



This is the **published version** of the bachelor thesis:

Cano Marcos, Oriol; Sánchez Pujadas, Javier Sánchez, tut. Ajudes semiautomàtiques per fer ground truth de segmentació de regions en imatges. 2025. (Enginyeria Informàtica)

This version is available at <https://ddd.uab.cat/record/308751>

under the terms of the  license

Ajudes semiautomàtiques per fer ground truth de segmentació de regions en imatges

Oriol Cano Marcos

Resum—Actualment, les eines de segmentació d'imatges són molt utilitzades per aplicacions que necessiten reconèixer o classificar objectes a partir d'imatges. Aquesta segmentació cal que sigui precisa i per tant és una tasca que requereix força temps, això fa que siguin necessàries eines que permetin una segmentació precisa d'imatges fent un procés semiautomàtic per poder fer aquestes tasques necessàries d'una forma més senzilla. Per aquest motiu, aquest treball es centrarà en facilitar la creació de ground truths on s'etiqueten regions d'imatges desenvolupant una eina de selecció que apliqui un algoritme de selecció per inundació, juntament amb altres filtres i algorismes de tractament d'imatges, fent que aquestes tasques no requereixin tant temps, sense perdre precisió en la selecció i fent seleccions més acurades.

Paraules clau— C++, Ground truth, tractament d'imatges, segmentació d'imatges, visió per computador, GTCreator, watershed, detecció de vores, Sobel, MLSEC, Harris, desviació estàndard

Abstract— Currently, image segmentation tools are widely used in applications that require the recognition or classification of objects from images. This segmentation needs to be precise, making it a time-consuming task. Therefore, there is a need for tools that facilitate accurate image segmentation through semi-automatic processes, simplifying these tasks. For this reason, this work will focus on facilitating the creation of ground truths, where image regions are labeled, by developing a selection tool that applies a flood-fill selection algorithm, along with other filters and image processing algorithms. This will make these tasks less time-consuming, without losing precision in the selection and achieving more accurate results.

Index Terms— C++, Ground truth, image processing, image segmentation, computer vision, GTCreator, watershed, edge detection, Sobel, MLSEC, Harris, standard deviation



1 INTRODUCCIÓ

Aquest treball és basa en la proposta del professor i tutor del treball Javier Sánchez Pujadas.

El desenvolupament d'algoritmes precisos per la segmentació de regions en imatges és una tasca molt important en camps com la visió per computació i el processament d'imatges. Aquests algoritmes són la base de moltes aplicacions com poden ser la detecció d'objectes a fotografies i la conducció autònoma. Un clar exemple sobre la importància d'això és Google Lens ^[1], una eina que permet identificar objectes i ubicacions del temps real a partir d'una imatge o utilitzant la càmera del mòbil. Perquè Google Lens sigui capaç de reconèixer objectes amb precisió, és necessari comptar amb un model d'aprenentatge automàtic entrenat amb una gran quantitat de dades etiquetades, és a dir, on cada regió de la imatge ha sigut assignat a una classe específica (persona, cotxe, edifici...). Aquestes dades d'etiquetatge es coneixen com a ground truth.

No obstant, la creació manual d'aquestes dades pot ser una tasca llarga i propensa a errors, especialment quan es treballa amb una gran quantitat d'imatges i si els objectes són petits i de formes irregulars.

En aquest context, les ajudes semiautomàtiques són una eina molt valuosa amb la que es pot agilitzar aquest procés i millorar la precisió.

Aquest treball consisteix en el desenvolupament d'eines que ajudin a la segmentació de regions d'imatges per fer el ground truth i si és possible acabar implementant aquestes eines a l'aplicació "GTCreator"^[2].

2 ESTAT DE L'ART

Pel que fa a aplicacions similars he centrat la meua cerca en dos tipus d'aplicacions diferents, les aplicacions d'edició d'imatge i les aplicacions per fer ground truth de segmentació de regions en imatges.

2.1 Aplicacions d'edició d'imatges

He buscat informació i he provat aplicacions d'edició d'imatges com "Photoshop" i "GIMP" ^[3] ja que aquest tipus d'aplicacions, tot i no ser exactament el mateix tipus

- E-mail de contacte: 1599727@uab.cat
- Menció realitzada: Computació
- Treball tutoritzat per: Javier Sánchez Pujadas, Departament de Computació
- Curs 2024/25

que l'aplicació GTCreator, també tenen un gran varietat d'eines de selecció de regions diferents amb les que he pogut experimentar i agafar idees per la eina de selecció que caldrà utilitzar durant el desenvolupament del projecte. Les eines de les aplicacions d'edició d'imatge que crec que podrien ser una aproximació al que finalment s'ha desenvolupat durant el projecte és el llac magnètic, que és una selecció semiautomàtica en la que mentre que l'usuari va marcant la zona desitjada, el llac magnètic va adaptant la selecció a les vores més properes a la selecció real de l'usuari.

2.2 Aplicacions de ground truth d'imatges

A part de les diferents aplicacions d'edició d'imatges també he buscat i provat aplicacions que serveixen per fer ground truth de segmentació d'imatges com "Supervisely"^[4] i "Labelbox"^[5] per veure quines eines de selecció ofereien.

