Community Science with the VVV

Universidad Tecnológica Nacional
Facultad Regional Córdoba

Juan Bautista Cabral, Sebastian Gurovich, John Good, Ricardo Medel, Diego Garcia Lambas, Eduardo Amôres, Juan Jose Claria, Dante Minniti + CEPRAM & VVV science teams

IATE CONICET UNC

1 Abstract

We are developing a web framework for community-science projects with the Vista Variables in the Via Lactea Survey. The first phase of the framework, the main subject of this poster is to classify background galaxies and detect possible supernovae within them. This framework and its analysis is being developed as part of Juan Bautista Cabral's PhD thesis.

2 The Vista Variables in the Via Lactea Survey

VVV: The Vista Variables in the Via Lactea Public ESO Survey (Minniti et al 2010) is a ~ 7 year multi-epoch, multi-band near-IR mapping of the galactic bulge and adjacent disk section of The Galaxy in the broad-band Y, Z, J, H & Ks filters. The first epoch of observations are complete for all bands and all tiles however the multi-epoch campaign that is ~ half way complete is expected to be completed by ~2017. In the multi-epoch campaign all pawprints/tiles are being re-observed in the Ks-band and it is expected that most fields will have about 100 epochs and that the cadence of overlapping tile regions will be ~2-3 times higher. The VVV mapping with data rates of ~ 300 GiB/night is providing a rich database for time-series analysis of galactic and extragalactic variable sources some of which include transient phenomena.

Community Science: This consists in facilitating collaborators be they researchers or laymen to access scientific data and to participate in the analysis and discovery of relevant scientific information. The data being produced by the VISTA Data Flow System Emerson et al. (VDFS: 2004) includes both catalogue and image data products. The sheer volume and nature of these products offer a possibility for the development of community science type applications that we are considering that include:

i) Identification and Classification of candidate galaxies
ii) Monitoring of candidate galaxies for transients
iii) Classification of light curve
iv) Monitoring of the galactic centre
v) Monitoring of the ecliptic plane

Phase 1: We now focus on developing applications i) and ii): Several types of galaxies and supernovae (SNae) are prevalent in the universe and marked correlations exists between them including SNae rates and host galaxy properties. The three most common SNae are the core collapsed SNae that include type II, 1b and 1c and the non-core collapsed SNae of type Ia. In a recent study Hakobyan et al. (2011) showed that all SNae types exhibit characteristics which depend on the properties of the host including colour, age, metallicity, morphology, and even structural components. The rate of core-collapse (CC) SNae for example has been confirmed to be closely related to the Star Formation Rate (SFR) of the prominent stellar population in galaxies and only indirectly to the total host mass. Interestingly for SNe Ia, 15% of Ia SNae rates are found in spiral galaxies with relatively young stellar populations. We wish to contribute understanding to this field, with our framework.

3 VVV Galaxies

Phase 1 (cont.): This project involves the eyeball detection/ classification of galaxies and the monitoring of supernovae within them. Since we wish to make eye-ball detections of galaxies in the VVV Tile images, to increase sensitivity, the production of colour composite RGB images is first required so we create colour composite RGB images with the Montage toolkit, and with the Aplpy python module. In the case of Montage the colour composite images have been made by re-projecting the J, H and Ks image data to the same image projection as defined by the Ks tile frames. In this process we use a fast tangent plane to plane gnomic projection algorithm. Three individual re-projections are required (1 for each band). A schematic depiction of the algorithm for just one re-projection is shown in Figure 1. In Figure 1 an arbitrary input image pixel in band J is re-projected to an output pixel with geometry defined by Ks-band. In this process the total input pixel flux is accurately projected as it is weighted by the sky area of the overlap. The output pixel value is the area-weighted average of the images being combined. We choose a colour table to be sensitive from 0 signal to maximum with a Gaussian-log scaling.

