Received: 29 May 2022 | Accepted: 1 July 2022

DOI: 10.1111/2041-210X.13984

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

APPLICATION

imageseg: An R package for deep learning-based image segmentation

Jürgen Niedballa[1] | Jan Axtner[1] | Timm Fabian Döbert[2] | Andrew Tilker[1,3] | An Nguyen[1] | Seth T. Wong[1] | Christian Fiderer[4,5] | Marco Heurich[4,5,6] | Andreas Wilting[1]

[1]Department of Ecological Dynamics, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany; [2]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada; [3]Re:wild, Austin, Texas, USA; [4]Bavarian Forest National Park, Grafenau, Germany; [5]Albert-Ludwigs Universität Freiburg, Freiburg, Germany and [6]Inland Norway University of Applied Science, Koppang, Norway

Correspondence
Jürgen Niedballa
Email: niedballa@izw-berlin.de

Funding information
Bundesamt für Naturschutz, Grant/Award Number: 3518830200; Bundesministerium für Bildung und Forschung, Grant/Award Number: 01LN1301A; Commonwealth Scientific and Industrial Research Organisation; Deutsche Forschungsgemeinschaft, Grant/Award Number: 491292795; Leibniz Institute for Zoo and Wildlife Research (IZW) in the Forschungsverbund Berlin eV Berlin, Berlin, DE; Point Defiance Zoo and Aquarium; United States Agency for International Development, Grant/Award Number: 72044020CA00001; University of Western Australia

Handling Editor: Sarah Goslee

Abstract

1. Convolutional neural networks (CNNs) and deep learning are powerful and robust tools for ecological applications, and are particularly suited for image data. Image segmentation (the classification of all pixels in images) is one such application and can, for example, be used to assess forest structural metrics. While CNN-based image segmentation methods for such applications have been suggested, widespread adoption in ecological research has been slow, likely due to technical difficulties in implementation of CNNs and lack of toolboxes for ecologists.
2. Here, we present R package imageseg which implements a CNN-based workflow for general purpose image segmentation using the U-Net and U-Net++ architectures in R. The workflow covers data (pre)processing, model training and predictions. We illustrate the utility of the package with image recognition models for two forest structural metrics: tree canopy density and understorey vegetation density. We trained the models using large and diverse training datasets from a variety of forest types and biomes, consisting of 2877 canopy images (both canopy cover and hemispherical canopy closure photographs) and 1285 understorey vegetation images.
3. Overall segmentation accuracy of the models was high with a Dice score of 0.91 for the canopy model and 0.89 for the understorey vegetation model (assessed with 821 and 367 images respectively). The image segmentation models performed significantly better than commonly used thresholding methods, and generalized well to data from study areas not included in training. This indicates robustness to variation in input images and good generalization strength across forest types and biomes.

4. The package and its workflow allow simple yet powerful assessments of forest structural metrics using pretrained models. Furthermore, the package facilitates custom image segmentation with single or multiple classes and based on colour or grayscale images, for example, for applications in cell biology or for medical images. Our package is free, open source and available from CRAN. It will enable easier and faster implementation of deep learning-based image segmentation within R for ecological applications and beyond.

KEYWORDS
canopy density, canopy hemispherical photography, computer vision, convolutional neural network, forest monitoring, machine learning, UNet, vegetation density

1 | INTRODUCTION

Technological advances over the last few decades have led to unprecedented increases in the amount of data available for ecological research, necessitating efficient and reproducible analytical workflows (Christin et al., 2019). Methodological developments in machine learning paired with improvements in computational power have enabled an increasingly widespread use of deep learning methods to make use of these data in a multitude of ecological applications, particularly for visual recognition (Kattenborn et al., 2021).

Convolutional neural networks (CNN) have been at the core of many of these innovations for their ability to analyse complex nonlinear data and performance at near-human or even exceeding human capabilities in specific tasks (Grace et al., 2018; Mnih et al., 2015). CNNs are multilayered artificial neural networks that consist of an input layer, stacked hidden layers (e.g. convolutional and pooling layers) and the final output layer (Minaee et al., 2021). The number of the hidden layers defines the depth of the network. Convolutional layers use filters to set every pixel in relation to its surrounding pixels, producing so-called feature maps. Feature maps are then downsampled and generalized using pooling layers (LeCun et al., 2015). Pooling reduces the dimensions of feature maps and therefore the number of parameters to learn. Convolutional layers introduce translation equivariance, whereas pooling layers introduce translation invariance.

One prominent domain of CNNs is image segmentation, which classifies each pixel of an image, providing locality information for each label class. Here, we primarily focus on two ecological applications of image segmentation in the context of forest structural metrics from colour photographs, the assessment of (i) tree canopy density (canopy closure and canopy cover) and (ii) understorey vegetation density. Moreover, we provide an outlook on how the package imageseg can be used for grayscale images or more complex multiclass applications.

Canopy density and understorey vegetation density are important metrics of forest structure and are therefore important for a number of ecological applications from biodiversity surveys to forest monitoring and ground-truthing remote sensing data (Dupuy & Chazdon, 2008; Latifi et al., 2016). Canopy density is related to canopy architecture, forest condition, carbon and stand density, and affects light regime and microhabitat structure (Jennings et al., 1999). Understorey vegetation is composed of immature trees, shrubs and herbaceous vegetation, and is thus linked to forest regrowth and plant succession (e.g. after disturbances), while also providing essential resources and shelter for wildlife (McShea & Rappole, 1992; Nilsson & Wardle, 2005). Both measures are thus related to carbon sequestration and storage, biodiversity conservation and ecosystem regeneration. Furthermore, due to their importance for forestry and agriculture, both measures are also economically relevant.

