

Utilizzo di Alberi Decisionali per la classificazione di aree urbanizzate

Steffenino Sara (*), Angeluccetti Irene (*), Disabato Franca (**)

(*) DIST, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino,
011 19751874, sara.steffenino@polito.it, irene.angeluccetti@polito.it

(**) ITHACA, Via Pier Carlo Boggio 61, 10138 Torino, 011 19751854, franca.disabato@ithaca.polito.it

Riassunto

Il presente articolo descrive uno studio preliminare per la classificazione di aree urbanizzate condotta per mezzo di Alberi Decisionali. Lo studio è stato effettuato sull'area di Torino, utilizzando immagini acquisite dal sensore Landsat ETM+ e la Cartografia Tecnica Regionale per la validazione. In particolare è stato valutato il variare dell'accuratezza di classificazione a seconda dei diversi attributi di input dell'algoritmo, quali il livello di pre-elaborazione radiometrica applicato, il numero di classi considerato, l'estensione temporale del training set e l'utilizzo di indici derivati da informazioni spettrali. I risultati hanno dimostrato che si possono raggiungere accuratèzze globali intorno all'80% e che gli indici sono l'attributo che permette di raggiungere tali accuratèzze.

Abstract

The paper describes a preliminary study on the urban classification accuracies obtained by means of the Decision Tree classifier. The study was conducted over the area of Turin (Italy), with Landsat ETM+ imagery and with an official regional map (Cartografia Tecnica Regionale) used as ground truth. In particular the variation of the accuracies was evaluated, depending on the changing of the algorithm input attributes such as the level of applied radiometric pre-processing, the considered number of classes, the temporal extent of the training set and the use of spectral indexes. Results show that overall accuracies of 80% can be achieved and that spectral indexes are the type of attribute that affect most these accuracies.

Introduzione

L'aggiornamento automatico di dati cartografici di base è di fondamentale importanza per una pluralità di applicazioni, dalle emergenze ambientali alla pianificazione territoriale. La classificazione di aree urbanizzate è una delle problematiche tuttora aperte nel campo dell'estrazione di tematismi da immagini satellitari. Nello specifico la presenza di pixel misti rende particolarmente difficoltoso il confronto tra diverse tipologie di ambiente urbano e l'identificazione di caratteristiche comuni che rendano possibile la creazione di procedure automatizzabili. L'utilizzo di tecniche di *data mining* permette di individuare regole per l'estrazione del suddetto tematismo, applicabili ad altre immagini acquisite dalla stessa tipologia di sensore.

Il presente lavoro si svolge nell'ottica di fornire un supporto alla produzione di cartografia speditiva per la logistica delle emergenze, costituendo un primo passo per la definizione di una metodologia di estrazione automatica delle aree urbanizzate a livello globale.

In particolare ci si è concentrati sulla valutazione dell'efficacia dell'utilizzo di Alberi decisionali per la classificazione di dati multispettrali, in termini di variazione dell'accuratezza in funzione di diversi attributi di input. Al fine di poter applicare la metodologia a scala sinottica, sono state utilizzate immagini a media risoluzione acquisite dal satellite Landsat 7. Viene nel seguito illustrata

la prima fase dello studio, in cui sono stati effettuati dei test sull'area di Torino, per la quale sono disponibili dei dataset di riferimento per la validazione.

Le accuratezze di classificazione sono state derivate per mezzo di una validazione che utilizza come riferimento una maschera estratta dalla Cartografia Tecnica Regionale Numerica (CTRN) a scala 1:10000.

La classificazione tramite alberi decisionali

L'Albero Decisionale è un classificatore non parametrico che non richiede nessuna assunzione aprioristica circa la distribuzione del campione. Tale classificatore è rappresentato da una struttura ad albero dove le foglie contengono gli elementi del campione classificati, mentre le ramificazioni, costituite da diversi nodi foglia, rappresentano l'insieme degli attributi che portano a quelle classificazioni (Shen et al., 2011).

