

Interactive Visual Feature Search

Devon Ulrich^{1*} and Ruth Fong² | ¹Jane Street ²Princeton University *Work done at Princeton University



CVPR
JUNE 17-21, 2024

Introduction

Reverse image search tool to easily and interactively visualize intermediate features for any computer vision model & dataset.

Designed for quick, interactive use in Jupyter Notebooks.

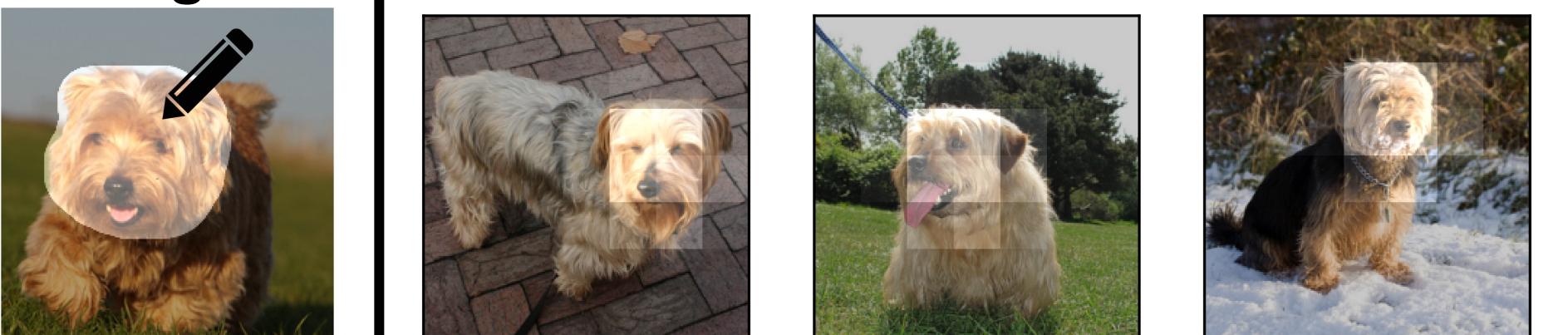
```
import visualsearch as vfs
model = torchvision.models.vgg16().features

cache = vfs.caching.precompute(dataset, model, ...)
search_tool = vfs.CachedSearchTool(model, cache)

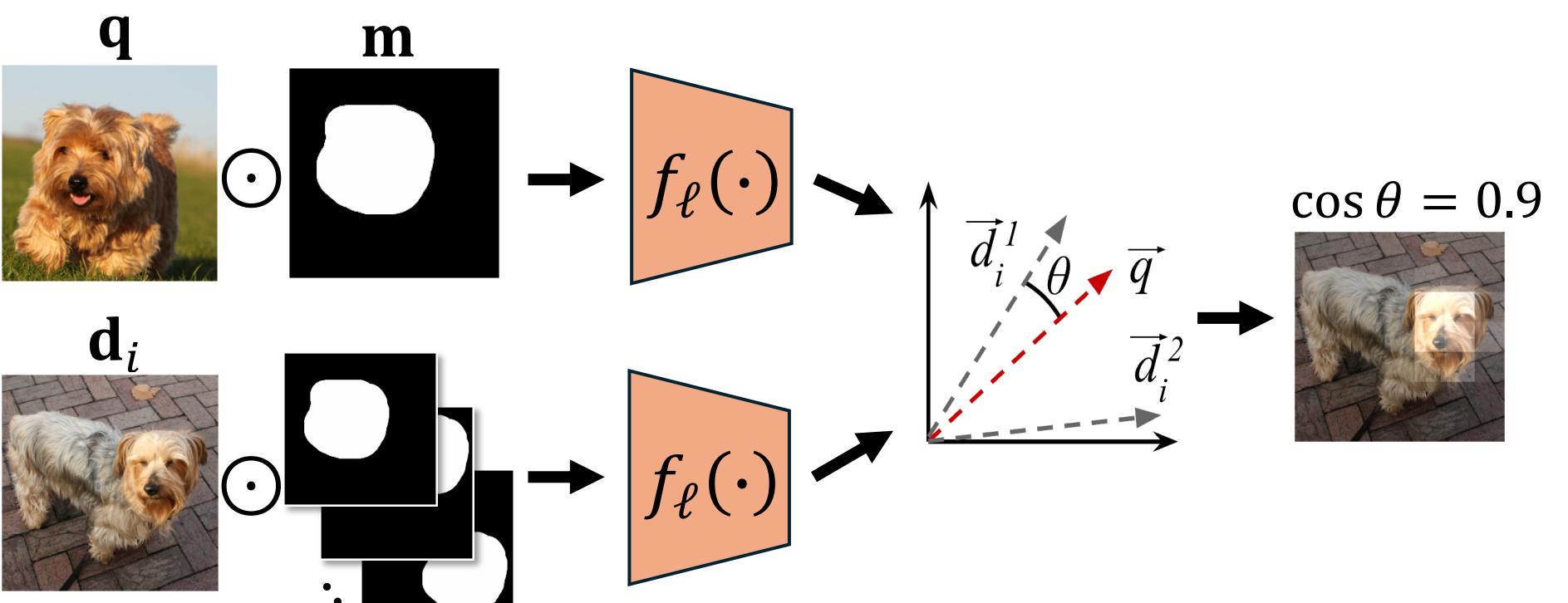
display(vfs.widgets.HighlightWidget(query_img))

outputs = search_tool.compute(
    query_img, get_selected_region())
display(vfs.plot_search_results(outputs, ...))
```

