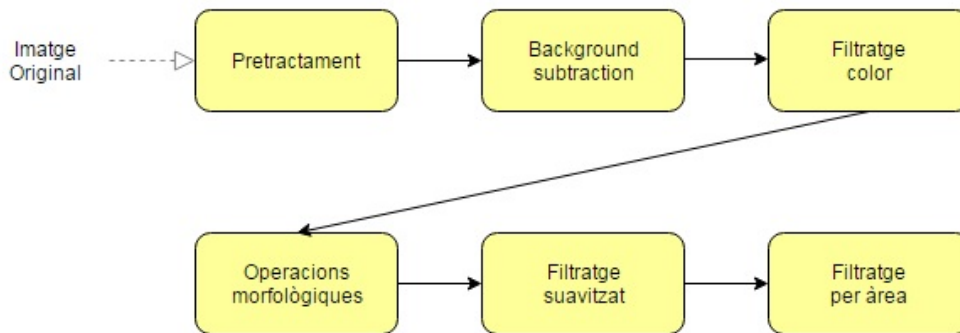


Fig. 1: Esquema dels procediments per detectar mans

5. Filtratge de suavitzat.
6. Filtratge per àrea.
7. Tractament dels labels per extreure l'etiquetatge.
8. Anàlisis de resultats.

Els quals es poden visualitzar a l'esquema de la figura 1.

4.1 Pretractament de les imatges

Un dels passos que més repercussió tindrà durant tot el projecte és el pretractament de les imatges. Una mala elecció del tracte que li volem donar ens condicionarà durant tota la programació.

D'una banda hi han les imatges del data set del projecte de l'any passat, on apareixen mans sobre un volant negre. Aquestes imatges no necessiten cap mena de pretractament, ja que no tenen excés de soroll com les del data set LISA.

D'altra banda, en les imatges del data set LISA apareix una escena amb un volant de tons grisos amb parts marrons una mica similars a la tonalitat de la pell, les mans al volant (quan hi apareguin) i un quadre de comandament típic de cotxe. D'aquesta escena el que ens interessa només és l'àrea on està el volant, i com sabem que la càmera del cotxe és fixa, retallarem les imatges amb la funció `imcrop`[7] i només ens quedarem amb la regió que ens interessa, així traiem soroll.

Una vegada tenim les imatges preparades per ser processades, s'hauria de fer una extracció dels píxels amb el valor del color de la pell perquè es filtrin les imatges. Amb el data set del projecte de l'any passat funciona aquest plantejament, però amb el data set LISA no, ja que hi ha diversos elements de la imatge que comparteixen un valor similar a la tonalitat de la pell, tal com es pot veure a les figures 2 i 3, és per això que es realitzarà un procés més robust per a poder filtrar bé les mans d'aquest data set.

4.2 Background subtraction

Es realitzarà un background subtraction tenint per referència com a fons un frame del vídeo on no hi apareguin les mans a sobre. Amb aquest frame es realitzarà el background subtraction amb les imatges que volem filtrar.

Si es realitzés un background subtraction sense tractar prèviament les imatges, és a dir, restant les dues imatges

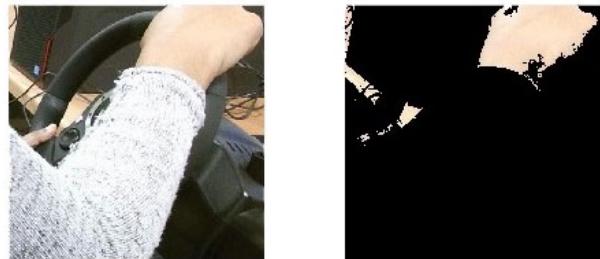


Fig. 2: Extracció color pell amb data set inicial



Fig. 3: Extracció color pell amb data set LiSA

directament, quedaria una cosa semblant a la figura 4. Per aquest motiu és important tractar bé les imatges.

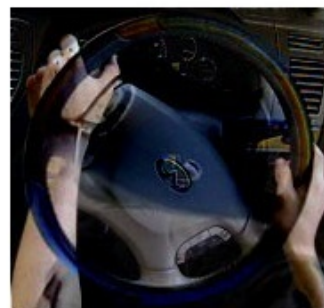


Fig. 4: Background subtraction sense tractament

Perquè sigui exitós aquest procés, primerament es passaran les imatges a escala de grisos, tant les que es volen filtrar com la que farà de background amb la funció `im2gray()`. A partir d'aquí, es normalitzaran les imatges perquè els valors dels píxels de les imatges estiguin entre 0 i 1.

