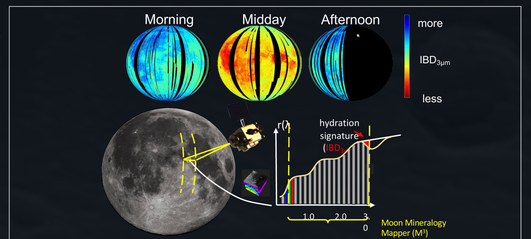
A Foundational Multimodal Model for Lunar Surface Correlation Discovery: Uncovering Hydration and Beyond

Tejas Panambur, Mario Parente, Kay Wohlfarth, Melissa Martinot, Tom Sander, Simone Zorzan, Janice L. Bishop

① Why Lunar Hydration Matters

"Cracking the Moon's Daily OH/H_2O Cycle"

- Since 2009, M³ has revealed that the Moon hosts a complex OH/H_2O cycle, despite its airless, arid regolith
- The 2.7–3 μm absorption band is a direct proxy for hydroxyl (OH) and water (H_2O) on lunar soils
- That band is deep at sunrise, diminishes at midday, then deepens again in the afternoon. The amplitude and "constant component" vary with location
- We still don't know whether these diurnal swings are driven primarily by solar wind implantation, mineralogy, or thermally activated diffusion
- **Why it matters**: Pinpointing these drivers is critical for future in-situ resource utilization (ISRU) and for interpreting upcoming data from Lunar Trailblazer



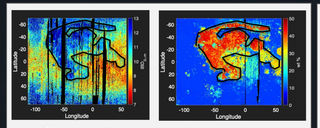
Top maps: water band depth for different lunar day periods. Bottom left: illustration of M³ scan and resulting spectrum. Bottom right, color-coded illustration of the hydration signature with respect to the maps on the top.

② Initial Insights from M³ Data

"Where Minerals Meet Water: Early Correlations"

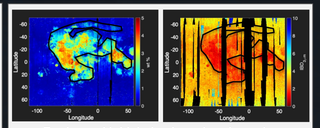
During the initial phases of data collection and analysis, we identified interesting relationships between mineralogic data and hydration signatures.

Morning $IBD_{3\mu m}$/Pyroxene map



Regions rich in pyroxene have a higher hydration signature in the morning.

Titanium map/Midday hydration ($IBD_{3\mu m}$)



Regions with high titanium abundance have a lower hydration signature during midday.

③ What We've Achieved

"Building a Lunar Foundation & Validating Key Hydration Links"

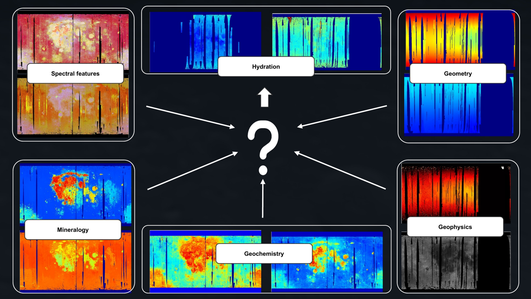
- **Developed Per-Modality Tokenizers:** Trained a set of dedicated tokenizer for each data source (spectral, mineralogical, elemental, topographic) to convert 128×128 patches into discrete tokens.
- **Built the first Lunar Foundation Model:** Integrated all tokenizers into a single multimodal masked autoencoder, fusing spectral, mineral, elemental, and topo tokens to predict any modality (or set of modalities) from any other modality (or set of modalities).
- **Validated Core Modal Correlations**
  - Pyroxene ↔ IBD_3 μm: Strong, qualitatively significant link confirmed
  - Olivine ↔ IBD_3 μm: Weaker but notable relationship observed

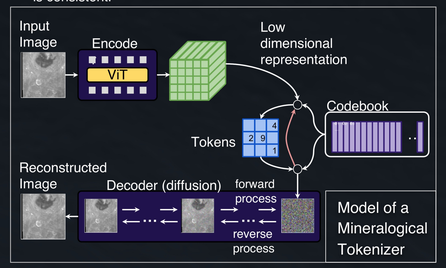
④ Our Lunar Multimodal ML Pipeline

"From Pixels to Predictions: The ABCs of Multimodal Training"

