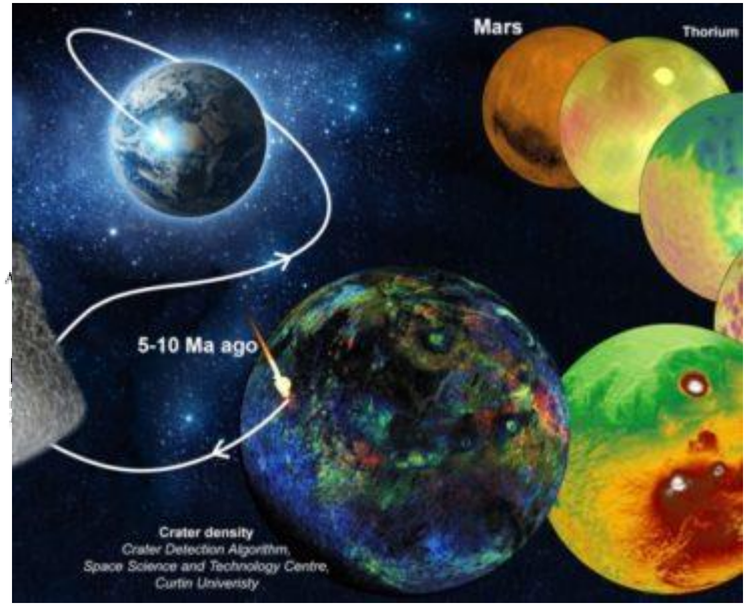
Who threw that rock?

Tracing the path of Martian meteorites back to the crater of origin using ML



K. Servis-Nussbaum

A. Lagain

G. Benedix

J. Fairweather

Scientific objectives of Crater Detection Algorithm

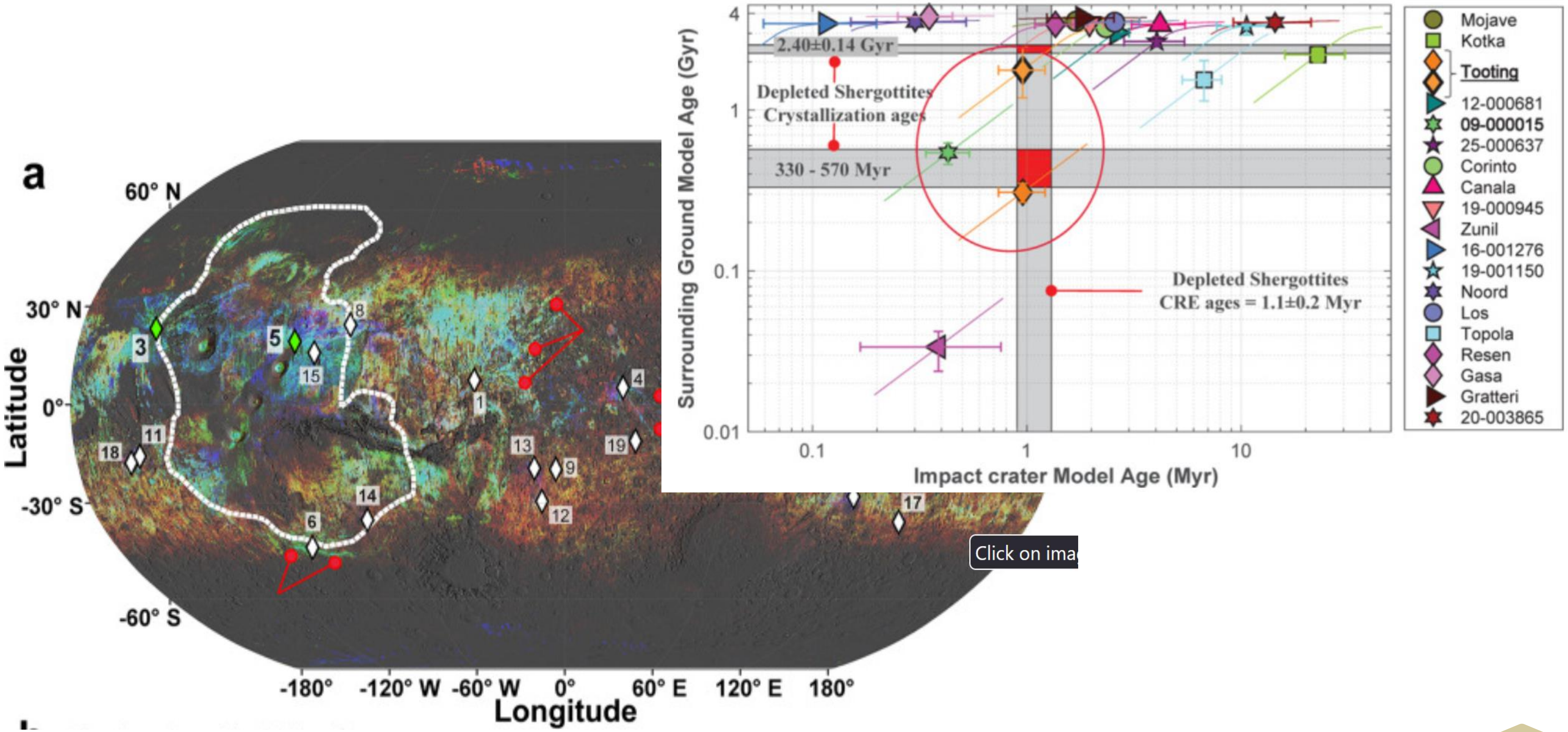
Why we do crater counting?

