MULTI-STAGE SEMANTIC SEGMENTATION TO MAP SMALL AND SPARSELY DISTRIBUTED HABITATS

Thijs L. van der Plas
The Alan Turing Institute
London, NW1 2DB, UK
t.vanderplas@turing.ac.uk

Simon T. Geikie
Peak District National Park Authority,
Bakewell, DE45 1AE, UK
simon.geikie@peakdistrict.gov.uk

David G. Alexander*
Peak District National Park Authority,
Bakewell, DE45 1AE, UK
david.alexander@peakdistrict.gov.uk

Daniel M. Simms*
Applied Remote Sensing Group, Cranfield University
Cranfield, MK43 0AL, UK
d.m.simms@cranfield.ac.uk

ABSTRACT

Land cover (LC) maps are used extensively for nature conservation and landscape planning, but low spatial resolution and coarse LC schemes typically limit their applicability to large, broadly-defined habitats. In order to target smaller and more specific habitats, LC maps must be developed at high resolution and fine class detail, using methods that can handle strong class imbalance. In this work, we present a new aerial photography data set with 12.5 cm ground resolution, annotated using a detailed, hierarchical land cover schema. We show that splitting up the semantic segmentation process into multiple stages critically improves the predictive performance, in particular by including the rare LC classes. We then apply this method to create a new LC map of the Peak District National Park (1439 km^2), England, at 12.5 cm resolution.

1 INTRODUCTION

In areas under pressure from high population density and intensive agriculture, habitats that support high biodiversity are often small and thereby sparsely distributed. This leads to a very imbalanced land cover (LC) class distribution, where the most relevant classes (for biodiversity research) occupy a smaller area than others (Waldner et al., 2019; Tuia et al., 2023). Therefore, to be able to accurately map these habitats, we require very-high-resolution remote sensing data to identify small areas of habitat, and methods that can deal with imbalanced data sets. We address this challenge by the following two contributions:

- The creation of a newly annotated, very-high-resolution (12.5 cm) data set of RGB aerial photography and land cover. We used a hierarchical land cover schema, consisting of 23 classes, specifically designed for UK National Parks (Taylor et al., 1991b;a; 2000).
- The development of multi-stage semantic segmentation, where the semantic segmentation task is split up in different stages corresponding to the hierarchical land cover schema. We show that this improves the segmentation of imbalanced data sets.

*: Co-senior author

This conference paper is based on our recent publication (Van der Plas et al., 2023b). When referring to this work, please cite Van der Plas et al. (2023b) instead of this conference paper.

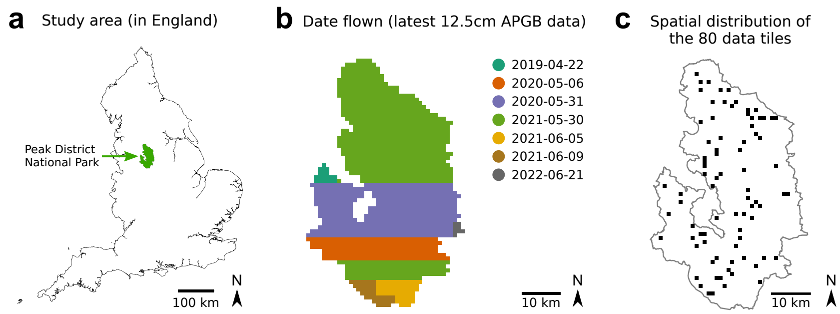

Figure 1: Study area details. **a)** Our area of study is the Peak District National Park. **b)** RGB image capture date per square kilometre tile. **c)** Distribution of 80 tiles used to create the data set.

We then applied this method to create a LC map at 12.5 cm resolution of the Peak District National Park (PDNP, 1439 km^2), providing the first updated very-high-resolution LC map of the PNDP since the manually annotated LC map by Taylor et al. (1991b).

