NOVELTY OR ANOMALY HUNTER - DRIVING NEXT-GENERATION SCIENCE AUTONOMY WITH LARGE HIGH-QUALITY DATASET COLLECTION

Nick Read (1), Mark Woods (1)
Spyros Karachalios (1)

(1) SCISYS, 23 Clothier Road, Bristol, BS4 5SS, UK, nicholas.read @scisys.co.uk
(1) SCISYS, 23 Clothier Road, Bristol, BS4 5SS, UK, mark.woods @scisys.co.uk
(1) SCISYS, 23 Clothier Road, Bristol, BS4 5SS, UK, spyros.karachalios @scisys.co.uk

ABSTRACT

The ESA Noah Or Novelty Hunter (NOAH) project follows on from the prior MASTER [1] project in working towards a robust autonomous science capability for European robotic exploration missions, and ground based data processing. Where MASTER was directed toward both rover and Earth Observation (EO) imagery, NOAH is focussed on Martian Rover Navcam images. As well as algorithmic improvement, a major component of the work is to curate a large annotated Martian image dataset. This dataset will drive the innovation into novel machine learning approaches, enable rigorous evaluation and is the focus of this paper. We review the drive behind this activity, details of the implementation and annotation approach and summarise the wider NOAH project to come.

1. INTRODUCTION

The observation of landforms, outcrops and small features within a (Martian) landscape is key to the understanding of its geologic past as well as present environmental conditions. Studies of such features have – for example – revealed the nature of streambeds at Gale Crater [2], and allowed to study Aeolian bedforms as they were encountered by the Curiosity [3]–[5], and Spirit rovers [6], [7]. With two active rovers (Opportunity, Curiosity) currently on Mars, and two more to be launched in 2020 (ExoMars, Mars2020), the imaging datasets are a huge, growing resource, which could potentially,

Bandwidth and the Earth-Mars distance limitations rule-out real-time control of rover navigation systems, payload instruments and this combined with a prolonged data relay complicates the prospect of achieving serendipitous scientific discovery. For planetary rovers there is a trade-off between detailed observation to ensure important targets are not missed, which requires slow traverses to downlink all the data, and maintaining sufficient progress to visit many science targets. In the current operational paradigm, we rely on a posteriori analysis of low resolution images and thumbnails to ensure that important targets are not missed as the rovers have limited or no ability to detect salient features in the terrain. Examples from actual missions such Opportunity's discovery of the Block Island meteorite and operational trials [8] have shown that this constraint does raise the possibility of missing science targets completely despite being in relatively close proximity to the rover path.

One solution to this problem is to equip rovers with some level of science autonomy to support the detection of potentially interesting science targets. The ESA NOAH project follows on from previous work such as MASTER [1] and CREST [15], has three key objectives:

- Develop a generic web-based dataset annotation tool.
- Create a large crowd-sourced Martian rover image dataset.
- Improve on the previous ESA MASTER [1] project to advance the TRL of planetary science autonomy, with improved algorithms and more flight-like implementation.

LabelMars (www.labelmars.net) is a citizen science activity to collect geological annotations of Martian rover navigation camera images. As part of the ESA NOAH project it's objective is to provide a large, high quality dataset to develop state-of-the-art machine vision algorithms for autonomous science detection, targeted at future rover missions. Large well-annotated datasets are of exceptional value to the computer vision community, and in this case the Mars exploration community. The goal of the LabelMars activity is to achieve five thousand annotated images – this will make it the largest geologically annotated Martian rover imagery database for vision research and development. In the wider image processing community there have been huge advances kick started by the availability of large datasets – most famously with ImageNet [9]. Of course, we do not aim for the scale of ImageNet's million-plus images, the scale is smaller, but the ambition is grand within the restricted scope of autonomous planetary science.

As well as the activity to produce this large annotated dataset, the NOAH project also advances the TRL of our science autonomy solution. Through a program of algorithmic research and development, combined with implementation of a Prototype Flight Detector we will advance the autonomous scientist concept toward the goal of a flight experiment on ExoMars or subsequent rover missions.

