remote sensing

MDPI

Article

African Lovegrass Segmentation with Artificial Intelligence Using UAS-Based Multispectral and Hyperspectral Imagery

Pirunthan Keerthinathan [1,2,*], Narmilan Amarasingam [1,2], Jane E. Kelly [2], Nicolas Mandel [2], Remy L. Dehaan [2], Lihong Zheng [2], Grant Hamilton [3] and Felipe Gonzalez [1]

1 QUT Centre for Robotics, School of Electrical Engineering and Robotics, Faculty of Engineering, Queensland University of Technology (QUT), 2 George Street, Brisbane City, QLD 4000, Australia; narmilan.amarasingam@hdr.qut.edu.au (N.A.); felipe.gonzalez@qut.edu.au (F.G.)
2 Gulbali Institute for Agriculture Water and Environment, Charles Sturt University, Boorooma Street, Wagga Wagga, NSW 2678, Australia; janekelly@csu.edu.au (J.E.K.); nmandel@csu.edu.au (N.M.); rdehaan@csu.edu.au (R.L.D.); lzheng@csu.edu.au (L.Z.)
3 School of Biology and Environmental Science, Faculty of Science, Queensland University of Technology (QUT), 2 George Street, Brisbane City, QLD 4000, Australia; g.hamilton@qut.edu.au
* Correspondence: pirunthan.keerthinathan@hdr.qut.edu.au; Tel.: +61-432386254

Abstract: The prevalence of the invasive species African Lovegrass (*Eragrostis curvula*, ALG thereafter) in Australian landscapes presents significant challenges for land managers, including agricultural losses, reduced native species diversity, and heightened bushfire risks. Uncrewed aerial system (UAS) remote sensing combined with AI algorithms offer a powerful tool for accurately mapping the spatial distribution of invasive species and facilitating effective management strategies. However, segmentation of vegetations within mixed grassland ecosystems presents challenges due to spatial heterogeneity, spectral similarity, and seasonal variability. The performance of state-of-the-art artificial intelligence (AI) algorithms in detecting ALG in the Australian landscape remains unknown. This study compared the performance of four supervised AI models for segmenting ALG using multispectral (MS) imagery at four sites and developed segmentation models for two different seasonal conditions. UAS surveys were conducted at four sites in New South Wales, Australia. Two of the four sites were surveyed in two distinct seasons (flowering and vegetative), each comprised of different data collection settings. A comparative analysis was also conducted between hyperspectral (HS) and MS imagery at a single site within the flowering season. Of the five AI models developed (XGBoost, RF, SVM, CNN, and U-Net), XGBoost and the customized CNN model achieved the highest validation accuracy at 99%. The AI model testing used two approaches: quadrat-based ALG proportion prediction for mixed environments and pixel-wise classification in masked regions where ALG and other classes could be confidently differentiated. Quadrat-based ALG proportion ground truth values were compared against the prediction for the custom CNN model, resulting in 5.77% and 12.9% RMSE for the seasons, respectively, emphasizing the superiority of the custom CNN model over other AI algorithms. The comparison of the U-Net demonstrated that the developed CNN effectively captures ALG without requiring the more intricate architecture of U-Net. Masked-based testing results also showed higher F1 scores, with 91.68% for the flowering season and 90.61% for the vegetative season. Models trained on single-season data exhibited decreased performance when evaluated on data from a different season with varying collection settings. Integrating data from both seasons during training resulted in a reduction in error for out-of-season predictions, suggesting improved generalizability through multi-season data integration. Moreover, HS and MS predictions using the custom CNN model achieved similar test results with around 20% RMSE compared to the ground truth proportion, highlighting the practicality of MS imagery over HS due to operational limitations. Integrating AI with UAS for ALG segmentation shows great promise for biodiversity conservation in Australian landscapes by facilitating more effective and sustainable management strategies for controlling ALG spread.

check for updates

Citation: Keerthinathan, P.; Amarasingam, N.; Kelly, J.E.; Mandel, N.; Dehaan, R.L.; Zheng, L.; Hamilton, G.; Gonzalez, F. African Lovegrass Segmentation with Artificial Intelligence Using UAS-Based Multispectral and Hyperspectral Imagery. *Remote Sens.* **2024**, *16*, 2363. https://doi.org/10.3390/rs16132363

Academic Editor: John Dymond

Received: 20 May 2024
Revised: 20 June 2024
Accepted: 25 June 2024
Published: 27 June 2024

Copyright: © 2024 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/).

Keywords: non-native; UAV; drone; wildfire; fire-prone

1. Introduction

African Lovegrass (ALG), also as known as Eragrostis curvula, is a perennial C_4 grass indigenous to Africa [1]. The species is found in most states and territories throughout Australia and contributes to both economic and environmental degradation of landscapes in these areas [2]. The initial introduction of this species to Australia was aimed at meeting the needs of soil preservation and pastoral use for livestock [3]. However, unpalatability and low crude protein content soon rendered the species unfavorable for livestock production [1]. The natural tendency of ALG to form dense swards enhances its competitive edge over other species. This often results in its dominance within grasslands and poses significant challenges, including agricultural losses, diminished diversity among native species, and most importantly, the escalation of bushfire risks attributed to its role as a fuel source [4]. Grasslands that are overgrown with invasive grasses are often more susceptible to fire. Grasslands primarily occupied by non-native grass species contain twice as much fuel as grasslands infiltrated by indigenous grass species. It can exhibit fire intensities three times higher than native species and show significant variability in fire duration [4]. It is therefore advantageous to keep native grasslands free of invasive ALG to protect the environment and to reduce the risk of fire.

Various existing methods for controlling ALG have been trialed, include grazing, burning, and herbicide application [2]. However, for these control measures to be effective, species locations must be must accurately identified. Precise mapping of ALG spatial distribution and determination of its severity or density is essential for informed decision-making and successful implementation of control strategies. Remote sensing (RS) offers promising capabilities for delivering timely and precise data regarding invasive weed infestations. Among the various RS methods, uncrewed aerial systems (UAS) demonstrate significant capability for acquiring data with high spatial and temporal resolution [5]. Therefore, using imagery from UAS may be used effectively for mapping the spatial distribution of the rapidly spreading ALG.

Mounted sensors on UAS offer diverse datasets for vegetation segmentation, mainly sourced from Red-Green-Blue (RGB) and multispectral (MS) cameras and a limited number using hyperspectral (HS) cameras. In the context of invasive species segmentation, improved spectral and spatial resolution significantly enhances the capabilities of artificial intelligence (AI) algorithms in RS and image analysis [6]. While Amarasingam et al. (2024) [7] summarized a number of research studies that focused on the segmentation of weeds using aerial imagery with classical machine learning (ML) and deep learning (DL) techniques, Table 1 shows the performance of models with respect to the imagery type and the spatial resolution in sparse invasive vegetation identification. MS-based ML models tend to demonstrate higher accuracy compared to RGB models in instances of lower spatial resolution because they offer more spectral information [7,8]. The fusion of MS and RGB imagery also shows great promise for enhancing the accuracy detection [9], yet further investigation is warranted to determine the value of HS and MS fusion for detecting a broader range of species.

Despite the availability of studies investigating the use of UAS imagery to detect invasive plants in recent years, there has been little focus on ALG, particularly in relation to the performance of ML algorithms for ALG detection. In addition, the potential impact of using HS imagery over MS imagery in identifying key feature bands and improving detection models for fire-prone grass species in mixed and sparse ecosystems remains unexplored [5]. To address these research gaps, this study used classical supervised ML algorithms and UAS-generated MS and HS imagery to detect and map the spatial distribution of ALG at four infested sites near Cooma, New South Wales, Australia. The three main objectives of this study were to (1) compare the performance of supervised

AI models for ALG segmentation from MS imagery, including extreme gradient boosting (XGBoost), random forests (RF), support vector machine (SVM), a customized convolutional neural network (CNN), and a state-of-the-art deep learning algorithm, U-Net [7], which has demonstrated the best results among recent invasive species detection studies, as shown in Table 1; (2) develop ALG segmentation models for two different seasonal conditions in different data collection settings; and (3) conduct a comparative analysis between HS and MS imagery in the segmentation of ALG, aiming to elucidate the strengths and limitations of each imaging modality for vegetation classification within grassland ecosystems.

Table 1. Recent studies on weed and invasive plant detection using UAS-generated imagery and AI algorithms.

Invasive Species	Imagery Type	ML Models	Performance	Spatial Resolution (cm/Pixel)	Reference
Bitou bush	RGB	MLPR	OA: 82%	3	[10]
Bitou bush	RGB	ANN	OA: 88–97%	1–2	[8]
Weeds in rice field	RGB	Magenet	OA: 77.5%	0.3	[11]
Bitou bush	MS	U-net	OA: 98%	2.2	[7]
Milk thistle weed	MS	SVM	OA: 96%	50	[12]
Amaranth, Pigweed, and Mallow weed	MS	NN and OBIA	OA: 92%	0.543	[6]
Weeds in sugar beet fields	MS	DeepMultiFuse	F1 score: 85.6–99%	1	[13]
Weeds in lettuce field	MS	YOLOv3, and R–CNN	OA: 89%	0.22	[14]
Weeds in sugar beet fields	MS	SegNet	OA: 57.6–86.3%	0.85–1.181	[15]
blackgrass weed	MS	RF	OA: 93%	1.16	[16]
barnyard grass and velvetleaf	MS and RGB fusion	DCNN	OA: 81.1–92.4%	0.41	[9]
Johnsongrass	HS	SAM and SMA	OA: 60–80%	2000	[17]
Common milkweed	HS	SVM and ANN	OA: 92.95–99.61%	40	[18]
Bitou bush	HS	SVM and XGB	OA: 86%	3.5	[7]

ML: machine learning, MS: multispectral, HS: hyperspectral, RGB: Red-Green-Blue, ANN: artificial neural network, SVM: support vector machine, RF: random forest, XGB: eXtreme Gradient Boosting, NN: neural network, OA: overall accuracy.

2. Materials and Methods

2.1. Experimental Design

The experimental design, depicted in Figure 1, comprises five integral components: data acquisition, preprocessing, pixel-wise labeling, MS-based prediction, and MS and HS comparison. Aerial surveys via UAS were conducted across four sites using MS and HS sensors, accompanied by ground truth collection within the regions of interest within these sites. MS-based ALG prediction was conducted separately for two seasonal conditions using five AI algorithms. For each season, a different AI model was developed to estimate the spatial distribution and the predicted ALG proportion was evaluated against the ground truth assessment. A comprehensive evaluation of HS and MS imagery for ALG identification using the benchmarked algorithm was then conducted at a single site where both HS and MS data were available. This facilitated a comparison of both imagery across various dimensions, including accuracy, cost-effectiveness, and operational constraints.