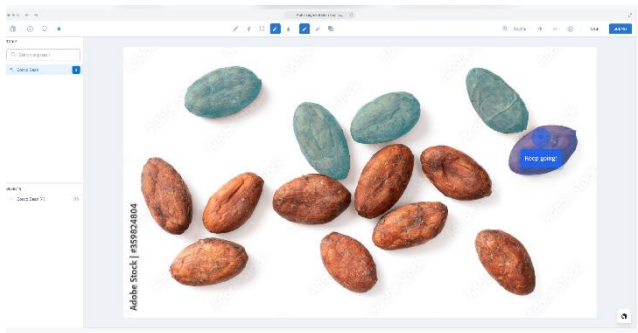


Figura 1. Eina de selecció semiautomàtica de Labelbox

La principal eina de selecció de regions que he trobat en aquest tipus d'aplicacions ha sigut una semblant al que seria la eina de la galleda, que fent click a un objecte, l'aplicació et seleccionava automàticament l'objecte aproximat que volies seleccionar. Tot i que aquesta eina pot ser molt útil i ràpida, no fa unes seleccions massa acurades si els objectes no es poden diferenciar clarament del fons, com és el cas de les imatges que es volen aconseguir segmentar al final d'aquest projecte i amb el que es faran totes les proves.

Per aquest motiu he decidit que la eina que serà finalment desenvolupada durant el projecte serà més semblant a la eina de selecció mencionada al subapartat anterior, el llac magnètic, que a les eines que ofereixen les alternatives a GTCreator.

A part d'aquestes alternatives, també he fet proves per familiaritzar-me amb l'aplicació GTCreator per veure les opcions que té i familiaritzar-me amb la interfície i les diferents eines de selecció de les que ja disposa l'aplicació, perquè com ja he mencionat a la introducció del treball, l'objectiu final d'aquest projecte serà la implementació de l'eina desenvolupada a aquesta aplicació per poder veure els resultats obtinguts i poder fer una comparació i una demostració real.

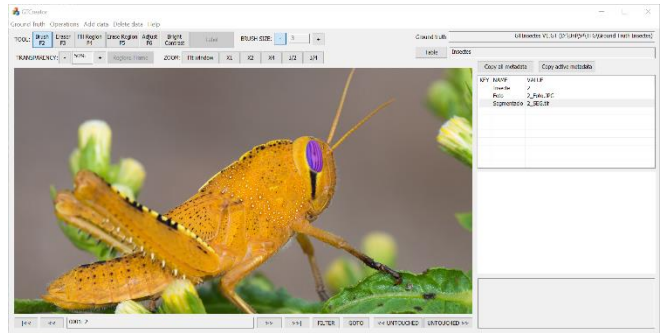


Figura 2. Aplicació GTCreator

3 OBJECTIUS

He dividit els objectius del projecte en dos grups diferenciats, els objectius principals que són els essencials per la finalització del projecte i els objectius extres que són millores respecte els objectius principals i que he anat desenvolupant en els moments que he tingut bloquejos durant el desenvolupament dels objectius principals i un cop finalitzats aquests.

Objectius principals:

- Objectiu 1: Dissenyar una eina de selecció que permeti que l'usuari marqui una zona de fora de la selecció i una zona de l'interior i es seleccioni automàticament l'interior seguint la vora desitjada.
- Objectiu 2: Implementar l'eina anterior a l'aplicació de segmentació d'imatges per fer ground truth GTCreator.

Objectius secundaris:

- Extra: Millorar la precisió de la eina de selecció.
 - Reduir el temps necessari per utilitzar la eina.
 - Millorar la selecció obtinguda.

4 METODOLOGIA

Per aconseguir els objectius proposats utilitzaré la metodologia Scrumban^[6]. Aquesta metodologia es una metodologia híbrida que combina elements de Scrum i Kanban. Aquesta metodologia va sorgir arran de la necessitat d'algunes empreses de fer la transició des de Kanban, àmpliament utilitzat en aquell moment, cap a Scrum quan aquesta metodologia va començar a guanyar popularitat. Algunes empreses van adoptar Scrum de forma completa, però altres van optar per Scrumban, una metodologia que recull característiques de totes dues, resultant més adequada per a certs projectes.

Scrumban és una metodologia que estructura el treball en cicles anomenats sprints, durant els quals s'han de completar les tasques programades, de la mateixa manera que la metodologia Scrum. A més, com a la metodologia

Kanban, es fa servir un tauler visual dividit en les columnes “To Do”, “In Progress” i “Done”, on es col·loquen les tasques del projecte, oferint una manera visual de seguir l'estat de cada tasca i el progrés general del projecte.

Per fer això, utilitzaré la web “Jira”^[7], una eina d'Atlasian que s'utilitza per la gestió de projectes que utilitzen una metodologia Agile.

4.1 Modificacions

Donat que aquesta metodologia està pensada per projectes on hi participa més d'una persona i no s'acaba d'adaptar bé a aquest projecte, la metodologia que utilitzaré és aquesta però amb alguna modificació:

1. Els sprints no seran tots de la mateixa duració, ja que al final de cada sprint caldria fer una entrega parcial del projecte per poder rebre feedback i no totes les parts d'aquest projecte tenen la mateixa duració ni poden ser entregades per separat. És necessari fer la integració a GTCreator de cada part per poder veure els resultats de la part feta. Per tant els Sprints acabaran cada cop que estigui planificada la finalització d'una part que pot ser entregada com a Demo.
2. Scrumban pot tenir rols, però donat que aquest projecte serà desenvolupat per 1 única persona no hi ha la necessitat d'utilitzar diferents rols.

5 PLANIFICACIÓ

Pel que fa a la planificació del projecte, si que hi ha hagut diferents modificacions, ja que els objectius principals del projecte van ser modificats al poc temps de començar amb el projecte per correccions del tutor. A part d'aquestes modificacions també hi ha hagut algunes variacions respecte a les dates estimades de finalització i hores de treball d'algunes de les parts.

Part del projecte	Duració	Data estimada de finalització	Hores de treball
Inici del projecte	4 setmanes	06/10/24	30h
Preparació de l'entorn de proves	4 setmanes	03/11/24	30h
Objectiu 1	4 setmanes	01/12/24	70h
Objectiu 2	5 setmanes	05/01/25	100h
Extra	5 setmanes*	09/02/25	100h

* Com l'objectiu extra és millorar tot el possible l'eina de selecció, la duració, data de finalització i hores de treball son orientatives ja que ocuparà tot el temps disponible un cop acabats els objectius principals del projecte.