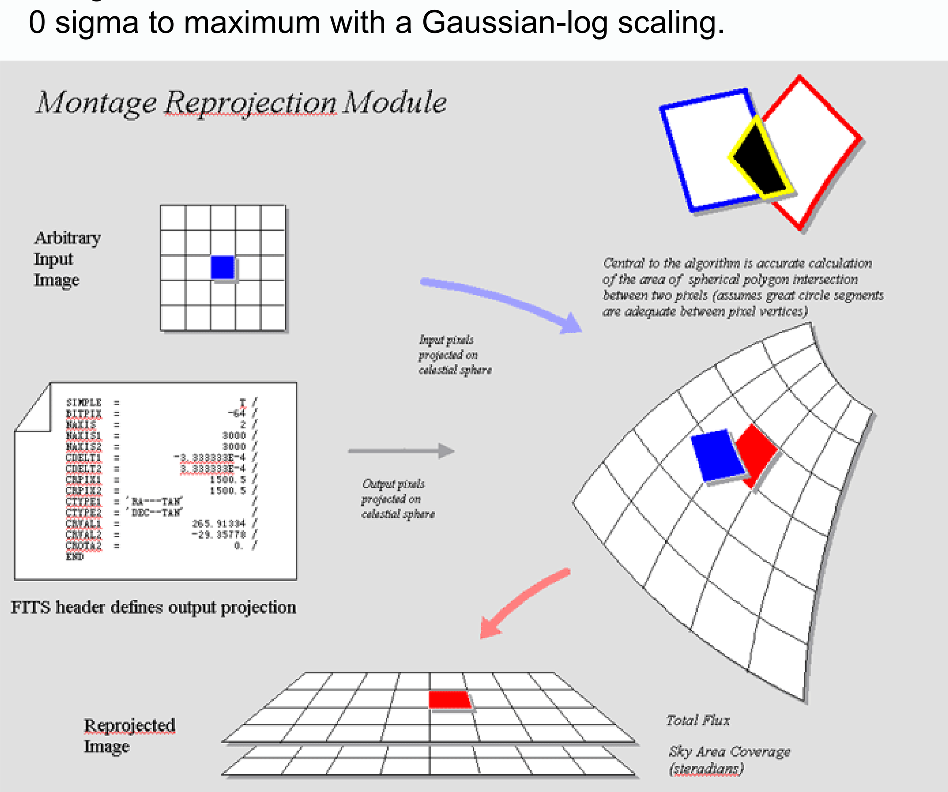

Figure 1: Reprojecting J image to the K image projection (adapted from Montage web-site)

Human Classification: In leading towards the web implementation of our framework, to improve our method, we are working with a group of people from the Centre For the Promotion of Adult Education (CEPRAM for the castellano acronym.) The CEPRAM folk have attended lectures given by our group and have decided with us on suitable classification methods.

4 Making Science Accessible

Training set: To date, students have studied images of galaxies from the Sloan Digital Sky Survey as well as from the "Atlas de las galaxias australes"(Sersic et al. 1968). This course focused on galaxy morphology and photometric properties. As a continuation of the course we are initiating supplementary community science activities and have selected a representative sample of VVV galaxies from our colour composite RGB VVV Tile images (disk and bulge) with clear galaxy morphological signatures. Of these selected galaxies, rectangular cut-out images are been collated as a training set to provide a representative sample for eye-ball detection and classification using simple instructions for the detection campaign of galaxies.

Expanding the set: Amôres et al. (2012) in studying the VVV disk tile d003 estimated that in the VVV footprint there may be ~30000 background galaxies. Most of these have larger size distributions than stars and in color-color space tend to lie away from the main stellar locus (see Figures 2,3). So to add to our sample, we intend to include into the training set some of these sources, mainly those that show hints of galaxy morphology.

We note: Until elements of our framework are sufficiently mature, we are using Aladin (Bonnarel et al 2000) as our visualization tool.

