

A REVIEW AND TEST OF SHORELINE EXTRACTION TECHNIQUES

R. Angelini¹, E. Angelats², G. Luzi²

¹ Department of Civil and Environmental Engineering, University of Florence, Italy - (riccardo.angelini)@unifi.it

² Geomatics Research Unit. Department of Signal Theory and Communication (TSC), Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA) – (guido.luzi, eduard.angelats)@cttc.cat

Parole chiave: estrazione linea di riva, classificazione, machine learning, multispettrale, radar.

Keywords: shoreline extraction, classifier, machine learning, multispectral, radar

Abstract

Le zone costiere rappresentano luoghi di elevato valore paesaggistico, economico e culturale. In un lembo di terra che va da qualche centinaio di metri al chilometro, si concentrano alte densità abitative, aree naturalistiche protette o luoghi di interesse commerciale e turistico. L'indicatore più immediato per monitorare la fascia costiera è rappresentato dalla linea di riva, intesa come confine tra mare e terraferma. Il monitoraggio della linea di riva è tradizionalmente effettuato con metodi topografici classici mediante l'utilizzo di ricevitori GNSS e/o laser scanner. Tuttavia, la sempre maggiore diffusione e distribuzione di dati satellitari, consente un monitoraggio più frequente e a costi contenuti, senza la necessità di una campagna terrestre di acquisizioni. L'obiettivo di questo studio è confrontare diverse metodologie di estrazione automatica della linea di riva da dati satellitari sia ottici che radar a media risoluzione e gratuiti, forniti dal programma Copernicus dell'Agenzia Spaziale Europea. Come caso di studio è stata scelta la spiaggia della Feniglia in provincia di Grosseto. Le tecniche utilizzate prevedono l'utilizzo di indici spettrali come Normalized Difference Water Index (NDWI) per esaltare nelle immagini il confine tra mare e terraferma. A tale scopo sono stati utilizzati per confronto algoritmi di classificazione supervisionata come Random Forest (RF), Minimum Distance (MD), Maximum Likelihood (ML) e non-supervisionata come k-means. Per verificare l'accuratezza delle diverse metodologie e fonti di dati è stata considerata una linea di riva di riferimento digitalizzata manualmente a partire da un ortomosaico frutto di un rilievo aerofotogrammetrico distribuito dalla Regione Toscana. I risultati migliori hanno raggiunto un'accuratezza di 3.75m corrispondente a meno della metà della dimensione del pixel a terra dei dati Sentinel-1 e Sentinel-2 (10m).

Shoreline represents the boundary between land and sea, and its accurate extraction is of utmost importance because of the economic and ecological value of coastal areas. Nowadays, satellite remote sensing is widely used for monitoring the natural environment. Indeed, satellite remote sensing data are cost-effective and periodically available over large areas at a relatively high spatial resolution. Hence, the automatic shoreline extraction from satellite images is a fundamental task for coastal monitoring and management. Shoreline extraction methods are usually applied to satellite remote sensing data. The goal of this study is to compare the performance of different shoreline extraction methods, such as thresholding and more complex classification approaches, such as Random Forest (RF), Minimum Distance (MD), Maximum Likelihood (ML), and K-means, using both optical and radar images. The considered case study area is the shallow basin of the Orbetello Lagoon and one of its arms called Feniglia. The data supplier is the Copernicus program, which, through the Sentinel-1 and Sentinel-2 missions, provides medium-resolution, open-access products. The accuracy of the obtained results from both methodologies is checked by validating the extracted shoreline using an aerial orthomosaic and, subsequently, a manually extracted shoreline. A preliminary accuracy assessment was performed for image classification, focusing on extracting four classes: water, soil, urban, and forest, using manual segmentation as a reference. In terms of deviation from the reference shoreline, the results obtained through the analysed methodologies achieved an accuracy of 3.75 m, less than half of the pixel size of the Sentinel-1 and Sentinel-2 used products.

1. Introduzione

Gli ambienti costieri sono fondamentali per la loro biodiversità e per le attività umane, sia ludiche che economiche. La zona costiera rappresenta un sistema altamente dinamico che subisce continue e molteplici trasformazioni. Le cause dei cambiamenti nelle zone costiere possono essere raggruppate in: geologiche e geomorfologiche, idrodinamiche, biologiche, climatiche e antropogeniche. (Łabuz, 2015).

Il monitoraggio dell'ambiente costiero viene solitamente condotto definendo diversi indicatori, tra cui il più comune e ampiamente utilizzato è la linea di costa. La definizione più semplice di "linea di costa" è il confine tra terra e mare. (Anders and Byrnes, 1991). Tuttavia, l'elevata variabilità temporale e spaziale negli ambienti costieri rende necessaria una definizione più precisa di questo indicatore. Secondo (Zollini et al, 2020) è

possibile individuare la linea di costa secondo tre modalità principali:

- caratteristiche visibili da un operatore su un'immagine aerea o di telerilevamento;
- l'intersezione tra un riferimento di marea e un modello digitale del terreno o un profilo costiero;
- caratteristiche delle immagini identificate da algoritmi automatici che non sono necessariamente visibili a un operatore senza ausili.

La metodologia adottata in questo articolo prevede un uso sistematico di dati provenienti dalla terza modalità confrontati con il "ground truth" proveniente dal primo tipo, avente una risoluzione spaziale di un ordine di grandezza superiore rispetto ai risultati da validare.

La raccolta tradizionale di dati sul campo, tipicamente eseguita manualmente da un operatore su un'ampia area costiera, risulta spesso onerosa in termini di tempo, costosa e influenzata dalla

capacità e dalle scelte dell'operatore stesso. Per contro, il telerilevamento si è affermato come uno strumento prezioso per il monitoraggio costiero grazie alla sua capacità di fornire dati su larga scala, a cadenza regolare, frequenti e con un rapporto costo-opportunità vantaggioso.