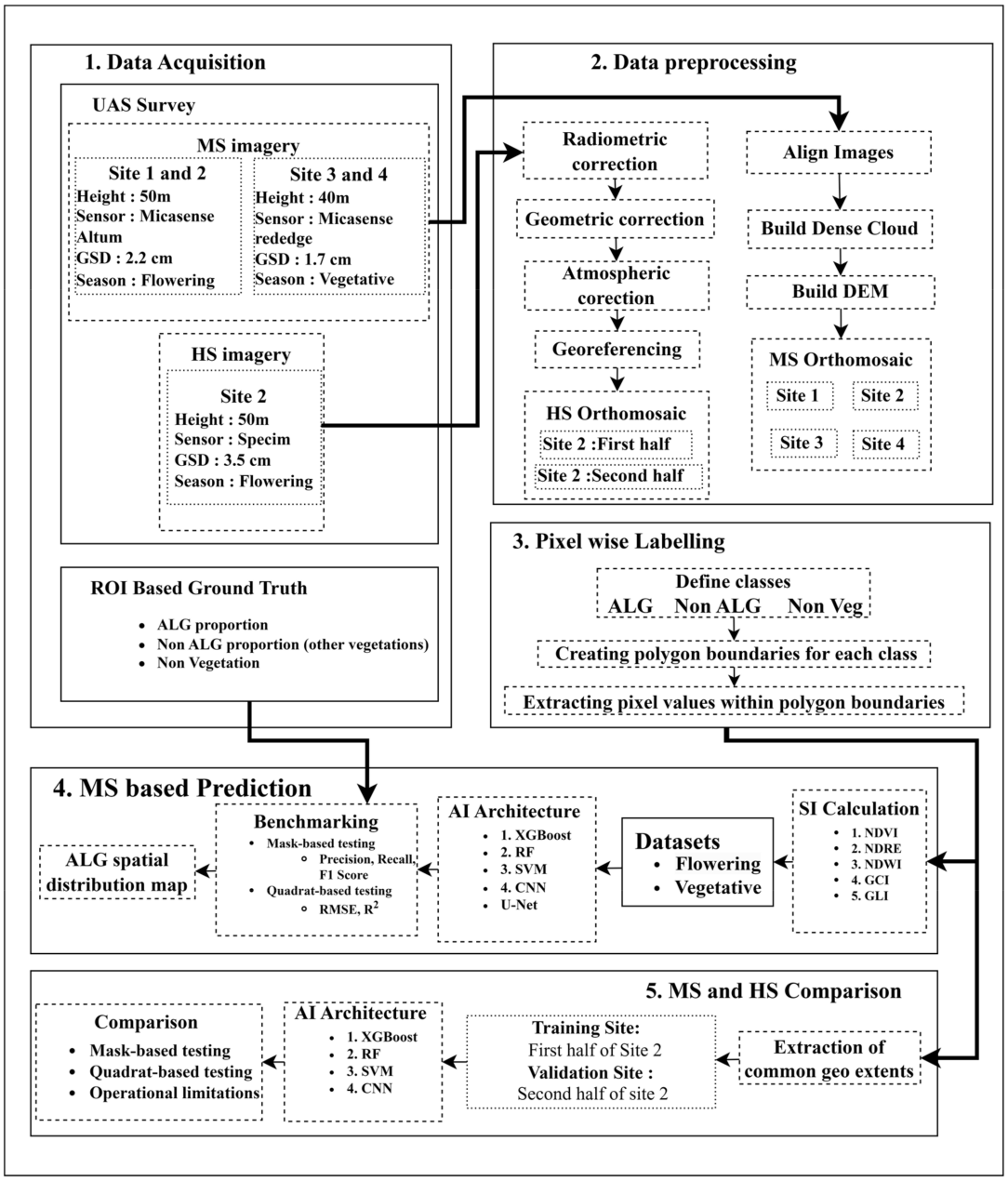


Figure 1. Overview of the study methodology, illustrating the key steps in data acquisition, data preprocessing, pixel-wise labeling, multispectral-based prediction, and multispectral and hyperspectral comparison.

2.2. Study Site

Two sites from each of two locations at Cooma, New South Wales, Australia (Figure 2), were selected for each survey (entitled "Bunyan" and "Cooma" sites). Specifically, Site 1 and Site 2 represent Bunyan sites, while Site 3 and Site 4 correspond to Cooma sites. These locations feature mixed grassland ecosystems with coverage extending over an area of interest measuring 8.8 hectares. The Bunyan sites were comprised of ALG and a range of other species, including *Vulpia* ssp. (Rats Tail Fescue), *Hordeum glaucum* (Barley grass) and *Rumex acetosa* (Sorrel). The botanical composition of the grassland at the Cooma sites also included ALG, in addition to *Stipa* ssp., *Poa annua, Bothriochloa macra* (Redgrass), and *Avena fatua* (Wild Oats). At Site 1 and Site 4, separate areas of both ALG and non-ALG vegetation were observed, often in distinct patches. Site 3 displayed fewer areas dominated solely

by ALG, with sparse occurrences. At Site 2, ALG was observed to be widely distributed, dominating the landscape.

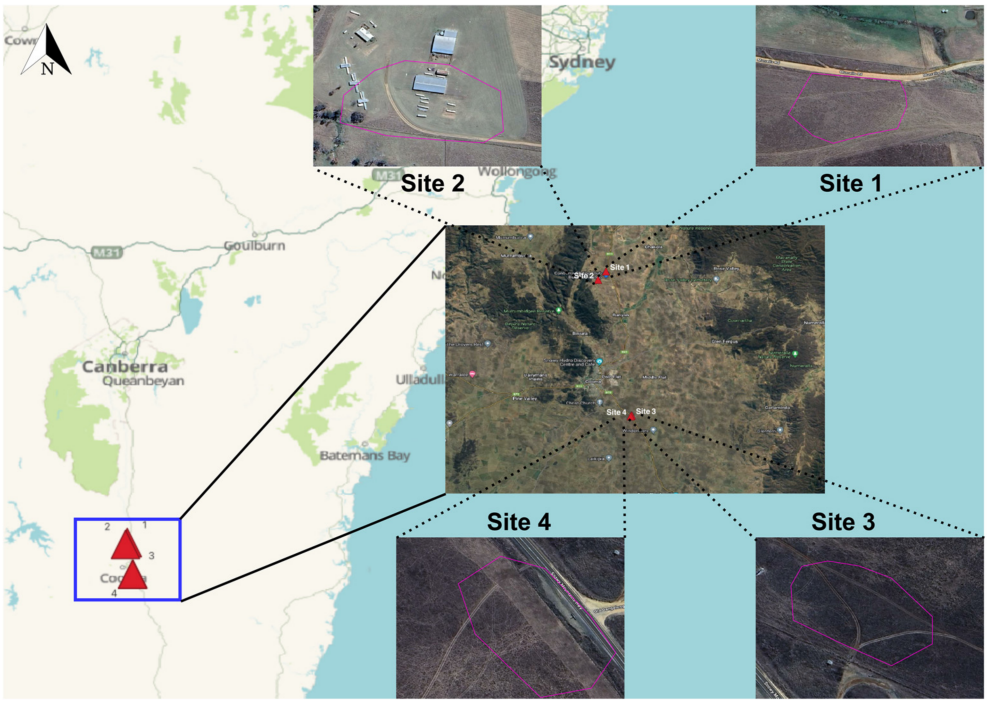


Figure 2. Map of the study sites. Site 1 and Site 2 correspond to Bunyan sites, while Site 3 and Site 4 correspond to Cooma sites, located in in New South Wales, Australia.

Sites were chosen to capture contrasting seasonal states at the time of data acquisition and were representative of infested paddocks seen in the region in relation to botanical composition and management. The Bunyan site data represents the flowering season. During this period, the target plants exhibit fully developed floral structures, often appearing as purplish or brownish hues in the imagery. Conversely, the Cooma site data, collected in December 2023, depicts the vegetative season. This season is characterized by the absence of flowering structures and the dominance of green foliage.

2.3. Data Acquisition

2.3.1. UAS Survey

UAS missions employed a MicaSense Altum and MicaSense RedEdge MS sensors mounted on a DJI Matrice 300 to capture MS imagery in the Bunyan and Cooma sites, respectively, and a Specim AFX VNIR sensor mounted on a DJI M600 for HS imagery collection at Site 2 only. Both MS sensors yield data across five bands, spanning the wavelengths of red, green, blue, red edge, and near-infrared (NIR). Conversely, the Specim AFX VNIR provides HS data, covering a continuous electromagnetic spectrum ranging from 400 to 1000 nm and comprising 448 distinct bands.

Data was acquired at the Bunyan and Cooma Sites during December 2022 and 2023, respectively. The MS sensing flights at the Bunyan sites were conducted at an altitude of 50 m above ground level, resulting in a ground sample distance (GSD) of 2.2 cm/pixel. At the Cooma sites, the UAS mission was flown at an altitude of 40 m above ground level, producing a GSD of 1.7 cm/pixel. HS sensing at Site 2 was performed at an altitude of 50 m, yielding a GSD of 3.5 cm/pixel. Table 2 presents the details of the collected data and the weather conditions recorded during each data collection period.

Table 2. Sensor, flight details, and weather conditions recorded during the UAS missions at Bunyan and Cooma sites.

Sites	Bunyan: Site 1 and Site 2	Cooma: Site 3 and Site 4
Seasonal Condition	Flowering: Cool, Wet, High Wind	Vegetative: Warm, Sunny, High Wind
Date and Time	Site 1: 13 December 2022, 1:00 p.m.–1:19 p.m. Site 2: 14 December 2022, 9:42 a.m.–12:57 p.m.	Site 3: 5 December 2023, 3:42 p.m.–3:45 p.m. Site 4: 5 December 2023, 3.51 p.m.–4.00 p.m.
Data Source	MS: MicaSense Altum HS: Specim AFX VNIR	MS: MicaSense RedEdge
Flight altitude (m)	50	40
Resolution (cm/pixel)	MS: 2.2 HS: 3.5	MS: 1.7
Temperature (°C)	6–12	28
Average Wind Speed ($m \cdot s^{-1}$)	12	8
Total Precipitation (mm)	3	0
Cloud Cover (%)	75	10

MS: multispectral, HS: hyperspectral.

2.3.2. Ground Truth

Square quadrats (1 m × 1 m) were randomly positioned around each study site to identify sample areas for ground truth assessment. The quadrats (sampled area) were constructed by placing PVC frames and distinctive striped tape to form squares measuring 1 m in length. A total of 10 quadrats were used for recording ground truth assessments at sites 1 and 2, while 19 and 20 quadrats were used for sites 3 and 4, respectively, to capture the wider variety of species. The quadrats were inspected by weed experts on site prior to the UAS surveys. Ground truth assessments consisted of recording species presence, average plant height, and percentage cover of ALG and other species. The proportion of area covered by ALG within each quadrat serves as the ground truth for model evaluation. This value is estimated by experts who analyze the top view of the quadrats. Their analysis considers the area occupied by ALG alongside other plant species and non-vegetation zones. Figure 3 illustrates the species diversity among the quadrats from all four sites. Close-up photos of each identified species within the quadrats were captured to facilitate later labeling of imagery.

2.4. Data Pre-Processing

2.4.1. Orthomosaic Generation

The orthomosaic generation process utilized a general methodology presented in other studies [7]. Both MS and HS images taken from UAS with geotags underwent preprocessing to generate a georeferenced reflectance orthomosaic. MS images were converted to reflectance using in-field calibration panels. Point clouds, digital elevation models (DEMs), and reflectance images were created using Agisoft Metashape 1.6.6 (Agisoft LLC, Petersburg, Russia). HS data followed a separate workflow. Radiometric calibration converted the raw digital numbers into radiance values using SPECIM Caligeo Pro. This calibration relied on on-board calibration files obtained during image capture. Following this, the radiance data was transformed into reflectance using DROACOR physical atmospheric correction. The MS orthomosaic of Site 1 covered a 58,915 m^2 area with approximately 121,734,836 pixels, Site 2 covered 19,863 m^2 with approximately 41,036,647 pixels, Site 3 covered 4296 m^2 with 14,865,678 pixels, and Site 4 covered 9570 m^2 with approximately 33,116,387 pixels. The HS orthomosaic of Site 2 covered a 7139 m^2 area with 5,829,387 pixels.

