



Article

Segmentation of Sandplain Lupin Weeds from Morphologically Similar Narrow-Leafed Lupins in the Field

Monica F. Danilevicz ¹, Roberto Lujan Rocha ², Jacqueline Batley ¹, Philipp E. Bayer ³, Mohammed Bennamoun ⁴, David Edwards ^{1,*} and Michael B. Ashworth ²

¹ Centre for Applied Bioinformatics, School of Biological Sciences, University of Western Australia, Perth, WA 6009, Australia

² Australian Herbicide Resistance Initiative, School of Agriculture and Environment, The University of Western Australia, Perth, WA 6009, Australia

³ Minderoo Foundation, Perth, WA 6009, Australia

⁴ Department of Computer Science and Software Engineering, University of Western Australia, Perth, WA 6009, Australia

* Correspondence: dave.edwards@uwa.edu.au

Abstract: Narrow-leafed lupin (*Lupinus angustifolius*) is an important dryland crop, providing a protein source in global grain markets. While agronomic practices have successfully controlled many dicot weeds among narrow-leafed lupins, the closely related sandplain lupin (*Lupinus cosentinii*) has proven difficult to control, reducing yield and harvest quality. Here, we successfully trained a segmentation model to detect sandplain lupins and differentiate them from narrow-leafed lupins under field conditions. The deep learning model was trained using 9171 images collected from a field site in the Western Australian grain belt. Images were collected using an unoccupied aerial vehicle at heights of 4, 10, and 20 m. The dataset was supplemented with images sourced from the WeedAI database, which were collected at 1.5 m. The resultant model had an average precision of 0.86, intersection over union of 0.60, and F1 score of 0.70 for segmenting the narrow-leafed and sandplain lupins across the multiple datasets. Images collected at a closer range and showing plants at an early developmental stage had significantly higher precision and recall scores (p -value < 0.05), indicating image collection methods and plant developmental stages play a substantial role in the model performance. Nonetheless, the model identified 80.3% of the sandplain lupins on average, with a low variation ($\pm 6.13\%$) in performance across the 5 datasets. The results presented in this study contribute to the development of precision weed management systems within morphologically similar crops, particularly for sandplain lupin detection, supporting future narrow-leafed lupin grain yield and quality.

Keywords: narrow-leafed lupin; sandplain lupin; *Lupinus angustifolius*; *Lupinus cosentinii*; image segmentation; deep learning; precision agriculture; herbicide application; weed management



Citation: Danilevicz, M.F.; Rocha, R.L.; Batley, J.; Bayer, P.E.; Bennamoun, M.; Edwards, D.; Ashworth, M.B. Segmentation of Sandplain Lupin Weeds from Morphologically Similar Narrow-Leafed Lupins in the Field. *Remote Sens.* **2023**, *15*, 1817. <https://doi.org/10.3390/rs15071817>

Academic Editors: Muhammad Moshir Rahman and Andrew Robson

Received: 20 February 2023

Revised: 22 March 2023

Accepted: 24 March 2023

Published: 29 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sandplain lupin (*Lupinus cosentinii*) is a highly competitive weed species that significantly reduces the grain yield and quality of narrow-leafed lupin (*Lupinus angustifolius*) crops [1–3]. Initially, sandplain lupins were introduced as a leguminous pasture species for cattle production due to their rapid growth on infertile sandy or loamy calcareous soils, and high protein content [1,2]. However, with the increase in cropping intensity, sandplain lupins became problematic, decreasing the yield and quality of narrow-leafed lupin crops through interspecific competition and acting as a source of anthracnose infection [4]. Narrow-leafed lupin is an important protein crop, with 75% of global production in Australia [5]. Grain lupin produces seeds with up to 44% protein content and is used as flour or a food supplement [5,6]. Currently, there are no options to chemically control sandplain lupins among narrow-leafed lupins due to their biological similarity; therefore, precision application of herbicide to sandplain lupin individuals is required to limit crop damage [7].

Precision agriculture considers the intra-field variability to provide tailored treatments for each region of the field [8]. This includes the development of weed maps to inform robotic weeding or targeted herbicide application, decreasing herbicide use while improving weed control [9,10]. Multiple studies have proposed methods for building weed maps and weed detection systems using images captured through unoccupied aerial vehicles (UAVs), ground vehicles, and hand-held devices [7,11–13]. Red–Green–Blue (RGB) images are most commonly used for weed detection due to easy accessibility. Nonetheless, multispectral and hyperspectral cameras are becoming increasingly common, as they provide more features that might be used to discriminate between weed and crop species [11]. The images are often analysed using handcrafted features to discriminate between crop and weed pixels. Handcrafted features are defined by the algorithm developer, using simplified shapes, or from plant spectral variation, based on differences in the canopy structure and leaf morphology [11,14–16]. However, using handcrafted features can introduce bias, which limits the method from being applied under different environmental conditions, such as varying light intensity [15].

Deep learning algorithms have emerged as an alternative approach to discriminate between crop and weed species without using handcrafted features. Deep learning algorithms learn directly from labelled training datasets, automatically extracting the relevant features that can be used to discriminate between objects [17]. The main deep learning algorithms applied for image-based weed detection are based on convolutional neural networks (CNNs) and transformers [18–22]. A study using CNNs obtained 97% accuracy in identifying grass and broadleaf weeds among soybeans (*Glycine max*) [18]. Additional studies using YOLOv3, a CNN-based architecture, successfully detected hedge bindweed (*Convolvulus sepium*) weed in sugar beet (*Beta vulgaris*) crops, with an average precision of 76–89% [19]. The YOLOv3 algorithm also achieved a precision of 71% and recall of 78% for common purslane (*Portulaca oleracea*) weed treatment by autonomously controlling the smart sprayer prototype, demonstrating the value of CNN-based architectures for precision agriculture [23]. Other CNN-based architectures have also been used for weed segmentation, allowing researchers to estimate the weed density and its likely impact on crop yield [20–22,24]. Segmentation models are reported to discriminate weeds from sunflower (*Helianthus annuus*) crops with 90% accuracy at an early stage [24]; rice (*Oryza sativa*) seedlings from three-leaf arrowhead (*Sagittaria trifolia*) weeds with 92.7% accuracy [22]; and chamomile (*Matricaria chamomilla*), common poppy (*Papaver rhoeas*), ivy-leaved speedwell (*Veronica hederifolia*), and field pansy (*Viola arvensis*) weeds from winter wheat (*Triticum aestivum*) with 94% accuracy [20]. Although CNN-based architectures are relatively successful at detecting and segmenting weeds in the field, deep learning models present high variability in performance (ranging from 76% to 94% accuracy), indicating that detection capability is associated with crop and weed targets [20–22,24].

In the case of sandplain and narrow-leafed lupins, the similarity between the species presents a challenge for implementing weed detection in the field. Similar weed and crop species present fewer discriminating features to enable accurate discrimination between plants, especially under varying field conditions, such as altitude/height, luminosity, plant density, soil appearance, and tillage [19]. In this study, we trained and assessed the performance of a U-net deep learning model with a pretrained Resnet18 model backbone [25] to identify sandplain lupins growing among morphologically similar narrow-leafed lupin crops. The model accuracy was assessed for weed segmentation; measuring the impact of the image and plant conditions; and appraising the model efficiency in locating sandplain lupin plants for robotic weed control and automated herbicide application.

2. Materials and Methods

The data collected for this study are available in WeedAI “2022—WA Sandplain and Narrow-leafed lupin” and Figshare at <https://figshare.com/articles/dataset/21746669> (accessed on 10 April 2022). The scripts developed are available on GitHub at <https://github.com/mdanilevicz/WeedDetectionML> (accessed on 10 April 2022). The docker files and singularity images [26] used were downloaded from <https://hub.docker.com/layers/osgeo/>

[gdal/alpine-small-latest/images/sha256-640f4dfba9d7d48b6f66a4ca3436ab0913f7473fad6033125dcb3da940227038?context=explore](https://github.com/OSGeo/gdal/releases/tag/3.12.0) (accessed on 10 April 2022) and https://docs.nvidia.com/deeplearning/frameworks/pytorch-release-notes/rel_21-07.html (accessed on 10 April 2022). A persistent overlay filesystem was used to dynamically install the required Python libraries as the study was developed. Unless otherwise stated, the scripts indicated in the methodology refer to the GitHub repository.

