

POLITECNICO DI TORINO

A.A. 2019/2020

TESI DI LAUREA MAGISTRALE IN INGEGNERIA
BIOMEDICA

QUANTIFICAZIONE SEMI-AUTOMATICA DI
CARATTERISTICHE ULTRASONOGRAFICHE
DEL NODULO TIROIDEO



**POLITECNICO
DI TORINO**

Autore:

Simone Pagliarulo

Relatore:

Filippo Molinari

Correlatore:

Bruno De Santi

Questi ringraziamenti sono per me la parte finale della mia tesi e del mio cammino universitario. Dopo molte risate e pianti, alti e bassi, 30 e 18, sono giunto anche io al termine di questo periodo di studi. Il mio pensiero va a chi ha reso possibile tutto questo, a chi non mi ha dato il piatto caldo ma ha fatto in modo che io imparassi a cucinare, a chi non ha mai smesso di credere in me, a chi ahimè non può più vedermi ma ci sarà sempre, a chi mi ha sostenuto, a chi mi ha sopportato, a chi mi ha guidato e soprattutto a chi mi ha dato la forza, la pazienza e la determinazione per raggiungere questo traguardo.

Vorrei ringraziare il professor Molinari, relatore di questa tesi di laurea, per l'aiuto fornitomi in tutti questi anni e la grande conoscenza che mi ha donato.

Vorrei ringraziare il correlatore De Santi oltre che per l'aiuto tecnico e didattico, per la disponibilità e ma soprattutto per la pazienza dimostratemi. Senza loro due questo lavoro non avrebbe preso vita in quanto si sono dimostrati ottime guide durante tutto il periodo di stesura.

Un ringraziamento sentito va alla mia famiglia che mi ha educato, cresciuto e iniziato alla vita. Loro sono stati la fonte di tutto. Mamma, Papà e i due bambocci, Manuel e Ludovica, spero siate orgogliosi di me.

Vorrei ringraziare i parenti tutti, che mi hanno sempre appoggiato nelle mie scelte e hanno contribuito, ognuno a proprio modo, a sostenermi in questi anni.

Vorrei ringraziare gli amici che mi hanno aiutato nei momenti di difficoltà e che mi hanno dato una mano nella carriera universitaria.

Vorrei ringraziare coloro che hanno rivolto un pensiero a me durante una loro qualsiasi giornata.

Infine vorrei ringraziare anche te per vabbè superfluo esprimersi... Grazie

Un sentito GRAZIE a tutti voi, se sono arrivato fin qui ricordatevi che è anche merito vostro.

Indice

1	Introduzione	7
1.1	Tiroide	8
1.2	Patologie Tiroidee	11
1.3	Imaging Ad Ultrasuoni	15
1.3.1	Immagini Ecografiche	15
1.3.2	Immagini Doppler	16
1.3.3	Immagini Elastosonografiche	17
1.3.4	Struttura di un Ecografo	19
1.3.5	Tecniche di Esame	21
1.4	Caratteristiche ultrasonografiche	23
1.4.1	Ecogenicità	23
1.4.2	Struttura	23
1.4.3	Margini	25
1.4.4	Presenza di Cisti	26
1.4.5	Forma	26
1.4.6	Vascolarizzazione	27
1.4.7	Elasticità	28
1.5	Classi TIRADS	30
1.6	Iter Diagnostico	32
1.7	Stato dell'Arte	34
2	Materiale e Metodi	37
2.1	Immagini Ecografiche	37
2.2	Segmentazione manuale	38
2.3	Valutazione delle Caratteristiche Ultrasonografiche	39
2.4	Principi di Statistica	40
2.4.1	Correlazione	40
2.4.2	Mann-Whitney U-Test	41
2.4.3	Analisi della Componenti Principali	41
2.4.4	Manova	43
2.5	Quantificazione Automatica delle Caratteristiche Ultrasonografiche	45
2.5.1	Ecogenicità	46
2.5.2	Struttura	52
2.5.3	Margini	56
2.5.4	Cisti	61
2.5.5	Forma	64
2.5.6	Vascolarizzazione	66
2.5.7	Elastosonografia	70
2.6	GUI	73
2.6.1	Struttura Graphic User Interface	73
2.6.2	Utilizzo Graphic User Interface	75
2.6.3	Caratterizzazione Graphic User Interface	76
3	Validazione e Risultati	81
3.1	Ecogenicità	81
3.2	Struttura	84
3.3	Margini	85
3.4	Cisti	86
3.5	Forma	87
3.6	Vascolarizzazione	88
3.7	Elastosonografia	89
3.8	Classificazione Caratteristiche Ultrasonografiche	90
4	Discussioni	95

1 Introduzione

L'imaging ad ultrasuoni ricopre un ruolo chiave nella diagnostica e nella gestione dei noduli tiroidei. Le caratteristiche numeriche per quantificare e esaminare le trame, intese come particolarità di una superficie, sono utilizzate in diversi ambiti applicati. Abitualmente, nell'imaging della tiroide la "trama" si riferisce alla comparsa ecografica di un nodulo, che può essere definito come per esempio omogeneo, eterogeneo, solido, cistico, ipoecogeno, isoecogeno, misto, iperecogeno. Tuttavia, la definizione di queste caratteristiche risulta essere operatore dipendente ed è influenzata dall'esperienza dell'operatore stesso. Negli ultimi anni, sono stati fatti sforzi per rendere più obiettiva la valutazione della struttura dell'eco dei noduli tiroidei (eliminando i potenziali pregiudizi dell'operatore) al fine di standardizzare la segnalazione, e meglio definire il rischio di malignità per evitare biopsie non necessarie. *Obiettivo*

Questa tesi nasce con lo scopo di quantificare in maniera semi-automatica delle caratteristiche ultrasonografiche di noduli tiroidei e classificarli. Alcune di queste caratteristiche ultrasonografiche sono costantemente predittive di malignità e sono utilizzate come criteri per l'FNA (fine needle aspiration). In tale tesi si propone un algoritmo semi-automatico in grado di selezionare solo alcune delle features, focalizzandosi sulle più rappresentative, in modo da valutare il rischio di malignità. Per migliorare la qualità dell'estrazione delle caratteristiche, si è provveduto ad ottimizzare l'immagine a seconda della feature da valutare, e per accentuarne la praticità si è creata una GUI (graphical user interface); i risultati qualitativi sono riportati nei capitoli seguenti. Il seguente elaborato si pone l'obiettivo primario di descrivere ed implementare l'impiego di nuove metodologie semi-automatiche per l'estrazione e la quantificazione di caratteristiche ultrasonografiche, in grado di discriminare le lesioni tiroidee nella classificazione. In altre parole si cerca di migliorare le potenzialità diagnostiche tenendo conto del costo limitato che l'impiego di questa metodologia offre. Gli step compiuti e l'interfaccia grafica descritti di seguito possono essere integrati con qualsiasi strumento ecografico diagnostico fornito di monitor, gravando però sul costo computazionale e aumentando leggermente i tempi di diagnosi. L'elaborato si compone di tre macro sezioni:

1. La prima parte riassume i concetti fondamentali alla base dell'indagine ecografica, fornendo le nozioni teoriche ed esplicando i principi fisici che governano l'imaging ad ultrasuoni.
2. La parte centrale illustra lo sviluppo e gli step compiuti per giungere all'estrazione delle features, soffermandosi anche sul miglioramento e sull'impiego tramite interfaccia grafica (GUI) del lavoro svolto.
3. L'ultima parte analizza i risultati e valuta la qualità dei risultati, soffermandosi sulla validità degli stessi, analizzandone limiti e punti di forza ma proponendo anche nuove strategie per il miglioramento della quantificazione delle caratteristiche.

1.1 Tiroide

Oggi giorno l'indagine diagnostica iniziale per il riscontro della presenza di lesioni tiroidee è ancora la palpazione manuale. L'esame clinico offre al medico una misura tattile delle proprietà meccaniche del tessuto in esame, ovvero ne delinea una grandezza qualitativa. La proprietà meccanica di primaria importanza che emerge da questa indagine è la rigidità. Tuttavia questa metodologia ha diversi limiti come la difficoltà di interazione con i tessuti in profondità e la dimensione delle lesioni rilevabili, rendendo così invisibili lesioni profonde e microscopiche. La tiroide, in Fig.1, è la più grande ghiandola del corpo umano ed è situata nella regione anteriore del collo. Si presenta con una forma "a farfalla" ed è costituita da due lobi collegati tra di loro da un istmo. Presenta una straordinaria variabilità di peso e dimensioni che dipendono dall'età, dal sesso e dall'ambiente in cui si vive.

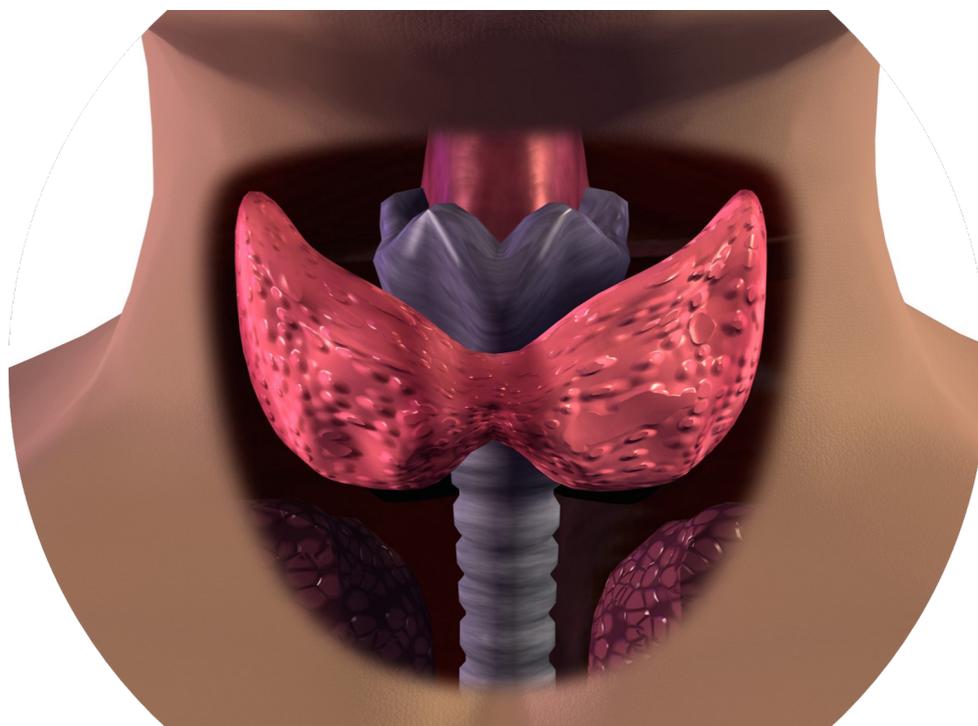


Figura 1: Posizione anatomica e forma della tiroide (tratto da [4])

I legamenti che costituiscono la guaina peritiroidea fissano la ghiandola nella parte anteriore del collo. A causa della sua posizione essa è in stretto contatto con la laringe, una porzione della trachea e con la carotide comune, la vena giugulare interna e il nervo vago. La ghiandola è avvolta in una guaina che comprende una rete vascolare che divide la superficie interna della guaina dalle quattro ghiandole paratiroidi che compongono l'organo. Da questa guaina nascono legamenti che tengono salda la tiroide alla cartilagine della laringe. L'unica fascia muscolare ad evitare il contatto tra la tiroide e la cute nella regione mediana è la fascia cervicale che in combinazione con i muscoli sternocleidioideo, sternotiroideo e omoioideo copre parzialmente la ghiandola. Il muscolo sternocleidomastoideo e il muscolo platisma omolaterali con la loro posizione anatomica cooperano per rivestire la ghiandola nella regione laterale. La tiroide è vascolarizzata principalmente da un complesso di quattro arterie, cioè le arterie tiroidee superiori e inferiori (separate

in due per lato), e drenata da un complesso di sei vene, cioè le vene tiroidee superiori, medie e inferiori (divise in tre per lato). Il parenchima tiroideo è rivesto da una capsula di tessuto connettivo, che invia all'interno della ghiandola tralci connettivali che fungono da confine per la divisione dell'organo in aree irregolari. Il parenchima presenta una struttura follicolare incorporando delle vescicole. I follicoli tiroidei sono delle vescicole chiuse la cui parete, l'epitelio follicolare, è formata dai tireociti. Queste cellule sono le principali cellule a cui spetta il compito della produzione degli ormoni. La colloide un materiale amorfo è incapsulato all'interno di queste vescicole. La produzione di calcitonina, ormone immischiato nel riassorbimento del calcio osseo, è specifica delle cellule parafollicolari le quali sono localizzate nello stroma della tiroide.

La sintesi di ormoni contenenti iodio triiodotironina (T3), tiroxina (T4) e della calcitonina dell'ormone peptidico è il principale scopo della ghiandola. Gli ormoni prodotti influenzano diversi aspetti della fisiologia e della vita umana [36]:

- **Metabolismo:** gli ormoni tiroidei aumentano il metabolismo basale avendo effetti su quasi tutti i tessuti del corpo. Ricoprono un ruolo fondamentale nell'assorbimento di sostanze nutritive nell'intestino, incrementandolo. Aiutano il ricambio cellulare. Incrementano la scissione del glucosio. Provocano la decomposizione dei grassi. Incrementano il numero di acidi grassi liberi, tuttavia influenzano anche i livelli di colesterolo nel sangue.
- **Cardiovascolare:** la velocità e la forza di contrazione del miocardio sono influenzate da ormoni tiroidei. Giocano un ruolo fondamentale nell'aumento della frequenza respiratoria e quindi nell'utilizzo di ossigeno. Incrementano anche l'attività dei mitocondri. La combinazione di queste attività aumenta il flusso sanguigno e per diretta conseguenza la temperatura corporea.
- **Sviluppo:** gli ormoni di principale utilizzo sono i T3 e T4, che hanno lo scopo di aumentare lo sviluppo dei neuroni e la crescita dei giovani. Sono di fondamentale importanza per quanto concerne il normale sviluppo. Gli ormoni tiroidei ricoprono una funzione basilare nella maturazione della rete neuronale durante lo sviluppo del feto
- **Altro:** gli ormoni tiroidei hanno il compito di mantenere una normale funzione sessuale, del sonno e dei modelli di pensiero. L'elevata presenza di ormoni è sinonimo di maggiore velocità di generazione del pensiero ma con minore attenzione. Il mantenimento di un regolare ciclo mestruale e la libido, vengono regolate da questi ormoni.

Nel sistema cardiocircolatorio è presente solo un minima percentuale di ormoni tiroidei. L'attività ormonale è comandata esclusivamente dallo 0,03% di T4 e lo 0,3% di T3 che si spostano liberamente nel sangue. La maggior parte è legata alla tiroxina (circa il 70%), alla transtiretina (10%) e all'albumina (15%). L'azione degli ormoni tiroidei è simile al riconoscimento delle proteine: avviene tramite il legame di recettori specifici (TR- α 1, TR- α 2, TR- β 1 e TR- β 2) sovra espressi dalla membrana cellulare che catturano gli ormoni che a loro volta agiranno sulla trascrizione del DNA modificandola in base alle esigenze. L'azione degli ormoni è anche intracellulare, in particolare reagiscono con enzimi all'interno del citoplasma quali ATPasi di calcio, adenilato ciclasi e con i trasportatori di glucosio catalizzando le reazioni. Una volta entrati in circolo, gli ormoni possono legarsi a proteine plasmatiche. Il legame più diffuso, 70-75% degli ormoni, è con la globulina che ne assicura

una vita più lunga trasformandoli in riserva se i livelli ormonali sanguigni sono troppo elevati. La concentrazione costante di ormoni nei tessuti è basata proprio su questo legame il quale viene scisso nel caso di carenza ormonale in una specifica regione del corpo, come per esempio in un organo. L'elemento alla base della produzione ormonale della tiroide è lo iodio. All'interno dei tireociti è sintetizzata una glicoproteina, la tireoglobulina, che viene depositata all'interno di una vescicola insieme ad una perossidasi. La tireoglobulina è una proteina costituita in media da 5000 amminoacidi, circa il 10% della massa totale sono carboidrati, mentre meno dell'1% è costituito dallo iodio. Contemporaneamente, il torrente ematico provvede al rifornimento di ioduri i quali sono introdotti nelle vescicole e ossidati dalla perossidasi e infine uniti ai residui di tirosina della tireoglobulina. Unendo un atomo di iodio si viene a creare una molecola di monoiodiotirosina (MIT) se invece si legano due atomi di iodio alla molecola si crea la diiodiotirosina (DIT). Grazie a queste molecole è possibile l'assemblaggio degli ormoni T3 (MIT+DIT) e T4 (DIT+DIT), che impiegano rispettivamente 3 o 4 atomi di iodio. Tutti questi processi, i quali portano alla sintesi di degli ormoni tiroidei, avvengono all'interno delle vescicole. Il contenuto delle vescicole viene poi riversato all'interno del lume del follicolo ed è accumulato sotto forma di colloide. Grazie alla presenza del colloide la tiroide è l'unica ghiandola endocrina del corpo umano che possiede la capacità di immagazzinare ormoni da rilasciare in sede extracellulare prima che esso venga secreto nel sangue. L'adeguata quantità di iodio permette una corretta sintesi ormonale da parte della tiroide: il fabbisogno giornaliero di iodio per la sintesi degli ormoni tiroidei è di circa 150 microgrammi, che aumentano notevolmente durante la gravidanza poiché sono necessari per lo sviluppo del feto. Anche un altro elemento, il selenio, è di fondamentale importanza poiché è un cofattore delle deiodinasi (enzimi responsabili della conversione della T4 in T3) e aiuta alla salvaguardia funzionale della tiroide [36].

1.2 Patologie Tiroidee

Un non adeguato funzionamento nella produzione ormonale della tiroide è spesso causato dalla presenza di nodi e tumori benigni o maligni, che solitamente intaccano le proprietà funzionali della ghiandola. Questi disordini comportano alcune patologie come la tiroidite, un processo infiammatorio. Il così detto "gozzo", un rigonfiamento del collo e l'allargamento della ghiandola, è spesso causato da patologie disfunzionali della tiroide con sovrapproduzione di ormoni. La localizzazione delle lesioni è spesso la ghiandola stessa. I noduli riscontrati sono comunemente asintomatici e non cancerosi. Nonostante la poca incidenza di malignità, i noduli che includono carcinomi follicolari, carcinomi papillari, carcinomi midollari e metastasi provenienti da altri siti si classificano come maligni, tuttavia si mostrano solo in una percentuale molto bassa (5%) rispetto al totale. Risulta più probabile riscontrare la presenza di un nodulo nelle donna, in color che sono frequentemente esposti a radiazioni ionizzanti e in coloro che hanno un inadeguato consumo di iodio. L'ipertiroidismo, ovvero l'eccessiva produzione di ormoni e una grande attività della ghiandola, è frequentemente il prodotto della malattia di Graves. Tuttavia ci possono essere altre cause che conducono a un'attività elevata della tiroide quali un gozzo multinodulare tossico, un adenoma o una semplice infiammazione. In opposizione, l'ipotiroidismo, ovvero la carenza di attività della ghiandola e la poca produzione di ormoni, è regolarmente causata da carenza di iodio. La patologia che induce ipotiroidismo è la tiroide di Hashimoto, una malattia autoimmune. Esistono numerose patologie tiroidee e si dividono in autoimmuni, tiroide Hashimoto, tiroide transitoria sporadica e tiroide transitoria post-partum, e non autoimmuni, come la tiroide di Riedel, tiroide acuta infettiva, tiroide sub-acuta di De Quervaine.

La caratteristica più comune nei noduli maligni, o cancri, tiroidei è la crescita eccessivamente rapida. A causa di ciò, la valutazione del rischio di malignità risulta di primaria importanza nei pazienti con noduli sospetti, al fine di selezionare coloro che dovrebbero sottoporsi a una biopsia con ago aspirato (FNA), adottato come Ground Truth. Anche in Italia il riscontro delle patologie è alto: si presentano circa 15300 nuovi casi, quali 11000 sono donne [36]. La grande maggioranza delle lesioni è rappresentata da neoplasie benigne, con una percentuale molto bassa di lesioni maligne. L'associazione italiana dei registri tumori (AIRTUM) ha registrato dal 2008 al 2013 i seguenti dati di tumore della tiroide nella popolazione italiana:

- 82% carcinoma papillare
- 7% carcinoma follicolare
- 3% carcinoma midollare
- 1% carcinoma anaplastico

nonostante l'alta probabilità di incidenza del tumore, la mortalità ad esso legato non è alta. La possibilità di sopravvivere è correlata con l'età alla diagnosi. Il tumore alla tiroide è il quinto cancro più comune nella donna, sotto i 45 anni, in tutto il mondo e il secondo in Italia.

Si definisce formazione di un nodulo la crescita anomala di un gruppo di cellule che compongono la tiroide stessa. Il tumore alla tiroide riscontra un'incidenza molto elevata, in oltre il 50% della popolazione mondiale adulta, la maggior parte dei noduli è benigna e solo meno del 7% è maligna [1]. La Fig.2 mostra , la frequenza di incidenza, la prognosi e le forme tumorali le quali posso essere identificate in [34]:

- Carcinoma Papillare, è causa del 50% dei tumori maligni tiroidei ed è costituito da papille più o meno ramificate con un asse connettivale rivestito da cellule cubiche
- Carcinoma Follicolare, costituito da piccoli follicoli, può essere incapsulato, oppure invadere la capsula presentando precoci metastasi
- Carcinoma Anaplastico, è costituito da una massa dura infiltrante i tessuti circostanti ed ha un alto grado di malignità
- Carcinoma Midollare, è costituito da un nodulo esteso fatto di cellule poligonali e produce una grande quantità di calcitonina, ormone che regola il ricambio di calcio nelle ossa e spesso causa metastasi per via linfatica ed ematica.

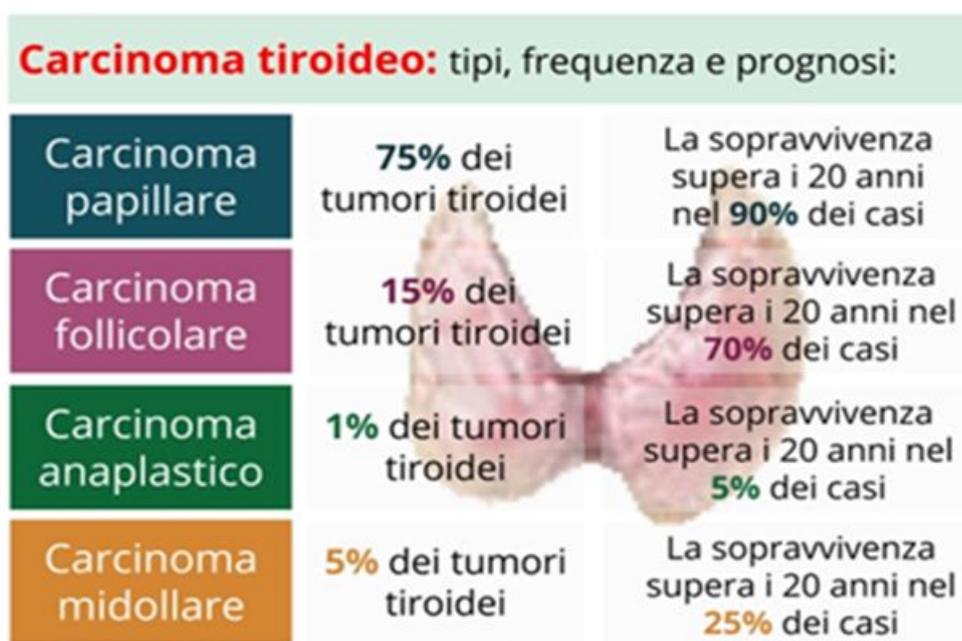


Figura 2: Rappresentazione tipologie, frequenza di incidenze e prognosi del carcinoma tiroideo(tratto da [7])

Solitamente la neoplasia si presenta in maniera asintomatica, soprattutto nella stadiazione iniziale. La crescita del nodulo non deve obbligatoriamente sviluppare caratteristiche di iper- o ipo-tiroidismo. A causa della mancanza di sintomi, come non avviene nelle patologie infiammatorie, è difficile riscontrare focolai attivi e potenzialmente pericolosi. La crescita di un nodulo tiroideo, può variare da pochi mesi (nel caso di tumori più aggressivi) fino ai dieci anni (nel caso di noduli benigni). L'ingrandimento del nodulo può intaccare il normale funzionamento delle strutture limitrofe, le quali devono affrontare una riduzione di spazio funzionale, e questi intaccamenti fungono da sintomatologia tipica della patologia. Tuttavia l'iperplasia della ghiandola invadendo le strutture circostanti può danneggiare le strutture nervose laringee, situate a stretto contatto con la tiroide, inducendo cambi nel tono di voce. Nei casi più gravi può anche interessare il sistema respiratorio, provocando dispnea, e nei casi più estremi coinvolgere il sistema digerente, comprimendo l'esofago causando disfagia. Il coinvolgimento di tessuti extra-tiroidei e altri organi è spesso dovuto alla natura della neoplasia. Indagando la caratteristica istologica del nodulo è emerso che l'istotipo papillifero è più propenso a creare metastasi per via linfatica e raramente per via ematica, invece l'istotipo follicolare ha tendenze opposte preferendo la metastatizzazione a

distanza sfruttando maggiormente il torrente ematico a discapito del linfatico. Oggigiorno per avere la sicurezza e la precisione di diagnosi si utilizza il Fine Needle Aspiration (FNA), cioè l'aspirazione con ago sottile. Le masse sospette, spesso legate a patologie, vengono analizzate con questa tecnica, mostrata in Fig.3. Essa prevede l'impiego di un ago sottile (diametro 23-25mm) che trapassa la capsula tumorale ed aspira una porzione di tessuto. Il tessuto asportato viene successivamente sottoposto a reagenti coloranti che ne delineano la natura. L'esame termina con la diagnosi dopo aver monitorato il tessuto al microscopio. Oggigiorno le biopsie per aspirazione con ago sottile sono definite procedure chirurgiche minori e hanno raggiunto un elevato grado di sicurezza e accuratezza.

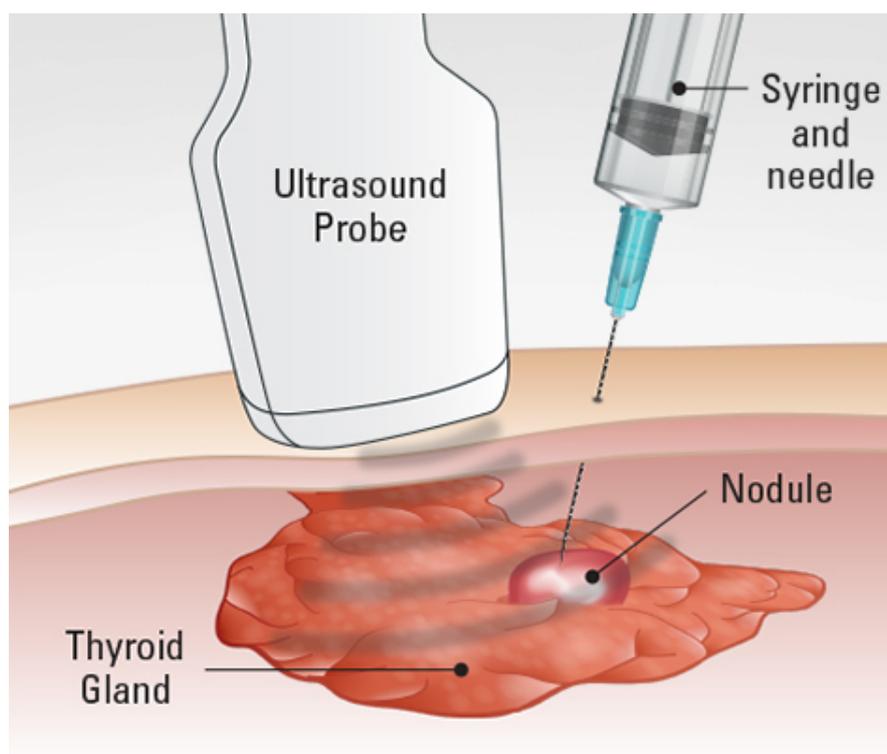


Figura 3: Esempio di esecuzione tecnica Fine Needle Aspiration tramite ago e sonda ad US, FNA (tratto da [29])

La biopsia di aspirazione dell'ago sta prendendo ampiamente piede e sta sostituendo la biopsia chirurgica (escissionale o aperta) poichè evita la necessità di ricovero ospedaliero per un esame clinico. Il FNA è principalmente utilizzata nella diagnosi di cancro e condizioni infiammatorie. I principali effetti collaterali della nuova tecnica sono dolore e lividi per il paziente. Invece per il medico le complicanze che si possono presentare sono la perdita del volume di prelievo, essendo minimo si potrebbero perdere delle cellule, con conseguente falso negativo, ma anche l'eventualità di prelievo di cellule che non consentano una diagnosi definitiva. Lo scopo principale dell'agoaspirato è distinguere i pazienti che devono essere sottoposti a trattamento chirurgico da quelli a cui necessitano solo di monitoraggio [3]. L'esame citologico consente di attribuire ad ogni paziente una categoria diagnostica ben definita e identificabile con un codice numerico, secondo la classificazione SIAPEC 2014 (Classificazione della Società Italiana di Anatomia Patologica e Citopatologia Diagnostica) [20], in Fig.4.

Codice	Categoria diagnostica	Rischio di malignità atteso (%)	Azione clinica suggerita
TIR1	Non diagnostico	Non definito	Ripetere FNA con guida ecografica
TIR1C	Non diagnostico cistico	Basso, variabile in base al quadro clinico	Valutare nel contesto clinico eventualmente ripetere FNA
TIR2	Non maligno/benigno	< 3	Follow up
TIR3A	Lesione indeterminata a basso rischio	< 10	Ripetere FNA/follow up
TIR3B	Lesione indeterminata ad alto rischio	15-30	Exeresi chirurgica
TIR4	Sospetto di malignità	60-80	Exeresi chirurgica con eventuale esame intraoperatorio
TIR5	Maligno	95	Exeresi chirurgica Approfondimento diagnostico in casi selezionati

Figura 4: Classificazione SIAPEC 2014 per l'esame citologico (tratto da [11])

1.3 Imaging Ad Ultrasuoni

1.3.1 Immagini Ecografiche

Quando si parla di ultrasonografia in ambito biomedicale si tende ad associare il termine a tecniche di imaging spesso utilizzate come strumento diagnostico. Questa è una metodologia operatore-dipendente, è in grado di acquisire immagini in tempo reale, utilizzando radiazioni non ionizzanti, non presentando quindi rischi per i pazienti ed è economica rispetto ad altre modalità di imaging. Un'immagine ecografica deriva dalla combinazione di diversi fenomeni fisici. In particolare, sono impiegati dei particolari cristalli per generare un fascio di onde meccaniche che si propaga nei tessuti. La generazione delle onde acustiche è dovuta all'effetto piezoelettrico. La velocità di propagazione degli ultrasuoni nei tessuti varia a seconda dell'elasticità del materiale e ha la seguente relazione con la frequenza di emissione [23]:

$$v = \lambda f$$

dove v è la velocità di propagazione, f è la frequenza e λ è la lunghezza d'onda. Un parametro che concorre alla creazione delle immagini è l'impedenza acustica. Questo parametro fondamentale è definito come la misura dell'opposizione di un sistema materiale all'onda acustica. Può essere calcolato per ogni tessuto con la seguente relazione:

$$Z = \rho v$$

dove v è la velocità di propagazione e ρ la densità media del tessuto. La velocità di propagazione media dei tessuti biologici è convenzionalmente pari a 1540m/s [23]. Il comportamento dell'onda sonora che interagisce con i tessuti è l'insieme di diversi fenomeni fisici. Prima di tutto, l'onda perde continuamente energia seguendo la relazione :

$$A(\zeta) = A_0 e^{-\alpha \zeta}$$

dove $A(\zeta)$ è l'ampiezza dell'onda alla profondità ζ , A_0 è l'ampiezza dell'onda iniziale e α è il coefficiente di assorbimento della materia che è proporzionale alla frequenza. Le onde acustiche non solo sono continuamente attenuate, ma subiscono altri due fenomeni, riflessione e trasmissione, quando incontrano una variazione di impedenza acustica. Queste interazioni sono associate a due coefficienti:

$$R = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2} \right)^2$$

$$T = 1 - R$$

Le onde riflesse sono rilevate elettronicamente dagli stessi cristalli emettitori. L'intensità luminosa delle immagini è proporzionale alla forza di ritorno dell'eco e il tempo di ritorno è convertito in profondità, secondo la relazione:

$$d = \frac{1}{2} c \Delta t$$

dove c è uguale alla velocità media di propagazione nei tessuti [23]. Esistono diverse tipologie di immagini ad ultrasuoni ottimizzate per applicazioni differenti, alcune delle quali adottano anche una codifica a colori. Le più comuni sono le immagini B-mode, basate sull'intensità luminosa pertanto, l'intensità luminosa dei pixel può classificarli come iperecogeni, ipoecogeni e anecogeni e sono una chiara rappresentazione della forza di riflessione del fascio sonoro. In questa tesi si sono ampiamente utilizzate immagini ecografiche B-mode, in Fig.5.

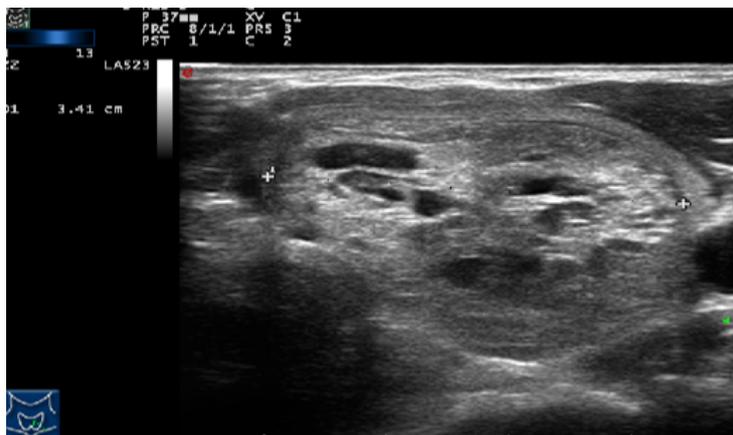


Figura 5: Immagine ultrasonografica di lesione tiroidea B-Mode

1.3.2 Immagini Doppler

L'ecografia Doppler è una tecnica di imaging non invasiva e comodamente ripetibile molto utilizzata per indagare la condizione vascolare, sia anatomica che funzionale, di un determinato tessuto per comprenderne la sua attività. L'effetto Doppler, che caratterizza queste immagini, è un fenomeno fisico che consiste nel cambiamento apparente dell'acutezza, intesa come frequenza, di un suono quando la sorgente è in movimento o l'osservatore stesso è in movimento verso la sorgente. I globuli rossi rappresentano le interfacce sulle quali si generano gli echi di ritorno. Se i flussi sono in avvicinamento rispetto alla sonda si riscontra un apparente aumento della frequenza o si riscontra una diminuzione apparente di frequenza se il flusso è in allontanamento. La velocità è calcolata tramite la relazione:

$$v = f_D \frac{c}{2f_0 \cos\theta}$$

in cui f_0 è la frequenza di emissione della sonda, c è la velocità media di propagazione nei tessuti, f_D è la frequenza doppler e θ rappresenta l'angolo di insonazione della sonda, la miglior risoluzione è tra 40 e 60. Le informazioni fornite dal segnale doppler caratterizzano il regime di velocità all'interno del vaso, una stima della distribuzione e il numero degli eritrociti, il regime di moto del flusso ematico e il suo profilo di velocità. Esistono diverse tipologie di immagini doppler. In questa tesi ci si è focalizzati su due tipologie: le immagini color Doppler e power Doppler, in Fig.6. In un'immagine Color Doppler ci si serve di una mappa di colore sovrapposta all'immagine B-mode che rappresenta il flusso sanguigno. Il colore è codificato in maniera tale che il rosso è assegnato ad un flusso discorde con la direzione di propagazione della radiazione mentre il blu ad un flusso concorde con la direzione di propagazione della radiazione sonora. La luminosità

del colore indica il modulo della velocità. Questa modalità non è in grado di dare informazioni quantitative ma fornisce esclusivamente informazioni qualitative poiché il flusso è rappresentato come media delle velocità. La modalità Power Doppler invece fornisce informazioni quantitative: calcolando l'integrale dell'energia è possibile visualizzare con una mappa di colori l'intensità del segnale Doppler senza tener conto della direzione in cui lo spostamento avviene, si pone l'accento sul numero degli eritrociti. Si assegna una scala di colore che va dal giallo al rosso magenta. Questa modalità è utilizzata per analizzare flussi sanguigni tenui con velocità ridotte o che con il Color Doppler risultano difficili da visualizzare.

