

Università degli studi di Padova

Facoltà di Ingegneria

Corso di laurea Triennale in Ingegneria Biomedica

*Classificazione Automatica di immagini
endoscopiche di esofago di Barrett*

Laureando : Ceron Roberto

Relatore: Grisan Enrico

Correlatore: Veronese Elisa

1	Introduzione	1
1.1	<i>La malattia</i>	1
1.1.1	Anatomia dell'esofago.....	2
1.1.2	La mucosa gastrica	4
1.1.3	Mucosa intestinale del colon	7
1.1.4	Esofago di Barrett.....	9
1.2	<i>Strumentazione endoscopica</i>	13
1.2.1	Principi di microscopia confocale	13
1.3	<i>Obiettivo del progetto</i>	17
2	Materiali	19
3	Metodi	21
3.1	<i>Procedura di Cropping</i>	21
3.2	<i>Riconoscimento parte sana</i>	23
3.2.1	Edge detection	23
3.2.2	Quantificazione dei bordi	26
3.3	<i>Classificazione malate : metodo "corr"</i>	27
3.3.1	Introduzione metodo.....	27
3.3.2	Sintesi del codice utilizzato	28
3.3.3	Descrizione del codice utilizzato	28
3.4	<i>Sintesi tra il risultato di classificazione e riconoscimento della parte sana</i>	35
4	Risultati	37
4.1	<i>Riconoscimento parte sana e parte malata</i>	37
4.2	<i>Metodo di classificazione tipo di metaplasia</i>	42
5	Conclusione e sviluppi futuri del metodo "corr"	47

1 Introduzione

1.1 La malattia

E' doverosa un'introduzione prettamente biologica sia per capire la natura della malattia dell'Esofago di Barrett e gli organi e i tessuti che questa coinvolge sia per non confondere tale patologia con una per certi aspetti simili, ossia la METAPLASIA GASTRICA.

Partiamo col chiarire due termini:

- Per METAPLASIA si intende quel fenomeno che consiste nella **trasformazione reversibile** di una cellula o di un tessuto adulto e completamente differenziato, in un tessuto di altro tipo, pure differenziato, in risposta a stimolazioni o a condizioni anomale. Rappresenta cioè l'espressione della capacità, latente nelle singole cellule, di differenziarsi in modo vario in conseguenza di stimoli .
Talora, se le cause che determinano la metaplasia persistono per lunghi periodi, si può verificare la trasformazione neoplastica (lesione pre-cancerosa).
- Con DISPLASIA si descrive un processo proliferativo anomalo, esuberante, circoscritto, che comporta un aumento di volume ed una alterazione nell'architettura della zona colpita. Tutti i processi displastici sono suddivisi, in base alla gravità del sovvertimento strutturale, in grado lieve, medio e severo. La displasia di grado severo prelude allo sviluppo di una forma tumorale maligna. Le displasie sono in genere provocate da stimoli irritativi cronici. Nel caso di epitelio di rivestimento si parla di displasia, indipendentemente dalla presenza di una proliferazione del tessuto, quando è presente un'alterazione dei processi di maturazione e differenziazione cellulare, con perdita dell'orientamento reciproco delle cellule, modificazioni nella forma e nelle dimensioni delle cellule e dei nuclei cellulari, disordine nella successione di strati cellulari procedendo dalle parti profonde a quelle superficiali dell'epitelio.

1.1.1 Anatomia dell'esofago

L'esofago è come tutti sappiamo un tubo muscolare lungo circa 25 cm negli adulti e con una larghezza di 2-3 cm. Collega la faringe allo stomaco e presenta la duplice funzione di consentire i passaggi del cibo fino allo stomaco e di impedirne il reflusso nel senso inverso insieme ad acido e succhi gastrici.

Il lume esofageo è una cavità virtuale, infatti, a riposo presenta una forma stellata per la presenza di pliche¹ longitudinali ossia sollevamenti della lamina muscularis mucosae (MM) o della sottostante sottomucosa SM (Vedi Figura 1.)

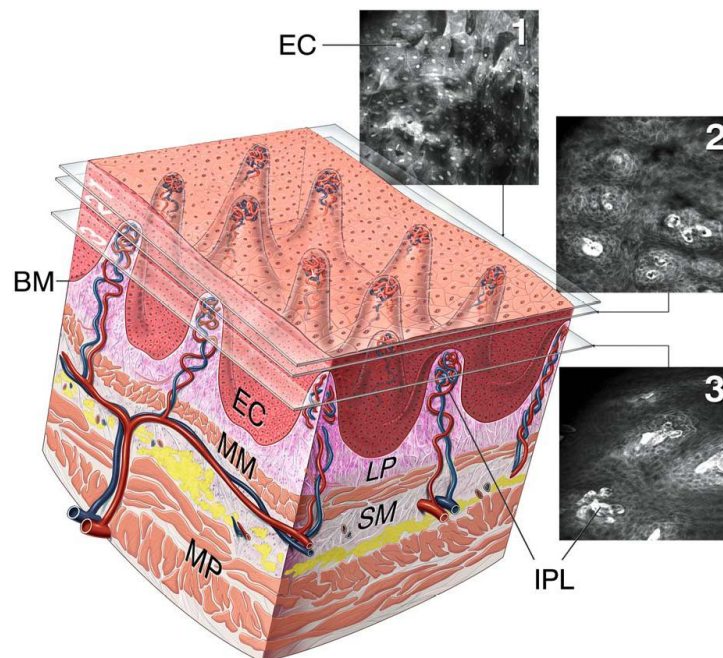


Figura 1 *EC* cellule epiteliali; *BM* membrana basale; *LP* lamina propria; *MM* lamina muscularis mucosae; *SM* sottomucosa; *MP* tonaca muscularis propria; *IPL* villosi intrapapillari.

Sempre riferendoci a Figura 1 le pareti dell'esofago sono formate da strati, o tonache, sovrapposte:

- mucosa
- sottomucosa
- tonaca muscolare
- tonaca avventizia

¹ Ripiegamenti, sollevamenti o inspessimenti del tessuto

La mucosa esofagea è uno strato spesso, di colore grigio-rosato o rossastro nella giunzione gastro-esofagea ; a riposo presenta numerosi solchi e creste, tali da occluderne il lume. L'epitelio della mucosa esofagea (EC) è pavimentoso pluristratificato, non cheratinizzato. Ha uno spessore di 400-500 μm , ed è interdigitato con la lamina propria sottostante (LP). Le cellule epiteliali sono squamose ed appiattite, provviste di nucleo tondeggianti (vedi Figura 2, frecce rosse) e sono unite dalla membrana basale (BM).



Figura 2

La **mucosa** a sua volta è costituita da tre strati:

- l'epitelio di rivestimento (EC)
- la lamina propria (LP) riccamente vascolarizzata
- la muscularis mucosae (MM)

Lo strato più esterno della parete esofagea invece, è **muscolare** (MP), ed è costituito da fibre muscolari di tipo striato (volontarie), da fibre muscolari lisce (involontarie) e miste. Tali fibre sono necessarie durante la deglutizione quando si contraggono, spingendo il cibo nello stomaco. Sono presenti inoltre ghiandole esofagee (*ghiandole cardiaci*), che versano la loro secrezione nel lume dell'esofago necessaria per mantenere umido l'esofago.

La **sottomucosa** (SM) presenta connettivo lasso con ghiandole e fibre nervose che regolano la secrezione ghiandolare e la motilità della muscularis mucosae.

Tutto questo è presente nel caso di una condizione non patologica.

1.1.2 La mucosa gastrica

Come detto inizialmente l'esofago di Barrett non deve essere confuso con un altro tipo di metaplasia che vede il tessuto esofageo trasformarsi in tessuto di tipo gastrico, vediamo quindi alcuni concetti riguardanti lo stomaco e il tessuto che lo costituisce.

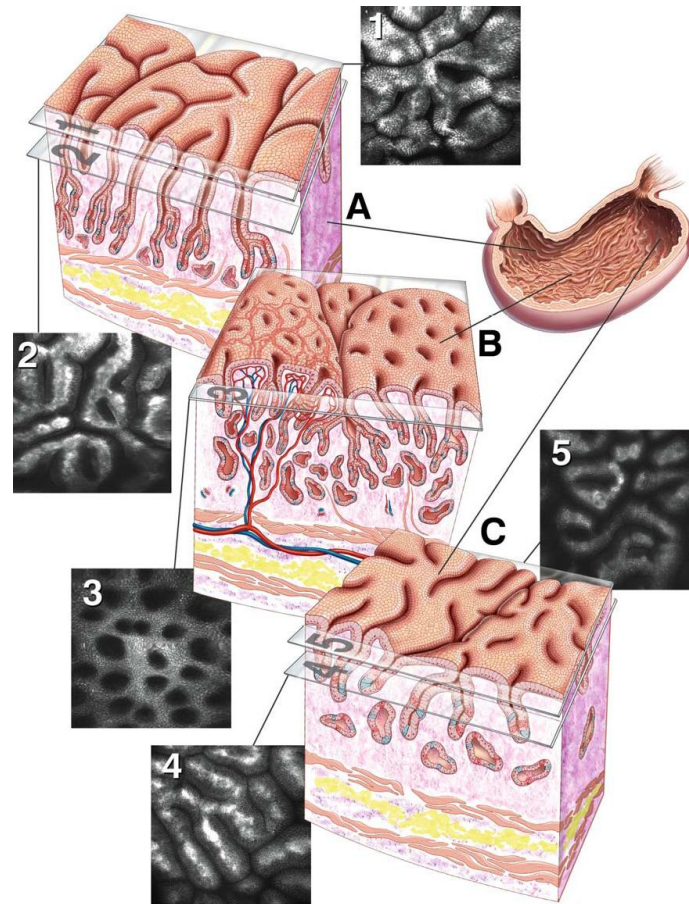


Figura 3: A antro; B corpo; C fondo.

Lo stomaco può essere suddiviso in quattro parti:

- una parte più distale, l'antro (A), è caratterizzata da epitelio foveolare² e ghiandole mucoidi basali
- una parte prossimale, detta corpo (B), tipicamente contiene una piccola porzione di epitelio foveolare assieme ad epitelio ghiandolare presente in quantità maggiore
- una zona di transizione dove il tessuto dell'antro si confonde con quello del corpo
- una seconda zona di transizione, quella del cardias, ove le mucose di esofago distale e stomaco prossimale s'incontrano.

² Con fossette, cioè depressioni sulla superficie di organo

È chiaro che proprio nel cardias si definisce il tratto di transizione di nostro interesse, infatti, se in condizioni normali l'estensione di tessuto tipico dello stomaco prossimale misura pochi millimetri, in casi patologici tale estensione cresce ulteriormente determinando in alcuni casi fenomeni metaplasici.

Il tessuto dello stomaco comunque, come tutti gli organi del tratto gastrointestinale, possiede quattro tonache, rispettivamente, dalla più interna alla più esterna: mucosa, sottomucosa, muscolare e sierosa. Si noti la corrispondenza strutturale con l'esofago.

Seguendo la Figura 1 e con convenzioni simili a quelle usate per descrivere la mucosa esofagea abbiamo che:

- la mucosa dello stomaco è formata da un epitelio superficiale che è a contatto con lume dell'organo, da una lamina propria di tessuto connettivo e dalla muscolaris mucosae. La mucosa si solleva in pliche di forma differente a seconda della zona dello stomaco considerata, alcune di queste sono temporanee altre sono permanenti (per esempio nel corpo), essendo veri e propri ripiegamenti della mucosa.
 - L'epitelio di rivestimento della mucosa è costituito da un epitelio **cilindrico semplice monostratificato con funzione secernente e di rivestimento**, similmente a quello che si ha per la mucosa del colon ma totalmente diverso dalla mucosa esofagea normale. L'epitelio presenta numerose invaginazioni del diametro di circa 0,2 mm e dal lume irregolare, dette fossette gastriche, alla base delle quali si aprono le ghiandole gastriche dello stomaco, che poi si approdano nella lamina propria e riversano il loro secreto nelle cripte delle areole gastriche. In realtà vi è più di un tipo di ghiandole gastriche ed esse differiscono per la struttura a seconda delle diverse regioni dello stomaco che si considerano.
 - La lamina propria è formata da tessuto connettivo lasso e vi si trovano un grande numero di capillari sanguigni. È separata dall'epitelio da una membrana basale.
 - La muscolaris mucosae dello stomaco è uno strato di cellule muscolari lisce, situato al di sotto della lamina propria.
- la *tonaca sottomucosa* dello stomaco è costituita prevalentemente di connettivo lasso, con numerose fibre elastiche e fibre collagene.
- la tonaca muscolare dello stomaco possiede una complessa struttura muscolare. La contrazione muscolare è regolata da una rete di fibre nervose amieliniche localizzate tra gli strati muscolari.
- la tonaca sierosa copre quasi completamente lo stomaco esternamente.

Di seguito sono riportate due immagini endoscopiche che mostrano mucosa gastrica dell'antro e mucosa gastrica del corpo dello stomaco ed effettivamente richiamano immagini endoscopiche di metaplasia gastrica (vedi Figura 4).

