

---

# Cirurgia assistida mitjançant visió per computador

*Treball de final de grau en Enginyeria Informàtica,  
especialitat en Computació.*

---

Autor: Sergi Coté i Rojas

Director: Joan Climent i Vilaró

Barcelona, a 20 de Juny de 2017

Facultat d'Informàtica de Barcelona (FIB)

Universitat Politècnica de Catalunya (UPC) - BarcelonaTech



# Índex general

## Capítol 0: Resum del projecte

0.1 Resum en català	8
0.2 Resum en castellà	9
0.3 Resum en anglès	10

## Capítol 1: Introducció

1.1 Context	11
1.2 Estat de l'art	12

## Capítol 2: Plantejament

2.1 Formulació del problema	15
2.2 Objectius del projecte	16

## Capítol 3: Gestió del projecte

3.1 Metodologia i rigor	18
3.1.1 Metodologia	18
3.1.2 Eines de seguiment	18

3.1.3 Validació	19
3.2 Planificació temporal	19
3.2.1 Fases del projecte	19
3.2.2 Especificació de tasques	19
3.2.3 Recursos	21
3.2.4 Gestió del temps	22
3.2.5 Planificació temporal final	22
3.2.6 Diagrama de Gantt	23
3.3 Pressupost	24
3.3.1 Pressupost de recursos humans	24
3.3.2 Pressupost de recursos de software	25
3.3.3 Pressupost de recursos de hardware	26
3.3.4 Costos indirectes	26
3.3.5 Pressupost total	27
3.3.6 Control de gestió	27

## **Capítol 4: Informe de sostenibilitat**

4.1 Dimensió econòmica	28
4.2 Dimensió social	28
4.3 Dimensió ambiental	29
4.4 Taula de sostenibilitat	29

## **Capítol 5: Plantejament de la solució**

5.1 Plantejament de la solució	30
--------------------------------	----

## **Capítol 6: Implementació**

6.1 El·laboració d'un data-set propi	32
6.2 Extracció de característiques	41
6.3 Construcció del classificador	41
6.4 Preprocessament i filtrat d'imatge	43
6.5 Obtenció del resultat	50
6.6 Marcació del resultat en la imatge i generació de vídeo	52
6.7 Anàlisi de resultats	56

## **Capítol 7: Conclusions**

7.1 Conclusions generals	60
7.2 Futures implementacions	61

## **Capítol 8: Referències**

8.1 Referències	62
-----------------	----

# Índex de figures

1. Diagrama de Gantt definitiu I	23
2. Diagrama de Gantt definitiu II	23
3. Imatge utilitzada en la construcció del data-set	33
4. Imatge utilitzada en la construcció del data-set	33
5. Imatge utilitzada en la construcció del data-set	34
6. Imatge binària de l'eina	35
7. Imatge binària de l'eina	36
8. Imatge binària de la classe negativa	36
9. Imatge binària de la classe negativa	37
10. Imatge binària de la classe positiva	38
11. Imatge binària de la classe positiva	39
12. Imatge binària de la classe negativa	39
13. Imatge binària de la classe negativa	40
14. Model SVM en 2 dimensions	42
15. Resultat d'aplicar Closing sobre la imatge	44
16. Resultat d'aplicar Closing sobre la imatge	44
17. Càlcul de gradients sobre la imatge	45
18. Càlcul de gradients sobre la imatge	46
19. Representació del color en HSV	47
20. Component de saturació de la imatge	48
21. Component de saturació de la imatge	48
22. Imatge preprocessada i filtrada per un llindar	49
23. Imatge preprocessada i filtrada per un llindar	50
24. Imatge resultat ponderada amb les puntuacions retornades pel model	51
25. Imatge resultat ponderada amb les puntuacions retornades pel model	52
26. Aplicació de Closing sobre la imatge resultat	53
27. Reducció del resultat a una línia	54
28. Aplicació de dilació sobre la línia obtinguda	54

29. Eliminació de píxels interiors	55
30. Imatge definitiva	55
31. Detecció correcte obtinguda per l'última versió de l'algoritme	57
32. Detecció correcte obtinguda per l'última versió de l'algoritme	58
33. Detecció correcte obtinguda per l'última versió de l'algoritme	58
34. Detecció correcte obtinguda per l'última versió de l'algoritme	59
35. Detecció correcte i un fals positiu obtinguts per l'última versió de l'algoritme	59

# Índex de taules

1. Resum de les característiques dels diferents data-sets utilitzats en diversos estudis per a la detecció d'instrumental mèdic	13
2. Resum dels conjunts de característiques utilitzades per diversos estudis per a la detecció d'instrumental mèdic	14
3. Gestió del temps	22
4. Pressupost de recursos humans	24
5. Pressupost de recursos de software	25
6. Pressupost de recursos de hardware	26
7. Costos indirectes	26
8. Pressupost total	27
9. Taula sostenibilitat	29
10. Presentació de resultats obtinguts	56



# Capítol 0

## Resum del projecte

### 0.1 RESUM EN CATALÀ

En aquest projecte s'ha desenvolupat una solució informàtica que, a partir de l'enregistrament en vídeo d'una cirurgia mínimament invasiva, és capaç de detectar i marcar sobre aquest vídeo les eines utilitzades per el cirurgià.

Concretament, el projecte s'ha desenvolupat per reconèixer l'instrumental mèdic utilitzat en una laparoscòpia per a la reparació d'hèrnies, però la solució desenvolupada és fàcilment exportable a d'altres tipus de cirurgies.

Per portar-lo a terme, s'ha discriminat l'eina basant-se en característiques de forma i s'han utilitzat tècniques de segmentació sense un gran cost computacional, habilitant així que la detecció pugui ser computada en temps real.

Per últim, cal remarcar que la solució desenvolupada assoleix la detecció de l'instrumental mèdic sense necessitat de cap modificació, ni de la sala d'operacions ni de l'eina a reconèixer.

## **0.2 RESUM EN CASTELLÀ**

En este proyecto se ha desarrollado una solución informática que, partiendo del enregistramiento en vídeo de una cirugía mínimamente invasiva, es capaz de detectar y marcar sobre este vídeo las herramientas usadas por el cirujano.

Concretamente, el proyecto se ha desarrollado para reconocer el instrumental médico utilizado en una laparoscopia para la reparación de hernias, pero la solución es fácilmente exportable a otros tipos de cirugía.

Para realizarlo, se ha discriminado la herramienta basándose en características de forma y se han utilizado técnicas de segmentación sin un gran coste computacional, habilitando así que la detección pueda ser computada en tiempo real.

Por último, es necesario remarcar que la solución desarrollada detecta el instrumental médico sin necesidad de ninguna modificación, ni de la sala de operaciones ni de la herramienta a reconocer.

### **0.3 RESUM EN ANGLÈS**

In this project an informatic solution has been developed which, given a recording video of a minimal invasive surgery, it is able to detect and point on this video the surgical tools that has been used.

Precisely, the project has been developed to recognize surgical tools used in a laparoscopic hernia repair, but the proposed solution is easy to export for other types of surgeries.

To build it, surgical tool's discrimination has been based on shape features and segmentation techniques with no big computational cost has been used, making possible to perform real-time detection.

Finally, it is necessary to point that the developed solution detects surgical tools with no needed modifications neither for medical box or for surgical tools to detect.

# Capítol 1

## Introducció

### 1.1 CONTEXT

Els avenços tecnològics dels darrers temps han permès a la medicina millorar molts dels seus procediments. El disseny d'instrumental mèdic més especialitzat, la reducció del tamany d'aquests instruments, l'especialització de les càmeres i microscòpis utilitzats en les operacions i la il·luminació proporcionada via fibra òptica o díodes LED, entre d'altres factors, han permès desenvolupar el que es coneix com a cirurgia mínimament invasiva<sup>1</sup>.<sup>[4][6]</sup>

Aquesta tècnica amb la qual es persegueix la reducció del trauma post-operatori, s'aplica principalment en laparoscòpies i endoscòpies. Ofereix uns evidents beneficis per als pacients, però requereix d'equips mèdics especialment entrenats per dur-les a terme.<sup>[5]</sup>

Els equips mèdics que realitzen aquest tipus de cirurgies també es troben sovint amb oclusions que afecten directament a la qualitat de la imatge, com poden ser les produïdes per el vessament de sang o el fum resultant de cremar teixits.

Per això, en l'última dècada han sorgit diversos sistemes de visió per computador que intenten assistir la tasca d'aquests equips mèdics, mitjançant la detecció de l'instrumental mèdic utilitzat en aquest tipus de cirurgies.

---

<sup>1</sup> MIS, Minimally Invasive Surgery.

## 1.2 ESTAT DE L'ART

La majoria d'estudis per a l'assistència a la cirurgia mitjançant la detecció de l'instrumental mèdic s'han desenvolupat principalment per a les cirurgies MIS, encara que també existeixen estudis que han desenvolupat sistemes per a altres tipus de cirurgies com la microcirurgia ocular o la neurocirurgia.<sup>[2][13]</sup>

Un dels punts claus a l'hora de desenvolupar aquests sistemes és l'elecció del *data-set*<sup>2</sup> que s'utilitzarà per a la detecció, la *Taula 1* ens mostra perfectament les diferències entre els *data-sets* utilitzats per diversos estudis.

En referència al tipus d'enregistrament, podem diferenciar en 4 els tipus emprats: simulat, *phantom*, *ex-vivo* i *in-vivo*. Mentre que la simulació ens aporta un control total sobre la situació que enregistrem però un realisme molt baix sobre l'ambient clínic, l'*in-vivo* ens aporta un realisme total sobre l'ambient clínic però cap control sobre la situació que enregistrem.<sup>[8]</sup>

Pel que fa al mètode d'adquisició de les dades, podem veure a la *Taula 1* com es diferencien per el tipus de càmera utilitzada per l'enregistrament, el tipus d'imatge, la resolució d'aquesta, així com el nombre d'imatges i el de seqüències de vídeo diferents utilitzades.

En quant a les dades de referència manuals, s'han utilitzat principalment caixes englobants de l'eina<sup>3</sup>, l'extrem de l'instrument mèdic o bé, la posició de l'eina.

Finalment, també s'analitzen les condicions adverses que tenen en compte aquests *data-sets*, així com la disponibilitat d'aquests a la xarxa.

La falta de *data-sets* a l'abast de tothom i la no estandarització d'aquests, fan difícil la seva comparació i combinació, fets que són coll d'ampolla per a l'avenç en la matèria d'estudi que comprèn aquest projecte. Aquests dos factors dificulten les tasques d'elaboració d'una classificació de sistemes que obtenen millors resultats o la robustesa a la que és possible arribar, per la dificultat de combinar diverses fonts sense una prèvia estandarització.

---

<sup>2</sup> Conjunt de dades, en aquest cas d'imatges.

<sup>3</sup> Bounding boxes.

	Study conditions				Data acquisition					Data reference				Challenging cond.				Availability	
	Surg. spec.	Data type				Rec. device	Image type	Resolution	Size		Manual				Lighting	Occlusion	Smoke		Motion blur
		Simulation	Phantom	Ex-vivo	In-vivo				# Images	# Seq.	Bound. box	Tip pos.	Pose	System					
(Allan et al., 2013)	M	✓	✓	✓	✓	-	M M M	720 × 288	-	15		✓	✓	✓				✓	
(Allan et al., 2014)	M			✓		E	S	1920 × 1080	400	1			✓					✓	
(Allan et al., 2015)	M			✓		E	S	720 × 576	1000	1			✓					✓	
(Alsheikhali et al., 2015)	R				✓	M	M	1920 × 1080	400	1			✓	✓				✓	
(Bouget et al., 2015)	N					M	M	612 × 460	2476	14	✓				✓	✓		✓	
(Burschka et al., 2005)	E M	✓	✓			-	S	-	< 500	1								✓	
(Cano et al., 2008)	M		✓			E	M	-	550	2								✓	
(Doignon et al., 2005)	M				✓	E	M	640 × 480	52	1					✓				
(Doignon et al., 2007)	M		✓		✓	E	M	-	30	1					✓				
(Haase et al., 2013)	M			✓		E T E T	S - S -	640 × 480 64 × 50 640 × 480 64 × 50	30	2		✓				✓			
(Kumar et al., 2013b)	M				✓	E	M	-	3k2	16	✓				✓	✓	✓	✓	
(Li et al., 2014)	E M				✓	M	M	640 × 480	1k5	4		✓	✓			✓		✓	
(McKenna et al., 2005)	M				✓	E	M	720 × 576	> 835	2	✓								
(Pezzementi et al., 2009)	M E M	✓	✓			- M E	S	640 × 480	> 100 > 65 > 175	1 1 1									
(Reiter & Allen, 2010)	M				✓	-	-	-	> 137	1					✓			✓	
(Reiter et al., 2012c)	M				✓	E	S	-	> 1k6	5	+				✓				
(Reiter et al., 2012a)	M				✓	E	S	-	> 1k6	5	+				✓				
(Richa et al., 2011a)	E		✓			M	S	1k6 × 1k2	1k1	2					✓	✓		✓	
(Rieke et al., 2015)	E				✓	M	M	1920 × 1080	800	4	✓		✓		✓				
(Speidel et al., 2006)	M				✓	E	M	-	> 60	1	✓								
(Speidel et al., 2008)	M			✓		E	S	768 × 576	560 100	6 100	✓	✓			✓	✓			
(Speidel et al., 2014)	M				✓	E	S	320 × 240	542	6		✓				✓	✓		
(Sznitman et al., 2012)	E M				✓	-	M	640 × 480	1k5 1k	4 1	✓				✓	✓		✓	
(Sznitman et al., 2013)	E		✓			C	M	1k6 × 1k2	> 400 850	1 1		✓	✓	✓	✓		✓	✓	
(Sznitman et al., 2014)	M N				✓	-	M	640 × 360 320 × 240	1390 472	2 1	✓	✓				✓	✓	✓	
(Voros et al., 2007)	M			✓		E	M	200 × 100	6 45	6 1		✓			✓	✓	✓		
(Wolf et al., 2011)	M	✓	✓			- E	M	- 768 × 576	1050 1050	1 1			✓	✓					
(Zhou & Payandeh, 2014)	M		✓		✓	E	M	-	2000	10					✓	✓	✓	✓	

Taula 1: Resum de les característiques dels diferents data-sets utilitzats en diversos estudis per a la detecció d'instrumental mèdic.<sup>[8]</sup>

Un altre punt clau a l'hora de desenvolupar un sistema de detecció com el que es proposa en aquest projecte, és l'elecció del conjunt de característiques que faran possible la descripció de l'instrumental mèdic al classificador. La Taula 2 ens mostra els diferents conjunts de característiques que s'han seleccionat per diversos estudis.