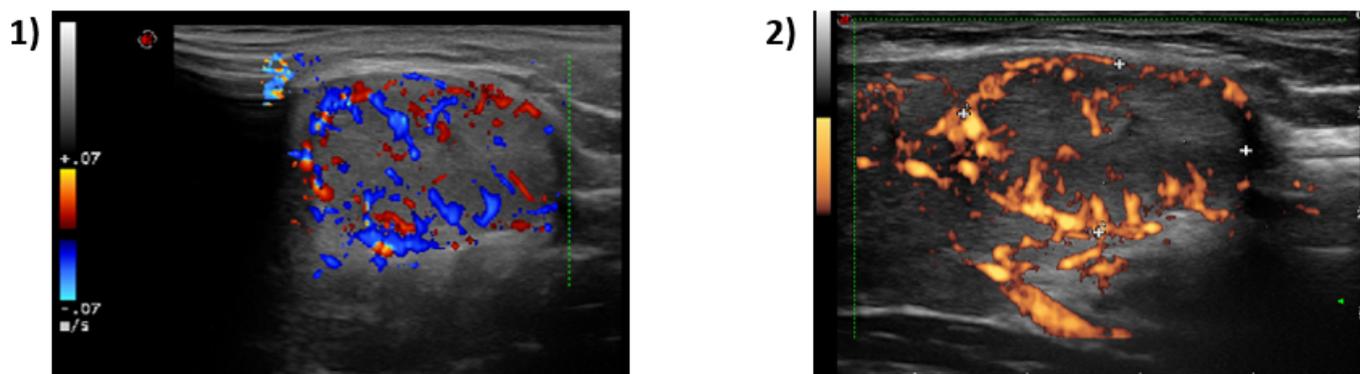


Figura 6: Immagine ecografica di lesione tiroidea (1)Color Doppler (2)Power Doppler

1.3.3 Immagini Elastosonografiche

L'elastosonografia rappresenta un esame complementare e di supporto all'ecografia. L'indagine elastografica evidenzia le proprietà elastiche del materiale, andando ad arricchire l'informazione pervenuta attraverso l'ecografia tradizionale. L'elastosonografia è una pratica innovativa che mira a prendere il posto dell'esame di palpazione, cercando di eliminare così la soggettività dell'esame stesso dovuta all'operatore, in altre parole cerca di minimizzare l'errore dovuto ad interpretazioni umane. Tuttavia grazie a questa tecnica si possono interrogare tessuti profondi o tessuti non accessibili tramite palpazione. La metodologia alla base dell'eccitazione dei tessuti è la medesima che si utilizza nell'ecografia, infatti le sonde elastosonografiche sono facilmente integrabili con le sonde ecografiche tradizionali. Lo scopo dell'elastografia è la stima del modulo di Young, un parametro fisico che definisce la rigidità di un materiale. La misura di questo parametro è tradotta in una mappa a colore sovrapposta all'immagine ecografica. La traduzione del colore, il quale va dal rosso al blu in maniera graduale, è associata alla composizione del tessuto, dall'elastico al rigido rispettivamente, come mostrato in Fig.7:

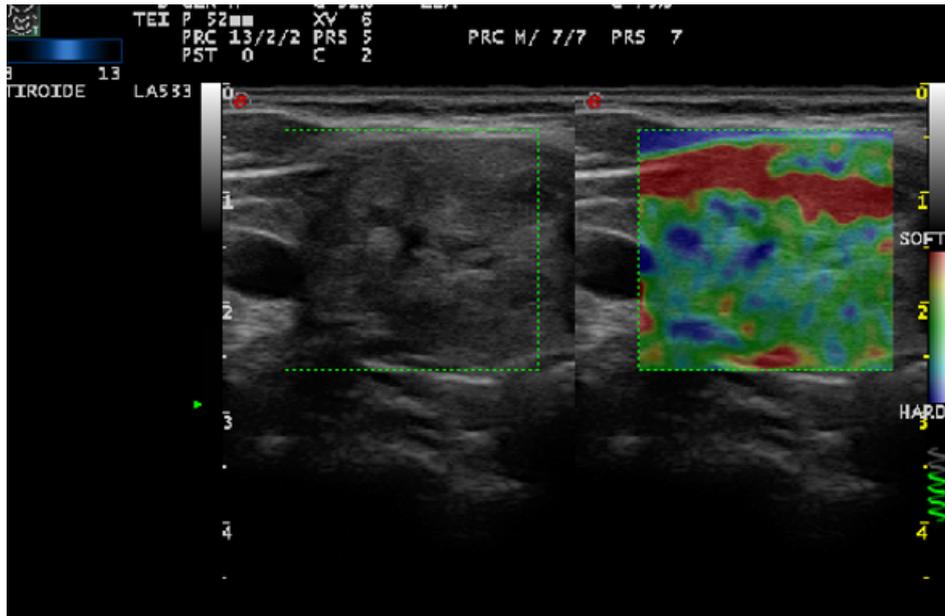


Figura 7: Immagine elastosonografica di lesione tiroidea

Con questa particolare tecnica sono possibili due tipologie di indagine [28]:

1. Indagine qualitativa. Basata sull'analisi delle deformazioni a partire dai dati relativi allo spostamento delle sorgenti degli echi. Il nome della metodologia deriva dalla tipologia di sorgente della deformazione impiegata: se la sorgente è la compressione della sonda in maniera manuale si parla di elastosonografia quasi-statica, altrimenti se la sorgente è un fascio ultrasonico si tratta di elastosonografia dinamica. Tuttavia d'interesse non è la sorgente, ma gli spostamenti longitudinali i quali sono misurati attraverso ultrasuoni. Questo è di fondamentale importanza poichè le onde trasversali in tessuti molli sono completamente attenuate. La deformazione è quindi data, in prima approssimazione, da:

$$\epsilon = \frac{\Delta L}{L_0} = \frac{L-L_0}{L_0}$$

dove ΔL è la variazione di profondità dello scatter tra due rilevazioni e L_0 la profondità iniziale. La relazione fondamentale utilizzata in tale tipo di analisi è la legge di Hooke per sforzi longitudinali:

$$\sigma = E\epsilon$$

in cui σ rappresenta la tensione applicata, E rappresenta il modulo di Young ed ϵ la deformazione misurata. Per ricavare in maniera analitica il modulo di Young, che rappresenta la grandezza d'interesse, vi è necessità di conoscere l'entità dello sforzo applicato. Si ricorre perciò ad una misura relativa, permettendo di conoscere in termini di differente rigidità zone continue ma senza quantificarne la reale misura. Per giungere a conclusione analitica alla sforzo (σ) di solito si assegna un valore comodo (ad esempio 1). La ricerca in questa metodologia mira a convergere gli sforzi in un metodo di stima dello sforzo e quindi per diretta conseguenza del modulo di Young.

2. Indagine quantitativa. L'indagine si basa su una stima della propagazione delle onde trasversali nel mezzo. La bassa frequenza ($< 200\text{Hz}$) che contraddistingue queste onde permette di controllarne l'attenuazione e quindi di seguirle durante la loro trasmissione nei tessuti d'interesse. Le possibilità di generazione delle onde trasversali è data da un dispositivo esterno, come nelle tecniche di elastosonografia transiente, oppure possono derivare dal fenomeno della forza di radiazione acustica. L'indagine quantitativa è definita tale poichè si risale al modulo di taglio, quindi al modulo di Young, conoscendo esclusivamente la velocità di propagazione delle onde trasversali. In particolare si sfrutta la relazione:

$$G = \frac{E}{2(1-\eta)}$$

in cui η rappresenta il modulo di poisson. Infatti, la velocità delle onde trasversali è data da:

$$V_{trsv} = \sqrt{\frac{G}{\rho}}$$

in cui ρ rappresenta la densità del mezzo. In base agli spostamenti delle sorgenti ecogene, si cerca di derivare una stima della misura di velocità di propagazione. Il limite principale di questa indagine è rappresentato proprio da questa stima.

1.3.4 Struttura di un Ecografo

Le apparecchiature ecografiche moderne sono costituite da tre elementi fondamentali: la sonda o trasduttore, il sistema centrale di elaborazione dati/immagini e il monitor, come mostrato in Fig.8.



Figura 8: Struttura completa di ecografo portatile provvista di monitor e sonde (tratto da [13])

L'elemento sensibile ed a diretto contatto con il paziente è la sonda. La sonda gioca anche un ruolo di ricezione, ma è impiegata anche per la creazione, del fascio ultrasonico. La generazione delle onde vede l'impiego di corrente di rete (corrente alternata 50Hz) per creare ultrasuoni tramite effetto piezoelettrico. Gli ultrasuoni vengono generati o in modo continuo (emissione continua) oppure introducendo delle pause tra una generazione e l'altra (emissione pulsata) [23]. Queste due modalità di emissione hanno scopi differenti ed vengono anche prodotti da sonde differenti. La modalità continua prevede l'impiego di una sonda con un singolo cristallo piezoelettrico e il suo utilizzo è spesso riservato all'oftalmologia. In opposizione l'emissione pulsata, con la generazione di impulsi successivi sfruttando un insieme di array di cristalli piezoelettrici, vede molteplici applicazioni ed è introdotta ormai in quasi tutte le sonde. I componenti essenziali della sonda sono [28]:

- Cristalli piezoelettrici. Sono gli elementi "senzienti", hanno la naturale particolarità di tradurre un segnale elettrico in onda ultrasonica e viceversa. Tale proprietà è la piezoelettricità. Quando si investe un cristallo con un'onda pressoria esso produce una differenza di potenziale tra le sue due facce (fenomeno piezoelettrico diretto) e allo stesso modo se si applica una differenza di potenziale a un cristallo esso produce

un'onda pressione (fenomeno piezoelettrico inverso).

- Strato di smorzamento. Risulta basilare l'impiego di un generatore di impulsi elettrici alla frequenza desiderata al fine di garantire la possibilità di emissione pulsata dei cristalli piezoelettrici, i quali devono emettere impulsi di breve durata. La durata dell'impulso elettrico non può essere troppo breve altrimenti si rischia che il cristallo non risponda all'impulso con conseguente mancata generazione del fascio ultrasonico e quindi perdita di risoluzione dell'ecografo. Lo strato di smorzamento ha la funzione di smorzare rapidamente le vibrazioni del cristallo eccitato. Il cristallo è eccitato da impulsi più lunghi del normale per evitare che qualche cristallo non risponda all'eccitazione e lo strato di smorzamento pone un fine alla vibrazione presente evitando una sovrapproduzione di ultrasuoni con conseguente perdita di risoluzione.
- Strato di accoppiamento. A causa dell'elevata differenza di impedenza acustica tra cute e trasduttore si incorre nel problema di non riuscire a raggiungere la profondità desiderata per la riflessione degli ultrasuoni. Lo strato di accoppiamento ha lo scopo di intervenire per evitare il problema: esso ricopre il ruolo di adattatore, cioè è caratterizzato da una ampiezza intermedia dell'interfaccia cute-trasduttore mitigando così l'effetto di riflessione.

Il sistema centrale di elaborazione comprende due elementi distinti: la sezione dedicata alla trasmissione/ricezione degli ultrasuoni e la sezione dedicata all'elaborazione del segnale. In particolare la sezione dedicata alla trasmissione/ricezione degli ultrasuoni funge da formazione del fascio ultrasonico. Per generare ultrasuoni pulsati il sistema gestisce i ritardi che eccitano i cristalli in maniera regolare. La temporizzazione e la sincronizzazione sono coordinate tramite clock il quale regola tutta la dinamica dei ritardi. La sezione dedicata all'elaborazione invece prevede degli step intermedi: preamplificazione, compensazione del guadagno, demodulazione, conversione A/D, memorizzazione e conversione in segnale video [28]. Successivamente alla fase di elaborazione è possibile rappresentare sul monitor l'esito dell'elaborazione.

1.3.5 Tecniche di Esame

Le tecniche di esame sono utili per analizzare le caratteristiche della tiroide e valutarne diversi aspetti ecografici. Tutte le strutture anatomiche del collo dovrebbero essere in primo luogo identificate tramite palpazione, per verificarne l'anatomia superficiale corretta, mentre in seguito si dovrebbe procedere a una valutazione più dettagliata delle strutture profonde sfruttando un'analisi ecografica. La scelta della sonda deve ricadere su sonde che sfruttano frequenze alte perciò si impiegano sonde lineari ad alta frequenza (7-13 MHz). Al fine di ottenere una risoluzione spaziale ottimale è fondamentale posizionare correttamente i fuochi del fascio sonoro alla profondità di scansione desiderata (focalizzazione dinamica). L'esame non richiede nessuna preparazione preliminare, l'unico fattore da tener presente per l'operatore è condurre l'esame con il paziente in posizione supina e con il collo in iperestensione per esporre la tiroide e facilitarne l'indagine. E' possibile indagare la tiroide con due direzioni di scansione differenti, in Fig.9, che mettono in risalto solo alcuni aspetti ecografici. Le scansioni trasversali sono condotte per piani successivi, partendo dall'apice e scendendo fino alla fine della ghiandola. L'utilità di queste scansioni è la valutazione della simmetria dei lobi tiroidei, stabilire l'orientamento anatomico generale della regione cervicale e stimare i rapporti con le altre strutture anatomiche. Le scansioni longitudinali servono per studiare l'omogeneità della struttura del parenchima tiroideo e

per una maggiore ricerca di piccole patologie. Le scansioni longitudinali servono per valutare caratteristiche ecografiche, quali la forma, in maniera da evitare delle incomprensioni per la valutazione finale. Esistono anche delle scansioni oblique che aiutano la diagnosi delle altre due tipologie di indagine.

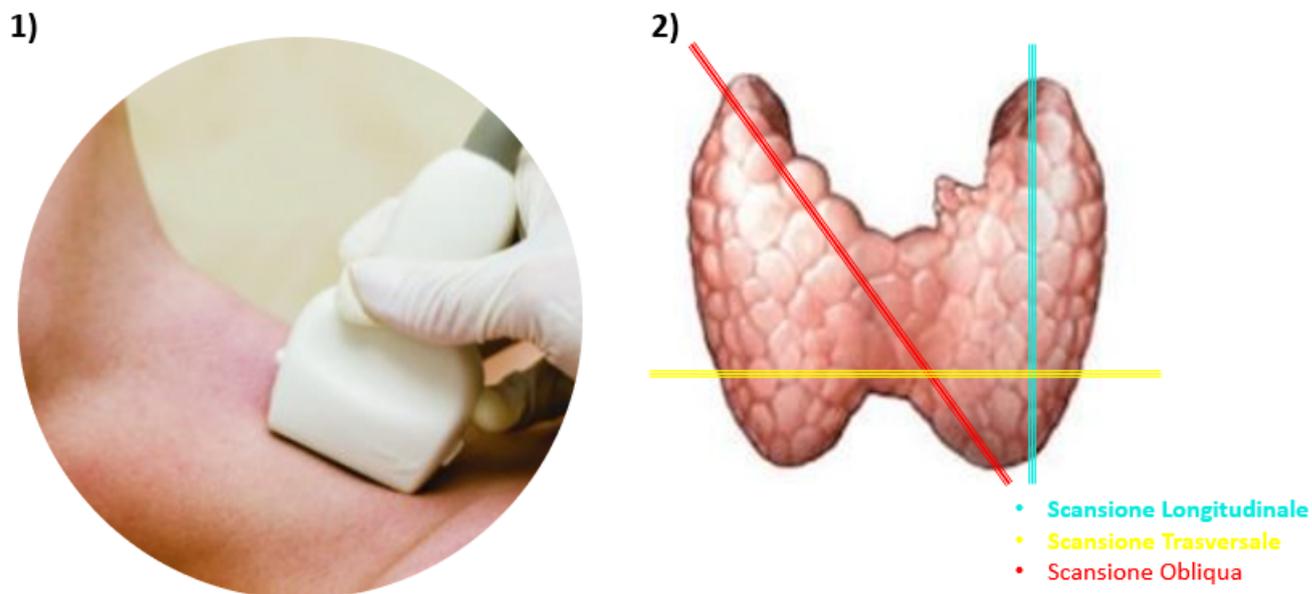


Figura 9: (1)Tecnica di esame con collo in iperestensione(2)Direzioni delle scansioni per acquisizione immagini (tratto da [24])

L'esame ecografico tiroideo ricopre un ruolo principale nella diagnosi. Tuttavia il suo valore diagnostico non è da intendersi come veritiero, poiché potrebbe portare ad errori nel conoscere la stadiazione di un nodulo. Quindi in seguito a un esame ecografico si ha un'idea iniziale, con una diagnosi "primaria", del grado di malignità, e grazie a questa grossolana quantificazione che il medico riesce a creare un programma specifico di trattamento del paziente. Le tecniche di indagine ecografiche oggi sono utilizzate come strumento diagnostico, poiché si raggiunge un livello di accuratezza elevato, ma esistono casi in cui non si riesce a definire la stadiazione di un nodulo quindi l'esame non è ancora una valida sostituzione all'esame citologico. Ciò nonostante le tecniche con cui si interroga un nodulo tiroide sono un supporto affidabile e vantaggioso ma comportano alcune limitazioni.

1.4 Caratteristiche ultrasonografiche

Le caratteristiche ultrasonografiche provenienti da immagini ad ultrasuoni sono molteplici. In questa tesi sono trattate solo alcune delle caratteristiche, in particolare:

1. Ecogenicità
2. Struttura
3. Margini
4. Presenza di cisti
5. Forma
6. Vascolarizzazione
7. Elasticità

1.4.1 Ecogenicità

L'ecogenicità è definita come la capacità di una struttura organica di riflettere ultrasuoni. Un tessuto è definito ipoecogeno se si presenta più scuro, mentre iperecogeno se si mostra più chiaro[18]. La riflettività di un tessuto, tuttavia dipende anche da parametri di scansione come il guadagno, la frequenza e la compressione. Perciò è utile confrontare più immagini ottenute variando le impostazioni di scansione. Una lesione tiroidea può essere classificata come iperecogena, isoecogena, ipoecogena se paragonata ai tessuti circostanti, come mostrato in Fig.10.

I noduli ecoprivi o anecogeni, che riflettono completamente il fascio ultrasonoro, sono indice di benignità in quanto considerabili come cisti. Un nodulo è considerato cistico se la componente anecogena costituisce almeno il 50% del suo volume. L'ecogenicità del nodulo è valutata per confronto rispetto a quella del parenchima tiroideo ma anche rispetto a quella del muscolo: si parla di lesione isoecogena se ha intensità pari al tessuto circostante, ipoecogena se più scura, iperecogena se è più chiara. Il livello di ecogenicità specifico del nodulo è associato ad entrambe le lesioni maligna e benigna. Noduli altamente ipoecogeni hanno una bassa sensibilità ma alta specificità, solitamente l'ipoecogenicità è associata a un nodulo maligno, in contrapposizione con l'iperecogenicità associata ad un nodulo benigno[19]. Se un nodulo ha natura mista può essere descritto come predominantemente ipo-,iso-, iper-ecogeno.

1.4.2 Struttura

Per struttura, o ecostruttura, di un nodulo si intende il grado di omogeneità della lesione[18]. Questa caratteristica è stata interpretata come un indice di dispersione ovvero una stima della variabilità dell'insieme di pixel di cui è composto il nodulo. Una lesione può presentare una struttura omogena o disomogena, un esempio in Fig.11

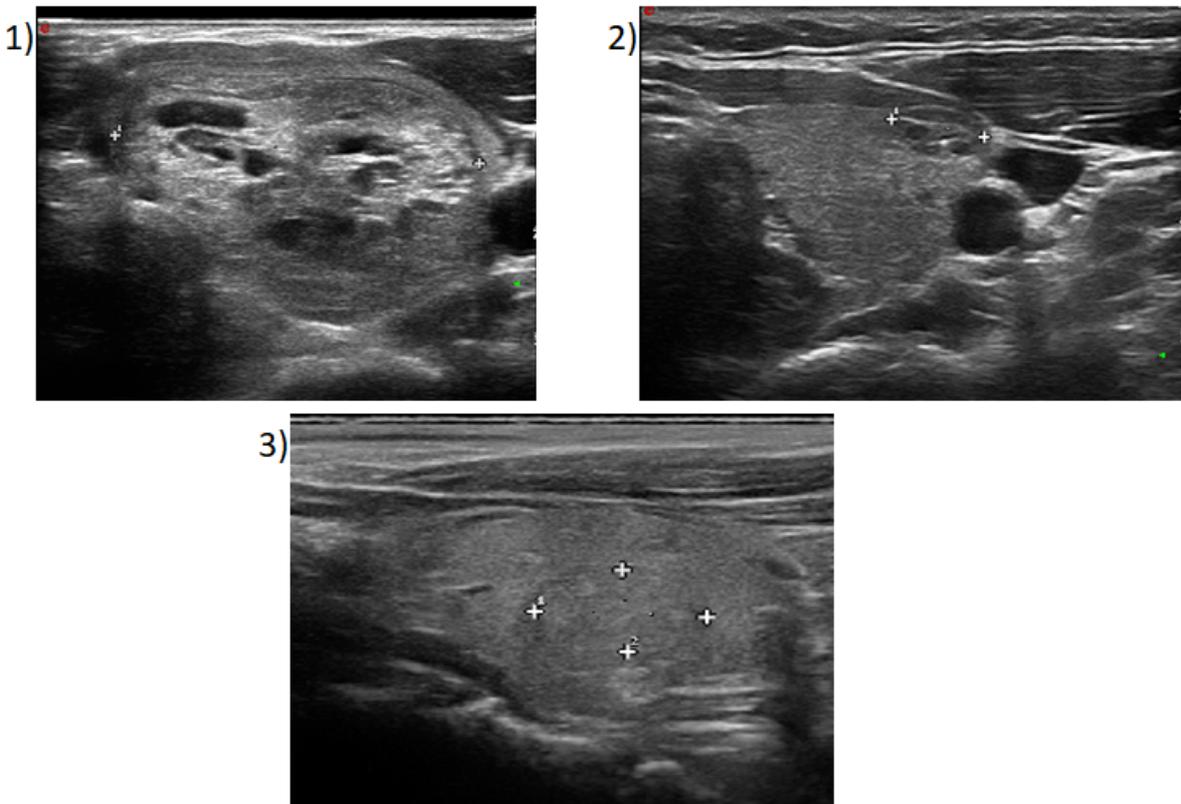


Figura 10: Immagini ultrasonografiche B-mode che mostrano (1)Nodulo iperecogeno(2)Nodulo ipoecogeno(3)Nodulo isoecogeno

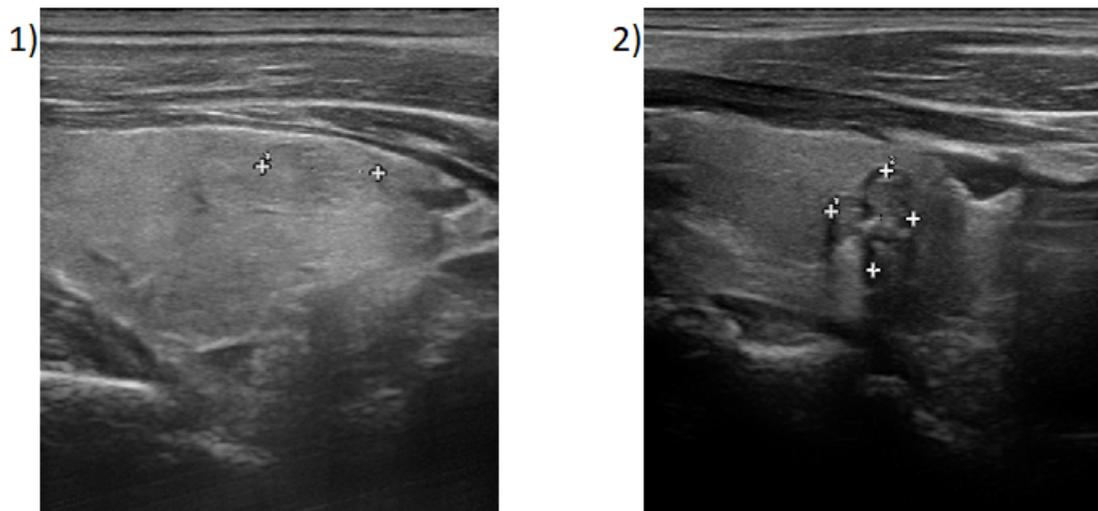


Figura 11: Immagini ultrasonografiche B-mode che mostrano (1)Nodulo con struttura omogenea (2)Nodulo con struttura disomogenea

In altre parole per struttura omogenea si intende una variabilità delle intensità dei pixel che compongono il nodulo molto ridotta, invece per disomogenea si intende una variabilità alta. Una struttura omogenea è associata ad un grado di benignità maggiore rispetto ad una non omogenea[19].

1.4.3 Margini

Per margine di una lesione si intende l'interfaccia con il tessuto adiacente. Un margine regolare ("smooth") presenta un'interfaccia uniforme e graduale[18]. In presenza di irregolarità o di intrusione della componente solida nel tessuto circostante, si classifica come irregolare, un esempio presentato in Fig12.

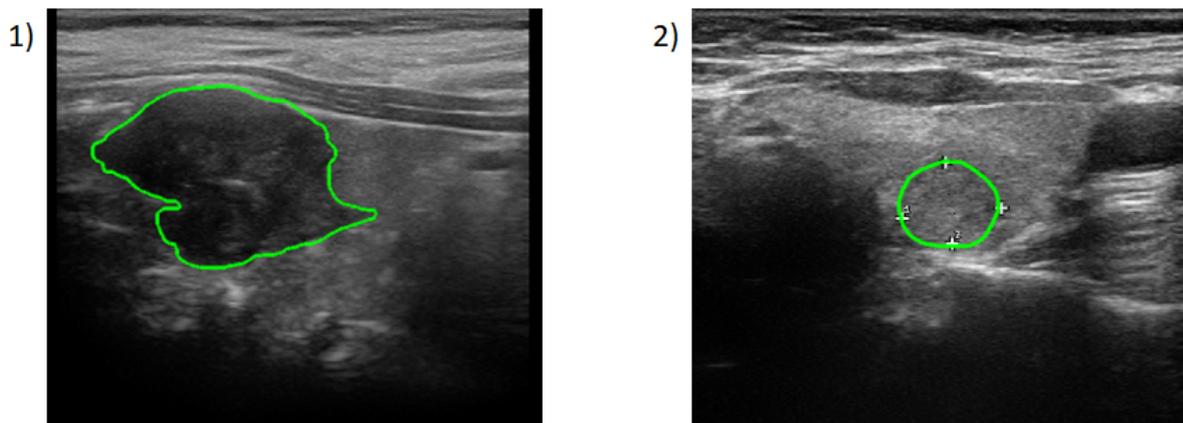


Figura 12: Immagini ultrasonografiche B-mode che mostrano (1)Margini irregolari(2)Margini regolari

Un margine regolare è più comune in noduli benigni, ma tra il 33% e il 93% di noduli maligni presentano margine regolare. Margini irregolari e lobulati sono caratteristici di malignità e sono altamente specifici. Questi bordi sono considerati rappresentativi di un

pattern di crescita aggressivo. Nonostante ciò, i margini definiti malati hanno mostrato di essere statisticamente significativi associati alla malignità e non è comune trovarli in noduli benigni. L'estensione della capsula di un nodulo nei tessuti adiacenti indica malignità invasiva. L'estensione oltre il confine tiroideo è classificata come estesa o minima. Un'ampia estensione extra tiroidea (ETE) caratterizzata dall'invasione schietta dei tessuti molli adiacenti e/o delle strutture vascolari è un segno altamente affidabile di malignità, un segno prognostico sfavorevole. Un'estensione extra tiroidea minima può essere sospetta ecograficamente parlando in presenza di abutment del bordo, rigonfiamento del contorno o perdita del bordo tiroideo ecogenico[19]. La presenza di un alone non è né discriminativa né si escludono a vicenda con altri tipi di margine; pertanto, si è deciso di ometterlo e non trattarlo nella tesi.

1.4.4 Presenza di Cisti

Una cisti è una cavità o sacca, normale o patologica, chiusa da una membrana ben distinta, contenente un materiale liquido o semisolido[18]. La presenza di cisti delinea la composizione interna di un nodulo; un nodulo è considerato cistico quando più del 50% del suo volume è costituito da una componente cistica, come mostrato in Fig.13

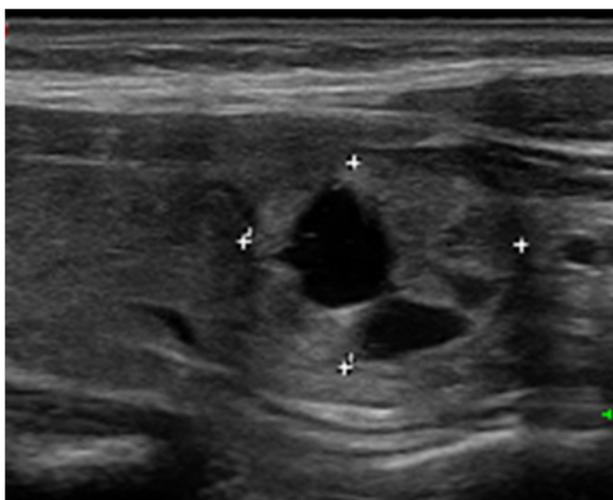


Figura 13: Immagine ultrasonografica B-mode con presenza di cisti caratterizzata evidenziata dal colore nero

La componente cistica si evidenzia sull'immagine data l'anecogenicità del materiale liquido, ovvero una cisti è caratterizzata dal colore nero. Ad un nodulo cistico è quasi universalmente associata una condizione benigna. La presenza di medie e micro cisti non assolve il nodulo da una condizione di malignità[19].

1.4.5 Forma

Si definisce forma di un nodulo la proporzione tra la sua altezza e la sua larghezza[18]. L'altezza è la misura della dimensione anteroposteriore mentre la larghezza si riferisce alla sua dimensione trasversale. Si preferisce rilevare le grandezze su immagini ottenute per scansione longitudinale. Una lesione può essere identificata sostanzialmente in 3 forme: "wider-than-tall" (più largo che alto), "taller-than-wide" (più alto che largo) e sferica, come esposto in Fig.14.

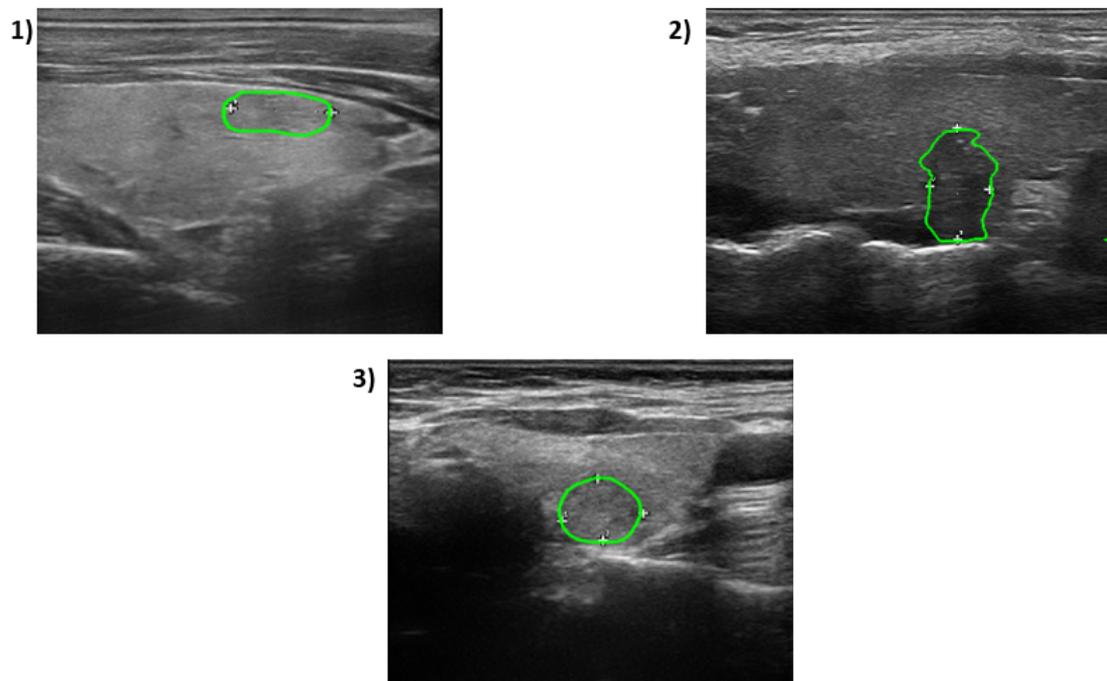


Figura 14: Immagini ultrasonografiche B-mode che mostrano forma (1)Wider than tall(2)Taller then wide(3)sferica

Una forma taller-than-wide è un indicatore altamente specifico di malignità. Invece forme sferiche e wider-than-tall sono associati alla natura benigna del nodulo[19].

1.4.6 Vascolarizzazione

Si definisce vascolarizzazione la distribuzione dei vasi ematici di un organo o di un apparato. Un tessuto è vascolarizzato quando al suo interno è provvisto di vasi sanguigni[18]. La vascolarizzazione indica anche l'entità dell'apporto sanguigno, che può variare notevolmente per effetto di meccanismi regolatori fisiologici, o per cause patologiche. La vascolarizzazione di un nodulo è indagata tramite immagini doppler, mostrata in Fig.6. La vascolarizzazione è divisa in base alla presenza di segnale doppler in aree predefinite del nodulo: area periferica e area intranodulare. L'assenza di segnale doppler è indice di benignità. Tuttavia la presenza pattern vascolari aumenta il sospetto di malignità in base alla quantità di segnale (Fig.15):

- Pattern perinodulare senza flusso intranodulare
- Pattern perinodulare e intranodulare con vasi sottili e regolari
- Pattern perinodulare e intranodulare con vasi spessi e irregolari

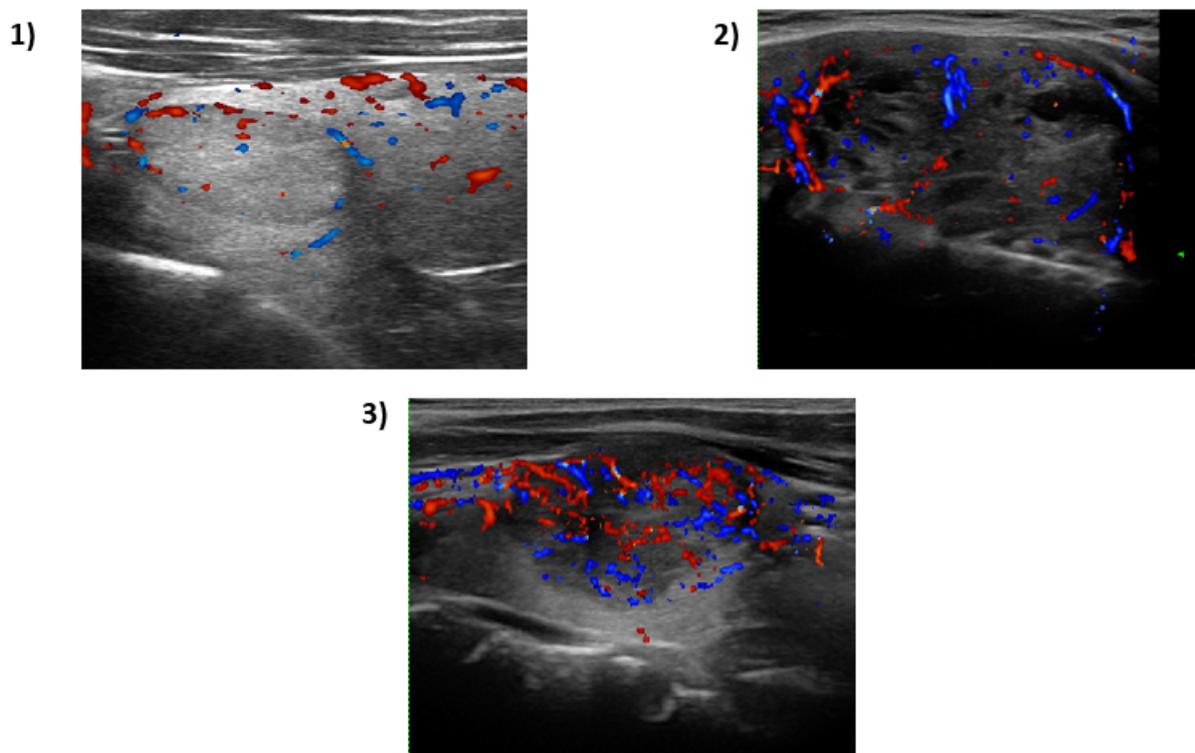


Figura 15: Immagini doppler che mostrano (1) Pattern perinodulare senza flusso intranodulare(2) Pattern perinodulare e intranodulare con vasi sottili e regolari(3)Pattern perinodulare e intranodulare con vasi spessi e irregolari

La vascolarizzazione è associata al metabolismo del nodulo elevato, perciò più il nodulo è attivo maggiore sarà il suo grado di sviluppo [19]. Tuttavia quando si è in presenza di noduli altamente maligni si può incorrere in assenza di vascolarizzazione poichè il tessuto indagato è un tessuto necrotico. Un nodulo altamente vascolarizzato è un nodulo con attività di crescita importante.