Si definiscono attributi quelle caratteristiche dei dati stessi che vengono utilizzate per assegnare ciascun dato ad una delle classi predefinite (*target*).

L'Albero Decisionale, tra i classificatori esistenti, presenta alcuni vantaggi utili per classificare tematismi a livello globale: in primo luogo consente di generare delle regole che possono poi essere applicate a diversi dataset, secondariamente può gestire attributi continui determinando uno *split point* che divide in due l'intervallo dei valori.

L'algoritmo del classificatore utilizza dei dati di *training* per costruire, tramite un procedimento induttivo, un modello costituito da regole di classificazione. Ciascun percorso radice-foglia rappresenta perciò una regola di classificazione.

Il modello così costruito viene successivamente applicato a dati definiti di *test*. Tali dati vengono quindi classificati secondo le regole contenute nel modello.

Tra i diversi algoritmi esistenti si è scelto di utilizzare il C4.5 (Quinlan, 1993), di riconosciuta efficacia (Shen et al., 2011). L'algoritmo determina il miglior attributo di *split* del campione sulla base del criterio GainRatio, che privilegia la generazione di classi pure e di numero contenuto (Kohavi & Quinlan, 2002). In particolare per questo studio è stato utilizzato WEKA, software open source per il *data mining*, che implementa l'algoritmo C4.5 con il nome J48.

La classificazione è influenzata in primo luogo dall'algoritmo utilizzato e dagli attributi di input ma anche da alcuni parametri impostabili dall'operatore quali: il numero minimo di oggetti che ciascuna "foglia" dovrà contenere (*minNumObj*) e il fattore di confidenza (*confidenceFactor*), da cui dipende il numero di ramificazioni che vengono eliminate dall'albero (*pruning*) per ridurne la complessità ed aumentarne l'accuratezza tramite la diminuzione dell'*overfitting*. All'aumentare del *minNumObj* diminuisce il numero di ramificazioni dell'albero; l'aumentare del *confidenceFactor* risulta invece in una diminuzione del *pruning* (Bouckaert et al., 2008).

Caso studio

La scelta dell'area di indagine è ricaduta sulla città di Torino: la città presenta una buona copertura temporale di immagini acquisite dai satelliti Landsat e dispone di una CTRN in scala 1:10000 con cui effettuare la validazione.

Per lo studio in questione sono state utilizzate sei scene Landsat della collezione *Enhanced Thematic Mapper* (ETM+) *slc-on* ed una scena della collezione *Global Land Survey* (GLS) 2005 (Tabella 1).

Il GLS contiene, per ogni area geografica, almeno un'immagine acquisita in quell'anno appartenente alla collezione TM, oppure alla collezione ETM+ *slc-off*; in quest'ultimo caso viene corretta dagli effetti indotti all'immagine dal malfunzionamento dello scanner.

Le sezioni della CTRN utilizzate per creare la maschera di validazione sono le seguenti: 155080, 155110, 155120, 155150, 155160, 156050, 156060, 156090. Tutte le sezioni sono aggiornate al 31 dicembre 2005 e disponibili gratuitamente in rete in formato vettoriale (www.webgis.csi.it/Ctrig/main.asp).

Data Set	Data acquisizione	Path/row	Cloud cover (%)
L7 ETM+ slc-on	06/10/1999	194/29	0,54
L7 ETM+ slc-on	30/11/1999	195/29	3,02
L7 ETM+ slc-on	05/06/2000	195/29	1,34
L7 ETM+ slc-on	30/07/2001	195/29	6,80
L7 ETM+ slc-on	24/08/2001	194/29	4,41
L7 ETM+ slc-on	28/09/2002	194/29	13,99
GLS 2005	02/07/2005	194/29	-

Tabella 1. Scene utilizzate per il caso studio.

Metodologia

L'obiettivo di questo primo lavoro effettuato sull'area di Torino è quello di valutare in che modo gli attributi utilizzati come input nell'Albero Decisionale possono influenzare l'accuratezza raggiungibile nella classificazione.

