



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola d'Enginyeria de Barcelona Est

TREBALL FI DE GRAU

Grau en Enginyeria Biomèdica

**CLASSIFICACIÓ AUTOMÀTICA D'IMATGES DIGITALS DE
FORMES ERITROCITÀRIES ANÒMALES EN SANG PERIFÈRICA**



Memòria i Annex

Autor: Núria Torrens Garriga
Director: Santiago Alférez Baquero
Co-Director: José Rodellar Benede
Ponent: Anna Merino González
Convocatòria: Gener 2017

Resum

L'anèmia hemolítica és una condició en la qual l'esperança de vida dels glòbuls vermells es veu disminuïda i, com a conseqüència, la massa d'aquests en sang és inferior a l'esperada. La característica principal d'aquesta alteració és la producció d'eritròcits amb una morfologia anòmala.

Es pretén crear una eina automàtica de detecció i classificació dels tipus de glòbuls vermells que permeti diagnosticar el perfil d'anèmia a partir de l'anàlisi dels frotis sanguinis. Aquest projecte es centra a identificar l'esferòcit i el drepanòcit vers el glòbul vermell sa. L'esferòcit es caracteritza per tenir una forma esferoïdal i estar mancat de la zona de pal·lidesa central de l'eritròcit normal. El drepanòcit, en canvi, es disposa en una forma semilunar.

S'obtenen 600 imatges d'hematies de 30 frotis de sang perifèrica procedents de 17 pacients diferents. Es processen les imatges i els eritròcits que les formen es sotmeten a una segmentació. Per a dur a terme aquest procés s'utilitza el programa MATLAB. Es determinen una sèrie de descriptors (de mida, forma i color) necessaris per a la identificació de les diverses variacions dels eritròcits i es classifiquen els descriptors segons la importància que tenen en la diferenciació d'esferòcits i drepanòcits.

Finalment, es realitza una anàlisi estadística dels diferents descriptors per a la comparació dels tres tipus d'eritròcits analitzats. Segons els resultats obtinguts, es pot afirmar que els algoritmes creats permeten diferenciar els esferòcits i els drepanòcits dels glòbuls vermells normals.

procés

Paraules clau:

Eritròcits, esferòcits, drepanòcits, citologia, anàlisi digital d'imatges.

Resumen

La anemia hemolítica es una condición en la que la esperanza de vida de los glóbulos rojos se ve disminuida y, como consecuencia, su masa en sangre adquiere un valor inferior al esperado. La principal característica de esta alteración es la producción de eritrocitos con una morfología anómala.

Se pretende crear una herramienta automática de detección y clasificación de los tipos de glóbulos rojos, la cual permita diagnosticar el perfil de anemia a partir de los frotis sanguíneos. Este proyecto se centra en identificar el esferocito y el drepanocito del glóbulo rojo sano. El esferocito se caracteriza por tener una forma esferoidal y carecer de la zona de palidez central del eritrocito normal. El drepanocito, en cambio, se dispone en una forma semilunar.

Se obtienen 600 imágenes de hematíes de 30 frotis de sangre periférica procedentes de 17 pacientes diferentes. Se procesan las imágenes y los eritrocitos que las forman se someten a una segmentación. Para llevar a cabo este proceso se utiliza el programa MATLAB. Se determinan una serie de descriptores (de tamaño, forma y color) necesarios para la identificación de las diversas variaciones de los eritrocitos y se clasifican los descriptores según su importancia en la diferenciación de esferocitos y drepanocitos.

Finalmente, se realiza un análisis estadístico de los diferentes descriptores para la comparación de los tres tipos de eritrocitos analizados. Según los resultados obtenidos, se puede afirmar que los algoritmos creados permiten diferenciar los esferocitos y los drepanocitos de los glóbulos rojos normales.

Palabras clave:

Eritrocito, esferocito, drepanocito, citología, análisis digital de imágenes.

Abstract

Hemolytic anemia is a condition in which red blood cells' life expectancy is decreased and as a consequence the mass of these in blood acquires a value lower than expected. The main feature of this alteration is the production of erythrocytes with abnormal morphology.

The aim is to create an automatic tool for detection and classification of the types of red blood cells, which allows diagnosing anemia from blood smears. This project focuses on identifying spherocytes and sickle cells from the healthy red blood cells. Spherocytes are characterized by having a spherical shape and a lack of the central pallor. Sickle cells use to have a half-moon shape.

A total of 600 images of erythrocytes from 30 peripheral blood smears from 17 different patients are obtained. Images are processed and subjected to a segmentation process. MATLAB program is used to carry out these processes. It is needed to have some descriptors determined (size, shape and colour) for the identification of different types of erythrocytes. These descriptors are classified according to their importance in the differentiation of spherocytes and sickle cells.

A statistical analysis of the different descriptors is conducted in order to compare the three types of erythrocytes mentioned. According to the results obtained, it can be said that the algorithm allows differentiating the spherocytes and the sickle cells from the normal red blood cells.

Key words:

Erythrocyte, spherocyte, sickle cell, cytology, digital image analysis

Agraïments

En primer lloc m'agradaria agrair l'ajuda i suport del meu tutor, Santiago Alférez, durant el transcurs de tot el projecte. Gràcies als seus amplis coneixements en programació i, concretament, en l'anàlisi digital d'imatges, m'he pogut enriquir conceptual i pràcticament en aquest àmbit.

També m'agradaria agrair al laboratoria CORE de l'Hospital Clínic de Barcelona la facilitat donada en l'obtenció de frotis sanguinis per a l'adquisició de les imatges. En concret, agraeixo al citòleg Àngel Molina el seu treball en la supervisió de la correcta extracció de les imatges i del procés d'etiquetatge.

Agraeixo també a la meva família l'ajuda indirecta i suport donat, i a totes aquelles persones que han posat un gra de sorra durant el transcurs d'aquest projecte.

Glossari

Hb	Hemoglobina
AH	Anèmia hemolítica
EH	Esferocitosi hereditària
BW	Imatge binària
WT	Watershed Transformation, Transformada Watershed
OI	Objecte d'interès
ROI	Region of interest, regió d'interès
CH	Convex Hull, Casc convex
Std	Desviació estàndard



Índex

RESUM	I
RESUMEN	II
ABSTRACT	III
AGRAÏMENTS	IV
GLOSSARI	V
ÍNDIX DE FIGURES	X
ÍNDIX DE TAULES	XIV
CAPÍTOL 1: PREFACI	1
1.1 Origen del treball	1
1.2 Motivació	1
1.3 Requeriments previs	2
1.4 Diagrama de Gantt	2
CAPÍTOL 2: INTRODUCCIÓ	3
2.1 Objectius del treball	4
CAPÍTOL 3: LA SANG	5
3.1 Eritròcits	7
3.2 Leucòcits	8
3.3 Les plaquetes	9
3.4 El plasma sanguini	9
CAPÍTOL 4: L'ANÈMIA	11
4.1 Esferocitosi	13
4.2 Drepanocitosi	14
4.3 L'hemoglobina	15
CAPÍTOL 5: ADQUISICIÓ DE LES IMATGES	17
5.1 Material utilitzat	17
5.2 Procediment d'adquisició	17
5.2.1 Preparació de la mostra	18
5.2.2 Col·locació de la mostra al microscopi	18

5.2.3	Captura de la imatge.....	19
CAPÍTOL 6: PROCESSAT DE LES IMATGES _____		21
6.1	Lectura de la imatge	22
6.2	Preprocessat i segmentació.....	22
6.2.1	Segmentació de les plaquetes	22
6.2.2	Preprocessament	23
6.2.3	Segmentació dels eritròcits d'interès	26
6.2.4	Filtrat dels esferòcits.....	27
6.3	Procés d'etiquetatge.....	28
6.4	Extracció de propietats morfològiques	29
6.4.1	Característiques físiques	29
6.4.2	Altres característiques	32
CAPÍTOL 7: ANÀLISI ESTADÍSTICA _____		35
7.1	Test de normalitat i d'homoscedasticitat	35
7.2	Estudi de rellevància dels descriptors	38
7.3	Anàlisi dels descriptors	39
7.3.1	Eix de major longitud	40
7.3.2	Mitjana component verda	42
7.3.3	Excentricitat	43
7.3.4	Circularitat.....	45
7.3.5	Mitjana component vermella	47
7.3.6	Eix de menor longitud.....	49
7.3.7	Asimetria estadística component blava	51
7.3.8	Variància circular.....	53
7.3.9	Perímetre	55
7.3.10	Mitjana component blava.....	56
7.4	Discussió dels resultats estadístics	58
CAPÍTOL 8: CLASSIFICACIÓ _____		61
8.1	Marc teòric.....	61
8.2	Metodologia de classificació	63
8.3	Discussió dels resultats de la classificació	65
CAPÍTOL 9: ANÀLISI DE L'IMPACTE AMBIENTAL _____		67
CAPÍTOL 10: CONCLUSIONS _____		69
10.1	Conclusions	69

10.2 Perspectives de futur	71
CAPÍTOL 11: PRESSUPOST	73
11.1 Costos de personal.....	73
11.2 Costos de material	73
11.3 Costos informàtics.....	74
11.4 Pressupost final.....	75
BIBLIOGRAFIA	77
ANNEX A	79
Estadística	79

Índex de figures

Figura 1. Diagrama de Gantt _____	2
Figura 2. Elements formats en la medul·la òssia. (Font: AEAL) _____	6
Figura 3. Procés formació glòbuls vermells. (Font: Probiomed) _____	8
Figura 4. Imatge d'un esferòcit. És la cèl·lula central de color més fosc. (Font: Imatge adquirida dels frotis cedits per l'Hospital Clínic.) _____	13
Figura 5. Imatge d'un drepanòcit. És la cèl·lula de forma allargada semilunar. (Font: Imatge adquirida dels frotis cedits per l'Hospital Clínic.) _____	14
Figura 6. Eritròcits que han perdut la seva forma natural, és una zona poc adequada per l'observació (x1000) _____	18
Figura 7. Imatge d'hematies normals juntament amb esferòcits (x1000) _____	18
Figura 8. Imatge d'un frotis de sang perifèrica mostnt els hematíes (x1000) _____	22
Figura 9. Components del color de la imatge original. Es pot veure que la component verda és la que resalta més les cèl·lules respecte el fons. _____	23
Figura 10. Comparació de l'histograma abans i després d'aplicar la funció "imadjust" _____	24
Figura 11. Component verda de la imatge original abans (esquerra) i després (dreta) de filtrar. _	24
Figura 12. Imatge binaritzada sense plaquetes _____	25
Figura 13. Eliminació objecte en contacte amb les cantonades i aplicació operador morfològic "imopen" _____	26
Figura 14. Resultat de la segmentació Watershed on s'han separat els eritròcits. _____	27
Figura 15. Transformada de la distància amb mètode "Chessboard". _____	27
Figura 16. Abans i després del filtratge segons àrea, excentricitat i circularitat. _____	28
Figura 17. Etiquetatge dels eritròcits amb l'aplicació Image Labeler _____	29

Figura 18. Gràfica d'una distribució normal _____	¡Error! Marcador no definido.
Figura 19. Gràfic de caixes de l'eix de major longitud _____	40
Figura 20. Gràfica de comparació múltiple amb Kruskal-wallis de l'eix de major longitud _____	40
Figura 21. Gràfic de caixes de la mitjana verd _____	42
Figura 22. Gràfica de comparació múltiple amb Kruskal-wallis de la mitjana verd _____	42
Figura 23. Gràfic de caixes de l'excentricitat _____	44
Figura 24. Gràfica de comparació múltiple amb Kruskal-wallis de l'excentricitat _____	44
Figura 25. Gràfic de caixes de la circularitat _____	46
Figura 26. Gràfica de comparació múltiple amb Kruskal-wallis de la circularitat _____	46
Figura 27. Gràfic de caixes de la mitjana vermell _____	47
Figura 28. Gràfica de comparació múltiple amb Kruskal-wallis de la mitjana vermell _____	48
Figura 29. Gràfic de caixes de l'eix de menor longitud _____	50
Figura 30. Gràfica de comparació múltiple amb Kruskal-wallis de l'eix de menor longitud _____	50
Figura 31. Gràfic de caixes de l'asimetria estadística blau _____	51
Figura 32. Gràfica de comparació múltiple amb Kruskal-wallis de l'asimetria estadística blau _____	52
Figura 33. Gràfic de caixes de la variància circular _____	53
Figura 34. Gràfica de comparació múltiple amb Kruskal-wallis de la variància circular _____	54
Figura 35. Gràfic de caixes del perímetre _____	55
Figura 36. Gràfica de comparació múltiple amb Kruskal-wallis del perímetre _____	55
Figura 37. Gràfic de caixes de la mitjana blau _____	57
Figura 38. Gràfica de comparació múltiple amb Kruskal-wallis de la mitjana blau _____	57

Figura 39. Hiperplans que separen un conjunt de dades. (Font:Wikipedia)	62
Figura 40. Mapeig de les dades a un espai de majors dimensions. (imatge de la dreta) (Font: Scielo [27.]	62
Figura 41. Gràfica de comparació múltiple amb Kruskal-Wallis de l'extensió	79
Figura 42. Gràfica de caixes de l'extensió	79
Figura 43. Gràfica de caixes de la rodonesa	80
Figura 44. Gràfica de comparació múltiple amb Kruskal-Wallis de la rodonesa CH	81
Figura 45. Gràfica de comparació múltiple amb Kruskal-Wallis de l'àrea	82
Figura 46. Gràfica de caixes de l'àrea	82
Figura 47. Gràfica de caixes de l'entropia de la component vermella	83
Figura 48. Gràfica de comparació múltiple amb Kruskal-Wallis de l'entropia component vermella	84
Figura 49. Gràfica de caixes de la desviació estàndard de la component vermella	85
Figura 50. Gràfica de comparació múltiple amb Kruskal-Wallis de la desviació estàndard component vermella	85
Figura 51. Gràfica de caixes de la desviació estàndard de la component verda	86
Figura 52. Gràfica de comparació múltiple amb Kruskal-Wallis de la desviació estàndard component verda	87
Figura 53. Gràfica de caixes de l'entropia de la component blava	88
Figura 54. Gràfica de comparació múltiple amb Kruskal-Wallis de l'entropia component blava	88
Figura 55. Gràfica de caixes de la curtosis de la component blava	89
Figura 56. Gràfica de comparació múltiple amb Kruskal-Wallis de la curtosis component blava	90
Figura 57. Gràfica de caixes de de l'entropia de la component verda	91
Figura 58. Gràfica de comparació múltiple amb Kruskal-Wallis de l'entropia component verda	91

Figura 59. Gràfica de comparació múltiple amb Kruskal-Wallis de la desviació estàndard component blava _____	93
Figura 60. Gràfica de caixes de de la desviació estàndard de la component blava _____	93
Figura 61. Gràfica de caixes de l'asimetria estadística de la component blava _____	94
Figura 62. Gràfica de comparació múltiple amb Kruskal-Wallis de l'asimetria estadística component vermella _____	95
Figura 63. Gràfica de caixes de l'asimetria estadística de la component verda _____	96
Figura 64. Gràfica de comparació múltiple amb Kruskal-Wallis de l'asimetria estadística component verda _____	96
Figura 65. Gràfica de caixes de la solidesa _____	97
Figura 66. Gràfica de comparació múltiple amb Kruskal-Wallis de la solidesa _____	98
Figura 67. Gràfica de caixes de la variància el·líptica _____	99
Figura 68. Gràfica de comparació múltiple amb Kruskal-Wallis de la variància el·líptica _____	99
Figura 69. Gràfica de caixes de la curtosis de la component verda _____	100
Figura 70. Gràfica de comparació múltiple amb Kruskal-Wallis de la curtosis component verda _____	101
Figura 71. Gràfica de caixes de la convexitat _____	102
Figura 72. Gràfica de comparació múltiple amb Kruskal-Wallis de la convexitat _____	102
Figura 73. Gràfica de caixes de la curtosis de la component vermella _____	103
Figura 74. Gràfica de comparació múltiple amb Kruskal-Wallis de la curtosis component vermella _____	104
Figura 75. Gràfica de caixes de l'elongació _____	105
Figura 76. Gràfica de comparació múltiple amb Kruskal-Wallis de l'elongació _____	105

Índex de taules

Taula 1. Descriptors de morfologia _____	32
Taula 2. Descriptors de color _____	34
Taula 3. Comparació entre diferents mètodes de l'estudi de rellevància _____	39
Taula 4. Percentils de l'eix de major longitud _____	41
Taula 5. Estudi de la significança de l'eix de major longitud _____	41
Taula 6. Valor dels percentils de la mitjana en la component verda _____	43
Taula 7. Estudi de la significança segons la mitjana de la component verda _____	43
Taula 8. Valor dels percentils de l'excentricitat _____	45
Taula 9. Estudi de la significança segons l'excentricitat _____	45
Taula 10. Valor dels percentils de la circularitat _____	46
Taula 11. Estudi de la significança segons la circularitat _____	47
Taula 12. Valor dels percentils per la mitjana d'intensitats de la component vermella _____	48
Taula 13. Estudi de la significança segons la mitjana de la intensitat de la component vermella. _____	49
Taula 14. Valor dels percentils per l'eix de menor longitud _____	50
Taula 15. Estudi de la significança segons l'eix de menor longitud _____	51
Taula 16. Valor dels percentils de l'asimetria estadística _____	52
Taula 17. Estudi de la significança segons l'asimetria estadística de la component blava. _____	52
Taula 18. Valors dels percentils de la variància circular _____	54
Taula 19. Estudi de la significança segons la variància circular _____	54
Taula 20. Valors dels percentils pel perímetre _____	56

Taula 21. Estudi de la significança segons el perímetre _____	56
Taula 22. Valor dels percentils de la mitjana d'intensitat de la component blava _____	57
Taula 23. Estudi de la significança segons la mitjana de la component blava _____	58
Taula 24. Exactitud dels mètodes de classificació _____	64
Taula 25. Matriu de confusió SVM _____	64
Taula 26. Anàlisi de l'impacte ambiental _____	67
Taula 27. Costos de personal _____	73
Taula 28. Costos de material _____	74
Taula 29. Costos informàtics _____	74
Taula 30. Cost total _____	75
Taula 31. Taula de percentils extensió _____	80
Taula 32. Estudi de la significança segons l'extensió _____	80
Taula 33. Taula de percentils rodonesa CH _____	81
Taula 34. Estudi de la significança segons la rodonesa CH _____	81
Taula 35. Taula de percentils àrea _____	83
Taula 36. Estudi de la significança segons l'àrea _____	83
Taula 37. Taula de percentils entropia vermell _____	84
Taula 38. Estudi de la significança segons l'entropia component vermella _____	84
Taula 39. Taula de percentils desviació estàndar vermell _____	86
Taula 40. Estudi de la significança segons la desviació estàndar component vermella _____	86
Taula 41. Taula de percentils desviació estàndar verd _____	87

Taula 42. Estudi de la significança segons la desviació estàndard component verda _____	87
Taula 43. Taula de percentils entropia blau _____	89
Taula 44. Estudi de la significança segons l'entropia component blava _____	89
Taula 45. Taula de percentils curtosis component blava _____	90
Taula 46. Estudi de la significança segons la curtosis component blava _____	90
Taula 47. Taula de percentils entropia component verda _____	92
Taula 48. Estudi de la significança segons l'entropia component verda _____	92
Taula 49. Taula de percentils desviació estàndard component blava _____	94
Taula 50. Estudi de la significança segons la desviació estàndard component blava _____	94
Taula 51. Taula de percentils asimetria estadística component vermella _____	95
Taula 52. Estudi de la significança segons l'asimetria estadística component vermella _____	95
Taula 53. Taula de percentils asimetria estadística component verda _____	97
Taula 54. Estudi de la significança segons l'asimetria estadística component verda _____	97
Taula 55. Taula de percentils solidesa _____	98
Taula 56. Estudi de la significança segons la solidesa _____	98
Taula 57. Taula de percentils variància el·líptica _____	100
Taula 58. Estudi de la significança segons la variància el·líptica _____	100
Taula 59. Taula de percentils curtosis component verda _____	101
Taula 60. Estudi de la significança segons la curtosis component verda _____	101
Taula 61. Taula de percentils convexitat _____	103
Taula 62. Estudi de la significança segons la convexitat _____	103

Taula 63. Taula de percentils curtosis component vermella _____	104
Taula 64. Estudi de la significança segons la curtosis component vermella _____	104
Taula 65. Taula de percentils elongació _____	106
Taula 66. Estudi de la significança segons l'elongació _____	106

CAPÍTOL 1:

PREFACI

1.1 Origen del treball

L'anèmia afecta a persones d'arreu del món amb un percentatge significatiu, pel que es creu necessari trobar una eina que permeti facilitar i accelerar el diagnòstic del pacient. Això és el que pretén aquest treball, aconseguir una eina alternativa i automàtica que mitjançant el processat d'imatges de mostres sanguínies es detecti la causa de l'anèmia.

Aquest treball és una continuació del projecte del treball de final de grau de l'alumna Clara Pladevall Viladecans [1]. El seu projecte donava una primera visió molt ben treballada sobre la Drepanocitosi (tipus d'anèmia faciforme) i els descriptors de forma dels eritròcits que la caracteritzen. En el meu treball s'ha volgut fer un estudi més a fons sobre cada tipus de descriptor trobat, concretant-los i extraient-ne de nous per obtenir més valors de classificació. Alhora s'ha centrat en l'estudi de l'Esferocitosi (un altre tipus d'anèmia hemolítica), el que ha permès contrastar i alhora precisar els resultats.

1.2 Motivació

La motivació sobre el processat d'imatges em va sorgir al cursar l'assignatura de Processat d'Imatges Biomèdiques en la titulació d'enginyeria biomèdica a la UPC, on vaig poder conèixer aquest camp de treball i algunes de les funcions bàsiques del tractat d'imatges. És per això que volia ampliar els meus coneixements i seguir descobrint noves eines i resultats.

El fet de continuar una investigació i un treball ja començats aporta una motivació extra ja que a partir dels resultats dels que parteixes, tens l'oportunitat de treballar-los, comparar-los per adquirir-ne de millors, i ampliar el camp de treball obtinguen així uns resultats més amplis i amb més significat.

1.3 Requeriments previs

Per la realització d'aquest treball cal que l'autor/a tingui una base en el software MATLAB ja que és l'eina utilitzada per el processat i la posterior extracció de característiques de les imatges. En el meu cas he treballat amb aquest programa tots els anys del grau d'Enginyeria Biomèdica que he cursat, però tot i així, durant la realització del treball, he ampliat molt els meus coneixements ja que m'ha permès aprofundir amb noves i interessants funcions del programa.

Alhora és important adquirir coneixements en citologia i sobretot dels tipus d'eritròcits sobre els que treballem, pel que és necessària una recerca teòrica prèvia. També cal una base en estadística per a dur a terme la posterior anàlisi estadístic dels descriptors.

1.4 Diagrama de Gantt

En aquest apartat s'exposa en forma de diagrama el temps de dedicació necessari per a la realització de cada una de les parts del treball.

DIAGRAMA DE GANTT									
Part	Inici	Fi	12-30 Set	1-15 Oct	16-31 Oct	1-15 Nov	16-30 Nov	1-15 Des	16-31 Des
Redacció de la memòria	01/11/2017	31/12/2017							
Recerca d'informació mèdica	12/09/2017	31/12/2017							
Recerca d'informació tècnica	01/10/2017	30/11/2017							
Adquisició imatge microscòpiques	01/10/2017	30/11/2017							
Algoritme preprocessament i segmentació	01/10/2017	15/11/2017							
Algoritme extracció de característiques de forma	16/10/2017	30/11/2017							
Algoritme extracció d'altres característiques	16/10/2017	30/11/2017							
Algoritme d'etiquetatge i relació de dades	01/11/2017	16/12/2017							
Adquisició i interpretació dels resultats	16/11/2017	31/12/2017							
Anàlisi estadístic	01/12/2017	31/12/2017							
Pressupost	01/12/2017	31/12/2017							
Conclusions	16/12/2017	31/12/2017							

Figura 1. Diagrama de Gantt

CAPÍTOL 2:

INTRODUCCIÓ

Segons la Organització Mundial de la Salut, l'anèmia afecta al 24,8 % de la població [2], el que implica que 1 de cada 4 persones pateixen aquesta malaltia. L'anèmia és una malaltia causada per la disminució de la massa d'eritròcits i d'hemoglobina en sang. En l'anèmia hemolítica una part dels eritròcits adquireixen una forma anòmala i es destrueixen abans dels 120 dies mitjans de vida. [3]

L'anèmia és una patologia estesa arreu del món que afecta a països desenvolupats i a països en desenvolupament. Es pot detectar tant en nens com en adults i en casos extrems pot dur a la persona a la mort. És per això que es considera de gran importància el diagnòstic de l'anèmia a temps.

Les tècniques de diagnòstic són varies i es basen en l'anàlisi de la morfologia i quantitat dels eritròcits. Un dels mètodes més utilitzats és l'observació en microscòpia, el qual requereix un professional ben format i temps de dedicació. Al llarg dels anys les tècniques d'anàlisi han evolucionat i actualment el diagnòstic es pot realitzar mitjançant l'anàlisi digital automàtic d'imatges d'eritròcits, un exemple n'és l'equip CellaVision DM96[4].

L'objectiu d'aquest treball és aconseguir un mètode automàtic de classificació de les cèl·lules sanguínies, concretament esferòcits i drepanòcits, a partir d'imatges. Aquestes seran obtingudes de frotis sanguinis de sang perifèrica gràcies al microscopi òptic Olympus DP73 connectat a la càmera BX43. Per assolir l'objectiu es treballarà amb mètodes de processat digital d'imatges, com es menciona en el treball de Mohammed, et al. [5] la segmentació ha de ser acurada per obtenir resultats fiables. Seguidament s'extrauran descriptors que permetran la diferenciació entre cèl·lules. L'estudi estadístic dels descriptors ens permetrà detectar quins seran de més rellevància.

Aquest mètode permetrà reduir el temps de dedicació dels professionals i dur a terme un anàlisi morfològic eritrocitari més objectiu, el que representa una millora del diagnòstic inicial dels pacients amb anèmies.

Els algorismes utilitzats en aquest projecte es poden trobar a la plataforma GitHub amb el nom TFG-NuriaTorrens [6].

2.1 Objectius del treball

Objectiu principal:

- Aconseguir un mètode per a fer possible la **classificació automàtica dels eritròcits** obtinguts de frotis de sang perifèrica, concretament esferòcits, drepanòcits i eritròcits normals.

Objectius secundaris:

- Millorar els algorismes de processat i segmentació.
- Extreure nous descriptors per a la identificació de les formes eritrocitàries objecte d'estudi.
- Ampliar coneixements mèdics en el camp de la citologia de sang perifèrica i en l'ús del material utilitzat.
- Adquirir coneixements en algorismes de programació i en el programa MATLAB

CAPÍTOL 3:

LA SANG

La sang és un teixit líquid que circula pels vasos sanguinis dels vertebrats. Està formada per una part líquida que inclou el plasma sanguini, i per una part sòlida que conté els glòbuls vermells, els glòbuls blancs i les plaquetes [7].

El cos humà adult té entre 4,5 i 6 litres de sang, que correspon al 7% del pes corporal. El 55% d'aquesta és plasma, compost per aigua, sals minerals i proteïnes mentre que el 45% restant es divideix en 40% de glòbuls vermells i 5 % de glòbuls blancs i plaquetes [8].

La majoria de cèl·lules que formen la sang es creen a la medul·la òssia, la qual es troba al teixit esponjós dels ossos plans i als canals medul·lars dels ossos llargs. Però també trobem la creació de limfòcits tipus T i B als ganglis limfàtics i a la melsa.

Els glòbuls sanguinis es produeixen a partir de les cèl·lules mare pluripotencials localitzades a la medul·la òssia, aquest procés s'anomena hematopoesis. Quan la cèl·lula mare es divideix, origina eritròcits immadurs, limfòcits immadurs o cèl·lules productores de plaquetes que maduraran fins a convertir-se en la part sòlida que compon el teixit sanguini. La sang és un teixit renovable del cos humà, pel que la medul·la òssia tindrà la funció de producció de cèl·lules sanguínies durant tota la vida. La duració útil d'aquest tipus de cèl·lules varia en funció de les condicions de l'organisme i de les necessitats de salut d'aquest, però com a valors generals podem afirmar que els glòbuls vermells tenen una vida de 120 dies, els glòbuls blancs una duració que varia entre unes hores i uns dies, i les plaquetes una vida de fins a 10 dies.

En la imatge següent es mostren les cèl·lules sanguínies que es formen a la medul·la òssia.

