

DEEP LEARNING BASATO SU IMMAGINI DA UAV PER LA SUPER RISOLUZIONE DI DATI SATELLITARI



Nives Grasso¹, Andrea Maria Lingua¹, Marina Casabona¹

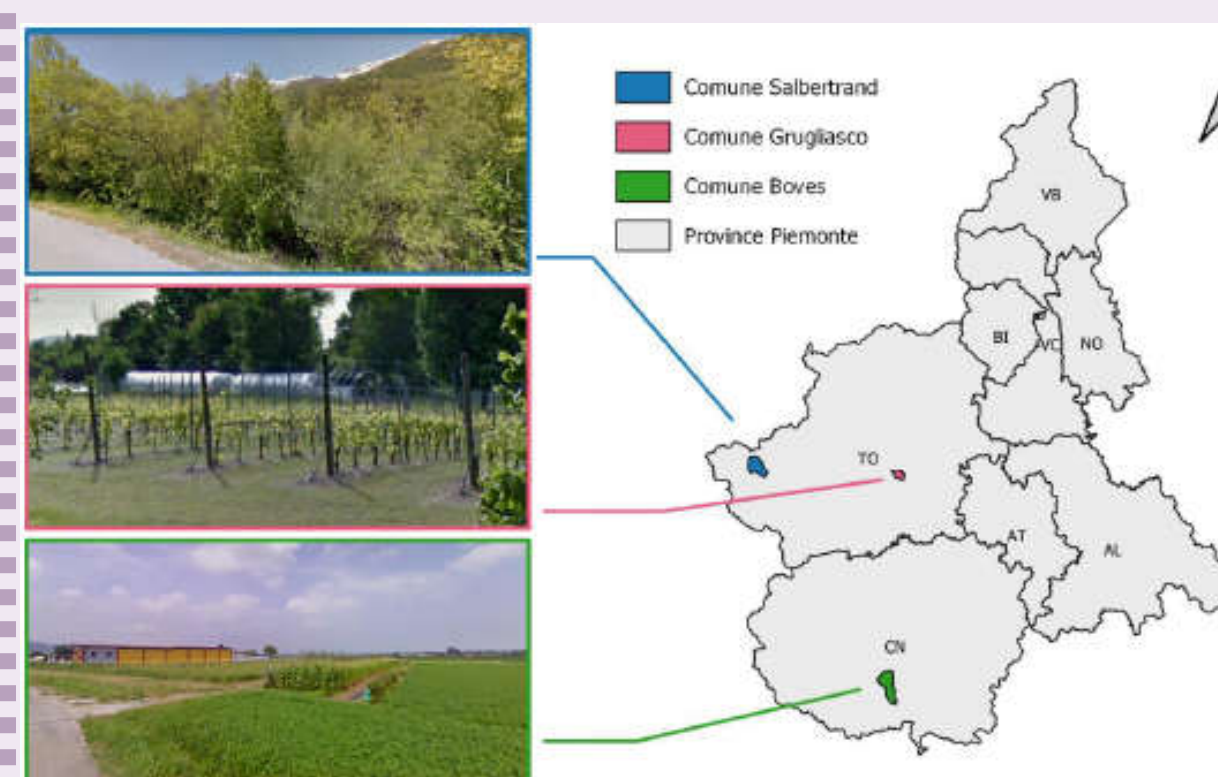
¹Politecnico di Torino, DIATI-Dipartimento di Ingegneria dell'Ambiente, del Territorio e delle Infrastrutture, nome.cognome@polito.it

