

La cartografia di Cesare Battisti: sperimentazione di una procedura di digitalizzazione automatica

Angelo Besana ^(a), Marco Ciolli ^(b), Stefano Gobbi ^(c), Paolo Zatelli ^(d)

^(a) DIST - Dipartimento Interateneo di Scienze, progetto e politiche del territorio, Politecnico e Università di Torino, Castello del Valentino, viale Mattioli, 39, 10125, Torino, tel. 011 0907446, fax 011 0907499, angelo.besana@unito.it.

^(b) DICAM - Dipartimento di Ingegneria Civile, Ambientale e Meccanica, Università degli Studi di Trento, via Mesiano, 77, 38123, Trento, tel. 0461 882625, fax 0461 882672, marco.ciolli@unitn.it

^(c) DICAM, tel. 0461 882686, stefano.gobbi@unitn.it

^(d) DICAM, tel. 0461 882618, paolo.zatelli@unitn.it

Introduzione

La figura di Cesare Battisti quale geografo e cartografo è stata di recente riscoperta e rivalutata per la sua originale posizione scientifica nel contesto storico della geografia accademica italiana e per l'attualità metodologica della sua analisi applicata, in grado di "integrare" compiutamente tra loro le componenti fisico-naturali e quelle antropiche nell'indagine e rappresentazione dell'organizzazione spaziale regionale (Dai Prà, 2018). Lo studio della produzione scientifica di Cesare Battisti può dunque rappresentare un contributo fondamentale per la storia del pensiero geografico (Quaini, 2018; Rombai 2016) ma anche un'interessante occasione per la sperimentazione di nuovi strumenti e metodologie di cartografia storica applicata (Dai Prà, 2010). In particolare, per questa seconda, specifica, finalità può risultare di notevole utilità la predisposizione di una procedura automatica di digitalizzazione e classificazione delle carte storiche.

La procedura implementata a tal fine si basa su di una metodologia di classificazione ad oggetti (*Object-based Image Analysis – OBIA*) in tre passi, che verrà descritta in dettaglio più avanti.

L'intera procedura è stata realizzata con GRASS GIS 7.6, automatizzando tutte le operazioni e, in particolare, effettuando la classificazione dei segmenti mediante *machine learning* in R, con l'utilizzo di tutti i classificatori disponibili. L'esito è una carta in formato digitale (sia in formato raster che vettoriale) che può così essere utilizzata direttamente in ambiente GIS.

Cesare Battisti cartografo

Battisti è stato senza soluzione di continuità geografo e cartografo, in quanto intese le carte geografiche sia come strumento indispensabile di analisi e di riflessione, sia come mezzo efficace di rappresentazione e di comunicazione di concetti e problematiche (Dai Prà e Rossi, 2016). La sua più importante opera cartografica è costituita da *Il Trentino. Cenni geografici, storici, economici* (1915), un atlante tematico composto da 19 tavole cartografiche. L'opera rappresenta uno dei primi studi monografici regionali in Italia in cui la regione trentina è studiata e rappresentata rispetto alle sue caratteristiche fisico-naturali, storico-culturali, socio-demografiche ed economiche, secondo la prospettiva teorica della cosiddetta "regione integrale" della scuola geografica fiorentina di Giovanni Marinelli. I progressi dell'epoca nei metodi di raccolta, classificazione e rappresentazione dei vari fenomeni geografici avevano portato alla realizzazione di inventari sistematici delle diverse variabili, la cui elaborazione e rappresentazione cartografica poteva consentire, in primo luogo, l'individuazione delle rispettive "regioni elementari", esplicative della diffusione e caratterizzazione geografica dei singoli fenomeni (il clima, la flora, la fauna, l'idrografia, la demografia, gli insediamenti, la toponomastica, ...). Da queste, quindi, potevano essere derivate le "regioni complesse" (climatiche, fitogeografiche, zoogeografiche, antropiche, ...) quale risultato della combinazione dei singoli fenomeni e della sovrapposizione delle corrispondenti regioni elementari. Sintesi conclusiva di tutti i fenomeni e i livelli di analisi considerati, espressione evidente di unità geografica era, infine, la "regione integrale" (Proto, 2015). Il Trentino di Battisti costituisce una prima evidenza applicativa del modello regionale marinelliano seppure con un importante spunto di originalità, in quanto la sovrapposizione tra le ideali regioni "naturale", sostanzialmente statica, ed "antropica", tendenzialmente dinamica a seguito delle vicende storiche (politiche, belliche, sociali, ...), porta all'individuazione di una pluralità di regioni geografiche (sei) che nel loro insieme compongono solidalmente il Trentino e l'Alto Adige quale regione "integrale" (*Tavola I, Confini geografici, storici ed etnografici*).