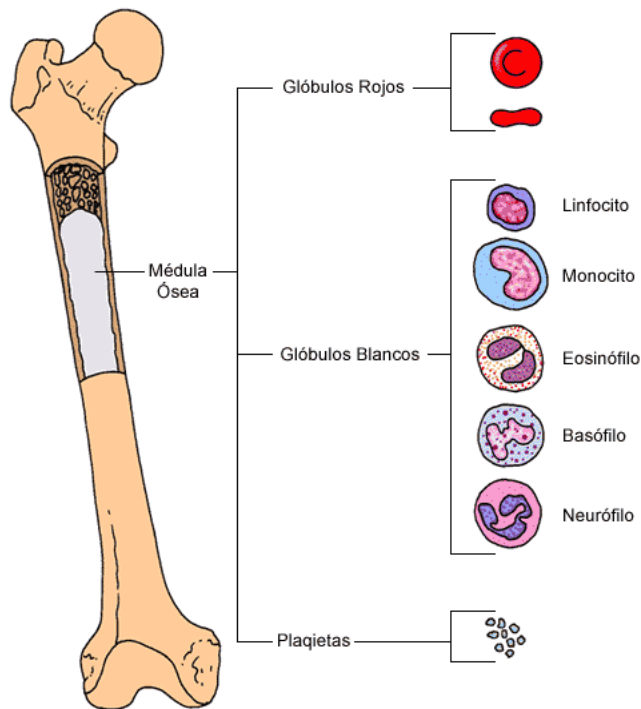


Figura 2. Elements formats en la medulla òssia. (Font: AEAL)

Les funcions de la sang són transport, regulació i protecció:

- **Transport:** La sang transporta oxigen des dels pulmons fins als teixits i òrgans del cos, i mou el diòxid de carboni en sentit contrari, per exhalar-lo en la respiració. També porta nutrients essencials des del tracte gastrointestinal i hormones des de les glàndules endocrines cap a les cèl·lules. Transporta calor i productes de rebuig cap a diferents òrgans per a la seva expulsió.
- **Regulació:** La sang regula el pH corporal (lleugerament bàsic, entre 7,35 i 7,45) gràcies als *buffers* o substàncies reguladores, les quals contenen i alliberen una espècie àcida o bàsica en funció de les necessitats de regulació. També ajuda a mantenir el nivell d'homeòstasi dels líquids de l'organisme i a regular la temperatura corporal.
- **Protecció:** Protecció davant de malalties gràcies a la capacitat de fagocitosi dels glòbuls blancs i de les diverses proteïnes sanguínies. Capacitat de coagulació de les plaquetes per evitar pèrdua excessiva de sang en les ferides.

Cada tipus de component sanguini té una funció específica que s'explica en els següents apartats.

3.1 Eritròcits

Els eritròcits, glòbuls vermells o altrament dits hematies són cèl·lules caracteritzades pel seu color vermellós i la seva forma de disc bicòncav amb el centre més enfonsat que la resta de l'estructura cel·lular. El seu diàmetre oscil·la entre 5 i 7,5 μm i la seva grossor és d'1 μm . El percentatge de volum total de sang ocupat per glòbuls vermells s'anomena hematòcrit.

A diferència de moltes altres cèl·lules, els eritròcits no tenen nucli, el que juntament amb la seva forma bicòncava els aporta una major flexibilitat i els permet modificar lleugerament la seva forma per poder adaptar-se a través dels diferents vasos sanguinis pels quals circulen [9]. En els mamífers, els glòbuls vermells no tenen mitocondris, el que els impedeix fer la respiració cel·lular. És per aquest motiu que utilitzen la glucosa per produir energia a través del procés de glucòlisi seguit per la fermentació làctica.

La funció dels eritròcits és el transport de l'oxigen des dels pulmons fins als diferents teixits del cos, així com l'eliminació del diòxid de carboni dels teixits per dur-lo als pulmons com a substància residual. L'hemoglobina (Hb) és la proteïna dels eritròcits encarregada de transportar l'oxigen. Constitueix el 35 % del volum dels eritròcits i la seva tonalitat vermella és el que els dona el color. Els nivells normals d'hemoglobina en adults estan entre 120 i 160g/L de sang, nivells inferiors podrien desenvolupar anèmies i per tant, dèficit de transport d'oxigen en sang[7].

Els eritròcits sobreviuen aproximadament 120 dies. Un cop transcorreguts, perden la membrana cel·lular i són eliminats de la sang per les cèl·lules fagocítiques de la melsa, del fetge i la medul·la òssia. L'hemoglobina és degradada per aquestes cèl·lules i en són reutilitzades les proteïnes per a altres funcions.

3.1.1 Eritropoesi

La formació d'eritròcits rep el nom d'eritropoesi, ja que es du a terme gràcies a l'hormona eritropoetina (del grec erythros - vermell i poiein – que facilita la creació). Com s'ha explicat anteriorment, els eritròcits s'originen a partir de les cèl·lules mare de la medul·la òssia. Aquestes donen lloc als proeritroblasts, unes cèl·lules amb un nucli voluminós i un diàmetre que equival dues vegades al d'un eritròcit. La seva funció serà emmagatzemar ferro per fabricar hemoglobina posteriorment. Dels proeritroblasts en deriven els eritroblasts; en aquest impàs les cèl·lules adopten mides més petites i pateixen diverses divisions fins a perdre el nucli. Aquests contenen ribosomes per a la formació de l'hemoglobina i un cop els nivells d'hemoglobina són els necessaris, la cèl·lula redueix el nivell de ribosomes i forma els reticulòcits, eritròcits immadurs. Els reticulòcits maduren en un període aproximat de 6 dies i s'abocuen a la sang convertint-se finalment en glòbuls vermells.

L'eritropoetina és segregada pel ronyó i activada a través del fetge i el nivell de producció d'aquesta hormona varia segons les necessitats del cos, és a dir, segons el dèficit o excés d'eritròcits[10].

En la següent imatge es mostra el procés de l'eritropoesi:

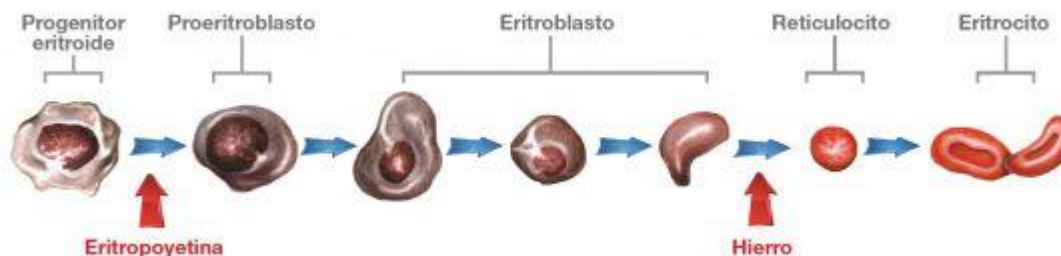


Figura 3. Procés formació glòbuls vermells. (Font: Probiomed)

3.2 Leucòcits

Els leucòcits o glòbuls blancs són les cèl·lules de defensa del sistema immunitari i tenen la funció de combatre i destruir microorganismes o cossos estranys que ataquin el cos, són cèl·lules mòbils i utilitzen els vasos sanguinis i la sang per arribar als punts d'infecció arreu de l'organisme. Es produeixen a la medul·la òssia i a la sang trobem valors entre 4.000 i 10.000 leucòcits/mm³, variable segons les condicions patològiques i fisiològiques.

A diferència dels eritròcits, els leucòcits tenen nucli, mitocondri i orgànuls cel·lulars. El seu diàmetre varia entre 8 i 20µm, per tant són més grans que els glòbuls vermells, i al no tenir hemoglobina, el citoplasma no és de color vermellós. Aquestes dues característiques seran essencials per poder diferenciar-los a través del microscopi òptic.

Els leucòcits es classifiquen en diferents tipus. El grup més comú són els neutròfils, ja que formen el 60 % del volum total de leucòcits del cos. Defensen l'organisme contra infeccions bacterianes i fongs i són el tipus de leucòcit més ràpid en generar una resposta.

El segon grup més comú són els limfòcits. Hi ha dues varietats d'aquest tipus de cèl·lula: els limfòcits T que ajuden a regular la funció d'altres cèl·lules immunitàries i ataquen a cèl·lules infectades i els limfòcits B, els qual produeixen anticossos (glicoproteïnes capaces d'identificar i neutralitzar elements estranys).

3.3 Les plaquetes

A diferència dels eritròcits i leucòcits, les plaquetes o trombòcits no són cèl·lules, sinó que es defineixen com a petits fragments citoplasmàtics irregulars, sense nucli, amb un diàmetre de 2 a 3µm. La seva creació també es du a terme a la medul·la òssia a partir de la fragmentació del citoplasma d'un tipus de cèl·lula precursora anomenada megacariòcit.

La seva funció és taponar les àrees danyades dels vasos sanguinis per evitar l'hemorràgia i facilitar la coagulació de la sang. Un nombre molt elevat de plaquetes pot ser contraproductiu, ja que pot causar coàguls a la sang, això es dona en determinades malalties hematològiques proliferatives.

3.4 El plasma sanguini

El plasma sanguini és la porció líquida de la sang on estan immerses les cèl·lules sanguínies i compon el 55 % del volum total de la sang. Està format en major part per aigua i conté altres substàncies com nutrients, sals minerals, hormones i proteïnes. El plasma és essencial per la coagulació de la sang, la irrigació dels teixits, la defensa immunitària i per mantenir la viscositat de la sang per que pugui circular per les artèries.



CAPÍTOL 4:

L'ANÈMIA

L'anèmia és una malaltia del sistema circulatori causada per un dèficit de glòbuls vermells sans en la sang, el que implica una concentració d'hemoglobina inferior a la normal, i impedeix un bon transport de l'oxigen necessari pel l'organisme. En aquesta afecció una part dels glòbuls vermells es destrueixen abans dels 120 dies normals de vida i la medul·la òssia no és capaç de produir-ne prou ràpidament per aconseguir uns nivells necessaris. Hi ha diferents tipus d'anèmia, quan està relacionada amb la vida mitja dels glòbuls vermells s'anomena **anèmia hemolítica**. [11]

Podem classificar l'anèmia hemolítica (AH) segons diferents aspectes:

-Segons l'etiologia:

- AH congènita: Els trets genètics defineixen la malaltia. La forma i funció de l'eritròcit depèn de la interacció entre tres unitats cel·lulars com són la membrana de l'eritròcit, els elements solubles intracel·lulars i l'hemoglobina.
 - Alteracions de la membrana: L'exemple més comú d'anèmia per alteracions a la membrana és l'**esferocitosi**. La membrana de l'eritròcit és la responsable de les propietats mecàniques i de moltes de les seves funcions fisiològiques. En el cas dels esferòcits hi ha un defecte en la unió entre el citoesquelet i la bicapa lipídica, el que comporta una disminució de la relació superfície/volum i per tant una pèrdua d'elasticitat del glòbul vermell.
 - Deficiències enzimàtiques: Apareix quan hi ha una manca d'enzims a l'espai intracel·lular de l'eritròcit. La més comú és la deficiència de la glucosa-6-fosfat deshidrogenasa la qual es transmet hereditàriament lligada al cromosoma X.
 - Hemoglobinopaties: En un eritròcit normal hi conviuen diversos tipus d'hemoglobina. Les hemoglobinopaties es produeixen per trastorns en la síntesi de la cadena que forma l'hemoglobina. Exemples són les talassèmies (producció de

menys glòbuls vermells sans i menys hemoglobina) i l'anèmia falciforme o **drepanocitosis**.

- AH adquirida: El medi on es troben immersos els hematies i els factors i condicions que afecten a la persona són els causants de la destrucció prematura.
 - AH immune: Trastorn immune on diversos tipus d'anticossos estan dirigits directament contra els antígens dels eritròcits, el que disminueix la seva supervivència. Quan els anticossos són produïts per el propi individu s'anomena anèmia hemolítica autoimmune.
 - AH no immune: Pot ser causada per agents físics, infeccions, fragmentació dels eritròcits, entre d'altres.

-Segons els **mecanismes de desintegració dels hematies**: Els eritròcits no tenen nucli ni orgànuls pel que no es podem reparar i moren quan es desgasten. La tonicitat del medi on es troben té una gran influència en el seu desgast. El procés de desintegració s'anomena hemòlisis.

- Extravascular: El 85% dels eritròcits es destrueixen extravascularment, sense alliberar la seva hemoglobina al plasma sanguini. Aquesta destrucció es pot produir a la melsa i amb menor mesura al fetge i a la medul·la òssia.
- Intravascular: L'hemòlisi intravascular és la destrucció de l'eritròcit dins de la circulació alliberant així l'hemoglobina al plasma sanguini. [12]

-Segons la **duració de l'anèmia**:

- Agudes: Quan l'inici i el fi de la malaltia estan ben definits. Generalment els símptomes apareixen ràpidament i el diagnòstic és més fàcil i urgent que en les cròniques. Les anèmies intravasculares acostumen a ser agudes.
- Cròniques: Quan l'anèmia és de llarga duració i de progressió lenta. El diagnòstic pot ser més lent ja que la manifestació dels símptomes no sempre és clara. Les anèmies hemolítiques extravasculares acostumen a ser cròniques.

4.1 Esferocitosi

L'esferocitosi és una malaltia caracteritzada per una anèmia hemolítica amb un trastorn en la membrana dels eritròcits que comporta que aquests adoptin forma esferoïdal i es descomponguin de manera prematura. Va ser descrita al 1871 per dos metges belgues, Vanlair i Masius.[13] Els eritròcits deformats s'anomenen esferòcits i es troben en sang perifèrica.

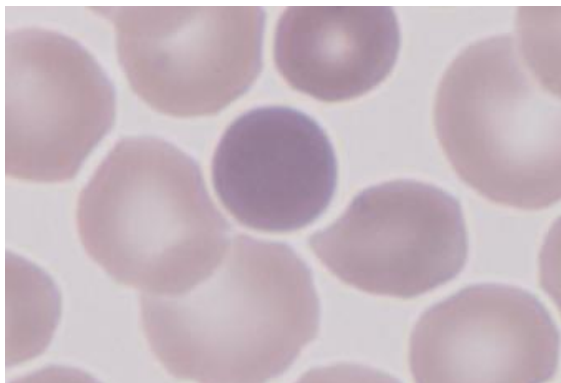


Figura 4. Imatge d'un esferòcit. És la cèl·lula central de color més fosc. (Font: Imatge adquirida dels frotis cedits per l'Hospital Clínic.)

L'esferocitosi pot ser adquirida o hereditària. Ambdues comporten la producció anormal de glòbuls vermells i es tracten de formes similars. L'esferocitosi adquirida apareix quan una persona amb eritròcits sans està exposada de manera prolongada, per exemple, a toxines o a anticossos, com a conseqüència, els glòbuls vermells perden la seva forma i adquireixen forma d'esfera.

En el cas de l'esferocitosi hereditària (EH), els gens són els culpables que la membrana no disposi de les proteïnes necessàries per adquirir una forma sana. L'EH és l'anèmia hemolítica més freqüent al món amb una prevalença de 1:2000 persones en alguns països europeus. El 75% dels casos són autosòmics dominants, i es pot observar en diferents races i grups ètnics [14].

La membrana del glòbul vermell està formada per una bicapa lipídica on predominen els fosfolípids i el colesterol i amb una menor mesura, els glicolípids i els aminofosfolípids. La membrana és la responsable de les propietats mecàniques i de moltes funcions fisiològiques de la cèl·lula. Les mutacions genètiques afecten a les proteïnes de membrana (espectrina i anquirina) necessàries per mantenir la forma normal de disc bicòncav de l'eritròcit, i fan que l'eritròcit adquireixi una forma més esferoïdal. Això comporta una pèrdua de l'elasticitat i de la capacitat de deformació de la cèl·lula, i alhora, produeix un augment en la concentració mitjana d'hemoglobina, pel que les cèl·lules adopten un color més vermellós que els eritròcits sans.

El diagnòstic de la malaltia es pot dur a terme a través d'una anàlisi de sang i de la visualització d'aquesta al microscopi. La fragilitat osmòtica del glòbul vermell és la prova més utilitzada en aquest diagnòstic. Aquesta prova s'utilitza tant per l'esferocitosi hereditària com per la talassèmia i consisteix en submergir els glòbuls vermells en una solució que els fa inflar-se, el resultat sorgirà de la mesura del temps que aguantin o tardin en descompondre's[15]. En cas de nadons o de pacients amb símptomes no gaire definits, és interessant fer un estudi genètic dels familiars.

Pel que fa al tractament, depenent de la evolució poden ser necessàries transfusions de sang o la realització d'una esplenectomia. Aquesta tècnica consisteix en l'extirpació de la melsa del pacient i es du a terme sols en casos extrems. És el tractament més efectiu en el control de l'anèmia. Després de l'esplenectomia, l'esferocitosi persisteix però els nivells de reticulòcits s'ajusten als normals i el temps de vida dels hematies augmenta[16].

4.2 Drepanocitosi

La drepanocitosi o anèmia de cèl·lules falciformes és una malaltia que afecta la sang dels éssers humans produint glòbuls vermells de forma anormal. És una anèmia hemolítica congènita que juntament amb la talassèmia es classifica en el grup de les hemoglobinopaties, malalties que afecten a l'hemoglobina.

En les persones amb drepanocitosi, alguns dels eritròcits adopten forma semilunar i s'anomenen drepanòcits. Això dificulta la seva funció i la mobilitat a través dels vasos sanguinis i provoca una disminució de la vida dels glòbuls vermells [17].

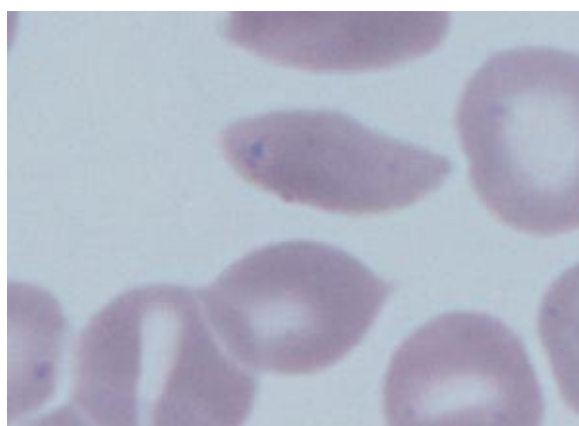


Figura 5. Imatge d'un drepanòcit. És la cèl·lula de forma allargada semilunar. (Font: Imatge adquirida dels frotis cedits per l'Hospital Clínic.)

És una malaltia d'origen genètic on es substitueix un aminoàcid per un altre a la cadena de globina (proteïna que forma part de l'hemoglobina beta)[18]. S'hereta d'ambdós progenitors i en cas que només un d'ells porti el gen, la malaltia no es manifesta en el pacient. És una de les malalties genètiques més freqüents a nivell mundial i es dona amb més freqüència en persones de ascendència africana.

La variació de l'hemoglobina beta dona lloc a un tipus d'hemoglobina conegut com a hemoglobina S. Els eritròcits resultants presenten més rigidesa podent obstruir els vasos sanguinis petits i reduir el flux de sang, el que comporta una disminució de l'aportació d'oxigen. [17]

En general la drepanocitosi es manifesta a partir dels cinc mesos de vida i la seva gravetat varia segons la persona. Conforme passa el temps la malaltia tendeix a agreujar-se, és per això que hi ha tractaments que redueixen aquesta evolució i ajuden a augmentar la vida del pacient.[17] El trasplantament de medul·la òssia és l'únic tractament definitiu ja que consisteix en substituir la medul·la òssia danyada per una que aporti la formació d'eritròcits sans. Tot i així, és una solució que no es pot donar a tots els pacients i de la qual cal valorar-ne els riscos.

4.3 L'hemoglobina

L'hemoglobina (Hb) és una proteïna de color vermellós que es troba exclusivament en els glòbuls vermells de la sang. La seva funció és transportar oxigen dels pulmons fins als teixits del cos i el diòxid de carboni dels teixits als pulmons per a la seva exhalació.

L'Hb està formada per 4 cadenes de globina, una proteïna globular. El tipus d'hemoglobina que trobem més present en adults és del tipus A (HbA) i està formada per dos globines alfa i dues beta. Quan aquestes cadenes beta pateixen una alteració genètica estructural apareix l'hemoglobina de tipus S, que és la causant de la drepanocitosi.

La detecció de l'hemoglobina S es du a terme a través de l'electroforesi (tècnica de separació de biomolècules sotmeses a un camp elèctric) o de manera cromatogràfica[19].



CAPÍTOL 5: ADQUISICIÓ DE LES IMATGES

L'objectiu del treball és obtenir una eina de classificació automàtica capaç de donar suport als metges en el procés de diagnòstic d'anèmies, concretament drepanocitosi i esferocitosi contribuint a reduir la part subjectiva d'aquest.

S'ha treballat a partir d'imatges obtingudes de 10 frotis sanguinis els quals corresponen a 10 pacients diferents diagnosticats amb esferocitosi. Els frotis són portaobjectes que contenen les mostres de fluid, en aquest cas sang perifèrica i els quals han estat cedits per l'Hospital Clínic de Barcelona. De cada frotis s'ha extret una mitja de 30 imatges, obtenint així un total de 300 sobre les quals treballar. Aquestes juntament amb les 300 imatges de les que disposàvem del treball de final de grau de Clara Pladevall [1], sumen **600 imatges**, suficients per donar validesa al resultat.

5.1 Material utilitzat

Per a l'adquisició de les imatges s'ha utilitzat la càmera Olympus DP73 [20] connectada al microscopi òptic Olympus BX43 [21] a 1000 augments i alhora connectada al PC. Per a capturar les imatges s'utilitza el programa d'ordinador CellSens Entry. El procés s'ha dut a terme a les instal·lacions de l'Escola d'Enginyeria de Barcelona Est (EEBE).

Al col·locar el frotis sanguini al microscopi, la càmera projecta la visió que es percep pels oculars a la pantalla de l'ordinador i amb el programa CellSens es pot veure la projecció en directe i permet extreure les imatges.

5.2 Procediment d'adquisició

El procés d'adquisició comença al col·locar la mostra a la platina del microscopi i acaba al capturar la imatge.

5.2.1 Preparació de la mostra

Les mostres es preparen a l'Hospital Clínic. Es diposita una petita mostra de sang perifèrica sobre el portaobjectes i s'estén, amb un mètode de fixació com calor o productes químics es fixa sobre el portaobjectes. Es du a terme el procés May Grünwald-Giemsa per la tinció de la mostra, ja que és una de les millors tècniques per diferenciar qualitativament i quantitativa els components sanguinis. Aquest mètode combina dues tècniques de tinció, la tècnica May-Grünwald que conté colorant eosina i blau de metilè dissolts en metanol i la Giemsa, que conté també eosina, blau de metilè i alguns productes d'oxidació.

Aquesta tècnica de coloració tenyeix els components sanguinis amb una escala que varia de colors blaus i rosats fins a violeta i púrpura, depenent del tipus de mostra i del seu pH. En general els eritròcits quedaran tenyits amb tons rosats o beix, els limfòcits amb tons blavosos i el plasma sanguini amb un to rosa més clar.

5.2.2 Col·locació de la mostra al microscopi

Es comprova que la platina del microscopi estigui el més avall possible i es col·loca la mostra subjecta a les pinces d'aquesta. S'escull la intensitat de llum d'interès i es comença observant per l'objectiu més petit del microscopi, el de 10x. Es mou el cargol macromètric fins que es pot veure la mostra a través de l'ocular i per veure la imatge més nítida i acabar d'enfocar s'utilitza el cargol micromètric.

Amb l'objectiu de 10x els glòbuls vermells es veuen petits i és molt difícil la seva diferenciació a simple vista, per això caldrà utilitzar objectius de més augments. De totes maneres aquest objectiu tant petit ens permet localitzar les zones d'interès del frotis, ja que cal buscar una zona de la mostra no gaire interior on els eritròcits no estiguin molt aglomerats i alhora no molt exterior perquè no hagin perdut la seva forma natural. A la Figura 6 veiem eritròcits que a causa del tractat de la mostra han perdut la zona pàl·lida central que els caracteritza, pel que seria una mala zona on extreure les imatges. En canvi a la Figura 7 veiem diferents tipus d'eritròcits amb la seva forma natural.

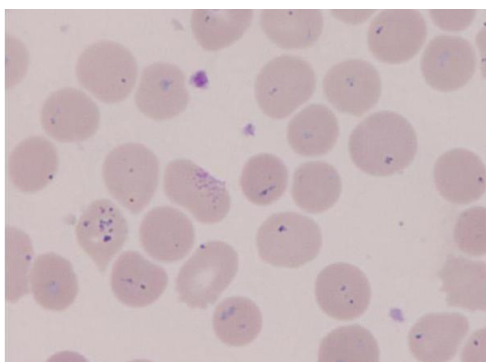


Figura 6. Eritròcits que han perdut la seva forma natural, és una zona poc adequada per l'observació (x1000)

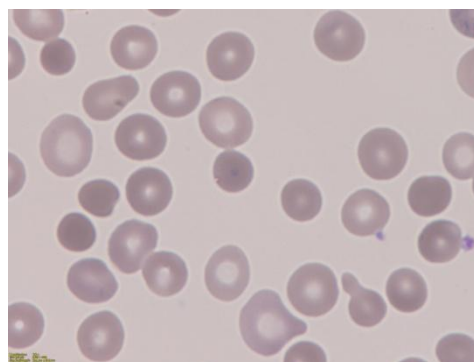


Figura 7. Imatge d'hematies normals juntament amb esferòcits (x1000)



Un cop localitzada la zona d'interès baixem altre vegada la platina i girem el revòlver per col·locar l'objectiu de 50x. Seguim el mateix procediment d'enfocament que per l'objectiu de 10x i utilitzant els cargols que permeten moure la platina verticalment i horitzontal, trobem la millor zona de visualització dels eritròcits. Aquest pas és important ja que al passar a l'objectiu de 100x el camp visual serà molt reduït.

El tercer i últim objectiu és el de 100 augments o també anomenat objectiu d'immersió. Baixem la platina i en el moment de girar el revòlver dipositem una petita gota d'oli d'immersió sobre el punt de la mostra on hi enfoca la llum. Aquest oli té el mateix índex de refracció que el vidre i s'utilitza per reduir la dispersió de la llum i incrementar la resolució de la imatge. Un cop s'ha posat la gota, col·loquem l'objectiu de 100x de manera que entri en contacte amb l'oli. Enfoquem amb el cargol macromètric en cas necessari i amb el micromètric i busquem les cèl·lules que ens interessin.

Ja tenim la mostra preparada per a l'adquisició de les imatges.

5.2.3 Captura de la imatge

Amb els dos cargols que mouen la platina es busca l'àrea on es trobin varis esferòcits i un cop localitzats a través dels oculars, passem a mirar el monitor de l'ordinador. Gràcies al programa CellSens Entry podem veure el que s'observa pels oculars i ens permet triar diferents paràmetres de les imatges com la resolució i l'objectiu que estem utilitzant.

Cal fer un últim enfocament a través de les eines del programa ja que varia una mica del microscopi i un cop enllestit es pot capturar la imatge. Per facilitar el processat, extraïem la imatge en dos formats diferents. La guardem amb format *.tif* (amb 12 bits i 2400x1800 de resolució) i amb format *.jpg*, ambdues amb 1000 augments. La imatge extreta en el primer format no es podrà veure al PC sense una prèvia conversió a 8 o 16 bits.

Un cop finalitzat tot el procés d'adquisició d'imatges, es treu la mostra del microscopi i es neteja suaument amb un paper l'oli d'immersió per no fer malbé la mostra.



CAPÍTOL 6:

PROCESSAT DE LES

IMATGES

El processat i tractat de les imatges es du a terme utilitzant el programa Matlab R2017b. Per a l'emmagatzematge dels algorismes i dades s'ha utilitzat la **plataforma GitHub**. Aquesta plataforma facilita el treball en equip, permet guardar el codi de forma ordenada i fer còpies de totes les versions anteriors que posteriorment s'han modificat, de manera que sempre es poden recuperar.

L'objectiu principal dels algorismes creats és aconseguir la diferenciació automàtica entre els diversos components de la sang perifèrica. Les imatges sanguínies contenen eritròcits, limfòcits i plaquetes principalment, però en nostre cas només ens centrem en la classificació dels primers esmentats. Els eritròcits detectats es separaran en normals o anormals, i es realitzarà una posterior classificació dels anormals en esferòcits, drepanòcits, reticulòcits o eritròcits deformats.