- Estimating the age of planetary surfaces (the more craters, the older the surface) (Fassett 2016; Hartmann 2005).
    - This can provide insights into geological processes of planetary bodies and beyond (late heavy bombardment, recent volcanicity and past presence of liquid water on Mars…).
- Secondary carter mapping is key to identifying the ejection site of Martian meteorites (Lagain et al. 2021):
    - A large enough impact that can cause material to be ejected with enough velocity to escape Mars would also cause a ray pattern of secondaries (craters formed by failed meteorites). By examining the patterns of small craters we can identify candidate primary craters.
    - In conjunction with other constraints (Cosmic Ray Exposure age and crystallisation age), we can narrow down the ejection site candidates to one-two for some samples, enabling us to improve our understanding of Mars and its differences to the Earth.
- Similarly, we can learn more about the evolution of other targets, such as Mercury and Ceres.

Discovering details about Mars's turbulent past

Analysis

- By combining the Mars crater database and resulting crater density map, information about TOF and composition we can identify the source of a major class of Martian meteorites.
- Images/Ref: Lagain, A., Benedix, G. K., Servis, K., Baratoux, D., Doucet, L. S., Rajšic, A., ... & Miljković, K. (2021). The Tharsis mantle source of depleted shergottites revealed by 90 million impact craters. *Nature Communications, 12*(1), 6352.

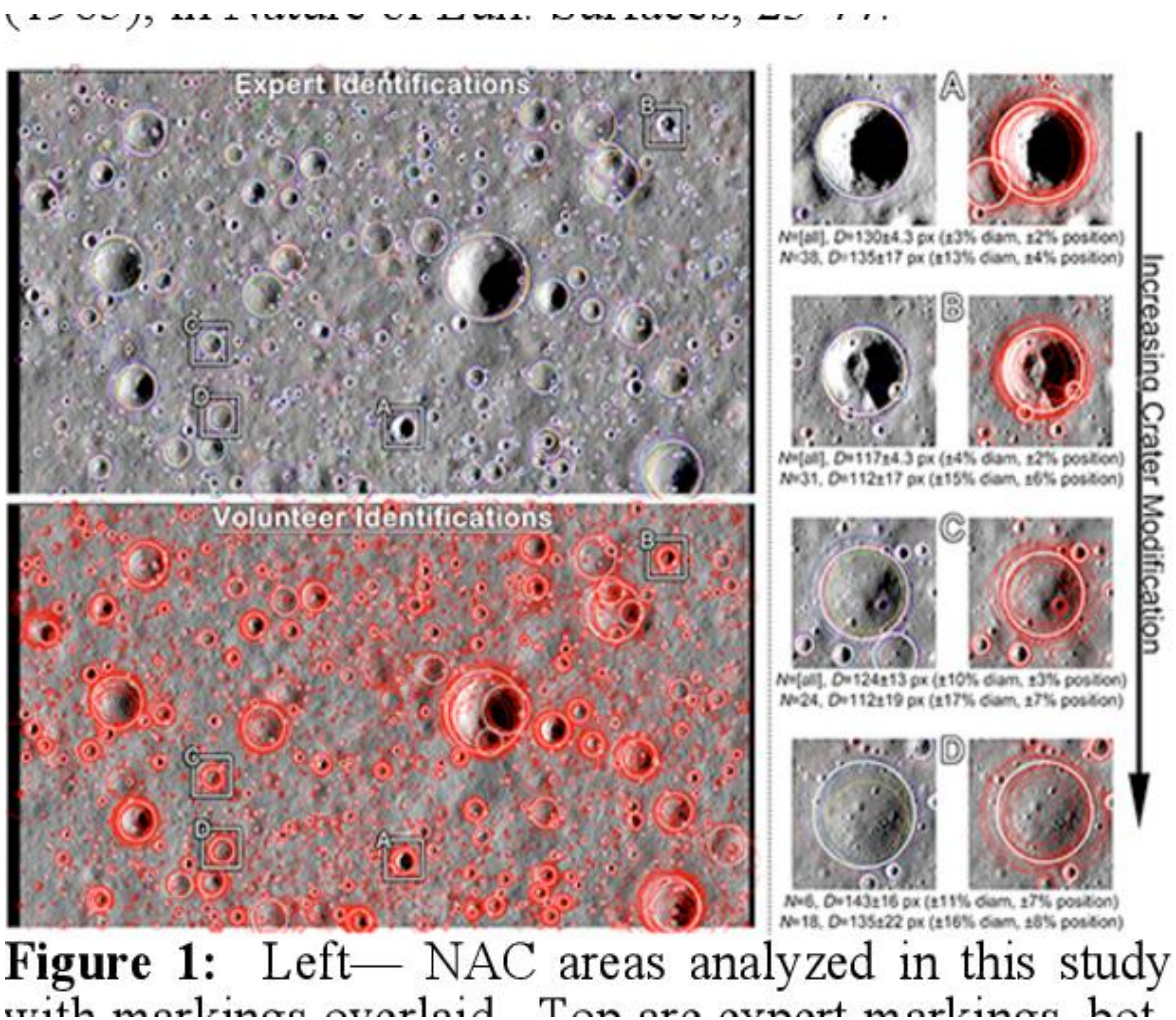


https://hive.curtin.edu.au/research/CDA-94M-release/

Challenges for CDA

Evaluation challenges

- Owing to crater degradation and depending on pixel size manually counted craters from imagery by experts may produce different results (Robbins et al. 2014) e.g.:
    - At least 20% difference in the number of craters amongst experts in the above study, and 30% of difference in the crater size.
- These results indicate that validation and training can be challenging but an ML based approach can be useful but we need to keep in mind:
    - Traditional measures of precision and recall will be low and :
    - The final result will need to be validated independently using an independently labelled dataset.