5.1 Planificació detallada dels objectius

Per aconseguir els objectius proposats, he trobat diferents algorismes utilitzats en la segmentació d'imatges que hem poden ser d'utilitat pel desenvolupament de les eines.

Pel primer objectiu he utilitzat un algorisme “watershed”^{[8][9]}. Aquest algorisme es basa en tractar una imatge com si fos un mapa topogràfic on l'algorisme fa una inundació del paisatge tenint en compte uns marcadors que indiquen els pics i les valls de la imatge. D'aquesta forma es pot fer que l'usuari marqui una regió exterior de la zona desitjada, que representaria el pic, i una regió de dins de la selecció, que seria la vall, i amb això es pot aconseguir que l'algorisme marqui la zona desitjada.

Per poder aplicar l'algorisme de “watershed” caldria utilitzar un algorisme que remarqui els contorns de la imatge, com Sobel^{[13][14]}.

Finalment per acabar amb els objectius principals, caldria implementar aquesta eina a l'aplicació GTCreator per poder utilitzar l'eina desenvolupada amb aquesta aplicació i poder-la utilitzar amb qualsevol imatge.

5.2 Seguiment de la planificació

Pel que fa al compliment de les dates proposades inicialment, s'ha pogut seguir satisfactòriament i sense endarreriments durant l'inici del projecte, la preparació de l'entorn de proves i el primer objectiu.

Cap al final del segon objectiu van sorgir imprevistos amb el desenvolupament d'aquest objectiu que van causar l'endarreriment de la data de finalització del segon objectiu gairebé una setmana, del 05/01 al 10/01.

Tot i aquests imprevistos durant el desenvolupament del segon objectiu, l'endarreriment total del projecte s'ha pogut mitigar gràcies a que mentre aquesta part del desenvolupament estava bloquejada, s'ha anat avançant i desenvolupant parts que correspondrien a l'objectiu extra.

Amb aquestes modificacions, les dates reals d'inici i finalització de cada part del projecte han sigut les següents:

Part del projecte	Data d'inici	Data de finalització
Inici del projecte	09/09/24	06/10/24
Preparació de l'entorn de proves	07/10/24	03/11/24
Objectiu 1	04/11/24	01/12/24
Objectiu 2	02/12/24	10/01/25
Extra	15/12/24	09/02/25

6 DESENVOLUPAMENT

Tot el desenvolupament d'aquest projecte s'ha fet en el llenguatge C++ i la única llibreria de tractament d'imatges utilitzada ha sigut la llibreria desenvolupada pel tutor del treball Javier Sánchez Pujadas anomenada GIPL ja que era la llibreria que utilitza l'aplicació GTCreator i utilitzant una llibreria que ja està inclosa al projecte es podia evitar bastants problemes.

6.1 Aplicació de proves

Per desenvolupar el primer objectiu, la creació de la eina de selecció, he creat una petita aplicació que llegia una imatge qualsevol, una màscara de l'exterior de la zona que es vol seleccionar i una màscara de l'interior.

Amb això vaig començar a aplicar diferents filtres per remarcar els contorns i reduir el soroll de les imatges, com Sobel, un filtre de mediana, un filtre gaussià i un filtre laplacià, i fer diverses proves per veure com quedaria una primera versió d'aquesta eina de selecció i veure diferents paràmetres i opcions de filtres que podia aplicar-hi per obtenir millors resultats.

Un cop fetes les proves i comparar els resultats obtinguts utilitzant diferents filtres i combinacions amb diferents paràmetres, he decidit utilitzar per la eina de selecció la opció que aplica un filtre Sobel per obtenir les vores de la imatge, un filtre de mediana per reduir el soroll i un filtre gaussià per ressaltar més les vores de la imatge.



Figura 3: Intersecció insecte-fons

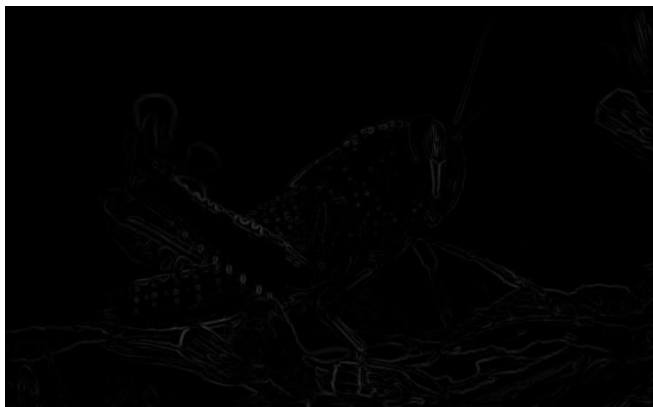


Figura 4: Vores de la imatge



Figura 5: Imatge amb la vora de l'insecte marcada amb l'eina

6.2 Implementació a GTCreator

Una cop ja estava feta una versió inicial de l'eina de selecció semiautomàtica vaig passar a l'aplicació GTCreator.

Durant els primers dies de començar amb el desenvolupament corresponent al segon objectiu, vaig familiaritzar-me amb l'ús d'aquesta aplicació així com amb el desenvolupament de interfícies gràfiques en C++ ja que implementar l'eina de selecció dins d'aquesta aplicació també suposava la edició actual de l'aplicació per poder incloure aquesta nova opció.

Pel que fa a la part de la interfície gràfica de la implementació, he creat un nou botó al diàleg principal de l'aplicació que quan l'usuari te seleccionada una màscara d'una imatge pot fer-hi click. Això obre un nou diàleg en el que es mostrarà la mateixa imatge que s'estava editant al diàleg principal però sense cap màscara.