Il mercato dei droni ha subito un forte sviluppo dovuto soprattutto all'innovazione dei componenti tecnologici. Il loro utilizzo è ormai consolidato in svariati ambiti tra cui quello dell'agricoltura di precisione, data la possibilità di coniugare l'acquisizione di dati su vasta area con risoluzioni elevate ed, eventualmente, con informazioni multispettrali. Tuttavia, il corretto svolgimento delle operazioni e le risoluzioni dei dati acquisiti non dipendono soltanto dal mezzo aereo e dai sensori a bordo, ma anche da una corretta pianificazione del volo, dall'altezza di acquisizione del dato e dal tipo di caso studio oltre che dall'esperienza dell'operatore. D'altra parte, diverse agenzie spaziali rendono disponibili i dati satellitari gratuitamente. Questi ultimi hanno una risoluzione temporale regolare che determina una disponibilità del dato in più fasi della stagione, che però potrebbe risultare inutilizzabile in presenza di nuvole sovrastanti l'area d'interesse. Inoltre, le immagini satellitari vengono acquisite con una risoluzione spaziale non minore di qualche metro. Risulta quindi necessario effettuare un'analisi dei costi in funzione della scala di applicazione e delle specifiche condizioni dell'oggetto di studio per valutare quando il rilevamento con UAV (Unmanned Aerial Vehicle) risulti più conveniente rispetto all'utilizzo di altre strumentazioni di Remote Sensing come le acquisizioni satellitari. In questo contesto si inserisce l'utilizzo il concetto di Super-Risoluzione, che può trovare un'interessante applicazione nei casi in cui non prevalga, in termini di convenienza, una specifica tipologia di acquisizione da remoto.

OBIETTIVI DELLA RICERCA:

- Studio dell'applicabilità della rete neurale convoluzionale VDSR (Very Deep Super Resolution) nel contesto del telerilevamento, utilizzando le immagini da drone come dati di base per allenare la rete e le immagini satellitari come immagini di input su cui migliorare la risoluzione;
- Applicazione nel campo del monitoraggio della vegetazione, per cui riveste particolare importanza l'utilizzo di dati multispettrali dai quali calcolare indici o elaborare classificazioni;
- La valutazione dell'accuratezza del risultato effettuata tramite il calcolo di metriche, il confronto con dati territoriali e l'elaborazione di classificazioni con relativo esame quantitativo dato dall'analisi della matrice di errore.

CASO STUDIO



Si dispone di immagini acquisite con drone in differenti località del Piemonte (Rivoira-Comune di Boves, Grugliasco e Salbertrand) e in diversi giorni nei mesi di giugno e luglio, mesi di fase attiva del ciclo vegetativo. I sensori utilizzati permettono di disporre di dati nelle bande RGB e NIR, ad eccezione dell'acquisizione nel sito di Salbertrand per la quale si è utilizzata solamente la banda RGB. Tutte le aree sono accomunate dalla presenza di vegetazione, seppur di diversa tipologia; ciò permette di riconoscere una maggiore variabilità spettrale nelle immagini sottoposte a test. A partire da questi dati si costruisce il dataset di immagini ad alta risoluzione per il training della rete neurale. Vengono effettuati tre differenti test sul sito di Rivoira incrementando l'estensione dell'area in esame al fine di valutare l'efficacia del metodo al variare della scala:

Test 01: 100 m x 200 m circa (2 ettari)
Test 02: 500 m x 500 m circa (25 ettari)
Test 03: 2 km x 2 km circa (400 ettari)

DATI UAV UTILIZZATI PER IL TRAINING DELLA RETE

LOCALITÀ	DATA	SENSORE	BANDE	RISOLUZIONE
Rivoira	26/07/2019	Mapir	RGB, Red, Green, NIR	2 cm
Rivoira	22/07/2020	Dji Phantom 4 Multispectral	RGB, NIR, Red Edge, Red, Green, Blue	3 cm
Grugliasco	25/06/2019	Slant Range	RGB, Red, Red Edge, NIR	2 cm
Salbertrand	17/06/2020		RGB	4 cm

- Generazione di ortomosaici multi-banda tramite algoritmi di SfM;
- Sottocampionamento dell'ortomosaico a 1m;
- Estrazione di porzioni quadrate di circa 70 m di lato, al fine di ottenere un campione di 479 immagini da utilizzare per il training