N=[all], D=130±4.3 px (±3% diam, ±2% position)
N=38, D=135±17 px (±13% diam, ±4% position)

N=[all], D=117±4.3 px (±4% diam, ±2% position)
N=31, D=112±17 px (±15% diam, ±6% position)

N=[all], D=124±13 px (±10% diam, ±3% position)
N=24, D=112±19 px (±17% diam, ±7% position)

N=6, D=143±16 px (±11% diam, ±7% position)
N=18, D=135±22 px (±16% diam, ±8% position)

Figure 1: Left— NAC areas analyzed in this study

Challenges for CDA

Fast labelling

- Experts would ordinarily take a few seconds to mark a crater but in order to reach the desired scale using automated methods, billions of detections are necessary.
- Given the variable results between experts, speed of inference is of primary concern.

Nextflow version of CDA

Previous versions of CDA:

- Monolithic container based version used with Themis data (Benedix et al. 2020)
- Shell script based version used for Jezero crater mapping (Servis et al. 2020)

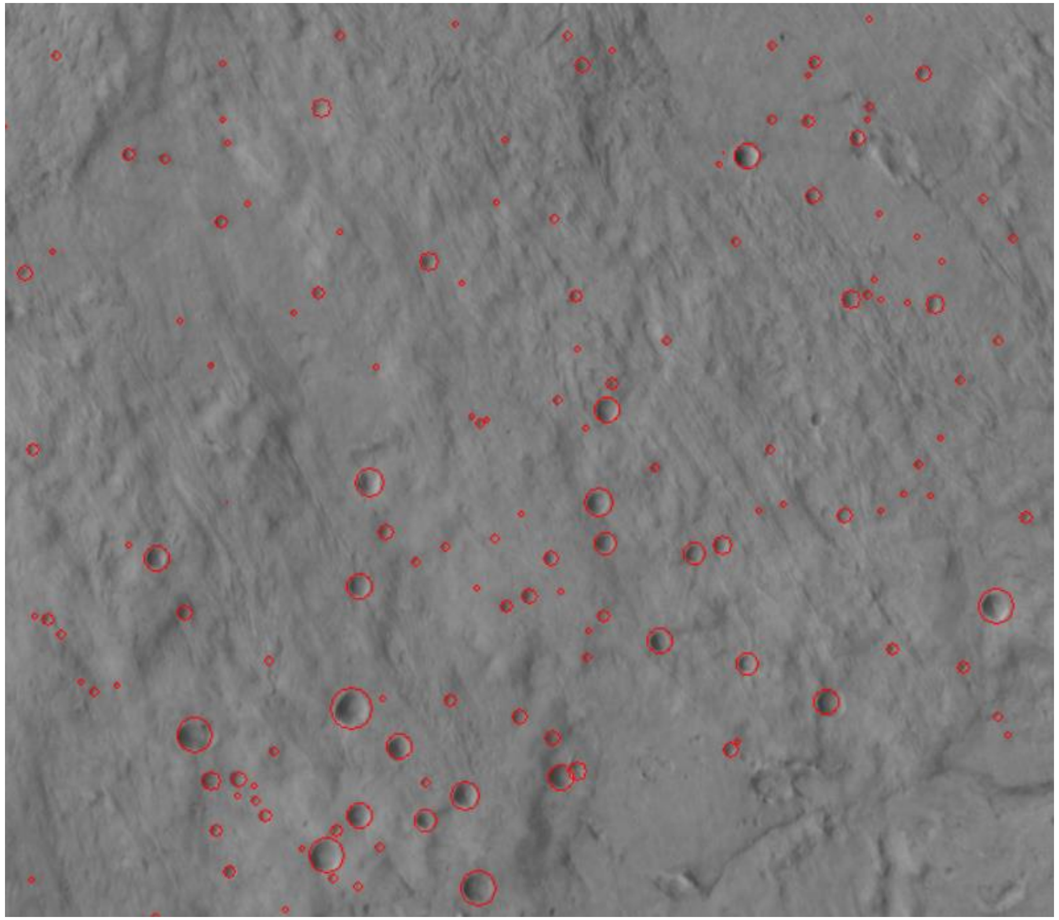
Advantages of nextflow version:

- Modularity
- Reusability
- Configuration injection
- Readability
- More target platforms
- Visualisation
- Log and trace

Workflows of CDA

There are several high-level tasks that are coded as workflows:

- Georeferencing: Taking a raw image from an orbiter and producing a georeferenced tile.
- Training: Using labelled tiles to train the algorithm
- Inference: Taking a trained network and producing a crater database.
- Annotation: Taking an image and a crater database and producing an annotated image.



ML workhorse (YOLO)

Main characteristics of the architecture (Redmon et al. 2016):

- Uses Convolutional Neural Network (CNN) blocks.
- Object detection as regression, instead of classification:
  - The network gives bounding box coordinates and class confidence (in this case we use only one class i.e. "crater")
  - Internally there are a number of anchor points in the pixel coordinates and an offset and size are produced as regression, along side the class confidence.
- Intended inference performance is real-time applications, such as self-driving cars, making it suitable for potentially millions of images (tiles in this case) that need to be evaluated.