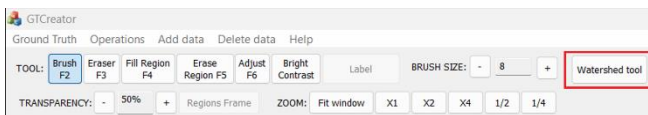


Figura 6: Botó per obrir el nou diàleg

Aquest nou diàleg permet seguir fent seleccions utilitzant el pinzell normal i l'eina d'esborrar. La finestra conté inicialment un botó per activar el pinzell de selecció normal, un botó per seleccionar el borrarador i un selector de mida pel pinzell o borrarador. A part d'aquests botons, que ja estaven a la finestra principal, hi ha un nou botó, "Apply watershed", que serveix per, un cop tens les zones necessàries seleccionades, aplicar l'algoritme desenvolupat i deixar seleccionada la zona d'interés. Un cop l'usuari estigui satisfet amb la selecció, pot donar click al botó d'acceptar per aplicar la nova selecció a la selecció que ja tenia feta a la finestra principal de l'aplicació, o en cas contrari, fer click a cancel·lar per tirar enrere les modificacions.

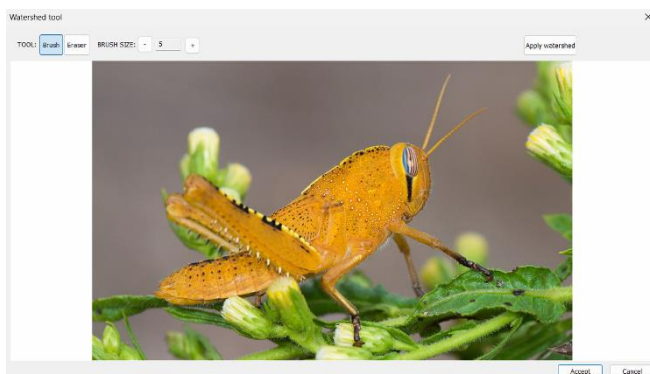


Figura 7: Diàleg de la nova eina

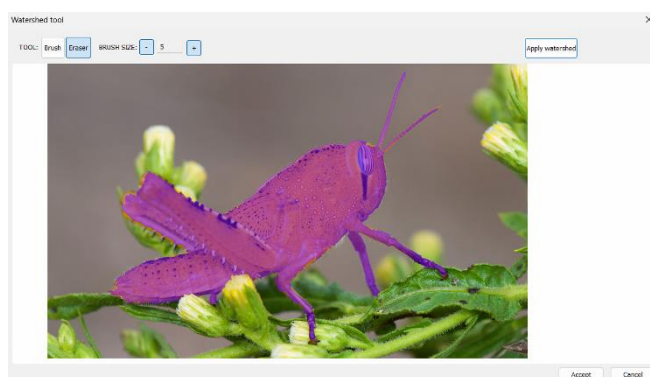


Figura 8: Selecció feta amb la eina integrada

6.3 Altres millores

Durant el desenvolupament del segon objectiu, la part que correspondria al subapartat anterior, vaig tenir alguns bloquejos causats per la falta de documentació de l'aplicació GTCreator i, per tant, la necessitat de fer diverses reunions amb el tutor i desenvolupador tant de GTCreator com de la llibreria utilitzada, GIPL.

Aquests períodes en els que l'objectiu principal quedava aturat, he aprofitat per anar fent millores a la eina desenvolupada anteriorment.

6.3.1 Millora de la selecció de màscares inicials

La primera millora, i que té un impacte més important en el temps necessari, ha estat la eliminació de la necessitat de crear dues màscares, una de la regió exterior de la zona desitjada i una de la interior. Aquesta modificació ha canviat el comportament de l'eina ja que en comptes d'utilitzar aquestes dues màscares per aplicar l'algoritme watershed s'utilitza una única màscara que indica on està la intersecció entre les regions interiors i exteriors de la zona objectiu.

Això ho he pogut fer passant la imatge binària proporcionada per l'usuari, que seria la màscara de la intersecció, a una imatge on tots els valors que estan dins de la zona objectiu tinguin un valor 1, els que estan fora tinguin valor 2 i els que estan a la intersecció tinguin valor 0. D'aquesta

manera, tenim tota la informació necessària per aplicar l'algoritme de watershed dins de la mateixa màscara.

Aquesta millora implica una reducció significativa al temps necessari per utilitzar la eina semiautomàtica desenvolupada, i a la vegada, també elimina possibles errors humans que podien tenir lloc si l'usuari creava dues màscares que es superposaven en alguna zona, ja que en aquests casos, no hi ha cap zona d'intersecció entre l'interior i l'exterior de la zona que es vol seleccionar i per tant no es pot seleccionar un perímetre que divideixi aquestes dues regions.

6.3.2 Parametrització de l'aplicació de filtres

Els filtres que s'apliquen a les imatges abans d'utilitzar l'algoritme de watershed utilitzen uns paràmetres que inicialment, després de fer proves amb diverses imatges i valors, se'ls havia assignat un valor que era el que mostrava uns resultats millors en totes les proves. Aquest valor era fixe i no modificable per l'usuari.

Aquesta millora permet a l'usuari modificar aquests valors utilitzats per poder seleccionar diferents paràmetres segons les necessitats de cada imatge, tot i que, si l'usuari no té la necessitat de modificar aquests valors, els pot deixar amb els paràmetres per defecte, que són els valors que estaven anteriorment fixats.

Amb aquesta millora ha implicat una modificació de la interfície gràfica per tal d'afegir unes caixes de text on l'usuari pot modificar els paràmetres que s'utilitzen a l'hora d'aplicar els filtres de mediana i filtre gaussià.