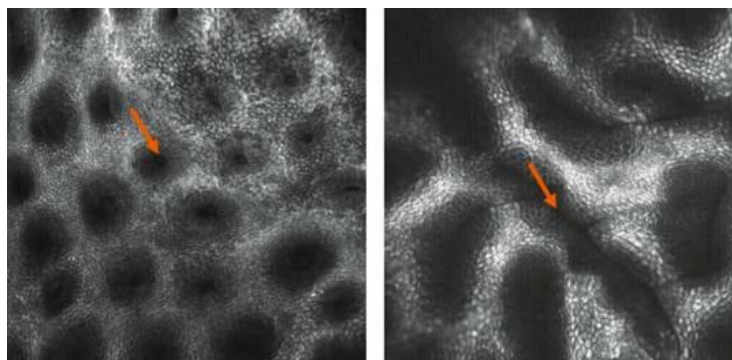


Figura 4

1.1.3 Mucosa intestinale del colon

Se siamo in presenza di Esofago di Barrett, ci troveremo a dover analizzare un tessuto di tipo metaplastico intestinale che si rifà essenzialmente alla struttura tipica della mucosa del colon (ossia una parte dell'intestino crasso). Vediamo quindi la sua struttura tipica.

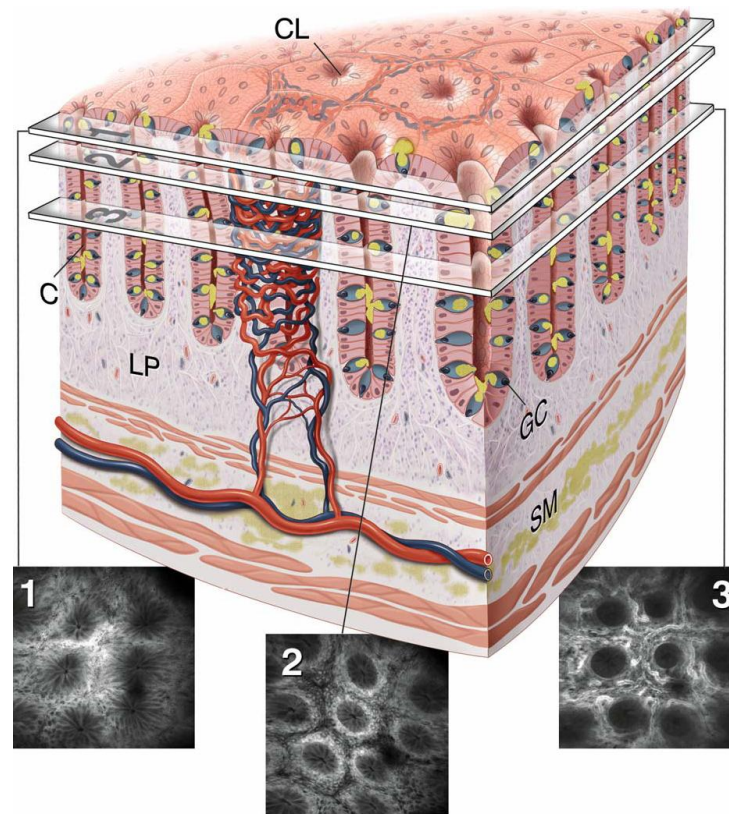


Figura 5. C Cripta, LP lamina propria, GC cellule a forma di calice, SM sottomucosa.

La mucosa superficiale del colon è composta da migliaia di cripte (C) o ghiandole orientate come strutture tubolari che si estendono verso la muscularis mucosae (tra LP e SM). Queste ghiandole sono supportate da un tessuto connettivo chiamato lamina propria (LP) che contiene vascolarizzazioni e innervazioni.

Le cellule epiteliali che rivestono la superficie interna del colon sono totalmente differenti da quelle squamose tipiche della mucosa esofagea normale e possono essere di due tipi: **cellule colonnari atte all'assorbimento di liquidi** e cellule caliciformi (GC) atte alla secrezione di muco necessario a lubrificare la parete interna del colon. Il contenuto delle cellule caliciformi si riversa nelle stesse cripte a forma d'imbuto e passa quindi effettivamente nel lume interno.

Ai fini diagnostici è molto importante osservare come questo tipo di mucosa sia caratterizzato da una visibile vascolarizzazione superficiale che proviene dalla lamina propria e forma una sorta di rete attorno ad ogni cripta (vedi Figura 6.)



Figura 6

1.1.4 Esofago di Barrett

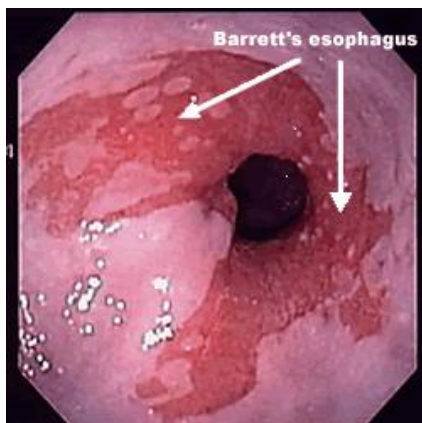
Partendo dalle semplici nozioni viste in precedenza, è possibile dare una sintesi della patologia che interessa questo lavoro; si sottolinea che le informazioni che saranno di seguito fornite provengono direttamente dal sito dello IOV³.

L'esofago di Barrett, dal nome del chirurgo inglese che descrisse questa patologia negli anni 50 dello scorso secolo, è una condizione pre-cancerosa che riguarda il rivestimento dell'esofago. I casi di **cancro esofageo** nel mondo occidentale sono quintuplicati negli ultimi 30 anni e molti di essi pare siano derivati dalla patologia di Barrett che inciderebbe per il 35% nel rischio di cancro esofageo. In **Italia** si contano circa 2.000 casi ogni anno. Numerosi pazienti soffrono di una condizione chiamata Malattia da reflusso gastroesofageo (MRGE). Questa è una malattia in cui l'acido normalmente presente nello stomaco può causare lesioni alla mucosa esofagea, producendo sintomi come il "bruciore" di stomaco, rigurgito, e dolore toracico. In alcuni pazienti con MRGE, le cellule normali dell'esofago sono danneggiate.

Nel corso del tempo, questi danni possono causare un'inflammatione cronica e cambiamenti genetici che inducono le cellule a divenire alterate. Il tessuto di rivestimento superficiale dell'esofago assume un aspetto diverso e microscopicamente non è più il normale tessuto di rivestimento dell'esofago, ma diventa un tessuto differente: se è simile alla mucosa gastrica e si parla di METAPLASIA GASTRICA, se è più simile alla mucosa intestinale (più precisamente del colon) si parla invece di METAPLASIA INTESTINALE o Esofago di Barrett appunto.

Per la corretta diagnosi di Esofago di Barrett è necessario che il paziente sia sottoposto ad un'esofagogastroduodenoscopia con biopsie.

Il tessuto dell'esofago di Barrett appare nelle immagini endoscopiche di un colore diverso, che indica dove eseguire la biopsia del tessuto per la valutazione anatomico-patologica (microscopica).



L'acido presente nello stomaco refluisce verso l'esofago provocando dei danni al rivestimento dell'esofago. L'area rosa intenso rappresenta tessuto di Barrett.

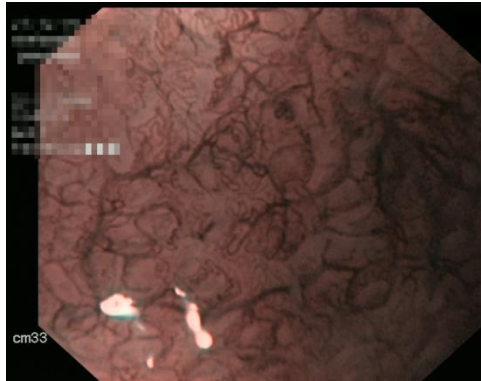
Figura 7

Il riscontro di cellule intestinali nell'esofago (metaplasia intestinale) conferma la diagnosi di esofago di Barrett.

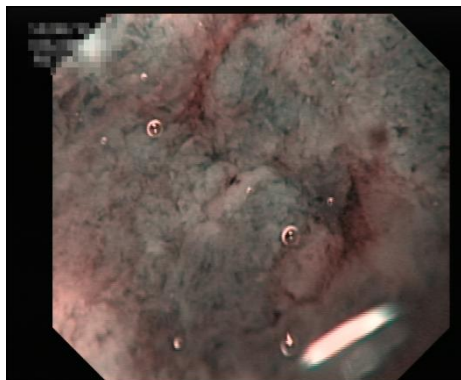
³ www.ioveneto.it

Esistono diversi tipi o "gradi" di esofago di Barrett che possono essere diagnosticati dalle biopsie. Questi "gradi" comprendono: la metaplasia intestinale (IM) senza displasia, la metaplasia intestinale con displasia di basso grado, e la metaplasia intestinale con displasia di alto grado.

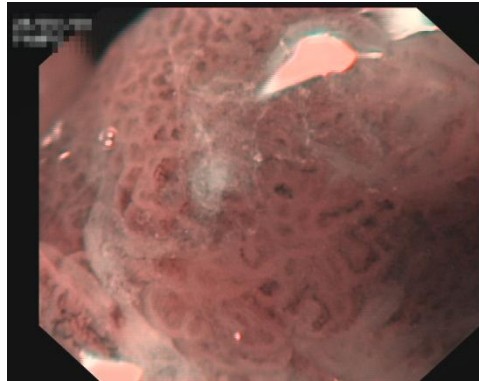
1. Metaplasia intestinale (IM) in cui le cellule dei tessuti hanno cominciato a modificarsi geneticamente e il tessuto è più simile a quello del rivestimento intestinale rosso piuttosto che al normale esofago. In questa fase, una persona ha già l'esofago di Barrett, ma non ha sviluppato displasia.



2. Displasia di basso grado (LGD) in cui meno del 50% delle cellule sono anomale ed hanno iniziato a cambiare in termini di dimensioni, forma o organizzazione. E' inoltre possibile riscontrare un aumento della velocità di crescita. Le cellule sono contenute all'interno del tessuto di rivestimento superficiale dell'esofago e non si sono diffuse in altre zone.
3. Displasia di alto grado (HGD). Come per la LGD, le cellule anomale risiedono all'interno del rivestimento dell'esofago. Le anomalie sono presenti in oltre il 50% di queste cellule e si riscontra un aumento ancora maggiore del tasso di crescita anormale e dell'organizzazione cellulare.



4. Adenocarcinoma (cancro esofageo) – Questa condizione avviene quando le cellule anormali hanno un tasso di crescita rapida e incontrollata. Le cellule invadono anche gli strati più profondi del vostro esofago e possono diffondersi altrove.



Sorveglianza

Le diverse società mediche mondiali raccomandano che i pazienti con esofago di Barrett siano sottoposti a un'esofagogastroduodenoscopia (EGDS) con biopsie ad intervalli regolari per il resto della loro vita. Poiché l'esofago di Barrett è una malattia **senza sintomi**, è molto importante seguirne con una rigorosa continuità l'eventuale evoluzione verso quadri patologici più severi che, pur essendo rari, esistono.

Possibilità terapeutiche

Oltre alle strategie di sorveglianza endoscopica per l'esofago di Barrett, vi sono altre opzioni terapeutiche in grado di eliminare completamente il "tessuto" di Barrett. Alcuni pazienti con displasia di alto grado possono essere sottoposti ad una procedura endoscopica operativa per rimuovere il tessuto malato mediante asportazione del tessuto con "mucosectomia".

In casi particolarmente severi di displasia di alto grado vi è l'indicazione a sottoporsi ad un'esofagectomia (asportazione dell'esofago o parte di esso) per evitare la progressione del tumore esofageo.

Un'ulteriore opzione è offerta dalla coagulazione o "ablazione" che consiste nel riscaldare il tessuto da rimuovere fino a quando non è più vitale. La tecnologia di ablazione a radiofrequenza è un tipo molto specifico di ablazione, in cui viene erogata l'energia termica in modo molto preciso e controllato.

DIAGNOSI CONFOCALE	ARCHITETTURA VASCOLARE	ARCHITETTURA CRIPTE
<i>EPITELIO DI TIPO GASTRICO</i>	Capillari con una forma regolare sono visibili solamente nelle parti più profonde dello strato di mucosa.	Epitelio colonnare regolare con aperture ghiandolari circolari e un'apparenza tipicamente ciottolante
<i>EPITELIO DI BARRETT</i>	Capillari sub epiteliali con una forma regolare sotto l'epitelio colonnare sono visibili nella parte superiore e nella parte più profonda dello strato mucoso.	Epitelio colonnare con presenza intermittente di mucina (una glicoproteina) nelle cellule caliciformi nella parte superiore dello strato mucoso. Nelle parti più profonde sono presenti cellule epiteliali di Barrett, villose, scure, di forma regolare cilindrica
<i>NEOPLASIA</i>	Capillari irregolari visibili nella parte superiore e nelle parti più profonde dello strato mucoso. Perdita di vasi conduce ad un'eterogenea e brillante intensità del segnale all'interno della lamina propria.	Celle in nero con bordi apicale e distale irregolari e forme, con un forte contrasto scuro nei confronti del tessuto circostante

1.2 Strumentazione endoscopica

L'endomicroscopia confocale laser è una nuova tecnica endoscopica che integra la video-endoscopia convenzionale con la microscopia confocale laser. Tale integrazione permette l'identificazione in - vivo delle microstrutture cellulari e subcellulari e consente, quindi, una diagnosi istologica di tessuto in - vivo e in tempo reale. Queste caratteristiche rendono la microscopia confocale endoscopica potenzialmente utile nella diagnosi precoce di lesioni tumorali o displastiche, così come nell'ottimizzazione delle biopsie e del trattamento endoscopico resettivo mirato.