1.1. Related Work

Annotated and pre-screened scientific datasets are of large benefit to the scientific community too, because they allow research on data bases which cannot be created within the financial and personnel base of a single project. There are numerous examples, mostly under the envelope of 'citizen science' which have demonstrated the value of such datasets. For a set of examples from space science, see [10]. Two examples of such overall classification have returned especially prominent results. The 'Galaxy Zoo' project used classification of the currently available very large number of images for volunteer-based morphological classification of galaxies [e.g.[11]; this project even led to the identification of new objects (with the volunteers being acknowledged in the publications [12]). In a similar way, the 'stardust@home' project asks volunteers to find the tiny tracks particles made in the aerogel of the Stardust mission [13], which by 2016 had resulted in the identification of over 200 impacts in the aerogel tiles. This was made possible by over 30,000 volunteers, who carried out more than 108 searches [14], enabling the science that follows.

For rover based images of the Martian surface, labelled images would allow – similar to the Galaxy Zoo project – the identification of similar textures and objects at different landing sites and use statistical methods for comparison – as well as enabling topical work across the different current and future landing sites.

By adding autonomous capabilities to detect novel or scientifically interesting phenomena in images we can enhance the scientific return of robotic exploration missions. The many images acquired for navigation during traverse operations that are normally discarded can be inspected, and those scoring highly selected for downlink or even further sensing of the target. Of course more powerful rover AI and autonomy will not replace ground-based scientists, but it offers the potential to lower the probability of missing proximal science targets. This has been demonstrated in practice in NASA's AEGIS [15] system deployed on MER and MSL, and in prior ESA and UKSA funded work investigating new methods in science autonomy such as the NOAH precursor MASTER [1] and CREST [16] projects.

2. DATASET COLLECTION

Registered users will be presented with Navcam images to annotate, which they will do with a combination of drawing polygons to indicate areas of interest and confirming, or labelling polygons drawn by other users. The Martian rover imagery is ideal to provide a narrative appeal combined with a visual indicator of progress by offering participants images from sols in sequence (or within a sol) with the idea that they are "seeing where the rover goes next". Participants can then visualize their progress by seeing pins on a map of Mars. This also provides a global progress indicator to show all users the overall annotation progress. Additionally, there will be a financial prize awarded for the completion of high quality annotations, as defined by agreement between labels produced by multiple participants.

2.1. System Design

The principal software component of the image annotation effort is known as the Dataset Annotation Tool (DAT). This is based on the Oxford University Zooniverse codebase [17]. Zooniverse is the largest and most popular citizen-science platform on the internet, so was a natural choice to base our efforts on. Some existing Zooniverse-based projects also focus on Martian imagery – namely the set of Planet Four [18] labelling efforts that are focused on orbital HiRISE imagery. For LabelMars, we learn from previous Zooniverse projects [19] which found amongst other things that a narrative appeal can lend an element of gamification and increase engagement. Following the traverse of a rover across the Martian surface provides a natural narrative incentive in this case, and information as to the context of the images increases the quality of interpretation for labelling. Zooniverse is based upon a Ruby-on-Rails server-side application called Panoptes. This provides an OAuth secured API that can be consumed by a compliant client application. Zooniverse.org uses a Javascript client called Panoptes-Front-End. Both components are customised and are used as basis for the LabelMars implementation.

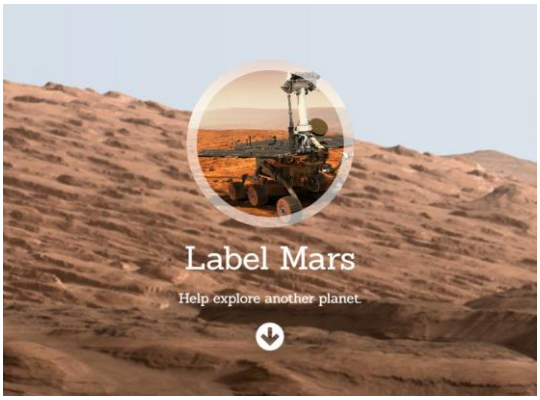


Figure 1 - Label Mars splash screen

The LabelMars implementation provides the ability for users to create crowd sourced projects where a large participant user-base can be utilised to classify datasets of varying size.