Els eritròcits normals són estructures altament circulars amb una pal·lidesa central que reflexa la seva forma de disc bicòncav. En el nostre cas, ens centrarem en els esferòcits i els drepanòcits com a eritròcits anormals. Com s'ha explicat anteriorment, l'esferocitosi es caracteritza per eritròcits de forma casi esferoïdal i una tonalitat més fosca, el que comporta una mancança de la pal·lidesa central que caracteritza als eritròcits. Pel que fa als drepanòcits, es caracteritzen per tenir una forma allargada semilunar degut a una mutació genètica.

Per poder obtenir una classificació automàtica fiable d'ambdós cal primer tractar la imatge i millorar-ne tant les característiques de color com de forma. Seguidament es segmenten i detecten els eritròcits d'interès i se n'extreuen els descriptors, els quals ens donaran els valors per a la diferenciació. Es realitza un etiquetatge manual de les cèl·lules i se n'extreuen els resultats. A continuació, en aquest capítol, es dona una explicació més detallada de cada un dels passos seguits per a la obtenció de resultats.

6.1 Lectura de la imatge

L'obtenció de les imatges del microscopi es du a terme a través del programa CellSens Entry, com s'ha explicat en l'apartat anterior, el qual ens retorna imatges de 12 bits. Per poder visualitzar-les i treballar amb elles, cal convertir-les a 16 bits a través de la següent transformació.

(Eq.6.1)

$$NovaImatge = ImatgeOriginal * ((2^{16} - 1) / (2^{12} - 1))$$

6.2 Preprocessat i segmentació

El preprocessat de la imatge inclou totes aquelles modificacions que cal fer a la imatge introduïda per obtenir una imatge final més nítida i neta sobre la qual sigui més fàcil treballar. El procés de segmentació, com bé el seu nom diu, consisteix en segmentar aquells objectes que ens interessa extreure de la imatge. En el nostre cas, seran els eritròcits, per tal de poder-los tractar individualment i saber-ne la localització.

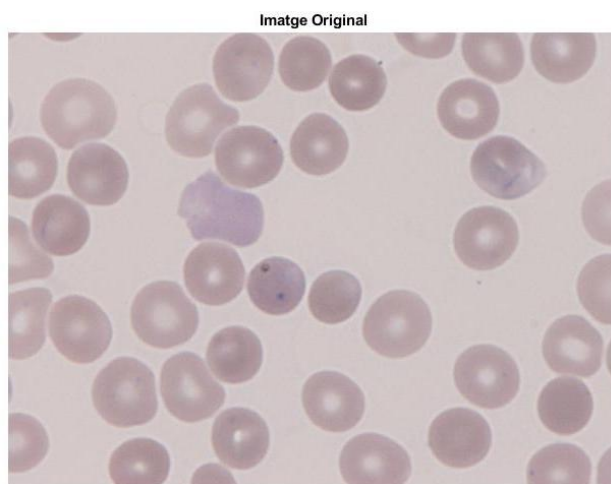


Figura 8. Imatge d'un frotis de sang perifèrica mostnt els hematies (x1000)

6.2.1 Segmentació de les plaquetes

L'objectiu d'aquesta part de la segmentació és la creació d'una màscara binària que elimini les plaquetes discriminant-les per la seva petita mida.

Primer convertim la imatge obtinguda de l'Eq.6.1 a l'espai de color *Lab*. Aquest espai ens permet treballar amb tots els colors de l'espectre, inclús amb els que no percebem a simple vista. Ens retorna la imatge en tres dimensions; *L* per la lluminositat, i *a* i *b* per les dimensions del color. Treballarem sobre la dimensió *b* ja que és la que millor diferencia les plaquetes pel color que tenen i així no perdem informació sobre el color de la imatge.

Creem una màscara binària que contingui els objectes amb dimensions inferiors a 1050 píxels per englobar només les plaquetes. Més endavant, durant el preprocessat, aquesta màscara es restarà de la màscara de la imatge original, eliminant així les plaquetes i deixant només els objectes amb àrees més grans. Depenent del frotis del qual s'hagin extret les imatges s'apreciaran més o menys nombre de plaquetes.

6.2.2 Preprocessament

Ens interessa obtenir una imatge nítida on hi sigui fàcil treballar i amb només els elements que ens puguin ser d'utilitat per extreure'n descriptors. És per això que separem la imatge en les seves components RGB (R pel vermell, G pel verd i B pel blau) i decidim treballar sobre la component verda. Aquesta decisió és deguda a que la component G és on es diferencien millor els objectes del fons i d'on en podem extreure més bons resultats.

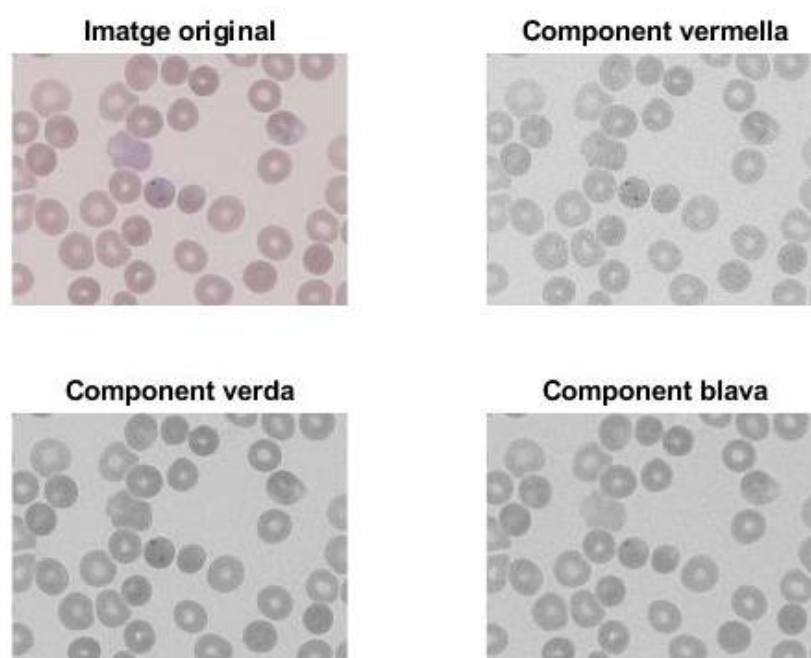


Figura 9. Components del color de la imatge original. Es pot veure que la component verda és la que resalta més les cèl·lules respecte el fons.

Apliquem un ajust de contrast amb la funció de Matlab *imadjust*. El contrast és la diferència del to d'intensitat entre punts de la imatge. Aquesta funció ajusta els valors d'intensitat al llarg de tot l'histograma i els reparteix de forma homogènia.

L'histograma d'una imatge és la representació en forma de gràfica dels diferents nivells d'intensitat o color dels píxels que conté, en relació amb la freqüència que hi apareixen. En la Figura 10 es pot apreciar el canvi al que ha estat sotmès l'histograma després d'aplicar la funció d'ajust de contrast.

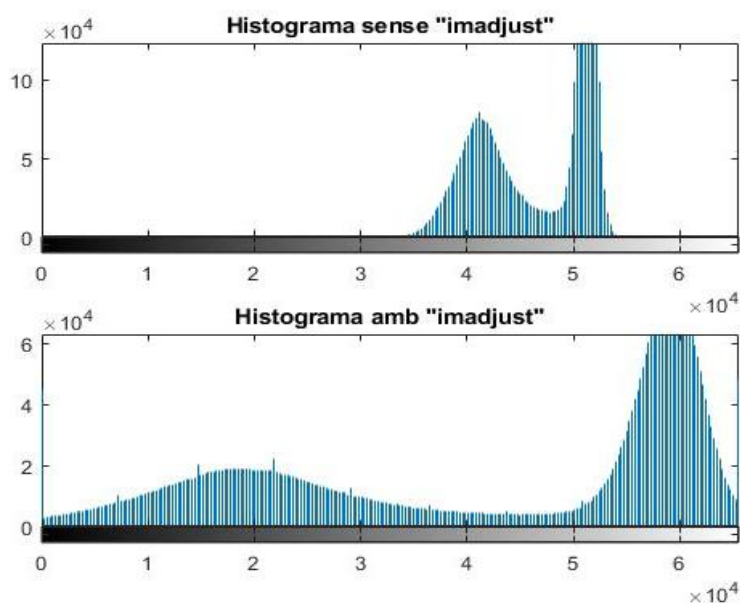


Figura 10. Comparació de l'histograma abans i després d'aplicar la funció "imadjust"

Seguidament s'aplica un filtre gaussià a la imatge. Aquest filtre s'utilitza per eliminar el soroll de la imatge ja que la suavitza. Per fer-ho, analitza els píxels donant més pes al píxel central i menys als píxels llunyans de manera progressiva, i així creant l'efecte de difuminat.

En la Figura 11 es compara la component verda sense filtres amb la component verda un cop se l'hi ha aplicat l'ajust d'histograma i el filtre gaussià. Es pot apreciar com en la imatge filtrada es ressalten més els eritròcits.

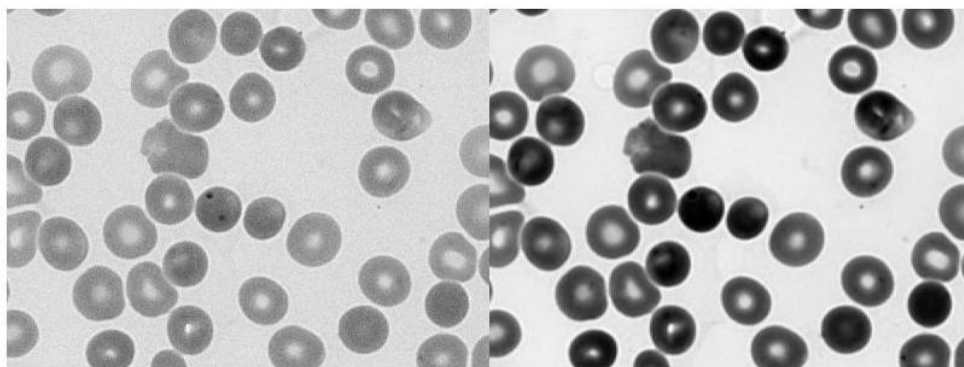


Figura 11. Component verda de la imatge original abans (esquerra) i després (dreta) de filtrar.

Binaritzem la imatge i eliminem les plaquetes gràcies a la màscara creada anteriorment. La binarització consisteix en convertir tots els píxels superiors al valor llindar establert a 1, i els inferiors a 0, de manera que obtenim una imatge amb només valors 1 i 0 (blanc i negre). El valor llindar o "threshold" s'obté a partir del mètode d'Otsu, el qual minimitza la variància dels píxels blancs i negres.

El resultat és una imatge binària on es diferencien les cèl·lules del plasma sanguini, i on ja no hi consten les plaquetes (veure Figura 12)

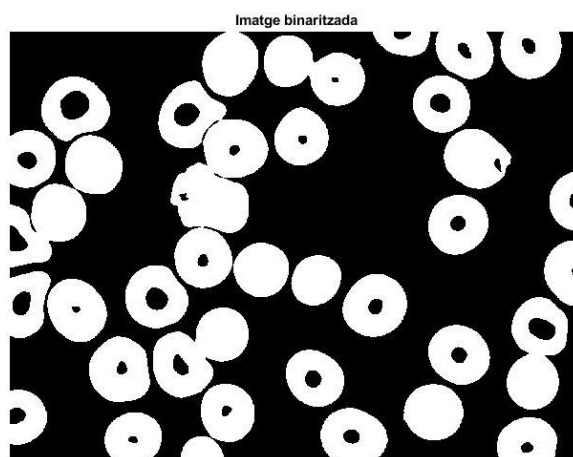


Figura 12. Imatge binaritzada sense plaquetes

Procedim a eliminar elements que no són del nostre interès. Eliminem els objectes connectats amb una àrea menor a 14500 píxels i aquells que estan en contacte amb les cantonades de la imatge, ja que significarà que no són eritròcits sencers.

Aplicant un operador morfològic d'obertura (*imopen*), eliminem les ramificacions entre objectes i arrodonim contorns, igual que reduïm el soroll del fons de la imatge. Aquest operador és el resultat d'aplicar-ne dos seguits: operador d'erosió i operador de dilatació, utilitzant el mateix element estructural per ambdós. L'element estructural ens servirà per determinar el tipus d'obertura que volem i el radi de píxels on interessa que s'apliqui.

L'erosió, com el seu nom indica, erosiona el contorn dels objectes eliminant els contorns irregulars i ajudant a separar alguns objectes. L'operador de dilatació augmenta el contorn dels objectes. Al aplicar-los de manera consecutiva obtenim una imatge amb uns objectes de dimensions molt similars als originals però amb els contorns arrodonits i sense caps. (veure Figura 13)

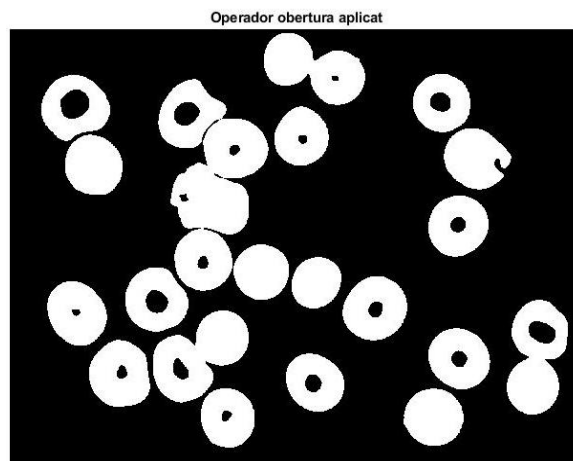


Figura 13. Eliminació objecte en contacte amb les cantonades i aplicació operador morfològic "imopen"

6.2.3 Segmentació dels eritròcits d'interès

El mètode de segmentació que millor funciona en el nostre cas és la segmentació Watershed.

Partim de la imatge binaritzada (BW) dels eritròcits d'interès, i aquests, com a tals, tenen una pal·lidesa central que els caracteritza. Això fa que s'aprecii una zona de zeros (negra) al mig de la zona d'uns (blanca) dels eritròcits. Cal eliminar aquests forats amb una variant de la funció *imfill* ('holes') per obtenir una segmentació correcta. El resultat és una imatge amb el fons negre i els eritròcits en formes arrodonides blanques, sense la zona central diferenciada. (Figura 14)

Per aplicar la transformada *watershed* cal primer crear una modificació de la imatge binaritzada trobant la transformada de la distància Euclidiana d'aquesta. Per cada píxel de la imatge BW, la transformada de la distància assigna un nombre que equival a la distància entre aquell píxel i el píxel amb valor diferent a zero més proper. Hi ha diferents mètodes que especifiquen la forma de càlcul de la distància; nosaltres utilitzem el mètode "chessboard".(Figura 15)

La transformada *Watershed* analitza la imatge obtinguda com si fos una superfície. Els píxels clars els representa com a elevacions d'aquesta superfície i els píxels negres com a valls. D'aquesta manera és capaç de classificar els píxels en regions, que després utilitza per una segmentació adequada. El resultat es mostra en la Figura 14.

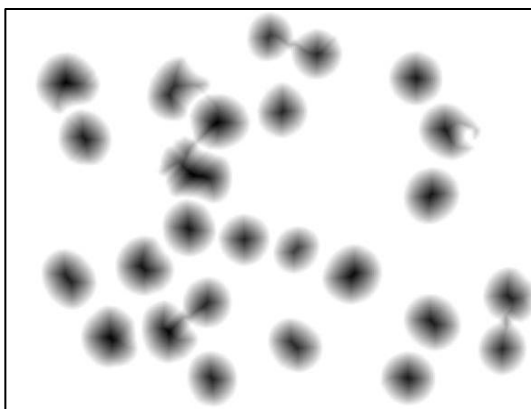


Figura 15. Transformada de la distància amb mètode "Chessboard".

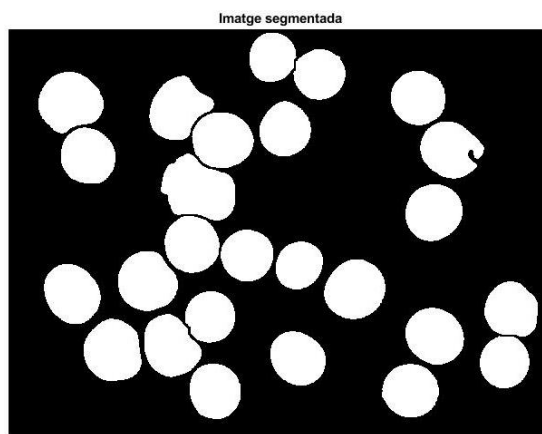


Figura 14. Resultat de la segmentació Watershed on s'han separat els eritròcits.

6.2.4 Filtrat dels esferòcits

El processat i la detecció dels esferòcits és un pèl més complexa que la dels drepanòcits. Sabem que els esferòcits tendeixen a ser més circulars que els eritròcits sans i que el seu color és més fosc. La principal dificultat és que els esferòcits, com s'aprecia en Figura 14, en una imatge binaritzada són difícils de diferenciar dels eritròcits. Per tant cal treballar amb altres descriptors diferents als de forma per la seva detecció.

En aquesta part del processat es pretén acabar de filtrar els possibles trossos d'eritròcits que hagin resultat de la segmentació i els eritròcits amb formes que no són del nostre interès. Per fer-ho es tenen en compte propietats com l'excentricitat i la circularitat, per tant no és un processat aplicable a imatges amb drepanòcits, ja que els desestimaria per la seva forma allargada.

Primerament, calculem i posem en una taula algunes propietats de cada objecte de la imatge com són l'àrea, el centroid i l'excentricitat. Posteriorment, a partir dels valors d'aquestes variables, apliquem les següents condicions.

- 1ra condició: L'àrea dels objectes d'interès (OI) ha de ser inferior a 65.000 píxels, ja que s'ha comprovat experimentalment que en cas contrari es tracta de limfòcits o eritròcits sobreposats que no han estat ben segmentats.
- 2na condició: L'excentricitat dels OI ha de ser inferior a 0.75, ja que sinó, en la majoria dels casos, no es tracta d'eritròcits d'interès, sinó de trossos d'eritròcits mal segmentats. Com més propera a 0, més circular és l'objecte.
- 3ra condició: La circularitat cal que tingui un valor superior a 0.75. Com més proper a 1 és el valor de l'objecte, més circular és.

El resultat de l'aplicació d'aquestes condicions es mostra en la Figura 16. Més endavant, a l'Apartat 6.4 s'expliquen amb detall les tres variables utilitzades per la selecció.

Un cop aplicades aquestes condicions obtenim una nova imatge amb ROIs preparades per extreure'n descriptors.

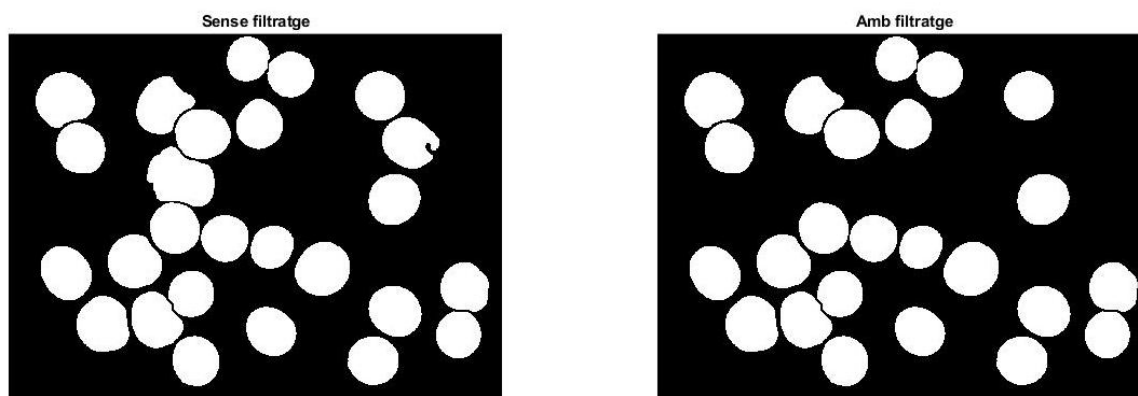


Figura 16. Abans i després del filtratge segons àrea, excentricitat i circularitat.

6.3 Procés d'etiquetatge

Un cop obtenim la imatge amb tots els eritròcits processats i delimitats, cal etiquetar cada cèl·lula per després poder-ho comparar i saber el percentatge d'encert del nostre algoritme.

El procés d'etiquetatge és manual i es du a terme juntament amb el citòleg de l'Hospital Clínic Àngel Molina Borrás. S'utilitza l'aplicació de MATLAB anomenada Image Labeler. L'aplicació permet donar un nom a cada cèl·lula, la qual quedarà delimitada per un rectangle (ROI - regió d'interès) que serà la seva etiqueta. L'objectiu d'aquest procés és tenir tots els eritròcits de les imatges etiquetats segons el tipus que són.

En el cas de l'esferocitosi, s'han tingut en compte 4 tipus de cèl·lules: eritròcits normals, reticulòcits (eritròcits immadurs), esferòcits i eritròcits deformats (Figura 17). Les ROIs etiquetades com a eritròcits deformats, en la majoria de casos, corresponen a eritròcits trencats o sobreposats, pel que aquesta etiqueta ens facilita la classificació en cas que després ens interessen desestimar-los.

S'ha seguit aquest procés per cada un dels eritròcits de les 300 imatges adquirides, el que correspon a més de 6.000 etiquetes, i s'hi ha sumat els eritròcits etiquetats en el treball previ de C. Pladevall.[1]

Un cop tenim l'etiquetatge complet, cal establir una relació entre les etiquetes manuals i els objectes de les imatges processades. Aquesta relació es troba gràcies a les components x i y del centroid de

les cèl·lules, de manera que cada ROI correspon a una cèl·lula concreta. En cas que una etiqueta no tingui relació amb cap de les dades dels centroides es classifica com a "Desconegut". Això pot ser degut a una mala segmentació de dues cèl·lules sobreposades.



Figura 17. Etiquetatge dels eritròcits amb l'aplicació Image Labeler

6.4 Extracció de propietats morfològiques

Obtinguda la imatge processada, segmentada i etiquetada, passem al següent pas: extracció de les característiques de cada eritròcit. Es calculen diversos descriptors morfològics que seran utilitzats per la posterior classificació automàtica dels glòbuls vermells en els tipus corresponents. Aquests descriptors es diferencien a continuació.

6.4.1 Característiques físiques

- Centroide: Retorna informació sobre on es localitza el centre de masses de l'objecte en la imatge. Retorna un vector amb dues components, la primera correspon a la posició "x" de l'objecte, i la segona a la posició "y".
- Bounding Box: Dimensions del requadre més petit que conté l'objecte al seu interior retornat en forma de vector. El vector és [cantonada amplada] on cantonada dóna les coordenades de la posició "x" i "y" de la cantonada superior esquerra del rectangle, i amplada és un altre vector amb les dimensions de l'amplada del Bounding Box en direcció "x" i les dimensions de l'amplada en direcció "y".

- Àrea: Espai que es conté dins de la membrana exterior de cada eritròcit. El valor de l'àrea es retorna amb píxels, ja que calcula el nombre de píxels continguts dins del contorn de cada objecte.
- Excentricitat: És la raó entre la distància focal d'una el·lipse (distància mitjana del centre de l'el·lipse fins a un dels seus focus) i la longitud del seu eix major. El valor de l'excentricitat sempre varia entre 0 i 1. Si l'excentricitat s'aproxima al valor 0 l'objecte tendirà a una forma més arrodonida.
- Solidesa: Proporció dels píxels a l'envolupant convexa que també són continguts a la regió. Es calcula com a raó de l'àrea de l'objecte i la seva àrea convexa, i retorna un nombre escalar. La solidesa variarà amb valors entre 0 i 1. Obtindrem solidesa igual a 0 quan la cèl·lula adopti forma convexa.

(Eq. 6.2)

$$\text{Solidesa} = \frac{\text{Àrea cèl} \cdot \text{lula}}{\text{Àrea Convexa}}$$

- Extensió: Relació dels píxels que engloba l'objecte en referència als continguts dins del Bounding Box.

(Eq. 6.3)

$$\text{Extent} = \frac{\text{Àrea cèl} \cdot \text{lula}}{\text{Àrea Bounding Box}}$$

- Perímetre: Retorna un nombre escalar que equival a la distància mesurada rodejant el contorn de cada eritròcit. És calcula a través dels píxels. El problema d'aquest paràmetre, igual que l'àrea, és que al treballar sobre el nombre de píxels de la imatge, al canviar els nivells de zoom, els valors varien.
- Major Axis Length: Valor escalar que retorna la longitud (en nombre de píxels) de l'eix major de la el·lipse que té el mateix moment central que la regió.
- Minor Axis Length: Longitud (en nombre de píxels) de l'eix menor de l'el·lipse amb el mateix moment central que la regió. Retornada en valor de píxels.
- Circularitat: És un descriptor que relaciona l'àrea cel·lular amb el perímetre. La circularitat d'un cercle equival a 1. Com més s'apropi el valor de la circularitat a 1, més circular serà la cèl·lula analitzada.

(Eq. 6.4)

$$Circularitat = 4 * \pi * \frac{\text{\textit{Àrea cèl} \cdot \text{\textit{lula}}}{\text{\textit{Perímetre cèl} \cdot \text{\textit{lula}}^2}}$$

- Elongació: Relació entre Major Axis Length i Minor Axis Length. Com més similars siguin els dos valors, i per tant, més circular sigui la cèl·lula, més elevada serà l'elongació.

(Eq. 6.5)

$$\text{\textit{Elongació}} = 1 - \frac{\text{\textit{Major Axis Length}}}{\text{\textit{Minor Axis Length}}}$$

- Rodonesa CH: Pel seu càlcul s'utilitza el perímetre del casc convex (Convex Hull) i l'àrea de l'objecte d'interès. És un paràmetre semblant a la circularitat, i igual que aquesta, té valors similars a 1.

(Eq. 6.6)

$$\text{\textit{Rodonesa CH}} = 4 * \pi * \frac{\text{\textit{Àrea cèl} \cdot \text{\textit{lula}}}{\text{\textit{Perímetre convex}}^2}$$

- Convexitat: S'obté també a partir del càlcul del perímetre convex. Estableix una relació entre el perímetre convex i el perímetre general.

(Eq. 6.7)

$$\text{\textit{Convexitat}} = \frac{\text{\textit{Perímetre convex}}}{\text{\textit{Perímetre}}}$$

- Variància circular: Similitud de l'objecte d'interès a la forma d'un cercle. Per calcular-la es relaciona la desviació típica de l'objecte amb la mitjana de la distància radial (distància del centroide als punts del contorn exterior de l'objecte que s'ajusta al cercle) segons la següent expressió:

(Eq. 6.8)

$$\text{\textit{Variància circular}} = \frac{\text{\textit{Desviació típica}}}{\text{\textit{Mitjana distància radial}}}$$

- Variància de l'el·lipse: Es calcula de la mateixa forma que la variància circular però la mitjana de la distància radial farà referència a la distància del centroide als punts exteriors on l'objecte s'ajusta a la forma d'el·lipse.

(Eq. 6.9)

$$\text{Variància de l'el·lipse} = \frac{\text{Desviació típica}}{\text{Mitjana distància punts el·lipse}}$$

Amb els resultats dels càlculs de tots els descriptors esmentats es crea una taula on es recullen totes les dades. En la figura següent es mostren algunes d'aquestes dades com a exemple.

etiquetas	Area	Eccentricity	Solidity	Extent	Perimeter	MajorAxisLength	MinorAxisLength
Esferocito	32553	0,380181227	0,995139398	0,790639497	637,278	211,7050593	195,8085528
Normal	34956	0,280420298	0,987736649	0,774116396	668,622	215,7303421	207,0746612
Drepanocits	43702	0,925214631	0,982531082	0,629938739	884,25	383,3906679	145,4753578
Drepanocits	18102	0,947652623	0,93117284	0,649375807	633,396	276,8675642	88,40468773
Drepanocits	45782	0,929690608	0,989389061	0,54582305	926,224	401,3792897	147,8446025
Normal	48116	0,696895319	0,991081176	0,763746032	795,24	292,4275849	209,721128
Normal	36617	0,294879607	0,994648775	0,784980813	677,744	220,9835087	211,1573477
Esferocito	31250	0,378874436	0,992315509	0,782502003	626,792	207,4629487	191,9961889
Drepanocits	53726	0,732693113	0,974497569	0,755639944	904,76	317,7360402	216,2381801
Esferocito	32455	0,326785996	0,994972255	0,791508146	636,062	209,1216042	197,6404854
Esferocito	35847	0,46757342	0,991371443	0,806057744	674,756	227,5576784	201,1505744
Normal	32026	0,73527162	0,991701245	0,627591613	677,4	248,5670518	168,4719273

Taula 1. Descriptors de morfologia

6.4.2 Altres característiques

Les característiques que es mostren a continuació tenen relació amb la intensitat dels píxels de la imatge. Es separa aquesta en les seves corresponents components vermella, verda i blava (RGB) i s'estableix una relació entre els píxels que formen cada ROI i les intensitats que corresponen a les seves posicions.