A tal fine sono stati valutati i seguenti attributi, descritti nel seguito: tipo di pre-elaborazione radiometrica, estensione temporale del training set (una o più immagini), numero di classi considerato e informazioni contenute nel pixel (radiometriche e/o indici derivati).

Per quanto riguarda i parametri di input che usa l'algoritmo per generare il classificatore, il confidenceFactor è stato impostato pari a 0,25, come di default, mentre al minNumObj è stato assegnato il valore 10 in tutti i casi tranne quello in cui vengono utilizzati solo gli indici per addestrare il classificatore, caso in cui è stato posto pari a 20.

Le combinazioni di parametri che sono state testate sono riportate in Tabella 2.

NUMERO DI CLASSI	ESTENSIONE TEMPORALE DEL TRAINING SET	INDICI E INFORMAZIONI SPETTRALI	PRE-ELABORAZIONI RADIOMETRICHE	NOME CASO
2 CLASSI / 5 CLASSI	Immagine singola (02/07/2005)	indici	Dark Subtraction	index_DS
			Calibrazione in riflettività	index_calib
			FLAASH	index flaash
		info spettrali	Dark Subtraction	spectral_DS
			Calibrazione in riflettività	spectral_calib
			FLAASH	spectral flaash
	indici + info spettrali	Dark Subtraction	spectral_index_DS	
		Calibrazione in riflettività	spectral_index_calib	
		FLAASH	spectral_index flaash	
	Set multitemporale	indici	Dark Subtraction	index_DS
			Calibrazione in riflettività	index_calib
			FLAASH	index flaash
		info spettrali	Dark Subtraction	spectral_DS
			Calibrazione in riflettività	spectral_calib
			FLAASH	spectral flaash
indici + info spettrali		Dark Subtraction	spectral_index_DS	
		Calibrazione in riflettività	spectral_index_calib	
		FLAASH	spectral_index flaash	

Tabella 2. Casistiche considerate per l'analisi.

Questi attributi sono stati utilizzati per generare diversi classificatori, applicati poi alla sola immagine del 02/07/2005: scelta dovuta alla necessità di rispettare la coerenza temporale tra la classificazione e il dataset di validazione. Le diverse classificazioni sono state validate tramite l'utilizzo della CTRN aggiornata al 2005, la quale ricopre una porzione della scena Landsat.

Pre elaborazioni radiometriche (calibrazione e correzione atmosferica)

Per analizzare lo stato delle superfici e i loro cambiamenti nel corso del tempo in modo quantitativo occorre basarsi su una serie di misure confrontabili.

Una prima riduzione della variabilità fra le diverse scene acquisite può essere ottenuta attraverso una normalizzazione rispetto alla radiazione incidente, passando dalla Radianza spettrale alla Riflettività apparente; questa pre-elaborazione radiometrica, nota come Calibrazione radiometrica in riflettività, è solo il primo passo per un confronto di tipo quantitativo tra i dati ottenuti dai diversi satelliti in tempi diversi.

Per ottenere la riflettività (albedo) della superficie a terra è necessario depurare il segnale anche dagli effetti dell'atmosfera, attraverso tecniche di correzione atmosferica più o meno complesse.

In questo caso studio sono stati utilizzati sia una correzione atmosferica di tipo semplificato, nota come *Dark Subtraction*, che una di tipo fisico rigoroso, nota come *Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes* (FLAASH).

Il metodo della Dark Subtraction elimina l'effetto della componente diffusiva dell'atmosfera sottraendo a tutti i pixel dell'immagine, banda per banda, il valore di DN corrispondente al minimo valore di riflessione. FLAASH, invece, è uno strumento di correzione atmosferica che utilizza MODTRAN4 per risolvere l'equazione di trasferimento radiativo. I parametri che entrano in gioco nell'equazione dipendono dall'elevazione media della superficie, angolo di vista del satellite ed elevazione solare, e presuppongono la conoscenza del modello atmosferico, del tipo di aerosol e della visibilità, ricavati da informazioni desumibili dalle scene.

I classificatori sono stati addestrati con training set estratti da immagini a cui sono stati applicati questi 3 tipi di pre-elaborazioni radiometriche.