2 DATA

We created a new data set of LC annotated RGB images at 12.5 cm resolution, which we have made publicly available (Van der Plas et al., 2023a). Our study area is the Peak District National Park (PDNP), England, which totals 1439 km^2 (Fig 1). Orthorectified aerial digital RGB photography of the entire PDNP (1439 km^2) was obtained at 12.5 cm ground resolution, at seven different imaging dates (Fig 1b), through the Aerial Photography Great Britain (APGB) agreement for UK public sector bodies (Bluesky, 2022). The image data were available as 1 km $\times$ 1 km *tiles*, which we split into 64 m $\times$ 64 m (i.e., 512 pixels $\times$ 512 pixels) *patches* for input to the Convolutional Neural Network (CNN) models. We then created a data set of 1027 non-overlapping patches, sampled from 80 tiles distributed across the PDNP (see Appendix A.1), which covered the variety in LC classes, the variability within LC classes and the variability in image acquisition across different regions (caused by different flight dates). We adapted the land cover schema from Taylor et al. (1991b;a), specifically designed for mapping UK habitats, with the addition of a new wetland vegetation class (F3d, wet grassland and rush pasture, also see Appendix A.2) and an extra subclass of upland heath (D1b, peaty soil upland heath): see Table A.1. This is a hierarchical land cover schema, as illustrated in Fig 2c. To facilitate further use of these data and this schema, we have created an online interpretation key with additional detail and example images of each class (Alexander et al., 2023).

The 1027 patches (64 m $\times$ 64 m) selected for training and testing the models were annotated manually by visual image-interpretation according to the LC schema of Table A.1. Annotations were first done by one human expert interpreter, and afterwards, they were all verified (and corrected where necessary) by a second expert interpreter. Finally, field visits were conducted to resolve any remaining ambiguities in the annotation, leading to a highly detailed and accurate data set representative of upland UK landscapes. We used publicly available woodlands data to aid and speed up the manual annotation of large woodlands in image patches (Forestry Commission, 2022). Further, OS NGD data were used for mapping the F2, G2a and H classes (Figure 2 and Appendix A.3, (Ordnance Survey, 2023)).

This data set was split randomly into 70% for training and 30% for testing. (The training set was used as validation to optimise the method). Spatial stratification was not applied, because patches were non-overlapping and spatial stratification was likely to bias the distributions of land cover, neighbouring land cover, image capture date and seasonality.

3 METHODS

3.1 U-NET MODEL TRAINING

We used Resnet50 U-Net models for all classifiers, all trained for 60 epochs using a batch size of 10, after which the model that achieved minimal training loss was selected. We considered two loss functions (cross entropy and focal loss with $\gamma = 0.75$), and optimised five CNNs for each loss function using an Adam optimizer for all classifier types (main, C, D and E and single-stage baseline, Table A.2). The best (*i.e.,* maximum accuracy) model of each type was then selected for the final predictions.

CNNs received RGB image patches of 512×512 pixels as input. To avoid edge effects, we used a padding of 22 pixels (meaning neighbouring image patches slightly overlapped). Input images were z-scored, and during training, data were augmented by random horizontal and/or vertical flipping. For training the main classifiers, LC annotations were relabelled to their corresponding main class (*e.g.,* C1 was relabelled to C). For training the detailed classifiers, LC annotations that were not relevant to the classifier (*e.g.,* C1 is not relevant to the D classifier) were blanked out and did not contribute to the loss during training/testing.

Predictions were evaluated by pixel-wise comparison between the human-annotated LC labels and the model-predicted LC labels. The following evaluation metrics were used, where TP = true positive, FP = false positive and FN = false negative predicted pixels, and c indexes the LC class (e.g., C1, C2, ...):

$$\text{sensitivity}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FN}_c}, \quad \text{precision}_c = \frac{\text{TP}_c}{\text{TP}_c + \text{FP}_c}, \quad \text{accuracy} = \frac{\sum_c \text{TP}_c}{\sum_c (\text{TP}_c + \text{FN}_c)} \tag{1}$$

3.2 MULTI-STAGE SEMANTIC SEGMENTATION

We developed a multi-stage approach because of the hierarchical LC schema, the high number of classes (23), the strong non-uniformity of the class distribution (Fig 2b) and the intra-class variance caused by the large area of interest (1439 km^2). The classification process was split into four stages (Fig 2a). First, one CNN model was used to predict the main classes directly from the RGB data. Second, existing topography data were used to overwrite these predictions with any F2 (open water), G (rock) or H (developed land) LC class (Ordnance Survey, 2023). Third, three separate CNN classifiers were trained for the prediction of the detailed sub-classes. These detailed predictions were then masked using the combined classes from the previous step. For example, the output of the C-classifier would predict, directly from the RGB data, detailed C1, C2, C4 and C5 classes at locations classified as C by the main classifier. Fourth, existing soil data (Natural England, 2022) were used to disambiguate between subclasses of D (moorlands) with or without peaty soil: *e.g.,* D1a or D1b (also see Appendix A.4).