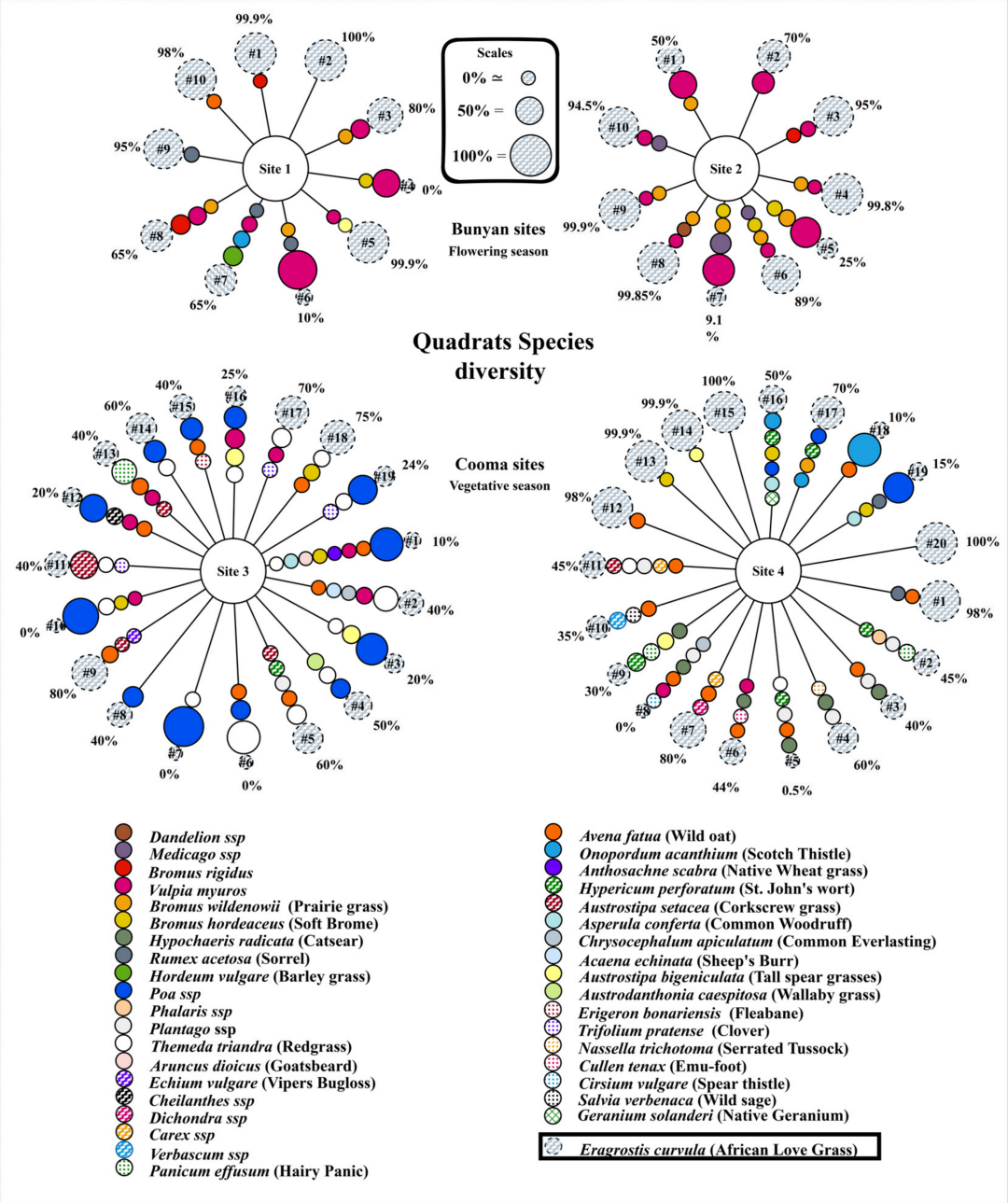


Figure 3. Illustration of quadrat species diversity at Bunyan and Cooma sites.

2.4.2. Georeferencing

Georeferencing ensures that both MS and HS geospatial data are accurately positioned on the Earth's surface, enabling consistent labeling and extraction of information for training AI algorithms. The methodology involves manual georeferencing of the HS orthomosaic from each site by aligning it with their corresponding MS orthomosaic using QGIS 3.2.8 Firenze software. All data was captured using the world geodetic system (WGS-84) datum at all sites. Ground Control Points (GCP) were marked on both MS and HS by using the corners of quadrats placed randomly throughout the sites. The cubic spline interpolation method was employed to align and transform spatial data, providing a smoother and more flexible fit to the GCPs. This method better accommodates irregularities in the data [19], especially as it comes from various flights and sensors.

2.5. Pixel-Wise Labeling

Pixel-wise labeling was carried out across three classes, ALG, non-ALG, and non-vegetation, using both manual and semi-automated approaches [7]. Weed experts collaborated on manual labeling, focusing on the MS orthomosaic maps due to their higher GSD compared to HS maps. Due to the challenge of visually separating ALG from the background using 2.2 cm GSD MS imagery alone, quadrat-based ground truths, along with close-up images taken during the survey, were used for labeling the MS imagery. The close-up images, with a resolution of 0.3 mm, were inspected by labelers. Both the top and side view images helped identify the locations of ALG, non-ALG, and non-vegetation classes in each quadrat. Regions containing ALG, other species, and non-vegetated areas were marked by geo vector polygons over the MS orthomosaic in a Geographic Information Systems (GIS) environment (QGIS 3.28 Firenze) (Figure 4). The raster pixels of the georeferenced MS orthomosaic imagery within the extracted polygons were used to obtain the spectral band values for each class. Manual labeling is a challenging task, even for experts, due to the high visual similarity with other grass species. Therefore, only areas where there was confidence in identifying ALG, non-ALG, and non-vegetation were labeled. In addition to manual labeling, a semi-automatic labeling approach was implemented, applying a normalized difference vegetation index (NDVI)-based threshold on quadrats where only ALG or non-ALG was present to exclude bare land and prevent mislabeling. Regions where the NDVI value exceeded 0.35 were considered vegetation (ALG or non-ALG). If no ALG was recorded within the quadrat, all filtered pixels were assigned to the non-ALG class.

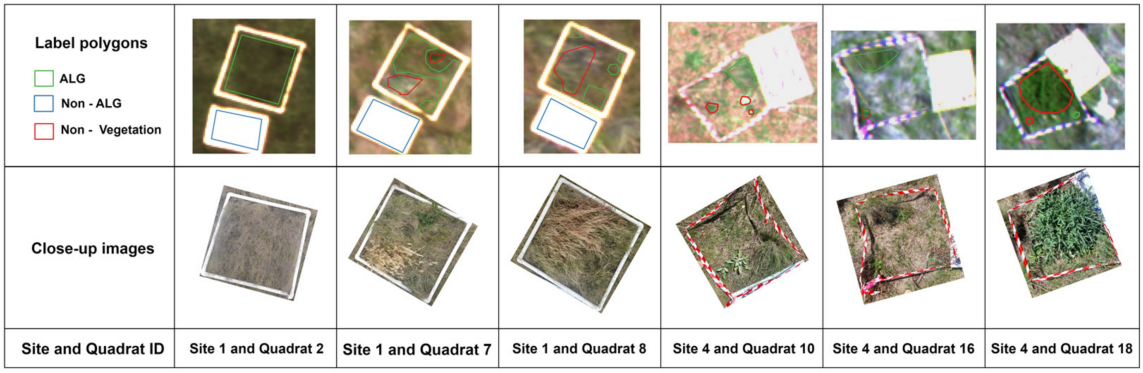


Figure 4. Labeled polygons of three randomly selected quadrats from Sites 1 and 4, along with their corresponding close-up images.

2.6. Model Training for Multispectral-Based African Lovegrass Segmentation

Labeled images exhibited limited pattern-texture variation for ALG and other grass species at the resolution of MS imagery. Grass species often appear intertwined, posing challenges for pattern-based texture analysis. MS imagery captures the reflectance of different wavelengths of light, providing valuable insights into vegetation composition and proportion, even in cases where texture is not clearly defined. Given the limitations of pattern texture analysis in this context, prioritizing spectral information is crucial. Thus, this study utilized spectral indices including NDVI [20], normalized difference water index (NDWI) [21], normalized difference red edge index (NDRE) [22], green chlorophyll index (GCI) [23], and green leaf index (GLI) [24]. Additionally, each of the five bands was individually divided by the sum of all bands, resulting in five additional spectral indices, namely ALGB, ALGG, ALGR, ALGRE, and ALGNIR. Altogether, 10 spectral indices listed in Table 3 were used to develop models. Figure 5 shows the spectral signature differences for these spectral indices.

Table 3. Spectral Indices used for model development.

Channels	Spectral Indices	Equation
SI1	NDVI	$\frac{NIR-R}{NIR+R}$
SI2	NDWI	$\frac{G-NIR}{G+NIR}$
SI3	GCI	$\frac{NIR}{G}-1$
SI4	GLI	$\frac{2*G-R-B}{2*G+R+B}$
SI5	NDRE	$\frac{NIR-RE}{NIR+RE}$
SI6	ALGB	$\frac{B}{R+G+B+RE+NIR}$
SI7	ALGG	$\frac{G}{R+G+B+RE+NIR}$
SI8	ALGR	$\frac{R}{R+G+B+RE+NIR}$
SI9	ALGRE	$\frac{RE}{R+G+B+RE+NIR}$
SI10	ALGNIR	$\frac{NIR}{R+G+B+RE+NIR}$

NDVI: normalized difference vegetation index, NDWI: normalized difference water index, NDRE: normalized difference red edge index, GCI: green chlorophyll index, GLI: green leaf index, ALGB: ALG blue index, ALGG: ALG green index, ALGR: ALG red index, ALGRE: ALG red-edge index, ALGNIR: ALG near infra-red index. Spectral bands: R: red, G: green, B: blue, NIR: near-infrared, RE: red-edge.

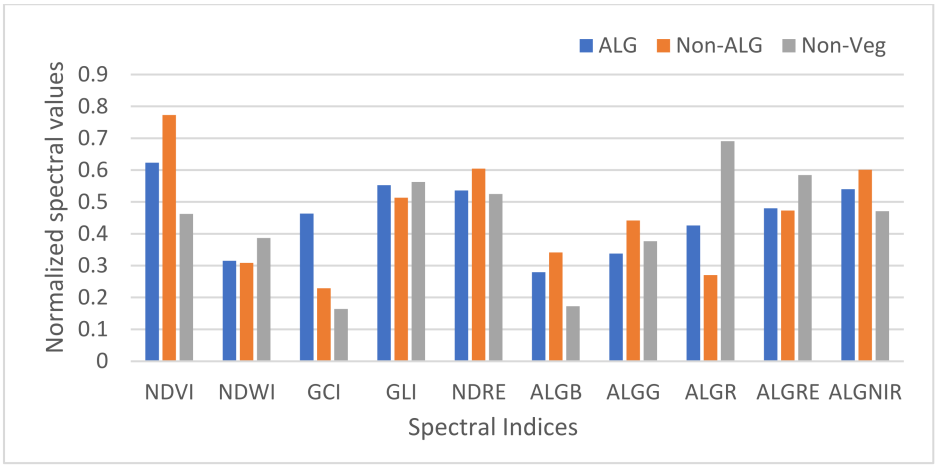


Figure 5. Spectral signature differences for spectral indices.

Our study deployed a customized CNN with a simple architecture, a U-Net model from the study of Amarasingam et al. (2024) [7], and classical ML models for ALG segmentation. This study investigated seasonal variations by developing models for each season (vegetative and flowering) using separate datasets. Additionally, a common model incorporating data from both seasons was developed. Models for the flowering season were developed using Site 1, while models for the vegetative season were developed using Site 4. The common model was developed using Site 1 and Site 4. All three models were then evaluated on test sites from both flowering (Site 3) and vegetative (Site 2) seasons for a comprehensive analysis. Details of these datasets can be found in Table 4. Each pixel labeled under three classes from the sites orthomosaic reflectance maps was integrated with the bands of surrounding pixels and organized into tabular files, comprising 9 pixels and 10 bands per pixel, resulting in 90 features for training. Subsequently, each data point underwent augmentation through 8 rotations (Figure 6). The dataset for model development was partitioned into a training set and a validation set in a ratio of 4:1. A total of 80% of the data was randomly selected for training, and the remaining 20% was allocated to validation. The validation set was evaluated through weighted average precision, recall, and F1 score metrics [25].

Table 4. Details of seasonal condition-based datasets.

