
HubbleCLIP: Associating Astronomical Observations and Natural Language with Multi-modal Models

Siddharth Mishra-Sharma

smsharma@mit.edu

The NSF AI Institute for Artificial Intelligence and Fundamental Interactions

Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Department of Physics, Harvard University, Cambridge, MA 02138, USA

Yiding Song

ydsong@mit.edu

The NSF AI Institute for Artificial Intelligence and Fundamental Interactions

Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Jesse Thaler

jthaler@mit.edu

The NSF AI Institute for Artificial Intelligence and Fundamental Interactions

Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

We present a multi-modal model which associates astronomical observations imaged by the *Hubble* Space Telescope (HST) with natural language. The model is fine-tuned from a base CLIP model using summarized proposal abstracts corresponding to HST observations. We show that the model embodies a meaningful joint representation between observations and text through experiments targeting observation retrieval (i.e., retrieving most relevant set of observations using natural language queries) and description retrieval (i.e., querying the astrophysical object classes and science use cases most relevant to a given observation). The model demonstrates the potential for using generalist rather than task-specific models for astrophysics research, in particular by leveraging text as an interface.

1 Introduction

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

2 Dataset and Processing

2.1 Summarization via guided generation

3 Methodology

3.1 Language-Image Pre-training

3.2 Pre-trained CLIP Model

3.3 Fine-tuning Objectives

4 Results and Discussion

4.1 Fine-tuned Retrieval Accuracy

4.2 ‘Zero-shot’ Hypothesis and Object Retrieval

4.3 Text-to-Image Retrieval

Hinton et al. (2006)

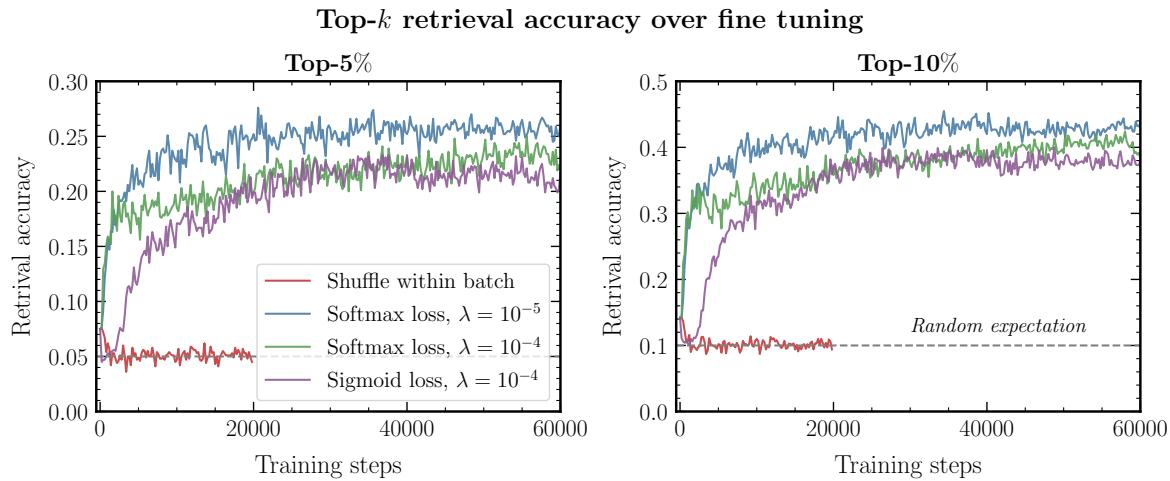


Figure 1: Retrieval accuracy

Broader Impact Statement

In this optional section, TMLR encourages authors to discuss possible repercussions of their work, notably any potential negative impact that a user of this research should be aware of. Authors should consult the TMLR Ethics Guidelines available on the TMLR website for guidance on how to approach this subject.

Author Contributions

If you’d like to, you may include a section for author contributions as is done in many journals. This is optional and at the discretion of the authors. Only add this information once your submission is accepted and deanonymized.