2.1. Experimental Field and Data Collection

Field experiments of narrow-leafed lupin naturally infested with sandplain lupin were conducted under rainfed field conditions in Mingenew, Western Australia (AU). The narrow-leafed lupin crop was sown on 9 May 2022 at a seeding rate of 95 kg/ha to a 3 cm depth over a 0.59 ha area. The field study had a stubble cover of 45% and was treated with pre-emergence herbicide before sowing. Images collected in the field trial site were named field-1 and field-2. Grower field images were collected at a cropping field in Mingenew, Western Australia. The images collected at the grower's field were named grow-1 and grow-2. The "field" and "grow" datasets present different farm management strategies, with varying seeding rates and crop row distance.

The images for the field-1, field-2, grow-1, and grow-2 datasets were collected using a DJI Phantom 4 unoccupied aerial vehicle (UAV) RGB camera. The images were collected between 12 and 2 PM under overcast or clear sky conditions. The details for the image collection can be seen in Table 1. The images were collected with 75% side overlap and 80% front overlap, with five ground control points distributed across the field to increase the GPS accuracy. Additionally, 217 images with 4879 sandplain lupins labelled among narrow-leafed lupins were downloaded from the Weed-AI database [27]. The images from ext-1 were collected on 12 July 2019 in multiple locations in Geraldton, Western Australia (AU), using an iPhone XS rear camera at an approximately 1.5 m height and 90 degree angle. The plant growth stage in each dataset was estimated using the Lupin Growth and Development report [28], which uses a crescent decimal score from 0–5.9 to indicate plant growth from dry seeds for sowing to harvest ripe, respectively.

Table 1. Description of the image datasets for weed detection and segmentation. GSD means ground sample distance, and the flight height is indicated in metres.

ID	Field Type	Platform	Collection Date	Growth Stage	GSD (cm/px)	Flight Height (m)	Total Images	Total Labels
field-1	Trial site	UAV	16 July 2021	2–3.3	0.27	10	101	1602
field-2	Trial site	UAV	11 August 2021	3–4	0.55	20	97	840
grow-1	Grower	UAV	16 July 2021	2–3.3	0.11	4	88	462
grow-2	Grower	UAV	19 August 2021	2–4	0.55	20	292	207
ext-1	Grower	Smartphone	12 July 2019	1–2.5	0.01	1.5	217	4879

2.2. Image Data Processing

The images in the "field" and "grow" dataset were processed following the method previously detailed in [29] to prepare the images for an orthomosaic and model input. The orthomosaics were assembled using Metashape (v1.8.0, Agisoft), and the plot shapefiles were generated using the plotshpcreate R library [30], as shown in the R script "generate_shapefile.Rmd". The shapefiles were used to extract the plots from the orthomosaic using Gdal (v3.2) and the "extract_plots.sh" script written in bash, and using the container image downloaded from <https://hub.docker.com/layers/osgeo/gdal/alpine-small-latest/images/sha256-640f4dfba9d7d48b6f66a4ca3436ab0913f7473fad6033125dcb3da940227038?context=explore> (accessed on 10 April 2022). The images gathered from the grower sites and obtained from Weed-AI did not undergo this processing, as they were not collected in a continuous overlap, as were the field experiment images.

The Python script (“improcessing.ipynb”) was employed to perform the following image processing steps in the field-1, field-2, grow-1, and grow-2 datasets. The GeoTiff images were converted to NumPy arrays and rotated to fit the same orientation. The array pixel values in each image were normalised to fit the 0–1 range, standardising the pixel values between datasets, and a copy of the images was converted to the jpeg format for drawing the bounding box labelling.

2.3. Bounding Box Labelling and Segmentation Masks

After processing the images from the “field” and “grow” datasets, the Colour Index of Vegetation (CIVE) detailed in Equation (1) [31] and the Otsu Threshold from Opencv2 were used to discriminate between soil and plant pixels [32,33] in the “improcessing.ipynb” custom Python script (Figure 1B). Makesense.ai tool was used to manually draw the bounding boxes around sandplain lupin plants identified in the field-1, field-2, grow-1, and grow-2 datasets, as shown in Figure 1C [34]. The ext-1 dataset had been previously labelled by the dataset owners. The coordinates from the bounding boxes around the sandplain lupin plants were overlaid on the plant pixels to label narrow-leafed and sandplain lupin pixels, using “improcessing.ipynb” to generate the segmentation masks used to train the deep learning model, as shown in Figure 1D. The images and masks were split into the standard size of 500 by 500 pixels using the “resize_images.ipynb” custom script to accelerate training the deep learning model.

$$\text{CIVE} = 0.441 \times \text{Red Band} - 0.881 \times \text{Green Band} + 0.385 \times \text{Blue Band} + 18.78745 \quad (1)$$

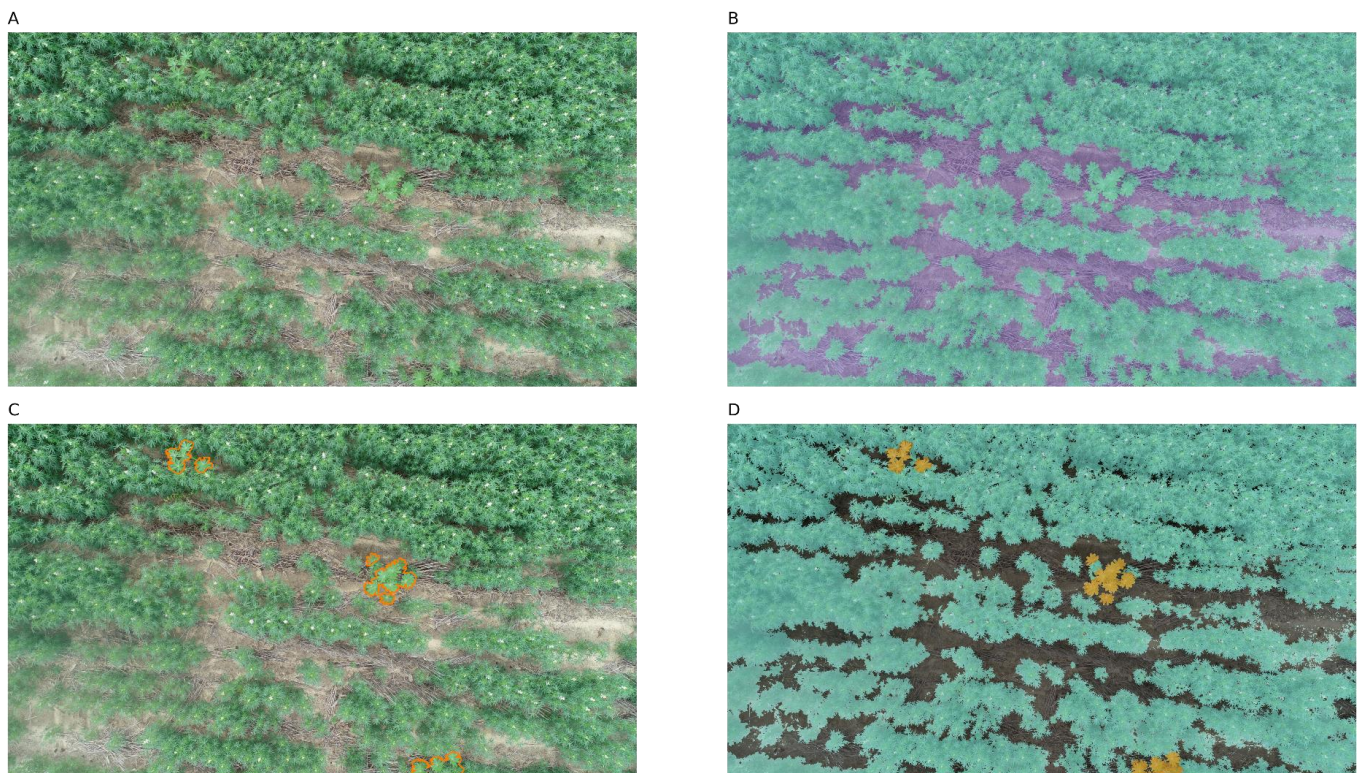


Figure 1. Representation of the image processing steps. (A) Original RGB image, (B) CIVE vegetation index mask separating soil pixels are coloured in purple and plant pixels are green, (C) Bounding box labelling over sandplain lupins overlaid on RGB image, (D) Ground-truth segmentation mask overlaid on the RGB image.

2.4. Segmentation Model Architecture

The deep learning model consisted of a feature extraction and a semantic segmentation module based on the U-Net architecture using a pretrained Resnet18 backbone [25,35], as

shown in Figure 2. The model was implemented and trained using the fastai library [36] in the custom script “model_kfold.py”. The model architecture used Pixel shuffle during segmentation [37]; self-attention [38], with DICE loss due to the unbalanced segmentation classes [39]; and ADAM as the optimisation function, and the learning rate was set to 0.001. The model implementation and fivefold validation were developed using an NVidia container available at [40], https://docs.nvidia.com/deeplearning/frameworks/pytorch-release-notes/rel_21-07.html (accessed on 10 April 2022).