Traditionally, tree canopy density was often assessed using manual methods such as spherical densiometers or photometric techniques requiring specialist equipment (see Paletto & Tosi, 2009 for comparison of methods). Traditional methods for quantification of understorey vegetation density include vegetation density boards or profile boards (Nudds, 1977). These methods are slow, labor-intensive and often require special equipment. Moreover, manual approaches are rarely reproducible, can be subject to observer bias and dependent on experience. Remote sensing technologies such as airborne or terrestrial laser scanning have been successfully used to determine forest structural attributes (Campbell et al., 2018; Latifi et al., 2016; Li et al., 2021; Wing et al., 2012; Zong et al., 2021). These methods, however, whilst powerful and accurate, remain constrained by logistical challenges and often prohibitive costs (Hummel et al., 2011).

Deep learning methods have been developed to help address these shortcomings and streamline analyses of forest structural metrics using easy-to-collect digital photographs (Abrams et al., 2019; Díaz et al., 2021; Li et al., 2020). The approaches outlined in Abrams et al. (2019) and Li et al. (2020) apply image segmentation based on the U-Net architecture introduced by Ronneberger et al. (2015), whereas the model of Diaz et al. (Díaz et al., 2021) uses deep learning regression to assess plant area index (PAI) and is thus not an image segmentation approach (i.e. it does not retain pixel-level information). These deep learning methods achieve higher accuracy than manual methods whilst being much faster and more cost-efficient (Abrams et al., 2019; Li et al., 2020). Uptake of deep learning methods in ecological applications, however, has been constrained by technical difficulties in implementation and the lack of a readily available workflow or software package to easily apply these methods. Furthermore, training datasets are often relatively homogeneous (i.e. from a single forest ecosystem), potentially impairing applicability of such models is across ecosystems.

Here, we present the R package imageseg, which implements a complete workflow for binary and multiclass image segmentation using deep learning models implemented in TensorFlow (Abadi et al., 2016). It works with grayscale as well as colour input images. We illustrate the workflow with new models for two forest structural metrics, canopy density and understorey vegetation density. Both models were built using the largest and most diverse dataset for these applications to date. The canopy density model is the first to be explicitly trained on both canopy closure and canopy cover images. Both models yielded highly accurate predictions. The pretrained models for canopy and understorey vegetation density, along with the simple workflow of imageseg, will enable faster and more reproducible assessment of these important forest structural metrics from easily obtainable data.

The imageseg R package presents a critical advancement in deep learning-based image segmentation. While imageseg and the pretrained models are particularly relevant to applied research in forestry, ecology and conservation, we believe the accessibility and versatility of the package and workflow can facilitate the development of custom image segmentation workflows in other fields such as microscopy or cell biology.

2 | PACKAGE DESCRIPTION

2.1 | Overview

We present imageseg, an R package which implements a general-purpose image segmentation workflow based on convolutional neural networks using the U-Net architecture (Ronneberger et al., 2015) or the U-Net++ architecture (Zhou et al., 2018). The package is suitable for binary and multiclass image segmentation for both grayscale and colour images. Models are implemented in R via Keras (Allaire & Chollet, 2021; Chollet et al., 2015) using a TensorFlow backend (Abadi et al., 2016).

The workflow covers data processing, model training and predictions for image segmentation tasks applied to digital images. In total, the package contains eight main functions for image processing, model creation, predictions, as well as the extraction of relevant information from predictions (see Table 1). The package workflow is explained in a detailed vignette and illustrated in Figure 1.

Our package is intended for two main purposes:

1. General purpose model training for image segmentation, including both de novo training with custom model architectures and continued training of existing models.
2. Model predictions using pretrained models that allow users to easily perform image segmentation on digital images. We provide pretrained models for two types of forest structural metrics (tree canopy density and understorey vegetation density) along with respective training data for download.

Here, we demonstrate the utility of the package for image segmentation and illustrate its application for forest structural metrics, namely a model for canopy density (differentiating canopy and sky) and a model for understorey vegetation density. Both of these applications perform binary classifications of all pixels in input images. In addition, the imageseg package also supports multiclass image segmentation (see the example showing the detection of bacteria, erythrocytes and background in dark-field microscopy images in Figure 2).

2.2 | Data processing

The imageseg package accepts jpg, tif and png images as input. Image processing in R is mainly performed by the magick package (the R interface to ImageMagick, Ooms (2021)), allowing for easy viewing of images in R.

Data processing is divided into a logical sequence of steps, implemented in select functions (see Table 1 and Figure 1). *resizeImages* resizes original images to the dimensions expected by the models while also allowing flexibility to prepare input for custom models. The function can optionally crop images or remove uninformative image

TABLE 1 List of functions in the imageseg package

Function	Description
resizeImages	Resize and save images
findValidRegion	Subset image to valid (informative) region
loadImages	Load image files with magick
dataAugmentation	Image data augmentation: rotating and mirroring images, adjusting colours
imagesToKerasInput	Convert magick images to keras input (4D arrays)
u_net	Create a U-Net architecture
u_net_plusplus	Create a U-Net++ architecture
loadModel	Load TensorFlow model and custom objects from hdf5 file
imageSegmentation	Model predictions from images based on TensorFlow model