4 RESULTS

We first trained the main classifier to classify the RGB images into the three main (natural) classes of our LC schema: woodlands (C), moorlands (D) and grasslands (E), reaching 95% accuracy overall, with sensitivity and precision above 0.9 for all classes (Table 1). Next, the C, D and E classifiers achieved 92%, 72% and 87% accuracy, respectively, with precision and sensitivity values ranging from 0.7 to 1 for most predicted classes (Table 1).

We found that our multi-stage semantic segmentation method (*i.e.*, stacking the main and detailed classifiers) outperformed conventional, single-stage CNN classification (*i.e.*, directly classifying the detailed classes, see Appendix A.5). Specifically, single-stage models could only predict 8/12 detailed classes, failing to learn the low-density data classes. Instead, multi-stage semantic segmentation predicted all 12 classes successfully, while performance was similar to single-stage segmentation on the more prevalent LC classes (Table 2).

Finally, we used our multi-stage approach to predict the land cover of the whole PDNP (Fig A.1a). This LC map provides, to the best of our knowledge, the first very-high-resolution LC map of the PNDP using the LC schema designed for UK National Parks since 1991 (Taylor et al., 1991b). Fig A.1b demonstrates the high resolution of predicted LC for three 1 km^2 RGB image tiles.

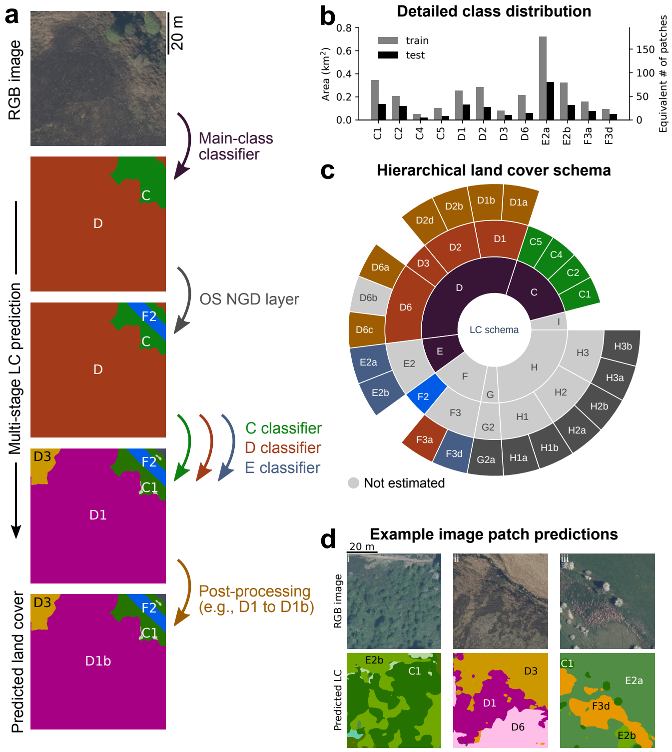

Figure 2: Multi-stage semantic segmentation approach, where **a)** land cover (LC) is predicted from RGB images in multiple stages: classification of the main class, overlaying the OS NGD layer, classification of the detailed subclasses, and post-processing to further split up subclasses (based on soil data). **b)** Distribution of data per detailed class, where the train and test data consisted of 70% and 30% of the data, respectively. **c)** The LC schema is hierarchical, and multiple classifiers and post-processing layers are used to predict all subclasses. Colours correspond to the steps in panel a. **d)** Three examples RGB image patches (top) and their predicted LC (bottom) are shown, demonstrating the final predictions by using the main class predictions as a mask for applying the combination of detailed class classifiers. LC codes are explained in Table A.1.

