

## Motivation

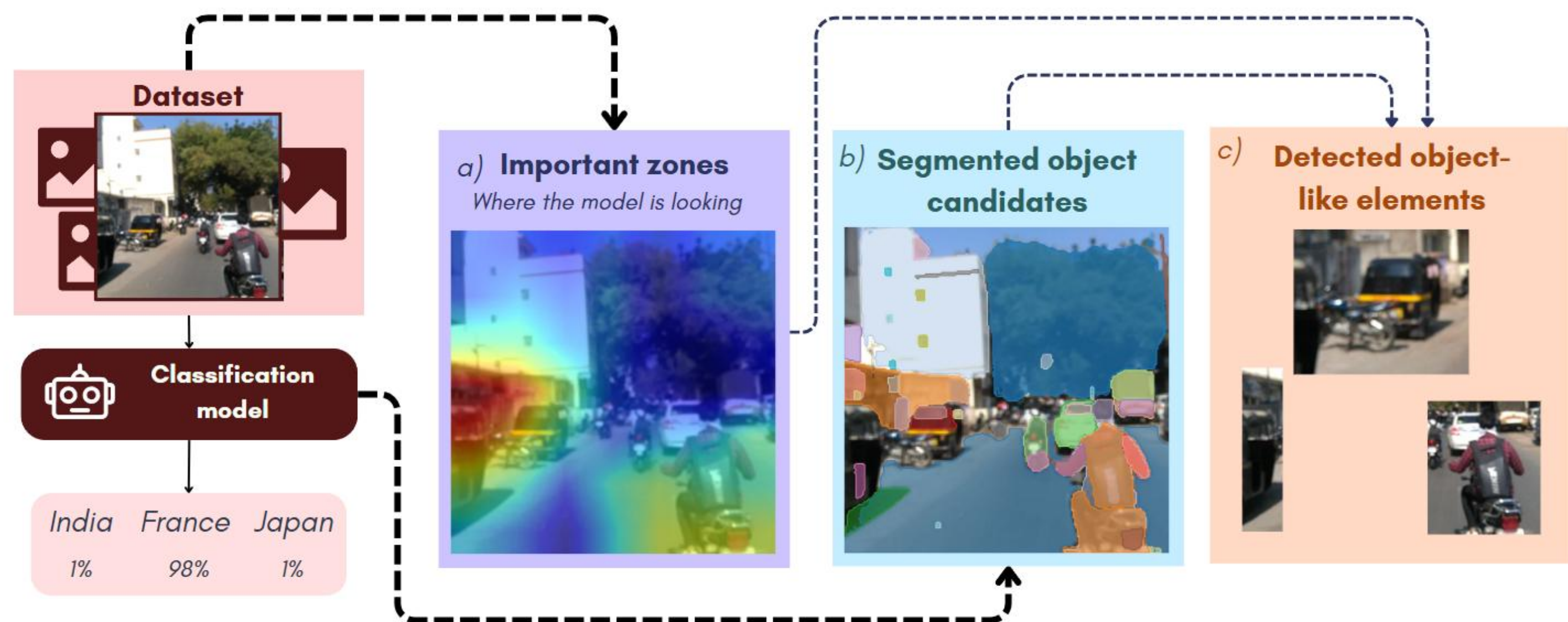
- Humans solve geolocation tasks such as GeoGuessr [1] using **concrete visual cues**, including road markings, vehicles, vegetation, street signs, and language.



- Although **modern geolocation models** achieve strong performance, it remains **unclear whether they rely on similar cues**.

