

Ricostruzione di mappe d'estinzione interstellare basata su Machine Learning non supervisionato



Relatori

Prof. Giuseppe Longo

Dr. Massimo Brescia (INAF)

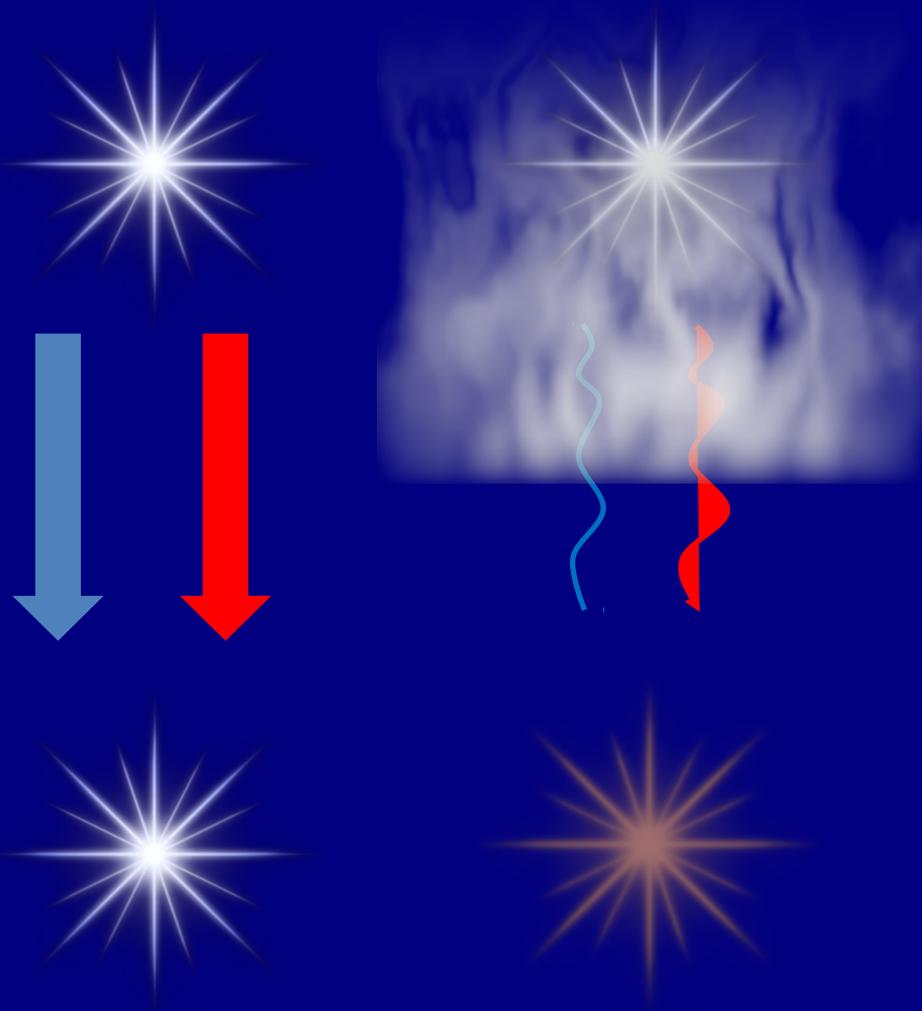
Candidato

Giuseppe Angora

N85000442

ANNO ACCADEMICO 2014/2015

Arrossamento interstellare per estinzione



L'estinzione è il termine con il quale si indica l'assorbimento e lo scattering del flusso fotonico nei gas e polveri delle regioni di formazione stellare (nebulose) posti tra osservatore e sorgente.

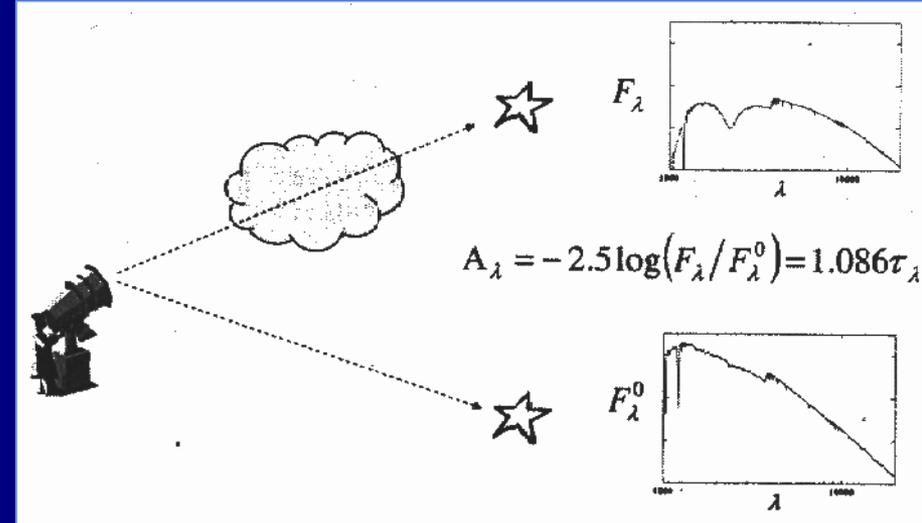
Un effetto tipico è il maggiore assorbimento della luce blu, che rende l'oggetto più rosso.

La quantità di estinzione è eterogenea, dipendendo dalla natura della regione di formazione stellare.



Metodo tradizionale di calcolo

Misurare la curva di estinzione della stella.

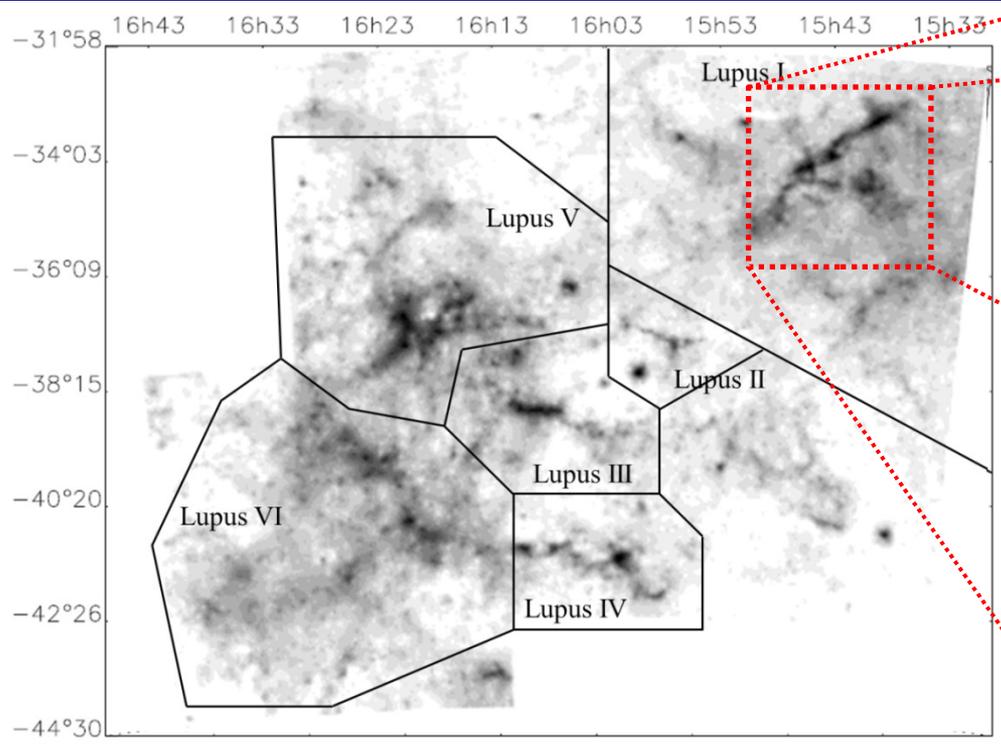


Relazione tra l'estinzione totale A_V , misurata in magnitudini, e la cosiddetta *column density* (densità di colonna) degli atomi di idrogeno N_H , misurata in cm^{-2} :

$$\left\{ \begin{array}{l} \left[\frac{N_H}{A_V} \right]_A = 1.37 \times 10^{21} \\ \left[\frac{N_H}{A_V} \right]_C = 1.87 \times 10^{21} \end{array} \right. \begin{array}{l} \longrightarrow \\ \longrightarrow \end{array} \begin{array}{l} \text{Alcal\`a 2016, INAF-OAC (com. priv.)} \\ \text{Cambr\`esy 1999, A\&A 345} \end{array}$$