1.4.7 Elasticità

L'elastosonografia è un sistema di misurazione non invasivo della rigidezza di un tessuto. In particolare si valuta l'elasticità di parti di tessuti molli o duri paragonandola con l'elasticità delle aree circostanti. Un tessuto è definito elastico se si mostra di un colore rosso, rigido se si mostra di un colore blu o intermedio se si mostra di colori intermedi. Tuttavia la codifica colori è accompagnata da un color bar che ne definisce la corretta associazione mostrata in Fig.16.

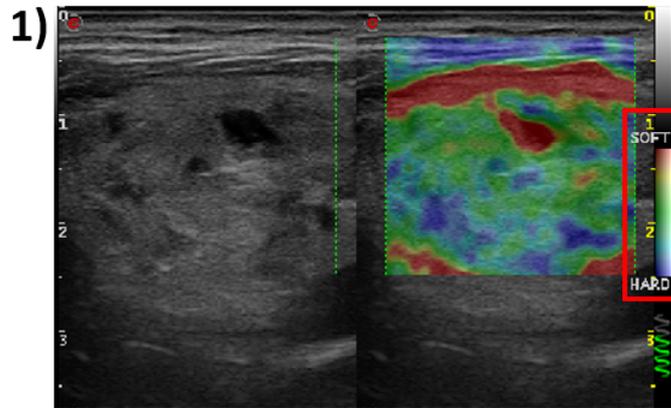


Figura 16: (1) Immagine elastosonografica con colorbar nel riquadro rosso che mostra la composizione elastica di un nodulo

Un nodulo iperelastico (soft o elastico) o isoelastico (intermedio) è spesso descrittore di un nodulo benigno, in questa caso si va dal rosso al verde. Il sospetto di malignità aumenta quando si indaga un nodulo ipoelastico (rigido), cioè di colore verde scuro, fino ad arrivare alla definizione di un nodulo duro di colore blu [19].

1.5 Classi TIRADS

In presenza di un carcinoma, le possibilità di guarigione aumentano se esso è scoperto precocemente. Al fine di accrescere queste possibilità una corretta diagnosi risulta di primario rilievo. La presenza di un nodulo viene normalmente individuata con l'esame fisico, riscontrando la disomogeneità della sua massa alla palpazione, ma per i noduli di minori dimensioni o di localizzazione interna la palpazione non è sufficiente. I noduli iperplastici e di grandi dimensioni possono causare le problematiche descritte precedentemente, anche se la maggior parte risultano asintomatici e vengono scoperti in maniera totalmente casuale. L'esame primario e meno specifico è l'ecografia, in quanto consente di valutare alcune caratteristiche ecografiche del nodulo, della ghiandola tiroidea e dei tessuti circostanti, che consentono di distinguere, solo in primo luogo, una lesione benigna da una maligna. L'esame ultrasonografico, tuttavia, fornisce soltanto informazioni morfologiche e strutturali, non funzionali. Se dall'esame ecografico non emerge la natura del nodulo, ci si avvale di esami più specifici. Recentemente l'European Thyroid Association (ETA) ha emanato delle linee guida per la descrizione e la standardizzazione ecografica del rischio di malignità del nodulo tiroideo basandosi sul modello del TIRADS (Thyroid Imaging and Reporting Data System) che prevede le seguenti categorie [37]:

- TIRADS 1: assenza di noduli tiroidei
- TIRADS 2: noduli benigni (rischio di malignità 0%); l'agoaspirato non è indicato
- TIRADS 3: noduli a basso rischio (rischio di malignità 2-4%); l'agoaspirato è consigliato solo per noduli maggiori di 20 mm.
- TIRADS 4: noduli a rischio intermedio (rischio di malignità 6-17%), l'agoaspirato è consigliato per noduli maggiori di 15 mm.
- TIRADS 5: noduli a rischio elevato (rischio di malignità 26-87%); l'agoaspirato è consigliato per noduli maggiori di 10 mm

Le caratteristiche ultrasonografiche TI-RADS sono classificate come benigne, minimamente sospette, moderatamente sospette, altamente sospette per malignità. I punti sono attribuiti per ogni caratteristica del nodulo, con caratteristiche più sospette vengono assegnate punti aggiuntivi. Nel valutare un nodulo, si seleziona una caratteristica da ciascuna categoria e si sommano tutti i punti attribuiti a ogni caratteristica. Il punto totale determina il livello TI-RADS del nodulo, che varia da TR1 (benigno) a TR5 (alto sospetto di malignità). Un esempio di come avviene l'attribuzione dei punti TIRADS è riportata in Fig.17.

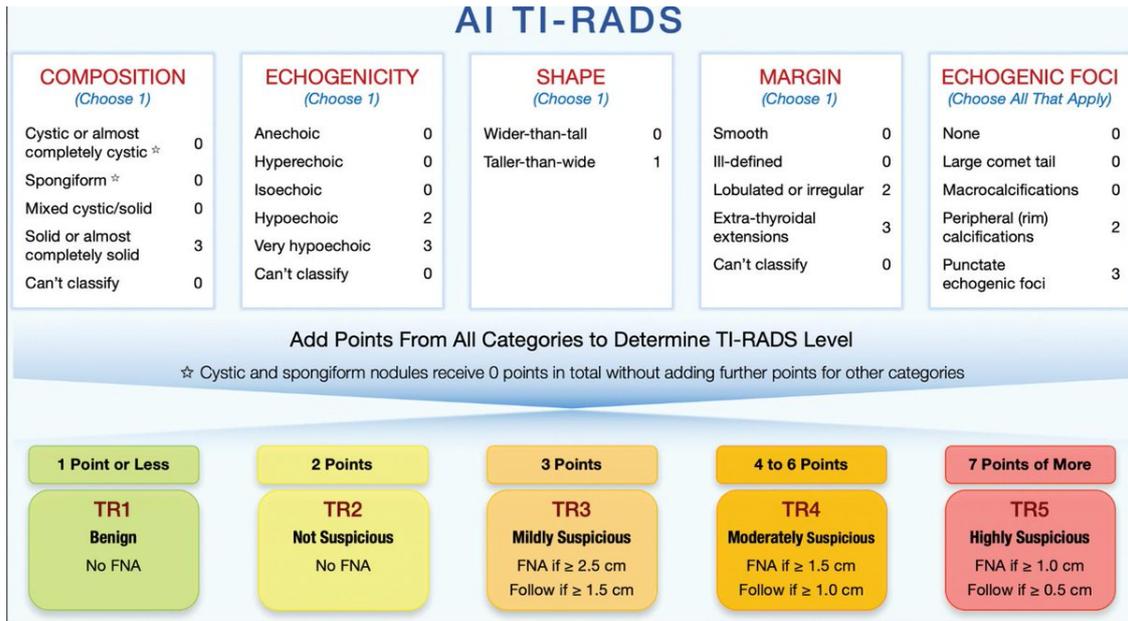


Figura 17: Esempio di attribuzione punteggio TI-RADS, basato sulle caratteristiche ecografiche (tratto da [16])

Nel TI-RADS, raccomandazioni per FNA o il follow-up degli ultrasuoni si basa sul livello TIRADS di un nodulo e sul suo diametro massimo. Per i livelli di rischio da TR3 a TR5, la Fig.67 presenta una soglia dimensionale sopra la quale dovrebbe essere raccomandato un FNA. Il TIRADS è progettato per bilanciare il vantaggio di identificare tumori clinicamente importanti contro il rischio e costo di sottoporre pazienti con noduli benigni o tumori indolenti alla biopsia e al trattamento. Le raccomandazioni provenienti dagli ultrasuoni di follow-up mitigano sostanzialmente la possibilità che significative neoplasie maligne rimangano inosservate nel tempo e sono in accordo con la crescente tendenza verso la sorveglianza attiva (attesa vigile) per il carcinoma tiroideo a basso rischio.

1.6 Iter Diagnostico

Il corretto iter diagnostico di un nodulo tiroideo è stato codificato nelle più recenti “Linee Guida per la gestione dei pazienti con nodulo tiroideo e carcinoma differenziato della tiroide”, prodotte dalla American Thyroid Association, e ratificato da molte altre società internazionali. La palpazione, passo iniziale dell’esame, è accompagnata dalla prescrizione di esami ematologici in grado di verificare la funzionalità della ghiandola tiroidea. Dai risultati di questi esami è successivamente scelta la metodica di imaging più appropriata per differenziare il nodulo sospetto in senso di benignità o malignità. La visualizzazione ecografica permette di confermare la presenza di un nodulo della ghiandola, identificando alcune caratteristiche del nodulo e localizzando correttamente la lesione. Le caratteristiche ecografiche quali l’ipoecogenicità del nodulo rispetto al parenchima circostante, la vascolarizzazione intranodulare, i margini irregolari, la presenza di microcalcificazioni, l’assenza di aloni periferico e la forma più alta che larga nella scansione trasversale sono associati a una più elevata probabilità di malignità. Inoltre, ecograficamente è possibile identificare alcuni elementi che possono distinguere la variante papillifera da quella follicolare: la prima si presenta più spesso ipoecogena e solida, spesso con margini irregolari e aumento della vascolarizzazione. Al contrario, il carcinoma follicolare è più spesso iso-iper ecogeno, con un alone spesso e irregolare, senza microcalcificazioni, e quando inferiore ai 2 cm non è mai associato a metastasi. Come ulteriore elemento di valutazione, oltre al color-doppler che identifica i pattern di vascolarizzazione, è impiegata l’elastosonografia nella diagnosi differenziale per valutare l’elasticità del tessuto. Per rafforzare la diagnosi dell’esame ultrasonografico ci si avvale anche di altre metodologie di esame per valutare le caratteristiche funzionali e anatomiche oltre che l’attività della ghiandola. La scintigrafia, in Fig.18, tiroidea è un esame che si esegue in particolari casi che richiedono un livello diagnostico approfondito. Questa tecnica di imaging si utilizza per valutare soprattutto la funzionalità, ma anche la forma e il corretto posizionamento della ghiandola e prevede la somministrazione di radiofarmaco per via orale o endovenosa. Sono impiegati solitamente tre tipologie di radiotracciante (Iodio131, Iodo123 e Tecnezio99) scelti in base alle peculiarità del paziente [30].

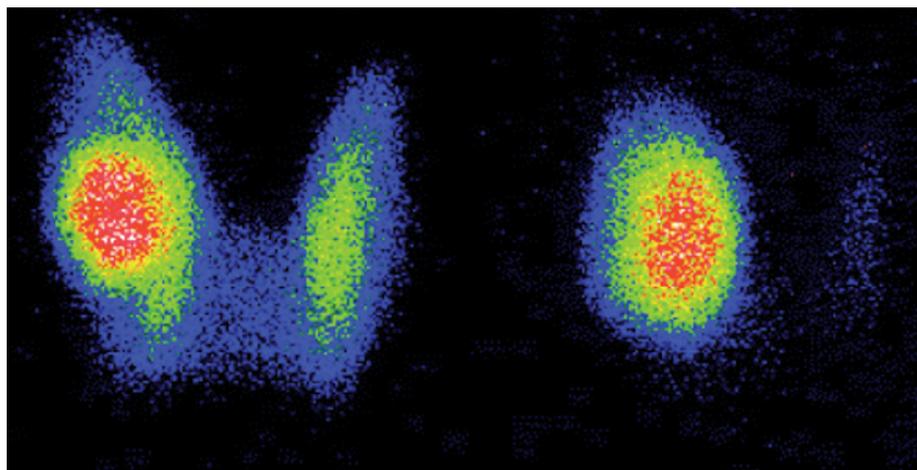


Figura 18: Imaging con Scintigrafia che mostra la funzionalità elevata del nodulo tiroideo (tratto da [35])

In presenza di un nodulo sospetto, individuato tramite imaging, ma che lascia dubbi sulla sua natura si procede con l’esame citologico. L’agoaspirato tiroideo(FNA) è una tecnica mini invasiva che consente solo il prelievo di alcune cellule contenute nel nodulo, di

conseguenza l'analisi del tessuto è di tipo citologico. Lo scopo principale dell'agoaspirato è distinguere i pazienti che devono essere sottoposti a trattamento chirurgico da quelli che necessitano solo di monitoraggio [3].

In contemporanea all'ecografia, le linee guida ATA prevedono anche l'esecuzione di un esame citologico con tecnica Fine-Needle-Aspiration. La FNAC eco-guidata è definita il golden standard nella diagnosi dei noduli tiroidei, permettendo una accurata identificazione delle lesioni. L'esame citologico andrebbe eseguito su tutti i noduli con diametro maggiore di 1 cm, venendo sconsigliata nei casi di lesioni sub-centimetriche. Tradizionalmente, i risultati della FNAC sono divisi nelle 5 categorie descritte precedentemente in Fig.4.

1.7 Stato dell'Arte

Oggi giorno la quantificazione delle caratteristiche ultrasonografiche dei noduli tiroidei deriva dall'acquisizione di immagini ecografiche e quindi dalla lettura delle stesse da parte di un esperto. L'esperto analizza il nodulo e valuta delle caratteristiche ultrasonografiche, in base alla sua esperienza. In particolare la grandezza quantitativa d'interesse può essere predittiva di malignità o meno, e solo il referto citologico fornisce successivamente la verità. Per comprendere in quale stadio si trova il nodulo e come comportarsi rispetto ad esso, l'esperto deve fornire la classificazione del nodulo basandosi solo su un'indagine visiva dell'immagine ecografica. Per implementare una metodologia di indagine innovativa, inizialmente è stata condotta una ricerca bibliografica sull'anatomia, la fisiologia e le patologie della tiroide. Da questa sono emerse diverse caratteristiche quantificabili presenti nelle immagini ad ultrasuoni. Così è stata condotta un'ulteriore ricerca bibliografica per trovare tecniche utili e innovative al fine di quantificare queste caratteristiche. Per la strategia di ricerca sono stati impiegati i seguenti termini discriminati e caratteristici: thyroid features, elasticity, quantification of the US features, malignant symptoms of thyroid nodule e suspicious thyroid nodules. Le caratteristiche riscontrate e trattate in letteratura sono risultate numerose, tuttavia alcune sono state trattate singolarmente altre sono state combinate tra di loro. Le strade per valutare la natura di un nodulo tiroideo sono essenzialmente due: una quantitativa e l'altra qualitativa. Sono due approcci differenti che si prefissano lo scopo di giungere al medesimo risultato.

L'approccio qualitativo prevede di correlare le features ultrasonografiche con un grado di malignità; ad esempio in [22], sono stati analizzati 309 pazienti, di cui solo 30 maschi, per un totale di 329 noduli tiroidei, si sono estratte 5 caratteristiche, margini, forma, struttura, ecogenicità e calcificazioni, che hanno condotto a una sensibilità del 86.5% per i noduli non follicolari e 18.2% per i follicolari e una specificità del 92.3% e 88.7% rispettivamente. In uno studio clinico si è valutato l'accuratezza diagnostica di un sistema di classificazione per la differenziazione tra noduli benigni e maligni. Ad esempio in [31] sono stati analizzati 191 pazienti, di cui 52 maligni, 82 benigni, 16 sospettosi di malignità, 23 borderline e 18 probabilmente benigni, che hanno condotto ad un'accuratezza di diagnosi ultrasonografica pari a 86%, 92%, 95%, 91% e 92% rispettivamente quando benigni, probabilmente benigni e i borderline venivano classificati come benigni. Limiti principali di questa tipologia di approccio è la valutazione delle features in maniera qualitativa. Quindi la valutazione è affetta da errori quali la variabilità intraoperatore, ma anche interoperatore, in quanto un esperto potrebbe dare due valutazioni differenti in due momenti differenti dello stesso nodulo. Nonostante l'ampia mole di materiale presente in letteratura non sono stati trovati articoli e materiale che trattino la quantificazione di tutte queste caratteristiche nella medesima maniera.

Il secondo approccio prevede di valutare una grandezza quantitativa indiretta del nodulo attuando infine una classificazione. Ad esempio in [26], in cui sono state analizzate 2450 immagini di noduli benigni e 2557 di noduli maligni, provenienti da 276 pazienti, e i valori di sensibilità, valore predittivo positivo, valore predittivo negativo e l'accuratezza del sistema di diagnosi sono stati 90.5%, 95.22%, 80.99% e 90.31% rispettivamente per i noduli maligni. Un altro approccio è stato presentato in [33], in cui sono stati indagati 40 noduli, 20 benigni e 20 maligni, in cui sono state estratte numerose features con diversi approcci, le features sono state in seguito classificate utilizzando metodi quali leave-one-out e cross-validation che hanno condotto a performance 92.42% e 94.64% rispettivamente. Un altro esempio di quantificazione indiretta è fornita da [25], in questo caso è stato proposto un

sistema CAD, che acquisisce e migliora un gruppo di caratteristiche generalmente usate dai radiologi, e comparato un con un sistemi CAD presenti nello stato dell'arte che hanno condotto a un'accuratezza di detection e diagnosi pari a 97.5% e 97.1% rispettivamente. Il limite principale che affligge questa è che le caratteristiche estratte non sono direttamente correlate con le caratteristiche che il medico valuta tuttavia forniscono una buona stabilità sulla ripetibilità dei test.

Questa tesi nasce con l'obiettivo di ovviare alle limitazioni che questi due approcci pongono ovvero valutare in maniera quantitativa delle features direttamente correlate con le caratteristiche ultrasonografiche analizzate dal medico. Nonostante l'uso di queste tecniche sia già stato adottato per aumentare l'efficienza per la detection di una singola caratteristica o l'associazione della stessa alla malignità/benignità del nodulo, non è stato rintracciato nello stato dell'arte analizzato un algoritmo semi-automatico pensato per quantificare l'effettivo stato di avanzamento del nodulo che prevedesse l'utilizzo di più caratteristiche contemporaneamente. Lo spirito di innovazione risiede nel quantificare con metodologie non presenti in letteratura e in maniera automatica la stadiazione dei noduli tiroidei.

2 Materiale e Metodi

2.1 Immagini Ecografiche

Lo studio è stato condotto su un database contenente 410 casi clinici di lesioni tiroidee, delle quali 350 benigne e 60 maligne. Sono stati utilizzati tutti i casi clinici, anche quelli in cui non era presente il referto citologico, poichè era sempre presente la classificazione del medico. Per ogni singolo caso è stata messa a disposizione almeno un'immagine ecografica; per alcuni casi sono stati messi a disposizione anche immagini doppler ed elastosonografiche. Il database iniziale comprendeva una struttura divisa in paziente, immagine ecografica, classificazione del nodulo, dimensioni del nodulo, referto citologico e già delle maschere binarie, create sulle immagini di partenza, le quali sono state usate per la segmentazione dei noduli e per la segmentazione del parenchima tiroideo. Il database è stato successivamente ampliato con la creazione di nuove maschere per la segmentazione del muscolo. Per la valutazione della vascolarizzazione e dell'elasticità, è stato messo a disposizione un nuovo database contenente i medesimi pazienti ma con ulteriori immagini specifiche. Queste due caratteristiche sono state valutate separatamente secondo altri passaggi descritti nei paragrafi successivi. E' stata messa a disposizione anche la classificazione fatta dal medico, adottata come Ground Truth. Tuttavia non è stata adottata la medesima classificazione, poiché risulta ampiamente ostico raggiungere lo stesso grado di specificità. E' stato così preferito ricorrere ad una classificazione delle caratteristiche semplificata, la quale si limita ad un grado di specificità minore. In particolare si definisce una nuova classificazione come mostrato in Fig.19 .

Classificazione reale	0a1b, 1a1b, 2b2a, 3b3h, 4d1, 5a1, 6a, 7c3b, 8c2
Abbreviazione	0a 1b, 1a 1b, 2b 2a, 3b 3h, 4d 1, 5a 1, 6a , 7c 3b, 8c 2
Classificazione abbreviata	0a,1a, 2b, 3b, 4d, 5a, 6a, 7c, 8c

Figura 19: Esempio di abbreviazione stringa

Questa nuova classificazione è stata introdotta poichè raggiungere un tale livello di specificità in maniera automatica è ampiamente arduo. Le caratteristiche estratte automaticamente descrivono solo la natura del nodulo in questione, senza delinearne una grandezza quantitativa.

2.2 Segmentazione manuale

La segmentazione manuale delle maschere, già presenti nel database, è stata utilizzata per diversi scopi: non esclusivamente per segmentare ma si sono anche estratte caratteristiche dalle maschere stesse. In questo caso il concetto di segmentazione è stato un concetto di identificazione e di unione di pixel che nell'immagine rappresentano il medesimo tessuto. Al fine di evitare che il riconoscimento del nodulo risulti errato è stata tracciata ogni maschera binaria manualmente, per il riconoscimento della forma, da un operatore esperto.

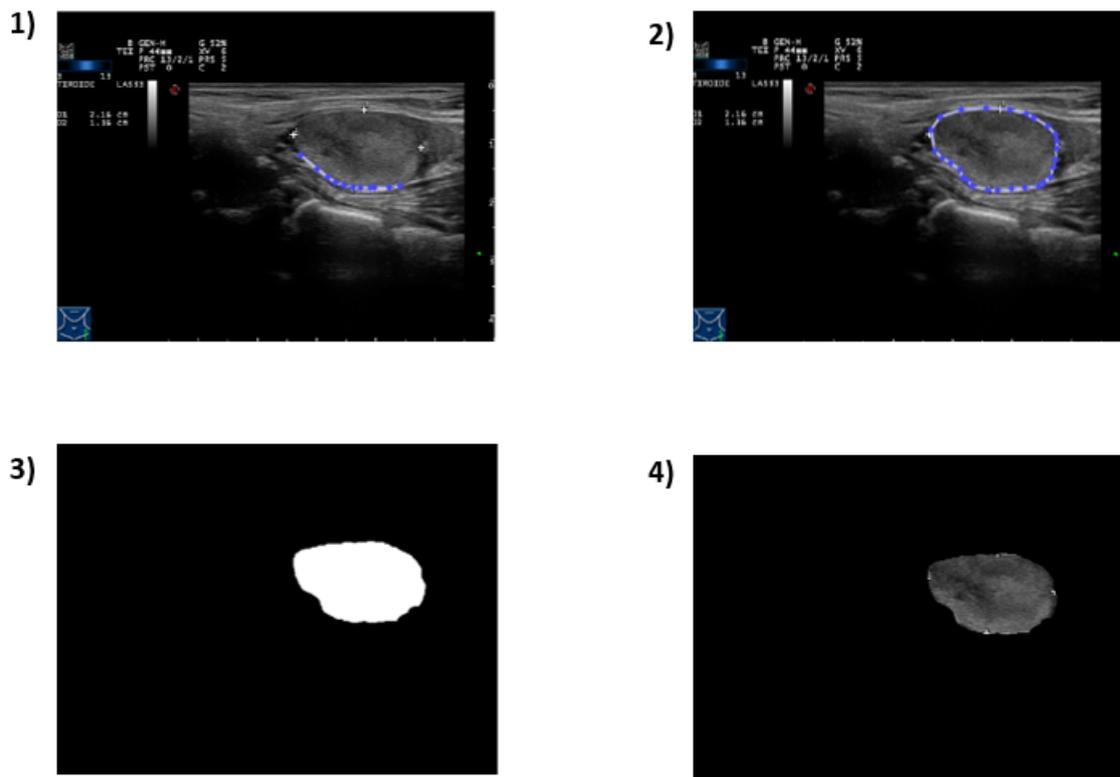


Figura 20: Esempio di come avviene la segmentazione manuale di un nodulo secondo lo schema(1)Tracciamento maschera(2)Maschera tracciata(3)Maschera binaria(4)Segmentazione nodulo tramite maschera

Nella Fig.20 sono stati riportati i passi che l'operatore ha compiuto per segmentare il nodulo. Le maschere così create sono state salvate e inserite nel database. Questa operazione è stata eseguita per ogni nodulo, identificando la porzione di tessuto d'interesse. L'operatore, oltre il nodulo, ha creato una ulteriore maschera per la segmentazione del parenchima tiroideo, seguendo gli stessi step; anche questa maschera è stata inserita nel database iniziale. Sono state eseguite altre operazioni di segmentazione manuale, nello specifico è stato segmentato il muscolo nelle immagini B-mode, e sono state tracciate le maschere per le immagini doppler ed elastosonografiche. Infine tutte le nuove segmentazioni sono state salvate ed inserite all'interno del database. L'operazione di segmentazione manuale è stata condotta anche su immagini doppler ed elastosonografiche per identificare la corretta posizione del nodulo.

2.3 Valutazione delle Caratteristiche Ultrasonografiche

Un nodulo tiroideo deve essere valutato focalizzando l'attenzione su diversi aspetti clinici quali la grandezza, la densità, la posizione anatomica e le caratteristiche ultrasonografiche. Per giungere ad una classificazione completa e specifica bisogna ottenere un quadro diagnostico complesso ma completo. Per valutare le caratteristiche ultrasonografiche bisogna tener presente quali sono i parametri da valutare e come è possibile associare la valutazione a dei parametri anatomici. I medici nel guardare, dedurre e valutare la stadiazione di un nodulo tiroideo utilizzano una misura quantitativa relativa e molto specifica, che tiene conto di più prove, ad esempio più acquisizioni ecografiche, e la sua affidabilità è legata all'esperienza nella diagnosi del medico. La valutazione delle caratteristiche ultrasonografiche è stata condotta da più esperti. Nello specifico sono state sottoposte ad ispezione di medici competenti tutte le immagini presenti nel database. Per ottenere una valutazione solida e robusta l'indagine è stata inizialmente gestita da due differenti medici, che incarnano il ruolo degli esperti, i quali sono stati chiamati a valutare le caratteristiche del nodulo. Ogni nodulo preso in esame è stato così classificato in due maniere: le due valutazioni potevano essere in accordo o in contrasto, per esempio il primo medico poteva definire il nodulo isoecogeno mentre il secondo poteva definirlo molto ipoecogeno. Giungere a una diagnosi spesso non è semplice, poichè il medico deve anche tener conto che alcune caratteristiche sono sintomi di diverse lesioni e non forniscono delle indicazioni univoche della causa. Tuttavia i medici possono ricorrere ad una diagnosi differenziale, in cui sono confrontate diverse caratteristiche per avere un quadro clinico più completo e quindi una diagnosi migliore. Per la corretta valutazione delle caratteristiche bisogna tener conto di diversi aspetti, come l'anamnesi del paziente o l'età, perciò i medici hanno bisogno di un ulteriore aiuto per la diagnosi. La valutazione finale, in questo caso, è stata regolata da un ulteriore medico, il più esperto, che poteva confermare un parere o l'altro.

CLASSIFICAZIONE:

1a1, 2b2, 3a2h,4d1,6a,7a1,8b

Nodo solido, debolmente ipoecogeno, ad ecostruttura moderatamente disomogenea, con margini regolari, con halo sign spesso completo, con alcune piccole calcificazioni senza cono d'ombra, di forma ovalare, con vascolarizzazione periferica incompleta, pattern elastosonografico intermedio

Figura 21: Esempio di valutazione specialistica del nodulo

La Fig.21 mostra come la classificazione sia accurata, specifica e descrittiva del nodulo in esame. Poichè l'indagine è stata gestita da più esperti e dato il suo il buon grado di specificità, la valutazione risultante può essere impiegata come Ground Truth ed è stato possibile basarsi su di essa per successive analisi e valutazioni. Le osservazioni dirette che hanno fornito una classificazione accurata sono di centrale importanza per valutare la qualità del lavoro condotto rapportandola ad una misura affidabile e sicura.

2.4 Principi di Statistica

L'indagine statistica si propone come obiettivo di valutare la qualità dei risultati ottenuti. L'importanza di questa indagine si evince dalla necessità di attribuire un valido significato agli esiti derivanti dalla quantificazione delle caratteristiche e di quanto essi possano essere attendibili. I test statistici si conducono indagando tutti i noduli presenti nel database. Inizialmente si tratta ogni immagine singolarmente e si ripercorrono i passi descritti precedentemente. I risultati sono mostrati graficamente nei paragrafi successivi e sono ottenuti per ogni singola feature considerando tutti i pazienti presenti nel database. Infine si è condotta un'analisi statistica per valutare l'efficacia nella distinzione delle varie classi per ogni feature. I test descritti di seguito, ad esclusione della MANOVA, non sono stati impiegati per tutte le caratteristiche ma solo per alcune per aiutare e migliorare i risultati finali. Questo paragrafo descrive le nozioni alla base dei diversi test statistici utilizzati e la loro applicazione.

2.4.1 Correlazione

La correlazione è una relazione tra due variabili tale che a ciascun valore della prima corrisponda un valore della seconda, seguendo una certa regolarità[32]. Per un'ulteriore verifica si è proceduto al calcolo della correlazione per ogni singola feature, utilizzando tutti i noduli presenti nel database. In particolare si è creato un vettore, chiamato vettore correlazione, che associa la classe corretta ad ogni nodulo, un esempio mostrato in Fig.22, assunto come Ground Truth.

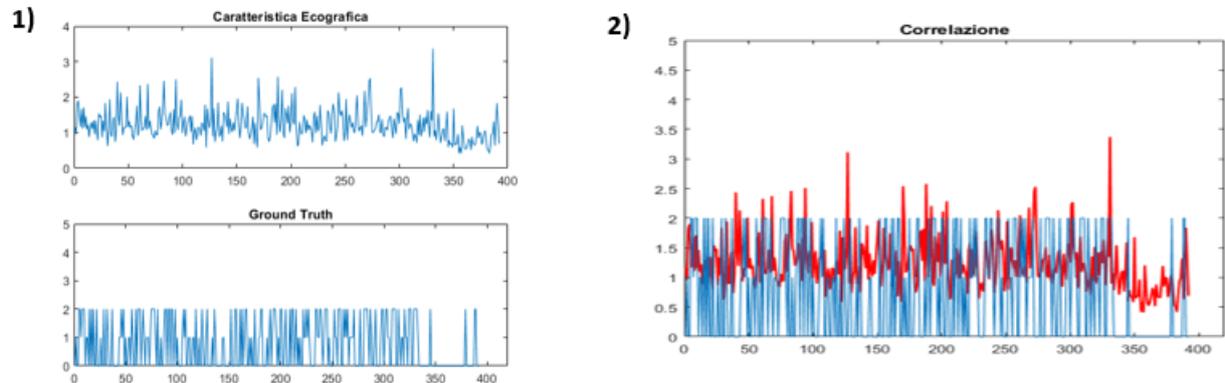


Figura 22: Esempio di come avviene un'analisi di correlazione usando (1) Caratteristica ecografica e Ground Truth (2) Rappresentazione di correlazione tra le due variabili; si nota la sovrapposizione tra il vettore Ground truth (in azzurro) e la caratteristica ecografica (in rosso)

Si è proceduto in seguito a una misura di correlazione tra il vettore in questione e i valori di ecogenicità. In particolare si è utilizzata il coefficiente di correlazione lineare di Pearson, il quale è il coefficiente di correlazione lineare più comunemente usato. Il calcolo della correlazione segue la seguente formulazione:

$$R = \frac{\sum_{i=1}^n (X_{a,i} - \bar{X}_a)(Y_{b,i} - \bar{Y}_b)}{\sqrt{\sum_{i=1}^n (X_{a,i} - \bar{X}_a)^2 \sum_{i=1}^n (Y_{b,i} - \bar{Y}_b)^2}}$$

in cui \bar{X}_a è la media dei campioni della prima variabile, \bar{Y}_b è la media dei campioni della seconda variabile. Il valore risultante dalle misure di correlazione è indice della qualità

della caratteristica. Questa concorre con gli altri test statistici per rafforzare e irrobustire la valutazione finale.

2.4.2 Mann-Whitney U-Test

Il test U di Mann-Whitney è un test non parametrico per verificare, in presenza di valori ordinali provenienti da una distribuzione continua, se due campioni statistici provengono dalla stessa popolazione [27]. Il test assume l'uguaglianza delle mediane di popolazione di due campioni indipendenti tra loro e le osservazioni all'interno di ogni campione sono indipendenti. Il test restituisce il numero di volte in cui un campione del secondo vettore precede uno del primo in una disposizione ordinata degli elementi nei due campioni indipendenti. È correlata alla statistica della somma dei ranghi Wilcoxon (W) nel modo seguente:

$$U = W - \frac{n_x(n_x-1)}{2}$$

in cui n_x rappresenta la lunghezza del vettore. Mann-Whitney determina la possibilità di ottenere maggiori osservazioni in una popolazione rispetto ad un'altra. L'ipotesi nulla nel test è quella che i due campioni siano tratti da una popolazione singola, e che dunque per questa ragione le loro distribuzioni di probabilità siano eguali. L'ipotesi alternativa è che uno dei campioni sia più grande in maniera stocastica. Questo richiede che i due campioni siano statisticamente indipendenti e l'osservazione ordinale o misurazioni continue. Questo test è stato utilizzato solo nella caratteristica margini, in particolare quando è stata calcolata la dimensione frattale del nodulo. Questo test è servito esclusivamente a verificare se sono stati selezionati due campioni indipendenti tra popolazioni aventi stessa distribuzione. successive analisi e valutazioni.

2.4.3 Analisi della Componenti Principali

L'analisi delle componenti principali (PCA) è una tecnica utilizzata nell'ambito della statistica multivariata per la semplificazione dei dati d'origine. Lo scopo primario di questa tecnica è la riduzione di un numero più o meno elevato di variabili (rappresentanti altrettante caratteristiche del fenomeno analizzato) in alcune variabili latenti[5]. Ciò avviene tramite una trasformazione lineare delle variabili che proietta quelle originarie in un nuovo sistema cartesiano nel quale le variabili vengono ordinate in ordine decrescente di varianza. Pertanto, la variabile con maggiore varianza viene proiettata sul primo asse, la seconda sul secondo asse e così via, un esempio è mostrato in Fig.23. La riduzione della complessità avviene limitandosi ad analizzare la varianza delle nuove variabili principali. In questa tecnica sono gli stessi dati che determinano i vettori di trasformazione. La PCA è una tecnica statistica adoperata in molti ambiti poiché quando ci si trova a semplificare un problema, riducendo la dimensione dello spazio di rappresentazione, si ha allo stesso tempo una perdita dell'informazione contenuta nei dati originali. La PCA consente di controllare egregiamente il "trade-off" tra la perdita di informazioni e la semplificazione del problema, scegliendo il numero appropriato di autovettori.