Training YOLO (Mars case) (Lagain et al. 2021b)

- Training set from High-Resolution Imaging Science Experiment (25cm/px) (McEwen et al. 2023) on Mars Reconnaissance Orbiter :
  - On such a small scale there are many non-crater features that we can train the network to avoid.
  - Used the Jezero crater site (E77-5-N18-0) where 2142 craters were manually marked of which 550 were held out for validation.
  - Extensive use of augmentation to expand the training dataset (rotation, shear, scaling and translation).
  - For Mars YOLOv3 was used.
- Evaluation of the final results was by labelling the intended target data set from Context Camera also on the Mars Reconnaissance Orbiter:
  - Manually mapping 2000 craters on the CTX on different geological units.
  - CTX resolution is 6m/px so a 10px diameter corresponds to 60m which was the lower limit evaluated.
  - It was noted that mid- and high-latitude (>50 deg) performance is lower due to the higher degree of crater degradation and the presence of glacial features (e.g. geysers, mud volcanoes…).
  - Overall the F1 score was 0.75 for the evaluation using CTX bearing in mind that we set the intersection over union for considering some crater as the same at 0.3.

Inference execution of CDA for Mars

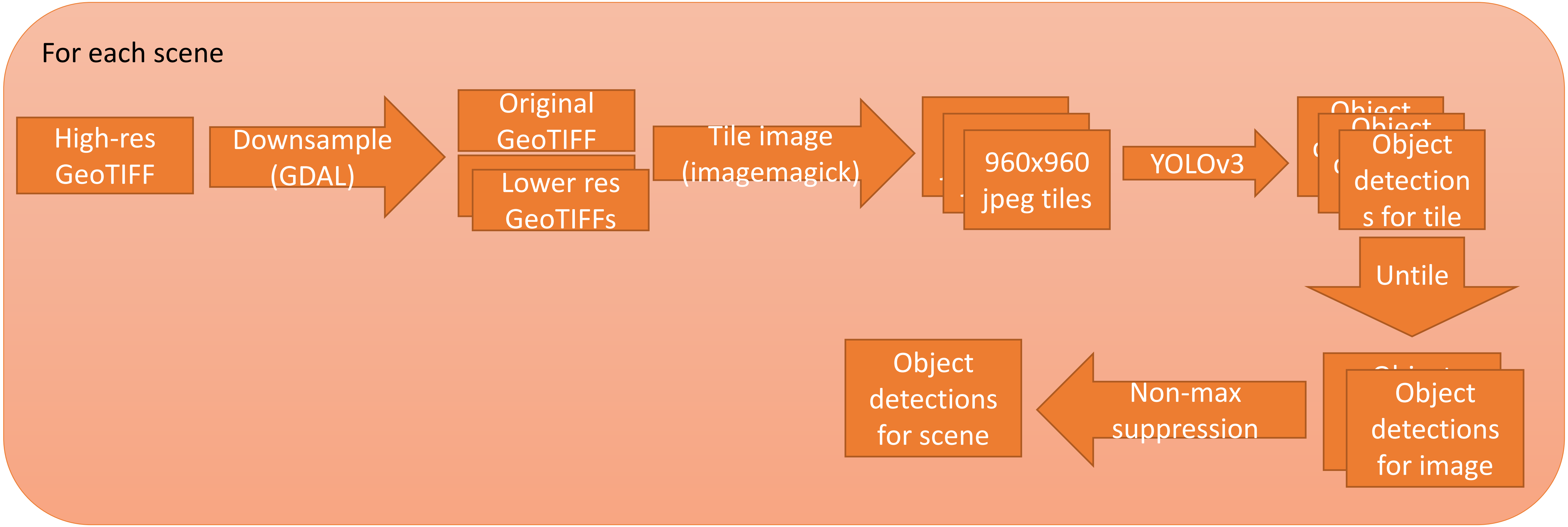
Covering the entire surface of Mars to 5m/px

- ~15000 original size images 5m/px from Murray Lab
- Downsample each to 40m/px and 160m/px = ~45000 images
- Tile each (8-2000 tiles) = a few tens of millions of tiles
- Mark each tile producing ~0-100 detections on each
- “Untile” each group of detections
- Run Non-max suppression on groups of nine adjacent scenes (target plus surrounds)

Additional considerations

- Executions fail but we don’t want to start from scratch after a fix.
- Execution needs to be done in groups, otherwise there would be too many jobs on the cluster, but groups need to be identifiable for debugging and individually repeatable.
- Dev/Test on local docker but deployment on slurm/singularity

Inference workflow of CDA



For each scene

High-res GeoTIFF

Downsample (GDAL)

Original GeoTIFF

Lower res GeoTIFFs

Tile image (imagemagick)