**Question: Do geolocation models rely on similar object-level visual cues than humans in geolocation?**

## Approach



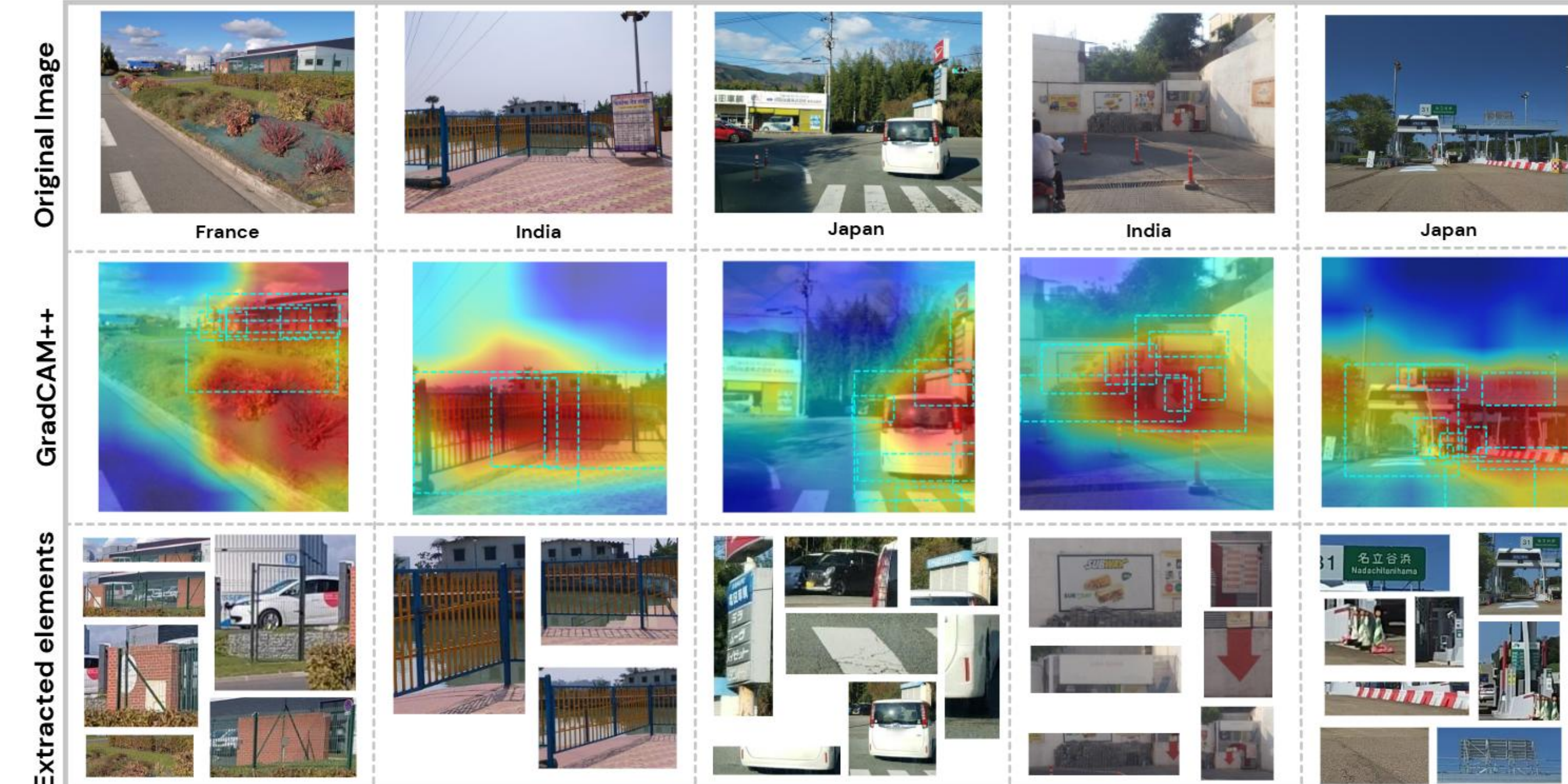
- Starting from a geolocation model that predicts the country of origin of a ground-level image:
- We extract salient regions using attribution methods → **Where is the model looking?**
  - We segment candidate object-like regions in the image → **What visual elements are present?**
  - We retain only segments overlapping salient regions → **Which elements contribute to the prediction?**

## Implementation

- We use the following settings:
- GradCAM++ [2] for attribution
  - MobileSAM [3] for segmentation