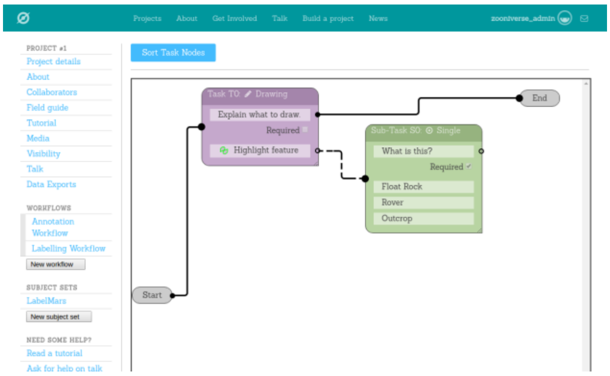


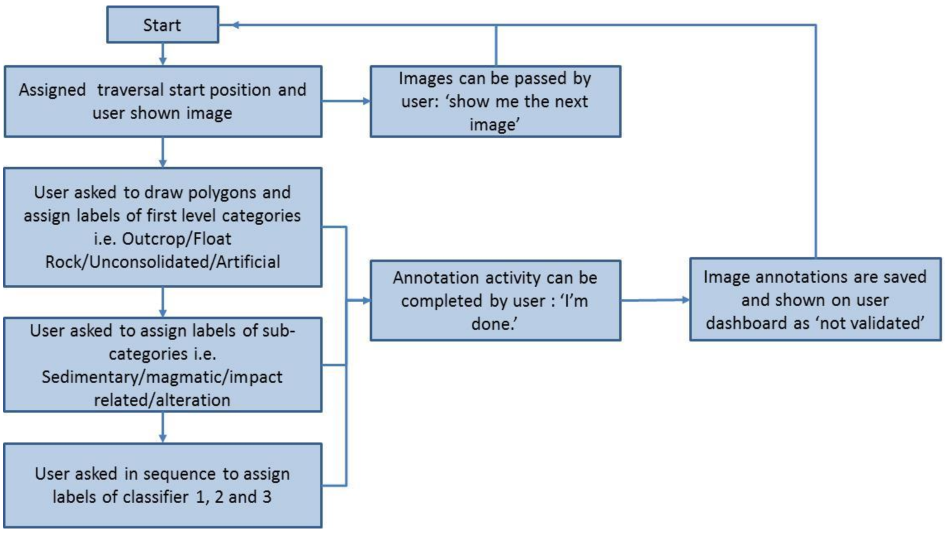
Figure 2 - Creation of a workflow

Administrative users, who control the setup of a labelling activity, have the ability to create custom

Figure 3 - Annotation Workflow

workflows that are presented to participants that allow them to highlight the desired features for a dataset that can then be exported by the project administrator. Figure 2 shows the visual tool that is used to assemble a user workflow. By providing these features the tool is adaptable to a wide range of image types and annotation tasks. This ensures its suitability for re-use in other annotation tasks.

The Javascript front-end provides a set of customisable annotation tools as in Figure 4. This allows the project administrator to determine the type, shape and labels of annotations that participants can perform. There is a careful balance to strike between fast efficient



annotation and high-quality data.

Although Zooniverse's software provides a lot of functionality, a requirement of the project is to ensure high quality annotation data, therefore certain features need to be extended for the LabelMars project.

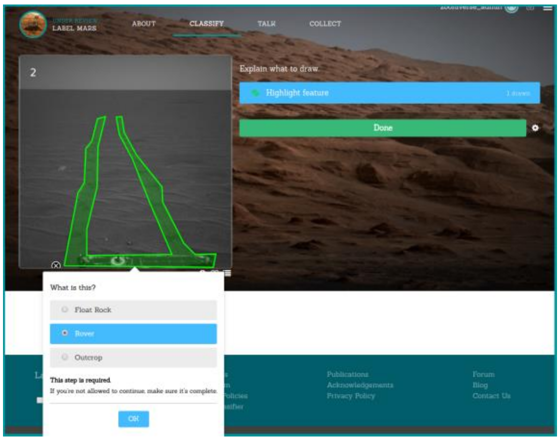


Figure 4 - Annotation tool

Participants are kept motivated by being presented with their usage statistics (such as 'number of annotations performed') and being credited in scientific publications. Motivation is an important part of any workflow as image annotation can be a tedious task. Despite this, with NOAH we have a huge advantage; the chance to ride along with a real rover on its track across Mars is a tantalising one.

A custom feature of Label Mars is the tailoring to obtain a large set of image segmentation data, with some images containing many segments. Participants are shown validated segments that they cannot change, so they have to segment other regions of the subject. This increases the granularity of the segmented image, as participants are encouraged to direct attention at unannotated areas. When a pre-annotated segment is highlighted, the participant is presented with a finite choice of classification options. Finally, any unlabelled areas are explicitly marked “background”. This workflow is outlined in Figure 3 - Annotation Workflow.