960x960 jpeg tiles

YOLOv3

Object detections for tile

Untile

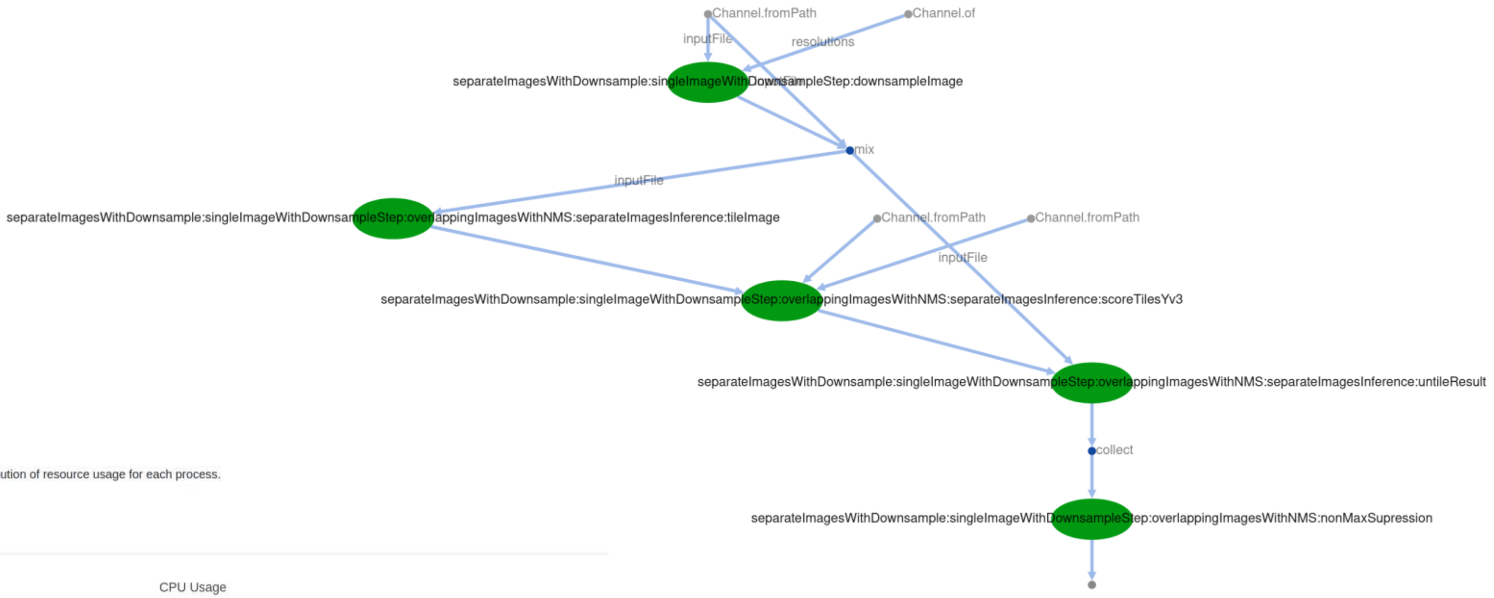
Object detections for image

Non-max suppression

Object detections for scene

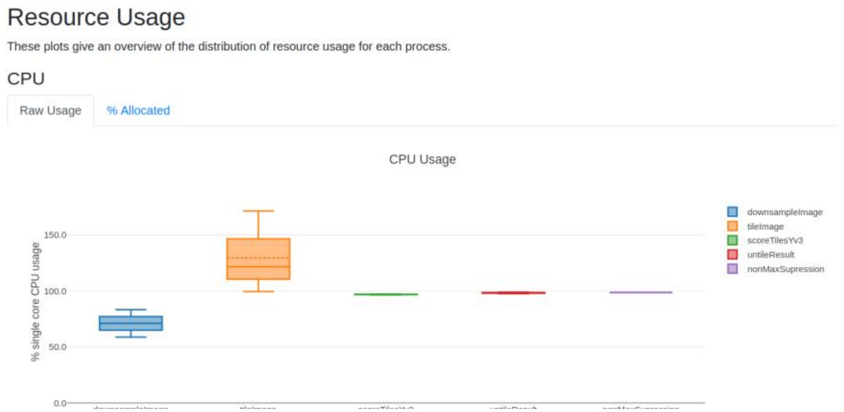
- Downsampling is needed in order to detect craters that are not visible due to being too large and only a small portion of them being on the highest resolution .

Nextflow version of CDA



Execution graph

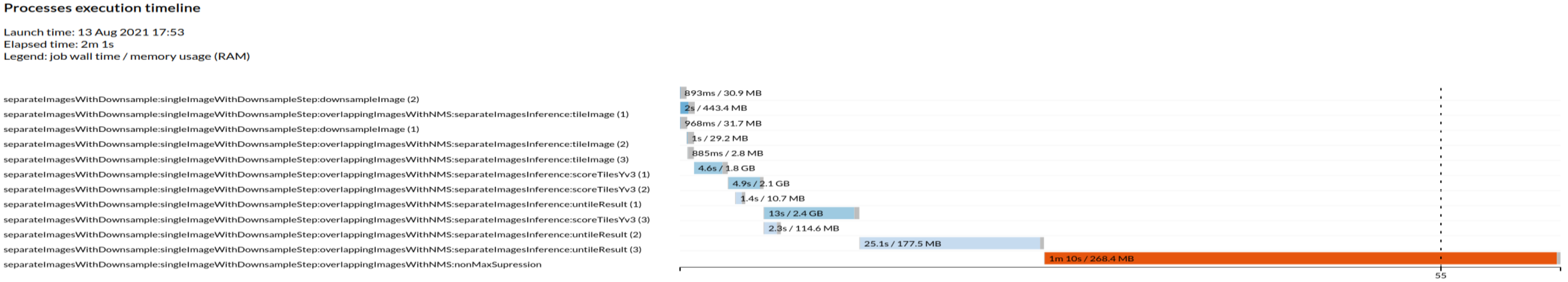
- Execution graph
- Report
- Timeline



Resource Usage

These plots give an overview of the distribution of resource usage for each process.