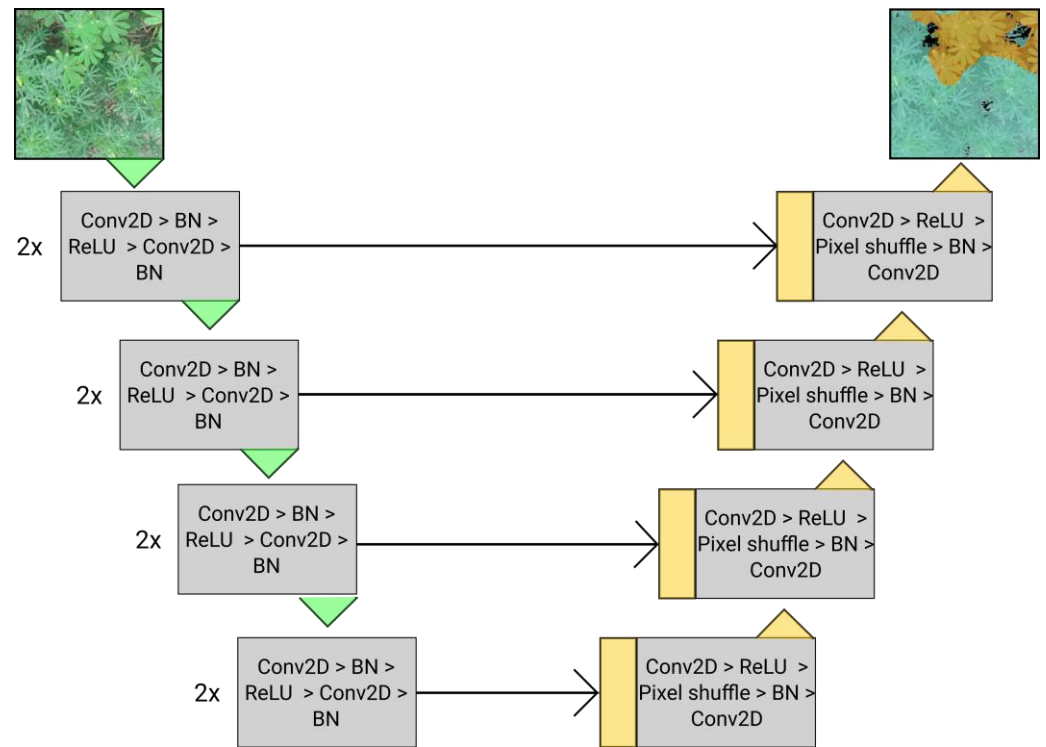


Figure 2. Scheme of the weed segmentation model based on the U-Net architecture. The feature extraction module is indicated by the green arrows, and the semantic segmentation module is indicated by the yellow arrows. The yellow box indicates the copied feature maps used in the segmentation module.

2.5. Pixel-Wise Evaluation Metrics

The metrics used to evaluate the pixel-wise segmentation performance were Precision, Recall, Intersection over Union (IoU), and Macro F1, using the scikit-learn library [41] and implemented through the “prediction_analysis.ipynb” script. The Precision and Recall metrics are detailed in Equations (2) and (3). The Intersection over Union, also known as the Jaccard Index (Equation (4)), calculates the area overlap between the predicted segmentation and the ground mask. Macro F1, shown in Equation (5), calculates the arithmetic mean over the individual F1 scores of the different target classes, and it is more robust toward unbalanced datasets [42].

$$\text{Precision} = \text{True positive} / (\text{True positive} + \text{False positive}) \quad (2)$$

$$\text{Recall} = \text{True positive} / (\text{True positive} + \text{False negative}) \quad (3)$$

$$\text{IoU} = \text{Area of overlap} / \text{Area of union} \quad (4)$$

$$\text{Macro F1} = 1/n \sum (2 \text{ Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (5)$$

2.6. Object-Wise Weed Detection

The predicted sandplain lupin estimator generated from the segmentation mask was obtained using the custom script “prediction_analysis.ipynb”. The contour of the sandplain lupin labels was extracted from the predicted segmentation mask, and objects with a total area smaller than 10×10 pixels were removed. The contours from the predicted mask and the ground-truth mask were superimposed to count which objects were identified in both masks and which sandplain lupins were mispredicted by the model.

2.7. Weed Map Construction from Predicted Masks

The reconstruction of the predicted masks into a weed map enables overlaying the predicted sandplain lupin position on the orthomosaic to extract its geospatial position. The original RGB orthomosaic from the field-1 dataset generated in Section 2.2 was converted to JPEG and cut into blocks of 500×500 pixels. Each image block was named according to its position in the orthomosaic (i.e., row00_column00.jpeg). The block images were fed to the trained deep learning model for sandplain lupin segmentation. The predicted segmentation masks were assembled back into position, based on their ID and standard size of 500×500 pixels, and overlaid in the orthomosaic. The general workflow is illustrated in Figure 3, in which the sandplain lupin segmentation mask on the right can be mapped back to the GPS-marked orthomosaic, guiding the implementation of the weed management strategies. The Python script used for the weed map assembly is shown in “plotting_results/weed_map_prediction.ipynb”.

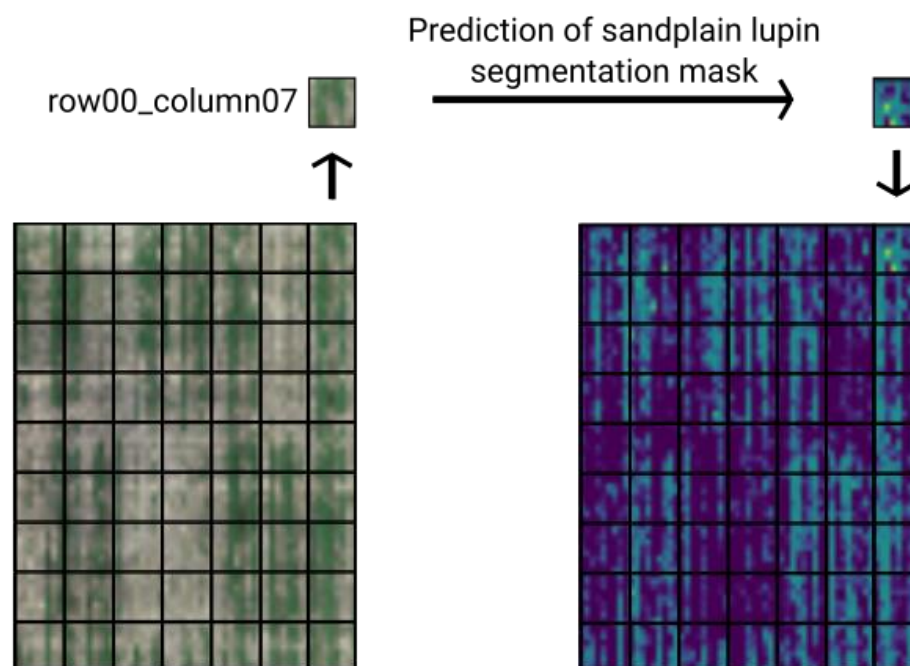


Figure 3. Sandplain lupin prediction in the field-1 dataset. The orthomosaic from the field-1 dataset was split into multiple smaller RGB images, which were fed to the deep learning model for prediction and reassembled into a weed map.

3. Results

3.1. Segmentation Performance for Sandplain and Narrow-Leafed Lupins

The segmentation model converged after 100 epochs using the mixed training dataset, with no noticeable improvement observed when training for longer periods. High segmentation performance was achieved for pixel-wise labelling of sandplain and narrow-leafed lupins in the field, with an average precision of 0.86, recall of 0.95, IoU of 0.60, and Macro F1 of 0.70. A detailed performance evaluation for each dataset condition is presented in Table 2. Narrow-leafed lupin segmentation was slightly more accurate in identifying the

pixels associated with the crop, providing a segmentation mask more like the ground-truth label, as indicated by the IoU and Macro F1 metrics. The precision metric was similar for both plant targets, showing a high proportion (0.85) of pixels labelled as narrow-leafed or sandplain lupins was true. Recall indicates the model's ability to detect pixels related to each class, showing that most of the narrow-leafed lupin pixels were detected (0.96), and more than half of sandplain lupin pixels were detected (0.62). The highest sandplain lupin segmentation performance was observed with the ext-1 and grow-1 datasets, which depict the plants at a higher resolution and at earlier developmental stages (Figure 4). Sandplain and narrow-leafed lupins are heliotropic plants, with the leaves moving in response to the sunlight direction to maximise absorption, as depicted in ext-1.