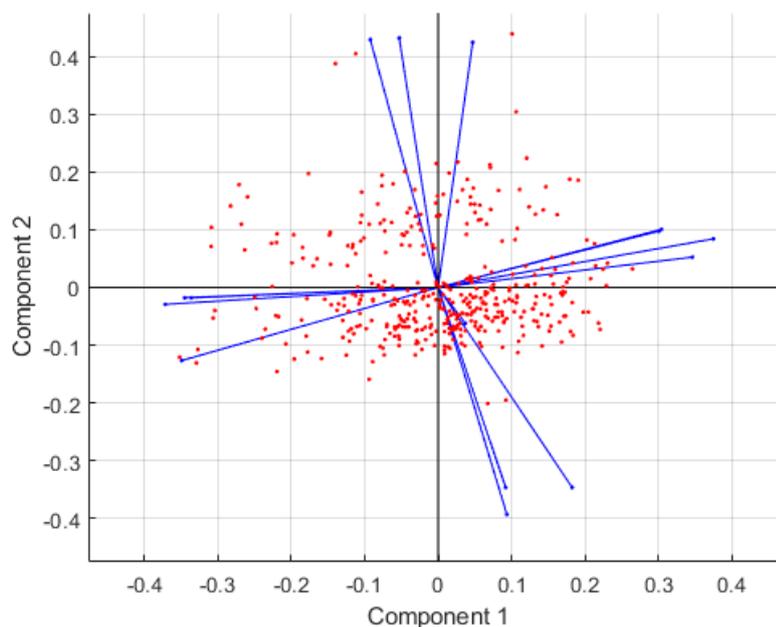


Figura 23: Rappresentazione grafica della disposizione dei dati secondo le componenti principali (PCA)

Per identificare le componenti principali ogni campione è descritto da nuove coordinate rispetto allo spazio delle componenti principali; queste coordinate prendono il nome di scores. In altre parole le componenti principali sono costruite come combinazione lineare delle variabili originali. I coefficienti lineari che descrivono le componenti principali in termini di variabili prendono il nome di loadings e permettono di determinare il contributo delle variabili originali al modello PCA. Un esempio di rappresentazione degli scores e dei loadings è riportato in Fig. 24

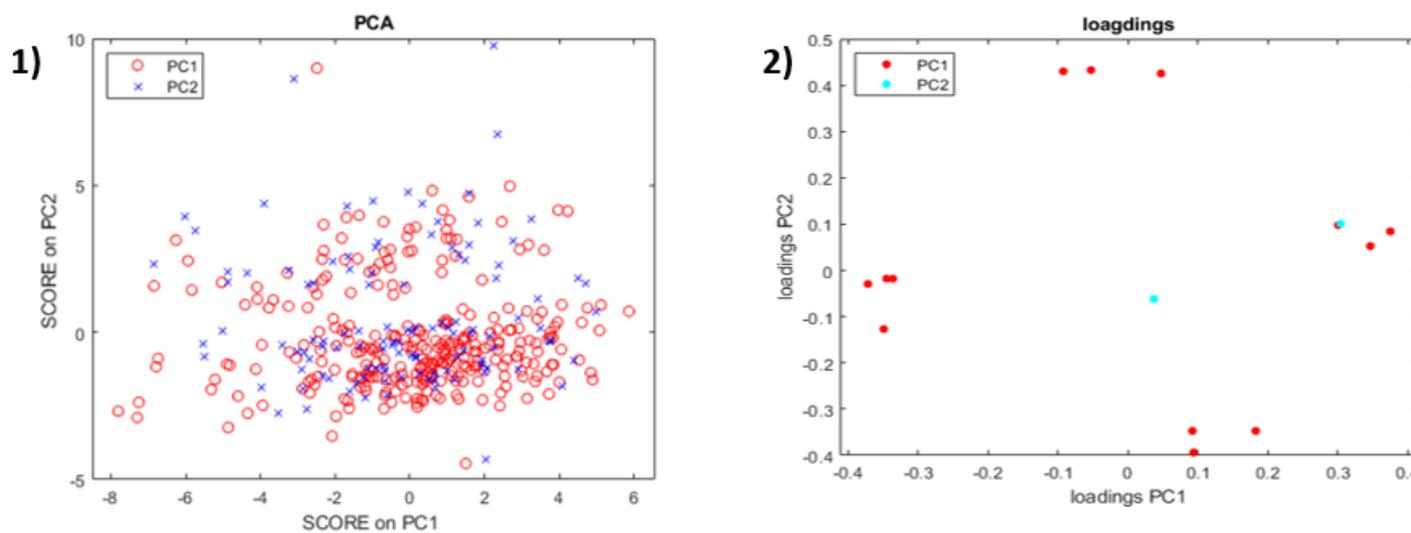


Figura 24: Rappresentazione delle coordinate e dei pesi per implementare la PCA (1) Rappresentazione degli scores (2) Rappresentazione dei loadings

Dal confronto tra il grafico degli scores e quello dei loadings si procede con l'interpretazione dei risultati. In particolare questa tecnica di analisi si applica solo ad alcune delle

caratteristiche ecografiche estratte, poichè non è sempre necessaria una semplificazione dei dati di origine.

2.4.4 Manova

L'analisi della varianza è un insieme di tecniche statistiche facenti parte della statistica inferenziale che permettono di confrontare due o più gruppi di dati analizzando la variabilità interna a questi gruppi con la variabilità tra i gruppi [38]. Si basa su un'ipotesi nulla la quale solitamente prevede che i dati di tutti i gruppi abbiano la stessa origine, ovvero la stessa distribuzione stocastica, e che le differenze osservate tra i gruppi siano dovute solo al caso. L'analisi della varianza si usa per determinare se più gruppi possono essere in qualche modo significativamente diversi tra loro (la varianza intergruppo contribuisce significativamente alla varianza totale; il fenomeno è legato a caratteristiche intrinseche del gruppo) o risultano omogenei (la varianza intra gruppo contribuisce significativamente alla varianza totale; il fenomeno è legato a caratteristiche relative a tutti i gruppi). In altre parole, il confronto si basa sull'idea che se la variabilità interna ai gruppi è relativamente elevata rispetto alla variabilità tra i gruppi, allora probabilmente la differenza tra questi gruppi è soltanto il risultato della variabilità interna. La MANOVA, in particolare, è un'estensione della più comune analisi della varianza (ANOVA). Nell'ANOVA si studiano le differenze tra una variabile. Nella MANOVA il numero delle variabili è maggiore. Quando si utilizza questo test si assumono i seguenti presupposti

- Le variabili di risposta sono continue
- I residui seguono la distribuzione di probabilità normale multivariata con medie pari a zero
- Le matrici varianza-covarianza di ciascun gruppo di residui sono uguali
- Le variabili sono indipendenti

I principali limiti e vantaggi sono elencati di seguito. In particolare MANOVA:

- è robusta per una modesta quantità di asimmetria nei dati. Gli outliers possono causare problemi di non-normalità nella distribuzione che neanche la robustezza del test può superare. Questo test è estremamente sensibile agli outliers, quindi si deve inizialmente valutare la presenza di outliers e determinare se l'assunzione della normalità risulta ragionevole.
- assume relazioni lineari tra le variabili dipendenti. Il verificarsi di relazioni curvilinee nei grafici a dispersione riduce il potere del Test.
- ipotizza che le matrici di covarianza all'interno del gruppo siano uguali. Se c'è un numero uguale di osservazioni per ogni gruppo il design del test è bilanciato e la robustezza è garantita.

Lo scopo principale del test è analizzare le caratteristiche estratte. L'analisi prende una serie di dati raggruppati e determina se la media di una variabile differisce significativamente tra i gruppi, determinando se l'intero set delle medie è diverso da un gruppo all'altro. Nello specifico si cerca di distinguere varie suddivisioni. Dall'applicazione del test si rappresentano le prime due variabili canoniche con un grafico a dispersione il quale utilizza coordinate cartesiane per visualizzare valori, tipicamente di due variabili, per un

insieme di dati, in Fig.25. Il test restituisce diverse informazioni, quali le variabili canoniche, i pesi delle variabili, la somma dei quadrati e la matrice dei prodotti incrociati e numerosi altri ancora. In particolare ci si è soffermato sulle variabili canoniche, variabili che sottoposte a cambiamenti di coordinate consentono di conservare la descrizione del sistema dinamico mediante la funzione hamiltoniana. L'informazione aggiuntiva fornita dal test è una stima della dimensione dello spazio contenente i mezzi del gruppo; il test verifica l'ipotesi nulla, cioè che i mezzi di ciascun gruppo siano lo stesso vettore multivariato n-dimensionale e che qualsiasi differenza osservata nel campione corrente sia dovuta a casualità. Se il parametro assume per empio un valore $d = 0$, non ci sono prove per respingere tale ipotesi, se $d = 1$, allora è possibile rifiutare l'ipotesi nulla al 5%, ma non è possibile rifiutare l'ipotesi che il multivariato significhi giacere sulla stessa linea. Allo stesso modo, se $d = 2$ i mezzi multivariati possono trovarsi sullo stesso piano nello spazio n-dimensionale, ma non sulla stessa linea. Il p-value rappresenta la robustezza del test, un $p\text{-value} < 0.05$ indica che la probabilità che la differenza misurata fra le medie possa dipendere solo dal caso per valori minori del 5%. L'output del test, il p-value, è un vettore di valori per una sequenza di test, il primo p-value verifica se la dimensione è 0, il successivo se la dimensione è 1 e così via. In questo caso entrambi i valori di p sono piccoli. I pesi delle variabili sono di fondamentale importanza, poichè rappresentano quanto una determinata feature ha influito sul calcolo delle variabili canoniche; si riesce così a distinguere e determinare le features più importanti in base al loro peso.

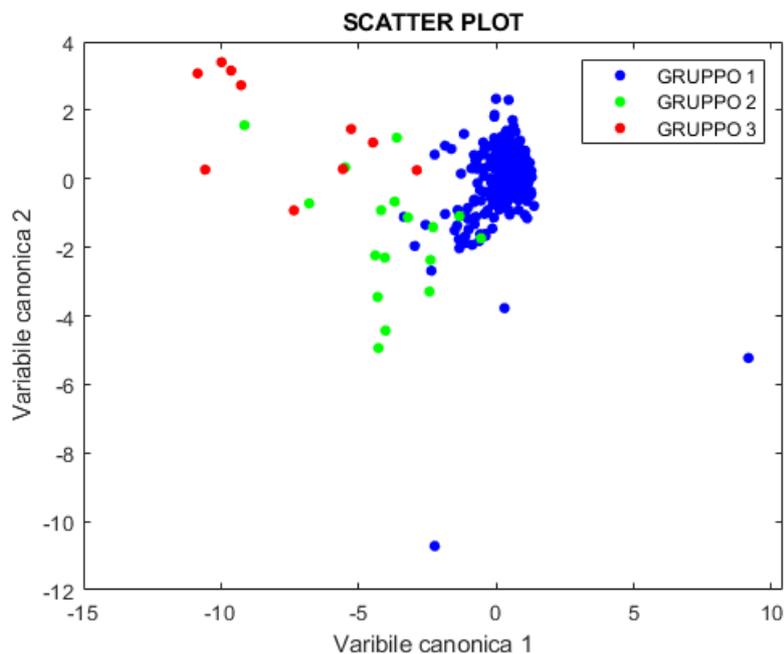


Figura 25: Scatter plot rappresentante le prime due variabili canoniche con codifica a colore che evidenzia tre classi differenti

Si rappresentano solo le prime due variabili canoniche poichè risultano essere le più discriminanti e le più accurate in termini di raffigurazione. I risultati della MANOVA applicata a tutte le features sono descritti nello specifico nei paragrafi seguenti.

2.5 Quantificazione Automatica delle Caratteristiche Ultrasonografiche

Per identificare e calcolare delle caratteristiche che consentono di descrivere alcuni aspetti ecografici rilevanti al fine di poter classificare una lesione tiroidea sono stati compiuti una serie di passaggi semi-automatici, come mostrato in Fig.26. I principali step affrontati sono:

1. Calcolo maschera per rimozione artefatti
2. Calcolo maschera cisti
3. Calcolo ecogenicità
4. Preprocessing delle immagini
5. Calcolo struttura
6. Calcolo forma
7. Calcolo margini
8. Calcolo vascolarizzazione
9. Calcolo elasticità

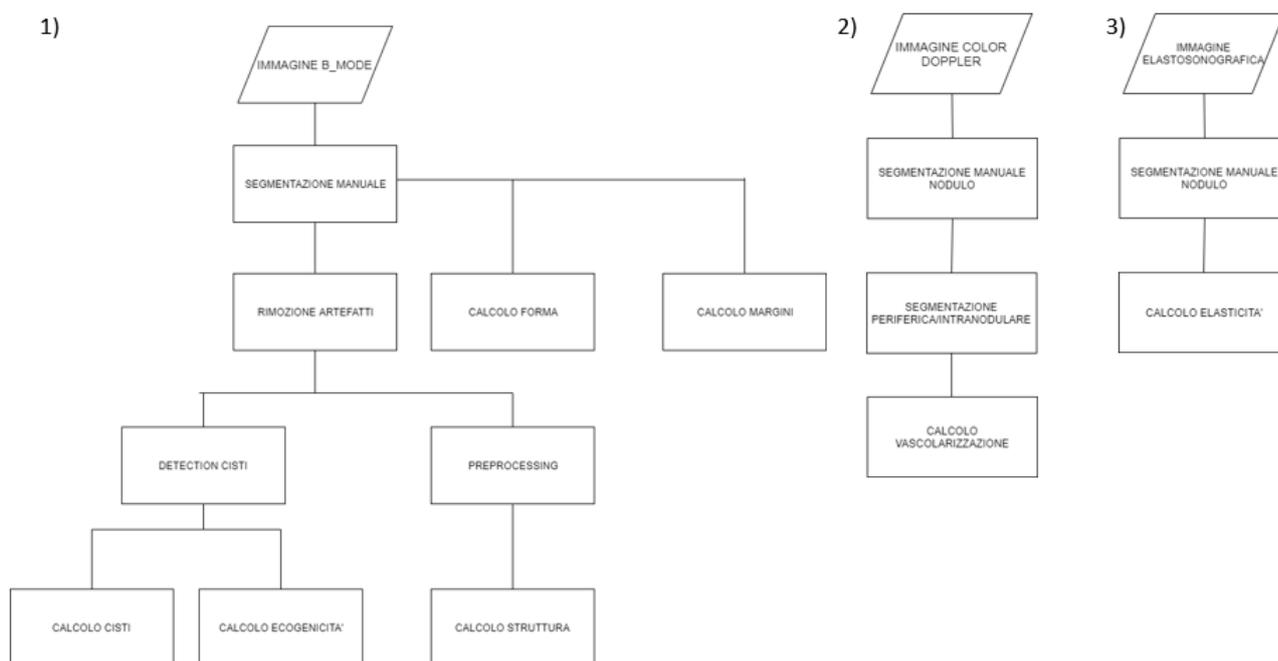


Figura 26: (1)Step per le immagini b-mode(2)Step per le immagine color doppler(3)Step per le immagini elastosonografiche

2.5.1 Ecogenicità

L'ecogenicità è definita come la capacità di una struttura organica di riflettere ultrasuoni. In termini digitali, si può assumere l'ecogenicità come l'intensità media dei pixel costituenti il nodulo. L'operazione preliminare per il calcolo dell'ecogenicità è stata la segmentazione del nodulo adoperando la maschera binaria già presente del database, in Fig.27, individualizzazione della ROI (region of interest).

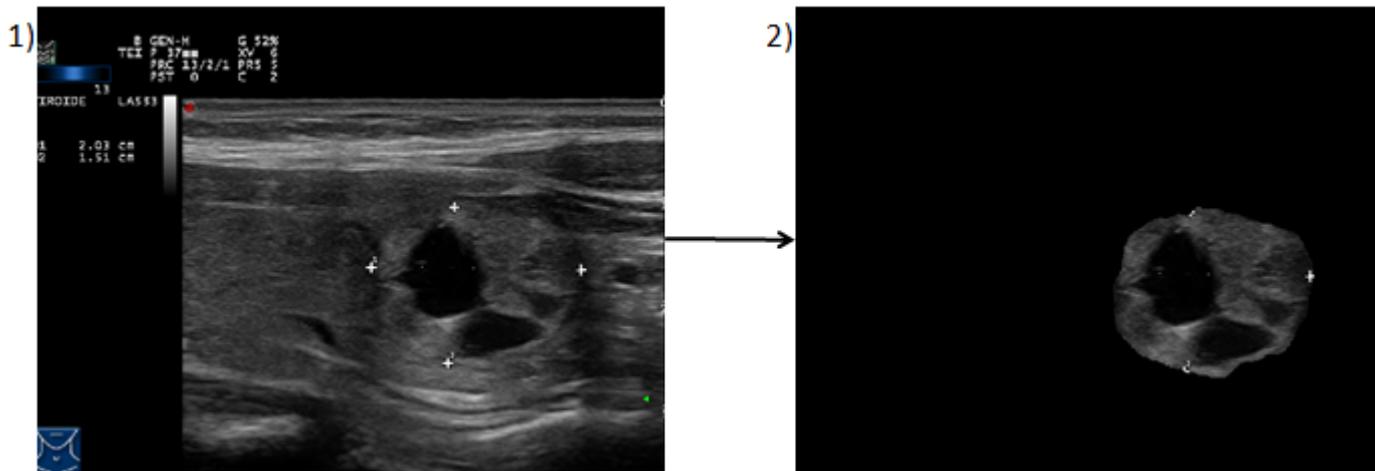


Figura 27: (1)Immagine totale non segmentata(2)Immagine nodulo segmentato con maschera binaria, ROI

Una volta identificato il nodulo e la sua posizione nell'immagine ecografica è stata calcolata la maschera degli artefatti. Inizialmente sono stati identificati gli artefatti introdotti dalla sonda con cui sono state acquisite le immagini. In particolare sono stati riconosciuti i quadratini e le croci, numerate e non numerate, un esempio riportato in Fig.28.

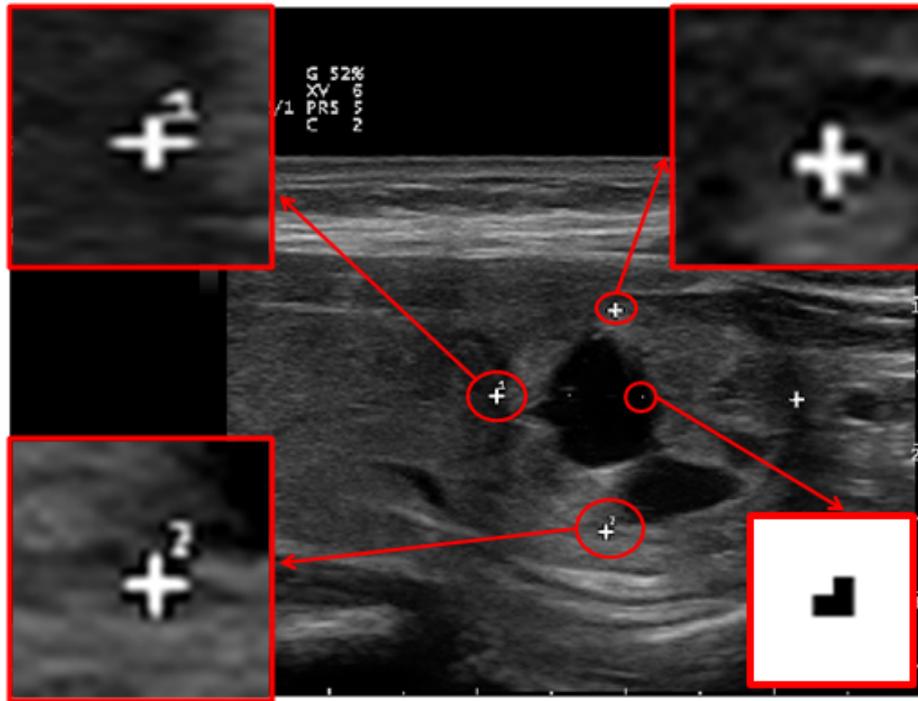


Figura 28: L'immagine mostra la posizione degli artefatti da riconoscere e la loro forma è rappresentata nei riquadri rossi

Gli artefatti sono presenti in tutte le immagini adoperate. Quindi a causa della necessità di pulizia delle immagini sono stati adottati per ogni caratteristica i seguenti step:

- Riconoscimento quadratino
- Caricamento Template
- Cross correlazione dei Template sull'immagine
- Ricerca indici
- Creazione quadrato per rimozione artefatti
- Rimozione quadratini

In particolare il riconoscimento dei quadratini è stato basato sulla somma delle intensità del quadratino stesso, essendo nota sia la dimensione che il valore di ogni singolo pixel che lo compone, quindi è stato ricercato e annesso alla maschera degli artefatti. Sono stati creati tre nuovi template: una croce con il numero 1, una croce con il numero 2, una croce bianca, i quali sono risultati notevolmente di supporto al template standard sia per il riconoscimento di forme più complesse come le croci numerate sia a causa di tessuto iperecogeno che interferiva con il riconoscimento degli artefatti, rendendo più ostico avendo a disposizione solo un template su sfondo nero. La Fig.29 mostra i template utilizzati:



Figura 29: (1)Template standard (2)Template con numero uno(3)Template con numero 2 (4)Template bianco

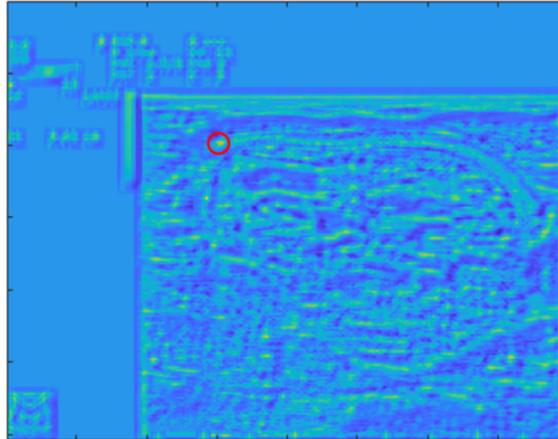


Figura 30: Immagine della cross correlazione per il riconoscimento degli artefatti. Massimo derivante dalla cross correlazione cerchiato in rosso che indica la presenza dell'artefatto

L'immagine derivante dalla cross correlazione, presentata in Fig.30, è stata inizialmente ridimensionata, per evitare i problemi dovuti alla cross correlazione del template sull'immagine, e adottata poi anche come immagine di supporto, cioè è stata sommata all'immagine iniziale per aumentarne il valore e per semplificare così la ricerca degli indici. La ricerca degli indici è stata gestita tramite soglie: le soglie sono state impostate in base alla qualità della cross correlazione e dei template stessi. Nello specifico è stata scelta una soglia pari al 86% del massimo per le croci, pari al 81.5% del massimo per le croci numerate e pari al 96% del massimo per la croce bianca, questi valori di soglia sono stati scelti in base a numerose prove in modo da minimizzare errori di riconoscimento. Non si sono presentate immagine in cui gli artefatti non sono stati riconosciuti, tuttavia circa nell'1% dei casi è stato introdotto un ulteriore quadratino, a causa del template bianco, dove non si presentano artefatti. Questo errore non ha influito sul calcolo dell'ecogenicità, tuttavia riguarda principalmente noduli iperecogeni ragion per cui non intacca la classificazione finale. In particolare sono stati ricercati quei punti di un colore giallo intenso corrispondenti ai massimi della cross correlazione, evidenziati in Fig.30. Successivamente è stato inserito un controllo in modo da evitare un numero cospicuo di indici: la soglia è stata variata in base al numero di indici trovati, lo scopo è stato quello di ottenere lo stesso numero di indici per i vari template in posizione differente e/o medesima, sovrapponendosi alla collocazione dell'artefatto. A seguire sono stati creati dei quadrati intorno le posizioni degli artefatti riconosciuti di dimensione 14x14 pixel. I quadrati creati e le posizioni dei quadratini trovati precedentemente sono stati inglobati nella maschera di segmentazione del nodulo creando la maschere degli artefatti, un esempio in Fig.31.

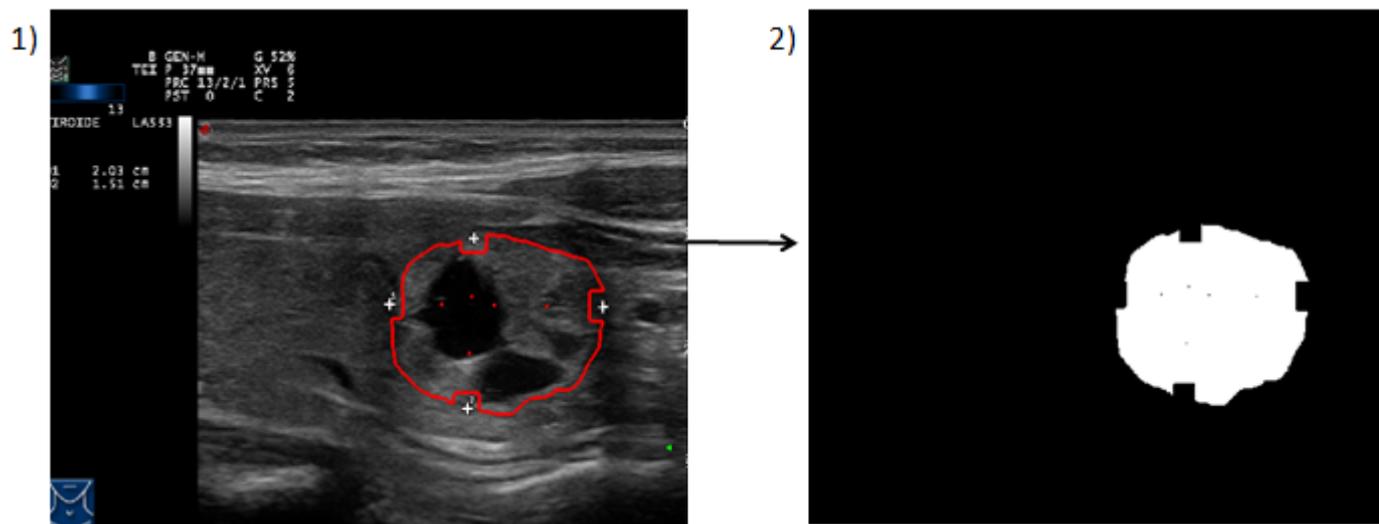


Figura 31: (1)Maschera artefatti sovrapposta al nodulo in rosso(2)Maschera artefatti binaria

Seguentemente alla maschera degli artefatti è stata calcolata la maschera delle cisti. Questa ulteriore maschera è risultata indispensabile al fine di evitare che pixel anecogeni componenti le cisti corrompessero il calcolo finale dell'ecogenicità, dato che la valutazione di questa caratteristica si basa solo sui pixel costituenti la porzione di tessuto costituente il nodulo. Le operazioni primarie per il riconoscimento delle cisti sono state riscontrare il minimo valore di intensità di grigio e calcolare la deviazione standard, il tutto all'interno della ROI. Successivamente è stata definita la soglia di riconoscimento delle cisti come:

$$L = GR_{min} + l * \sigma_{ROI}$$

dove GR_{min} rappresenta il valore minimo in termini di scala di grigio, σ_{ROI} rappresenta la deviazione standard della ROI, l rappresenta il moltiplicatore della σ_{ROI} [8]. La scelta del valore del parametro l è ricaduta su un valore pari a 0.5. Il tuning del parametro l ha visto la creazione di diverse maschere con l'utilizzo di valori di l da 0.3 a 1, a passi di 0.025. In Fig.32 sono mostrate alcune maschere in base ad un determinato valore del parametro.

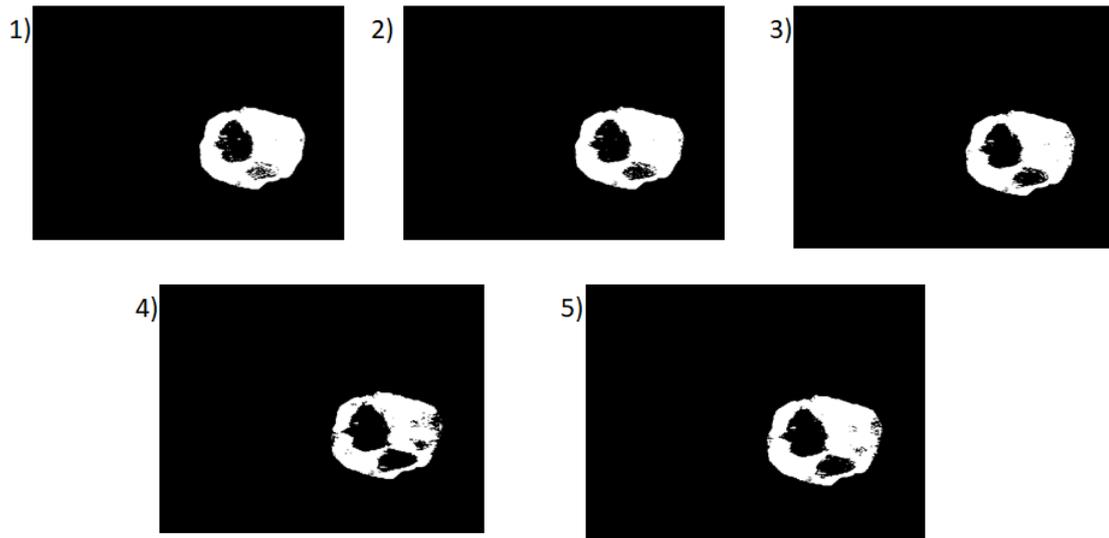


Figura 32: Tuning del parametro l e maschere derivanti per (1) $l=0.3$ (2) $l=0.4$ (3) $l=0.45$ (4) $l=0.5$ (5) $l=0.6$

E' stato notato in Fig.32, come all'aumentare del parametro l aumenti anche la dimensione della maschera, riferendosi ai pixel scuri. E' stato constatato che il valore più rappresentativo della morfologia della cisti fosse pari a $l=0.5$. I pixel che sono stati inglobati nella maschera presentavano un valore minore della soglia a causa dell'anecogenicità della cisti, in Fig.33 un esempio di maschere delle cisti. In seguito alla creazione delle maschere è stata valutata l'ecogenicità, non tenendo in considerazione le zone cistiche e gli artefatti. L'esclusione di questi parametri è stata fondamentale per il calcolo dell'ecogenicità evitando che zone anecogene e artefatti influenzassero negativamente il risultato. L'ecogenicità di un nodulo è stata calcolata come media dei valori all'interno delle maschere create, cioè solo i pixel che lo rappresentano, ed è stata rapportata rispetto all'ecogenicità del parenchima tiroideo e anche all'ecogenicità del muscolo. L'ecogenicità del muscolo e del parenchima tiroideo sono state calcolate con il medesimo criterio ma nelle rispettive ROI relative ad una porzione di tessuto specifico.

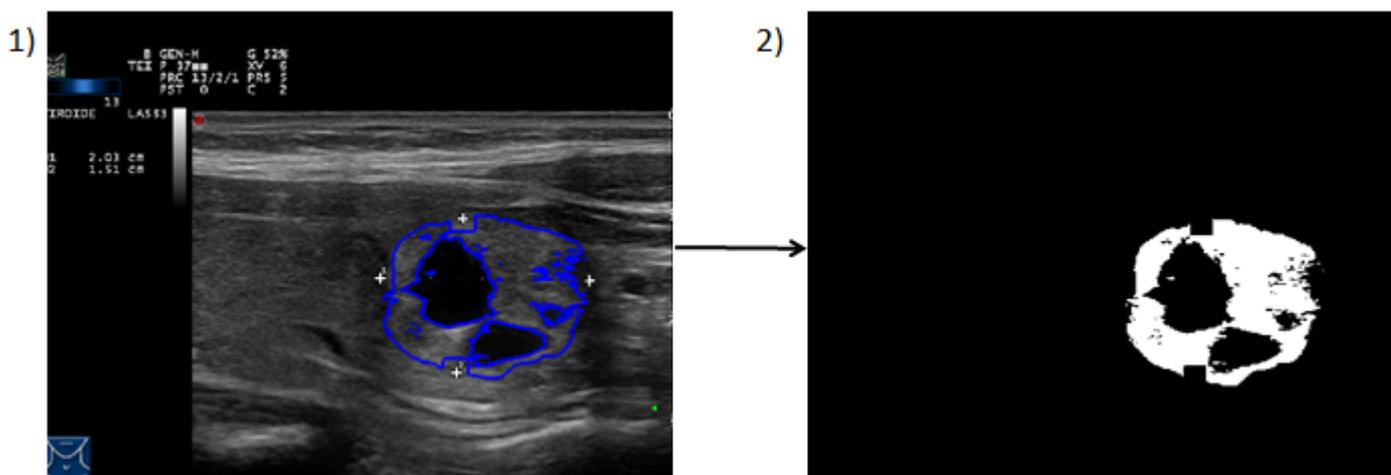


Figura 33: (1)Maschera che incorpora le cisti sovrapposta al nodulo in blu(2) Maschera cisti binaria

La decisione di rimozione degli artefatti e delle cisti è stata presa in base alla scelta che l'esperto compie durante l'indagine dell'ecogenicità. Infatti il suo parere si fonda esclusivamente sull'ecogenicità del tessuto di cui si compone il nodulo, non tenendo conto della presenza di cisti, pixel anecogeni, e degli artefatti. E' stato deciso di condurre le prove di correlazione rispetto alla classificazione presente nel database. La caratteristica più rappresentativa risulta l'ecogenicità rispetto alla tiroide con una correlazione pari a 0.489. Questo valore è tra i più alti ottenuti il che rende l'ecogenicità rispetto alla tiroide una delle caratteristiche più descrittive della natura del nodulo rispetto alle successive restanti.