CPU



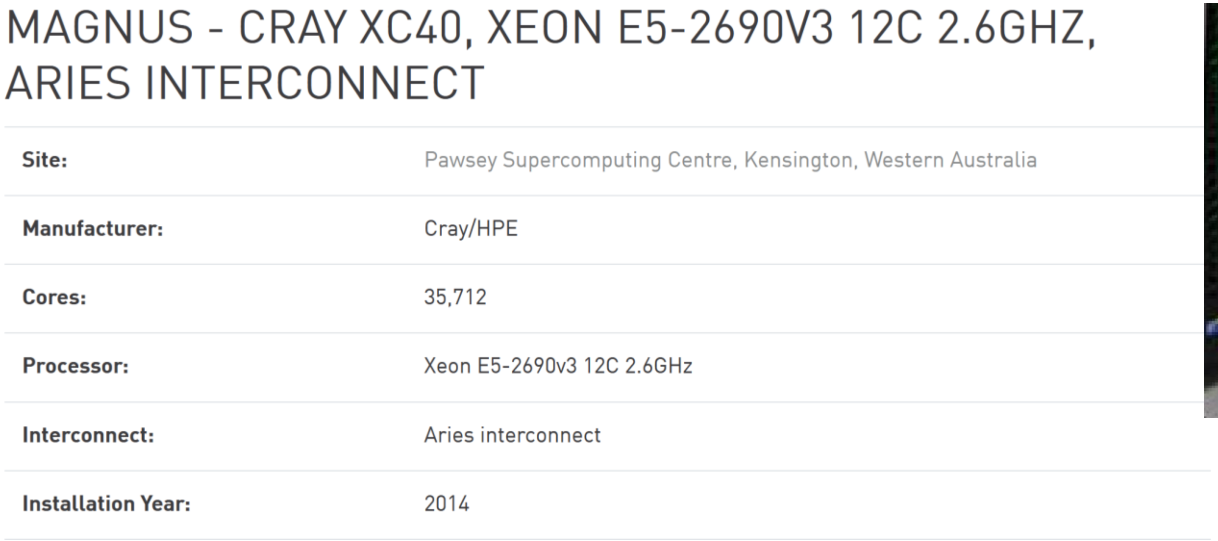
Processes execution timeline

Launch time: 13 Aug 2021 17:53
Elapsed time: 2m 1s
Legend: job wall time / memory usage (RAM)

separateImagesWithDownsample:singleImageWithDownsampleStep:downsampleImage (2)	893ms / 30.9 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:tileImage (1)	2s / 443.4 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:downsampleImage (1)	968ms / 31.7 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:tileImage (2)	1s / 29.2 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:tileImage (3)	885ms / 2.8 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:scoreTilesYv3 (1)	4.6s / 1.8 GB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:scoreTilesYv3 (2)	4.9s / 2.1 GB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:untileResult (1)	1.4s / 10.7 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:scoreTilesYv3 (3)	13s / 2.4 GB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:untileResult (2)	2.3s / 114.6 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:separateImagesInference:untileResult (3)	25.1s / 177.5 MB
separateImagesWithDownsample:singleImageWithDownsampleStep:overlappingImagesWithNMS:nonMaxSupression	1m 10s / 268.4 MB

Magnus/Zeus (now decommissioned)

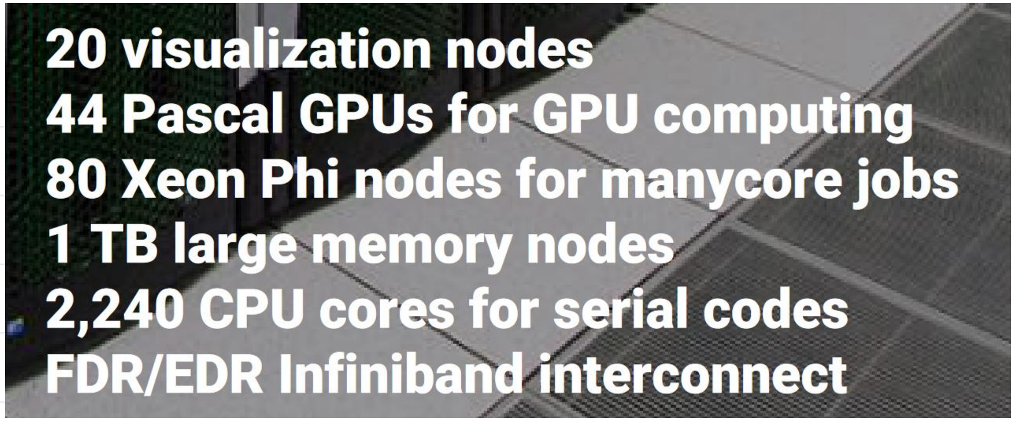
Magnus



MAGNUS - CRAY XC40, XEON E5-2690V3 12C 2.6GHZ, ARIES INTERCONNECT

Site:	Pawsey Supercomputing Centre, Kensington, Western Australia
Manufacturer:	Cray/HPE
Cores:	35,712
Processor:	Xeon E5-2690v3 12C 2.6GHz
Interconnect:	Aries interconnect
Installation Year:	2014