Table 2. Segmentation performance comparison between the hold-out datasets for each condition using fivefold cross-validation. Metrics are presented per class: narrow-leafed lupin (NLL), sandplain lupin (SL), and weighted average (Avg), which averages the metric between the classes considering the number of true instances in each class.

Dataset	Precision			Recall			IoU			Macro F1		
	NLL	SL	Avg	NLL	SL	Avg	NLL	SL	Avg	NLL	SL	Avg
field-1	0.81	0.82	0.81	0.95	0.70	0.93	0.51	0.45	0.51	0.64	0.57	0.64
field-2	0.89	0.82	0.88	0.91	0.43	0.89	0.51	0.26	0.50	0.58	0.37	0.57
grow-1	0.96	0.84	0.96	0.99	0.85	0.99	0.87	0.64	0.87	0.93	0.76	0.92
grow-2	0.95	0.83	0.95	1.00	0.32	0.99	0.72	0.29	0.72	0.82	0.42	0.82
ext-1	0.68	0.88	0.69	0.97	0.78	0.97	0.41	0.54	0.42	0.56	0.67	0.57

3.2. Target Accuracy for Detecting Individual Sandplain Lupin Weeds

Identifying sandplain lupin infestations in the field is a primary requirement for implementing targeted weed management practices. Object-wise detection at the plant level is commonly used to indicate the model's capacity to locate weeds in the field [43], as opposed to pixel-wise segmentation (Section 3.1), which measures the segmentation mask completeness. The weed objects were obtained by leveraging the segmentation mask contours, considering each independent contour object as an individual target, as shown in Figure 5. The predicted sandplain lupin location (indicated in orange) may not cover the whole weed area, but it can guide weed management decisions if a minimum area threshold is defined (Figure 5). In this case, each sandplain lupin contour label had an area larger than 100 pixels, removing low confidence regions, as most sandplain lupin leaves would occupy an area above the defined threshold.

The performance of the object-wise identification varied depending on the dataset, with an average of 80.3% of the sandplain lupins being accurately detected, as shown in Table 3. Although the number of sandplain lupin targets varied substantially between the datasets, a high percentage of sandplain lupin regions were identified in each condition. In the ext-1 dataset, each leaf was considered an independent object because of its canopy structure during an early developmental stage, causing ext-1 to present an inflated sandplain lupin count. Overall, 4 datasets presented a sandplain lupin identification accuracy above 77%, except for the grow-2 dataset, in which the images were collected at a 20 m height at a later developmental stage, causing a loss in image resolution and constituting a more complex challenge for the segmentation model (Table 3). The total number of sandplain lupins present in each image was correlated with the number of predicted sandplain lupins, as indicated in Figure 6. The field-1 and ext-1 datasets presented a higher density of weeds per image. However, the variation in weed infestation levels did not affect the model accuracy.

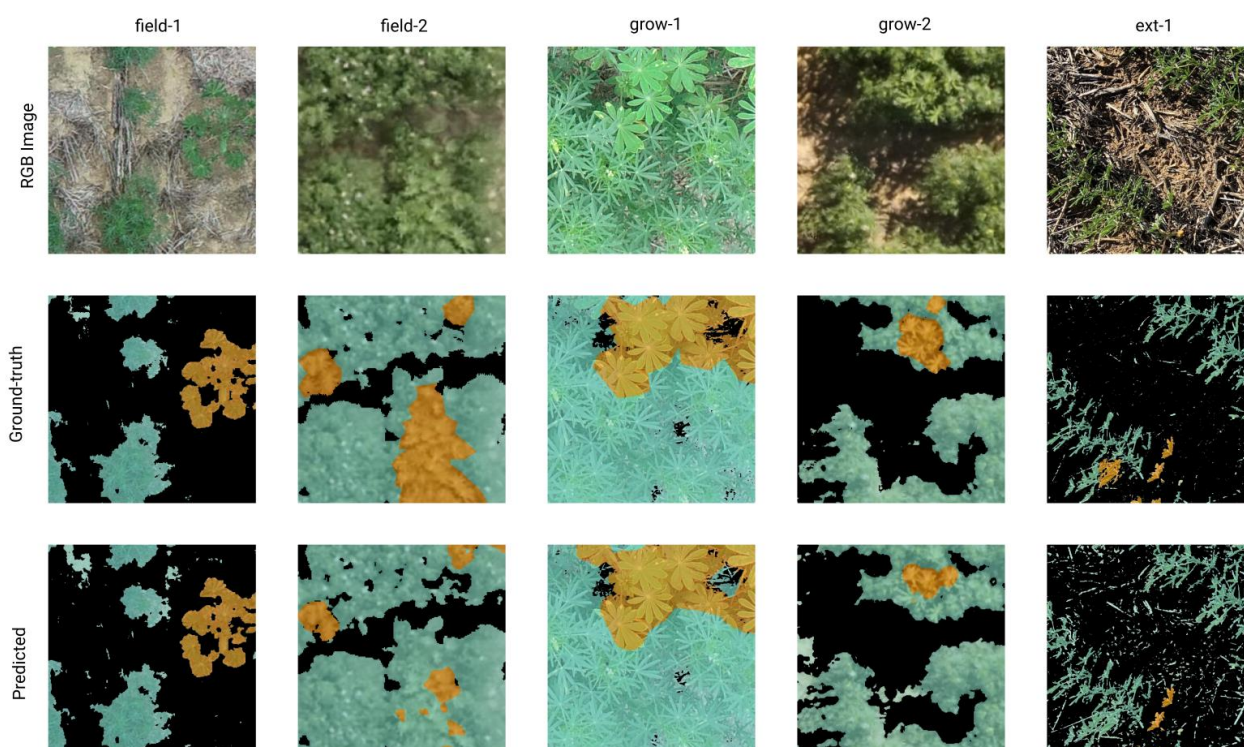


Figure 4. Predicted mask from the segmentation model. Each column shows an image sample from a specific dataset (field-1, field-2, grow-1, grow-2, and ext-1). The first row is the RGB image, followed by the ground truth and predicted masks.

3.3. Effect of Environmental Conditions on Sandplain Lupin Detection

A comparison between the sandplain lupin segmentation performance across the different datasets indicates that field management conditions, plant development stage, and UAV flight altitude play a substantial role in the model performance (Figure 7). The precision metrics indicate what proportion of sandplain lupin pixels predicted is true, whereas recall shows whether the model could find all sandplain lupin pixels. Most datasets presented a similar trend in precision performance, with the model showing significantly higher precision scores in ext-1 in comparison to field-2 (p -value 0.01). Regarding the recall performance, the model showed strong variation, depending on the dataset observed. A one-on-one comparison revealed that the recall metrics varied significantly for all datasets (p -value < 0.05), although ext-1, field-1, and grow-1 present similar recall values between 0.70 to 0.85, which indicates these datasets have the majority of the sandplain lupin pixels detected. It is possible to observe that in Figure 7B, there was a higher dispersion of the recall value across the different image samples within the same dataset, as some datasets also presented a higher number of images in the hold-out dataset due to a larger initial dataset. The field-2 and grow-2 datasets presented the lowest recalls observed, but they were still significantly different (p -value < 0.05).

3.4. Weed Mapping Increases Herbicide Use Efficiency

The reconstruction of the predicted sandplain lupin images into a field orthomosaic enables visualisation of the weed-infested regions for both targeted herbicide application and measuring the total area covered by sandplain lupin and narrow-leaved lupin. Here, the reconstruction of the field trial dataset (field-1) in Figure 8 shows that the predicted sandplain lupins compose 5.4% of the total area associated with plant pixels, with the remaining 94.6% labelled as belonging to narrow-leaved lupin. The field trial shown in Figure 8 is approximately 5952 m², but only 1.5% of the area is covered by sandplain lupin weeds, meaning a spraying area reduction of 98.5%, in which the model can target 79.9% of the weeds in the field for the field-1 dataset (Table 2).