5 Conclusion and Discussion

We have created a data set of LC-annotated patches of RGB images that we have made publicly available online (Van der Plas et al., 2023a). This also includes an interpretation key with image examples and written descriptions of all classes (Alexander et al., 2023). We have annotated over 1000 patches of size 64 m $\times$ 64 m at the level of the detailed LC schema, spanning over 20 LC classes (Table A.1). Patches were sampled across the PDNP (Figure 1), which consequently led to variation in the flight dates (and associated variations in light conditions, seasonality effects on vegetation, time of day, etc.). This presented an additional challenge to the model and led to some misclassification because of spectral differences between some training and test data. This is an

Class name	Code	Sensitivity	Precision	Density	Classifier
Wood and Forest Land	C	0.91	0.96	24.2%	Main
Moor and Heath Land	D	0.97	0.93	34.6%	Main
Agro-Pastoral Land	E	0.93	0.93	41.3%	Main
Broadleaved High Forest	C1	0.73	0.92	9.9%	C
Coniferous High Forest	C2	0.99	0.78	9.9%	C
Scrub	C4	0.46	0.77	1.7%	C
Clear Felled/New Plantings in Forest Areas	C5	0.96	0.92	2.7%	C
Upland Heath	D1	0.84	0.80	10.8%	D
Upland Grass Moor	D2	0.70	0.82	8.8%	D
Bracken	D3	0.52	0.83	3.6%	D
Heather/Grass/Blanket Peat Mosaic	D6	0.64	0.48	4.6%	D
Improved Pasture	E2a	0.92	0.88	26.8%	E
Rough Pasture	E2b	0.67	0.76	10.6%	E
Wetland, Peat Bog	F3a	0.77	0.72	6.3%	D
Wetland, Wet Grassland and Rush Pasture	F3d	0.85	0.86	4.1%	E

Table 1: Results of the individual main, C, D and E classifiers on the test set.

Class name	Code	Sens. SS	Prec. SS	Sens. MS	Prec. MS
Broadleaved High Forest	C1	0.90	0.65	0.62	0.87
Coniferous High Forest	C2	0.81	0.89	0.99	0.80
Scrub	C4	Not pred.	Not pred.	0.38	0.61
Clear Felled/New Plantings in Forest Areas	C5	Not pred.	Not pred.	0.95	0.94
Upland Heath	D1	0.91	0.75	0.83	0.78
Blanket Peat Grass Moor	D2	0.55	0.55	0.67	0.72
Bracken	D3	Not pred.	Not pred.	0.48	0.75
Upland Heath/Blanket Peat Mosaic	D6	0.33	0.38	0.63	0.47
Improved Pasture	E2a	0.89	0.93	0.90	0.85
Rough Pasture	E2b	0.70	0.53	0.62	0.67
Wetland, Peat Bog	F3a	0.92	0.48	0.77	0.70
Wetland, Wet Grassland and Rush Pasture	F3d	Not pred.	Not pred.	0.67	0.73

Table 2: Effective sensitivity and precision values for single-stage (SS) and multi-stage (MS) classifiers. Not pred., not predicted.

inherent challenge of large-scale applications of LC prediction, and we hope that by making our train and test data sets publicly available, this can further be addressed by the broader research community. Given the large number of classes, the focus on natural LC instead of urban LC, the spatial distribution and the variety both within and between classes, we believe that this data set is a valuable resource to the community and is representative of much of the upland landscape found in the UK.

We then developed a multi-stage approach for classifying hierarchical LC schemes with large variations in the density of each class. Deconstructing the classification process into multiple steps achieved high accuracy on a large number of LC classes (95% accuracy on main classes, 92% on C, 72% on D and 87% on E), outperforming single-stage semantic segmentation for uneven class distributions. LC was predicted at high resolution (12.5 cm), enabling the identification of small habitat patches such as individual trees, heather patches and scrub (Fig 2d). The multi-stage approach was also able to handle complex cosmopolitan habitats such as wet grassland and rush pasture, which occurs both within moorlands and grasslands, by including it in more than one detailed classifier.

Our approach can be used to detect a wide range of habitats from the same aerial image data, from those with a broad species mix and mosaics to single species. This has wide-ranging applications in landscape ecology and biodiversity monitoring, especially in regions where important habitats are small and mixed. For example, these include fire risk modelling (Millin-Chalabi et al., 2023), climate change vulnerability assessments (Santos et al., 2021), tree planting planning (Peak District National Park Authority, 2022), biodiversity monitoring (Tuia et al., 2022; Rolnick et al., 2022) and habitat mapping (O'Connell et al., 2015). Finally, as the interest and demand for habitat corridors grows in the UK (UK Department for Environment, Food & Rural Affairs, 2018; Bailey et al., 2022), detailed LC maps will be crucial to comprehend the extent and role of these mosaics and collections of fragmented habitat patches (Van der Plas et al., 2023b).