Highlight Widget

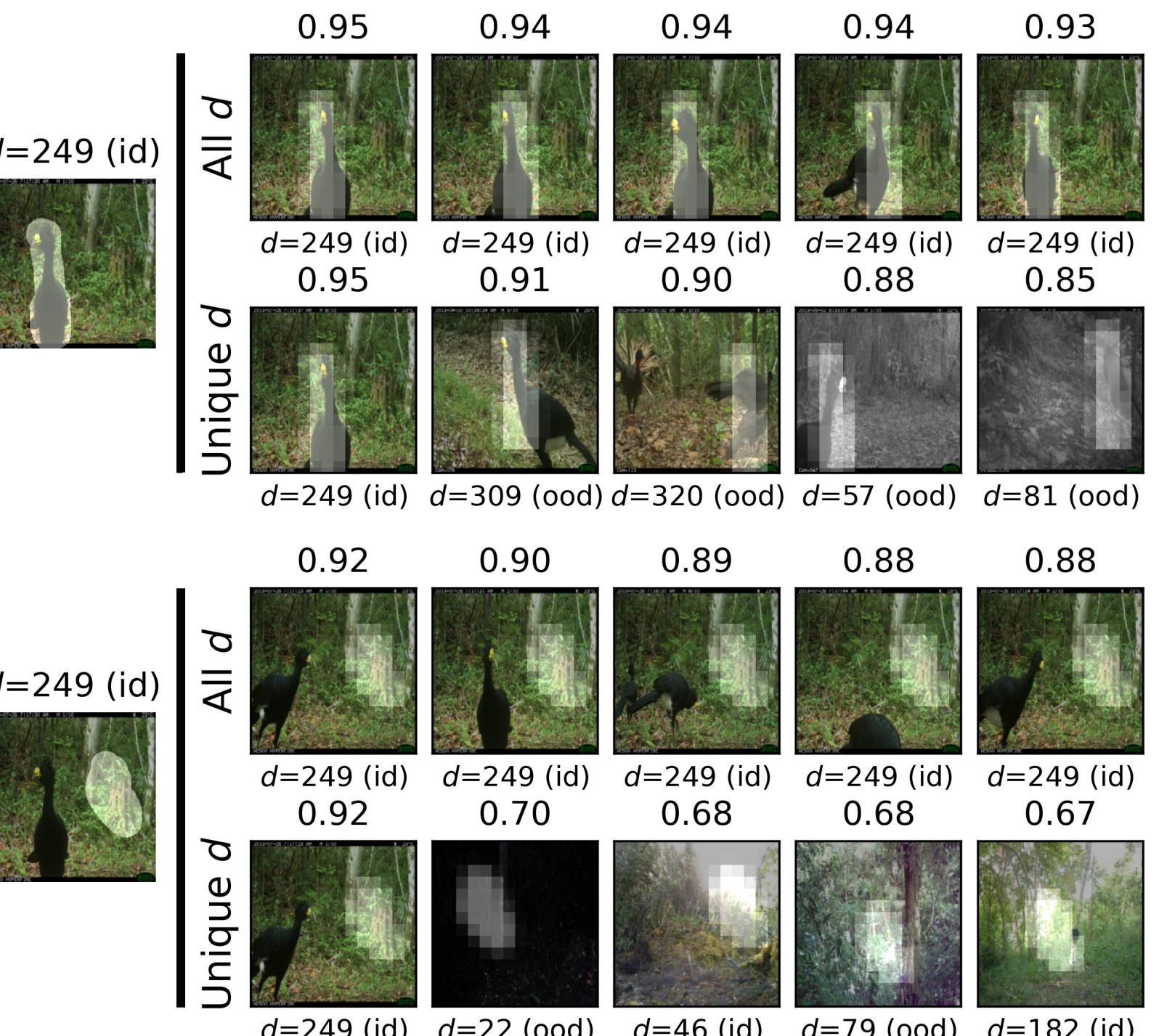


Method

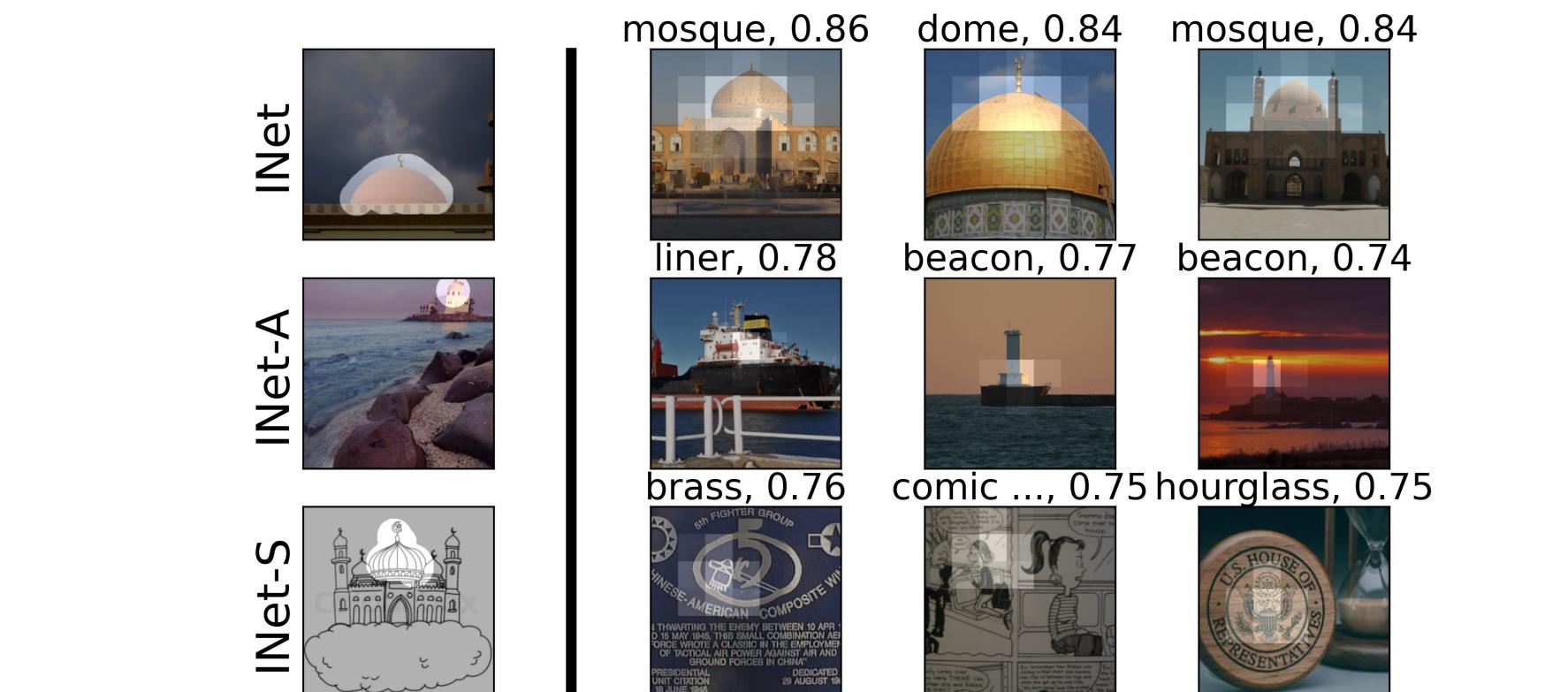


User selects region m in query image q ; our tool returns the top k most similar images from dataset \mathcal{D} using CNN features after layer ℓ (f_ℓ).