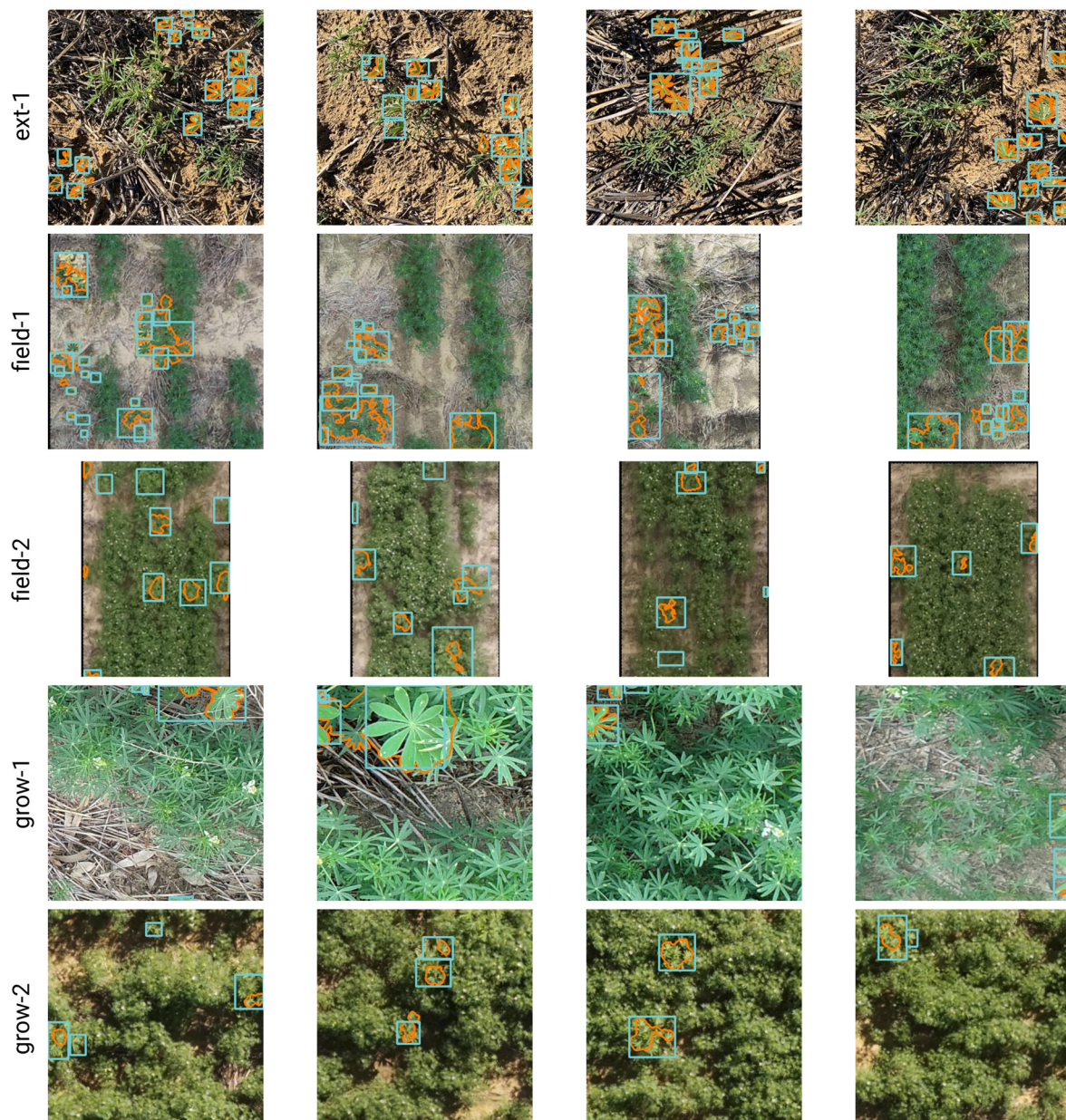


Figure 5. Object-wise detection of sandplains by the model on different datasets. Each line corresponds to images from a single dataset; the order from top-bottom is ext-1, field-1, field-2, grow-1, and grow-2. The sandplain lupin bounding boxes represent the ground-truth weed labelling, whereas the orange contour shows the weed region identified by the model.

Table 3. Percentage of identified sandplain lupin regions. The count was performed using model predictions on the hold-out dataset compared to the ground-truth labels.

Dataset	Number of Sandplain Lupins	Predicted Sandplain Lupins	Percentage Identified (%)	R ²
field-1	737	589	79.91	0.76
field-2	143	111	77.62	0.64
grow-1	87	75	86.20	0.76
grow-2	31	23	74.19	0.73
ext-1	938	785	83.37	0.91
Total	1936	1583	80.32	0.76

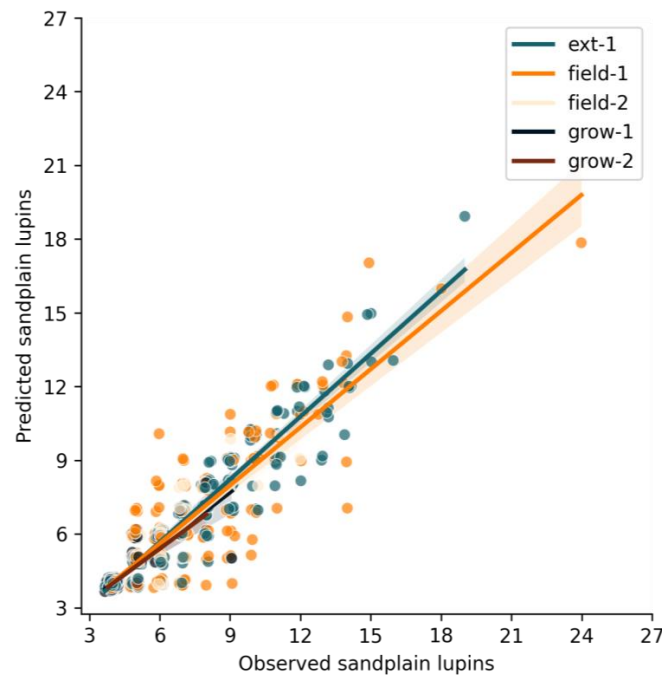


Figure 6. Comparison of predicted and observed sandplain lupins. Each point corresponds to the number of sandplain lupins identified in the 500×500 -pixel image. The x -axis indicates the number of sandplain lupins identified in the ground-truth labelling, and the y -axis shows the number of sandplain lupins detected by the model. The prediction R^2 values for each dataset are ext-1 (0.91), field-1 (0.76), field-2 (0.64), grow-1 (0.76), and grow-2 (0.73).

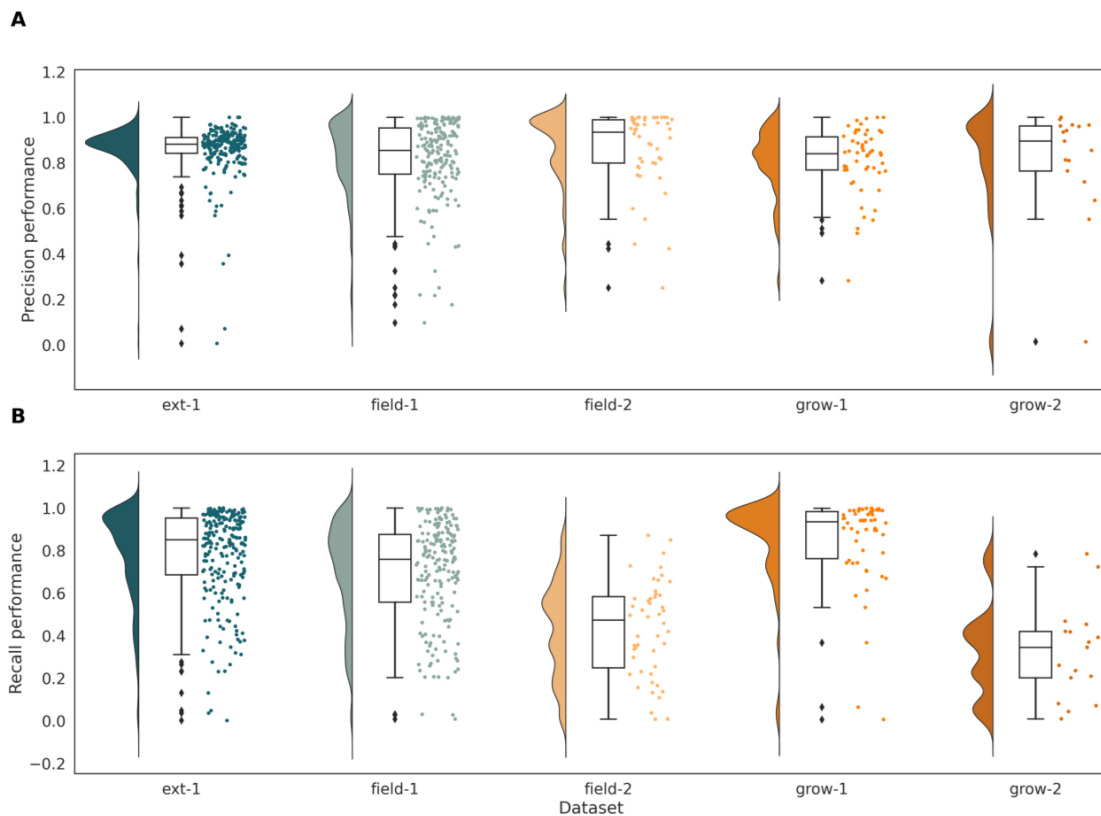


Figure 7. Comparison of precision and recall metrics per image sample distributed across five datasets with varied conditions. (A) Shows the precision score for sandplain lupin segmentation on each dataset; (B) Shows the recall score for sandplain lupin segmentation on each dataset.