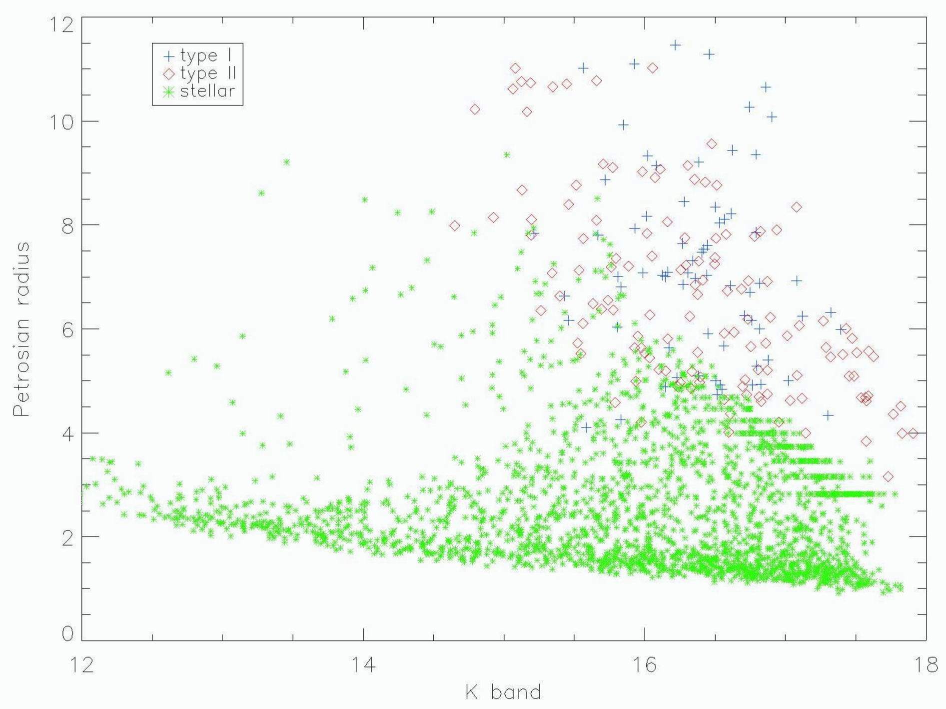

Figure 2: Size distribution for Tile d003, showing galaxy candidates, from Amôres et al. (2012).

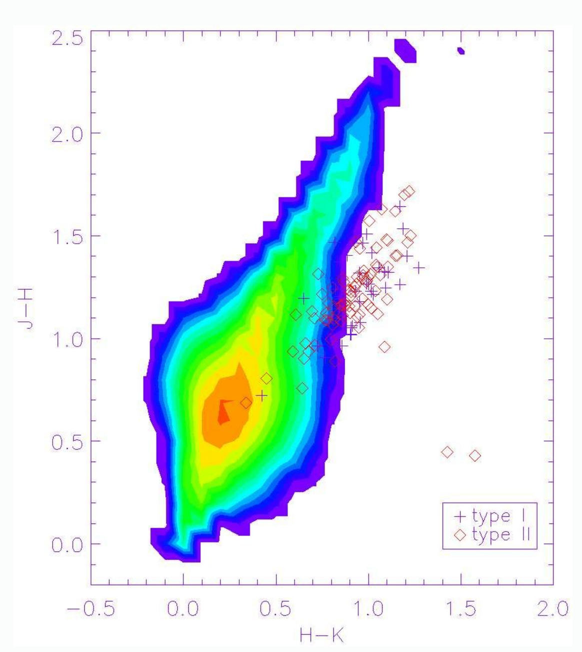

Figure 3: Colour-colour distribution, showing galaxy candidates, from Amôres et al. (2012).

5 Phase 1 of the Framework

Database: The database for Phase 1 (i) will contain all the ingested RGB cutout images (5 x 5 arcmin), hereby called "fraction", and additional meta data. both from the re-projected header, including: filename, ra,dec (WCS), tile number, filter name, obs-mjd, radii, etc., for all bands. In Figure 4, a simplified database is shown.

Functionality: As part of the web framework, after the user authenticates his/her identity they will be presented with a random "fraction" from the data base.

1) If the analyst believes this random "fraction" contains a galaxy, they will be prompted to draw a box around the galaxy(ies).
2) If the analyst is confident that their "fraction" does not contain any galaxy then will be able to flag it as such.
3) It the analyst prefers not to flag that "fraction", but to leave it for a future analysis they will be able to ignore it.
4) The software internally registers each of these choices into a table. When a galaxy is found, the coords of the galaxy will be registered.
5) Later a new cutout image will be presented to the user.

