

# Breve relazione inerente l'esercizio 1 del 13 Giugno 2011

Paola Gasparini -Matricola 755386

Nell'esercizio 1 viene fornita un'immagine di cellule plasmatiche in cui sono presenti cellule sane e cellule malate, quest'ultime presentano nucleo ingrandito e colorazione non uniforme e più chiara.

In figura 1 è mostrata l'immagine originale.

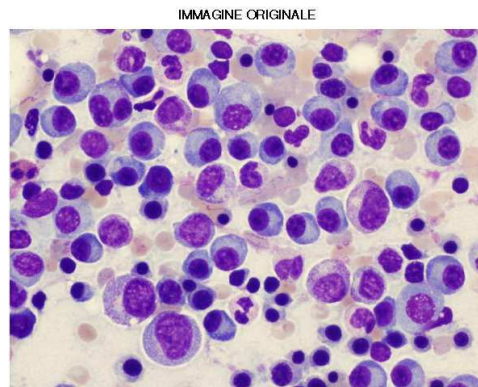


Fig 1. Immagine originale

L'esercizio chiede che vengano valutati i parametri clinici quali numero di cellule, area e forma (esprese come  $\text{media} \pm \text{dev st}$ ) per le cellule normali e per quelle atipiche contenute nell'immagine. Come specificato nel forum, è ritenuto corretto calcolare i parametri per i nuclei tralasciando i citoplasmi.

Il primo passo è stato suddividere l'immagine nei due gruppi, cellule sane e cellule malate, attraverso dell'algoritmo di clusterizzazione k-means. Per poter utilizzare questo algoritmo è stato necessario trasformare l'immagine in double ed in scala di grigi. L'algoritmo è stato utilizzato creando 4 clusters, l'inizializzazione dei centri di ogni cluster è avvenuta distribuendo i centri uniformemente all'interno dei valori dell'istogramma.

In figura 2 viene mostrato l'istogramma suddiviso in 4 clusters realizzato attraverso l'algoritmo delle k-means.

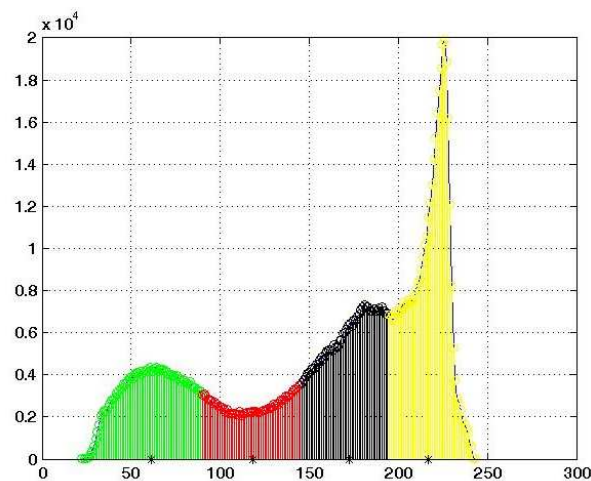


Figura 2. Istogramma ottenuto con l'algoritmo k-means con k uguale a quattro.

In figura 3 viene invece mostrata l'immagine binarizzata delle cellule sane (primo cluster) utilizzando questo algoritmo.

IMMAGINE BINARIZZATA DI CELLULE SANE

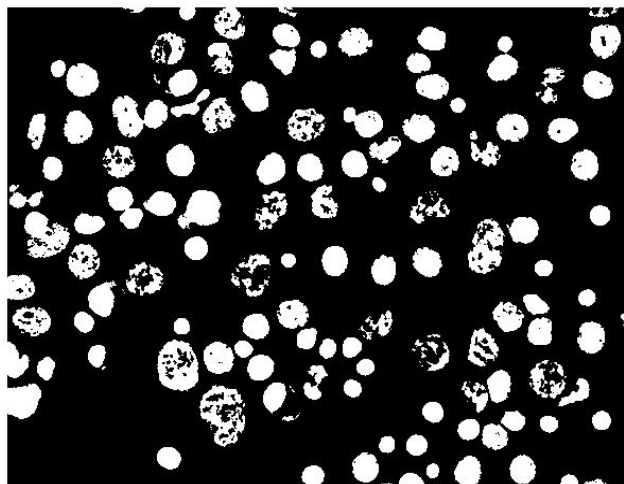


Figura 3. Immagine binarizzata delle cellule sane usando l'algoritmo k-means a 4 clusters.