Domain Generalization

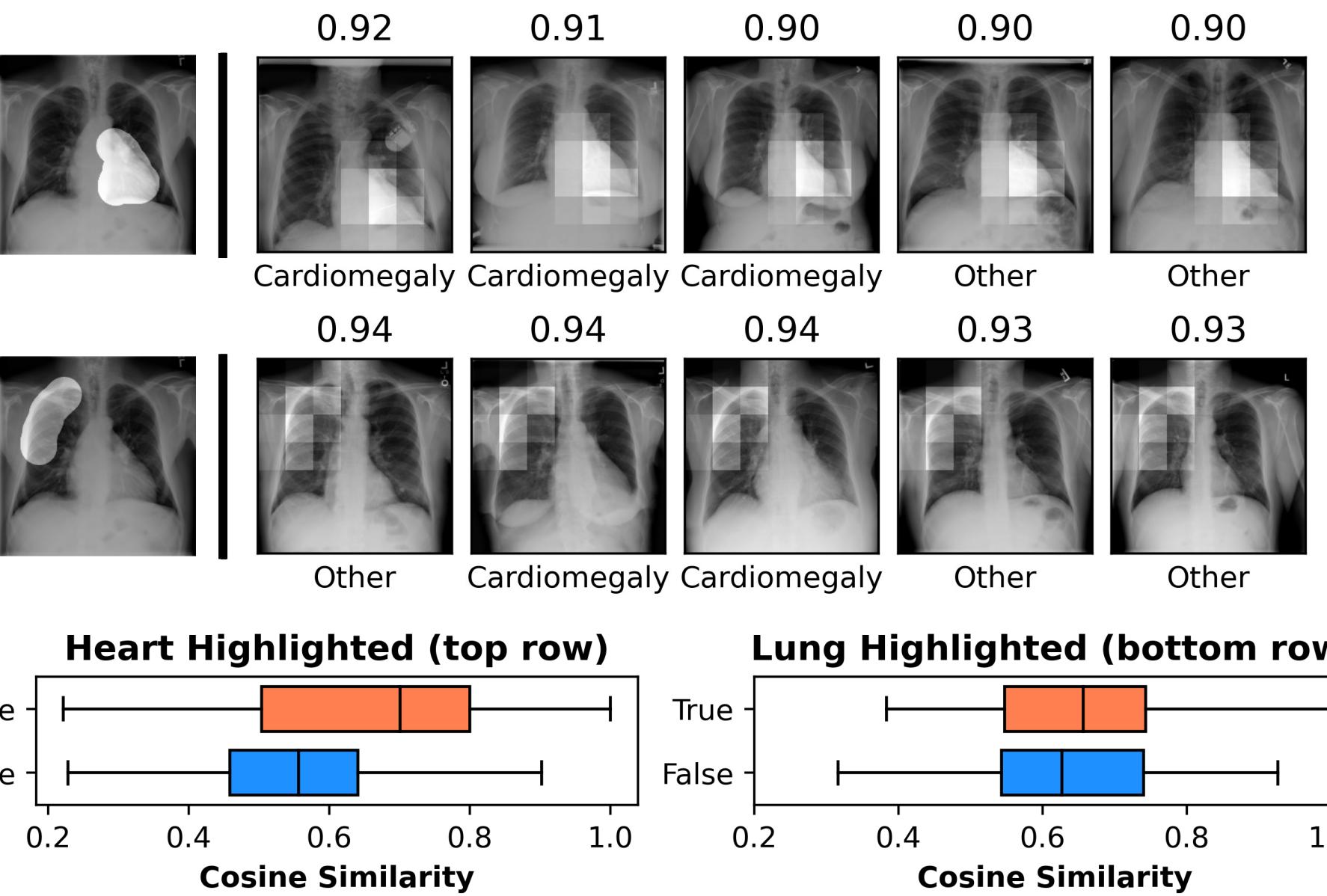


- iWildCam dataset [1] has in-domain locations (e.g., $d=249$) and out-of-domain locations not in training set (e.g., $d=57$).
- When bird is highlighted (top rows), o.o.d. images with same bird species are returned with high similarity (vs. background location highlighted in bottom rows).



- ImageNet classifier queried on mosques from ImageNet (top row, in-domain), ImageNet-A [2], and ImageNet-Sketch [3] (middle and bottom rows, both o.o.d.)
- In-domain search results contain other mosques, while o.o.d. results instead include regions with similar textures

Chest X-Ray Classifiers



- Classifier trained on ChestXray-14 dataset [4]. Query image is labelled with cardiomegaly (enlarged heart).
- When the heart (with cardiomegaly) is highlighted (top row), most similar images also have cardiomegaly (vs. lung highlighted, bottom row).

Editing a Classifier



- CNN editing [5] improves accuracy for classifying vehicles on snowy surfaces.
- When snow below a car is highlighted, edited classifier results include other vehicles on asphalt (bottom row) vs. other objects on snowy surfaces (top row).

PASS Training



Faces are blurred to protect the privacy of photographed individuals.

- PASS dataset [6] is an ImageNet alternative that does not contain photos of humans.
- Searching for a face yields comparable results for an ImageNet model (top row) vs. PASS model (bottom row).

Code and Demos

Code and demos available on GitHub.

Designed for Jupyter and Google Colab notebooks.



References

- P. W. Koh et al., ICLR 2022. WILDS: A Benchmark of in-the-Wild Distribution Shifts.
- D. Hendrycks et al., CVPR 2021. Natural adversarial examples.
- H. Wang et al., NeurIPS 2019. Learning robust global representations by penalizing local predictive power.
- X. Yang et al., CVPR 2017. ChestX-ray8.
- S. Santurkar et al., NeurIPS 2021. Editing a classifier by rewriting its prediction rules.
- Y. M. Asano et al., NeurIPS 2021. PASS.