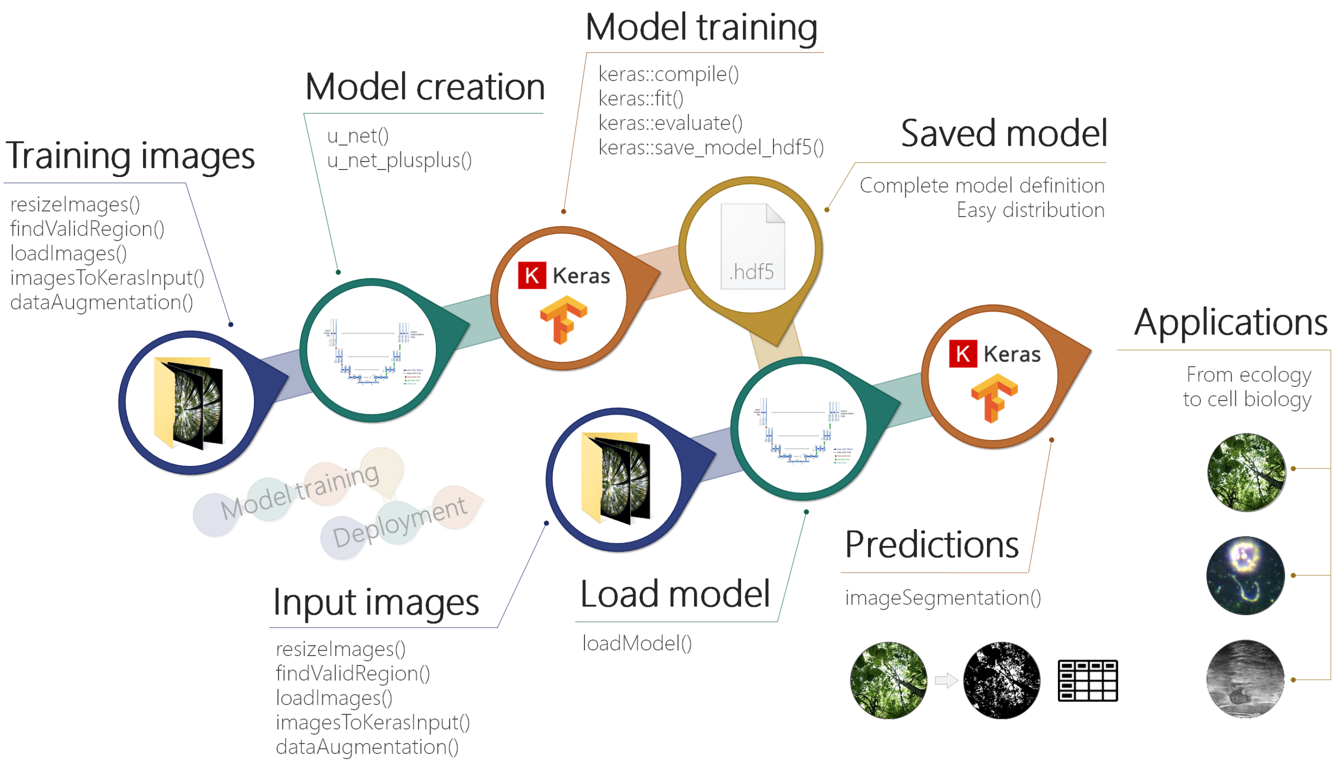

FIGURE 1 Workflow of the R package imageseg showing the main steps and their relevant functions. The steps at the top show the model training workflow and the steps at the bottom show the model deployment workflow.