A partir de la imatge binària filtrada obtenim un llistat del conjunt de píxels que formen cada objecte. Es relaciona el valor d'aquests píxels amb una de les components de la imatge i es busca quina intensitat té cadascun d'aquest píxels. A partir d'aquestes dades és d'on s'extreuen les següents característiques. Cada una d'elles tindrà tres valors diferents, un valor per la imatge de la component vermella, un per la component verda i un per la blava.

- Mitjana: Partim de tres vectors amb els valors d'intensitat dels píxels, un vector per a cada component R,G i B. Per a calcular la mitjana cal tractar cada vector per separat. Es sumen totes les dades obtingudes i es divideixen entre el nombre píxels S'obté la mitjana de la intensitat dels píxels que conté cada objecte.

Eq. 6.10

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i$$

on X és el vector, i N el número d'observacions.

- Desviació estàndard: Retorna un nombre escalar que correspon a l'índex numèric de la dispersió dels valors dels píxels. Es defineix com $\sqrt{\sigma^2}$ on σ^2 és la variància de la variable. S'utilitzen els valors d'intensitat de gris dels píxels, pel que la desviació estàndard ens donarà nocions d'intensitat de color.

Eq. 6.11

$$\sigma = \sqrt{\left(\frac{1}{N-1} \sum_{i=1}^N |A_i - \mu|^2\right)}$$

on μ és la mitjana

- Asimetria estadística o Skewness: Indica el grau de simetria que presenta la distribució de probabilitat d'una variable sense necessitat d'obtenir la seva representació gràfica. Es considera el valor de la mitjana i es passa una recta paral·lela a l'eix d'ordenades en aquest punt. Si hi ha el mateix nombre de valors a la dreta que a l'esquerra de la recta a la gràfica, es considera una distribució simètrica. En cas contrari és una distribució asimètrica.

Eq. 6.12

$$S = \frac{E(x - \mu)^3}{\sigma^3}$$

On μ és la mitjana, σ és la desviació estàndard i $E(t)$ és l'esperança de t.

- Curtosis: Calcula el grau de concentració dels valors de la variable al voltant de la zona central de la mostra. Aquest valor, juntament amb l'asimetria estadística s'utilitzen per contrastar si es pot acceptar que una distribució estadística determinada segueix una distribució normal.

Eq. 6.13

$$k = \frac{E(x - \mu)^4}{\sigma^4}$$

on μ és la mitjana, σ és la desviació estàndard i $E(t)$ és l'esperança de t .

- Entropia: Es considera com a entropia el desordre d'un sistema. S'utilitzen els valors d'intensitat de gris de cada píxel i es retorna un escalar, el qual caracteritzarà la textura de la imatge.

Eq. 6.14

$$-\text{sum}(p.* \log_2(p))$$

on p conté l'histograma normalitzat retornat de la funció "imhist"

Totes aquestes característiques s'afegeixen als resultats dels descriptors de la taula anterior (Taula 1). A continuació es mostra una part de les dades dels nous descriptors extrets.

Mean_Green	Mean_Red	Mean_Blue	Std_Green	Std_Red	Std_Blue	Skewness_G
36054,71496	41537,84069	40768,10813	2871,271517	2871,271517	2871,271517	0,903984402
41022,19064	45480,50212	43248,26548	2508,539002	2508,539002	2508,539002	1,343675901
40585,27445	41572,95318	44535,91177	1717,481818	1717,481818	1717,481818	0,352915606
40826,26008	40478,58049	45528,76986	1844,437397	1844,437397	1844,437397	0,838552159
40160,63746	41221,04093	44647,0232	1865,815126	1865,815126	1865,815126	-0,055647917
41902,8704	46331,84471	45995,59498	3014,34062	3014,34062	3014,34062	0,869571377
38536,64211	44314,70366	40258,45744	2894,056812	2894,056812	2894,056812	-0,016534542
40728,5545	44281,66758	44449,61331	2680,223867	2680,223867	2680,223867	0,368539285
40098,44798	39675,11566	45286,59733	2283,950898	2283,950898	2283,950898	1,18630618
36787,23303	42653,24369	41614,6566	3159,834183	3159,834183	3159,834183	1,639318774
38599,60432	43150,12671	41473,49033	2242,052036	2242,052036	2242,052036	0,866697303
43517,56598	48560,50309	46903,68526	3052,318798	3052,318798	3052,318798	1,087533951
42347,70092	45325,18239	43132,83305	2574,050438	2574,050438	2574,050438	0,041765318

Taula 2. Descriptors de color

CAPÍTOL 7:

ANÀLISI ESTADÍSTICA

La finalitat de l'anàlisi estadística és determinar quins són els descriptors extrets que millor classifiquen els eritròcits en la categoria que els correspon. Primerament, es duen a terme els testos de normalitat i homoscedasticitat per a cada un dels descriptors per tal de saber si presenten una distribució normal i si hi ha homogeneïtat en la variància. Seguidament s'implementen tres mètodes diferents de classificació. Cada mètode ens donarà informació sobre l'ordre en el qual els descriptors presenten més diferències en les dades entre cèl·lules. Per acabar s'analitzen amb més profunditat els 10 descriptors millor classificats segons el mètode mRMR., el qual utilitza teoria de la informació amb un mètode de màxima rellevància i mínima redundància.

Obtenim més de 9000 dades amb diferents tipus d'eritròcits. Seleccionem els drepanòcits, els esfèrics i els normals i desestimem els restants. Per a obtenir un correcte ordre de classificació és recomanable igualar el nombre d'eritròcits de cada tipus. Disposem de 552 drepanòcits i més quantitat dels altres dos tipus de cèl·lules, per tant aquest seran els que ens limitaran el nombre de dades a analitzar. Seleccionem de manera aleatòria 552 dades de cada un dels tres tipus i les ajuntem en una matriu. Els testos que s'expliquen a continuació es duen a terme amb aquestes dades.

7.1 Test de normalitat i d'homoscedasticitat

El **test de normalitat** s'utilitza per determinar si un conjunt de dades segueixen o no una distribució normal. Les distribucions normals es caracteritzen per representar les dades gràficament en forma de campana, amb la major probabilitat que les dades tinguin el valor del mig d'aquesta i menor probabilitat conforme s'allunyen de la zona central. Les distribucions normals es poden descriure amb la mitjana i la desviació estàndard, ja que dues distribucions normals amb la mateixa mitjana i desviació estàndard seran exactament de la mateixa forma. [22]

El test parteix d'una hipòtesi nul·la que afirma que el conjunt de dades formen una distribució normal. Es calcula el p-valor amb un valor de significança del 5%. Quan el p-valor sigui inferior a 0.05, el valor de h serà 1, el que significa que per aquell conjunt de dades es rebutja la hipòtesi nul·la al

valor de significança del 5%. En canvi, quan $h=0$, no podem afirmar que no formin una distribució normal.

Per a dur a terme aquests càlculs s'utilitza el test Chi-square Goodness-of-fit.

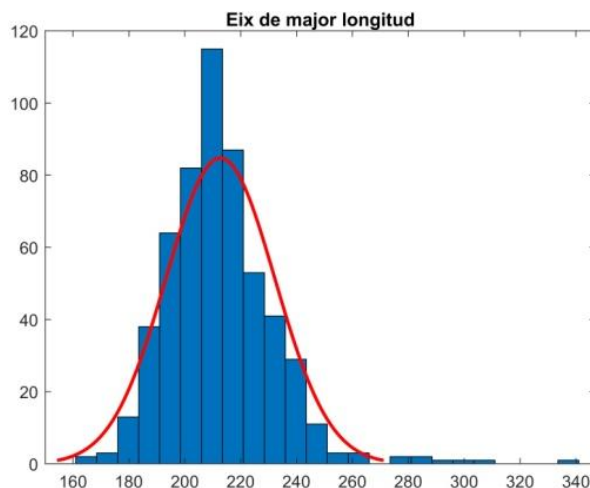


Figura 18. Gràfica d'una distribució normal

El **test d'homoscedasticitat** o d'homogeneïtat de variàncies pretén determinar si diferents grups de dades presenten una distribució normal i tenen la mateixa variància, es cas afirmatiu els classificarem com a homoscedàstics. Parteix de la hipòtesi nul·la que afirma que la variància és igual entre els diferents grups analitzats, el que implica que la desviació estàndard tampoc variarà entre ells. Es calcula el p-valor, si aquest és baix o igual a 0 indica que la funció rebutja la hipòtesi nul·la i dóna suport a la afirmació que com a mínim una columna té una variància diferent a les altres [22].

Utilitzem el test Bartlett per aconseguir els resultats.

Resultats del test de normalitat

S'ha realitzat el test de normalitat a les dades de cada un dels descriptors per a cada tipus de cèl·lula. Pel **esferòcits** s'ha trobat que per tots els descriptors excepte dos s'ha de rebutjar la hipòtesi nul·la que les dades provenen d'una distribució normal. En canvi, els resultats obtinguts pels descriptors d'eix de menor longitud i convexitat indiquen que no es pot rebutjar la hipòtesi nul·la, pel que les dades poden provenir d'una distribució normal.

Pel que fa als **drepanòcits**, no es pot rebutjar la hipòtesi nul·la pels següents descriptors: solidesa, eix de major longitud, circularitat, rodonesa CH, convexitat, mitjana component verda i component

blava, asimetria estadística component blava, curtosis component verda i component vermella, entropia component verda i component blava.

Referent als eritròcits etiquetats com a **normals**, es pot rebutjar la hipòtesi nul·la en tots els descriptors excepte per l'eix de menor longitud, la convexitat, l'asimetria estadística de la component blava, la curtosis component verda i de la vermella i l'entropia de la component vermella i de la blava. Per aquest anomenats les dades formarien una distribució normal.

Resultats test d'homoscedasticitat

El test d'homoscedasticitat ens permet treballar amb els tres grups de cèl·lules alhora. Pels següents descriptors no podem rebutjar la hipòtesi nul·la, pel que les dades poden provenir d'una distribució normal amb la mateixa variància: la mitjana de la component blava, la desviació estàndard de les tres components de color i l'asimetria estadística de la component verda.

Anàlisi dels resultats

Veiem que hi ha variació en els resultats entre descriptors, el que indica que no tots segueixen el mateix tipus de distribució. Per a analitzar cada descriptor individualment caldria realitzar diferents testos en funció dels resultats de normalitat i homoscedasticitat obtinguts per a cada descriptor.

En el nostre cas, per tal de simplificar el procediment, aplicarem el test **Kruskal-Wallis** per a tots els descriptors per obtenir una estimació de la comparació entre els tres grups de cèl·lules per a cada descriptor.

Aquest és un test no paramètric, el que significa que no assumeix que les dades provenen d'una distribució que pot ser descrita per la mitjana i la desviació estàndard, és a dir, no assumeix que les dades provenen d'una distribució normal [22].

A més a més, aquest test assumeix que les dades són homoscedastàtiques, és a dir, que els diferents grups tenen la mateixa variància. Aquesta condició no es compleix per totes les dades que analitzarem. Tot i així, segueix essent un bon mètode de comparació i l'utilitzarem tenint en compte la manca de fiabilitat total que representarà per alguns descriptors.

El mètode de comparació interactiva que utilitzem representa en una gràfica les medianes de les dades del descriptor, de manera que s'aprecia la diferència o similitud entre elles.

7.2 Estudi de rellevància dels descriptors

Aquest apartat pretén classificar els descriptors analitzant-los segons la importància que tenen en la correcta diferenciació dels eritròcits. Per fer-ho s'implementen diferents mètodes i es comparen:

- Mètode Relieff: mètode de classificació que no té en compte la redundància entre descriptors. Es tracta d'un algoritme de selecció aleatòria d'instàncies per al càlcul del pes de característiques. L'algoritme adopta la selecció aleatòria de les instàncies segons l'estimació del pes. S'adapta bé tractant mostres amb soroll i és robust a les interaccions de característiques [23].
- Mètode Neighborhood Component Analysis (NCA): és un mètode de classificació no lineal i no paramètric en el qual la qualitat millora a mesura que augmenta la quantitat de dades a analitzar [24].
- Mètode Minimum Redundancy Maximum Relevance (mRMR): Els descriptors poden ser classificats i ordenats segons el paper que tenen en la correcta diferenciació de cèl·lules, aquest procés s'anomena màxima rellevància. Aquests descriptors contenen informació que és rellevant però alhora redundant, per això el mètode en qüestió intenta resoldre aquest problema eliminant aquells que són redundants. En la classificació final obtindrem en primera posició aquell mètode que diferencia millor les cèl·lules i alhora no presenta redundància amb cap altre. En canvi, en darrera posició trobarem aquell mètode que ens dona menys informació sobre la classificació, ja que és el que presenta menys rellevància i més redundància [25].

A continuació es mostra una taula amb els descriptors diferenciats en mètodes i ordenats de major a menor segons la funció que duen a terme en la classificació:

Mètode Relief	Mètode NCA	Mètode mRMR
Elongació	Asimetria estadística Verd	Eix de major longitud
Excentricitat	Mitjana Verd	Mitjana Verd
Variància circular	Excentricitat	Excentricitat
Eix de major longitud	Àrea	Circularitat
Rodonesa CH	Mitjana Blau	Mitjana Vermell
Circularitat	Mitjana Vermell	Eix de menor longitud
Perímetre	Eix de menor longitud	Asimetria estadística Blau
Mitjana Vermell	Asimetria estadística Vermell	Variància circular
Àrea	Perímetre	Perímetre
Eix de menor longitud	Variància circular	Mitjana Blau
Mitjana Verd	Elongació	Extensió
Mitjana Blau	Entropia Blau	Rodonesa CH
Extensió	Entropia Verd	Àrea
Asimetria estadística Blau	Desviació estàndard Blau	Elongació
Entropia Vermell	Eix de major longitud	Curtosis Vermell
Desviació estàndard Vermell	Desviació estàndard Verd	Entropia Vermell
Desviació estàndard Verd	Convexitat	Solidesa
Entropia Blau	Circularitat	Variància el·líptica
Curtosis Blau	Desviació estàndard Vermell	Entropia Blau
Entropia Verd	Entropia Vermell	Curtois Verd
Desviació estàndard Blau	Rodones CH	Convexitat
Asimetria estadística Vermell	Asimetria estadística Blau	Asimetria estadística Vermell
Asimetria estadística Verd	Curtosis Verd	Curtosis Blau
Solidesa	Solidesa	Desviació estàndard Vermell
Variància el·líptica	Variància el·líptica	Asimetria estadística Verd
Curtosis Verd	Extensió	Desviació estàndard Blau
Convexitat	Curtosis Blau	Entropia Verd
Curtosis Vermell	Curtosis Vermell	Desviació estàndard Verd

Taula 3. Comparació entre diferents mètodes de l'estudi de rellevància

La taula mostra com cada mètode classifica de forma diferent els descriptors. Es continua el treball segons l'ordre de descriptors donat pel mètode mRMR, ja que es considera que té més validesa al optimitzar els resultats interactuant amb la mínima redundància i màxima rellevància dels descriptors.

7.3 Anàlisi dels descriptors

S'analitzen en detall els 10 descriptors millor classificats segons el mètode mRMR. Els resultats estadístics dels descriptors restants s'afegeixen a l'Annex.

7.3.1 Eix de major longitud

L'eix de major longitud és el descriptor que presenta una major rellevància i una menor redundància dels 28 descriptors trobats. Com s'ha explicat anteriorment, els drepanòcits es caracteritzen per tenir una forma semilunar, pel que el seu eix de major longitud tindrà un valor més elevat que en el cas dels esferòcits i els eritròcits normals. En el gràfic de caixes s'aprecia la diferència. Pel que fa als esferòcits, són més circulars i petits que els eritròcits normals, i per tant, tal i com es mostra en el gràfic, seran els que tindran l'eix de major longitud amb valors més baixos.

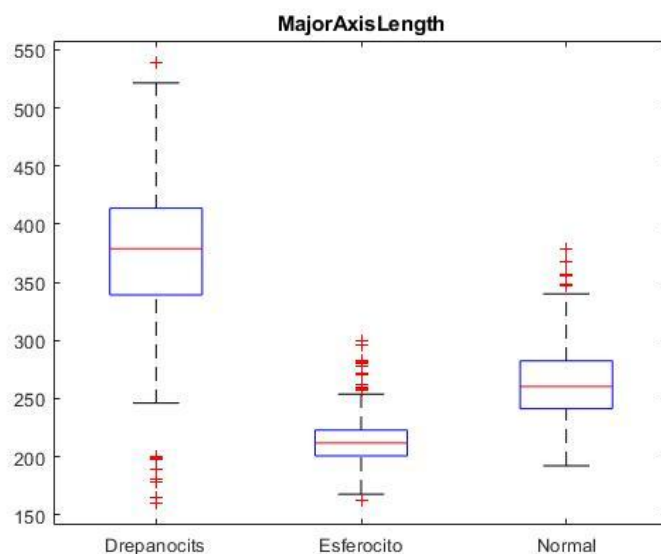


Figura 19. Gràfic de caixes de l'eix de major longitud

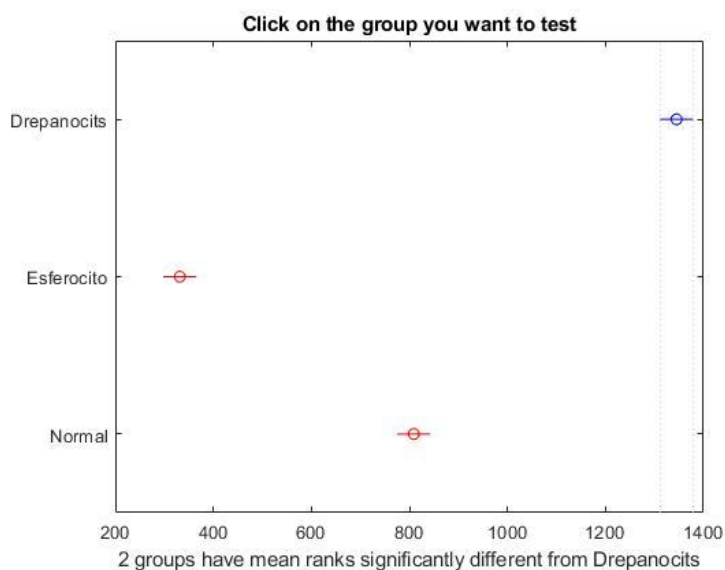


Figura 20. Gràfica de comparació múltiple amb Kruskal-wallis de l'eix de major longitud

El gràfic interactiu de comparació múltiple mostra les posicions de les medianes (amb una rodona) de cada tipus d'eritròcit pel descriptor de l'eix de major longitud. Les medianes entre dues classes seran significativament diferents quan la seva representació en el gràfic no es solapi. En aquest cas veiem com els valors de les medianes per cada eritròcit són significativament diferents.

En la Taula 4 es mostren els percentils per a cada tipus d'eritròcit on P₂₅ representa el primer quartil al 25%, P₅₀ el segon quartil al 50% i P₇₅ el tercer quartil al 75%.

Percentils eix de major longitud	Esferòcits	Drepanòcits	RBC normals
P ₂₅	200.7690	339.1218	241.3068
P ₅₀	211.7714	379.0040	260.3774
P ₇₅	222.9511	413.7470	282.4679

Taula 4. Percentils de l'eix de major longitud

La multicomparació entre els eritròcits amb el test Kruskal-Wallis ens permet calcular el p-valor de cada comparació. Com s'ha explicat anteriorment, si el p-valor és superior al valor de significança (en aquest cas del 5%), indicarà que no es pot rebutjar la hipòtesi nul·la. En cas contrari, si el p-valor és inferior al 5% de significança, indicarà que podem rebutjar la hipòtesi nul·la que les tres mostres provenen de la mateixa distribució. Com més baix sigui aquest valor, més forta serà l'evidència contra la hipòtesi nul·la.

Eix de major longitud	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.7459 \cdot 10^{-10}$	<0.05	✓

Taula 5. Estudi de la significança de l'eix de major longitud

Els p-valors obtinguts per cada una de les comparacions són molt propers a 0. Això dona una evidència major al fet que aquest descriptor hagi estat seleccionat com el de major rellevància i menor redundància, ja que rebutja la hipòtesi nul·la. Els resultats del p-valor tant inferiors al valor de significança afirma que les dades provenen de diferents tipus d'eritròcits.

7.3.2 Mitjana component verda

En aquest cas no és tant evident la diferència entre els tres tipus de cèl·lules al gràfic de caixes. Tot i això, el valor de la mitjana de la intensitat dels píxels en la component verda és el segon descriptor en la classificació de menor redundància i major rellevància. Les cèl·lules que es diferencien més són els eritròcits normals, amb una mitjana d'intensitat dels píxels major que els altres dos tipus, les mitjanes dels quals són bastant similars.

En canvi, en el gràfic de comparació múltiple s'aprecia perfectament la diferència entre els valors de les medians dels tres tipus de cèl·lules. Igualment, són els eritròcits normals els que prenen el valor major.

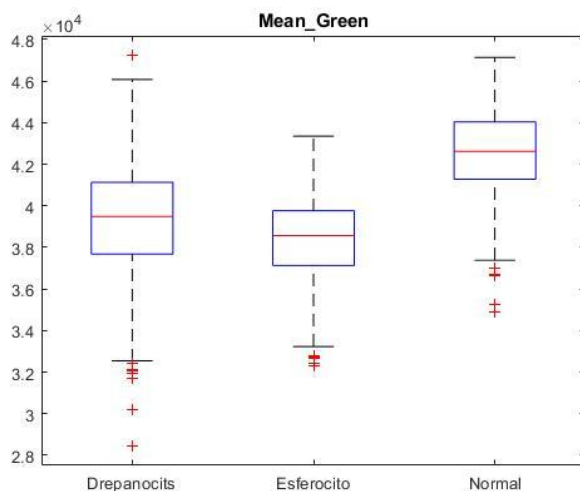


Figura 21. Gràfic de caixes de la mitjana verd

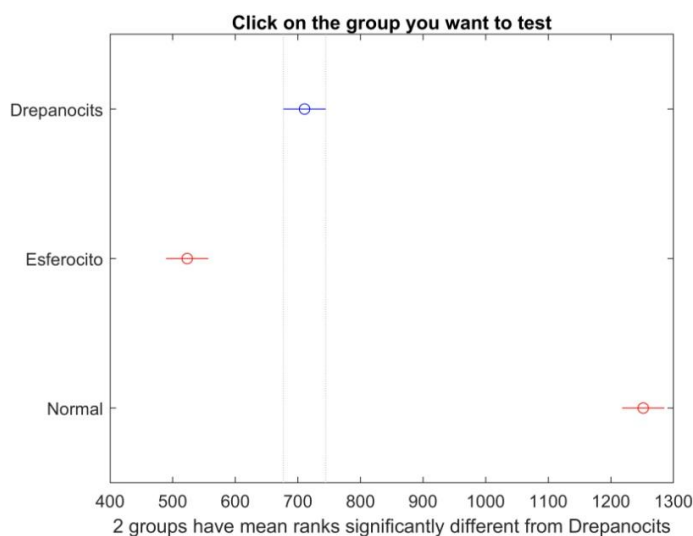


Figura 22. Gràfica de comparació múltiple amb Kruskal-wallis de la mitjana verd

A la taula següent es mostren els percentils de la mitjana d'intensitat de píxels de la component verda per a cada tipus de cèl·lula:

Mitjana_Green	Esferòcits	Drepanòcits	RBC normals
P ₂₅	3.7134*10 ⁴	3.7683*10 ⁴	4.1281*10 ⁴
P ₅₀	3.8571*10 ⁴	3.9489*10 ⁴	4.2620*10 ⁴
P ₇₅	3.9764*10 ⁴	4.1132*10 ⁴	4.4038*10 ⁴

Taula 6. Valor dels percentils de la mitjana en la component verda

Els valors de p obtinguts de la comparació entre tipus de cèl·lules ens indiquen que les seves dades no tenen valors similars pel que podem afirmar que provenen de mostres diferents. Per tant la mitjana de la intensitat dels píxels en la component verda serà un descriptor capaç de diferenciar entre els tres grups d'eritròcits.

Mitjana Verd	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	2.0369*10 ⁻⁹	<0.05	✓
Drepanòcits - Normals	9.5606*10 ⁻¹⁰	<0.05	✓
Esferòcits - Normals	9.5606*10 ⁻¹⁰	<0.05	✓

Taula 7. Estudi de la significança segons la mitjana de la component verda

7.3.3 Excentricitat

L'excentricitat ocupa la tercera posició en la llista de mRMR. Com s'ha explicat anteriorment, aquesta varia entre 0 i 1, sent 0 en circumferències i 1 pel formes diferents. Analitzant les dades observem que pels drepanòcits l'excentricitat pren valor propers a 1, això és degut a la seva forma semilunar, la qual s'allunya de la forma d'una circumferència perfecta. Als esferòcits els corresponen els valors més baixos, inferiors a 0.5, ja que tenen una forma esferoïdal, com el seu nom indica. Per últim, referent als eritròcits normals, aquests es troben en valors entre 0.55 i 0.35, allunyant-se dels drepanòcits però propers als valors dels esferòcits.

En la gràfica de comparació es plasma aquesta diferència de valors en els drepanòcits, ja que la posició on es troben dista de la dels esferòcits i dels eritròcits normals.

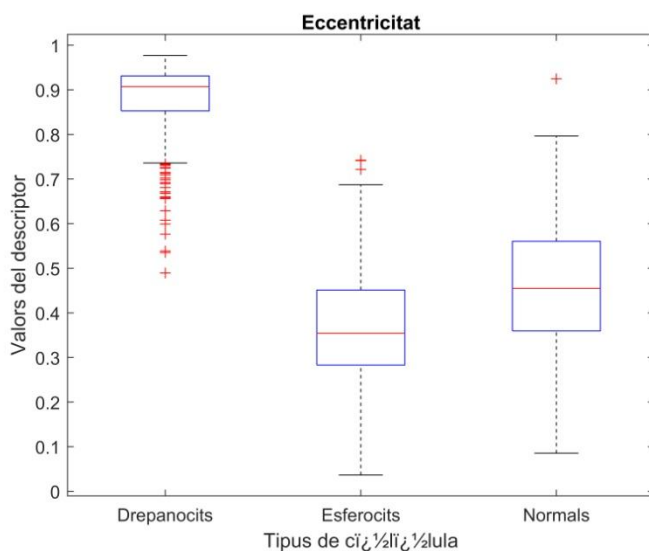


Figura 23. Gràfic de caixes de l'excentricitat

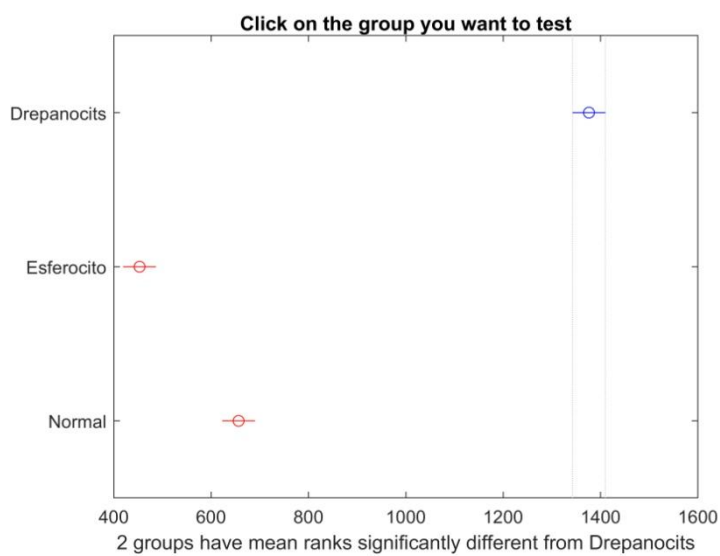


Figura 24. Gràfica de comparació múltiple amb Kruskal-wallis de l'excentricitat

A la següent taula es mostren els percentils per cada grup de cèl·lules:

Percentils excentricitat	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.2821	0.8527	0.3592
P ₅₀	0.3601	0.9072	0.4473
P ₇₅	0.4409	0.9312	0.5493

Taula 8. Valor dels percentils de l'excentricitat

En la taula de valors de p, podem corroborar les dades que ens donava la gràfica de comparació. Veiem que els valors que assoleix p al comparar amb els drepanòcits són molt baixos, en canvi al fer la comparació entre esferòcits i eritròcits normals, el valor és més alt. Tot i ser un a valor més alt, segueix sent inferior al valor de significança pel que podem afirmar que les dades dels tres tipus de mostres es diferencien correctament.

