****

**RCC and OI Team Up to Use Modern AI to Survey the Ancient World**

The part of the world now known as Afghanistan was home to some of the earliest human communities, dating back as far as 50,000 years. But today, the country’s ancient sites are under constant threat, as wars, religious extremism, agriculture, and development lead to the destruction of irreplaceable artifacts and architecture.

In 2015, the US Department of State and the US Embassy in Kabul awarded the University of Chicago Oriental Institute a grant to launch the [Afghan Heritage Mapping Partnership (AHMP)](https://oi.uchicago.edu/research/camel/afghan-heritage-mapping-partnership), which uses satellite imagery to discover and monitor sites of cultural importance in the country. Recently, members of the [Research Computing Center](http://rcc.uchicago.edu) have joined that collaboration, bringing in the latest computer vision approaches to help the team meet its goal of surveying Afghanistan’s more than 650,000 square kilometers.



With the RCC’s help, the AHMP has sped up their process fivefold, finding tens of thousands of potential new archeological sites and automating the monitoring of areas in particularly severe danger. In just two months, the AI-based site detection method has added 11,900 candidate sites to the AHMP dataset, narrowing the areas that the human researchers must subsequently validate.

“What we found is that it’s so much faster to check a site that has already been proposed, than it is to look at a whole region,” said [Gil Stein](https://nelc.uchicago.edu/faculty/stein), Professor of Near Eastern Archaeology in the Department of Near Eastern Languages and Civilizations and at the Oriental Institute and Principal Investigator for the Oriental Institute’s cultural heritage preservation grants in Afghanistan. “We’re hugely excited about this collaboration.”

In mid-May, researchers from the OI and RCC presented these early results to visitors from the Department of State. Teodora Szasz, Computational Scientist for Image Analysis, Visualization, and Machine Learning at the RCC, used the tiled wall display of the Data Visualization Lab to fly across the Afghanistan landscape in Google Earth, zooming in on satellite images of mounds, caravanserai, forts, and qanat — a pipeline system used for water transport and irrigation in ancient agriculture.



Thousands of years later, the computational pipeline built by the RCC team intakes several different types of imagery used by AHMP, including declassified spy plane photographs from the Cold War, hand-drawn topographical maps, and modern satellite and LIDAR scans. Trained on 3,200 sites labeled by human researchers, the pipeline utilizes high-performance computing resources to scan the entirety of Afghanistan, cut up into 70 million individual tiles, in roughly a week, finding similar structures.

Each pass of the model generates areas of interest for manual validation, which creates new labeled data to improve the model even further. In some cases, the model even goes beyond human ability, using the LIDAR data to recognize evidence of ancient sites that wouldn’t have been visible to the naked eye.



“This has been a great collaboration, one that I could not have imagined before,” said [H. Birali Runesha](https://researchinnovation.uchicago.edu/leadership/h-birali-runesha/), Associate Vice President for Research Computing and Director of the Research Computing Center. “It brings together expertise from computer vision, AI, GIS and archaeology, and the team has achieved great results. But more importantly for me, this collaboration has great potential for broader impact beyond my expectation. I can envision how other researchers in archaeology and other disciplines can make use of the methods and pipeline the team has developed.”

The RCC collaboration with AHMP was not the only OI project presented during the visit that uses modern technology to study the ancient world. [Another collaboration](https://oi.uchicago.edu/research/projects/oi-nma-partnership), between the University of Chicago and the [National Museum of Afghanistan](https://www.nationalmuseum.af/), uses 3D imaging, hologram projection, and digital archives to help reconstruct statues from the Hadda archeological site shattered into thousands of fragments during the Afghanistan Civil War. Even after the setbacks of the COVID pandemic and the Taliban’s return to power, the international partnership continued through virtual conferences and computational tools.

“The Taliban in a weird way did us a favor that we can now see inside these sculptures,” Stein said. “We can see impressions on the insides of the sculpture, reconstructing the production steps for how these ancient craftsmen built the sculptures of Hadda. That helps us understand these artifacts not as individual sculptures, but as products of people in a community, interacting with monks and the monastery.”

===== just notes below here

* Impressions on insides of sculpture, TR on reconstructing the sequence of production steps for how these ancient craftsman built sculptures of Hadda
* Built up from bundle of sticks
* If you understand steps of how they produced these things, can get to the craftsman who created the sculptures…understanding not as individual sculptures but products of people in community interacting with monks and monastery

<https://oi.uchicago.edu/research/camel/afghan-heritage-mapping-partnership>

Gil Stein

Prof of Near Eastern Archeology

Orientall Institute

04:25 - importance of interdisciplinary collaboration

04:50 - working with you opened our eyes to what is possible

Fabio Columbo

Alejandro

Elisa

Fabio

Hadda Sculptural Project

Collab with NMA (afghanistan)

Started in 2016, ends this year

Reconstruct art destroyed during civil war

Alejandro - Hadda excavations

First explorations in 1834

Excavations start in 1923

13 different sites

Stupas

1998, site and museum destroyed and looted

Fabio

Feb/March 2001, Taliban destroyed many sites

7,600 fragments in 12 crates

2016 - UChicago & NMA assess feasibility of assembling structures

Archival info also destroyed

2019 found photos and information

1. Documentation of each fragment
2. Conservation treatment of fragments
3. Assess original extent of collection and losses
4. Training of local staff

2016-2019 on site

2020-2022 virtual missions

“We started with a blind situation, giant 3d puzzle w/o any pictures or documentation, without knowing which kind of sculpture was smashed.”

