

A1 Titolo dell'attività di ricerca

AstroInformatics & Machine Learning

A2 Responsabile

(aggiungere eventuale referente del Dipartimento se il Responsabile non è un afferente ad esso)

Responsabile: Giuseppe Longo

A3 Personale Dipartimento di Fisica (Professori e Ricercatori)

Prof. Ordinari: 1 (G. Longo)

Prof. Associati: 1 (M. Paolillo)

Ricercatori universitari: 1 (G. Covone)

RTDA: 1 (S. Cavuoti)

RTDB: 0

A4 Collaborazioni con altri enti

INAF, INFN, Caltech – Center for Data Driven Discovery (USA), Heidelberg Institute for Theoretical Studies (D), Harvard Smithsonian Center for Astrophysics (USA)

A5 Personale strutturato ricercatore o tecnologo altri enti convenzionati

Massimo Brescia (INAF)

A6 Altro personale di ricerca (Assegnisti, Borsisti)

Assegnisti: 0

Borsisti Post-doc: 0

Borsisti: 0

A7 Dottorandi di Ricerca

Valeria Amaro

Oleksandra Yu. Razim (Marie Curie PhD Fellowship – SUNDIAL)

B1 Breve descrizione della linea di ricerca

(max 1000 caratteri)

L'avvento delle survey digitali ha da oltre un decennio reso l'astronomia una "data rich science", aprendo la strada all'utilizzo intensivo di nuove tecnologie per l'analisi automatica e la visualizzazione di big data complessi. Il gruppo di Napoli è stato tra i pionieri di questo nuovo settore (astro informatica) ed ha acquisito una notevole visibilità internazionale. Il responsabile è stato chairperson del KDD (Knowledge Discovery Interest Group dell'IVOA), del working group della missione Euclid sul machine learning, delegato nazionale delle azioni europee COST TU1208 e TD1304, membro fondatore dell'interest Group on Astronomical data mining dell'IEEE.

Al momento sono attive proficue collaborazioni con oltre una decina di istituti stranieri tra cui il California Institute of Technology, l'Institute for Theoretical Studies di Heidelberg, l'Harvard Smithsonian Center for Astrophysics, l'Osservatorio di Ginevra (CH), ec.. Il responsabile è chairperson permanente

della serie di convegni internazionali ASTROINFORMATICS ed è spesso co-chair di altri convegni nel settore organizzati da IEEE, EU etc. Uno dei membri del gruppo (SC) ha vinto il premio dell'International Astrostatistics Association per la migliore pubblicazione.

Le metodologie sviluppate trovano applicazione in pressoché tutte le ricerche svolte dalla sezione in area astrofisica e cosmologica, alle cui schede si rimanda per ulteriori approfondimenti.

B2 Descrizione attività svolta nel triennio 2014-2016

(max 2000 caratteri)

L'attività svolta nel triennio può essere così riassunta:

- 1 – Messa a punto di metodi innovativi (MLPQNA, GAME e LEMON) basati sul machine learning, per la risoluzione di problemi di regressione e, in particolare, del software METAPHOR per la derivazione di Probability Distribution Functions con metodi di ML.
- 2 – Mantenimento della piattaforma di data mining DAME sviluppata in collaborazione con INAF e Caltech;
- 4 – Messa a punto di algoritmi di deep learning per l'identificazione automatica di eventi di strong lensing in dati di survey;
- 5 - Utilizzo delle metodologie per la derivazione di redshift fotometrici per oltre 60 milioni di galassie nella Data Release 10 della Sloan Digital Sky Survey e di svariati milioni di galassie dalle survey VST-VOICE e KiDS. Partecipazione ad alcuni data challenge internazionali per la derivazione di photometric redshift (PHAT, Euclid DC1, Euclid DC2) con risultati di eccellenza. Ciò ha fruttato la leadership del gruppo di Photometric Redshift con machine learning per la missione spaziale Euclid che sarà lanciata nel 2019 dall'ESA (scheda di Cosmologia osservativa).
- 6 – Partecipazione al gruppo di lavoro su Automatic Data Quality per la missione spaziale Euclid
- 7 - Applicazione delle metodologie di cui sopra allo studio di fenomeni transienti e realizzazione della pipeline STRADIWA per la simulazione di oggetti variabili in immagini astronomiche (scheda di Astrofisica delle Alte Energie).
- 8 – Applicazione delle metodologie di clustering allo studio della distribuzione di ammassi globulari nell'ammasso di Fornace (scheda di Cosmologia osservativa).
- 9 – Utilizzo di tecniche di clustering per definire la mappatura spettroscopica dello spazio dei parametri fotometrici di Euclid.
- 10 - Applicazione delle metodologie alla diagnostica precoce dell'Alzheimer.
- 11 - Applicazioni a tecniche di visualizzazione di dati complessi con utilizzo di mondi virtuali, Oculus Rift e Hololens.