$[cm^{-2} mag^{-1}]$

Nebulosa di Lupus



Mappa d'estinzione di Lupus osservata nella banda B (coordinate J2000; Cambrèsy 1999)

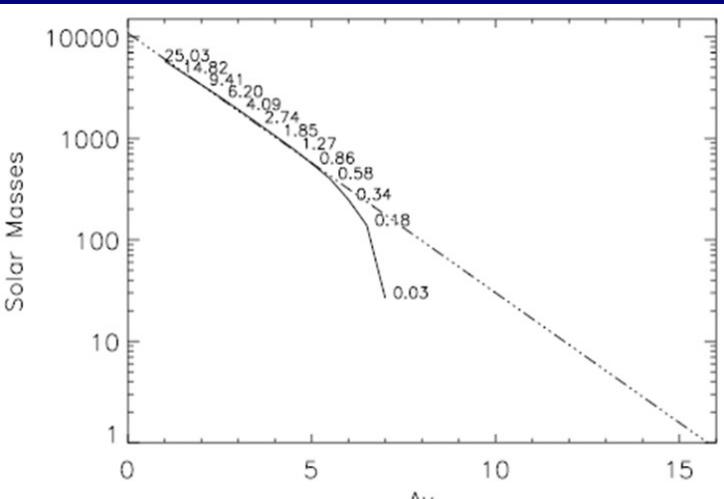


Immagine osservata da USNO-PMM di Lupus I nella banda B (coordinate J2000, 5x5 gradi DEC, 7000 x 7000 pixel di risoluzione)

La misurazione è stata effettuata attraverso un metodo di conteggio di stelle usando una griglia adattiva formata da una decomposizione basata su filtri wavelet applicata ad immagini ottiche ad alta risoluzione (Cambrèsy 1999).

Thresholding e rapporto estinzione/massa

Osservazioni di nubi molecolari su ampi intervalli di masse suggeriscono il comportamento frattale della struttura delle nebulose [Elmegreen (1996)]. La massa misurata è contenuta in un cilindro con raggio di base L e altezza indefinita H . Poiché l'estinzione A_v è correlata proprio alla grandezza H , è possibile scrivere la relazione tra massa ed estinzione:

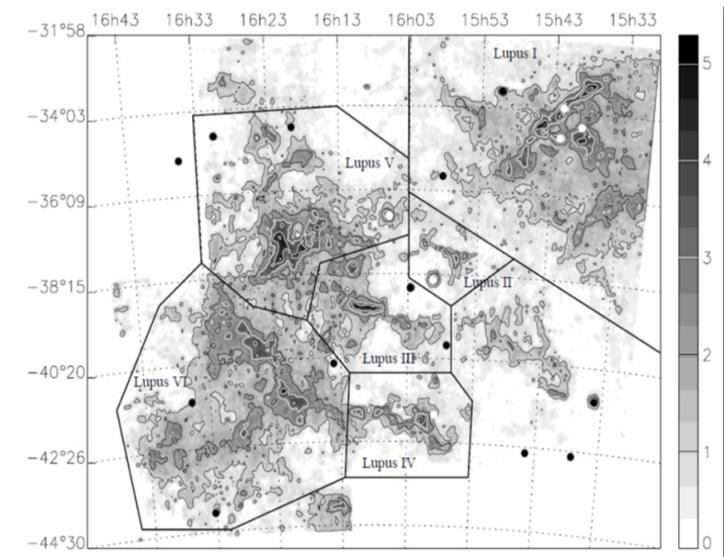


$$\log M = \log M_{tot} + slope \cdot A_v \quad [\text{Cambrèsy}(1999)]$$

$$M(A_v) = l_{pix}^2 \frac{N_H}{A_v} \mu \frac{m_H}{M_\odot} \sum_i A_v(i) \quad [\text{Dickman}(1978)]$$

$$[\mu]_A = 2.02 \quad [\mu]_C = 2.4$$

$$\frac{m_H}{M_\odot} = \frac{1.66}{1.98} \cdot 10^{-57} M_\odot$$



Regione	distanza (pc)	$A_{v,m}$	$A_{v,e}$	M_{tot}	slope
Lupus I	150	5.3	7.1	10^4	-0.56
Lupus II	150	3.8	5.7	80	-0.33
Lupus III	150	4.9	7.6	1150	-0.40
Lupus IV	150	5.3	7.0	630	-0.40
Lupus V	150	5.2	10.6	2500	-0.32
Lupus VI	150	4.8	7.0	10^4	-0.57

Problemi

Osservazioni
astronomiche



- Immagini/spettri a basso S/N
- Proiezione 2D, no profondità, foreground/background noise
- No verità assoluta (modelli/simulazioni) , **unsupervised**

Estinzione



- «effetto cataratta» non eliminabile
- Conteggio statistico affetto da errore (soprattutto distanza regione)
- Contorni regione non facilmente definibili



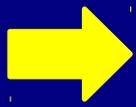
Obiettivi

Segmentazione dell'estinzione (**clustering**)