	Overall		Features								Pose estim.			Prior knowledge			Tracking			
	Strategy	Tracking	Color	Gradients	HOG	Texture	Shape	Motion	Depth	Semantic Labels	Discriminative	Generative	Ad-hoc	Tool shape	Tool location	User assist.	Kinematics	Bayes.	Particle	Initialisation
(Allan et al., 2013)	G		✓																	
(Allan et al., 2014)	G	✓	✓			✓				✓			✓	✓	✓			✓		
(Allan et al., 2015)	G	✓	✓					✓					✓	✓	✓					
(Alsheekhali et al., 2015)	A		✓	✓																
(Bouget et al., 2015)	D		✓	✓	✓	✓				✓										
(Burschka et al., 2005)	A D		✓	✓						✓			✓				✓			
(Cano et al., 2008)	A	✓	✓	✓									✓	✓	✓					✓
(Doignon et al., 2005)	A		✓	✓									✓	✓	✓					
(Doignon et al., 2007)	G		✓	✓									✓	✓	✓					
(Haase et al., 2013)	A		✓	✓									✓	✓	✓					
(Kumar et al., 2013b)	D	✓			✓	✓		✓												✓
(Li et al., 2014)	D	✓				✓										✓				+
(McKenna et al., 2005)	G	✓	✓							✓				✓	✓				✓	
(Pezzemanti et al., 2009)	G		✓			✓							✓	✓						
(Reiter & Allen, 2010)	G	✓	✓			✓							✓		✓					✓
(Reiter et al., 2012c)	G	✓	✓			✓							✓				✓			
(Reiter et al., 2012a)	D		✓	✓						✓			✓	✓			✓			
(Richa et al., 2011a)	D	✓	✓							✓							✓			✓
(Rieke et al., 2015)	D	✓	✓		✓					✓				✓						✓
(Speidel et al., 2006)	G	✓	✓											✓					✓	
(Speidel et al., 2008)	A		✓										✓							
(Speidel et al., 2014)	A	✓	✓					✓					✓	✓						✓
(Sznitman et al., 2012)	D	✓		✓											✓					+
(Sznitman et al., 2013)	G D	✓	✓															✓		
(Sznitman et al., 2014)	A			✓						✓			✓	✓						
(Voros et al., 2007)	A			✓									✓	✓						
(Wolf et al., 2011)	G	✓		✓									✓	✓	✓				✓	
(Zhou & Payandeh, 2014)	A	✓		✓									✓	✓					✓	

Taula 2: Resum dels conjunts de característiques utilitzades per diversos estudis per a la detecció d'instrumental mèdic.<sup>[8]</sup>

Podem veure a la Taula 2 com s'han utilitzat bàsicament 3 estratègies per a la detecció de les eines i l'estimació de la seva posició. L'estratègia discriminativa es basa en la utilització de classificadors discrets, mentre que la generativa es basa en trobar un model de regressió per al conjunt de característiques seleccionat. En canvi, l'estratègia ad-hoc es basa en la combinació empírica de diversos models simples per a la detecció.

En quant a les característiques més seleccionades destaca l'ús del color en diferents espais, ja que l'elecció correcta d'aquest espai ens permet discriminar millor entre eines metàl·liques i teixit orgànic. Una altra característica a destacar són els gradients, àmpliament utilitzats en l'àmbit de la visió per computador.

Pel que fa al coneixement previ, el més utilitzat per aquests sistemes són la forma de l'eina i la seva posició. I pel que fa al seguiment de l'instrument, la tècnica majoritària ha estat la coneguda com *tracking by initialization*, que es basa en la posició detectada en el *frame* anterior per a fer aquest seguiment.

# Capítol 2

## Plantejament

### 2.1 FORMULACIÓ DEL PROBLEMA

Tenint en compte el context i l'estat de l'art exposats en el capítol anterior, és clar que existeix la necessitat o utilitat de brindar una certa assistència als equips mèdics que porten a terme cirurgies mínimament invasives.

Així doncs, inicialment es va plantejar el projecte com el desenvolupament d'un sistema d'assistència a les cirurgies *MIS* mitjançant la detecció de l'instrumental mèdic utilitzat en aquest tipus de cirurgies. Per al desenvolupament de la solució es van establir certes característiques que hauria de complir, així com l'abast que la solució hauria d'abarcara.

Pel que fa a les característiques, es va definir que la solució hauria de dotar-se d'un *data-set* propi del que servir-se alhora de construir el classificador, així com basar-se en característiques de forma per discriminar l'eina.

Pel que fa a l'abast, es van tenir en compte dues consideracions:

En primer lloc, el *data-set* cobrirà situacions d'oclusió de l'instrument, ja siguin causades per algun òrgan o pel fum resultant de cremar teixits. Però no les situacions en que l'objectiu de la càmera queda tacat de sang, ja que es considera que la limitació temporal del projecte ho fa inassumible.

En segon lloc, no es pretén fer un sistema que sigui capaç de detectar tot l'instrumental mèdic utilitzat en les *MIS*, sinó un que sigui capaç de detectar més d'una eina, però amb un rang reduït de variacions entre aquestes. Aquesta decisió es justifica per dos factors: Hi ha diversos tipus de *MIS* i cadascuna utilitza un rang d'eines diferent; S'utilitzaran característiques de forma per discriminar els instruments, per tant, s'assumeix que no es pot abarcar un rang molt variat d'eines a reconèixer. A més a més, si tenim en compte el cost relativament baix de desenvolupar sistemes diferents per cirurgies diferents, el qual



es creu que és objectivament assumible, la decisió queda més que justificada.

Finalment, es pot afirmar amb rotunditat que la solució que es presenta en aquest treball resol en gran mesura el problema plantejat. Ja que compleix amb les característiques establertes inicialment, les quals s'expliquen amb molt més detall al capítol 6 d'aquesta memòria. Així com amb l'abast al que aspirava el plantejament inicial de la solució.

## 2.2 OBJECTIUS DEL PROJECTE

Tal hi com ja s'ha exposat a la secció anterior, el projecte soluciona el problema plantejat inicialment en bona mesura, ja que s'ha estat capaç de desenvolupar un sistema de detecció de l'instrumental mèdic utilitzat en una cirurgia mínimament invasiva. Aquest sistema a més, compleix en bona mesura els objectius que havien estat plantejats inicialment. A continuació, s'explica el grau de compliment d'objectius inicials al que s'ha arribat:

S'ha elaborat un *data-set* propi que cobreix la majoria de situacions que es poden donar. A més, s'han cobert també situacions difícils, com pot ser la oclusió de l'eina a causa del fum resultant de cremar teixits. Encara que, lamentablement, no ha estat possible que aquest *data-set* estigués dotat d'imatges de diferents fonts.

És possible la detecció en temps real de l'instrument, ja que la solució és potencialment paral·lelitzable, però la implementació de paral·lelisme en la solució no s'ha pogut realitzar a causa de la limitació temporal del treball.

Per a la detecció de les eines s'han utilitzat característiques de forma, concretament s'ha fet ús dels moments de Hu, que són un subconjunt dels moments invariants de la imatge.

No s'ha pogut provar que la solució sigui capaç de reconèixer dos o més tipus d'instruments, ja que les fonts escollides tant per generar les imatges del *data-set*, com per provar la solució, el cirurgià només utilitza una eina. El criteri per escollir aquestes fonts ha estat prioritzar el nombre de situacions que cobrien més enllà de les eines que hi poguessin aparèixer. No obstant, és necessari recalcar que la eina utilitzada té forma de pinça en el seu extrem, per tant, les seves formes varien bastant en funció de si la pinça és oberta o tancada. Així doncs, si tenim en compte que es detecta l'eina utilitzant

característiques de forma, el fet de que s'utilitzi aquest tipus d'eina ha afegit dificultat a la detecció d'aquesta.

# Capítol 3

## Gestió del projecte

### 3.1 METODOLOGIA I RIGOR

#### 3.1.1 METODOLOGIA

Per a la realització del projecte s'ha utilitzat la metodologia de treball *Scrum*. Així doncs, s'ha prioritzat tenir una primera versió amb unes funcionalitats mínimes i s'ha iterat sobre la solució, de manera que aquesta ha anat adquirint més qualitat. Aquestes iteracions, anomenades *sprints*, s'han portat a terme de forma bisetmanal.

#### 3.1.2 EINES DE SEGUIMENT

Pel que fa a les eines de seguiment, principalment s'ha utilitzat el correu electrònic per a què el director fes el seguiment de quin era l'estat del projecte, tot i que en cas de trobar alguna dificultat que impedís de seguir desenvolupant la solució, s'ha demanat una reunió amb director del projecte per discutir sobre possibles alternatives de solució i definir el nou pla d'acció.

### **3.1.1 VALIDACIÓ**

A partir d'un conjunt d'imatges de prova que cobreixen gran part de les situacions que es poden donar durant una laparoscòpia. Per a cada *sprint*, s'ha evaluat la qualitat de la solució obtinguda basant-se en el nombre d'errors de detecció d'aquesta. Fent especial èmfasi en el nombre de falsos positius i negatius obtinguts. I analitzant el perquè d'aquests errors, a fi de poder definir les tasques del següent *sprint* de manera que aquests es minimitzessin.

## **3.2 PLANIFICACIÓ TEMPORAL**

### **3.2.1 FASES DEL PROJECTE**

Inicialment, es va dividir el projecte en diverses fases, les quals s'enumeren a continuació:

- Documentació de la fita inicial i recollida de dades.
- Anàlisi i preprocessament de dades.
- Extracció de característiques.
- Obtenció de la funció descriptora.
- Filtratge de la imatge i detecció.
- Validació de resultats.
- Documentació final.

### **3.2.2 ESPECIFICACIÓ DE TASQUES**

Per a cada fase de les enumerades a la secció anterior es van especificar les diverses tasques que s'havien de portar a terme. A continuació, s'exposen:

### **Documentació de la fita inicial i recollida de dades:**

Recollida de *data-sets*, estudi de l'estat de l'art de la temàtica del projecte i la seva planificació temporal, així com l'adquisició del coneixement necessari per portar a terme les diferents fases del projecte. També s'hi va incloure l'anàlisi econòmic i de sostenibilitat i la resta de requeriments necessaris per assolir la fita inicial.

### **Anàlisi i preprocessament de dades:**

Anàlisi dels *data-sets* recollits per determinar la seva inclusió o exclusió del projecte. També s'hi va incloure la segmentació manual de l'instrumental mèdic que aparegui en les diverses fonts recollides.

### **Extracció de característiques:**

Implementació del càlcul de característiques de forma invariants, bàsicament es va decidir centrar el conjunt d'aquestes característiques en diversos moments de Hu i de Zernike per a la seva posterior selecció.

### **Obtenció de la funció descriptora:**

Aplicació d'un classificador capaç de ponderar les diferents característiques extrems en la fase anterior, per tal d'obtenir una funció lineal o no lineal que pugui ser utilitzada en el filtratge de la imatge. En aquest cas, es perfilaven dos tipus de classificadors: SVM<sup>4</sup> i CNN<sup>5</sup>.

### **Filtratge de la imatge i detecció:**

Aplicació de la funció descriptora sobre la imatge amb un llindar que caldria estimar utilitzant força bruta i comparant els resultats obtinguts. Així com programar la marcadura dels resultats de la filtració en la imatge, per tal de fer efectiva la detecció a nivell visual.

### **Validació de resultats:**

Per un anàlisi dels resultats que s'obtenen, posant especial èmfasi en la taxa de falsos positius i falsos negatius, així com en les situacions on aquestes taxes són més altes. Finalment, s'hauria de tornar a iterar sobre les diferents fases per tal d'aconseguir una

---

<sup>4</sup> Support Vector Machine.

<sup>5</sup> Convolutional Neural Network.

detecció correcta en aquestes situacions, tenint en compte el risc de caure en *overfitting*<sup>6</sup> que això comporta.

#### **Documentació final:**

Redacció de la memòria definitiva del projecte.

### **3.2.3 RECURSOS**

Els recursos utilitzats durant l'elaboració del projecte són els següents:

- MacBook Air 2012 amb sistema operatiu OSX El Capitan.
- Matlab en versió R2015a, per al desenvolupament del codi del projecte.
- Google Drive, per a la gestió documental.
- Google Docs, per a l'elaboració de documents.
- Gantter, per a la planificació temporal.
- Els coneixements adquirits al grau, per a desenvolupar els diferents rols que requereix el projecte.

---

<sup>6</sup> Ajustar extremadament el model a un conjunt de dades en particular que deriva en una pèrdua total de generalitat.

### 3.2.4 GESTIÓ DEL TEMPS

A la taula que es mostra a continuació, s'indiquen les hores que s'han dedicat a cada fase:

<b>FASE</b>	<b>DEDICACIÓ</b>
Documentació i recollida de dades	98h
Anàlisi i preprocessament de dades	21h
Extracció de característiques	20h
Obtenció de la funció descriptora	12h
Filtratge de la imatge i detecció	36h
Validació de resultats	83h
Documentació final	30h
<b>TOTAL</b>	<b>300 hores</b>

*Taula 3: Gestió del temps*

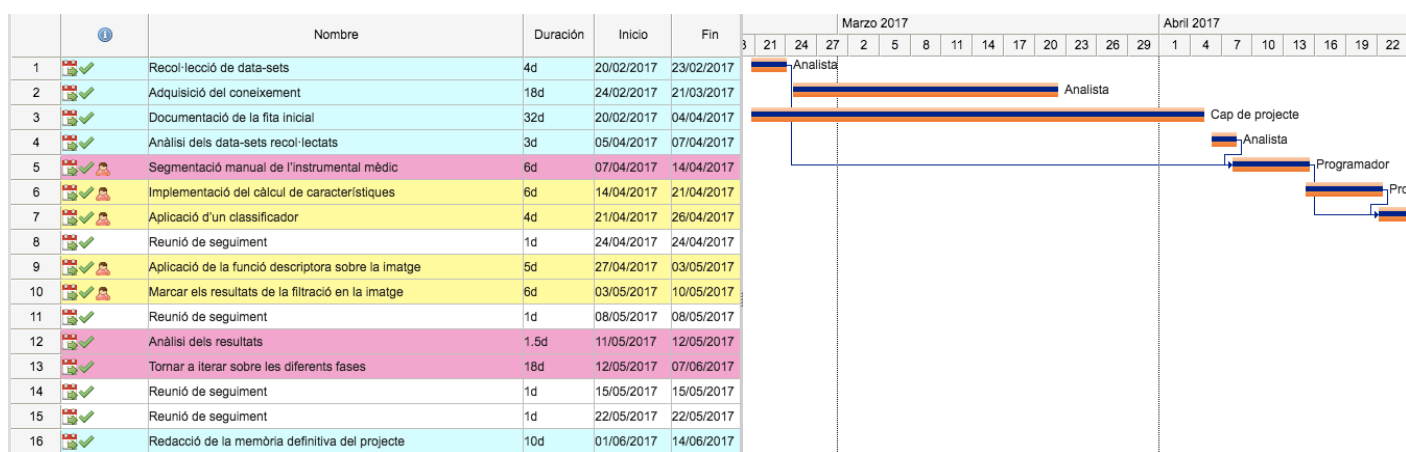
### 3.2.5 PLANIFICACIÓ TEMPORAL FINAL

Inicialment, la duració del projecte estava estimada en 16 setmanes, del 20 de febrer de 2017 fins al 14 de juny de 2017. Finalment, però, la duració del projecte ha estat de 17 setmanes, del 20 de febrer de 2017 fins al 20 de juny de 2017, dia escollit per la universitat per presentar la memòria final del projecte.

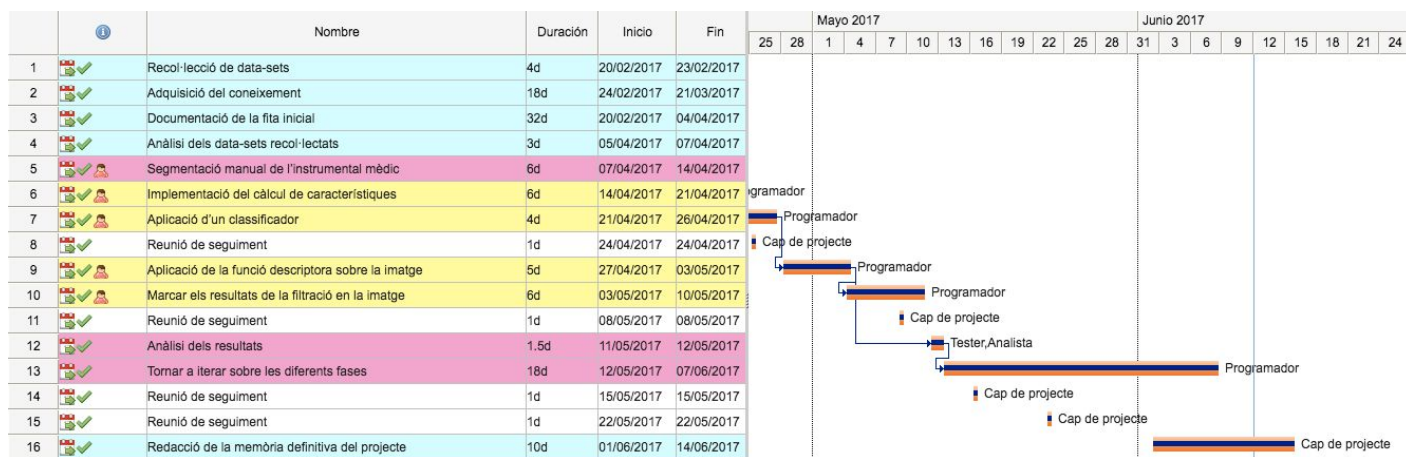
Així doncs, pel que fa a la planificació del projecte, no s'han produït grans canvis respecte de la planificació inicial, més enllà d'un petit endarreriment en la reiteració de les diferents fases per millorar la qualitat de la solució. Aquest endarreriment però, s'ha compensat afegint 15 hores de treball aprofitant els 6 dies que s'havien guanyat en la entrega de la memòria del projecte.