Figura 9: Nou diàleg modificat

6.3.3 Noves opcions de filtres

Els filtres utilitzats inicialment són un filtre Sobel per obtenir les vores de la imatge, un filtre de mediana per reduir el soroll i un filtre gaussià per ressaltar més les vores de la imatge. Però un cop acabat el projecte vaig fer proves amb més filtres diferents, com el gradient morfològic, la desviació estàndard, el laplacià i el detector de cantonades de Harris, i vaig trobar una combinació que, en alguns casos, donava uns millors resultats que la combinació que estava utilitzant.

Aquestes noves opcions de filtratge substitueixen l'ús del filtre Sobel per uns altres.

Un dels filtres és el filtre desenvolupat per Antonio Manuel López Peña a la seva tesi doctoral "Multilocal methods for ridge and valley delineation in image analysis"^[19] i implementat a la llibreria GIPL per Javier Sánchez Pujadas. Aquest és un filtre anomenat MLSEC utilitza el tensor estructural de la imatge per així detectar canvis en la imatge en direccions oposades, això indicaria la presència d'una vall o cresta. Aquesta nou mètode, a diferència

d'altres detectors de vores com Sobel, segueix proporcionant informació en situacions en les que el gradient es zero, cosa especialment útil en imatges on hi ha variacions subtils o sense una vora clarament marcada.

A part d'aquesta opció, també s'han implementat les opcions d'utilitzar els filtres de la desviació estàndard i la opció del filtre de detecció de cantonades Harris^[20].

Aquesta millora, d'igual forma que l'anterior, també fa necessària una modificació de la interfície gràfica per permetre a l'usuari seleccionar quin filtre vol aplicar entre "Harris", "Std deviation", "Sobel" i "MLSEC". A part de la selecció del filtre a utilitzar, l'usuari també pot modificar dos paràmetres que variaran, en alguns casos més que en altres, la selecció feta per l'eina donant així una gran flexibilitat per poder utilitzar aquesta eina en qualsevol tipus d'imatge sense problemes.

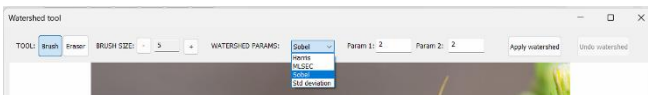


Figura 10: Opcions de filtres

6.3.4 Opció per desfer la selecció

Com a última millora del treball s'ha afegit un botó al diàleg que, un cop l'usuari ha utilitzat l'eina per fer una selecció, permet a l'usuari tornar enrere tornant així a la selecció que havia fet l'usuari inicialment per permetre-li acabar d'ajustar els paràmetres i utilitzar els diferents filtres per acabar d'ajustar la selecció. A part d'això, quan s'utilitza aquest nou botó també queda marcada la zona d'intersecció que ha utilitzat l'algoritme per fer la selecció fent més fàcil a l'usuari la correcció d'una selecció mal feta, ja que en cas de que l'algoritme seleccioni una regió equivocada, l'usuari pot veure quin contorn ha seguit aquesta eina i pot corregir-la modificant una zona de la màscara marcada inicialment per evitar la zona errònia.

7 RESULTATS

Finalment s'han pogut completar satisfactòriament el primer objectiu que era el desenvolupament d'una eina de selecció semiautomàtica que a partir d'unes seleccions de l'usuari fos capaç de seleccionar la zona desitjada automàticament. Això finalment s'ha aconseguit mitjançant un algoritme de selecció per inundació sobre una imatge filtrada amb diferents algoritmes.

A part d'això també s'ha pogut assolir el segon objectiu del projecte, que era la implementació d'aquesta eina de selecció desenvolupada a l'aplicació GTCreator. Aquesta part s'ha desenvolupat creant un nou botó a la finestra principal de l'aplicació que obre un nou diàleg dissenyat exclusivament per l'ús d'aquesta nova eina. Des de aquest nou diàleg, l'usuari pot marcar la zona d'intersecció de la regió que es vol seleccionar, que és la única màscara necessària per utilitzar l'eina desenvolupada, i pot triar entre diferents opcions i paràmetres per utilitzar i acabar d'ajustar

l'eina per tal de millorar tot el possible la selecció feta per aquesta, ja que segons cada imatge poden ser millors uns altres valors que no són els seleccionats per defecte.

Pel que fa amb el tercer objectiu, l'objectiu que estava proposat com a extra d'aplicar millores sobre la eina desenvolupada tant per reduir el temps necessari per utilitzar-la com per augmentar la precisió d'aquesta també es podria donar com a assolit satisfactòriament ja que s'han fet tres modificacions respecte al que serien els dos primers objectius que tenen implicacions en la reducció del temps i la millora de precisió i donen més llibertat de personalització per l'usuari.