L'opera di Battisti aveva anche un importante significato politico e tale risultato di sintesi consentiva di sostenere, sul piano scientifico, la presunta italianità della regione, fondata sulla considerazione fondamentale che i vari movimenti di popolazioni, prevalentemente italiane, nel corso del tempo avevano portato ad una preponderante diffusione dell'elemento culturale italiano (Dai Prà e Rossi, 2016; Proto, 2015). La tavola III, *Distribuzione etnico-linguistica della popolazione*, consente di raffigurare il principale elemento su cui poggiare tale considerazione. Essa riprende esattamente la carta Achille Dardano che illustra la distribuzione della popolazione nell'Italia nord-orientale (Istria compresa) sulla base della lingua parlata, pubblicata sempre dall'Istituto De Agostini all'inizio del 1915 (Proto, 2015).

È questa la carta tematica dell'atlante battistiano che sarà utilizzata per testare la procedura automatica di digitalizzazione e classificazione qui proposta. Il

complesso delle carte de *Il Trentino* può essere scaricato in formato immagine dalla sezione cartografica dell'articolo di Matteo Proto (2015) più sopra citato [http://storicamente.org/proto_battisti_trentino#nt-5 - visitato il 28/7/2019].

Digitalizzazione e classificazione della carta

La carta, originariamente in scala 1:500 000, è disponibile come immagine JPEG, con dimensione 2234x2500 pixel, 72 DPI, YcbCr color space a 8 bit.