Excentricitat	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$1.0683 \cdot 10^{-9}$	<0.05	✓

Taula 9. Estudi de la significança segons l'excentricitat

7.3.4 Circularitat

El descriptor de circularitat ocupa la quarta posició segons l'ordre mRMR. La circularitat, igual que l'excentricitat, oscil·la en valors entre 0 i 1. En aquest cas, com més proper a 1 sigui el valor, més s'aproxima l'objecte a la forma d'un cercle. Altra vegada veiem una clara diferenciació dels drepanòcits respecte els altres dos tipus de cèl·lules, degut a la seva forma més allargada. Igual que en l'excentricitat, els esferòcits són els que tenen assignats els valors que indiquen la tendència a la forma circular.

En la gràfica de comparació múltiple s'observen uns valors de mediana significativament diferents pels drepanòcits i més propers entre els esferòcits i els eritròcits normals.

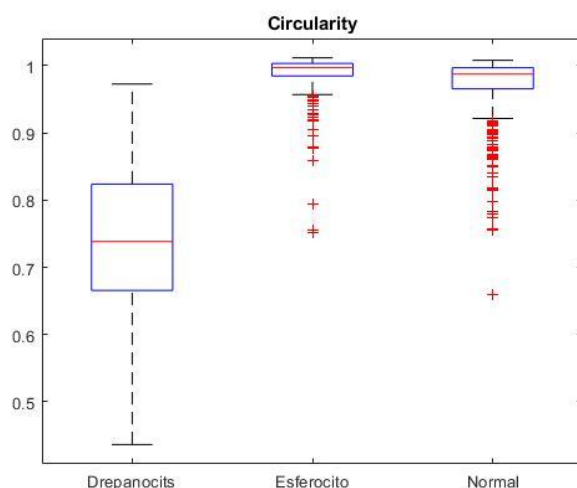


Figura 25. Gràfic de caixes de la circularitat

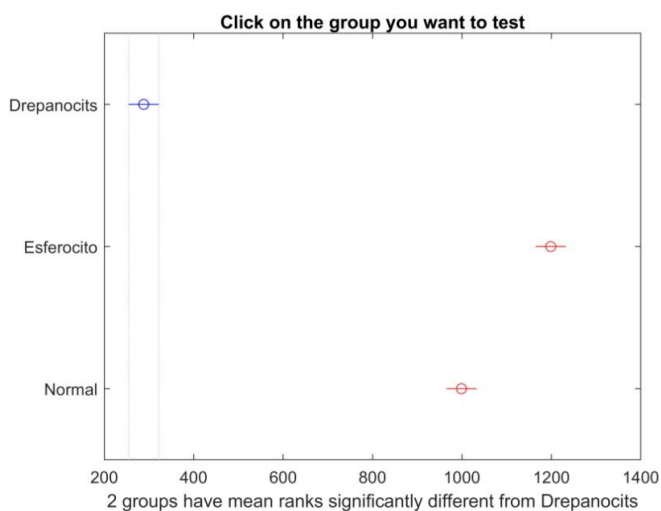


Figura 26. Gràfica de comparació múltiple amb Kruskal-wallis de la circularitat

Seguidament es mostra la taula de percentils corresponent a la circularitat per cada tipus de cèl·lula:

Percentils circularitat	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.9848	0.6656	0.9655
P ₅₀	0.9973	0.7382	0.9875
P ₇₅	1.0033	0.8238	0.9970

Taula 10. Valor dels percentils de la circularitat

En la taula de l'estudi de significança segons la circularitat es comparen els p-valors entre grups de cèl·lules. El p-valor entre drepanòcits i normals i drepanocits i esferòcits és el mateix i és significativament baix, el que indica una bona diferenciació d'aquest tipus de cèl·lules entre elles. Pel

que fa als esferòcits i als eritròcits normals, com em vist en la gràfica de caixes, aquest prenen valors que podrien considerar-se similars, però el p-valor segueix sent inferior al valor de significança.

Circularitat	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$1.0265 \cdot 10^{-9}$	<0.05	✓

Taula 11. Estudi de la significança segons la circularitat

Tots els p-valors són inferiors al valor ed significança pel que es considera que la circularitat és un descriptor útil per a la diferenciació dels tres tipus d'eritròcits.

7.3.5 Mitjana component vermella

Seguint l'ordre, la mitjana d'intensitat dels píxels de la component vermella serà el cinquè descriptor a analitzar.

El gràfic de caixes mostra com cada tipus d'eritròcit pren un valor mitjà diferent. S'observa que els eritròcits normals són els que engloben un rang de valors més gran, el que significa que les intensitat dels píxels que componen aquest tipus de cèl·lules varien més que en les altres. En canvi, els esferòcits tenen un interval inferior de valors d'intensitat, pel que trobarem valors d'intensitat més similars en tots ells.

Observant el gràfic de comparació múltiple es pot anticipar que els tres tipus d'eritròcits es diferenciaran correctament. Per corroborar-ho caldrà calcular els p-valors.

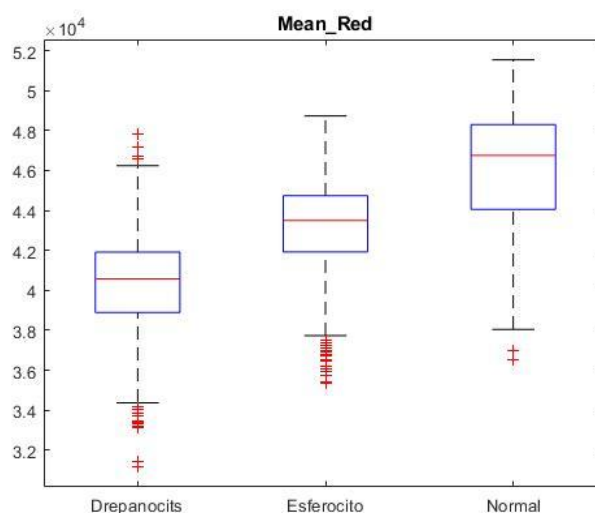


Figura 27. Gràfic de caixes de la mitjana vermella

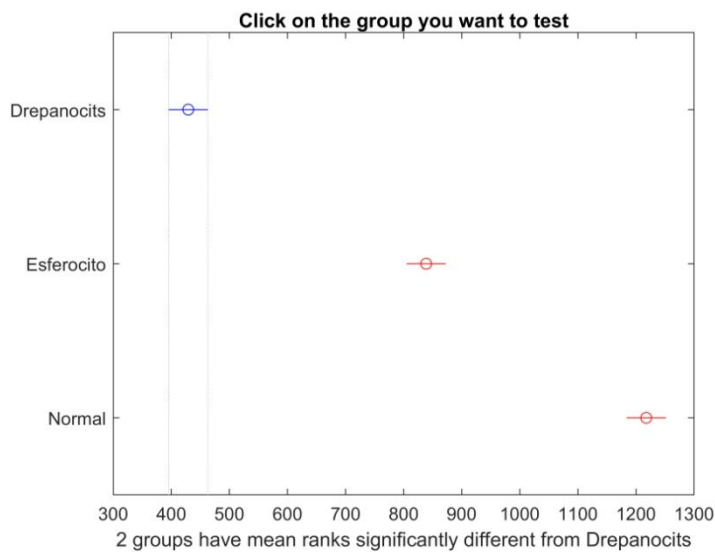


Figura 28. Gràfica de comparació múltiple amb Kruskal-wallis de la mitjana vermell

A continuació es mostra la taula de percentils de la mitjana d'intensitats de la component vermella per a cada tipus de cèl·lula.

Percentils mitjana vermell	Esferòcits	Drepanòcits	RBC normals
P ₂₅	4.1922*10 ⁴	3.8896*10 ⁴	4.4054*10 ⁴
P ₅₀	4.3500*10 ⁴	4.0574*10 ⁴	4.6755*10 ⁴
P ₇₅	4.4737*10 ⁴	4.1918*10 ⁴	4.8297*10 ⁴

Taula 12. Valor dels percentils per la mitjana d'intensitats de la component vermella

En la taula Taula 13 es mostra com els tres valors de p obtinguts són molt propers a 0 i inferiors al valor de significança del 5%, pel que el descriptor de la mitjana d'intensitat de la component vermella permetrà una bona classificació de les cèl·lules.

Mitjana vermell	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	9.5606*10 ⁻¹⁰	<0.05	✓
Drepanòcits - Normals	9.5606*10 ⁻¹⁰	<0.05	✓
Esferòcits - Normals	9.5606*10 ⁻¹⁰	<0.05	✓

Taula 13. Estudi de la significança segons la mitjana de la intensitat de la component vermella.

7.3.6 Eix de menor longitud

L'eix de menor longitud pren el 6è valor en l'escala de descriptors, a diferència de l'eix de major longitud al qual se li assigna la primera posició. Amb aquest descriptor es torna a donar importància a l'avaluació de la forma de la cèl·lula quan es tracten els drepanòcits.

Com s'ha comentat anteriorment, els drepanòcits tenen forma allargada, implicant un eix de major longitud més gran que els altres tipus de cèl·lules, però alhora un eix de menor longitud inferior en valor als altres eritròcits.

La conclusió pels esferòcits i eritròcits normals és molt similar a la comentada en l'anàlisi de l'eix de major longitud. Ambdós tenen forma circular, pel que els dos eixos seran molt similars entre ells, amb la diferència que els esferòcits tendeixen a tenir unes dimensions inferiors a les dels eritròcits normals.

Totes aquestes afirmacions es poden apreciar en el gràfic de caixes que es mostra a continuació. Altra vegada trobem una gràfica de comparació múltiple amb valors que s'allunyen molt de sobreposar-se.

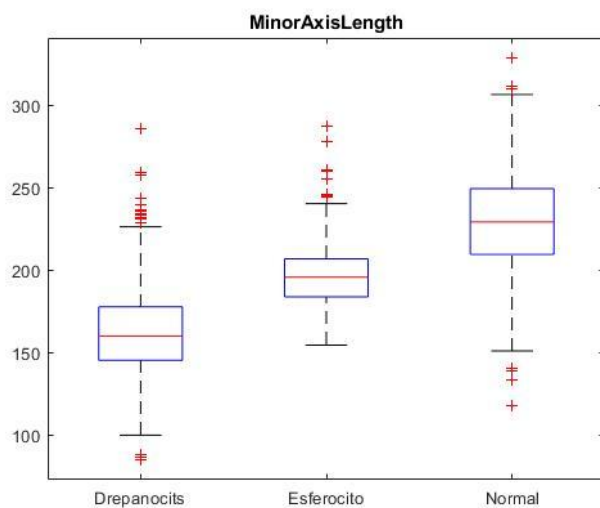


Figura 29. Gràfic de caixes de l'eix de menor longitud

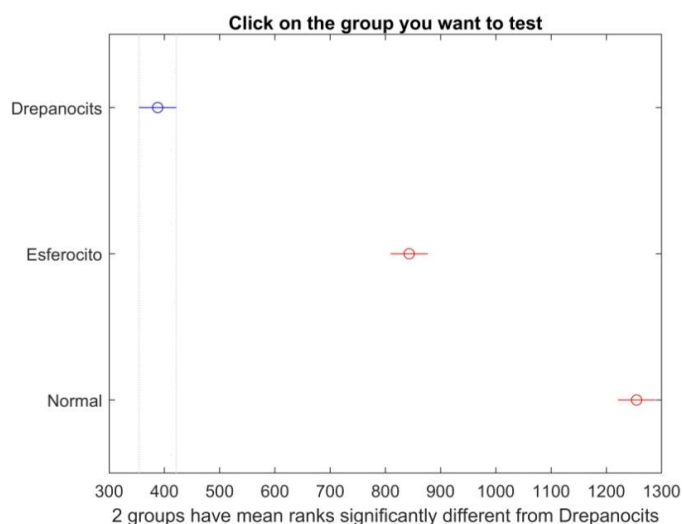


Figura 30. Gràfica de comparació múltiple amb Kruskal-wallis de l'eix de menor longitud

En la taula següent es calculen els percentils per l'eix de menor longitud per a cada tipus d'eritròcit.

Percentils eix de menor longitud	Esferòcits	Drepanòcits	RBC normals
P ₂₅	183.9259	145.4570	209.6318
P ₅₀	195.8586	160.2464	229.2841
P ₇₅	206.9684	177.8759	249.4756

Taula 14. Valor dels percentils per l'eix de menor longitud

Analitzant la taula següent, afirmem que l'eix de menor longitud és un descriptor capaç de diferenciar els tres tipus d'eritròcits segons els p-valors calculats, ja que tots tenen un valor proper a 0 i són inferiors al valor de significança.

Eix de menor longitud	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 15. Estudi de la significança segons l'eix de menor longitud

7.3.7 Asimetria estadística component blava

En el 7è descriptor de la classificació ens trobem en la mateixa situació que en la mitjana d'intensitat de la component verda, on la diferenciació de les cèl·lules segons els valors representats en el gràfic de caixes no és tant evident com en altres descriptors. Segons el gràfic es pot interpretar que la simetria que presenten les distribucions de probabilitat de cada una de les cèl·lules prenen valors similars.

Tot i això, la gràfica de comparació múltiple mostra clarament com les medianes dels valors de cada eritròcit són significativament diferents.

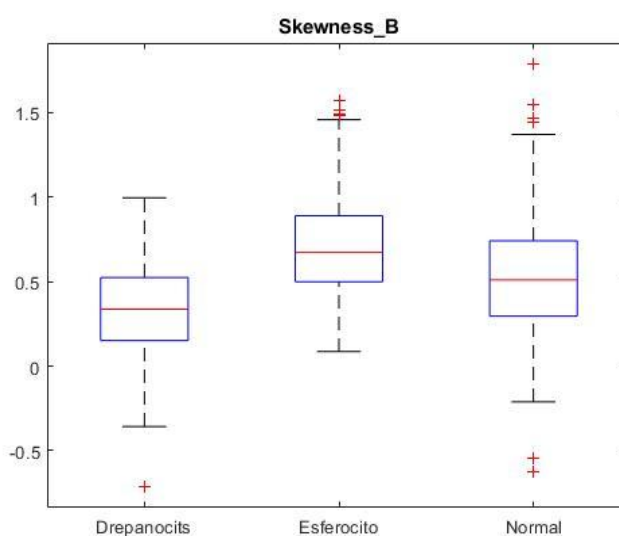


Figura 31. Gràfic de caixes de l'asimetria estadística blava

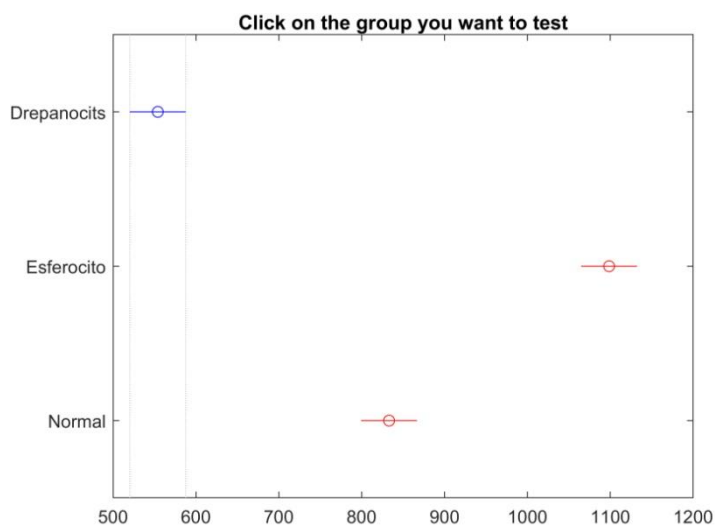


Figura 32. Gràfica de comparació múltiple amb Kruskal-wallis de l'asimetria estadística blava

En la següent taula es mostra el valor dels percentils per a l'asimetria estadística de la component blava.

Percentils asimetria estadística blava	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.4997	0.1517	0.2956
P ₅₀	0.6723	0.3373	0.5101
P ₇₅	0.8900	0.5248	0.7414

Taula 16. Valor dels percentils de l'asimetria estadística

Com es mostra en la taula següent, els tres valors de p són molt inferiors al valor de significança, pel que es pot concloure que l'asimetria estadística de la component blava és un descriptor capaç de diferenciar els tres tipus d'eritròcits.

Asimetria estadística blava	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 17. Estudi de la significança segons l'asimetria estadística de la component blava.

7.3.8 Variància circular

La variància ocupa la vuitena posició en l'ordre de mínima redundància i màxima rellevància dels descriptors. Amb la simple observació del gràfic de caixes es pot suposar que la diferenciació dels drepanòcits serà molt positiva, mentre que la diferenciació entre esferòcits i eritròcits normals podria portar més dificultats.

Quan es tracta de diferenciació de drepanòcits, els descriptors de forma són els més rellevants. Tant els esferòcits com els eritròcits normals s'aproximen molt a la forma d'un cercle, pel que es situen en valors propers a 0. En canvi els drepanòcits s'allunyen del valor 0.

Una altra observació a remarcar és l'interval de valors que representa cada eritròcit. Els drepanòcits representen en un interval major que les altres cèl·lules. Aquest tenen una forma semilunar com s'ha comentat, però poden tenir dimensions diferents entre ells. És per això, que al buscar la relació de la seva forma amb la d'un cercle, es poden obtenir resultats variants.

En la gràfica de la comparació múltiple s'observa com els drepanòcits són els que tindran una millor diferenciació.

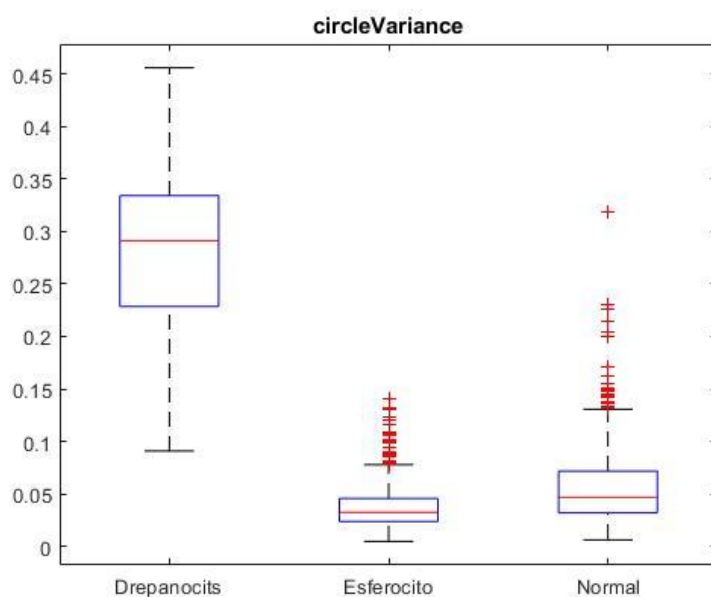


Figura 33. Gràfic de caixes de la variància circular

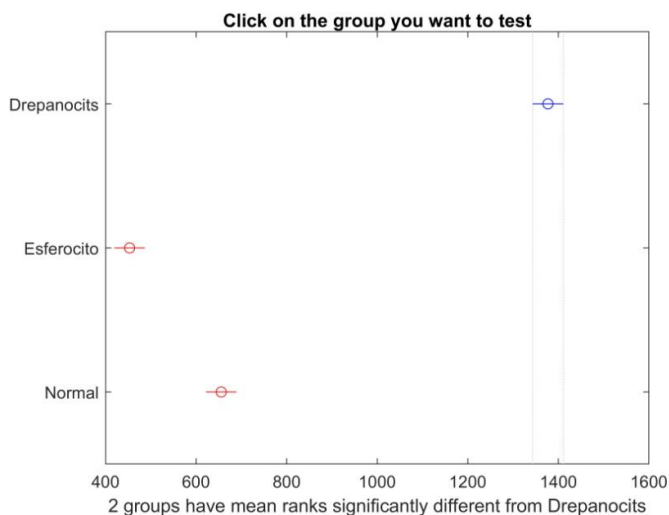


Figura 34. Gràfica de comparació múltiple amb Kruskal-wallis de la variància circular

En la taula següent els mostren els percentils per a la variància circular.

Percentils variància circular	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.0238	0.2284	0.0320
P ₅₀	0.0325	0.2911	0.0469
P ₇₅	0.0458	0.3340	0.0718

Taula 18. Valors dels percentils de la variància circular

El valor de p de la comparació entre els esferòcits i els eritròcits normals és major que els altres dos valors. Aquest p-valor major és el que es representa a la gràfica de comparació situant aquets dos tipus de cèl·lules més propers entre ells que respecte els drepanòcits. Tot i així, tots els p-valors són inferiors al valor de significança del 5%.

Variància Circular	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$1.0592 \cdot 10^{-9}$	<0.05	✓

Taula 19. Estudi de la significança segons la variància circular

L'anàlisi dels p-valors ens permet afirmar que amb el descriptor de la variància circular serà possible la diferenciació dels tres tipus d'eritròcits.

7.3.9 Perímetre

En el novè lloc trobem el perímetre, aquest és un descriptor de la forma de la cèl·lula. Els esferòcits són les cèl·lules més petites dels tres tipus, pel que són les que trobem en els valors inferiors del gràfic de caixes. Els drepanòcits, al disposar d'una forma allargada, tenen un perímetre major que els eritròcits normals, els quals es queden situats amb valors de perímetre entre els dels drepanòcits i els dels esferòcits.

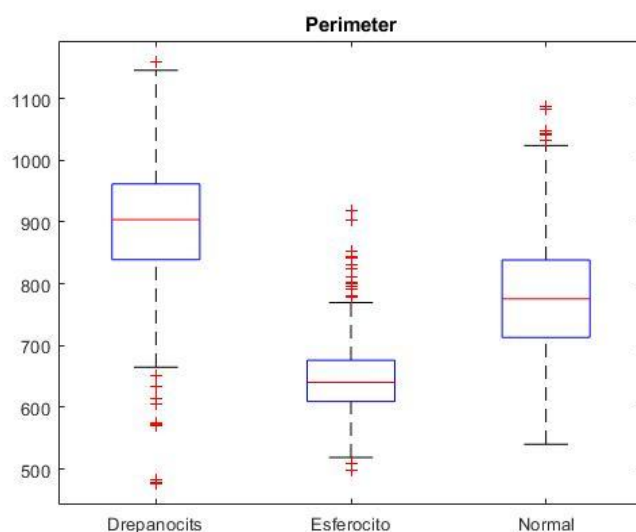


Figura 35. Gràfic de caixes del perímetre

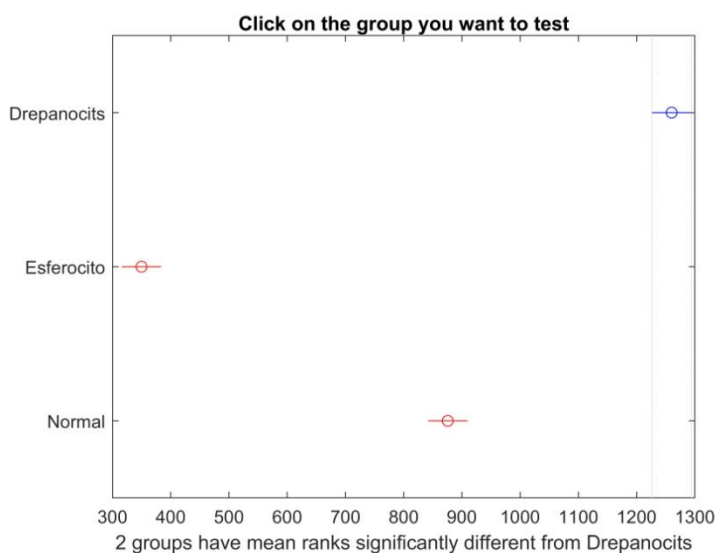


Figura 36. Gràfica de comparació múltiple amb Kruskal-wallis del perímetre

En la taula següent es mostren els valors dels percentils per aquest descriptor:

Percentils perímetre	Esferòcits	Drepanòcits	RBC normals
P₂₅	609.4315	539.0530	713.1940
P₅₀	640.4530	503.9745	775.5040
P₇₅	676.3135	561.6960	538.6380

Taula 20. Valors dels percentils pel perímetre

Els tres p-valors iguals i inferiors al valor de significança indiquen una perfecta diferenciació entre els tres grups de cèl·lules.

Perímetre	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 21. Estudi de la significança segons el perímetre

7.3.10 Mitjana component blava

En desè lloc de la llista de 28 descriptors hi trobem el càlcul de la mitjana de la intensitat de píxels de la component blava.

Per aquest descriptor prenen valors més alts els eritròcits normals i valors més baixos els esferòcits. Els intervals de valors dels tres tipus de cèl·lules són bastant similars entre ells.

En la gràfica de comparació múltiple s'observa com són també els eritròcits normals els que tenen un valor de mediana superior. Cap dels tres tipus es solapa sinó que disten considerablement.

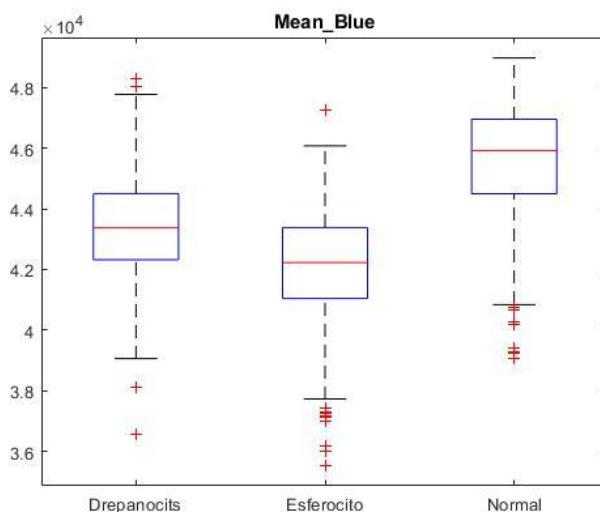


Figura 37. Gràfic de caixes de la mitjana blau

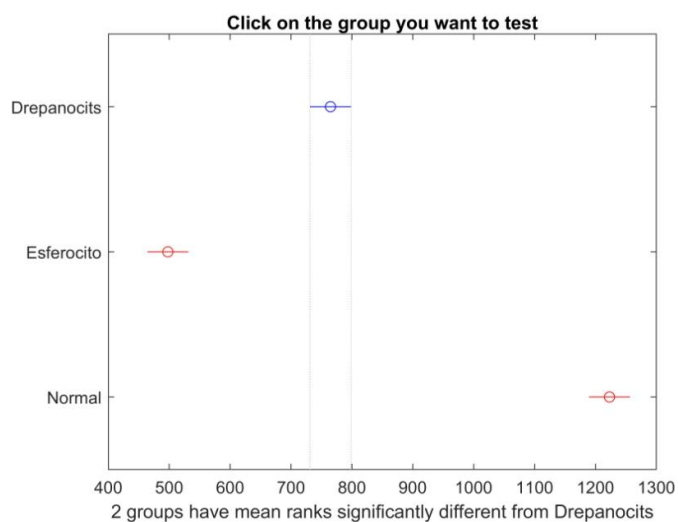


Figura 38. Gràfica de comparació múltiple amb Kruskal-wallis de la mitjana blau

A continuació es calculen els percentils per al valor de la mitjana d'intensitat de la component blava per a cada tipus de cèl·lula.

Percentils mitjana blava	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.0406	0.4777	0.0667
P ₅₀	0.0671	0.5793	0.1056
P ₇₅	0.1025	0.6354	0.1646

Taula 22. Valor dels percentils de la mitjana d'intensitat de la component blava

En la següent taula es mostren els p-valors per la comparació segons la mitjana de la component blava dels eritròcits.

Els tres p-valors prenen valors molt inferiors al valor de significança del 5%.

Mitjana blava	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 23. Estudi de la significança segons la mitjana de la component blava

Es pot afirmar que l'anàlisi de la mitjana d'intensitat de la component blava és un descriptor capaç de diferenciar els tres tipus de cèl·lules.

La resta de descriptors que no han estat detallats en aquest apartat es poden trobar el l'Annex.

7.4 Discussió dels resultats estadístics

Un dels objectius del treball és aconseguir una correcta extracció de descriptors, els quals siguin útils per diferenciar entre els tres tipus d'eritròcits sobre els que treballem.

Podem afirmar que els 10 descriptors analitzats compleixen aquest objectiu. Tots i cada un d'ells tenen la capacitat de diferenciar els tres tipus de cèl·lules entre elles, ja que el p-valor d'aquests és molt inferior al valor de la significança del 5%.

Pel que fa als drepanòcits, els descriptors més rellevants en la diferenciació entre ells i els esferòcits o els eritròcits normals són els de forma, com l'excentricitat, la circularitat, l'eix de major longitud i la variància circular.