Estensione temporale del training set

Considerando l'obiettivo finale di ottenere un algoritmo in grado di classificare in modo automatico le aree urbane, che possono essere derivate da scene differenti, si è scelto di valutare l'effetto dell'utilizzo di informazioni radiometriche e/o indici derivati da una singola scena o da una media dei valori di un set multitemporale.

Gli input sono quindi, in un caso (immagine singola) le firme spettrali acquisite solo sull'immagine del 02/07/2005, nell'altro (set multitemporale) le firme spettrali medie delle scene riportate in Tabella 1.

Indici e informazioni spettrali utilizzati per il training set

Indici derivati da informazioni spettrali vengono spesso utilizzati nel telerilevamento per identificare elementi di interesse.

In questo studio sono stati valutati alcuni indici già utilizzati in letteratura per l'estrazione dell'urbano e ne è stato proposto uno nuovo.

Gli indici considerati, le cui equazioni sono riportate in Tabella 3, sono i seguenti:

- NDBI (Normalized Difference Built-up Index), utilizzato per la mappatura di aree urbane insieme all'NDVI in studi precedenti (Zhang et al., 2003) con accuratezze del 92%;
- NDBBBI (Normalized Difference Blue Band Built-up Index), utilizzato in Baraldi et al. (2006), tiene conto della componente di blu presente nelle aree urbane e nei terreni non coltivati;
- NDVI (Normalized Difference Vegetation Index), utilizzato in studi sulla vegetazione ed utile per separare le classi acqua, suolo e vegetazione;
- BUI (Built-Up Index), utilizzato in Lee, Lee, and Chi (2010);
- TEST, un indice proposto dagli autori sulla base dell'analisi delle firme spettrali di aree campione, che presenta valori di poco superiori ad 1 (range approssimativo: 1.1 ÷ 1.2) in corrispondenza delle aree urbane e valori più alti in corrispondenza delle altre classi analizzate.

INDICE	NDBI	NDBBBI	NDVI	BUI	TEST
EQUAZIONE	$\frac{TM5-TM4}{TM5+TM4}$	$\frac{TM1-TM5}{TM1+TM5+0.001}$	$\frac{TM4-TM3}{TM4+TM3}$	NDBI-NDVI	$TM5 + \frac{TM4}{TM7}$

Tabella 3. Indici utilizzati come attributi in input.

In input al classificatore sono state utilizzate alternativamente solo le informazioni spettrali, oppure solo gli indici, oppure entrambe. Tra questi l'algoritmo determina i migliori attributi di *split*, che costituiranno i nodi foglia di ciascun classificatore.

Numero di classi considerato nel training set

Il training set utilizzato per addestrare il classificatore è composto da informazioni relative alla classe di urbano e da informazioni relative alle altre classi.

In un caso (5 classi) sono state estratte ROI di urbano, aree industriali, vegetazione, acqua e suolo nudo, l'albero decisionale classifica secondo questi 5 target e i tematismi vengono poi aggregati per la validazione in modo da analizzare solamente la classe di urbano (aree urbane ed aree industriali) e non urbano (acqua, vegetazione, suolo nudo).

Nel secondo caso (2 classi) queste stesse ROI vengono aggregate a monte del processo di classificazione; il classificatore pertanto definisce solo una classe di urbano e una di non urbano.

In questo modo è possibile valutare se l'utilizzo di diverse classi nel training set affina il processo di classificazione oppure se è sufficiente basarsi su due macro classi di urbano e non urbano.

Maschera di validazione

Le diverse classificazioni sono state validate con una maschera a due valori (urbano/non urbano) ottenuta dalle sezioni aggiornate al 2005 della CTRN.