Zeus

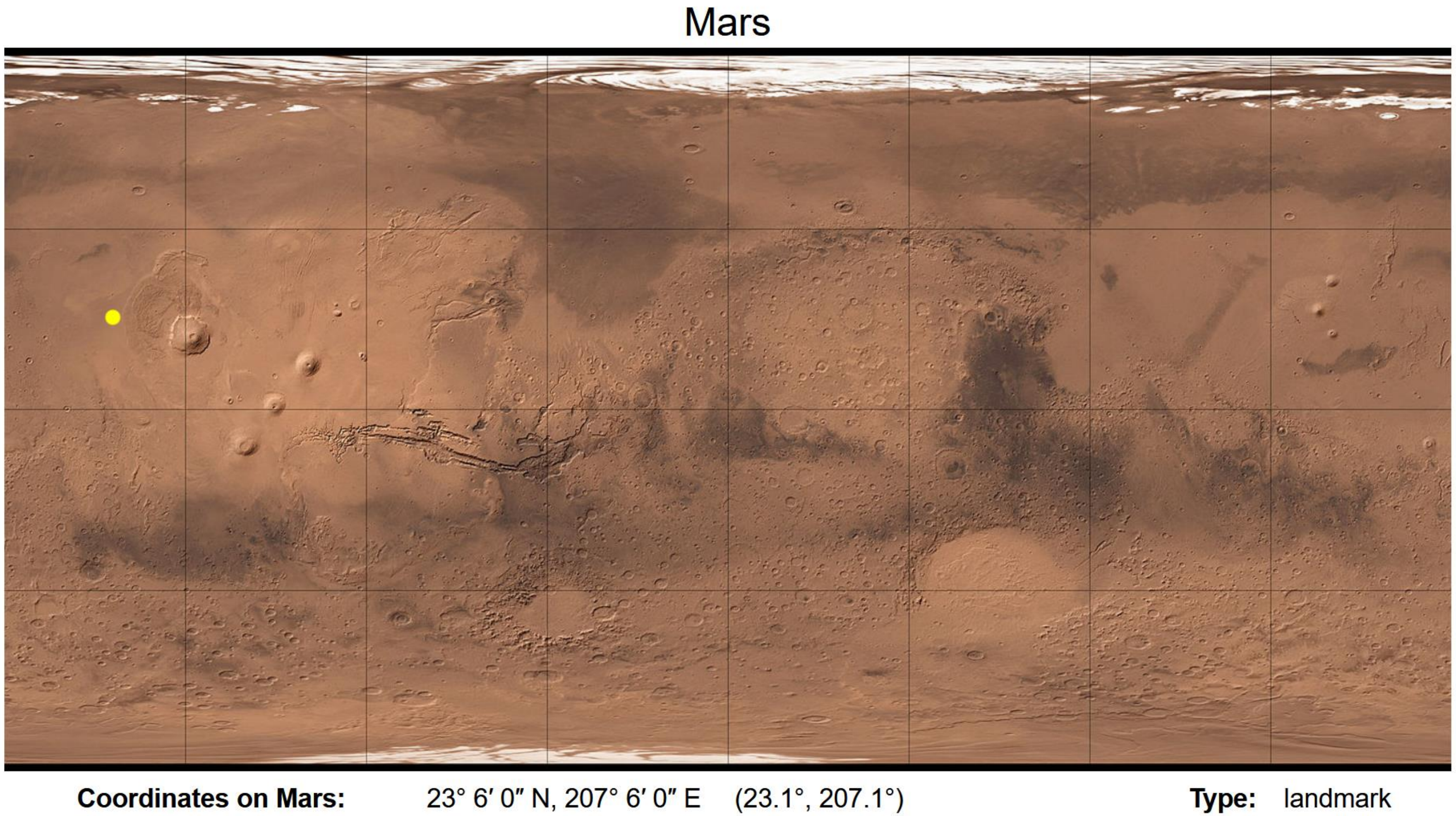


20 visualization nodes
44 Pascal GPUs for GPU computing
80 Xeon Phi nodes for manycore jobs
1 TB large memory nodes
2,240 CPU cores for serial codes
FDR/EDR Infiniband interconnect

You dropped something!

Long story short

- Here is where we believe the depleted shergottites most likely came from.



Coordinates on Mars: 23° 6′ 0″ N, 207° 6′ 0″ E (23.1°, 207.1°) **Type:** landmark