Metodo

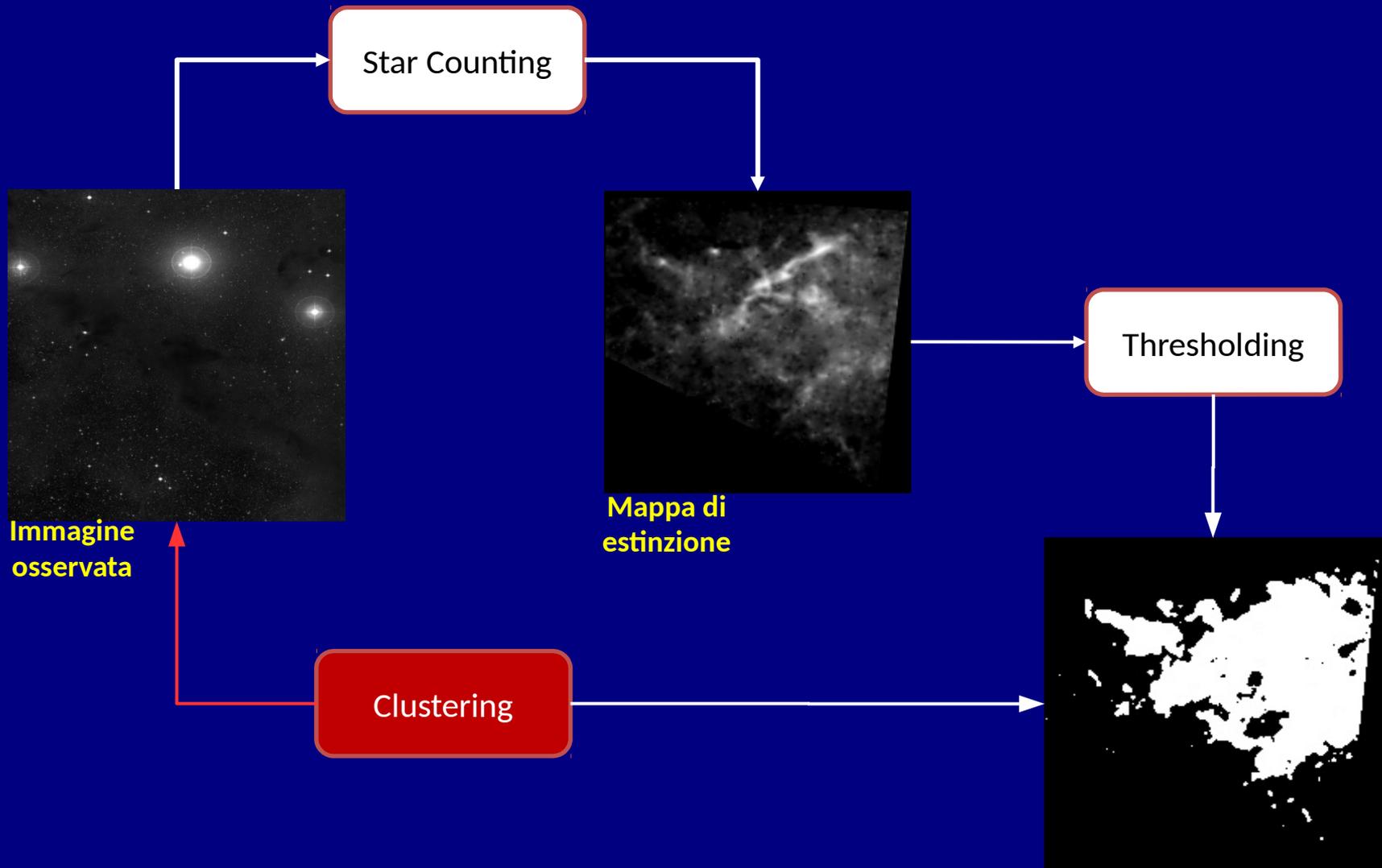
Machine Learning unsupervised



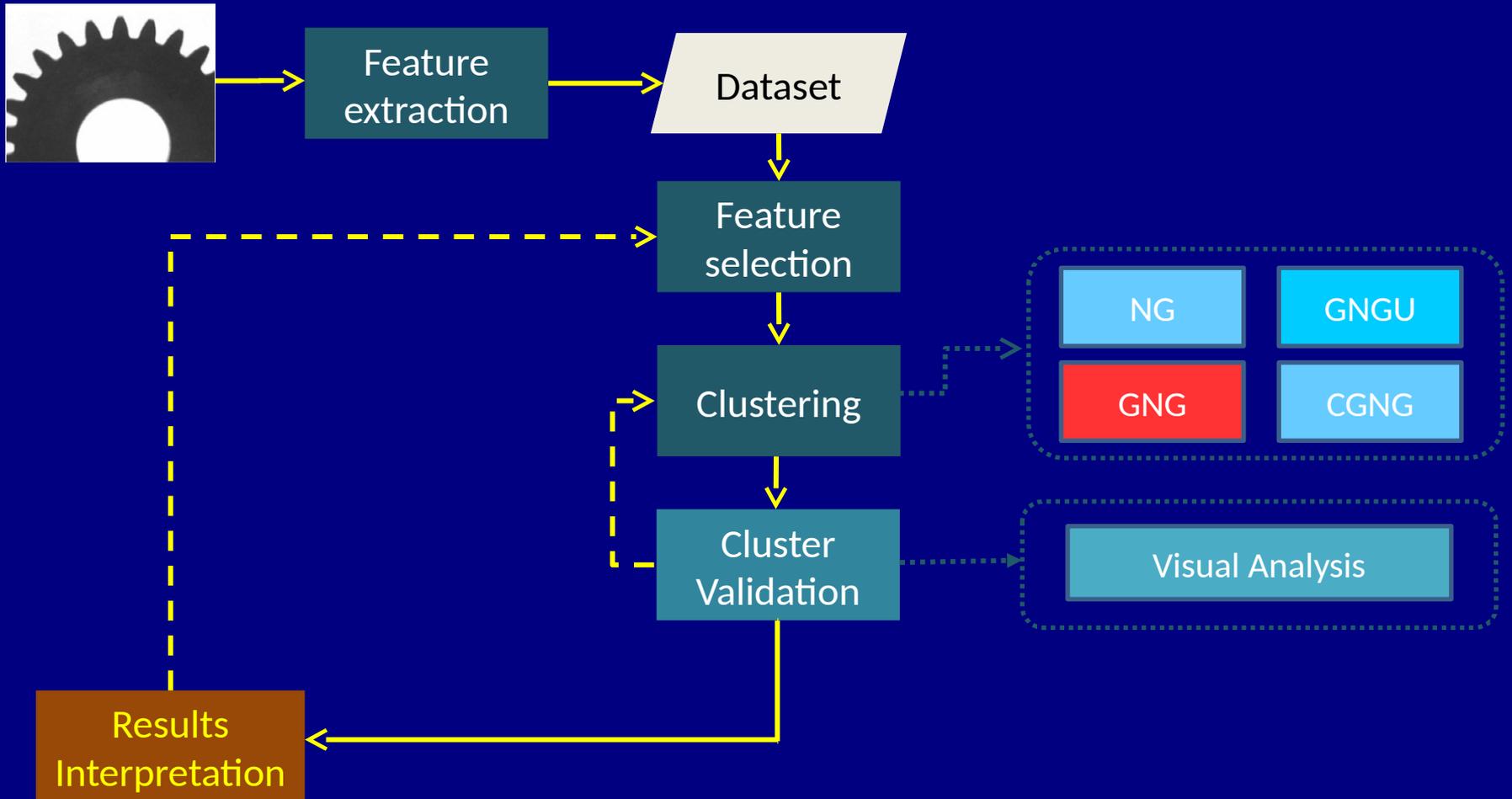
Procedura

Immagini sintetiche	Validazione numerica
Mappe d'estinzione	Validazione scientifica
Nuove immagini HD	Applicazione