Le immagini satellitari provenienti sia da sensori remoti passivi che attivi possono essere utilizzate per estrarre la linea di costa. I sensori passivi, come quelli multispettrali, si basano sulle proprietà di riflettanza naturale delle strutture naturali o antropiche presenti nelle aree costiere, per la porzione dello spettro elettromagnetico considerata, tipicamente dal visibile all'infrarosso. Le immagini ottiche offrono un metodo semplice per estrarre le linee di costa basate sulla differenza di riflettanza spettrale tra la terraferma e l'acqua. Purtroppo, i dati multispettrali satellitari sono sensibili alle condizioni meteorologiche. Invece, le tecniche di telerilevamento attive, come Synthetic Aperture Radar (SAR), permettono di raccogliere misure utilizzabili indipendentemente dalla presenza di nuvolosità. Questo comporta un vantaggio evidente rispetto ai sensori passivi. Tuttavia, le immagini SAR richiedono una fase di preelaborazione non necessaria per quelle ottiche.

I metodi automatici per l'estrazione della linea di riva possono essere raggruppati in tre categorie (Toure et al., 2019):

- metodi basati sull'utilizzo di una soglia opportuna, detti anche di "thresholding";
- metodi basati su tecniche di classificazione delle immagini;
- approcci fondati su algoritmi di "edge detection", che mirano, cioè, a determinare i bordi tra le regioni di interesse presenti all'interno dell'immagine.

L'utilizzo di un metodo di thresholding rappresenta il metodo più semplice e veloce tra le tecniche di segmentazione. Tuttavia, potrebbe non essere una tecnica sufficientemente accurata in alcuni contesti influenzati da una elevata variabilità di illuminazione. Per questo motivo, vi sono stati numerosi tentativi per migliorare i risultati ottenibili tramite questo metodo. Tra le varianti sviluppate si possono elencare per esempio: la possibilità di stabilire un valore soglia ottimale tra due classi utilizzando il metodo di Otsu (Otsu, 1979), approcci più complessi e flessibili basati sull'utilizzo di soglie adattative (Aedla et al., 2015) o tramite l'utilizzo di soglie multiple (Jishuang and Wang, 2002). I metodi di classificazione cercano di rappresentare i dati in modo più semplice raggruppandoli in sottogruppi, o classi, in modo tale che gli elementi all'interno di una classe siano il più possibile simili tra loro, secondo determinati criteri, e il più possibile dissimili dagli elementi delle altre classi. È possibile distinguere tra due differenti tipologie di classificazione: supervisionata, prevede l'allenamento di un modello mediante una serie di dati etichettati a priori, o non supervisionata, nella quale il modello cerca di dividere l'immagine in gruppi (cluster) sulla base delle loro caratteristiche, senza avere a disposizione informazioni a priori. I secondi si contraddistinguono per la loro semplicità; in particolare, il metodo k-means è il più comunemente utilizzato (García-Rubio et al., 2015). La classificazione supervisionata permette generalmente di gestire fenomeni più complessi (Tsekouras et al., 2018a). Gli algoritmi di classificazione più comunemente usati al giorno d'oggi sono il Maximum Likelihood (ML), il Random Forest o il Minimum Distance (MD) (Bamdadinejad et al., 2021). Indipendentemente dal classificatore utilizzato, per ottenere risultati affidabili è essenziale una corretta fase di training, che preveda l'utilizzo di un set di dati rappresentativo. Lo scopo dovrebbe essere quello di utilizzare dati provenienti da diverse missioni satellitari e diverse aree con una varietà di tipologie di spiagge (Abdelhady et al., 2022) al fine di includere più variabilità possibile. Includendo nel dataset di allenamento una grande varietà di tipologie di spiagge

e dati provenienti da diverse aree geografiche, si può aumentare la probabilità che il modello sia in grado di classificare correttamente nuovi dati, non visti durante la fase di allenamento. Al fine di migliorare l'individuazione delle caratteristiche ricercate nelle immagini, molti autori hanno testato con successo indici spettrali per l'individuazione della vegetazione o dell'acqua come il Normalised Difference Water Index (NDWI) or Normalised Difference Vegetation Index (NDVI) (Ozturk and Sesli, 2015). Questi possono essere utilizzati sia nelle metodologie di "thresholding" che in fase preliminare in quelle di classificazione.

Gli approcci basati sul rilevamento dei bordi delle strutture dell'immagine "Edge detection" sfruttano le variazioni di colore, intensità dei livelli di grigio o texture di un'immagine per determinare i confini tra le aree di interesse, nel caso specifico, la linea di costa. Molti autori hanno proposto diversi approcci come il "Canny Edge Detection" (Tonye et al., 2000), Snakes (Klinger et al., 2011) e il Level Set Algorithm (Ouyang et al., 2010).

La maggior parte degli approcci appena menzionati sono stati implementati per dati derivanti da sensori passivi come quelli multispettrali. Pochi sono stati adattati ai sensori attivi e, verosimilmente, ancor meno tentativi sono stati compiuti nella loro integrazione. Alla luce di queste considerazioni, il presente studio si propone di analizzare e confrontare diverse tecniche di estrazione della linea di costa a partire da immagini radar e multispettrali. L'obiettivo specifico di questa ricerca è di mettere a confronto metodologie basate sulla scelta di una soglia con le tecniche di classificazione, sia supervisionate che non supervisionate.

I metodi scelti sono stati testati su immagini radar Sentinel-1 e sulle immagini multispettrali Sentinel-2, entrambe con una risoluzione spaziale di 10m. L'area pilota è la costa della Feniglia, un tombolo di 6km che chiude a Sud la Laguna di Orbetello, in Toscana (Italia). L'accuratezza dei risultati ottenuti è stata verificata confrontando la linea di costa estratta con una tracciata manualmente, derivante da un ortomosaico aereo con risoluzione spaziale di 0.2m fornito dal geoscopio WMS della Regione Toscana. L'accuratezza della classificazione è stata valutata utilizzando un "ground truth", determinato attraverso l'interpretazione visiva delle classi presenti nell'immagine.