7.1 Mostra i comparació de resultats

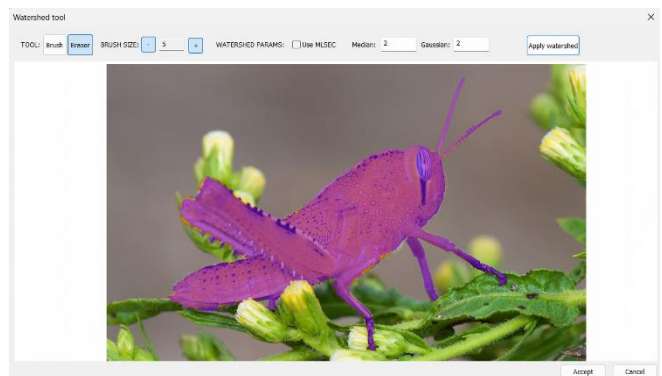


Figura 11: Selecció imatge de proves 1

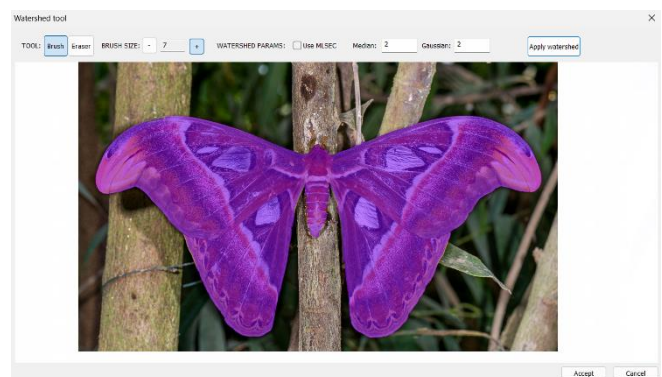


Figura 12: Selecció imatge de proves 2

Com es pot veure a les dues figures anteriors (figura 11 i figura 12), la selecció que es pot veure està feta utilitzant únicament la eina desenvolupada. Tot i que la selecció feta en els dos casos és força bona, es pot veure com influeix la complexitat de la imatge. Tot i que les dues imatges tenen zones complexes, la figura 11 amb moltes vores i petits detalls i la figura 12 amb poca diferència entre els colors del fons i de l'insecte, a la primera imatge es pot veure com la selecció no és tant exacta principalment a la regió superior de la pota de l'insecte, errors causats principalment per l'alta presència de vores en aquella regió. Cal destacar també que els dos casos anteriors han sigut seleccionats amb els paràmetres per defecte i que les màscares

utilitzades en les dues situacions anteriors han sigut creades únicament utilitzant el pinzell i el borrador de l'aplicació i que han sigut creades en poc temps (10 minuts en el cas de la figura 11 i 3 minuts en el cas de la figura 12).

Els resultats d'aquestes seleccions es poden millorar, principalment a la imatge de proves 1 (figura 11), amb una selecció inicial més exacte en la que la vora més marcada de dins de la selecció de l'usuari sigui la vora que delimita l'interior de la zona desitjada i l'exterior.

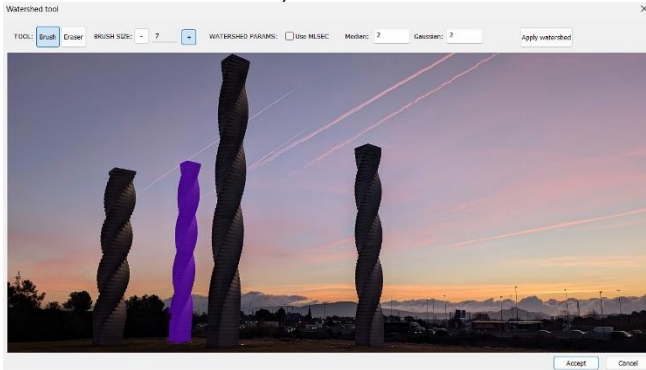


Figura 13: Selecció feta amb filtre Sobel

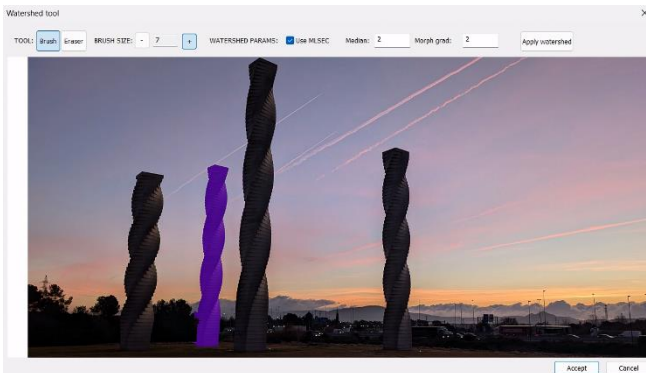


Figura 14: Selecció feta amb filtre MLSEC

En les dues imatges anteriors (figura 13 i figura 14) podem observar com les dues tenen el mateix problema, la part superior, en la que es poden diferenciar sense problemes el fons i l'objecte que es vol seleccionar, la eina no té cap problema en fer la selecció, però en la zona inferior de la imatge, on tant el fons com l'objecte són molt foscos i no es pot diferenciar tant bé, la eina té dificultats al seleccionar la zona correctament, clar que també cal tenir en compte que la selecció inicial de la intersecció entre l'objecte i el fons ha estat una selecció molt burda i feta molt ràpidament, en menys d'un minut, a part de deixar els paràmetres per defecte, que ajudarien a fer una millor segmentació.