B3 Descrizione attività programmata nel triennio 2017-2019

(max 2000 caratteri)

Si deve far fronte a numerosi impegni assunti dal gruppo nei confronti della comunità internazionale e, in particolare:

- 1) Implementazione della pipeline METAPHOR nella pipeline di analisi dati di Euclid e distribuzione dei prodotti alla comunità Euclid (oltre 1800 persone in Europa e USA).
- 2) Parallelizzazione e porting sotto GPU dei principali algoritmi sviluppati nel triennio precedente;
- 3) Implementazione di una pipeline per il riconoscimento automatico di eventi di lensing gravitazionale forte usando deep learning (KiDS, Euclid)
- 4) Completamento di programmi di osservazione spettroscopica di sorgenti selezionate per completare la mappatura dello spazio dei parametri fotometrici del satellite Euclid.
- 5) Applicazione dei metodi ai dati delle survey KiDS e VST-VOICE nelle loro varie data release ed utilizzo dei dati prodotti per una serie di indagini sulla struttura cosmica su grande scala

C1 Pubblicazioni scientifiche nel triennio 2014-2016

(indicare il numero complessivo nel triennio e elencare le più significative (max 10))

N. complessivo: 30 (di cui 21 referate)

1. Cavuoti, S.; Amaro, V.; Brescia, M.; et al; 2017, *A cooperative approach among methods for photometric redshifts estimation: an application to KiDS data*, Monthly Notices of the Royal Astronomical Society Volume 466, Issue 2, p. 2039-2053
2. Riccio, G.; Brescia, M., Cavuoti, S.; et al; 2017, *C 3, A Command-line Catalog Cross-match Tool for Large Astrophysical Catalogs*, Publications of the Astronomical Society of the Pacific, Volume 129, Number 972
3. Cavuoti, S.; Amaro, V.; Brescia, M.; et al; 2017, *METAPHOR: a machine-learning-based method for the probability density estimation of photometric redshifts*, Monthly Notices of the Royal Astronomical Society Volume 465, Issue 2, p.1959-1973
4. D'Isanto, A.; Cavuoti, S.; Brescia, M.; et al.; 2016, *An analysis of feature relevance in the classification of astronomical transients with machine learning methods*, Monthly Notices of the Royal Astronomical Society, Volume 457, Issue 3, p.3119-3132
5. Tortora, C.; La Barbera, F.; Napolitano, et al.; 2016, *Towards a census of supercompact massive galaxies in the Kilo Degree Survey*, Monthly Notices of the Royal Astronomical Society, Volume 457, Issue 3, p.2845-2854
6. Masters, D.; Capak, P.; Stern, D.; et al.; 2015, *Mapping the Galaxy Color-Redshift Relation: Optimal Photometric Redshift Calibration Strategies for Cosmology Surveys*, The Astrophysical Journal, Volume 813, Issue 1, article id. 53, 15 pp.
7. Mercurio, A.; Merluzzi, P.; Busarello, G.; et al.; 2015, *Shapley Supercluster Survey: construction of the photometric catalogues and i-band data release*, Monthly Notices of the Royal Astronomical Society, Volume 453, Issue 4, p.3685-3698
8. Cavuoti, S.; Brescia, M.; Tortora, C.; et al.; 2015, *Machine-learning-based photometric redshifts for galaxies of the ESO Kilo-Degree Survey data release 2*, Monthly Notices of the Royal Astronomical Society, 452 3, 3100-3105
9. de Jong, J. T. A.; Verdoes K. G. A.; Boxhoorn, D. R.; et al.; 2015, *The first and second data releases of the Kilo-Degree Survey*, Astronomy & Astrophysics, Volume 582, id.A62, 26 pp.
10. Brescia, M.; Cavuoti, S.; & Longo, G.; 2015, *Automated physical classification in the SDSS DR10. A catalogue of candidate quasars*, Monthly Notices of the Royal Astronomical Society, 450, 4, 3893-3903