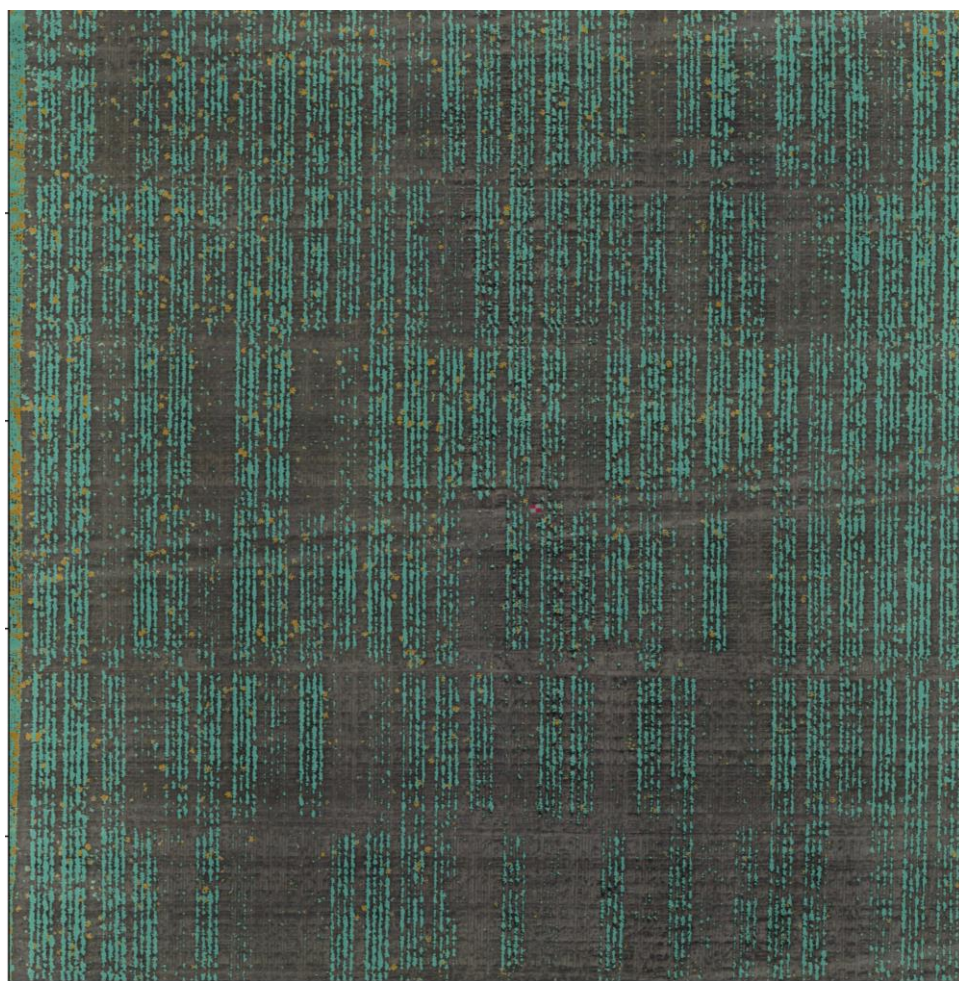


Figure 8. Reconstruction of the weed map for visualisation of the sandplain lupin infestation on the field-1 dataset. In this figure, the RGB image was overlaid with the model prediction mask, in which green represents the predicted narrow-leafed lupin, and the sandplain lupins are in orange.

4. Discussion

Weed detection among morphologically similar crops is a challenge for effective weed control, as fewer features can be employed by deep learning models to distinguish between species [44,45]. In this study, we employed pixel-wise and object-wise metrics to evaluate the model efficiency for detecting sandplain lupins among narrow-leafed lupins in the field. Object-wise detection is particularly effective in indicating the proportion of weeds that could be targeted by real-time robotic weeding or precision herbicide application [44]. In this study, an average of 80.3% of the sandplain lupin objects were successfully identified by the model, with a low variation ($\pm 6.13\%$) across the 5 datasets assessed. As reported in Gao [19], an accuracy of 70% is sufficient for a weed detection model to be efficient for field operation in order to achieve meaningful reductions in herbicide application.

The model presented in this study achieved high pixel-wise segmentation performance, with the precision (0.82 to 0.88) and recall scores (0.32 to 0.85) showing variation across the 5 datasets analysed, which were collected under distinct conditions. The object-wise and pixel-wise metrics indicate the model was able to identify most of the sandplain lupins across the five datasets, although the segmentation masks are likely to only partially cover the weed canopy area. The model's resultant performance is comparable to previous studies tackling the detection of weeds among similar crops [44,46,47]. For example, a canola and wild radish classification model achieved an average of 90.9% accuracy under a controlled environment using LBP handcrafted features [46]. Another study reports an F-score of 93.3% for Italian ryegrass (*Lolium perenne*) detection among wheat crops [47]. Even though

these studies present high-performance metrics, the detection of these morphologically similar species relied on handcrafted features, which may limit the model's applicability in the field environment. In contrast, our study measured the model performance across diverse field conditions.

Developing a weed detection model that achieves robust performance under diverse field and plant conditions remains a challenge [43], as the plant developmental stage, density, and image conditions themselves significantly impact the model performance, as shown in this study. Here, the model presented the highest weed detection performance (precision > 0.82, recall > 0.70, object-wise detection > 79.9%) in the field-1, grow-1, and ext-1 datasets for identifying pixels associated with sandplain lupin weeds. The images collected in these datasets present high resolution (GSD below 0.3 cm/pix), which may have contributed to distinguishing the morphologically similar weed and crop species across the different sites. Previous studies pointed out that image collection methods may impact the leaf spectral reflectance, playing a central role in the detection performance [15,47,48]. Using the field-1, grow-1, and ext-1 datasets, the plants were presented at early developmental stages, showing reduced canopy overlap and plant density. The background complexity and canopy overlap are known issues for weed detection in the field [45]; a previous study focusing on the detection of morphologically similar ryegrass within wheat fields observed a variation in the model performance, depending on the plant developmental stage [44]. Altogether, our results indicate using images collected at high resolution from young plants are more suitable for the detection of sandplain lupin weeds among a morphologically similar crop. This find is corroborated by previous studies carried out using images collected at 2–3 m height for distinguishing combinations of crops and weeds with similar morphology [44,46,47].

Narrow-leafed and sandplain lupins present heliotropism, which changes the plant's leaf direction and canopy structure relative to the position of the sun, enhancing solar absorption and affecting evapotranspiration [49,50]. Heliotropism is depicted in the ext-1 dataset, with both species' leaves pointing sideways. Besides showing the heliotropic movement of the plants, the ext-1 dataset presented a complex environment, with plants at the seedling stage and surrounded by tillage and the residue of the previous crop. The ext-1 dataset was the only instance in which the sandplain lupin detection precision was superior to narrow-leafed lupins by 0.20, potentially due to the thinner leaves of narrow-leafed lupin when seen from above, making them difficult to spot in the complex background. Heliotropism in young narrow-leafed lupin may positively affect the precision of sandplain lupin identification, as the highest value was achieved in the ext-1 dataset (0.88) compared to the other datasets with no observed heliotropism (<0.84). The variation in canopy structure imposes an additional factor for the identification of morphologically similar crops and may affect models for weed detection among other crops presenting heliotropism, such as the common bean (*Phaseolus vulgaris*), pea (*Pisum sativum*), sunflower (*Helianthus annuus*), and soybean (*Glycine max*) [51–54]. To circumvent this challenge, this study aimed to image the plants around midday and/or under overcast conditions. However, when covering large field areas, heliotropism is unavoidable and should be included in the model's training dataset. As shown in this study, a representative training dataset is important for detecting sandplain lupins among narrow-leafed lupins.