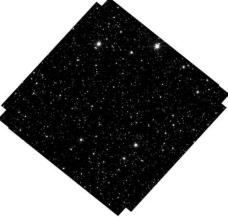
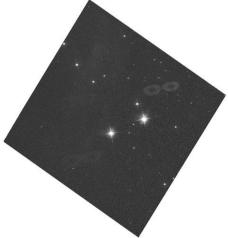
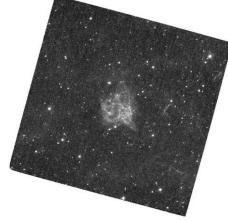
Image	Obs. cycle	Prop. ID	Summarized abstract
	26	15513	isolated black holes, old stellar population, background stars, Galactic bulge, microlensing events, light deflection, astrometric microlensing, 100-day events, low-mass stars, relative proper motions; long-duration events, T>300-day events, observed durations, BHs with mass > 10 Msun, T>300 days, very-long duration events, OGLE detects six T>300 days events each year, monitoring a few T>300-day events
	19	12577	interstellar dust, Cas A outburst, historical supernova SN, supernova remnant (SNR), astrophysical objects, light echoes, dust in the Milky Way, cooling envelope, shock breakout, progenitor star; observing its light echoes, studying large numbers of objects, connecting how a single object should change with time, time series of spectra, spatially resolved light curve, measuring the properties of the cooling envelope, estimating the radius of the progenitor star, connecting the progenitor star to the explosion to the SN to the SNR
	7	7340	oxygen-rich supernova remnants, LMC, SMC, O III λ 5007 emission-line, S II $\lambda\lambda$ 6724 emission-line, O III $\lambda\lambda$ 3727 emission-line, ionization structure, chemically peculiar debris, metal-rich plasmas, supernova explosions, active pulsar, synchrotron nebula, young supernova remnant, extended complex of young stars ; models for nucleosynthesis in massive stars, excitation mechanisms in extremely metal-rich plasmas, the dynamics of supernova explosions, validating observed phenomena in astrophysical contexts
	22	13757	type Iax supernovae, white dwarfs, luminous blue systems, S1, companion stars, accretion onto white dwarfs, massive stars, unrelated stars, HST pre-explosion data, NGC 1309; progenitor systems, thermonuclear white dwarf supernovae, white dwarf supernovae, confirmation, characterization

Table 1: Data overview

Acknowledgments

Use unnumbered third level headings for the acknowledgments. All acknowledgments, including those to funding agencies, go at the end of the paper. Only add this information once your submission is accepted and deanonymized.

References

Geoffrey E. Hinton, Simon Osindero, and Yee Whye Teh. A fast learning algorithm for deep belief nets. *Neural Computation*, 18:1527–1554, 2006.