### 3.2.6 DIAGRAMA DE GANTT

A continuació, es mostra el diagrama de Gantt del projecte amb la planificació definitiva de les tasques i el rol que les ha portat a terme. Així com, el risc que es preveia que tindria cadascuna d'elles inicialment<sup>7</sup>. El diagrama ha estat dividit en dos parts per facilitar la seva visualització.



Figures 1: Diagrama de Gantt definitiu I



Figures 2: Diagrama de Gantt definitiu II

<sup>7</sup> Rosa: Risc Alt. Groc: Risc moderat. Blau: Risc baix.



### 3.3 PRESSUPOST

A continuació es detallen els pressupostos per als recursos humans, de software i de hardware que han estat necessaris per portar a terme les tasques presents al diagrama de Gantt. Així com també, els costos indirectes d'aquestes.

#### 3.3.1 PRESSUPOST DE RECURSOS HUMANS

Per a l'elaboració del projecte ha estat necessari assignar 4 rols, que han estat desenvolupats per una mateixa persona, ja que el projecte s'ha portat a terme de forma individual. A continuació es detallen les hores i tasques desenvolupades per cada rol i el seu cost:

ROL	TASQUES ASSIGNADES	HORES	PREU	PREU TOTAL
Cap de projecte	<ul style="list-style-type: none"><li>• Documentació de la fita inicial.</li><li>• Reunions de seguiment.</li><li>• Redacció de la memòria definitiva del projecte.</li></ul>	90	50€/h	4500€
Analista	<ul style="list-style-type: none"><li>• Recol·lecció de data-sets.</li><li>• Adquisició del coneixement.</li><li>• Anàlisi dels <i>data-sets</i> recollits.</li><li>• Proves de rendiment i anàlisi dels resultats.</li></ul>	59	40€/h	2360€
Programador	<ul style="list-style-type: none"><li>• Segmentació manual de l'instrumental mèdic.</li><li>• Implementació del càlcul de característiques.</li><li>• Aplicació d'un clasificador.</li><li>• Aplicació de la funció descriptora sobre la imatge.</li><li>• Marcar els resultats de la filtració en la imatge.</li><li>• Tornar a iterar sobre les diferents fases.</li></ul>	131	35€/h	4585€
Tester	<ul style="list-style-type: none"><li>• Proves de rendiment i anàlisi dels resultats.</li></ul>	20	20€/h	400€
<b>TOTAL</b>		<b>300</b>		<b>11845€</b>

Taula 4: Pressupost de recursos humans

### 3.3.2 PRESSUPOST DE RECURSOS DE SOFTWARE

Per a l'elaboració del projecte han estat necessàries diverses eines de software que es detallen a continuació, juntament amb el preu, les tasques implicades, la vida útil d'aquestes i l'amortització que s'en deriva.

PRODUCTE	PREU	TASQUES IMPLICADES	VIDA ÚTIL	AMORTITZACIÓ
Google Docs	0€	<ul style="list-style-type: none"> <li>• Documentació de la fita inicial.</li> <li>• Redacció de la memòria definitiva del projecte.</li> </ul>	-	0€
Google Drive	0€	<ul style="list-style-type: none"> <li>• Documentació de la fita inicial.</li> <li>• Redacció de la memòria definitiva del projecte.</li> </ul>	-	0€
Gantter	0€	<ul style="list-style-type: none"> <li>• Documentació de la fita inicial.</li> <li>• Redacció de la memòria definitiva del projecte.</li> </ul>	3 anys	0€
MATLAB Suite <sup>8</sup>	69€	<ul style="list-style-type: none"> <li>• Segmentació manual de l'instrumental mèdic.</li> <li>• Implementació del càlcul de característiques.</li> <li>• Aplicació d'un classificador.</li> <li>• Aplicació de la funció descriptora sobre la imatge.</li> <li>• Marcar els resultats de la filtració en la imatge.</li> <li>• Proves de rendiment i anàlisi dels resultats.</li> </ul>	3 anys	7€
GIMP	0€	<ul style="list-style-type: none"> <li>• Segmentació manual de l'instrumental mèdic.</li> </ul>	3 anys	0€
OSX El Capitan	0€	<ul style="list-style-type: none"> <li>• Totes.</li> </ul>	3 anys	0€
<b>TOTAL</b>	<b>69€</b>			<b>7€</b>

Taula 5: Pressupost de recursos de software

<sup>8</sup> <https://www.mathworks.com/pricing-licensing.html?prodcode=ML&intendeduse=student>

### 3.3.3 PRESSUPOST DE RECURSOS DE HARDWARE

A la taula que es mostra a continuació s'indiquen els diferents recursos de hardware que s'han utilitzat, així com el seu preu, les tasques implicades, la vida útil i l'amortització prevista d'aquests.

PRODUCTE	PREU	TASQUES IMPLICADES	VIDA ÚTIL	AMORTITZACIÓ
MacBook Air 13'' <sup>9</sup>	1099€	Totes	5 anys	73€
Disc Dur Toshiba 1TB <sup>10</sup>	20€	<ul style="list-style-type: none"><li>• Recol·lecció de data-sets.</li><li>• Anàlisi del <i>data-sets</i>.</li></ul>	5 anys	1€
<b>TOTAL</b>	<b>1119€</b>			<b>74€</b>

Taula 6: Pressupost de recursos de hardware

### 3.3.4 COSTOS INDIRECTES

Per al càlcul dels costos indirectes s'ha contemplat únicament el consum d'electricitat. Inicialment, s'havien previst també el possible augment de recursos humans en cas de desviació, i les eventuais reparacions que poguessin requerir els recursos de hardware. En el primer cas, si ha estat necessari un augment d'hores de feina de programador, el cost de les quals ha estat afegit al pressupost de recursos humans. Per contra, no ha estat necessària cap reparació de hardware.

PRODUCTE	PREU TOTAL
Electricitat	2100€
<b>TOTAL</b>	<b>2100€</b>

Taula 7: Costos indirectes

<sup>9</sup> <http://www.apple.com/es/shop/buy-mac/macbook-air>

<sup>10</sup> <http://www.toshiba.eu/hard-drives/portable/canvio-basics/>

### 3.3.5 PRESSUPOST TOTAL

Finalment, tenint en compte els diversos recursos i costos identificats, el pressupost total del projecte serà el següent:

<b>CONCEPTE</b>	<b>COST</b>
Recursos humans	11845€
Recursos de software	69€
Recursos de hardware	1119€
Costos indirectes	2100€
<b>TOTAL</b>	<b>15133€</b>

*Taula 8: Pressupost total*

### 3.3.6 CONTROL DE GESTIÓ

El control de gestió del projecte s'ha basat en el seguiment dels temps especificats en el diagrama de Gantt i en les reunions de seguiment amb el director del projecte. A més, el fet d'utilitzar una metodologia de treball *Scrum* i la utilització d'*sprints* bisetmantals, ha permès detectar les eventuais desviacions de manera ràpida i eficient.

En el cas de la desviació en la fase de validació, s'ha utilitzat el diagrama de Gantt per reorganitzar les tasques i s'ha augmentat el pressupost de recursos humans, per tal de complir amb les restriccions temporals del projecte.

# Capítol 4

## Informe de sostenibilitat

### 4.1 DIMENSIÓ ECONÒMICA

En els pressupostos mostrats anteriorment s'han evaluat els costos tant de recursos materials com humans. Així com el cost de les desviacions o reparacions que puguin sorgir durant el desenvolupament del projecte. El correcte ajustament dels recursos i els temps necessaris per portar-lo a terme fan d'aquest un projecte competitiu.

### 4.2 DIMENSIÓ SOCIAL

L'aplicació de la visió per computador s'ha començat a aplicar a la medicina en els últims 10 anys, tot i que actualment ja existeixen robots que són capaços d'operar de manera autònoma<sup>11</sup>. El problema és que els costos d'aquests robots o aplicacions solen ser molt cars i, per tant, a l'abast de molt pocs. Aquest projecte va en la línia d'afavorir la utilització de la visió per computador en operacions mèdiques amb uns costos més assumibles.

De la seva utilització se'n beneficiarà el pacient, ja que millorarà la seva qualitat de vida i reduirà el temps de recuperació post-operatori. Així com, l'hospital, ja que veurà reduït els costos de medicació i residència del pacient, i amb un cost d'implementació del sistema relativament petit, si tenim en compte que no cal modificar ni l'instrumental mèdic ni la sala d'operacions.

---

<sup>11</sup> Robot *Da Vinci*.

A més, el projecte pot ajudar als cirurgians a desenvolupar la seva feina amb més comoditat i precisió. Així com a reduir els possibles errors en les operacions i, per tant, a prevenir als pacients dels perjudicis que els puguin ocasionar aquests errors.

### 4.3 DIMENSIÓ AMBIENTAL

L'utilització de recursos amb una petjada medioambiental destacable en el projecte es redueix a els recursos de hardware utilitzats i a l'electricitat consumida.

### 4.4 TAULA DE SOSTENIBILITAT

<b>SOSTENIBILITAT</b>	<b>ECONÒMICA</b>	<b>SOCIAL</b>	<b>AMBIENTAL</b>
<b>PLANIFICACIÓ</b>	Apartat 4.1	Apartat 4.2	Apartat 4.3
<b>VALORACIÓ</b>	7	9	8

*Taula 9: Taula de sostenibilitat*

# Capítol 5

## Plantejament de la solució

### 5.1 PLANTEJAMENT DE LA SOLUCIÓ

La solució plantejada finalment és la següent:

- Donat el video d'una laparoscòpia per a la reparació d'hèrnies i un model de classificació, generat prèviament a partir del *data-set*.
- S'extreuen els *frames*<sup>12</sup> i es guarden en una carpeta del sistema d'arxius.
- Per a cada *frame*:
  - ◆ Preprocessat del *frame* mitjançant el color, operadors morfològics i gradients.
  - ◆ Filtrat del *frame* per diversos llindars.
  - ◆ Per cada imatge binària resultant de la filtració del *frame* per un llindar determinat:
    - Classificació de cada component connexa mitjançant el model prèviament generat.
    - Generació de la imatge resultat ponderada a partir de les puntuacions obtingudes per cada component connexa en la classificació.
  - ◆ Obtenció de la imatge resultat mitjançant la suma de totes les imatges resultat ponderades obtingudes per cada filtració del *frame*.
  - ◆ Marcació del resultat obtingut en el *frame* fent ús d'operadors morfològics.
  - ◆ Guardar el *frame* "marcat" en una carpeta del sistema d'arxius.
- Un cop marcats tots els *frames*, es carreguen de la carpeta on han estat desats i es genera el vídeo resultat.

---

<sup>12</sup> Imatges que formen el vídeo.

El procés que s'ha seguit per arribar a la solució, així com, els detalls d'implementació i les explicacions teòriques de les tècniques mencionades aquí, es troben a la secció corresponent del capítol 6.



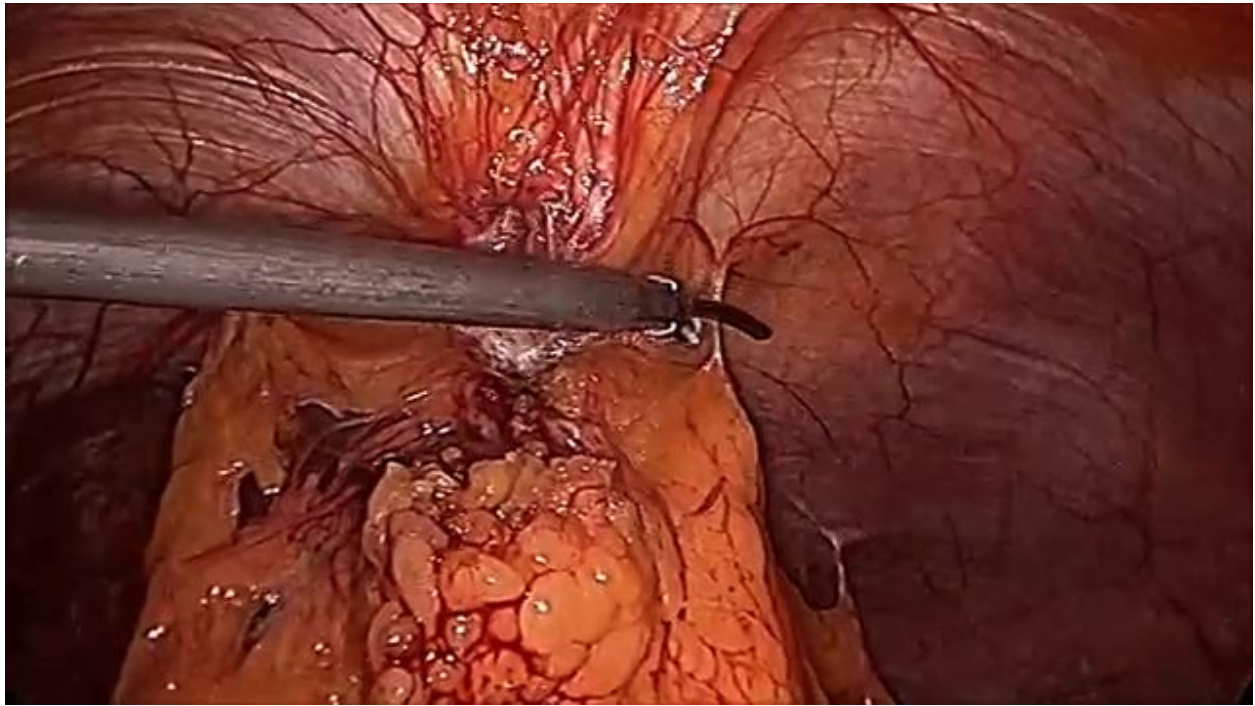
# Capítol 6

## Implementació

### 6.1 EL-LABORACIÓ D'UN DATA-SET PROPI

Primerament, es van recopilar diversos *data-sets* d'altres estudis i es va analitzar la seva possible reutilització. Alguns es van descartar perquè havien estat recopilats utilitzant imatges d'altres cirurgies mínimament invasives com endoscòpies o microcirurgies oculars. Finalment, els que si que utilitzaven imatges de laparoscòpia o que, com a mínim, utilitzaven eines comuns que també s'utilitzen en la cirurgia del nostre interès, van acabar essent descartades perquè la segmentació manual dels instruments no es feia sobre el seu contorn, sinó mitjançant una capça englobant sobre l'eina o sobre parts d'aquesta.<sup>[1][2][13]</sup>

Descartada la reutilització de *data-sets*, es van seleccionar les imatges més representatives de la font escollida, segons el criteri de cobrir la major part de situacions que es poden donar en una laparoscòpia.