Using the model proposed here, it is estimated that the area requiring herbicide application was reduced by 98% compared to broadcast application, accurately targeting 79.9% of the sandplain lupins in the field-1 dataset. However, depending on the attributes of the datasets, the model could successfully identify 74% to 86% of the sandplain lupins, having the potential to further increase the herbicide application efficacy. Another study using UAV for weed detection showed a 20–60% increase in the herbicide application efficiency, targeting 74% of broadleaf weeds growing among grasses [10]. The efficiency of herbicide use, when comparing precision versus broadcast application, varies depending on the weed density and distribution in the field, with patchy weed infestations being more efficiently controlled using precision herbicide application [10,55]. In addition, machinery

technology factors, such as the machinery travel speed and time required to achieve operating pressure at the nozzle to initiate a full spray distribution, will also affect the final area to be treated with herbicide. In our study, the predicted reduction in the herbicide-treated area is due to the sparseness and heterogeneity of the sandplain lupin infestation in this field trial. Although there is no specific literature published on the dormancy of sandplain lupin, other wild species of hard-seeded lupin, such as *Lupinus articus*, have been found to remain viable for an estimated 10,000 years [56]. To control the seed bank of these highly dormant species, weed maps, as shown in Figure 5, can be used to monitor the weed density, assess the effectiveness of weed control strategies over multiple seasons, and apply weed control measures that will exhaust the weed seed bank.

It is important to highlight that the prediction mask may not cover the whole weed canopy and that overlapping plant canopies may lead to underestimating the weed density in each area. In addition, further model development and training are required before deployment, as the model was trained solely using images obtained in a limited region in Western Australia. Future studies can use multiyear weed maps to finetune the model, assessing whether it will present increased detection accuracy for areas with persistent sandplain lupin infestations.

5. Conclusions

Sandplain lupins are a problematic weed in narrow-leafed lupin crops, causing a decrease in the yield and quality. The lack of selective herbicide treatments requires the development of advanced weed management techniques for crop protection and the reduction of the weed seed bank. This study presents an effective model for identifying and mapping sandplain lupins among morphologically similar narrow-leafed lupins, generating a weed map that can inform spatial weed control strategies, such as spot spraying herbicide treatments. Although the similar morphology between sandplain and narrow-leafed lupins poses a challenge for weed identification, this study on average achieved a 70% pixel-wise F1 score and 80.3% object-wise accuracy for the identification of sandplain lupins across 5 datasets depicting distinct field conditions. The results also indicate that high-resolution images and imaging plants at early developmental stages may be more suitable for weed identification. Over a longer timeframe, sandplain lupin maps generated at multiple time points can be compared to assess the effectiveness of the weed control strategy and to detect regions recalcitrant to treatment.

Author Contributions: R.L.R. designed and maintained the field trial. M.F.D. and R.L.R. collected the image data. M.F.D. processed the data and developed the model. J.B., P.E.B., M.B., D.E. and M.B.A. provided additional analysis and data interpretation, and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to thank the Pawsey Supercomputing Centre for computation resources. The Australian Government supported this work through the Australian Research Council (Projects DP210100296, DP200100762, and DE210100398) and the Grains Research and Development Corporation (Projects 9177539, 9177591, and UWA2007-002RTX). Monica F. Danilevicz was supported by the Research Training Program scholarship and the Forrest Research Foundation.

Data Availability Statement: The image data and labels generated for this study are available at <https://figshare.com/articles/dataset/21746669> (accessed on 10 April 2022). The implementation of the model and the custom scripts used for data processing are available at the GitHub repository <https://github.com/mdanilevicz/WeedDetectionML> (accessed on 10 April 2022).

Acknowledgments: The authors would like to acknowledge Ken Flower and Frank D'Emden for their support in the data collection phase.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. DPIRD. Early History of Lupins in Western Australia | Agriculture and Food. Available online: <https://www.agric.wa.gov.au/lupins/early-history-lupins-western-australia> (accessed on 21 June 2022).
2. Brand, J.D.; Tang, C.; Rathjen, A.J. Screening rough-seeded lupins (*Lupinus pilosus* Murr. and *Lupinus atlanticus* Glads.) for tolerance to calcareous soils. *Plant Soil* **2002**, *245*, 261–275. [CrossRef]
3. Megirian, G. Review Investigates Control Options for Blue Lupin and Weeds in the West. *Groundcover* 2020, Issue 147, July–August 2020. Available online: <https://groundcover.grdc.com.au/weeds-pests-diseases/weeds/tackling-the-problematic-lupin-and-weeds-that-give-wa-growers-the-blues> (accessed on 21 June 2022).
4. Thomas, G. DAW665-Advanced Management Strategies for Control of Anthracnose and Brown Spot in Lupins-GRDC. Available online: <https://grdc.com.au/research/reports/report?id=376> (accessed on 21 June 2022).
5. Lucas, M.M.; Stoddard, F.L.; Annicchiarico, P.; Frías, J.; Martínez-Villaluenga, C.; Sussmann, D.; Duranti, M.; Seger, A.; Zander, P.M.; Pueyo, J.J. The future of lupin as a protein crop in Europe. *Front. Plant Sci.* **2015**, *6*, 705. [CrossRef] [PubMed]
6. Pollard, N.J.; Stoddard, F.L.; Popineau, Y.; Wrigley, C.W.; MacRitchie, F. Lupin flours as additives: Dough mixing, breadmaking, emulsifying, and foaming. *Cereal Chem.* **2002**, *79*, 662–669. [CrossRef]
7. López-Granados, F.; Torres-Sánchez, J.; De Castro, A.-I.; Serrano-Pérez, A.; Mesas-Carrascosa, F.J.; Peña, J.M. Object-based early monitoring of a grass weed in a grass crop using high resolution UAV imagery. *Agron. Sustain. Dev.* **2016**, *36*, 67. [CrossRef]
8. Zhang, N.; Wang, M.; Wang, N. Precision agriculture—A worldwide overview. *Comput. Electron. Agric.* **2002**, *36*, 113–132. [CrossRef]
9. Dammer, K.-H.; Wartenberg, G. Sensor-based weed detection and application of variable herbicide rates in real time. *Crop. Prot.* **2007**, *26*, 270–277. [CrossRef]
10. Hunter, J.E.; Gannon, T.W.; Richardson, R.J.; Yelverton, F.H.; Leon, R.G. Integration of remote-weed mapping and an autonomous spraying unmanned aerial vehicle for site-specific weed management. *Pest Manag. Sci.* **2020**, *76*, 1386–1392. [CrossRef]
11. Che'Ya, N.N.; Dunwoody, E.; Gupta, M. Assessment of weed classification using hyperspectral reflectance and optimal multispectral UAV imagery. *Agronomy* **2021**, *11*, 1435. [CrossRef]
12. Huang, Y.; Lee, M.A.; Thomson, S.J.; Reddy, K.N. Ground-based hyperspectral remote sensing for weed management in crop production. *Int. J. Agric. Biol. Eng.* **2016**, *9*, 98–109.
13. Shahbazi, N.; Flower, K.C.; Callow, J.N.; Mian, A.; Ashworth, M.B.; Beckie, H.J. Comparison of crop and weed height, for potential differentiation of weed patches at harvest. *Weed Res.* **2021**, *61*, 25–34. [CrossRef]
14. Bosilj, P.; Duckett, T.; Cielniak, G. Analysis of Morphology-Based Features for Classification of Crop and Weeds in Precision Agriculture. *IEEE Robot. Autom. Lett.* **2018**, *3*, 2950–2956. [CrossRef]
15. Sanders, J.T.; Jones, E.A.L.; Minter, A.; Austin, R.; Roberson, G.T.; Richardson, R.J.; Everman, W.J. Remote Sensing for Italian Ryegrass [*Lolium perenne* L. ssp. *multiflorum* (Lam.) Husnot] Detection in Winter Wheat (*Triticum aestivum* L.). *Front. Agron.* **2021**, *3*, 687112. [CrossRef]
16. Zhang, Y.; Gao, J.; Cen, H.; Lu, Y.; Yu, X.; He, Y.; Pieters, J.G. Automated spectral feature extraction from hyperspectral images to differentiate weedy rice and barnyard grass from a rice crop. *Comput. Electron. Agric.* **2019**, *159*, 42–49. [CrossRef]
17. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
18. Dos Santos Ferreira, A.; Freitas, D.M.; da Silva, G.G.; Pistori, H.; Folhes, M.T. Weed detection in soybean crops using ConvNets. *Comput. Electron. Agric.* **2017**, *143*, 314–324. [CrossRef]
19. Gao, J.; French, A.P.; Pound, M.P.; He, Y.; Pridmore, T.P.; Pieters, J.G. Deep convolutional neural networks for image-based *Convolvulus sepium* detection in sugar beet fields. *Plant Methods* **2020**, *16*, 29. [CrossRef]
20. de Camargo, T.; Schirrmann, M.; Landwehr, N.; Dammer, K.-H.; Pflanz, M. Optimized deep learning model as a basis for fast UAV mapping of weed species in winter wheat crops. *Remote Sens.* **2021**, *13*, 1704. [CrossRef]
21. Lottes, P.; Behley, J.; Chebrolu, N.; Milioto, A.; Stachniss, C. Robust joint stem detection and crop-weed classification using image sequences for plant-specific treatment in precision farming. *J. Field Robot.* **2020**, *37*, 20–34. [CrossRef]
22. Ma, X.; Deng, X.; Qi, L.; Jiang, Y.; Li, H.; Wang, Y.; Xing, X. Fully convolutional network for rice seedling and weed image segmentation at the seedling stage in paddy fields. *PLoS ONE* **2019**, *14*, e0215676. [CrossRef]
23. Partel, V.; Kakarla, S.C.; Ampatzidis, Y. Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence. *Comput. Electron. Agric.* **2019**, *157*, 339–350. [CrossRef]
24. Fawakherji, M.; Youssef, A.; Bloisi, D.; Pretto, A.; Nardi, D. Crop and Weeds Classification for Precision Agriculture Using Context-Independent Pixel-Wise Segmentation. In Proceedings of the 2019 Third IEEE International Conference on Robotic Computing (IRC), Naples, Italy, 25–27 February 2019; pp. 146–152. [CrossRef]
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
26. Kurtzer, G.M.; Sochat, V.; Bauer, M.W. Singularity: Scientific containers for mobility of compute. *PLoS ONE* **2017**, *12*, e0177459. [CrossRef] [PubMed]
27. Weed-AI. Available online: <https://weed-ai.sydney.edu.au/> (accessed on 25 January 2022).
28. Walker, J.; Hertel, K.; Parker, P.; Edwards, J. *Lupin Growth and Development*; Munroe, A., Ed.; Industry & Investment NSW: Sydney, NSW, Australia, 2011.
29. Danilevicz, M.F.; Bayer, P.E.; Boussaid, F.; Bennamoun, M.; Edwards, D. Maize yield prediction at an early developmental stage using multispectral images and genotype data for preliminary hybrid selection. *Remote Sens.* **2021**, *13*, 3976. [CrossRef]