En el cas dels esferòcits, els descriptors que estan relacionats amb la circularitat de la cèl·lula també juguen un paper molt important. En aquest cas remarcaríem la circularitat i l'excentricitat, tot i que no serien del tot discriminants dels eritròcits normals.

El fet que els esferòcits i els eritròcits normals siguin de formes tant similars, dificulta la classificació. És aquí on cal treballar sobre descriptors de color com les mitjanes d'intensitat.

És molt interessant el fet que el mètode mRMR ha col·locat entre els 10 descriptors amb menor redundància i major rellevància els descriptors que fan referència a la mitjana d'intensitat de píxels per cèl·lula de cada una de les tres components de color. Per tant, és interessant l'estudi d'aquesta en casos on la forma no és determinant per la classificació, ja que tant la mitjana d'intensitat de les cèl·lules en la component vermella, com en la component blava, com en la verda permeten diferenciar els eritròcits en els seus grups corresponents.



CAPÍTOL 8:

CLASSIFICACIÓ

La classificació és un procés en el qual s'estableix una relació entre els descriptors calculats i el procés d'etiquetatge manual amb la finalitat de classificar de forma automàtica les dades entrades. És possible gràcies a l'aprenentatge automàtic. En aquest cas es realitza amb l'aplicació de MATLAB anomenada *Classification Learner*. L'algoritme analitza les dades de cada un dels descriptors i busca una relació entre elles per classificar-les en les classes que es desitgi, en aquest cas en 3: esferòcits, drepanòcits i normals.

8.1 Marc teòric

Validació creuada de 10 iteracions

La validació creuada de 10 *folds* o iteracions divideix les dades entrades en 10 subconjunts. Un d'aquests subconjunts s'utilitza com a prova i els altres 9 com a entrenant. El procés es repeteix durant 10 iteracions i s'extreu la mitjana total dels resultats de cada una de les iteracions.

Màquines de suport vectorial

Les màquines de suport vectorial (**SVM**) són un conjunt d'algoritmes d'aprenentatge supervisat que es basen en la representació dels punts de la mostra a l'espai separant-los per un hiperplà o per un conjunt d'hiperplans, depenent del tipus de mostra entrada. Analitza cada dada i la situa en una de les bandes existents que separa l'hiperplà, classificant-la segons les seves característiques.

Aquest tipus d'algoritme intenta buscar els hiperplans que separin de manera òptima les dades i ho fa disposant l'hiperplà de manera que tingui la màxima distància amb els punts que estan més a prop d'ell. A la Figura 39 es pot observar com tant H2 com H3 separen els dos grups de mostres, però és H3 el que ho fa deixant la màxima distància entre ell i els punts més propers, és per això que l'hiperplà utilitzat serà el representat per H3.

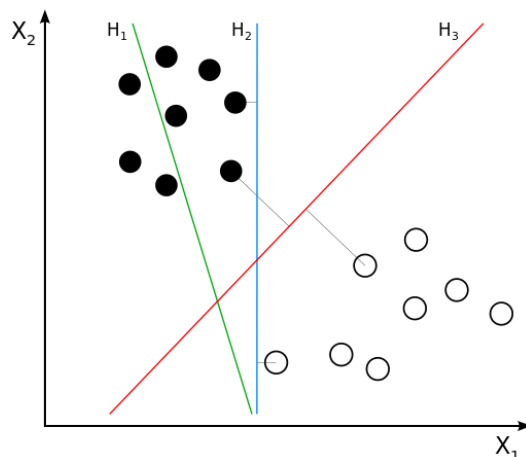


Figura 39. Hiperplans que separen un conjunt de dades. (Font:Wikipedia)

La funció Kernel s'utilitza en casos on trobem més de dos conjunts de dades de classificació o en aquells on un hiperplà de línia recta no és capaç de separar completament els grups de dades [26].

En aquest cas cal augmentar les dimensions de l'espai on estem disposant les dades i així augmentar la capacitat de computació de la màquina d'aprenentatge lineal. Les funcions Kernel ens permeten convertir un problema de classificació no lineal en l'estai dimensional original en un problema de classificació lineal en un espai dimensional major. [27]

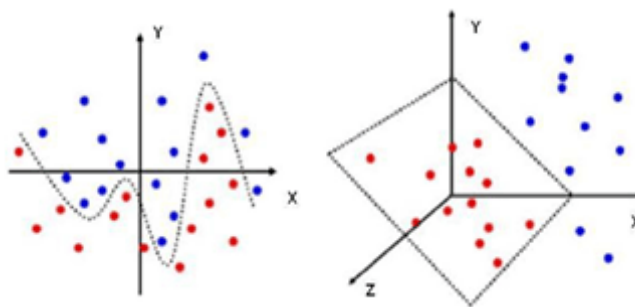


Figura 40. Mapeig de les dades a un espai de majors dimensions. (imatge de la dreta) (Font: Scielo [27.])

8.2 Metodologia de classificació

En el nostre cas s'entrena una classificació amb validació creuada de 10 *folds* on hi entraran en joc tots els descriptors trobats i anteriorment explicats i on es provaran diferents mètodes de classificació.

Un cop ha classificat a cada eritròcit en un dels tres grups segons les dades dels descriptors, compara els resultats obtinguts amb les etiquetes entrades comparant-ho i calculant el percentatge d'incert. Hi ha diversos models de classificació pel que podem automatitzar el procés per trobar el que millor treballa amb les nostres dades. Els models de classificació amb els que treballem són l'arbre de decisió (DT), l'anàlisi discriminant (DA), la màquina de suport vectorial (SVM), el mètode de veïns més propers (KNN) i la classificació de conjunts (EC). Cada un d'aquests mètodes té diversos tipus de càlcul que varien en resultats.

Es parteix d'un total de 552 cèl·lules de cada un dels tipus, el que equival a 1656 eritròcits, com s'ha explicat a l'inici de l'apartat anterior.

S'analitzen els **outliers**, que són aquells descriptors que contenen uns valors extrems en relació a la majoria, és a dir, són representats per una distància major al gràfica respecte els altres valors. Als gràfics de caixes són representats amb una creu vermella.

S'aplica un algoritme per filtrar-los i eliminar-los. Es determinen els valors extrems superiors i inferiors dels descriptors calculant 1.5 vegades el rang interquartil i es calcula quants descriptors amb outliers presenta cada cèl·lula. Si les cèl·lules sobrepassen un lílindar establert, es consideren cèl·lules extremes i seran eliminades. El valor lílindar es fixa igual al valor de la mitjana de l'outlier per a cada observació, més dues vegades la desviació estàndard d'aquest.

Al tenir un nombre elevat de mostres la diferència de dades és mínima. En el nostre cas s'eliminen 93 cèl·lules del total, pel que la classificació s'aplicarà sobre 1563 eritròcits. Eliminar aquestes cèl·lules ens donarà resultat més encertats i ens ajudarà a millorar la qualitat de la classificació.

En la Taula 24 es mostra el millor resultat per a cada mètode de classificació:

Mètode de classificació	Millor mètode	Exactitud (%)
Arbre de decisió (DT)	Medium Tree	91.7
Anàlisi discriminant (DA)	Linear Discriminant	93.4
Màquina de suport vectorial (SVM)	Quadratic SMV	94.7
Veïns més propers (KNN)	Weighted KNN	92.9
Classificació de conjunts (EC)	Bagged Trees	93.0

Taula 24. Exactitud dels mètodes de classificació

El mètode de classificació que dona una millor exactitud és el mètode quadràtic de la **màquina de suport vectorial (SVM) amb un 94.7%**.

Al provar la classificació amb el mètode multiclasse “one vs all” enlloc de “one vs one” millora la exactitud per la màquina de suport vectorial a **95.6%**

A continuació es mostra la matriu de confusió de cada tipus de cèl·lula: esferòcit, drepanòcit i eritròcits normals. S'analitzen un total de 1563 eritròcits. En cada casella de la matriu de confusió es mostra el percentatge de cèl·lules i el nombre d'aquestes.

Classe verdadera	Drepanòcit	98% 540	0% 0	2% 12
	Esferòcit	<1% 1	97% 535	3% 16
	RBC Normal	2% 9	7% 40	92% 503
		Drepanòcit	Esferòcit	RBC Normal
Classe predictiva				

Taula 25. Matriu de confusió SVM

Seguidament es calculen paràmetres estadístics sobre la matriu de confusió. S'utilitzen les sigles VP pels valors veritaders positius, VN pels veritaders negatius, FP pels falsos positius i FN pels falsos negatius.

(Eq. 0.1)

$$Sensibilitat = \frac{VP}{VP + FN}$$

(Eq. 0.2)

$$Especificitat = \frac{VN}{VN + FP}$$

Paràmetre	Drepanòcits	Esferòcits	Normals
Exactitud	98.21%	97.05%	92.30
Sensibilitat	97.46%	95.47%	91.12%
Especificitat	93.28%	94.29%	96.47%

8.3 Discussió dels resultats de la classificació

Aquest apartat dona validesa a la resta de treball ja que mostra un valor d'exactitud del 95.6% dels resultats obtinguts. Aquest resultats provenen d'un bon processat, una bona segmentació i una bona extracció de característiques, sense els quals no s'hagués arribat a una exactitud tant elevada.

Analitzant el percentatge d'encert en cada tipus d'eritròcit, podem trobar lògic que el valor donat a l'encert en els drepanòcits sigui el més elevat, del 98%. Com bé sabem, són els que es diferencien més en forma, pel que segurament són els més senzills de detectar dels tres. Tot i així, els drepanòcits que no s'han classificat com a tals, ho han fet com a eritròcits normals, i en cap cas com a esferòcits.

El segon millor valor d'encert se l'enduen els esferòcits, amb un 97%. Altra vegada els esferòcits no classificats correctament s'han col·locat al grup de normals, deixant-ne només un mal classificat com a drepanòcit.

Per últim i amb el valor més baix d'encert trobem els eritròcits normals, amb un 92%. Tot i ser el valor més baix es considera un bon percentatge d'encert. En aquest cas, trobem un 7% d'eritròcits normals classificats com a esferòcits, ja que és el tipus més similar.

Es compara l'exactitud obtinguda amb la primera part del treball [1], i es veu una mínima diferència. Aquesta pot ser deguda al fet que en aquest projecte es treballa amb tres tipus de cèl·lules i no amb dues, com és el cas de l'anterior.

Per a millorar els resultats es proposa aprofundir en el processat i la segmentació i trobar nous descriptors que siguin capaços de diferenciar amb més exactitud les cèl·lules, més enllà dels de mida i color.

CAPÍTOL 9:

ANÀLISI DE L'IMPACTE AMBIENTAL

En aquest capítol es pretén avaluar l'efecte que té el nostre projecte sobre el medi ambient i les conseqüències que pot dur la seva implementació.

En la següent taula es mostren dos possibles impactes relacionats amb el nostre projecte:

	Causa	Conseqüències professional	Conseqüències pacient
Impacte 1	Reducció del temps de dedicació de l'expert	-Inversió del temps en altres aspectes -Estalvi econòmic en personal	-Reducció del nombre de proves
Impacte 2	Automatització del diagnòstic	-Estalvi econòmic en recursos -Diagnòstic més precís, sense subjectivitat.	-Reducció del nombre de proves -Diagnòstic més precís

Taula 26. Anàlisi de l'impacte ambiental

El fet de poder reduir el temps de dedicació de l'expert significa que aquest serà útil en algun altre aspecte del laboratori. Estem optimitzant els recursos. A més, serà positiu pel pacient ja que no seran necessàries tantes proves. Això implica un estalvi en material de laboratori.

Per altra banda, automatitzar els resultats també va relacionat amb el primer impacte. Aquest automatització disminuirà el factor de subjectivitat que aportava l'expert, tractant totes les mostres per igual i amb menys error.

CAPÍTOL 10: CONCLUSIONS

El principal objectiu a assolir en aquest projecte era la classificació automàtica de glòbuls vermells a partir d'imatges extretes de frotis de sang perifèrica. Com s'ha explicat al llarg del treball, s'han seguit diferents passos per assolir l'objectiu satisfactòriament.

Primerament, cal recordar que aquest projecte està relacionat amb el treball de final de grau de Clara Pladevall [1], d'ampliant-lo a l'estudi de noves cèl·lules i intentant donar sortida a les propostes de millora que es proposaven.

El treball comença amb el procés d'extracció de les imatges al microscopi òptic, continua treballant en els algorismes de processat i segmentació de les imatges per a l'obtenció de cada un dels eritròcits que les componen. El següent pas és l'extracció de descriptors de forma, mida i color per fer possible la diferenciació dels glòbuls vermells, analitzant la rellevància d'aquests en la diferenciació cel·lular. I finalitza amb la classificació dels eritròcits de manera automàtica, avaluant el percentatge d'encert del mètode.

10.1 Conclusions

És cert que per a realitzar aquest treball són necessaris coneixements en processat digital d'imatges i aprenentatge automàtic, ja que és la base d'una gran part del projecte, però també són necessaris coneixements en citologia, en aquest cas, sobretot en drepanòcits i esferòcits. La bona adquisició d'imatges és una condició que està estrictament relacionada amb el bon ús del material i amb el coneixement de les cèl·lules que conté. L'adquisició de les imatges de sang perifèrica sobre les que es treballa en aquest projecte ha estat realitzada per mi, pel que m'ha calgut aprendre el procediment a seguir per fer-ho correctament. Una imatge mal adquirida pot implicar errors en els resultats. En aquest projecte s'han aconseguit els frotis gràcies a l'estreta col·laboració amb el laboratori CORE de l'Hospital Clínic de Barcelona.

En el processat i la segmentació s'han realitzat modificacions per assolir millor els resultats. S'ha treballat sobre unes components de color diferents. S'han modificat els valors de filtratge d'acord amb les necessitat de la nova cèl·lula afegida a l'estudi. I s'ha millorat la segmentació discriminant

aquells eritròcits mal segmentats. Tot i la millora, es considera que una de les majors limitacions del treball es troba en la segmentació de cèl·lules sobreposades. En alguns casos la superposició és tal que, tot i intentar una segmentació, les cèl·lules perden les característiques pròpies que es necessiten per la posterior implementació de descriptors.

Un procés d'etiquetatge correcte és primordial per a poder dur a terme la classificació i la posterior comparació dels resultats. Cal la supervisió d'un citòleg expert per donar validesa als resultats. Tot i així, hi ha glòbuls vermells que poden presentar característiques ambigües i fer dubtar a l'expert entre dos tipus diferents, sobretot si sortim de la classificació de drepanòcits i esferòcits. En aquest cas caldria contrastar els coneixements per evitar la subjectivitat.

S'extreuen nous descriptors, arribant a un total de 28 i es sotmeten als testos de normalitat i homoscedasticitat per a saber el tipus de distribució de la mostra. En aquest punt s'ha treballat sobre una estimació utilitzant el mètode Kruskal-Wallis, suposant que totes les mostres disposen d'una variància similar. Per a treballar de manera correcta caldria utilitzar el mètode Welch's anova [23].

Es classifiquen els descriptors segons la seva rellevància en la diferenciació dels tipus de cèl·lules. S'estudien els mètodes Relieff, el mètode neighborhood Component Analysis i el mètode mRMR. És interessant la comparació entre ells ja que els resultats varien significativament. S'analitzen els resultats obtinguts pel mètode mRMR i es qualifiquen de molt satisfactoris. S'estudien els 10 descriptors de màxima rellevància i mínima redundància, deixant en els primers llocs l'eix de major longitud, la mitjana de la intensitat dels píxels en la component verda i l'excentricitat. Els drepanòcits són els que, en general, es diferencien més fàcilment per la seva forma allargada. La diferenciació entre esferòcits i eritròcits normals no és tant obvia, ja que la seva silueta és més similar; cal buscar altres tipus de descriptors com són els de color. Es considera interessant que tant la mitjana de la intensitat dels píxels de la component verda, com de la component blava, com de la vermella es trobin en els 10 primers descriptors. Cal tenir-la en compte en comparacions on la forma de la cèl·lula no és determinant.

La comparació amb els resultats obtinguts en la primera part d'aquest treball [1] no és directa, ja que es tracta amb diferents tipus de cèl·lules. Tot i així, es coincideix en 5 dels 10 descriptors de mRMR, com són els eixos de menor i major longitud.

L'últim pas i més important és la classificació de les dades. El classificador SVM mostra una exactitud de la classificació del 95.6 %, la qual es considera satisfactòria. Aquest percentatge elevat prové dels bons resultats en cada un dels passos seguits durant el projecte, sense els quals no seria possible tant encert en la classificació. En la comparació amb el projecte previ [1], es nota una diferència mínima dels valor d'exactitud degut al fet que en aquest projecte es treballa amb tres tipus de cèl·lules diferents i no amb dues.

Per concloure, es considera l'anàlisi digital d'imatges i l'aprenentatge automàtic amb un gran potencial per ajudar al citòleg en la caracterització morfològica de les cèl·lules sanguínies i en el diagnòstic de malalties.

10.2 Perspectives de futur

- Millora en el procés de segmentació d'aquelles cèl·lules sobreposades per a aconseguir una segmentació total el més òptima possible i així poder reduir la desestimació d'eritròcits.
- Ampliar la classificació a més tipus d'eritròcits. Pot donar resultats interessants en la classificació.
- Seguir amb l'extracció de nous descriptors. Aquesta perspectiva de futur va lligada amb l'anterior ja que al afegir nous tipus de cèl·lules, segurament serà necessari buscar nous descriptors òptims per aquestes.
- En el cas concret d'aquest projecte, una perspectiva de futur seria aconseguir més descriptors capaços de diferenciar completament entre esferòcits i eritròcits normals. En la classificació final un 7% dels eritròcits normals són classificats com es esferòcits. Aquest valor és millorable.



CAPÍTOL 11:

PRESSUPOST

En aquest apartat es fa una estimació dels costos del projecte de final de grau realitzat. Aquest podrien variar en funció de les facilitat donades a l'autor.

11.1 Costos de personal

Costos de personal engloba els costos en €/hora dels experts que han col·laborat en aquest projecte.

Tipus de cost	Cost (€/h)	Hores (h)	Total (€)
<i>Dedicació de l'estudiant</i>	8	768	6.144
<i>Dedicació del tutor del projecte UPC</i>	60	60	3.600
<i>Dedicació del citòleg Hospital Clínic</i>	50	8	400
<i>Dedicació del codirector del projecte</i>	50	6	300
TOTAL			10.444€

Taula 27. Costos de personal

11.2 Costos de material

En aquest apartat es tenen en compte els materials utilitzats durant el projecte i l'estimació del valor del seu possible desgast.

Tipus de cost	Cost	Quantitat (unitat)	Total (€)
<i>Utilització microscopi i càmera DP73</i>	100€	-	100
<i>Oli d'immersió</i>	5€	1 unitat	5
<i>Portaobjectes</i>	8€	Pack de 50 unitats	8
<i>Desplaçament</i>	15 desplaçaments	8€/desplaçament	120
<i>Impressió i enquadernació del treball</i>	25€	1	25
TOTAL			258€

Taula 28. Costos de material

11.3 Costos informàtics

A continuació es mostren els costos relacionats amb els programes informàtics utilitzats durant la realització del treball i amb el material de caire Informàtic que ha estat indispensable.

Tipus de cost	Cost	Quantitat (unitat)	Total (€)
<i>Matlab 2017b</i>	69€	1	69
<i>Microsoft Office</i>	7€/mes	4 mesos	28
<i>Dropbox Plus 1T</i>	8,25€/mes	4 mesos	33
<i>Plataforma GitHub privada</i>	6€/mes	4	24
<i>Ordinador portàtil hp</i>	540€	1	540
TOTAL			694€

Taula 29. Costos informàtics

11.4 Pressupost final

Tipus de cost	Cost (€)
<i>Costos de personal</i>	10.444
<i>Costos de material</i>	258
<i>Costos informàtics</i>	694
TOTAL	11.396 €

Taula 30. Cost total

El cost estimat del treball realitzat és d'**11.396€**.



BIBLIOGRAFIA

- [1] C. Pladevall Viladecans, "Anàlisi digital d'imatges de formes eritrocitàries anòmales," Universitat Politècnica de Catalunya, 2017.
- [2] "OMS | Prevalencia mundial de la anemia y número de personas afectadas," *WHO*, 2013.
- [3] E. McLean, M. Cogswell, I. Egli, D. Wojdyla, and B. de Benoist, "Worldwide prevalence of anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993-2005," *Public Heal. Nutr.*, vol. 12, no. 4, pp. 444–454, 2009.
- [4] A. Digital and C. Morphology, "Automated Digital Cell Morphology Efficiency , Proficiency , Connectivity and Collaboration."
- [5] E. A. Mohammed, M. M. A. Mohamed, B. H. Far, and C. Naugler, "Peripheral blood smear image analysis: A comprehensive review.," *J. Pathol. Inform.*, vol. 5, no. 1, p. 9, 2014.
- [6] N. Torrens, "TFG-NuriaTorrens," 2017. [Online]. Available: <https://github.com/nuriatorrens/TFG-NuriaTorrens>.
- [7] W. Rosell Puig, C. Dovale Borjas, and I. Álvarez Torres, *Morfología humana*, Lic. Ilean. La Habana: Editorial Ciencias, 2002.
- [8] MedLine Plus, "Blood." [Online]. Available: https://medlineplus.gov/blood.html#cat_42.
- [9] A. S. of Hematology, "Blood Basics," 2017. [Online]. Available: <http://www.hematology.org/Patients/Basics/>.
- [10] "Enciclopedia Salud: Definición de Eritropoyesis." [Online]. Available: <http://www.enciclopediasalud.com/definiciones/eritropoyesis>.
- [11] A. . Fallis, "Anemia," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.
- [12] J. A. Clinton Hidalgo, "SÍNDROME DE LA ANEMIA HEMOLITICA," *Rev. Medica Costa Rica y Centroam.*, vol. 583, no. 583, pp. 85–90, 2008.
- [13] S. Perrotta, P. G. Gallagher, and N. Mohandas, "Hereditary spherocytosis," *Lancet*, vol. 372, no. 9647, pp. 1411–1426, Oct. 2008.
- [14] "Esferocitosis hereditaria. Revisión. Parte I. Historia, demografía, etiopatogenia y diagnóstico," *Arch. Argent. Pediatr.*, vol. 113, no. 1, pp. 69–80, Feb. 2015.
- [15] MedLine Plus: enciclopedia médica, "Examen de la fragilidad osmótica." [Online]. Available: <https://medlineplus.gov/spanish/ency/article/003641.htm>.
- [16] M. H. García and M. E. Del Cueto, "Esferocitosis hereditaria: Aspectos clínicos, bioquímicos y

moleculares,” *Rev. Cuba. Hematol. Inmunol. y Hemoter.*, vol. 18, no. 1, pp. 7–24, 2002.

- [17] Medlineplus, “Anemia falciforme.”
- [18] Genagen, “Drepanocitosis - Genagen.” [Online]. Available: <http://www.genagen.es/area-pacientes/informacion-genetica-y-enfermedades-hereditarias/enfermedades-geneticas-mas-frecuentes/drepanocitosis/>.
- [19] Z. Bustamante and R. Garcia, “Genetica, características de la hemoglobina s, anemia falciforme y haplotipos,” *Univ. Mayor San Simon Boliv.*, no. fig 1, pp. 1–6, 2002.
- [20] “BX43 | Olympus Life Science.” [Online]. Available: <https://www.olympus-lifescience.com/es/microscopes/upright/bx43/>.
- [21] “DP73 - Olympus 17.28 MP Peltier Cooled Microscope Digital Cameras.” [Online]. Available: <https://www.olympus-ims.com/es/microscope/dp73/>.
- [22] J. McDonald, “Homoscedasticity - Handbook of Biological Statistics.” [Online]. Available: <http://www.biostat handbook.com/homoscedasticity.html>.
- [23] J. McDonald, “Kruskal–Wallis test - Handbook of Biological Statistics.” [Online]. Available: <http://www.biostat handbook.com/kruskalwallis.html>.
- [24] J. Goldberger, S. T. Roweis, G. E. Hinton, R. Salakhutdinov, S. T. Roweis, and R. Salakhutdinov, “Neighbourhood Components Analysis,” *Adv. Neural Inf. Process. Syst.*, pp. 513–520, 2004.
- [25] S. F. Rosario and K. Thangadurai, “RELIEF : Feature Selection Approach,” *Int. J. Innov. Res. Dev.*, vol. 4, no. 11, pp. 218–224, 2015.
- [26] J. Resendiz Trejo, “Las maquinas de vectores de soporte para identificación en línea,” 2006.
- [27] J. M. Bernal-de Lázaro, A. Prieto-Moreno, O. Llanes-Santiago, and E. García-Moreno, *Revista científica de ingeniería mecánica.*, vol. 14, no. 2. Instituto superior politécnico José Antonio Echeverría, 2011.

ANNEX A

Estadística

Ampliació de l'apartat d'anàlisi estadística de cada tipus de descriptor.