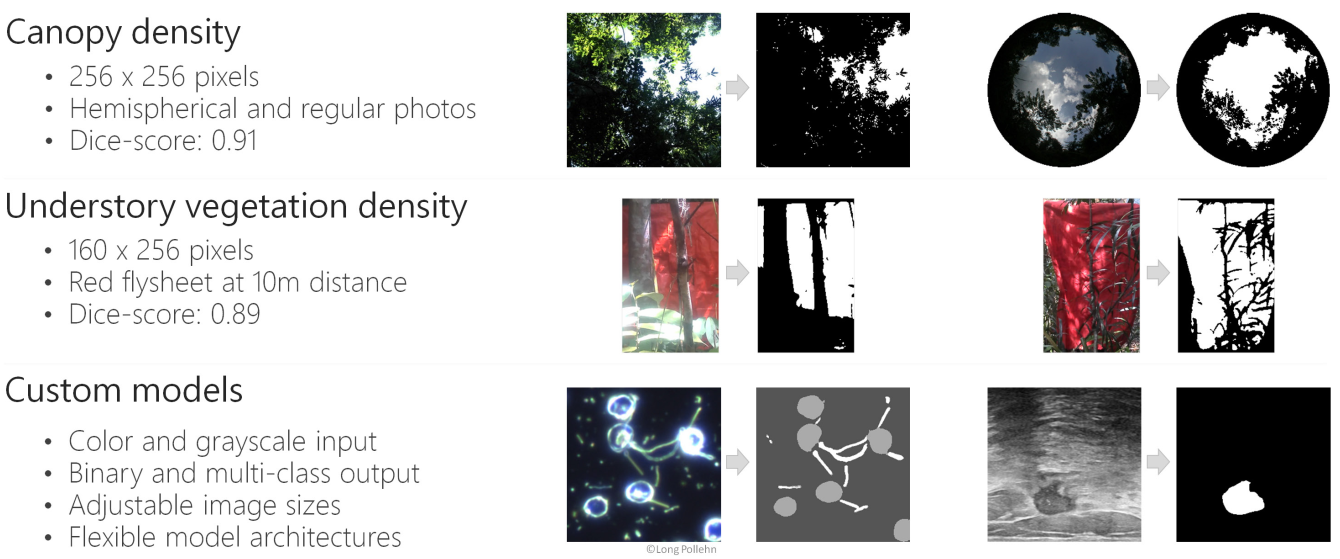

FIGURE 2 Example input and output of imageseg models, with two examples per model type (input left, output right). The canopy and understorey vegetation density models convert colour images to binary predictions of sky/vegetation and red flysheet/vegetation respectively. The package workflow is furthermore suitable for custom image segmentation tasks. The examples at the bottom show a multiclass image segmentation of a (colour) microscopy image, and a binary segmentation of a (grayscale) ultrasound breast cancer scan image (see package vignette for code and data sources).

regions. The *findValidRegion* function can help identify the valid (informative) image region for cropping of images (e.g. for removing uninformative black borders from hemispherical canopy photos taken with fisheye lenses) by converting images to grayscale, thresholding and identifying informative rows/columns above the threshold. *loadImages* loads images into R as magick objects, allowing for easy viewing of images in R. *dataAugmentation* performs optional and adjustable data augmentation (rotation, mirroring, modulating brightness, saturation and hue) to increase the amount and variability of available training data and thus the robustness and invariance of the

CNN (only relevant for model training, not for model predictions). *imagesToKerasInput* converts the images to model input required by the CNN (4D array with dimensions: image–row–column–channel).

2.3 | Model training

Users who wish to train their own model can create a U-Net-based architecture with the *u_net* function, which was slightly modified from the original function in the platypus package (Maj, 2020). It supports flexible image sizes, data types, number of U-Net blocks, layers per block, convolution filters, the optional inclusion of dropout layers and supports binary and multiclass classifications (thus also allowing general purpose image segmentation as shown in Figure 2).

Furthermore, users can create U-Net++ architectures with the *u_net_plusplus* function. It was ported from Python code on GitHub (Sokol, 2022).

As mentioned above, the *dataAugmentation* function can perform various data augmentation tasks prior to model training to increase the amount and diversity of available training data. Users who wish to only apply pretrained models can skip the model training steps.

2.4 | Model predictions

loadModel loads a pretrained model from a hdf5 file (containing the entire architecture and training state of the model in a single file) and ensures the relevant custom objects are loaded (custom loss functions and metrics). *imageSegmentation* is the main function for predictions and performs image segmentation using the model and images provided by the user. It returns the pixelwise probabilities for the predictions, the classified image segmentation masks based on the input images, and summary tables. In binary predictions (such as the vegetation structure models we present), these summaries represent the desired vegetation metrics (openness/gap fraction) along with their inverse (canopy closure/cover, or understorey vegetation density). It allows for adjustable masking of hemispherical canopy images (e.g. circular masks to exclude uninformative corners in hemispherical images). In models with multiple output classes, it returns classified image segmentation masks, probability layers for each class, and summarizes prevalence and probabilities of all predicted classes.

3 | VEGETATION STRUCTURE MODELS

We demonstrate the utility of the imageseg package for image segmentation of two types of forest structural metrics: a model for tree canopy density (differentiating canopy and sky in canopy cover images and hemispherical canopy closure images) and a model for understorey vegetation density. Both models use colour photographs as input and were trained on the largest and most diverse training dataset for these forest structural metrics to date.

Here, we follow Jennings et al. (1999) in defining canopy closure as the proportion of sky hemisphere obscured by vegetation when viewed from a single point. It is commonly recorded using canopy hemispherical photography (CHP), for example, via fisheye lenses. In contrast, canopy cover is the percent forest floor occupied by the vertical projection of tree crowns. It is recorded using canopy cover photography (CCP) (Chianucci, 2016; Macfarlane et al., 2007). While cameras cannot technically record the exact vertical projection of tree crowns (due to the camera sensor being essentially a single point), the narrower field of view and higher focal length ensure the canopy cover images approximate actual canopy cover well.

We define understorey vegetation density analogously to canopy cover as the percent area covered by the lateral obstruction from understorey vegetation. It is assessed using a red flysheet of 1 × 1.5 m held at a distance of 10 m from the camera. The red background colour was chosen to provide a strong colour contrast with vegetation. The method is similar to the cover board photography method in Campbell et al. (2018).

3.1 | Model architecture and training

Both models (canopy and understorey vegetation density) share the same underlying U-Net architecture (Ronneberger et al., 2015). They use colour images as input and only differ in the dimensions of the input images. Both models perform binary classification with one output class. The model architecture is a standard U-Net with 32 feature maps in the first convolution (the number of feature channels doubles at every downsampling step along the contracting path). We chose U-Net over more recent or more complex architectures like U-Net++ because it is still a capable and flexible model architecture which is well suited to the relatively simple segmentation tasks at hand and has lower memory requirements during training than U-Net++ models of comparable size.

During model training, we used BCE-Dice loss (sum of binary cross-entropy and soft-Dice loss) and the 'Adam' model optimization algorithm (Kingma & Ba, 2017) with the following parameters: initial learning rate = 0.001, Beta_1 = 0.9, Beta_2 = 0.999, decay rate = 0. To avoid overfitting, we reduced the learning rate by a factor of 5 when validation loss stopped decreasing (with patience 3), employed an early stopping strategy with patience 5 and used the model weights from the epoch with the lowest validation loss for the final model (Caruana et al., 2000). We used a batch size of 12 for the canopy model and 16 for the understorey model. Maximum number of epochs was set to 100, but early stopping led to shorter model training.