1.2.1 Principi di microscopia confocale

Il sezionamento ottico di un sistema biologico consiste nella raccolta di una serie d'immagini di piani paralleli, spostando il fuoco dell'obiettivo lungo un asse che generalmente coincide con l'asse di propagazione della luce.

Per ottenere una perfetta rappresentazione di un singolo piano del campione, si dovrebbe idealmente raccogliere soltanto la luce proveniente da quel particolare piano; poiché tuttavia, anche i piani sovrastanti e sottostanti emettono luce, vi è una perdita di nitidezza dell'immagine.

La chiave del successo della tecnica confocale consiste nella rimozione delle interferenze provenienti dai piani adiacenti a quello ove si è focalizzati, mediante l'uso del cosiddetto pinhole. La microscopia confocale usa, per eccitare le molecole, una sorgente luminosa monocromatica molto intensa, il laser della lunghezza d'onda di 488 nm. La luce emessa dai fluorocromi eccitati dal laser viene catturata dalle lenti dell'obiettivo, attraversa lo specchio dicroico e raggiunge il fotomoltiplicatore, che trasforma l'intensità luminosa in un segnale elettrico di intensità proporzionale. Tra lo specchio dicroico ed il fotomoltiplicatore, il fascio luminoso attraversa un diaframma, o pinhole, che impedisce alla luce proveniente dalle zone fuori fuoco di raggiungere il fotomoltiplicatore. In questo modo solo il segnale luminoso relativo al piano di fuoco viene registrato e utilizzato nella formazione dell'immagine finale. Il risultato è un'immagine poco disturbata dalla diffusione della luce delle zone non a fuoco.

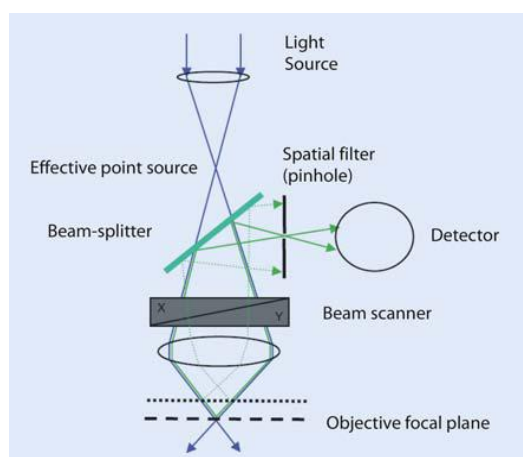


Figura 8

Di fatto, oltre al pinhole per la luce emessa, viene utilizzato anche un pinhole per la luce di eccitamento, in modo da illuminare solo una porzione microscopica del campione, aumentando il contrasto. Per ottenere la rappresentazione non di una porzione microscopica del campione ma di un intero piano, si muove il fascio di luce lungo il campione di punto in punto, in modo che tutto il piano situato alla profondità voluta venga illuminata dal fascio di luce secondo una precisa sequenza. Questo processo viene detto scansione. Per aumentare la velocità di acquisizione delle immagini, alcuni microscopi muovono il fascio di luce mediante specchi mobili che dirigono la luce incidente verso il campione in una scansione regolare. Questi specchi rendono possibile la ricostruzione dell'immagine in meno di un secondo. In genere, sono acquisite e mostrate sul monitor 0,8 o 1,6 immagini endoscopiche per secondo. La necessità di minimizzare le dimensioni della sonda endoscopica ha condotto all'utilizzo di una singola fibra ottica che si faccia carico della trasmissione sia della luce laser proiettata sul campione bersaglio, sia del segnale luminoso emesso dal campione bersaglio illuminato, il quale tramite la stessa fibra ottica raggiunge il fotomoltiplicatore, dove viene convertito nel segnale elettrico corrispondente. Il segnale elettrico in uscita dal fotomoltiplicatore viene quindi digitalizzato ed inviato ad un computer che registra i valori di intensità misurati per ogni punto. Questi valori vengono utilizzati per ricostruire l'immagine: ogni punto corrisponde ad un pixel dello schermo, e l'intensità luminosa del punto verrà rappresentata da una corrispondente tonalità di grigio. L'accostamento di tutti i singoli pixel corrispondenti a punti scanditi dal fascio laser nel campione darà così l'immagine finale.

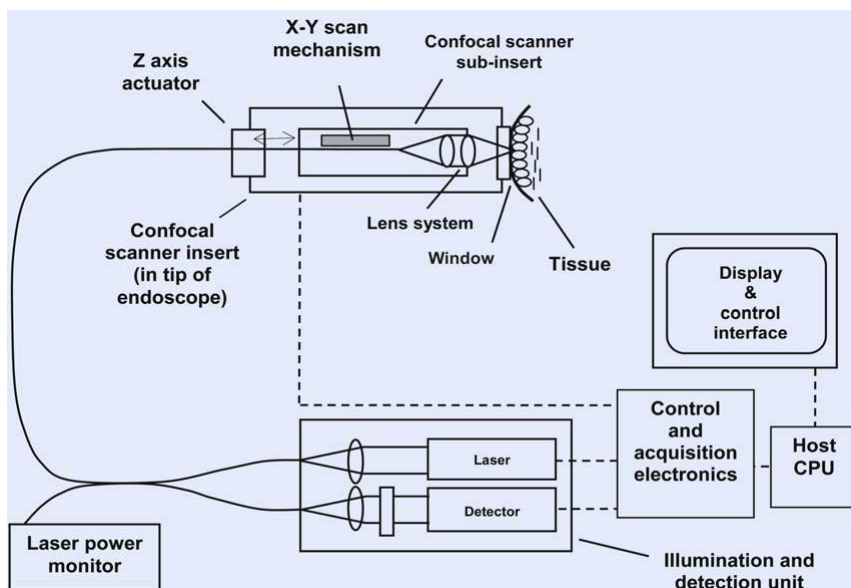


Figura 9

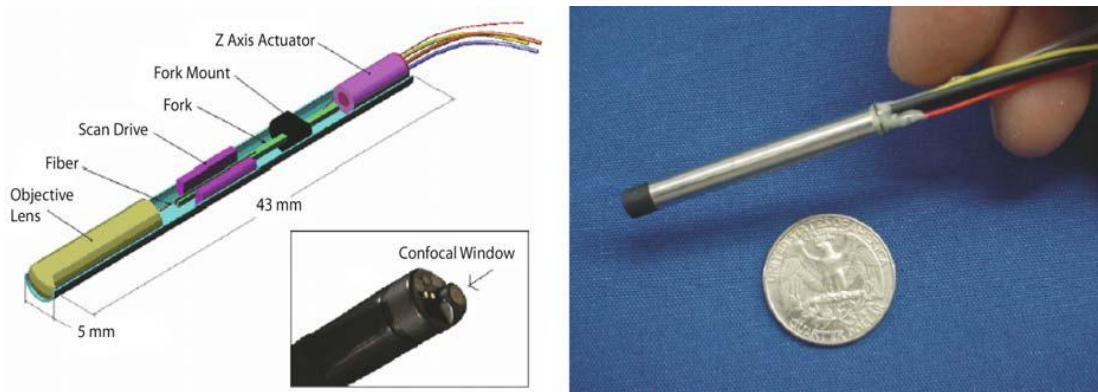


Figura 10

Spostando lungo l'asse verticale il campione dopo ogni scansione, è possibile eseguire una serie di scansioni successive corrispondenti ai piani focali via via più profondi all'interno del campione.

Queste scansioni prendono il nome di sezioni ottiche e la loro sovrapposizione ordinata, eseguita via software, consente di ricostruire un'immagine complessiva dell'intero volume scandito, in cui tutti i piani sono contemporaneamente a fuoco. I programmi per l'elaborazione di immagini non registrano, quindi, soltanto la luminosità di ciascun punto, ma anche la sua localizzazione nel campione, cioè la sua posizione in un piano e la sua profondità: i punti definiti dalle tre coordinate, (x, y, z) , detti voxel, costituiscono l'equivalente tridimensionale dei pixel di un'immagine bidimensionale.

L'area di mucosa esplorata è di 500 per 500 μm con una risoluzione di 0,7 μm . La scansione può essere spinta dalla superficie della mucosa sino a 250 μm in profondità. La lunghezza del terminale rigido dell'endoscopio è di 48 mm. Il canale operativo è di 2,8 mm. Manopole e pulsanti di comando sul manipolo sono gli stessi di un endoscopio convenzionale, con l'aggiunta di due pulsanti addizionali per gestire i singoli piani di scansione confocale dello spessore di 7 μm .

La procedura endoscopica confocale si svolge in modo analogo a quella tradizionale. Lo studio preliminare della mucosa con coloranti non determina alcuna interferenza con l'autofluorescenza del viscere. Durante l'acquisizione delle immagini endomicroscopiche il terminale dello strumento deve essere appoggiato delicatamente sulla mucosa/lesione da indagare. L'esplorazione microendoscopica necessita di un'approfondita conoscenza dell'architettura della mucosa dei vari distretti digestivi.

In base al disegno dell'architettura ghiandolare e vascolare sono stati definiti pattern endomicroscopici di normalità, aspetto rigenerativo e neoplastico (Fig. 3).



Figura 11

Le immagini endomicroscopiche sono generate mediante l'uso di un agente di contrasto fluorescente. La fluoresceina sodica, già largamente utilizzata in oftalmologia ed angiografia, è il mezzo di contrasto più adoperato per il basso costo e l'assenza di potenzialità mutagenica. Dopo la somministrazione endovenosa di 5-10 mL di fluoresceina sodica al 10%, le cellule, il sistema vascolare e il tessuto connettivo possono essere ben differenziati. Per le sue proprietà farmacocinetiche, i nuclei cellulari non sono chiaramente visibili. L'acriflavina idrocloride può essere applicata per via topica per colorare il nucleo e il citoplasma cellulare. Questa sostanza è assorbita in pochi secondi e la sua azione è circoscritta agli strati superficiali della mucosa. Il suo uso è da limitare per una possibile attività mutagenica. Fluoresceina e acriflavina possono essere utilizzate simultaneamente. Lo studio endomicroscopico consente di esplorare una superficie più vasta della mucosa rispetto alle biopsie random e di ottimizzare e mirare il campionamento bioptico e il trattamento endoscopico resettivo.

1.3 Obiettivo del progetto

Il lavoro che verrà illustrato qui di seguito è frutto di una lunga collaborazione e riporta una parte dei risultati conseguiti .

Il codice implementato ha come finalità quella di tentare una classificazione tra due tipi d'immagini:

- immagini con tessuto di “esofago di Barrett” e quindi di metaplasia intestinale
- immagini mataplasiche gastriche sempre provenienti dalla zona distale dell'esofago

Oltre a questo si è reso necessario un codice che riconoscesse la presenza di eventuale tessuto sano adiacente a tessuto malato nelle immagini a nostra disposizione.

2 Materiali



Il materiale utilizzato e' stato gentilmente fornito dallo IOV - Istituto Oncologico Veneto - I.R.C.C.S. (Istituti di Ricovero e Cura a Carattere Scientifico).

Come si può leggere nel sito ufficiale⁴ esso "...è il primo Istituto del Veneto specificatamente destinato alla prevenzione, diagnosi e cura dei tumori ed alla ricerca sul cancro: la sua missione è infatti quella di fornire l'assistenza più avanzata ai malati neoplastici e svolgere nello stesso tempo ricerca biomedica, essenziale per il progresso delle conoscenze e il trasferimento ai pazienti delle cure più innovative".

Il centro esegue circa 2700 endoscopie l'anno facendo uso delle più moderne tecnologie disponibili utilizzate per la diagnosi ed il follow-up di neoplasie e lesioni e condizioni precancerose.

In particolare di nostro interesse e' la diagnosi dell'Esofago **di Barrett** eseguita con la tecnologia NBI- Zoom, AFI ed endoscopia confocale.

I file contenenti le immagini endoscopiche a noi pervenuti a partire dal mese di Novembre 2010 comprendono sostanzialmente immagini clinicamente suddivise in metaplasiche e displasiche. Ai fini del progetto sono state utilizzate immagini metaplasiche di tipo gastrico e intestinale oltre a due immagini di tessuto esofageo sano.

Il primo passo e' stato quello di selezionare le immagini che non presentassero disturbi tali da compromettere il riconoscimento della struttura del tessuto. Non sono perciò state comprese nello studio immagini che presentavano grossi difetti di focalizzazione o "bubble noise"⁵ in quantità considerevole.