Image-to-Text Retrieval

Image	Top classes (fine-tuned)	Top classes (base)	Abstract
	<ul style="list-style-type: none"> 1. dark matter 2. Einstein rings 3. galaxy mergers 4. gravitational lensing 5. galaxy clusters 	<ul style="list-style-type: none"> 1. galaxy clusters 2. ultra diffuse galaxies 3. dwarf galaxies 4. gravitational lensing 5. crowded stellar field 	<p>Category: COSMOLOGY. We propose to study the physical nature of dark matter by using massive, merging clusters of galaxies. As shown with the Bullet Cluster (1E0657-56), such massive well-measured systems are critical for our understanding of dark matter. By more than doubling the number of clusters in the sample and obtaining systems at different observation angles, impact parameters, geometrical arrangements, and merger velocities, the systematic uncertainties in the dark matter cross section calculations can be improved substantially, allowing us to move from rough order of magnitude estimates to measurements with quantifiable uncertainties that can be compared usefully with the predictions from numerical simulations, and the constraints on alternate gravity models become unambiguous. Our proposed targets are three extraordinary, merging galaxy clusters with X-ray and optical offsets that are placed at ideal redshifts for such a study: A520, A1758N, and A2163. To pin down the position of the dark matter component we require high resolution, absolutely calibrated mass maps. High resolution gravitational lensing data is needed to attain this goal, which can only be achieved with the excellent resolving power of the HST.</p>
	<ul style="list-style-type: none"> 1. dark matter 2. galaxy mergers 3. Einstein rings 4. gravitational lensing 5. dark energy 	<ul style="list-style-type: none"> 1. ultra diffuse galaxies 2. galaxy clusters 3. gravitational lensing 4. high-redshift quasars 5. dwarf galaxies 	<p>Category: COSMOLOGY. We propose to study the physical nature of dark matter by using massive, merging clusters of galaxies. As shown with the Bullet Cluster (1E0657-56), such massive well-measured systems are critical for our understanding of dark matter. By more than doubling the number of clusters in the sample and obtaining systems at different observation angles, impact parameters, geometrical arrangements, and merger velocities, the systematic uncertainties in the dark matter cross section calculations can be improved substantially, allowing us to move from rough order of magnitude estimates to measurements with quantifiable uncertainties that can be compared usefully with the predictions from numerical simulations, and the constraints on alternate gravity models become unambiguous. Our proposed targets are three extraordinary, merging galaxy clusters with X-ray and optical offsets that are placed at ideal redshifts for such a study: A520, A1758N, and A2163. To pin down the position of the dark matter component we require high resolution, absolutely calibrated mass maps. High resolution gravitational lensing data is needed to attain this goal, which can only be achieved with the excellent resolving power of the HST.</p>
	<ul style="list-style-type: none"> 1. crowded stellar field 2. supernova remnants 3. compact stellar remnants 4. primordial black holes 5. pre-main sequence stars 	<ul style="list-style-type: none"> 1. stellar abundances 2. stellar populations 3. interstellar medium 4. pre-main sequence stars 5. Cepheid variables 	<p>Category: RESOLVED STELLAR POPULATIONS. Exploiting the full power of the Wide Field Camera 3 (WFC3), we propose deep panchromatic imaging of four fields in the Galactic bulge. These data will enable a sensitive dissection of its stellar populations, using a new set of reddening-free photometric indices we have constructed from multiple filters across UV, optical, and infrared wavelengths. These indices will provide the most accurate and complete catalogues for the properties of thousands of individual bulge stars. Proper motions of these stars derived from multiepoch observations will allow separation of the stellar populations. Using these photometric and astrometric tools, we will reconstruct the detailed star formation history as a function of position within the bulge, and thus differentiate between rapid and slow star formation. This will allow us to probe the connection between metallicity and star formation rate, revealing how the characteristic mass of star formation varies with chemistry. Our sample of bulge stars with accurate metallicities will extend the range of stellar populations that have been studied. In addition, as the WFC3 has a very large field of view, our measurements will extend this knowledge to a remote environment with a very distinct chemistry. Our proposal also includes observations of six well-studied globular and open star clusters; these observations will serve to calibrate our photometric and astrometric tools, and to provide a testbed for the use of these tools in the outer disk and halo of the Galaxy. Besides enabling our own program, these products will provide powerful new tools for a host of other stellar-population investigations with HST/WFC3. We will deliver all of the products from this Treasury Program to the community in a timely fashion.</p>
	<ul style="list-style-type: none"> 1. low surface brightness galaxies 2. star formation histories 3. galaxy formation 4. ultra diffuse galaxies 5. circumgalactic medium 	<ul style="list-style-type: none"> 1. gravitational lensing 2. high-redshift quasars 3. brown dwarfs 4. trans-Neptunian objects 5. Kuiper Belt objects 	<p>Category: Stellar Populations and the Interstellar Medium. Observations of the ultra-faint dwarfs (UFDs), as relics of the epoch of reionization, allow us to probe the earliest epochs of star formation (SF). In particular, the UFDs in low density environments for most of their lives provide unique tools to probe the effects of early environmental conditions on the SF history of SFHs of the UFDs and their stellar populations. We have been able to identify a set of UFDs that have the best record of early local environments. We propose to obtain deep ACS and UVIS imaging in F606W and F814W for 2 LMC-like UFDs at low redshift. This program is designed to identify systematic differences in the stellar populations of recently captured UFDs vs. long-term MW satellites (data available from previous programs) by using high-fidelity color-magnitude diagrams constrained by the same set of filters. The main goals of this program are: (1) to determine whether the SFH of the UFDs is constant or changes with time, and whether the SFH is constant or changes with environment; (2) to determine whether SF is quenched at different times with different rate in UFDs in low density environment at early times, probing the patchiness of reionization by directly comparing with theoretical predictions; (3) identify variations in the sub-Solar IMF across UFDs born in different environments; (4) pave the way for a more accurate constraint on the MW halo mass.</p>
	<ul style="list-style-type: none"> 1. supernovae 2. Cepheid variables 3. star clusters 4. star forming galaxies 5. dust 	<ul style="list-style-type: none"> 1. high-redshift quasars 2. gravitational lensing 3. Kuiper Belt objects 4. compact stellar remnants 5. ultra diffuse galaxies 	<p>Category: RESOLVED STELLAR POPULATIONS. We propose to test two of the clearest predictions of the theory of evolution of massive-star evolution: 1) the formation of Wolf-Rayet stars depends strongly on these stars' metallicity (Z), with relatively fewer WR stars forming at lower Z, and 2) Wolf-Rayet stars die as Type Ib or Ic supernovae. To carry out these tests we propose to use the HST to study the stellar populations of the M101 supercluster. This supercluster is an important, well-studied system that has been extensively observed with ground-based telescopes. Our previous HST survey of the Hill sphere of the S101 complex found that the WR population in the central parts of the supercluster is dominated by the recent star-forming history of these large complexes: young cores of O and WR stars are surrounded by older halos containing older stars. We propose to use the HST to study the WR populations in the outer parts of the supercluster. The main goal of this proposal is to directly test this paradigm in a single galaxy, M101 being the ideal target. The abundance gradient across M101 (a factor of 20) suggests that relatively many more WR will be found in the outer parts of the galaxy than in the inner parts. We will use the HST to study the WR populations in the outer parts of the supercluster and compare them to their inner parts as a function of radius. The WR population in M101 may be abundant enough for one to erupt as a Type Ib or Ic supernova within a generation. The clear WR populations in the outer parts of the supercluster are likely to be dominated by WR stars that are relatively heavily populated by WR stars, are common in M101. It is widely claimed that such superclusters produce the integrated luminosity of thousands of Wolf-Rayet stars located in hundreds of M101 superclusters, and correlate those numbers against the supercluster sizes and luminosities. It is likely (but far from certain) that Supercluster sizes and emission-line luminosities are driven by their Wolf-Rayet star content. Our sample will be the largest and best-ever Supercluster/Wolf-Rayet sample, an excellent local proxy for characterizing starburst galaxies in superclusters.</p>

Figure 2: Retrieval accuracy

A Appendix

You may include other additional sections here.

