# CO-REGISTRAZIONE DI IMMAGINI MULTISPETTRALI E TERMICHE ASINCRONEON THE CO-REGISTRATION OF ASYNCHRONOUS MULTI-SPECTRAL AND THERMAL IMAGES

I. Cortesi<sup>1, 2</sup>,\*A. Masiero<sup>1</sup>, G. Tucci<sup>1</sup>

 <sup>1</sup> Department of Civil and Environmental Engineering, University of Florence, Florence 50139, Italy - (irene.cortesi, andrea.masiero, grazia.tucci)@unifi.it
<sup>2</sup> Department of Geodesy and Geoinformation, TU Wien, 1040 Vienna, Austria

Key words: Co-registration, Multi-spectral camera, Thermal camera, Plastic detection, UAV

Parole chiave: Co-registrazione, Camera multispettrale, Camera Termica, Individuazione di plastica, VAPR

## Abstract

Plastic pollution has a severe impact on the ecosystem, altering its natural equilibrium and causing serious health issues to both flora and fauna. Several actions have already been undertaken in order to reduce the plastic litter dispersion in the environment, both in terms of changing the human behavior, reducing the use of plastics and avoiding their dispersion, and of implementing methods for detecting and collecting the already dispersed ones. This paper focuses on the latter, and, in particular, on plastic litter detection on the fluvial environment. To this aim, an Unmanned Aerial Vehicle, provided with a multi-spectral and a thermal camera, have been used, in order to: (i) allow affordable periodic monitoring of relatively long river reaches, (ii) detect even quite small macro-plastics, based on their spectral signature. More specifically, since the cameras deployed in our data collection campaigns are not synchron-ized, this work aims at presenting the developed strategy for the co-registration of the acquired imagery, which results to be quite challenging given the few amount of visual features recognizable on the images acquired flying at a limited altitude over a river. The proposed methodology, which is based on the correlation maximization between multi-spectral and thermal images, provided reas-onable results on the considered case study. The obtained values of normalized intersection over union of plastic areas are over 90%.

L'inquinamento da parte dei rifiuti di plastica ha un grave impatto sull'ecosistema, alterando il suo equilibrio naturale e causando seri problemi alla flora e alla fauna. Negli ultimi anni sono già state intraprese diverse azioni per ridurre la dispersione di rifiuti di plastica nell'ambiente sia in termini di cambiamento del comportamento umano (riducendone l'uso ed evitandone la dispersione) sia nell'implementazione di metodi capaci di rilevare e raccogliere i materiali abbandonati. Questo articolo si concentra su quest'ultimo aspetto e, in particolare, sul rilevamento di rifiuti di plastica nell'ambiente fluviale. In questo contesto è stato utilizzato un Veicolo Aereo a Pilotaggio Remoto (VAPR) dotato di una camera multispettrale e una termica, al fine di: (i) consentire un monitoraggio periodico di tratti di fiume relativamente lunghi, (ii) rilevare anche macroplastiche di dimensioni ridotte, basandosi sulla loro firma spettrale. Poichè le camere impiegate nelle campagne di raccolta dati non sono sincronizzate, questo lavoro si propone di presentare la strategia sviluppata per la co-registrazione delle immagini acquisite, che si rivela piuttosto complessa a causa della scarsa quantità di visual features riconoscibili sulle immagini acquisite volando a bassa quota su un fiume. La metodologia proposta, basata sulla massimizzazione della correlazione tra immagini multispettrali e termiche, ha fornito risultati ragionevoli nel caso di studio considerato. I valori ottenuti del Normalized Intersection over Union delle aree di plastica superano 90%.

#### 1. Introduzione

L'inquinamento dovuto alla plastica è una delle minacce più discusse e tristemente note del nostro tempo. Ogni ecosistema sulla Terra è colpito da questo grave problema causato dall'uomo. L'uso costante di materiale plastico e il suo errato smaltimento mettono seriamente a rischio la salute di vari ecosistemi, compresi gli animali che vi abitano (Plastics Europe, 2022, United Nations Environment Program, 2018). Oltre alle consolidate cattive abitudini umane, la diffusione della pandemia di COVID-19 ha aumentato questo fenomeno incoraggiando l'utilizzo di dispositivi monouso confezionati singolarmente (Peng et al., 2021).