In all models, 70% of the original images were used for training, 10% for model validation during training and 20% of images were used as test data for model evaluation. Training images were subject to data augmentation to increase robustness of the network and invariance (e.g. to lighting conditions and image geometry). Model accuracy was evaluated using two similarity scores, Dice similarity coefficient and Jaccard index. See Table 2 for details on model

	Canopy model	Understorey vegetation model
Original images	4109 (CHP: 1310, CCP: 2799)	1835
Training images[a]	2877 (17262)	1285 (2570)
Validation images[b]	411	183
Test images[b]	821	367
Image size (W × H)	256 × 256	160 × 256
Data augmentation		
Rotation	90°, 180°, 270°	—
Mirroring	Horizontal and vertical axes	Vertical axis
Dice score	0.91 (CHP: 0.91, CCP: 0.91)	0.89
Jaccard index	0.86 (CHP: 0.85, CCP: 0.86)	0.84

[a]Before (after) data augmentation.

[b]Without data augmentation (only original images).

TABLE 2 Overview of training data and results of the canopy density and understorey vegetation density models

training and results. We assessed the generalization strength of the canopy and understorey models by holding out data from individual study areas, fitting models with the remaining data and evaluating model performance against the withheld data (see Supporting Information S3 for details).

Training was conducted in R 4.1.2 with keras 2.4.3 and TensorFlow 2.2.0 on a Windows workstation with 128 Gb of RAM and an Nvidia GeForce GTX1660 Super GPU.

3.2 | Training data

Compared to previous approaches, our models use a larger and more diverse training dataset. Both canopy and understorey data have broader geographical scope by including training data from Viet Nam, Laos, Sabah (Malaysian Borneo), and in the case of CCP data, Germany. The canopy model was trained on CCP and CHP images. Images were taken under varying illumination conditions. Image masks were created manually in the GNU Image Manipulation Program (GIMP; The GIMP Development Team, 2019). See Table 2 and Supporting Information S1 for more details on the training data used in both models and their processing.

3.2.1 | Canopy model

For the canopy model, we used a total of 4109 canopy photographs (and their respective segmentation masks). Of these, 1310 were hemispherical CHP images (from Sabah, Malaysian Borneo) and 2799 were CCP images (from Malaysia, Viet Nam, Laos and Germany). Sixty-nine of the CCP images were sky photos for which we assessed segmentation accuracy separately.

CHP images cover multiple habitat types along a gradient from primary tropical rainforest, over secondary forest subjected to varying degrees of logging, to oil palm plantations within the Stability of Altered Forest Ecosystems (SAFE) project area in Sabah, Malaysian Borneo (Ewers et al., 2011). CCP images cover primary and secondary tropical rainforests in three countries (along an elevational gradient from 60 m to 1420 m a.s.l.), and also include temperate coniferous and mixed mountain forests of the Bavarian Forest National Park, Germany (covering an elevational gradient from 600 to 1450 m a.s.l; Cailleret et al., 2014).

All canopy images were resized to 256 × 256 pixel resolution while preserving the original aspect ratio (thus slightly cropping the rectangular original images to obtain square images). We increased the number of images available for training and validation via data augmentation, which included image rotation (90, 180 and 270 degrees) and mirroring along the horizontal and vertical axes, resulting in a total of 17,262 images for training and 411 for validation (see Table 2). We used 821 unaugmented images for model evaluation.

For reference, we compared the results of the canopy model with threshold-based image segmentation results from the R packages image.Otsu and coveR (Chianucci et al., 2022; Wijffels et al., 2020).

3.2.2 | Understorey model

For the understorey vegetation density model, we used 1835 images and their respective segmentation masks (866 from Malaysia, 842 from Viet Nam and 127 from Laos). Images were taken of a 1 × 1.5 m red flysheet held at a distance of 10 m from the cameras and cropped to contain only the extent of the flysheet (and the vegetation covering it). Cropping could not be automated due to high variation in images and was thus done manually (the need for cropping can be avoided if images are taken zoomed in and the flysheet fills the image frame). All understorey images were resized to 160 × 256 pixels (aspect ratio 5:8). Data augmentation included mirroring images along the vertical axis (to preserve image orientation), resulting in a total of 2570 images for training and 183 images for validation. Model evaluation was done with 367 unaugmented images (see Table 2).

3.3 | Model results

The canopy model achieved an overall Dice similarity coefficient of 0.91 and Jaccard index of 0.87. The respective scores were 0.91 and 0.85 for CHP images, and 0.91 and 0.86 for CCP images. Both indices were 1 for sky photos. These scores are considerably higher than those achieved in the image segmentation methods implemented in the R packages coveR and image.Otsu, which had overall Dice coefficients of 0.80 and 0.66 and Jaccard indices 0.66 and 0.49 respectively. See Supporting Information S4 for more details on the method comparison.

The understorey vegetation density model achieved an overall Dice coefficient of 0.89 and Jaccard index of 0.84 (Table 2). See Figure 2 and Supporting Information S2 for additional examples of model input and output of both models.

The models in which individual study areas were withheld as a test of generalization strength performed well on the withheld data, with Dice scores only 0.02–0.03 units lower than in the global model, even if the model was only trained with tropical rainforest images and predicted to images from temperate forests in Germany. Even for CHP data, Dice score was only 0.10 lower, despite this model only being trained on CCP data. This suggests overall good generalization strength to unseen data, tolerance to variable image geometry and no overfitting. See Supporting Information S3 for details.