Data and code availability

Data supporting this study (train and test data set of RGB images and land cover annotations) are openly available from the Cranfield Online Research Data (CORD) repository at `https://doi.org/10.17862/cranfield.rd.24221314`. All code, including example notebooks to load the data and models, is available at `https://github.com/pdnpa/cnn-land-cover`.

Acknowledgements

TLvdP was supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC) [grant number BB/M011224/1]. This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/Y028880/1], funding from The Alan Turing Institute via the internship scheme Turing Internship Network, and the Peak District National Park Authority.

References

David G Alexander, Thijs L Van der Plas, and Simon T Geikie. Interpretation Key of Peak District Land Cover Classes. https://reports.peakdistrict.gov.uk/interpretation-key/docs/introduction.html, 2023. (Accessed 2 April 2024).

J. J. Bailey, C. A. Cunningham, D. C. Griffin, G. Hoppit, C. A. Metcalfe, C. M. Schéré, T. J. P. Travers, R. K. Turner, J. K. Hill, P. Sinnadurai, R. Stafford, D. Allen, N. Isaac, B. Ross, D. Russi, B. Chamberlain, N. Harvey Sky, and S. McKain. Protected areas and nature recovery. achieving the goal to protect 30% of uk land and seas for nature by 2030, 2022. URL www.britishecologicalsociety.org/protectedareas.

Bluesky. Aerial Photography for Great Britain, © Bluesky International Limited and Getmapping Plc [2022], 2022. URL https://www.apgb.co.uk/.

Forestry Commission. National forest inventory woodland gb 2020. Forestry Commission Open Data Publication, last updated 30 June 2022., 2022. URL https://data-forestry.opendata.arcgis.com/datasets/eb05bd0be3b449459b9ad0692a8fc203_0/about.

Aaron E. Maxwell, Timothy A. Warner, and Luis Andrés Guillén. Accuracy assessment in convolutional neural network-based deep learning remote sensing studies—part 2: Recommendations and best practices. *Remote Sensing*, 13, 7 2021. ISSN 20724292. doi: 10.3390/rs13132591.

Gail Millin-Chalabi, Pia Labenski, Ana Maria Pacheco Pascagaza, Gareth Clay, and Fabian Ewald Fassnacht. Dynamic fuel mapping in the south pennines using a multitemporal intensity and coherence approach. In *European Space Agency-Fringe 2023*, 2023.

Natural England. Peaty soils location (england). https://naturalengland-defra.opendata.arcgis.com/datasets/1e5a1cdb2ab64b1a94852fb982c42b52_0/about (Accessed 2 February 2024), 2022.

Ordnance Survey. OS National Geographic Database ©, retrieved 10 january 2023., 2023. URL https://osdatahub.os.uk/.

Jerome O'Connell, Ute Bradter, and Tim G Benton. Wide-area mapping of small-scale features in agricultural landscapes using airborne remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 109:165–177, 2015.

Peak District National Park Authority. The wooded landscapes plan: increasing tree and scrub cover in the peak district national park landscapes (2022 – 2032). https://www.peakdistrict.gov.uk/__data/assets/pdf_file/0027/447255/Wooded-Landscapes-Plan-Final-Draft-July-22.pdf (Accessed 2 February 2024), 2022.

David Rolnick, Priya L Donti, Lynn H Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, et al. Tackling climate change with machine learning. *ACM Computing Surveys (CSUR)*, 55(2):1–96, 2022.

Maria J Santos, Adam B Smith, Stefan C Dekker, Maarten B Eppinga, Pedro J Leitão, David Moreno-Mateos, Naia Morueta-Holme, and Michael Ruggeri. The role of land use and land cover change in climate change vulnerability assessments of biodiversity: a systematic review. *Landscape Ecology*, pp. 1–16, 2021.

JC Taylor, Alan C Bird, TR Brewer, MA Keech, and MJ Stuttard. *Landscape Change in the National Parks of England and Wales: Final Report, Volume II Methodology*. Silsoe College, Bedford, UK, 1991a. ISBN 1 871564 16 6.

JC Taylor, Alan C Bird, MA Keech, and MJ Stuttard. *Landscape Change in the National Parks of England and Wales: Final Report, Volume I Main Report*. Silsoe College, Bedford, UK, 1991b. ISBN 1 871564 15 8. https://publications.naturalengland.org.uk/publication/5216333889273856.