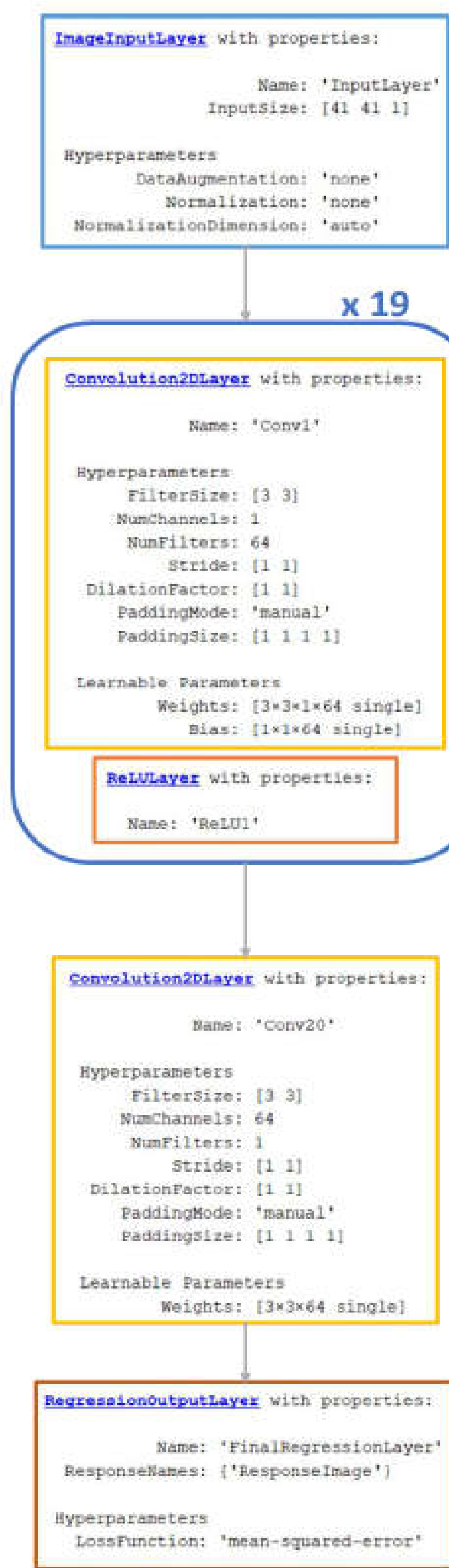
DATO SATELLITARE DA MIGLIORARE

LOCALITÀ	DATA	SENSORE	RISOLUZIONE
Rivoira	26/07/2019	Sentinel-2	10 m

- Bande TCI, B03, B04, B08
- Ricodifica immagine a 8 bit

Scelta operativa per la rete

La procedura si basa sulla determinazione dei residui tra un'immagine di riferimento a maggiore risoluzione e una a minor risoluzione scalata per portarla alla stessa dimensione dell'immagine di riferimento tramite interpolazione bicubica. L'immagine dei residui contiene le informazioni relative alla componente in alta frequenza. Nella VDSR i residui sono determinati sulla luminanza. Una volta che la rete VDSR impara a calcolare l'immagine dei residui, è possibile ricostruire un'immagine a risoluzione migliorata sommando i residui stimati all'immagine scalata a minor risoluzione. I canali della crominanza (Cb e Cr) sono aumentati di scala tramite interpolazione bicubica. L'estrazione delle informazioni risulta più efficace incrementando il numero di layers, e applicando filtri di ridotte dimensioni in sequenza molte volte.



LAYERS DELLA RETE

- **layers convoluzionali:** 19 livelli costituiti da 64 filtri di dimensione 3x3 e un livello con singolo filtro 3x3x64 per la ricostruzione dell'immagine;
- **layers ReLU (Rectified Linear Unit):** introducono non linearità nel sistema imponendo una soglia che azzeri tutti i valori negativi.

DATI DI TRAINING

Le immagini subiscono un ridimensionamento, in modo da corrispondere alle dimensioni attese dalla rete, e una rotazione, riflessione in direzione x e variazione di scala per aumentare la variabilità dei dati di training a disposizione.