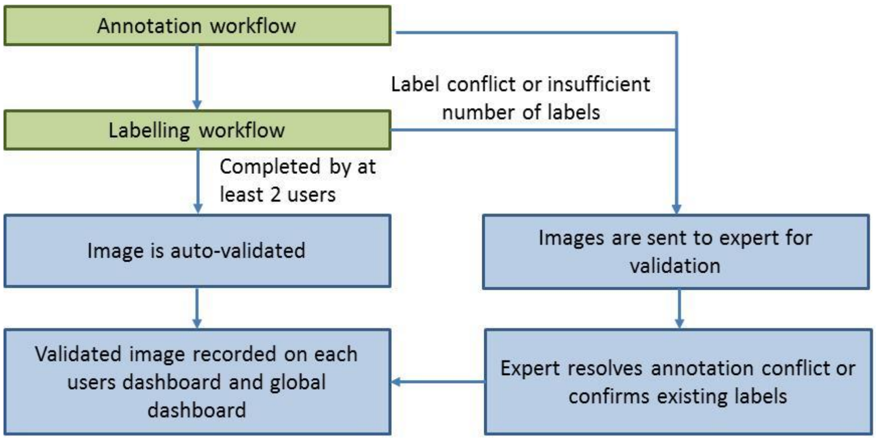


Figure 5 – Conflict Resolution Workflow

Another custom feature will allow for images to be auto-validated. When a number of participants have annotated similar regions Label Mars will flag this segment as ‘validated’. This segment can no longer be annotated and is highlighted when participants are presented the subject image. To validate the annotations, the intersection-over-union area measure is used to compare similarity. Any annotations that have conflicting annotations will be flagged for an expert user to validate. This workflow is shown in Figure 5

For the duration of the project the Label Mars software is deployed using Amazon Web Services (AWS) cloud infrastructure – see Figure 6. By utilising this infrastructure Label Mars benefits from quick content delivery and failsafe/recovery measures. The subject set is stored in an AWS S3 bucket and CloudFront cache to quickly deliver the image content to the participant. This reduces lag time between images. The API resides on two identical Elastic Compute (EC2) instances using an Elastic Load Balancer (ELB) to direct requests to them. Using a load balancer increases performance of the API as they are processing requests in parallel, it also increases reliability as API requests are processed across multiple servers instead of a potential single point of failure. The EC2 instances have access to a Relational Database Service (RDS) instance that provides persistent storage for segmentation data collection. The RDS instance has automated backups as a failsafe.

2.2. Annotation Campaign Execution

The main decisions running the annotation campaign centre on:

- The type of audience to attract;
- The class labelling scheme to use;
- The selection of images to label.

To some extent these are inter-related questions. If the labelling scheme requires detail Martian geological knowledge, then that naturally restricts the audience. The selection of images is driven by some of the project requirements, but also guided by geological advice and expertise.

The class labelling scheme was chosen to acquire an analysis based on a geological interpretation of the images. So, for example, it includes classes such as outcrop and float rock and within these attributes such as sedimentary or magmatic, coarse or fine grain and structured or homogeneous.
The class labelling used also defines our target audience. We target undergraduate educated geologists (or better) but do not require specific planetary geology expertise. This allows for a good balance of a reasonably large target audience and yet capturing a good level of useful information on the images.

The images are drawn from the Navigation Camera (NAVCAM) images from the Spirit, Opportunity and Curiosity rovers. This is to ensure the data are representative of the sorts of images that an on-board autonomous scientist agent could expect to have available for analysis. The complete set of full resolution images numbers nearly sixty thousand and we sample from these to reduce to a more manageable five thousand.