*Figura 3: Imatge utilitzada en la construcció del data-set*



*Figura 4: Imatge utilitzada en la construcció del data-set*



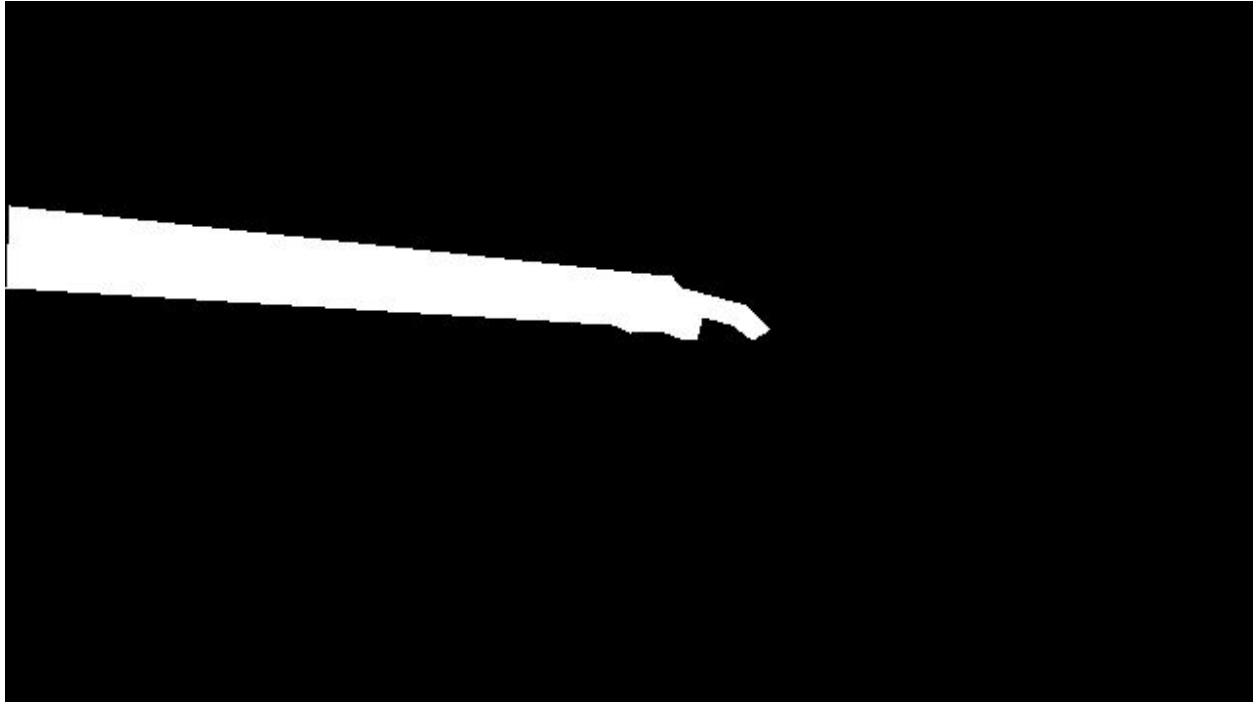
*Figura 5: Imatge utilitzada en la construcció del data-set*

Un cop escollits, va ser necessari segmentar manualment les imatges seleccionades, generant així un conjunt de 80 imatges binàries de l'eina a reconèixer. Aquesta segmentació manual, juntament amb un conjunt d'imatges binàries obtingudes seguint el següent algoritme:

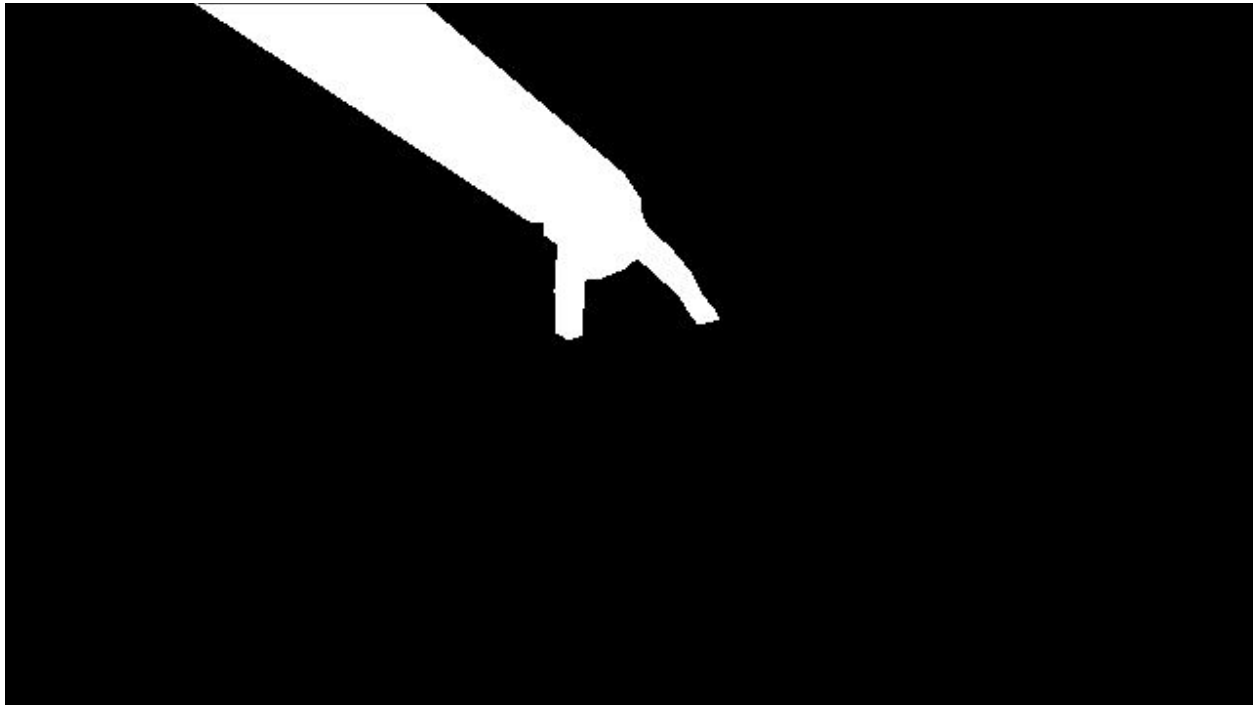
→ Per cada imatge original que ha estat segmentada manualment:

- ◆ Conversió a escala de grisos de la imatge original.
- ◆ Filtrat de la imatge mitjançant la binarització d'aquesta en diversos llindars. Per a cadascuna de les imatges resultants de binaritzar per un determinat llindar :
  - Eliminar les components connexes amb àrees inferiors a 500 píxels.
  - Per a cada una de les components connexes de la imatge:
    - Si la component connexa té un cert percentatge de píxels en comú amb l'eina segmentada manualment:
      - ◆ Guardar la imatge d'aquesta component connexa al conjunt de classe negativa.

representaven les classes positiva i negativa, respectivament, que van ser utilitzades com a primer *data-set* per construir el classificador. El funcionament del classificador utilitzat en aquest projecte està ampliament explicat a la secció 6.3 d'aquest capítol.



*Figura 6: Imatge binària de l'eina*



*Figura 7: Imatge binària de l'eina*



*Figura 8: Imatge binària de la classe negativa*



*Figura 9: Imatge binària de la classe negativa*

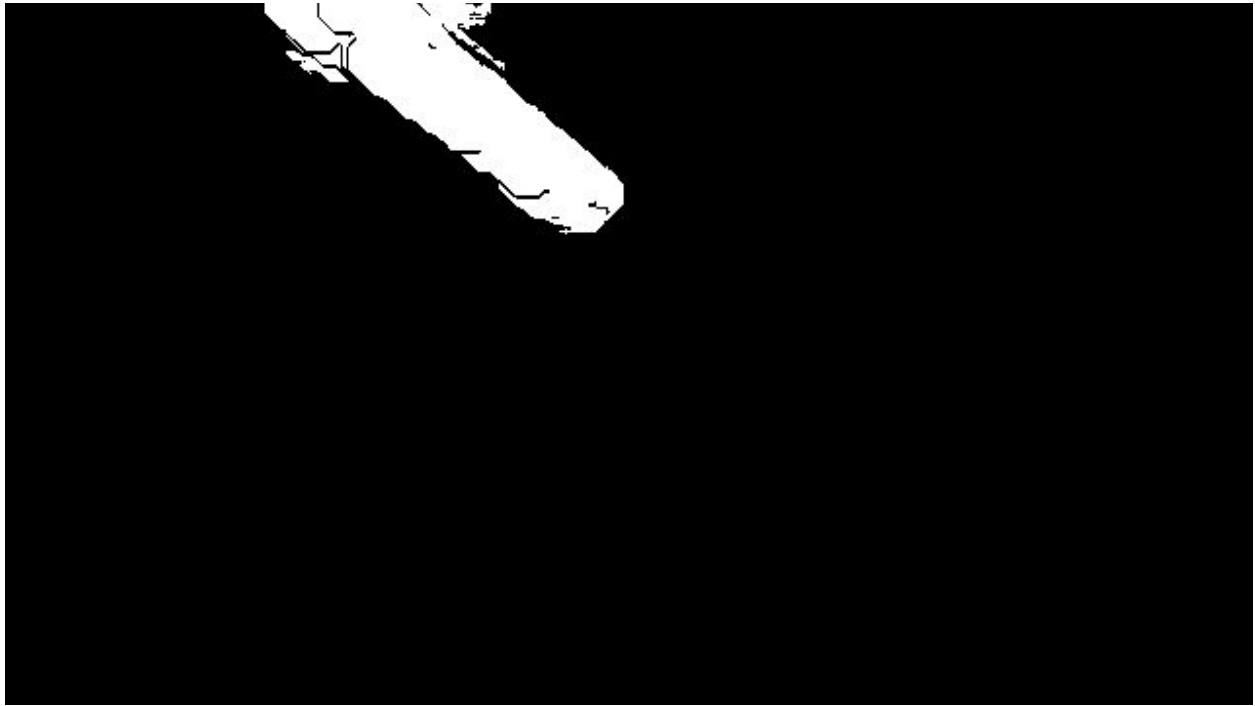
De seguida es va veure que seria necessària la generació automàtica d'un nou *data-set*, basant-se en la informació que es podia extreure de la segmentació manual. Ja que aquest, al haver estat obtingut manualment, li mancava realisme, és a dir: línies menys rectes, forats enmig de l'eina... En definitiva, al *data-set* li faltaven les imperfeccions de segmentar una regió de forma automàtica. Així doncs, es va generar les noves classes positiva i negativa del *data-set* seguint el següent algoritme:

- Per cada imatge original que ha estat segmentada manualment:
  - ◆ Conversió a escala de grisos de la imatge original.
  - ◆ Preprocessament de la imatge mitjançant operadors morfològics, gradients i color. Aquest pas i el següent, s'expliquen amb l'extensió que correspon a la secció 6.4 d'aquest capítol.
  - ◆ Filtrat de la imatge mitjançant la binarització d'aquesta en diversos llindars. Per a cadascuna de les imatges resultants de binaritzar per un determinat llindar :
    - Eliminar les components connexes amb àrees inferiors a un cert llindar.
    - Per a cada una de les components connexes de la imatge:

- Si la component connexa té un cert percentatge de píxels en comú i una determinada relació d'àrea amb l'eina segmentada manualment:
  - ◆ Guardar la imatge d'aquesta component connexa al conjunt de classe positiva.
- En cas contrari:
  - ◆ Guardar la imatge d'aquesta component connexa al conjunt de classe negativa.



*Figura 10: Imatge binària de la classe positiva*



*Figura 11: imatge binària de la classe positiva*



*Figura 12: imatge binària de la classe negativa*





*Figura 13: imatge binària de la classe negativa*

Un cop implementat l'algoritme, es van fer diverses proves per tal d'acotar els diferents llindars que determinen quan la imatge d'una component connexa s'ha d'incloure en el conjunt d'imatges de la classe positiva o negativa.

Finalment, es va optar per utilitzar els següents llindars per determinar com ha de ser classificada la imatge d'una component connexa:

- Classe positiva: 80% de píxels en comú amb l'eina i una àrea d'entre el 25 i el 120% de la de l'eina segmentada manualment.
- Classe negativa: menys del 60% de píxels en comú amb l'eina segmentada manualment.

Aquesta decisió es dotarà d'una justificació més extensa a la secció 6.7 d'aquest capítol, on s'analitzen els resultats obtinguts per a cada *data-set*.

## 6.2 EXTRACCIÓ DE CARACTERÍSTIQUES

Inicialment, el plantejament del projecte era descriure l'eina mitjançant característiques de forma. Així doncs, es va escollir un conjunt de moments de la imatge, també coneguts com a moments de Hu, com a característiques descriptores de l'eina.

El moments de Hu són moments de la imatge, és a dir, una mitjana ponderada de les intensitats dels píxels de la imatge o una combinació d'aquestes. I la seva principal característica és que aquests són invariants a la l'escalat, la rotació i la translació, raó per la qual es van escollir.<sup>[9][15]</sup>

Per a la implementació del seu càlcul, es va aprofitar la implementació que dona un popular llibre de visió per computador.<sup>[16]</sup>

## 6.3 CONSTRUCCIÓ DEL CLASSIFICADOR

Es buscava un model d'aprenentatge supervisat que fos capaç de classificar l'eina del cirurgià mitjançant una funció descriptora, ja que l'extracció d'aquesta funció permetria exportar l'algoritme de detecció a d'altres llenguatges de programació sense necessitat d'exportar el model sencer.

Inicialment, es van contemplar dos models que complien aquestes característiques, SVM i CNN. Es va escollir SVM perquè existia un coneixement previ del seu funcionament, però cal remarcar, que la comparació dels dos models hagués aportat positivament al projecte, no obstant, es va descartar perquè sobrepassava l'abast del projecte plantejat en un principi.

SVM és un model d'aprenentatge supervisat capaç de classificar entre dues classes d'elements definits per una sèrie de característiques computables. Per fer-ho, defineix un hiperplà en l'espai de característiques, el qual separa les dues classes d'elements de manera que, la distància que hi ha entre els elements de les classes més pròxims a l'hiperplà i aquest sigui la més gran possible.<sup>[10]</sup>

Els nous elements a classificar són representats en l'espai de característiques i es determina a quina classe pertanyen depenent del cantó del l'hiperplà en que es troben.

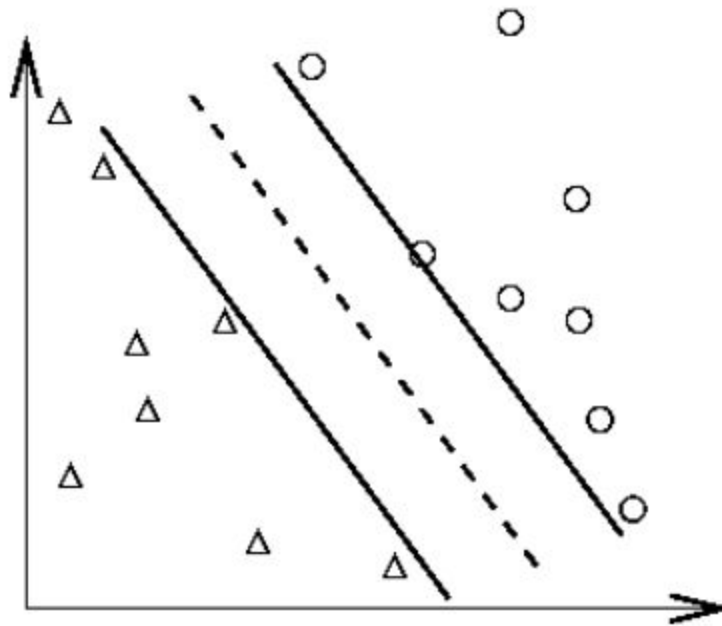


Figura 14: Model SVM en 2 dimensions<sup>[10]</sup>

Així doncs, un cop escollit el model es van començar a fer les primeres proves. Aquestes van consistir, bàsicament, en determinar quin tipus de *Kernel* seria l'adequat per al problema a resoldre. El *Kernel* és una funció de transformació entre espais dimensionals que SVM utilitza quan les dues classes no són linealment separables en l'espai de característiques. Llavors, mitjançant el *Kernel* es transforma aquest espai dimensional augmentant la seva dimensionalitat i facilitant així la separació entre classes.<sup>[10]</sup> Es van fer proves amb dos tipus de *Kernel*.

En primer lloc, es va provar un *Kernel* lineal, que, per les seves característiques, és adequat utilitzar-lo quan es pretèn classificar entre dues classes ben definides, per exemple, per classificar entre la classe "salmó" i la classe "sardina". Degut a la naturalesa del problema a resoldre, en el qual es persegueix la classificació entre la classe "eina" i la resta. Aquest *Kernel* donava resultats molt dolents.