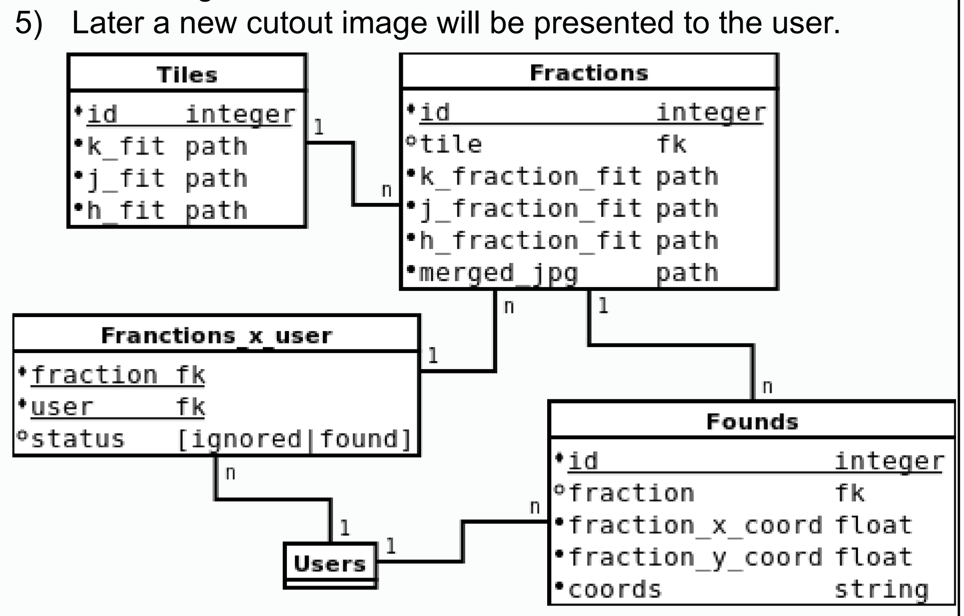

Figure 4: Extended database for the detection of galaxies

6 RGB VVV colour cutouts

Cutouts or Fractions: Figure 5 shows examples of "fractions" for the tile b204 with angular scale-size (in red). These obviously selected "fractions" have mostly bright central galaxies. We chose these to show that for the larger galaxies, basic morphological information will be available that will be useful when determining SNae-galaxy correlations with part ii) of our phase 1 experiment.