Le immagini sono di dimensioni differenti pur assumendo valori prossimi a 700x550 e contengono un bordo nero che fa da contorno all'effettiva immagine di tessuto, ossia l'informazione utile . Sono inoltre presenti informazioni di acquisizione dello strumento che si presentano in sovraimpressione ,alcune volte anche sovrapposte all'informazione utile .

Molto spesso poi le immagini non sono totalmente di tipo gastrico o di tipo intestinale ma, possono presentare anche ampie porzioni di tessuto sano . Ecco perché il lavoro che verrà successivamente illustrato consta di un algoritmo per la distinzione sano/malato e un algoritmo di riconoscimento del tipo di metaplasia. L'implementazione di tutte le procedure nonché la maggior parte dei grafici utilizzati per illustrare i risultati sono stati ottenuti con l'utilizzo di un potente e versatile software : MATLAB.

⁴ www.ioveneto.it

⁵ Si intendono le bolle gassose che compaiono alcune volte nelle immagini

3 Metodi

3.1 Procedura di Cropping

Il primo passo affrontato è stato quello di costruire due insiemi contenenti matrici RGB rispettivamente di metaplasia gastrica e metaplasia intestinale dai quali poi sono stati ricavati gli insiemi di training e di test. I due insiemi dovevano contenere matrici delle stesse dimensioni e possibilmente quadrate dato che generalmente queste sono più facili da gestire .

Tali insiemi sono stati ottenuti con un processo di *cropping* che prevede innanzitutto il prelievo della sola parte realmente contenente l'informazione dall'immagine di ingresso . Infatti, come si può vedere in Fig.4-a le immagini a nostra disposizione contengono un ampio bordo nero dovuto proprio alla modalità con cui lo strumento acquisisce l'informazione .

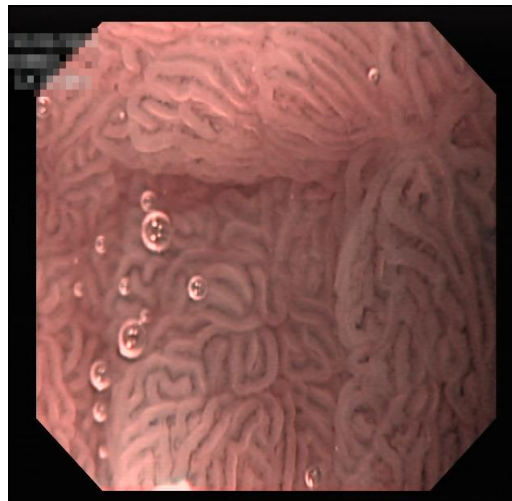
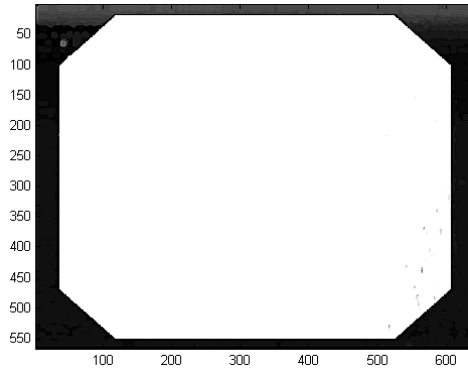
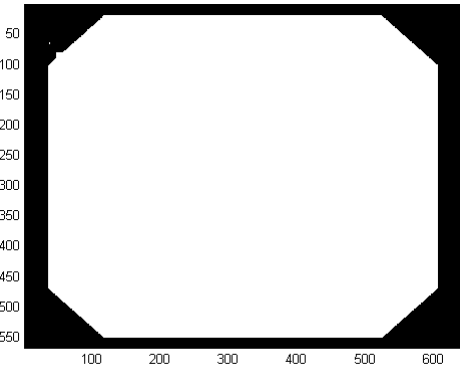


Fig.4-a

Una volta eliminato il bordo ,attraverso una brutale ma efficace applicazione di un *filtro sharpen* seguita da una potente chiusura morfologica , si è proseguito con un ritaglio finalizzato appunto ad ottenere una matrice RGB con delle precise dimensioni . Dopo una serie di prove è stata fissata la dimensione di 400x400x3 che garantisce per tutte le immagini in nostro possesso un ritaglio ottimale ossia la minor perdita di informazione per tutte le immagini originali che ricordiamo essere di dimensioni diverse .Le immagini qui di seguito (Figura 12) chiariscono il processo per ottenere le effettive matrici di lavoro , aggiungiamo a scopo informativo che i *filtri sharpen* vengono utilizzati per aumentare i bordi degli oggetti e per questo possono essere utilizzati come rivelatori di bordo, come nel nostro caso.

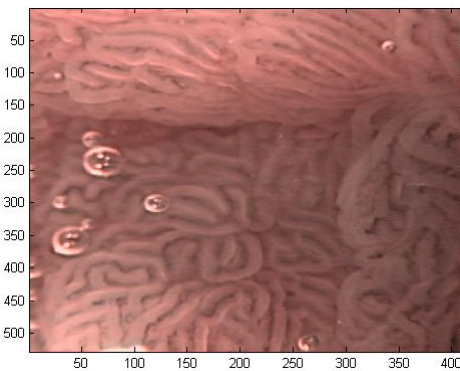


Uso del filtro *sharpen*



Uso di sogliatura e operazione di chiusura . Ora i bordi neri son ben definiti e si può procedere ad un primo ritaglio

IMMAGINE DEFINITIVA



400x400x3

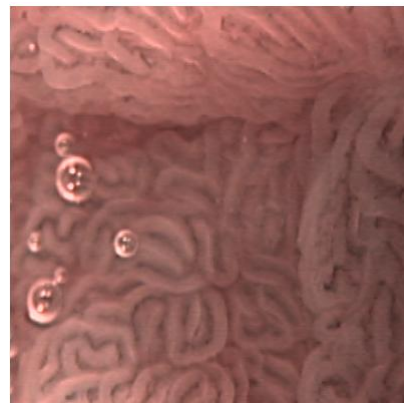


Figura 12

3.2 Riconoscimento parte sana

Per identificare la parte è stata utilizzata una procedura che prevede in primis una divisione dell'immagine di partenza ,che ricordiamo essere una matrice quadrata 400x400, in 16 sottomatrici 100x100 . Per tatare il sistema di riconoscimento della parte sana sono stati quindi creati degli insiemi di training contenti delle immagini 100x100 sia di solo tessuto sano sia di solo tessuto metaplasico (senza badare alla distinzione tra gastrico e intestinale).

3.2.1 Edge detection

La caratteristica che più diversifica il tessuto sano da quello metaplastico è l'omogeneità del primo e la complessità strutturale del secondo, il quale presenta un'architettura tissutale piuttosto complessa e diversificata. Ne consegue che il tessuto metaplastico in una immagine in scala di grigi presenta molti più bordi di quello sano.

L'algoritmo di classificazione opera sulla base di questa distinzione: mediante le funzioni di edge detencion, messe a disposizione da MATLAB, è possibile ottenere un immagine binaria che riporta i bordi presenti nell'immagine grayscale.

A tal fine si utilizza il metodo di Canny, il quale trova i bordi cercando i massimi locali del gradiente dell'immagine in scala di grigi. Il gradiente è calcolato utilizzando la derivata di un filtro Gaussiano. Il metodo utilizza due soglie, una per i bordi più evidenti, l'altra per individuare i bordi più deboli, i quali vengono inclusi nell'immagine di output solo se sono connessi a quelli più evidenti. Prima di applicare Canny però è stata fatta un omogeneizzazione della luminosità dell'immagine 100 x 100 in questo modo : a ciascun pixel della matrice 100x100 viene sottratta la media locale calcolata nel suo intorno, usando cioè una maschera di media di dimensione 5x5 . Il valore di tale media sarà basso nelle zone scure ed elevato nelle zone chiare. Il risultato della sottrazione è che tutte le matrici 100x100, sia quelle appartenenti a zone scure che a zone chiare (dell'intera immagine), presentano ora un istogramma distribuito su un range di valori confrontabile . (vedi Figura 13)

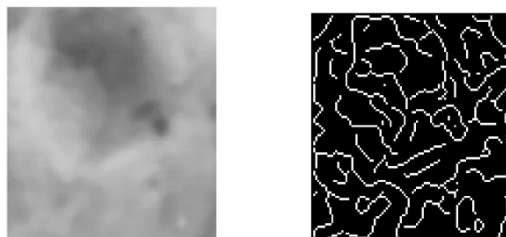


Figura 13 A sinistra una matrice 100x100 pixel di tessuto sano, prima dell'operazione di omogeneizzazione della luminosità. A destra la corrispondente immagine binaria dei bordi trovati (con soglia 0.15).

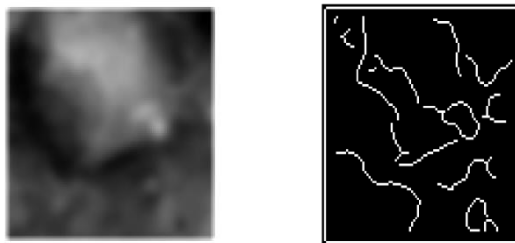


Figura 14 A sinistra matrice 100x100 pixel di tessuto sano, dopo l'operazione di omogeneizzazione della luminosità. A destra la corrispondente immagine binaria dei bordi trovati (con soglia 0.15). Si nota che la quantità di bordi trovati nell'immagine sana, in seguito all'omogeneizzazione della luminosità, è notevolmente ridotta.

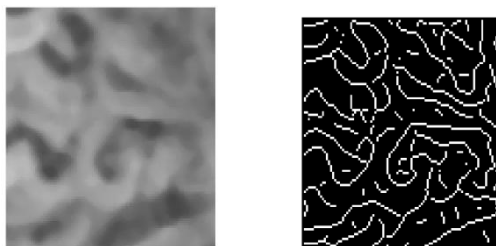


Figura 15 A sinistra matrice 100x100 pixel di tessuto metaplastico prima dell'operazione di omogeneizzazione della luminosità. A destra la corrispondente immagine binaria dei bordi trovati (con soglia 0.15).

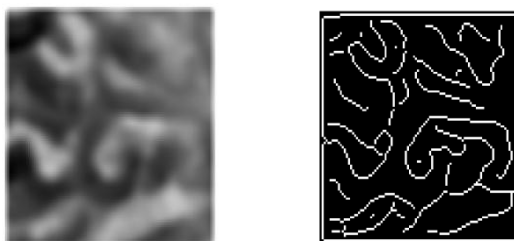


Figura 16 A sinistra matrice 100x100 pixel di tessuto metaplastico, dopo l'operazione di omogeneizzazione della luminosità. A destra la corrispondente immagine binaria dei bordi trovati (con soglia 0.15). Si nota che la quantità di bordi trovati nell'immagine metaplastica, in seguito all'omogeneizzazione della luminosità, è ridotta di poco.

I valori delle soglie sono fissi per tutte le immagini. La soglia alta (*thresh*) è stata definita tramite un'indagine statistica eseguita su un insieme di addestramento, costruito con matrici 100x100 pixel; la soglia bassa è $0.4 * thresh$. Il valore di *thresh* è relativo al più grande dei valori della dimensione del gradiente dell'immagine.

Si spiega di seguito brevemente il criterio di selezione del valore di soglia fissa.

A partire dal valore di soglia (*auto-thresh*) assegnato automaticamente da Matlab nell'applicazione del metodo di Canny, è stato effettuato lo studio dell'immagine binaria dei bordi, al variare della soglia in un intervallo di valori dell'intorno di *auto-thresh* (entro limiti dell'intervallo che permettano di visualizzare risultati significativi). Per distinguere il tessuto sano da quello metaplastico, si è stabilito infine il valore della soglia che permettesse di massimizzare la distinzione tra i due raggruppamenti sulla base del parametro di classificazione finale(illustrato nel paragrafo successivo).

3.2.2 Quantificazione dei bordi

Il parametro che stabilisce se una immagine presenta tessuto sano o metaplastico è il numero di pixel di valore 1 presenti nell'immagine binaria restituita dall'algoritmo di *edge detection*, che quantifica il numero bordi presenti e quindi stabilisce l'anatomia tessutale della parete dell'esofago.

Anche in questo caso dall'analisi sulle immagini dell'insieme di training, è possibile stabilire una soglia fissa che sulla base del numero di pixel di valore 1, classifica ogni matrice 100x100 pixel come sana o metaplastica.



Figura 17 Le due immagini binarie evidenziano la differenza tra la quantità di bordi trovata in un immagine con tessuto sano (a sinistra) e in una con tessuto metaplastico (a destra).