Figure 6- Label Mars deployment architecture

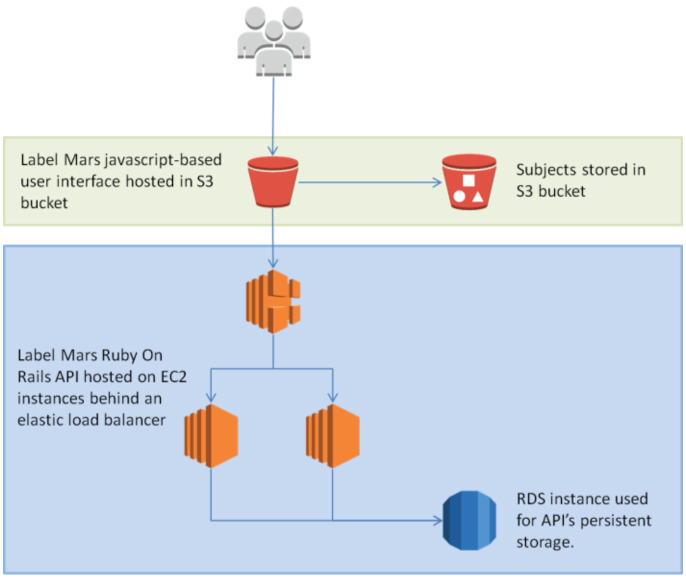
Images were selected so that they would include a wide sampling of the different terrains and phenomena of interest across all three rover missions, with a bias toward likely terrain to be encountered by the ExoMars Rover. Engineering images primarily of the rovers themselves, or images of the sky, were filtered out also using the embedded pointing-angle metadata. Despite this, many images still contain sections of the rover hardware – it is unrealistic to always expect it not to intrude and in the case of large platform like Curiosity which dexterous arms removing all images with hardware included would be very restrictive.

To motivate users the order in which the subject set is presented to a Label Mars participant is the same order as the Navcam images of a rover were captured. This order is preserved because this provides a natural narrative sequence: each subject is a step further in the rover's traversal.

3. PRELIMINARY RESULTS & DATASET EXPLOITATION

The initial labelling phase has been progressed with the competition being open for a period of several months. It is now progressing to a second more targeted phase. Preliminary results from the open labelling competition highlighted the complexity of the task in terms of individual image variability and volume. The competition attracted a range of dedicated labellers who spent a significant period of time on the various tasks. Survey based feedback has shown that that in general the competition and paid aspect where not important to the users. Contribution to the science base was the primary driver. We will continue to analyse this feedback and report on our findings once the labelling process is complete.

The immediate use for the dataset will be to advance the state of the art in autonomous science discovery. The original MASTER work had a far smaller dataset, which imposed restrictions in the performance of machine learning approaches and the rigour of evaluation. A larger dataset would improve the approaches used previously as well as opening up the potential of the new, state of the art data driven approaches to machine vision such as deep learning [20]–[22]. Deep convolutional neural networks show such promise as in several domains they have been shown to out-perform human efforts, and their utility for practical problems sees massive investment into new algorithms and implementations that can be exploited in the space domain. The engine driving this improvement is large



datasets, and whilst by modern standards a “mere” five-thousand images is not a large set, it represents a significant step in the right direction.

This dataset also has additional value for image processing and machine learning challenges. Not only is it Martian data with geological interpretation, but it is part of a larger unlabelled dataset that could be exploited and has rich metadata. Associated with each image is all the rich context in the configuration of the spacecraft, plus the implicit information in the sequencing of the Navcams, and a wealth of additional information could be associated such as the commands sent to the rovers or satellite imagery.

3.1. Future Work

The Dataset Annotation Tool is designed to be re-usable for a variety of annotation tasks – and this is proven with the flexibility of the core Zooniverse platform being adapted for this task. Future work has already been initiated using the tool to support labelling to label HiRISE imagery in order to support the development of an autonomous, semantic segmentation, based terrain analysis for site understanding and rover navigation. Perhaps the most natural extension of this work would be to increase the size of the annotated dataset – label a larger portion of the 60,000+ Navcam images or extend to Pancams also.

The immediate future work planned as part of the NOAH project is described in the previous section – this dataset will drive the development of more advanced, higher TRL science autonomy. This could of course be taken further. NOAH is directed toward on-board science autonomy as the ultimate goal, but there is also the use-case of ground support tools to process rover imagery and help direct human attention. It is easy to envisage some future integration of the DAT and the science autonomy, which could produce an online collaborative tool for scientists to explore and annotate Martian data products whilst also receiving suggestions from the science autonomy system. Such a system with the ability to understand the content of images would also enable tantalising content-based searches – imagine selecting an interesting rock in an image and being presented with other captured images of similar looking rocks.

4. SUMMARY

This paper has described the development of a Dataset Annotation Tool (DAT) which is part of the NOAH project. The tool is being used to label a large volume of Mars rover imagery to allow the evaluation of machine learning based science autonomy algorithms. This work has already shown to have application beyond the original NOAH activity.

