

**UNIVERSITÀ DEGLI STUDI DI PADOVA**  
**FACOLTÀ DI INGEGNERIA**



**CORSO DI LAUREA IN INGEGNERIA**  
**BIOMEDICA**

*Identificazione non supervisionata di lesioni  
di immagini del fundus*

**Relatore:** prof. Enrico Grisan

**Laureanda:** Maria Serena Crocetta

**Anno accademico 2010-2011**



# **INDICE**

## **SOMMARIO**

## **INTRODUZIONE**

## **CAPITOLO 1 – L’OCCHIO, LA RETINA E LE RETINOPATIE**

1.1 – L’occhio

1.2 – La retina

*1.2.1 – La struttura della retina*

*1.2.2 – Le malattie della retina*

## **CAPITOLO 2 - ALGORITMO PER LA SEGMENTAZIONE DI IMMAGINI RETINICHE**

2.1 – Segmentazione e clustering

2.2 – Algoritmo k-means Iterative Fisher (KIF)

*2.2.1 – K-means*

*2.2.2 – Iterative Fisher Linear Discriminant (iFld)*

*2.2.3 – Binary Hierarchical Kif*

## **CAPITOLO 3 – APPLICAZIONE DELL’ALGORITMO**

3.1 – Le immagini utilizzate e l’oftalmoscopia

3.2 – Applicazione dell’algoritmo

3.3 – Selezione delle classi

## **CAPITOLO 4 – RISULTATI**

## **CONCLUSIONI**

## **BIBLIOGRAFIA**

## **APPENDICE**



# SOMMARIO

L'obiettivo di questo elaborato è il riconoscimento automatico di alcune lesioni della retina, in particolare emorragie, microaneurismi, drusen, essudati duri e molli.

Tale riconoscimento viene effettuato tramite l'applicazione ad immagini retiniche di un algoritmo non supervisionato di clustering, il k-means, nella versione rivista e modificata da Clausi, chiamata k-means Iterative Fisher, che permette di identificare automaticamente il numero di clusters.

Nel primo capitolo vengono proposte considerazioni anatomiche sull'occhio e sulla retina, e viene fatta una panoramica sulle principali retinopatie.

Nel secondo capitolo viene presentato l'algoritmo nelle sue diverse parti, ovvero il k-means, l'Iterative Fisher Linear Discriminant e la scelta della soglia con il Binary Hierarchical Kif.

Nel terzo capitolo vengono presentate le immagini della retina messe a disposizione per questo elaborato dall'Università di Padova, Udine e Trieste, con qualche cenno all'oftalmoscopia. Viene poi mostrato come l'algoritmo opera in dettaglio nelle immagini.

Nel quarto capitolo vengono riportati i risultati dell'applicazione del codice sulle immagini a disposizione, e vengono calcolati degli indici statistici da cui possiamo trarre conclusioni sull'efficienza dell'algoritmo.



# INTRODUZIONE

La storia delle immagini biomediche, iniziata poco più di cento anni fa con la scoperta dei raggi X da parte di W. C. Roentgen, è stata oggetto negli ultimi venti anni di una notevole spinta innovativa con lo sviluppo dell' Angiografia Digitale, della Tomografia Assiale Computerizzata, dell'Ecografia, della Risonanza Magnetica Nucleare e in ultimo della Radiologia Digitale. L'interpretazione delle immagini biomediche, demandata all'occhio esperto di Radiologi e Medici Nucleari, dipende spesso dalla visibilità di importanti dettagli contenuti nell'immagine.

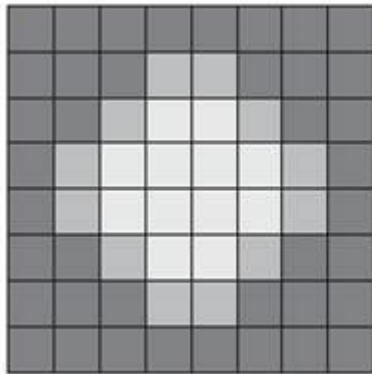
Questi ultimi a loro volta sono soggetti a limiti sia oggettivi che soggettivi. Tra i primi ricordiamo soprattutto la percezione di un determinato reperto, che varia considerevolmente in rapporto all'esperienza del radiologo, alle condizioni tecniche di registrazione dell'immagine e a quelle ambientali di illuminazione, e il rendimento dell'osservatore, che risulta tutt'altro che costante ed estremamente sensibile al variare delle condizioni di osservazione. Tra i parametri oggettivi che intervengono nel limitare la visibilità citiamo soprattutto la risoluzione spaziale, cioè la capacità di una metodica di registrare come separate due strutture situate a breve distanza l'una dall'altra e la risoluzione di contrasto, che è rappresentata dalla possibilità di visualizzare, con diversi livelli di grigio nell'immagine in bianco e nero, due strutture caratterizzate da valori simili di interazione con le radiazioni elettromagnetiche, ionizzanti o con gli ultrasuoni.

Tra i limiti oggettivi dobbiamo includere anche il nostro organo della vista. Infatti, l'occhio umano, a differenza di quanto si possa pensare, non costituisce un sistema di visione perfetto, quindi il suo "modo di vedere" le cose può, in alcuni casi non consentire l'identificazione di particolari importanti nell'interno di una immagine. Fattori sia oggettivi che soggettivi condizionano pesantemente la corretta interpretazione delle immagini biomediche e ne riducono di conseguenza l'accuratezza diagnostica. Da qualche anno c'è stato un ricorso sempre più marcato all'utilizzo della tecnologia informatica e alla progressiva introduzione delle immagini digitali.

La parola digitale deriva dal vocabolo inglese "digit" che vuol dire cifra, numero, perciò "immagine digitale" significa letteralmente immagine "numerica". L'immagine digitale è formata da piccoli quadratini, denominati "pixels" (da "picture elements" = elementi dell'immagine); ciascun quadratino presenta nel suo interno un'unica tonalità di grigio, che varia per i diversi pixels a seconda dei punti dell'immagine. Tutti i pixels uniti assieme formano una matrice bidimensionale o tridimensionale che dà origine all'immagine finale. Ad ogni pixel viene assegnato un numero, il cui valore assoluto corrisponde ad una determinata tonalità nella scala dei grigi. L'utilizzo di immagini digitali consente l'analisi oggettiva ed automatica di diversi parametri attraverso algoritmi di elaborazione (image processing) che facilitano l'acquisizione delle informazioni in esse contenute o, in alcuni casi, le rendono evidenti all'esame dell'occhio umano.

Numerosi algoritmi sono stati sviluppati nel corso degli anni in rapporto alle differenti esigenze applicative. Alcuni di essi tendono a modificare dei parametri intrinseci per consentire l'identificazione di alcune componenti, mentre altri consentono l'estrazione automatica dei valori assunti da determinate grandezze, altri ancora riducono lo spazio occupato in memoria da ogni singola immagine per velocizzare la visualizzazione e la trasmissione a distanza.

Qualunque sia il tipo di algoritmo di elaborazione, la procedura operativa è simile per tutti: i numeri che codificano il livello di grigio di ciascun pixel vengono modificati mediante l'applicazione di una funzione matematica e formano una immagine diversa da quella di partenza.



10	10	10	10	10	10	10	10	10	10
10	10	10	50	50	10	10	10	10	10
10	10	50	90	90	50	10	10	10	10
10	50	90	90	90	90	50	10	10	10
10	50	90	90	90	90	50	10	10	10
10	10	50	90	90	50	10	10	10	10
10	10	10	50	50	10	10	10	10	10
10	10	10	10	10	10	10	10	10	10



# CAPITOLO 1

## L'OCCHIO, LA RETINA E LE RETINOPATIE

### 1.1 – L'occhio

L'occhio è un complesso organo di senso, derivato per evoluzione dalle macchie fotosensibili sulla superficie del corpo degli invertebrati. All'interno della sua struttura protettiva ciascun occhio possiede uno strato di recettori, un sistema di lenti che mette a fuoco la luce sui recettori ed un sistema di neuroni che conduce gli impulsi dai recettori al cervello.

L'occhio è un organo pari e simmetrico, di forma ovoidale situato nella parte anteriore della cavità orbitale; i diametri verticale e trasversale misurano rispettivamente 23 e 23,5 mm. Il peso del bulbo oculare è di 6,5 – 7,5 g; la sua consistenza è duro-elastica. Le principali strutture dell'occhio sono riportate in figura 1.1.1.

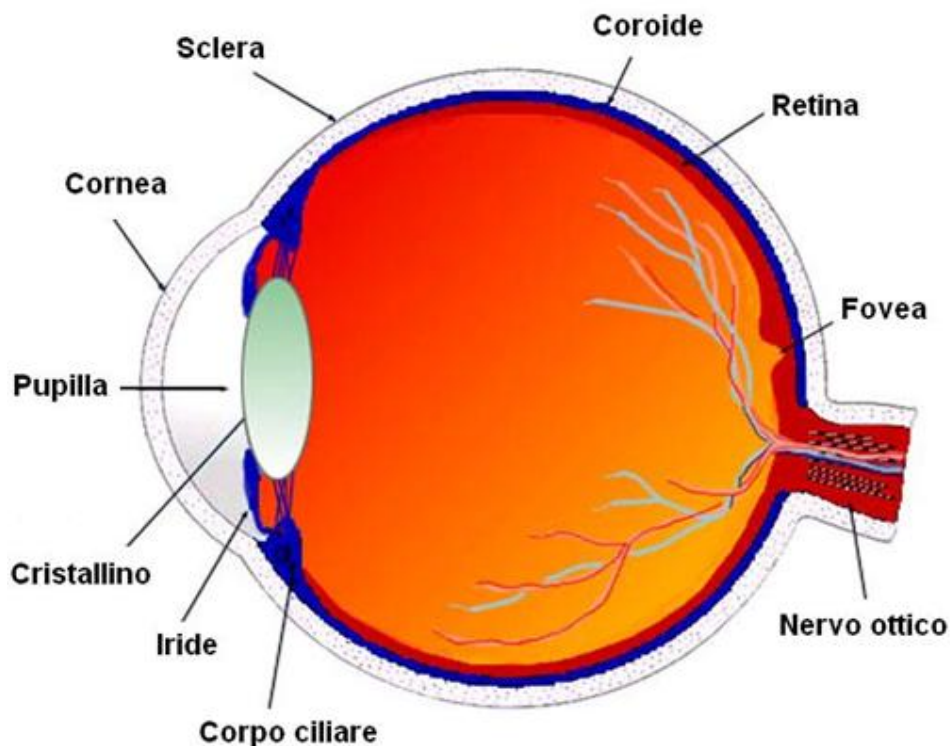


Figura 1.1 – struttura dell'occhio

La **sclera**, strato protettivo esterno del globo oculare, nella parte anteriore costituisce la **cornea trasparente**, attraverso la quale i raggi luminosi entrano nell'occhio. All'interno della sclera si trova la **coroide**, lo strato contenente molti dei vasi sanguigni che nutrono le strutture del globo oculare. I due terzi posteriori della coroide sono rivestiti dalla **retina**, tessuto nervoso contenente le cellule recettrici.

Il **cristallino** è una lente trasparente mantenuta in posizione da un legamento circolare, detto il legamento del cristallino, connesso al **corpo ciliare**, il bordo anteriore ispessito della coroide. Il corpo ciliare contiene fibre muscolari circolari e longitudinali che si inseriscono in prossimità della giunzione sclero-corneale.

Davanti alla lente si trova l'**iride**, struttura opaca e pigmentata che costituisce la parte colorata dell'occhio e che delimita la **pupilla**. L'iride contiene fibre muscolari circolari che restringono la pupilla e fibre muscolari radiali che la dilatano.

Lo spazio compreso tra il cristallino e la retina è riempito principalmente da una sostanza gelatinosa incolore detta **umor vitreo**. L'**umor acqueo**, invece, è un liquido trasparente che nutre la cornea e il cristallino, e viene prodotto dal plasma a livello del corpo ciliare per diffusione e trasporto attivo. Fluisce attraverso la pupilla andando a riempire la camera anteriore dell'occhio e viene normalmente riassorbito attraverso una rete di trabecole nel canale dello Schlemm, un canale venoso situato a livello della giunzione tra l'iride e la cornea.

## 1.2 – La retina

### 1.2.1 – La struttura della retina

La retina è un sottile foglietto spesso circa 0.4 millimetri, che riveste la porzione interna della parte posteriore dell'occhio e si estende in avanti fino alla porzione ottica dell'occhio cioè cornea, iride e cristallino. La retina è composta da 5 tipi diversi di neuroni. I fotorecettori, come dice la parola stessa, sono stimolati dalla luce. La via più semplice e comune che porta le informazioni dall'occhio al cervello passa attraverso i fotorecettori, le cellule bipolari e le cellule ganglionari. Ogni cellula ganglionare ha un lungo prolungamento chiamato assone che attraversa la superficie della retina, arriva al disco ottico per entrare dentro al nervo ottico. Le cellule amacrine e orizzontali servono per arricchire e migliorare il rendimento della retina.

Le cellule dei bastoncelli e dei coni rappresentano i fotorecettori, cioè cellule nervose modificate che sono in grado di recepire gli stimoli luminosi, trasdurli in impulsi nervosi e trasmetterli alle cellule bipolari con cui fanno sinapsi.