Extensió

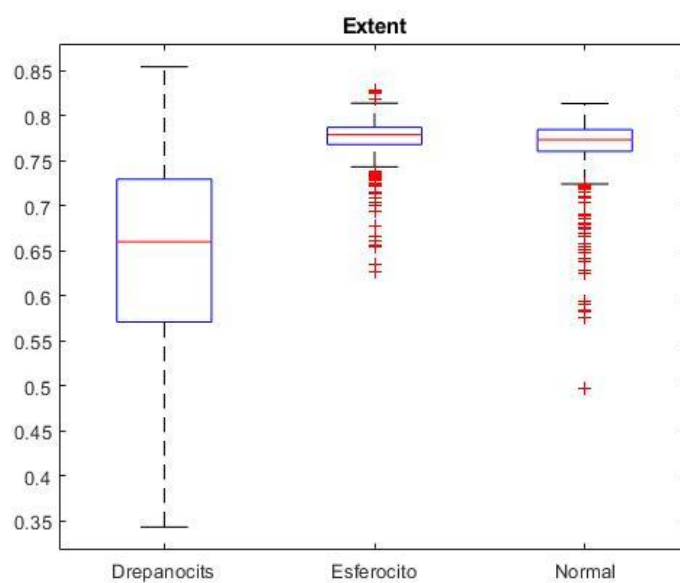


Figura 42. Gràfica de caixes de l'extensió

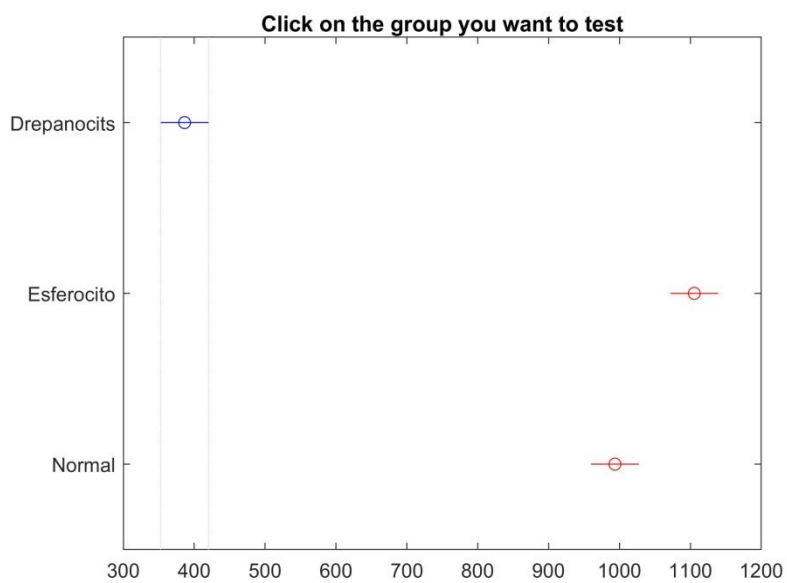


Figura 41. Gràfica de comparació múltiple amb Kruskal-Wallis de l'extensió

Percentils extensió	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.7678	0.5708	0.7601
P ₅₀	0.7786	0.6598	0.7730
P ₇₅	0.7867	0.7293	0.7844

Taula 31. Taula de percentils extensió

Extensió	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$2.9800 \cdot 10^{-4}$	<0.05	✓

Taula 32. Estudi de la significança segons l'extensió

Rodonesa CH

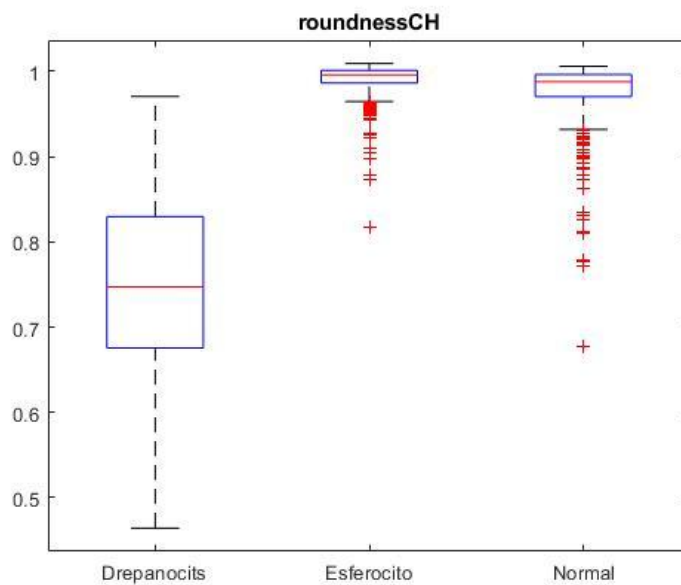


Figura 43. Gràfica de caixes de la rodonesa

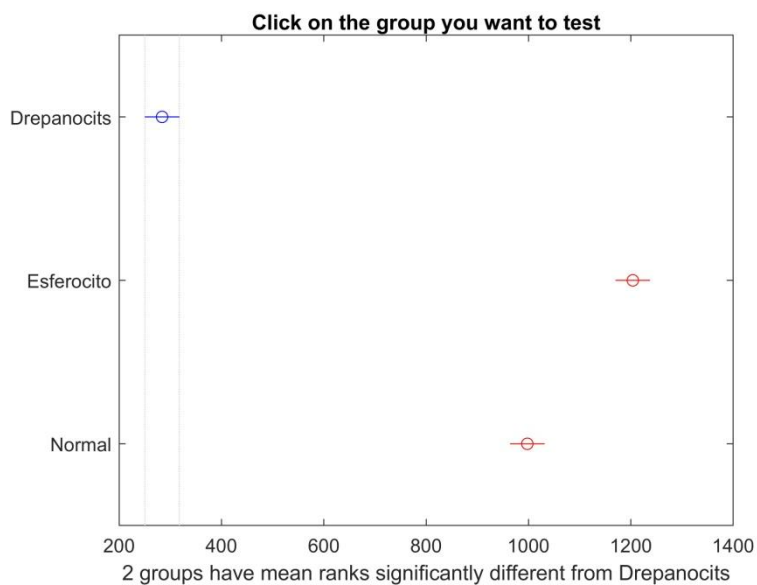


Figura 44. Gràfica de comparació múltiple amb Krusal-Wallis de la rodonesa CH

Percentils rodonesa CH	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.9865	0.6757	0.9704
P ₅₀	0.9960	0.7473	0.9879
P ₇₅	1.0012	0.8298	0.9965

Taula 33. Taula de percentils rodonesa CH

Rodonesa CH	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5825 \cdot 10^{-10}$	<0.05	✓

Taula 34. Estudi de la significança segons la rodonesa CH

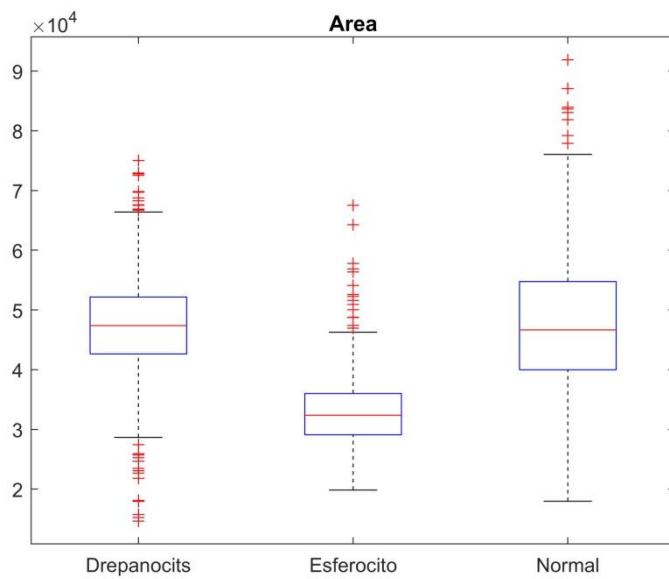
Àrea

Figura 46. Gràfica de caixes de l'àrea

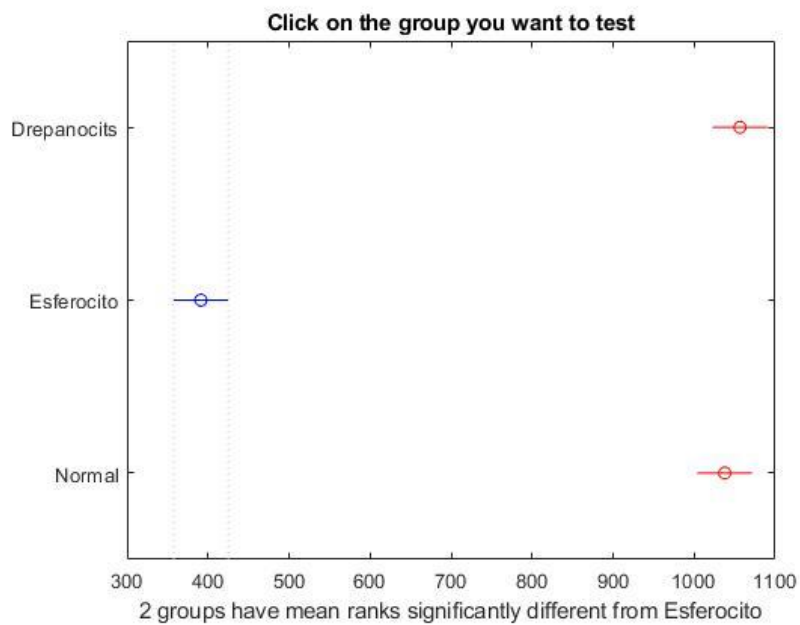


Figura 45. Gràfica de comparació múltiple amb Kruskal-Wallis de l'àrea

Percentils àrea	Esferòcits	Drepanòcits	RBC normals
P ₂₅	29095	4.2635*10 ⁴	3.9975*10 ⁴
P ₅₀	32368	4.7383*10 ⁴	4.6654*10 ⁴
P ₇₅	36009	5.2166*10 ⁴	5.4760*10 ⁴

Taula 35. Taula de percentils àrea

Àrea	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	9.5606*10 ⁻¹⁰	<0.05	✓
Drepanòcits - Normals	0.7895	>0.05	✗
Esferòcits - Normals	9.5606*10 ⁻¹⁰	<0.05	✓

Taula 36. Estudi de la significança segons l'àrea

Entropia component vermella

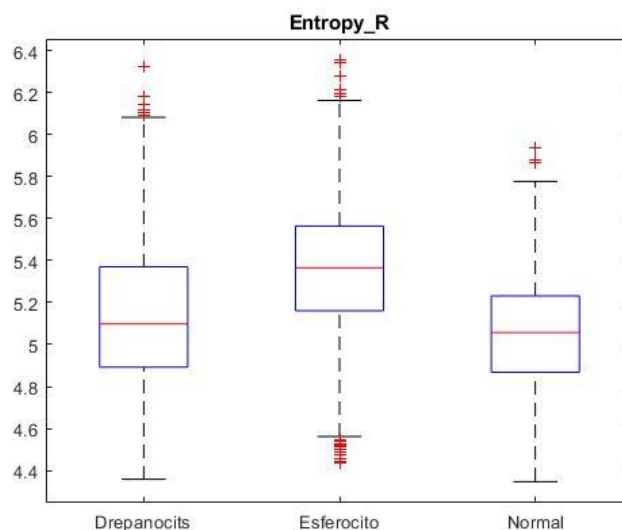


Figura 47. Gràfica de caixes de l'entropia de la component vermella

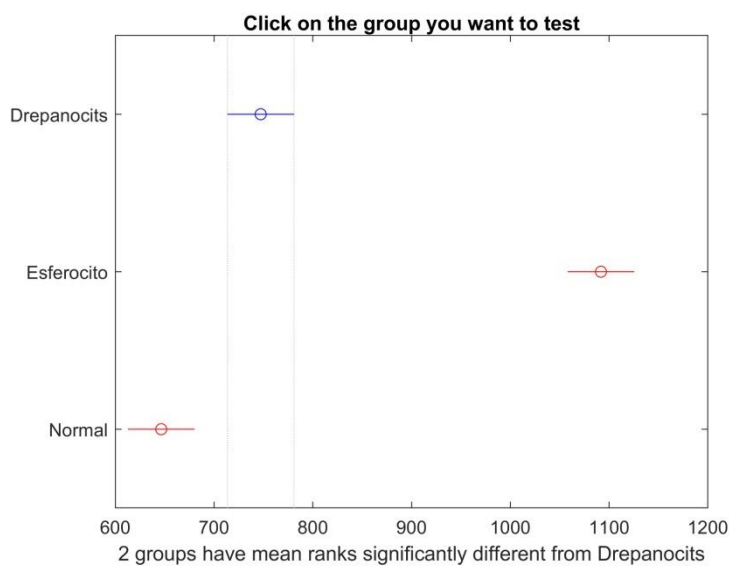


Figura 48. Gràfica de comparació múltiple amb Kruskal-Wallis de l'entropia component vermella

Percentils entropia vermell	Esferòcits	Drepanòcits	RBC normals
P ₂₅	5.1604	4.8921	4.8678
P ₅₀	5.3651	5.982	5.0567
P ₇₅	5.5644	5.3697	5.2319

Taula 37. Taula de percentils entropia vermell

Entropia vermell	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	0.0014	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 38. Estudi de la significança segons l'entropia component vermella

Desviació estàndard component vermella

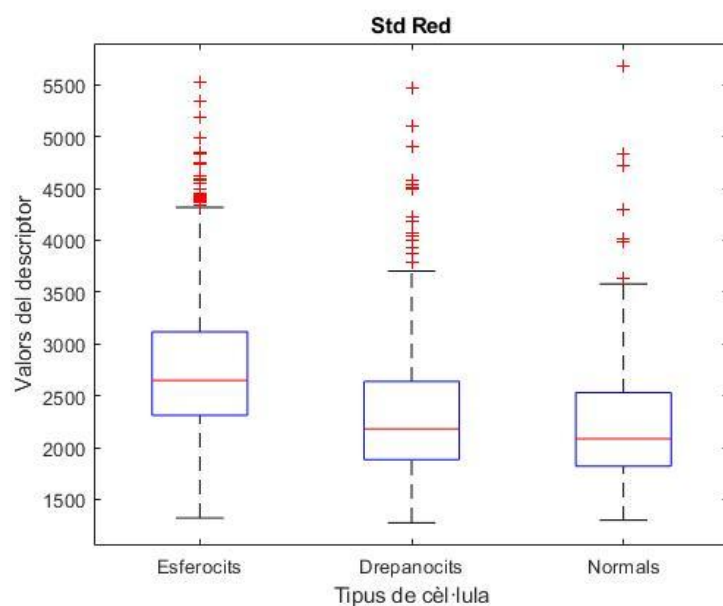


Figura 49. Gràfica de caixes de la desviació estàndard de la component vermella

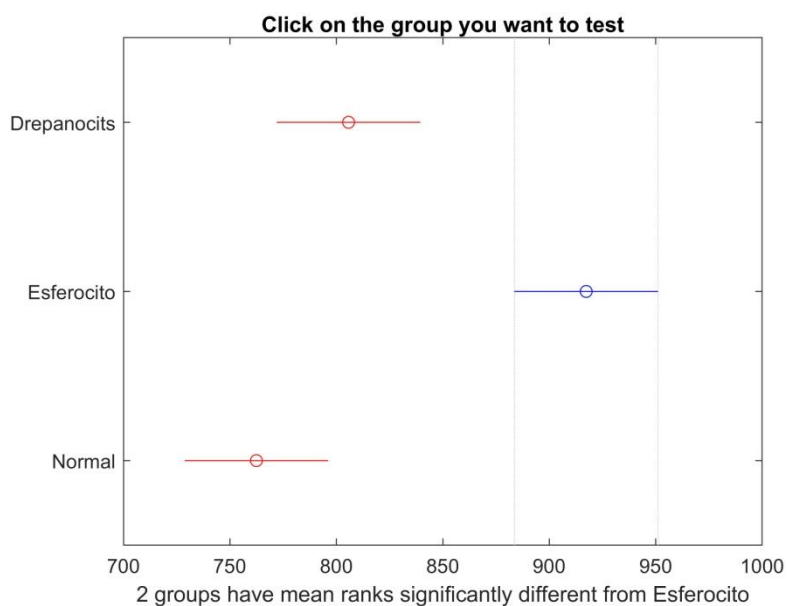


Figura 50. Gràfica de comparació múltiple amb Kruskal-Wallis de la desviació estàndard component vermella

Percentils std vermell	Esferòcits	Drepanòcits	RBC normals
P ₂₅	2.3110*10 ³	1.8839*10 ³	1.8210*10 ³
P ₅₀	2.6491*10 ³	2.1796*10 ³	2.0843*10 ³
P ₇₅	3.1162*10 ³	2.6372*10 ³	2.5299*10 ³

Taula 39. Taula de percentils desviació estàndar vermell

Std vermell	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	3.1019*10 ⁻⁴	<0.05	✓
Drepanòcits - Normals	0.2891	>0.05	✗
Esferòcits - Normals	2.2178*10 ⁻⁷	<0.05	✓

Taula 40. Estudi de la significança segons la desviació estàndard component vermella

Desviació estàndard component verda

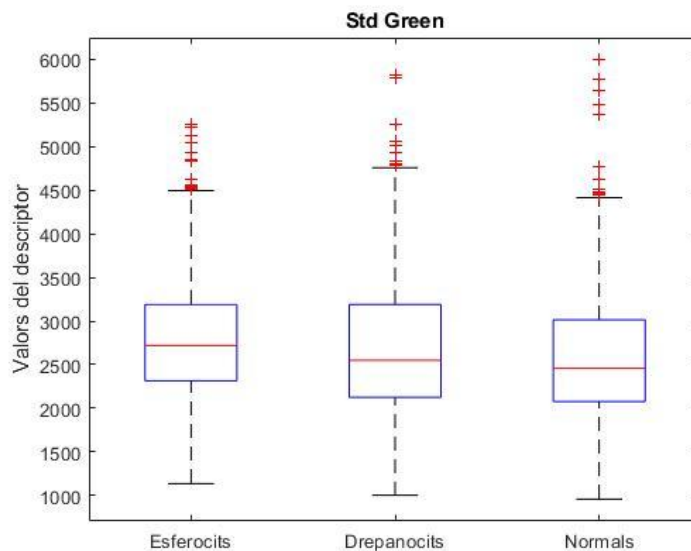


Figura 51. Gràfica de caixes de la desviació estàndard de la component verda

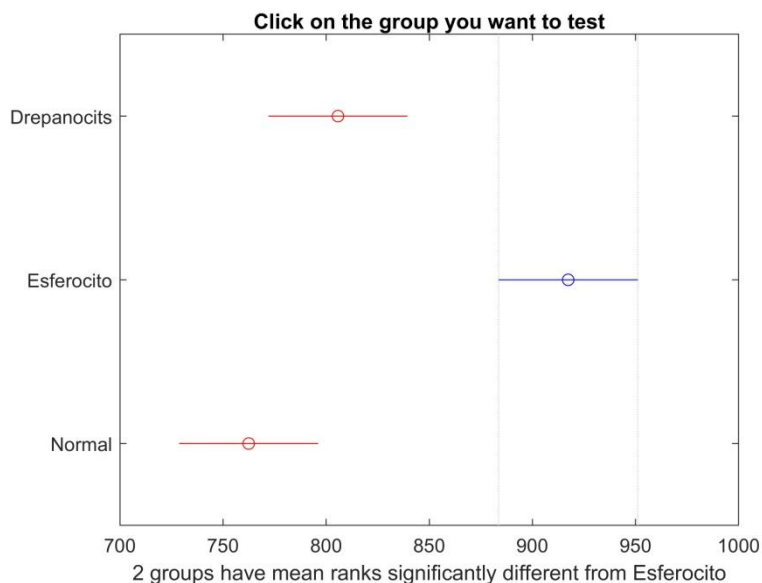


Figura 52. Gràfica de comparació múltiple amb Kruskal-Wallis de la desviació estàndard component verda

Percentils std verd	Esferòcits	Drepanòcits	RBC normals
P ₂₅	2.3159*10 ³	2.1267*10 ³	2.0773*10 ³
P ₅₀	2.7198*10 ³	2.5506*10 ³	2.4577*10 ³
P ₇₅	3.1888*10 ³	3.1893*10 ³	3.0144*10 ³

Taula 41. Taula de percentils desviació estàndard verd

Std verd	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	3.1019*10 ⁻⁴	<0.05	✓
Drepanòcits - Normals	0.2891	>0.05	✗
Esferòcits - Normals	2.2178*10 ⁻⁷	<0.05	✓

Taula 42. Estudi de la significança segons la desviació estàndard component verda

Entropia component blava

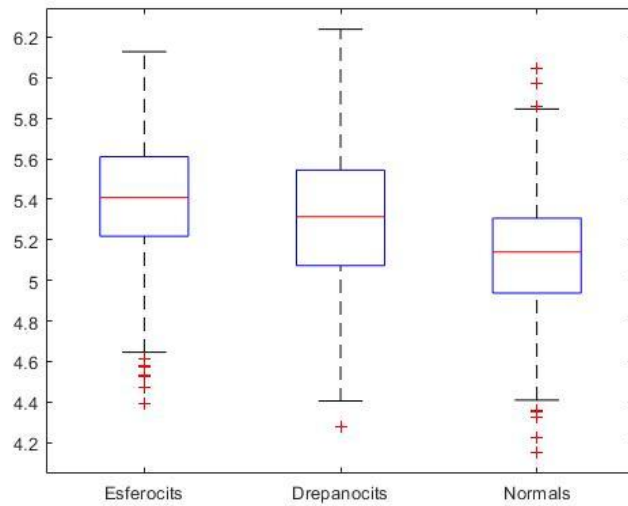


Figura 53. Gràfica de caixes de l'entropia de la component blava

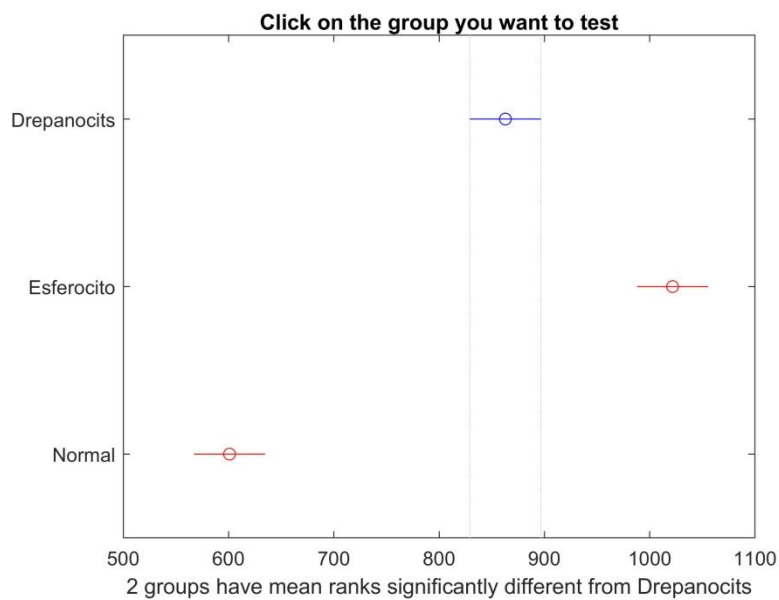


Figura 54. Gràfica de comparació múltiple amb Kruskal-Wallis de l'entropia component blava

Percentils entropia blau	Esferòcits	Drepanòcits	RBC normals
P ₂₅	5.2180	5.0733	4.9379
P ₅₀	5.4087	5.3141	5.1403
P ₇₅	5.6107	5.5442	5.3067

Taula 43. Taula de percentils entropia blau

Entropia blau	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$1.0395 \cdot 10^{-7}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 44. Estudi de la significança segons l'entropia component blava

Curtosis component blava

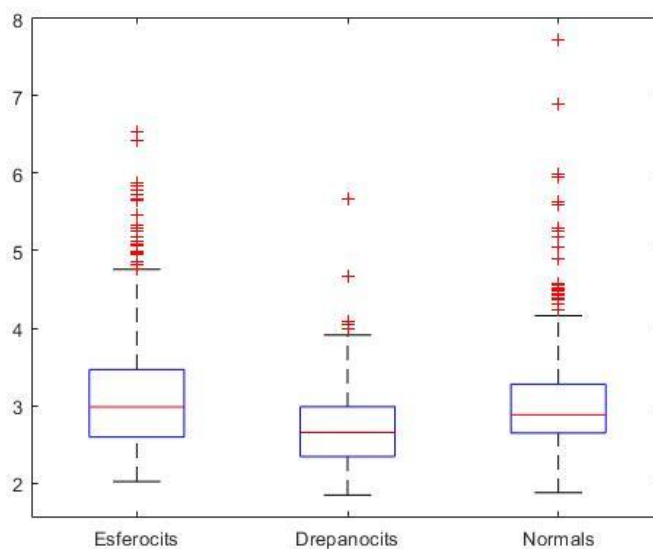


Figura 55. Gràfica de caixes de la curtosis de la component blava

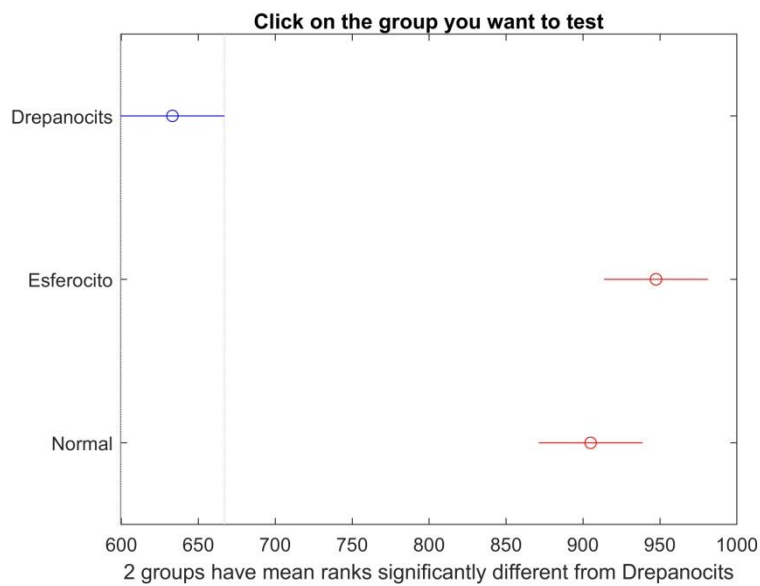


Figura 56. Gràfica de comparació múltiple amb Kruskal-Wallis de la curtosís component blava

Percentils curtosís blau	Esferòcits	Drepanòcits	RBC normals
P ₂₅	2.5957	2.3446	2.6460
P ₅₀	2.9835	2.6547	2.8822
P ₇₅	3.4649	2.9856	3.2769

Taula 45. Taula de percentils curtosís component blava

Curtosis blau	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	0.3022	>0.05	✗

Taula 46. Estudi de la significança segons la curtosís component blava

Entropia component verda

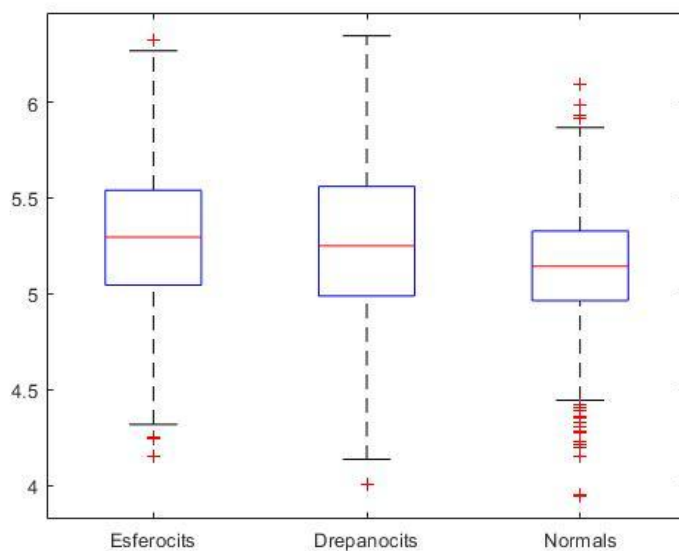


Figura 57. Gràfica de caixes de de l'entropia de la component verda

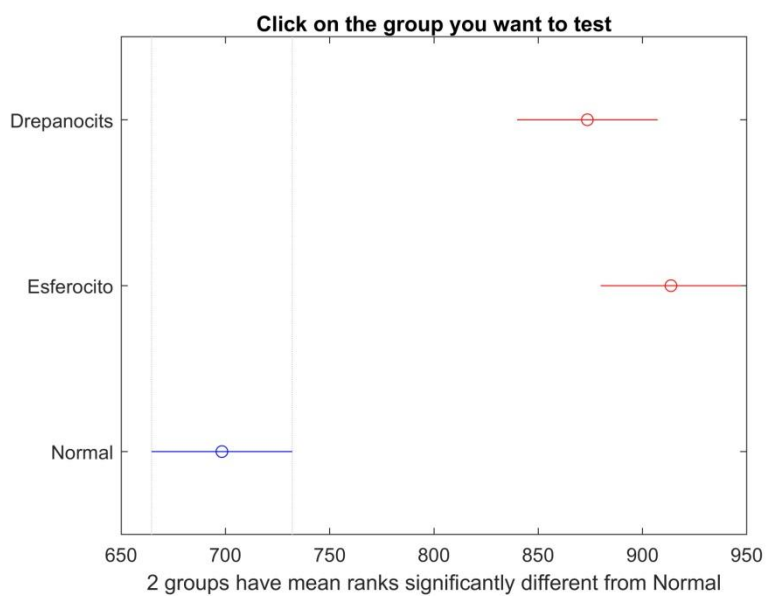


Figura 58. Gràfica de comparació múltiple amb Kruskal-Wallis de l'entropia component verda

Percentils entropia verd	Esferòcits	Drepanòcits	RBC normals
P ₂₅	5.0460	4.9891	4.9648
P ₅₀	5.2958	5.2503	5.1442
P ₇₅	5.5408	5.5605	5.3273

Taula 47. Taula de percentils entropia component verda

<i>Entropia verd</i>	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	0.3446	>0.05	✗
Drepanòcits - Normals	$4.3583 \cdot 10^{-9}$	<0.05	✓
Esferòcits - Normals	$9.5625 \cdot 10^{-10}$	<0.05	✓

Taula 48 .Estudi de la significança segons l'entropia component verda

Desviació estàndard component blava

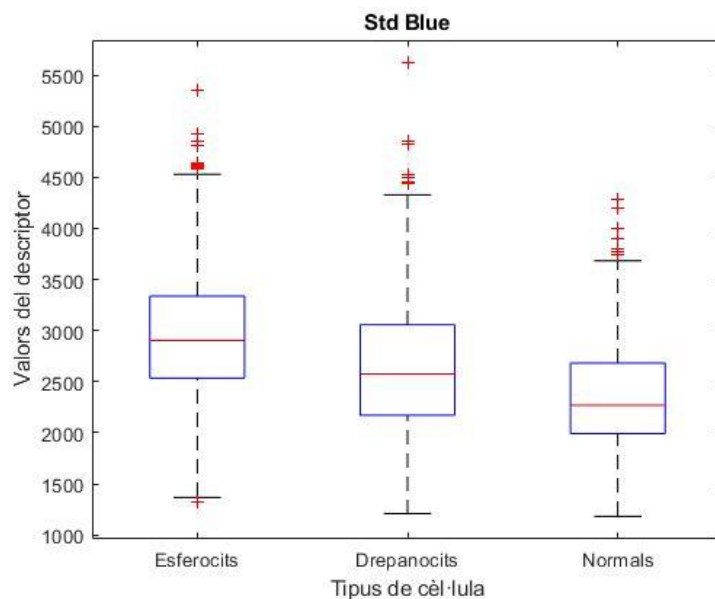


Figura 60. Gràfica de caixes de de la desviació estàndard de la component blava

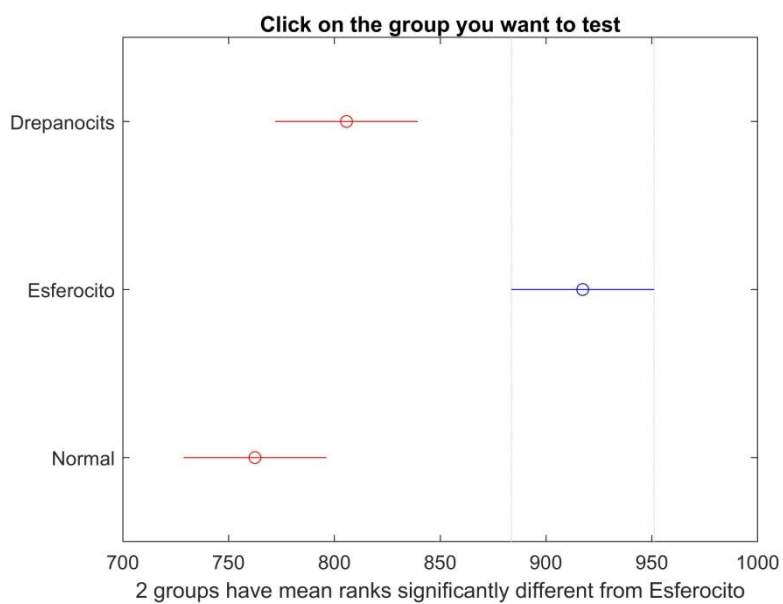


Figura 59. Gràfica de comparació múltiple amb Kruskal-Wallis de la desviació estàndard component blava