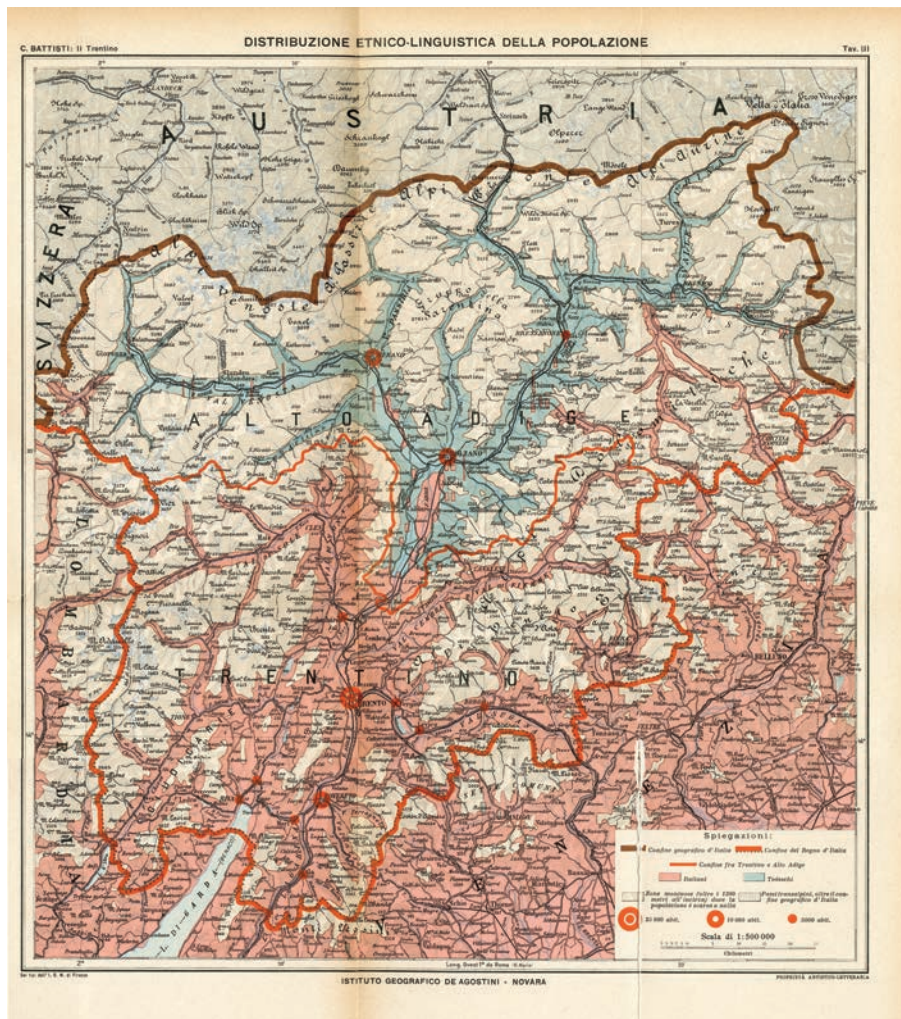


Figura 1 – “Carta della distribuzione etnico-linguistica della popolazione”, carta originale.

La classificazione è eseguita con classificazione ad oggetti (*Object-based Image Analysis – OBIA*) in 3 passi:

1. formazione dei *segmenti*, cioè aree dell’immagine ritenute omogenee in termini di significato e quindi classificate come un singolo oggetto
2. classificazione di ogni segmento in una delle classi della carta

3. ripulitura dei segmenti appartenenti a classi non significative per l'uso della carta

Tutta l'elaborazione è fatta con GRASS GIS 7.6, e l'intera procedura è stata automatizzata tramite *script*. Prima dell'elaborazione la legenda è stata mascherata.

La formazione dei segmenti è effettuata con approccio *region-growing*: i segmenti (inizialmente i singoli pixel) sono uniti in un segmento se

1. la loro differenza è la minore rispetto alla differenza con altri segmenti
2. la loro differenza è più piccola di un valore di soglia prefissata

Si devono scegliere i due parametri:

1. *threshold*, che controlla la soglia di similarità tra i segmenti. Maggiore è il valore e maggiore è la differenza (in termini di valori sui pixel/bande dell'immagine) che devono avere i segmenti e quindi maggiore è la loro dimensione; varia da 0 a 1
2. il numero minimo di pixel perché un segmento esista

La combinazione ottimale di questi due parametri è quella che massimizza l'omogeneità interna in ogni segmento e l'eterogeneità tra segmenti. I valori ottimali si possono determinare con una *unsupervised segmentation parameter optimization (USPO)*, in GRASS con il modulo *i.segment.uspo*.

L'applicazione alla carta in esame per 3 regioni attorno a Bolzano, Trento e Schio, rappresentative per tipo e geometria di segmenti, ha fornito:

Region	Thresh	Minsize	Optimization
Schio	0.40	10	1.57905345316
Trento	0.40	10	1.40471623656
Bolzano	0.35	10	1.65160867683

Tabella 1. Parametri di ottimizzazione per la segmentazione secondo USPO su tre regioni significative.

L'uso del valore 0.4 per il threshold comporta la non formazione di segmenti per molte lettere delle scritte più piccole: è stato quindi diminuito a 0.05.

La classificazione è eseguita con 11 classi, aggiungendo alle voci della legenda le classi relative a idrografia, strade e testo: (1) Italiani, (2) Tedeschi, (3) Zone montuose, (4) Paesi transalpini, (5) Idrografia, (6) Strade, (7) Confine Regno d'Italia, (8) Confine Trentino, (9) Confine geografico, (10) Città, (11) Testo.

La classificazione dei segmenti è stata effettuata mediante *machine learning* in R. Sono stati usati tutti i classificatori disponibili, *Support Vector Machine (with a radial kernel) (SVMradial)*, *Random Forest (RF)*, *Recursive partitioning (Rpart)*, e *k-Nearest Neighbors (kNN)*. Allo stesso modo, tutti i sistemi di voto disponibili, *Simple Majority Vote (SMV)*, *Simple Weighted Vote (SWV)*, *Best*

Worst Weighted Vote (BWWV) e *Quadratic Best Worst Weighted Vote (QBWWV)* sono stati usati.

Per la classificazione è stata aggiunta una banda di *texture*, per distinguere la classe "Zone montuose" dalla classe "Paesi transalpini": hanno infatti colore simile ma la seconda classe presenta una retinatura, evidente in una mappa di *texture*. È stata calcolata la *texture* sulla banda del rosso, che presenta il contrasto maggiore, con finestra 7x7 e metodo della correlazione.