Així doncs, es va provar d'utilitzar un *Kernel* gaussià, que, per les seves característiques, és adequat utilitzar-lo quan es pretèn classificar entre una classe i la resta, com és el cas

del problema que tracta aquest projecte. Per tant, aquest *Kernel* donava uns resultats molt millors i va ser finalment l'escollit.

## 6.4 PREPROCESSAMENT I FILTRAT D'IMATGE

Com ja s'ha explicat a la secció 6.1 d'aquest mateix capítol, ben d'hora es va veure que seria necessari un preprocessament de la imatge abans de que aquesta s'enviés al classificador per tal de reconèixer l'eina.

En primera instància, es va intentar aplicar operadors morfològics sobre la imatge original en escala de grisos, concretament, es van provar d'aplicar *Opening* i *Closing* sobre aquesta.

Donada una imatge i un element estructurant, que es pot entendre com una màscara que s'aplica sobre cada píxel de la imatge. Els operadors morfològics són un conjunt d'operacions que s'apliquen sobre la imatge, mitjançant la combinació d'aquesta i l'element estructurant utilitzant els operadors bàsics de conjunts<sup>13</sup> o una combinació d'aquests.<sup>[17]</sup>

Un *Closing* és un operador morfològic compost que consisteix en, donat un element estructurant, aplicar primer una dilació sobre la imatge, seguida d'una erosió. Ambdues utilitzant el mateix el element estructurant. Això permet eliminar els petits forats interiors de les diferents regions de la imatge mitjançant l'expansió dels tons de gris de més intensitat.

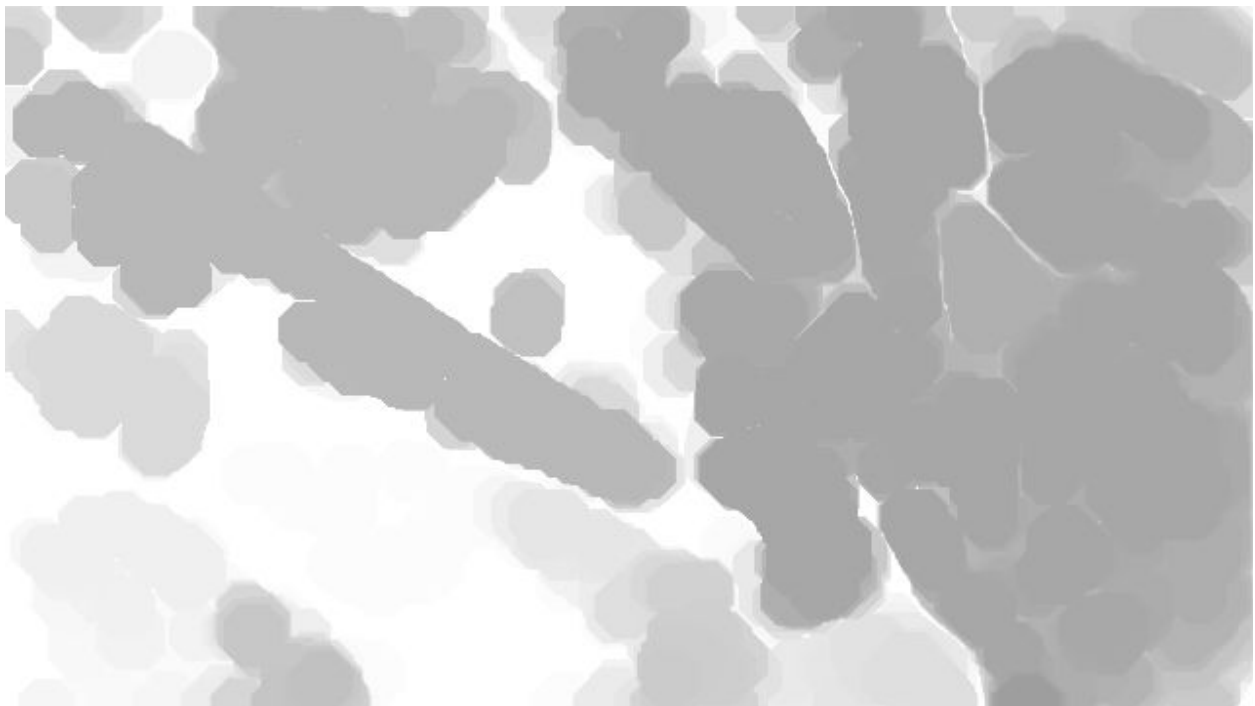
En canvi, un *Opening* consisteix en aplicar primer una erosió sobre la imatge i posteriorment una dil·lació. Obtenint com a resultat l'eliminació de les diferents regions de la imatge que tenen una àrea menor a la de l'element estructurant i que el seu entorn tingui una intensitat de gris menor, mitjançant l'expansió dels tons de gris de menys intensitat. L'objectiu de l'ús d'*Opening* o *Closing* en la imatge era netejar soroll d'aquesta i homogenitzar l'interior de l'eina a reconèixer.

---

<sup>13</sup> Intersecció, unió, inclusió i complement.



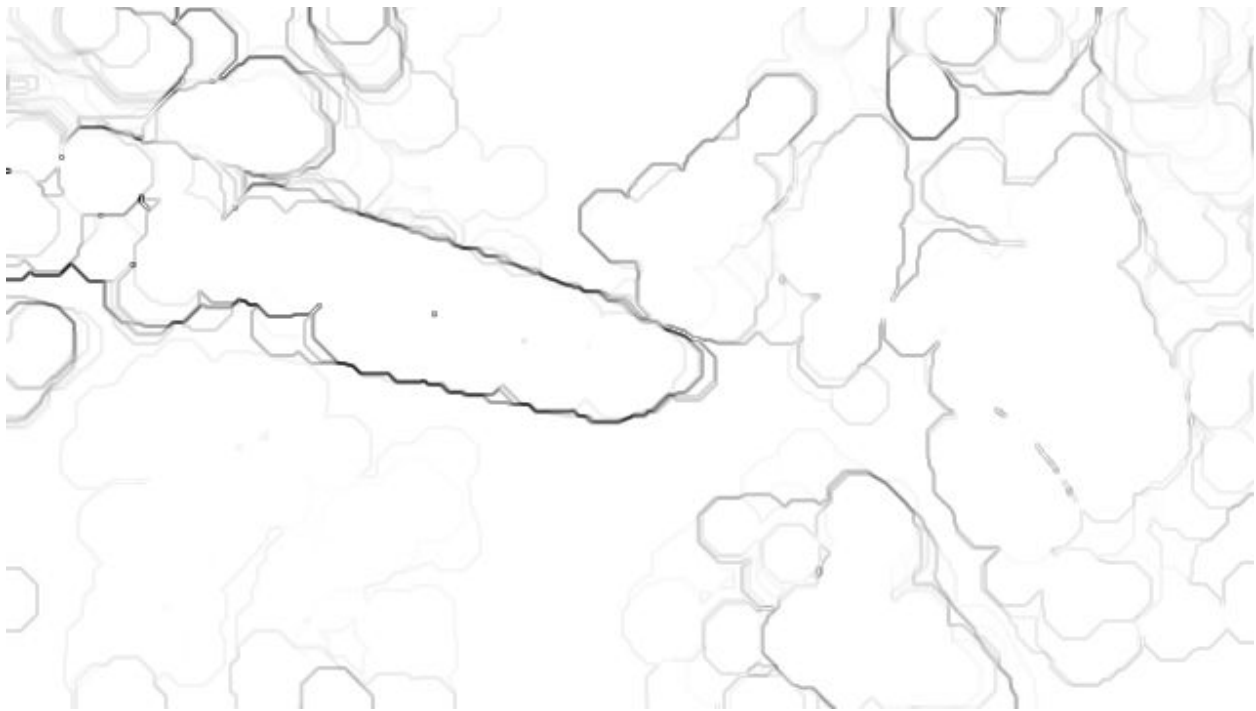
*Figura 15: Resultat d'aplicar Closing sobre la imatge*



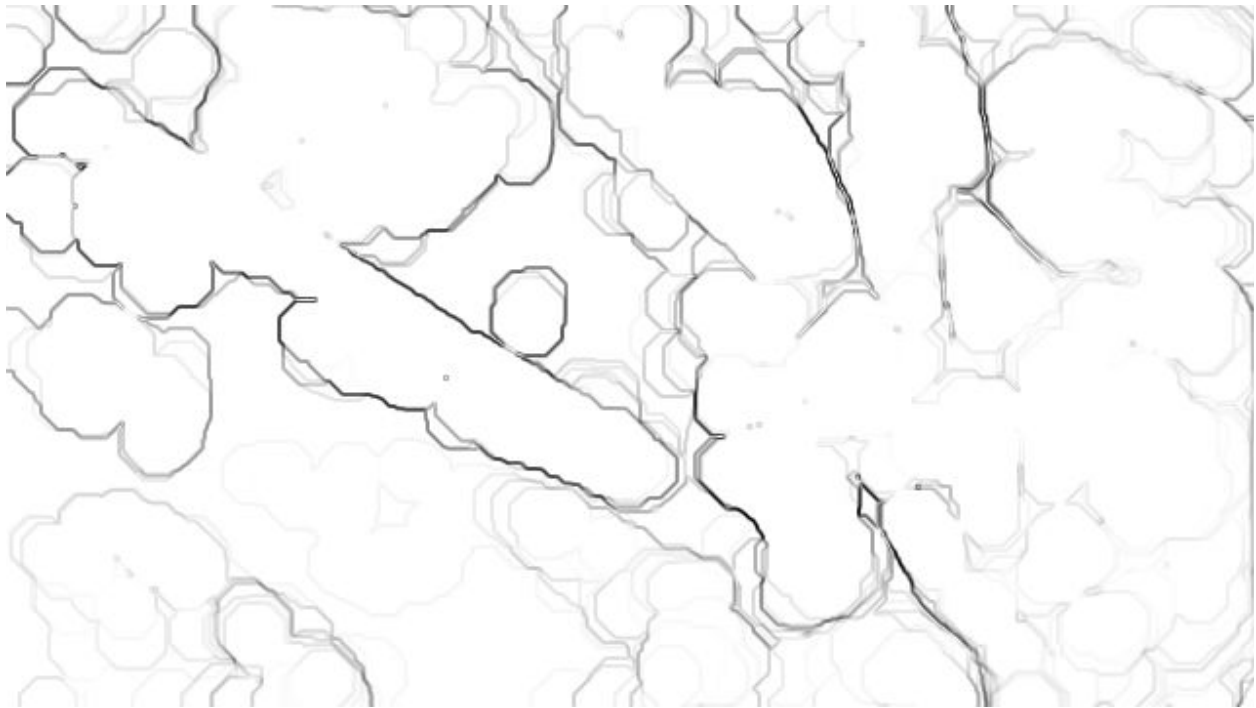
*Figura 16: Resultat d'aplicar Closing sobre la imatge*

Amb aquest preprocessament i el filtratge de la imatge, el funcionament del qual s'explicarà més endavant en aquesta mateixa secció, es van començar a obtenir les primeres deteccions d'eina, però la qualitat d'aquesta seguia essent molt pobre, per tant, es va decidir tenir en compte el mòdul dels gradients de la imatge alhora de preprocessar-la.

Calcular el mòdul del gradient d'una imatge significa assignar a cada píxel d'aquesta la pendent que hi ha entre el seva intensitat i la dels seus veïns. Partint de la idea que aquest càlcul assignaria als píxels pertanyents al contorn de l'eina un valor molt més alt que als pertanyents al seu interior, es va decidir invertir la imatge resultant per tal d'aconseguir que al binaritzar per a un determinat llindar, els valors pertanyents a l'interior de l'eina fossin d'un valor més alt que no pas els pertanyents al contorn de l'eina, de manera que amb el llindar adequat s'obtidria una única component connexa de l'eina que contindria tots els seus píxels interiors.



*Figura 17: Càlcul de gradients sobre la imatge*



*Figura 18: Càlcul de gradients sobre la imatge*

Malauradament, la il·luminació de la càmera i les reflexions que provoca, produeixen en alguns casos una gran distorsió dels gradients de la imatge, impossibilitant així la segmentació de l'eina.

Així doncs, es va veure que seria desitjable tenir en compte informació més relativa a la puresa del color que no pas a la intensitat d'aquest. Per tal d'aconseguir-ho, es va descartar la conversió de la imatge original a escala de grisos i es va optar per transformar la representació del color de RGB<sup>14</sup> a HSV<sup>15</sup>. Un cop transformat, es va pendre com a imatge amb la que treballar les operacions morfològiques i els gradients la component de saturació de la imatge original, la qual determina la puresa del color i no té en compte ni quin color és, ni com de brillant és aquest color.<sup>[7]</sup>

Posteriorment, es van analitzar els valors dels píxels pertanyents a l'eina en aquesta imatge i se li va aplicar un llindar per tal de descartar d'entrada diverses regions de la imatge on era improbable que s'ubiqués l'eina.

---

<sup>14</sup> Red, Green, Blue.

<sup>15</sup> Hue, Saturation, Value.

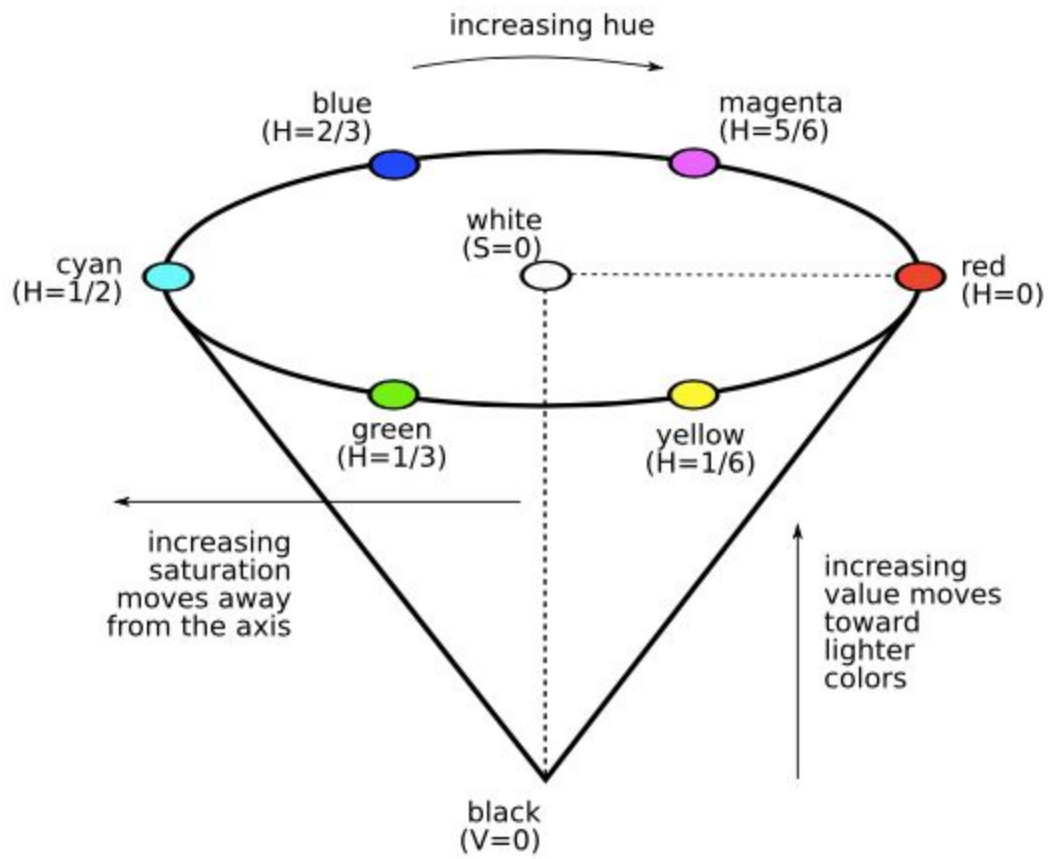
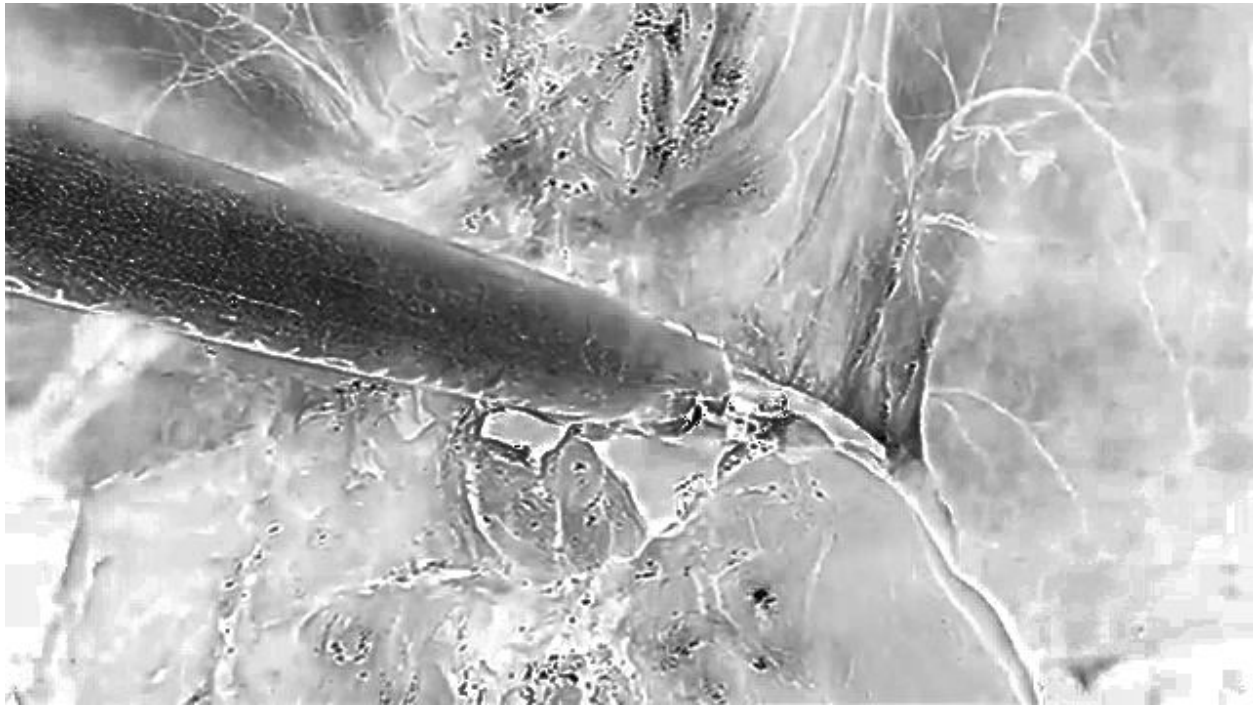