2.5.2 Struttura

Per struttura si intende, in termini di immagini, la differente intensità dei pixel che compongono il nodulo. In particolare una struttura omogenea ha pixel con intensità simili e disomogenea con intensità differenti. I passi compiuti per giungere ad una stima della struttura sono:

- Preprocessing dell'immagine
- Definizione dei livelli di grigio
- Estrazione Haralick features
- Estrazione Galloway features
- Calcolo struttura

Il primo passo compiuto, per quanto concerne il preprocessing, è stato il lo scalamento dell'immagine, in particolare è stato implementato un min-max scaling dell'immagine mascherata. Nell'immagine mascherata sono presenti i quadratini delle maschere degli artefatti; al fine di evitare che questi artefatti entrino a far parte del calcolo della struttura e conducano così ad errori grossolani, è stato implementato un filtro mediano con cui processare l'immagine. Successivamente è stato definito il numero di livelli di grigio sui quali estrarre le features. Le Haralick features sono state successivamente calcolate a partire dalla "Gray Level Co-occurrence Matrix" (GLCM), mostrata in Fig.34. Il GLCM è una matrice quadrata con una dimensione pari al numero di livelli di grigio nell'immagine. L'elemento $GLCM(i, j)$ misura il numero di volte in cui un pixel di un determinato livello di grigio i si trova adiacente a un pixel di livello di grigio j . Poiché due pixel possono essere adiacenti in verticale, orizzontale e nelle due direzioni diagonali, sono stati calcolati i valori della matrice per i quattro angoli 0, 45, 90 e 135. Sono state quindi estratte le features di primo ordine: media, deviazione standard, varianza, skewness, kurtosis ed entropia. Anche le features di secondo ordine vengono estratte per ogni angolazione e mediate tra di loro, in particolare sono stati ricavati 4 descrittori medi: contrasto, correlazione, energia, omogeneità. Sono state quindi ricavati 10 Haralick features. In una struttura grossolana, le lunghe corse intese come sequenza di livelli di grigio possono esistere più frequentemente rispetto che in una trama fine. Ciò implica che possono presentarsi sequenze più lunghe di pixel con la stessa intensità nelle immagini grossolane, ma brevi corse in trame sottili. Le caratteristiche di Galloway sono quindi descrittori matematici che si basano sulla matrice della lunghezza di esecuzione RLM(RUN LENGTH MATRIX), mostrata in Fig.34, cioè sulla lunghezza di queste sequenze. In un RLM, il pixel $R(i,j)$ contiene il numero di pixel con lunghezza di esecuzione "j" e intensità "i" in una determinata direzione. Sono stati determinati i valori della matrice per i quattro angoli 0, 45, 90 e 135 e mediati tra di loro

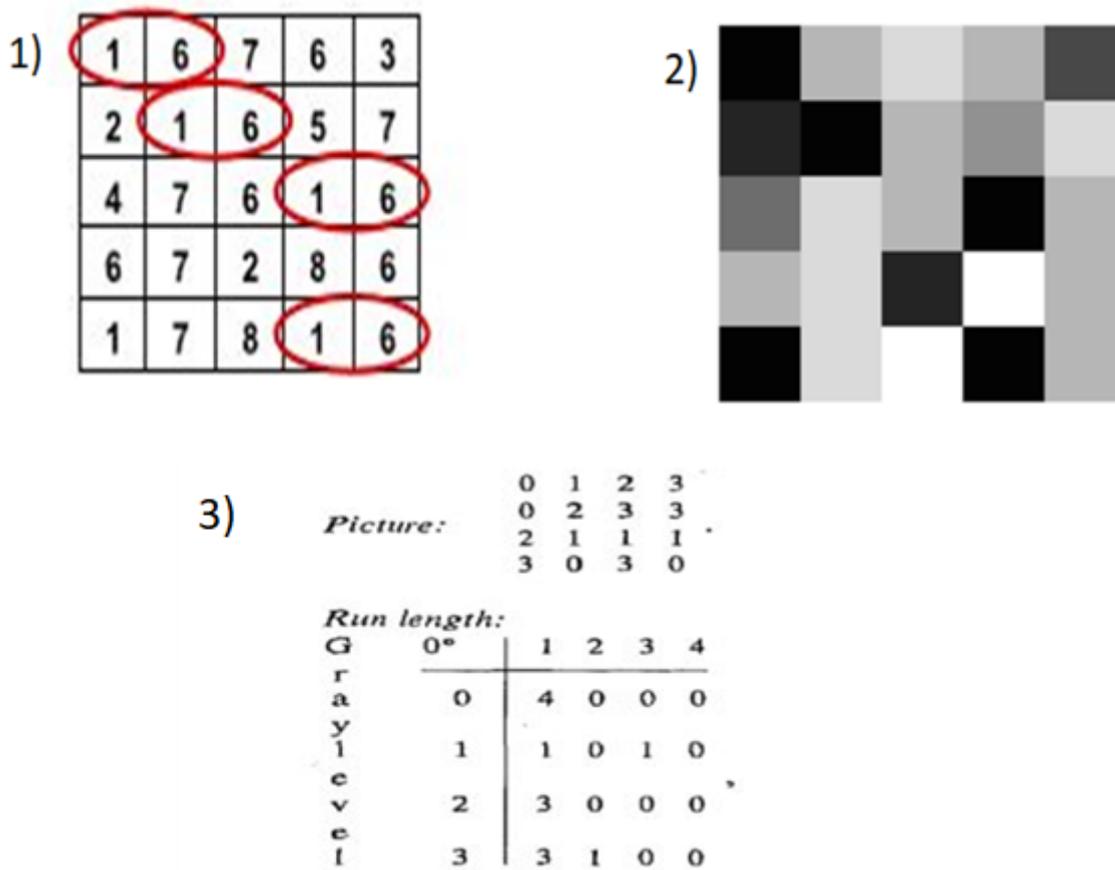


Figura 34: Esempi di matrici utilizzate (1)GLCM numerica(2) GLCM in scala di grigi(3) RLM numerica, (tratto da [2])

Sono state estratte 5 features in totale: Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray-Level Nonuniformity (GLN), Run Length Nonuniformity (RLN) e Run Percentage (RP). Le caratteristiche estratte sono 15, a cui sono state aggiunte il calcolo della deviazione standard calcolata dai pixel del nodulo. La struttura è stata così descritta da un totale di 16 features.

La scelta del numero dei livelli di grigio e della dimensione del filtro è stata determinata dalla messa a punto dei parametri e perciò avvenuta in base ai risultati derivanti da successive prove adottando combinazioni di parametri diverse. Per decretare la combinazione migliore dei parametri è stato osservato il p-value proveniente dalla varie prove. E' stata adottata la combinazione più efficace, filtro mediano di dimensione 3x3 e numero di livelli di grigio pari a 128, poichè il p-value derivante dai parametri, pari a , è stato il più basso raggiunto. Gli intervalli dei valori di utilizzo per ciascun parametro sono:

1. *Filtro Mediano* [3 5 7]:

E' stata variata la dimensione del filtro mediano usato nel preprocessing. Il filtro mediano è responsabile dell'introduzione dell'effetto di media che affligge l'immagine e ne consegue una maggiore omogeneità in base alla sua dimensione.

2. *Livelli di Grigio* [32 64 128]:

Sono stati variati i livelli di grigio utilizzati nel calcolo delle caratteristiche. I livelli scandiscono la risoluzione della caratteristica e ne deriva l'aumento della risoluzione all'aumentare del numero dei livelli impiegato.

Sono state indagate nove combinazioni totali. L'analisi statistica adottata per questa caratteristica è derivante da un approccio ibrido, PCA-MANOVA. Una volta calcolate le features, per ogni singola combinazione, è stata eseguita una PCA. L'utilizzo della PCA si è dimostrato di sostanziale importanza, poichè sono stati ottenuti numerosi descrittori per la struttura e la rimozione di caratteristiche correlate tra loro si è dimostrata migliorativa per quanto concerne i risultati. Si è successivamente proseguito con la MANOVA. La Fig.35 mostra il grafico finale a dispersione discendente dai due test:

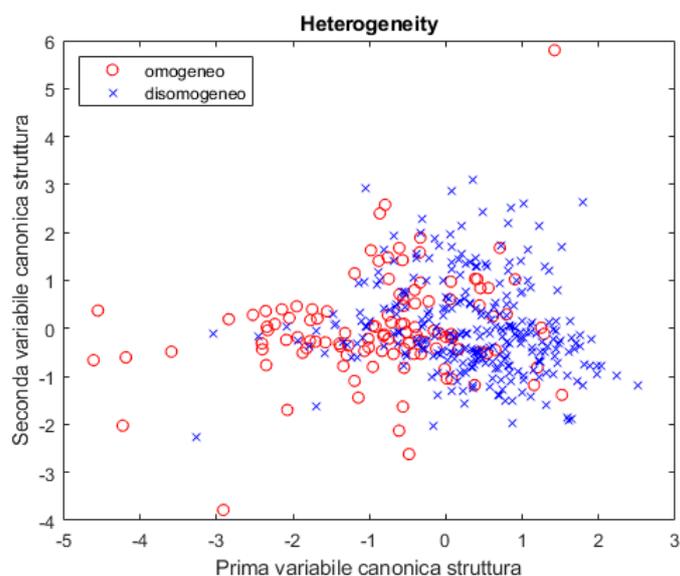


Figura 35: Scatter plot della prima e della seconda variabile canonica rappresentante la struttura dei noduli divisi nelle due categorie: omogeneo e disomogeneo

In questo caso sono state calcolate media, deviazione standard, varianza, skewness, kurtosis, entropia, contrasto, correlazione, energia, omogeneità, SRE, LRE, GLN, RLN, RP utilizzate per il test. E' stato ricavato il grafico soprastante riportando sull'asse delle ascisse e su quella delle ordinate, rispettivamente, il valore ottenuto utilizzando la prima (parametro 1) e la seconda (parametro 2) combinazione lineare più efficiente.

Dalla Fig.35, la distinzione tra le due classi non risulta possibile visivamente, ma il valore $d=1$ e il $p\text{-value}=8.35 \times 10^{-15}$, derivanti dal test tuttavia affermano l'opposto, cioè è possibile distinguere le due classi con buona affidabilità in base al parametro uno.

Di seguito sono stati così estratti i pesi delle variabili più rappresentative, le quali sono state scelte in base al contributo maggiore che forniscono nel calcolo della prima variabile canonica della MANOVA.

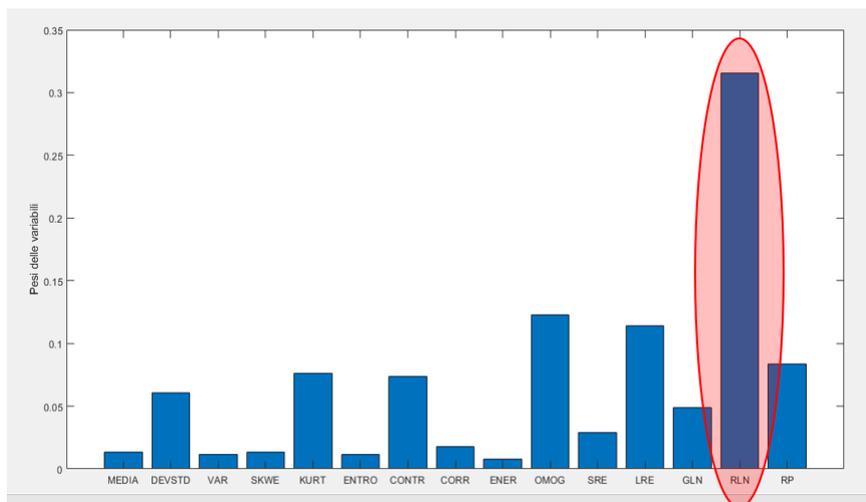


Figura 36: Immagine dei pesi che corrispondono in ordine a MEDIA, DEVSTD, VAR, SKWE, KURT, ENTRO, CONTR, CORR, ENER, OMOG, SRE, LRE, GLN, RLN, RP. Il peso evidenziato corrisponde alla caratteristica di maggior importanza nel calcolo della prima variabile canonica

La Fig.36 mostra il peso posseduto da ogni caratteristica nel calcolo della prima variabile canonica. Infine la scelta per estrarre un singolo descrittore rappresentativo della caratteristica in questione è ricaduta sulla variabile con peso maggiore evidenziata in Fig.36. In questo caso è stata selezionata la Run Length Non-Uniformity, la quale è di principale rilievo nella composizione della prima variabile canonica e in assoluto detiene il peso maggiore. A causa del notevole contributo apportato alla quantificazione della caratteristica è stata utilizzata come descrittore principale della struttura. La mancata ottimizzazione dei parametri può portare a delle perdite di informazione o a degli errori nel calcolo delle caratteristiche, quindi il tuning di questi due parametri ricopre un ruolo principale nella quantificazione della struttura.

2.5.3 Margini

Per questa caratteristica sono state percorse due differenti strade. Per il primo percorso è stato adottato un approccio “boundary-based”, che descrive i contorni del nodulo invece di privilegiare le informazioni relative al contenuto, ovvero un approccio “region-based”. In particolare sono stati realizzati i seguenti step:

- Individuazione del contorno dell’oggetto
- Calcolo proprietà globali
- Individuazione punto più distante
- Calcolo distanza
- Trasformata di Fourier

Il primo passo per l’estrazione dei descrittori è stato l’individuazione del contorno. In seguito è stato individuato il punto più lontano e ne è stata calcolata la distanza euclidea (Farthest point distance, FPD). La funzione FPD si ricava dal contorno di un’immagine, cioè per ogni punto “a” $(x(u),y(u))$ si individua il punto “b” (x_{fp}, y_{fp}) più lontano da esso e si calcola la distanza euclidea di ciascuno di essi dal centroide (x_c, y_c) . La funzione $FPD(u)$ è definita come la somma tra le due distanze rappresentata graficamente in Fig.37, ed espressa dalla formula:

$$FDP(u) = \sqrt{[x(u) - x_c]^2 + [y(u) - y_c]^2} + \sqrt{[x_{fp}(u) - x_c]^2 + [y_{fp}(u) - y_c]^2}$$

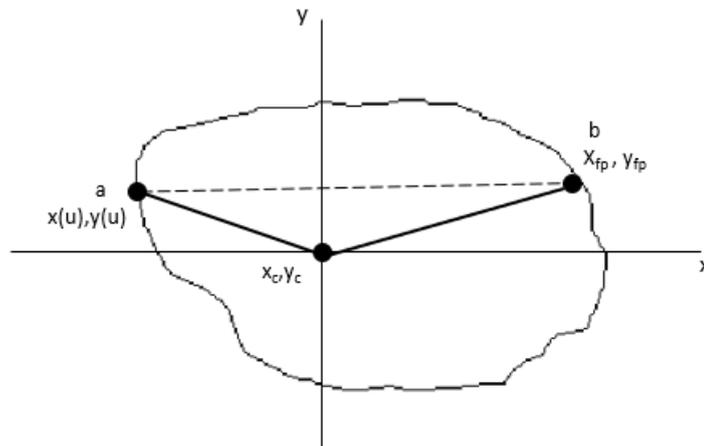


Figura 37: Esempio pratico di calcolo Farthest point distance su un nodulo tiroideo

Applicando la trasformata discreta di Fourier alla FPD, sono stati ricavati i descrittori di Fourier. L’utilizzo della trasformata di Fourier consente di ottenere alcune proprietà necessarie per un’analisi corretta della forma. È stato considerato esclusivamente il modulo di ciascun coefficiente trascurandone la fase, in modo da eliminare la dipendenza dalla rotazione. La prima metà dei descrittori è stata normalizzata rispetto alla componente continua, la quale contiene informazioni relative all’energia media del segnale e solitamente ha un modulo maggiore agli altri, in modo così da assicurare l’invarianza rispetto alla scala. Infine, poiché i contorni sono stati calcolati rispetto al centroide ne consegue che

i descrittori ottenuti risultano invarianti anche rispetto alla traslazione [17]. I valori dei coefficienti si sono rivelati simmetrici rispetto al punto medio. Per tale ragione ne sono stati calcolati 21 per utilizzarne i primi 10 (esclusa la componente continua). In aggiunta a questi sono stati calcolati anche 4 descrittori globali ovvero circolarità, eccentricità, solidità e proporzioni.

La seconda strada percorsa adotta un approccio totalmente diverso basato sulla dimensione frattale [6]. In geometria frattale la dimensione frattale è una quantità statistica che dà un'indicazione di quanto completo appare un frattale per riempire lo spazio. Per calcolare e implementare la dimensione frattale di un nodulo sono risultati necessari i seguenti step:

- Crop della maschera
- Definizione perimetro
- Copertura oggetto con la griglia di square cell di lunghezza nota
- Conteggio delle cell in cui si trova l'oggetto
- Ripetizione copertura e conteggio con square cell di dimensioni decrescenti

Il crop della maschera ha permesso di identificare l'oggetto in una regione limitata senza dover ricoprire con la griglia tutta l'immagine. La dimensione frattale adotta la sotto sequenza di divisione della struttura originale per generare una struttura frattale. La figura seguente mostra un esempio di creazione di dimensione frattale utilizzando una sottostruttura:

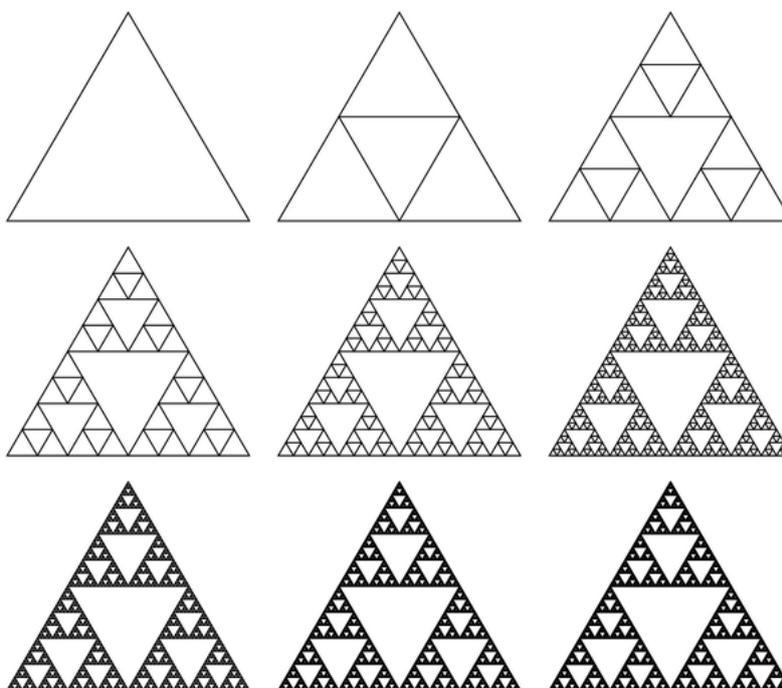


Figura 38: Evoluzione del triangolo di Sierpiński, (tratto da 38)

Il triangolo mostrato in Fig.38 è un esempio di frattale: cioè è un insieme generato da un pattern che si ripete uguale a se stesso su scale differenti. I valori frattali calcolati

sono stati presentati su un grafico in scala logaritmica, in Fig.39, dove l'asse orizzontale rappresenta la lunghezza del lato di una cella e l'asse verticale rappresenta il numero di mesh contenenti gli elementi dell'oggetto. La linea di regressione è stata contrassegnata in base a punti calcolati. Il valore della pendenza della retta di regressione è direttamente correlato al valore della quota di conteggio delle caselle.

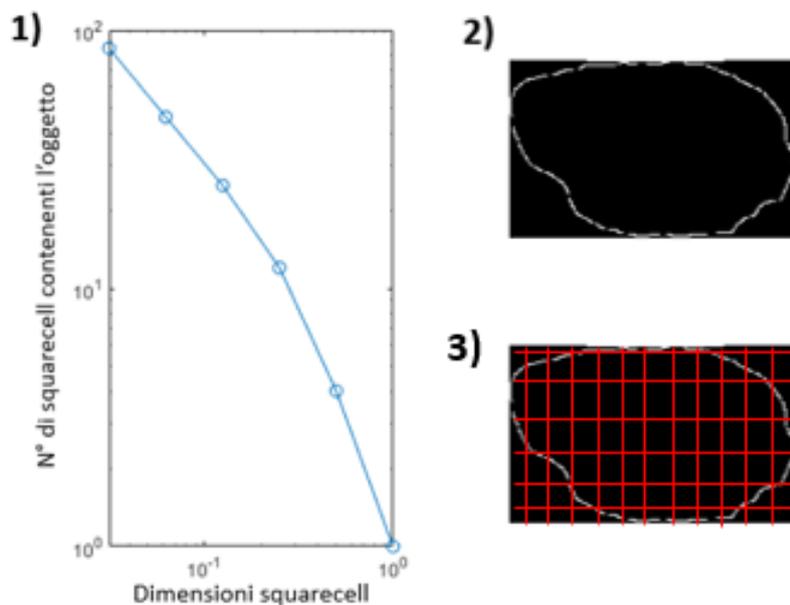


Figura 39: Rappresentazione di step per il calcolo delle dimensioni frattali (1) Esempio di rappresentazione in scala logaritmica del numero di square cell rispetto alle dimensioni (2) Esempio di crop della maschera e perimetro (3) Esempio copertura oggetto con griglia evidenziata in rosso

In questo caso è stato effettuato uno dei test statistici più semplici, il Mann-Whitney test, per verificare se le dimensioni frattali potesse essere rappresentative della caratteristica. Il p-values derivante dal test è pari a 0.045, che è significato in termini statistici tuttavia al limite dall'accettabilità. Un criterio maggiormente restrittivo sul p-values lancerebbe dei dubbi sulla validità del test e quindi sull'impiego delle dimensioni frattali come unico descrittore.

Il calcolo dei margini, come per la struttura, è stato anch'esso discendete da un approccio ibrido, PCA-MANOVA. Successivamente all'ottenimento dei valori, è stato compiuto uno sfolgimento dei dati, conservando una buona qualità dei dati originali, tramite PCA. Di seguito è stata condotta la MANOVA sugli scores provenienti dalla PCA. In Fig.40 è mostrata la distribuzione delle prime due variabili provenienti dai test:

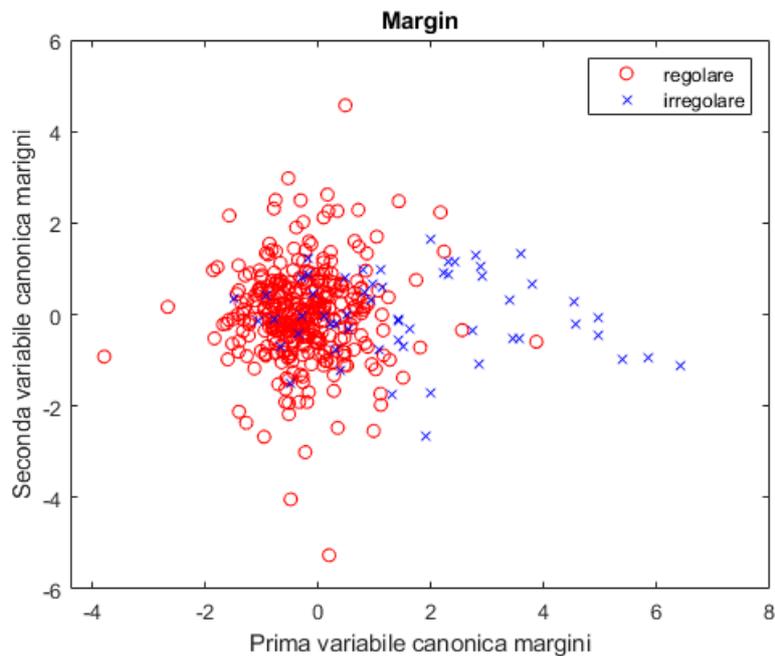


Figura 40: Scatter plot della prima e della seconda variabile canonica rappresentante i margini dei noduli divisi nelle due categorie: regolare e irregolare

Per gli scores, e quindi per la MANOVA, si sono utilizzate 15 features: eccentricità, solidità, proporzioni, circolarità, 10 descrittori di Fourier e le dimensioni frattali. Si è ottenuto un valore $d=1$, anche se dall'immagine soprastante non risulta una netta distinzione visiva.

Tuttavia è stato possibile discriminare tra le due classi in funzione del parametro uno anche se la distribuzione dei valori per i noduli con margini irregolari si è dimostrata più dispersiva. La solidità invece del test è rappresentata da un $p\text{-value} = 3.3 * 10^{-24}$.

Per rivelare la caratteristica più significativa e delineare così una variabile descrittiva vantaggiosa è stata performata l'estrazione dei pesi delle caratteristiche. E' stata così ricercata la caratteristica che fornisce un contributo maggiore per il calcolo della prima variabile canonica. La Figura seguente mostra i pesi delle caratteristiche:

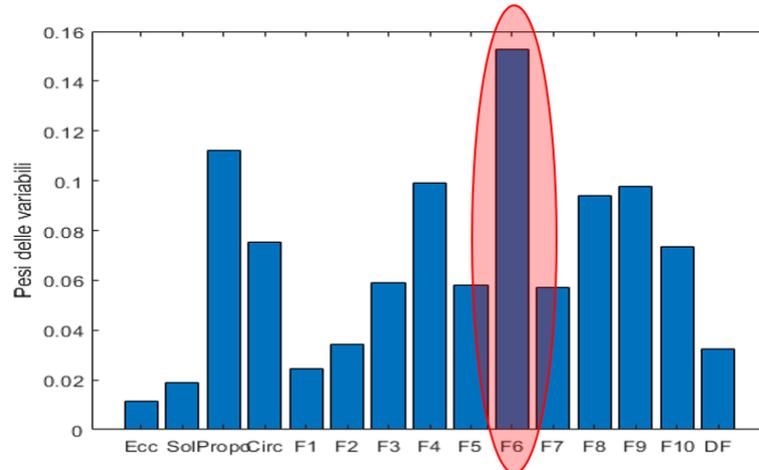


Figura 41: Istogramma dei pesi della prima variabile canonica dei margini. La barra evidenziata mostra la caratteristica con peso maggiore

La variabile di interesse è stata identificata nel sesto descrittore di Fourier, evidenziata in Fig.41. Nonostante l'importante peso posseduto da questa variabile le altre non si sono dimostrate di importanza notevolmente minore, infatti anch'esse sono state impiegate nella definizione della prima variabile canonica con un discreto margine di influenza. Tutto questo si traduce in una difficoltà di quantificazione della caratteristica corretta poichè a causa dei pesi più o meno bilanciati non è risultato agevole discriminare i margini utilizzando unicamente una variabile.

2.5.4 Cisti

Una cisti è una cavità o sacca, normale o patologica, racchiusa da una membrana ben distinta. L'identificazione della cisti ricopre un ruolo chiave nella descrizione della natura del contenuto interno del nodulo. Un nodulo è considerato solido se la componente cistica occupa meno del 20% del volume totale [9]. Per la detection delle cisti sono stati seguiti i successivi step:

- Creazione maschera cisti
- Definizione contorni maschera
- Applicazione operatori morfologici
- Identificazione proprietà della regione
- Esclusione aree
- Calcolo dimensione cisti

Nello specifico la maschera delle cisti è stata la medesima impiegata in precedenza per il calcolo dell'ecogenicità. Uno step fondamentale per la detection è risultato essere l'utilizzo di operatori morfologici. Poichè le proprietà della regione sono state calcolate con un insieme 8-connesso è stato necessario identificare il profilo della ROI e legare tra di loro i pixel non connessi, rimuovere i pixel isolati, riempire i buchi presenti e applicare un apertura. In seguito sono state identificate le proprietà della regione, in particolare sono estratti come proprietà della regione il centroide, i rettangoli di selezione, i quali sono stati posizionati intorno alle componenti cistiche, e le aree dei rettangoli, indice della dimensione delle cisti stesse, come mostrato in Fig.42.

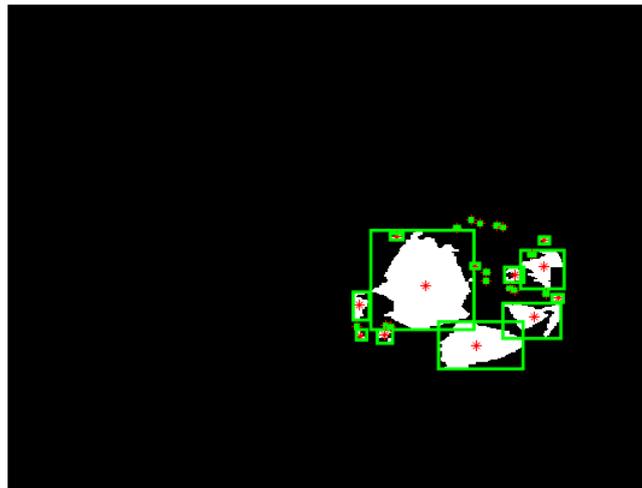


Figura 42: Esempio di maschera binaria con rettangoli di selezione e relativi centroidi per l'identificazione delle cisti

Infine la dimensione delle cisti è stata definita come la media delle aree considerate. Per verificare quale percentile avrebbe permesso di descrivere in maniera più opportuna

la distribuzione delle aree sono state condotte delle prove supplementari modificando il valore dei percentili. In particolare l'operazione della variazione delle aree da scartare è stata condotta per:

- *Percentili* [50 60 70 80]:

L'impiego dei percentili gioca un ruolo rilevante nel calcolo della dimensione delle cisti, poichè le dimensioni medie sono influenzate da quante aree presenti sono rilevate all'interno della maschera. Questo parametro cerca di dare una descrizione della distribuzione delle aree delle cisti.

Per la valutazione del percentile più accurato sono stati paragonati i p-value, provenienti dalla MANOVA. Tuttavia questi valori non si sono dimostrati rilevanti, per tutte le prove $p > 0.05$ e $d=0$, perciò per sancire il percentile più vantaggioso è stata svolta un'analisi di correlazione. La correlazione più conveniente è stata ottenuta utilizzando l'80° percentile con un valore pari a 0.411. Il tuning dei percentili, analizzato tramite MANOVA, ha visto come parametri utilizzati nel test statistico: area proveniente dal percentile corrente, deviazione standard, kurtosis e skewness calcolati rispetto ad ogni percentile presente nell'intervallo di confidenza, per un totale di tredici variabili per ogni percentile analizzato. La Fig.43 mostra il risultato del test tramite grafico a dispersione del percentile più idoneo. L'obiettivo del percentile è stato la descrizione della distribuzione delle aree delle cisti nel nodulo.

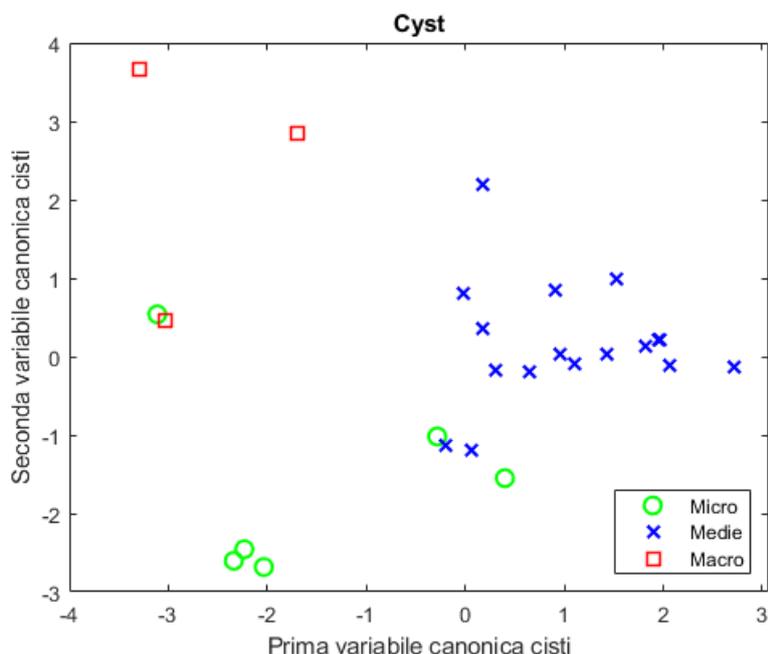


Figura 43: Scatter plot della prima e della seconda variabile canonica rappresentante la presenza di cisti nei noduli divisi nelle tre categorie: micro cisti, medie cisti e macro cisti

L'esiguo numero di noduli con cisti non è stato di sostegno al test ragion per cui il grafico soprastante non è risultato pienamente affidabile. L'aiuto al test è giunto dai valori di correlazione che sono risultati comunque solidi per ogni percentile impiegato, infatti il minore è derivato dal 50° percentile ed è stato pari a 0.32. Ciò nondimeno, per delineare una caratteristica vantaggiosa, atta alla quantificazione della caratteristica, sono

stati estratti i pesi della prima variabile canonica, in Fig.44. La variabile di principale importanza, anche dall'estrazione dei pesi, si è dimostrata la media, evidenziata in Fig.44.

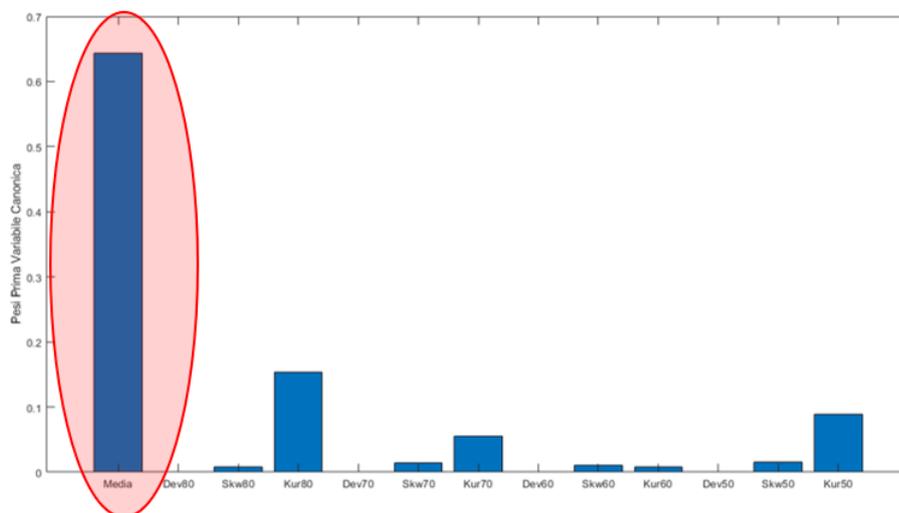


Figura 44: Istogramma dei pesi delle cisti. La variabile evidenziata è la media ed è la variabile di maggiore peso nella costruzione della prima variabile canonica

Questa variabile, rappresentativa dell'area media delle cisti, ha influenzato largamente l'esito del test statistico. Lo sbilanciamento dei pesi, indice di una robustezza della variabile, è stata una conferma: la media delle aree delle cisti è stata accettata come una variabile idonea alla descrizione dell'andamento della presenza di cisti.

2.5.5 Forma

Si definisce forma del nodulo la proporzione tra la sua altezza e la sua larghezza. Il riconoscimento e la valutazione della forma si basa sui seguenti step:

- Identificazione baricentro
- Creazione retta iniziale
- Creazione retta perpendicolare
- Calcolo rapporto

Il primo passo per l'estrazione della caratteristica forma è stato il calcolo del baricentro della maschera. Esso è estratto dalle proprietà della regione come centroide. In seguito sono state calcolate l'altezza e la larghezza del nodulo. Per individuare queste caratteristiche sono stati ricercati i pixel bianchi, costituenti la maschera, sfruttando le coordinate del baricentro. Nello specifico si sono contati i pixel bianchi della riga e della colonna passanti per il baricentro. Infine definite le grandezze è stata calcolata la forma come rapporto tra le due. Per il calcolo di questa caratteristica non si è tenuto conto dell'orientazione del nodulo, come mostrato in Fig.45.

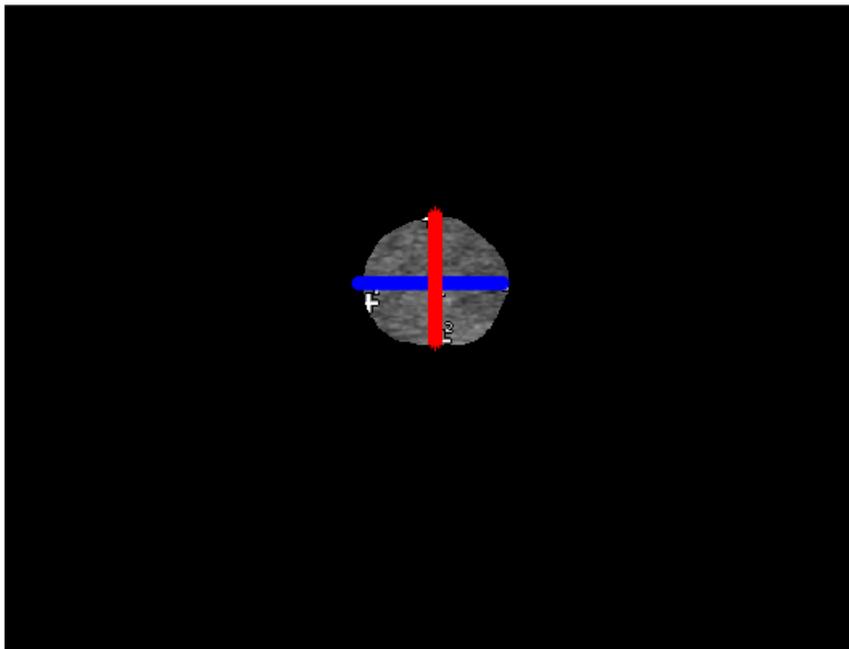


Figura 45: Esempio di posizionamento rette. La retta rossa è la retta originale, la blu la retta perpendicolare. L'intersezione tra le due è il baricentro

La caratteristica forma è una delle caratteristiche di minor costo computazionale grazie alla sua semplicità di calcolo. Il valore di correlazione derivante è pari a 0.407 perciò non è stato ritenuto necessario condurre delle ulteriori analisi statistiche. Il valore vantaggioso di correlazione ha dimostrato come è possibile associare, in maniera accurata e semplice,

alla forma di un nodulo il grado di malignità. Nonostante il valore proficuo di correlazione la quantificazione della forma non ha tenuto conto di un fattore ovvero l'orientazione del nodulo. Tuttavia sono state condotte numerose prove calcolando l'orientazione con diverse metodologie, per esempio derivarla come proprietà della regione, e non sono stati notati miglioramenti in termini di correlazione. Infine è stata adottata come caratteristica per descrivere la forma il rapporto tra altezza e larghezza come descritte in precedenza.