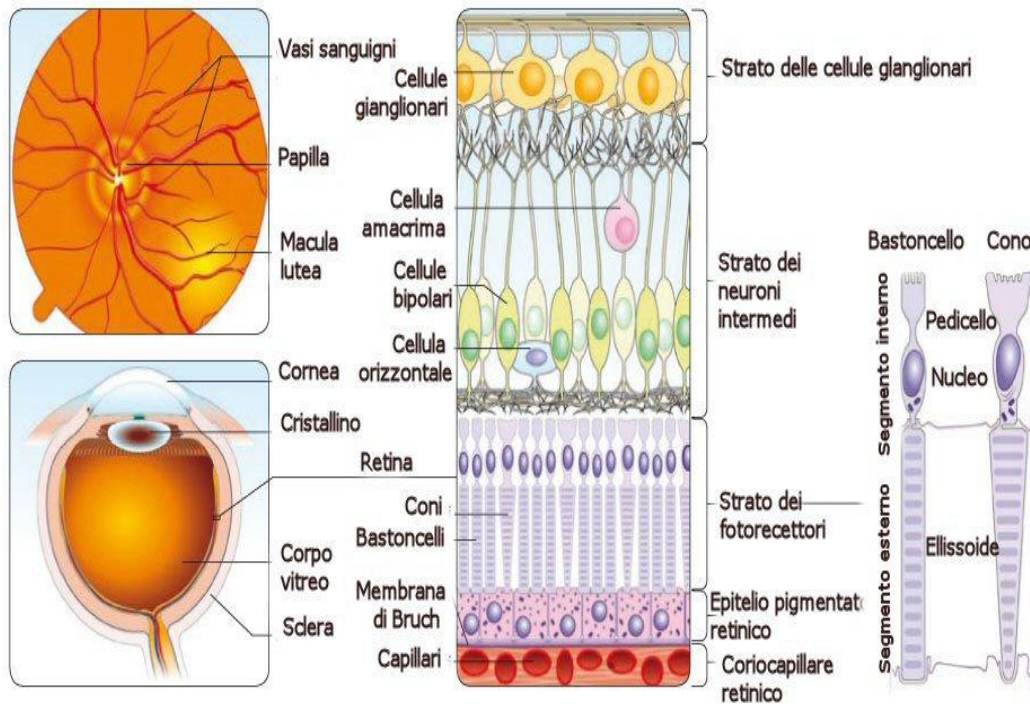
I bastoncelli contengono come pigmento visivo la rodopsina, sono sensibili agli stimoli luminosi deboli e pertanto sono implicati nella visione crepuscolare e in bianco e nero. Non consentono una visione discriminata, in quanto gli impulsi trasdotti da più cellule dei bastoncelli convergono sulla stessa cellula bipolare e vengono quindi trasmessi all'encefalo attraverso la stessa fibra nervosa.

Le cellule dei coni vengono stimulate in condizioni di elevata luminosità e sono implicati nella visione cromatica in quanto possono contenere tre tipi di pigmenti visivi, sensibili a stimoli luminosi di differenti lunghezze d'onda e quindi a tre differenti colori: rosso, verde e blu. La percezione dei colori dipende dalla stimolazione combinata dei tre tipi di coni. Inoltre, i coni sono responsabili della discriminazione visiva in quanto un unico cono si articola con una cellula bipolare e questa con una cellula gangliare; da questa origina una fibra del nervo ottico che, pertanto, conduce all'encefalo un'informazione visiva unitaria.

Le cellule bipolari, con cui i fotorecettori fanno sinapsi, a loro volta fanno sinapsi con quelle gangliari, i cui assoni convergono e lasciano l'occhio formando il nervo ottico. Le cellule orizzontali connettono tra loro i fotorecettori nello strato plessiforme esterno. Le cellule amacrine connettono tra loro le cellule gangliari nello strato plessiforme interno; non hanno assone e i loro processi formano connessioni sia pre-sinaptiche che post-sinaptiche con gli elementi nervosi vicini.

Poiché lo strato ricettivo della retina poggia sull'epitelio pigmentato, contiguo alla coroide, i raggi luminosi devono attraversare lo strato delle cellule gangliari e quello delle cellule bipolari per arrivare ai coni e ai bastoncelli.

La struttura della retina è illustrata in figura 1.2.



**Figura 1.2 – struttura della retina**

L'epitelio pigmentato assorbe i raggi luminosi, evitandone la riflessione, che renderebbe confuse le immagini visive, attraverso la retina.

Gli elementi neuronali sono tenuti insieme da cellule gliali dette di Muller, i cui processi formano una membrana limitante interna sulla superficie interna della retina ed una membrana limitante esterna nello strato dei recettori.

Il punto di uscita del nervo ottico e di entrata dei vasi sanguigni retinici è situato mediamente a 3 mm e un po' al di sopra del polo posteriore del globo oculare. Questa zona, visibile all'oftalmoscopio, si chiama papilla o disco ottico; è priva di fotorecettori e pertanto è cieca (punto cieco). Al polo posteriore dell'occhio si trova una piccola area pigmentata giallastra, detta macula lutea e al centro di questa vi è la fovea centrale, una parte della retina meno spessa, priva di bastoncelli e ricca di coni, molto addensati tra loro, presente nell'uomo e nei primati. In quest'area ci sono pochissime cellule e nessun vaso sanguigno sovrastanti i recettori e di conseguenza la fovea è il punto della retina dove l'acuità visiva è massima. Quando un oggetto attrae la nostra attenzione, gli occhi si muovono in modo che i raggi luminosi provenienti dall'oggetto vadano a cadere sulla fovea.

Le arterie, le arteriole e le vene che si trovano negli strati superficiali della retina più vicini al vitreo, possono essere viste all'oftalmoscopio. Poiché questo è l'unico luogo dell'organismo dove le arteriole sono facilmente visibili, l'esame oftalmoscopico è di grande valore nella diagnosi e nella valutazione del diabete mellito, dell'ipertensione, e di altre malattie che possono alterare i vasi

sanguigni. I vasi retinici irrorano le cellule bipolari e gangliari; i recettori invece sono nutriti in gran parte dal plesso capillare della coroide, e per questo soffrono molto in caso di distacco della retina.

## 1.2.2 – Le malattie della retina

### **RETINOPATIA DIABETICA**

La retinopatia diabetica è una patologia retinica causata dalle alterazioni dei piccoli vasi ematici, a cui vanno incontro i pazienti affetti da diabete. La sua incidenza è andata crescendo in relazione all'aumento della durata media della vita del diabetico ed è passata dal 15% di 50-60 anni fa al 50% e più degli ultimi anni.

La malattia diabetica produce alterazioni delle pareti dei capillari retinici, con fuoriuscita di liquido dalle pareti alterate ed alterazioni occlusive, con conseguente sofferenza dei tessuti per ridotto apporto ematico (ischemia retinica).

La retinopatia diabetica (R.D.) si può schematicamente suddividere in non proliferante e proliferante.

La R.D. non proliferante è caratterizzata da varie lesioni retiniche che nella fase iniziale della malattia costituiscono la cosiddetta *retinopatia diabetica background*:

- microaneurismi: sono delle estroflessioni della parete dei capillari della retina e hanno l'aspetto di piccoli punti rossi. A seguito del riflesso rosso della retina, solo i microaneurismi più grandi possono essere visti all'oftalmoscopio
- emorragie retiniche: le forme più comuni sono le petecchie emorragiche situate negli strati più profondi della retina e che di solito si raggruppano ai margini di aree ischemiche. Un altro tipo di emorragia sono quelle a fiamma, situate a livello dello strato delle fibre nervose, nella parte più superficiale della retina
- edema retinico: col tempo la parete dei microaneurismi si assottiglia e può fuoriuscire liquido che si accumula nello spessore della retina che prende il nome di edema retinico
- essudati duri: sono depositi costituiti da fibrina e lipidi che fuoriescono dai capillari dilatati e dai microaneurismi; hanno un colore giallastro e i bordi 'duri', netti. Sono di solito organizzati a formare degli archi o anelli e sono solitamente localizzati ad una certa distanza dalla macula ma col tempo le si avvicinano. Dato che sono il risultato di una fuoriuscita di liquido dai capillari nello spessore della retina, si associano di solito a edema retinico

In questa fase della R.D. si possono avere alterazioni della regione maculare (la parte centrale della retina, molto importante per la visione), che causano una riduzione visiva più o meno grave: la maculopatia edematosa (l'edema può essere localizzato o diffuso) e la maculopatia ischemica.

Quando le alterazioni ischemiche si fanno più accentuate si ha la cosiddetta *retinopatia diabetica pre-proliferante* caratterizzata da:

- noduli cotonosi (detti anche essudati molli): sono delle aree grigio-bianche a margini mal definiti. La presenza di essudati duri è la conseguenza del restringimento e occlusione delle piccole arteriole della retina
- anomalie microvascolari intraretiniche: sono delle dilatazioni dei capillari retinici in aree di retina dove c'è una diffusa occlusione dei capillari
- alterazioni venose: vene di calibro irregolare
- infarti emorragici: caratterizzati da emorragie retiniche rotondeggianti di colore scuro

L'ischemia retinica porta alla produzione di sostanze che stimolano la crescita di vasi anomali (neovasi): la R.D. proliferante. E' caratterizzata dalle seguenti alterazioni:

- neovascolarizzazione: si manifesta come conseguenza dell'ischemia retinica ed è l'aspetto più importante della retinopatia diabetica proliferante. I vasi neoformati crescono in modo disordinato e hanno una parete fragile che li rende soggetti a sanguinamenti. I neovasi si possono accompagnare anche ad una proliferazione fibrosa che cresce nella cavità vitreale (lo spazio presente dentro all'occhio) provocando distacchi di retina
- emorragie vitreali: il sanguinamento dei neovasi può essere improvviso e abbondante causando un rapido e grave calo della vista; il sangue può riassorbirsi spontaneamente, ma la retrazione del sangue coagulato può peggiorare l'effetto di trazione sulla retina
- distacco retinico: le trazioni esercitate sulla retina dal tessuto fibrovascolare che avvolge i neovasi e dai coaguli vitreali, producono sollevamenti del tessuto retinico (distacco retinico trazionale) e possono lacerare la retina stessa (distacco retinico regmatogeno).

## **OCCLUSIONE ARTERIOSA E VENOSA**

L'occlusione di un vaso arterioso può avere diverse cause: un ateroma (deposito di grasso che ispessisce la parete interna delle arterie) o un'inflammatione del vaso, un embolo di varia natura che parte dal cuore o dalla carotide, un aumento repentino della pressione oculare (anche per pressione accidentale sul bulbo durante interventi chirurgici).

L'occlusione di un vaso venoso può essere causato da: diabete, aumento della pressione oculare, inflammatione di un vaso o alterazioni del sangue; spesso avviene in persone anziane.

L'ostruzione dell'arteria centrale della retina provoca:

- notevole riduzione della vista (talvolta cecità improvvisa), spesso preceduto da amaurosi fugace(cecità temporanea)
- i vasi arteriosi della retina risultano essere ristretti e la retina assume un aspetto biancastro sul quale spicca il colore rosso della macula "macula rosso ciliegia" (la macula è la zona per la visione distinta)

L'ostruzione della vena centrale della retina provoca:

- diminuzione della vista che raramente raggiunge la gravità dell'occlusione arteriosa (può esserlo se l'ischemia retinica o l'edema della macula è notevole)
- i vasi venosi della retina diventano tortuosi e dilatati con presenza di emorragie retiniche; è presente qualche piccolo infarto retinico (sotto forma di essudato cotonoso), edema della papilla (porzione intra-oculare del nervo ottico) ed edema maculare

## **DEGENERAZIONE MACULARE LEGATA ALL'ETA'**

E' una malattia degenerativa, tipica dell'età avanzata, che colpisce la macula. La degenerazione maculare legata all'età è, nei paesi industrializzati, la principale causa di cecità legale nei soggetti con età superiore a 50 anni. Con l'invecchiamento, questa malattia determina delle alterazioni anatomiche a carico della macula, che rappresenta quella parte centrale della retina che ci permette di vedere nitidamente gli oggetti e le persone che ci circondano. Esistono due forme di AMD: quella atrofica (non neovascolare) e quella essudativa (neovascolare). Entrambe, anche se con diversi meccanismi patogenetici, possono determinare una drammatica riduzione della funzione visiva, con perdita della visione centrale.

- Degenerazione Maculare Atrofica: caratterizzata dalla presenza di depositi giallastri (drusen) nella macula, rappresenta l'85% delle forme di AMD. Con l'avanzare della malattia, si determina un'atrofia dell'epitelio pigmentato retinico (REP) ed una perdita dei

fotorecettori (coni e bastoncelli) tale da determinare una progressiva e grave riduzione della visione centrale con perdita della capacità di lettura

- Degenerazione Maculare Essudativa: può rappresentare un'evoluzione della forma atrofica oppure esordire improvvisamente senza alcun preavviso. Essa rappresenta il 15% delle forme di AMD e consiste nello sviluppo di alcuni neovasi che sollevano la retina sensoriale e provocano una distorsione delle immagini, immediatamente percepita dal paziente, che aumenta progressivamente con l'avanzare della malattia. Quando la forma essudativa evolve in una forma emorragica, il paziente può perdere completamente, e nel giro di pochi secondi, la visione centrale. Con il passare del tempo, se non trattata adeguatamente, la degenerazione maculare di tipo essudativo evolve in una forma cicatriziale che, alterando definitivamente la complessa struttura retinica, rende impossibile il recupero di una visione utile.



## CAPITOLO 2

# ALGORITMO PER LA SEGMENTAZIONE DI IMMAGINI RETINICHE

### 2.1 - Segmentazione e clustering

Con il termine segmentazione (o classificazione) si intende una procedura che permette di associare ciascun oggetto (che sia un'immagine, un pattern, un dato numerico) appartenente a un generico spazio multidimensionale, a una o più etichette corrispondenti alle possibili classi di cui può far parte un oggetto. L'algoritmo che viene approfondito in questo elaborato mira alla suddivisione del dominio  $R : \{x,y\}$  di una funzione discreta di valori di grigio  $G(x,y)$ , ovvero l'immagine di partenza, in  $n$  regioni disgiunte  $R_i$  ( $i = 1 \dots n$ ) in base ad un criterio di omogeneità  $E$  prefissato, in modo tale da soddisfare ai seguenti requisiti:

- $\bigcup_{i=1}^n R_i = R$
- ogni  $R_i$  è spazialmente connessa
- $E(R_i)$  è soddisfatto in ogni  $R_i$
- $E(R_i \cup R_j)$  non è soddisfatto per nessuna coppia di  $R_i$  ed  $R_j$  confinanti

Per ottenere una segmentazione si usano le informazioni riguardanti alcuni tratti salienti (features, caratteristiche) degli oggetti in esame, e li si confronta all'interno di un training set. Nel caso in cui questo training set sia etichettato, e quindi le possibili classi sono note e ogni campione nel set è già associato alla classe di appartenenza, si parla di apprendimento supervisionato; in caso contrario, cioè quando le possibili classi vanno direttamente ricavate dai dati stessi, come nel nostro caso di interesse, abbiamo a che fare con apprendimento non supervisionato.

Con il termine clustering si denota un famiglia di metodi non supervisionati in grado di individuare raggruppamenti intrinseci (cluster) tra i dati nello spazio multidimensionale, e di creare in corrispondenza di tali raggruppamenti le classi. Queste classi sono inizialmente incognite, e non è noto nemmeno il loro numero, cosa che rende il problema decisamente complesso.

I metodi di clustering possono essere ricondotti a due famiglie:

- Clustering gerarchico: attraverso operazioni di tipo bottom-up o top-down organizza i dati in una sequenza innestata di gruppi che possono essere visualizzati come un albero.
- Clustering partizionale: individua (solitamente attraverso algoritmi euristici iterativi) delle partizioni che minimizzano un dato criterio di clustering, ad esempio la somma dei quadrati delle distanze. Di questi metodi fa parte quello che verrà trattato in questo elaborato, il k-means.