Season	Site Location	Model Development Sites	Test Site
Flowering	Bunyan	Site 1	Site 2
Vegetative	Cooma	Site 4	Site 3

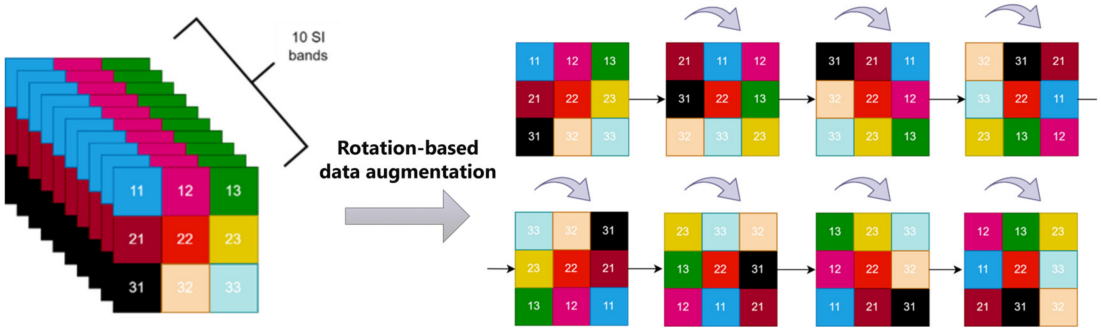


Figure 6. Modelling and augmentation of data points during ALG model development.

2.6.1. Classical Machine Learning Models

This study compared the performance of three classical ML algorithms: SVM, RF, and XGBoost. The ML algorithms used in this study are non-parametric supervised classifiers that have commonly shown promising results in recent invasive species segmentation studies (Table 1). SVM, introduced by Vapnik [26], can perform classification and regression tasks based on transforming the inputs into high-dimensional separatable feature space by kennel functions. SVM excels at finding separation boundaries in high-dimensional data, especially with limited samples, but can be computationally expensive for very large datasets. RF is built from numerous randomized decision trees, providing robustness against overfitting and less sensitivity to noisy data [27]. However, training in a large forest can be resource-intensive. XGBoost gained popularity in recent years due to its high accuracy and efficiency, combining numerous decision trees and employing gradient boosting techniques to enhance performance [28]. Boosting methods achieve high accuracy with features like handling missing values but can be more intricate to tune and interpret compared to the other two algorithms [29].

In developing the ML models, we utilized a combination of Python libraries and tools. The 'xgboost' library provided a powerful implementation of gradient boosting, while the 'SVC' class from 'sklearn.svm' was employed for SVM. Additionally, we developed a RF model using the 'RandomForestClassifier' class from 'sklearn.ensemble'. We evaluated the performance of the models using functions from scikit-learn's 'metrics' module. For data manipulation and analysis, we utilized the 'pandas' library, and the training curve was visualized using 'matplotlib'. Hyperparameter tuning was conducted, and the learning curves were analyzed to determine the optimal stopping point for training, aimed at preventing overfitting and improving the generalization ability of the model to unseen data.

2.6.2. Deep Learning Models

State-of-the-art models for image segmentation and object detection tasks, such as U-net, SegNet, YOLO, are typically not appropriate when images lack clear patterns or structures. These models need a large and varied set of training samples to learn patterns effectively and make accurate predictions. Insufficiently diverse training data, particularly data containing limited structural patterns, can result in suboptimal model performance because the model struggles to capture the complete range of features required for accurate classification [30,31]. Therefore, a customized CNN model was developed, with a focus on incorporating both spectral and spatial information, with higher priority given to spectral relationships, to reduce complexity and enhance applicability to ALG segmentation (Figure 7). This architecture comprises two convolutional layers followed

by two fully connected (dense) layers. The activation function used was leaky-ReLU, and the optimization algorithm chosen was softmax. A 1 × 1 convolutional filter was used to emphasize spectral details while minimizing computational demands needed to capture broader spatial patterns. This dimensionality reduction approach is useful in deep networks to manage computational complexity [32]. Each 1 × 1 convolutional filter combines information across 10 spectral index channels, enabling the network to learn spectral representations. The model was built using the Sequential API from TensorFlow.keras, incorporating convolutional, flattening, and dense layers. Flattening layers were employed to convert data into one-dimensional arrays, and dense layers were used for fully connected connections. The model with the best validation accuracy was saved for further use.

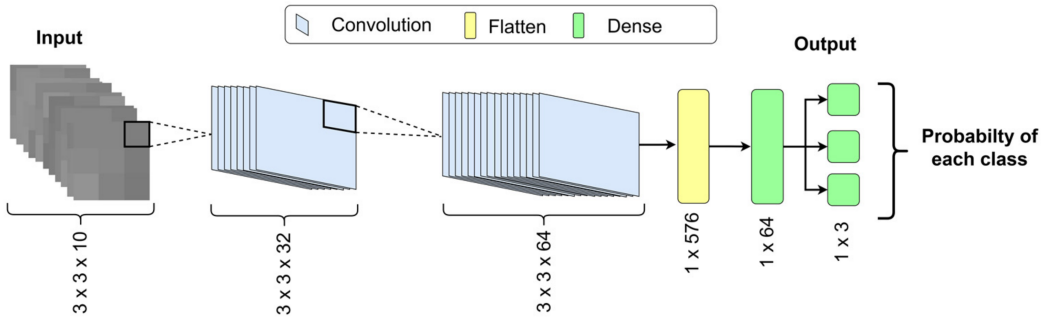


Figure 7. Custom CNN model architecture for MS-based ALG segmentation. For MS and HS comparison, the third dimension of the first layer captures the channel depth, which is 5 for MS and 448 for HS imagery. The remaining dimensions are unchanged.

To compare the developed custom CNN model's performance, we implemented a state-of-the-art DL U-Net model with an architecture based on the work of Amarasingam et al. [7], which has demonstrated superior performance among the recent studies on invasive species detection in mixed environments, as listed in Table 1. We adapted the architecture (Figure 8) by upsampling the input data size (originally 3 × 3) to align with the reference architecture. Additionally, we modified the final layers of the U-Net architecture, adding a last layer with an output depth of three (representing the three classes), followed by a SoftMax activation function to achieve three-class classification.

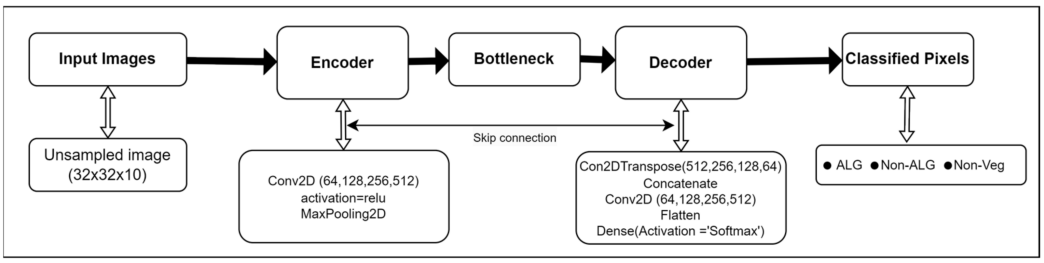


Figure 8. U-Net architecture used for ALG classification.

2.7. Model Training for Hyperspectral-Based African Lovegrass Segmentation

Four AI models, including SVM, RF, XGBoost, and the custom CNN model, were used for HS-based segmentation development. HS orthomosaic imagery from Site 2 was divided into two regions, with labeled data from the first half used for model development and the second half for testing. This study incorporated the surrounding pixels, which encompassed 9-pixel data for each data point, similar to MS-based model development. However, instead of utilizing spectral indices, all band channels, including 448 bands, were normalized using the min-max approach [33] and used for training. The input layer dimensions of the models used in the MS-based study were adjusted to accommodate the differences in input data dimensions. Specifically, the model designed for MS-based segmentation was configured with 10 input channels, whereas the model designed for

HS-based segmentation utilized 448 channels. The remaining architecture of the model was unchanged. The dataset for model development was divided into a training set and a validation set in a ratio of 4:1, similar to the MS-based ALG segmentation study described in Section 2.6.

2.8. Comparison of Prediction Using Multispectral and Hyperspectral Imagery

For comparison purposes, the same models and labeled polygons chosen for HS-based ALG detection were used to develop a separate MS-based model. This comparison model was trained solely on the same labeled georeferenced polygons from Site 2 MS imagery only. The same region used for developing and testing the HS-based ALG segmentation model, as described in Section 2.7, was employed for training and testing this comparison model. The study incorporated surrounding pixels, encompassing nine pixels of data for each data point. Spectral indices were not employed; instead, the five bands from the MS imagery were normalized using a min-max approach and utilized for training, mimicking the exact approach used in HS-based ALG segmentation. Specifically, 5 channels were used for the MS model and 448 channels for the HS model.

2.9. Evaluation Metrics

The testing was conducted using two approaches to facilitate an evaluation of the generalization capability of the models across different data collection settings. The first approach is based on masked regions where ALG and other classes can be confidently differentiated. These regions were used with known labels and the metrics precision, recall, and F1 score [25] were used to evaluate the classification. These metrics were calculated on a per-class basis. The per-class evaluation focused specifically on the ALG class. The second approach involved quadrat-based ALG proportion prediction to assess the performance of the models. This testing is to assess the models' performance in mixed environments, where quadrats lack clear borders for ALG, non-ALG, and non-vegetation areas. The quadrats from the test sites were used to predict the ALG. The ratio of ALG pixels classified from the total pixels was then compared with the ground truth ALG proportion. The root mean square error (RMSE) [34] and correlation (R^2) [35] values for the quadrats in the sites were generated to quantify the robustness of models. Model testing involved comparing the model predictions with the actual ground truth observations within specific quadrats.

3. Results

3.1. Multispectral-Based African Lovegrass Segmentation

3.1.1. Validation Dataset Performance

Table 5 shows the classification report for AI models during the validation. Evaluation metrics, including precision, recall, and F1 score, indicate that the XGBoost and the custom CNN outperformed RF and SVM, achieving 99% overall accuracy for both seasons. U-Net achieved good performance but lagged behind XGBoost and the custom CNN in the vegetative season. RF performed less accurately compared to other models, with the lowest overall accuracy among both seasonal datasets. SVM shows competitive performance during the flowering season, with 98% accuracy; however, its performance decreases during the vegetative season. The models trained on combined seasonal data performed similarly to how they ranked when trained on individual seasons.

Table 5. Classification report for different classical machine learning models including support vector machine (SVM), random forest (RF), U-Net, and the developed custom CNN model under flowering and vegetative seasonal conditions within the validation dataset for ALG. The highest accuracy values for each dataset are highlighted in bold.

Seasonal Dataset	Metrics	RF	SVM	XGB	Custom CNN	U-Net
Flowering	Precision (%)	95.4	97.8	99.8	99.8	99.6
	Recall (%)	94.7	97.8	99.8	99.8	99.6
	F1 Score (%)	94.7	97.8	99.8	99.8	99.6
	Accuracy (%)	95	98	**99**	**99**	**99**
Vegetative	Precision (%)	90.5	95.1	98.4	98.5	97.5
	Recall (%)	90.4	94.9	98.4	98.5	97.3
	F1 Score (%)	90.5	94.9	98.4	98.5	97.4
	Accuracy (%)	90.4	95	**98**	**98**	97.4
Flowering and Vegetative	Precision (%)	90.8	92.3	99.2	99.2	97.6
	Recall (%)	88.2	91.6	99.2	99.1	97.4
	F1 Score (%)	88.2	91.3	99.2	99.2	97.5
	Accuracy (%)	88	92	**99**	**99**	**98**

SVM: support vector machine, RF: random forest, XGB: EXtreme gradient boosting, CNN: convolutional neural network.