Figura 19: Representació del color en HSV<sup>[7]</sup>





*Figura 20: Component de saturació de la imatge*

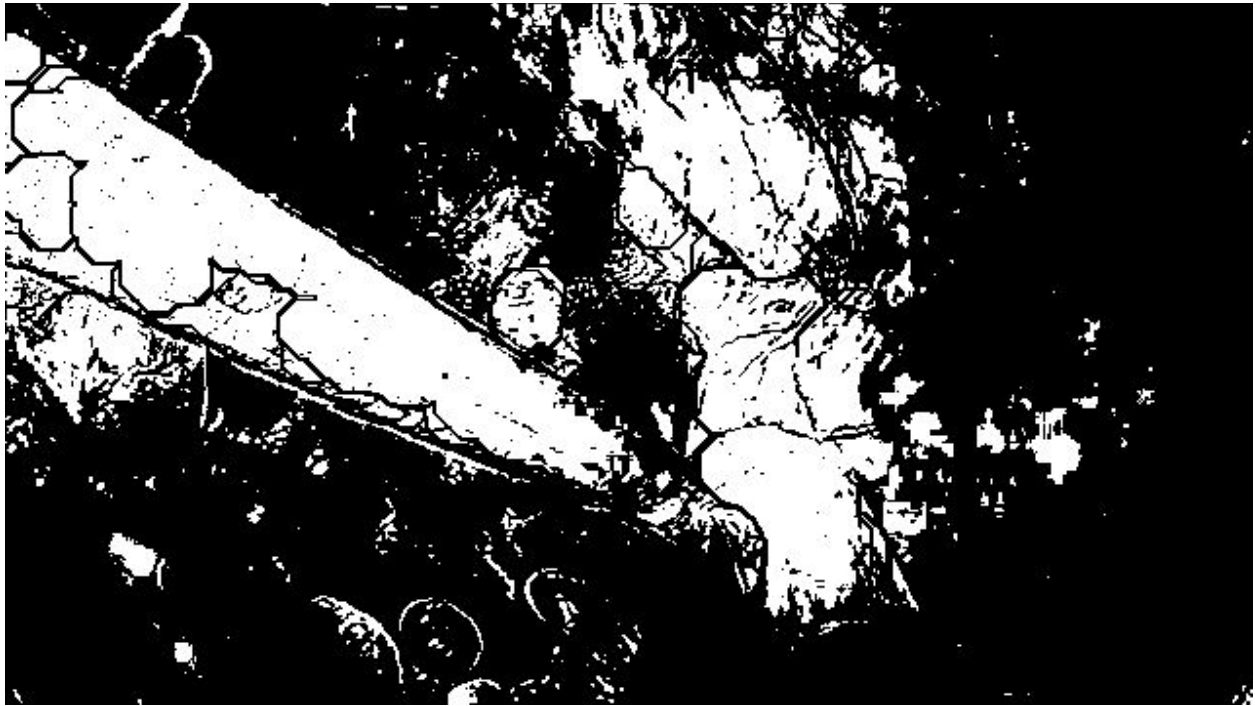


*Figura 21: Component de saturació de la imatge*

Un cop preprocessada la imatge, és necessari binaritzar-la de manera adequada i extreure les components connexes candidates per tal de que el model les classifiqui. Per fer-ho, s'ha definit un conjunt de llindars que cobreixen tot l'espectre de valors de la imatge preprocessada i que van des de l'1% dels valors més alts fins al 90%, centrant-se en la forquilla que va des de l'1% de valors més alts fins al 10%, on hi ha més possibilitats de trobar la millor segmentació de l'eina. Després, per cadascun d'aquests llindars, s'han seleccionat les components connexes de la imatge binaritzada i s'han enviat al model per tal de que aquest les classifiqui.



*Figura 22: Imatge preprocessada i filtrada per un llindar*



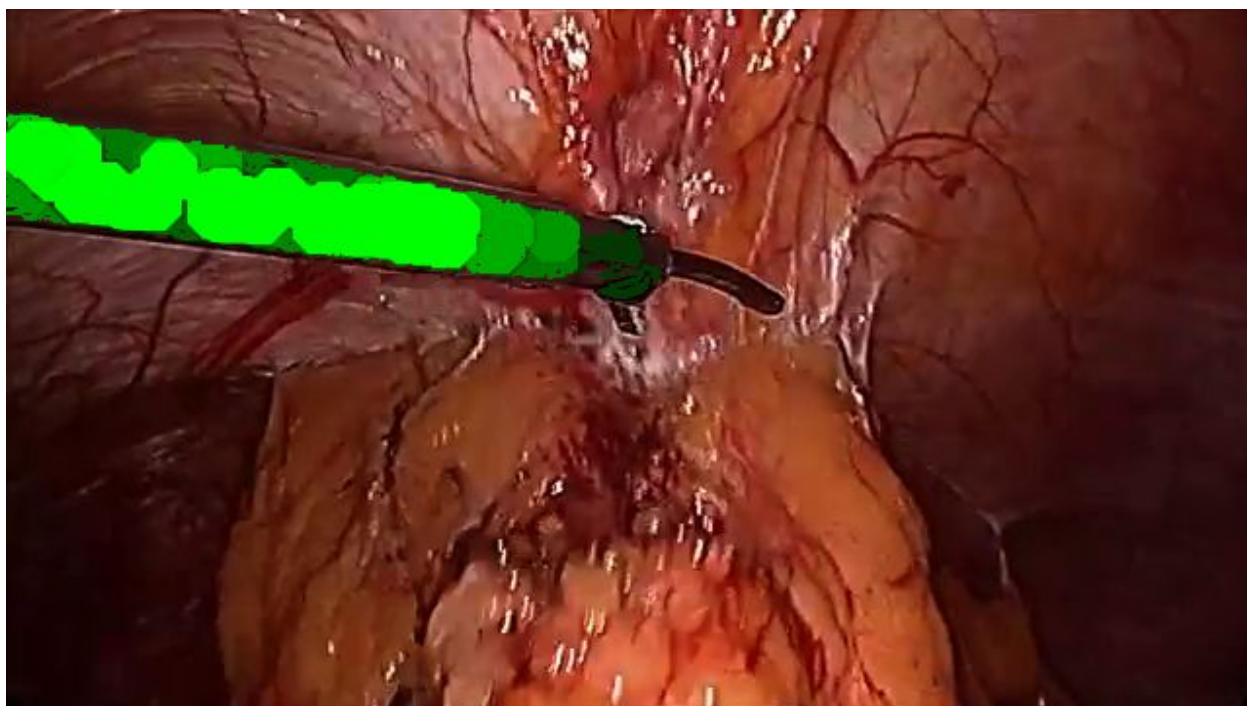
*Figura 23: Imatge preprocessada i filtrada per un llindar*

## 6.5 OBTENCIÓ DEL RESULTAT

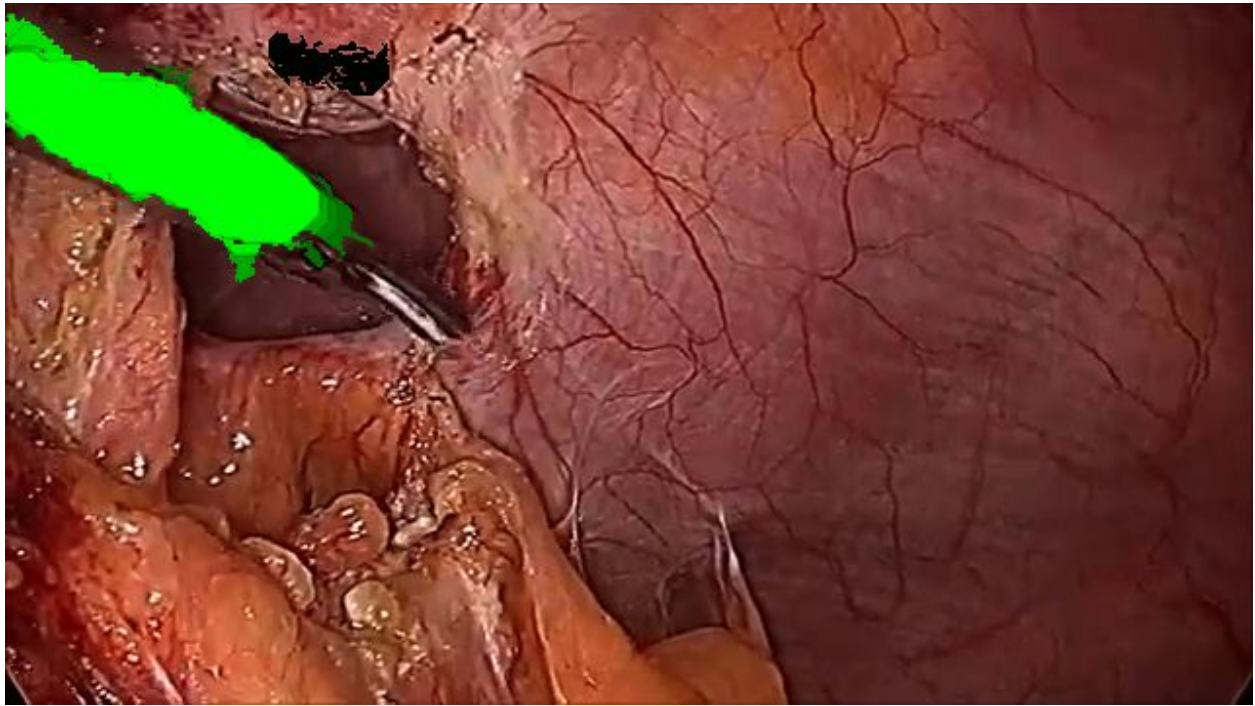
A continuació, es detalla l'algoritme complet per a l'obtenció del resultat:

- Per cada imatge de vídeo on detectar l'instrumental mèdic:
  - ◆ Conversió a escala de grisos de la imatge.
  - ◆ Preprocessament de la imatge mitjançant operadors morfològics, gradients i color.
  - ◆ Definició d'imatge resultat amb les mateixes mides que la imatge preprocessada, però amb tots els valors a 0.
  - ◆ Filtrat de la imatge preprocessada mitjançant la binarització d'aquesta en diversos llindars. Per a cadascuna de les imatges resultants de binaritzar per un determinat llindar :
    - Eliminar les components connexes amb àrees inferiors a un cert llindar.

- Definició d'imatge resultat provisional amb les mateixes mides que la imatge preprocessada, però amb tots els valors a 0.
- Per a cada una de les components connexes de la imatge:
  - Càlcul de moments invariants de la imatge de la component connexa.
  - Classificació de la component connexa:
    - ◆ Si el model la classifica com a classe “eina”:
      - Sumar la imatge de la component connexa ponderada per la puntuació obtinguda en el model i la imatge resultat provisional.
  - Sumar la imatge resultat i la imatge resultat provisional.
- ◆ Marcar el resultat en la imatge basant-se en la imatge resultat obtinguda. Aquest pas s'explica detalladament en la següent secció d'aquest mateix capítol.



*Figura 24: Imatge resultat ponderada amb les puntuacions retornades pel model*



*Figura 25: Imatge resultat ponderada amb les puntuacions retornades pel model*

## **6.6 MARCACIÓ DEL RESULTAT EN LA IMATGE I GENERACIÓ DE VÍDEO**

Un cop obtinguda la imatge ponderada de resultats, per a marcar l'eina en la imatge s'han seguit el següent passos:

En primer lloc, es binaritza la imatge ponderada per a un determinat llindar i s'aplica un *Closing* sobre la imatge per tal de juntar les petites parts de l'eina que hagin pogut quedar separades i eliminar forats interiors.

Seguidament, s'eliminen les components connexes d'aquesta que tinguin una àrea menor a un determinat llindar, amb l'objectiu de netejar els falsos positius que pugui haver-hi.