Gli elementi della CTRN sono ripartiti in 11 gruppi in funzione della tipologia di entità cartografica che questi rappresentano (es. Strade e Ferrovie, Vegetazione, etc) e, all'interno di ogni gruppo, sono ulteriormente suddivisi per categorie (es. Tronco di Strada Asfaltata, Piazza e Parcheggi, etc) (www.regione.piemonte.it/sit/argomenti/pianifica/cartografia_sit/dwd/manuale_ctr/). Le entità che sono state prese in considerazione per la costruzione della maschera sono solo quelle di tipo areale; in particolare, nella classe urbano sono state incluse le entità appartenenti ai gruppi 01 (Strade e ferrovie) e 02 (Fabbricati, arredo urbano e manufatti tecnici). Alcuni affinamenti di questa attribuzione sono tuttavia stati operati in ambiente GIS, ed in particolare:

- gruppo 01: le "Aree urbane non pavimentate", categoria molto ampia comprendente una varietà di entità (dalle aree verdi di proprietà privata alla pavimentazione asfaltata circostante agli edifici), sono state attribuite alla classe urbano se adiacenti agli edifici residenziali (buffer di 10 m), e alla categoria non urbano negli altri casi. Gli "Impianti sportivi", invece, sono stati attribuiti alla classe non urbano ad esclusione delle loro coperture e fabbricati;
- gruppo 05 (Divisioni del terreno): le entità appartenenti alla tipologia "Cimitero" sono state assegnate alla classe urbano ad esclusione di alcuni cimiteri più vegetati.

La nuova *feature* poligonale così ottenuta è stata quindi convertita in un raster con risoluzione di 30 m, assegnando ad ogni pixel la classe presente con un'area maggiore nella *feature* di partenza. Il processo è schematizzato in Figura 4.

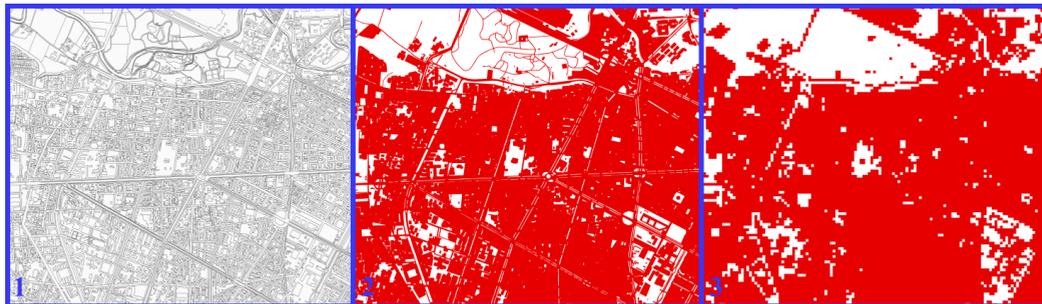


Figura 4. Fasi di creazione della maschera di validazione: (1) stralcio di cartografia numerica di Torino; (2) attribuzione di una classe (urbano in rosso, non urbano in bianco) a ciascun elemento; (3) conversione da formato vettoriale a raster.

Risultati

Le classificazioni ottenute sono state validate con la maschera creata a partire dalla CTRN; in totale sono stati validati 323864 pixel, ricavando il numero di pixel correttamente e non correttamente classificati. Una procedura tradizionale di verifica dell'accuratezza della classificazione si ottiene dall'analisi della matrice di confusione (Gomasasca, 2004), dalla quale si estraggono i seguenti parametri di accuratezza: *Overall Accuracy* (OA), per descrivere l'accuratezza globale, *Producer's Accuracy* (PA), *User's Accuracy* (UA) per descrivere il comportamento del classificatore per ogni singola classe.

Si rimanda all'ampia letteratura già presente sull'argomento per la definizione dei vari parametri (Congalton, 1991; Foody, 2002).

I parametri estratti da ciascuna matrice sono presentati in Tabella 5.

Analizzando globalmente i risultati in termini di OA, si può affermare che i classificatori che utilizzano set multitemporali forniscono generalmente risultati migliori (OA = 76.6%) rispetto a quelli che utilizzano l'immagine singola (OA = 66.2%), e che i training set costituiti da 5 classi distinte (OA = 72.6%) invece di 2 (OA = 70.2%) consentono di individuare meglio il tematismo urbano.