Text-to-Image Retrieval: Base Model

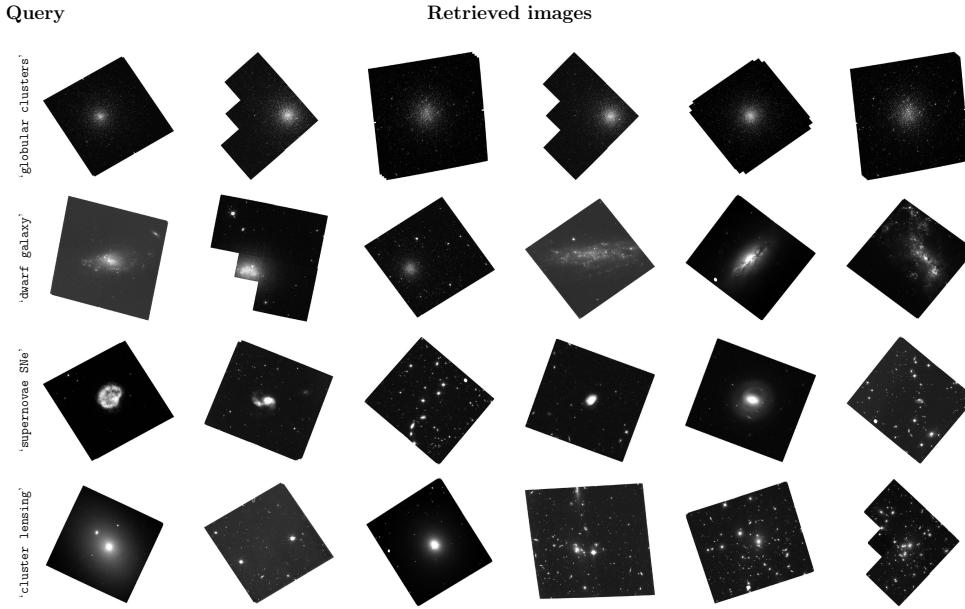


Figure 3: Retrieval accuracy

Text-to-Image Retrieval: Fine-Tuned Model

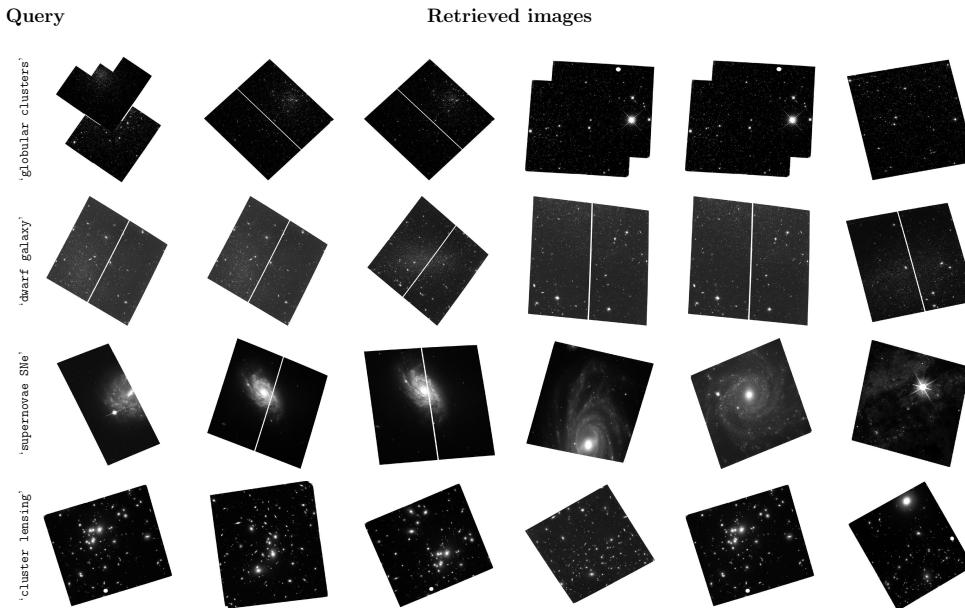


Figure 4: Retrieval accuracy