3.3 Classificazione malate : metodo “corr”

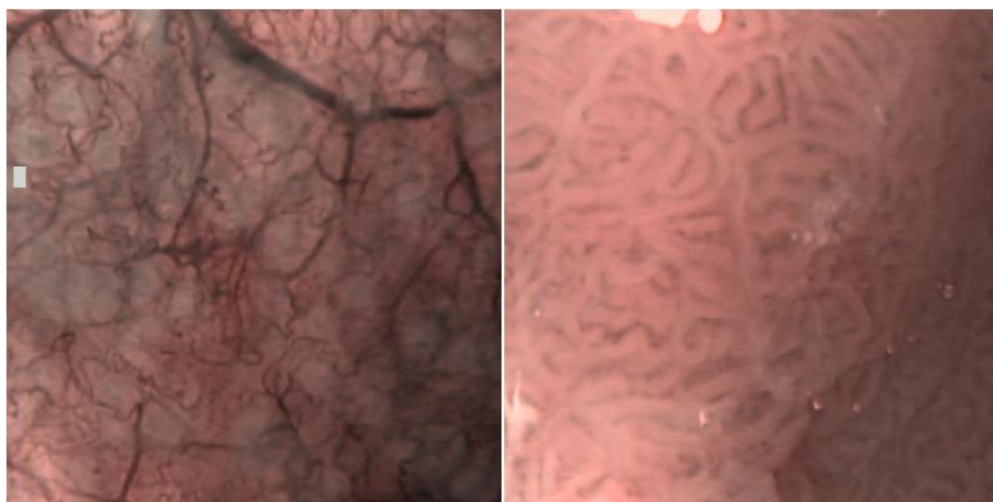


Figura 18

Figura 19

3.3.1 Introduzione metodo

Se chiedessimo ad una persona di guardare un immagine di metaplasia intestinale (Figura 18) e un immagine di metaplasia gastrica (Figura 19) e invitassimo poi la stessa a dirci quale differenza ha notato di primo acchito , nella maggior parte dei casi la risposta sarebbe legata alla presenza consistente di vasi sanguigni nella prima e alla loro quasi totale assenza nella seconda. Questa banale osservazione è in realtà il cuore del metodo di classificazione che andremo a spiegare qui di seguito .

Le strutture vascolari tipiche di immagini di metaplasia intestinale hanno valori bassi nella scala dell' intensità , rispetto ovviamente alla totalità dell'immagine.

Anche nelle immagini di metaplasia gastrica sono presenti regioni scure e quindi è immediato pensare che la sola presenza di aree scure non possa fornire un principio di classificazione .E' necessario perciò discernere tra regioni vascolari scure e regioni scure ma non vascolari .

3.3.2 Sintesi del codice utilizzato

L'algoritmo implementato, diversamente da quello per il riconoscimento della parte sana, usa le immagini nella loro totalità cioè lavora con matrici 400x400 e consta principalmente di due blocchi:

1. un'elaborazione di principio, che ha il compito di far emergere le regioni scure
2. una seconda parte che seleziona e mantiene solo le parti scure che sono sedi vascolari.

Il risultato di classificazione si raggiunge passando attraverso il metodo di MATLAB *corr2()* che fornisce un indice di correlazione tra l'immagine in uscita dal primo blocco di elaborazione e quella successiva al blocco di selezione. In pratica se la selezione ha mantenuto la maggior parte delle strutture ci aspetteremo un indice di correlazione alto, in caso contrario l'indice di correlazione sarà basso. Così stabilendo una soglia di confronto per tale indice sarà possibile dire di che tipo di metaplasia si tratta l'immagine elaborata.

3.3.3 Descrizione del codice utilizzato

3.3.3.1 Procedura "Black&White" o primo blocco

Una volta ottenute le effettive immagini su cui lavorare (vedi procedura di cropping) il passo successivo è stato quello di trovare un algoritmo adeguato per dividere chiaramente le regioni scure da quelle più chiare in modo da poterle quindi isolare e ottenere un'immagine binaria che contesse le strutture di nostro interesse ossia i vasi per le immagini metaplasia intestinale e le aree scure (ma non di vaso) per le immagini di metaplasia gastrica. Vedi rispettivamente Figura20 e Figura21

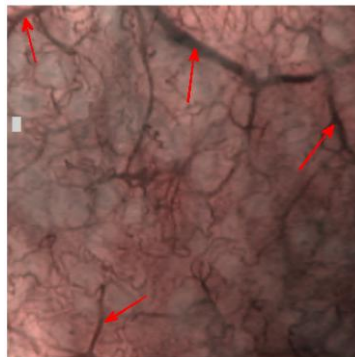


Figura 20

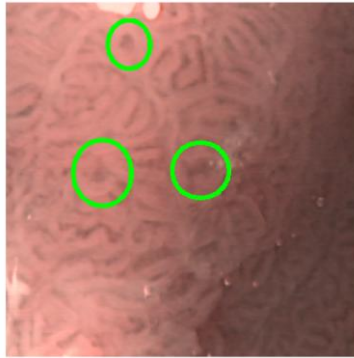


Figura 21

Contrariamente a quanto fatto nella fase iniziale di sperimentazione di tale metodo, si è deciso di sfruttare l'informazione proveniente da ogni singolo canale della matrice RGB invece di convertire quest'ultima in una matrice bidimensionale d'intensità. Il canale R, il canale G ed il canale B sono trattati con lo stesso algoritmo, che chiameremo "Black&White", in modo separato e in pratica come se fossero delle matrici d'intensità indipendenti.

I risultati di ogni singolo canale sono quindi sommati banalmente con un'operazione di tipo OR .

La procedura che tratta ogni canale usa delle funzioni già predisposte da MATLAB come l'utile *adapthisteq*() e la funzione *imfilter*() con cui si sono potuti applicare alcuni dei filtri necessari, come un *filtro media* e un *filtro sharpen*, in maniera semplice ed immediata. I filtri usati sono richiamabili in MATLAB rispettivamente con i comandi *fspecial*('average') e *fspecial*('unsharp').

La funzione *imfilter*() è ,banalmente, una procedura di convoluzione tra l'immagine di partenza ed il filtro e ,in maniera intuitiva, non fa altro che ritornare un'immagine in cui i singoli pixel sono la somma, pesata dai valori della matrice di filtro, dei pixel adiacenti . Se per esempio si usa *il filtro media* questo riporterà nel pixel in cui si centra la maschera di filtro il valore medio dei pixel adiacenti ; ovviamente le dimensioni della maschera di filtro decidono il risultato finale dell' operazione.

Per quanto riguarda *adapthisteq*() il concetto di base è che avviene una modifica a livello dell'istogramma dell' immagine ai fini di migliorarne il contrasto ; l'algoritmo prevede un *modus operandi* locale e non globale come fa ,per esempio ,*histeq*().

Le figure seguenti mostrano il canale R prima e dopo *adapthisteq*(), rispettivamente Figura 22 e Figura 23 .

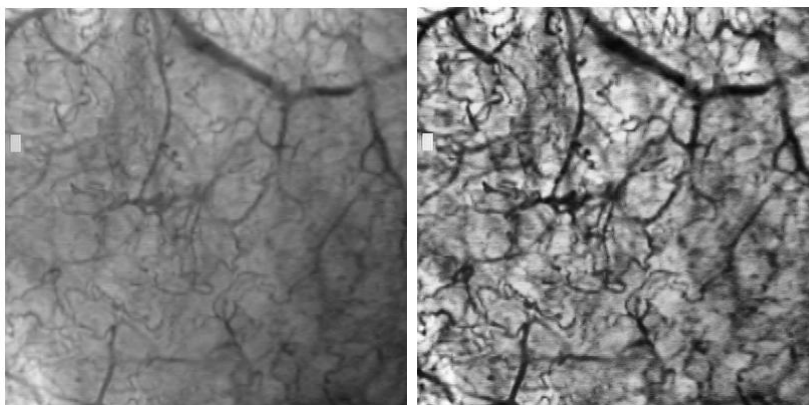


Figura 22

Figura 23

Oltre alle funzioni sopracitate è stato creato un algoritmo ad hoc molto semplice. Questo è usato all'inizio della procedura e ha lo scopo di stabilire una media fissa per tutte le immagini in ingresso. E' logico, infatti, pensare che immagini con una media alta saranno più luminose d'immagini con una media bassa. Il codice funziona come un semplicissimo sistema con controllo a retroazione per cui finché la media dell'immagine non è quella desiderata si continua a traslare l'istogramma dell'immagine verso valori chiari o scuri a seconda della necessità. Per traslare l'istogramma si usa la funzione di MATLAB *imadjust* () che, come riporta la guida, mappa un intervallo a piacere di valori di un'immagine, data in ingresso, in un nuovo intervallo di valori sempre a scelta. La legge che governa la mappatura può essere variata agendo su un parametro di metodo, il parametro *GAMMA*; a seconda del valore di *GAMMA* l'istogramma si sposta verso valori chiari o scuri , ottenendo così l'effetto voluto .

Alla fine della procedura "*Black&White*" sostanzialmente si esegue un taglio dell'immagine attraverso l'uso di una soglia fissa e si restituiscono in una matrice binaria le strutture di cui si parlava in precedenza ossia regioni scure (di vaso e non).

Più precisamente l'istogramma dell'immagine ,che precedentemente era stata tratta con le funzioni prima citate , viene "quantizzato" in soli cinque valori cioè si raggruppano più valori d'istogramma in un unico livello. Ora, per capirci, si dispone di un'immagine formata da cinque layers e il layer di livello minore contiene l'informazione sulle strutture che sono di nostro interesse . Non si deve far altro che prelevare tale layer ed eventualmente ripulirlo dall'effetto "sale e pepe " dovuto a pixel isolati . Tale pulizia può essere fatta con un filtro mediano o ,come nel nostro caso, dalla funzione *bwareopen* (). Dopo l'operazione di somma dei risultati di ogni canale, si è rivelata utile anche una semplice chiusura morfologica in modo da collegare aree facenti chiaramente parte dello stesso vaso ma che per inefficienze di procedura risultano disconnesse.

Il risultato complessivo dei tre canali è quello nella figura sotto riportata (Figura 24).

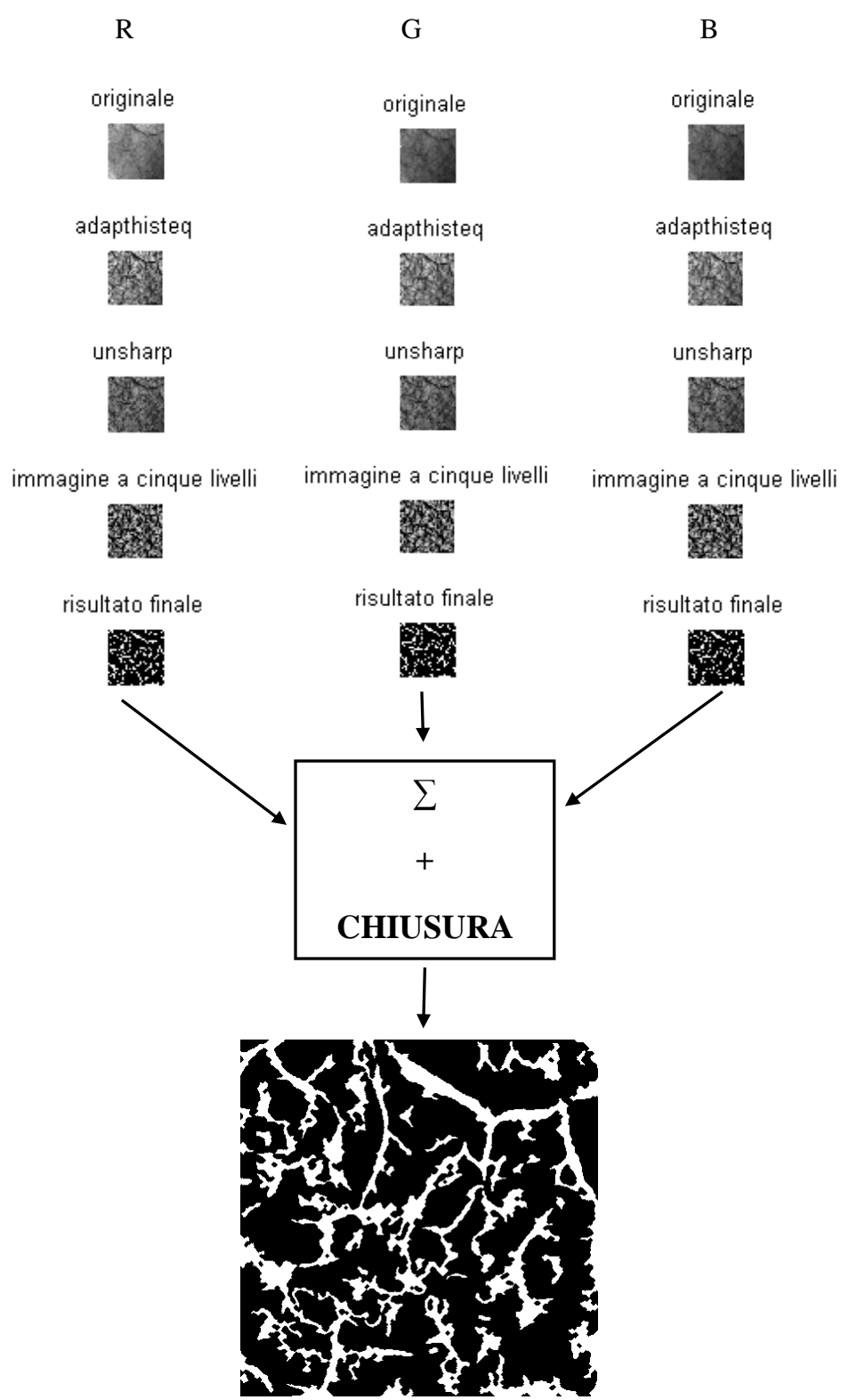


Figura 24

Lo stesso processo vale per un'immagine di tipo metaplasia gastrica e quindi avremo le seguenti immagini .