JC Taylor, TR Brewer, and AC Bird. Monitoring landscape change in the national parks of england and wales using aerial photo interpretation and gis. *International Journal of Remote Sensing*, 21 (13-14):2737–2752, 2000. doi: 10.1080/01431160050110269.

Devis Tuia, Benjamin Kellenberger, Sara Beery, Blair R Costelloe, Silvia Zuffi, Benjamin Risse, Alexander Mathis, Mackenzie W Mathis, Frank van Langevelde, Tilo Burghardt, et al. Perspectives in machine learning for wildlife conservation. *Nature communications*, 13(1):792, 2022.

Devis Tuia, Konrad Schindler, Begüm Demir, Gustau Camps-Valls, Xiao Xiang Zhu, Mrinalini Kochupillai, Sašo Džeroski, Jan N van Rijn, Holger H Hoos, Fabio Del Frate, et al. Artificial intelligence to advance earth observation: a perspective. *arXiv preprint arXiv:2305.08413*, 2023.

UK Department for Environment, Food & Rural Affairs. A green future: Our 25 year plan to improve the environment. `https://www.gov.uk/government/publications/25-year-environment-plan` (Accessed 2 February 2024), 2018.

Thijs L. Van der Plas, Simon T. Geikie, David G. Alexander, and Daniel M. Simms. Very high resolution aerial photography and annotated land cover data of the Peak District [dataset]. `https://doi.org/10.17862/cranfield.rd.24221314`, 2023a. (Accessed 27 March 2024).

Thijs L Van der Plas, Simon T Geikie, David G Alexander, and Daniel M Simms. Multi-stage semantic segmentation quantifies fragmentation of small habitats at a landscape scale. *Remote Sensing*, 15(22):5277, 2023b.

François Waldner, Yang Chen, Roger Lawes, and Zvi Hochman. Needle in a haystack: Mapping rare and infrequent crops using satellite imagery and data balancing methods. *Remote Sensing of Environment*, 233:111375, 2019.

A APPENDIX

LC80-main	**LC80**	**LC20**	**Name**	**New?**
C (Wood and forest land)	C1	C1	Broadleaved high forest	-
C	C2	C2	Coniferous high forest	-
C	C4	C4	Scrub	-
C	C5	C5	Clear felled/newly planted trees	-
D (Moor and heath land)	D1	D1a	Upland heath	-
D	D1	D1b	Upland heath, peaty soil	Yes
D	D2b	D2b	Upland grass moor	-
D	D2d	D2d	Blanket peat grass moor	-
D	D3	D3	Bracken	..
D	D6a	D6a	Upland heath/grass mosaic	-
D	D6c	D6c	Upland heath/blanket peat mosaic	-
E (Agro-pastoral land)	E2a	E2a	Improved pasture	-
E	E2b	E2b	Rough pasture	-
F (Water and Wetland)	F2	F2	Open water, inland	-
F	F3a	F3a	Peat bog	-
F	D2/E2	F3d	Wet grassland and rush pasture	Yes
G (Rock and coastal land)	G2	G2	Inland bare rock	-
H (Developed land)	H1a	H1a	Urban area	-
H	H1b	H1b	Major transport route	-
H	H2a	H2a	Quarries and mineral working	-
H	H2b	H2b	Derelict land	-
H	H3a	H3a	Isolated farmsteads	-
H	H3b	H3b	Other developed land	-
I (Unclassified land)	I	I	Unclassified land	-

Table A.1: Land cover schema adapted from (Taylor et al., 1991b;a; 2000). LC80 is the original schema, LC20 is the updated schema that we have created. LC80-main denotes the main class. Only classes that are present in PDNP are included.