2. Area di studio

L'area di test si trova nella parte meridionale della regione Toscana. Essa comprende la Laguna di Orbetello e uno dei tomboli che la separano dal mare chiamato Feniglia. La separazione dell'ambiente lagunare dal mare tramite la formazione dei due tomboli che la cingono è stata provocata dal combinato disposto di innalzamento del livello del mare Olocenico e dai fenomeni di diffrazione delle onde marine sulla paleo-isola del Monte Argentario (Barnes, 1980). Il tombolo della Feniglia è una striscia di terra di larghezza compresa tra i 700m e il chilometro e lunga 6km. Ad esso è associata un'area protetta istituita nel 1971. Oltre a rappresentare una riserva naturale, costituisce una zona di grande interesse economico e turistico. In passato, alcuni studi hanno già identificato le tendenze evolutive della linea di costa, la presenza di mega-cusps e l'evoluzione delle superfici vegetate (Cipriani et al., 2004). Alla luce dei cambiamenti climatici e dell'innalzamento del livello del mare, sono necessarie ulteriori ricerche.



Figura 1. Il tombolo della Feniglia con la caratteristica forma ad arco in direzione Ovest-Est. La mappa è nel sistema di riferimento WGS84-UTM32.

3. Materiali e metodologia

Questa sezione ha lo scopo di presentare i dati (sottosezione 3.1) e i metodi utilizzati in questo lavoro; ovvero tecniche di definizione delle soglie “thresholding” (sottosezione 3.2) e tecniche di classificazione delle immagini (sottosezione 3.3) Sono state utilizzate immagini open-source sia ottiche che radar, fornite dal programma Copernicus dell’Agenzia Spaziale Europea (ESA). Per l’elaborazione dei dati sono stati utilizzati software open-source come il Sentinel Application Platform (SNAP) e QGIS 3.4

3.1 Dati

In questo studio sono state utilizzate un’immagine satellitare multispettrale Sentinel-2 e un’immagine satellitare SAR Sentinel-1, come mostrato nella Tabella 1. Sentinel-1 e 2 sono missioni satellitari sviluppate dall’Agenzia Spaziale Europea (ESA) nell’ambito del Programma Copernicus I dati satellitari sono liberamente disponibili a diversi livelli di elaborazione per entrambe le piattaforme. I dati satellitari sono stati ottenuti attraverso il portale Copernicus Open Access Hub.

Piattaforme	Data [gg/mm/aaaa]	Ora [hh:mm]	Marea [m]	Onda [m]
S-2	23/07/2019	10:10	0.01	0.31
S-1	18/07/2019	05:19	0.07	0.43
Aereo	19/07/2019	12:15	0.08	0.41

Tabella 1. Immagini utilizzate nello studio e relative condizioni di meteomarine al momento della presa.

I satelliti Sentinel 2A e 2B sono dotati di un sensore multispettrale (MSI) che acquisisce immagini in 13 bande comprendenti il visibile, l’infrarosso vicino e l’infrarosso ad onde corte. La risoluzione spaziale varia tra 10m e 60m in funzione della banda spettrale considerata. In questo studio sono stati utilizzati i prodotti di livello 2A, orto rettificati, con correzione atmosferica e maschera sulla nuvolosità.

La missione Sentinel-1 offre diverse modalità di acquisizione per le immagini. In questo lavoro è stata usata la modalità di acquisizione Interferometric Wide Swath (IW) che fornisce immagini di dimensione 250 km * 250 km ad una iniziale risoluzione spaziale di 5 m x 20 m. Tra i diversi tipi di prodotto messi a disposizione è stato scelto il Ground Range Detected (GRD), un dato corretto geometricamente e per la curvatura terrestre. Inoltre, tale immagine è proiettata nel sistema di

riferimento World Geodetic System 84 (WGS-84) con il corretto fuso Universal Transverse Mercator (UTM) per l’area geografica di riferimento e una risoluzione finale a maglia quadrata di 10 m x 10 m.

Allo stato attuale, la ricerca ha visto l’utilizzo di immagini Sentinel-1 e Sentinel-2 separatamente. Gli sviluppi futuri si rivolgeranno al miglioramento dei risultati attuali, muovendosi inoltre nella direzione della fusione delle informazioni provenienti da ambedue le piattaforme per l’estrazione della linea di riva, al fine di ottenere risultati più accurati e completi.

Al fine di valutare le performance dei metodi implementati è stato utilizzato un ortomosaico, con risoluzione spaziale 0.2 m x 0.2 m fornito gratuitamente dalla Regione Toscana.

Per offrire ulteriori informazioni sulle condizioni ambientali a contorno sono stati presi in considerazione anche dati oceanografici come altezza d’onda e di marea. Questi sono essenziali nelle indagini riguardanti l’area costiera e in particolare quelle riguardanti la linea di riva. Infatti, la posizione della stessa cambia continuamente sotto l’influenza dei cicli di marea e del moto ondoso. L’utilizzo di dati provenienti da stazioni limitrofe alla zona di interesse è volto a migliorare l’accuratezza delle analisi permettendo di considerare gli effetti locali di questi cambiamenti, come sottolineato da (Liu et al., 2016).

Per le indagini sono stati considerati i dati sulle onde e le maree provenienti dalla boa di Giannutri, sotto la gestione della Regione Toscana, e dal mareografo di Civitavecchia, a cura dell’Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), essendo le più vicine al sito di studio. Dopo una verifica di tali informazioni, sono state considerate non rilevanti in questa fase, poiché caratterizzate da oscillazioni centimetriche, trascurabili rispetto all’ordine di grandezza della risoluzione delle immagini e dunque degli errori attesi. È importante sottolineare che questa ricerca è incentrata sulle metodologie di estrazione della linea di costa, piuttosto che su un monitoraggio puntuale o temporale della linea di riva. Di conseguenza, si è optato per l’utilizzo di dati in intervalli temporali ristretti e vicini tra loro, al fine di rendere minima l’influenza delle naturali modifiche della linea di riva a seguito anche di deboli mareggiate.

3.2 Thresholding

Nel diagramma in Figura 2 sono mostrati i dettagli della metodologia “thresholding” seguita nella definizione delle soglie più opportune.