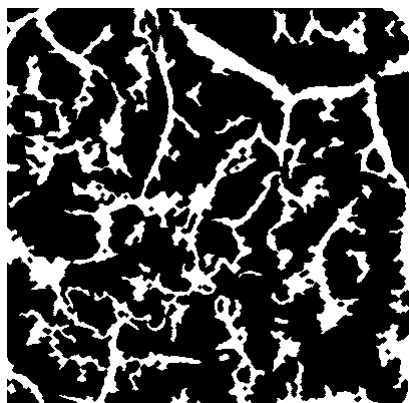


Figura 25



Figura 26

Ciò che si nota è la forte presenza di elementi disconnessi di varia forma e dimensione nell'immagine di metaplasia gastrica, Figura 26, e un insieme di strutture invece molto più connesse per quanto concerne l'immagine di metaplasia intestinale , Figura 25 . Inoltre in quest'ultima le strutture sono logicamente di tipo allungato, danno cioè l'idea di estendersi secondo una direzione preferenziale in accordo con la natura della loro origine .

3.3.3.2 Parte di selezione o secondo blocco

Questa è la seconda parte importante del metodo di classificazione cioè quella che seleziona e mantiene solo le strutture vascolari dell'immagine in uscita dalla procedura "Black&White".

In realtà non si è giunti a una soluzione che possa dirsi definitiva in quanto risulta difficile tradurre a livello procedurale una distinzione ottimale tra le strutture che caratterizzano le due immagini.

La soluzione adottata , e quella più immediata e intuitiva, è quella di distinguere vasi e non-vasi secondo il parametro dell'area. Se analizziamo, infatti, ogni singola struttura come una regione indipendente questa sarà dotata di una certa area facilmente calcolabile attraverso il metodo *regionprops*().

Questo metodo fornisce tra tante informazioni disponibili della regione che si vuole esaminare anche appunto l'area della stessa intesa come numero di pixel di cui è costituita. Dato che le regioni vascolari generalmente si estendono su gran parte dell'immagine è pensabile che la loro area sia maggiore rispetto a quella delle regioni tipiche della metaplasia gastrica. Così scegliendo un opportuno valore di soglia per l'area si può selezionare e mantenere solo quelle aree che rispettano la condizione imposta.

Una volta ottenuta quindi un'immagine in cui sono state asportate le regioni la cui area è minore del valore di soglia imposto, si procede con una semplice correlazione tra quest'immagine e quella originale in uscita dalla procedura "Black&White".

In realtà la funzione di correlazione che viene utilizzata non fornisce una matrice ma direttamente l'indice di correlazione (che assume valori tra [-1;1]) e nel nostro caso equivale ad esprimere quanto uguali sono le due matrici a confronto.

Questo indice⁶ si ottiene attraverso la formula

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}}$$

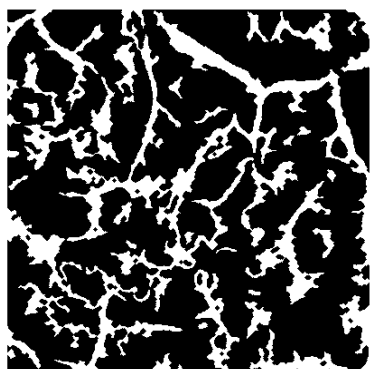
dove A e B sono matrici della stessa dimensione, \bar{A} e \bar{B} sono le medie delle stesse matrici.

L'indice di correlazione in realtà nel nostro caso può essere solamente positivo. Infatti affinché si presenti un indice negativo dovrei avere paradossalmente un'immagine con valori negativi ma ciò non è possibile in quanto le immagini in ingresso al metodo sono binarie e con due soli valori possibili 0 e 1.

Per trovare il valore minimo di area sono stati analizzati differenti grafici che plottando il valore dell'indice di correlazione per ogni immagine degli insiemi di training variando di volta in volta la soglia dell'area minima. È stata poi scelta quella che disponeva gli insiemi delle due tipologie di malattia in zone distinte del grafico, avendo cura che le immagini di metaplasia intestinale si presentassero in una zona appartenente a valori alti di correlazione e le immagini di metaplasia gastrica in una zona relativamente bassa.

⁶ Ricordiamo che in generale l'indice di correlazione tra due segnali è utile per capire se al variare di uno dei due anche l'altro varia. Se variano nello stesso senso si ha un indice di correlazione positivo al contrario se il senso è inverso si ha un indice di correlazione negativo; l'indice di correlazione è 0 qualora vi sia una mancanza totale di correlazione.

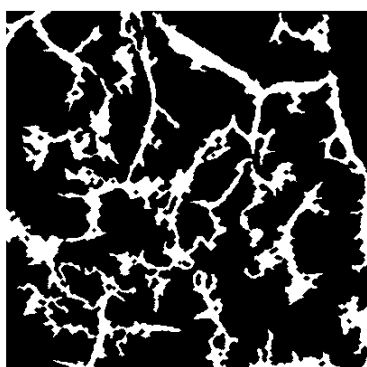
Dopo la procedura di selezione abbiamo quindi il seguente risultato⁷ rispettivamente per metaplasia intestinale e metaplasia gastrica.



MENO



=



MENO



=



⁷ è stata usato un valore di area minima pari a 1000 , vedere RISULTATI

3.4 Sintesi tra il risultato di classificazione e riconoscimento della parte sana

L'informazione proveniente dalla procedura che riconosce la parte sana, ai fini del codice di classificazione di cui fin qui abbiamo parlato, non è in alcun modo utilizzata. Tale informazione, infatti, è presa in considerazione solamente a posteriori cioè dopo aver già deciso di che tipo di metaplasia si tratta l'immagine data in ingresso. Ciò che si fa quindi è una semplice operazione di AND e quindi avendo a disposizione la maschera della parte sana (attraverso l'utilizzo del codice per il riconoscimento sano/malato) essa potrà essere utilizzata per specificare quali parti dell'immagine non sono malate del tipo di metaplasia indicato dall'indice di correlazione in uscita dal metodo di classificazione prima descritto.

4 Risultati

4.1 Riconoscimento parte sana e parte malata

Il codice che implementa l'algoritmo di classificazione è stato realizzato lavorando su un insieme d'immagini selezionate, scelte per definire due insiemi di training:

- l'insieme di training delle immagini con tessuto sano, contiene 13 matrici di dimensione 100x100 pixel
- l'insieme di training delle immagini con tessuto metaplastico, contiene 210 matrici di dimensione 100x100 pixel.

La classificazione delle immagini in due insiemi passa attraverso la definizione di due valori di soglia fissi:

- la *soglia di Canny*, che trova i bordi nell'immagine in scala di grigi cercando i massimi locali del gradiente dell'immagine.
- la *soglia di quantificazione dei bordi*, che in base al numero di pixel che assumono valore 1 nell'immagine binaria restituita dall'algoritmo di edge detection, stabilisce se si tratta d'immagine sana o metaplastica.

La prima soglia è stata selezionata nell'intervallo di valori $0.01 < th < 0.35$. Per questi valori è, infatti, possibile individuare un numero di bordi sufficiente a distinguere le due tipologie di tessuto. Si è quindi valutata la differenza tra il minimo numero di pixel di valore 1 ricavato dall'insieme di training delle immagini con tessuto sano, e il massimo numero di pixel di valore 1 ricavato dall'insieme di training delle immagini con tessuto metaplastico, al variare della soglia th . Il valore di th che rende massima tale differenza, e che perciò consente di distinguere al meglio i due insiemi di classificazione, è 0.15.

Il numero di pixel di valore 1, che definisce la soglia di quantificazione dei bordi per $th = 0.15$, è individuabile nell'intervallo 1028 – 1036.

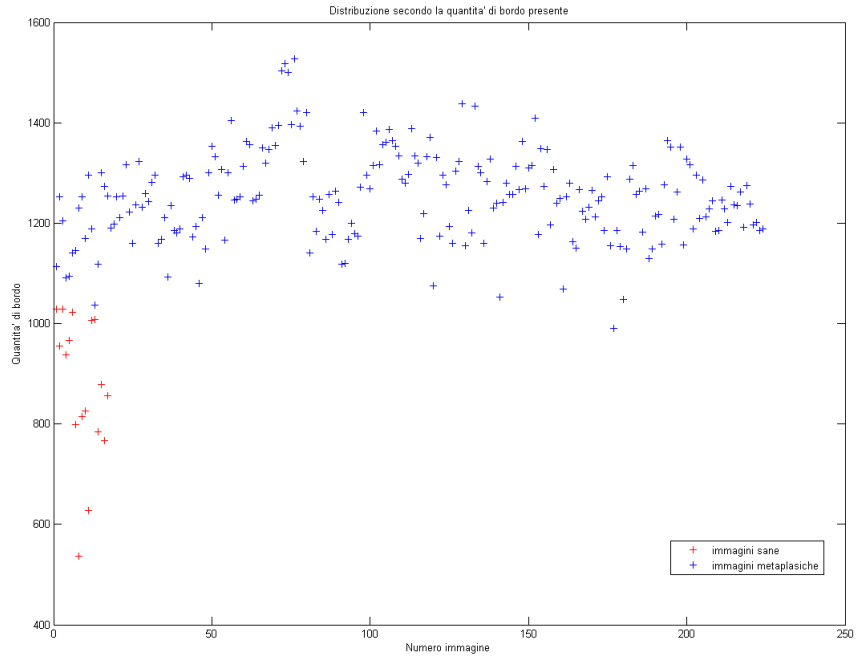


Figura 27 Immagini con tessuto sano e con tessuto metaplastico a confronto: il grafico mostra la distribuzione dei valori relativi al numero di pixel di valore 1 presenti nelle immagini che compongono i due insiemi di training. Il valore della “soglia di Canny” è 0.15.

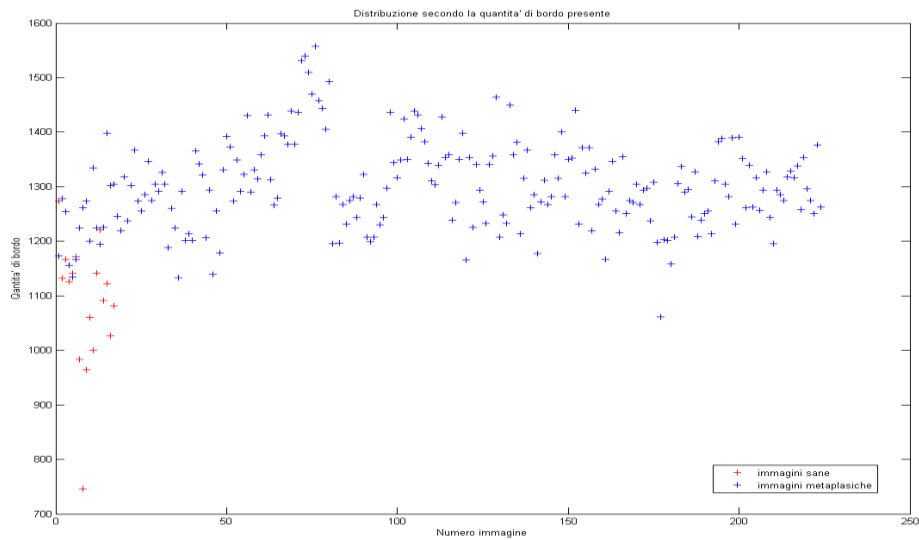


Figura 28 Grafico della distribuzione dei valori relativi al numero di pixel di valore 1 presenti nelle immagini che compongono i due insiemi di training. Il valore della “soglia di Canny” è in questo caso 0.24.

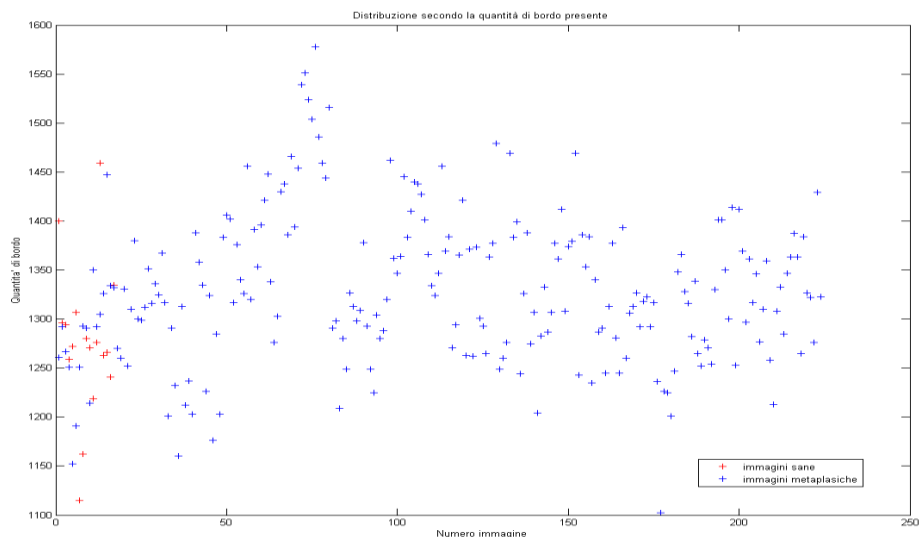


Figura 29 Grafico della distribuzione dei valori relativi al numero di pixel di valore 1 presenti nelle immagini che compongono i due insiemi di training. Il valore della “soglia di Canny” è in questo caso 0.07.