Date l'entità e la gravità di questa minaccia, diversi approcci sono stati proposti dalla comunità del telerilevamento durante l'ultimo decennio al fine di rilevare dei rifiuti di plastica diffusi nell'ambiente. In particolare, nel corso degli anni sono stati utilizzati diversi strumenti e metodi per rilevare e monitorare oggetti di plastica galleggianti in acqua dolce e salata. Alcuni ricercatori hanno utilizzato censimenti visivi in situ per il loro lavoro (Galgani et al., 2013, Geraeds et al., 2019), ma la maggior parte ha utilizzato immagini o ortofoto ottenute da satelliti (Tasseron et al., 2021, Themistocleous et al., 2020, Topouzelis et al., 2020) o immagini acquisite con Veicoli Aerei a Pilotaggio Remoto (VAPR) (Cortesi et al., 2022, Jakovljevic et al., 2020,

<sup>\*</sup> Corresponding author

Iordache et al., 2022, Cortesi et al., 2023).

L'utilizzo dei droni consente non solo una grande flessibilità nella scelta dell'area di studio e nella pianificazione delle tempempistiche, ma offre anche una vasta gamma di sensori disponibili, che possono migliorare le prestazioni complessive dell'individuazione della plastica. In particolare, (Cortesi et al., 2022) suggerisce che la combinazione di dati multispettrali (nella gamma di lunghezze d'onda dello spettro elettromagnetico da 433 a 875 nm) e dati termici potrebbe essere utilizzata al fine di ridurre i falsi positivi.

In accordo con quanto appena detto, il lavoro presentato in questo articolo, che fa parte di un progetto che affronta il problema dell'individuazione dei rifiuti di plastica in ambienti fluviali, mira a generare un'immagine multibanda che includa bande dello spettro elettromagnetico nel visibile, nel vicino infrarosso e nel termico. L'immagine multispettrale prodotta mapperà la scena come se fosse acquisita per ogni banda con lo stesso orientamento della camera e al medesimo istante. Successivamente, il set di immagini risultanti dalla procedura sviluppata sarà utilizzato in un lavoro futuro al fine di testare e migliorare l'approccio di rilevamento dei rifiuti di plastica precedentemente proposto in (Cortesi et al., 2022, Cortesi et al., 2021).

# 2. Materiali

La strumentazione utilizzata in questa ricerca include la camera multispettrale MAIA-S2 (prodotta da SAL Engineering and EOPTIS, Italia) e la camera termica DJI H20T equipaggiate sul drone DJI Matrice 300, come è mostato in Figura 1. La raccolta dati è avvenuta attraverso un volo sopra l'area di studio, una porzione del fiume Ombrone (vicino Grosseto, Italia, Figura 2). L'area mappata da una singola immagine solitamente include pochi elementi: l'acqua del fiume, i campioni di plastica (sei in totale, differenti per dimensione, colore, texture e polimero), e in alcuni casi la riva del fiume e la vegetazione ripariale.



Figura 1. Camera multispettrale MAIA-S2, camera termica H20T DJI, antenna GNSS MAIA, Sensore ILS e antenna GNSS aggiuntiva montati sul Matrice 300 DJI.

MAIA-S2, che è stata montata sul Matrice 300 attraverso un giunto rigido, acquisisce immagini multibanda ad alta risoluzione (velocità di acquisizione massima 1 Hz) in diverse lunghezze d'onda nelle regioni del visibile (VIS) e del vicino infrarosso (NIR) attraverso un array di 9 camere (SAL Engineering and EOPTIS, 2018). Le caratteristiche della camera sono mostrate nella Tabella 1.



