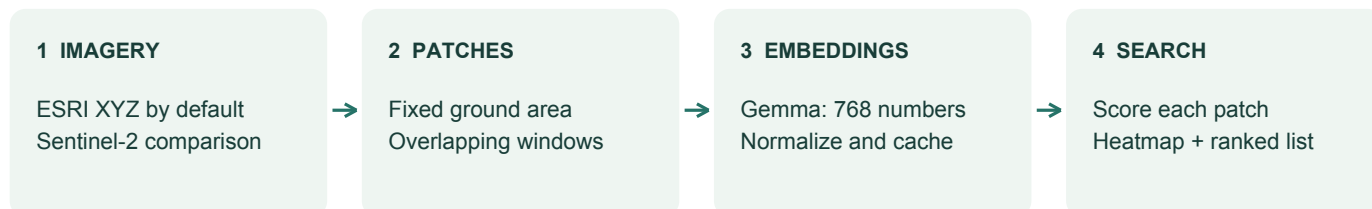


Find kilns by words or examples

The app ranks overhead image patches and places their similarity scores on a map. It supports discovery and human review; a high score is not a confirmed kiln.

ONE SHARED PIPELINE



The same encoder places words and images in a shared space. Similar items have similar unit-length vectors. Once images are embedded, new queries reuse the cached vectors.

THREE WAYS TO DEFINE WHAT YOU WANT

01 Text to image

Describe the target: "A satellite image of a brick kiln."

Encode the sentence with SearchQuery. Rank image vectors by cosine similarity to the text vector.

$$\text{score} = \cos(\text{text}, \text{patch})$$

"brick
kiln"

02 Image to image

Select one kiln image, or several positive examples.

Use its image vector, or normalize the mean of the positive image vectors. Match the same visual pattern.

$$\text{score} = \cos(\text{positive prototype}, \text{patch})$$

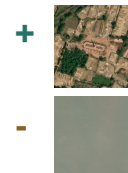


03 Positive + negative to image

Provide kilns and backgrounds you want to reject.

Build separate positive and negative prototypes. Reward resemblance to kilns and subtract resemblance to negatives.

$$\text{score} = \cos(\text{positive}, \text{patch}) - \cos(\text{negative}, \text{patch})$$



No fine-tuning is needed. All three modes work on the map. Click + Positive on a kiln or - Negative on a lookalike, then search again. Prepared image vectors are reused; moving the threshold needs no model call.

FROM SCORES TO A MAP AND A REVIEW QUEUE



Default extent: 31 x 31 windows, 320 m footprint, 50% overlap: a 5.12 x 5.12 km envelope. Reference-overlapping windows are excluded. Nearby kilns can merge into one candidate; large sites can create more than one.

Existing labels: the AlphaEarth app source saves three kiln example coordinates. Other clicked points are session state. Export/import their CSV to preserve them. SentinelKilnDB supplies the separate labelled evaluation set.

How this differs from AlphaEarth

Both systems compare embeddings. The crucial difference is what an embedding represents, and whether a text query can be compared directly with it.

INPUTS AND UNIT OF COMPARISON

ALPHAEARTH SPACE		GEMMA KILN SEARCH
Annual multi-sensor observations 64D vector for each 10 m pixel Reference locations -> pixel similarity		One RGB image patch at a time 768D vector for the whole patch Words or example images -> patch similarity
Query	Selected locations / reference embeddings. The linked app has no direct text-query encoder.	Text, one image, multiple positives, or positive-minus-negative examples.
What it sees	Annual surface patterns from multiple Earth-observation sources.	Appearance in the chosen RGB imagery. No annual time series in this prototype.
Heatmap	Dense scores on the 10 m embedding grid. Context extends around each pixel.	Scores from overlapping image windows. Smooth-looking colours do not imply precise boundaries.
Changing imagery	A satellite basemap change does not change the underlying AlphaEarth embeddings.	Map: full ESRI or ESRI resampled to 10 m, re-embedded. Regional tests separately compare ESRI and original Sentinel-2.
Scale and use	Precomputed geospatial features suit large-area screening.	New areas require tile fetching and inference. Cached patches support fast reranking.

MEASURED PILOT / NOT AN ALPHAEARTH HEAD-TO-HEAD

432 sites across four longitude bands: 96 support, 96 calibration, 240 test. Each target band is held out from support and calibration; each test band has 30 positives and 30 negatives. All selected sites are at least 5 km apart.

Average precision (AP) on held-out test scenes; higher is better

Test band	S2 / text	S2 / pos-neg	ESRI / text*	ESRI / pos-neg*
West (<75E)	0.980	0.995	0.936	1.000
Central (75–82E)	0.928	0.946	0.672	0.982
East (82–88E)	0.948	0.923	0.802	0.999
Far east (≥88E)	0.972	1.000	0.999	1.000

*ESRI results measure agreement with transferred Sentinel labels. Capture dates and alignment are not independently validated. Balanced convenience samples and easy negatives can inflate AP. These numbers do not measure kiln-count accuracy or establish superiority over AlphaEarth.

Next comparison: use identical areas, dates and independently reviewed labels for both systems; report precision at a fixed review budget and recall of known kilns by region. A useful hybrid is AlphaEarth screening -> Gemma reranking -> human review.