3.4 | Performance

Model predictions can be created quickly with standard hardware and on CPU alone. A laptop with an Intel i7-7500U CPU created predictions for about 100 canopy images per minute. A workstation with a dedicated graphics card created predictions for >1000 images per minute.

4 | APPLICATION EXAMPLES

Our pretrained forest structure models are suitable for quickly deriving habitat metrics from easy-to-collect data using powerful CNNs. The diversity of the model training datasets allows for greater flexibility and generalization strength than previous models. They can be applied to record canopy cover and vegetation density in biodiversity studies (Asad et al., 2020; Tilker et al., 2020), for ground-truthing in remote sensing applications (Campbell et al., 2018; Hansen et al., 2019), or in monitoring of forest pathology, forest conditions or stresses (e.g. under drought conditions). Researchers in crop sciences might also find such models helpful for assessments of crop growth using photographic data.

Furthermore, the package and workflow are not restricted to forest structural metrics and binary classifications, but are universal and can also be applied to image segmentation tasks in other domains, including for multiclass image segmentation (e.g. in microscopy or cell biology, see Figure 2). The package and its workflow were designed to be accessible and simple to use in order to facilitate simple and quick experimentation without the need for in-depth technical knowledge about CNNs.

The generalization strength of CNNs ensures that the methods are robust and not dependent on optimal lighting conditions, thus allowing greater flexibility in data collection, for example, under field conditions in which waiting for optimal conditions is often not an option, or for cellular and microscopic images which are often highly variable.

5 | CONCLUSION

This paper presents an R package for image segmentation using deep convolutional neural networks based on the powerful and flexible U-Net architecture (Ronneberger et al., 2015). The package implements a complete workflow including image processing, model creation and prediction. It was designed to be accessible and provide a simple and intuitive user experience that does not require specialist-level technical understanding of deep learning models and CNNs. imageseg constitutes a major advancement toward making deep learning-based image classification available to non-specialists in a convenient and easy-to-use R package.

We demonstrate the usefulness of the package for deriving forest structural metrics (canopy and understorey vegetation density) from colour photographs and provide pretrained models for these metrics. The models were trained on large and diverse training datasets and achieved high segmentation accuracy (Dice similarity coefficient of 0.91 for canopy images and 0.89 for understorey images). Beyond this immediate application, the workflow and functions we provide are flexible and allow for general purpose image segmentation, making the package a versatile tool for simple implementation of complex image segmentation workflows in R.

AUTHOR CONTRIBUTIONS

Jürgen Niedballa, Jan Axtner and Andreas Wilting conceptualized the idea. Jürgen Niedballa and Jan Axtner developed the package and conducted model training. Timm Fabian Döbert provided the canopy hemispherical photography data. Andrew Tilker, An Nguyen and Seth T. Wong provided the canopy cover and understorey vegetation training data from Sabah, Vietnam and Laos. Christian Fiderer and Marco Heurich provided canopy cover training data from Bavarian Forest National Park. All authors contributed to writing the paper.

ACKNOWLEDGEMENTS

We thank the Sabah Biodiversity Center for issuing research permits and the Sabah Forestry Department for its support. We thank Azlan bin Mohamed, John Mathai, local research assistants and others involved in fieldwork for data collection and processing of canopy cover and understorey photos from Sabah. In Vietnam and Laos, we thank the staff of the WWF-CarBi project for providing extensive logistical support; Bach Ma NP, the Thua Thien Hue and

Quang Nam Saola NRs and Xe Sap NPA for providing permissions and personnel to conduct this research; and our field team leaders. This project received financial support from the German Federal Ministry of Education and Research (BMBF FKZ: 01LN1301A), the United States Agency for International Development (USAID) through Cooperative Agreement No. 72044020CA00001 'USAID Biodiversity Conservation' Activity, Point Defiance Zoo and Aquarium, San Francisco Zoo and Leibniz-IZW. T.F.D. was supported by a University of Western Australia PhD scholarship and a CSIRO top-up scholarship. We thank the SAFE project, South East Asia Rainforest Research Programme (SEARRP) and IDEA Wild for logistical support and equipment. Marion Pfeifer assisted with canopy hemispherical photo processing. Data collection strongly benefitted from local traditional knowledge. We thank the German Federal Agency for Nature Conservation for funding research in Bavarian Forest National Park (FKZ: 3518830200). The Bavarian Forest National Park thanks all field assistants who took the canopy pictures and processed them. We are grateful to Oliver R. Wearn and Adam F. Smith for their networking efforts which connected the collaboration partners in this project. Open Access funding was enabled and organised by Projekt DEAL and was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation—Projektnummer 491292795).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

imageseg is a free and open source R package. The latest release version is available on CRAN (https://CRAN.R-project.org/package=imageseg). A tutorial vignette is included in the package. Source code and the development version are available from GitHub (https://github.com/EcoDynIZW/imageseg). The package version used in model training is available from Zenodo at https://doi.org/10.5281/zenodo.7041052 (Niedballa, 2022). Pretrained models, classification examples, R code and data for model training and testing are available for download. Data deposited in the Dryad repository at https://doi.org/10.5061/dryad.x0k6djhnj (Niedballa et al., 2022).