Figura 2. Drone in volo sopra l'area di studio (Istia d'Ombrone, Grosseto, Italia).

La camera MAIA-S2 può comunicare con un ricevitore GNSS (RTK) esterno montato sul drone, con correzioni provenienti da una stazione base GNSS posizionata vicino ( $\approx 100$  m) all'area di volo.

Gli output raw prodotti dalla camera multispettrale richiedono delle correzioni radiometriche e geometriche e una coregistrazione tra le nove bande. La correzione radiometrica avviene attraverso un software propretario di image-processing (dato in dotazione da SAL Engineering), sfruttando le misure del Sensore di Irraggiamento Luminoso (ILS), posizionato sulla parte superiore del DJI Matrice 300 (Figura 1).

La DJI H20T è una camera multisenore, che acquisisce sia immagini RGB che termiche (Tabella 1). Le immagini RGB e termiche sono acquisite a diversi istanti (siano  $t_{RGB}$  e  $t_{th}$ i tempi di acquisizione di una coppia di immagini RGB e termiche, quindi  $t_{RGB}/= t_{th}$ ; in entrambi i casi 2 Hz è la massima velocità di acquisizione), e sfortunatamente asincroni rispetto alle acquisizioni di MAIA-S2 (ad esempio, se  $t_{ms}$  è il tempo di acquisizione dell'immagine multispettrale temporalmente più vicina alle immagini RGB e termiche precedentemente menzionate, allora  $t_{ms}/=t_{RGB}$  e  $t_{ms}/=t_{th}$ ). Inoltre, DJI H20T essendo un sensore ufficiale DJI, possiede gimbal in grado di compensare (parzialmente) le vibrazioni e che consente di effettuare acquisizioni (prossime) al nadir. L'accuratezza della sincronizzazione temporale tra le acquisizioni H20T e il ricevitore GNSS (RTK) Matrice 300 non è garantita e pertanto le posizioni delle esposizioni basate sul GNSS sono solitamente determinate con un errore di livello inferiore al metro.

Il data-set iniziale includeva 654 immagini multispettrali e 414 immagini termiche acquisite a diverse quote tra i 10 e gli 80 metri sopra il livello del fiume. Le coppie di immagini selezionate sono formate dall'immagine termica e da quella multispettrale acquisite nell'istante temporale più simile, in base ai timestamp forniti da MAIA-S2 e DJI H20T. Poiché MAIA-S2 acquisisce le immagini a 1 Hz, l'intervallo di tempo più lungo

MAIA-S2			
Lunghezza focale	7.5 mm		
Diagonale FoV	43°		
Risoluzione immagine	agine $1208 \times 960$ pix		
Banda spettrale	433 - 875 nm		
Camera termica infrarosso			
Lunghezza focale	13.5 mm		
Diagonale FoV	40.6°		
Risoluzione immagine	640 × 512 pix		
Banda spettrale	$8000 - 14000 \ nm$		
Wide RGB Camera			
Lunghezza focale	4.5 mm		
Diagonale FoV	82.9°		
Risoluzione immagine	4056 × 3040 pix		
Banda spettrale	400 - 700 nm		

Tabella 1. Caratteristiche delle camere MAIA-S2 e DJI H20T.

possibile tra le due acquisizioni di una coppia di immagini è di 0,5 s.

Si assume, inoltre, che gli oggetti di plastica siano (abbastanza) statici. Vale la pena notare infatti che quest'ultima è un'ipotesi piuttosto inaffidabile in condizioni di lavoro generali, tuttavia è abbastanza realistica in questo caso di studio. Le plastiche sono state legate con lenze da pesca trasparenti per consentire il recupero di tali campioni introdotti artificialmente nel fiume solo per questo test. Pertanto, nonostante le plastiche si siano sicuramente mosse durante la raccolta dei dati, il loro movimento durante l'intervallo di tempo tra le acquisizioni delle immagini multispettrali e termiche corrispondenti è stato di solito piuttosto ridotto.