Esiste poi un terzo insieme di immagini che contiene 25 matrici 100x100 pixel di tessuto in parte sano e in parte metaplastico, i cui valori si distribuiscono nell’uno o nell’altro dei due insiemi di classificazione in base alla quantità di bordi trovati nell’immagine. Non è stato, infatti, possibile in questo caso distinguere nettamente il numero di pixel di bordo per le immagini di tale insieme da quello degli altri due insiemi, in quanto i valori risultano nella maggior parte dei casi sovrapposti.

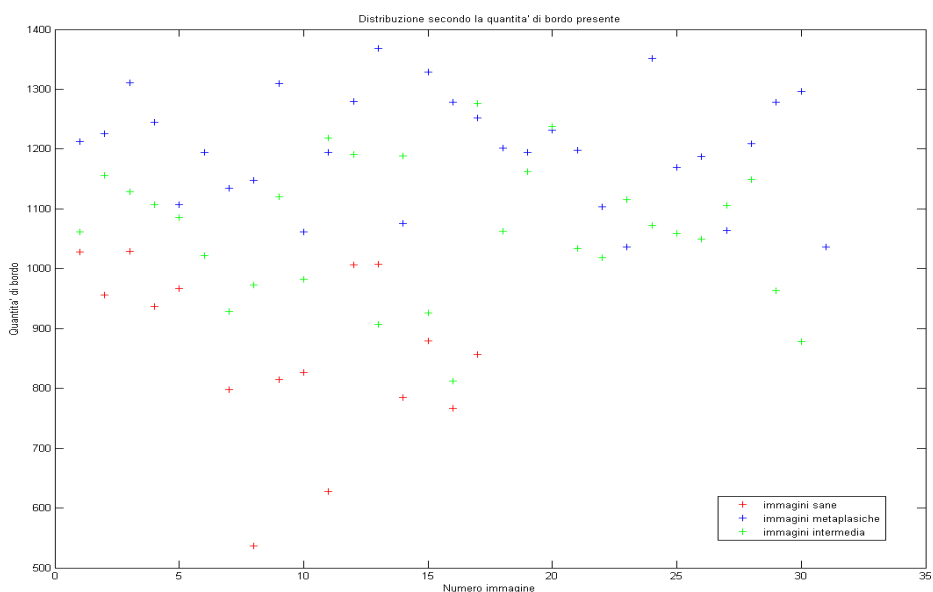


Figura 30 Nel grafico sono comprese le zone di confine tra tessuto sano e metaplastico, cioè le matrici che contengono entrambi i tipi di tessuto (crocette verdi). Il valore della “soglia di Canny” è 0.15. Si nota come, a seconda della quantità di tessuto dell’uno o dell’altro tipo in ciascuna matrice, le crocette verdi si distribuiscono ora nell’uno ora nell’altro dei due insiemi di valori.

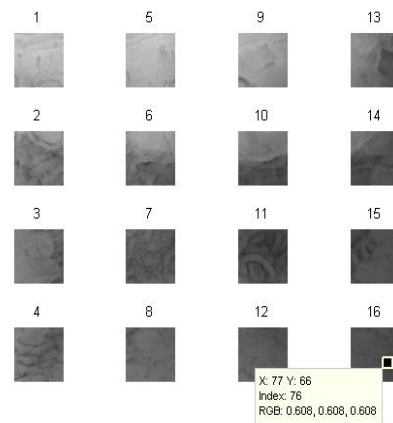
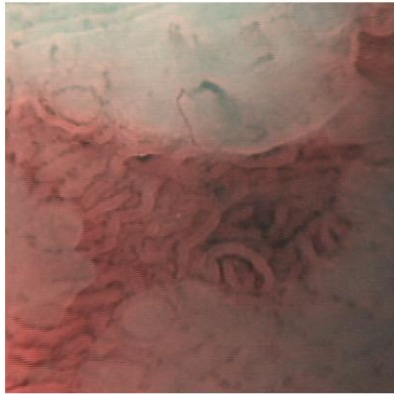


Figura 31

Figura 32

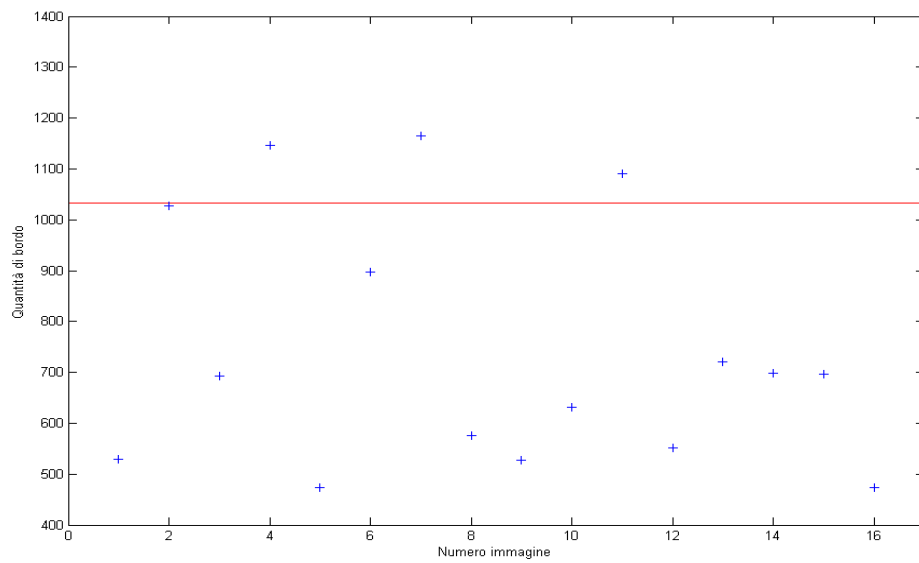


Figura 33

La Figura 31 mostra un'immagine 400x400 pixel, scelta dall'insieme di test, in cui è presente sia tessuto sano che metaplastico. La Figura 32 mostra la suddivisione dell'immagine in 16 matrici 100x100. La classificazione delle matrici è raffigurata nel grafico di figura Figura 33, dove la linea rossa rappresenta il valore della "soglia di quantificazione dei bordi", fissato a 1032.

L'algoritmo è stato infine testato sulle 176 immagini dell'insieme di test, costituito da matrici 100x100 pixel di tessuto sano e metaplastico. Nell'immagine sottostante sono riportati i risultati della classificazione, effettuata stabilendo la "soglia di Canny" a 0.15 e la "soglia di quantificazione dei bordi" a 1030.

L'unico errore di classificazione certo è relativo all'immagine numero 129 (indicata dalla freccia), che viene erroneamente classificata come sana. I valori che si distribuiscono, invece, molto vicini alla linea rossa (che identifica la "soglia di quantificazione dei bordi") sono relativi ad immagini di confine tra il tessuto sano ed il tessuto metaplastico, cioè che presentano entrambi i tipi di tessuto. Per questo tipo d'immagini la classificazione non è precisa perché dipende dalla quantità dell'uno o dell'altro tipo di tessuto presente nell'immagine.

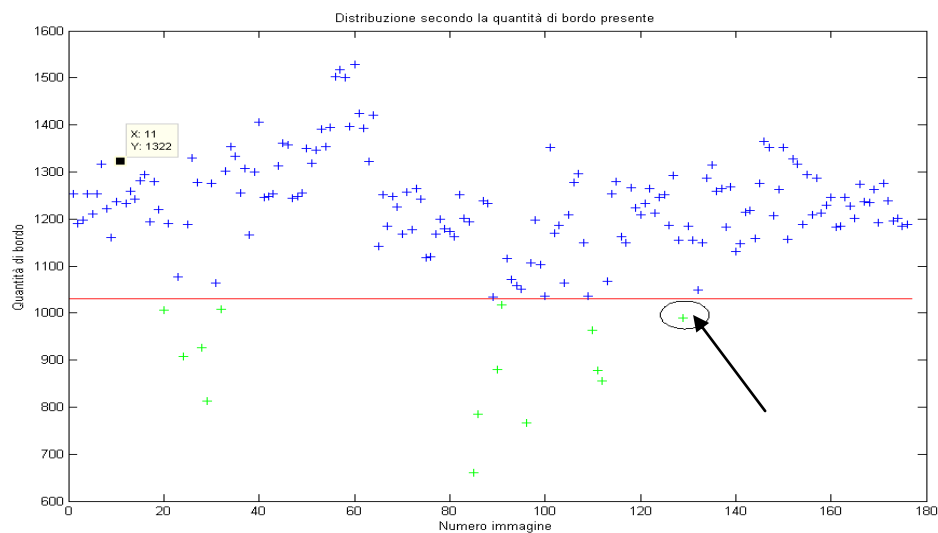


Figura 34 Classificazione delle immagini dell'insieme di test. Le crocette verdi indicano le immagini che la procedura ha riconosciuto come sane, mentre le crocette blu sono immagini riconosciute come malate.

4.2 Metodo di classificazione tipo di metaplasia

Il metodo “corr” precedentemente trattato è stato “calibrato” attraverso un insieme di training costituito da :

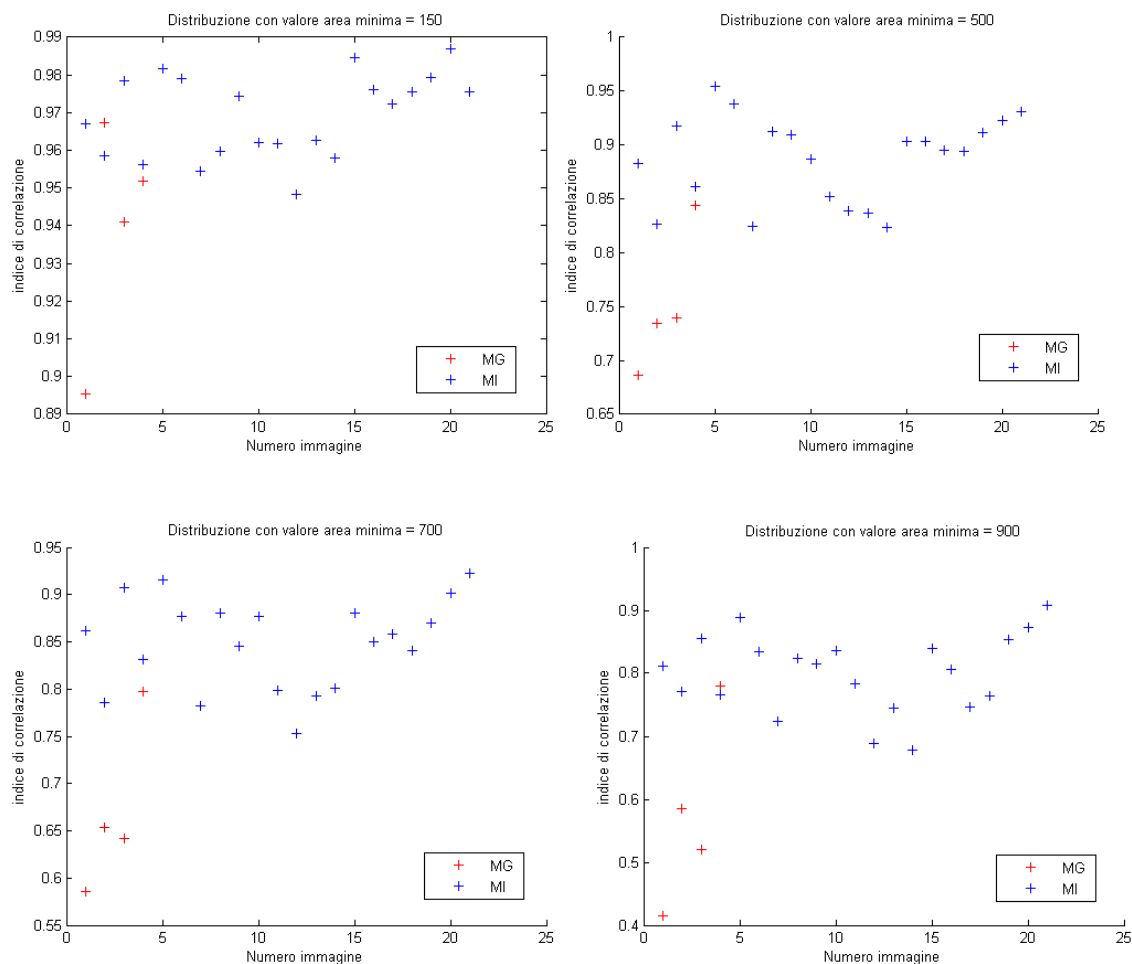
- 4 immagini di metaplasia gastrica
- 21 immagini di metaplasia intestinale