Le immagini Sentinel-2 sono state scaricate al livello 2A che fornisce un’immagine orto rettificata e corretta dai disturbi atmosferici. L’indice NDWI è stato calcolato utilizzando la formula seguente:

$$NDWI = \frac{B3 - B8}{B3 + B8} \quad (1)$$

ove $B3$ = Banda del verde
 $B8$ = Banda dell’infrarosso vicino

A causa della natura del dato SAR, il procedimento per l’estrazione della linea di riva dall’immagine Sentinel-1, ha dovuto comprendere alcuni passaggi di preelaborazione, eseguiti tramite il software open-source SNAP di ESA:

- utilizzo delle orbite ricalcolate tramite i successivi passaggi del satellite dalla data di acquisizione per ottenere una migliore coregistrazione;
- rimozione del contributo di rumore termico: questo passaggio rimuove l'effetto del rumore al sensore radar;
- calibrazione: questa fase converte i valori del segnale radar in unità di backscatter misurando l'intensità del segnale riflettente riducendo dunque l'effetto degli errori nel sistema di misurazione e acquisizione;
- filtro per lo "Speckle": questa fase reduce il tipico rumore gaussiano presente nell'immagine radar generato dall'interazione del segnale con la superficie ruvida terrestre. È stato utilizzato un filtro di Lee per la sua efficacia nel mantenere le informazioni sui bordi e ridurre il rumore;
- correzione del terreno: questa fase corregge le distorsioni nell'immagine causate dalla topografia della superficie terrestre. L'immagine è stata georeferenziata nel sistema di riferimento WGS-84 UTM-32N utilizzando il modello digitale del terreno Shuttle Radar Topography Mission (SRTM).

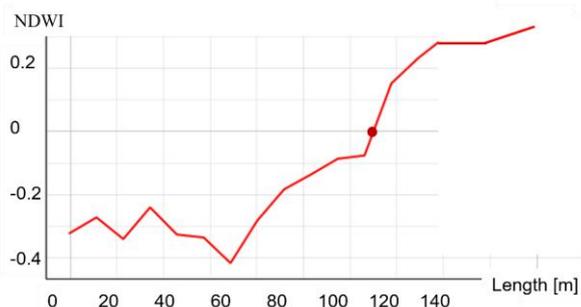
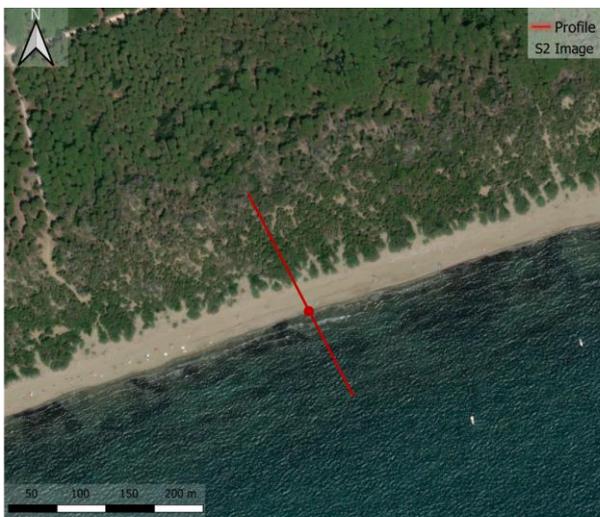


Figura 2. valori dell'indice NDWI lungo un transetto. Il punto rosso rappresenta il valore dell'indice NDWI utilizzato per la separazione di acqua e terraferma.

Sono state adottate diverse polarizzazioni e dimensioni della finestra per il filtro di Lee. Ad una maggiore dimensione della finestra corrisponde una maggiore superficie di interpolazione di dati con una possibile perdita di informazioni. Infine, sono state selezionate le seguenti configurazioni, :

- polarizzazione VH con finestra per il filtro di Lee di dimensione 7 * 7px;

- polarizzazione VV con finestra per il filtro di Lee di dimensione 7 * 7px;
- polarizzazione VV con finestra per il filtro di Lee di dimensione 5 * 5px

Per binarizzare l'immagine SAR in un contesto specifico di spiaggia sabbiosa è stata adottata una soglia di -20dB, seguendo le indicazioni di (Chen et al., 2022).

Invece, per separare le aree di terra e acqua nell'immagine raster rappresentante l'NDWI, è stato adottato un valore di soglia di 0.

Le immagini risultanti dall'analisi radar e multispettrale sono state vettorializzate, in formato linea, e in seguito smussate in QGIS per migliorare il risultato scendendo ad un livello sub pixel, come illustrato nel diagramma della Figura 3.

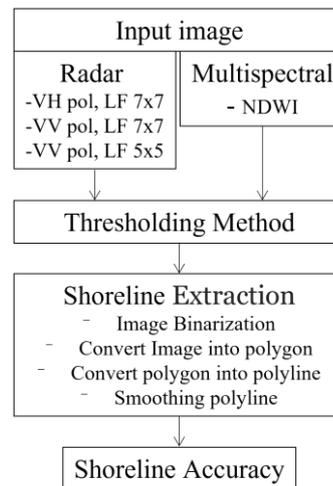


Figura 3. Diagramma rappresentante la metodologia di "thresholding" nell'estrazione della linea di riva.

3.3 Metodi di Classificazione

I metodi di classificazione considerati per l'estrazione della linea di costa sono stati implementati nel software SNAP. Sono state utilizzate sia immagini radar che multispettrali, testati differenti approcci e sperimentati diversi metodi per evidenziare le differenze spettrali nelle immagini. Per quanto riguarda i dati radar, è stato calcolato il valore assoluto della polarizzazione VH moltiplicato per la polarizzazione VV. In relazione ai dati ottici, sono stati testati l'indice spettrale NDWI e diverse combinazioni di bande.

La metodologia per estrarre la linea di costa dopo la classificazione dell'immagine è illustrata nella Figura 4. Le strategie di classificazione possono essere suddivise in diverse categorie in base a:

- classificazione basata sui singoli pixel (pixel-based) o per oggetti (object-oriented);
- Classificazione supervisionata (necessita di una fase di training) o non supervisionata.