Per l'addestramento del classificatore sono stati campionati 209 punti distribuiti in modo uniforme tra le classi.



Figura 2 - "Carta della distribuzione etnico-linguistica della popolazione", classificazione con *Best Worst Weighted Vote (BWWV)*; i colori delle classi sono scelti in modo che siano simili all'originale.

Rimozione delle categorie non di interesse

La rimozione delle categorie non di interesse per le successive elaborazioni, tutte quelle oltre la classe 4, è effettuata con un modulo di GRASS realizzato appositamente, *r.fill.category*. Questo modulo applica iterativamente un filtro

passa basso, usando l'operatore moda con finestra 33x33 pixel, ai soli pixel della classe da rimuovere. Per la rimozione delle 7 classi è stato creato uno script che applica ricorsivamente *r.fill.category*. Per la determinazione della dimensione del filtro da applicare è stato sviluppato un secondo modulo di GRASS *r.object.thickness* che determina larghezza minima, media e massima di oggetti di una data categoria su una carta raster. Entrambi i moduli sono automaticamente installabili come *Add-on* di GRASS.

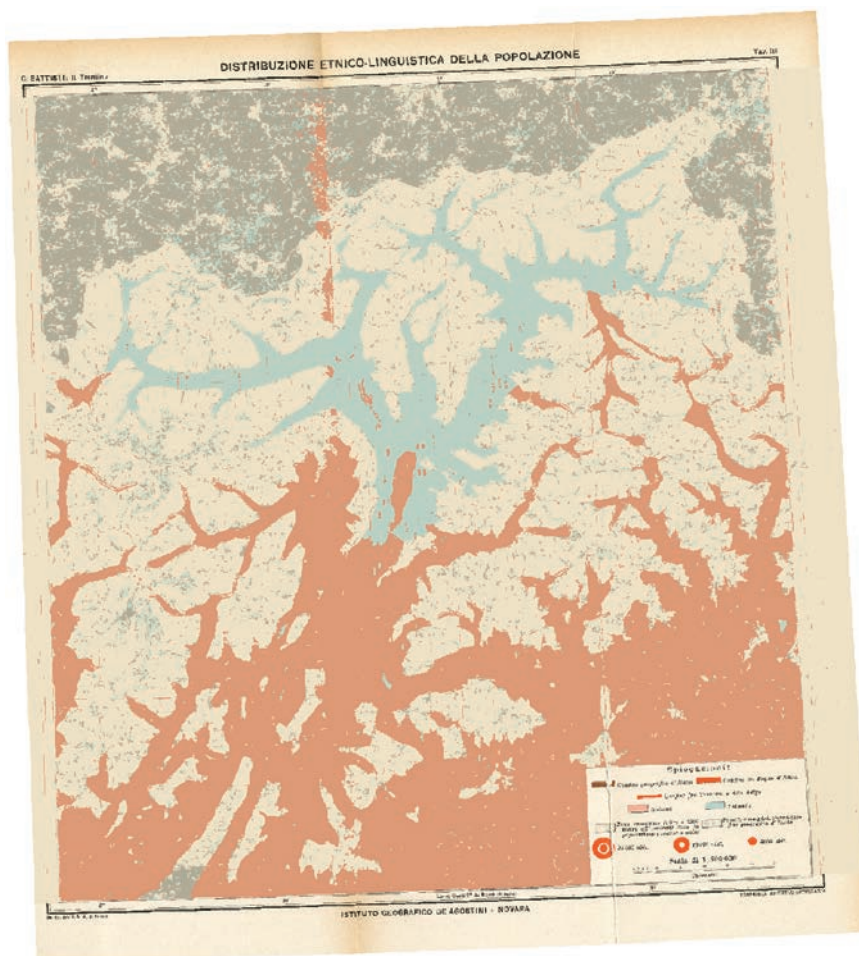


Figura 3 - "Carta della distribuzione etnico-linguistica della popolazione", classificazione con Best Worst Weighted Vote (BWWV) dopo l'eliminazione delle 7 classi non di interesse.

Valutazione dei risultati

Per la valutazione dei risultati è stata creata una carta di controllo contenente 24 aree digitalizzate manualmente. Per ogni sistema di voto è stata determinata la matrice di confusione, l'accuratezza complessiva (percentuale di pixel classificati correttamente) e la *K* di Cohen.

