Article

Detection of Parthenium Weed (*Parthenium hysterophorus* L.) and Its Growth Stages Using Artificial Intelligence

Benjamin Costello [1,2], Olusegun O. Osunkoya [3,*], Juan Sandino [1], William Marinic [1], Peter Trotter [4], Boyang Shi [3], Felipe Gonzalez [1] and Kunjithapatham Dhileepan [3]

1 Institute of Future Environments, Queensland University of Technology, 2 George Street, Brisbane, QLD 4000, Australia
2 Northrop Grumman Australia, Level 2/100 Brookes Street, Brisbane, QLD 4006, Australia
3 Invasive Plant & Animal Science Unit, Department of Agriculture and Fisheries, Biosecurity Queensland, 41 Boggo Road, Dutton Park, QLD 4102, Australia
4 Aspect UAV Imaging, 16 Spoonbill Street, Peregian Beach, QLD 4573, Australia
* Correspondence: olusegun.osunkoya@daf.qld.gov.au

Abstract: Parthenium weed (*Parthenium hysterophorus* L. (Asteraceae)), native to the Americas, is in the top 100 most invasive plant species in the world. In Australia, it is an annual weed (herb/shrub) of national significance, especially in the state of Queensland where it has infested both agricultural and conservation lands, including riparian corridors. Effective control strategies for this weed (pasture management, biological control, and herbicide usage) require populations to be detected and mapped. However, the mapping is made difficult due to varying nature of the infested landscapes (e.g., uneven terrain). This paper proposes a novel method to detect and map parthenium populations in simulated pastoral environments using Red-Green-Blue (RGB) and/or hyperspectral imagery aided by artificial intelligence. Two datasets were collected in a control environment using a series of parthenium and naturally co-occurring, non-parthenium (monocot) plants. RGB images were processed with a YOLOv4 Convolutional Neural Network (CNN) implementation, achieving an overall accuracy of 95% for detection, and 86% for classification of flowering and non-flowering stages of the weed. An XGBoost classifier was used for the pixel classification of the hyperspectral dataset—achieving a classification accuracy of 99% for each parthenium weed growth stage class; all materials received a discernible colour mask. When parthenium and non-parthenium plants were artificially combined in various permutations, the pixel classification accuracy was 99% for each parthenium and non-parthenium class, again with all materials receiving an accurate and discernible colour mask. Performance metrics indicate that our proposed processing pipeline can be used in the preliminary design of parthenium weed detection strategies, and can be extended for automated processing of collected RGB and hyperspectral airborne unmanned aerial vehicle (UAV) data. The findings also demonstrate the potential for images collected in a controlled, glasshouse environment to be used in the preliminary design of invasive weed detection strategies in the field.

Keywords: drones; image-processing; machine-learning; spectral-signature; weed-detection

1. Introduction

Biological invasion is a global environmental change phenomenon, often leading to disruptions of ecosystem goods and services. Biological invasion begins with the introduction of organisms from their native range into a new environment followed by periods of sequential adjustments of colonisation/extinction/recolonisation, establishment and naturalisation [1–5]. Where conditions are favourable the exotic species will increase dramatically in abundance and geographical distribution to become invaders with significant impact on nature conservation and agriculture. In Australia, like most regions around the world ~30,000 species have been introduced of which ~10% have become invaders [2,5].

One of such invaders, with its origin in the Americas, is the annual plant, (*Parthenium hysterophorus* L.) (henceforth referred to as parthenium weed).