Percentils std blau	Esferòcits	Drepanòcits	RBC normals
P ₂₅	2.5351*10 ³	2.1713*10 ³	1.9929*10 ³
P ₅₀	2.9036*10 ³	2.5764*10 ³	2.2707*10 ³
P ₇₅	3.3385*10 ³	3.0276*10 ³	2.6827*10 ³

Taula 49. Taula de percentils desviació estàndard component blava

Std blau	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	3.1019*10 ⁻⁴	<0.05	✓
Drepanòcits - Normals	0.2891	>0.05	✗
Esferòcits - Normals	2.2178*10 ⁻⁷	<0.05	✓

Taula 50 .Estudi de la significança segons la desviació estàndard component blava

Asimetria estadística component vermella

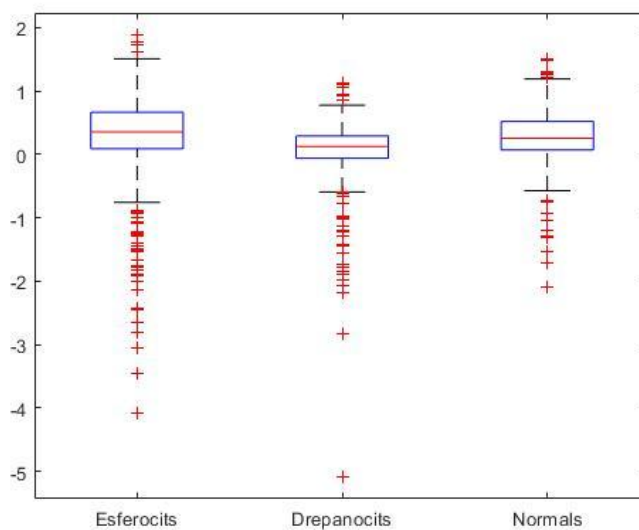


Figura 61. Gràfica de caixes de l'asimetria estadística de la component blava

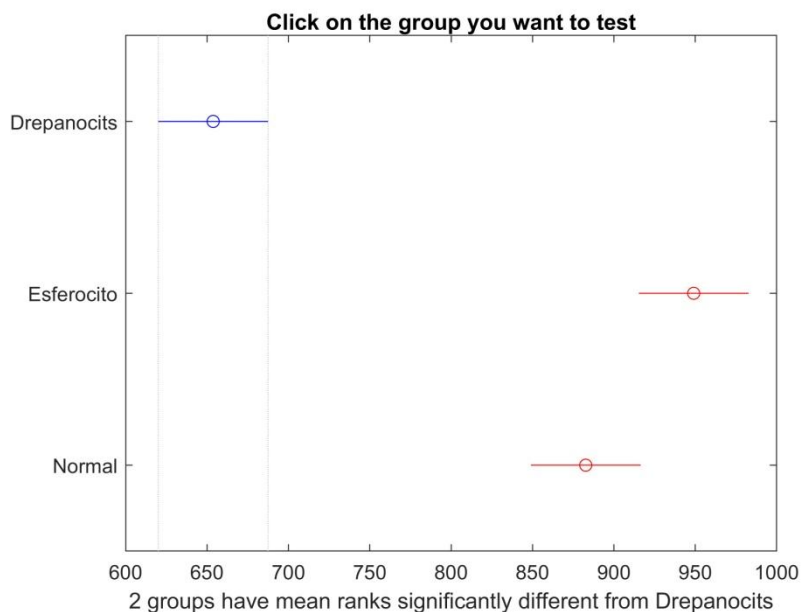


Figura 62. Gràfica de comparació múltiple amb Kruskal-Wallis de l'asimetria estadística component vermella

Percentils asimetria estadística vermell	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.0912	-0.0601	0.0713
P ₅₀	0.3574	0.1279	0.2562
P ₇₅	0.6678	0.2936	0.5221

Taula 51. Taula de percentils asimetria estadística component vermella

Asimetria estadística vermell	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	0.0552	>0.05	✗

Taula 52. Estudi de la significança segons l'asimetria estadística component vermella

Asimetria estadística component verda

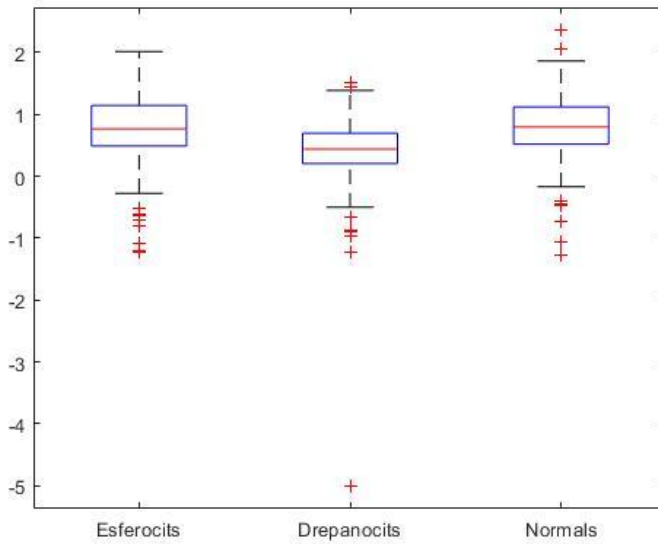


Figura 63. Gràfica de caixes de l'asimetria estadística de la component verda

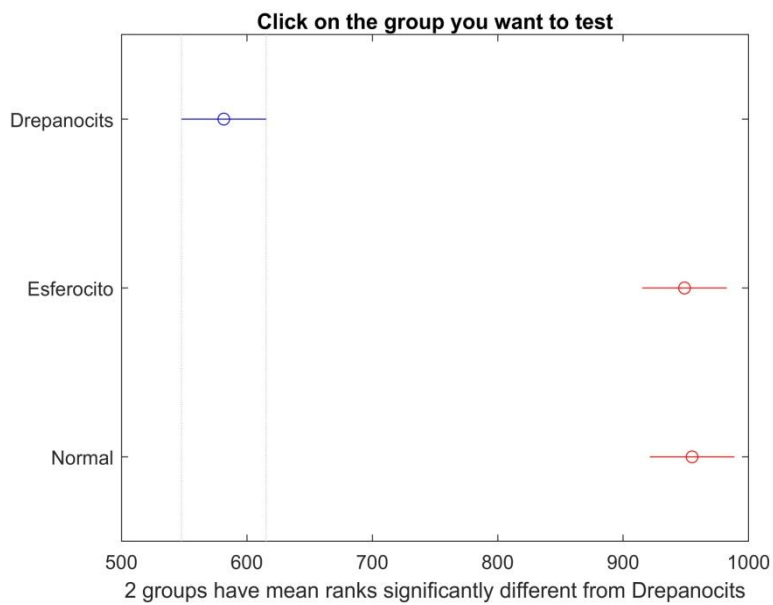


Figura 64. Gràfica de comparació múltiple amb Kruskal-Wallis de l'asimetria estadística componetn verda

Percentils asimetria estadística verd	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.4880	0.2040	0.5175
P ₅₀	0.7653	0.4375	0.7940
P ₇₅	1.1424	0.6938	1.1190

Taula 53. Taula de percentils asimetria estadística component verda

Asimetria estadística verd	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	0.9760	>0.05	✗

Taula 54. Estudi de la significança segons l'asimetria estadística component verda

Solidesa

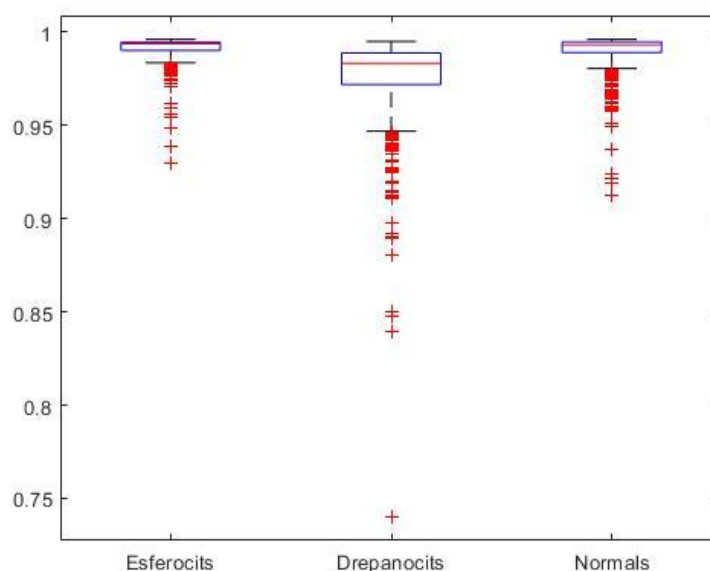


Figura 65. Gràfica de caixes de la solidesa

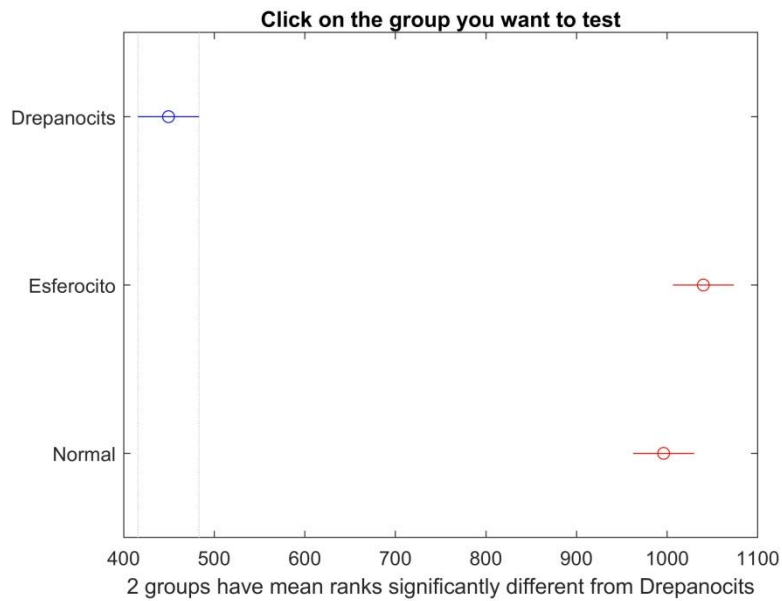


Figura 66. Gràfica de comparació múltiple amb Kruskal-Wallis de la solidesa

Percentils solidesa	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.9902	0.9719	0.9891
P ₅₀	0.9938	0.9833	0.9932
P ₇₅	0.9947	0.9889	0.9948

Taula 55. Taula de percentils solidesa

Solidesa	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	0.2798	>0.05	✗

Taula 56. Estudi de la significança segons la solidesa

Variància el·líptica

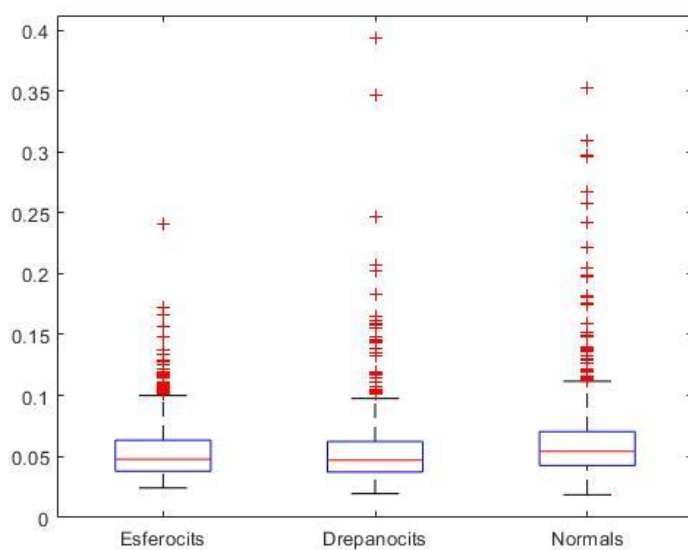


Figura 67. Gràfica de caixes de la variància el·líptica

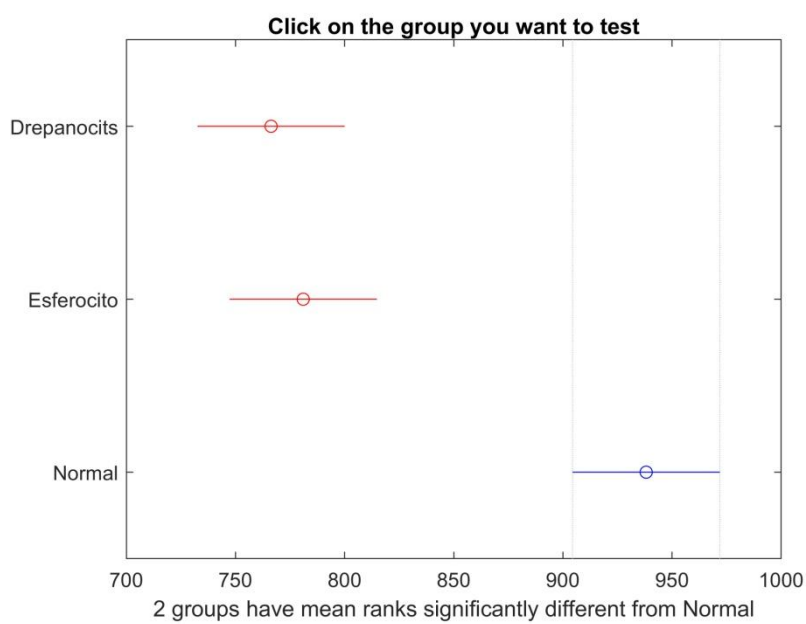


Figura 68. Gràfica de comparació múltiple amb Kruskal-Wallis de la variància el·líptica

Percentils variància el·líptica	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.0370	0.0370	0.0425
P ₅₀	0.0475	0.0468	0.0540
P ₇₅	0.0631	0.0621	0.0703

Taula 57. Taula de percentils variància el·líptica

Variància el·líptica	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	0.8656	>0.05	✗
Drepanòcits - Normals	$8.0084 \cdot 10^{-9}$	<0.05	✓
Esferòcits - Normals	$1.4446 \cdot 10^{-7}$	<0.05	✓

Taula 58. Estudi de la significança segons la variància el·líptica

Curtosis component verda

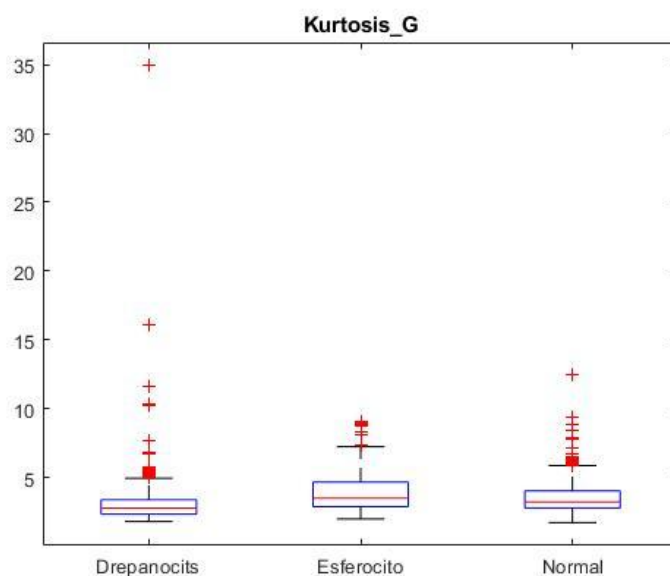


Figura 69. Gràfica de caixes de la curtosis de la component verda

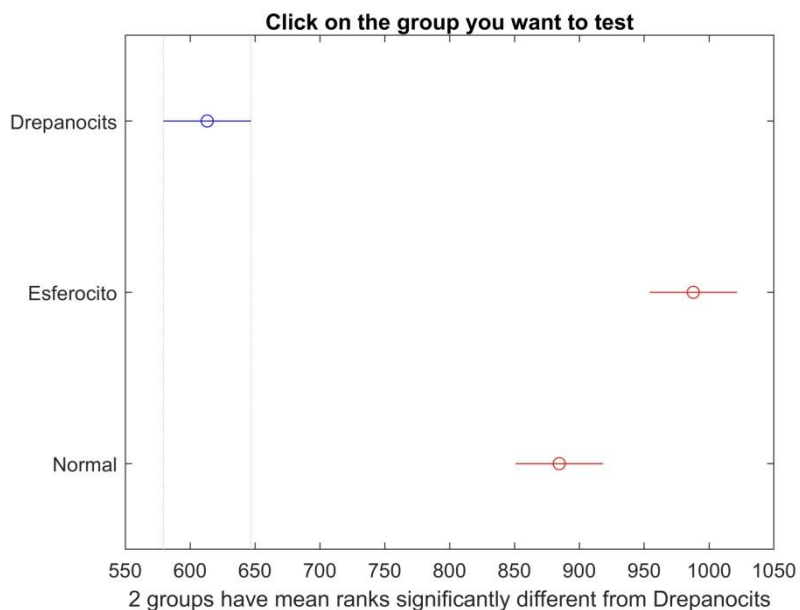


Figura 70. Gràfica de comparació múltiple amb Kruskal-Wallis de la curtosis component verda

Percentils curtosis verd	Esferòcits	Drepanòcits	RBC normals
P ₂₅	2.8781	2.3319	2.7958
P ₅₀	3.4927	2.7725	3.2352
P ₇₅	4.6624	3.3825	4.0270

Taula 59. Taula de percentils curtosis component verda

Curtosis verd	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓

Taula 60. Estudi de la significança segons la curtosis component verda

Convexitat

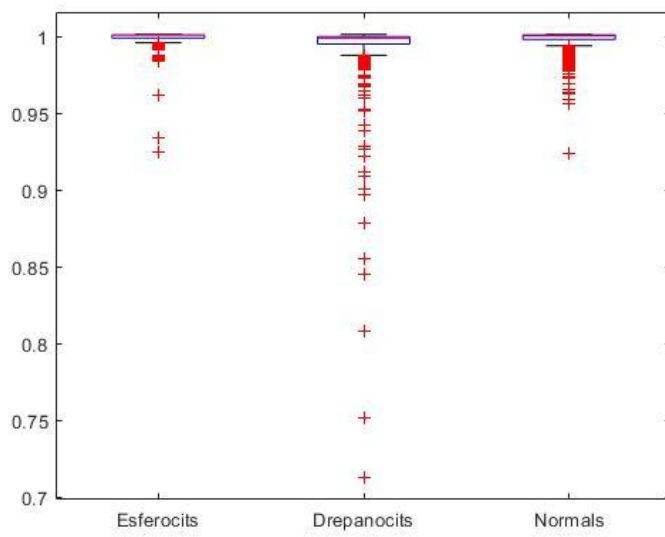


Figura 71. Gràfica de caixes de la convexitat

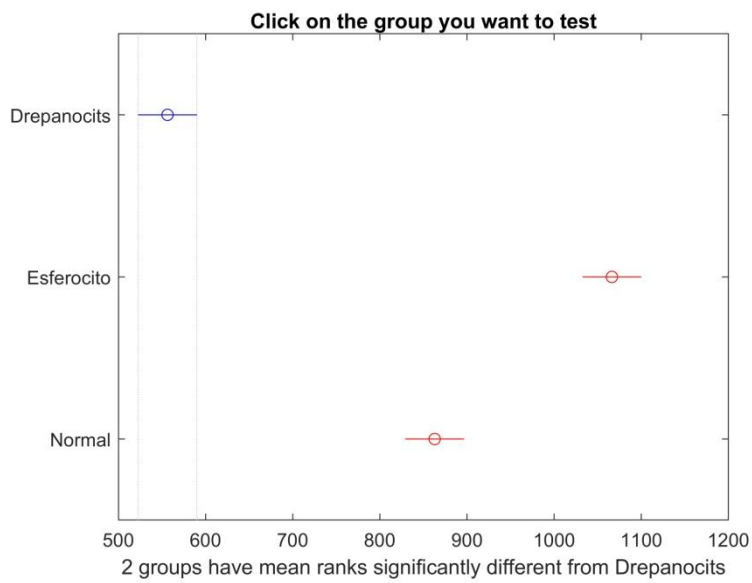


Figura 72. Gràfica de comparació múltiple amb Kruskal-Wallis de la convexitat

Percentils convexitat	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.9990	0.9951	0.9981
P ₅₀	1.0009	0.9986	1.0004
P ₇₅	1.0011	1.0000	1.0009

Taula 61. Taula de percentils convexitat

Convexitat	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.6053 \cdot 10^{-10}$	<0.05	✓

Taula 62. Estudi de la significança segons la convexitat

Curtosis component vermella

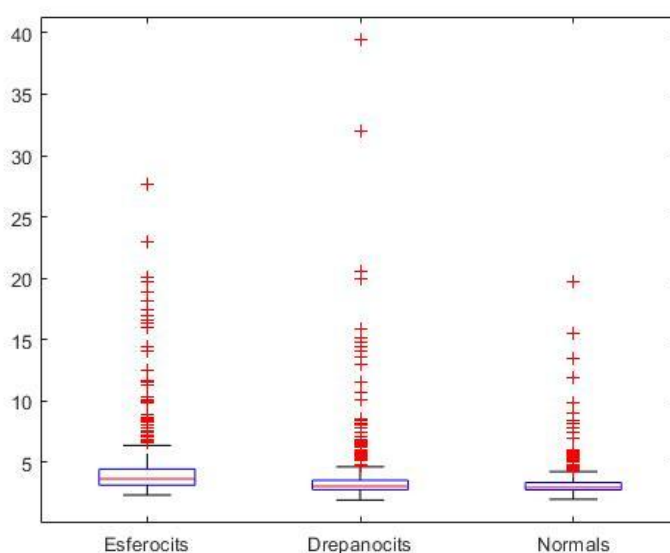


Figura 73. Gràfica de caixes de la curtosis de la component vermella

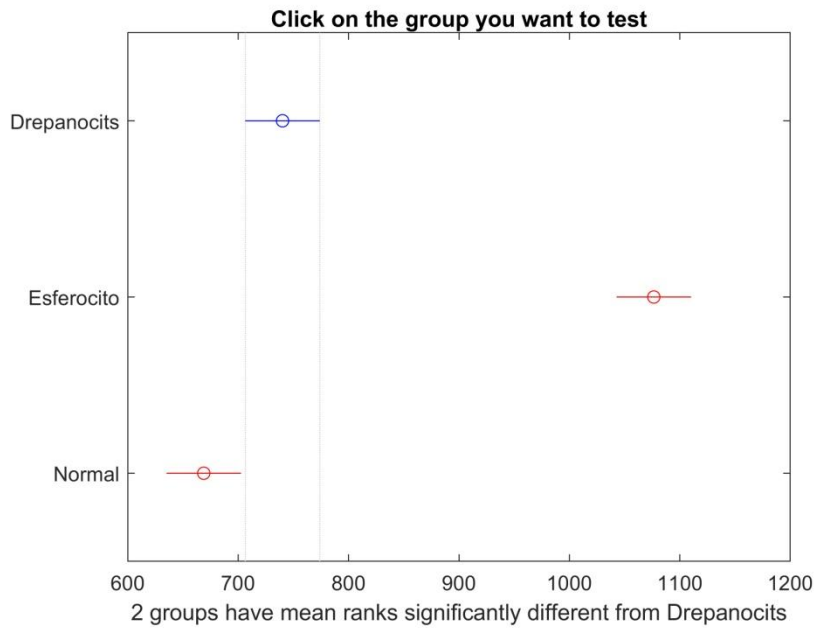


Figura 74. Gràfica de comparació múltiple amb Kruskal-Wallis de la curtosis component vermella

Percentils curtosis vermell	Esferòcits	Drepanòcits	RBC normals
P ₂₅	3.1121	2.7412	2.7208
P ₅₀	3.6300	3.0298	2.9287
P ₇₅	4.4266	3.5053	3.3264

Taula 63. Taula de percentils curtosis component vermella

<i>Curtosis vermell</i>	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	0.0352	<0.05	✓
Esferòcits - Normals	$9.6053 \cdot 10^{-10}$	<0.05	✓

Taula 64. Estudi de la significança segons la curtosis component vermella

Elongació

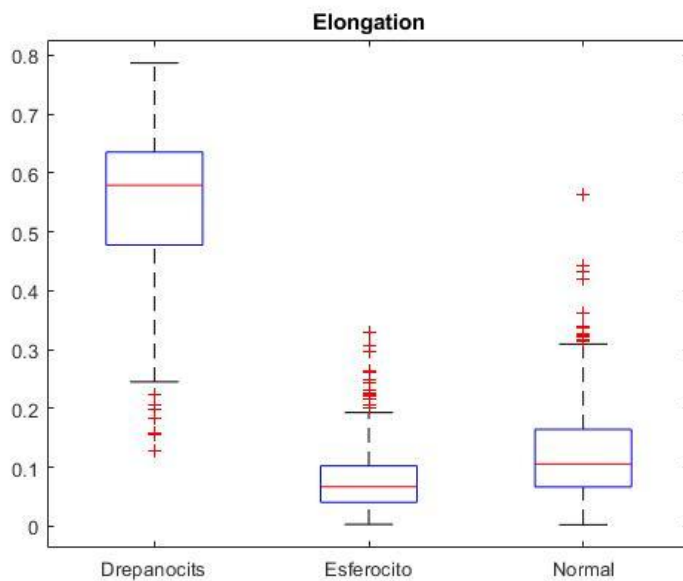


Figura 75. Gràfica de caixes de l'elongació

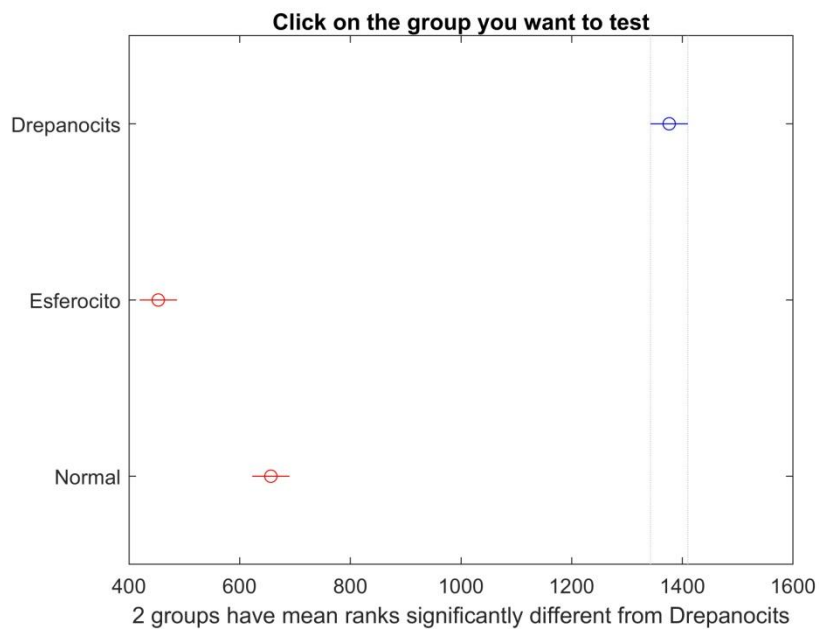


Figura 76. Gràfica de comparació múltiple amb Kruskal-Wallis de l'elongació

Percentils elongació	Esferòcits	Drepanòcits	RBC normals
P ₂₅	0.0409	0.4777	0.0668
P ₅₀	0.0649	0.5793	0.1094
P ₇₅	0.1075	0.6354	0.1718

Taula 65. Taula de percentils elongació

Elongació	P-Valor	Significança	Resultat
Drepanòcits - Esferòcits	$9.5606 \cdot 10^{-10}$	<0.05	✓
Drepanòcits - Normals	$9.5606 \cdot 10^{-10}$	<0.05	✓
Esferòcits - Normals	$9.6044 \cdot 10^{-10}$	<0.05	✓

Taula 66. Estudi de la significança segons l'elongació