TRAINING

Il training consente di stimare, attraverso la minimizzazione dello scarto quadratico medio tra l'immagine ricostruita e la corrispondente immagine ad alta risoluzione, i parametri della rete che definiscono i calcoli effettuati tramite ciascun livello. Tali parametri permettono poi di elaborare qualsiasi immagine di cui si vuole migliorare la risoluzione, a prescindere dal fatto che l'area di studio rientri tra quelle riprese dal dataset di training. Il tasso di apprendimento (learning rate) è impostato a un valore iniziale di 0.1, elevato rispetto ad altre CNN, al fine di velocizzare il training.

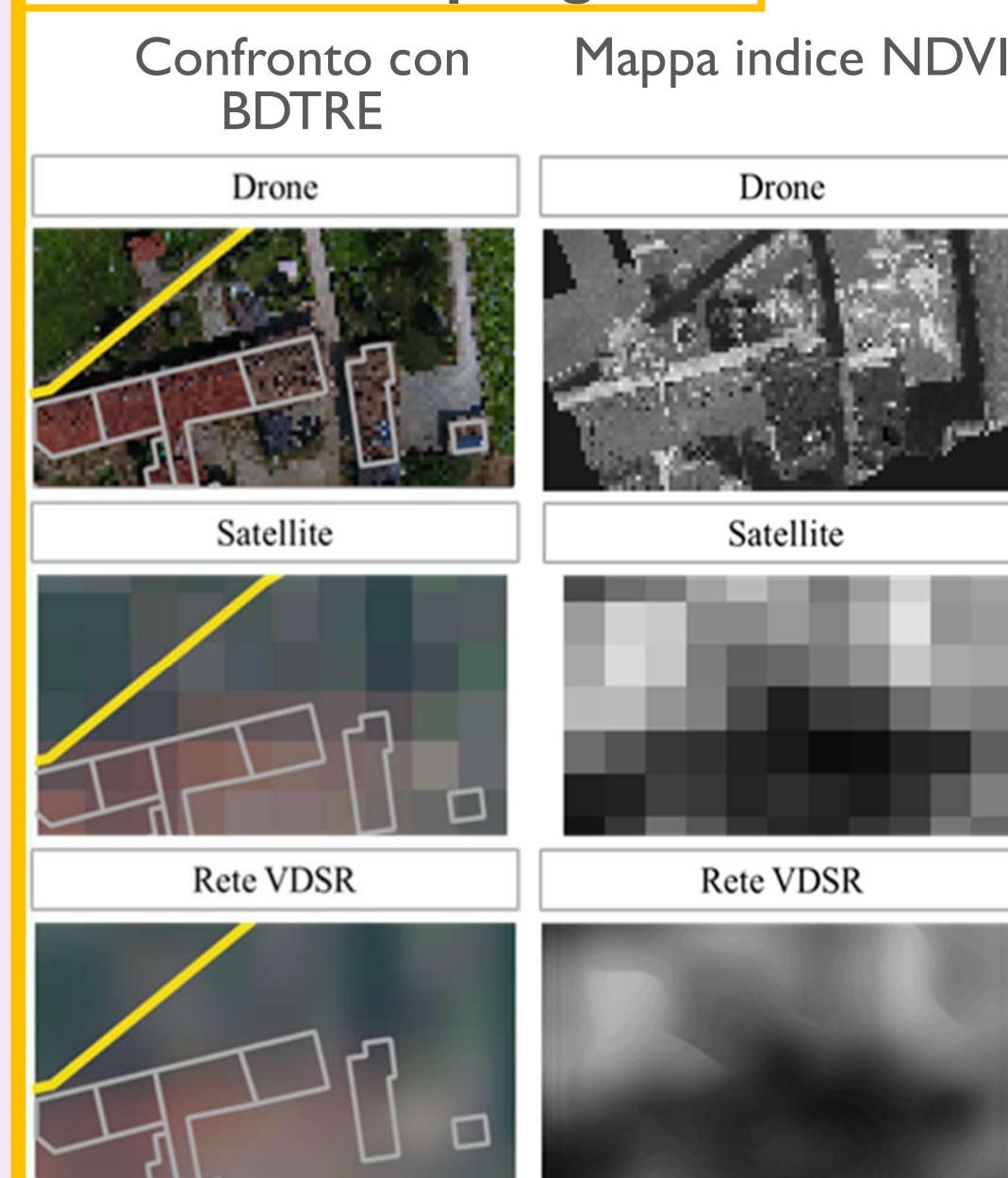
In questo lavoro è stato analizzato il risultato al variare di un singolo aspetto caratteristico alla volta in modo da individuare le impostazioni che permettano di ottenere l'accuratezza migliore.