## 2.2 – Algoritmo k-means Iterative Fisher (KIF)

### 2.2.1 - K-means

Il k-means è un metodo molto semplice computazionalmente e altrettanto semplice da implementare, sviluppato da MacQueen nel 1967, che basa il clustering sul criterio che minimizza i quadrati delle distanze dai centri delle classi (detto anche criterio di minima varianza). L'obiettivo dell'algoritmo è quello di suddividere gruppi di oggetti in K partizioni sulla base dei loro attributi.

Nella sua versione base può essere così descritto:

- generazione di  $k$  punti da usare come centri dei cluster
- assegnazione di ciascun punto al cluster il cui centro è più vicino, in base al criterio di distanza euclidea: per ogni punto della matrice di partenza  $X$  si calcola la distanza dai  $k$  centri individuati e delle  $k$  distanze ottenute si individua la minore:

$$d_j = \min_{i=1,\dots,k} d_i(x, C_i)$$

- calcolo dei centri dei cluster come media delle coordinate di tutti i punti che vi appartengono
- se le nuove medie sono identiche a quelle calcolate in precedenza, il processo termina, altrimenti si utilizzano le nuove medie come centri dei cluster e si ritorna al secondo punto

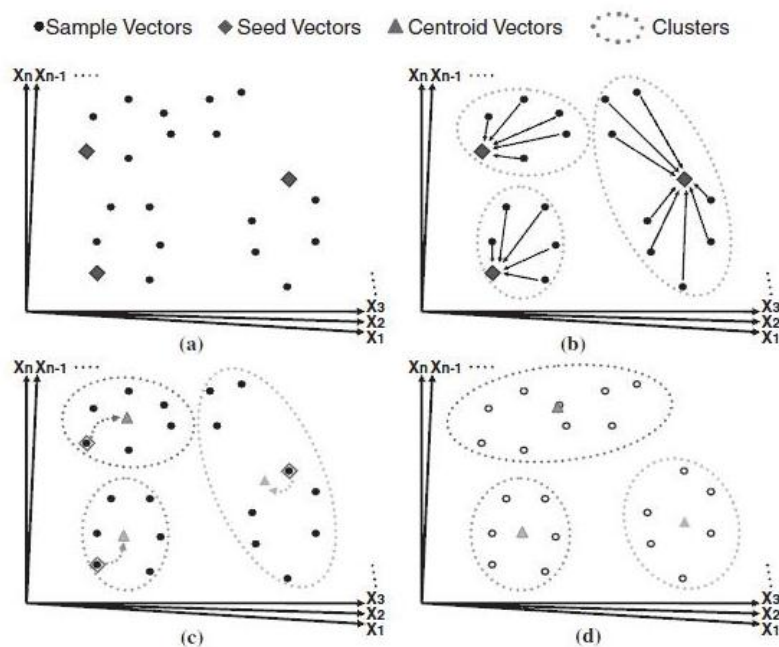


Figura 2.1 – schema di funzionamento del k-means tratto da [3]

In figura 2.1 vediamo i vettori rappresentati come punti nello spazio multidimensionale, in un esempio con  $k=3$ . In (a) vengono scelti casualmente alla prima iterazione i seed vectors, ovvero i centroidi iniziali. Poi vengono calcolate le distanze euclidee di ogni punto dai diversi centroidi, e i vari punti vengono così assegnati a un cluster (b). Successivamente si ricalcolano i centroidi come punti medio in ogni cluster (c). Si procede quindi iterando il processo utilizzando i nuovi centroidi.

Questo algoritmo tende a convergere abbastanza rapidamente e a fornire risultati piuttosto buoni, a patto di partire da una soluzione iniziale ragionevole. Il k-means ha alcuni svantaggi: innanzitutto utilizzando il criterio di metrica euclidea il risultato è accettabile solo nel caso di cluster sferici; inoltre il numero di classi  $k$  deve essere noto a priori.

Una possibile soluzione viene proposta da Clausi in [1], con l'algoritmo k-means Iterative Fisher (KIF), che si sviluppa in due passi: prima il K-means viene applicato ad un gruppo di pixel, poi le due classi che ne risultano sono usate per stimare parametri necessari a migliorare la segmentazione fatta dal K-means, attraverso un procedimento chiamato iterative Fisher Linear Discriminant (iFLD) che sarà descritto nel paragrafo 2.2.2.

Implementiamo inoltre una variante del KIF detta binary hierarchical KIF, che possa stimare anche il numero delle classi presenti dato che il numero di classi non è noto a priori.

### 2.2.2 - Iterative Fisher Linear Discriminant (iFLD)

L'ipotesi di sfericità dei clusters limita la bontà della separazione delle classi, come si può vedere in figura 2.2. Per tenere conto, sotto l'ipotesi di distribuzione normale, dell'intera matrice di covarianza per determinare l'iperspazio di separazione ottima, la stima dei parametri di media e covarianza derivati dal k-means vengono corretti applicando l'FLD, algoritmo proposto da Clausi in [1], iterando il procedimento risultano via via migliori.

In figura 2.2 la linea rossa rappresenta la divisione determinata inizialmente dal k-means, quella blu dopo l'applicazione iterativa del FLD, che assicura una migliore divisione tra i due clusters.

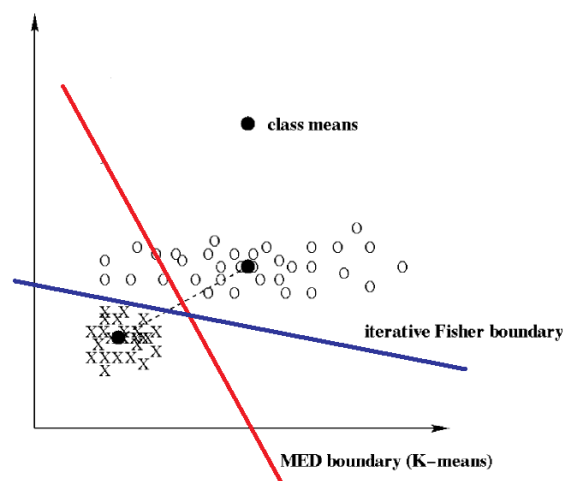


Figura 2.2 – divisione in classi operata da iFLD, tratto da [1]

La matrice di partenza  $X$  è di dimensioni  $n \times m$ , dove  $n$  corrisponde al numero di elementi ed  $m$  al numero delle caratteristiche. Inizialmente viene divisa in due classi per le quali vengono calcolati i parametri di dispersione  $S_b$  e  $S_w$  tra le due classi e media delle due classi, che sono determinabili come

$$S_b = (m_1 - m_2)(m_1 - m_2)^T$$

$$S_w = S_1 + S_2 \quad \text{con} \quad S_i = \frac{1}{N_i - 1} \sum_{x \in C_i} (x - m_i)(x - m_i)^T \quad i=1,2$$

dove  $m_1$  e  $m_2$  sono i centri delle due classi.

Possiamo così trovare  $\omega$ , vettore di dimensioni  $m \times 1$ , con  $m$  numero delle caratteristiche:

$$\omega = S_w^{-1}(m_1 - m_2)$$

Possiamo così calcolare l'indice di separazione delle classi  $\tau$  definito come

$$\tau = \frac{\omega^T S_b \omega}{\omega^T S_w \omega}$$

L'indice è direttamente proporzionale alla distanza tra i centri delle classi, e inversamente proporzionale alla loro dispersione attorno alle medie.

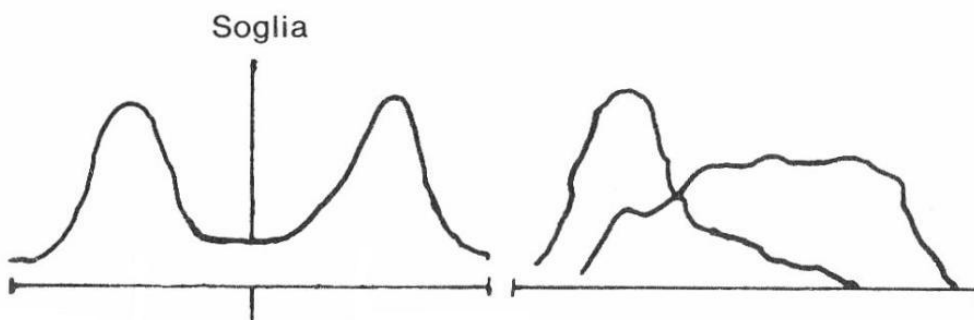


Figura 2.4 – separazione della classi

Nella parte sinistra di figura 2.4 le due classi sono ben distinte, e l'indice di separazione sarà quindi abbastanza elevato. Nella parte destra invece le due classi non sono ben separate, il risultato non è soddisfacente. Per migliorare l'indice  $\tau$  si può individuare la soglia che minimizza l'errore di classificazione considerando i valori di  $y$  minori di  $\tau$  appartenenti a una classe, gli altri appartenenti a un'altra classe; l'indice di separazione così ricalcolato dovrebbe essere maggiore del precedente e di conseguenza migliore perché significa che le classi sono più distanti. Questo procedimento viene iterato per ottenere un  $\tau$  sempre migliore, e si è osservato che con cinque iterazioni si ottiene un buon risultato.

### 2.2.3 - Binary Hierarchical Kif

Il criterio appena visto è un esempio di divisione dei dati in due classi; nel nostro caso però non sappiamo in quante classi dovremmo dividerli. Si ricorre quindi al binary hierarchical KIF, un algoritmo che costruisce un albero binario le cui foglie contengono i clusters individuati. Una classe viene considerata indipendente, e quindi viene associata a un nodo esterno, se l'indice di separazione  $\tau$  è minore di un indice di paragone già fissato  $\bar{\tau}$ ; in caso contrario la classe viene associata a un nodo interno e verrà divisa finché non si troverà che  $\tau < \bar{\tau}$ . Arrivati a questa condizione, il nodo diventa esterno e si procede con la visita pre-order dell'albero.

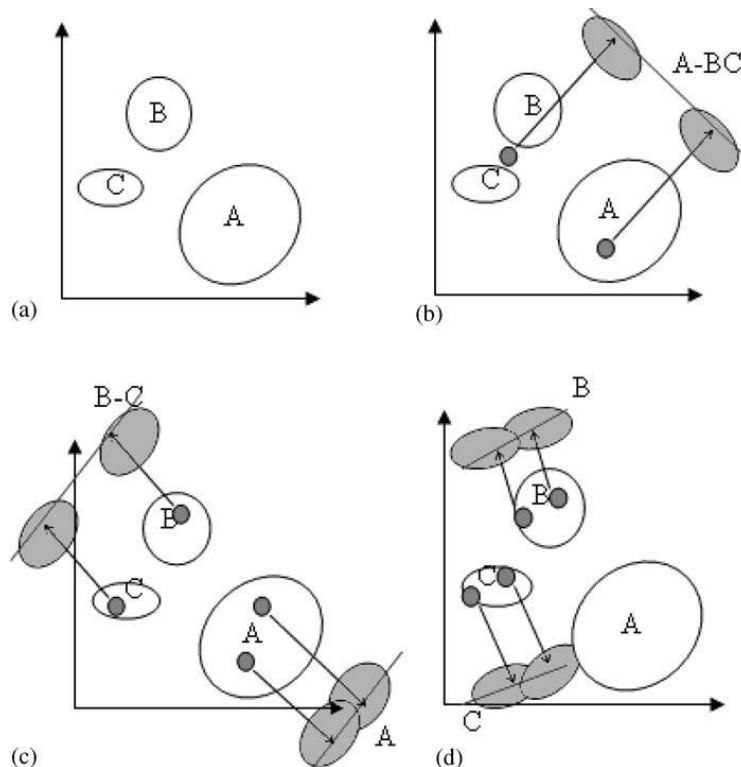


Figura 2.5 –divisione delle classi attraverso il binary hierarchical KIF, tratto da [2]

Vediamo in figura 2.5 il caso di un data-set di tre classi in uno spazio bidimensionale. Tutti i dati inizialmente vengono associati a una sola classe (a). Alla prima iterazione l'algoritmo genera due clusters A e BC (b); secondo l'indice di separazione i due clusters sono indipendenti e di conseguenza i dati vengono divisi in due classi. Quindi l'algoritmo viene applicato a entrambe le classi (c). Sempre considerando l'indice di separazione, A non risulta ulteriormente divisibile al contrario di BC che viene diviso in B e C. Il criterio viene applicato a B e C, che non risultano ulteriormente divisibili, e così il processo termina (d).

## CAPITOLO 3

### APPLICAZIONE DELL'ALGORITMO

#### 3.1 - Le immagini utilizzate e l'oftalmoscopia

Le immagini a disposizione per l'elaborato sono 60 immagini della retina, acquisite attraverso fundus camera nelle cliniche oftalmiche dell'Università di Padova, Udine e Trieste.

L'oftalmoscopia è la tecnica di indagine diagnostica basata sull'esame visivo diretto del fondo oculare effettuato mediante oftalmoscopio. Essa permette di valutare i caratteri della papilla ottica, della regione maculare e della porzione periferica della retina, le dimensioni e lo stato dei vasi retinici; l'eventuale presenza di emorragie, essudati, pigmentazioni o distacchi retinici la densità e la limpidezza dell'umor vitreo. È elemento cardinale della diagnostica oculistica e neurologica. La dilatazione della pupilla (*midriasi*) agevola l'esame del fondo oculare, permettendo di osservare anche la periferia retinica, tuttavia esistono oggi degli strumenti che offrono questa possibilità anche a pupilla stretta.

L'oftalmoscopia può essere diretta o indiretta:

Oftalmoscopia diretta: per esaminare il fondo oculare, è necessario che un fascio di raggi luminosi penetri nella pupilla dell'occhio osservato, si rifletta sulla sua retina e quindi venga raccolto dall'occhio dell'osservatore. Con i moderni oftalmoscopi portatili, come quello riportato in figura 3.1, che sono i più diffusi strumenti per l'esame del fondo oculare, la sorgente luminosa è costituita da una piccola lampadina a forte luminosità, posta ad un'estremità del manico dell'apparecchio, su cui è inserita la testina ottica. Questa è dotata di un foro d'osservazione, al davanti del quale può ruotare un dischetto contenente lenti di diverso potere per la messa a fuoco. La luce della lampadina viene inviata all'occhio osservato per mezzo di uno specchietto posto appena sotto il foro d'osservazione la luce esce da un'apertura ricavata sull'altra faccia della testina, in corrispondenza del foro per l'esaminatore. Con questo sistema la retina appare diritta e ingrandita di circa 15-20 volte, mentre l'ampiezza dell'angolo di retina esplorato è di circa  $10^\circ$ .