L'utilizzo di diversi livelli di pre-elaborazioni radiometriche non sembra influenzare le accuratèzze di classificazione, mentre le combinazioni di indici e informazioni spettrali le condizionano più marcatamente.

Alcune classificazioni presentano bassi valori di UA per la classe urbano (UA < 60%), dovuti ad una sovraclassificazione del tematismo: questo è il caso delle classificazioni a immagine singola in cui sono state utilizzate le informazioni spettrali. Nettamente diverso è il comportamento dei rimanenti classificatori a immagine singola che utilizzano solo gli indici come input.

				URBANO		NON URBANO				URBANO		NON URBANO	
NOME CASO		OA	PA	UA	PA	UA	OA	PA	UA	PA	UA	PA	UA
Immagine singola (02/07/2005)	index_DS	79.9	86.9	71.3	74.8	88.8	80.5	85.6	72.6	76.8	88.1	88.1	88.1
	index_calib	69.9	91.3	59.0	54.4	89.7	71.1	85.9	61.0	60.5	85.7	85.7	85.7
	index_flash	72.7	91.5	61.7	59.2	90.7	72.6	90.1	61.9	60.0	89.4	89.4	89.4
	spectral_DS	59.6	97.3	50.9	32.4	94.4	67.4	96.6	56.5	46.4	95.0	95.0	95.0
	spectral_calib	59.6	97.3	50.9	32.4	94.4	67.4	96.6	56.5	46.4	95.0	95.0	95.0
	spectral_flash	59.5	97.3	50.8	32.2	94.4	58.7	97.6	50.4	30.8	94.6	94.6	94.6
	spectral_index_DS	58.3	97.6	50.1	30.0	94.6	67.9	89.2	57.5	52.6	87.1	87.1	87.1
	spectral_index_calib	59.6	97.3	50.9	32.4	94.4	68.2	90.5	57.7	52.2	88.4	88.4	88.4
	spectral_index_flash	59.5	97.3	50.8	32.2	94.4	59.3	90.5	50.7	36.8	84.3	84.3	84.3
Set multitemporale	index_DS	78.6	72.1	75.5	83.2	80.6	78.4	75.0	73.7	80.8	81.8	81.8	81.8
	index_calib	77.0	87.4	67.4	69.6	88.5	77.7	66.8	76.9	85.6	78.2	78.2	78.2
	index_flash	77.4	69.9	74.4	82.7	79.3	79.8	71.0	78.6	86.1	80.5	80.5	80.5
	spectral_DS	63.8	98.0	53.7	39.2	96.5	72.6	92.6	61.5	58.3	91.6	91.6	91.6
	spectral_calib	75.5	68.7	71.7	80.5	78.1	78.6	69.1	77.3	85.4	79.4	79.4	79.4
	spectral_flash	77.4	81.3	69.8	74.7	84.8	77.5	79.7	70.4	75.9	83.9	83.9	83.9
	spectral_index_DS	78.6	72.1	75.5	83.2	80.6	73.2	92.5	62.6	58.9	91.3	91.3	91.3
	spectral_index_calib	79.2	65.3	81.3	89.2	78.1	78.6	69.1	77.3	85.4	79.4	79.4	79.4
	spectral_index_flash	77.4	69.9	74.4	82.7	79.3	77.3	79.8	70.1	75.6	83.9	83.9	83.9
							2 CLASSI						
									5 CLASSI				

Tabella 5. Accuratezza delle classificazioni (%) in termini di Overall Accuracy (OA), Producer's Accuracy (PA), User's Accuracy (UA).

Sono stati evidenziati in verde i tre migliori classificatori, con *OA* intorno all'80%, caratterizzati da regole costruite con i seguenti attributi di *split*:

- index_DS (2 classi, immagine singola): utilizza NDBI e NDBBBI;
- index_DS (5 classi, immagine singola) : utilizza NDBBBI ed NDVI;
- index_flaash (5 classi, set multitemporale): utilizza NDVI, TEST, NDBBBI e BUI.