2.5.6 Vascolarizzazione

La distribuzione dei vasi ematici di un organo o di un apparato è stata indagata tramite l'impiego di immagini doppler. Per il calcolo dell'indice di frazione vascolare si sono adottati i seguenti step:

- Caricamento maschera
- Creazione maschere periferica e intranodulare
- Creazione immagini con solo segnale doppler
- Calcolo indici di frazione vascolare

Le maschere vascolari sono state disegnate per ogni immagine presente nel nuovo database. In particolare le maschere sono state disegnate a mano sulle immagini doppler. E' stata tracciata inizialmente solo la maschera esterna. Per ottenere la maschera intranodulare e periferica con la medesima forma ma con dimensioni differenti sono stati adottati degli operatori morfologici. E' stata inizialmente performata una dilatazione per inglobare la presenza di segnale vascolare periferico esterno al nodulo e successivamente una erosione per creare le maschere. La Fig.46 mostra chiaramente come la maschera intranodulare combacia perfettamente con la porzione nera all'interno della maschera periferica. Infine le nuove maschere create sono state salvate in una struttura e utilizzate per il calcolo dell'indice di area di frazione vascolare, in Fig.46.

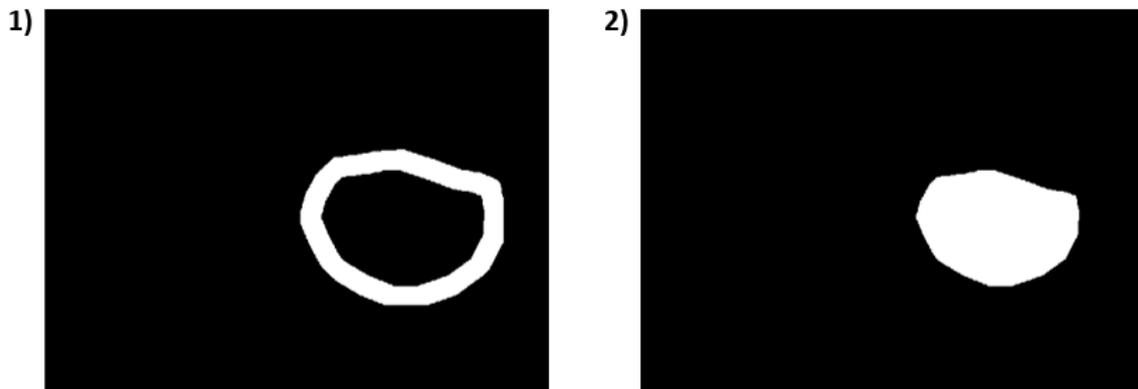


Figura 46: Esempio di maschere binarie create (1)Maschera periferica(2)Maschera intranodulare

Il primo passo per la creazione delle immagini con solo segnale doppler è stata l'identificazione dei pixel in toni di grigio nelle ROI d'interesse, ai quali è stato assegnato un valore nullo. Così facendo sono state create delle nuove immagini con sfondo nero e solo segnale doppler, come illustrato in Fig.47.

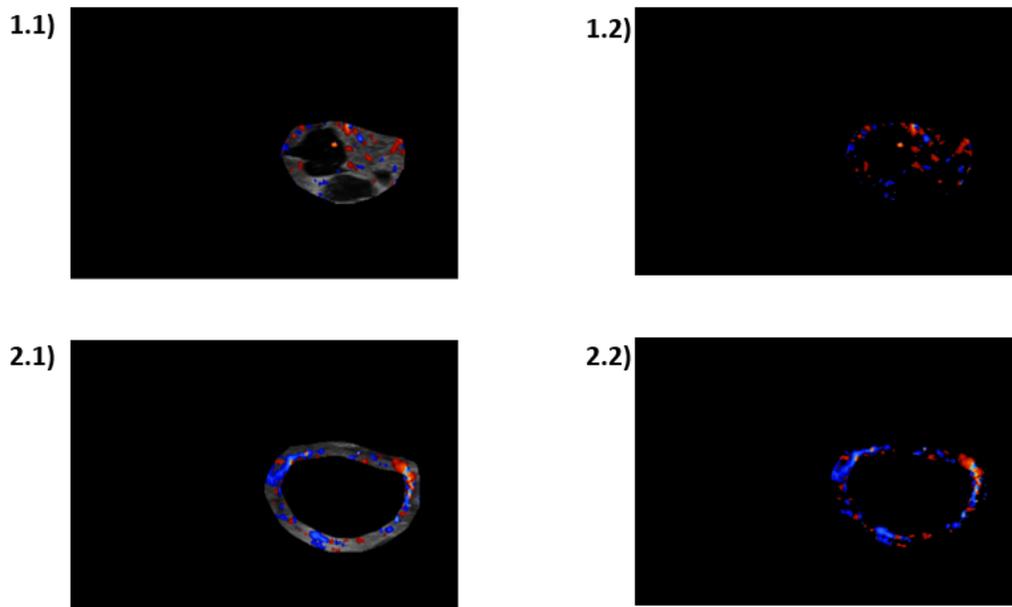


Figura 47: Esempio di utilizzo delle maschere create per identifica regione di interesse (o ROI) (1.1) ROI intranodulare,(1.2)ROI con la sola presenza di segnale doppler intranodulare(2.1)ROI periferica,(2.2)ROI con la sola presenza di segnale doppler periferico

Il calcolo degli indici di frazione vascolare è stato fatto tramite le maschere periferiche e intranodulari. Nello specifico è stata selezionata inizialmente la maschera periferica la quale è stata poi impiegata sia per l'immagine originale e sia l'immagine avente solo segnale. Infine il calcolo della frazione vascolare periferica è avvenuto tramite rapporto tra la lunghezza dei pixel della ROI in cui vi era solo segnale doppler e l'area intesa come lunghezza dei pixel della ROI stessa. Il calcolo della frazione di area vascolare intranodulare è stato analogo al precedente, ma invece dell'impiego della maschera periferica è stata utilizzata la maschera intranodulare. L'ottimizzazione dei parametri morfologici per la creazione delle maschere e quindi per riscontrare maggior quantità di segnale è stata condotta testando diversi parametri per quanto concerne la dilatazione e l'erosione. In particolare è stata variata la dimensione della maschera tramite un'azione morfologica, cioè sono stati provati diversi valori di *ratioDilate* e di *ratioPeriphery*, in particolare:

1. *ratioDilate* [0.05 : 0.2]: Il *ratioDilate* è definita come la percentuale da moltiplicare al raggio della regione, il quale è stato calcolato come raggio equivalente. L'operazione finisce la grandezza dell'elemento strutturale, in questo caso è stato utilizzato un disco, con il quale effettuare la dilatazione. Questo parametro influisce perciò sulla quantità di segnale periferico presente. Il parametro è stato variato nell'intervallo di confidenza a passi di 0.03.
2. *ratioPeriphery* [0.1 : 0.4]: Il *ratioPeriphery* è definita come la percentuale da moltiplicare al raggio della regione, il quale è stato calcolato anch'esso come raggio equivalente. Questo moltiplicazione definisce la grandezza dell'elemento strutturale con il quale effettuare l'erosione per creare la maschera intranodulare. Il parametro è stato variato nell'intervallo di confidenza a passi di 0.1 .

La combinazione dei parametri migliore è stata scelta in base alla misura di correlazione maggiore. In particolare sono stati valutate le correlazione per entrambe le maschere,

intranodulare e periferica, e sono stati riscontrati rispettivamente i valori di 0.2643 e 0.2034 , i più alti paragonati alle altre prove. La combinazione che ha generato i valori di correlazione è stata ratioDilate pari a 0.05 e ratioPeriphery pari a 0.4. Una volta definiti i parametri appropriati sono state create le maschere. In seguito alla dilatazione è stata eseguita un'erosione. La maschera intranodulare è derivata dal risultato dell'erosione. Successivamente per la creazione della maschera periferica sono state combinate il negativo della maschera intranodulare con la maschera proveniente dalla dilatazione. La seguente Fig.48 mostra gli effetti della dilatazione e dell'erosione rispetto alla maschera originale:

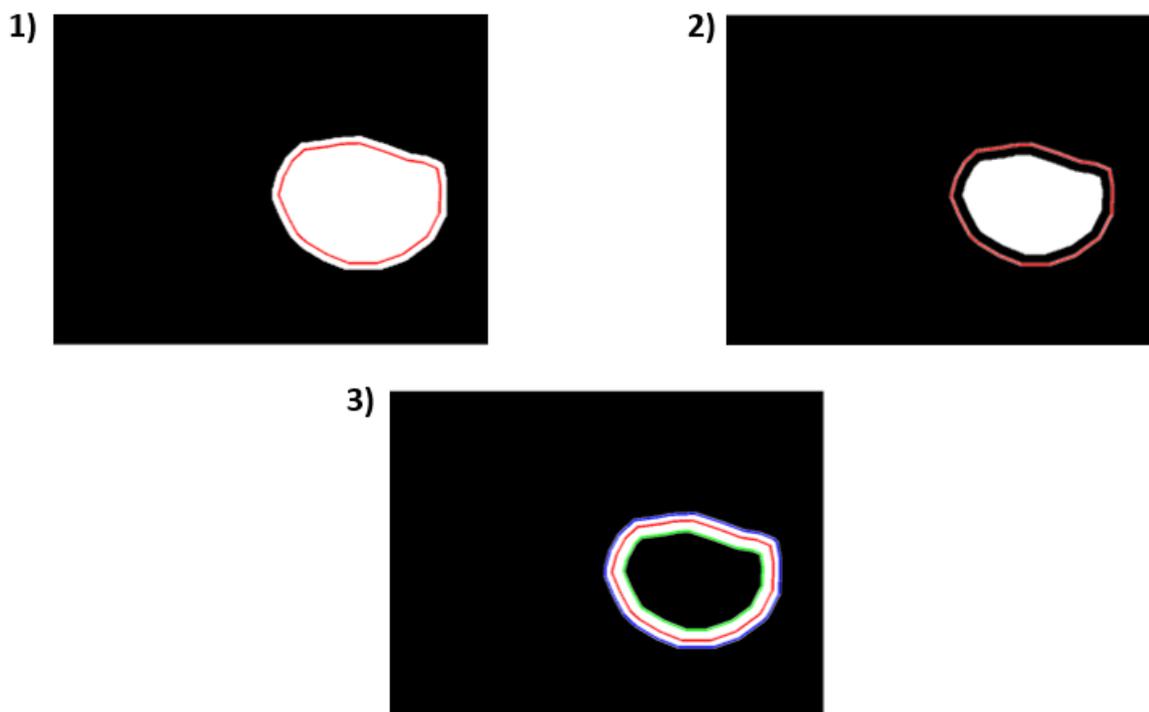


Figura 48: Esempio di applicazione di operatori morfologici, in rosso si evidenzia il contorno della maschera iniziale (1)Maschera derivante dalla dilatazione,(2)Maschera derivante dall'erosione(3)Maschera originale in rosso, effetto della dilatazione in blu, effetto dell'erosione in verde

Al fine di identificare la variabile maggiormente descrittiva per la vascolarizzazione è stata implementata la MANOVA. In questo singolare caso è stato fruttuoso unire due classi: vascolarizzazione intranodulare e vascolarizzazione peri-intranodulare. La nuova classe creata è stata definita con vascolarizzazione in entrambe le regioni. Questa unione è stata resa possibile poichè non sono state riscontrate immagini in cui la vascolarizzazione intranodulare fosse sprovvista di periferica. In seguito all'assunzione di sole due classi è stata focalizzata l'attenzione sul test statistico, in Fig.49 il risultato grafico. La fusione di due classi ha enfatizzato il risultato del test.

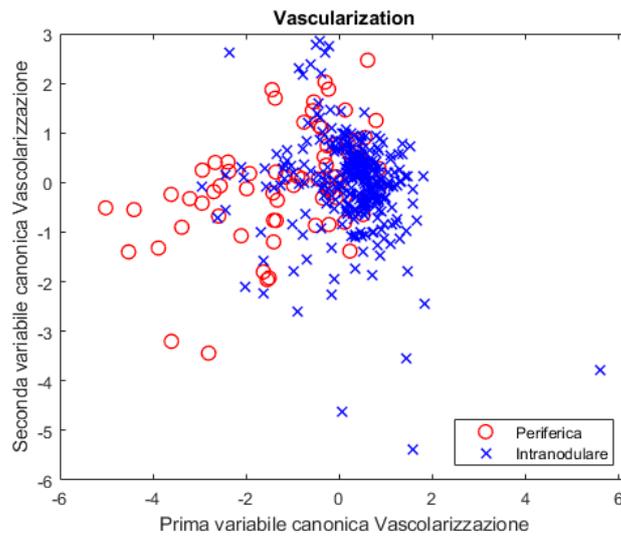


Figura 49: Scatter plot della prima e della seconda variabile canonica che rappresentano le classi unite per l'andamento di noduli vascolarizzati in zona periferica e intranodulare

In concomitanza ai valori di correlazione, si sono dimostrati più opportuni i parametri assunti in precedenza. Il test statistico migliore ha restituito come valori $d=1$ con un $p - value = 7.97 * 10^{-13}$. La variabile che maggiormente influisce sulla definizione della prima variabile canonica è stato l'indice di frazione vascolare intranodulare. L'azione dominante è stata quindi introdotta da questo indice ragion per cui è stato possibile impiegarlo nella valutazione della malignità di un nodulo.

2.5.7 Elastosonografia

Per valutare l'elasticità di un tessuto sono state utilizzate immagini elastosonografiche. Il calcolo della rigidità si compone dei seguenti punti:

- Creazione vettore COLOR
- Caricamento maschere elastosonografiche
- Calcolo distanza coseno di ogni pixel dal vettore COLOR
- Identificazione della posizione della distanza
- Calcolo della rigidità

La Fig.50 mostra come avvengono i passi per giungere al calcolo della rigidità. Le maschere elastosonografiche sono state disegnate per ogni immagine presente nel nuovo database. In particolare le maschere sono state tracciate manualmente sulle immagini elastosonografiche, identificando i contorni del nodulo. L'operazione primaria per il calcolo della "stiffness" di un tessuto è stata la creazione del vettore COLOR, un vettore simile al color bar. Nel vettore COLOR è stata inserita una scala a colori, in codifica RGB, la quale vedeva come punto di partenza il colore rosso andando a sfumare gradualmente fino al colore blu. L'identificazione della ROI è stata attuata tramite il caricamento delle maschere elastosonografiche. Sono stati indagati singolarmente tutti i pixel appartenenti alla ROI e ne è stata calcolata la distanza da ogni singola tripletta RGB presente nel vettore COLOR. Successivamente è stata selezionata la distanza minore per ogni pixel identificandone la posizione. Infine il calcolo della rigidità è stato condotto sulle posizioni, poichè la posizione rappresenta il valore più prossimo ad un determinato colore presente nel vettore COLOR. In altre parole è stato ricercato nel vettore COLOR il valore del pixel indagato, trasducendo così una misura di posizione in una codifica a colore. Successivamente sono state mediate tutte le posizioni rappresentanti le minime distanze coseno e sono state poi normalizzate rispetto alla lunghezza del vettore COLOR; è stata così definita una misura di rigidità media del nodulo in questione.

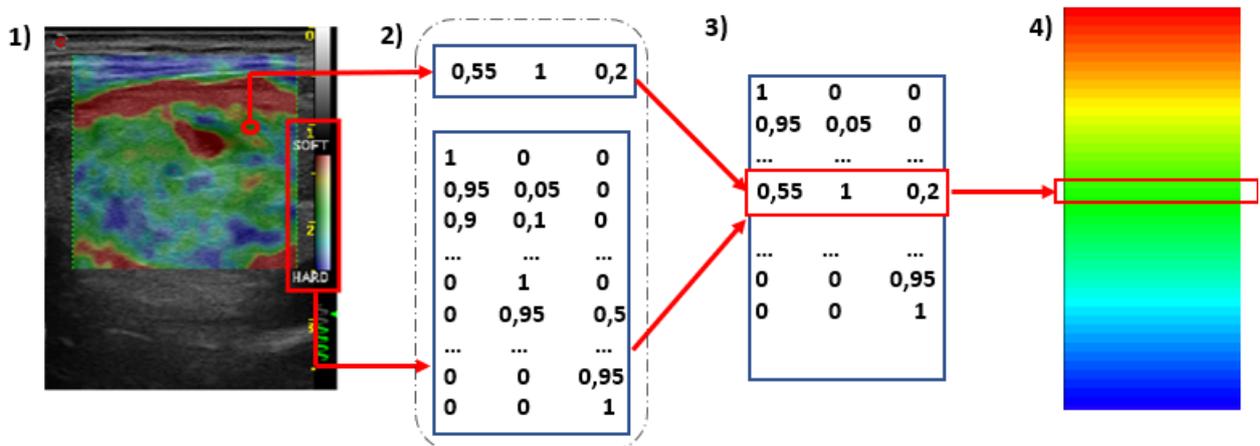


Figura 50: Esempio di calcolo dell'elasticità (1) Immagine elastosonografica con vettore COLOR (colorbar) e con pixel indagato nei riquadri rossi (2) Rappresentazione numerica del pixel e del vettore COLOR (3) Ricerca pixel con distanza minima nel vettore COLOR (4) Associazione distanza ad una codifica a colori nel vettore COLOR

Il tuning dei parametri per questa caratteristica ha riguardato esclusivamente la metodologia con cui calcolare la distanza. In particolare sono state utilizzate :

- *Distanza euclidea:*

E' la misura del segmento che ha come estremi i due punti in questione. Con questa distanza lo spazio euclideo si trasforma in uno spazio metrico.

- *Distanza chebichev:*

Nota come la distanza della scacchiera, è una misura su spazi vettoriali, cioè adotta il massimo valore della differenza lungo gli assi come distanza.

- *Distanza minkowski:*

E' una generalizzazione della distanza euclidea in uno spazio euclideo.

- *Distanza mahalanobis:*

E' una distanza per valutare quanto sia simile uno spazio campionario incognito rispetto ad uno spazio noto. Si basa sulla misura della correlazione dell'insieme dei dati, con la quale è possibile analizzare i differenti pattern.

- *Distanza coseno:*

La distanza coseno è la misura dell'angolo compreso tra i due vettori rispetto all'origine: questa distanza è sensibile alla rotazione dei vettori, quindi alla distanza che li divide e non alla loro lontananza rispetto all'origine degli assi.

La metodologia con cui valutare la distanza ha influenzato tutto la quantificazione della caratteristica, poichè la distanza è poi tradotta in una codifica a colore bisogna valutare la migliore in termini di rappresentazione grafica cioè la più simile all'elastsonografia. Al fine di valutare la distanza più idonea è stato impiegata nuovamente la MANOVA per ogni distanza utilizzata. Il grafico seguente, Fig.51, mostra la dispersione dei valori di elasticità.

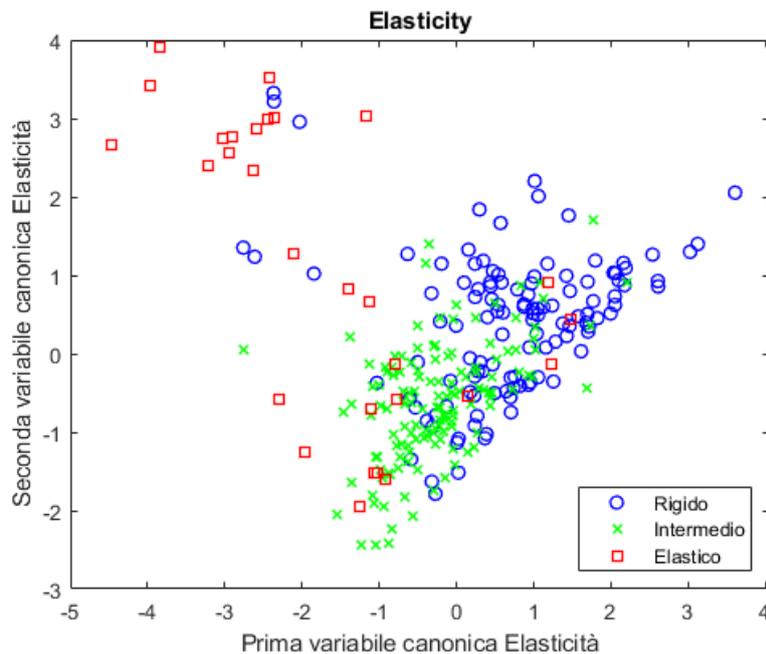


Figura 51: Scatter plot della prima e della seconda variabile canonica rappresentante l'elasticità dei un nodulo divisi nelle tre classi elastico, intermedio e rigido

Sono state calcolati ulteriori parametri per implementare la MANOVA quali: elasticità, varianza, mediana, skewness e kurtosis. La distanza migliore è stata scelta ricercando i valori correlazione maggiore. In particolare il miglior valore di correlazione è pari a -0.455 e comporta due valori di p-value pari rispettivamente a $1.6 * 10^{-36}$, $9.42 * 10^{-17}$ e sono ottenuti utilizzando la distanza coseno. Le distanze provate si sono comunque dimostrare valide poichè hanno generato valori accettabili, sia di correlazione che di p-value, tuttavia la distanza coseno è risultata la più accurata nella rappresentazione dell'immagine stessa, ovvero tutti i valori di distanza calcolati sono stati in seguito visualizzati e hanno riprodotto fedelmente la mappa a colori introdotta dall'elastsonografia. Alla correlazione è stato associato un valore negativo a causa della codifica che è stata attribuita alla quantificazione dell'elasticità poichè un livello alto di elasticità non sottintende un nodulo elastico bensì il contrario, cioè è stata una rappresentazione di un nodulo rigido. Il significato implicito è stato quello di attribuire ad un valore alto di elasticità la malignità poichè è stato inteso come punto di partenza per la benignità un valore di elasticità pari a zero.

2.6 GUI

L'interfaccia grafica (GUI, Graphical User Interface), mostrata in Fig.52, è una tipologia di interfaccia utente che consente l'interazione uomo-macchina in modo visuale utilizzando rappresentazioni grafiche piuttosto che utilizzando una interfaccia a riga di comando. Lo scopo della GUI è quello di consentire all'utente di interagire con il calcolatore manipolando graficamente le immagini in maniera semplificata e diretta, evitando l'esecuzione dei comandi guidata da istruzioni o dei comandi impartiti tramite tastiera. La praticità di utilizzo influenza notevolmente l'impiego della GUI, poichè consente all'utente di compiere gli step elencati precedentemente per ogni singola immagine selezionata. Pertanto il costo computazionale e il tempo di diagnosi automatica diminuiscono drasticamente.

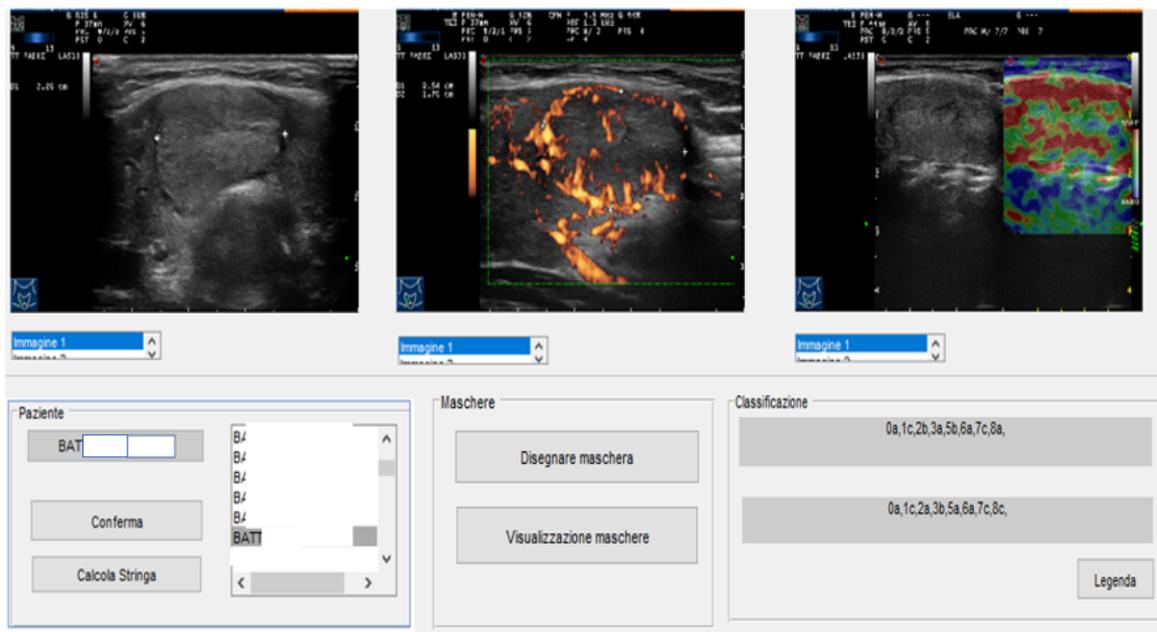


Figura 52: Esempio di utilizzo dell'interfaccia grafica, GUI

2.6.1 Struttura Graphic User Interface

Questa GUI utilizza una combinazione di tecnologie e dispositivi per fornire una piattaforma con cui gli utenti possono interagire, per le attività di raccolta e produzione di informazioni. Nello specifico la GUI si compone di:

1. Pannello Paziente, in Fig.53. Questa sezione di interfaccia grafica è dedicata alla scelta del paziente. Non si limita solo alla scelta del paziente, ma anche alla conferma e all'applicazione degli step per la quantificazione delle caratteristiche ultrasonografiche. I pulsanti che lo compongono sono:
 - Pulsante Conferma
 - Pulsante Calcola Stringa
 - L'elenco dei pazienti
 - Riquadro paziente
2. Pannello Immagini, in Fig.53. Questa sezione di interfaccia grafica è dedicata alla visualizzazione delle immagini. In particolare sono visualizzate le immagini B-mode,

vascolari e elastosonografiche di ogni paziente. Nel caso non fosse presente l'immagine relativa ad un dato paziente compare il messaggio 'Image not found' nel riquadro dove l'immagine è assente. Nel momento in cui non è ancora stato selezionato nessun paziente i riquadri presentano tutti il medesimo messaggio, come mostrato in Fig.53. Il pannello è fornito di una lista delle immagini, nel caso un paziente ne avesse a disposizione un vasto numero, permettendo così la possibilità di scelta dell'immagine all'utente, come mostrato in Fig.54.

3. Pannello Maschere, in Fig.53. Questo pannello è destinato alla creazione e alla gestione delle maschere adottate nell'algoritmo. In questa particolare parte risiede la possibilità di personalizzare le maschere da parte dell'utente. Il pannello si compone di due pulsanti:

- Pulsante Disegnare maschera
- Pulsante Visualizzazione maschere

4. Pannello Classificazione, in Fig.53. Questo particolare pannello fornisce esclusivamente informazioni di classificazione. In particolare si compone di due riquadri che mostrano rispettivamente la classificazione corretta, ogni qualvolta si sceglie un paziente, e la classificazione automatica, ogni qualvolta si richiede di calcolarla con l'apposito pulsante. Le due stringhe riportate sono state modificate: ovvero si ridimensiona la stringa corretta mostrando solo le prime due cifre della classificazione di ogni caratteristica. La stringa automatica invece si compone esclusivamente di due cifre per ogni caratteristica. In questo pannello vi è unicamente il pulsante tabella che corre in aiuto all'interpretazione della stringa mostrata a video.

La possibilità di personalizzazione primaria è rappresentata dalla facoltà di scelta delle immagini da parte dell'utente, in base alla sua esperienza e alla sue competenze. La scelta si basa sull'immagine che meglio evidenzia la caratteristica da quantificare. L'esperienza dell'utente è quindi di essenziale rilievo, poichè un utente impreparato potrebbe selezionare delle immagini non ottimali per la decisione finale, creando maggior discrepanza tra la classificazione originale e quella automatica. Eventualmente l'utente ha anche la facoltà di integrare e rendere ancora più personale la classificazione finale, tracciando nuove maschere maggiormente rappresentative del nodulo in questione o della regione specifica che si vuole indagare.

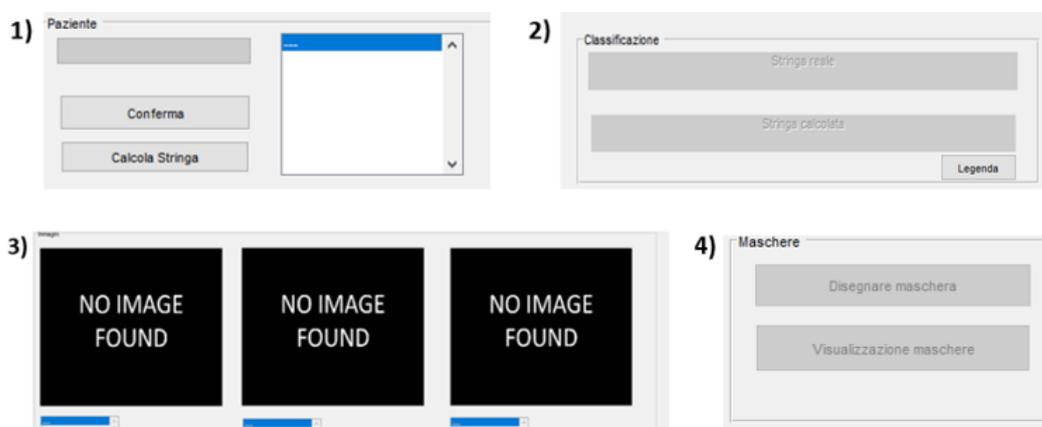


Figura 53: Esempio di pannelli di default non ancora in uso : (1) Pannello Paziente(2) Pannello Classificazione(3) Pannello Immagini(4) Pannello Maschere

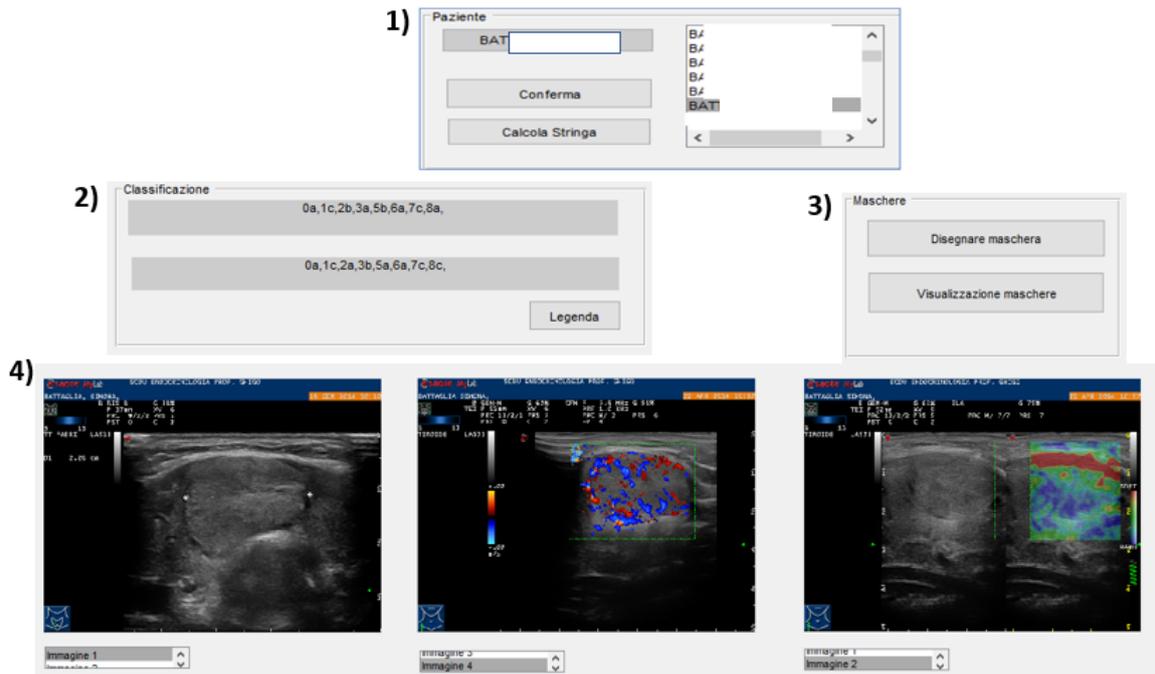


Figura 54: Esempio di pannelli in uso: (1) Pannello Paziente(2) Pannello Classificazione(3) Pannello Maschere(4) Pannello Immagine

2.6.2 Utilizzo Graphic User Interface

L'utilizzo ottimale della GUI prevede di seguire in maniera sequenziale i comandi elencati seguentemente, illustrati in Fig.55:

- Selezionare un paziente
- Confermare la scelta
- Selezionare le immagini migliori
- Calcolare la Stringa finale



Figura 55: Guida grafica per il corretto uso della GUI (1)Selezione Paziente, (2)Conferma, (3)Immagini migliori, (4)Calcolo stringa

Ogni pulsante utilizza una particolare algoritmo legato a un determinato comando. Tutti i pulsanti e tutte le finestre immagine sono collegati tra loro. Nello specifico una volta modificato o pigiato un pulsante verrà modificata l'informazione ad esso associata, cambiando così l'informazione inviata ai pulsanti collegati. L'utilizzo della GUI può essere influenzato dall'operatore. L'esperienza dell'utente incide sulla decisione finale automatica, la quale potrebbe essere migliorativa o peggiorativa, in base alle differenze che l'utente introduce rispetto agli step che ha compiuto per la quantificazione delle caratteristiche. Come descritto in precedenza la GUI ha carattere semiautomatico, quindi l'utilizzo ottimale prevede che l'operatore intervenga il meno possibile con l'algoritmo. Ciò non è sempre possibile in quanto il numeroso set di immagini e l'idea stessa che la GUI creata possa essere utilizzata impiegando un nuovo dataset, richiede l'intervento attivo da parte dell'utente al fine di colmare la mancanza di dati, cioè bisogna disegnare le maschere che non sono presenti. In altre parole più l'interferenza dell'utente è presente maggiormente i risultati finali risentono di tale intervento.

2.6.3 Caratterizzazione Graphic User Interface

L'algoritmo adottato dalla GUI è semiautomatico, cioè necessita di un intervento attivo da parte dell'operatore quando si verificano condizione di sbarramento al calcolo della stringa automatica. Ogni tasto sottintende l'esecuzione di uno specifico comando. I pulsanti sono raggruppati nei pannelli descritti in precedenza, i quali conferiscono praticità e forma alla GUI. Questi pannelli hanno un obiettivo ben delineato e preciso, permettendo solo le operazione conformi ad esse. In particolare ogni pulsante adempie ad uno specifico compito e gioca un ruolo di connessione con gli altri pulsanti. Lo scopo e la porzione di algoritmo eseguita da ogni pulsante è descritta in seguito:

- Lista pazienti. E' una listbox: ha il compito di listare in ordine alfabetico tutti i pazienti presenti nel database. Comunica con il riquadro paziente.
- Riquadro paziente. E' un edit-text: ha il compito di rappresentare il nome del paziente selezionato dalla lista pazienti. Nel caso si conoscesse già il nome del paziente

è possibile anche inserirlo manualmente all'interno di esso. Comunica con il tasto Conferma.