3.1.2. Test Dataset Performance

Quadrat-based testing indicates that the custom CNN outperformed XGBoost at the test sites, demonstrating a lower RMSE when compared to the recorded ground truth (Table 6). In the vegetative season, the RMSE of the XGBoost was approximately 1.3 times higher, whereas the custom CNN had three times less error for the flowering season compared to XGBoost. Similarly, the custom CNN demonstrated the best correlation to the prediction and ground truth values. It is evident that the developed custom CNN demonstrated superior performance across both seasons. This is shown by the higher prediction accuracy for the flowering season, displaying a 5.77% RMSE and a correlation coefficient of 0.989 compared to the vegetative season in the test dataset. The U-Net showed a slightly higher RMSE in both seasons, with 16.47% and 18.41% for the flowering and vegetative seasons, respectively. The evaluation across both seasons revealed that models trained on data from one season performed poorly when applied to the other season. Interestingly, models developed using flowering season data showed better adaptability to the vegetative season. Conversely, vegetative season models exhibited higher RMSE—reaching approximately 75% for all tested AI models—when tested on flowering season data. While models trained on combined datasets performed well in both seasons, their accuracy was lower in comparison to models specifically trained and tested on data from the same season.

The common models developed from both seasons maintained higher performance in both seasons. The developed custom CNN model achieved an RMSE of 8.23% and 14.02% for the flowering and vegetative seasons, respectively, while the XGBoost model achieved an RMSE of 19.75% and 15.25% for the flowering and vegetative seasons, respectively. The custom CNN model outperformed XGBoost in both seasons, showing a correlation of 0.966 and 0.821 with the ground truth assessment. The U-Net achieved good performance in both flowering and vegetative seasons, although its RMSE was slightly higher compared to XGBoost and the custom CNN. The SVM and RF also showed competitive results in the vegetative season. However, both SVM and RF models struggled in the flowering season, exhibiting significantly higher errors. Figure 9 shows the prediction maps of quadrats in test sites using the models from combined dataset. The ground truth information and the predicted ALG proportions for 10 quadrats during the flowering season (Site 2), as predicted by the models from the combined dataset, are shown in Figure 10. The ground truths and the predicted ALG proportions for 19 quadrats during the vegetative season

(Site 3), as predicted by the models from the combined dataset, are shown in Figure 11. Figure 12 shows the predicted spatial distribution of ALG over the test sites.

Table 6. Multispectral imagery-based ALG model performance on test sites during each season. RMSE and R^2 metrics correspond to quadrat-based testing. Precision (%), recall (%), and F1 score (%) are the weighted average evaluation metrics correspond to mask-based testing.

Season Used to Develop Models	Metrics	Season Used to Test Models									
		Flowering					Vegetative				
		SVM	RF	XGB	Custom CNN	U-Net	SVM	RF	XGB	Custom CNN	U-Net
Flowering	RMSE	37.32%	40.86%	17.56%	5.77%	16.47%	51.15%	63.91%	33.64%	22.33%	31.44%
	R^2	0.1990	0.0306	0.977	0.989	0.9660	0.4781	0.2335	0.596	0.574	0.6705
	Precision (%)	66.83	75.74	85.61	87.79	78.81	72.20	67.03	51.95	51.20	54.72
	Recall (%)	62.21	65.22	83.08	87.40	74.76	70.80	67.06	62.11	58.95	44.63
	F1 Score (%)	59.36	60.20	82.24	86.40	73.21	67.75	58.10	55.14	52.73	35.07
Vegetative	RMSE	77.02%	73.15%	71.03%	72.58%	74.72%	15.29%	19.11%	14.2%	12.9%	18.41%
	R^2	0.2672	0.3247	0.322	0.298	0.2648	0.8880	0.8403	0.831	0.851	0.8504
	Precision (%)	32.17	53.94	53.40	43.39	70.79	76.15	73.27	78.19	73.76	70.23
	Recall (%)	41.35	45.86	41.17	43.76	66.61	67.94	66.05	71.02	70.31	72.09
	F1 Score (%)	32.06	37.17	36.71	35.35	59.33	60.10	57.95	62.14	61.64	70.62
Flowering and Vegetative	RMSE	34.25%	32.08%	19.72%	8.23%	14.32%	15.42%	18.68%	15.25%	14.02%	18.47%
	R^2	0.3005	0.7527	0.978	0.966	0.9643	0.8511	0.8247	0.81	0.821	0.8555
	Precision (%)	69.41	77.43	83.33	87.63	76.79	76.49	76.92	75.36	71.12	69.6
	Recall (%)	70.80	72.58	81.62	87.47	72.3	66.05	65.60	70.50	70.25	70.04
	F1 Score (%)	66.07	66.34	80.02	87.06	50.33	56.87	56.77	61.19	62.62	69.5

SVM: support vector machine, RF: random forest, XGB: eXtreme gradient boosting, CNN: convolutional neural network.

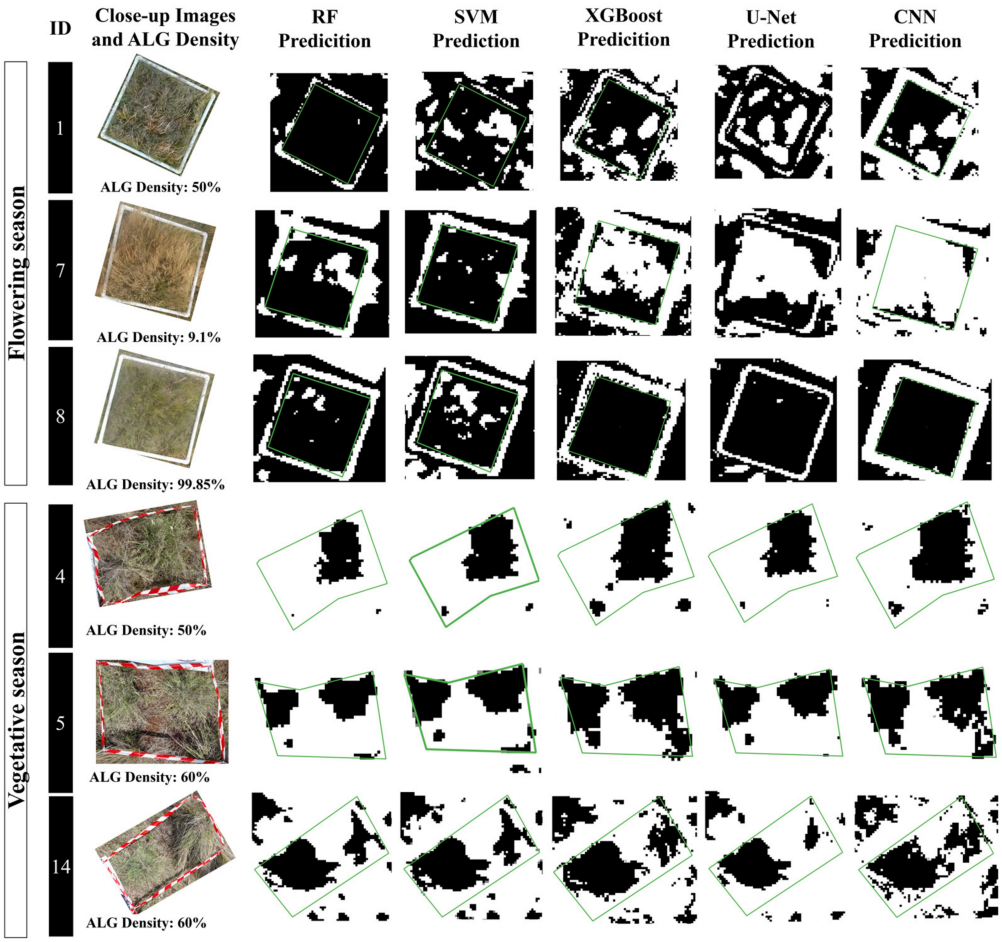


Figure 9. Multispectral-based prediction maps of three quadrats from test sites using the models developed from the combined seasonal dataset. The filled black regions represent the ALG.

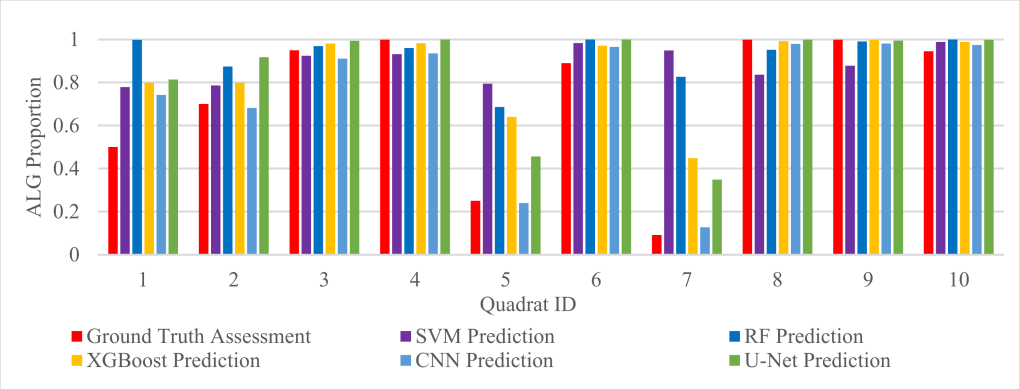


Figure 10. The ground truth and the predicted ALG proportion from the Bunyan test site (flowering) using the models developed from the combined seasonal dataset.

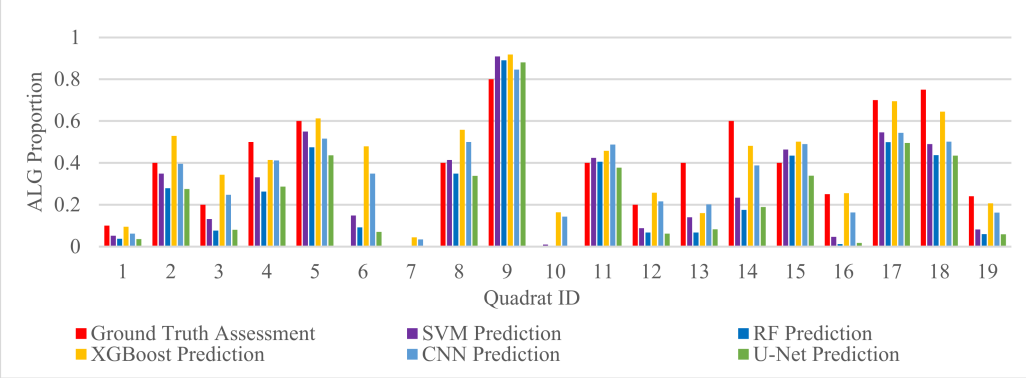


Figure 11. The ground truth and the predicted ALG proportion from the Cooma test site (vegetative) using the models developed from the combined seasonal dataset.

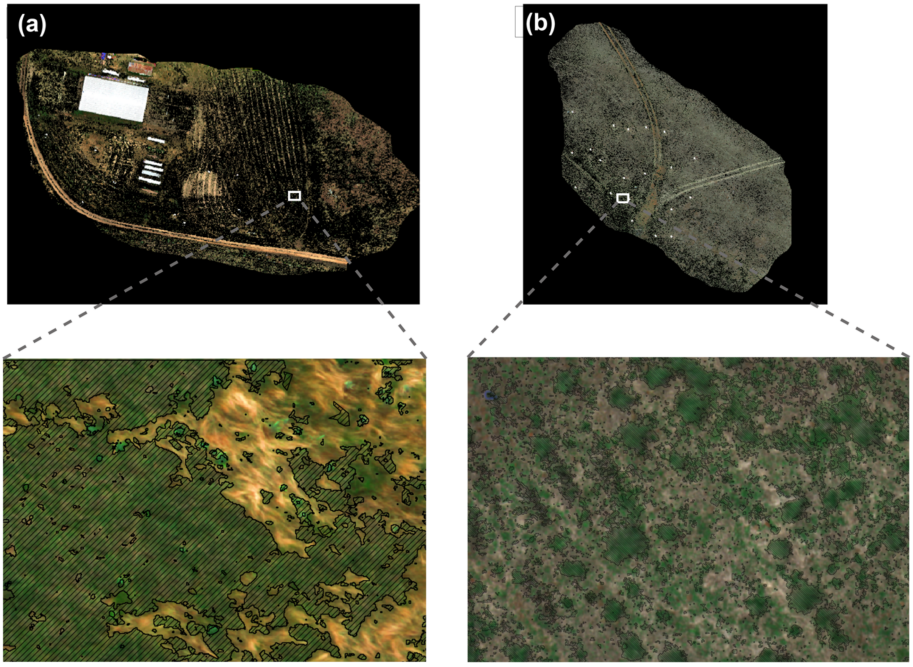


Figure 12. Multispectral-based segmented ALG spatial distribution map of test sites. (**a**) Cooma site; (**b**) Bunyan site. The hashed black polygon represents the ALG detected region.