Dal momento che le accuratzeze dai suddetti classificatori sono molto simili tra loro e che l'area classificata copre un'area più vasta della maschera di validazione, un'ulteriore analisi visiva è stata condotta anche al di fuori dell'area validata, al fine di meglio comprendere il comportamento di questi tre classificatori.

Alcuni esempi sono riportati in Figura 6.

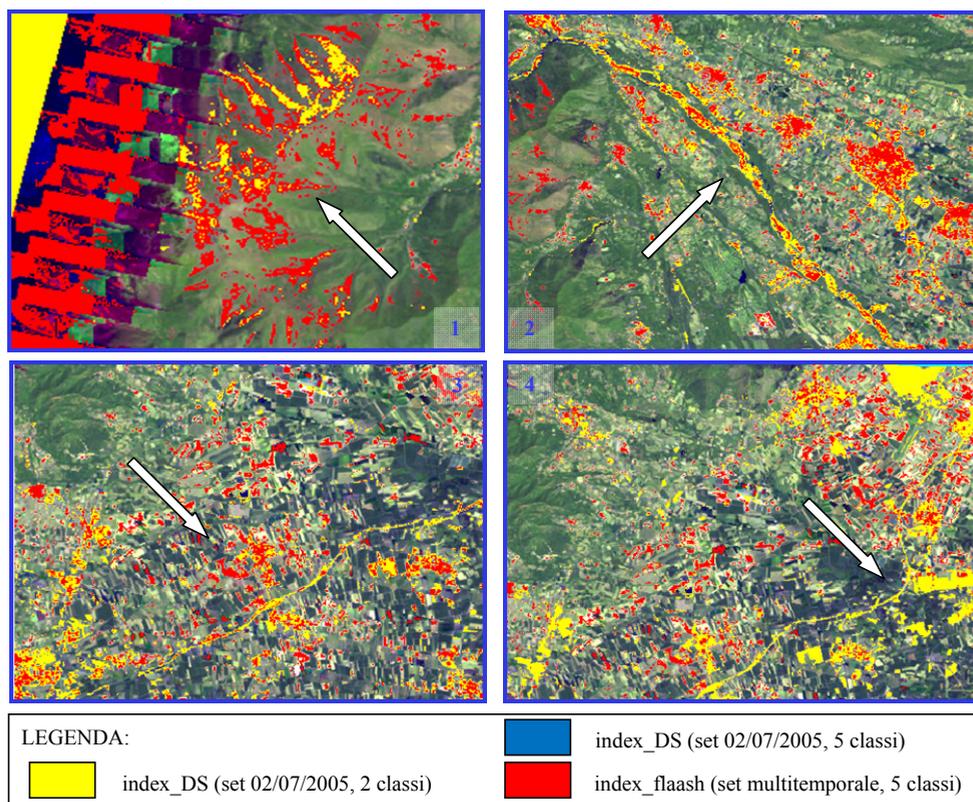


Figura 6. Comportamento dei 3 classificatori nelle zone d'ombra (1), nelle zone di suolo nudo (2), nelle aree vegetate (3) e lungo le strade (4).

I classificatori a singola immagine hanno comportamento simile tra di loro nelle zone analizzate e per questo in Figura 6 è visibile solo quello a 2 classi; dalla Figura 6 emerge che il classificatore addestrato con un set multitemporale classifica meglio le zone di suolo nudo (in Figura 6.2 le zone di suolo nudo sono quelle adiacenti al corso d'acqua) includendo meno urbano in questa classe, ma sovraclassifica maggiormente l'urbano nelle zone d'ombra (6.1) e in quelle vegetate (6.3). Per quanto riguarda le strade (6.4), queste non vengono classificate come urbano nel caso multitemporale.

Le regole generate dall'algorithmo che applicate all'immagine hanno fornito i risultati migliori, sono rappresentate nell'albero di Figura 7.