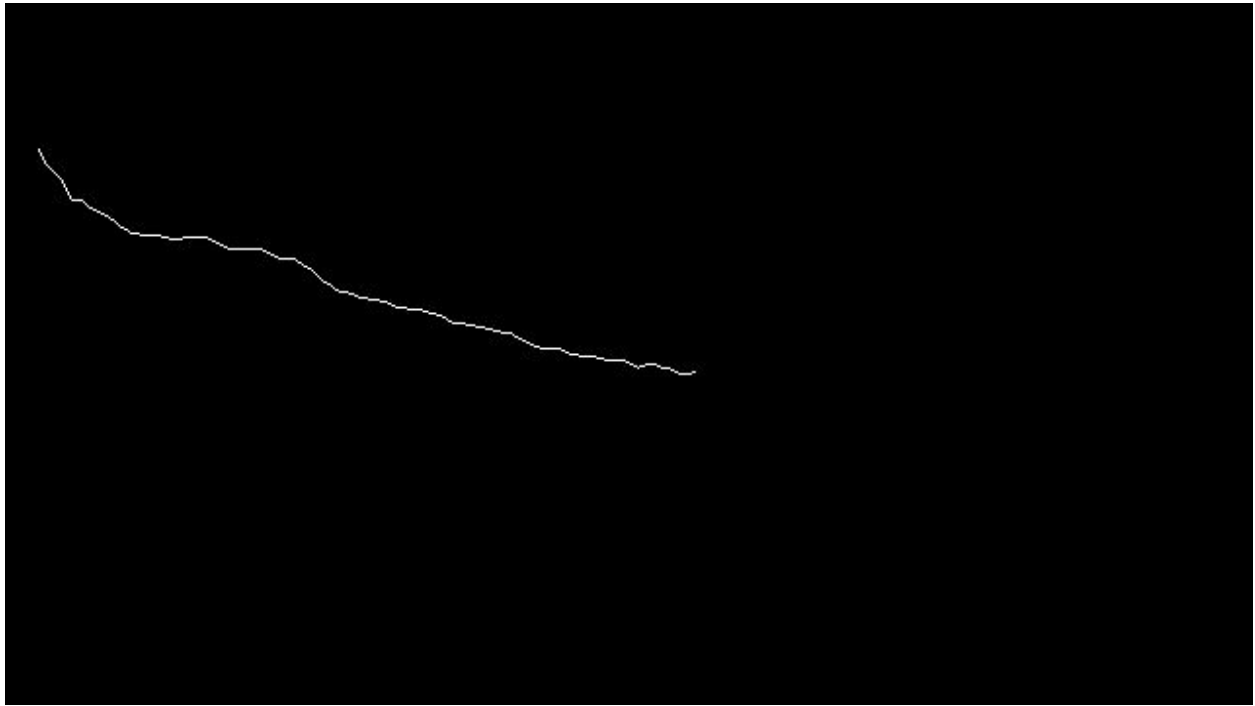
A continuació, mitjançant operacions morfològiques, es redueix el resultat a una línia i s'aplica una dilació sobre aquesta, a fi que la marcació englobi tota l'eina, s'eliminen els píxels interiors, deixant únicament el contorn, al qual se li aplica una dilació amb un

element estructurant petit, per tal de facilitar la seva visualització.

Finalment, es guarda la imatge marcada en el sistema d'arxius. Un cop totes les imatges han estat guardades, aquestes són utilitzades per a la generació del vídeo resultat.



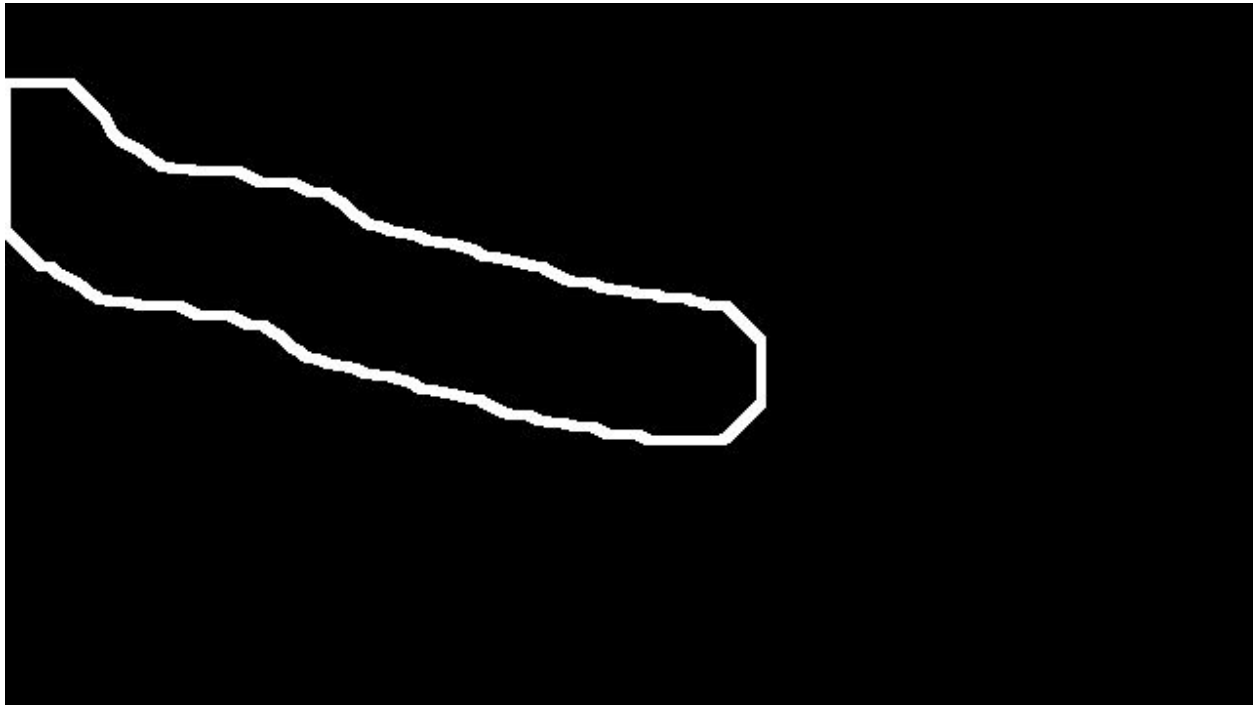
*Figura 26: Aplicació de Closing sobre la imatge resultat*



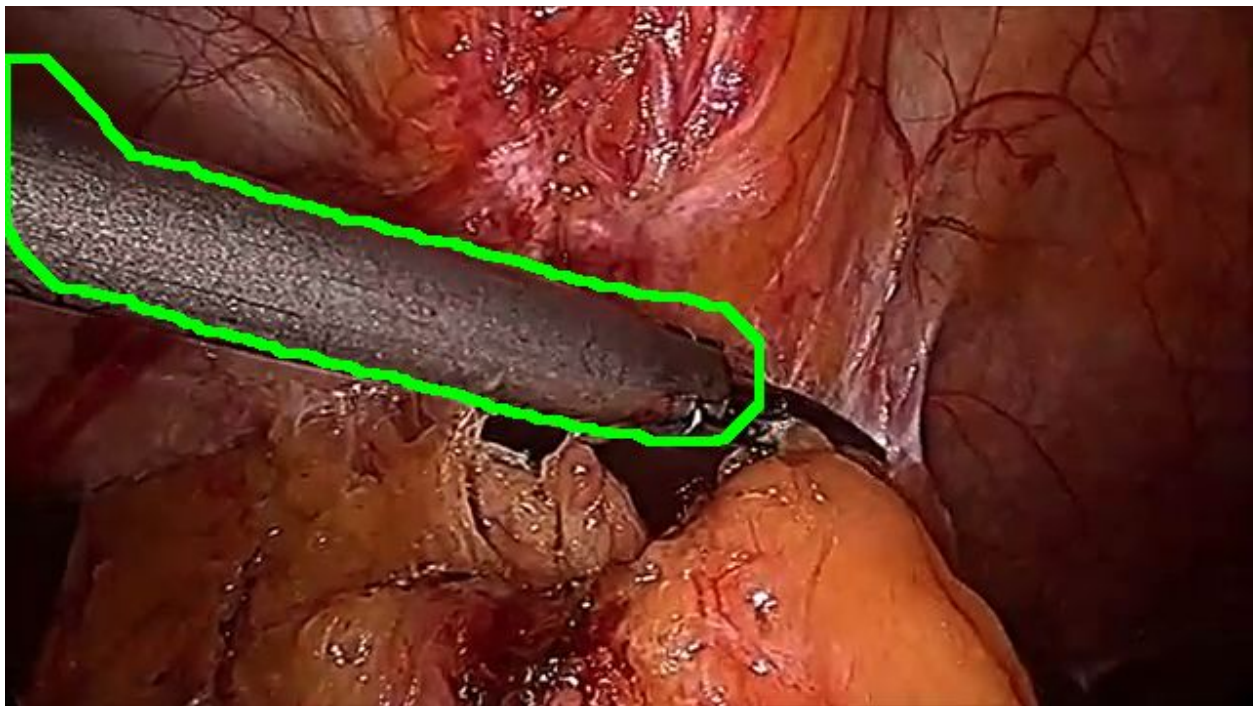
*Figura 27: Reducció del resultat a una línia*



*Figura 28: Aplicació de dilació sobre la línia obtinguda*



*Figura 29: Eliminació de píxels interiors*



*Figura 30: Imatge definitiva*



## 6.7 ANÀLISI DE RESULTATS

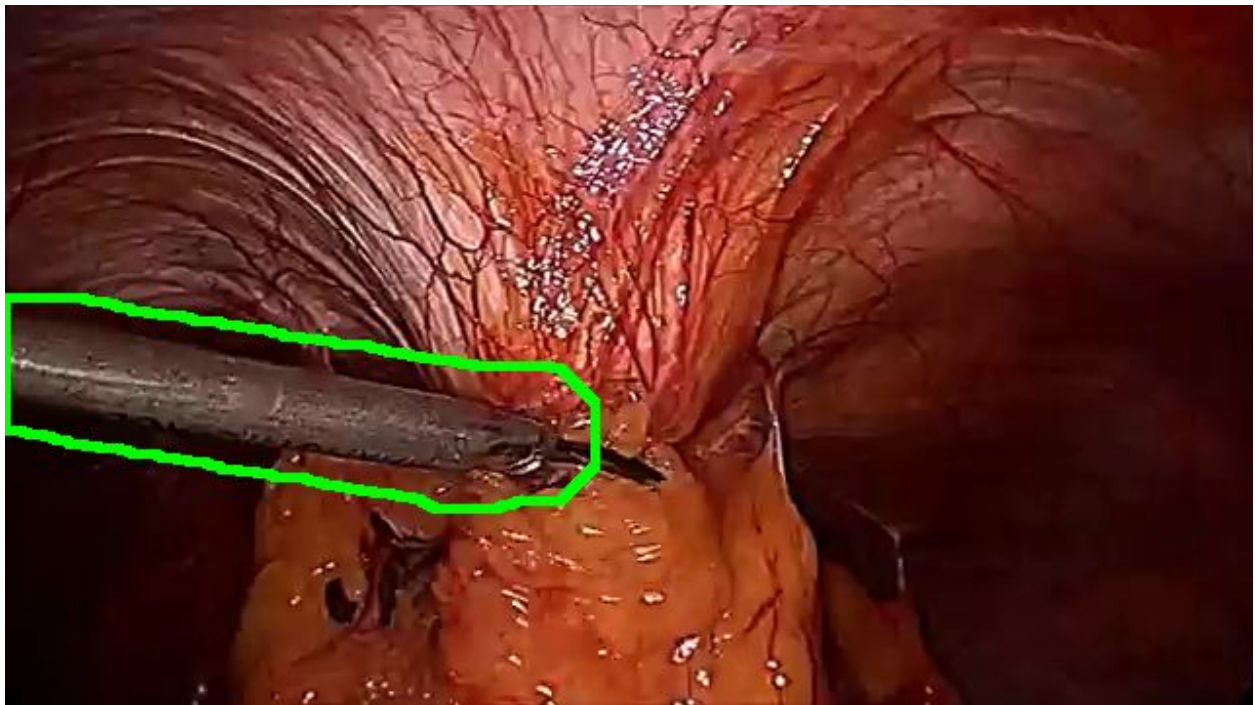
A continuació, es presenta la taula de resultats obtinguts en les diferents proves realitzades. Les columnes “Píxels en comú” i “Relació d'àrea” fan referència a la manera en que s'ha generat el *data-set*. La columna “Àrea mínima” fa referència al llindar utilitzat per eliminar components connexes menors a una certa àrea i la columna “Ponderació mínima” fa referència al llindar amb que s'ha binaritzat la imatge resultat ponderada pel model. En la columna “Preprocessat” s'hi indiquen les tècniques que s'han utilitzat per segmentar la imatge i a la columna “Resultats” es mostren els resultats obtinguts per cada versió de l'algoritme per a 100 imatges escollides aleatòriament.

Versió	Llindars					Preprocessat			Resultats			
	Píxels en comú		Relació d'àrea		Àrea mínima	Ponderació mínima	Color	Operacions morfològiques	Gradients	Correcte	Falsos positius	Falsos negatius
	P	N	P	N								
1	>75%	<60%	25-120%	-	0	0%	No	Open	Si	48%	91%	9%
2	>75%	<60%	25-120%	-	500	0%	No	Open	Si	48%	63%	37%
3	>75%	<60%	25-120%	-	1000	0%	No	Open	Si	46%	51%	49%
4	>75%	<60%	25-120%	-	2000	0%	No	Open	Si	39%	44%	56%
5	>85%	<20%	25-120%	-	1000	0%	Si	Close	Si	72%	24%	25%
6	>85%	<50%	25-120%	-	1000	0%	Si	Close	Si	67%	20%	30%
7	>75%	<50%	25-120%	-	1000	0%	Si	Close	Si	69%	22%	32%
8	>85%	<60%	25-120%	-	1000	0%	Si	Close	Si	75%	21%	23%
9	>85%	<70%	25-120%	-	1000	0%	Si	Close	Si	65%	19%	33%
10	>85%	<60%	25-120%	-	1000	5%	Si	Close	Si	73%	15%	23%
11	>85%	<60%	25-120%	-	1000	25%	Si	Close	Si	73%	13%	23%
12	>85%	<60%	25-120%	-	1000	50%	Si	Close	Si	73%	13%	23%

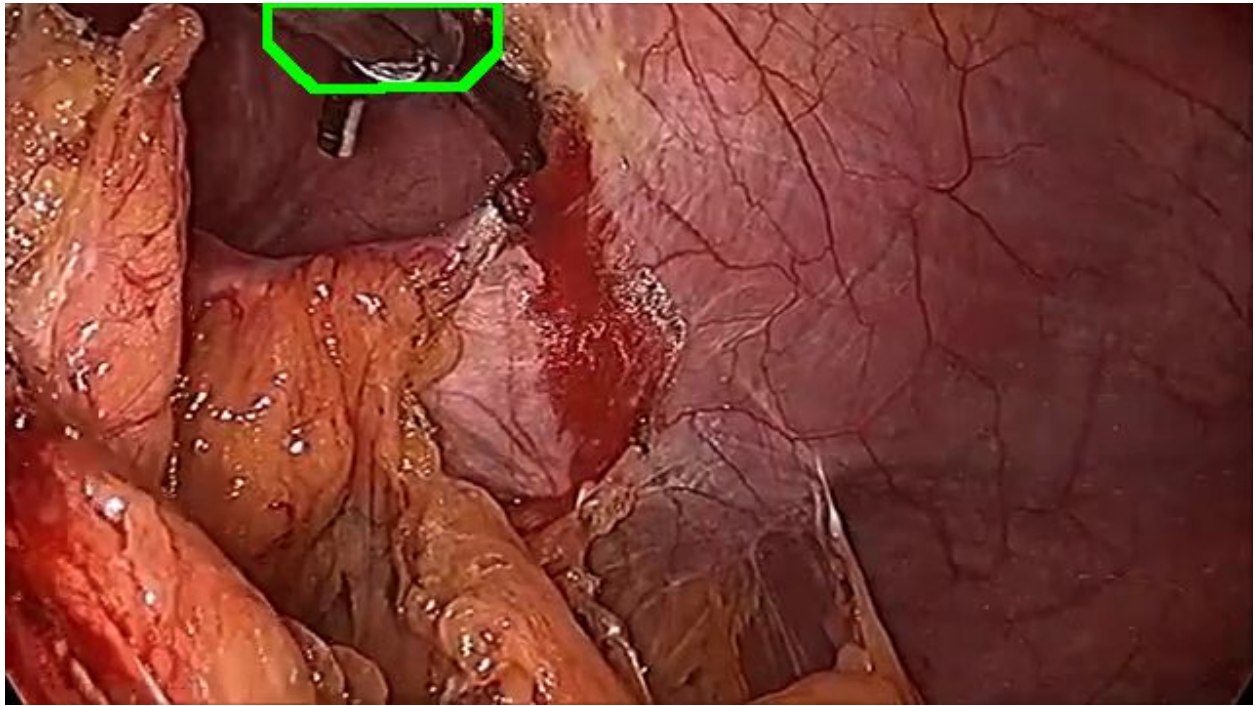
Taula 10: Presentació de resultats obtinguts

La taula mostra clarament el procediment que s'ha seguit en el desenvolupament d'aquest projecte. Es va prioritzar tenir una primera implementació completa que dongués resultats, llavors, mica en mica, s'ha anat afegint valor a la solució a mesura que es desenvolupaven més versions.