Table 7 presents the per-class evaluation metrics for ALG class using mask-based testing. Overall, the developed custom CNN model consistently demonstrated superior performance across both seasons. The custom CNN model exhibited superior performance during the flowering season, with a precision of 85.81%, recall of 98.41%, and an F1 score of 91.68%. Following the custom CNN, the XGBoost model achieved precision, recall, and F1 scores of 74.85%, 99.35%, and 85.37%, respectively. The U-Net model also performed well, particularly in recall (99.81%) and F1 score (84.38%). Both RF and SVM exhibited lower performance. During the vegetative season, the U-Net model led in precision with 95.56%, while XGBoost achieved the highest recall (95.89%) and F1 score (90.68%). The custom CNN model continued to perform strongly, while the RF and SVM also showed competitive performance in this season. Overall, the developed custom CNN model consistently demonstrated superior performance across both seasons.

Table 7. ALG class-based model performance on test sites during each season. Precision (%), recall (%), and F1 score (%) are the per-class evaluation metrics specifically for the ALG class. The highest values for each metric for each dataset are highlighted in bold.

ALG Class Metrics	Season Used to Test Models									
	Flowering					Vegetation				
	SVM	RF	XGB	Custom CNN	U-Net	SVM	RF	XGB	Custom CNN	U-Net
Precision (%)	72.72	72.79	74.85	**85.81**	73.08	90.31	92.31	86	89.10	**95.56**
Recall (%)	92.04	98.66	**99.35**	98.41	99.81	85.55	83.79	**95.89**	92.17	83.15
F1 Score (%)	81.25	83.77	85.37	**91.68**	84.38	87.86	87.84	**90.68**	90.61	88.93

SVM: support vector machine, RF: random forest, XGB: eXtreme gradient boosting, CNN: convolutional neural network.

Table 8 presents the development (training and validation) and testing times for the AI models. The SVM and U-Net exhibit the longest development times, at 3700.4 and 6749.7 s, respectively, significantly exceeding those of the other models. Even with the use of a high-performance computer (HPC) equipped with additional resources, U-Net still recorded the longest training time. In contrast, XGBoost is notable for its rapid training and testing times while maintaining competitive performance.

Table 8. Development, testing time, and computer specifications for the combined seasonal (vegetative and flowering) datasets.

Models	RF	SVM	XGB	Custom CNN	U-Net
Development time (s)	941.8	6749.7	54.4	737.5	3700.4
Testing time (s)	0.65	73	0.72	2.14	46.48
Computer specification	Processor: 12th Gen Intel(R) Core (TM) i7-1255U 1.70 GHz RAM: 16.0 GB (15.6 GB usable)				Processor: AMD EPYC 7713 64-Core RAM: 100 GB GPU: A100-SXM4-40GB

SVM: support vector machine, RF: random forest, XGB: eXtreme gradient boosting, CNN: convolutional neural network.

3.2. Comparison of Predictions Using Multispectral and Hyperspectral Imagery

Table 9 presents the prediction results derived from the models for MS and HS data. During model development, the validation performance shows that the custom CNN achieved the highest weighted average precision, recall, and F1 score for both imagery types compared to other models. All models performed well, with HS data yielding slightly higher accuracy than MS data. A similar performance trend is observed in the test results, where quadrat-based testing with the custom CNN showed that HS-based predictions achieved an RMSE of 19.58 while MS-based predictions had an RMSE of 20.51. XGBoost also showed competitive results, with an RMSE of 20.54 for MS and 24.46 for HS, while SVM

performed better than XGBoost and closer to the custom CNN performance, with an RMSE of 19.66. Mask-based testing based on the ALG class-based metrics achieved the highest results for the custom CNN, consistent with the trend observed in quadrat-based testing for the custom CNN. The ground truth assessment and the predicted ALG proportion by all tested models using MS and HS imagery for five quadrats from the test site are shown in Figure 13. Figure 14 shows the prediction map of quadrats in test sites using the AI models.

Table 9. Classification results from the models for multispectral and hyperspectral ALG imagery using the validation dataset from Site 2. The lowest RMSE values for each imagery are highlighted in bold.

Dataset	Metrics	Imagery							
		Multispectral				Hyperspectral			
		SVM	RF	XGB	Custom CNN	SVM	RF	XGB	Custom CNN
Validation data	Precision (%)	91.69	88.60	98.36	98.93	99.88	92.38	99.83	99.9
	Recall (%)	91.60	88.36	98.35	98.92	99.87	92.39	99.82	99.9
	F1 Score (%)	91.49	88.18	98.35	98.92	99.88	92.38	99.82	99.9
	RMSE	27.82%	28.4%	20.54%	**20.51%**	19.66%	33.34%	24.46%	**19.58%**
	R^2	0.6308	0.8234	0.9661	0.878	0.8532	0.4950	0.8079	0.962
Test data	Precision (%)	94.64	90.92	94.72	96.9	98.21	90.09	98.47	99.06
	Recall (%)	93.87	89.3	93.76	96.89	98.19	89.89	98.46	99.05
	F1 Score (%)	93.99	89.56	93.93	96.88	98.19	89.93	98.46	99.05

SVM: support vector machine, RF: random forest, XGB: eXtreme gradient boosting, CNN: convolutional neural network.

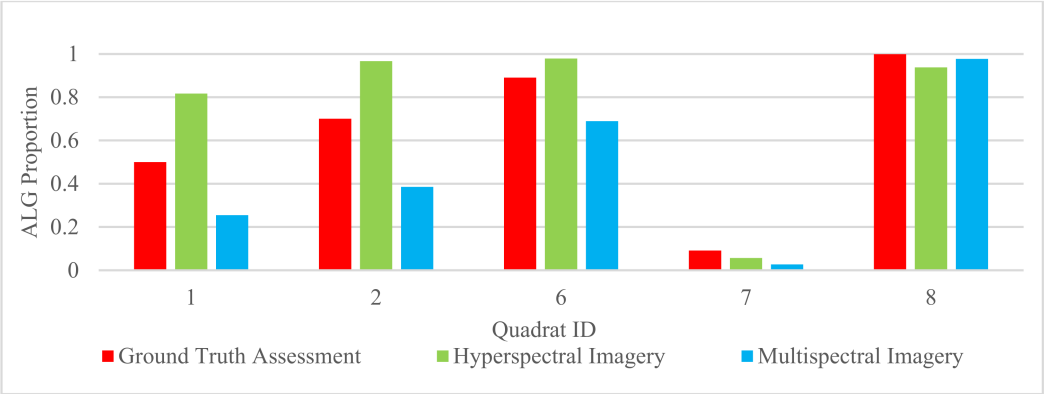


Figure 13. The ground truth and the predicted ALG proportion of the quadrats from the test region of Site 2 by the custom CNN model.

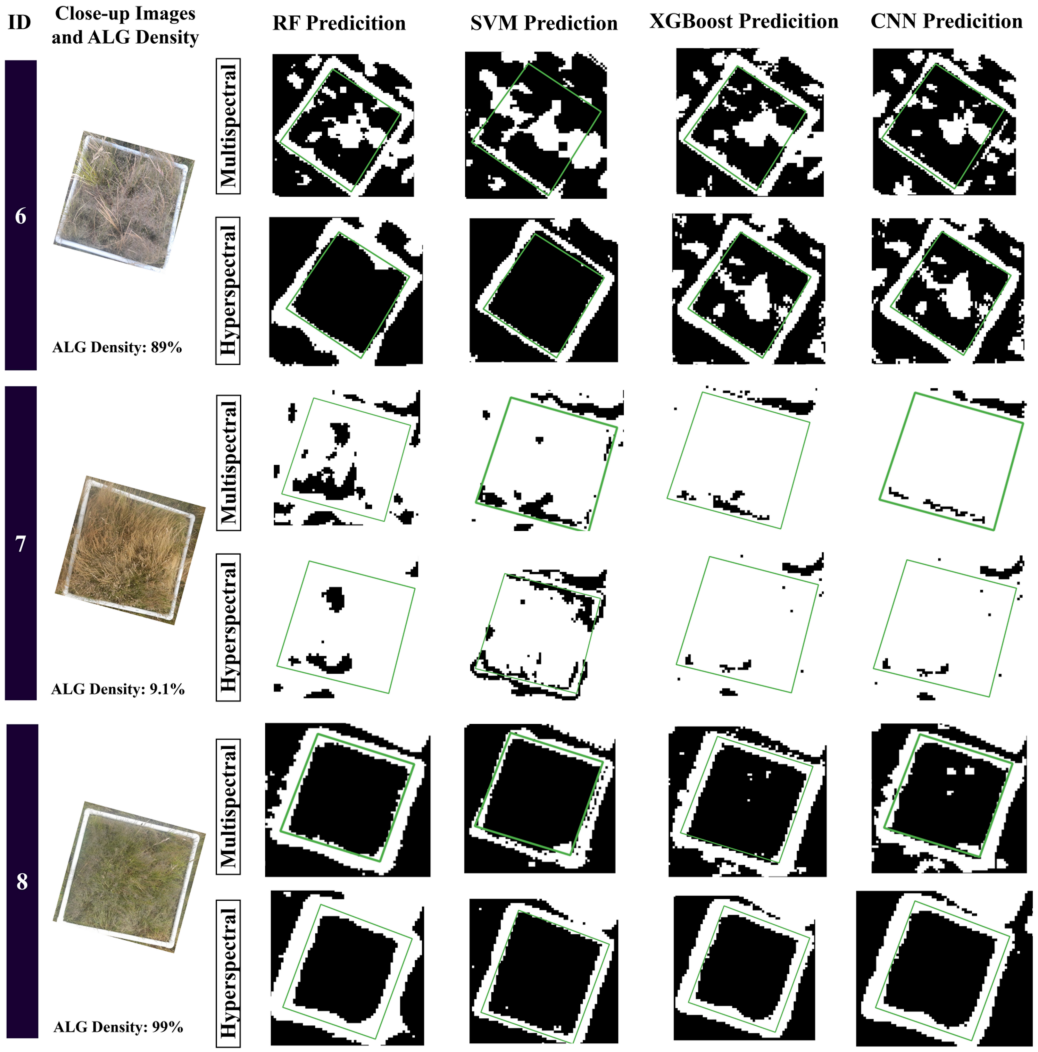


Figure 14. Comparison of multispectral and hyperspectral imagery-based models prediction maps of three quadrats from test sites. The filled black regions represent the ALG.

4. Discussion

The ability to detect ALG from remotely sensed imagery using AI approaches not only helps to improve fire management and biodiversity in ALG-infested environments but also contributes to the broader field of using AI for invasive plant surveillance and management. This research employed an AI-based ALG segmentation approach, one not previously explored in studies. It builds upon the established foundation of utilizing UAS RS for invasive plant segmentation [6,8,10,16,18]. The AI model testing was conducted using two approaches. The first approach involved quadrat-based ALG proportion prediction to assess the performance of the models in a mixed environment. Secondly, masked regions where ALG and other classes can be confidently differentiated were selected for testing, and the metrics precision, recall, and F1 score were used to evaluate all classifications. Our investigation demonstrated the superior performance of a custom CNN model in detecting ALG compared to classical ML models and a U-Net DL model during classification testing.