Classifier	**Loss function**	**Mean**	**Std**	**Max**	**Selected?**
C	Cross entropy	0.90	0.01	0.92	Yes
C	Focal loss	0.87	0.10	0.93	-
D	Cross entropy	0.67	0.03	0.72	Yes
D	Focal loss	0.70	0.02	0.72	-
E	Cross entropy	0.85	0.02	0.87	Yes
E	Focal loss	0.83	0.02	0.84	-
Main	Cross entropy	0.93	0.02	0.95	Yes
Main	Focal loss	0.92	0.01	0.94	-
Single-stage	Cross entropy	0.67	0.04	0.71	N/A

Table A.2: **Average CNN performance.** The loss function was varied for each of the four classifiers (main, C, D, E). Five runs were performed for each model, for 60 epochs, saving the model at the best epoch (in terms of train loss). Afterwards, the mean, standard deviation (std) and maximum (max) accuracy on test data were computed for each classifier. Focal loss was used with $\gamma = 0.75$. The 'Selected?' column indicates whether the maximum-performing model of that classifier type was used in further analyses. Additionally, five runs were performed for the single-stage model, which learns to classify the detailed classes directly, using the cross entropy loss function.

A.1 SELECTING IMAGE PATCHES FOR TRAINING AND TESTING

The image data were available as 1 km $\times$ 1 km *tiles*, which were split into 64 m $\times$ 64 m (i.e., 512 pixels $\times$ 512 pixels) *patches* for input to the CNN models. To sample patches evenly across