La fórmula de la normalització és la següent:

$$imgN = \frac{((imgG - \min(imgG(:)))}{(\max(imgG(:)) - \min(imgG(:)))}; \quad (1)$$

És important que a l'hora de normalitzar, les imatges estiguin en tipus de dada double, ja que originalment es troben en uint8 i aquest tipus de dada no accepta floats, per tant, a l'hora de normalitzar es quedaria una imatge completament en negre, ja que un valor que sigui 0.083, per exemple, passaria a ser 0.

Per fer-ho, s'utilitzarà la funció `im2double[8]` de Matlab. El que fa aquesta funció és convertir uint8 a double i normalitzar les dades entre 0 i 1, permetent números float.

Una vegada normalitzades les imatges, es podrà procedir a restar les imatges amb la imatge de background. Aquesta resta serà una resta absoluta. El resultat d'aquest procés es pot observar a la figura 5.

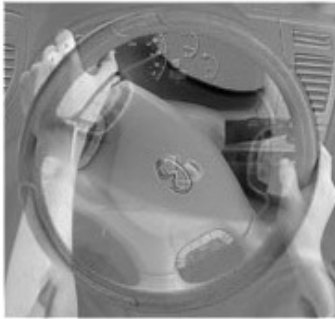


Fig. 5: Background subtraction normalitzat

Com es pot observar, normalitzant i restant les imatges entre si ens retorna massa informació, per tant, a la imatge resultant del background subtraction se li aplicarà un threshold per a quedar-nos amb els valors on es trobin les mans. Per seleccionar aquest threshold ens fixarem en l'histograma de la imatge.

I es binaritzarà la imatge anterior amb el valor del threshold que es cregui convenient a partir de l'histograma, tenint com a resultat una imatge binaritzada com la de la figura 6.



Fig. 6: Background subtraction binaritzat amb threshold

En aquest punt, a la imatge binaritzada se li retornarà el color original amb la funció `bsxfun [12]`, que multiplicarà la imatge binaritzada amb la imatge RGB original convertint

prèviament la binaritzada en el mateix tipus de dada que l'original (uint8).

El que s'aconsegueix amb aquest pas és tenir una imatge on només es veu a color el que s'ha estat filtrant fins ara, és a dir, les regions blanques de la imatge binaritzada. El resultat es pot comprovar a la figura 7.



Fig. 7: Background subtraction a color

Com es pot observar, aquest resultat millora l'anterior aconseguit només amb el background subtraction sense threshold.

Ara sí que es podrà passar a realitzar el filtratge de color de pell, ja que en aquest punt s'ha aconseguit filtrar certs elements que s'assemblaven al color de les mans.

4.3 Filtratge del color de la pell

Una vegada es tenen els resultats del background subtraction anterior, es passa al filtratge del color de la pell per extreure les mans.

En aquest moment el que es té és la imatge original que s'està analitzant menys les parts que no hi han canviat respecte al frame sense mans. El primer pas serà agafar una mostra de pell i aplicar-li una tolerància que serà un valor fix perquè trobi tota la pell possible.

Desglossarem la imatge a analitzar en els 3 canals RGB, i per a cadascun es buscaran els píxels que corresponguin amb la regió de color que busquem, és a dir, que els valors d'aquests píxels estiguin entre el valor mínim del color de la pell menys la tolerància i el valor màxim del color de la pell més la tolerància.

Amb els resultats, es farà una matriu binària on 1 serà on estigui el color buscat i 0 on no.

Finalment aquesta matriu binària es multiplicarà per la imatge original per a retornar el color original només a les parts on s'hagi trobat el valor de la pell.

Com es pot apreciar a la figura 8, en les imatges del data set LISA aquest filtratge funciona força bé sempre que no hi hagi canvis d'il·luminació a la imatge, però quan hi ha aquests canvis d'il·luminació no es poden reconèixer bé aquests valors i en moltes ocasions el centre del volant continua apareixent després d'aquest filtratge, fent que en comptes de només tenir les mans, tenir també una regió al centre que pertany al volant. Quan la llum li donava directament a les mans, tampoc és capaç de reconèixer-les correctament.