In Australia, parthenium weed grows as an annual, herbaceous/shrubby invasive species plant of national significance [1,2]. Parthenium weed has a deep-penetrating taproot system and an erect shoot system [1]. Young plants form a rosette of leaves close to the soil surface. Mature stems are greenish and longitudinally grooved, covered in small stiff hairs (hirsute), and become much-branched at maturity. The alternately arranged leaves are simple with stalks (petioles) up to 2 cm long and form a basal rosette during the early stages of growth. The lower leaves are relatively large (3–30 cm long and 2–12 cm wide) and are deeply divided (bi-pinnatifid or bi-pinnatisect) [3]. Leaves on the upper branches decrease in size and are also less divided than the lower leaves. As parthenium weed matures, the plant develops many branches on its upper half, and may eventually grow up to two metres. With sufficient rainfall and warm temperature, parthenium weed has the ability to germinate and establish at any time of the year. Flowering usually commences 4–6 weeks after germination and soil moisture seems to be the major contributing factor to flowering [3]. Parthenium weed is a prolific seed producer and a fully grown plant can produce up to 157,000 seeds in its lifetime, though the majority of plants produce less than 4000 seeds [3]. Seeds persist and remain viable in the soil for reasonably long periods. The prolonged longevity of seeds of parthenium weed in the soil, coupled with allelopathy effect of varying growth stages of the plant on native plants, efficient dispersal via water (especially during flood) contribute to its spread and hence high invasiveness [2,4,5]. Parthenium weed has an economic impact on Australia's agriculture and landscape of ~A\$129 million a year [1]. This includes control costs and production losses to the agricultural sector. Control options such as pasture management, use of biological control agents, and herbicides exist for managing the weed [6]. However, many of these tactics are limited by their potential environmental impact and the sheer scale of the infestations [1,7]. Blanket herbicide spraying has been shown to adversely affect non-target plant species, and can reduce soil fertility; such chemicals also have the potential to migrate into the surrounding environment, impacting animal species and disrupting food chains [7]. This method of control is also economically unfeasible. Site-specific weed management (SSWM) options using Unmanned Aerial Vehicle (UAV) imagery for weed detection and control have been shown to reduce herbicide use by 14 to 39.2% in Maize fields when compared to conventional uniform blanket spraying [8]. In this example, an accurate weed map was necessary for efficient SSWM which highlights the importance of an accurate and robust detection method. Autonomous detection methods using red, green, blue (RGB), and hyperspectral imagery (HSI) for weed detection have shown very promising results in both crop and pastoral environments [9–13].

Attaining significant accuracy in automated weed recognition is challenging, but is gaining wide currency [9–12]. Some work has been done and reported on automated parthenium weed identification and mapping in natural and agricultural settings. In South Africa, at the landscape level, the spatial and temporal distribution of parthenium weed in response to rainfall using multi-year SPOT (Satellite Pour l'Observation de la Terre) data and random forest classification algorithm (one of several machine learning approaches) was explored with the general finding that the satellite images should be combined with UAV to provide higher spatial resolution data for improved detection accuracy [10]. The use and importance of TPOT (Tree-Based Pipeline Optimization Tool) and other Convolutional Neural Network (CNN) as effective machine learning tools for selecting and fine-tuning algorithms for parthenium weed discrimination and monitoring in pastoral field has also been demonstrated [9]. Similarly, the use of hyperspectral data for identification of optimal wavelength that are sensitive to discrimination of parthenium weed have been explored [9,10]. In most of the above cited cases, overall identification and classification accuracies varied between 68–86%, implying that approaches that may yield higher classification accuracies are still necessary [see also 9] for parthenium weed. In addition, there are no reported studies that have attempted to use machine learning to

identify and classify the discernable growth stages (seedling, rosette, and reproductive adults) of parthenium weed.

RGB imagery, captured using a digital camera, imitates the way the human eye perceives colour [14]. A typical digital single-lens reflex (DSLR) camera is equipped with a Bayer filter system consisting of a mosaic of red, green, and blue filters embedded on the image sensor [15]. Spatial resolution is another important factor for RGB imagery, particularly for object identification, and refers to the area where individual measurements can be made [11]. This is measured by ground sampling distance (GSD) which is the area in the real reference plane that is covered by a single pixel in the image. By combining an image sensor with a high number of pixels and a small scene size (often in millimeters), several features of interest can be easily discerned [16].

Object detection is the procedure of determining the instance of the class to which the object belongs and estimating the location of the object by outputting the bounding box around the object [17]. This means an object detector needs to distinguish objects of certain target classes from backgrounds in the image with precise localization and correct categorical label prediction to each object instance [18]. Convolutional Neural Networks (CNNs) have the capacity to learn very complex functions from data with minimal domain knowledge. Combined with large-scale labeled data and graphic processing units (GPUs) with high computational capacity, CNNs have been particularly successful for object detection [19].

Hyperspectral imaging is an analytical technique based on the study of the chemical and physical behaviour of the reflected or scattered light coming from a specific surface [20]. This type of imaging goes beyond the conventional image processing and spectroscopy because of its capability to obtain spatially resolved spectra from an object [21].