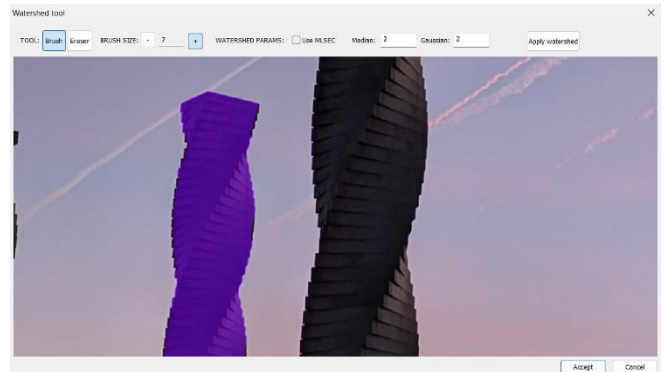


Figura 15: Ampliació de la zona superior de la figura 13

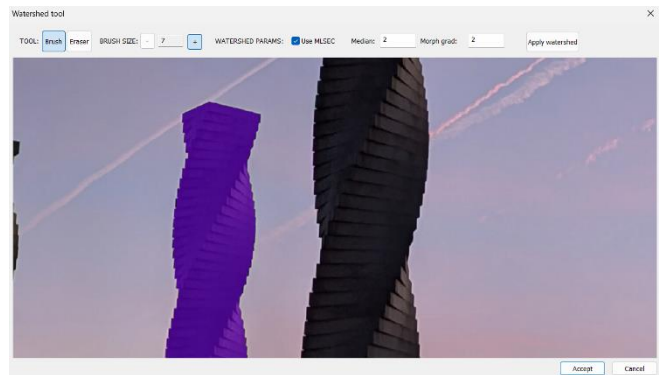


Figura 16: Ampliació de la zona superior de la figura 14

En aquestes dues últimes figures, en la zona superior ampliada de les seleccions fetes amb el filtre Sobel (figura 15) i amb el filtre MLSEC (figura 16) podem observar com tot i semblar iguals hi ha una millor selecció en la imatge que utilitza el filtre MLSEC ja que en la que utilitza Sobel es fa una selecció de l'objecte una mica menys precisa que arriba a seleccionar regions de fora de l'objecte.

Tot i això, es pot veure com en imatges que no són tant complexes com en les primeres figures mostrades en aquest apartat (figures 11 i 12), es fa una selecció pràcticament perfecta de l'objecte desitjat, utilitzant màscares creades en poc més d'un minut i els paràmetres per defecte, a excepció de la diferència entre l'ús del filtre Sobel i MLSEC.

7.2 Selecció amb i sense ús de la eina

Per poder demostrar la utilitat real d'aquesta eina i poder donar dades reals sobre la millora que suposa l'ús de la eina desenvolupada respecte una selecció feta sense eines semiautomàtiques he fet una prova de selecció amb i sense l'ús d'aquesta eina.

Per fer aquesta prova he escollit la imatge de la figura 12 ja que no és gaire complexa a l'hora de fer la selecció.



Figura 17: Selecció feta amb watershed



Figura 18: Selecció sense watershed

La figura 17 ha estat seleccionada utilitzant l'eina desenvolupada així com el pinzell i el borrador normals de l'aplicació GTCreator per acabar d'ajustar la selecció en aquelles zones on l'eina desenvolupada ha sigut més imprecisa, mentre que per fer la selecció de la figura 18 s'han utilitzat el pinzell i el borrador com a l'anterior figura i a part una eina que omple el contingut envoltat per una selecció feta, d'aquesta forma s'ha pogut fer la selecció marcant manualment la zona exterior de l'insecte i seleccionant l'interior amb aquesta última eina.

Com es pot veure a les anteriors figures, la selecció es pràcticament idèntica a simple vista tot i que si fem una ampliació a les zones d'intersecció de la selecció amb el fons podríem veure com la selecció feta amb l'ajuda de l'eina desenvolupada segueix molt millor la vora de la imatge que separa el fons de l'insecte seleccionat. Tot i això, la selecció que utilitza l'eina semiautomàtica s'ha fet en 3 minuts i 50 segons, mentre que l'altre selecció s'ha fet en 7 minuts i 26 segons, el que significa una reducció de temps de 216 segons, o el que és el mateix, una millora del 48.43% en el temps, que podria ser una diferència major en altres imatges més complexes.

8 CONCLUSIONS

Aquest treball ha permès assolir amb èxit tant els objectius principals com els complementaris proposats, consolidant així els resultats esperats. En primer lloc, s'ha dissenyat i desenvolupat amb èxit una eina de selecció semiautomàtica per a la segmentació de regions en imatges. Aquesta eina permet a l'usuari seleccionar zones d'interès amb major precisió i agilitat mitjançant l'aplicació d'un algorisme de selecció per inundació combinat amb filtres de processament d'imatges. L'algorisme watershed, juntament amb filtres com Sobel i MLSEC, ha demostrat ser eficaç en la detecció de vores i la definició de límits, millorant la qualitat de les segmentacions.

El segon objectiu, que consistia en la integració d'aquesta eina a l'aplicació GTCreator, també s'ha assolit satisfactòriament. La implementació ha inclòs la creació d'un nou botó i diàleg dins de la interfície gràfica de GTCreator, facilitant l'accés i l'operativitat de la nova eina. Aquesta integració permet als usuaris combinar la nova eina de selecció amb les funcionalitats ja existents de GTCreator, ampliant les possibilitats de segmentació i millorant els resultats finals.

Pel que fa als objectius complementaris, s'han introduït millores significatives que han optimitzat el rendiment de l'eina. La reducció del temps necessari per crear seleccions, la millora de la precisió en la segmentació, la possibilitat de personalitzar els paràmetres dels filtres han contribuït a fer que l'eina sigui més eficient i adaptable a diferents necessitats, i la possibilitat de tornar a la selecció inicial després d'utilitzar l'eina per poder fer correccions i modificar els paràmetres utilitzats per millorar els resultats. Especialment destacable és la substitució de la necessitat d'utilitzar dues màscares per una de sola, agilitzant el flux de treball i minimitzant els errors humans.

Finalment cal dir que la eina desenvolupada i millorada durant tot aquest treball és una eina flexible en diferents tipus d'imatge, gràcies a tots els paràmetres que l'usuari pot modificar manualment segons el que sigui més convenient, i amb molta utilitat en tasques de creació de ground truths on s'etiqueten imatges, ja que com s'explica a la introducció d'aquest treball, aquesta és una tasca molt necessària actualment i és una acció que requereix molt temps per fer-la correctament, i per tant, qualsevol eina que faci aquestes funcions més àgils són de molta utilitat. A més a més aquesta eina no només agilitza aquest treball sinó que també millora les seleccions fetes manualment gràcies a que ressegueix els contorns de la imatge d'una forma més precisa que les seleccions que es poden fer manualment.