Per resoldre aquest problema es va tractar d'executar el codi 'Adaptive Skin Detection' [13], un codi en C++ amb llibreries d'OpenCV que en principi hauria de funcionar bé



Fig. 8: Filtre color



Fig. 10: Operacions morfològiques

per a entorns amb il·luminació canviant. Després d'haver aconseguit instal·lar les llibreries d'OpenCV en Windows, que no és una tasca fàcil [14], es va procedir a executar el codi amb la funció MEX [9] de Matlab i no millorava en res els resultats obtinguts fins ara. Un exemple de resultat d'aquest codi es pot veure a la figura 9.



Fig. 9: Resultat codi Adaptive Skin Detection

A part d'aquest codi, també es va provar de realitzar aquest filtratge però en l'espai de color HSV, però amb els canvis d'il·luminació el valor H de les mans canviava massa provocant que no es pogués detectar. És a dir, després de provar diverses solucions per a les imatges on hi ha canvis d'il·luminació, no es va poder millorar els resultats obtinguts fins aquest moment.

4.4 Operacions morfològiques

El pas anterior que s'ha realitzat genera molt de soroll, ja que pot ser que hi hagi petites regions de píxels on per algun motiu s'hagi trobat aquesta tonalitat de color encara que no pertanyi a la regió de les mans. Aquesta informació de petites regions que no ens interessa pot ser filtrada amb operacions d'erosió i dilatació.

Per tant, es crearà una estructura de tipus disc, i es realitzarà l'operació d'erosió amb la imatge que s'està analitzant. Això el que farà serà fer més petites aquestes regions provocant que les més petites desapareguin.

Per a tornar la forma a la resta de la imatge, es realitzarà el pas invers, que és una dilatació amb la mateixa estructura de tipus disc que farà més grans les regions restants de la imatge.

El resultat d'aquestes dues operacions morfològiques es pot observar a la figura 10.

És a dir, amb aquestes dues operacions morfològiques s'aconsegueix eliminar part del soroll de l'anterior apartat.

4.5 Filtratge suavitzat

Un dels problemes trobats durant el background subtraction va ser que en restar les dues imatges entre si, la part del volant també s'eliminava, així com la part de la mà que estava justament a sobre de la regió del volant, ja que els colors eren massa similars i no diferenciava un objecte de l'altre.

Per tant, en aquest punt s'hauria de tenir una imatge en negre amb dues regions que pertanyen a les mans, encara que aquestes mans es troben tallades per la part del volant fet que no en tinguem 2 regions, sinó més.

Per solucionar aquest problema, el que es va pensar va ser realitzar un suavitzat de la imatge perquè les parts de les mans que estan a cada costat interior i exterior del volant s'ajuntin.

Aquest suavitzat s'aconsegueix mitjançant l'ús de les funcions fspecial [10] i imfilter [11] de tipus replicate. El resultat d'utilitzar-les es pot veure a la figura 11, on a més es pot veure la imatge binaritzada perquè es pugui apreciar millor el resultat d'aquesta operació.



Fig. 11: Filtratge suavitzat

Gràcies a aquest filtratge s'aconsegueix que les regions s'ajuntin tenint una màscara completa de les mans. És a dir, una regió sencera de les mans a l'esquerra i una altra a la dreta (sempre que estiguin les dues mans al volant).

Una altra alternativa que es va provar abans de realitzar aquest filtratge per tal de solucionar el problema de la regió de mans perduda pel volant, va ser realitzar una màscara del volant per agafar les mans d'aquesta regió i després unir-les al resultat de l'anterior apartat, però tenia un inconvenient i era que també tornava informació del volant que ja havíem filtrat anteriorment. Així que es va haver de descartar fins que finalment es va trobar la solució explicada en aquest apartat.

4.6 Filtratge per àrea

L'últim filtratge que es realitzarà per tal de tenir les dues regions de les mans el més filtrades possibles i sense soroll serà un filtratge per àrea amb la funció `regionprops` [15] de Matlab.

El primer pas a realitzar serà una binarització de la imatge, i una vegada s'obtingui una imatge com la de la dreta de la figura 11, s'etiquetaran els elements connectats de la imatge amb la funció `bwlabel` [16] i es calcularan les propietats dels objectes de la imatge amb la funció `regionprops` de matlab.

Amb les propietats obtingudes, es podran mostrar requadres al voltant de totes les regions trobades amb la funció `rectangle`. A la figura 12 es poden apreciar totes les regions identificades en color verd de dos frames diferents.