Aspetto analizzato	Impostazione
Risoluzione data set di training	1 m
Risoluzione immagine di test	10 m
Fattori di scala per le immagini di training	2, 3 e 4
Angolo di rotazione delle immagini di training	20°
Numero di immagini di training	479 immagini RGB da drone
Numero di filtri dei layer convoluzionali	128 filtri

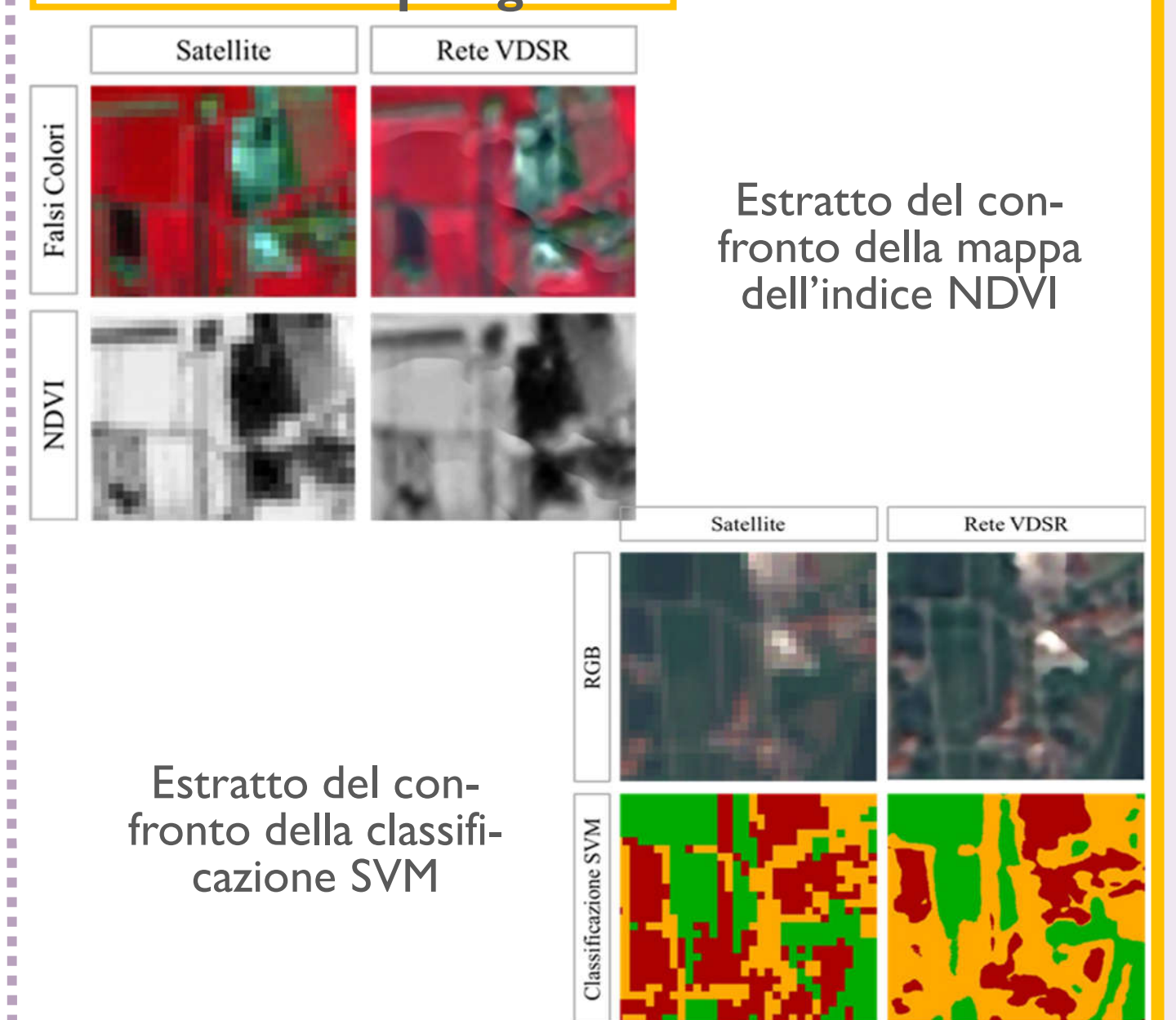
Risultati e valutazione dell'accuratezza