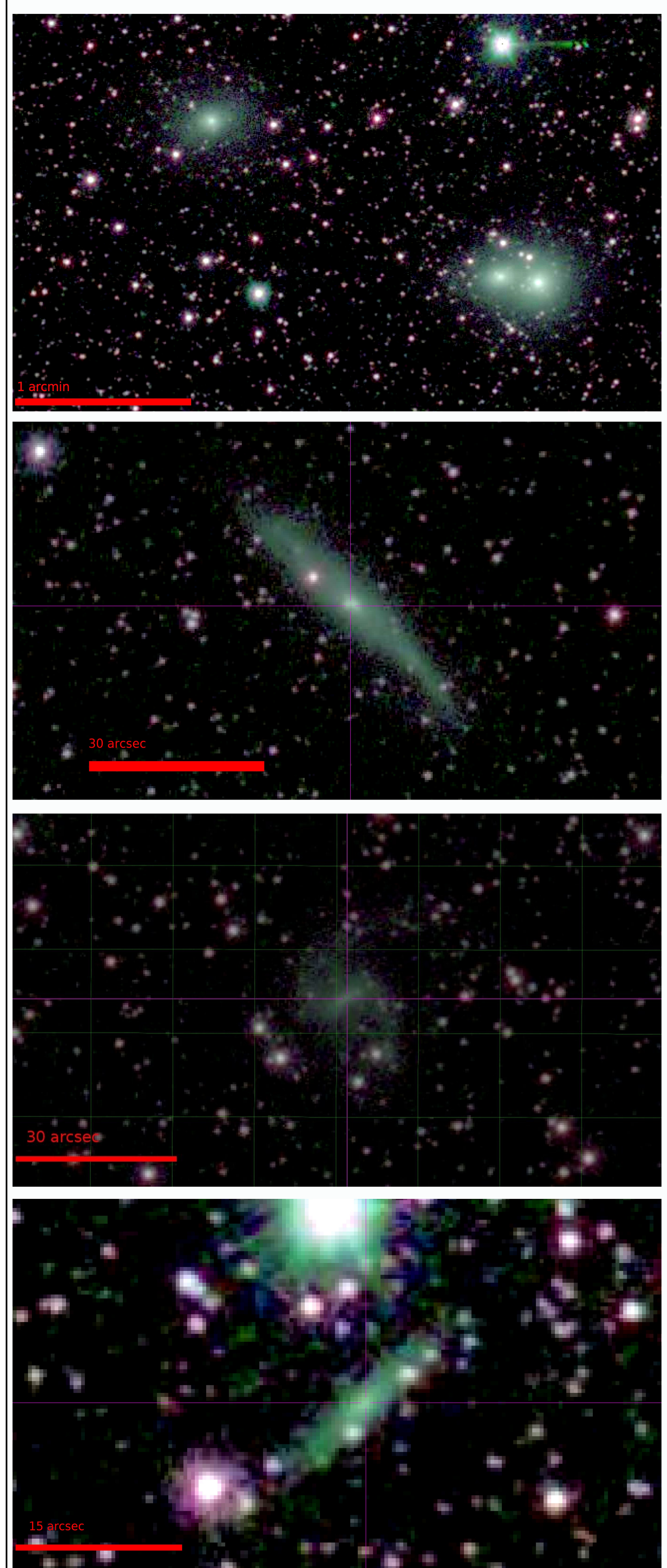

Figure 5: Example of VVV cutouts (or "fractions") each which contains at least one bright galaxy with angular scale as indicated in red.

Infrastructure: The infrastructure required for this project is shown in Figure 6. It consists of the Data Base which contains the images and meta data as described above, the phase 1 Application software which contains the interface and the Analyst who accesses the application and data through the web server that hosts both.

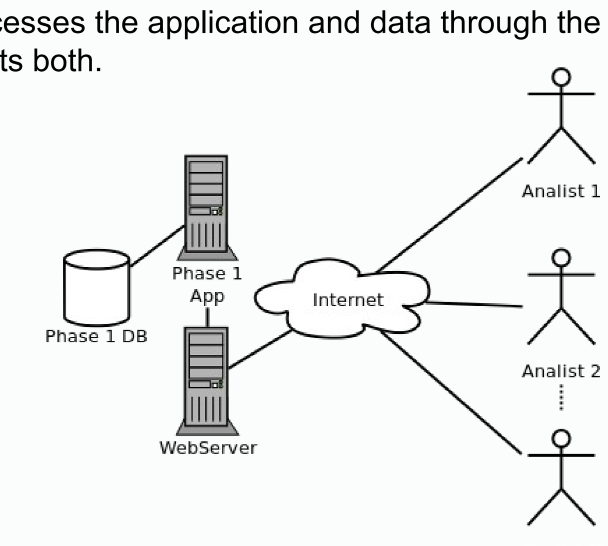

Figure 6: The above infrastructure is proposed for the deployment of the application.

Framework Modules: Below is the list of modules we are using for the development of our web framework.

Montage (http://montage.ipac.caltech.edu/) – *see Section 3*
Python (http://www.python.org/) - main application language
Astropy (http://www.astropy.org/) - Astronomical Library
Aplpy (http://aplpy.github.io/index.html) - Plotting Library
Django (https://www.djangoproject.com/) - Webframework
PostgreSql (http://www.postgresql.org/) - DBS
Apache Web Server (http://httpd.apache.org/) – application server
Nginx (http://nginx.org/) – to serve static archives

7 Interfacing galaxies with people!

Simple interface: A simple interface has been designed for our web-framework, as shown below in Figure 8 that includes:

1) Selection area tool
2) Button to activate and deactivate the selected area
3) Button to ignore momentarily this "fraction"
4) Button to ignore permanently this "fraction"
5) Toolbar that contains the history of classifications