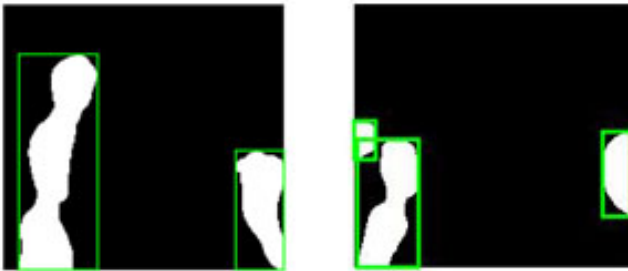


Fig. 12: Regions detectades en color verd

Com es pot observar, en una de les dues imatges apareixen més regions de les que ens interessa. Depenent de la imatge analitzada, hi poden aparèixer també restes del volant o d'altres elements i aquestes regions no ens interessa tenir-les. Per tant, el que es farà serà filtrar aquestes regions per a només tenir en compte les que tinguin una àrea major a 500. D'aquesta manera les regions petites que contenen informació que no ens interessa deixaran de tenir valor.

Per tant, es realitza una cerca de les regions amb àrea menor a 500 i es mostren de color vermell com a la figura 13.

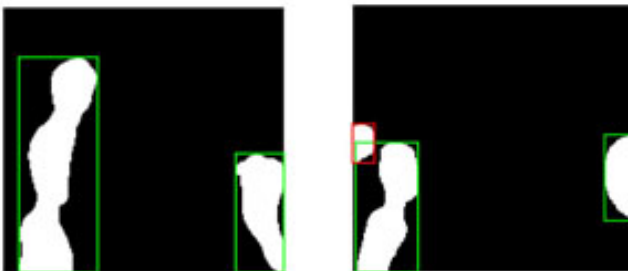


Fig. 13: Regions amb àrea major a 500 en verd, petites en vermell

El resultat que s'hauria de tenir en aquest moment és, una o dues regions verdes, depenent de les mans que hi hagi, i regions petites vermelles que correspondrien a la informació que no ens interessa.

Per tant, es procedeix a eliminar les regions amb àrea menor a 500 perquè les mans quedin totalment filtrades.

Aquest resultat es pot veure a la figura 14.

Com a part final del procés de filtratge, se li retornarà el color original a les parts on la imatge binària sigui blan-



Fig. 14: Resultat filtratge per àrea

ca. Aquesta funcionalitat es realitzarà amb la funció `bsxfun` [12] i es pot observar el seu resultat a la figura 15.



Fig. 15: Resultat del procés amb el seu color original

4.7 Tractament labels per extreure l'etiquetatge

Una vegada es tenen les imatges resultants, el que interessa d'aquestes imatges és saber quantes s'han detectat en cada frame, si una, dues o cap.

Per tal d'obtenir aquesta informació, tornarem a utilitzar la funció `bwlabel` de Matlab que retornarà dos paràmetres, el primer serà una matriu d'etiquetes de la imatge i el segon paràmetre correspondrà al nombre d'objectes que té la imatge.

Amb aquesta dada, es compararà amb les dades de l'etiquetatge i es podrà saber quantes mans s'han detectat correctament.

5 EXPERIMENTS

Durant el treball s'ha treballat amb 3 data sets diferents, els quals han sigut:

- Data set original: Data set de 1366 imatges realitzat per David i Malvika, els estudiants que van realitzar el projecte de l'any passat. 640 d'aquestes imatges són amb en David de conductor i 726 amb la Malvika. Aquestes imatges van ser preses amb una Kinect en un entorn controlat amb un volant Logitech sobre una taula dins un laboratori.
- Data set test LISA: 1065 imatges del Laboratory for Intelligent and Safe Automobiles capturades amb una Kinect. Aquestes imatges tenen diverses particularitats, com canvis d'il·luminació no molt constants i elements amb una tonalitat semblant al color de la pell.

- Data set train LISA: 858 imatges del Laboratory for Intelligent and Safe Automobiles capturades amb una Kinect. La gran diferència amb el data set de test és que les imatges d'aquest data set estan amb constants canvis d'il·luminació. Es podria dir que aquest data set és el més complex dels tres per aquest motiu.



Fig. 16: Mostra de cadascun dels data sets

Una mostra de cadascun es pot apreciar a la figura 16.