Figura 3.1 – oftalmoscopio portatile

Oftalmoscopia binoculare indiretta: si ottiene un'immagine rovesciata, che viene raccolta per mezzo di una lente condensatrice interposta tra occhio dell'osservatore e occhio esaminato. Oggi hanno trovato ampia diffusione a livello specialistico gli oftalmoscopi indiretti binoculari: in pratica, l'oculista indossa un caschetto, come quello di figura 3.2, sul quale sono montati un binocolo e un sistema d'illuminazione, e indirizza verso l'occhio esaminato il fascio di luce prodotto dallo strumento. Se non anteponesse all'occhio esaminato una lente positiva di potere 20-30 diottrie, vedrebbe la pupilla rossa, come in una fotografia scattata di fronte con il flash, ma la lente condensatrice consente di formare l'immagine. Questa è



invertita (cioè l'alto appare in basso e il lato destro a sinistra) ma l'angolo illuminato è molto più ampio che con l'oftalmoscopio diretto, e l'immagine, osservata con entrambi gli occhi attraverso il binocolo, è stereoscopica, permettendo di apprezzare la profondità.

**Figura 3.2 – oftalmoscopio indiretto binoculare**

Aggiungere all'oftalmoscopia la fotografia a colori del fundus o retinografia è utile dal punto di vista clinico e per lo screening in quanto fornisce una documentazione permanente della retinopatia osservata alla oftalmoscopia; risulta pertanto essenziale quando si debbano fare confronti nel tempo delle condizioni fondo oculare e discutere degli effetti del trattamento. La retinografia viene eseguita tramite il retinografo. Per la visualizzazione dei capillari retinici si utilizza un retinografo (fluorangiografo) e tramite l'utilizzo di una fotocamera digitalizzatrice si ottiene la digitalizzazione delle immagini, che in tempo reale vengono archiviate nella scheda PMCI inserita nel corpo della fotocamera digitalizzatrice collegata al retinografo. Successivamente le immagini scelte vengono trasferite nel disco rigido di un PC dedicato per essere elaborate. La digitalizzazione riveste grande importanza poiché consente di elaborare le immagini in modo di evidenziare con grande nitidezza le più piccole alterazioni della rete capillare. Inoltre è di grande vantaggio nei casi in cui la retinopatia necessita di documentazione in tempi molto brevi ai fini del trattamento LASER. Le moderne tecniche di trasferimento dei dati (modem, reti, ecc.), poi, consentono di osservare l'angiogramma in sedi diverse, da quella in cui l'esame è stato eseguito, in tempo reale. In figura 3.3 vediamo un esempio di retinografo digitale.

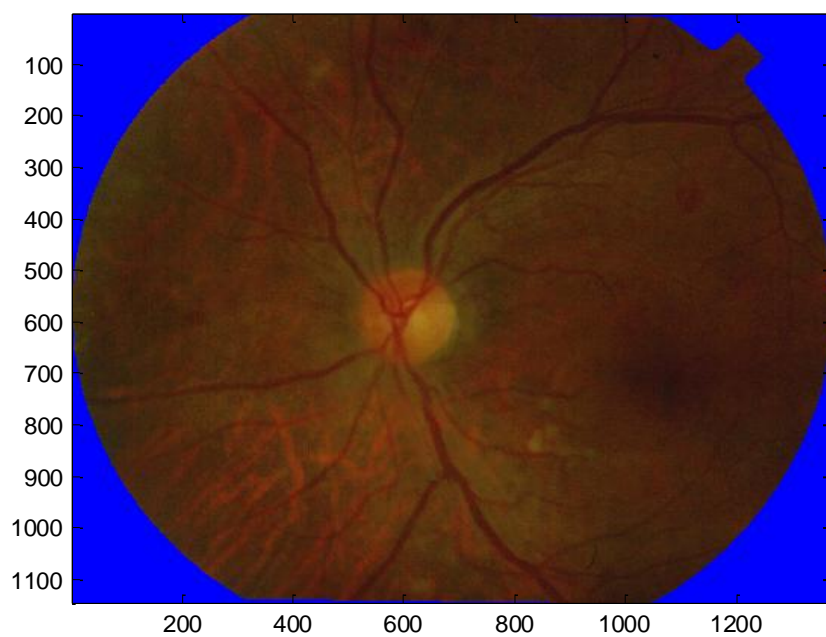




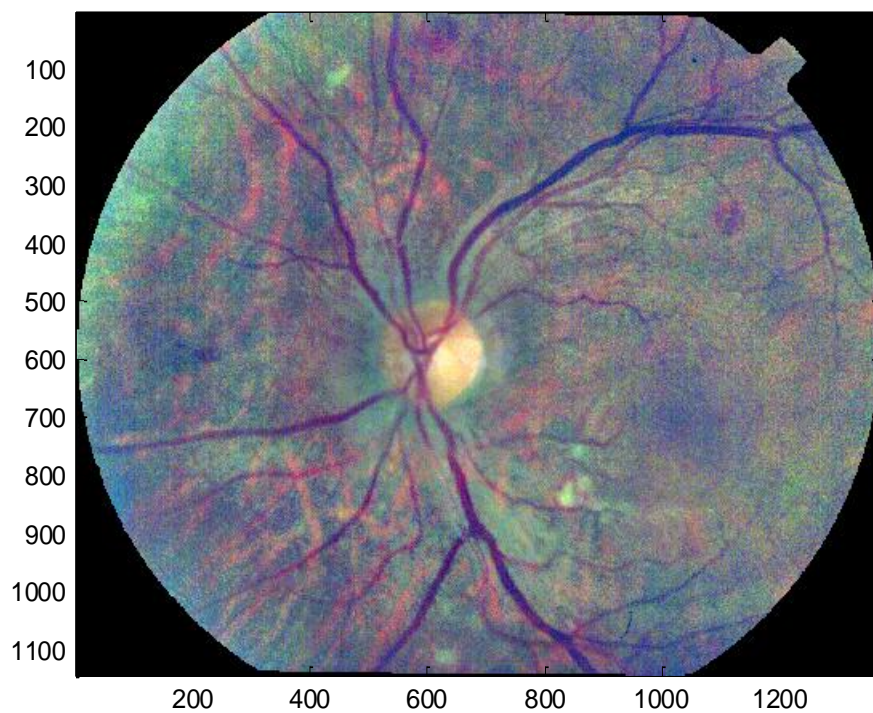
**Figura 3.3 – retinografo digitale**

Le immagini a nostra disposizione acquisite con queste tecniche, come quella di figure 3.4, sono state poi digitalizzate ed elaborate al Dipartimento di Ingegneria dell'Informazione, ottenendo un'immagine come quella di figura 3.5. E' stato necessario modificare illuminazione e contrasto normalizzandoli a valori standardizzati in quanto era presente sia del rumore di fondo che può crearsi in fase di acquisizione, sia variazioni di luminosità e contrasto locali dovuti per esempio all'alto potere riflettente del nervo ottico. Si sono dunque utilizzate per l'analisi queste immagini pre-elaborate.

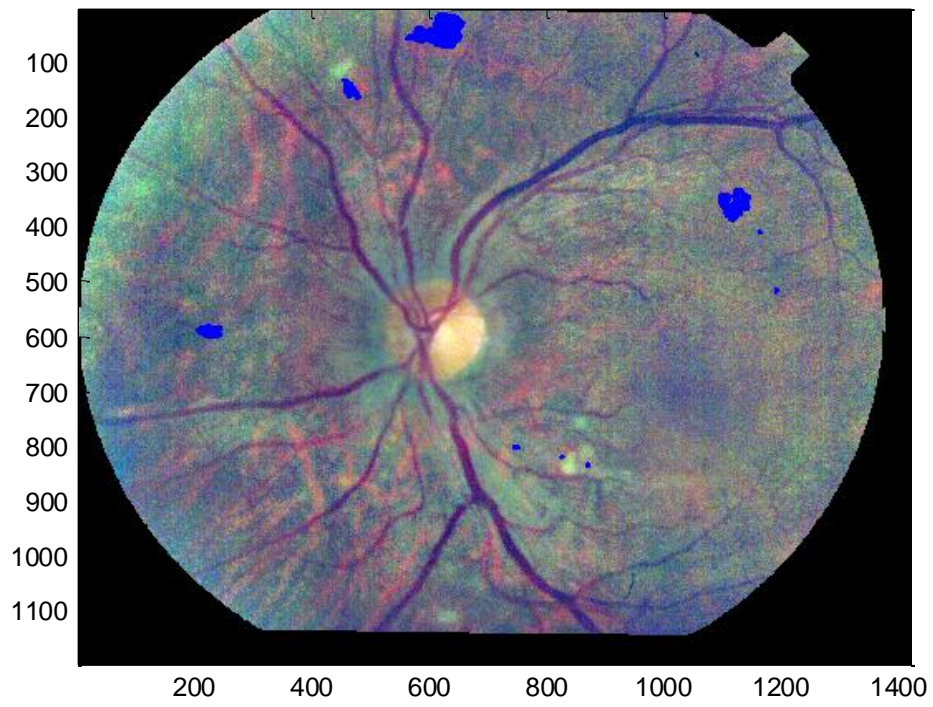
Sono state fornite per il confronto finale le stesse immagini in cui però le lesioni sono state evidenziate manualmente da un medico e delle maschere che rappresentano l'andamento dei vasi sanguigni, identificati automaticamente tramite l'algoritmo descritto nel paragrafo 3.3, come mostrato rispettivamente in figura 3.6 e 3.7.



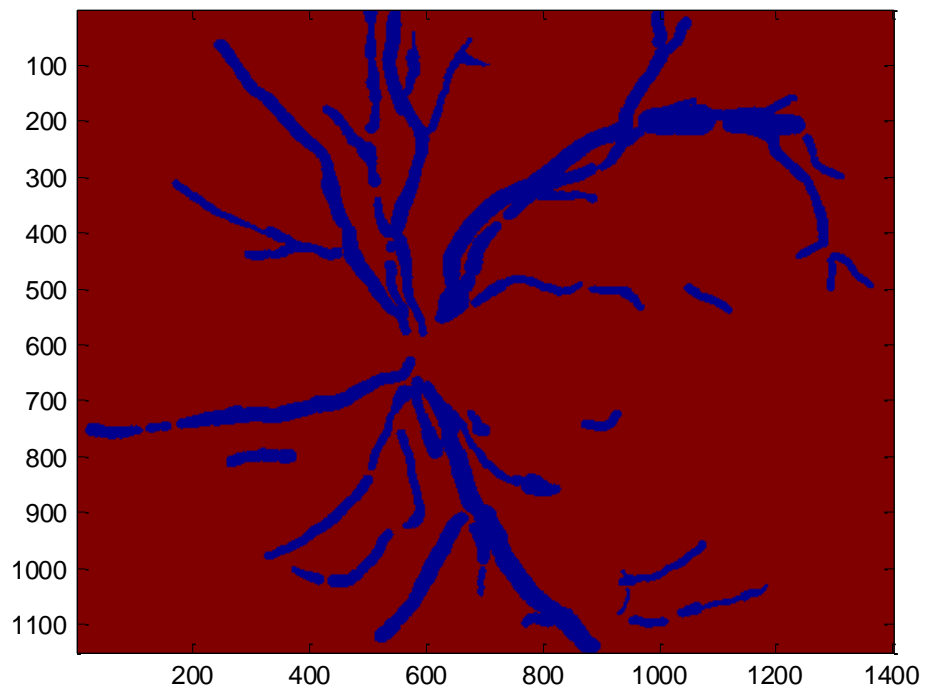
**Figura 3.4 – immagine della retina acquisita con oftalmoscopio e digitalizzata**



**Figura 3.5 – immagine pre-filtrata**



**Figura 3.6 – immagine della retina con lesioni segnate manualmente**



**Figura 3.7 – andamento dei vasi sanguigni**

## 3.2 - Applicazione dell'algoritmo

L'obiettivo del codice, che ha come ingresso un'immagine della retina, è quello di ottenere in uscita una corrispondente immagine che ne evidenzia le lesioni.

Le lesioni che vengono trattate in questo elaborato sono emorragie, microaneurismi, drusen, essudati duri ed essudati molli (detti anche noduli cotonosi). Osservando l'immagine 3.8 in livelli di grigio in scala [0,1], si può facilmente notare che emorragie (indicati in rosso) e microaneurismi, lesioni scure, corrispondono a un basso livello di grigio nel canale verde, mentre drusen, essudati duri (giallo) ed essudati molli (blu) ad alti livelli di grigio.