#### 3. Metodo

Questo lavoro si occupa del processo di co-registrazione di coppie di immagini multispettrali e termiche, nelle condizioni di lavoro menzionate nella sezione precedente.

Si presume che i parametri di orientamento interno di entrambe le camere (multispettrale e termica) siano pre-calibrati. Data l'acquisizione simultanea delle bande multispettrali, si assume, inoltre, che la loro relazione geometrica sia rigida e, di conseguenza, il sistema complessivo delle camere multispettrali possa essere pre-calibrato.

La co-registrazione tra immagini multispettrali e termiche risulta essere complicata a causa dalla loro relazione geometrica variabile, dovuta da (i) l'acquisizione asincrona, (ii) le diverse connessioni al drone (rigide/non rigide), (iii) il movimento del drone.

Considerando quanto appena descritto due opzioni alternative possono essere considerate per la co-registrazione delle immagini multispettrali e termiche:

 Inserire tutte le immagini termiche e multispettrali in un unico sistema di riferimento utilizzando un automatic bundle block adjustment. In questa maniera potrebbe essere possibile generare delle ortofoto multispettrali;

## 2. Lavorare con signole immagini.

La prima opzione è stata testata ma ha mostrato principalmente due problematiche. In primo luogo, questa applicazione può essere utilizzata solo con immagini acquisite ad altezze di volo elevate, in cui è visibile abbastanza texture (ad esempio, l'argine del fiume). Di conseguenza, la plastica nel fiume viene mappata solo su alcuni pixel. Inoltre, l'accuratezza nella registrazione tra il bundle block termico e multispettrale risulta di pochi pixel. Nonostante l'alto grado di automazione, questo approccio non poteva quindi essere utilizzato.

L'opzione 2, invece, è percorribile se vengono soddisfatte alcune ipotesi. Se è visibile solo l'acqua mappata o una piccola sezione dell'argine, la scena è essenzialmente piatta. Pertanto, le immagini di questa scena possono essere trasformate l'una nell'altra utilizzando un'omografia. In alternativa, come approssimazione, può essere utilizzata anche una trasformazione affine. In questo caso, l'accuratezza si è dimostrata essere più elevata. Una buona conoscenza dei parametri esterni della camera (ad esempio, nel caso della georeferenziazione diretta) o la disponibilità di alcune corrispondenze tra le due immagini possono essere sfruttate per stimare correttamente la trasformazione. La procedura adottata sarà descritta nel dettaglio nel seguito.

Il drone ha volato a bassa velocità e a diverse altezze, con la maggior parte delle immagini acquisite a quote piuttosto basse. È stato precedentemente osservato che un intervallo di altezza di volo compreso tra 20 m e 40 m è ragionevole per rilevare efficacemente i rifiuti di plastica (Cortesi et al., 2022). A tali quote, nella maggior parte delle immagini è visibile solo l'acqua (e i rifiuti di plastica, quando presenti) (Figura 3 e 4).

Poiché le immagini multispettrali e termiche sono state acquisite in modo asincrono, anche i movimenti dei rifiuti di plastica possono avere un impatto negativo sull'efficacia della co-registrazione. Tuttavia, date le condizioni di lavoro descritte nella sezione precedente, si presuppone che gli oggetti di plastica siano (quasi) statici durante l'intervallo di tempo (< 0, 5 s) tra l'acquisizione dell'immagine multispettrale e quella termica.

Date le caratteristiche del caso studio, le immagini rappresentano spesso una scena quasi piana. Pertanto, considerando un'immagine termica e quella multispettrale corrispondente (acquisita a meno di 0, 5 s prima o dopo l'immagine termica, come già menzionato) e assumendo una scena quasi statica, si presume che la distorsione sia già stata corretta su entrambe le immagini e che tutte le bande nell'immagine multispettrale siano già co-registrate, in modo tale da coniderare solo una di queste. Sarà presente una matrice di omografia, ovvero una matrice non singolare  $3 \times 3$ , tale che

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \cong H \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}$$
(1)

dove  $\begin{bmatrix} u & v & 1 \end{bmatrix}^{\top}$  sono coordinate omogenee di punti nell'immagine termica e  $\begin{bmatrix} u' & v' & 1 \end{bmatrix}^{\top}$  i corrispondenti nella multispettrale. L'uguaglianza approssimativa in (1) è chiaramente valida per un fattore di scala.