Per comprovar la precisió o valor predictiu positiu del procés en els diferents data sets, s'emprarà la fórmula de la precisió tal com es pot veure a la figura de l'equació 2, on el valor de True Positives (TP) serà el nombre d'imatges on les regions detectades correctament pertanyen a les mans, i False Positives (FP) serà el nombre d'imatges on les regions detectades com a mans no són realment mans:

$$Precisio = \frac{TP}{TP + FP} \quad (2)$$

6 RESULTATS

6.1 Data set original

Pel data set original dividirem els resultats en dos, ja que d'una banda tenim 640 imatges amb el David de conductor i de l'altre tenim 726 imatges amb la Malvika.

Per les imatges del data set d'en David, els resultats han sigut els que es poden veure a la taula 1.

TAULA 1: TAULA RESULTATS DATA SET ORIGINAL DAVID

	número	percentatge
detectades	500	78,12 %
no detectades	140	21,87 %

Amb aquestes imatges s'ha aconseguit gairebé un 80% d'encerts. Aquests resultats han sigut molt positius. Una representació gràfica del percentatge d'encerts al llarg del temps es pot veure a la figura 17.

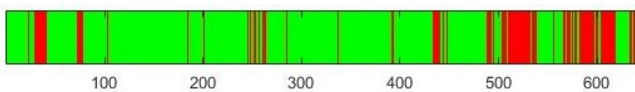


Fig. 17: Barra resultats data set original David

D'aquests resultats, es tenen els següents true positives i false positives de la taula 2:

Per tant, la precisió (o valor predictiu positiu), és del 99,4%.

Un exemple d'aquests resultats el veiem a la figura 18 .

TAULA 2: TRUE POSITIVES (TP) I FALSE POSITIVES (FP) DATA SET ORIGINAL DAVID

TP	FP
497	3



Fig. 18: Imatge exemple dels resultats data set original David

Per part del data set original realitzat per la Malvika, es troben els següents resultats de la taula 3.

TAULA 3: TAULA RESULTATS DATA SET ORIGINAL MALVIKA

	número	percentatge
detectades	474	65,28 %
no detectades	252	34,71 %

Aquests resultats han sigut una mica pitjors, ja que la Malvika porta una polsera i en principi les imatges es tallaven. Per solucionar el problema de què es tallessin, les imatges van haver de passar per un filtratge de suavitzat tal com està explicat en seccions anteriors.

En suavitzar-les, petites regions es van fer més grans pel suavitzat i van crear soroll, fent que en molts casos en comptes de detectar les 2 regions de les mans, detectés 3 regions, les dues mans i una regió petita més.

Una representació gràfica del percentatge d'encerts al llarg del temps es pot veure a la figura 19.

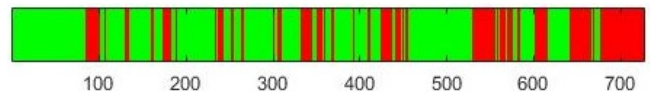


Fig. 19: Barra resultats data set original Malvika

Al final de la barra apareixen molts frames en vermell, això es deu al fet que en aquests frames el conductor s'aixecava del seient i creava interferències amb la càmera i el volant.

Els TP i FP d'aquest data set són els indicats en la taula 4.

Per tant, calculant la precisió o valor predictiu positiu d'aquest data set dona 99,78%.

TAULA 4: TP I FP DATA SET ORIGINAL MALVIKA

TP	FP
473	1

Un exemple de detecció de les mans d'aquest data set es pot veure a la figura 20.



Fig. 20: Imatge exemple dels resultats data set original Malvika

6.2 Data set test LISA

Els resultats d'haver realitzat tots els processos especificats en aquest projecte en el data set test LISA es poden veure a la taula 5.

TAULA 5: TAULA RESULTATS DATA SET TEST LISA

	número	percentatge
detectades	551	51,73 %
no detectades	514	48,26 %

Una representació gràfica del percentatge d'encerts al llarg del temps es pot veure a la barra de la figura 21.

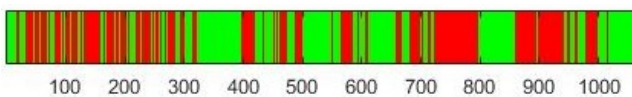


Fig. 21: Barra resultats data set test

D'aquests resultats, hi ha els següents TP i FP de la taula 6.

TAULA 6: TP I FP DATA SET TEST LISA

TP	FP
375	179

Per tant, la precisió d'aquest data set és del 68.05%.