30. Anderson, S.L.; Murray, S.C.; Malambo, L.; Ratcliff, C.; Popescu, S.; Cope, D.; Chang, A.; Jung, J.; Thomasson, J.A. Prediction of Maize Grain Yield before Maturity Using Improved Temporal Height Estimates of Unmanned Aerial Systems. *Plant Phenome J.* **2019**, *2*, 1–15. [[CrossRef](#)]
31. Kataoka, T.; Kaneko, T.; Okamoto, H.; Hata, S. Crop growth estimation system using machine vision. In Proceedings of the 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM 2003), Kobe, Japan, 20–24 July 2003; pp. b1079–b1083.
32. Xu, X.; Xu, S.; Jin, L.; Song, E. Characteristic analysis of Otsu threshold and its applications. *Pattern Recognit. Lett.* **2011**, *32*, 956–961. [[CrossRef](#)]
33. Bradski, G. The OpenCV library. *Dr. Dobb's J. Softw. Tools Prof. Program.* **2000**, *25*, 120–123.
34. Skalski, P. Make Sense. Available online: <https://github.com/SkalskiP/make-sense/> (accessed on 25 January 2022).
35. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Lecture Notes in Computer Science; Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Volume 9351, pp. 234–241.
36. Howard, J.; Gugger, S. Fastai: A Layered API for Deep Learning. *Information* **2020**, *11*, 108. [[CrossRef](#)]
37. Shi, W.; Caballero, J.; Huszar, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. *arXiv* **2016**. [[CrossRef](#)]
38. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
39. Sudre, C.H.; Li, W.; Vercauteren, T.; Ourselin, S.; Jorge Cardoso, M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Lecture Notes in Computer Science; Cardoso, M.J., Arbel, T., Carneiro, G., Syeda-Mahmood, T., Tavares, J.M.R.S., Moradi, M., Bradley, A., Greenspan, H., Papa, J.P., Madabhushi, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; Volume 10553, pp. 240–248.
40. PyTorch | NVIDIA NGC. Available online: <https://catalog.ngc.nvidia.com/orgs/nvidia/containers/pytorch/tags> (accessed on 15 December 2022).
41. Fabian, P.; Gael, V.; Alexandre, G.; Vincent, M.; Bertrand, T. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
42. Opitz, J.; Burst, S. Macro F1 and Macro F1. *arXiv* **2019**. [[CrossRef](#)]
43. Sa, I.; Popović, M.; Khanna, R.; Chen, Z.; Lottes, P.; Liebisch, F.; Nieto, J.; Stachniss, C.; Walter, A.; Siegwart, R. WeedMap: A Large-Scale Semantic Weed Mapping Framework Using Aerial Multispectral Imaging and Deep Neural Network for Precision Farming. *Remote Sens.* **2018**, *10*, 1423. [[CrossRef](#)]
44. Wu, Z.; Chen, Y.; Zhao, B.; Kang, X.; Ding, Y. Review of Weed Detection Methods Based on Computer Vision. *Sensors* **2021**, *21*, 3647. [[CrossRef](#)] [[PubMed](#)]
45. Su, D.; Qiao, Y.; Kong, H.; Sukkarieh, S. Real time detection of inter-row ryegrass in wheat farms using deep learning. *Biosyst. Eng.* **2021**, *204*, 198–211. [[CrossRef](#)]
46. Le, V.N.T.; Ahderom, S.; Alameh, K. Performances of the LBP Based Algorithm over CNN Models for Detecting Crops and Weeds with Similar Morphologies. *Sensors* **2020**, *20*, 2193. [[CrossRef](#)] [[PubMed](#)]
47. Sapkota, B.; Singh, V.; Neely, C.; Rajan, N.; Bagavathiannan, M. Detection of Italian Ryegrass in Wheat and Prediction of Competitive Interactions Using Remote-Sensing and Machine-Learning Techniques. *Remote Sens.* **2020**, *12*, 2977. [[CrossRef](#)]
48. Girma, K.; Mosali, J.; Raun, W.R.; Freeman, K.W.; Martin, K.L.; Solie, J.B.; Stone, M.L. Identification of optical spectral signatures for detecting cheat and ryegrass in winter wheat. *Crop Sci.* **2005**, *45*, 477–485. [[CrossRef](#)]
49. GRDC GROWNOTES: *Lupin Western*; GRDC: Sydney, NSW, Australia, 2017.
50. Prichard, J.M.; Forseth, I.N. Rapid leaf movement, microclimate, and water relations of two temperate legumes in three contrasting habitats. *Am. J. Bot.* **1988**, *75*, 1201. [[CrossRef](#)]
51. Fu, Q.A.; Ehleringer, J.R. Heliotropic leaf movements in common beans controlled by air temperature. *Plant Physiol.* **1989**, *91*, 1162–1167. [[CrossRef](#)]
52. Jaffe, M.J. On heliotropism in tendrils of *Pisum sativum*: A response to infrared irradiation. *Planta* **1970**, *92*, 146–151. [[CrossRef](#)]
53. Grant, R.H. Potential effect of soybean heliotropism on ultraviolet-b irradiance and dose. *Agron. J.* **1999**, *91*, 1017–1023. [[CrossRef](#)]
54. Atamian, H.S.; Creux, N.M.; Brown, E.A.; Garner, A.G.; Blackman, B.K.; Harmer, S.L. Circadian regulation of sunflower heliotropism, floral orientation, and pollinator visits. *Science* **2016**, *353*, 587–590. [[CrossRef](#)] [[PubMed](#)]
55. López-Granados, F.; Torres-Sánchez, J.; Serrano-Pérez, A.; De Castro, A.I.; Mesas-Carrascosa, F.-J.; Peña, J.M. Early season weed mapping in sunflower using UAV technology: Variability of herbicide treatment maps against weed thresholds. *Precis. Agric.* **2016**, *17*, 183–199. [[CrossRef](#)]
56. Porsild, A.E.; Harington, C.R.; Mulligan, G.A. *Lupinus arcticus* Wats. Grown from Seeds of Pleistocene Age. *Science* **1967**, *158*, 113–114. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.