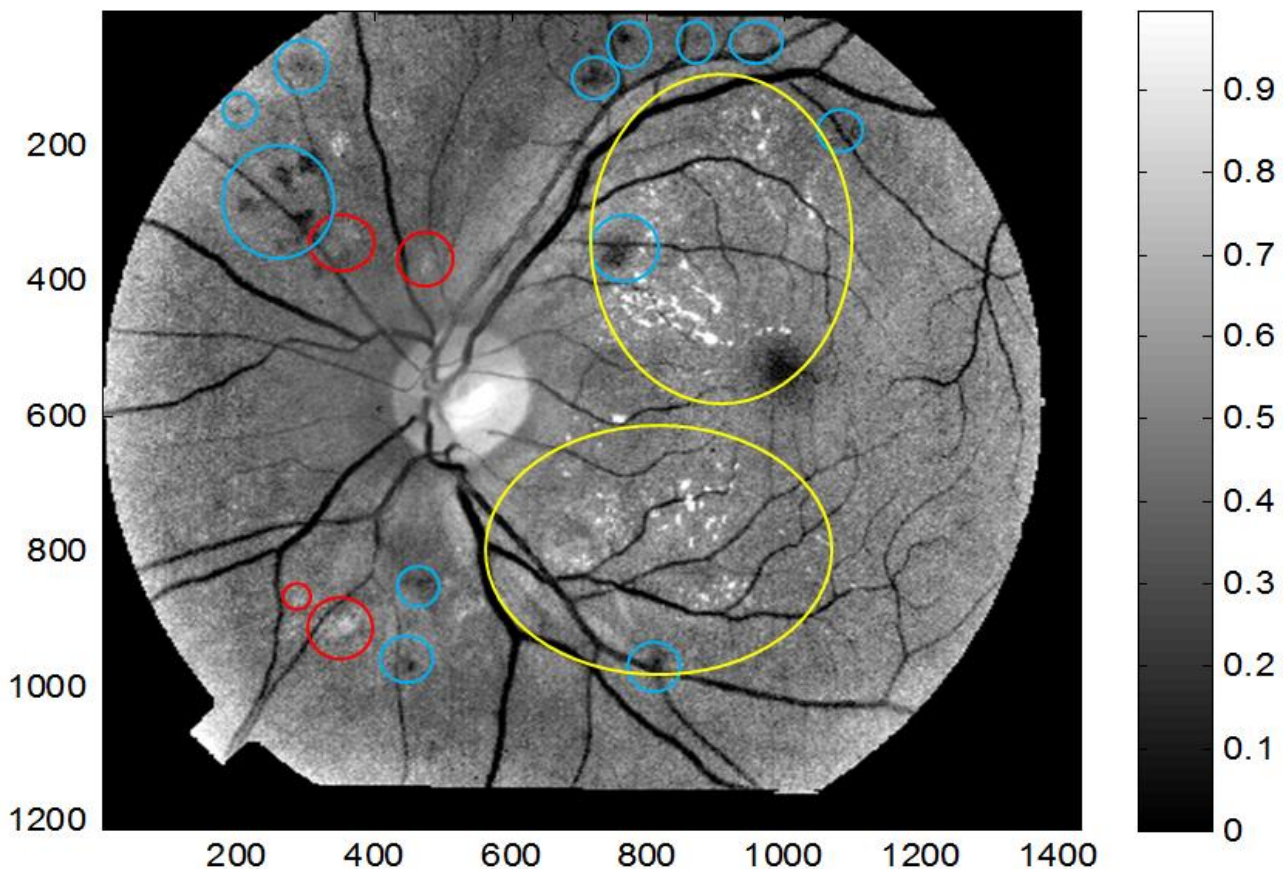


Figura 3.8 – immagine della retina in livelli di grigio

Possiamo notare che due pixel appartenenti a due regioni differenti, per esempio uno allo sfondo e l'altro a un'emorragia, hanno lo stesso livello di grigio. Questo suggerisce la necessità di utilizzare non solo informazioni globali ma anche locali, e per questo motivo abbiamo applicato l'algoritmo, oltre che alle immagini intere, che hanno dimensioni [1152,4113], anche a delle regioni più piccole, in particolare quadrati di lato 300 o 400 pixel contenenti lesioni. Abbiamo poi confrontato i risultati per capire se l'applicazione dell'algoritmo a regioni limitate migliora l'identificazione delle lesioni.



Analizziamo ora il codice che sarà caratterizzato in certi passaggi da un dualismo dovuto alla necessità di distinguere lesioni scure piuttosto che chiare. Riportiamo come esempio le immagini risultanti dall'applicazione del codice su una regione limitata di un'immagine della retina in cui sono presenti alcune emorragie.

### 3.2.1- Preprocessing

Dopo aver acquisito con Matlab l'immagine normalizzata, viene applicato un filtro di Wiener con lo scopo di ridurre il rumore presente. Vengono calcolate per ciascun pixel media  $\mu$  e covarianza  $\sigma$  e successivamente viene calcolata una nuova stima dell'intensità  $\hat{i}$ :

$$\mu = \frac{1}{MN} \sum_{n_1, n_2 \in \eta} i(n_1, n_2)$$

$$\sigma^2 = \frac{1}{MN} \sum_{n_1, n_2 \in \eta} i^2(n_1, n_2) - \mu^2$$

$$\hat{i} = \mu + \frac{\sigma^2 - v^2}{\sigma^2} (i(n_1, n_2) - \mu)$$

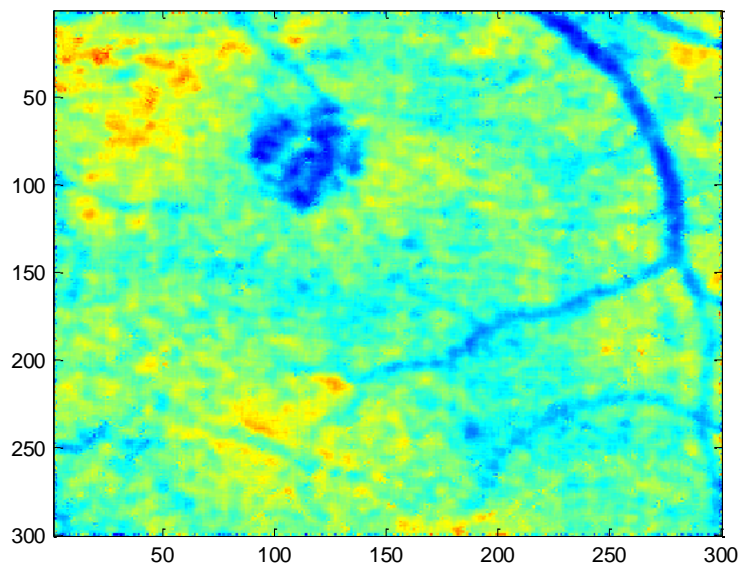


Figura 3.9 – immagine dopo l'applicazione del filtro di Wiener

dove  $v^2$  corrisponde alla varianza del rumore.

### 3.2.2- Features

Per evidenziare meglio le zone chiare o le zone scure delle immagini utilizziamo un filtro minimo/massimo adattivo. Tramite l'uso di una matrice di livelli mediani di grigio, ricaviamo poi un'immagine in cui le zone scure o chiare vengono poste a zero, a seconda del caso di interesse, mentre i pixel non nulli rappresentano di quanto ognuno di essi si discosta dal livello mediano.

Applichiamo filtro minimo/massimo adattivo, cioè un filtro che assegna a un pixel un valore pari al minimo  $m$  (massimo  $M$ ) dei valori dei pixel che rientrano nel suo intorno, scelto arbitrariamente come un quadrato di lato di 5 pixel, se la loro media è più vicina al minimo(massimo).

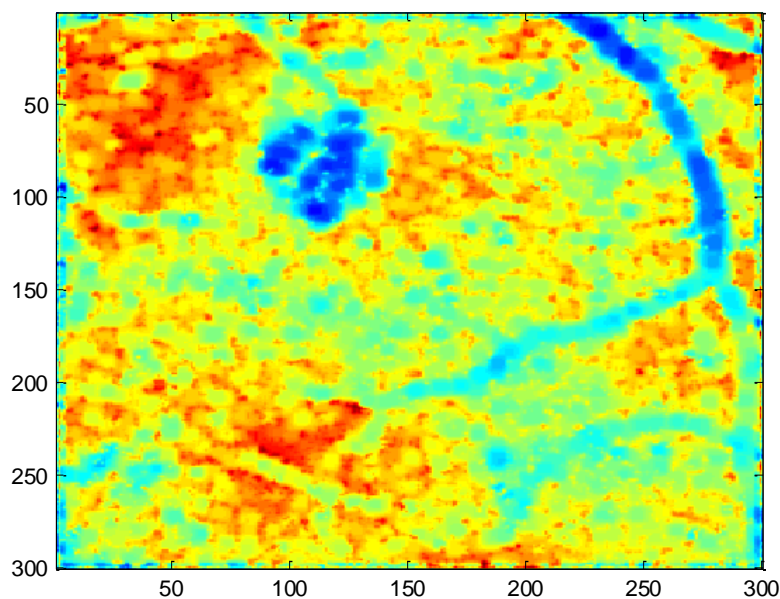
A ogni pixel verrà assegnato un nuovo valore pari a:

$$p^* = \begin{cases} m, & \text{se } p < \frac{M+m}{2} \\ M, & \text{altrimenti} \end{cases}$$

Considerando le immagini con lesioni scure, modifichiamo questo filtro per risaltarle meglio, assegnando il valore M solo ai pixel che hanno un livello di intensità superiore all'80% rispetto al massimo M, ottenendo un'immagine come quella in figura 3.10. Per le lesioni chiare vale il ragionamento contrario, ovvero viene assegnato il valore m ai pixel che hanno un livello di intensità inferiore al 20% rispetto al minimo m.

$$p^*_{\text{scure}} = \begin{cases} m, & \text{se } p < 0.8*(M - m) + m \\ M, & \text{altrimenti} \end{cases}$$

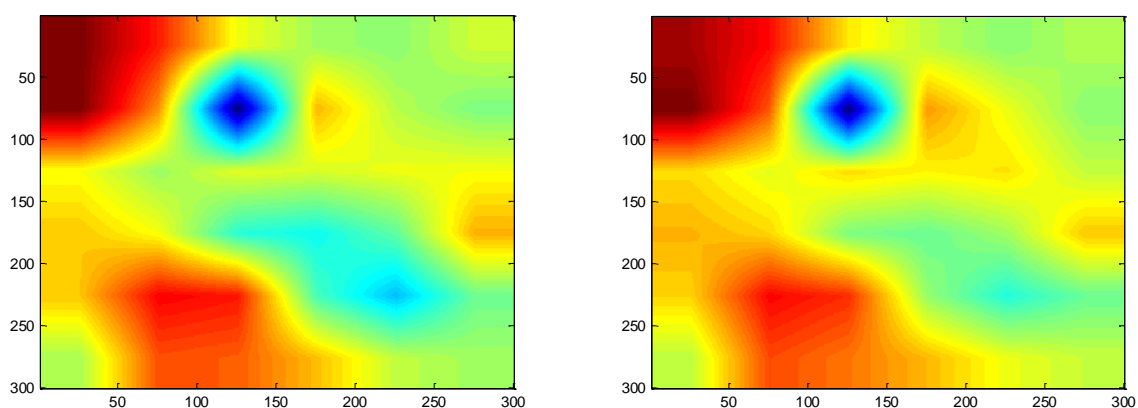
$$p^*_{\text{chiare}} = \begin{cases} m, & \text{se } p < 0.8*(m - M) + M \\ M, & \text{altrimenti} \end{cases}$$



**Figura 3.10 – immagine dopo l'applicazione del filtro minimo/massimo adattivo**

Le successive operazioni verranno applicate a entrambe le immagini in figura 3.9 e 3.10.

Per ogni finestra quadrata di lato 50 pixel dell'immagine, costruiamo una matrice quadrata contenente il livello di grigio mediano corrispondente, e costruiamo così una matrice di mediani come quelle in figura 3.11 e 3.12 che viene poi ridimensionata perché le dimensioni possano concordare. Possiamo in questo modo conoscere l'andamento della luminosità nell'intera area in analisi senza scendere nel dettaglio, quindi non tenendo traccia del singolo livello di grigio contenuto in un pixel ma fornendone un'idea più globale sul suo andamento.



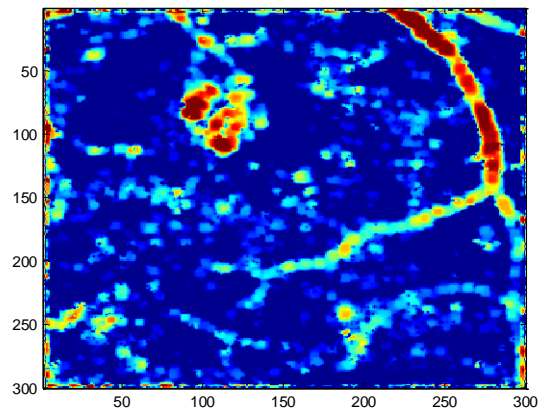
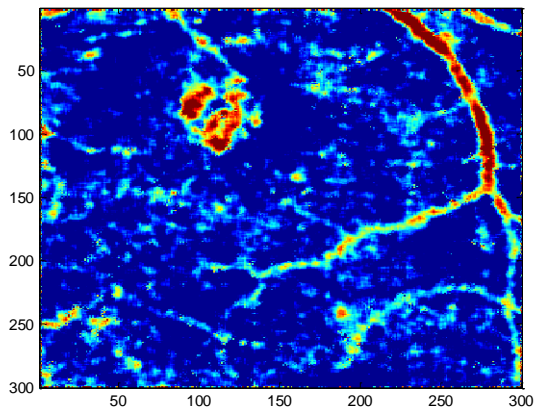
**Figura 3.11 –matrici con livello di grigio mediano - Figura 3.12**

Per evidenziare lesioni con basso livello di grigio si azzerano tutte le parti dell'immagine I che superano il livello  $I_{med}$  di mediano associato, e si assegnano alle altre un livello di intensità che corrisponde a quanto si discostano dal livello del mediano, ottenendo l'immagine di figura 3.13 lavorando sulla 3.9, e la 3.14 sulla 3.10.

$$\text{diff\_b} = \begin{cases} I_{med} - I & \text{se } I_{med} - I \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

Per evidenziare quelle con alto livello di grigio si opera con il processo contrario:

$$\text{diff\_a} = \begin{cases} I_{med} - I & \text{se } I_{med} - I < 0 \\ 0 & \text{altrimenti} \end{cases}$$



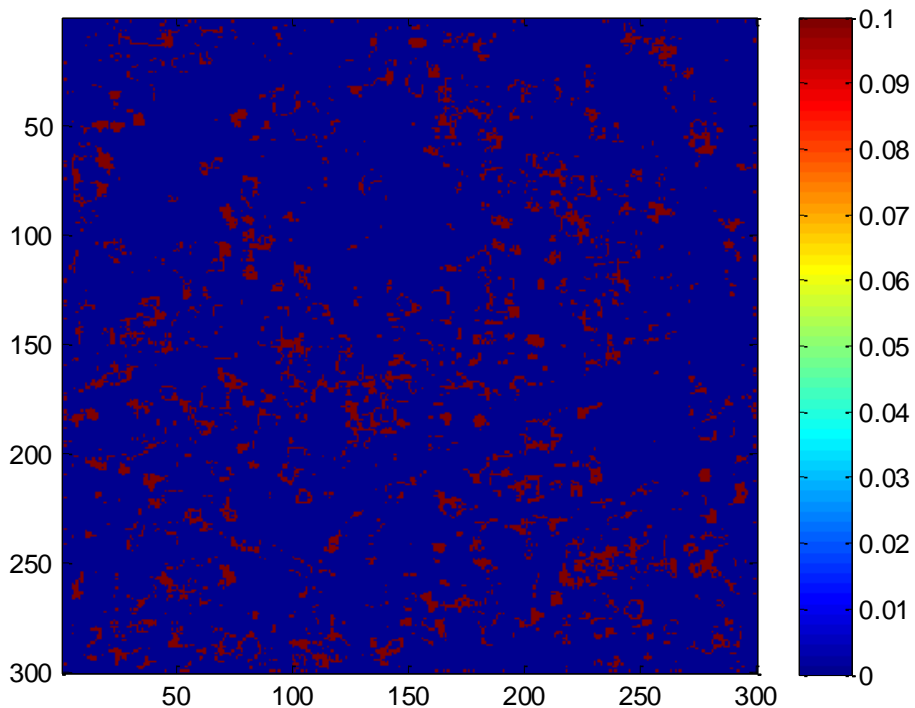
**Figura 3.13 – immagini che rappresentano quanto ogni pixel si discosta dal mediano - Figura 3.14**