In questo lavoro, la classificazione basata sui pixel è stata testata tramite il Maximum Likelihood, il Minimum Distance e il Random Forest, come metodi supervisionati. L'algoritmo K-means è stato scelto per la classificazione non supervisionata. Le immagini satellitari sono state ritagliate e comprendono un'area dove sono state scelte 4 classi: area urbana, area vegetata, suoli e acqua. I passaggi a valle della fase di classificazione, svolti per l'Estrazione della linea di riva, sono sovrapponibili a quelli già

descritte nella sottosezione 3.2 relativa alla metodologia di “thresholding”.

3.3.1 Maximum Likelihood (ML)

Il classificatore “Maximum Likelihood” è un metodo ampiamente utilizzato per la classificazione delle immagini, in cui le classi dei pixel vengono assegnate in base al criterio di massima verosimiglianza. La separazione tra le classi nello spazio decisionale è uno dei principali fattori che influenzano l'accuratezza della classificazione basata su ML (Ahmad e Quegan, 2012).

3.3.2 Minimum Distance (MD)

Il classificatore “Minimum Distance” si basa sulla disponibilità di valori centrali delle classi (ad esempio, medie, determinate nella fase di addestramento): le classi vengono assegnate ai pixel sulla base del criterio di distanza minima per ciascun valore centrale della classe (B. R. Shivakumar e S. V. Rajashekararadhya, 2017). Questo metodo viene spesso utilizzato nelle applicazioni di telerilevamento e analisi delle immagini grazie alla sua semplicità ed efficienza.

3.3.3 Random Forest (RF)

Il "Random Forest" è un metodo di apprendimento automatico che genera un insieme di alberi decisionali, ognuno costruito a partire da un campione casuale di dati. Ciascun albero fornisce una previsione di classe per un dato input, e la classe finale assegnata a tale input è quella che compare con maggiore frequenza nelle previsioni di tutti gli alberi. Questo approccio consente di aumentare l'accuratezza del modello e di prevenire il sovradattamento, ovvero la tendenza del modello ad adattarsi eccessivamente ai dati di addestramento, perdendo la capacità di generalizzare su nuovi dati non visti. L'output della classificazione di RF corrisponde alla moda statistica di numerosi alberi decisionali, risultando in un modello più robusto rispetto a un unico albero di classificazione ottenuto da un singolo ciclo del modello (Breiman, 2001). Nel contesto della regressione, l'output di RF rappresenta la media di tutti gli alberi di regressione sviluppati in parallelo, senza alcuna potatura. RF offre diverse potenzialità, come la stima interna degli errori, la capacità di valutare la rilevanza delle variabili e la competenza nel trattare variabili di scarso impatto. Prendendo in considerazione il problema di classificazione dell'area costiera, l'approccio RF si è dimostrato efficace quando i dati di input integravano la banda NIR (Dizaji, 2018).

3.3.4 K-means

Il K-means è un algoritmo di clustering che presuppone di conoscere a priori il numero delle classi, o cluster, (k) in cui si desidera suddividere il dataset. L'algoritmo assegna ogni pixel della scena al centro del cluster più vicino, che viene determinato come il valore medio di tutti i campioni appartenenti al cluster. La sua semplicità e l'efficienza computazionale lo rendono una scelta popolare per il clustering di grandi dataset. Tuttavia, non tiene conto delle variazioni di scala dei dati e delle correlazioni. Inoltre, minimizza la varianza intra-cluster, ma non è garantito il raggiungimento di un minimo globale. Nonostante queste limitazioni, è una delle scelte più comuni per l'analisi dei cluster.

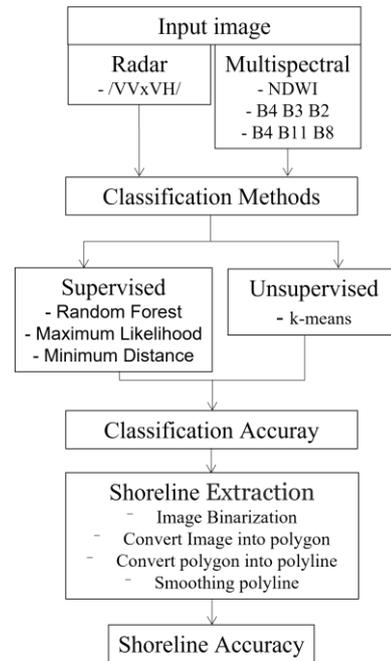


Figura 4. Diagramma rappresentante le metodologie di classificazione usate per l'estrazione della linea.

3.3.5 Valutazione dell'accuratezza della classificazione

La valutazione dell'accuratezza è stata ottenuta confrontando i risultati dei metodi considerati con una classificazione eseguita manualmente basata sui pixel. Il modo comune per riportare gli errori è attraverso la matrice di confusione, che evidenzia le classificazioni errate per ogni gruppo. La matrice degli errori è sempre presentata come una matrice quadrata, e gli elementi diagonali indicano il numero di campioni classificati correttamente. Al contrario, gli elementi fuori dalla diagonale forniscono informazioni sugli errori di omissione e commissione. Il criterio scelto per definire l'accuratezza è stata l'accuratezza complessiva (OA), definita sommando il numero di valori classificati correttamente e dividendo per il numero totale di campioni.

$$OA = \frac{(TP+TN)}{(TP+FP+TN+FN)} \quad (2)$$

La Tabella 2 mostra le performance di alcuni dei metodi di classificazione supervisionata esaminati e i rispettivi dati di partenza. La Figura 5 confronta i risultati della classificazione sui dati S1 in alcuni dei casi considerati.

Piattaforma	Dati	Metodo	OA
S1	/VVxVH/	ML	0.852
S1	/VVxVH/	MD	0.847
S1	/VVxVH/	RF	0.847
S2	NDWI	ML	0.169
S2	NDWI	MD	0.169
S2	NDWI	RF	0.169

Tabella 2. Accuratezza dei metodi di classificazione in termini del parametro Overall Accuracy (OA).