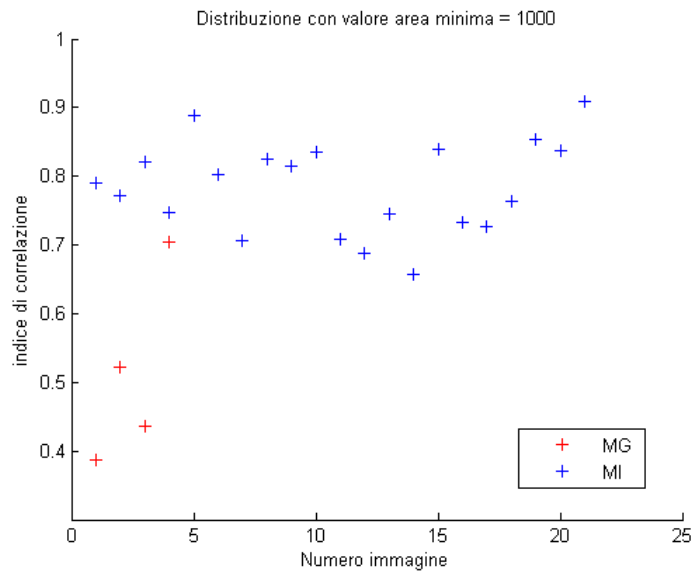
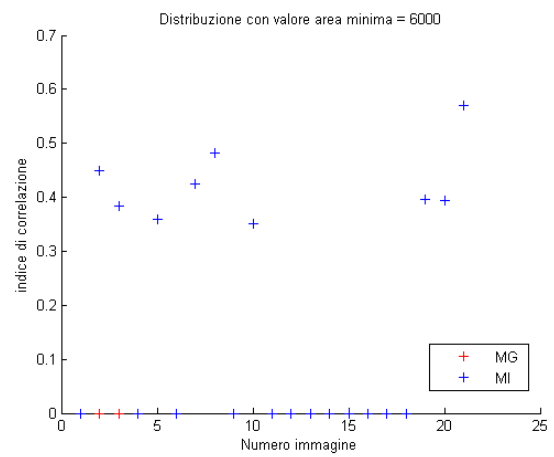
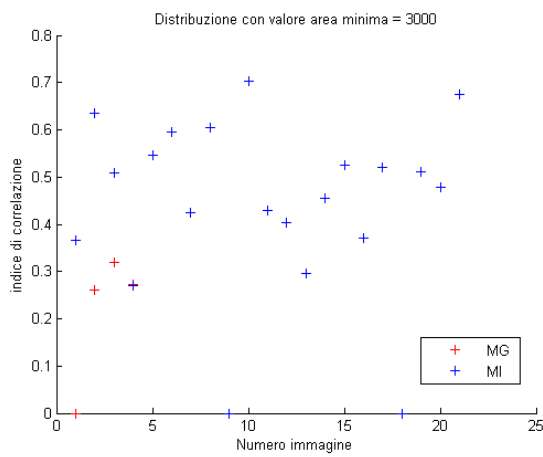
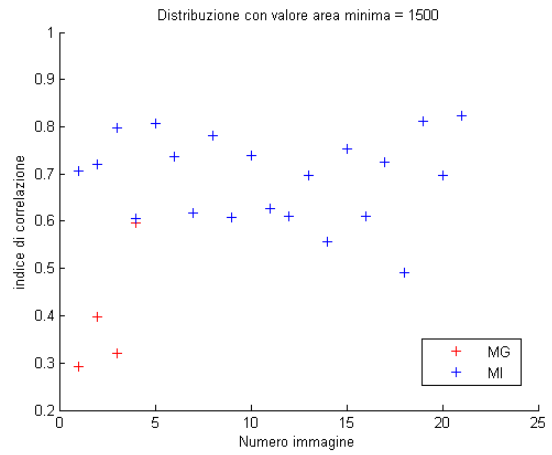
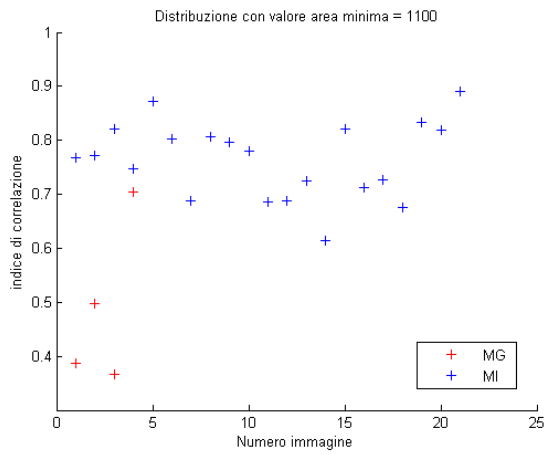
di modo che fosse possibile stabilire sia il valore di area minima necessario alla procedura di selezione sia un valore soglia per l'indice di correlazione in modo da poter classificare i due diversi tipi di malattia.

Si è poi applicato l'algorithm così tarato su un insieme di test che comprende:

- 2 immagini di metaplasia gastrica
- 10 immagini di metaplasia intestinale

Qui di seguito sono riportati i grafici che hanno permesso l'individuazione del valore di area minima necessario nella procedura di selezione.





Analizzando le varie distribuzioni si è giunti alla conclusione che il valore di area minima può essere fissato sul valore 1000.

E' stato inoltre deciso, valutando la distribuzione di valori nel grafico relativo al valore di area minima 1000, che una possibile soglia in grado di discriminare le due tipologie di malattia può essere fissata sul valore di 0.65.

Ciò ha portato al seguente risultato di test :

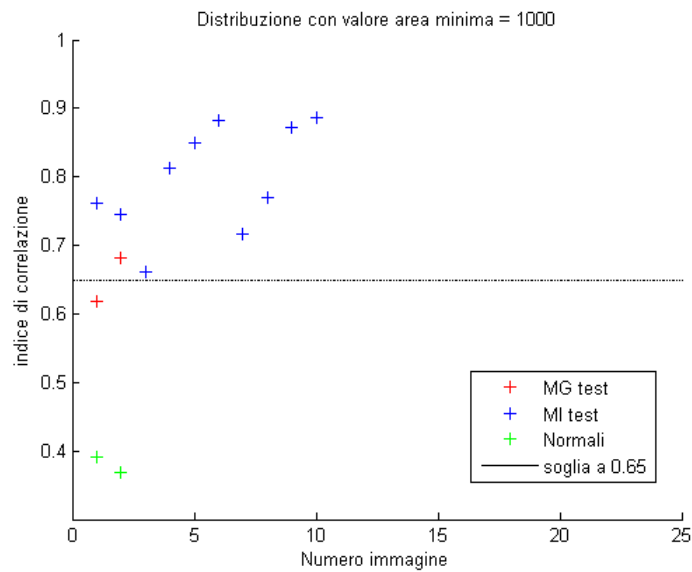


Figura 34

Per quanto possano essere consistenti i risultati, dato il numero esiguo di immagini sia di training che di test, si possono fare due considerazioni importanti:

- Le immagini di metaplasia intestinale sono classificate correttamente
- La seconda immagine di metaplasia gastrica è classificata in modo errato

La causa di questo errore di classificazione è stata attribuita all'ampia parte di quello che sembra essere tessuto sano (freccia rossa) e ad una "zona d'ombra" (freccia blu) come si vede nell' immagine in questione (Figura 35) .

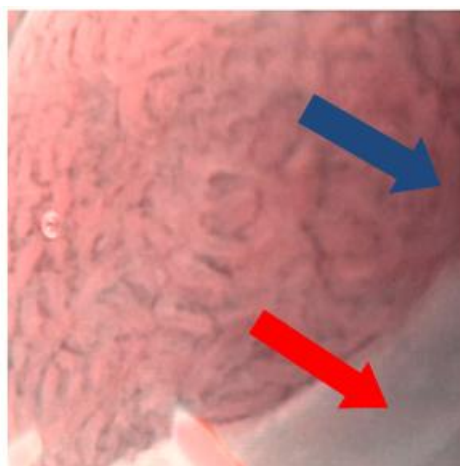


Figura 35

Il risultato della procedura per quest'immagine è illustrato qui di seguito (Figura 36) .

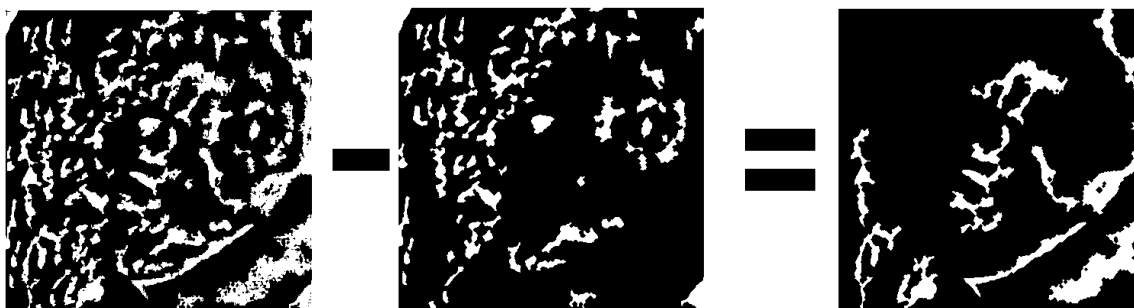


Figura 36

Si notano due grandi regioni connesse che alzano l'indice di correlazione e quindi il risultato in questo caso non è soddisfacente .

Un'ultima nota, riguardante il grafico della Figura 36 , è la presenza di due immagini (in verde) di tessuto normale . Esse vengono classificate come tessuto malato di tipo gastrico poi sarà il codice per rilevare la parte sana a chiarire la reale natura del tessuto.

5 Conclusione e sviluppi futuri del metodo “corr”

Al di là del fatto che il sistema sia basato su soglie fisse e quindi soggetto a tutti i limiti che queste causano soprattutto se si hanno insiemi di training ridotti, il metodo illustrato per distinguere immagini di metaplasia gastrica e immagini di metaplasia intestinale ha un buon margine di funzionamento se la qualità dell'immagine d'ingresso è buona.

Infatti, un possibile miglioramento per aumentarne l'efficacia è sicuramente quello di porre maggior attenzione sull'elaborazione iniziale dell'immagine in ingresso.

Ad esempio, problemi dovuti a “zone d'ombra” come quella vista in precedenza dovrebbero essere affrontati in maniera specifica da algoritmi con il preciso scopo di rendere omogenea la luminosità dell'immagine.

Tuttavia il limite effettivo, ben più stringente, di tale metodo è il fatto che si parte dall'ipotesi che le immagini in ingresso siano totalmente malate o con una limitata quantità di tessuto sano. Spesso questo può verificarsi ma, ci sarà sempre la possibilità che in ingresso si presenti un'immagine con una quantità di tessuto sano tale da ridurre di molto lo spazio effettivo che può ospitare le strutture tipiche di metaplasia intestinale e quelle tipiche di metaplasia gastrica.

Più l'area che ospita tali strutture si riduce più aumenta il rischio di una classificazione errata poiché regioni scure di vaso e regioni scure non-di-vaso divengono via via più confrontabili.

In conclusione volendo sviluppare effettivamente qualcosa di molto robusto ed efficace, è assolutamente necessario passare da un approccio d'insieme ,come quello visto, ad uno più puntuale ,locale ,che sia in grado di fornire risultati corretti indipendentemente dalla quantità di tessuto sano presente nell'immagine da classificare.

Bibliografia

- [1] Guyton & Hall (2001) *Fisiologia Medica*, Edises Edizioni Scientifiche ed Universitarie
- [2] Kiesslich, R & Galle, P & Neurath, M (2008) *Atlas Of Endomicroscop*, WW
- [3] Yung-Sheng Chen (2009) *Image Processing*, InTech Publication
- [4] www.intechweb.org
- [5] www.ioveneto.it
- [6] [www.mathworks.com help/toolbox/images](http://www.mathworks.com/help/toolbox/images)
- [7] <http://www.news-medical.net/health/Barretts-Esophagus-Pathology>
- [8] http://esofagobarrett.iannetti.it/esofagobarrett/esofagite_di_barrett
- [9] <http://it.wikipedia.org/wiki/Stomaco>

Ringraziamenti ... si ringrazia ...

Per la possibilità concessa e l'aiuto gentilmente dato: *Dr. Enrico Grisan , Dott.ssa Elisa Veronese.*

Per il materiale gentilmente concesso l' *Istituto Oncologico Veneto (IOV).*

Per la magnifica collaborazione *Giulio .*

Per il supporto morale : *i miei genitori , Chiara , Alessandro , Andrea , Marco , Giovanni, Filippo , Stefania e Viviana e ovviamente tutti gli amici .*

<<...per aspera ad astra.....>>