Aquesta precisió es deu als canvis d'il·luminació, ja que en comptes de detectar mans, es detecten regions del volant per què la tonalitat d'aquest canvia, així com la tonalitat de les mans que es tornen molt més blanques i les fa indetectables.

Un exemple de detecció de mans d'aquest data set es pot veure a la figura 22.



Fig. 22: Imatge exemple dels resultats data set test LISA

6.3 Data set train LISA

Finalment, en l'últim i més complex data set, els resultats obtinguts són els que es poden observar a la taula 7.

TAULA 7: TAULA RESULTATS DATA SET TRAIN LISA

	número	percentatge
detectades	290	33 %
no detectades	568	66 %

Una representació gràfica del percentatge d'encerts al llarg del temps es pot visualitzar a la barra de la figura 23.



Fig. 23: Barra de resultats data set train

On d'aquests resultats, tenim els següents true positives i false positives de la taula 8:

TAULA 8: TP I FP DATA SET TRAIN LISA

TP	FP
192	98

Per tant, la precisió d'aquest data set és de 66.20 %. Aquests resultats es deuen al sobre filtratge per culpa de la il·luminació excessiva, que fa que tot canviï de color i no es pugui filtrar correctament. En molts cassos de FP es detectava una de les mans i l'altre regió corresponia a una porció del volant.

Un exemple de detecció de mans en aquest data set es pot veure a la figura 24.

7 CONCLUSIONS I TREBALL FUTUR

L'objectiu principal de millora de resultats de la identificació de les mans respecte al projecte de l'any passat s'ha completat amb èxit amb un filtratge de pell i àrea.

Un dels objectius importants d'aquest projecte, que era saber quantes mans hi havia en cada moment, s'ha completat realitzant diversos mètodes per a aconseguir el major

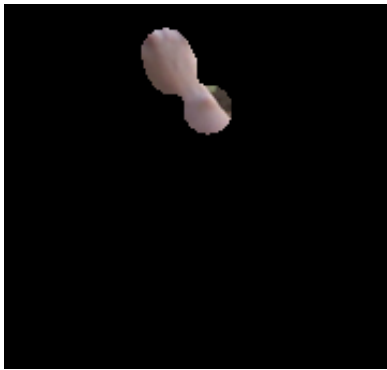


Fig. 24: Imatge exemple dels resultats data set train LISA

filtratge de la imatge perquè a l'hora de comptar regions, només hi haguessin les mans.

Aquesta tasca ha sigut especialment difícil, ja que al data set LISA hi havia regions de la imatge amb un color de píxel molt semblant al de la mà, per tant el mètode del filtratge per color de pell no era suficient per a aquest data set, mentre que pel data set d'imatges del projecte de l'any passat sí que ho era, ja que no tenien aquesta particularitat.

Com tampoc es disposa d'una imatge sense mans al volant per a tots els casos, un background subtraction tampoc era suficient, ja que quan es gira el volant amb les mans, el volant també gira i el centre del volant no està com el del frame que no té mans, per tant el volant també apareixia i dificultava el filtratge de les mans.

Finalment amb un background subtraction amb un threshold i seguidament un filtratge per color de píxel es va aconseguir un resultat bastant bo. I gràcies a aquest resultat i després d'operacions morfològiques i un filtre de suavitzat, es va poder filtrar per regions i en eliminar totes les regions amb una àrea menor a 500, ja teníem les mans filtrades.

Per tant, com hem pogut comprovar, tota la metodologia emprada per a la solució del projecte funciona molt bé quan les condicions lumíniques són normals, és a dir, no hi ha una sobre il·luminació donada pel sol ni ombres generades, i a més, com menys elements tingui el fons de la imatge, millor, ja que no s'haurà de filtrar de més.

En el cas del data set del projecte de l'any passat, els mètodes utilitzats funcionen bastant bé, ja que estan totes les imatges preses des d'un entorn controlat, mentre que amb el data set d'imatges LISA els mètodes funcionen bé sempre que no hi hagi sobre il·luminació no desitjada.

Això ho podem comprovar gràcies al fet que amb el data set de test tenim un millor resultat que amb el data set de train, sent l'única diferència entre els dos les condicions lumíniques.