Gli ultimi 2 sistemi di voto sono quelli che danno il risultato migliore e, come spesso accade, assegnano gli stessi voti ed ottengono quindi gli stessi valori di accuratezza e *K di Cohen*.

La carta è stata georeferenziata nel sistema ETRS89 con proiezione UTM fuso 32N. Come riferimento è stato usato il reticolato in gradi presente sulla scansione.

Sistema di voto	Accuratezza complessiva	K
SMV	89.135855	0.848647
SWV	95.405619	0.936135
BWV	95.539406	0.937987
QBWV	95.539406	0.937987

Tabella 2. Accuratezza complessiva e *K di Cohen* per i 4 sistemi di voto.

La georeferenziazione è stata fatta nel sistema ETRS89 non proiettato e la carta è stata successivamente riproiettata in ETRS89/UTM32N (EPSG 25832). Ne risulta una carta raster con risoluzione in nord-sud di 100.5 m ed est-ovest di 69.1 m.

Conclusioni

La classificazione con *Object-based Image Analysis* si è rivelata un buon strumento per la digitalizzazione automatica di mappe storiche, come indicato dai valori in Tabella 2. Sperimentazioni su ortofoto hanno indicato la possibilità di applicare procedure di classificazione *batch*, con applicazione della mappa di addestramento costruita su una immagine ad una collezione di immagini dal contenuto simile (Balasubramaniam et al., 2019). Per migliorare il risultato è possibile aumentare/spostare i punti di *training*, provare altre coppie di *threshold* e numero minimo pixel per la segmentazione, provare altre dimensioni della finestra di calcolo della *texture* ed altri operatori e modificare il numero di classi.

L'esito è una carta in formato digitale, raster o vettoriale a seconda delle caratteristiche dei tematismi originari e/o delle successive esigenze di analisi, attraverso cui poter indagare le caratteristiche della distribuzione spaziale delle categorie rappresentate secondo le funzioni e gli strumenti propri dell'ambiente GIS. Ad esempio, per i due principali gruppi etno-linguistici della regione (in questa carta i ladini sono considerati italofoeni) è agevolmente possibile calcolare l'estensione superficiale delle aree da questi rispettivamente occupate, come riportato nella Tabella 3.

Classe	Superficie [km ²]	Superficie [%]
Italiani	8405.998	28.27

Tedeschi	2617.452	8.80
Zone montuose	14761.037	49.64
Paesi transalpini	3954.040	13.30

Tabella 3. Superfici delle classi nella carta finale (Fig. 3).

Bibliografia

Balasubramaniam, R. , Namboodiri, S. , Subrahmanyam, G. , Nidamanuri, R. (2019) *Batch-Mode Active Learning-Based Superpixel Library Generation for Very High-Resolution Aerial Image Classification*, *Advances in Intelligent Systems and Computing*, 799, 307–318. DOI: 10.1007/978-981-13-1135-2_24

Battisti C. (1915), *Il Trentino. Cenni geografici, storici, economici con un'appendice sull'Alto Adige*, Istituto Geografico De Agostini, Novara

Dai Prà E. (2010), "L'attualità della cartografia storica fra convergenze disciplinari e nuove tecnologie", *GEOmedia*, 14: 6-9.

Dai Prà E. (a cura di) (2018), *Cesare Battisti geografo e cartografo di frontiera*, CISGE, Roma.

Dai Prà E., Rossi M. (2016), "Cesare Battisti, geografo e cartografo", in Dal Prà L. (a cura di), *Tempi della storia, tempi dell'arte. Cesare Battisti tra Vienna e Roma*, Provincia Autonoma di Trento, Trento, 111-122.

Proto M. (2015), "Geografie e cartografie di guerra: «Il Trentino. Cenni geografici, storici, economici. Con un'appendice sull'Alto Adige» di Cesare Battisti (1915)", *STORICAMENTE*, 11: 1–19. DOI: 10.12977/stor617

Quaini M. (2018), "Cesare Battisti: la più grande anomalia nella storia della geografia", in Dai Prà E. (a cura di), *Cesare Battisti geografo e cartografo di frontiera*, CISGE, Roma, 27-40.

Rombai L. (2016), *Cesare Battisti (1875-1916). Geografo innovatore*, Phasar