Combining data from both seasons did lead to a reduction in error compared to single-season models when tested on data from the alternate season for both testing approaches. This makes the combined model more versatile. The inclusion of data acquired from different seasons and with various data collection settings strengthened the combined model's robustness. According to the ALG class-based metrics from the masked-based testing approach, the five models (RF, SVM, XGBoost, custom CNN, and U-Net) showed varying performance across seasons. The custom CNN consistently achieved high precision

in both seasons, indicating its ability to accurately identify true positives and minimize false positives. While U-Net and XGBoost performed well, their lower average precision in comparison to the custom CNN suggests the potential for more false positives. Although RF and SVM exhibited good precision during the vegetative season, their lower performance during flowering possibly led to more false positives in those conditions. The high recall of U-Net and XGBoost across seasons suggests they effectively identified a high proportion of true positives and minimized false negatives. The strong recall of the custom CNN demonstrates its ability to capture most true positives. Contrastingly, SVM's lower recall, particularly in the flowering season, indicates it might miss a higher number of actual calls, leading to more false negatives.

The custom CNN emerged as the most robust model based on the F1 score, which considers both precision and recall. This suggests the custom CNN effectively balances identifying true positives while minimizing false positives and negatives across all seasons. Overall, the custom CNN's consistent performance across all metrics suggests its potential for accurate and robust species call prediction in various seasonal conditions. The developed custom CNN excelled at capturing hierarchical spectral features and spatial relationships within the data. This proficiency in feature extraction allows the custom CNNs to generalize better to unseen data compared to XGBoost in image classification. Although the U-net model also showed notable precision, especially in the vegetative season, the developed custom CNN architecture was superior to U-Net. In the context of classifying pixels in a mixed environment, the task of detecting grass species relies more on spectral information than on spatial patterns. Since U-Net is designed to leverage detailed spatial information for segmentation tasks, its complexity becomes redundant in this scenario. The simpler custom CNN architecture, which focuses on extracting essential spectral features, proves to be more effective, thereby explaining U-Net's comparatively lower performance in this specific application.

Quadrat-based testing revealed that among the classical ML methods, XGBoost achieved the highest accuracy in the MS-based prediction compared to ground truth ALG proportion. Its gradient boosting framework, which builds successive trees to correct errors of the previous trees, results in a stronger overall model. XGBoost offers a straightforward method for measuring feature importance, which is crucial for feature selection and understanding the model. XGBoost offers regularization and is capable of capturing complex non-linear relationships between features and the target variable [36], often achieving higher accuracy compared to SVM and RF in vegetation classification studies [7,37]. However, for image data, CNN generally outperforms XGBoost [38,39]. Studies on segmenting broad-leaved pepper [40] and bitou bush [7] showed similar performance, with XGBoost classifying target vegetation with higher overall accuracy than other classical methods such as SVM and RF. When comparing classical methods to deep learning networks, both studies demonstrated that the U-Net achieved significantly higher accuracy, with an increase of 10–20% over XGBoost. Other researchers have also incorporated deep learning techniques to enhance the overall accuracy of classification in complex natural environments [41]. However, the complexity and inference speed of models like U-Net are critical considerations for practical applications [40,42]. If the task does not require detailed spatial information and relies more on spectral data, the additional complexity of U-Net may become unnecessary and counterproductive. The developed custom CNN model extracts sufficient features without the overhead, making them more efficient and more accurate for these tasks.

Both the XGBoost and custom CNN models accurately predicted ALG proportions in most quadrats when the ALG proportion was very high or very low in both seasonal conditions. However, in vegetative season, the sixth quadrat from Site 3 resulted in inaccurate predictions for both seasonal models. This inaccuracy was due to the presence of *Bothriochloa macra* (Redgrass) in 70% of this quadrat, which shares visual similarities with ALG. This species was not well represented in the labeled data from the sites of vegetative season because it is limited in growth at Site 4, resulting in less labeled data

being incorporated for developing the models. Encountering difficulties in predicting invasive species segmentation when similar species are mixed together is a common challenge [43]. Incorporating UAS-labeled images with diverse vegetation types can enhance the model's robustness in detecting ALG across a broader range. However, the site characteristics of ALG, mixed with other grass species of similar height, made continuous labeling challenging for experts. The nature of the sites restricted labeling to clear ALG and non-ALG areas, which particularly limited the training data.

Comparatively higher errors were observed in the predictions when ALG was mixed with other species in similar proportions. This could be due to the inaccurate ground truth recordings, which may have occurred due to the human eye not capturing accurate proportions of species, a task especially challenging in mixed environments. This can be mitigated by carefully inspecting close-range images of quadrats, images which remain an important aspect of weed UAS surveys. It is also possible that lower resolution in the MS images may have led to assessments not clearly capturing the details of mixed vegetation. Distinguishing between similar-looking species based on color or texture poses challenges and limits the labeling process. Additionally, when it comes to predictions, lower resolution causes confusion in distinguishing each species. This may lead the models to exhibit lower accuracy in quadrats belonging to mixed environments. A low-altitude flight can improve resolution, but it may reduce coverage area and require multiple flights to survey the entire site [5], increasing time and cost.

It is important to note that this combined model's performance still was not as high as models specifically trained and tested on data from a single season. All AI models struggled when applied to a different season than they were trained on. The performance is highly influenced by plant phenology changes between seasons, the data source employed for image acquisition, and flight altitude. Interestingly, models developed using flowering season data showed better adaptability to the vegetative season. This may be the result of flowering landscapes retaining the characteristics of both flowering and vegetative season. Meanwhile, models trained on the vegetative season dataset exhibited higher error when tested on flowering season dataset. The presence of flowers in the testing data likely creates confusion for the models when those features are absent in the vegetative season. For the most accurate results, further exploration might involve creating an ensemble model [44] that combines the strengths of separate seasonal models. This could potentially be utilized to achieve better overall performance when dealing with variations in altitude, sensor type, and season.

The comparison between HS and MS imagery-based model performances revealed that HS achieved better accuracy in both the training and validation datasets. However, acquiring HS imagery is associated with high costs and heavy-weight sensors, which, in turn, reduce the endurance of UAS and limit the coverage area. It also increases data size and processing time and results in lower resolution compared to MS imagery [5,7]. Despite the limited data used to develop and test the models, MS imagery proves to be more advantageous, with satisfactory results compared with HS imagery due to its operational benefits.

This study successfully demonstrated the feasibility of incorporating AI models for achieving sufficient accuracy. However, there are limitations in the study, as yet-unmapped sites may display a variety of species that are not represented in the dataset and the models developed do not include the details of all species across all different seasonal conditions. This could constrain the use of these models to particular conditions and locations. Additionally, the comparison between HS and MS imagery was only conducted for Site 2, representing only flowering seasonal conditions. Future efforts will need to focus on comparing model performance across various grasslands, incorporating data from additional environments, and including more seasonal variation to enhance the models' robustness in unseen grasslands. Furthermore, a wider range of lightweight CNN architectures can be explored, potentially through techniques such as transfer learning or ensemble methods [45]. While both the XGBoost and the custom CNN offer good

performance for ALG detection, a hybrid approach like ConvXGB [46] could be even more beneficial. ConvXGB combines several CNN layers for feature extraction with XGBoost as the final layer for prediction, leveraging the strengths of both models for improved accuracy [47].

Promising findings for the classical XGBoost and the customized CNN models in this ALG segmentation study empower land managers and researchers to further explore this approach. With precise mapping, land managers can implement targeted interventions such as selective herbicide application, controlled burns, and mechanical removal, reducing the impact on non-target species and minimizing resource wastage. Regular monitoring using the developed models can ensure that ALG does not overtake pastures. Conservationists can focus on protecting and restoring native plant communities by precisely mapping ALG invasions, ensuring that interventions do not inadvertently harm native biodiversity. This approach ensures the protection of both economic interests and environmental health. The scalability of ALG segmentation in mixed environments presents an opportunity to integrate satellite-based RS, despite challenges posed by low temporal and spatial satellite resolution. These challenges can be addressed through advanced techniques such as satellite/UAS data fusion [48] and multi-scale modeling [49]. Utilizing predictions from UAS as the ground truth for satellite-based models can enable scaling predictions to larger geographical areas. By bridging the resolution gap between high-resolution UAS data and broader satellite coverage, this study holds immense potential to advance ALG prediction across broader geographical scales. The use of UASs for RS continues to be limited by factors like performance, cost, and availability. As a result, their contribution to practical grassland management decisions is not always satisfactory [43,50]. However, ongoing advancements in this field offer potential for enhanced segmentation of ALG and could significantly contribute to control, thereby reducing fire risk and improving agricultural productivity in areas infested by this species.

5. Conclusions

Infestations of ALG poses significant threats to Australian ecosystems including agricultural losses, reduced biodiversity, and increased fire risk. UAS equipped with RS technology, coupled with advanced AI algorithms, offer a promising approach for accurately mapping the spatial distribution of ALG. This study evaluated the performance of state-of-the-art AI algorithms, including XGBoost, RF, SVM, U-Net, and a customized CNN, in identifying and mapping ALG in the mixed Australian landscape. Study results revealed the superior performance of the custom CNN model in ALG segmentation, with high accuracy and lower error rates compared to classical ML methods. Training models on single-season data resulted in limited generalizability, evident from performance decline when applied to data with seasonal and collection variations. However, incorporating multi-season data during training significantly improved model performance across seasons. This suggests the effectiveness of multi-season data integration for enhancing generalizability and fostering a more versatile and robust model. Additionally, the study compared the effectiveness of MS and HS imagery for ALG segmentation, with MS imagery being more practical due to lower operational limitations. Lightweight models can thereby achieve sufficient accuracy rendering them suitable for ALG segmentation, despite certain limitations in terms of species representation and seasonal variability. Study results necessitate future research to focus on developing more robust AI models by comparing performance across diverse grassland species and incorporating data from various environments and seasons. Overall, these advancements are crucial for enhancing the effectiveness of AI-based solutions in grassland management and biodiversity conservation efforts.

Author Contributions: Conceptualization, P.K., N.A., J.E.K. and F.G.; data curation, R.L.D.; formal analysis, P.K.; funding acquisition, J.E.K.; investigation, P.K. and F.G.; methodology, P.K. and N.A.; project administration, J.E.K.; resources, J.E.K.; software, P.K.; supervision, L.Z., G.H. and F.G.; validation, P.K.; writing—original draft, P.K.; writing—review and editing, P.K., N.A., J.E.K., N.M., L.Z., G.H. and F.G. All authors have read and agreed to the published version of the manuscript.

Funding: This project is supported through funding from the Department of Agriculture, Fisheries and Forestry grant round (Grant number: 4-FY9LIKQ), Advancing Pest Animal and Weed Control Solutions, as part of the Established Pest Animal and Weeds Pipeline program and the ARC Discovery program (Project number: DP 220103233).

Data Availability Statement: Data relevant to this study can be provided upon request.

Acknowledgments: The authors gratefully acknowledge the support of Susannah Harper of the Snowy Monaro Regional Council, and Jane Tracy and David Eddy of Southeast Local Land Services NSW for their assistance with plant identification for labeling. Computational (and/or data visualization) resources and services used in this work were provided by the eResearch Office, Queensland University of Technology, Brisbane, Australia.

Conflicts of Interest: The authors declare that they have no conflicts of interest related to this research.