Estrazione mappa di estinzione

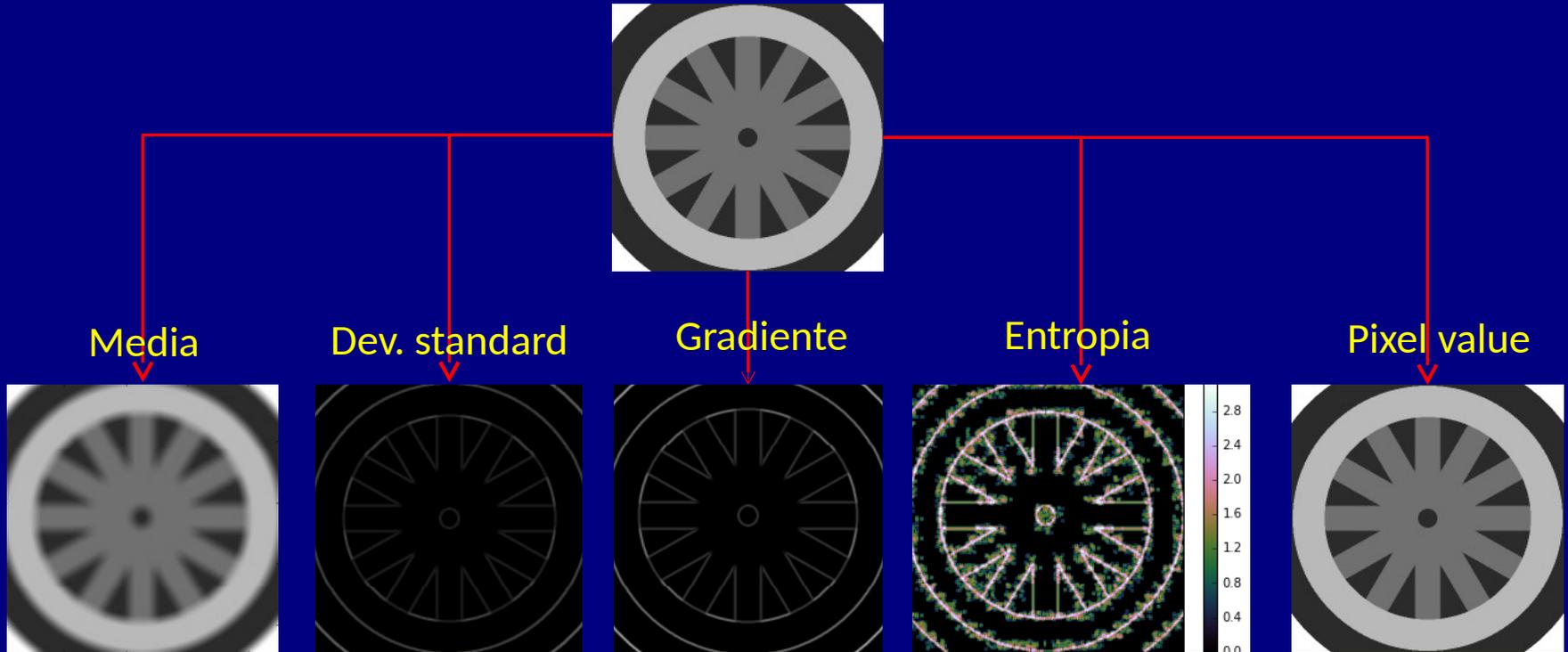


Metodo Proposto



Spazio dei parametri

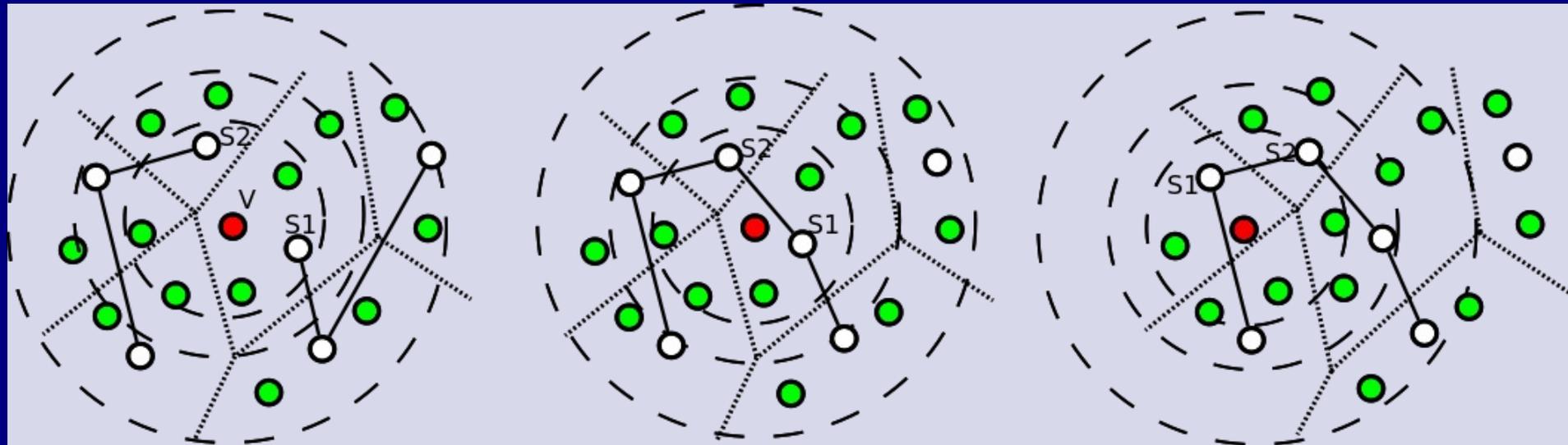
Tramite il processo di feature extraction si ottiene, per ogni pixel, un N-vettore che lo identifica nello spazio dei parametri (**PS**).



La rete Neural-Gas

Martinetz and Schulten (1994), Fritzke (1995)

Modello di quantizzazione vettoriale: definisce una topologia tra lo spazio degli input V e lo spazio dei vettori di riferimento W . Ogni vettore di input di input v è descritto dall'unità neurale vincente $S1$ rappresentata dal vettore di riferimento w , per il quale la misura della distorsione, sia minima. Lo spazio degli input risulta suddiviso nelle regioni dette di Voronoi.



Eccitazione sinaptica:

$$f(D_v) = h_\lambda(k_i(v, w)) = e^{-k_i/\lambda}$$

Regola di adattamento:

$$\Delta w_i = \epsilon \cdot f_i(D_v) \cdot (v - w_i)$$

Limitazioni della rete NG:

- Tempi di computazione elevati
- Parametri variabili nel tempo
- Apprendimento topologico del data set non sempre corretto
- Staticità del numero di unità neurali

Le reti Growing-Neural-Gas

Fritzke (1994,1995), Holmström (2002)

Le reti Growing Neural Gas sono caratterizzate da un numero di neuroni variabili durante l'apprendimento.

GNG:

$$\Delta error(s_1) = \|w_{s_1} - \xi\|^2$$

GNGU:

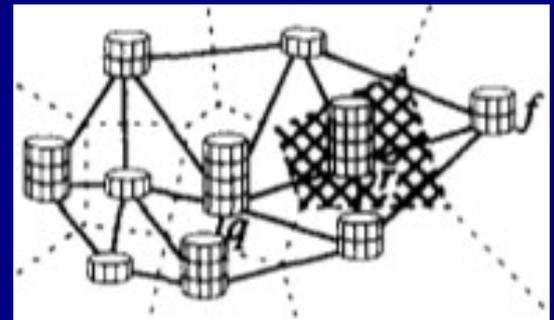
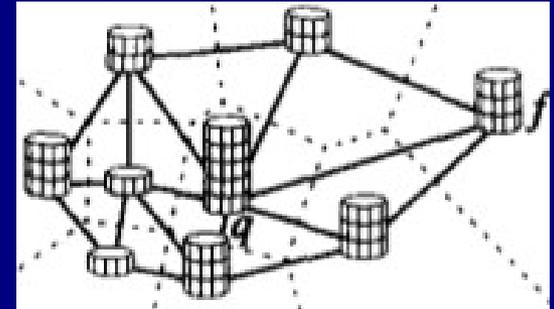
$$\Delta U_{s_1} = \|\xi - w_{s_2}\|^2 - \|\xi - w_{s_1}\|^2$$

$$\min(Utility) \cdot k < \max(localError)$$

CGNG:

$$h_c = \tau_c / \sum_{j \in W} \tau_j \quad \hat{p} = h_c / |F_c|$$

$$\bar{l}_c = \frac{1}{\text{len}(Neighbors(c))} \sum_{i \in N_c} \|w_c - w_i\|$$



L'impulso di input è ricevuto dall'unità vincente e dai suoi vicini topologici.