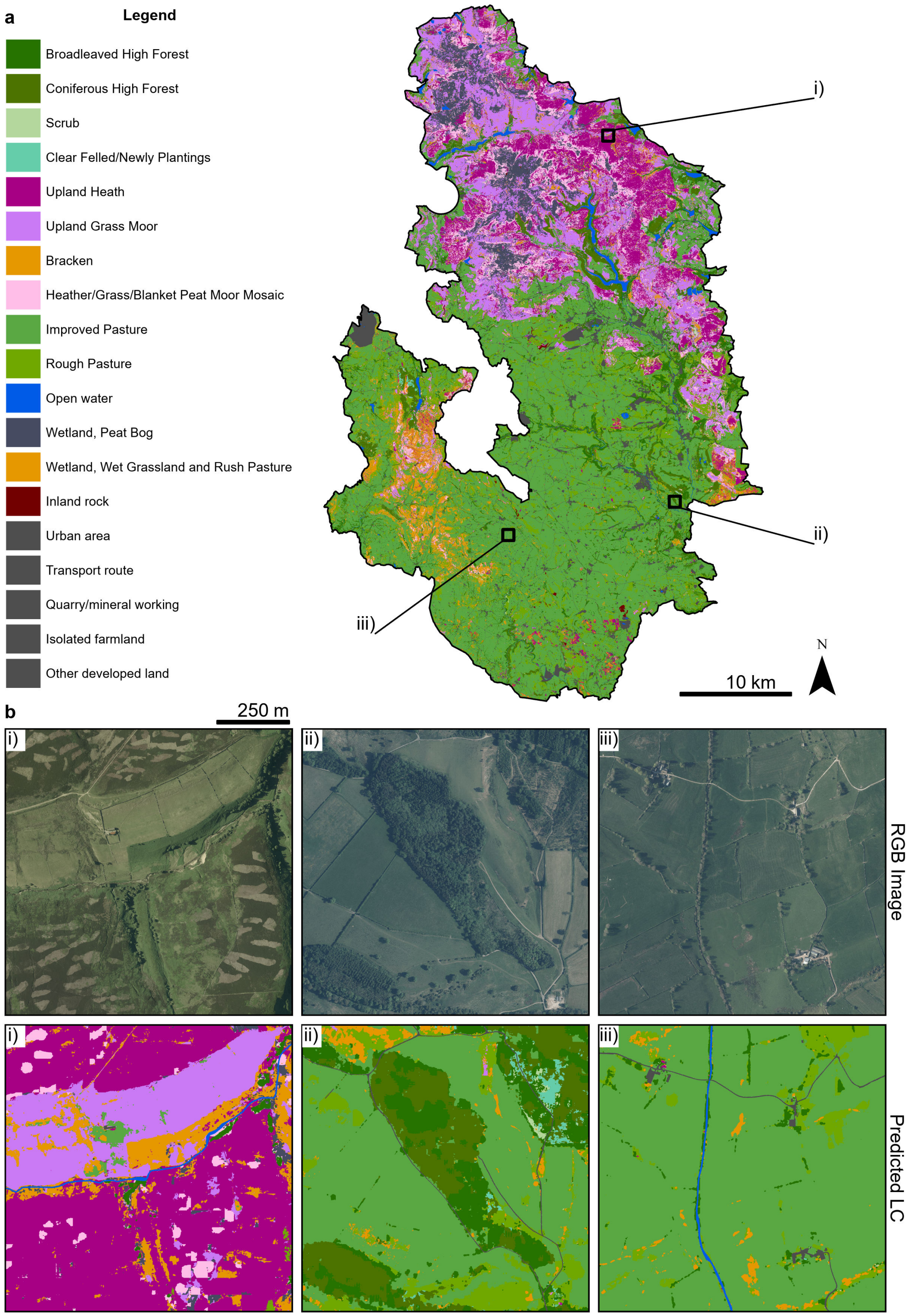

Figure A.1: **a)** Land cover predictions of entire Peak District National Park (1,439 km^2) at 12.5 cm resolution. **b)** Close-ups of the three insets shown in panel a). The RGB images (top) and model predictions (bottom) of three 1 km x 1 km example tiles are shown.

both space and land cover, while ensuring sufficient samples of the rare classes, we selected image patches across the PDNP with the following procedure:

First, we used the 1991 census data (Taylor et al., 1991b) to select 50 tiles (of 1 km^2 each) that were representative of the overall LC distribution (of 1991). To do so, we generated 50,000 random samples of 50 tiles, computed the L1 loss of the LC area distribution of the sample compared to the LC area distribution of the entire PDNP, and selected the sample with the lowest L1 loss. This resulted in a sample of 50 tiles that was spatially distributed evenly across the PDNP. From each of these tiles, we randomly selected nine image patches (of 64 m $\times$ 64 m)—one from each block of a 3 $\times$ 3 grid across the tile—resulting in 450 image patches. This approach was used to prevent bias in sample selection and to ensure the accuracy metrics used for validation were representative of the final mapped outputs Maxwell et al. (2021). However, some classes are much more prevalent than others, meaning that rare classes were very unlikely to be sampled sufficiently for model training. Therefore, an additional 577 patches were selected manually across the same 50 tiles, plus 30 extra tiles were selected to boost rare classes (Figure 1c).

A.2 New LC class F3d

The new LC class wet grassland and rush pasture (F3d) posed a challenge for the classifiers, as it typically occurs in small patches both within moorlands (D) and grasslands (E). As CNNs rely on the context of the RGB image for classification, these different types of habitat surroundings were initially found to confuse the CNN models. Therefore, we decided to: (1) include F3d as a category in *both* the detailed D classifier and the detailed E classifier and (2) for the purpose of training the main classifier only, remap any F3d polygons to D class. In other words, to the CNN classifiers, F3d was presented as a subclass of D (moorlands) while allowing the possibility to classify E (grasslands) into F3d given its presence in grasslands too.

A.3 Merger with OS Layer for Developed Land

Ordnance Survey (OS) data were used to map the water (F2), rock (G2) and developed land (H) classes, as these had already been accurately and recently mapped by Ordnance Survey (2023). After the main classifier predicted the main class of the land cover, these predictions were overwritten by the OS layer (*i.e.,* areas that contained OS polygons replaced the model-predicted polygons, Fig 2a). To do so, OS polygon classes were relabelled to our LC schema (the relabelling key is available online). Quarry (H2a) OS annotations were found to be inaccurate, and therefore, they were all manually verified and deleted if necessary.

A.4 Post-Processing of Model Predictions

Some detailed classes were distinguished based on secondary soil data (Fig 2a). Specifically, some D classes had peat-soil and non-peat-soil variants (D1a and D1b, D2b and D2d, and D6a and D6c). To identify these, the model predicted D1, D2 and D6 generally, and predictions were subsequently labelled as peat/non-peat based on the 'Peaty Soils Location' data set from Natural England (Natural England, 2022). For each predicted D1, D2 and D6 polygon, the intersection with the peaty soils layer polygons was computed, and it was then assigned the peat label if the intersection was greater than 50% of the area of the predicted polygon.

However, model predictions are reported and shown without this post-processing step, as it is a deterministic separation of some classes that does not change the performance or accuracy of any class (but does create extra classes).

A.5 Single-Stage Semantic Segmentation

For comparison with the multi-stage models, detailed LC classes were also predicted directly using conventional single-stage semantic segmentation. U-Nets were trained using exactly the same protocols and parameters as previously described. Again, five networks were trained, and the best-performing network was selected for further analysis (Table A.2). Networks were trained to predict the detailed LC classes directly. (Only those that need to be model-predicted, *i.e.*, from Table 2).