ORCID

Jürgen Niedballa https://orcid.org/0000-0002-9187-2116
Jan Axtner https://orcid.org/0000-0003-1269-5586
Timm Fabian Döbert https://orcid.org/0000-0002-1601-8665
Andrew Tilker https://orcid.org/0000-0003-3630-8691
An Nguyen https://orcid.org/0000-0003-0456-3866
Seth T. Wong https://orcid.org/0000-0001-8083-9268
Christian Fiderer https://orcid.org/0000-0001-9706-6265
Marco Heurich https://orcid.org/0000-0003-0051-2930
Andreas Wilting https://orcid.org/0000-0001-5073-9186

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2016). TensorFlow: Large-scale machine learning on heterogeneous distributed systems (arXiv:1603.04467). *arXiv*. http://arxiv.org/abs/1603.04467

Abrams, J. F., Vashishtha, A., Wong, S. T., Nguyen, A., Mohamed, A., Wieser, S., Kuijper, A., Wilting, A., & Mukhopadhyay, A. (2019). Habitat-Net: Segmentation of habitat images using deep learning. *Ecological Informatics*, *51*, 121–128. https://doi.org/10.1016/j.ecoinf.2019.01.009

Allaire, J., & Chollet, F. (2021). keras: R Interface to 'Keras'. R package version 2.6.1. Retrieved from https://CRAN.R-project.org/package=keras

Asad, S., Abrams, J. F., Guharajan, R., Sikui, J., Wilting, A., & Rödel, M.-O. (2020). Stream amphibian detectability and habitat associations in a reduced impact logging concession in Malaysian Borneo. *Journal of Herpetology*, *54*(4), 385–392. https://doi.org/10.1670/19-136

Cailleret, M., Heurich, M., & Bugmann, H. (2014). Reduction in browsing intensity may not compensate climate change effects on tree species composition in the bavarian forest national park. *Forest Ecology and Management*, *328*, 179–192. https://doi.org/10.1016/j.foreco.2014.05.030

Campbell, M. J., Dennison, P. E., Hudak, A. T., Parham, L. M., & Butler, B. W. (2018). Quantifying understory vegetation density using small-footprint airborne lidar. *Remote Sensing of Environment*, *215*, 330–342. https://doi.org/10.1016/j.rse.2018.06.023

Caruana, R., Lawrence, S., & Giles, C. L. (2000). Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In *Advances in neural information processing systems (NIPS)* (Vol. *13*, pp. 402–408). MIT Press.

Chianucci, F. (2016). A note on estimating canopy cover from digital cover and hemispherical photography. *Silva Fennica*, *50*(1), 1518. https://doi.org/10.14214/sf.1518

Chianucci, F., Ferrara, C., & Puletti, N. (2022, January 14). coveR: An R package for processing digital cover photography images to retrieve forest canopy attributes. *BioRxiv*, 2022.01.13.475850. https://doi.org/10.1101/2022.01.13.475850

Chollet, F. (2015). *Keras*. Retrieved from https://keras.io

Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, *10*(10), 1632–1644. https://doi.org/10.1111/2041-210X.13256

Díaz, G. M., Negri, P. A., & Lencinas, J. D. (2021). Toward making canopy hemispherical photography independent of illumination conditions: A deep-learning-based approach. *Agricultural and Forest Meteorology*, *296*, 108234. https://doi.org/10.1016/j.agrformet.2020.108234

Dupuy, J. M., & Chazdon, R. L. (2008). Interacting effects of canopy gap, understory vegetation and leaf litter on tree seedling recruitment and composition in tropical secondary forests. *Forest Ecology and Management*, *255*(11), 3716–3725. https://doi.org/10.1016/j.foreco.2008.03.021

Ewers, R. M., Didham, R. K., Fahrig, L., Ferraz, G., Hector, A., Holt, R. D., Kapos, V., Reynolds, G., Sinun, W., Snaddon, J. L., & Turner, E. C. (2011). A large-scale forest fragmentation experiment: The stability of altered forest ecosystems project. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1582), 3292–3302. https://doi.org/10.1098/rstb.2011.0049

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, *62*, 729–754. https://doi.org/10.1613/jair.1.11222

Hansen, A., Barnett, K., Jantz, P., Phillips, L., Goetz, S. J., Hansen, M., Venter, O., Watson, J. E. M., Burns, P., Atkinson, S., Rodríguez-Buritica, S., Ervin, J., Virnig, A., Supples, C., & De Camargo, R. (2019). Global humid tropics forest structural condition and forest structural integrity maps. *Scientific Data*, *6*(1), 232. https://doi.org/10.1038/s41597-019-0214-3

Hummel, S., Hudak, A. T., Uebler, E. H., Falkowski, M. J., & Megown, K. A. (2011). A comparison of accuracy and cost of LiDAR versus stand exam data for landscape management on the Malheur national forest. *Journal of Forestry*, *109*(5), 267–273. https://doi.org/10.1093/jof/109.5.267

Jennings, S., Brown, N., & Sheil, D. (1999). Assessing forest canopies and understorey illumination: Canopy closure, canopy cover and other measures. *Forestry: An International Journal of Forest Research*, *72*(1), 59–74. https://doi.org/10.1093/forestry/72.1.59

Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on convolutional neural networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *173*, 24–49. https://doi.org/10.1016/j.isprsjprs.2020.12.010

Kingma, D. P., & Ba, J. (2017). Adam: A method for stochastic optimization. *ArXiv:1412.6980 [Cs]*. Presented at the 3rd International Conference for Learning Representations ICLR 2015, San Diego, USA. Retrieved from http://arxiv.org/abs/1412.6980