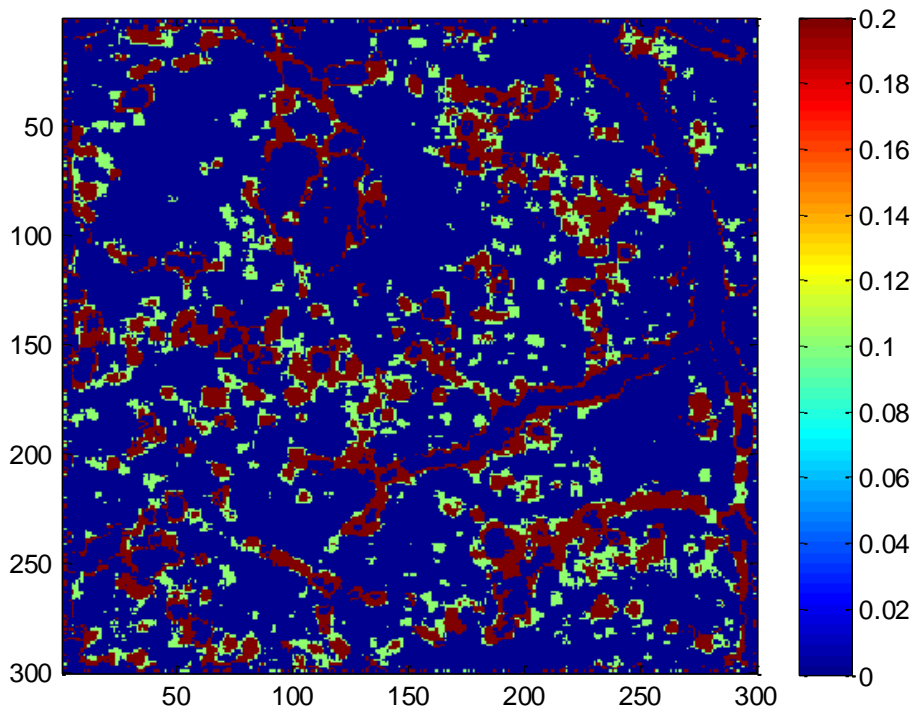
La matrice d'ingresso  $X$  del binary hierarchical kif è costituita da due colonne che contengono la prima i valori diversi da zero contenuti nella 3.13, ottenuta a partire dall'immagine in cui non è stato applicato il filtro adattivo, la seconda quelli diversi da zero presi dalla 3.14.

Si passa quindi all'applicazione di questo algoritmo, che divide i valori di  $x$  in varie classi; un gruppo di pixel viene considerato classe indipendente quando un'ulteriore suddivisione porterebbe ad un  $\tau < \bar{\tau}$ , e quando supera 5000 elementi. Questa condizione è stata aggiunta per evitare il formarsi di un numero elevato di classi, mentre  $\bar{\tau}$  è stato ricavato empiricamente e viene posto uguale a 5. Per rendere l'algoritmo più robusto nel caso in cui le classi siano molto numerose, nel caso non si ottenga alcuna segmentazione con  $\bar{\tau} = 5$ , si prova con  $\bar{\tau} = 6$ , ed eventualmente incrementandolo fino ad 8.

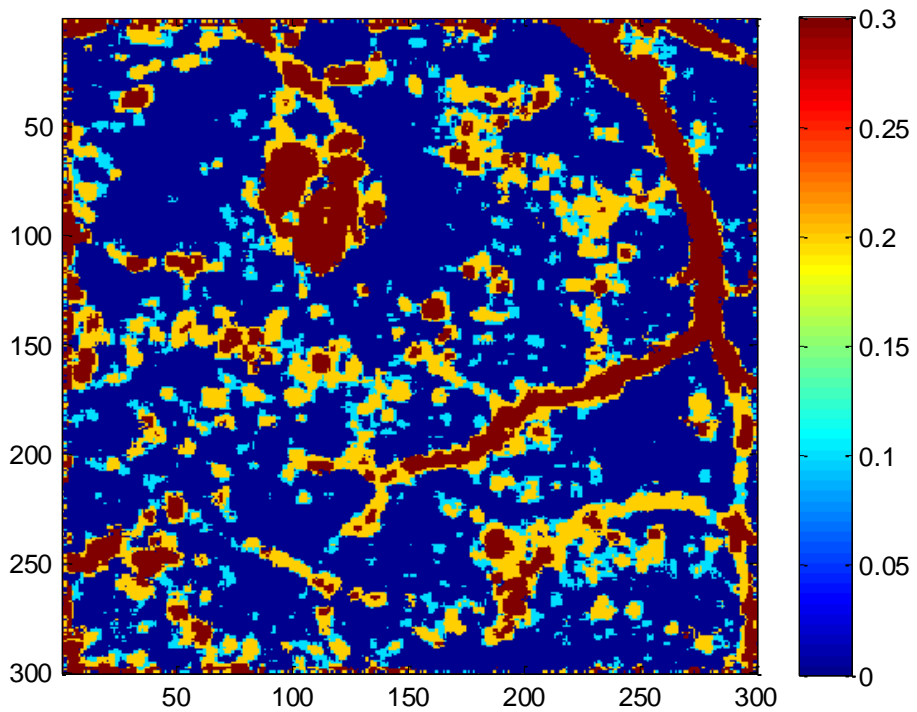
Nelle seguenti immagini possiamo seguire il processo di clustering che divide l'immagine in esame inizialmente in due classi, poi in tre, fino ad arrivare al risultato finale di quattro classi. Il livello di grigio, che parte da 0 per la classe più scura, aumenta progressivamente di 0.1.



**Figura 3.15 – divisione in due classi operata per  $\tau = 3.5$**



**Figura 3.16 – divisione in tre classi operata per  $\tau = 4.1$**



**Figura 3.17 – risultato finale: divisione in quattro classi operata per  $\tau = 5.7$**

### 3.3 - Selezione delle classi

Passando alla selezione delle classi contenenti lesioni è importante fare alcune precisazioni.

Per quanto riguarda la selezione di emorragie, e quindi di lesioni scure, saranno identificati come lesioni anche i vasi, in quanto i pixel corrispondenti hanno un simile livello di intensità. Per ovviare a questo problema viene creata una maschera binaria dei vasi e successivamente moltiplicata per l'immagine contenente le classi, in cui i vasi risulteranno quindi posti a zero.

Inoltre si preferisce evitare di trascurare delle lesioni, a discapito dell'aumento dei falsi positivi; infine i blob contenenti meno di 40 elementi vengono posti a zero in quanto si è riscontrato che solitamente non corrispondono a lesione.

Si considera quindi l'albero binario risultante dal binary hierarchical KIF i cui nodi esterni rappresentano le classi individuate e contengono i pixel ad esse appartenenti. Si è osservato sperimentalmente che le classi contenute nel nodo  $n$  più a destra del sottoalbero di sinistra, (nell'esempio di figura 3.18 in nodo 5) e nei nodi alla sua destra contengono lesioni, con alcune eccezioni:

- se il nodo  $n$  si trova ad un livello maggiore di tre allora viene preso come soglia il nodo  $n-1$ .
- se il numero di componenti contenuti nella classe  $k$  è maggiore di quelli in  $n+1$  allora è quest'ultimo il nodo che farà da soglia.

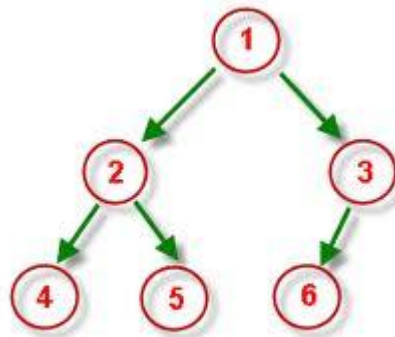
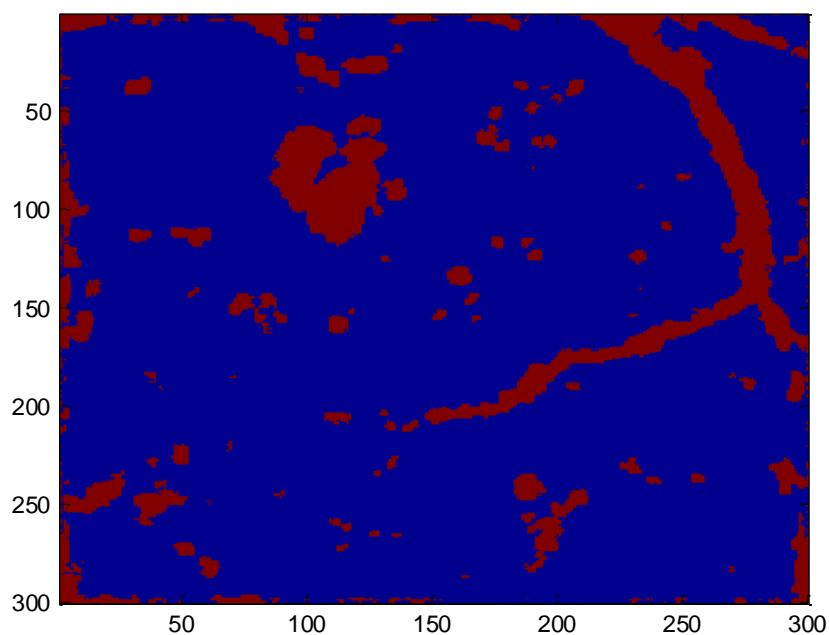


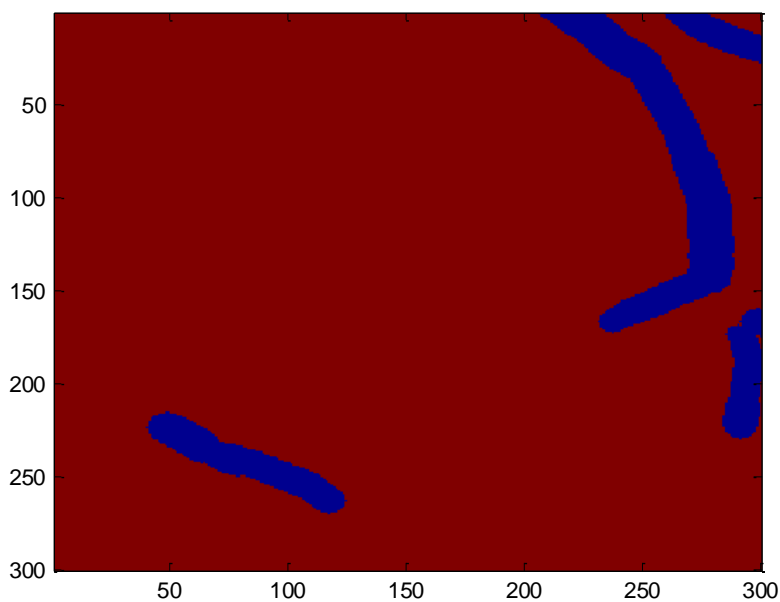
Figura 3.18 – albero binario

Considerando dunque tali classi come segmentazione dell'immagine, otteniamo un'immagine binaria come quella in figura 3.19, in cui vengono posti a 1 (in rosso) i pixel identificati come lesioni.

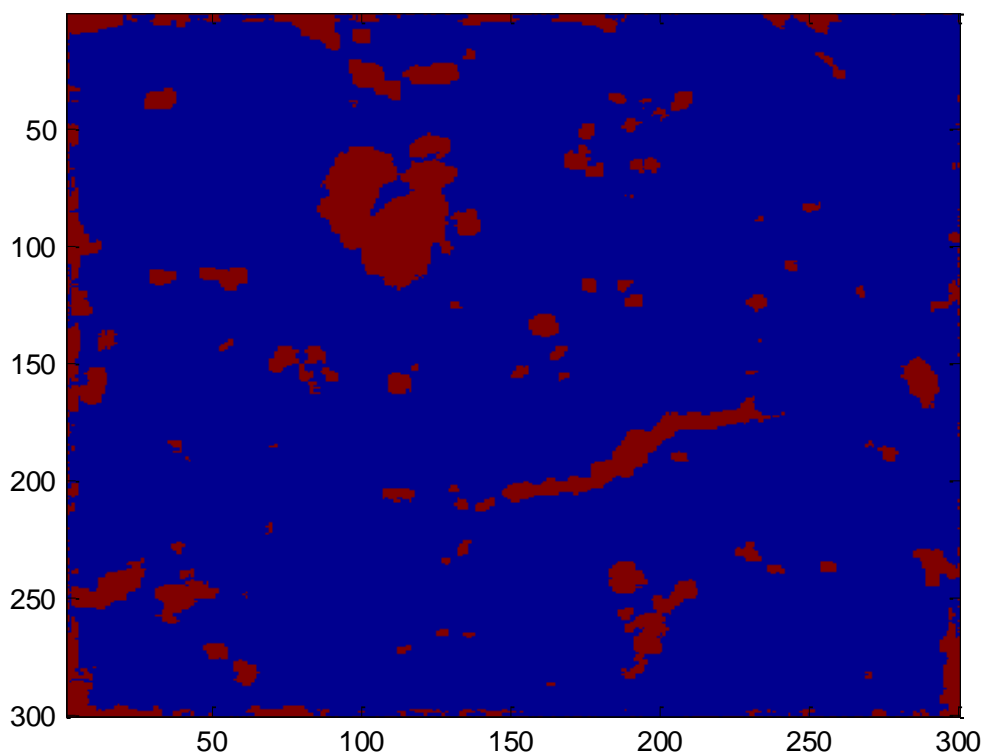


**Figura 3.19 – i pixel rossi rappresentano le lesioni riconosciute dall’algoritmo**

Come si può osservare, nel caso di lesioni scure risultano considerate lesioni anche i vasi sanguigni in quanto hanno un livello di grigio simile ad esse. Per questo motivo vengono utilizzate delle maschere dei vasi, come quella in figura 3.20, dove questi vengono posti a zero; moltiplicando questa immagine per l’immagine sogliata di figura 3.19 , otteniamo come in figura 3.21 l’identificazione delle sole lesioni.