Come si può vedere il risultato ottenuto non è stato però soddisfacente in quanto risultavano appartenere al primo cluster anche molti pixels di cellule malate.

Si è quindi deciso di aumentare il numero di clusters in modo da aumentare la selettività sull'intensità delle cellule sane.

Portando il numero di clusters a 9 il risultato è risultato migliore.

E' bastato poi utilizzare la funzione "imerode" sull'immagine binarizzata e successivamente "imdilate" sull'immagine erosa per eliminare le cellule malate dall'immagine.

L'elemento strutturale utilizzato è stato un disco di raggio 9. L'immagine binarizzata delle cellule sane è mostrata in figura 4.

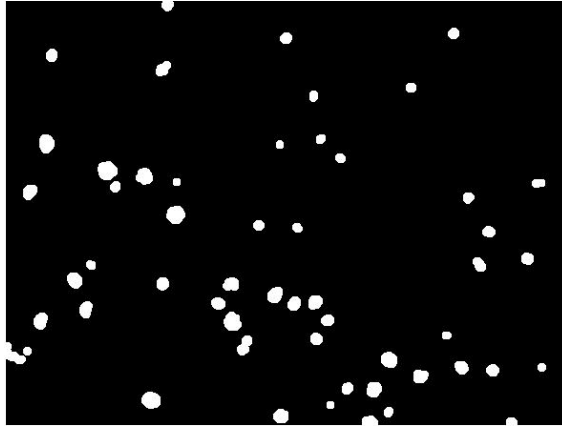


Fig.4 immagine binarizzata delle cellule sane usando algoritmo k-means a 9 cluster seguito dalle funzioni “imerode” e “imdilate”.

Qui di seguito è mostrata l’immagine ottenuta estraendo i contorni e sovrapponendola all’immagine originale è mostrata in figura 5.

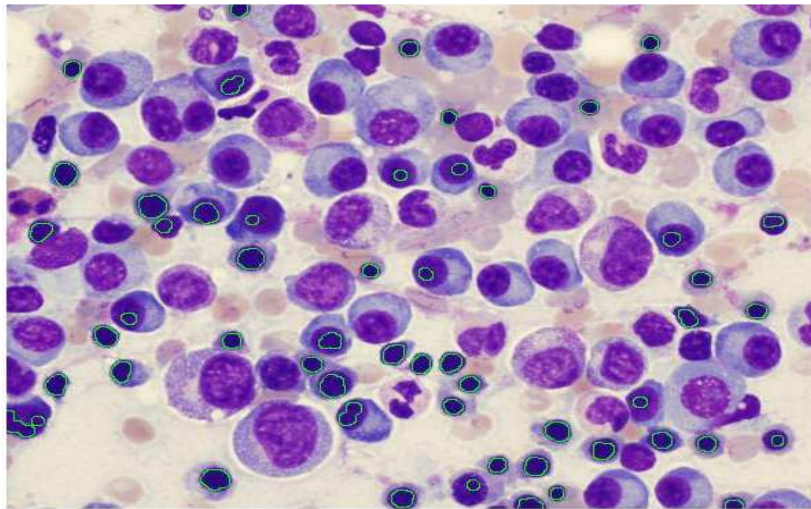


Figura 5. In verde i contorni delle cellule sane.

Per contare le cellule sane ho utilizzato la funzione “bwlabel” sull’immagine binarizzata in modo da applicare un numero differente ad ogni cellula, ho poi creato l’immagine usando “label2rgb”.