Gli oggetti di plastica sono in genere facilmente distinguibili nelle immagini multispettrali e termiche (escludendo sunglint,





Figura 3. Esempio di immagini multispettrale (rappresentata nelle bande Rosso, Verde e Blu, RGB) (a) termica (b). Si noti come la maggior parte delle immagini sia costituita da acqua.



Figura 4. Zoom della Figura 3 su campioni di plastica. (a) e (c) mostrano l'immagine multispettrale in una combinazione di bande RGB e (b) e (d) motrano l'immagine termica.

schiuma del fiume e rocce), come mostrato in Figura 3 e 4. Per tale motivo, solitamente è possibile utilizzare i metodi standard di segmentazione delle immagini per estrarre i pixel contenenti plastica dall'immagine, ad esempio con metodo di Otsu (Otsu, 1975).

Quando sono visibili almeno quattro oggetti di plastica nelle due immagini, è possibile ottenere facilmente una stima dell'omografia in (1) mediante il metodo della Trasformazione Lineare Diretta (DLT) (Hartley and Zisserman, 2003) utilizzando i centroidi delle plastiche come punti corrispondenti. Infatti, dal momento che le visual features non correlate agli oggetti di plastica sono raramente rilevabili e che il matching delle features tra le immagini multispettrali e termiche ha dato risultati non molto affidabili, l'utilizzo dei centroidi delle plastiche sembra attualmente il modo più affidabile per abbinare i punti nelle due immagini. Anche senza l'utilizzo di descrittori visivi, che in questo caso hanno purtroppo dato risultati non affidabili, le corrispondenze tra i centroidi delle plastiche nelle due immagini possono essere determinate automaticamente in alcune condizioni operative, ad esempio: (i) se sono disponibili buone stime dei parametri esterni della camera forniti dai sensori sulla piattaforma (caso di georeferenziazione diretta), (ii) sfruttando accurate posizioni di esposizione basate su GNSS e una procedura simile al Random Sample Consensus (RANSAC, (Fischler and Bolles, 1981)) per determinare le corrispondenze più affidabili, (iii) se le due viste delle camere sono molto simili, può essere utilizzato un approccio del nearest neighbor. Se i metodi automatici non funzionano, è necessario stabilire manualmente le corrispondenze.

In realtà, (i) non può essere utilizzato nel caso considerato, poiché non sono disponibili misurazioni affidabili dell'orientamento della camera. Inoltre, possono sorgere alcuni problemi anche nel caso (ii), in presenza di riflessi solari sull'acqua (sunglint), che possono introdurre molti punti non corrispondenti nelle regioni rilevate con una segmentazione automatica, poiché la firma spettrale del sunglint è simile a quella della plastica (Cortesi et al., 2022). RANSAC, o una delle sue varianti, può essere utilizzato per determinare gli outliers, le corrispondenze tra gli oggetti segmentati nelle due immagini e una stima della matrice di omografia H. Tuttavia, qualsiasi procedura simile a RANSAC probabilmente fallirà quando il numero di punti non corrispondenti sarà molto maggiore del numero di oggetti di plastica, come accade in alcune delle immagini acquisite (Cortesi et al., 2022). (iii) non può essere utilizzato perchè, dato che l'orientamento della camera MAIA-S2 non è compensato da un gimbal, le due viste delle camere possono essere diverse. Motivata da queste osservazioni, la procedura di segmentazione-associazione delle plastiche è attualmente assistita da un operatore umano nella nostra implementazione attuale.