The objective of image classification is to recognize and characterize the classes, as a unique colour or gray level, which represents some features. These dissimilar colour classes can also be used to show the spectral characteristics of every unique pixel within the class [22]. This analysis obtains a colour map that distinguishes materials with different spectral profiles.

In this study, deep-learning algorithms and decision-tree-based protocols are applied to the challenges of automated detection and growth stage classification of invasive alien plants, in our case, to parthenium weed. In this context, we explore the viability of RGB and HSI data of parthenium weed captured in a control environment to aid in its field mapping procedure. Thus, our aim is to provide a platform for more efficient and accurate detection and mapping of parthenium weed using its captured images (in various growth stages: seedlings, rosettes, and reproductive adults) and subjected them to artificial intelligence (AI) algorithms. If the trial is successful, it will lead potentially to the optimisation of management protocols as the weed control cost can be minimised and efficacy improved simultaneously. The findings also have application potential in more complicated pasture and cropping environments where parthenium weed (and other invasive alien plants) are causing significant problems.

2. Materials and Methods

2.1. Preamble

It is hypothesized that deep-learning algorithm of Yolov4 (You Only Look Once version 4) and decision-tree machine learning of XGBoost (Extreme Gradient Boosting version 1.3.3) protocols when applied to their respective data types (RBG and HIS, respectively) will allow for highly accurate detection and growth stage classification of parthenium weed in a controlled environment, and that if successful each technique has the potential to extend to farm-scale and landscape scenarios. In this context, parthenium weed of different growth stages and non-parthenium plants (grasses) were imaged using RGB and HSI cameras. Non-parthenium plants that occur naturally with the weed in the field, including mature kangaroo grass (*Themeda triandra* Forssk.), and desert bluegrass (*Bothriochloa ewartiana* (Domin) C.E. Hubb.) were either interspersed with weed and photographed

(RGB case) or photographed individually and interspersed digitally (HSI case). Plants were interspersed in this way to simulate how these images would look if captured in the field. There were 665 unique RGB images captured in this process, and these images were labeled with a class (a unique designation for a group of objects) for parthenium plants and a class for parthenium flowers. The combined HSI image was generated and digitally combined to create a combined image, and this was labeled with six classes (so as to separate unique materials)—three for the parthenium weed growth stages of seedling, rosette, and flowering adult, one for the non-parthenium plants (grasses), and two classes for the soil in the pots and concrete ground the pots sat on. Respective datasets were then used to train a Yolov4 detector and an XGBoost classifier (a decision-tree-based ensemble machine learning algorithm that uses a gradient boosting framework). The procedures for the capture and analysis of the RBG and HSI datasets are given below.