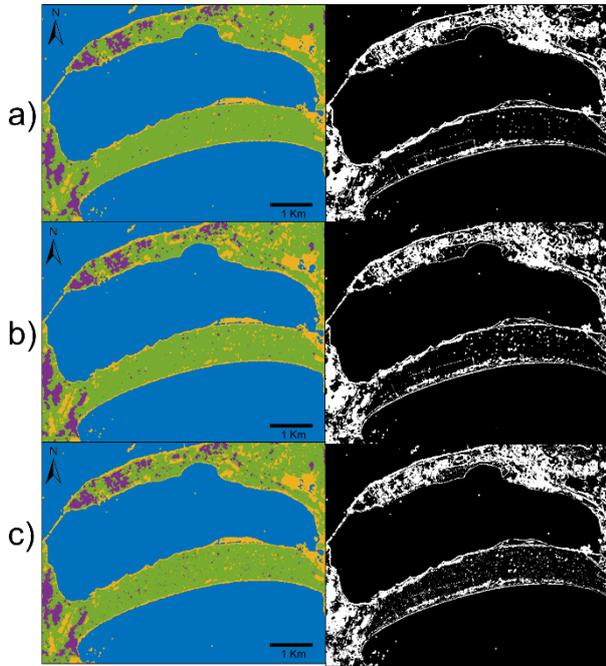


Figura 5. Risultati della classificazione (colonna sinistra) sui dati S1 e i corrispondenti pixel classificati erroneamente (colonna destra) per a) MD sulla banda /VVxVH/, b) ML sulla banda /VVxVH/, c) RF sulla banda /VVxVH/. Classi considerate: vegetazione (verde), aree urbane (viola), acqua (blu) e suolo (giallo).

4. Risultati e discussione

Le diverse linee di costa ottenute tramite le varie metodologie sono state confrontate con quella digitalizzata manualmente a partire dall'ortomosaico ad alta risoluzione menzionato nella sezione 3.1. Il metodo descritto da (Demir et al., 2017) è stato implementato in ambiente GIS per valutare l'accuratezza nell'estrazione automatica della linea di riva. Ciascuna linea di costa è stata campionata ogni 10m per un totale di circa 600 punti lungo tutta la lunghezza del tombolo della Feniglia. Successivamente, sono stati generati dei transetti per ciascun punto e su essi è stata misurata la distanza euclidea con la linea digitalizzata manualmente, come illustrato in Figura 6.



Figura 6. Valutazione dell'accuratezza nella generazione delle linee di riva, in termini di distanza euclidea da quella di riferimento.

Sono stati calcolati parametri statistici, tra cui distanza minima, massima, media e deviazione standard, al fine di valutare i risultati delle due metodologie, riportati nella Tabella 3.

Thresholding					
		Min dis [m]	Max dis [m]	Avg dis [m]	Std [m]
S1	VH LF 7x7	0.08	73.76	24.66	11.14
	VV LF 7x7	0.00	44.47	7.82	6.49
	VV LF 5x5	0.02	34.00	7.77	6.21
S2	NDWI	0.02	20.59	6.97	3.95
Classification					
		Min dis [m]	Max dis [m]	Avg dis [m]	Std [m]
S1 /VVxVH/	K-means	0.04	24.66	7.55	5.16
	RF	0.01	25.86	9.00	5.51
	ML	0.02	29.04	11.67	5.80
	MD	0.01	25.23	8.81	5.45
S2 B4B3B2	K-means	0.05	17.94	6.24	3.34
S2 B4B8B11		0.08	15.46	3.75	2.65
S2 NDWI	RF	0.02	20.32	4.95	3.13
	ML	0.02	20.32	4.95	3.13
	MD	0.02	20.32	4.95	3.13

Tabella 3. I risultati statistici sono stati derivati dalle misurazioni delle distanze tra la linea di costa stimata automaticamente e quella di riferimento estratta manualmente.

I risultati mostrano che l'algoritmo di classificazione non supervisionato k-means, applicato ai dati ottici con la combinazione di bande B4-B8-B11, ha ottenuto le migliori prestazioni. Inoltre, il trattamento preliminare dei dati tramite il calcolo dell'indice NDWI si è rivelata un'ottima strategia. Per quanto riguarda i dati Sentinel-1, l'algoritmo k-means applicato alla combinazione di polarizzazioni derivante dal valore assoluto del prodotto tra VV e VH, ha fornito le migliori prestazioni. Le combinazioni di bande, sensori e metodologie menzionate hanno raggiunto una notevole accuratezza dell'ordine del sub-pixel.

Inoltre, è stato effettuato un confronto in tre settori della spiaggia, ciascuno lungo circa 2 km, basato sulla distanza media dalla linea di costa di riferimento, al fine di valutare la presenza di eventuali differenze nella distribuzione spaziale delle distanze (come mostrato nella Figura 5).

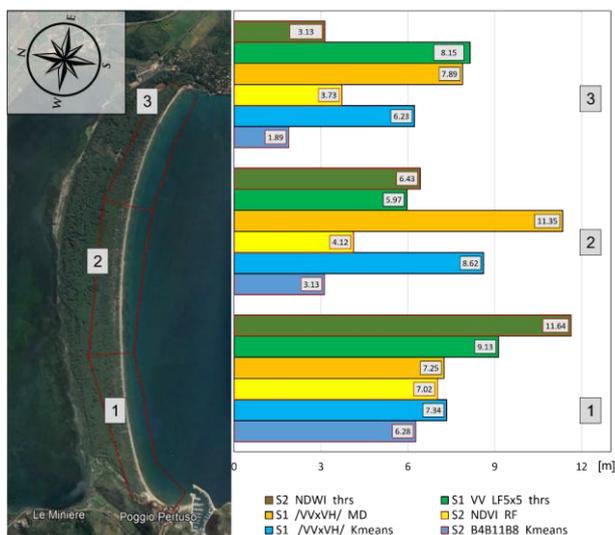


Figura 7. Distribuzione per settore delle distanze dalla linea di riva di riferimento di alcune metodologie selezionate.