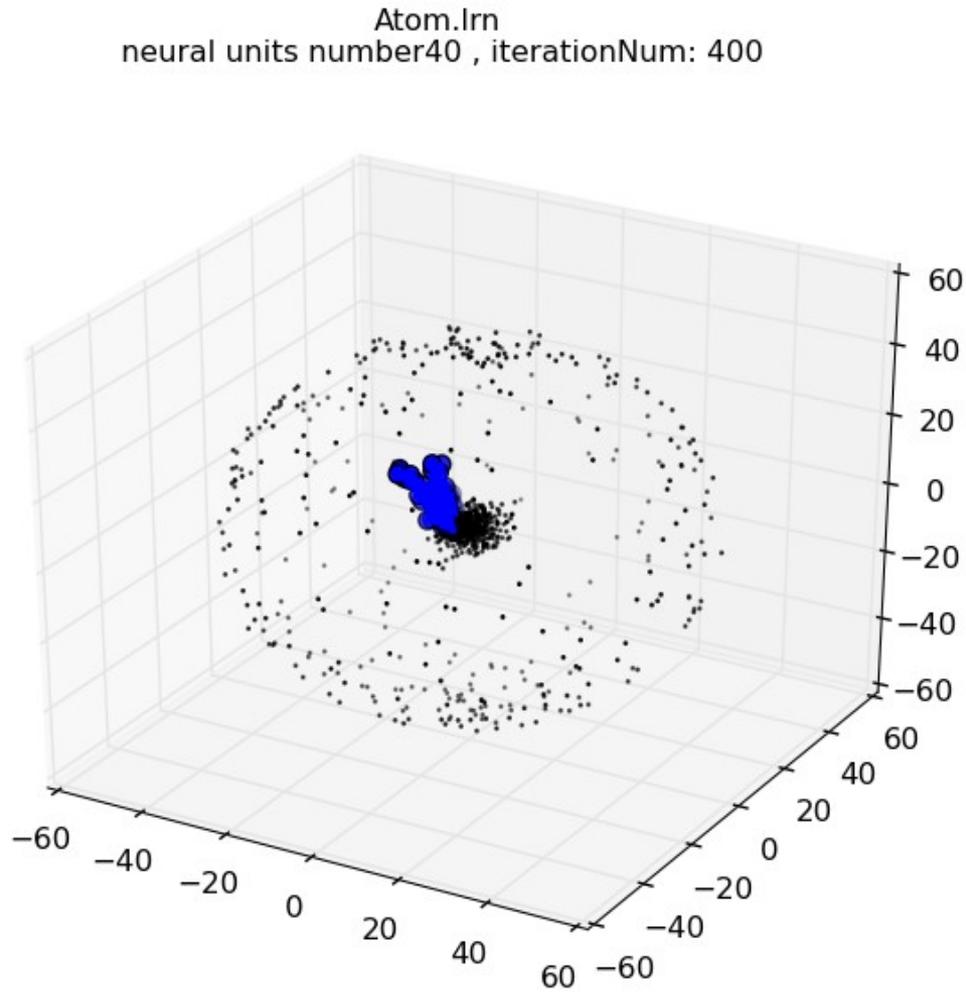


La regola di adattamento:

$$\Delta w_{s_1} = \epsilon_w (\xi - w_{s_1})$$

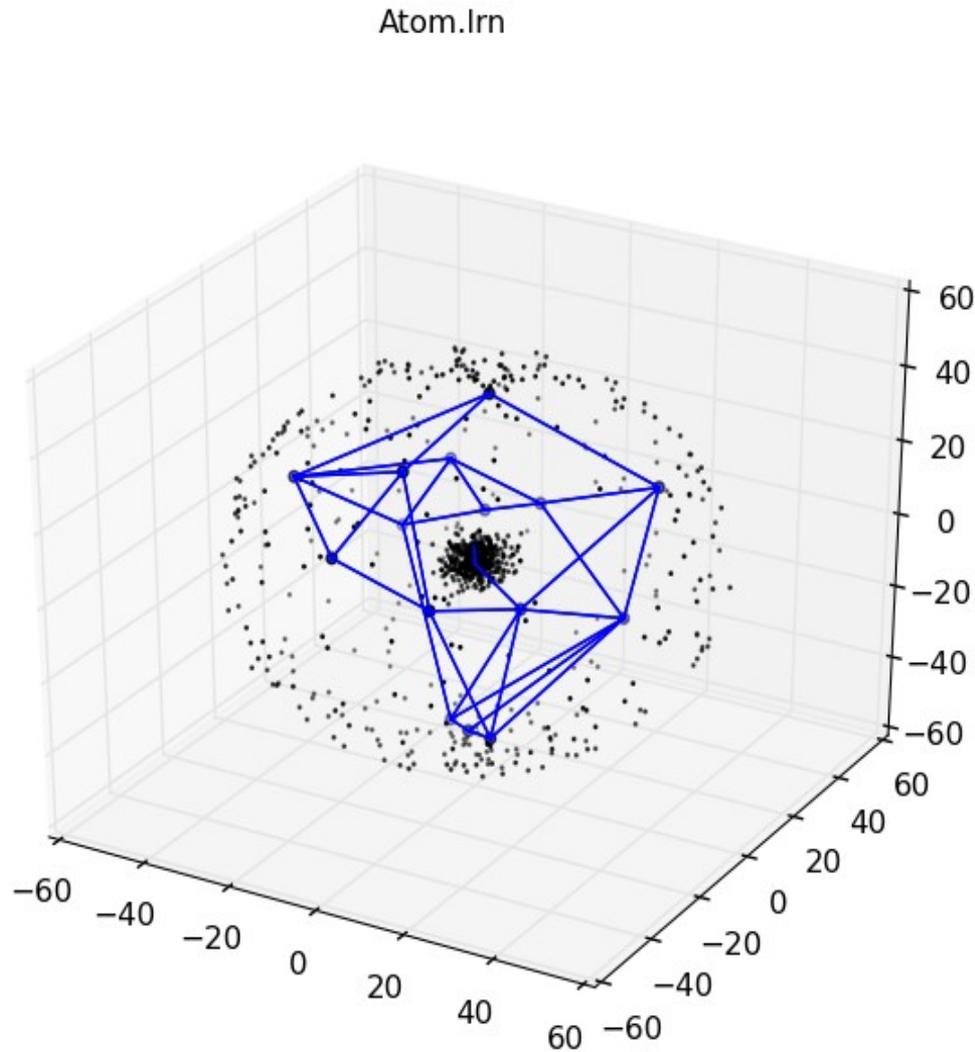
$$\Delta w_n = \epsilon_n (\xi - w_n) \quad \forall n \in Neighbors(s_1)$$

Apprendimento del data set sintetico "Atom" da parte della rete NG



Neural units	40
t_max	20000
ϵ_i	0.3
ϵ_f	0.05
λ_i	30
λ_f	0.01
T_i	20
T_f	30

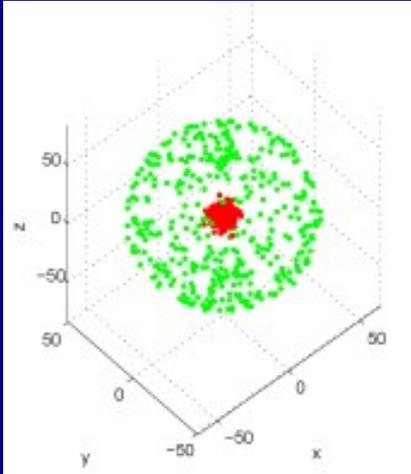
Apprendimento del data set sintetico "Atom" da parte della rete GNG



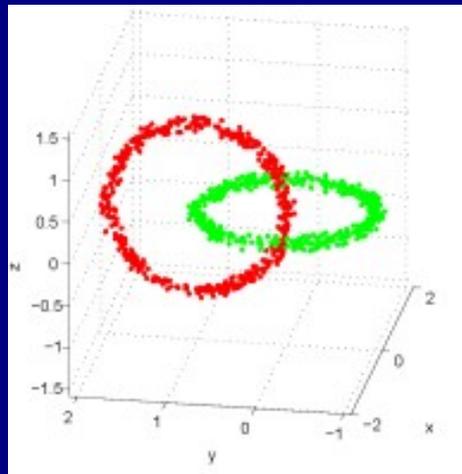
Neural units	45
ϵ_w	0.02
ϵ_n	0.0006
λ	350
age_{\max}	10
α_g	0.15
α_e	0.0005