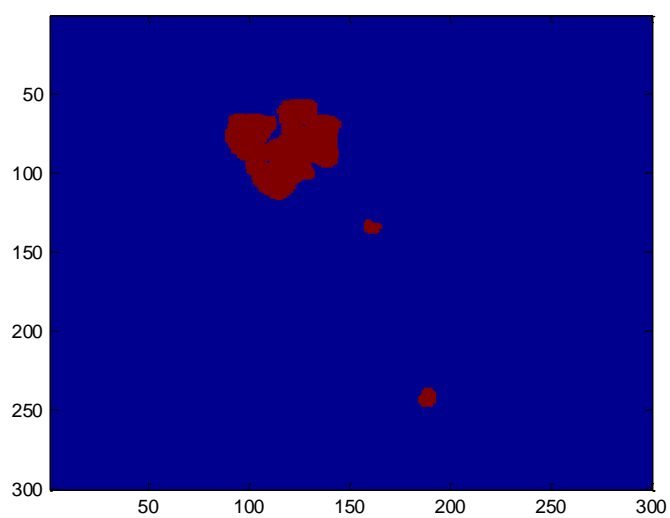


**Figura 3.20 – maschera dei vasi**

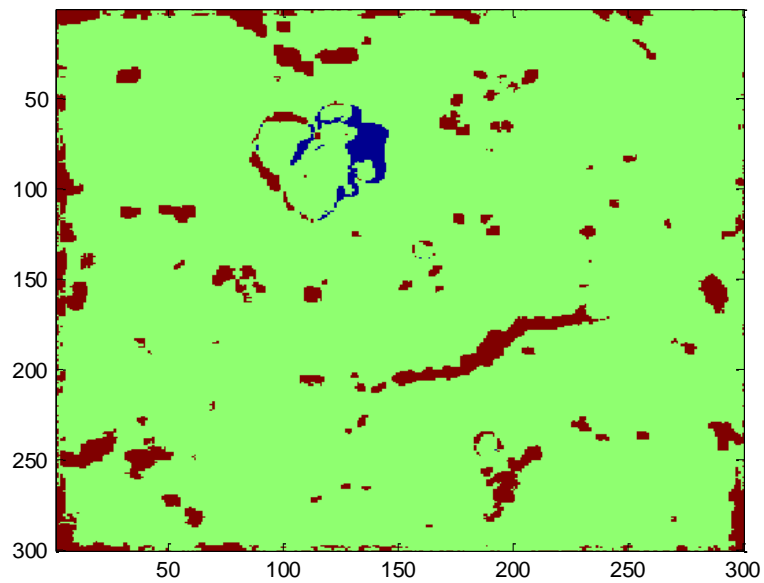


**Figura 3.21 – immagine delle lesioni riconosciute senza considerare i vasi**

S può confrontare quest'immagine con la maschera delle lesioni ottenuta a partire da immagini come in figura 3.22 in cui erano state segnate manualmente le lesioni presenti.



**Figura 3.22 – maschera delle lesioni segnate manualmente**



**Figura 3.23 – differenza tra lesioni riconosciute dall’algoritmo e lesioni realmente presenti**

Quest’ultima immagine, ottenuta come differenza tra l’immagine sogliata e la maschera delle lesioni, è composta da pixel verdi, di valore zero, che corrispondono a lesioni identificate o a zone di sfondo correttamente individuate; pixel rossi, di valore uno, corrispondono a zone di sfondo identificate erroneamente come lesioni; infine pixel blu, di valore meno uno, che corrispondono a lesioni non identificate.

Nel prossimo capitolo verrà presentata una statistica effettuata sulle immagini analizzate, tenendo conto delle considerazioni appena esposte.

## CAPITOLO 4

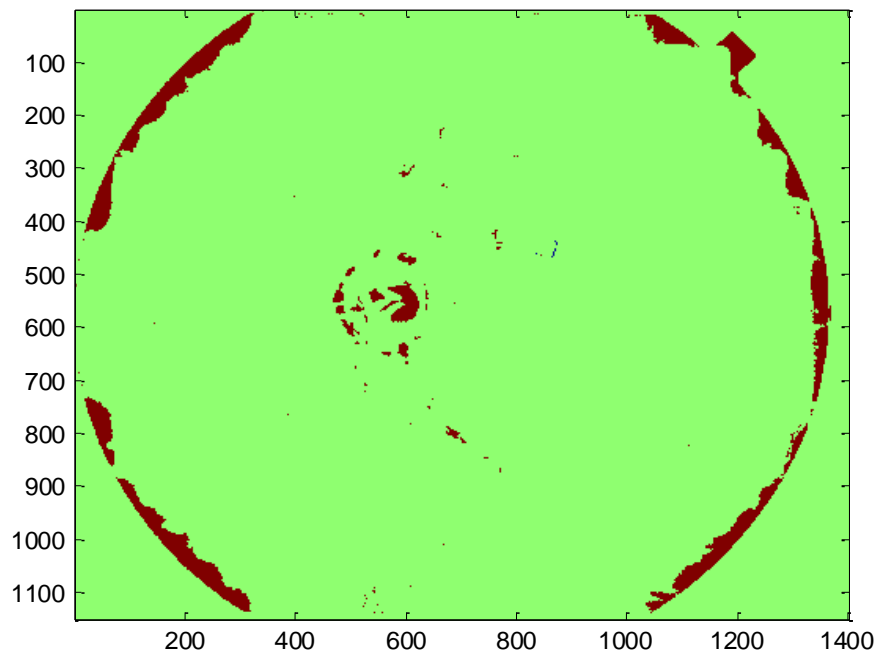
### RISULTATI

Il risultato finale dell'algoritmo è un'immagine binaria che evidenzia i blob che sono stati stimati come lesioni. Facciamo ora una statistica sulle immagini esaminate per capire quanto bene siano state identificate le lesioni, e di conseguenza dare una valutazione all'algoritmo applicato.

Confrontiamo l'immagine risultante dal codice con l'immagine della maschera della lesioni, ottenuta a partire da quella in cui erano state evidenziate manualmente le lesioni, e possiamo così identificare:

- Veri positivi: il blob risultante corrisponde effettivamente a una lesione
- Veri negativi: non risulta alcun blob in corrispondenza di una zona dello sfondo
- Falsi positivi: viene identificata come lesione una zona che non corrisponde a lesione
- Falsi negativi: una lesione non corrisponde a un blob, in altre parole non viene riconosciuta

Per il conteggio di questi indici non si è valutata la corrispondenza tra una lesione e un blob al 100%, ma intorno a una soglia del 60% di lesione riconosciuta. Inoltre per il conteggio dei falsi positivi non si sono considerati come tali i blob con area minore di 20 pixel e quelli sul bordo dell'immagine, particolarmente concentrati nel caso di lesioni chiare, in quanto costituivano evidentemente rumore (figura 4.1).



**Figura 4.1**

Per avere un'idea più precisa sull'area riconosciuta si passa a una valutazione pixel per pixel, ricordando che nell'immagine risultante dal codice i pixel riconosciuti come lesione hanno valore 1:

- Veri positivi: un pixel di valore 1 corrisponde effettivamente a una lesione
- Veri negativi: un pixel di valore 0 corrisponde a una lesione assente
- Falsi positivi: un pixel di valore 1 non corrisponde a una lesione
- Falsi negativi un pixel di valore 0 corrisponde a una lesione

Introduciamo ora gli indici che ci permettono di valutare quanto bene il codice ha identificato le lesioni.

La sensibilità indica la percentuale del rapporto tra le lesioni correttamente identificate e le lesioni totali, e si calcola come

$$\text{sensibilità blob} = \frac{VP}{VP+FN} * 100$$

In altre parole la sensibilità è la percentuale di lesioni correttamente testate come positive.

La specificità indica quanto sfondo viene correttamente segmentato come zona priva di lesione, e si calcola come

$$\text{specificità blob} = \frac{VN}{VN+FP} * 100$$

Essa è pertanto la percentuale di sfondo correttamente identificato come negativo.

Non è possibile determinare il numero di veri negativi, di conseguenza consideriamo solo la sensibilità al blob. Per precisare meglio l'area identificata calcoliamo la sensibilità e specificità d'area considerando non più i blob interi ma ogni pixel:

$$\text{sensibilità di area} = \frac{\text{vera area segmentata che corrisponde a lesione}}{\text{area totale delle lesioni presenti}}$$

$$\text{specificità di area} = \frac{\text{area totale del fondo segmentato come privo di lesioni}}{\text{area totale del fondo privo di lesioni}}$$



Calcoliamo anche il tasso di falsi positivi, che indica quante lesioni riconosciute dall'algoritmo lo sono realmente. Si precisa che si preferisce ottenere dei falsi positivi piuttosto che dei falsi negativi.

$$\text{tasso FP} = \frac{\text{area falsi positivi}}{\text{area veri positivi} + \text{area falsi positivi}}$$

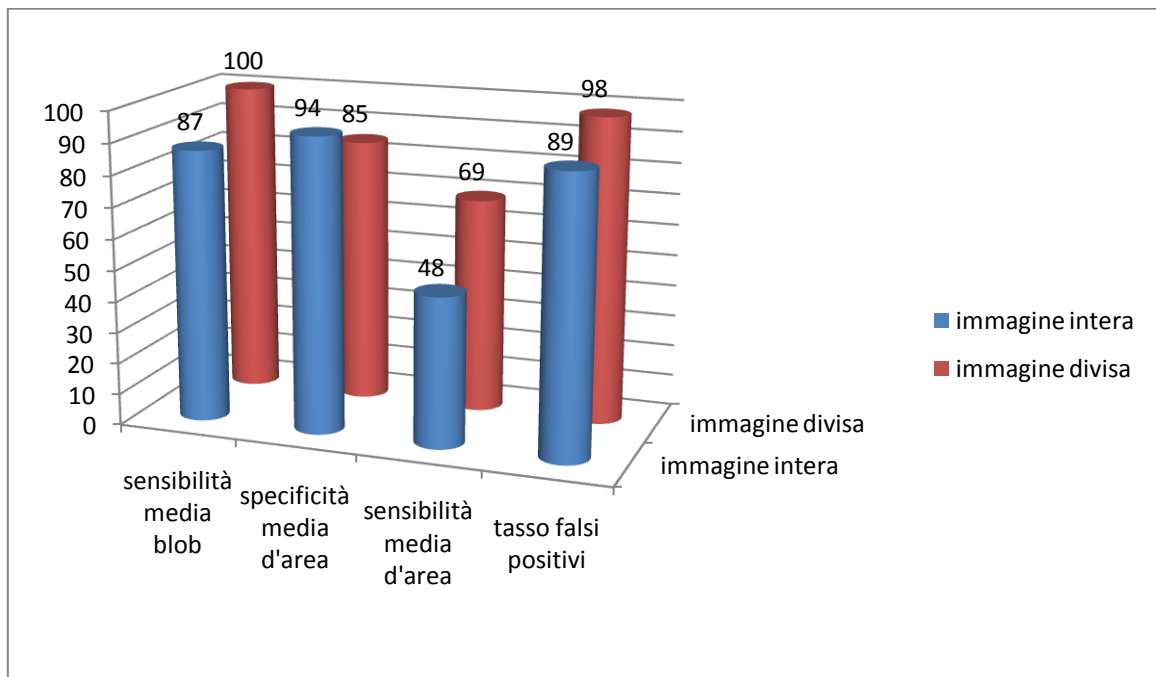
Per quanto riguarda la specificità, al denominatore dovrebbero comparire solo i veri negativi, mentre nella formula considerata compaiono anche i falsi negativi, che però sono in numero trascurabile rispetto ai primi e quindi consideriamo accettabile quest'approssimazione.

Riportiamo in appendice le tabelle in cui sono stati inseriti gli indici calcolati per ogni immagine, divisi per categoria di lesione, e da queste ricaviamo i valori medi di sensibilità del blob, della specificità e sensibilità d'area, e il tasso di falsi positivi:

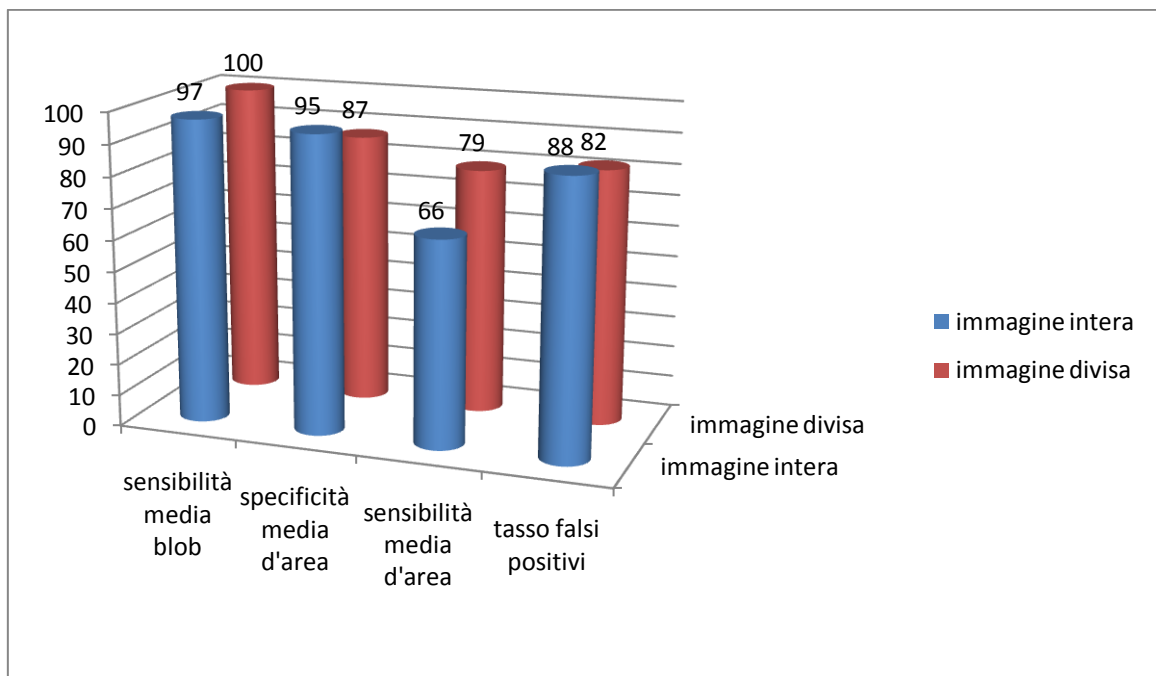
IMMAGINI DIVISE IN REGIONI 300 x 300				
	sensibilità media blob (%)	specificità media d'area (%)	sensibilità media d'area (%)	tasso di falsi positivi (%)
drusen	100	85	69	98
essudati molli	100	87	79	82
essudati duri	92	86	85	85
emorragie	88	97	66	56

IMMAGINI INTERE				
	sensibilità media blob (%)	specificità media d'area (%)	sensibilità media d'area (%)	tasso di falsi positivi (%)
drusen	87	94	48	89
essudati molli	97	95	66	88
essudati duri	87	96	60	87
emorragie	94	97	65	86