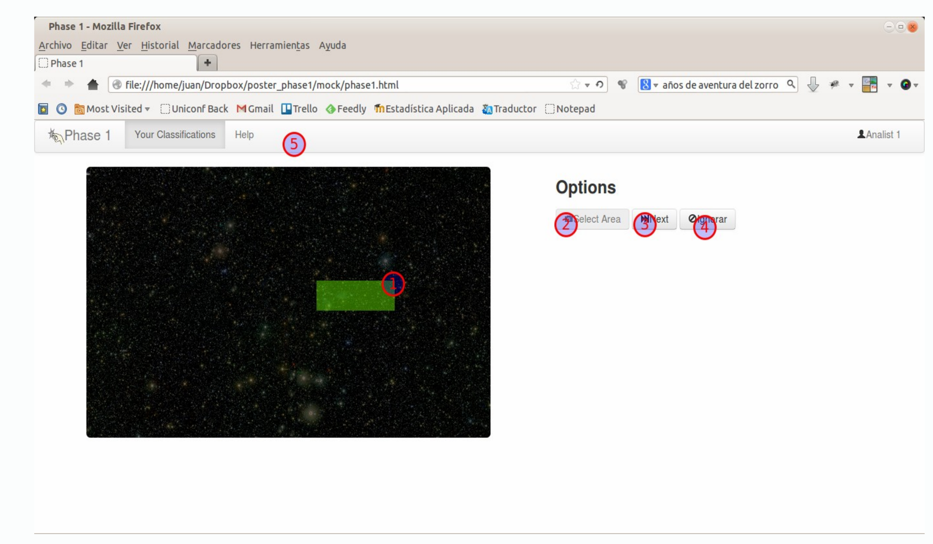

Figure 8: Interface for the deployment of Phase 1 i)

8 Interfacing SNae with people!

Re-projecting for SNae: After the results of phase 1 i) are internally verified then Phase 1 ii) will commence. The DB will be queried for positions, tile numbers, obs-mjd, etc for the galaxy 'fractions' in the multiple epoch frames that will be re-projected as per Section 3, with pixel intensity-scale optimized against changes in the background. The interface will be adapted to allow the inclusion of movies made for all galaxy 'fractions' in different epochs. Analyzers will be able to click on the frame and any noticeable transient in similar fashion as described in Section 7.

9 Finding Supernovae

Light curves and science: When a supernova is detected, the CASU 1.3 pawprint and tile catalogues will be queried at the detected position/epoch and the detection manually verified. We will construct basic light curves from the data at all epochs and will perform time series analysis. Our analysis will be used to estimate the type of SNae, its distance, stellar metallicity from the galaxy mass-metallicity relationship, galaxy stellar mass component and any SNae rate dependence with general galaxy photometric properties as derived by surface photometric analysis of this galaxy and stellar population synthesis methods (Bruzual and Charlote 2003). We also hope to correlate SNae rates with other macroscopic properties.

10 Conclusions

We are setting up community science applications with data from the Vista Variables in the Via Lactea survey to study galaxies and supernovae. Given the SNae rate is ~50 events per year for typical galaxies, that the rate of the light curve decay is a few hundred days, and that Amôres et al. (2012) predicted as much as 30000 galaxies in the VVV, we might reasonably expect to find of the order of 100 supernovae across the whole survey. Results of our Phase 1 study will be verified at all stages but essentially our effort will be used to learn about supernovae and any statistical relation between macroscopic galaxy parameters.

Acknowledgements:

We gratefully acknowledge use of data from the ESO Public Survey programmeID 179.B-2002 taken with the VISTA telescope, data products from the Cambridge Astronomical Survey Unit.

This research made use of Montage, funded by the National Aeronautics and Space Administration's Earth Science Technology Office, Computation Technologies Project, under Cooperative Agreement Number NCC5-626 between NASA and the California Institute of Technology. Montage is maintained by the NASA/IPAC Infrared Science Archive

References and collaborators

[1] Minniti, D. et al. 2010, NewA **15**, 433
[2] Emerson et al. 2004, SPIE, **5493** 401
[3] Hakobyan A. et al. 2011 Ap **54**, 301
[4] Sersic J.L, 1968 BAAA **6**, 41
[5] Amôres, E, B, et al. AJ, **144**, 127
[6] Bonnarel, F. et al 2000 A&AS **143**, 3
[7] Bruzual, G & Charlot, S, MRAS **344**, 100
[8] Jacob et al. IJCS&E 2009, **4**, 73
[9] M. J. Graham et al., astro-ph.IM (1206.4035 in press)