Validazione numerica

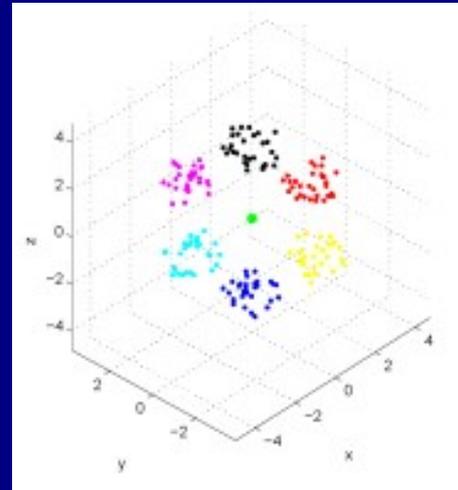
Data set "Atom":



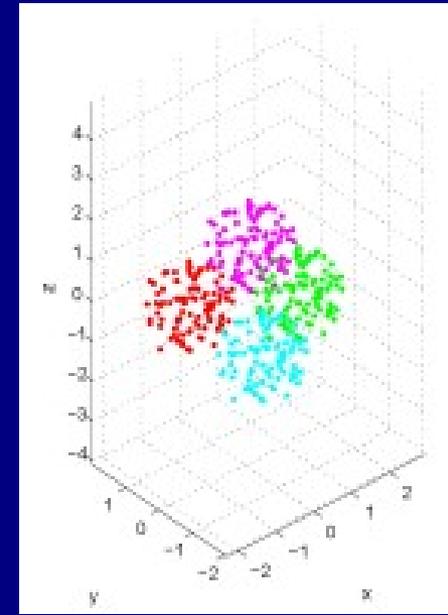
Data set "Chainlink":



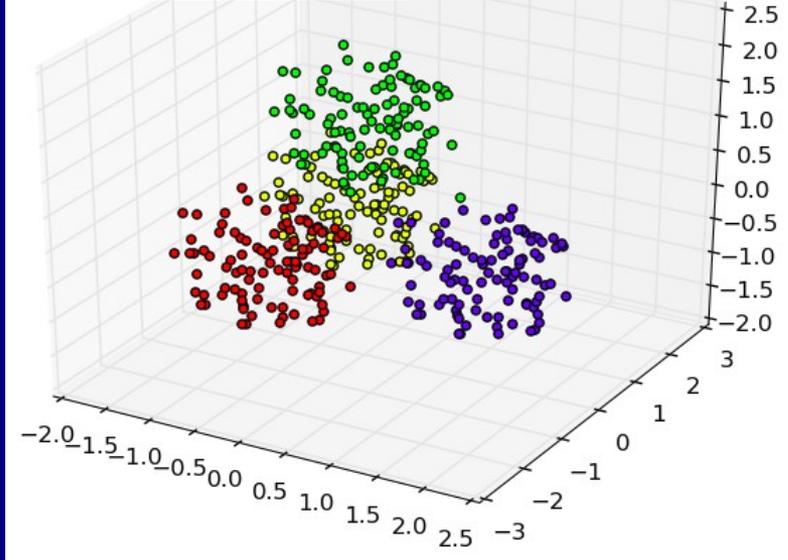
Data set "Hepta":



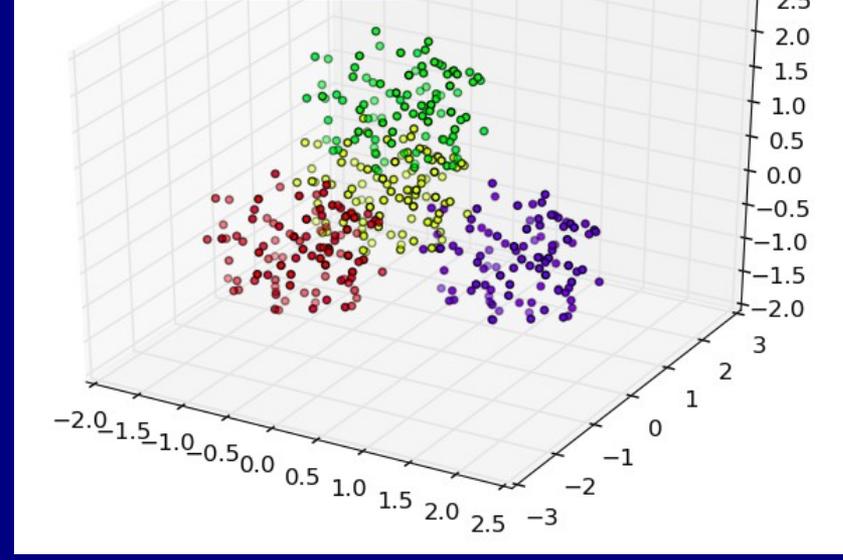
Data set "Tetra":



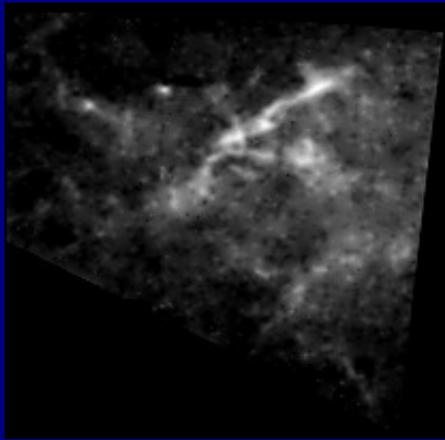
Apprendimento GNG
del data set "Tetra":



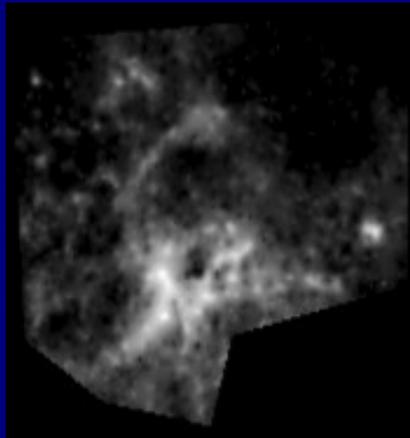
Apprendimento GNGU
del data set "Tetra":