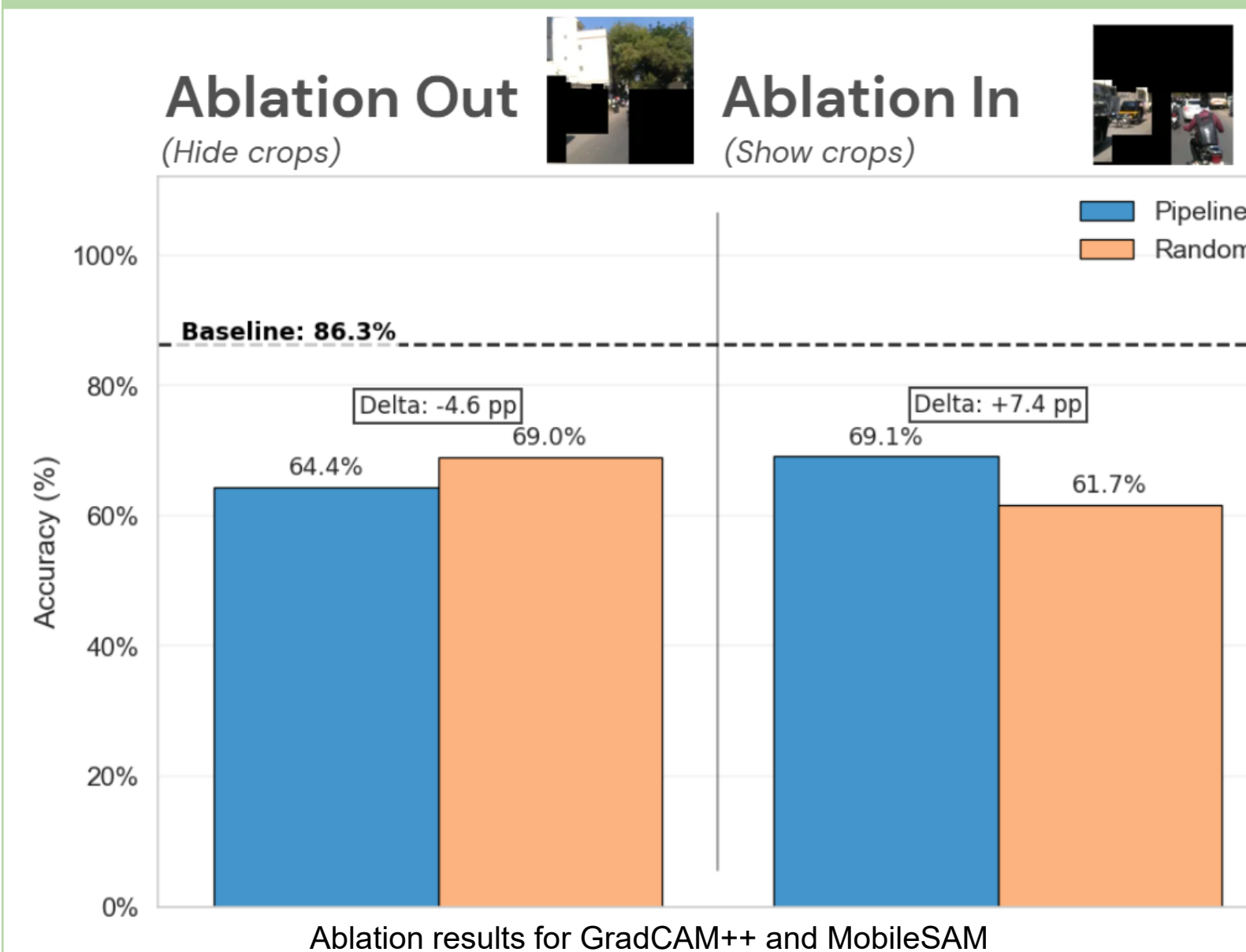
We apply the framework to a ResNet50-based geolocation model trained to classify images from 3 countries: **France, India, and Japan**.

We obtain **object-like elements** considered important for the model's prediction.



**Attribution-guided object-like regions retain significantly more predictive information than randomly selected regions.**

## Validation



We evaluate the extracted regions using ablation tests:

- Ablation Out:** Remove extracted regions and predict → We expect a **larger drop**
- Ablation In:** Keep only extracted elements and predict → We expect a **smaller drop**

Results show that **attribution-guided elements consistently outperform random regions**.

See paper for additional results with different attribution and segmentation methods.

## Discussion

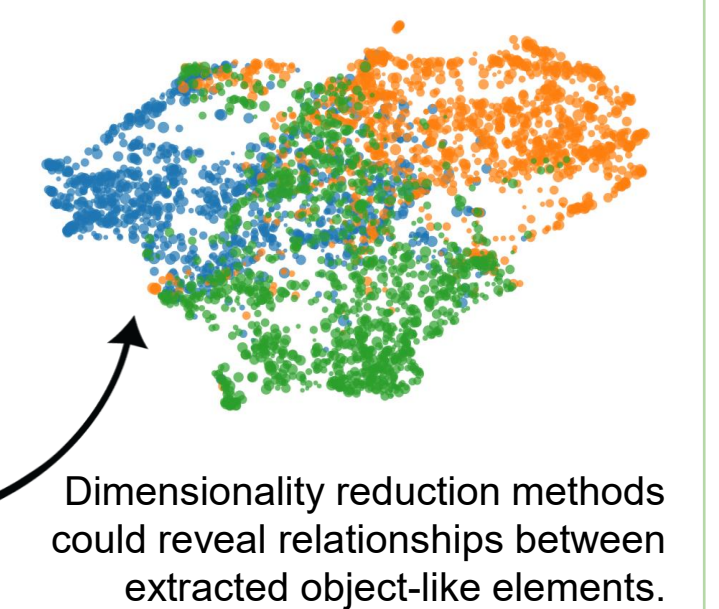
Our findings suggest that geolocation predictions can be **partially traced to object-like visual evidence**.

### Limitations

- Study limited to 3 countries
- CNN-based geolocation model
- No stronger non-attribution baselines

### Future Work

- Larger-scale geolocation benchmarks
- Modern geolocation architectures
- Analysis of country-specific cues
- Human evaluation



## References

- GeoGuessr website: <https://www.geoguessr.com/>
- A. Chattopadhyay et al., Grad-CAM, 2018 IEEE Winter Conference on Applications of Computer Vision
- Y. Liu et al., MobileSAM-Track, Remote Sensing

ArXiv paper



Test your geolocation skills (demo)