També cal dir que l'eina desenvolupada no es presenta com una solució independent, sinó com una eina complementària dins de GTCreator. La seva combinació amb altres eines ja existents potència el procés de creació de ground truth, permetent als usuaris obtenir segmentacions més precises i eficients. Aquesta complementarietat assegura que l'eina pugui adaptar-se a diferents tipus d'imatges i escenaris.

En resum, aquest treball ha aconseguit desenvolupar i integrar amb èxit una eina semiautomàtica que millora notablement el procés de segmentació d'imatges. Les proves realitzades han confirmat que aquesta eina redueix el temps de treball significativament i incrementa la qualitat de les seleccions fetes, contribuint significativament a l'eficiència del procés de creació de ground truth. Així, es pot concloure que els objectius del projecte s'han assolit de manera satisfactòria, oferint una eina robusta, flexible i de molta utilitat que enriqueix les funcionalitats de GTCreator i facilita el treball dels usuaris en tasques creació de ground truths on s'etiqueten regions d'imatges.

BIBLIOGRAFIA

- [1] "Google Lens: Real-time answers to questions about the world around you", Google Blog. [En línia]. [Accedit: 4-oct-2024] Disponible a: <https://blog.google/products/google-lens/google-lens-real-time-answers-questions-about-world-around-you/>
- [2] J. Bernal, A. Histace, M. Masana, Q. Angermann, C. Sanchez-Montes, C. RodriguezdeMiguel, M. Hammami, A. Garcia-Rodriguez, H. Cordova, O. Romain, G. Fernandez-Esparrach, X. Dray, i F. J. Sanchez, "GTCreator: a flexible annotation tool for image-based datasets," *Int. J. CARS*, vol. 14, pp. 191–201, 2019. [En línia]. [Accedit: 3-oct-2024] Disponible a: <https://doi.org/10.1007/s11548-018-1864-x>.
- [3] "Selecting with the Lasso tools", Adobe Photoshop Help. [En línia]. [Accedit: 24-set-2024] Disponible a: <https://helpx.adobe.com/photoshop/using/selecting-lasso-tools.html>
- [4] "Supervisely: Platform for AI-powered annotation and training", Supervisely. [En línia]. [Accedit: 3-oct-2024] Disponible a: <https://supervisely.com/>
- [5] "Labelbox: The leading training data platform for AI", Labelbox. [En línia]. [Accedit: 3-oct-2024] Disponible a: <https://labelbox.com/>
- [6] "Scrumban: A guide to the hybrid methodology", Atlassian. [En línia]. [Accedit: 28-set-2024] Disponible a: <https://www.atlassian.com/agile/project-management/scrumban>
- [7] "Jira Software: Herramienta de gestión de proyectos", Atlassian. [En línia]. [Accedit: 1-oct-2024] Disponible a: <https://www.atlassian.com/es/software/jira>
- [8] "Watershed (image processing)", Wikipedia. [En línia]. [Accedit: 26-set-2024] Disponible a: [https://en.wikipedia.org/wiki/Watershed_\(image_processing\)](https://en.wikipedia.org/wiki/Watershed_(image_processing))
- [9] "Watershed Algorithm", OpenCV Documentation. [En línia]. [Accedit: 26-set-2024] Disponible a: https://docs.opencv.org/4.x/d3/db4/tutorial_py_watershed.html
- [10] "Canny edge detector", Wikipedia. [En línia]. [Accedit: 24-set-2024] Disponible a: https://en.wikipedia.org/wiki/Canny_edge_detector
- [11] "Canny Edge Detection", OpenCV Documentation. [En línia]. [Accedit: 24-set-2024] Disponible a: https://docs.opencv.org/3.4/da/d5c/tutorial_canny_detector.html
- [12] "Dijkstra's Shortest Path Algorithm | Greedy Algo-7", GeeksforGeeks. [En línia]. [Accedit: 30-set-2024] Disponible a: <https://www.geeksforgeeks.org/dijkstras-shortest-path-algorithm-greedy-algo-7/>
- [13] "The Sobel Operator", HIPR2, School of Informatics, University of Edinburgh. [En línia]. [Accedit: 6-nov-2024] Disponible a: <https://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm>
- [14] "How the Sobel Operator Works", Automatic Addison. [En línia]. [Accedit: 6-nov-2024] Disponible a: <https://automaticaddison.com/how-the-sobel-operator-works/>
- [15] "C++ GUI: How to Create a Graphical User Interface in C++", Simplilearn. [En línia]. [Accedit: 30-oct-2024] Disponible a: <https://www.simplilearn.com/tutorials/cpp-tutorial/cpp-gui>
- [16] "Resource Files in Visual Studio (Windows)", Microsoft Learn. [En línia]. [Accedit: 30-oct-2024] Disponible a: <https://learn.microsoft.com/en-us/cpp/windows/resource-files-visual-studio?view=msvc-170>
- [17] "CDialog Class", Microsoft Learn. [En línia]. [Accedit: 4-des-2024] Disponible a: <https://learn.microsoft.com/en-us/cpp/mfc/reference/cdialog-class?view=msvc-170>
- [18] "VC++ / C++ MFC tutorial 1: Creating a Dialog box for user input", YouTube. [En línia]. [Accedit: 4-des-2024] Disponible a: <https://www.youtube.com/watch?v=Vp81ypIAyVs>
- [19] A. M. López Peña, "Multilocal methods for ridge and valley delineation in image analysis," *Tesi doctoral*, Dept. Computació, Universitat Autònoma de Barcelona, Bellaterra, Espanya, 28 de gener de 2000.
- [20] "Lecture 06: Edge Detection", Penn State CSE. [En línia]. [Accedit: 27-gen-2025] Disponible a: <https://www.cse.psu.edu/~rtc12/CSE486/lecture06.pdf>