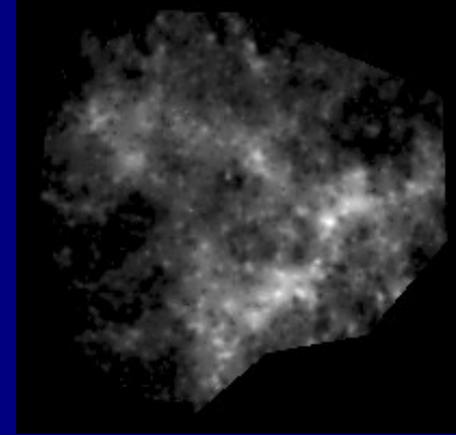
GNG su mappe di estinzione



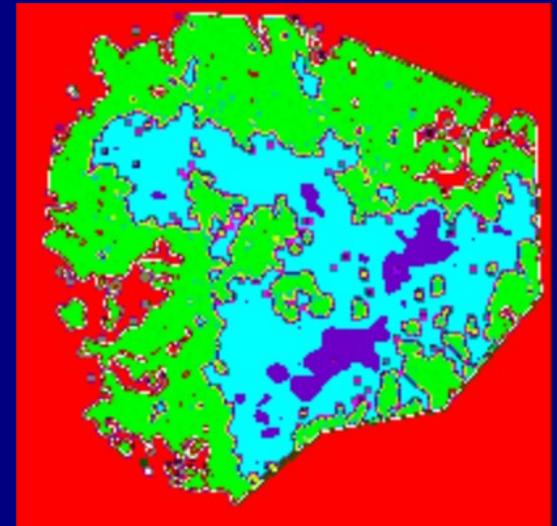
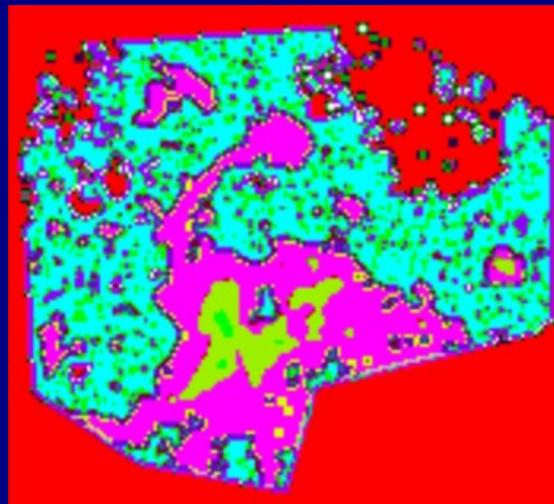
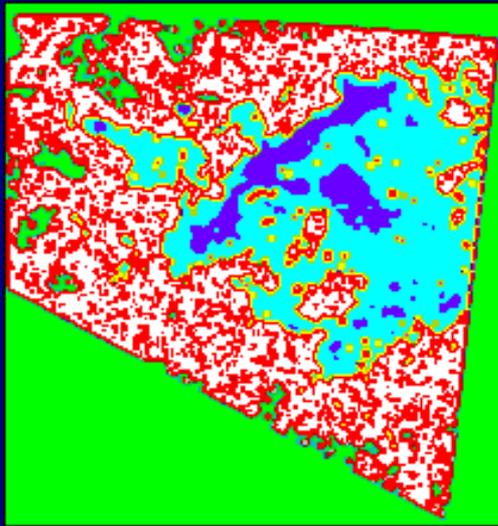
Mappa di estinzione di
Lupus I



Mappa di estinzione di
Lupus V

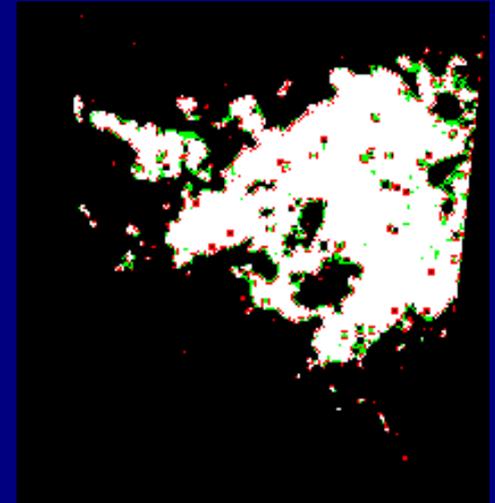
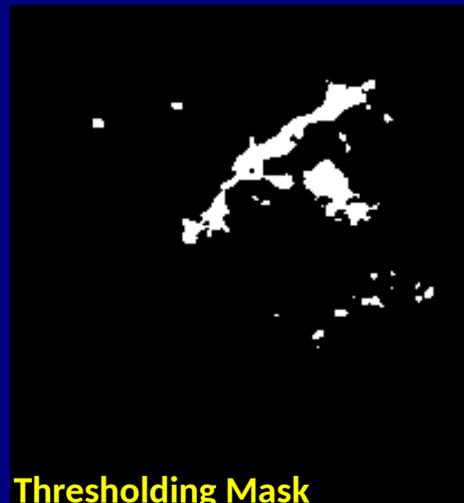


Mappa di estinzione di
Lupus VI

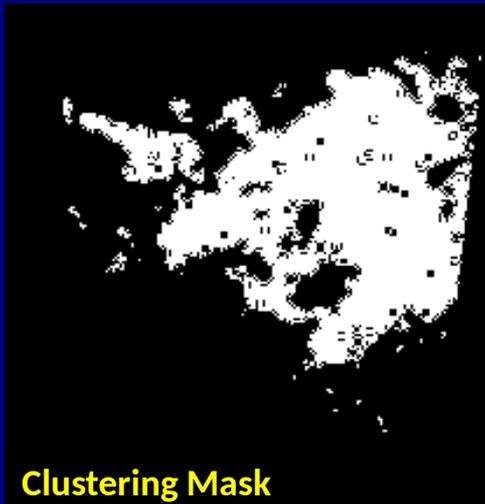
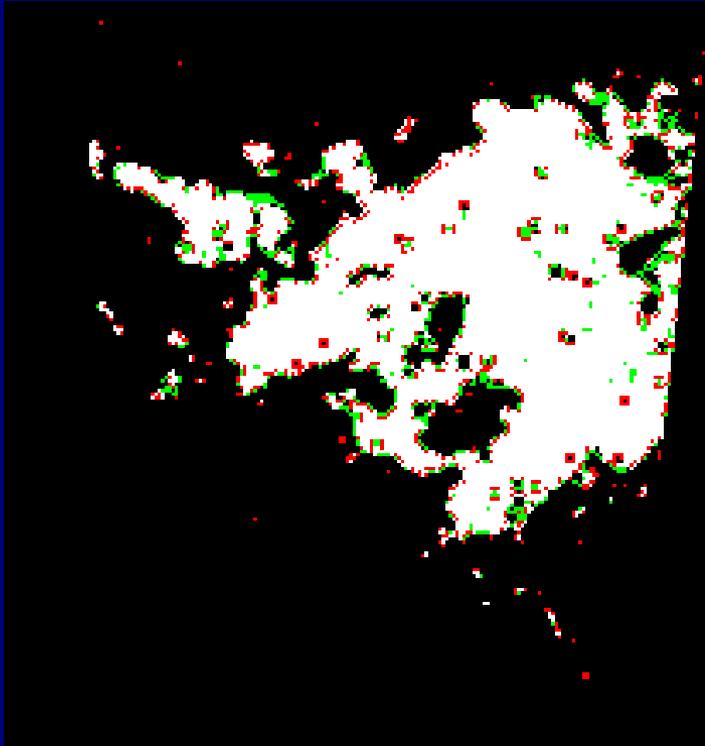


Stima della massa: validazione scientifica

Dagli apprendimenti della rete GNG è possibile estrarre delle maschere binarie identificative delle regioni di estinzione. Tali maschere possono essere confrontate con quelle ottenute tramite thresholding sulla mappa di estinzione, con valore soglia settato in modo che le maschere siano quanto più simili possibile.



Confronto tra regioni di estinzione



Clustering Mask



Thresholding Mask

Thresholding a **1.1**

Rapporto Massa
Clustering / Thresholding

M_{Cambresy}

$M_{\text{Alcalà}}$

0.66

0.66

**N° di pixel rispetto alla
maschera di thresholding**

In comune

9494

Aggiunti

828

Rimossi

986

**Variatione numero pixel
nella maschera**

1814

Estinzione media aggiunta

1.0

Conclusioni e sviluppi futuri

Multi-disciplinarietà: astrofisica + machine learning + image processing (sviluppo software in python)

Complessità: immagini ad alta risoluzione, basso S/N, foreground/background noise, no ground truth assoluta

Impatto scientifico: automazione nel processo di analisi dati in astrofisica; validità scientifica confrontabile con i metodi tradizionali, rispetto ai quali rivela vantaggi esclusivi: **miglioramento della capacità di circoscrivere e delimitare le sotto-regioni a densità di estinzione variabile; sostituzione del meccanismo arbitrario di scelta delle soglie di flusso e di conteggio delle regioni di estinzione, con tecniche automatiche.**

Sviluppi Futuri:

Approfondimento della rivelazione diretta dell'estinzione, estendendo l'analisi ad altre nebulose, raffinando la scelta dello spazio dei parametri, e la verifica dei parametri astrofisici derivati (ad es. la correlazione massa/luminosità).