5. REFERENCES

[1] I. Wallace and M. Woods, ‘MASTER: A Mobile Autonomous Scientist For Terrestrial and Extra-Terrestrial Research’, *13th Symp. Adv. Space Technol. Robot. Autom. ASTRA*, no. 3, 2015.

[2] R. M. E. Williams *et al.*, ‘Martian Fluvial Conglomerates at Gale Crater’, *Science*, vol. 340, no. 6136, pp. 1068–1072, 2013.

[3] M. G. A. Lapotre *et al.*, ‘Large wind ripples on Mars: A record of atmospheric evolution’, *Science*, vol. 353, no. 6294, pp. 55–58, 2016.

[4] M. Day and G. Kocurek, ‘Observations of an aeolian landscape: From surface to orbit in Gale Crater’, *Icarus*, vol. 280, pp. 37–71, 2016.

[5] R. E. Arvidson *et al.*, ‘Mars Science Laboratory Curiosity Rover Megaripple Crossings up to Sol 710 in Gale Crater’, *J. Field Robot.*, 2016.

[6] R. Greeley *et al.*, ‘Columbia Hills, Mars: Aeolian features seen from the ground and orbit’, *J. Geophys. Res. Planets*, vol. 113, no. E6, 2008.

[7] ‘Spirit at Gusev Crater: Plates’, *Science*, vol. 305, no. 5685, pp. 811–818, 2004.

[8] M. Woods and A. Shaw, ‘Simulating Remote Mars Rover Operations in the Atacama Desert for Future ESA Missions’, 2014.

[9] O. Russakovsky *et al.*, ‘ImageNet Large Scale Visual Recognition Challenge’, *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[10] B. J. H. Méndez, ‘SpaceScience@Home: Authentic Research Projects that Use Citizen Scientists’, in *EPO and a Changing World: Creating Linkages and Expanding Partnerships*, 2008, vol. 389, p. 219.

[11] C. Lintott *et al.*, ‘Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies’, *\mnras*, vol. 410, pp. 166–178, Jan. 2011.

[12] D. Clery, ‘Galaxy Zoo Volunteers Share Pain and Glory of Research’, *Science*, vol. 333, no. 6039, pp. 173–175, Jul. 2011.

[13] A. J. Westphal *et al.*, ‘Stardust@home: Virtual Microscope Validation and First Results’, in *37th Annual Lunar and Planetary Science Conference*, 2006, vol. 37.

[14] A. Westphal *et al.*, ‘A Massively Distributed Search for Impacts in Aluminum Foil on the Stardust Interstellar Collector’, in *Lunar and Planetary Science Conference*, 2016, vol. 47, p. 2275.

[15] T. A. Estlin *et al.*, ‘AEGIS Automated Science Targeting for the MER Opportunity Rover’, *ACM Trans. Intell. Syst. Technol. TIST*, vol. 3, no. 3, p. 50, 2012.

[16] M. Woods, A. Shaw, D. Barnes, D. Price, D. Long, and D. Pullan, ‘Autonomous science for an ExoMars Rover–like mission’, *J. Field Robot.*, vol.

26, no. 4, pp. 358–390, 2009.
[17] R. Simpson, K. R. Page, and D. De Roure, ‘Zooniverse: observing the world’s largest citizen science platform’, in *Proceedings of the 23rd international conference on world wide web*, 2014, pp. 1049–1054.
[18] Zooniverse Project, ‘Planet Four’, *Planet Four*. [Online]. Available: https://www.planetfour.org/. [Accessed: 19-Apr-2017].
[19] A. Eveleigh, C. Jennett, S. Lynn, and A. L. Cox, ‘“i want to be a captain! i want to be a captain!”: Gamification in the old weather citizen science project’, in *Proceedings of the first international conference on gameful design, research, and applications*, 2013, pp. 79–82.
[20] Y. Bengio, Y. LeCun, and others, ‘Scaling learning algorithms towards AI’, *Large-Scale Kernel Mach.*, vol. 34, no. 5, pp. 1–41, 2007.
[21] Y. Bengio and others, ‘Learning deep architectures for AI’, *Found. Trends® Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.
[22] Y. LeCun, Y. Bengio, and G. Hinton, ‘Deep learning’, *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

6. ADDITIONAL INFORMATION

The NOAH project is carried out under an R&D program of, and funded by, the European Space Agency.