- Pulsante conferma. E' un push button: questo pulsante ricopre un ruolo centrale. Ha il compito di confermare il paziente selezionato. Una volta premuto si lega con il pannello immagine, a cui riferisce il nome del paziente e la posizione dello stesso nel database. Comunica anche con il riquadro classificazione reale a cui fornisce la corretta classificazione, la quale è abbreviata rispetto all'originale presente nel database, da visualizzare.
- Riquadro immagine. E' la rappresentazione di due assi: ha il compito di gestire e visualizzare le immagini paziente presenti nel database. Comunica con il pulsante conferma e il pulsante calcola stringa, con i quali scambia informazioni sulle immagini correnti da rappresentare e utilizzare per i calcoli. Questa interfaccia è dotata di tre riquadri immagine, ognuno gestisce una tipologia differente di immagine, cioè B-mode, vascolarizzazione e elastosonografia, ma trasferisce le immagini nella medesima maniera.
- Lista immagine. E' una listbox: ha il compito di listare tutte le immagini presenti nel database. Comunica esclusivamente con il riquadro immagine ad essa dedicato e impone un cambio di immagine nel momento in cui è selezionata un'immagine differente dalla corrente.
- Pulsante disegna maschera. E' un push button: questo pulsante rafforza la semi-automaticità dell'algoritmo, cioè ha il compito di aprire l'immagine B-mode corrente e permettere all'utente di tracciare una maschera. Questa maschera è usata per l'estrazione delle caratteristiche del nodulo. Comunica con il pannello immagini, da cui prende l'immagine, con il tasto calcola stringa e con il tasto visualizzazione maschere, a entrambi fornisce la maschera.
- Pulsante visualizzazione maschere. E' un push button; ha il compito di mostrare tutte le maschere utilizzate per il calcolo delle caratteristiche. Comunica con il tasto calcola stringa dal quale recepisce le maschere.
- Pulsante calcola stringa. E' un push button, e rappresenta il cuore della GUI: ha il compito di calcolare la stringa automatica. Comunica con il pannello immagine dal quale riceve l'immagine per il calcolo della stringa e la visualizza nel pannello classificazione. La sezione semiautomatica dell'algoritmo è legata a questo pulsante, presentato in Fig.56. Il calcolo delle caratteristiche ultrasonografiche necessita di almeno una maschera, la quale non è sempre presente nel database. L'assenza della maschera comporterebbe l'interruzione dell'algoritmo, quindi si richiede all'utente di disegnare la maschera sull'immagine corrente nel momento in cui esse non è presente. La mancanza della maschera è palesata da un messaggio di warning, presentato in Fig.56; successivamente si apre una nuova finestra con l'immagine corrente sulla quale disegnare la maschera. L'algoritmo prevede la visualizzazione della maschera disegnata, seguita da una conferma della qualità della stessa. La conferma avviene tramite l'inserimento di una lettera Y/N (codifica Yes/No), come mostrato in Fig.56; l'algoritmo ripropone il disegno della maschera finchè non si raggiunge un risultato ottimale che l'operatore deve confermare. Il calcolo finale della stringa automatica avviene adottando tutti gli step descritti precedentemente. Dopo aver estratto tutte le caratteristiche automatiche del nodulo si crea la classificazione automatica

sfruttando delle soglie. Le soglie sono fisse e scelte tramite istogramma, riportati nel capitolo dei risultati e scelte in base alle medie e gli errori mostrati, delle variabile più significative della caratteristica in questione. Ad ogni valore numerico estratto dall'immagine si associa una codifica numerica e letterale, abbreviata come la classificazione reale, che simboleggia la natura del nodulo in esame. La stringa calcolata è rappresentata infine nel riquadro classificazione automatica.

- Riquadro classificazione reale. E' uno static text: ha il compito esclusivo di mostrare a video la classificazione reale, ma abbreviata, fornita nel database. Comunica con il pulsante conferma dal quale recepisce l'informazione riguardante la stringa modificata.
- Riquadro classificazione automatica. E' uno static text: ha il medesimo compito del riquadro della classificazione reale. L'unica differenza tra i due è la provenienza dell'informazione: questo riquadro comunica con il pulsante calcola stringa e quindi mostra a video la stringa automatica.
- Pulsante legenda. E' un push button: corre in aiuto alla classificazione finale. Nello specifico presenta in figura una tabella in cui sono mostrate le caratteristiche del medico e le caratteristiche automatiche esplicitate.

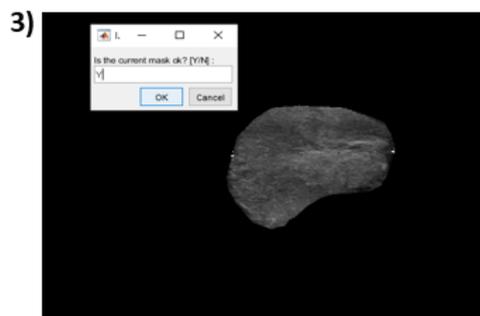
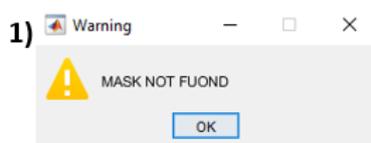


Figura 56: (1) Esempio di Warning box (2) Algoritmo Semiautomatico con intervento attivo: creazione maschera (3) Algoritmo Semiautomatico con intervento attivo: conferma qualità maschera creata

L'interfaccia grafica di cui sopra è stata creata rispetto al database di immagini fornito, in particolare adotta il database aggiornato, e anche il nuovo database formato dalle sole maschere per il calcolo della vascolarizzazione e dell'elasticità del nodulo. Tuttavia la GUI non prevede la possibilità di integrare un ulteriore dataset di immagini, ma se si volesse validare la capacità di classificazione automatica trattata con un altro dataset bisognerebbe intervenire manualmente nella sezione dedicata al caricamento del database aggiornandone esclusivamente il nome. I risultati provenienti dall'algoritmo possono quindi variare, ma l'utilizzo della GUI rimane il medesimo; in particolare sostituendo il

database si va in contro ad un aumento del costo computazionale e in termini di tempo poichè non essendoci le maschere già disegnate e preimpostate l'interfaccia prevede il tracciamento delle nuove maschere. Questo fattore grava maggiormente sulla capacità e sull'esperienza dell'utente di identificare correttamente il nodulo. E' severamente sconsigliato l'uso della GUI da parte di un utente inesperto, in quanto potrebbe vanificare l'efficacia, le performance dell'utilizzo dell'interfaccia sono le medesime riportate nel capitolo successivo.

3 Validazione e Risultati

I risultati grafici e numerici sono riportati di seguito e sono derivanti da tutte le caratteristiche analizzate singolarmente per tutti i 410 pazienti presente nel database. Si sono rappresentati unicamente i risultati successivi al tuning dei parametri, quindi con i parametri ottimizzati, ripercorrendo tutti gli step elencati nel capitolo 2 per ogni caratteristica. Si sono così ottenuti i risultati finali per tutti i pazienti valutando tutte la quantificazione di tutte le caratteristiche ultrasonografiche. Per alcune si è preferito mostrare anche la prima variabile canonica derivante dalla MANOVA, poichè portatrice di ulteriori informazioni, a volte più espressive, che non sempre sono emerse dalla caratteristica analizzata. Le variabili canoniche sono delle combinazioni lineari realizzate in modo da massimizzare la varianza tra le varie suddivisioni. Si sono raccolti questi risultati grazie anche al numero sufficiente di pazienti presenti nel dataset, cioè si è potuto condurre un'analisi statistica che ne dimostri la solidità.

3.1 Ecogenicità

Per quanto concerne l'ecogenicità si sono calcolati tre valori: ecogenicità (Echo), intesa come media delle intensità dei pixel costituenti il nodulo, ecogenicità rispetto alla tiroide (Echo-thy), intesa come media delle intensità del nodulo rapportata all'intensità dei pixel costituenti la tiroide, ed infine ecogenicità rispetto al muscolo, intesa come media delle intensità del nodulo rapportata all'intensità dei pixel costituenti il muscolo (Echo-mus). L'ecogenicità è una caratteristica altamente specifica rispetto alla natura del nodulo, infatti è il primo parametro quantitativo che il medico valuta durante un esame ecografico. Come ulteriore conferma per la scelta della caratteristica di efficacia superiore si è analizzato graficamente l'andamento di questi valori tramite box plot, mostrati in Fig.57:

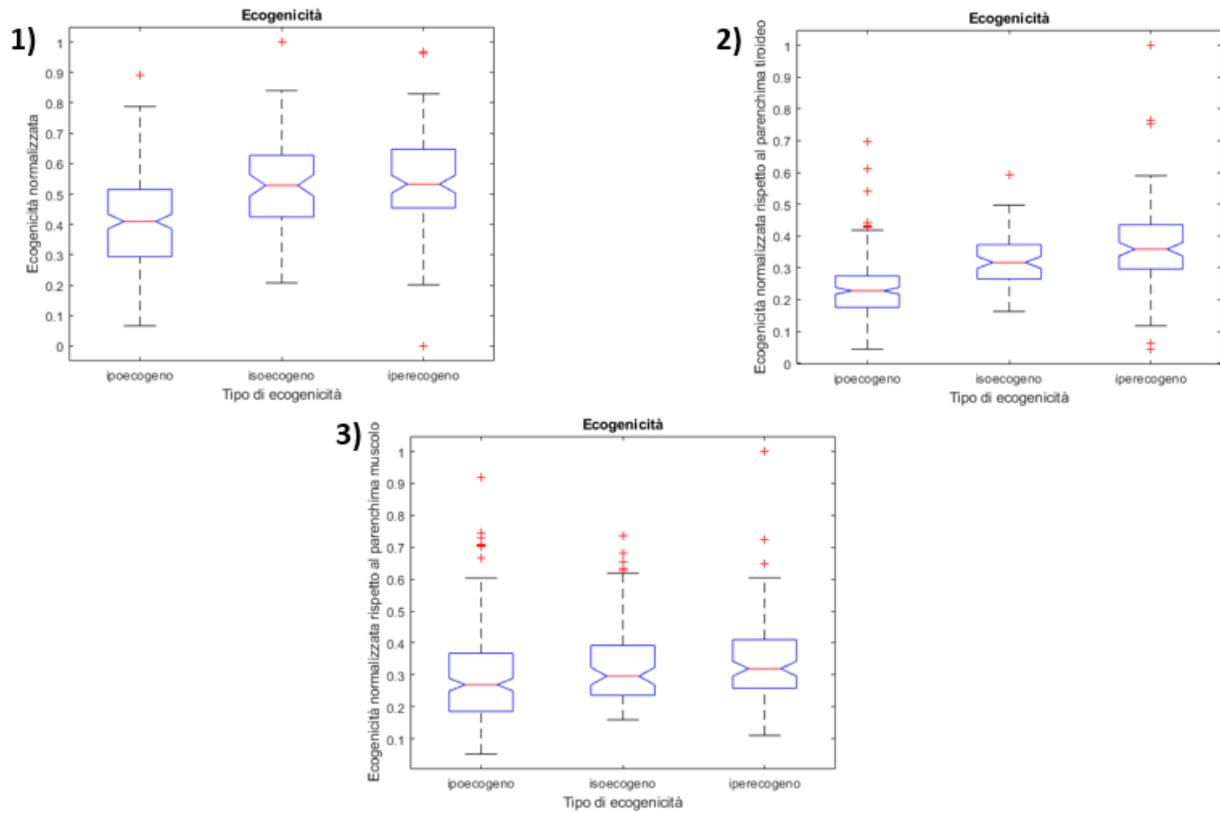


Figura 57: Box plot dei valori di (1)Ecogenicità (2)Ecogenicità rispetto alla tiroide(3)Ecogenicità rispetto al muscolo

Si nota dalla Fig.57 che gli intervalli notch dei boxplot che raffigurano i valori medi di Echo-thy non presentano sovrapposizioni ovvero questi valori sono statisticamente significativi e con valori medi ben distinti. Differentemente gli intervalli notch che rappresentano i valori di Echo ed Echo-mus si sovrappongono e perciò non rappresentano valori statisticamente significativi. Ci si è serviti di una supplementare validazione per confermare nuovamente che la Echo-thy incalza il ruolo di caratteristica più rappresentativa dell'ecogenicità, nonostante non fosse necessario poiché presenta un valore di correlazione pari a 0.49 . La Echo-thy si presenta come un descrittore affidabile, ciò rende possibile l'accostamento di un valore numerico a una grandezza anatomica del nodulo per identificarne in modo più o meno accurato la natura. Si è riportato di seguito l'istogramma che rappresenta le media dei valori di ecogenicità raggruppate per suddivisioni, Fig.58.

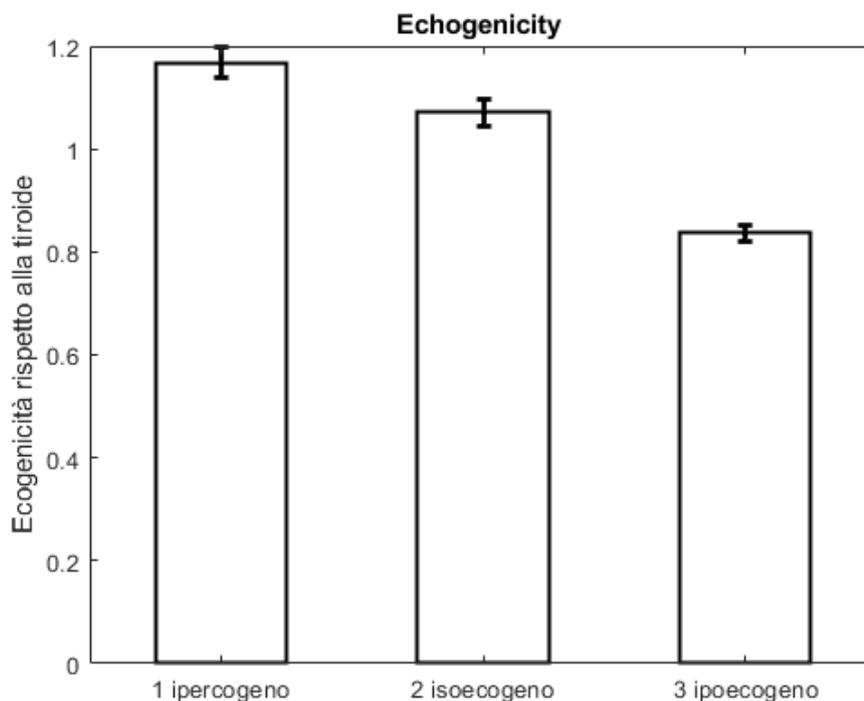


Figura 58: Istogramma che riporta medie ed errori standard dell'Echo-thy raggruppati per classi

Le due barre verticali mostrano il valore dell'errore standard il quale aiuta notevolmente l'interpretazione dei risultati, e assume valore pari a 0.03, 0.026 e 0.016 per le classi rispettivamente iper-,iso- e ipo-ecogeno. Le medie attorno le quali si creano i raggruppamenti valgono 1.17,1.07 e 0.83 rispettivamente per iper-,iso- e ipo-ecogeno. Normalmente un abbassamento del valore di ecogenicità si associa ad un aumento del rischio di malignità, come è riportato ampiamente in letteratura e in [14]. Tuttavia questa caratteristica non comporta una verità assoluta, ovvero è altamente specifica ma se valutata singolarmente la sua efficacia diminuisce notevolmente. Il valore di correlazione dimostra però come la Echo-thy sia relazionata in buon grado alle classi di rappresentazione. Una divisione netta delle classi permetterebbe di abbinare la caratteristica alla natura del nodulo con una specificità superiore. Le tre suddivisioni hanno uno specifico valore diagnostico, vale a dire che l'ipoecogenicità è la rappresentazione specifica per eccellenza di malignità, l'iso- e ipo-ecogenicità invece sono sintomi di benignità e comportano il medesimo livello diagnostico. Questa valutazione è rafforzata dal valore dell'errore standard rappresentato in Fig.58, poichè un errore maggiore nelle caratteristiche associate alla natura benigna del nodulo non scuote le fondamenta della diagnosi. Nonostante ciò l'errore minore riscontrabile si nota nell'Echo-thy ipoecogena, questa indicazione si è mostrata fortificante siccome un errore minimo è indice di un'attribuzione numerica appropriata. Si evince che le medie riportate sono ben distinte e quindi rappresentative dell'insieme delle caratteristiche. La valutazione che il medico fornisce dell'ecogenicità si accosta fedelmente alla quantificazione attribuitagli. Come descritto in [14] l'incidenza della malignità è solo del 4% per lesioni iperecogene; questa percentuale sale al 26% per lesioni isoecogene. In pratica la Echo-thy è una caratteristica ultrasonografica molto utile e specifica per differenziare i noduli maligni dai benigni, presentando una quantificazione inversamente proporzionale al grado di malignità.

3.2 Struttura

Per quanto riguarda l'eterogeneità della struttura del nodulo i risultati si sono mostrati fortemente caratteristici. Si sono calcolate quindici variabili totali e, in seguito al tuning del filtro mediano, dei livelli di grigio e all'identificazione della coppia di parametri ottimali, si è passati a una rappresentazione canonica delle variabili di interesse. In particolare, si è rappresentata tramite istogramma la prima variabile canonica, raggruppata per classi (Fig.59) la quale nasce dalla MANOVA, che ha generato un $p - value = 8 * 10^{-15}$, $d = 1$, affiancata dalla variabile di maggior peso, cioè la RLN. Le variabili canoniche sono combinazioni lineari delle variabili originali, scelte per massimizzare la separazione tra gruppi. La combinazione lineare più significativa del test è sempre la prima.

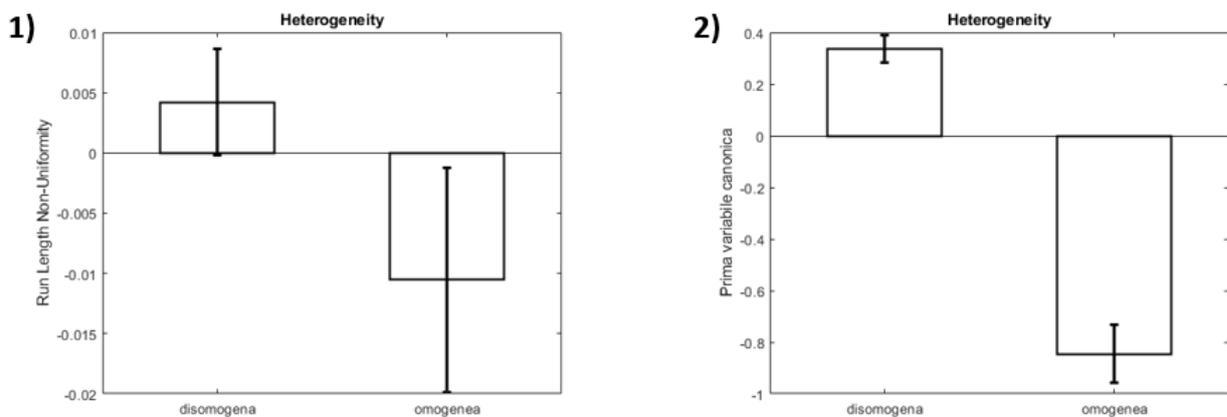


Figura 59: Istogrammi che riportano medie ed errori standard della struttura raggruppati per classi(1) Istogramma della variabile più rappresentativa (2)Istogramma della prima variabile canonica

Si nota dai grafici a barre che come la prima variabile canonica risulti più descrittiva sia in termini di medie, che assumono valori pari a 0.33 e -0.85 rispettivamente per disomogeneo e omogeneo, sia in termini di errori, che assumono valori pari 0.05 e 0.11 con il medesimo accostamento, quindi si è preferito adottare come soggetto dei risultati la prima variabile canonica. La correlazione proveniente dalla variabile in esame si è dimostrata superiore alle altre, pari a 0.47, e quindi adottare la prima variabile canonica come descrittore primario si traduce in una rappresentazione ottimale. Dalla rappresentazione della variabile canonica si nota chiaramente come noduli omogenei tendano maggiormente ad assumere mediamente valori negativi, in opposizione ai noduli disomogenei. La caratterizzazione dell'omogeneità mostra discreti risultati ed i valori assunti dalla prima variabile canonica sono mediamente differenti, nonostante ciò l'errore che riguarda l'omogeneità, pari a 0.11, si traduce in una grande dispersione dei valori della variabile, ovvero si ha grande difficoltà ad accomunare questi valori con la caratteristica in esame. L'errore della disomogeneità è minore, pari a 0.05, e implica un maggior raggruppamento delle variabili di interesse. L'eterogeneità di un nodulo non è una caratteristica altamente specifica, tuttavia una struttura disomogenea è spesso associata ad un aumento del rischio di malignità. Come dimostrato e riportato in [16] l'eterogeneità non è una caratteristica che si è evidenziata come una differenza significativa tra noduli benigni e maligni. La quantificazione di questa caratteristica non permette quindi di valutare la natura di un nodulo, ma se associata ad altre caratteristiche ecografiche può essere un utile e un consolidante contributo alla definizione del grado di malignità.

3.3 Margini

La classificazione presente nel database prevede la differenziazione dei margini in: lineare, corrugato, mammellonato e spicolato. Tuttavia solo la classe lineare, la quale definisce un margine regolare, è associata alla benignità, quindi si è preferito suddividere i margini in due classi: regolare e irregolare. L'assunzione fatta riscontra significato diagnostico poichè margini irregolari, che comprendono corrugato, mammellonato e spicolato, sono associati alla malignità. Per giungere ai risultati si è proceduto al calcolo delle proprietà globali della regione, dei descrittori di Fourier e delle dimensioni frattali. In seguito si è performata la MANOVA e si è definita la variabile di maggior rappresentazione; il test ha restituito un $p - value = 3.3 * 10^{-24}, d = 1$. Si sono riportati di seguito gli istogrammi rappresentanti la prima variabile canonica e il 6° descrittore di Fourier (Fig60) che mostrano le medie e gli errori per ogni suddivisione delle due variabili.

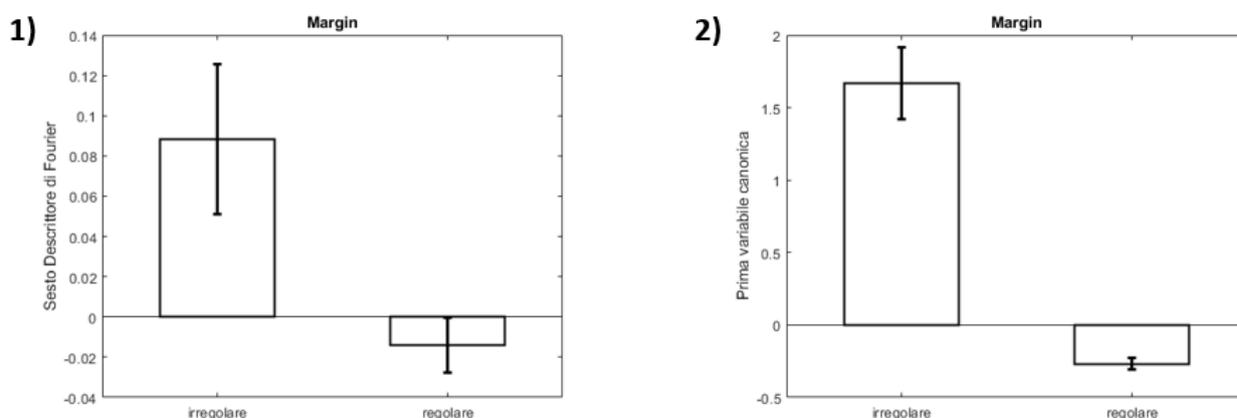


Figura 60: Istogrammi che riportano medie ed errori standard dei margini raggruppati per classi(1) Istogramma della variabile più rappresentativa (2)Istogramma della prima variabile canonica

Si evince dai grafici a barre che come la prima variabile canonica risulti più descrittiva sia in termini di medie, pari a 1.67 per gli irregolari e -0.27 per i regolari, sia di errori. Si nota come l'errore minore, pari a 0.04 per i regolari e pari 0.25 per gli irregolari, riesca a conferire maggior potere discriminatorio, ovvero grazie a ciò si riesce ad ottenere un buon raggruppamento dei margini con una discreta dispersione. I noduli con margini regolari sono soliti essere accostati a valori della variabile canonica minori rispetto agli irregolari. I margini sono una caratteristica molto specifica, quindi all'aumentare dell'irregolarità cresce anche il rischio di malignità, poichè un margine regolare racchiude un nodulo che non presenta caratteristiche infiltranti. Si è soliti attribuire, anche come si desume dai grafici (Fig.60), valori bassi alla regolarità dei margini. Il valore di correlazione della prima variabile canonica è pari a 0.55 e rappresenta una delle caratteristiche di maggior accuratezza in termini di rappresentazione delle classi. Non si sono riportati i grafici delle dimensioni frattali poichè non hanno influenzato notevolmente il risultato, tuttavia il Mann-Whitney test ha restituito un $p-value=0.044$, ai limiti dell'accettabilità. La quantificazione dei margini permette quindi una buona valutazione della natura del nodulo, tuttavia margini regolari non sono fonte di esclusione della malignità poichè anche in letteratura si sono riscontrati casi di noduli con margini regolari di natura maligna.

3.4 Cisti

Per quanto concerne le cisti si sono calcolate tredici variabili totali e, successivamente alla messa a punto dei parametri ottimali, si è giunti alla rappresentazione della caratteristica tramite istogrammi. La prima variabile canonica e l'area media delle cisti sono mostrate in Fig.61 tramite istogrammi. Come si evince dal grafico rappresentante la prima variabile canonica non sono presenti valori di macro cisti poichè il numero delle macro cisti presenti nel database è molto limitato. La MANOVA perde robustezza con questa caratteristica a causa della mancanza di dati a sufficienza, infatti $p\text{-value} = 0.097$, $p\text{-value} = 0.26$, $d = 0$ lanciano dubbi sulla validità del test. Quindi in questo caso si assume come caratteristica principale l'area media delle cisti, intesa come distribuzione delle aree. I risultati derivanti dalla presenza di cisti sono quindi basati sulle dimensioni medie delle aree.

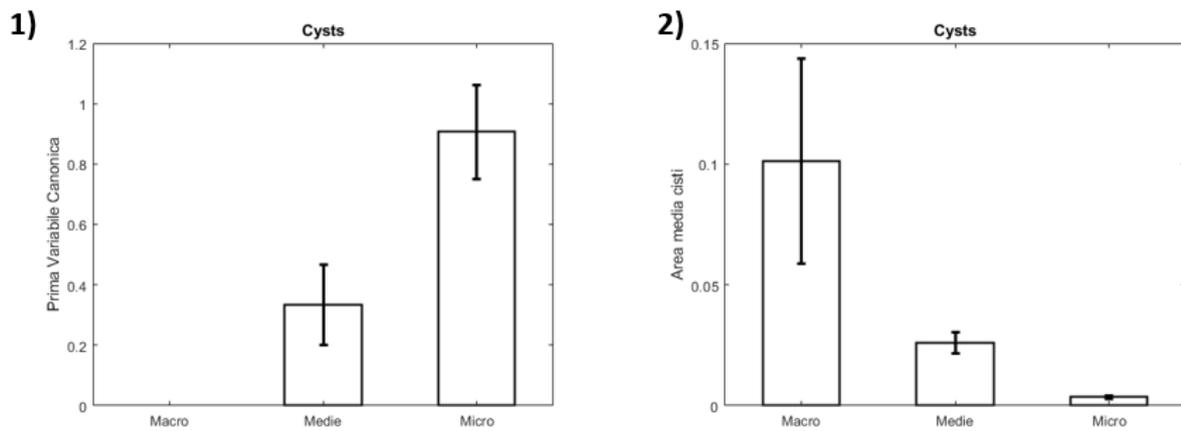


Figura 61: Istogrammi che riportano medie ed errori standard della presenza di cisti raggruppati per classi(1)Istogramma della prima variabile canonica(2) Istogramma della variabile più rappresentativa

Come si evince dall'istogramma le medie sono rappresentative di un buon grado di discriminazione tra le classi, e valgono 0.1, 0.026 e 0.003, tuttavia il notevole errore presente nei gruppi macro e medie cisti, pari rispettivamente a 0.04 e 0.004, non permette una distinzione ottimale nonostante le medie differenti. La classe delle micro cisti risulta robusta in termini sia di media che di errore. Il valore di correlazione della caratteristica è pari a 0.41 e indica un'accurata rappresentazione dell'insieme di dati e della robustezza della caratteristica stessa. Gli errori delle macro e medie cisti sono rilevanti, infatti le macro cisti dimostrano una dispersione elevata paragonata alle altre classi, questo è dovuto anche al loro esiguo numero. Tuttavia la presenza di micro cisti presenta un raggruppamento compatto e accurato, con un errore pari a $6 * 10^{-4}$, si prediligerebbe ottenere per questa suddivisione una media bassa con errore minimo. La quantificazione della presenza di cisti è solitamente rappresentativa di benignità infatti all'aumentare della dimensione delle cisti diminuisce il grado di malignità, come riportato in [12]. Questa caratteristica permette un'accurata valutazione per quanto riguarda la natura del nodulo, ma se stimata singolarmente non raggiunge un grado di specificità sufficiente alla classificazione finale, perciò è spesso valutata e supportata da altre caratteristiche ultrasonografiche.

3.5 Forma

Per quanto riguarda la forma si è proceduto al calcolo del rapporto tra altezza e larghezza del nodulo. Questa caratteristica data la sua semplicità di interpretazione è una caratteristica rappresentativa del nodulo, come mostra il suo valore di correlazione pari a 0.4.

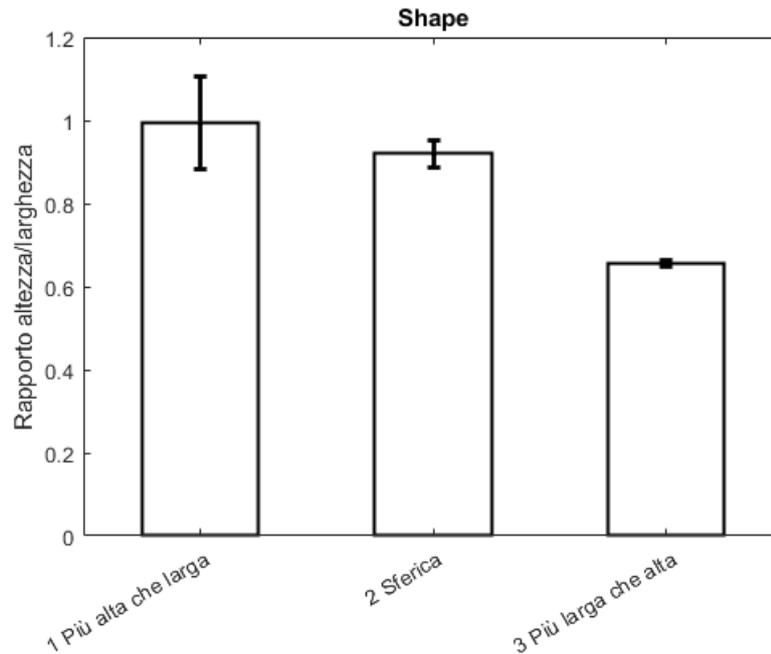


Figura 62: Istogramma che riporta medie ed errori standard della forma suddivisa per classi

Nell'istogramma, mostrato in Fig.62 si riportano le medie, più alta che larga (PAL) = 1, sferica (S) = 0.92 e più larga che alta (PLA) = 0.65, ed gli errori standard, pari a PAL = 0.11, S = 0.03 e PLA = 8×10^{-4} , della forma raggruppati per suddivisioni. Come si evince dal grafico l'errore maggiore affligge la suddivisione PAL, poichè questa classe è associata all'incremento del grado di malignità un errore notevole lancia dubbi sulla validità della caratteristica. Tuttavia nel database i noduli che entrano a far parte di questa classe sono in numero molto limitato, circa il 2%, perciò se si avesse a disposizione un numero maggiore di questi noduli l'errore diminuirebbe e questa caratteristica acquisterebbe ancora maggiore solidità in termini discriminatori tra classi. La medesima assunzione si può impiegare per la forma sferica, che nel database è presente circa nel 4% dei noduli, e si può notare come l'errore diminuisca notevolmente aumentando anche solo di una piccola percentuale il numero di campioni analizzati. L'errore minore che rende molto robusta e stabile la classe PLA deriva dal consistente numero di noduli presenti nel database e accomunati a questa suddivisione. Una forma più alta che larga è una caratteristica molto specifica ed è associata ad un elevato grado di malignità e le altre classi sono associate ad un decremento del grado di malignità. Questa caratteristica offre la possibilità di un'accurata valutazione per quanto concerne la natura del nodulo con un elevato grado di specificità. Tuttavia ci sono casi in cui noduli che presentano forma sferica o più larga che alta abbiano una natura maligna quindi anche questa caratteristica non può essere considerata a se stante per la classificazione finale.

3.6 Vascolarizzazione

La valutazione dei risultati per quanto riguarda la vascolarizzazione si conduce sulle immagini doppler. In questo specifico caso si è accorpata la classe intranodulare alla classe peri-intranodulare, poichè dalle immagini non risulta esserci una vascolarizzazione intranodulare in assenza di una periferica. I risultati ottenuti sono pertinenti alle zone di indagine cioè si valuta la presenza di segnale separatamente in sede periferica e in sede centrale del nodulo. Si è scelto di rappresentare solo l'indice di frazione vascolare calcolato con la maschera intranodulare che fornisce risultati più attendibili, poichè restituisce una correlazione pari a 0.188 rispetto a 0.162 per l'utilizzo della maschera periferica. Si è provveduto alla rappresentazione dell'indice di frazione vascolare intranodulare e della prima variabile canonica, proveniente dalla MANOVA (Fig.63). I risultati del test, il $p - value = 8 * 10^{-13}$, $d = 1$

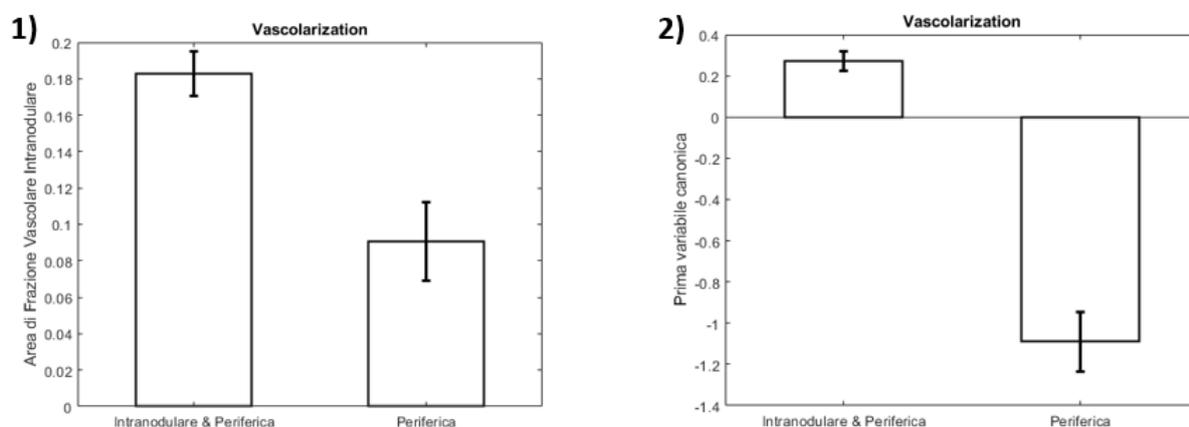


Figura 63: Istogrammi che riportano medie ed errori standard della vascolarizzazione suddivisi per classi(1) istogramma della variabile più rappresentativa(2)Istogramma della prima variabile canonica

Dalla figura di cui sopra si nota come entrambe le caratteristiche siano rappresentative dell'insieme dei dati, per cui si è posta l'attenzione sulla prima variabile canonica che mostra per le suddivisioni un errore minore, pari a 0.145 e 0.05, e medie con maggior differenza, pari a -1.09 e 0.27, rispettivamente periferica e peri-intranodulare. Si è soliti accomunare l'aumento della vascolarizzazione nodulare con un incremento di fattore di rischio di malignità, tuttavia più recentemente si è dimostrato che questa caratteristica non ha capacità predittive. Uno studio su quasi 700 neoplasie tiroidee ha rilevato l'assenza di vascolarizzazione intranodulare nel 63% dei casi maligni [10]. La vascolarizzazione è una misura dell'attività del nodulo, anche se si riesce a quantificarla in maniera discreta, non funge da perno centrale nella classificazione finale, ma è utile come supporto.

3.7 Elastosonografia

La misura dell'elasticità di un nodulo si basa su immagini elastosonografiche. Si traduce una misura di posizione in una codifica colori. Si riportano di seguito i valori di elasticità calcolati e la prima variabile canonica rappresentati tramite diagrammi a barre (Fig.64). Come per le altre caratteristiche la combinazione lineare rappresentata deriva dalla MANOVA a cui sono state fornite altre caratteristiche estratte dal nodulo e discusse in precedenza. Il test genera i seguenti valori: $p - value = 1.6 * 10^{-36}$, $p - value = 9.4 * 10^{-17}$, $d = 2$.