Un'analisi accurata dei risultati evidenzia un miglioramento dell'accuratezza muovendosi da ovest a est, dal settore 1 al settore 3. Questo andamento è stato rilevato in tutte le analisi condotte utilizzando dati ottici, mentre le prestazioni del radar mostrano una maggiore variabilità, senza una tendenza rilevabile. Una possibile spiegazione del trend riscontrato per i dati ottici potrebbe essere correlata al lieve aumento della pendenza della spiaggia e delle dimensioni granulometrica muovendosi dal settore 1 al settore 3. Tali considerazioni necessitano di ulteriori conferme.

L'area di studio selezionata si è rivelata particolarmente impegnativa a causa della ridotta profondità della spiaggia, variabile tra i 10 e i 50 metri e paragonabile alla dimensione del pixel dell'immagine satellitare. Un altro aspetto da considerare è la bassa pendenza della spiaggia. La rilevazione della linea di costa con il metodo proposto può essere impegnativa nelle aree di intermareali, dove il mezzo poroso presenta un grado di saturazione maggiore, portando a incertezze nella rilevazione della linea di costa (Spinosa et al., 2021b).

Un significativo fattore di incertezza è la riflettanza, che in ambienti con un alto contenuto di umidità risulta simile a quella dell'acqua. Infatti, alla saturazione, la lunghezza del percorso ottico nell'acqua è al suo massimo e alcune lunghezze d'onda specifiche possono essere completamente assorbite (Nolet et al., 2014). Un altro aspetto da considerare per migliorare le analisi in tale contesto è considerare pienamente le differenze di risalita causate da onde e maree. Per estendere la metodologia ad altre aree, è necessario anche considerare le diverse granulometrie presenti in diverse spiagge sabbiose.

Per quanto riguarda la metodologia di "thresholding", al fine di migliorare i risultati, varrebbe la pena esplorare metodi che prevedano soglie adattive a livello locale, in modo da ovviare alle differenze di luminosità e contrasto presenti in diverse porzioni dell'immagine analizzata (Liu e Jezek, 2004). In riferimento ai metodi di classificazione, nel caso non supervisionato, l'Analisi

delle Componenti Principali (PCA) può essere utilizzata per una determinazione ad hoc del numero di classi (Hannv et al., 2013). Nel caso si utilizzino algoritmi di classificazione supervisionata, grazie alla loro forte capacità di gestire fenomeni complessi, le reti neurali potrebbero migliorare l'accuratezza del rilevamento della linea di costa (Tsekouras et al., 2018b).

5. Conclusioni

L'acquisizione accurata della linea di costa riveste un ruolo importante in ambito scientifico, data la sua rilevanza rispetto a tematiche ambientali e socio-economiche. Fenomeni come l'erosione costiera, l'innalzamento del livello del mare e in generale la gestione del territorio costiero sono esempi emblematici di contesti in cui l'identificazione accurata della linea di costa assume un'importanza cruciale.

Nel presente studio, sono state implementate diverse metodologie per l'estrazione della linea di costa a partire da immagini satellitari, sia radar che multispettrali. Sono state testate tecniche che prevedono la scelta di una soglia ("thresholding") per la distinzione acqua/suolo e successive strategie di classificazione dell'immagine. L'accuratezza di entrambe le metodologie è stata valutata digitalizzando una linea di riva di riferimento tramite un'ortofoto ad alta risoluzione. Inoltre, l'accuratezza della classificazione dell'immagine è stata ulteriormente valutata mediante un riferimento ("Ground Truth") ottenuto attraverso una classificazione manuale. Ciò ha permesso una valutazione più completa dell'accuratezza delle metodologie e ha fornito una ulteriore validazione dei risultati. Entrambe le metodologie hanno raggiunto livelli di accuratezza inferiore al mezzo pixel (la dimensione del singolo pixel delle immagini satellitari utilizzate corrispondeva a 10x10m). Il risultato ottenuto è notevole, considerando la complessità dell'area esaminata.

Le metodologie implementate in questo studio presentano il vantaggio di essere eseguibili con software e dati gratuiti oltre ad essere riproducibili in scenari diversi.

L'utilizzo di immagini satellitari per l'estrazione della linea di costa offre benefici quali la capacità di coprire vaste aree a costi ridotti rispetto ai metodi tradizionali, come la fotografia aerea e i rilievi sul campo. La metodologia proposta ha il pregio di essere adattabile sia ai dati ottici che radar, che possono essere utilizzati in modo complementare. Tra i possibili sviluppi futuri si possono annoverare: le tecniche di fusione di dati e il perfezionamento dell'utilizzo di algoritmi di Machine Learning.

Bibliografia

Abdelhady, H.U., Troy, C.D., Habib, A., Manish, R., 2022. A Simple, Fully Automated Shoreline Detection Algorithm for High-Resolution Multi-Spectral Imagery. *Remote Sens.* 14, 557. <https://doi.org/10.3390/rs14030557>

Aedla, R., Dwarakish, G.S., Reddy, D.V., 2015. Automatic Shoreline Detection and Change Detection Analysis of Netravati-GurpurRivermouth Using Histogram Equalization and Adaptive Thresholding Techniques. *Aquat. Procedia, INTERNATIONAL CONFERENCE ON WATER RESOURCES, COASTAL AND OCEAN ENGINEERING (ICWRCOE'15)* 4, 563–570. <https://doi.org/10.1016/j.aqpro.2015.02.073>