Il conteggio delle cellule sane è risultato essere 54 (54 labels ottenuti) e la figura risultante è la seguente:

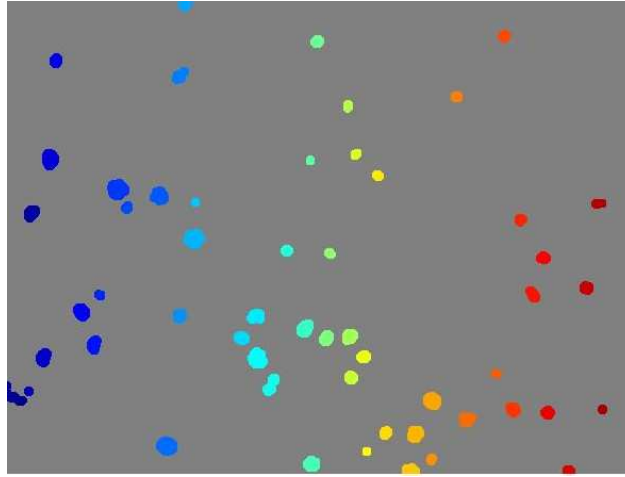


Figura 6. Cellule sane etichettate ognuna con un colore differente.

Un semplice modo per calcolare le cellule malate è stato quello di contare le cellule totali utilizzando l'immagine ottenuta attraverso la divisione in cluster con  $k$  uguale a 2 ed infine sottrarre al numero delle cellule totali il numero delle cellule sane.

L'immagine delle cellule totali etichettate è la seguente:

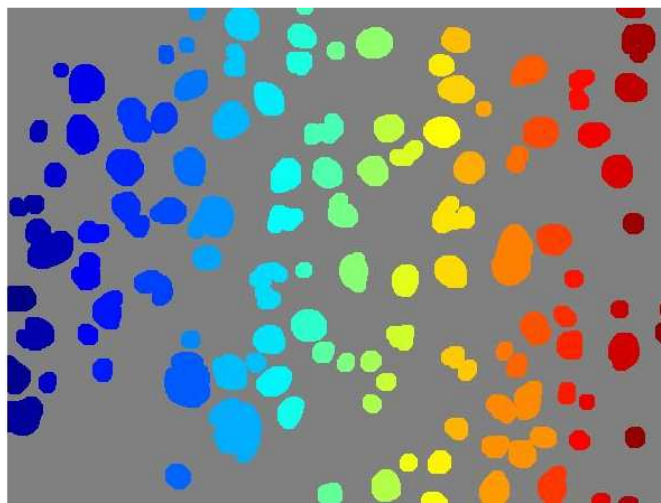


Fig. Cellule totali (sane e malate) comprese nell'immagine.

Per separare le cellule unite si è utilizzato la funzione “bwdist” sul negativo dell'immagine binarizzata e poi successivamente la funzione “watershed”. Il numero di cellule totali risulta essere 115, sottraendo 54 ottengo che il numero di cellule malate è 61. In questo calcolo bisogna comunque tener conto che alcune cellule sono rimaste fuori dalla classificazione.

Infine si è calcolata l'area media e l'eccentricità di entrambi i gruppi. L'area media di ogni gruppo viene calcolata come somma dei pixels totali del gruppo diviso il suo numero di cellule.

Per quanto riguarda le cellule malate l'area media di ognuna è pari a  $258.813/115$ , ossia 2250.54 pixels per ogni cellula. Le cellule sane invece presentano un'area pari a  $33723/54$ , ossia 624.50 pixels per ogni cellula. La deviazione standard delle cellule malate per quanto riguarda l'area

risulta essere 1.362, ossia una variabilità di  $1362/2250.54$ , ossia del 60%, mentre la std delle cellule sane è di 271.13, ossia una variabilità del 43%.

Per trovare l'eccentricità del gruppo di cellule sane e del gruppo di cellule malate ho utilizzato la funzione "regionprops" che fornisce in uscita uno struct array S sull'immagine binarizzata di ciascuno dei due gruppi. S.centroids fornisce l'ascissa e l'ordinata del centro di massa; il gruppo delle cellule sane presenta centro di massa in  $x=520.35$  e  $y=561.33$  mentre il gruppo delle cellule malate presenta centroide in  $x=571.26$  e  $y=441.79$ . Sempre usando "regionprops" si può calcolare che l'eccentricità del gruppo di cellule sane è invece 0.74 mentre per il gruppo di cellule malate è 0.68. L'eccentricità media di ogni cellula sana, calcolata dopo l'utilizzo di "bw2label", è 0.48, con una deviazione standard pari a 0.18, mentre per le cellule malate è 0.57, con una std pari a 0.18.