Per una futura millora del projecte i amb més temps, es podria millorar la part de la detecció de les mans amb un mètode adaptatiu a la llum en l'espai de color HSV per a que el mètode sigui més tolerant a canvis d'il·luminació. Com s'ha comentat durant aquest informe, es va intentar fer un mètode en l'espai de color HSV encara que sense èxit, considerant per tant que les imatges del data set LISA tenien massa particularitats per a poder abordar-les totes exitosament durant el projecte.

També, un altre punt interessant seria, una vegada es poguessin filtrar correctament totes les mans fins i tot amb canvis d'il·luminació, amb un Kalman Filter [17] o algun

altre mètode, es podria identificar cap a quina direcció giren les mans i, per tant, saber la direcció a la que està girant el cotxe en cada moment.

AGRAÏMENTS

Aquest projecte no podria haver-se realitzat sense el recolçament de la meva tutora, Aura Hernández, que ha estat atenta durant tot el projecte a més de donar-me suport durant el projecte.

També agraeixo a la meva família el suport que m'ha donat durant la realització del projecte, sense la seva ajuda tampoc hagués pogut tenir el temps necessari per realitzar-lo. Sense oblidar-me també dels ànims rebuts per part d'amics i companys.

REFERÈNCIES

- [1] "Los accidentes de tráfico se cobran la vida de 1.160 personas durante el año pasado", *DGT*, Gener 2017. [En línia]. Disponible a: <http://www.dgt.es/es/prensa/notas-deprensa/2017/20170103-Los-accidentes-de-trafico-se-cobran-la-vida-de-1160personas-durante-el-año-pasado.shtml> Consulta 18 de Febrer de 2017.
- [2] "Pursuit for Vehicle Safety, Driving Simulator", *Toyota*. [En línia]. Disponible a: http://www.toyota.com.cn/innovation/safety\technology/safety\measurements/driving_simulator.html Consulta 26 de Juny de 2017.
- [3] David Jegan Abishek D, Malvika SS, "Hands and foot activity analysis while driving", Juny 2016.
- [4] E. Ohn-Bar and M. M. Trivedi, Hand Gesture Recognition in Real-Time for Automotive Interfaces: A Multimodal Vision-based Approach and Evaluations, *IEEE Trans. on Intelligent Transportation Systems*, 2014. [En línia]. Disponible a: <http://cvrr.ucsd.edu/LISA/hand.html>
- [5] Minhas, Saad, et al. LLEE: A Photorealistic Virtual Environment for Assessing Driver-Vehicle Interactions in Self-driving Mode." *Computer Vision-ECCV 2016 Workshops*. Springer International Publishing, 2016.
- [6] Mathworks documentation, `imsubtract`. [En línia]. Disponible a: <https://es.mathworks.com/help/images/ref/imsubtract.html>
- [7] Mathworks documentation, `imcrop`. [En línia]. Disponible a: <https://es.mathworks.com/help/images/ref/imcrop.html>
- [8] Mathworks documentation, `im2double`. [En línia]. Disponible a: <https://es.mathworks.com/help/matlab/ref/im2double.html>

- [9] Mathworks documentation, mex. [En línia]. Disponible a: <https://es.mathworks.com/help/matlab/ref/mex.html>
- [10] Mathworks documentation, fspecial. [En línia]. Disponible a: <https://es.mathworks.com/help/images/ref/fspecial.html>
- [11] Mathworks documentation, imfilter. [En línia]. Disponible a: <https://es.mathworks.com/help/images/ref/imfilter.html>
- [12] Mathworks documentation, Bsxfun. [En línia]. Disponible a: <https://es.mathworks.com/help/matlab/ref/bsxfun.html>
- [13] pi19404, "Adaptive Skin Color Detection"*Code Project*, Agost 2014. [En línia]. Disponible a: <https://www.codeproject.com/Articles/808416/Adaptive-Skin-Color-Detection>
- [14] "Installation in Windows". *OpenCV*. [En línia]. Disponible a: [http://docs.opencv.org/2.4/doc/tutorials/introduction/windows_install/windows_install.html](http://docs.opencv.org/2.4/doc/tutorials/introduction/windows/_install/windows_install.html)
- [15] Mathworks documentation, RegionProps. [En línia]. Disponible a: <https://es.mathworks.com/help/images/ref/regionprops.html>
- [16] Mathworks documentation, bwlabel. [En línia]. Disponible a: <https://es.mathworks.com/help/images/ref/bwlabel.html>
- [17] Mathworks documentation, Kalman Filter. [En línia]. Disponible a: <https://es.mathworks.com/help/ident/ref/kalmanfilter.html>