C2 Presentazioni a Conferenze internazionali e nazionali

(solo se lo speaker è tra il personale elencato nel punto A3)

Oltre 30 interventi in totale tra cui i più significativi:

1. G. Longo, "Big Data From Space", Frascati, 2014
2. G. Longo, Astrominformatics 2015, Dubrovnik, Croazia
3. G. Longo, Astrominformatics 2014, Viña del Mar, Chile
4. S. Cavuoti: "A cooperative approach among methods for photometric redshifts estimation." October 2016, Sorrento (Italy) Astrominformatics 2016
5. S. Cavuoti: "Data Quality Common Tools status" May 2016, Lisbon (Portugal) Euclid Consortium Meeting 2016
6. S. Cavuoti: "Data-Rich Astronomy: Mining Sky Surveys with PhotoRAPToR" May 2014, Lisbon (Portugal) Statistical Challenges in Cosmology in the 21st Century
7. S. Cavuoti: "Photometric Redshifts for KIDS with MLPQNA" March 2014, Naples (Italy) KIDS Science meeting
8. S. Cavuoti: "Astroinformatics for dummies" January 2014 Melbourne (Australia) Galah Workshop, Monash University, Invited Lecture

C3 Presentazioni di brevetti internazionali e nazionali

nessuno

D1 Progetti di ricerca attivi

(Progetti di Enti di ricerca, Progetti Europei, Progetti MIUR, PON, POR, ...)

1. **EU Horizon 2020, Marie Curie ITN SUNDIAL.** (2017-2020). P.I. locale G. Longo
2. **FARO 2011:** M.Paolillo, responsabile del progetto "Time Domain Astronomy: automatic detection of astronomical transients in synoptic surveys"
3. **COST-Action TD1403.** G. Longo Delegato Nazionale
4. **COST Action TU 1208. G. Longo Delegato Nazionale**
5. **FP7-SPACE-2013-1,** S. Cavuoti membro del progetto "VIALACTEA - The Milky Way as a Star Formation Engine", financed by the Programme "European Commission 7th Framework Programme for Research"
6. **LSST,** G. Longo and S. Cavuoti membri del progetto "Supernovae Demography and Rates based on Machine Learning Classification"
7. **E-Rosita:** S. Cavuoti membro
8. **ESO project:** S. Cavuoti Co-Investigator of the: GCAV - Galaxy Clusters at Vircam;
9. **ESO project:** S. Cavuoti Co-Investigator of the: VST-GAME - Galaxy Assembly as a function of Mass and Environment with VST;
10. **ESO project:** S. Cavuoti Co-Investigator of the: KABS - KiDS-ATLAS Bridging Survey.
11. **KiDS:** S. Cavuoti e G. Longo membri
12. **Euclid:** G. Longo e S. Cavuoti membri EUCLID Mission System Team responsible of Data Quality Common Tools for the Science Ground

Segment. Members of EUCLID Legacy Science Working Group for transient discovery and Organization Units OU-PHZ for photometric redshifts.

13. **S. Cavuoti:** member of the team for the development of common scientific libs for Euclid