Nella maggior parte delle immagini acquisite, il numero di oggetti di plastica visibili raramente  $e \ge 4$ . Inoltre, anche quando il numero di oggetti è sufficiente per determinare una matrice di omografia  $\hat{H}$  con DLT, sfruttando i centroidi delle plastiche come menzionato in precedenza, è evidente che  $\hat{H}$  è una stima dell'omografia tra le due immagini e quindi solo ragionevolmente vicina a quella ottimale. Per questa ragione, è stato implementato un approccio diverso sia per migliorare la stima ottenuta come descritto sopra, sia per risolvere il problema generale quando il numero di oggetti di plastica è insufficiente per applicare la procedura descritta in precedenza.

La logica dell'approccio implementato consiste nel determinare l'omografia tra le immagini multispettrali e termiche non distorte che massimizza la loro correlazione. Presumendo che la scena sia quasi piana l'omografia può essere scritta nel seguente modo (Ma et al., 2004):

$$H = K_{th} \left( R + \frac{\mathbf{t} \ \mathbf{n}^{\mathsf{T}}}{d} \right) K_{ms}^{-1}$$
<sup>(2)</sup>

dove  $K_{th} \in K_{ms}$  sono le due matrici di proiezione delle camere, che si presume siano note ,  $K_{ms}$ -<sup>1</sup> e l'inversa della matrice  $K_{ms}$ , (R, t) e la trasformazione rigida (rotazione, traslazione) tra i due sistemi di riferimento delle camere, n è la normale del piano della scena e d è la distanza tra la camera e il piano della scena. Date le caratteristiche del volo, d è solitamente molto più grande della lunghezza del vettore di traslazione t, quindi il contributo di t  $\mathbf{n}_{d}^{T}$  è relativamente piccolo per altezze di volo abbastanza elevate. Di conseguenza,

$$H \approx H_{\infty} = K_{th} R K_{ms}^{-1} \tag{3}$$

dove *R* dipende solo dai tre angoli relativi tra le due immagini, che devono essere ottimizzati (essendo di valore assoluto relativamente piccolo). Oltre a  $H_{\infty}$ , la trasformazione considerata tra le due immagini viene ottenuta anche ottimizzando una piccola traslazione nel piano dell'immagine, portando a 5 parametri da stimare.

La banda multispettrale che risulta essere più simile a quella termica è l'ultima banda NIR della MAIA-S2, S9 (855-875 nm), ed è quindi quella selezionata nella procedura di ottimizzazione. Tuttavia, nonostante siano abbastanza simili, purtroppo sono ancora visibili alcune differenze notevoli. La Figura 5 confronta la rappresentazione multispettrale (a) e termica (b) di un campione di plastica. L'immagine multispettrale nella Figura 5(a) è stata ottenuta mostrando le tre bande NIR della MAIA-S2, ovvero S7 (773-793 nm), S8 (784,5-899,5 nm) e S9 (vedi (Cortesi et al., 2022) per un elenco completo delle bande spettrali). Le tre bande MAIA-S2 mostrate nella figura sono quelle con l'aspetto tipicamente più simile a quello termico. Si noti anche il diverso comportamento della temperatura della plastica nella Figura 5(b) rispetto a quella nella Figura 4(b).



Figura 5. Zoom di un campione di plastica: (a) mostrano nelle tre bande del vicino infrarosso (NIR) nell'immagine multispettrale, (b) immagine termica.

Motivata dall'ultima osservazione, la correlazione è massimizzata tra le maschere delle immagini termiche e multispettrali, in cui sono stati estratti solo gli oggetti di plastica.

Per riassumere, il workflow implementato è il seguente:

 segmentazione delle plastiche sulle immagini termiche e multispettrali;

- individuazione manuale tra le corrispondenze degli oggetti di plastica;
- utilizzo delle maschere degli oggetti di plastica e le cor
  - rispondenze stabilite al passaggio precedente per stimare in modo ottimale l'omografia, sfruttando il modello ridotto (3).