Discovering details about Mars's turbulent past

Significance

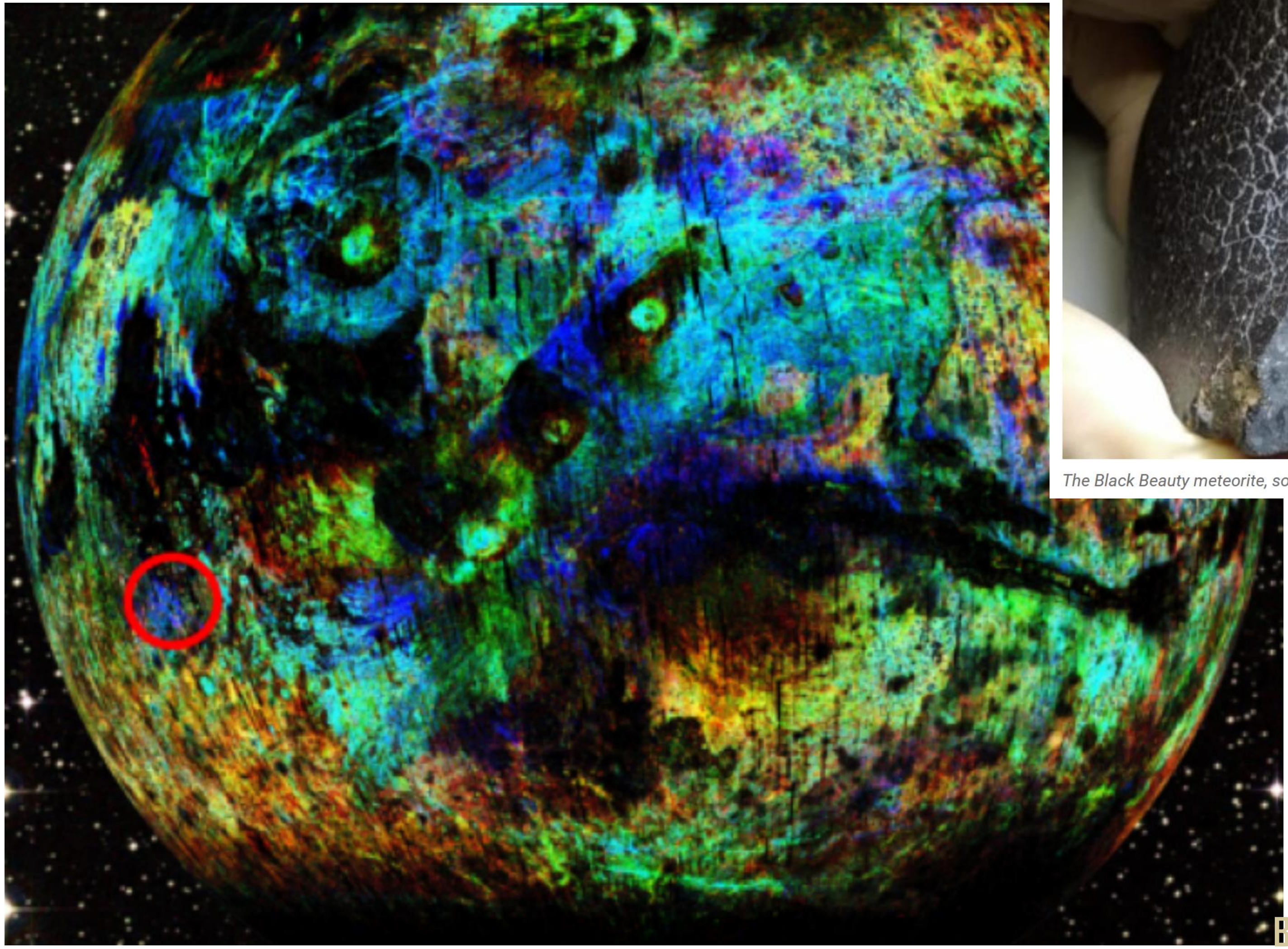
- The finding implies a major thermal anomaly (a plume), that has been active in the Tharsis region throughout the history of Mars.
- This is likely similar to the process underlying volcanism in Hawaii, but:
- Since Mars has no tectonic plates that plume of magma has been rising for billions of years undisturbed and caused the Tharsis region to form and grow.
- This activity stopped 340Ma ago (later activity may have existed but are not recorded in the group of meteorites considered here)

Discovering details about Mars's turbulent past

Another stray rock!

- Using our database and TOF analysis and some newer simulations we were also able to identify the source of the Black Beauty meteorite as the Karratha crater.
- Ref: Lagain, A., Bouley, S., Zanda, B., Miljković, K., Rajšić, A., Baratoux, D., ... & Bland, P. A. (2022). Early crustal processes revealed by the ejection site of the oldest martian meteorite. *Nature Communications*, *13*(1), 3782.

Karratha Crater on Mars, source NASA MRO.



The Black Beauty meteorite, source NASA.

Further targets

We have already used this workflow on other rocky bodies to calibrate model ages and understand whether the overall flux has changed

- Moon (Fairweather et al. 2022, 2023)
- Multiple bodies (Earth, Mars, Moon) (Lagain et al. 2022)
- Others pending publication

Thank you

- **Benedix, G. K., Lagain, A., Chai, K., et al. 2020, 2007**
- **Fairweather, J. H., Lagain, A., Servis, K., et al. 2022, Earth and Space Science, 9, e2021EA002177**
- **Fairweather, J. H., Lagain, A., Servis, K., & Benedix, G. K. 2023, Earth and Space Science, 10, e2023EA002865**
- **Lagain, A., Benedix, G. K., Servis, K., et al. 2021a, Nat Commun, 12, 6352**
- **Lagain, A., Benedix, G., Servis, K., et al. 2021b, https://meetingorganizer.copernicus.org/EPSC2021/EPSC2021-12.html**
- **Lagain, A., Bouley, S., Zanda, B., et al. 2022a, Nat Commun, 13 (Nature Publishing Group), 3782**
- **Lagain, A., Kreslavsky, M., Baratoux, D., et al. 2022b, Earth and Planetary Science Letters, 579, 117362**
- **Servis, K., Lagain, A., Benedix, G., et al. 2020, https://meetingorganizer.copernicus.org/EGU2020/EGU2020-6269.html**

- Fassett, C. I. 2016, Journal of Geophysical Research: Planets, 121, 1900
- Hartmann, W. K. 2005, Icarus, 174, 294
- McEwen, A. S., Byrne, S., Hansen, C., et al. 2023, Icarus, 115795
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. 2016, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA: IEEE), 779
- Robbins, S. J., Antonenko, I., Kirchoff, M. R., et al. 2014, Icarus, 234, 109x