Latifi, H., Heurich, M., Hartig, F., Müller, J., Krzystek, P., Jehl, H., & Dech, S. (2016). Estimating over- and understorey canopy density of temperate mixed stands by airborne LiDAR data. *Forestry: An International Journal of Forest Research*, *89*(1), 69–81. https://doi.org/10.1093/forestry/cpv032

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Li, K., Huang, X., Zhang, J., Sun, Z., Huang, J., Sun, C., Xie, Q., & Song, W. (2020). A new method for forest canopy hemispherical photography segmentation based on deep learning. *Forests*, *11*(12), 1366. https://doi.org/10.3390/f11121366

Li, S., Wang, T., Hou, Z., Gong, Y., Feng, L., & Ge, J. (2021). Harnessing terrestrial laser scanning to predict understory biomass in temperate mixed forests. *Ecological Indicators*, *121*, 107011. https://doi.org/10.1016/j.ecolind.2020.107011

Macfarlane, C., Grigg, A., & Evangelista, C. (2007). Estimating forest leaf area using cover and fullframe fisheye photography: Thinking inside the circle. *Agricultural and Forest Meteorology*, *146*(1), 1–12. https://doi.org/10.1016/j.agrformet.2007.05.001

Maj, M. (2020). platypus: Tools for computer vision in R. (Version 0.1.1) [R package]. Retrieved from https://github.com/maju116/platypus

McShea, W. J., & Rappole, J. H. (1992). White-tailed deer as keystone species within forest habitats of Virginia. *Virginia Journal of Science*, *43*, 177–186.

Minaee, S., Boykov, Y. Y., Porikli, F., Plaza, A. J., Kehtarnavaz, N., & Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(7), 3523–3542. https://doi.org/10.1109/TPAMI.2021.3059968

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533. https://doi.org/10.1038/nature14236

Niedballa, J. (2022). EcoDynIZW/imageseg: Imageseg v0.5.0 (v0.5.0). *Zenodo*. https://doi.org/10.5281/zenodo.7041053

Niedballa, J., Axtner, J., Döbert, T. F., Tilker, A., Nguyen, A., Wong, S. T., Fiderer, C., Heurich, M., & Wilting, A. (2022). Data from: Imageseg: An R package for deep learning-based image segmentation. *Dryad Repository*. https://doi.org/10.5061/dryad.x0k6djhnj

Nilsson, M.-C., & Wardle, D. A. (2005). Understory vegetation as a forest ecosystem driver: Evidence from the northern Swedish boreal forest. *Frontiers in Ecology and the Environment*, *3*(8), 421–428. https://doi.org/10.1890/1540-9295(2005)003[0421:UVAAFE]2.0.CO;2

Nudds, T. D. (1977). Quantifying the vegetative structure of wildlife cover. *Wildlife Society Bulletin (1973–2006)*, *5*(3), 113–117.

Ooms, J. (2021). magick: Advanced graphics and image-processing in R. (Version 2.7.3) [R package]. Retrieved from https://docs.ropensci.org/magick/

Paletto, A., & Tosi, V. (2009). Forest canopy cover and canopy closure: Comparison of assessment techniques. *European Journal of Forest Research*, *128*, 265–272. https://doi.org/10.1007/s10342-009-0262-x

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical image computing and computer-assisted intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Sokol, A. (2022). *Pneumothorax detection with U-Net/U-Net++ in Tensorflow 2.x* [Python]. Retrieved from https://github.com/albertsokol/pneumothorax-detection-unet (Original work published 2020)

The GIMP Development Team. (2019). GIMP. (Version 2.10.12). Retrieved from https://www.gimp.org

Tilker, A., Abrams, J. F., Nguyen, A., Hörig, L., Axtner, J., Louvrier, J., Rawson, B., Quang, H. A. N., Guegan, F., Nguyen, T. V., Le, M., Sollmann, R., & Wilting, A. (2020). Identifying conservation priorities in a defaunated tropical biodiversity hotspot. *Diversity and Distributions*, *26*(4), 426–440. https://doi.org/10.1101/790766

Wijffels, J., BNOSSAC, Balarini, J. P., & Nesmachnov, S. (2020). image.Otsu: Otsu's Image Segmentation Method (Version 0.1). Retrieved from https://CRAN.R-project.org/package=image.Otsu

Wing, B. M., Ritchie, M. W., Boston, K., Cohen, W. B., Gitelman, A., & Olsen, M. J. (2012). Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment*, *124*, 730–741. https://doi.org/10.1016/j.rse.2012.06.024

Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N., & Liang, J. (2018, July 18). *UNet++: A nested U-net architecture for medical image segmentation*. Presented at the 4th Deep Learning in Medical Image Analysis (DLMIA) Workshop. https://doi.org/10.48550/arXiv.1807.10165

Zong, X., Wang, T., Skidmore, A. K., & Heurich, M. (2021). Estimating fine-scale visibility in a temperate forest landscape using airborne laser scanning. *International Journal of Applied Earth Observation and Geoinformation*, *103*, 102478. https://doi.org/10.1016/j.jag.2021.102478

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Niedballa, J., Axtner, J., Döbert, T. F., Tilker, A., Nguyen, A., Wong, S. T., Fiderer, C., Heurich, M., & Wilting, A. (2022). imageseg: An R package for deep learning-based image segmentation. *Methods in Ecology and Evolution*, *13*, 2363–2371. https://doi.org/10.1111/2041-210X.13984