2.2. RGB Dataset Preparation

The RGB dataset was created with the aim of simulating what is obtainable in a pastoral environment as closely as possible. Plants were grown from seeds inside a secure temperature-controlled (27/22 ± 2 °C, day/night) glasshouse at the Ecosciences Precinct, Dutton Park (ESP), Brisbane, Australia by Biosecurity Queensland (BQ) staff of the Queensland Department of Agriculture and Fisheries. The plants included parthenium weed at three distinct growth stages, and control plants consisting of several growth stages of kangaroo grass (*Themeda triandra*), and desert bluegrass (*Bothriochloa ewartiana*). The parthenium weed growth stages used were seedling (<1-month-old), rosette (1–2 months old), and flowering (2–4 months old; see Figure 1). Seedlings are parthenium weed plants that had not developed the rosette structure, rosette included established parthenium weeds without flowers, and flowering plants included established reproductive parthenium weeds with flowers and/or seeds. In total 8 seedling plants, 6 rosette plants, and 12 flowering plants made up the dataset as well as the 12 individual plants of the grass species mentioned above.
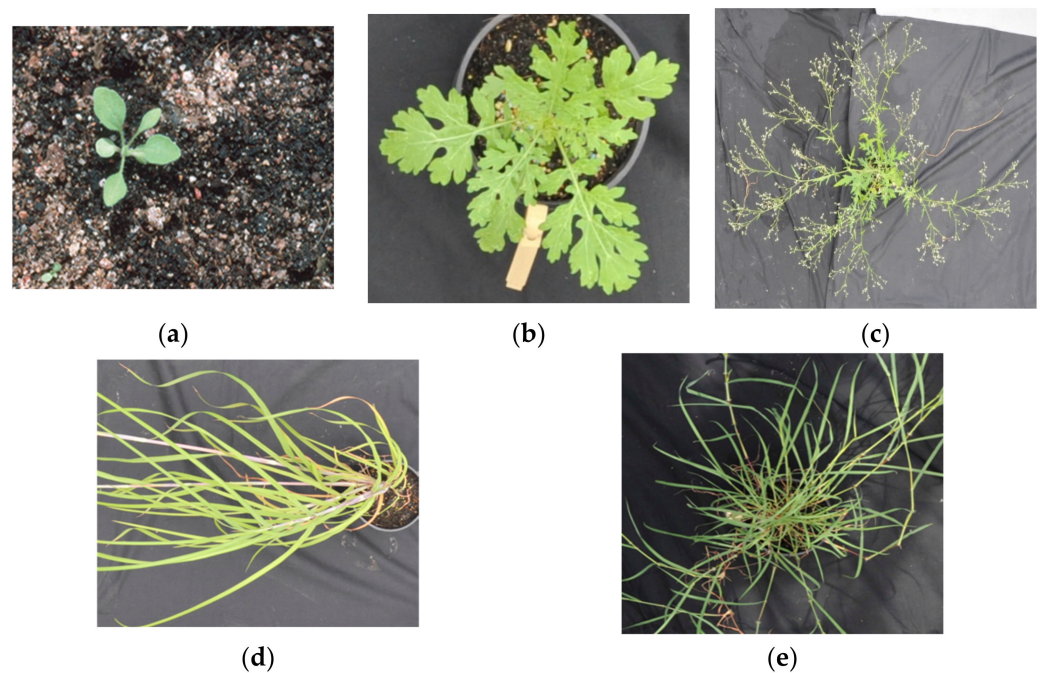
Figure 1. Aerial photos of the three parthenium weed growth stages of (**a**) seedling, (**b**) rosette, (**c**) flowering adult plant, and co-occurring monocot plants of (**d**) desert bluegrass (*Bothriochloa ewartiana*) and (**e**) kangaroo grass (*Themeda triandra*).

On 4 March 2021, a 1.5 m × 2.25 m area was set up with direct sun exposure outside the glass-house area of ESP, Dutton Park, Brisbane, Australia. Images were captured using

a Nikon D2 with a 24–70 mm lens and a focal length of 70 mm. Seedling and rosette plants were imaged first, beginning with no plants in the arrangement and gradually increasing up to five plants per growth stage for imaging. This step was then repeated, first with two non-parthenium or control plants (grasses) up to five control plants. This process was then repeated for flowering plants and finally for a combination of seedlings, rosettes, and flowering plants of the parthenium weed. After these structured arrangements were captured, more complicated arrangements with higher numbers of parthenium weeds of each growth stage and higher numbers of non-parthenium plants were imaged. These complicated arrangements also included high plant occlusion to mimic overlapping which is common in pastoral environments. The resulting dataset contained 665 unique images with parthenium weed and non-parthenium plants in unique arrangements. On 18 March the same process as above was used again to capture another RGB dataset that is completely unique to the training and validation dataset. This resulted in a unique test dataset of 94 images.

2.3. RGB Image Processing

Processing of the RGB images first required the images to be accurately labeled. Labeling was applied to only parthenium weeds with control plants remaining unlabeled. The label convention used was one that grouped flowering and non-flowering parthenium weeds into a single class and a second class for individual flowers. Plants of each growth stage were combined in this way due to their visually similar characteristics while the flower class was used to detect the presence of flowers which was then used to classify flowering and non-flowering plants. Parthenium seedlings remained unlabeled due to their visually indistinct features and very small size (see Figure 1).

Once labeled, the images were tiled or broken up into smaller images of 1024 × 1024 pixels so they could be used to train a Yolov4 detection network. This size was chosen to reduce the effect of down-sampling done by the detection network whilst ensuring the largest objects could still be represented in a single image, as shown in Figure 2. Consecutive rows and columns, seen below (Figure 2b), also have a 50% overlap. This was done so that each plant would be more likely to fall towards the center of the tile and so entire plants could be represented with a single bounding box. This also allowed for the removal of bounding boxes that fall on small fractions of a plant as a result of the tiling process.