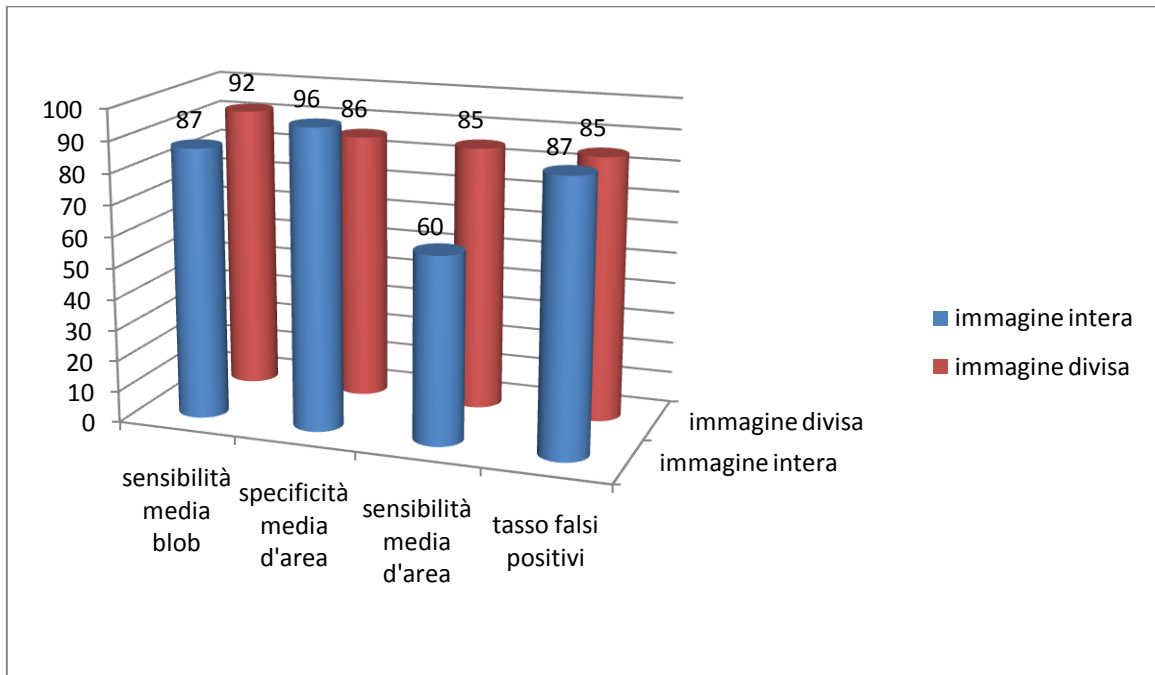
Rappresentiamo ora i risultati in un diagramma a barre:



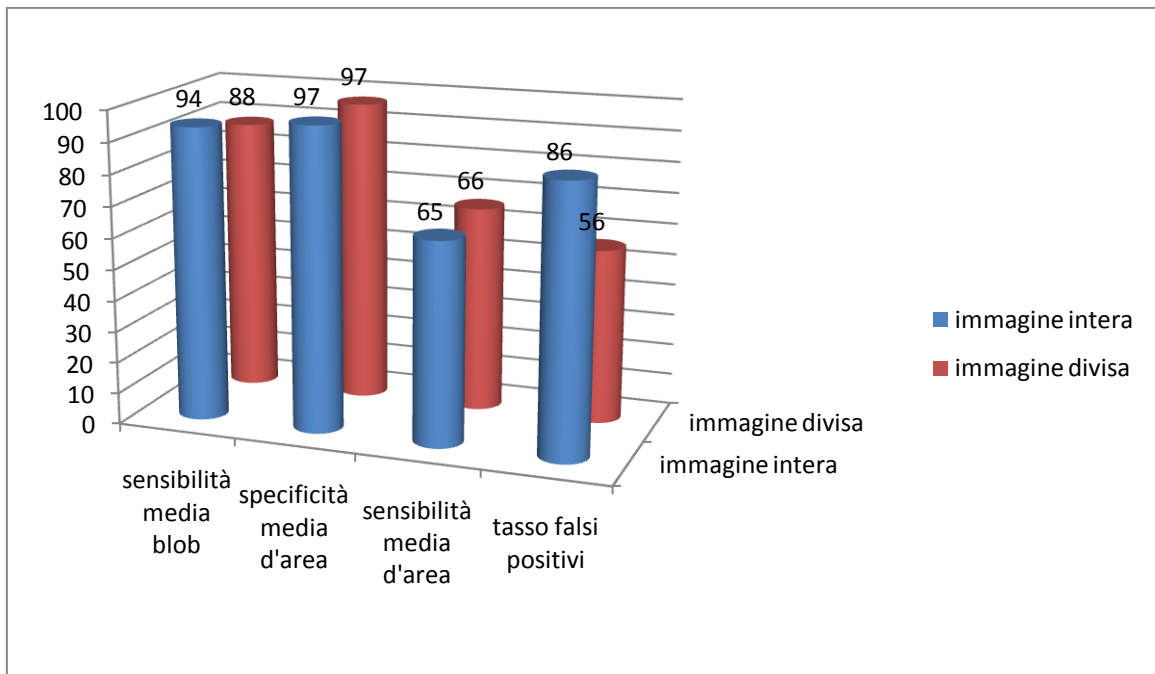
**drusen**



**essudati molli**



**essudati duri**



**emorragie**

La sensibilità è l'indice a cui viene prestata più attenzione, in quanto indica la percentuale di lesioni individuate rispetto alle totali realmente presenti. Notiamo osservando i grafici che, per quanto riguarda l'identificazione di drusen ed essudati molli e duri, tale indice migliora di una percentuale importante se l'algoritmo viene applicato a porzioni più piccole dell'immagine invece che all'immagine intera. Per quanto riguarda le emorragie invece, la sensibilità media del blob risulta un po' migliore se l'algoritmo viene applicato all'immagine intera. Probabilmente il nodo scelto come soglia non è adeguato e converrebbe includere, tra i pixel appartenenti a lesioni, anche quelli nel nodo subito a sinistra di quello selezionato. Il tasso di falsi negativi risulta piuttosto elevato in quanto si preferisce evitare di tralasciare lesioni.

# CONCLUSIONI

In questo elaborato è stata trattata la tecnica del riconoscimento automatico di alcuni tipi di lesioni in immagini della retina tramite l'utilizzo di un algoritmo non supervisionato di clustering, il K-means Iterative Fisher, sviluppato da Clausi.

L'algoritmo è stato applicato inizialmente alle immagini intere. Per ogni immagine si sono ricavati gli indici di sensibilità al blob, specificità e sensibilità d'area per valutare la sua efficienza, e si è calcolata poi una media di questi indici per ogni tipologia di lesione. Si è poi effettuato lo stesso procedimento applicando l'algoritmo su zone più ristrette delle immagini, in particolare quadrati di lato 300 pixel contenenti lesioni. Ogni riquadro è stato trattato come un'immagine indipendente dalle altre, e si è stimata una nuova media degli indici sopracitati per ogni tipo di lesione. Il risultato è che per quanto riguarda le lesioni chiare, la sensibilità è molto migliore nella segmentazione ottenuta a partire da porzioni di immagini. Per quanto riguarda le lesioni scure invece, la sensibilità d'area media risulta lievemente migliore quando il codice viene applicato a parti più piccole di immagini, la sensibilità al blob no.

Non essendo soddisfacenti i risultati riguardanti le lesioni scure, sarebbe interessante analizzare meglio la scelta della soglia tra i clusters contenenti pixel di sfondo e quelli con pixel di lesioni per le immagini con tali lesioni.

# BIBLIOGRAFIA

- [1] D.A. Clausi, **K-means Iterative Fisher (KIF) unsupervised clustering algorithm applied to image texture segmentation**, Pattern Recognition 35 (2002), pag. 1959-1972
- [2] Lap Keung Szeto, **Gene expression data clustering and visualization based on a binary hierarchical clustering framework**, Journal of Visual Languages and Computing 14 (2003), pag 341–362
- [3] Yitao Ma\_ and Tadashi Shibata, **A Binary-Tree Hierarchical Multiple-Chip Architecture for Real-Time Large-Scale Learning Processor Systems**, Japanese Journal of Applied Physics 49 (2010)
- [4] N. Schena, **Sviluppo di algoritmi per l'identificazione di segni diagnostici in immagini retiniche**, Tesi di laurea, Università di Padova, AA 2006-2007
- [5] D. Benozzo, **Segmentazione automatica di lesioni in immagini della retina**, Tesi di laurea, Università di Padova, AA 2008-2009
- [6] W. F. Ganong, **Fisiologia Medica**, Piccin
- [7] D. Perina, **Segmentazione di immagini retiniche sfruttando l'omogeneità cromatica e la connettività spaziale**, Tesi di laurea, Università di Padova, AA 2005-2006

# APPENDICE

Riportiamo le tabelle che riportano i dati delle immagini e gli indici calcolati. Nella prima colonna viene specificato il riquadro di pixel analizzato per le varie immagini, indicate in grassetto.

Riconoscimento di drusen:

	n blob lesioni	VP	FN	FP	sensibilità blob	specificità area	sensibilità area	Tasso di FP
<b>PD09-1</b>								
200-499x550-849	1	1	0	66	100	89	16	98
760-1059x 10-309	1	1	0	43	100	82	100	94
<b>PD17-1</b>								
400-699x350-649	6	6	0	62	100	83	74	98
<b>PD17-3</b>								
750-1049x550-849	3	3	0	75	100	84	97	95
<b>PD19-1</b>								
200-599 x 450-849	13	13	0	59	100	87	60	94

Riconoscimento di essudati retinici molli (o noduli cotonosi):

	n blob lesioni	VP	FN	FP	Sensibilità blob	Specificità area	Sensibilità aera	Tasso di FP
<b>8766-1</b>								
991-1290 x 120-419	5	5	0	9	100	98	48	66
1070-1369 x 350-649	3	3	0	6	100	97	72	62
<b>8766-6</b>								
300-599 x 120-419	2	2	0	46	100	87	80	90
200-499 x 700-999	2	2	0	37	100	88	87	91
<b>PD14-3</b>								
300-599 x 700-999	1	1	0	61	100	88	94	89
<b>PD56-1</b>								
350-649 x 20-319	3	3	0	39	100	89	78	76
950-1249 x 300-599	1	1	0	43	100	85	84	86
<b>PD56-3</b>								
300-599 x 50-349	2	2	0	54	100	88	88	89
450-749 x 835-1134	1	1	0	28	100	85	78	85
650-949 x 700-999	3	3	0	29	100	71	90	93
<b>PD68-1</b>								
300-599 x 650-949	1	1	0	26	100	85	72	82



Riconoscimento di essudati duri:

	n blob lesioni	VP	FN	FP	Sensibilità blob	Specificità area	Sensibilità aera	Tasso di FP
<b>8765-11</b>								
710-1009x150-449	12	12	0	26	100	88	81	89
780-1179x450-849	80	74	6	28	93	85	81	86
350-649 x 10-309	1	1	0	64	100	84	100	91
450-749 x 650-949	5	5	0	59	100	76	99	91
<b>8765-17</b>								
400-699 x 270-569	121	115	6	41	95	84	90	82
400-699 x 570-869	116	112	4	51	97	83	89	83
530-829 x 700-999	63	56	7	21	89	89	70	75
<b>8766-1</b>								
250-649 x 90-489	59	57	2	45	97	85	84	91
290-589 x 490-789	113	112	1	34	99	86	93	90
150-449x790-1089	21	21	0	49	100	86	79	87
<b>8766-6</b>								
770-1069x100-399	73	70	3	23	96	91	64	80
100-999 x 400-699								
640-940 x 600-899	63	61	2	9	97	97	88	81

Riconoscimento di emorragie e microaneurismi:

	n blob lesioni	VP	FN	FP	sensibilità blob	specificità area	sensibilità area	Tasso di FP
<b>8766-1</b>								
291-691x746-1146	8	3	5	4	38	100	38	55
500-899 x 7-406	5	5	0	46	100	95	72	61
950-1350 x 450-850	10	9	1	17	90	98	93	49
1000-1300x150-450	12	9	3	0	75	100	74	31
<b>8766-6</b>								
150-549 x 50-449	11	8	3	30	73	99	54	47
750-1050x800-1100	2	1	1	1	50	100	26	41
<b>PD14-1</b>								
250-550 x 500-800	1	1	0	56	100	95	96	52
<b>PD14-3</b>								
650-950 x 450-750	1	1	0	11	100	99	71	47
<b>PD56-1</b>								
350-650 x 15-315	3	3	0	13	100	99	42	70
350-650 x 200-500	3	3	0	12	100	98	77	77
350-650 x 700-1000	3	3	0	19	100	95	84	72
700-1000x800-1100	1	1	0	2	100	100	75	65
900-1300 x 500-900	4	3	1	47	75	97	39	78
<b>PD56-3</b>								
1000-1299x275-574	3	3	0	64	100	93	74	78
150-349 x 450-749	1	1	0	19	100	94	87	75
650-949 x 700-999	3	3	0	26	100	95	59	69

Esaminiamo ora i risultati del codice applicato questa volta alle immagini intere.

Drusen:

	n blob lesioni	VP	FN	FP	sensibilità blob	specificità area	sensibilità area	Tasso di FP
PD09-1	1	1	0	38	100	94	64	89
PD17-3	3	3	0	24	100	93	70	87
PD19-1	13	8	5	18	62	94	11	90

Essudati molli:

	n blob lesioni	VP	FN	FP	sensibilità blob	specificità area	sensibilità area	Tasso di FP
8766-1	8	8	0	29	100	95	75	89
8766-6	4	4	0	33	100	96	17	85
PD14-3	1	1	0	12	100	96	91	91
PD56-1	6	5	1	23	83	97	49	84
PD56-3	6	6	0	23	100	90	83	85
PD68-1	2	2	0	23	100	96	79	92

Essudati duri:

	n blob lesioni	VP	FN	FP	sensibilità blob	specificità area	sensibilità area	Tasso di FP
8765-11	96	69	27	15	72	97	46	90
8765-17	259	218	41	44	84	96	57	87
8766-1	288	284	4	56	99	93	82	81
8766-6	245	232	13	20	95	96	55	90

Emorragie:

	n blob lesioni	VP	FN	FP	sensibilità blob	specificità area	sensibilità area	Tasso di FP
8766-1	41	38	3	25	93	99	63	82
8766-6	33	26	7	26	79	98	54	75
PD14-1	1	1	0	14	100	99	66	89
PD14-3	1	1	0	19	100	99	71	62
PD56-1	21	19	2	61	90	90	70	87
PD56-3	9	9	0	83	100	95	63	94