#### 4. Risultati

Il metodo implementato è stato attualmente validato su 303 coppie immagini contenenti un totale di 494 istanze di oggetti di plastica.

Un esempio dei risultati ottenuti dalla procedura è mostrato nelle Figure 6 e 7, le quali riportano la co-registrazione delle immagini della Figura 3.



Figura 6. Risultato della co-registrazione delle immagini nella



Figura 7. Zoom della Figura 6 sui campioni di plastica.

Invece, la Tabella 2 e le Figure 8 e 9 mostrano i risultati ottenuti sull'intero dataset di validazione. Poiché' in alcuni casi l'aspetto delle plastiche nelle immagini multispettrali e termiche è piuttosto diverso, anche le corrispondenti regioni segmentate risultano talvolta abbastanza differenti. Per questa ragione la tabella e le figure mostrano i risultati ottenuti per quanto riguarda l'Intersection over Union (IoU) tra le regioni delle plastiche segmentate e coregistrate, ma tali valori sono stati normalizzati rispetto al loro massimo possibile nei casi considerati.

La Figura 8 mostra la distribuzione dei valori normalizzati di IoU sulle 494 istanze plastiche nel dataset. I risultati numerici corrispondenti (mediana, valore medio e Median Absolute

	Mediana	Media	MAD
Normalized IoU	93.92%	87.48%	6.08%
Mean normalized IoU	91.25%	89.71%	8.75%

Tabella 2. Risultati numerici della co-registrazione.

Deviation (MAD)) sono riportati nella prima riga della Tabella 2. Invece, la Figura 9 mostra la distribuzione della media dei valori normalizzati di IoU calcolati su ciascuna delle 303 immagini prese in considerazione (risultati numerici sulla seconda riga della Tabella 2).

Inoltre, occorre evidenziare che, come mostrano le Figure 10 e 11, i risultati ottenuti sono paragonabili anche su porzioni differenti del dataset, acquisite ad altezze molto diverse (15 e 80 metri nel caso delle figure in questione).

In generale i risultati ottenuti mostrano una performance ragionevole del metodo utilizzato (median normalized IoU > 90%), con poche istanze di plastica che presentano valori bassi di normalized IoU.



Figura 8. Distribuzione dei mean Normalized Intersection of Union values per tutte le plastiche presenti nel data-set considerato.



Figura 9. Distribuzione dei mean Normalized Intersection of Union values per tutte le immagini nel data-set considerato.



Figura 10. Distribuzione dei mean Normalized Intersection of Union values per tutte le immagini acquisite ad un altezza di volo di 15 metri.



Figura 11. Distribuzione dei mean Normalized Intersection of Union values per tutte le immagini acquisite ad un altezza di volo di 80 metri.

# 5. Conclusioni

Il lavoro presentato in questo articolo fa parte di un progetto volto a investigare la rilevazione di rifiuti di plastica in ambienti fluviali basata sulla loro firma spettrale, considerando bande nel visibile, nell'infrarosso vicino e termico. A tale scopo, è stata condotta un'acquisizione di dati utilizzando due camere, nello specifico la camera multispettrale MAIA-S2 e la camera termica DJI H20T, montate su un drone in grado di sorvolare l'area di studio. Sfortunatamente, le due camere hanno acquisito le immagini in modo asincrono e la loro relazione geometrica non è stata costante.

La presenza di poche visual features (principalmente legate agli stessi oggetti di plastica) nelle immagini acquisite e nelle condizioni di lavoro prese in considerazione ha spinto gli autori a implementare un processo di co-registrazione immagineimmagine, basato sulla massimizzazione della correlazione tra immagini multispettrali e termiche.