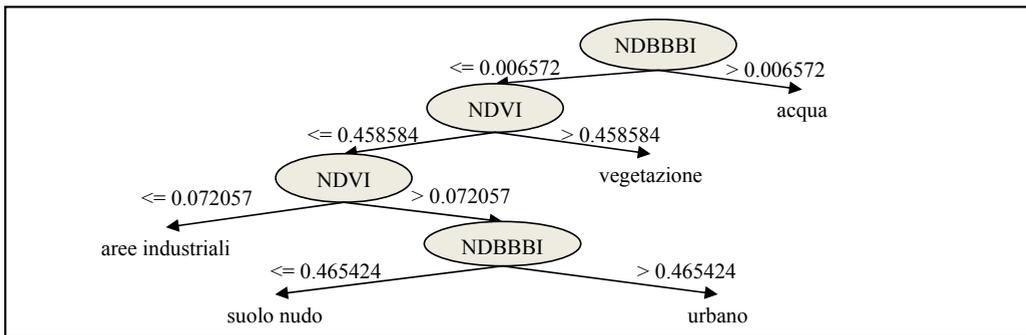


Figura 7. Regole estratte dal classificatore che ha ottenuto la miglior Overall Accuracy (index_DS: 5 classi, immagine singola del 02/07/2005).

Conclusioni e sviluppi futuri

In questo studio si è valutato in che modo i parametri utilizzati in input al classificatore considerato (Albero decisionale), influenzino l'accuratezza finale delle classificazioni.

I risultati ottenuti mostrano che, mentre alcuni parametri di input condizionano in modo trascurabile il risultato finale (es. il tipo di pre-elaborazione radiometrica applicata all'immagine), altri risultano di fondamentale importanza per il corretto riconoscimento del tematismo (es. indici).

La presenza dell'indice TEST nelle regole di uno dei classificatori risultati più performanti, incoraggia ulteriori investigazioni circa l'opportunità del suo utilizzo.

Lo studio proposto in questo articolo va considerato come il primo *step* di un lavoro più ampio volto all'individuazione automatica delle aree urbane da dati a media risoluzione.

In futuro tale metodologia verrà testata su aree differenti e verrà valutato l'utilizzo di altri parametri di input (es. informazioni tessiture), nonché di altre elaborazioni da applicare all'immagine (es. tecniche di segmentazione) per favorire l'estrazione di tale tematismo.

Riferimenti bibliografici

- Baraldi A., Puzzolo V., Blonda P., Bruzzone L., & Tarantino C. (2006), "Automatic Spectral Rule-Based Preliminary Mapping of Calibrated Landsat TM and ETM+ Images", *IEEE Transactions on Geoscience and Remote Sensing*, 44: 2563–2586
- Bouckaert R. R., Frank E., Hall M., Kirkby R., Reutemann P., Seewald A., & Scuse D. (2008), *WEKA Manual for version 3-6-0*, University of Waikato, Hamilton, New Zealand.
- Congalton R. G. (1991), "A review of assessing the accuracy of classifications of remotely sensed data", *Remote Sensing of Environment*, 37(1): 35–46.
- Foody G. M. (2002), "Status of land cover classification accuracy assessment", *Remote Sensing of Environment*, 80(1): 185–201.
- Gomarasca M. (2004), *Elementi di geomatica*, Edizioni AIT, 422-423
- Kohavi, R., & Quinlan, J. R. (2002), "Handbook of data mining and knowledge discovery", *Oxford University Press*
- Lee J. I. N. A., Lee S. S., & Chi K. H. (2010), "Development of an urban classification method using a built-up index", *Selected topics in power systems and remote sensing*, 39–43
- Quinlan J. R. (1993), *C4.5: Programs for Machine Learning*, M. Kaufmann Ed, California
- Shen W., Wu G., Sun Z., Xiong W., Fu Z., & Xiao R. (2011), "Study on classification methods of remote sensing image based on decision tree technology", *International Conference on Computer Science and Service System (CSSS)*, 4058–4061
- Zhang X., Friedl M. A., Schaaf C. B., Strahler A. H., Hodges J. C. F., Gao F., Reed B. C., et al. (2003), "Monitoring vegetation phenology using MODIS", *Remote Sensing of Environment*, 84: 471–475