- Ahmad, A., Quegan, S., 2012. Analysis of maximum likelihood classification on multispectral data. *Appl. Math. Sci.* 6, 64256436.
- Anders, F.J., Byrnes, M., 1991. Accuracy of shoreline change rates as determined from maps and aerial photographs. *Shore Beach* 59, 17–26.
- B R, Shivakumar., S V, Rajashekaradhyaa., 2017. Spectral Similarity for Evaluating Classification Performance of Traditional Classifiers. <https://doi.org/10.1109/WiSPNET.2017.8300111>
- Bamdadinejad, M., Ketabdari, M.J., Chavooshi, S.M.H., 2021. Shoreline Extraction Using Image Processing of Satellite Imageries. *J. Indian Soc. Remote Sens.* 49, 2365–2375. <https://doi.org/10.1007/s12524-021-01398-3>
- Breiman, L., 2001. Random Forests. *Mach. Learn.* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Barnes, R.S.K., 1980. Coastal Lagoons. CUP Archive.
- Boak, E., Turner, I., 2005. Shoreline Definition and Detection: A Review. *J. Coast. Res. J. Coast. Res.* 21, 688–703. <https://doi.org/10.2112/03-0071.1>
- Chen, C., Chen, H., Liang, J., Huang, W., Xu, W., Li, B., Wang, J., 2022. Extraction of Water Body Information from Remote Sensing Imagery While Considering Greenness and Wetness Based on Tasseled Cap Transformation. *Remote Sens.* 14, 3001. <https://doi.org/10.3390/rs14133001>
- Cipriani, L.; Ferri, S.; Iannotta, P.; Mannori, S.; Pranzini, E. Beach erosion and protection in Tuscany Monitoring results: a Capitalization tool from COASTGAP Project Erosione Costiera 2004, 75–92.
- Demir, N., Oy, S., Erdem, F., Şeker, D.Z., Bayram, B., 2017. INTEGRATED SHORELINE EXTRACTION APPROACH WITH USE OF RASAT MS AND SENTINEL-1A SAR IMAGES. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* IV-2-W4, 445–449. <https://doi.org/10.5194/isprs-annals-IV2-W4-445-2017>
- Dizaji, R., 2018. Coastline Extraction By Using Random Forest Method; A Case Study Of Istanbul, *Geomatik.* <https://doi.org/10.29128/geomatik.362179>
- García-Rubio, G., Huntley, D., Russell, P., 2015. Evaluating shoreline identification using optical satellite images. *Mar. Geol.* 359, 96–105. <https://doi.org/10.1016/j.margeo.2014.11.002>
- Hannv, Z., Qigang, J., Jiang, X., 2013. Coastline Extraction Using Support Vector Machine from Remote Sensing Image. *J. Multimed.* 8, 175–182. <https://doi.org/10.4304/jmm.8.2.175-182>
- Jishuang, Q., Wang, C., 2002. A MULTI-THRESHOLD BASED MORPHOLOGICAL APPROACH FOR EXTRACTING COASTAL LINE FEATURE IN REMOTE SENSED IMAGES.
- Jishuang, Q., Wang, C., 2002. A MULTI-THRESHOLD BASED MORPHOLOGICAL APPROACH FOR EXTRACTING COASTAL LINE FEATURE IN REMOTE SENSED IMAGES.
- Klinger, T., Ziem, M., Heipke, C., Schenke, H.W., Ott, N., 2011. Antarctic Coastline Detection using Snakes. *Photogramm. - Fernerkund. - Geoinformation* 421–434. <https://doi.org/10.1127/1432-8364/2011/0095>
- Labuz, T.A., 2015. Environmental Impacts—Coastal Erosion and Coastline Changes, in: The BACC II Author Team (Ed.), *Second Assessment of Climate Change for the Baltic Sea Basin, Regional Climate Studies.* Springer International Publishing, Cham, pp. 381–396. https://doi.org/10.1007/978-3-319-16006-1_20
- Liu, H., Jezek, K.C., 2004. Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods. *Int. J. Remote Sens.* 25, 937–958. <https://doi.org/10.1080/0143116031000139890>
- Liu, Z., Li, F., Li, N., Zhang, H., 2016. A Novel Region-Merging Approach for Coastline Extraction From Sentinel-1A IW Mode SAR Imagery. *IEEE Geosci. Remote Sens. Lett.* <https://doi.org/10.1109/LGRS.2015.2510745>
- Nolet, C., Poortinga, A., Roosjen, P., Bartholomeus, H., Ruessink, G., 2014. Measuring and Modeling the Effect of Surface Moisture on the Spectral Reflectance of Coastal Beach Sand. *PLOS ONE* 9, e112151. <https://doi.org/10.1371/journal.pone.0112151>
- Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* 9, 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Ouyang, Y., Chong, J., Wu, Y., 2010. Two coastline detection methods in Synthetic Aperture Radar imagery based on Level Set Algorithm. *Int. J. Remote Sens.* 31, 4957–4968. <https://doi.org/10.1080/01431161.2010.485142>
- Ozturk, D., Sesli, F.A., 2015. Shoreline change analysis of the Kizilirmak Lagoon Series. *Ocean Coast. Manag., Special Issue: Third International Symposium on Integrated Coastal Zone Management (ICZM): Towards Sustainable Coasts - “Recent developments and advancements”* 118, 290–308. <https://doi.org/10.1016/j.ocecoaman.2015.03.009>
- Spinosa, A., Ziemba, A., Saponieri, A., Damiani, L., El Serafy, G., 2021a. Remote Sensing-Based Automatic Detection of Shoreline Position: A Case Study in Apulia Region. *J. Mar. Sci. Eng.* 9, 575. <https://doi.org/10.3390/jmse9060575>
- Tonye, E., Akono, A., Nyongui, A., Nlend, C., Rudant, J.-P., 2000. Cartographie des traits de côte par analyse texturale d’images radar à synthèse d’ouverture ERS-1 et E-SAR. *Teledetection* 3, 183–204.
- Toure, S., Diop, O., Kpalma, K., Maiga, A.S., 2019. Shoreline Detection using Optical Remote Sensing: A Review. *ISPRS Int. J. Geo-Inf.* 8, 75. <https://doi.org/10.3390/ijgi8020075>
- Tsekouras, G.E., Trygonis, V., Maniatopoulos, A., Rigos, A., Chatzipavlis, A., Tsimikas, J., Mitianoudis, N., Velegrakis, A.F., 2018. A Hermite neural network incorporating artificial bee colony optimization to model shoreline realignment at a reef-fronted beach. *Neurocomputing, Applications of Neural Modeling in the new era for data and IT* 280, 32–45. <https://doi.org/10.1016/j.neucom.2017.07.070>
- Zollini, S., Alicandro, M., Cuevas-González, M., Baiocchi, V., Dominici, D., Buscema, P.M., 2020. Shoreline Extraction Based on an Active Connection Matrix (ACM) Image Enhancement Strategy. *J. Mar. Sci.* <https://doi.org/10.3390/jmse8010009> Eng. 8, 9.