I risultati, ottenuti su un set di dati di convalida composto da 303 coppie di immagini, dimostrano una performance ragionevole del metodo, come indicato dal median normalized Intersection over Union superiore all'90%. Sviluppi futuri saranno dedicati alla convalida del metodo su un set di dati molto più ampio, attualmente composto circa 800 coppie di immagini multispettrali e termiche acquisiti in aree di studio diverse. Inoltre, saranno testate anche tecniche basate sul Deep Learning al fine di individuare corrispondenze affidabili tra immagini multispettrali e termiche (Sarlin et al., 2020). Sarà anche investigata la possibilità di tracciare gli spostamenti delle plastiche nel tempo.

#### Bibliografia

Cortesi, I., Masiero, A., De Giglio, M., Tucci, G., Dubbini, M., 2021. Random Forest-Based River Plastic Detection with a Handheld Multispectral Camera. *The International Archives* of Photogrammetry, Remote Sensing and Spatial Information Sciences, 43, 9–14.

Cortesi, I., Masiero, A., Tucci, G., Topouzelis, K., 2022. UAVbased river plastic detection with a multispectral camera. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B3-2022, 855–861.

Cortesi, I., Mugnai, F., Angelini, R., Masiero, A., 2023. Mini Uav-Based Litter Detection on River Banks. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10, 117–122.

Fischler, M., Bolles, R., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.

Galgani, F., Hanke, G., Werner, S., De Vrees, L., 2013. Marine litter within the European marine strategy framework directive. *ICES Journal of marine Science*, 70(6), 1055–1064.

Geraeds, M., van Emmerik, T., de Vries, R., bin Ab Razak, M. S., 2019. Riverine plastic litter monitoring using unmanned aerial vehicles (UAVs). *Remote Sensing*, 11(17), 2045.

Hartley, R., Zisserman, A., 2003. *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Iordache, M.-D., De Keukelaere, L., Moelans, R., Landuyt, L., Moshtaghi, M., Corradi, P., Knaeps, E., 2022. Targeting Plastics: Machine Learning Applied to Litter Detection in Aerial Multispectral Images. *Remote Sensing*, 14(22), 5820.

Jakovljevic, G., Govedarica, M., Alvarez-Taboada, F., 2020. A deep learning model for automatic plastic mapping using

unmanned aerial vehicle (UAV) data. Remote Sensing, 12(9), 1515.

Ma, Y., Soatto, S., Kosecka, J., Sastry, S., 2004. An Invitation to 3-D Vision: From Images to Geometric Models, Interdisciplinary Applied Mathematics: Imaging, Vision, and Graphics, Vol. 26. Springer.

Otsu, N., 1975. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296), 23–27.

Peng, Y., Wu, P., Schartup, A. T., Zhang, Y., 2021. Plastic waste release caused by COVID-19 and its fate in the global ocean. *Proceedings of the National Academy of Sciences*, 118(47), e2111530118.

Plastics Europe, 2022. Marine litter https://plasticseurope.org/sustainability/circularity/wastemanagement-prevention/marine-litter/. (Accessed 12 january 2022).

SAL Engineering and EOPTIS, 2018. Maia, the multispectral camera. https://www.spectralcam.com/maia-tech/. (Accessed 12 January 2022).

Sarlin, P.-E., DeTone, D., Malisiewicz, T., Rabinovich, A., 2020. Superglue: Learning feature matching with graph neural networks. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 4938–4947.

Tasseron, P., Van Emmerik, T., Peller, J., Schreyers, L., Biermann, L., 2021. Advancing Floating Macroplastic Detection from Space Using Experimental Hyperspectral Imagery. *Remote Sensing*, 13(12), 2335.

Themistocleous, K., Papoutsa, C., Michaelides, S., Hadjimitsis, D., 2020. Investigating detection of floating plastic litter from space using Sentinel-2 imagery. *Remote Sensing*, 12(16), 2648.

Topouzelis, K., Papageorgiou, D., Karagaitanakis, A., Papakonstantinou, A., Ballesteros, M. A., 2020. Plastic litter project 2019: Exploring the detection of floating plastic litter using drones and sentinel 2 satellite images. *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 6329–6332.

United Nations Environment Program, 2018. The state of plastics. world environment day outlook.