References

1. Firn, J. African Lovegrass in Australia: A Valuable Pasture Species or Embarrassing Invader? *Trop. Grassl.* **2009**, *43*, 86–97.
2. Roberts, J.; Florentine, S.; van Etten, E.; Turville, C. Germination Biology, Distribution and Control of the Invasive Species Eragrostis Curvula [Schard. Nees] (African Lovegrass): A Global Synthesis of Current And Future Management Challenges. *Weed Res.* **2021**, *61*, 154–163. [CrossRef]
3. Johnston, W.H.; Aveyard, J.M.; Legge, K. Selection and testing of Consol Lovegrass for Soil Conservation and Pastoral Use. *J. Soil. Conserv.* **1984**, *40*, 38–45.
4. Walker, Z.C.; Morgan, J.W. Perennial Pasture Grass Invasion Changes Fire Behaviour and Recruitment Potential of A Native Forb in a Temperate Australian Grassland. *Biol. Invasions* **2022**, *24*, 1755–1765. [CrossRef]
5. Keerthinathan, P.; Amarasingam, N.; Hamilton, G.; Gonzalez, F. Exploring Unmanned Aerial Systems Operations in Wildfire Management: Data Types, Processing Algorithms and Navigation. *Int. J. Remote Sens.* **2023**, *44*, 5628–5685. [CrossRef]
6. 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]
7. Amarasingam, N.; E Kelly, J.; Sandino, J.; Hamilton, M.; Gonzalez, F.; L Dehaan, R.; Zheng, L.; Cherry, H. Bitou Bush Detection and Mapping Using UAV-Based Multispectral and Hyperspectral Imagery and Artificial Intelligence. *Remote Sens. Appl. Soc. Environ.* **2024**, *34*, 101151. [CrossRef]
8. Harris, S.; Trotter, P.; Gonzalez, F.; Sandino, J. Bitou bush surveillance UAV trial. In Proceedings of 14th Queensland Weed Symposium, Brisbane, Australia, 4–7 December 2017.
9. Xia, F.; Quan, L.; Lou, Z.; Sun, D.; Li, H.; Lv, X. Identification and Comprehensive Evaluation of Resistant Weeds Using Unmanned Aerial Vehicle-Based Multispectral Imagery. *Front. Plant Sci.* **2022**, *13*, 938604. [CrossRef] [PubMed]
10. Hamylton, S.M.; Morris, R.H.; Carvalho, R.C.; Roder, N.; Barlow, P.; Mills, K.; Wang, L. Evaluating Techniques for Mapping Island Vegetation from Unmanned Aerial Vehicle (UAV) Images: Pixel Classification, Visual Interpretation and Machine Learning Approaches. *Int. J. Appl. Earth Obs.* **2020**, *89*, 102085. [CrossRef]
11. Huang, H.; Lan, Y.; Deng, J.; Yang, A.; Deng, X.; Zhang, L.; Wen, S. A Semantic Labeling Approach for Accurate Weed Mapping of High Resolution UAV Imagery. *Sensors* **2018**, *18*, 2113. [CrossRef]
12. Alexandridis, T.K.; Tamouridou, A.A.; Pantazi, X.E.; Lagopodi, A.L.; Kashefi, J.; Ovakoglou, G.; Polychronos, V.; Moshou, D. Novelty Detection Classifiers in Weed Mapping: Silybum Marianum Detection on UAV Multispectral Images. *Sensors* **2017**, *17*, 2007. [CrossRef] [PubMed]
13. Khoshboresh-Masouleh, M.; Akhoondzadeh, M. Improving Weed Segmentation in Sugar Beet Fields Using Potentials of Multispectral Unmanned Aerial Vehicle Images and Lightweight Deep Learning. *J. Appl. Remote Sens.* **2021**, *15*, 034510. [CrossRef]
14. Osorio, K.; Puerto, A.; Pedraza, C.; Jamaica, D.; Rodríguez, L. A Deep Learning Approach for Weed Detection in Lettuce Crops Using Multispectral Images. *AgriEngineering* **2020**, *2*, 471–488. [CrossRef]
15. 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 Sen.* **2018**, *10*, 1423. [CrossRef]
16. Su, J.; Yi, D.; Coombes, M.; Liu, C.; Zhai, X.; McDonald-Maier, K.; Chen, W.-H. Spectral Analysis and Mapping of Blackgrass Weed by Leveraging Machine Learning and UAV Multispectral Imagery. *Comput. Electron. Agric.* **2022**, *192*, 106621. [CrossRef]
17. Martín, M.P.; Ponce, B.; Echavarría, P.; Dorado, J.; Fernández-Quintanilla, C. Early-Season Mapping of Johnsongrass (*Sorghum halepense*), Common Cocklebur (*Xanthium strumarium*) and Velvetleaf (*Abutilon theophrasti*) in Corn Fields Using Airborne Hyperspectral Imagery. *Agronomy* **2023**, *13*, 528. [CrossRef]
18. Papp, L.; van Leeuwen, B.; Szilassi, P.; Tobak, Z.; Szatmári, J.; Árvai, M.; Mészáros, J.; Pásztor, L. Monitoring Invasive Plant Species Using Hyperspectral Remote Sensing Data. *Land* **2021**, *10*, 29. [CrossRef]
19. Cao, J.; Fu, J.; Yuan, X.; Gong, J. Nonlinear Bias Compensation of ZiYuan-3 Satellite Imagery with Cubic Splines. *Isprs J. Photogramm.* **2017**, *133*, 174–185. [CrossRef]

20. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [CrossRef]
21. Gao, B.-C. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [CrossRef]
22. Barnes, E.; Clarke, T.; Richards, S.; Colaizzi, P.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T. Coincident Detection of Crop Water Stress, Nitrogen Status and Canopy Density Using Ground Based Multispectral Data. In Proceedings of Fifth International Conference on Precision Agriculture, Bloomington, MN, USA, 16–19 July 2000.
23. Kurbanov, R.K.; Zakharova, N.I. Application of Vegetation Indexes to Assess the Condition of Crops. *Agric. Mach. Technol.* **2020**, *14*, 4–11. [CrossRef]
24. Eng, L.; Ismail, R.; Hashim, W.; Baharum, A. The Use of VARI, GLI, and VIgreen Formulas in Detecting Vegetation in aerial Images. *Int. J. Technol.* **2019**, *10*, 1385. [CrossRef]
25. Goutte, C.; Gaussier, E. A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation. In Proceedings of European Conference on Information Retrieval, Santiago de Compostela, Spain, 21–23 March 2005; pp. 345–359.
26. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
27. Valero-Jorge, A.; González-De Zayas, R.; Matos-Pupo, F.; Becerra-González, A.L.; Álvarez-Taboada, F. Mapping and Monitoring of the Invasive Species Dichrostachys cinerea (Marabú) in Central Cuba Using Landsat Imagery and Machine Learning (1994–2022). *Remote Sens.* **2024**, *16*, 798. [CrossRef]
28. Chen, T.; Guestrin, C. Xgboost: A Scalable Tree Boosting System. In Proceedings of 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
29. Rudin, C.; Chen, C.; Chen, Z.; Huang, H.; Semenova, L.; Zhong, C. Interpretable Machine Learning: Fundamental Principles and 10 Grand Challenges. *Stat. Surv.* **2022**, *16*, 1–85.
30. Raniga, D.; Amarasingam, N.; Sandino, J.; Doshi, A.; Barthelemy, J.; Randall, K.; Robinson, S.A.; Gonzalez, F.; Bollard, B. Monitoring of Antarctica's Fragile Vegetation Using Drone-Based Remote Sensing, Multispectral Imagery and AI. *Sensors* **2024**, *24*, 1063. [CrossRef] [PubMed]
31. Krichen, M. Convolutional Neural Networks: A Survey. *Computers* **2023**, *12*, 151. [CrossRef]
32. Gavrikov, P.; Keuper, J. The Power of Linear Combinations: Learning with Random Convolutions. *arXiv* **2023**, arXiv:2301.11360.
33. Ma, H.; Huang, W.; Dong, Y.; Liu, L.; Guo, A. Using UAV-Based Hyperspectral Imagery to Detect Winter Wheat Fusarium Head Blight. *Remote Sens.* **2021**, *13*, 3024. [CrossRef]
34. Chai, T.; Draxler, R.R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE). *Geosci. Model. Dev. Discuss.* **2014**, *7*, 1525–1534.
35. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared is More Informative Than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *Peerj Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]
36. Wang, L.; Zhao, C.; Liu, X.; Chen, X.; Li, C.; Wang, T.; Wu, J.; Zhang, Y. Non-Linear Effects of the Built Environment and Social Environment on Bus Use among Older Adults in China: An Application of the XGBoost Model. *Int. J. Environ. Res. Public. Health* **2021**, *18*, 9592. [CrossRef]
37. Ramdani, F.; Furqon, M.T. The Simplicity of XGBoost Algorithm Versus the Complexity of Random Forest, Support Vector Machine, and Neural Networks Algorithms in Urban Forest Classification. *F1000Research* **2022**, *11*, 1069. [CrossRef]
38. Yu, F.; Zhang, Q.; Xiao, J.; Ma, Y.; Wang, M.; Luan, R.; Liu, X.; Ping, Y.; Nie, Y.; Tao, Z.; et al. Progress in the Application of CNN-Based Image Classification and Recognition in Whole Crop Growth Cycles. *Remote Sens.* **2023**, *15*, 2988. [CrossRef]
39. Mukhamediev, R.I.; Symagulov, A.; Kuchin, Y.; Yakunin, K.; Yelis, M. From Classical Machine Learning to Deep Neural Networks: A Simplified Scientometric Review. *Appl. Sci.* **2021**, *11*, 5541. [CrossRef]
40. Amarasingam, N.; Vanegas, F.; Hele, M.; Warfield, A.; Gonzalez, F. Integrating Artificial Intelligence and UAV-Acquired Multispectral Imagery for the Mapping of Invasive Plant Species in Complex Natural Environments. *Remote Sens.* **2024**, *16*, 1582. [CrossRef]
41. Lobo Torres, D.; Queiroz Feitosa, R.; Nigri Happ, P.; Elena Cué La Rosa, L.; Marcato Junior, J.; Martins, J.; Olã Bressan, P.; Gonçalves, W.N.; Liesenberg, V. Applying Fully Convolutional Architectures for Semantic Segmentation of a Single Tree Species in Urban Environment on High Resolution UAV Optical Imagery. *Sensors* **2020**, *20*, 563. [CrossRef]
42. Kislov, D.E.; Korznikov, K.A. Automatic Windthrow Detection Using Very-High-Resolution Satellite Imagery and Deep Learning. *Remote Sens.* **2020**, *12*, 1145. [CrossRef]
43. Amarasingam, N.; Hamilton, M.; Kelly, J.E.; Zheng, L.; Sandino, J.; Gonzalez, F.; Dehaan, R.L.; Cherry, H. Autonomous Detection of Mouse-Ear Hawkweed Using Drones, Multispectral Imagery and Supervised Machine Learning. *Remote Sens.* **2023**, *15*, 1633. [CrossRef]
44. Ma, C.; Wang, W.; Wang, H.; Cao, Z. Ensemble of Deep Convolutional Neural Networks for Real-Time Gravitational Wave Signal Recognition. *Phys. Rev. D* **2022**, *105*, 083013. [CrossRef]
45. Gupta, J.; Pathak, S.; Kumar, G. Deep Learning (CNN) and Transfer Learning: A Review. *J. Phys. Conf. Ser.* **2022**, *2273*, 012029. [CrossRef]
46. Thongsuwan, S.; Jaiyen, S.; Padcharoen, A.; Agarwal, P. ConvXGB: A New Deep Learning Model for Classification Problems Based on CNN and XGBoost. *Nucl. Eng. Technol.* **2021**, *53*, 522–531. [CrossRef]

47. Jiao, W.; Hao, X.; Qin, C. The Image Classification Method with CNN-XGBoost Model Based on Adaptive Particle Swarm Optimization. *Information* **2021**, *12*, 156. [CrossRef]
48. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Daloye, A.M.; Erkbol, H.; Fritschi, F.B. Crop Monitoring Using Satellite/UAV Data Fusion and Machine Learning. *Remote Sens.* **2020**, *12*, 1357. [CrossRef]
49. Sagan, V.; Maimaitijiang, M.; Sidike, P.; Maimaitiyiming, M.; Erkbol, H.; Hartling, S.; Peterson, K.; Peterson, J.; Burken, J.G.; Fritschi, F. UAV/Satellite Multiscale Data Fusion for Crop Monitoring and Early Stress Detection. In Proceedings of the 4th ISPRS Geospatial Week 2019, Enschede, The Netherlands, 10–14 June 2019.
50. Parvathi, S.; Tamil Selvi, S. Detection of Maturity Stages of Coconuts in Complex Background Using Faster R-CNN Model. *Biosyst. Eng.* **2021**, *202*, 119–132. [CrossRef]

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.