

Happy new year from AI-Hab

Happy New Year from the AI-Hab project team! Since our kickoff in 2024, we have made significant strides towards AI-powered habitat classification. In this newsletter, we'll share the latest project updates.

Many of you participated in our first workshop last April, where your feedback emphasised **accuracy, transparency, and robustness** as key priorities for the AI-Hab tool. We have taken these priorities to heart in developing our initial AI model.

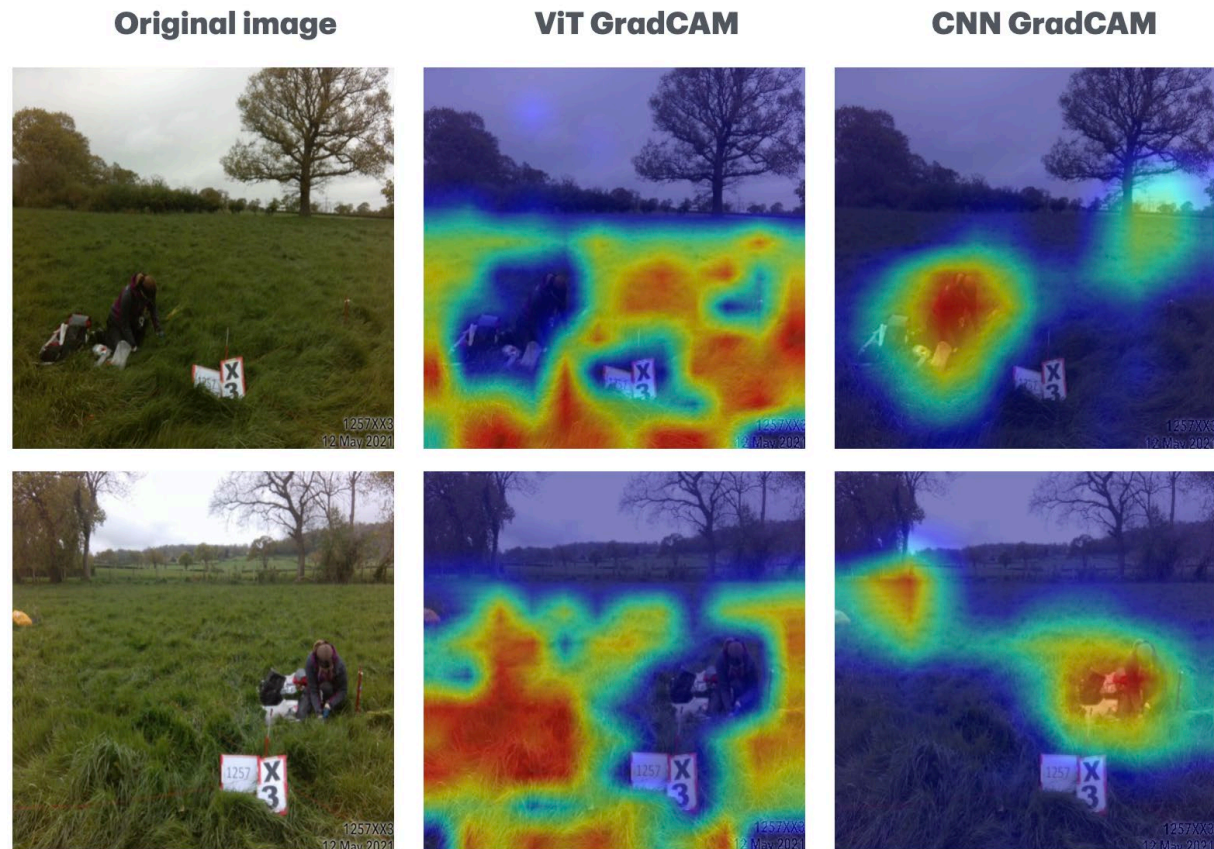
Model development

The first step in model development was to create a training data set. We did this by curating a dataset of 5,599 images across 20 broad habitat types at UKHab Level 3 from UKCEH's Countryside Survey (2019-2023).

We trained several models on these images. At this stage, our objective was to see how well different "off-the-shelf" models perform task of identifying habitat from images, without intensive fine tuning. We call these our baseline models, so that as we develop our models further, we can compare their performance to these baseline models to see how much better the new models are. We developed baseline models for habitat classification (at UKHab Level 3) based on state-of-the-art deep neural networks, including convolutional neural networks (CNNs) and vision transformers (ViT).

Once we've fitted models, we assessed how well they perform by calculating their accuracy. Beyond accuracy, visualisation techniques were particularly useful as it showed us which parts of the image the model is using to perform a habitat classification, and how the model regards different habitats internally. We did this using importance mapping (GradCAM) and feature dimensionality reduction (t-SNE).

- Successes:
 - ViTs consistently outperform CNNs of similar size/complexity (i.e., # of parameters)
 - ViTs consider global context better than CNNs, ignoring distractor objects, such as survey markers present in the image
 - Our best performing baseline (Swin-T) achieves 69.31% accuracy at UKHab Level 3
- Challenges:
 - Grassland is tricky; the lowest performance was observed for these categories
 - Swin-T is a large ViT model; it may not be suitable for inference on mobile devices currently



Heatmap visualisations produced using GradCAM, highlighting the extent to which different regions contribute to the overall classification of a habitat. CNNs tend to focus their attention on small regions of an image, including extraneous objects found in the Countryside Survey dataset (signs, bags, people). ViTs consider global context more consistently by design, while also largely ignoring distractor objects. ViTs also consistently outperform CNNs at habitat classification.

New team members

Dr Hongrui Shi

Hongrui Shi is a Post-Doctoral Research Associate in Computer Vision/Machine Learning at the University of Lincoln. Hongrui's PhD research was focused on federated learning, distributed machine learning, and addressing the challenges of training AI efficiently on remote devices. Hongrui is currently leading the development of AI for habitat classification on the AI-Hab project, under the supervision Dr James Brown, Dr Petra Bosilj, and Dr Lan Qie.

Dr Simon Rolph

Simon is a data scientist at the UK Centre for Ecology & Hydrology. Simon has taken over the day-to-day management of UKCEH's part of the project from Lisa Norton. Lisa will remain involved in the project. This includes integrating the AI model into the E-Surveyor platform. E-Surveyor is a free mobile app that helps surveyors assess the quality of habitats and already offers a suite of tools such as AI for identifying plant species. You can learn more about E-Surveyor here:

<https://www.ceh.ac.uk/data/e-tools/e-surveyor>

Other current team members (just so you don't forget about us):

[Dr Lan Qie](#)

[Dr James Brown](#)

[Dr Petra Bosilj](#)

[Dr Lisa Norton](#)

[Prof. Richard Pywell](#)

Plans for 2025

This year, we will continue training and refining the AI model and explore alternative training regimes, such as contrastive learning, while also incorporating the hierarchical nature of the UKHab classification system into the model. We will also begin developing the specifications for the API that will link the model to the user interface. We will make sure to incorporate your feedback on the desirable features for the user interface during this stage.

As promised, we will organise another stakeholder workshop, likely in late 2025 or early 2026. Stay tuned for more updates!

We thank you again for your engagement, and please feel free to contact us.

The AI-Hab team