Al fine di valutare l'accuratezza dei risultati ottenuti, sono state effettuate valutazioni di tipo analitico, tramite il calcolo di metriche (**PSNR (Peak Signal-to-Noise Ratio)** e **SSIM (Structural Similarity Index)** **NIQE (Naturalness Image Quality Evaluator)**) e valutazioni di quantificazione dell'accuratezza di rappresentazione di aspetti spaziali tramite il confronto con il dato vettoriale Base Dati Territoriale di Riferimento degli Enti (**BDTRE**) (verifica visiva della coerenza dell'elaborazione e il miglioramento di risoluzione ottenuto) e tramite l'elaborazione di classificazioni Support Vector Machine (**SVM**), applicate alle immagini con la composizione di bande blue, green, red, NIR e indice NDVI.

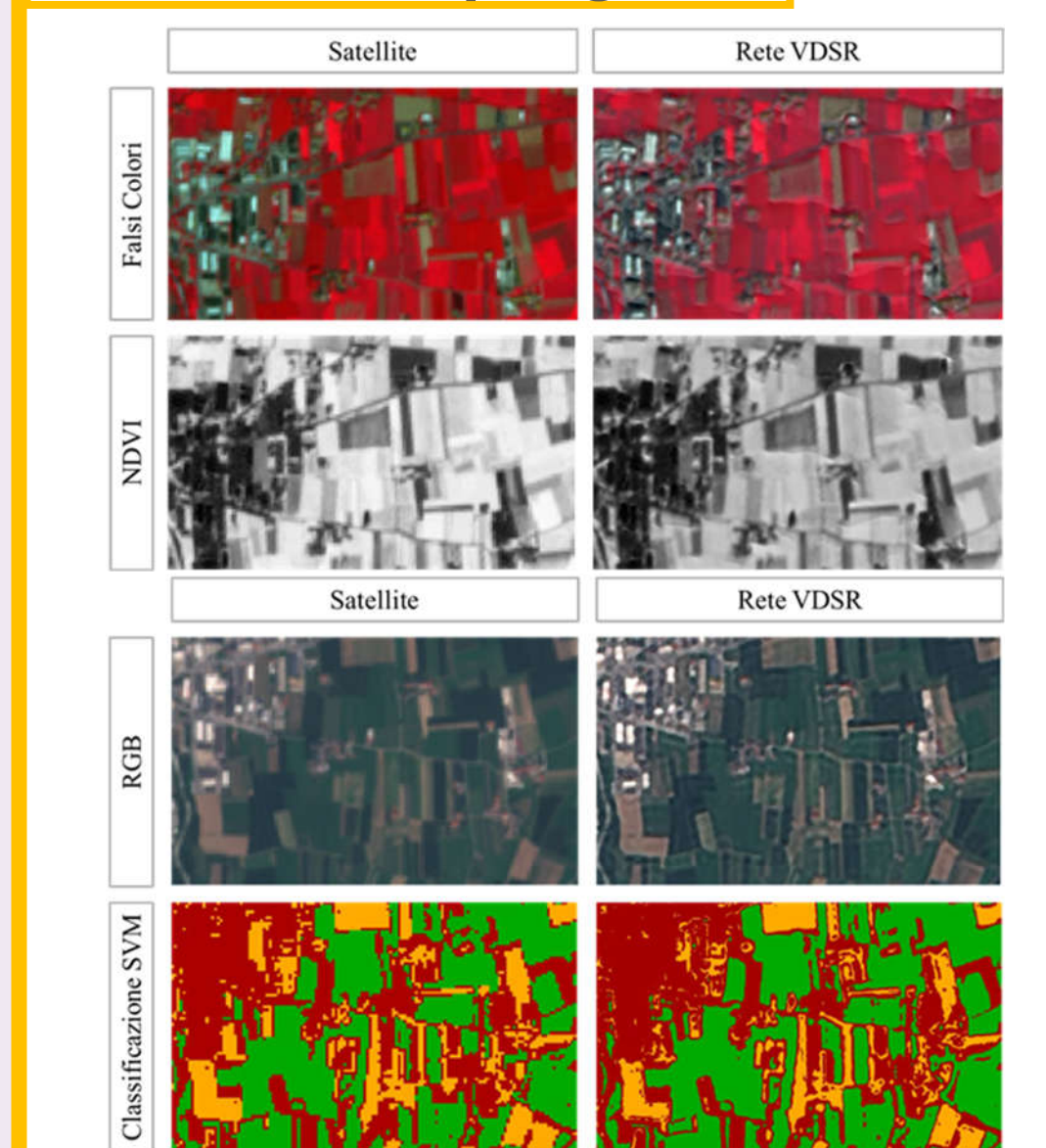
TEST 01—Output grafici



TEST 02—Output grafici



TEST 03—Output grafici



Statistiche relative alla classificazione SVM basata sull'immagine satellitare Sentinel-2

		Dati di riferimento			Area	Wi	Overall accuracy [%]	Kappa hat classification
		1	2	3				
Dati classificazione	1	0.1653	0	0.1033	72600	0.2686	76.9	0.65
	2	0	0.4314	0	116600	0.4314		
	3	0.1277	0	0.1724	81100	0.3000		
Totale		0.2930	0.4314	0.2757	270300			
Area		79188	116600	74512	270300			

Statistiche relative alla classificazione SVM basata sull'elaborazione della rete neurale VDSR

		Dati di riferimento			Area	Wi	Overall accuracy [%]	Kappa hat classification
		1	2	3				
Dati classificazione	1	0.1114	0	0.0104	32493	0.1218	81.2	0.71
	2	0.0000	0.371	0	98994	0.3710		
	3	0.1646	0.0134	0.3292	135342	0.5072		