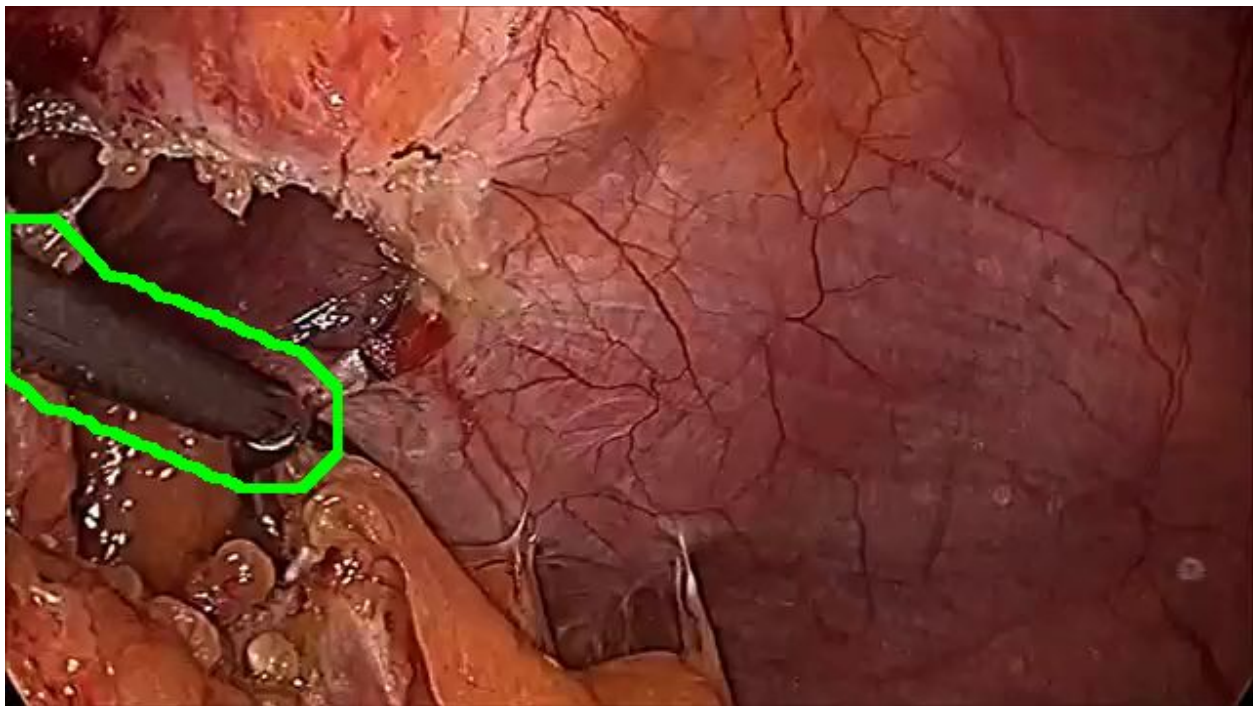
Aquestes millores, les podríem dividir en dos períodes, el primer, en que es buscava augmentar el percentatge de deteccions correctes, i un segon període, on es prioritzava la reducció dels falsos positius, ja que es considera que és preferible perdre percentatge de detecció, però que quan es detecti, aquesta marcació sigui correcta. Així doncs, finalment, la versió definitiva de l'algoritme ha estat la número 12.



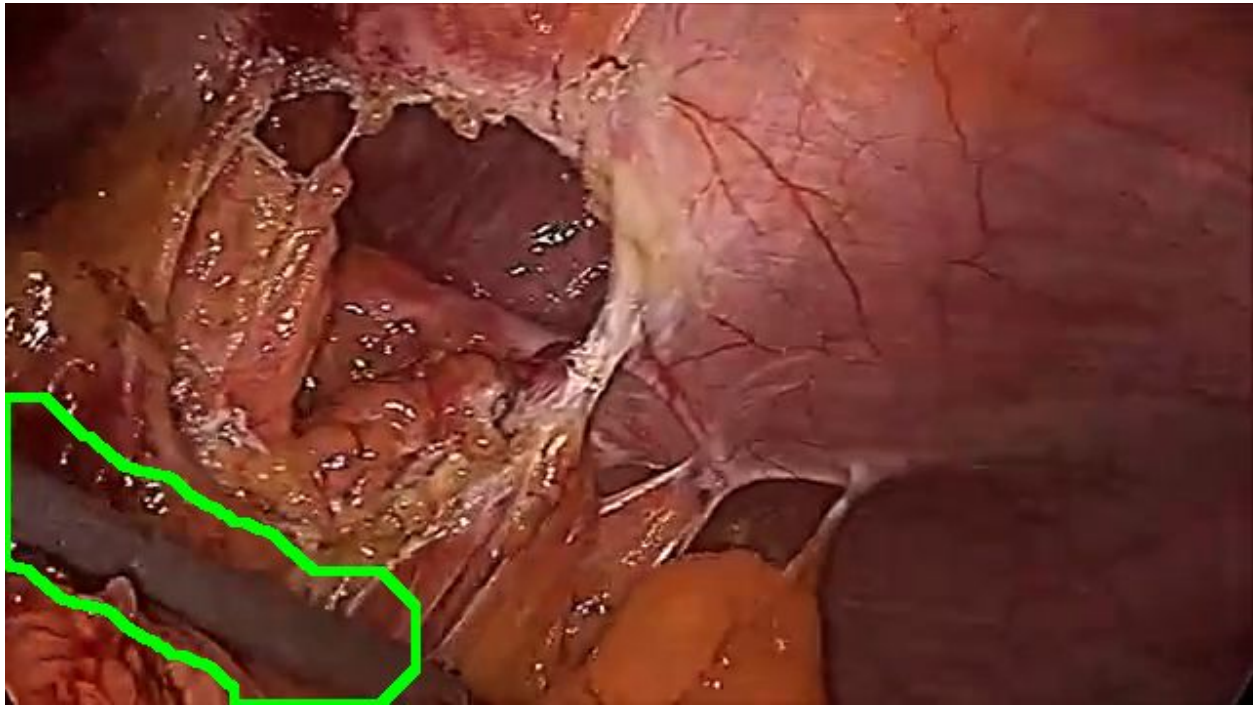
*Figura 31: Detecció correcte obtinguda per l'última versió de l'algoritme*



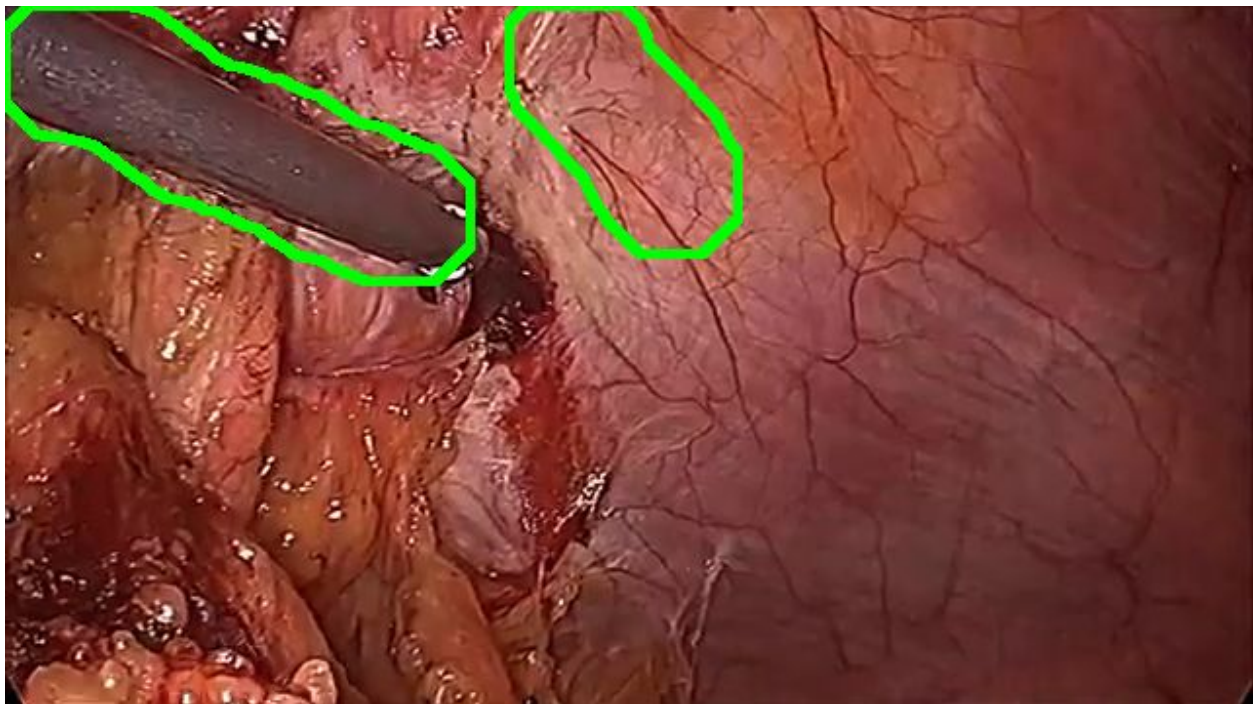
*Figura 32: Detecció correcte obtinguda per l'última versió de l'algoritme*



*Figura 33: Detecció correcte obtinguda per l'última versió de l'algoritme*



*Figura 34: Detecció correcte obtinguda per l'última versió de l'algoritme*



*Figura 35: Detecció correcte i un fals positiu obtinguts per l'última versió de l'algoritme*

# Capítol 7

## Conclusions

### 7.1 CONCLUSIONS GENERALS

Amb el desenvolupament d'aquest projecte, s'ha demostrat que és possible la detecció de l'instrumental mèdic utilitzat en cirurgies mínimament invasives. Basant-se en característiques de forma, utilitzant tècniques sense un gran requeriment computacional, habilitant així que aquesta detecció pugui ser computada en temps real, i sense necessitat de cap modificació ni de la sala d'operacions ni de l'eina a reconèixer.

Concretament, el projecte s'ha desenvolupat per reconèixer l'instrumental utilitzat en una laparoscòpia per a la reparació d'hèrnies, però la solució desenvolupada és fàcilment exportable a d'altres tipus de cirurgies.

També s'han pogut constatar les limitacions d'utilitzar característiques de forma per al reconeixement de l'eina. Ja que, per exemple, el fet de que l'eina acabi en forma de pinça que es va obrint i tancant, depenent del moment del vídeo en que es vol reconèixer l'eina, dificulta molt el reconeixement total d'aquesta. Així doncs, es pot aspirar a reconèixer diversos tipus d'eina però limitant-se únicament al tronc, que és la part comuna i invariable de totes.

El desenvolupament del projecte ha estat més que profitós per adonar-se del que realment significa implementar una solució robusta a un problema de certa complexitat. Així com, per ampliar coneixements en l'àmbit de la computació i la visió per computador.

## 7.2 FUTURES IMPLEMENTACIONS

A continuació, s'expliquen algunes futures implementacions de la solució que donarien aquesta d'un major valor:

- Utilitzar la detecció del *frame* anterior del vídeo per millorar la segmentació i detecció en el *frame* actual.
- Paral·lelitzar l'algoritme, fent especial èmfasi en el filtratge d'imatge i el processament de diversos *frames* en paral·lel. Reduint així el temps de computació d'aquest i assolint la detecció de l'instrumental mèdic en temps real.
- Pel que fa a l'extracció de característiques, s'ha seleccionat únicament els moments invariants de Hu com a característiques discriminatòries de l'eina. Es deixa per a un futur, la implementació del càlcul de moments de Zernike i les proves de selecció d'aquests, les quals s'han descartat per limitacions temporals.
- Insistir més en la optimalitat dels paràmetres que defineixen el *Kernel* del model utilitzat i, en un futur, plantejar-se el desenvolupament d'un *Kernel* propi.

# Capítol 8

## Referències

### 8.1 REFERÈNCIES

1. Sznitman, R., Richa, R., Taylor, R. H., Jedynek, B., & Hager, G. D. (2013). Unified detection and tracking of instruments during retinal microsurgery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(5), 1263–1273. <http://doi.org/10.1109/TPAMI.2012.209>
2. Sznitman, R., Basu, A., Richa, R., Handa, J., Gehlbach, P., Taylor, R. H., ... Hager, G. D. (2011). Unified detection and tracking in retinal microsurgery. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6891 LNCS(PART 1), 1–8. [http://doi.org/10.1007/978-3-642-23623-5\\_1](http://doi.org/10.1007/978-3-642-23623-5_1)
3. Kumar, S., Narayanan, M. S., Singhal, P., Corso, J. J., & Krovi, V. (2013). Product of tracking experts for visual tracking of surgical tools. *IEEE International Conference on Automation Science and Engineering*, 480–485. <http://doi.org/10.1109/CoASE.2013.6654037>
4. Dogangil, G., Davies, B. L., & Rodriguez y Baena, F. (2010). A review of medical robotics for minimally invasive soft tissue surgery. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(5), 653–679. <http://doi.org/10.1243/09544119JEIM591>
5. Van Der Meijden, O. A. J., & Schijven, M. P. (2009). The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: A current review. *Surgical Endoscopy and Other Interventional Techniques*, 23(6), 1180–1190. <http://doi.org/10.1007/s00464-008-0298-x>
6. Darzi, A., & Mackay, S. (2002). Clinical review Recent advances in minimal access surgery. *British Medical Journal*, 324(7328), 31–34. <http://doi.org/10.1136/bmj.324.7328.31>
7. Muller, T. (1986). Introduction to color theory. *The Indian Textile Journal*, 96,

- 78–79. <http://doi.org/10.1119/1.1991151>
8. Bouget, D., Allan, M., Stoyanov, D., & Jannin, P. (2017). Vision-based and marker-less surgical tool detection and tracking: a review of the literature. *Medical Image Analysis*, 35, 633–654. <http://doi.org/10.1016/j.media.2016.09.003>
  9. Zekovich, S., & Tuba, M. (n.d.). Hu Moments Based Handwritten Digits Recognition Algorithm, (44006), 98–103. Retrieved from <http://www.wseas.us/e-library/conferences/2013/CambridgeUK/AISE/AISE-15.pdf>
  10. Chih-Wei Hsu, Chih-Chung Chang, and C.-J. L. (2008). A Practical Guide to Support Vector Classification. *BJU International*, 101(1), 1396–400. <http://doi.org/10.1177/02632760022050997>
  11. Pudil, P., Novovieova, J., & Kittler, J. (1994). Pattern Recognition Letters. *Pattern Recognition Letters*, 15(June 1993), 1119–1125. <http://doi.org/10.1016/j.patrec.2010.09.010>
  12. Reiter, A., Allen, P. K., & Zhao, T. (2012). Articulated Surgical Tool Detection Using Virtually-Rendered Templates. *Computer Assisted Radiology and Surgery (CARS)*, 1–8.
  13. Bouget, D., Benenson, R., Omran, M., Riffaud, L., Schiele, B., & Jannin, P. (2015). Detecting Surgical Tools by Modelling Local Appearance and Global Shape. *IEEE Transactions on Medical Imaging*, 34(12), 2603–2617. <http://doi.org/10.1109/TMI.2015.2450831>
  14. Withey, D. J., & Koles, Z. J. (2007). Medical Image Segmentation: Methods and Software. *2007 Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging*, 140–143. <http://doi.org/10.1109/NFSI-ICFBI.2007.4387709>
  15. Li, D. (2010). Analysis of moment invariants on image scaling and rotation. *Innovations in Computing Sciences and Software Engineering*, (2010), 415–419. <http://doi.org/10.1007/978-90-481-9112-3-70>
  16. Thompson, M., Gonzalez, R. C. R., Wintz, P., Woods, R. E. R., & Masters, B. R. (2002). *Digital image processing. Leonardo* (Vol. 14). <https://doi.org/10.2307/1574313>
  17. Sedaaghi, M. H. (2002). Morphological operators. *Electronics Letters*, 38(22), 1333. <https://doi.org/10.1049/el:20020943>