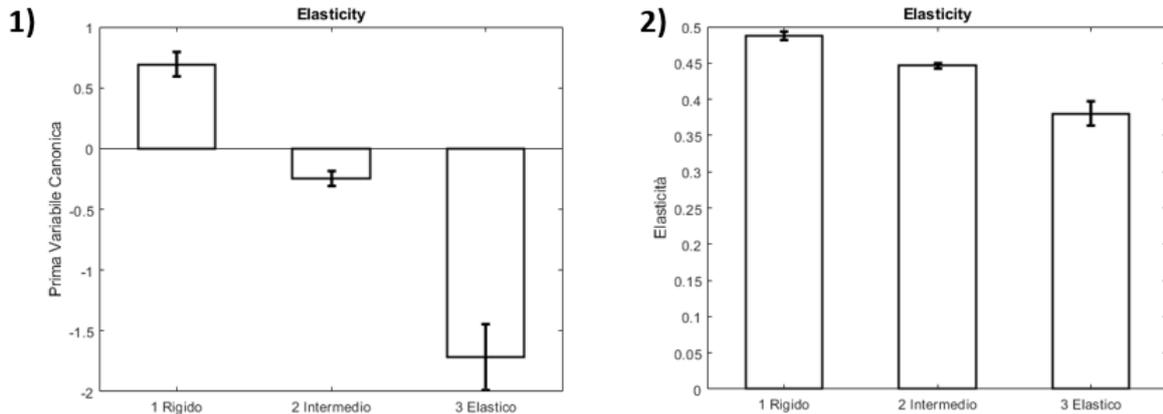


Figura 64: Istogrammi che riportano medie ed errori standard dell'elasticità suddivisi per classi(1) istogramma della prima variabile canonica(2)Istogramma dell'elasticità

Gli istogrammi riportano le medie delle tre suddivisioni, in particolare rigido, intermedio ed elastico e valgono rispettivamente 0.69, -0.245 e -1.71, e in aggiunta le barre dell'errore standard. In questo caso la prima variabile canonica risulta vantaggiosa in termini di rappresentazione, anche se mostra un errore maggiore, pari a 0.27, 0.063 e 0.1 per le classi rispettivamente rigido, intermedio ed elastico, ma dimostra anche una maggiore correlazione pari -0.583. La distinzione delle classi in base alle medie è efficace e si può giungere ad un accurato raggruppamento. L'elasticità di un nodulo è una caratteristica con un buon grado di accuratezza per quanto concerne la descrizione del grado di malignità. Come dimostrato e riportato in [21], solitamente una notevole presenza di colore blu è associata alla malignità. Questa caratteristica è in grado di descrivere in maniera accurata la natura di un nodulo, tuttavia è una metodologia non presente in tutti gli strumenti diagnostici ad ultrasuoni. La quantificazione dell'elasticità restituisce dei buoni risultati sia in termini di correlazione sia di suddivisione tra le classi. E' una caratteristica che a differenza delle altre, se ben analizzata, è altamente specifica e potrebbe già da sola valutare il grado di malignità del nodulo. Tuttavia esistono casi in cui anche questa metodologia perde la sua efficacia quindi è sempre preferibile che sia affiancata da altre caratteristiche.

3.8 Classificazione Caratteristiche Ultrasonografiche

Questa sezione è dedicata alla classificazione dei noduli del dataset in due classi: Benigno e Maligno. Si è ritenuto opportuno condurre una classificazione finale per valutare quanto è accurata la quantificazione delle caratteristiche ultrasonografiche e se potesse realmente ricoprire un ruolo di sostegno diagnostico. Inizialmente si è valutata la correlazione tra le caratteristiche utilizzate per la classificazione, in Fig.65 è riportato un esempio di matrice di correlazione.

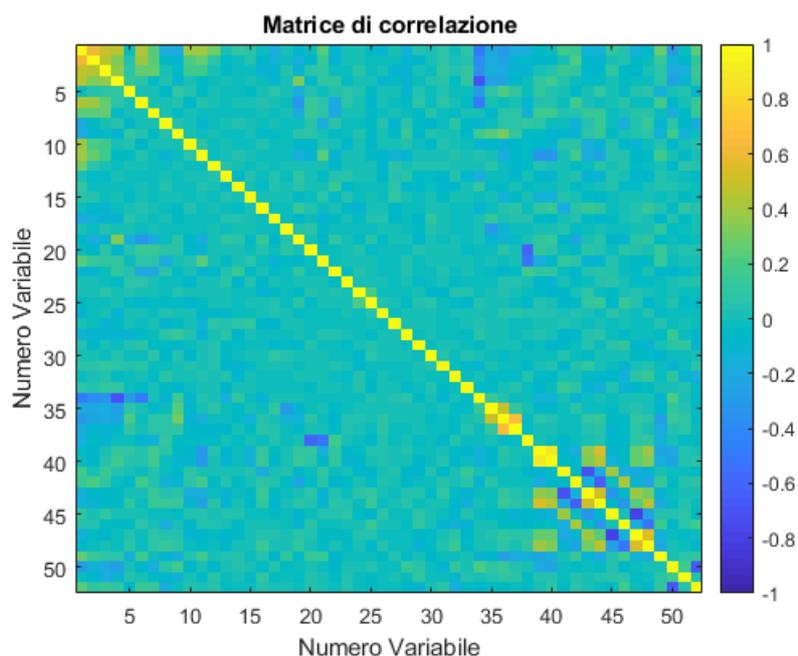


Figura 65: Matrice di correlazione delle caratteristiche utilizzate nella MANOVA per la classificazione Ben/Mal

Un coefficiente di correlazione massimo, pari a 1 e rappresentato da un colore giallo (Fig.65), denota che non è avvenuta nessuna diversificazione delle caratteristiche, quindi si sta analizzando ripetutamente la stessa variabile. Un coefficiente pari a 0, colore azzurro, indica totale indipendenza stocastica, cioè le features si muovono in modo autonomo. Un coefficiente minimo, pari a -1 ed evidenziato in blu scuro, mostra una totale decorrelazione, cioè le variabili si comportano in maniera opposta. La matrice permette di valutare la qualità della diversificazione, tanto meno le variabili sono correlate tra loro tanto maggiore sarà la robustezza del risultato finale. La matrice di correlazione è uno strumento utile per rafforzare la qualità delle features. Successivamente aver verificato la correlazione opportuna delle variabili si è condotta una MANOVA sul set di caratteristiche totali, sfruttando come Ground Truth la citologia presente nel database. La Fig.66 mostra la distribuzione della prima variabile canonica. Dal grafico si nota la buona distribuzione dei noduli per le due suddivisioni.

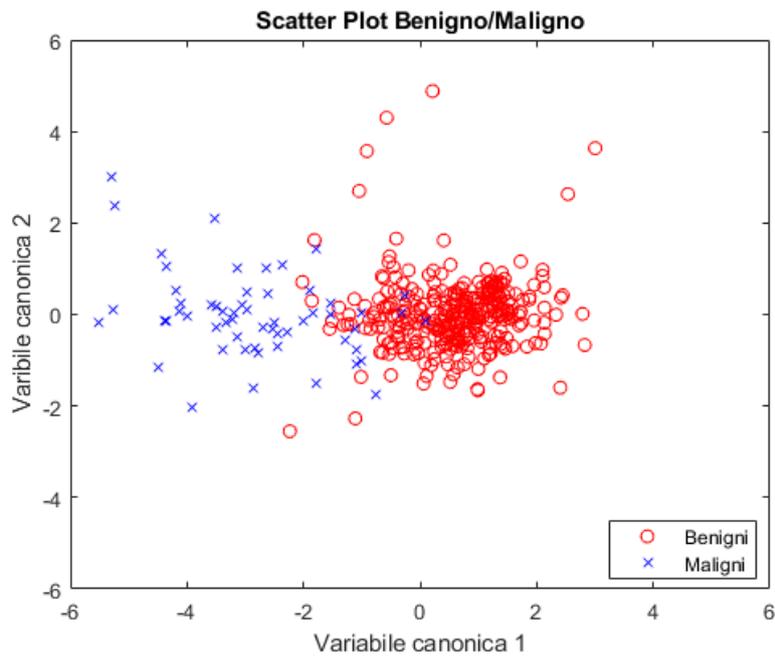


Figura 66: Grafico a dispersione della prima e della seconda variabile canonica e rappresentazione dei noduli suddivisi nelle due classi Benigna e Maligna

I risultati forniti dal test sono $d=1$ e un $p - value = 9.72 * 10^{-43}$ che dimostrano una buona robustezza sia del test sia del raggruppamento delle variabili in classi differenti. La scarsa dispersione dei dati in funzione della prima variabile canonica manifesta una discreta solidità in termini di distinzione. Per comprendere quali caratteristiche sono state utilizzate per il grafico a dispersione si sono estratti i pesi dalla MANOVA (Fig.67).

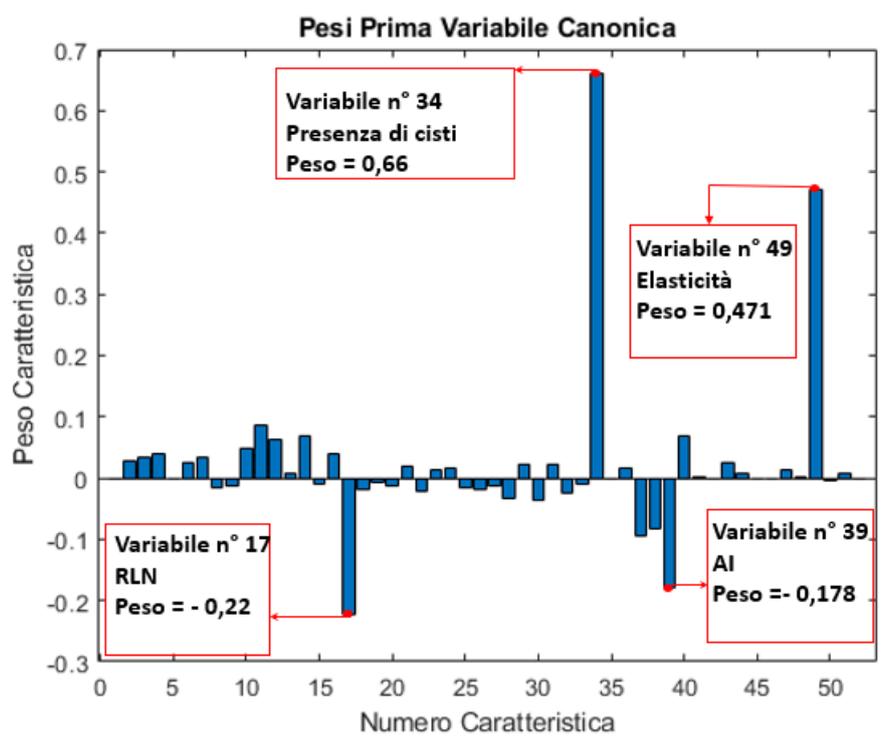
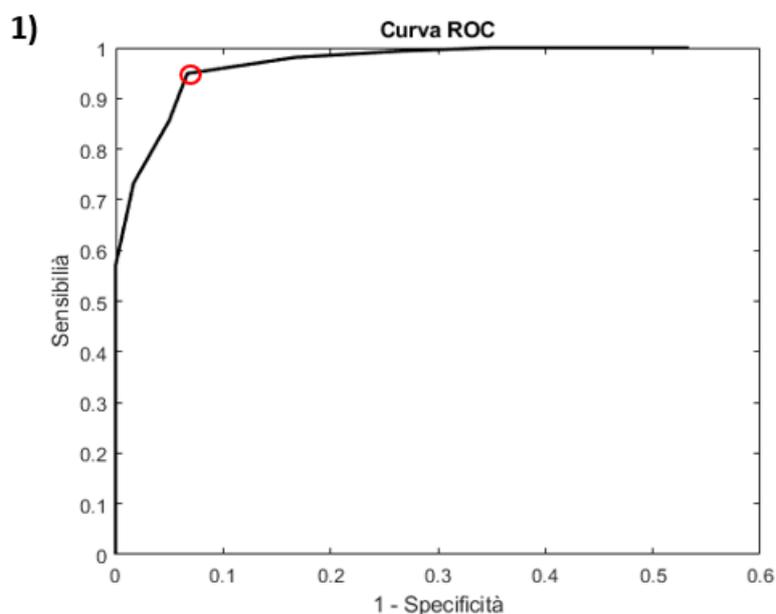


Figura 67: Diagramma a barre dei pesi delle caratteristiche utilizzate nella classificazione finale. Nei riquadri in rosso sono evidenziate i pesi e le caratteristiche più importanti

Il diagramma a barre rappresentato mostra l'influenza delle variabili nella rappresentazione della prima variabile canonica. La Fig.67 rivela come avviene la classificazione finale, ovvero quali caratteristiche riproducono, più o meno, fedelmente il processo decisionale che il medico percorre per giungere ad una conclusione e quindi ad una definizione del grado di malignità del nodulo. Si nota come la presenza di cisti e l'elasticità ricoprono un ruolo cardinale nella classificazione finale. I pesi negativi sono accostati a una riduzione della variabile, ciò implica un aumento del grado di malignità. I pesi positivi si comportano in maniera opposta, diminuiscono il grado di malignità.

Infine si è proceduto alla classificazione, nonostante lo sbilanciamento del database iniziale. Non si è utilizzato un classificatore complesso, anzi si è proceduto a classificare nella maniera più semplice possibile, ovvero utilizzando la prima variabile canonica, assunta come migliore combinazione lineare possibile di tutte le caratteristiche. Si è impostata una soglia pari a -1 e si è proceduto attribuendo la malignità ad un valore minore della soglia e la benignità a uno maggiore. La scelta della soglia è stata definita tramite curva ROC (Fig.68), in cui si sono calcolati Sensibilità e Specificità al variare della soglia. Si è definito l'intervallo di variazione da -2.6 a 0.15.



2)

	Benigno	Maligno
Benigno	298	4
Maligno	16	56

Figura 68: Risultati grafici classificazione (1)Curva ROC, sulla quale si evidenzia in rosso la posizione in cui si è scelta la soglia pari a -1(2) Confusion matrix relativa alla curva ROC e ai noduli analizzati

Si è evidenziata la posizione assunta sulla curva ROC. Questa posizione vantaggiosa ha portato alla definizione della soglia. In particolare con la soglia pari a -1 si ottengono dei valori di specificità pari al 93.3% e di sensibilità pari al 94.9%. L'accuratezza che deriva dalla curva risulta essere pari al 94.6%. Queste valutazioni e questi parametri ottenuti conferiscono validità alla quantificazione delle caratteristiche e lasciano presagire dei possibili sviluppi futuri partendo da solide fondamenta.

4 Discussioni

La gestione dei noduli tiroidei a citologia indeterminata è un problema non ancora risolto. Gli esami citologici che non forniscono esito rispetto alla natura del nodulo sono seguiti dall'intervento chirurgico esclusivamente per diagnosi. Successivamente alla tiroidectomia si verifica che la maggior parte dei noduli analizzati sono di natura benigna, perciò si è in costante ricerca di nuove metodologie per ridurre interventi chirurgici inutili. Oggigiorno la gestione dei noduli tiroidei a citologia ignota deve essere pianificata in funzione dei dati clinici e delle caratteristiche ecografiche. Con la stesura del presente elaborato si è cercato di descrivere e discriminare tramite l'utilizzo di immagini ad ultrasuoni la natura maligna di un nodulo da quella benigna. Si è trattata ogni caratteristica singolarmente come descritto in precedenza e si è provveduto ad ottimizzare le immagini per quella specifica caratteristica. I principali vantaggi dell'impiego della quantificazione semi-automatica sono il buon grado di specificità che si raggiunge, la facile implementazione, la possibilità di integrare il lavoro in un macchinario diagnostico e il supporto, oggettivo e numerico, fornito al medico durante l'esame strumentale. Tuttavia si sono riscontrate numerose problematiche che hanno influito sul risultato finale. Il limite principale che emerge da questa tesi è la difficoltà di quantificazione numerica in maniera precisa, ma con l'impiego di nuovi dati e una maggiore disponibilità di immagini ecografiche, si potrebbe giungere a risultati ottimi che potrebbero aiutare, anche rivestendo il ruolo di assistente, il medico nella sua diagnosi. Un altro suo limite risiede nella scarsa specificità raggiunta a livello diagnostico, a causa delle misure quantitative condotte non si è riuscito ad ottenere la stessa specificità che gli esperti adottano nel valutare le caratteristiche. L'implementazione rispetto ad un database ignoto, nel caso si utilizzi un nuovo database, potrebbe condurre a risultati differenti rispetto a quelli ottenuti, e potrebbe essere necessario apportare delle ulteriori modifiche tecniche e introdurre nuovi passaggi nella quantificazione delle caratteristiche, tuttavia la poca capacità di generalizzazione si risolve solo in minima parte tramite utilizzo dell'interfaccia grafica. Il calcolo delle caratteristiche invece ha riscontrato esigue difficoltà. Il limite legato al calcolo dell'ecogenicità è intrinseco all'immagine, poichè le maschere usate per l'identificazione delle regioni che rientrano nel calcolo sono state tracciate e verificate, tuttavia le intensità dei pixel dipendono fortemente dalla qualità stessa dell'immagine. L'ecogenicità rispetto alla tiroide, a differenza delle altre due, fornisce i risultati migliori poichè essa può essere accostato alla misura dell'ecogenicità del tessuto limitrofo e nel quale ha avuto origine il nodulo. La quantificazione della struttura e dei margini invece richiede più analisi, per cui si riesce per entrambe a ottenere dei risultati quasi ottimali. Una quantificazione della struttura è ostica poichè non emerge dall'immagine la struttura di un nodulo in termini numerici, infatti è molto più immediato valutare la struttura con occhio umano. La presenza di cisti incontra il suo principale limite nella descrizione della distribuzione delle aree presenti, ovvero risulta faticoso fornire un parametro ottimale che descriva accuratamente la disposizione delle cisti in un nodulo. In seguito al tuning dei percentili si è scelto il più accurato per la maggior parte dei noduli, tuttavia non sempre gli elementi utili vengono esaminati sotto la percentuale scelta, in altre parole si corre il rischio di valutare elementi non rilevanti o perdere elementi di interesse. Il numero limitato di cisti presenti nel database affligge ulteriormente la stima della caratteristica, nonostante ciò il vantaggio principale risiede nell'accostamento di benignità con la presenza di cisti, soprattutto nel caso di macro cisti. Il freno principale per il calcolo della caratteristica consiste nell'orientazione del nodulo, ciò nonostante si commette un errore del trascurabile. Questa caratteristica è molto valida data la sua semplicità. Le

altre caratteristiche quali la vascolarizzazione e l'elastasonografia non introducono problematiche intrinseche all'immagine, anzi introducendo una mappa colore sull'immagine, rappresentativa direttamente della grandezza in esame, non hanno riscontrato particolari difficoltà nella loro quantificazione. Tuttavia la vascolarizzazione si è dimostrata poco predittiva del grado di malignità [14] Il limite principale di questa caratteristica risiede nel calcolo delle maschere, ovvero quando si calcola il diametro equivalente poichè se un nodulo assume una forma schiacciata il raggio derivante, il quale fornisce dimensioni all'elemento strutturale, potrebbe introdurre una sovrastima o una sottostima del segnale vascolare. In seguito all'indagine statistica si evidenzia che le migliori caratteristiche, per la rappresentazione delle classi, sono l'ecogenicità rispetto alla tiroide e l'elastasonografia. Le altre sono meno rappresentative come la presenza di cisti e la vascolarizzazione, che forniscono una misura meno discriminante. Il processo decisionale del medico che attribuisce il grado di malignità ad un nodulo si può tradurre in un processo decisionale automatico in base alle features più significative che, una volta quantificate, forniscono una valida classificazione del nodulo. La classificazione del nodulo in benigni/maligni ha dimostrato come la quantificazione delle caratteristiche ultrasonografiche si riversi in una buona capacità decisionale. Gli errori commessi, che ricoprono circa il 5% delle decisioni finali, sono spesso causate da dei parametri anatomici ingannevoli in termini digitali, infatti la decisione finale compete sempre il medico e non lo strumento diagnostico, poichè anche nuove metodologie d'indagine non sono ancora oggi paragonabili alla decisione umana dettata dall'esperienza. La quantificazione delle caratteristiche ultrasonografiche in maniera semi-automatica è un'ardua sfida, ovvero attribuire un numero ad un parametro anatomico che ne descriva il grado di malignità. Un vantaggio nel valutare le caratteristiche in maniera semi-automatica è ottenere un parere più oggettivo tralasciando eventuali errori dovuti al condizionamento della persona. La possibilità di attribuire ad un nodulo una quantità numerica significa associare implicitamente a quel nodulo una classe nota, tuttavia la valutazione delle caratteristiche non garantisce il più delle volte che l'associazione adoperata utilizzando una sola caratteristica sia corretta, anzi il più delle volte si ricorre a diverse caratteristiche. Si denota come l'ipotesi di riuscire a quantificare le caratteristiche ultrasonografiche in maniera semi-automatica risulta un'ardua impresa, ciò nonostante si giunge a discreti risultati con tutti i limiti del caso. I buoni risultati lasciano intravedere delle successive e future migliorie per la quantificazione automatica ma le problematiche da affrontare sono molteplici.

5 Conclusione

Nella valutazione ecografica di una lesione tiroidea il medico si basa principalmente sull'ecogenicità, l'omogeneità e la qualità dell'immagine e le associano a un grado di malignità. Tuttavia esistono ancora molteplici vie interpretative e non tutte portano ad una diagnosi accurata, infatti la qualità della diagnosi è strettamente legate all'esperienza del medico. Il grado di malignità finale deriva dall'unione di tutte le caratteristiche ecografiche e dalla loro corretta quantificazione. Con la seguente tesi si cerca di fornire una descrizione approfondita dei noduli tiroidei adottando come punto di partenza la definizione delle caratteristiche ultrasonografiche associandole alla funzionalità e all'anatomia della tiroide, passando per l'estrazione e la quantificazione semi-automatica delle caratteristiche ultrasonografiche, e validandole tramite indagini differenti. La quantificazione delle caratteristiche ultrasonografiche, in maniera semiautomatica, si è performata per ecogenicità, struttura, margini, presenza di cisti, forma, vascolarizzazione ed elasticità. Si è estratta e trattata ogni caratteristica singolarmente. Per il calcolo delle caratteristiche si è provveduto a ottimizzare degli step specifici, i quali si sono dimostrati di fondamentale importanza per la valutazione di ogni caratteristica. Si è anche creata un'interfaccia grafica che agevola la quantificazione, permette di selezionare le immagini e modificare le maschere e mostra la classificazione finale rispetto a quella reale. Non è facile paragonare questo approccio rispetto ad altri, poichè non è possibile disporre di ricchi database pubblici, per performare e ottimizzare la quantificazione delle caratteristiche ultrasonografiche. La classificazione finale dei noduli nelle due classi ricopre il ruolo di test di qualità del lavoro svolto. Le buone performance del classificatore scaturiscono soprattutto dalla buona qualità di dati da classificare. Le successive e numerose azioni che si sono compiute per cercare di quantificare più parametri anatomici si riflettono infine in una robusta quantificazione delle caratteristiche ultrasonografiche. Questa metodologia si pone come finalità la discriminazione delle lesioni da trattare chirurgicamente da quelle che necessitano esclusivamente di monitoraggio. La trattazione di lesioni a citologia ignota può essere trattata tramite quantificazione delle features così facendo si cerca di evitare tiroidectomie e ricoveri ospedalieri per un semplice esame diagnostico riducendo notevolmente il costo dell'esame stesso e rendendo l'esame più confortevole anche per il paziente. La programmazione del trattamento medico a cui bisogna sottoporsi in caso di presenza di lesione tiroidea può essere progettata grazie anche all'aiuto del lavoro svolto in questa tesi ma sotto la supervisione di un medico esperto. Le sfide, lasciate ancora in sospese, sono rappresentate dalla miglioramento che si può introdurre nel calcolo delle caratteristiche e da nuove metodologie di quantificazione delle stesse, un esempio innovativo è la quantificazione dei margini tramite fractal dimension. Lo studio condotto offre delle solide fondamenta sulle quali basarsi per poter condurre nuovi studi riguardanti le caratteristiche ultrasonografiche. Per poter sfruttare al meglio il lavoro svolto in questa tesi per studi futuri si raccomanda di impiegare nuovi metodi per aumentare il potere descrittivo delle caratteristiche e renderle più specifiche possibili, ovvero raggiungere il medesimo livello diagnostico che il medico fornisce durante un esame specialistico. Nonostante avendo considerato un modesto set di pazienti che portano già a dei risultati sub-ottimali, le performance di classificazione possono essere affinate, utilizzando classificatori complessi magari tramite anche l'utilizzo di machine learning. Esistono tuttavia anche margini di miglioramento della quantificazione ultrasonografica stessa, by-passando le limitazione computazionali intrinseche che ogni caratteristica porta con sè come nei casi trattati nei capitoli precedenti. L'applicazione di un metodo scientifico a problemi deci-

sionali complessi implica la gestione e la messa a punto di sistemi in grado di fornire soluzioni, ottimali o subottimali, che meglio conducano ad una decisione rapida, semplice e corretta. Con la presente tesi si è cercato di creare un nuovo metodo quantitativo avanzato che mira a fornire ausilio ad un processo decisionale complesso, tramite l'utilizzo di strumenti matematici e statistici. La nuova metodologia può essere considerata come la formalizzazione di un problema di associazione numerica a delle caratteristiche anatomiche, la quale può essere assunta come modello primordiale su cui basarsi per prendere decisioni[**RAJENDRAACHARYA2012508**].

Bibliografia

- [1] U. Rajendra Acharya et al. «Non-invasive automated 3D thyroid lesion classification in ultrasound: A class of ThyroScanTM systems». In: *Ultrasonics* 52.4 (2012), pp. 508–520. ISSN: 0041-624X. DOI: <https://doi.org/10.1016/j.ultras.2011.11.003>. URL: <http://www.sciencedirect.com/science/article/pii/S0041624X11002472>.
- [2] «Advances In Quantitative Muscle Ultrasonography Using Texture Analysis Of Ultrasound Images». In: *Ultrasound in Medicine & Biology* 41.9 (9 settembre 2015). Accessed on: Nov.21,2019. [Online]. Available: <https://lnx.endocrinologiaoggi.it/2011/06/scintigrafia-tiroidea/>, pp. 2520–2532.
- [3] Dott. Massimiliano Andrioli. «Agos aspirato tiroideo». In: <https://www.endocrinologiaoggi.it/2011/06/agoaspirato-tiroideo/> (). consultato il 23.11.19.
- [4] S. Arcieri. In: *Wordpress* (28.AUG.2014). Accessed on: Nov.21,2019. [Online]. Available: <https://stefanoarcieri.wordpress.com/2014/08/28/a-vista-docchio-sara-colpa-della-tiroide/>.
- [5] Sergio Bolasco. «Analisi multidimensionale dei dati. Metodi, strategie e criteri d'interpretazione». In: 6^a ed., Roma, Carocci, (2014).
- [6] Marta Borowska et al. «Fractal dimension in textures analysis of xenotransplants». In: *Signal, Image and Video Processing* 11.8 (nov. 2017), pp. 1461–1467. ISSN: 1863-1711. DOI: 10.1007/s11760-017-1108-5. URL: <https://doi.org/10.1007/s11760-017-1108-5>.
- [7] «Carcinoma tiroideo». In: *medicina360* (). Accessed on: Nov.21,2019. [Online]. Available: <https://www.medicina360.com/wp-content/uploads/2019/08/carcinoma-tiroideo-il-tumore-alla-tiroide-640x480.jpg>.
- [8] Kuen-Yuan Chen et al. *Computerized Detection and Quantification of Microcalcifications in Thyroid Nodules*. 2011. DOI: <https://doi.org/10.1016/j.ultrasmedbio.2011.03.002>. URL: <http://www.sciencedirect.com/science/article/pii/S0301562911001372>.
- [9] Kuen-Yuan Chen et al. «Computerized Quantification of Ultrasonic Heterogeneity in Thyroid Nodules». In: *Ultrasound in Medicine & Biology* 40.11 (2014), pp. 2581–2589. ISSN: 0301-5629. DOI: <https://doi.org/10.1016/j.ultrasmedbio.2014.06.009>. URL: <http://www.sciencedirect.com/science/article/pii/S0301562914004190>.
- [10] Kuen-Yuan Chen et al. «Computerized Quantification of Ultrasonic Heterogeneity in Thyroid Nodules». In: *Ultrasound in Medicine & Biology* 40.11 (2014), pp. 2581–2589. ISSN: 0301-5629. DOI: <https://doi.org/10.1016/j.ultrasmedbio.2014.06.009>. URL: <http://www.sciencedirect.com/science/article/pii/S0301562914004190>.
- [11] «Classificazione Italiana 2014 della Citologia Tiroidea». In: (). Accessed on: Nov.21,2019. [Online]. Available: <http://www.societaitalianadiendocrinologia.it/html/pag/classificazione-italiana-2014-della-citologia-tiroidea.asp>.
- [12] National University Hospital Department of Diagnostic Imaging e Singapore Yong Loo Lin School of Medicine National University of Singapore. «Thyroid nodules: risk stratification for malignancy with ultrasound and guided biopsy.» In: *Endocrinol Metab Clin North Am Cancer Imaging* 11 (2011 Dec 28), pp. 209–23.
- [13] «Ecografo». In: (). Accessed on: Nov,21,2019. [Online]. Available: <http://www.vomark.com/product/mindray-dc-7/>.
- [14] Division of Endocrinology. «Clinical Diagnostic Evaluation of Thyroid Nodules». In: *Endocrinol Metab Clin North Am* 48(1) (2019 Mar), pp. 61–84.
- [15] «Evoluzione triangolo sierpinski». In: (2013). Accessed on: Nov.22,2019. [Online]. Available: <http://maddmaths.simai.it/2013/04/evoluzione-triangolo-sierpinski/>, p. 1.
- [16] Dr. R.G. Fonseca. «Clasificación ecográfica TI-RADS en nódulo tiroideo». In: <http://www.drlazzari.it/blog/2015/04/28/clasificacion-ecografica-ti-rads-en-nodulo-tiroideo/> (2015). Accessed on: Nov.21,2019. [Online]. Available: <https://www.ecosia.org/images?q=tiradsid=CD333B87>.
- [17] Akrem El-ghazal, Otman Basir e Saeid Belkasim. «Farthest point distance: A new shape signature for Fourier descriptors». In: *Signal Processing: Image Communication* 24.7 (2009), pp. 572–586. ISSN: 0923-5965. DOI: <https://doi.org/10.1016/j.image.2009.04.001>. URL: <http://www.sciencedirect.com/science/article/pii/S0923596509000393>.

- [18] Amogh Hegde Gopinathan Anil e F.H. Vincent Chong. «Thyroid nodules: risk stratification for malignancy with ultrasound and guided biopsy». In: *E-med* (Date accepted for publication 28 September 2011), pp. 209–223.
- [19] Jorge Luiz Gross. «Thyroid Ultrasound Features and Risk of Carcinoma: A Systematic Review and Meta-Analysis of Observational Studies Thyroid». In: *Creative Commons Attribution Noncommercial License 25.5* (2015), pp. 209–223.
- [20] «Italian consensus for the classification and reporting of thyroid cytology». In: *JOURNAL OF ENDOCRINOLOGICAL INVESTIGATION* 37 (6) (2014), pp. 593–599.
- [21] Ryoji Kagoya, Hiroko Monobe e Hitoshi Tojima. «Utility of elastography for differential diagnosis of benign and malignant thyroid nodules». In: *Otolaryngology - Head and Neck Surgery* 143.2 (2010), pp. 230–234. ISSN: 0194-5998. DOI: <https://doi.org/10.1016/j.otohns.2010.04.006>. URL: <http://www.sciencedirect.com/science/article/pii/S0194599810003736>.
- [22] Eisuke Koike et al. «Ultrasonographic Characteristics of Thyroid Nodules: Prediction of Malignancy». In: *Archives of Surgery* 136.3 (mar. 2001), pp. 334–337. ISSN: 0004-0010. DOI: 10.1001/archsurg.136.3.334. eprint: <https://jamanetwork.com/journals/jamasurgery/articlepdf/391154/soa0167.pdf>. URL: <https://doi.org/10.1001/archsurg.136.3.334>.
- [23] Dipartimento di Elettronica e Telecomunicazioni (DET) Politecnico di Torino Turin Italy Kristen M. Meiburger. «Quantitative Ultrasound and Photoacoustic Imaging for the Assessment of Vascular Parameters». In: *PoliTO Springer Series* 9 (2017), pp. 1–15.
- [24] Dr. F. Lazzari. «Esame ecografico tiroideo». In: <http://www.drlazzari.it/blog/> (). Accessed on: Nov.21,2019. [Online]. Available: <http://www.drlazzari.it/blog/>.
- [25] Tianjiao Liu et al. «Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks». In: *Medical Image Analysis* 58 (2019), p. 101555. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2019.101555>. URL: <http://www.sciencedirect.com/science/article/pii/S1361841519300970>.
- [26] Yun Lu. «Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network». In: *World Journal of Surgical Oncology* 17 (2019), p. 12.
- [27] Donald R. Mann Henry B.; Whitney. «Ön a Test of Whether one of Two Random Variables is Stochastically Larger than the Other: » In: *Annals of Mathematical Statistics*. 18 (1) ((1947)), pp. 50–60.
- [28] G. Marchionni. «Elastosonogra: principi di funzionamento e applicazioni». In: ((2012-2013)), available on: https://amslaurea.unibo.it/6908/1/marchionni_giulia_tesi.pdf, pp. 7–9, 15–19.
- [29] DR. A. MAS. «FNA of Thyroid (Fine Needle Aspiration of Thyroid Nodule)». In: (NOVEMBER 21, 2018.). Accessed on: Nov.21,2019. [Online]. Available: <https://thyroid4patients.com/fna-thyroid/>.
- [30] Dott.ssa M. Mazzola. «Scintigrafia Tiroidea: come si esegue, a cosa serve, precauzioni e contrindicazioni». In: *medicina360* (4.06.2019). <https://www.medicina360.com/articoli.html>.
- [31] MD6 Myung Ho Rho. «Differentiation between Benign and Malignant Solid Thyroid Nodules Using an US Classification System». In: *Korean J Radiol* 12(5) (2011 Sep-Oct), pp. 559–567.
- [32] Filippo Bottazzi e Giorgio Mortara Paolo Enriques. «CORRELAZIONE». In: *treccani* (1931).
- [33] Antonin Prochazka et al. «Patch-based classification of thyroid nodules in ultrasound images using direction independent features extracted by two-threshold binary decomposition». In: *Computerized Medical Imaging and Graphics* 71 (2019), pp. 9–18. ISSN: 0895-6111. DOI: <https://doi.org/10.1016/j.compmedimag.2018.10.001>. URL: <http://www.sciencedirect.com/science/article/pii/S0895611118301915>.
- [34] Articolo realizzato dalle redazione. «Sintomi dei problemi alla tiroide». In: (). <https://www.medicina360.com/sintomi>
- [35] «Scintigrafia tiroidea». In: (). Accessed on: Nov.21,2019. [Online]. Available: <https://lnx.endocrinologiaoggi.it/2011/01/scintigrafia-tiroidea/>.
- [36] «Trattato di anatomia umana vEdi. Erme». In: *Edi. Ermes* vol 2 (agosto 2007), pp. 375–544.
- [37] «European Thyroid Association Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules in Adults: The EU-TIRADS Eur Thyroid». In: *Eur Thyroid J* 6 (2017), pp. 225–237.

- [38] R. T. Warne. «Ä primer on multivariate analysis of variance (MANOVA) for behavioral scientists». In: *Practical Assessment, Research & Evaluation* 19 (17) (2014), pp. 1–10.