Totale		0.2760	0.3844	0.3396	266828			
Area		73638	102577	90614	266828			

CONCLUSIONI:

- I test effettuati sulle aree di diversa estensione mettono in evidenza la diversa possibilità di fruizione dell'immagine elaborata con la rete VDSR al variare dell'area di indagine. Per una porzione di territorio di dimensione dell'ordine di grandezza del centinaio di metri di lato, il dato acquisito con drone ha una risoluzione e un dettaglio che le altre immagini non riescono a rappresentare. Per un'area di circa 25 ettari, il miglioramento ottenuto con l'elaborazione viene esaltato; l'estensione risulta abbastanza elevata per valutare l'utilizzo del drone con attenzione, ma non così ampia da accettare il dettaglio dell'immagine satellitare. Per una zona di analisi di dimensione dell'ordine di grandezza di alcuni km di lato o maggiore, l'elaborazione offre un miglioramento, ma il dettaglio richiesto dallo studio può essere soddisfatto anche dall'utilizzo dell'immagine satellitare originale.
- Il tempo di elaborazione sia per il training della rete neurale, sia per la sua attivazione, risulta non trascurabile se non si dispone di un'attrezzatura adeguata.
- Dai test effettuati risulta un'evidente analogia dei risultati ottenuti con interpolazione bicubica e con rete neurale VDSR. Questo aspetto conferma il corretto funzionamento della rete, ma evidenzia come i due metodi offrano prestazioni paragonabili.
- Questo lavoro incentiva studi futuri di ottimizzazione dell'algoritmo stesso e il confronto con altri algoritmi di Deep Learning come SelfEX o SRCNN.