A – Abstraction: Tokenize Lunar Modalities

- **Input Sources (Modalities):** Information comes in different shapes and sizes
  - Spectral Data (M³ IBD_3 μm & full-band reflectance)
  - Mineralogical Maps (pyroxene, olivine, plagioclase)
  - Elemental Abundances (Ti, Fe, etc. from Gamma-Ray)
  - Geophysical/Topographic (DEM, albedo, slope)
  - Metadata (TimeStamp etc.)
- **Advanced Calibration:** M³ data are corrected for thermal & photometric effects so that every pixel is consistent.

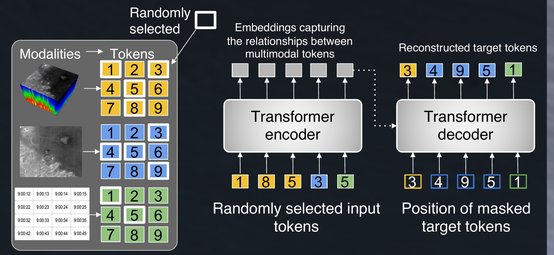




- Each data source (spectral, mineralogical, elemental, topographic, etc.) is equipped with its own fully trained tokenizer
- For every modality, each 128 × 128 patch is run through its modality's VQ encoder to generate a discrete token ID from its modality's unique codebook.
- The diffusion decoder uses that token's embedding to denoise noise into the original patch.

B – Bridging: Masked Autoencoder Learns Cross-Modal Representations

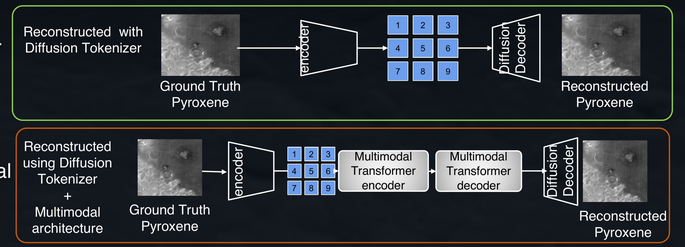
- Once the tokenizers are trained, we train a transformer encoder–decoder to learn cross-modal relationships.
- **Training:**
  - Randomly hide a subset of tokens across modalities
  - Visible tokens pass to the encoder; masked tokens become reconstruction targets
- Encoder embeddings captures relationships between multimodal tokens. Decoder receives (1) the encoder's embeddings and (2) the positional indices of the masked tokens. It reconstructs each masked token at its original position, regardless of which modality it belonged to.



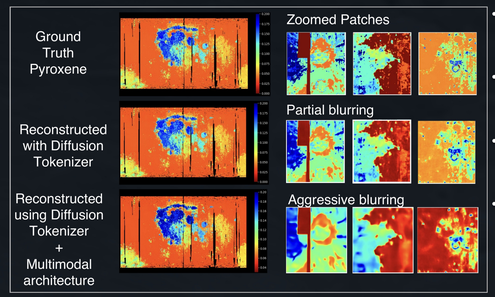
⑤ C - Causal Correlation Discovery

"Testing which Modalities Drive Hydration & Mineral Relationships"

Inference Pipeline



Within-Modality Fidelity ("Can We Rebuild What We Have?")



- **Sanity Check:** Confirm the model can reconstruct one modality accurately before testing causal links
- **Tokenizer-Only:** Rebuilds the Pyroxene map with minimal blur with fine textures
- **Multimodal Fusion:** Introduces extra smoothing around edges but retains the overall pyroxene footprint
- **Takeaway:** Since tokenizer-only is sharper, any cross-modal errors in bridging phase reflect causal complexity, not a broken autoencoder

Cross-Modal Inference ("Which Modalities Correlate with Hydration?")

For this work all experiments and validation were conducted solely among the three modalities: Pyroxene, Olivine, and Morning IBD_3 (hydration).

Primary Insights : Pyroxene abundance is the strongest causal predictor of morning IBD3 (Hydration) signals