Struttura frattale di nebulose

Un frattale è un oggetto geometrico dotato di omotetia interna: si ripete nella sua forma allo stesso modo su scale diverse.

La dimensione della distribuzione per le sotto-strutture all'interno di un frattale segue una relazione [Mandelbrot (1996)]:

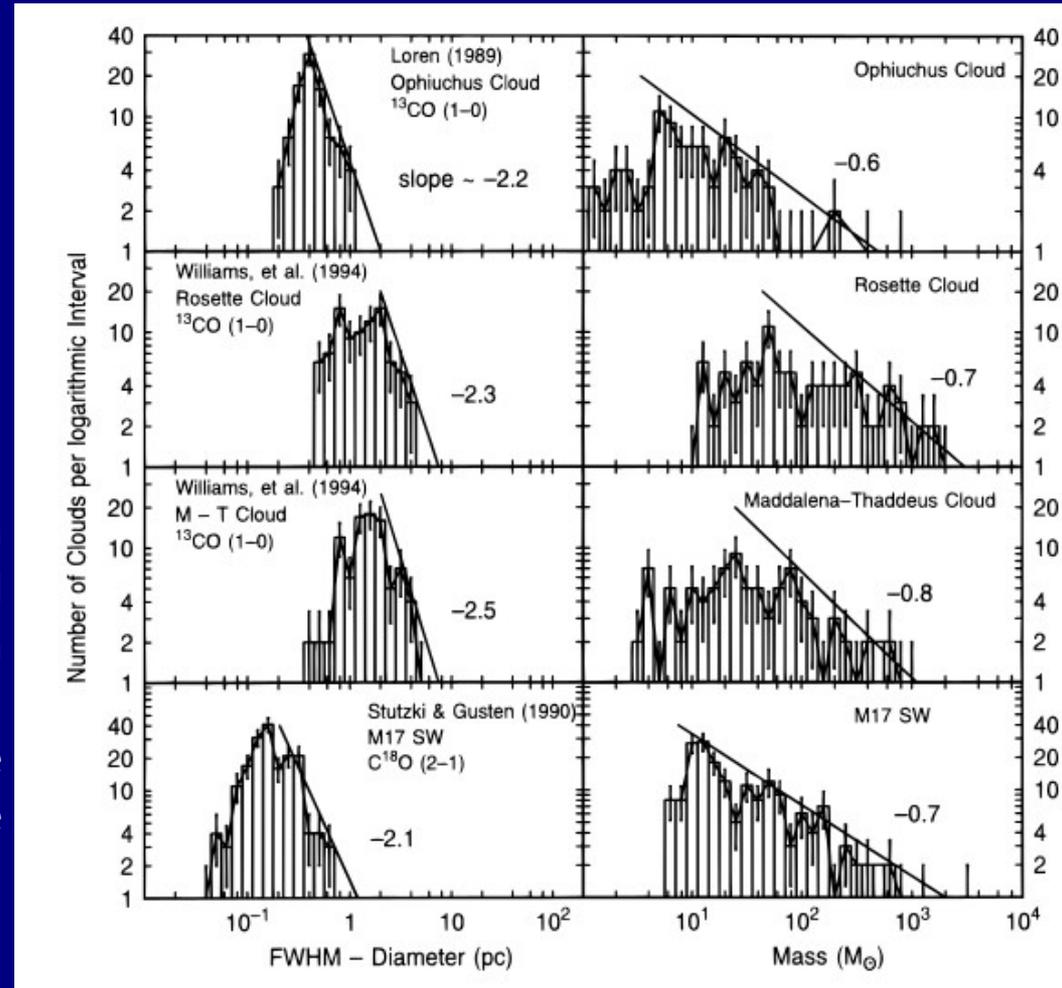
$$N(\lambda > L) \propto L^{-D}$$

Osservazioni di nubi molecolari su ampi intervalli di masse suggeriscono che esiste una legge di potenza tra la Massa e la dimensione L [Elmegreen (1996)]:

$$M \propto L^D$$

La massa misurata è contenuta in un cilindro di base L e altezza indefinita H (perché la profondità della nebulosa non è costante sulla superficie di base del cilindro), e poiché l'estinzione A_v è correlata proprio alla grandezza H, è possibile scrivere la relazione tra massa ed estinzione [Cambrésy1999]:

$$\log M = \log M_{tot} + slope \cdot A_v$$



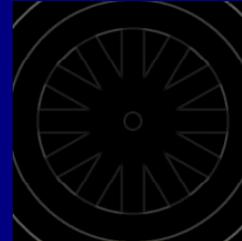
Features extraction

Ogni pixel è associato ad un vettore n-dimensionale, le cui componenti sono ottenute attraverso un'operazione realizzata su una griglia, intorno al pixel, di dimensioni 3x3, 5x5, 7x7, 9x9.

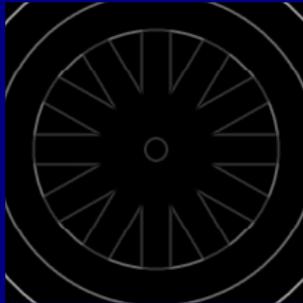
Media



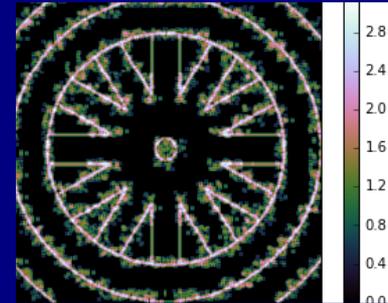
Dev. standard



Gradiente



Entropia



$$\begin{aligned} grad_{vert} &= \frac{Img(i, j) - Img(i + k, j)}{2k + 1} \\ grad_{horiz} &= \frac{Img(i, j) - Img(i, j + k)}{2k + 1} \\ grad_{diag} &= \frac{Img(i, j) - Img(i + k, j + k)}{2k + 1} \end{aligned}$$

$$entropy = - \sum_i p_i \log p_i$$

Considerazioni sulle reti sviluppate

- La struttura della rete è determinata automaticamente a partire dai dati input.
- Non è necessario imporre una dimensione massima per la rete, anzi, il processo di accrescimento può continuare indefinitamente.
- Tutti i parametri del modello sono costanti nel tempo, per cui non è necessario definire tassi di decadimento.
- L'inserimento delle nuove unità provoca una riduzione dell'errore di quantizzazione.

NG	Neural Units	ε_i	ε_f	λ_i	λ_f	T_i	T_f	T_{max}		
GNG	Neural Units	ε_w	ε_n	λ_g	α_g	α_e	age_{max}			
GNGU	Neural Units	ε_w	ε_n	λ_g	α_g	α_e	age_{max}	λ_d	α_e	k
CGNG	Neural Units	ε_w	ε_n	λ_g	α_e	η	age_{max}	λ_d		

Spazio di ricerca ampio.
Scelta dei parametri cruciale per l'apprendimento.
Risultati sub-ottimali anche per scelte coadiuvate dall'esperienza

Necessità di sviluppare un ottimizzatore per i parametri delle reti Growing

Algoritmi Evolutivi (EA)

Algoritmi Evolutivi

Beyer, Schwefel (2002); Donatti, Lomp, Würtz(2010)

Gli algoritmi evolutivi sono una classe di algoritmi di ricerca statistica basati sui principi di evoluzione darwiniana. Secondo questo paradigma di ottimizzazione, è inizializzata una popolazione i cui individui sono il risultato dell'apprendimento della rete neurale identificata da un determinato insieme di parametri denominati cromosoma.

Fitness function:

$$f_{GR}(I) = f_{SD}(I) + N \quad f_{SD}(I) = \sum_{t \in S_T} \sum_{k=1}^{\tilde{n}} d(c_k, t)$$