2017: sorting, reorganizing, and rehousing fragments in acid-free archival bags and boxes

2017-2018: mild cleaning, data collection, and stabilization of fragments (in NMA conservation lab)

2018-19: reassembling sculptural fragments

Some fragments retained categorization codes, could be matched against records

Feb 17 2020, last picture taken by OI team in Kabul

* Left tables covered in fragments

Eliza

2020-2022: shifted from on site to virtual missions, focused on study of collection and reassessment of all documentation collected earlier in the project

* Created more than 20 technical reports

2600 sculptures discovered during 1920s currently located outside of Afghanistan vs. 2922 currently in Afghanistan

Buddha heads, seated and standing buddhas, animals, figurines

Photographic database, 4753 files (out of 17808 total)

Conservation database: 483 files

Data on state of conservation

* Majority suffer from deterioration due to aging, followed by war and other events

Online workshops

Donated cameras and laptops to NMA

3D imaging and modeling

* Highlight damage to sculptures
* Virtual reassembling of fragments
* Holograms of 3D model fragments rotating

Loss assessment

* 4658 objects inventoried in 60s and 70s
* 2922 still remaining in Afghanistan
* 1737 lost or too damaged (37.35%)

Alejandro

Comparing numbers on objects to surviving cards and catalogues

Estimate that 54.84% of Hadda collection in Tissot has been lost

SD Q: has info been shared with Interpol?

* Gil: share with museums, law enforcement, auction houses, so nobody can say I didn’t know this was from national museum
* Assuming most were destroyed

Fabio

Further developments: papers, online conference, comprehensive catalogue, photo database

Gil: virtual conference on Hadda and early buddhist art (end of July)

* Don’t just marginalize as conservation, but bring together art history, archeologists, historians, conservators – real accelerator for research on Hadda
* Working on putting sculptures back into the rooms from field excavations
* Not just importance for preserving heritage, but intellectual and scholarly importance
* Focused on bringing researchers together, specialists working together to come up with new understanding of Hadda as monastic community
* Public communication

Q: unintended benefit of conservation work from them being broken up?

* Fabio: Incredible opportunity from bad situation
* Instead of taking samples, gave to us complete info of composition, how they did these sculptures
* Gil: The Taliban in a weird way did us a favor that we can now see inside these sculptures.
* Impressions on insides of sculpture, TR on reconstructing the sequence of production steps for how these ancient craftsman built sculptures of Hadda
* Built up from bundle of sticks
* If you understand steps of how they produced these things, can get to the craftsman who created the sculptures…understanding not as individual sculptures but products of people in community interacting with monks and monastery. (59:50)
* Not just chemical composition, but other things emerging
* Fabio:

==== new recording

Afghan Heritage Mapping Partnership

Andrew Wright, project manager

Started in 2015

* Catalog and map known and previously unknown archeologicall sites
  + Collab with RCC to create automated process to detect sites
* Detect and monitor looting
* Anallyze threats to heritage sites from modern development
* Train and provide professional support

Found 11,265 new sites and 17,000 hydrological features (such as irrigation systems)

Recently surpassed 50% of the land in Afghanistan, 1kmX1km at a time

Uses variety of imagery sources: spy imagery from 60s and 70s, commercial satellite imagery (DigitalGlobe), LIDAR, topographic maps

* Overlapping but not identical information

Site types: Mounds, Caravanserai, Forts, Qanat (underground water tunnels)

Looting Detection

* Monitor and record threats to archeological heritage sites
* Looting looks like pitting
* Military and police installations (esp on mounds)
* Development and agriculture
* Experimented with computer vision to detect

2019-2021: large scale bulldozing in Northern Afghanistan

* 175 mounded sites

Proposal to work with State Dept on assessing who was in power at different times of damage

Pipeline survey and monitoring

* 877 archeological sites within 5km of pipeline plans

AIA “Red List” of Threatened Heritage Sites

* Constantly checking whenever new imagery is available
* Monitor social media accounts
* One case so far: Gereshk Fortress bulldozed in Sept 2021 (“a misunderstanding”)

Workshops in Kabul and Chicago

* Remote training on open source GIS technology

36:30: Teodora

AI-Based Site Detection in Satellite Imagery

* OI and RCC collaboration

364,000 sq kilometers in 7 years (55.8)

Need to complete remaining 288,000 sq kilometers in 1.5 years (44.2) with RCC

* 5x speedup goal

Birali message (39)

* Great collab, one that not imaged before, brings together expertise from CV, AI, GIS, archeology and achieved great results
* Great potential for broader impact beyond my expectations; archeologists can make use of methods and pipeline the team has developed

Ask Dora for full team

Data Handling

* Initially developed for Lidar, adapted to other imagery
* 70 million tiles, filled up supercomputer memory

Impact of AI

* Trained with human labeled data
* Lidar can see what naked eye cannot in some casess

Collaborative Process

* Technical and archeological expertise
* A zig zag and continuous cycle validation testing processing and reiteration
* We train the model to become more and more intelligent
* Gil: The two teams are in constant dialogue, troubleshooting for each other, I’ve never seen anything ilke it, it’s really neat. (54:30)

3,200 sites human labeled

Over 2 months, increased dataset by 11,900 sites with AI (just a snapshot, even larger today)

Takes AI & supercomputer 1 week to go over entirety of Afghanistan

Human validation takes 6 months

* 57:30 Gil comment that this will allow us to meet deadline
* Dora: similar to medical imaging project
* Gil: it’s so much faster to check a site that’s already been proposed, than to look at a whole image and find one (10x faster)
* Dora: these models can be applied to other regions and areas

Gil: we’re hugely excited about this