After tiling each image, all tiles were checked to remove any inaccurate labels created during the tiling process. Augmentation- a technique of alternating the existing data to create more, unique data for the model training process- was implemented as it provides two benefits in the training stage: (1) it reduces overfitting probabilities by increasing the variability of the features used to make detections; and (2) it increases the size of the training dataset and hence the number of unique objects labeled. Geometric and radiometric augmentations (techniques that can be used to artificially expand the size of a training dataset by creating modified versions of images in the dataset) were applied randomly within predetermined bounds so that all augmented images do not misrepresent how the parthenium weed plants naturally occurs. The augmentation parameters and examples can be seen in Table 1 and Figure 3, respectively. The tiling process resulted in a dataset of 27,230 images. This set of images was randomly split using a ratio of 8:2 resulting in a training set of 21,784 images and a validation set of 5446 images. The training set of 21,784 images was then augmented which doubled the size of the training set to 43,568 images. This gives a training to validation ratio of 0.875:0.125, respectively. The second set of independent RBG images taken two weeks after the initial image capture (see above, i.e., *RBG data preparation* section) served as the test dataset (a total of 4230 once the images were tiled) and were used to benchmark the performance of the trained detector using images the detector has never seen before.
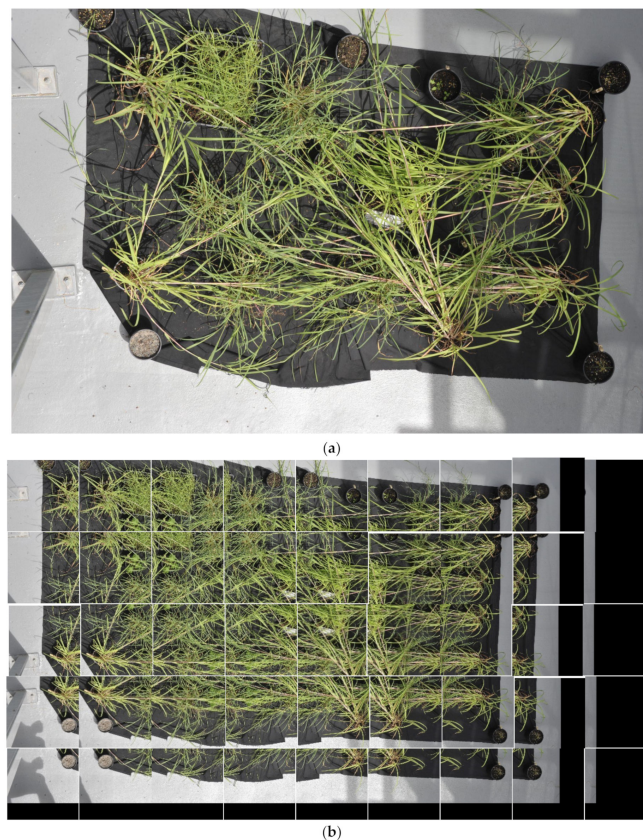
Figure 2. Example of a raw image captured (**a**), and the resulting 1024 × 1024 tiles with 50% overlap between consecutive rows and columns (**b**).

Table 1. A list of image augmentations applied to the dataset and their descriptions.

| Augmentation	Description
Brightness	Changes in how reflective each pixel appears.
Contrast	Changes in brightness between pixels.
Saturation	Changes in intensity and purity of a pixel.
Hue	Changes in the actual colour of a pixel.
Rotation	Changes in the angle of the image.
Reflection	Reflection of one half of the image to the other half about the horizontal or vertical axis.
Scale	Changes in image size.
Shear	Slants the shape of the image either in the x or y direction
Translation	Moves the image in the x and/or y direction.

2.4. Yolov4 Modelling

To test the performance of current state-of-the-art (SOTA) detectors, several benchmark datasets have been created. Of these, the Microsoft common objects in context (MS COCO) dataset was designed for the detection and segmentation of objects occurring in their natural context [23]. Yolov4 is a single-stage detector that boasts a mean average precision (mAP) of 65.7% at a detection speed of ~65 frames per second (FPS) when benchmarked on the MS COCO dataset [24]. Yolov4 is one of the few high-performing detectors capable of real-time object detection which is above 20fps and hence was used in this study [18].

YOLOv4 is a one-stage detector consisting of a CSPDarknet53 backbone, SPP additional block, and PANet path-aggregation neck with YOLOv3 dense predictor for the head as its architecture [24]. The structure of Yolov4 can be seen below in Figure 4.

Figure 3. Examples of augmented images with all augmentation types applied randomly within preset bounds.
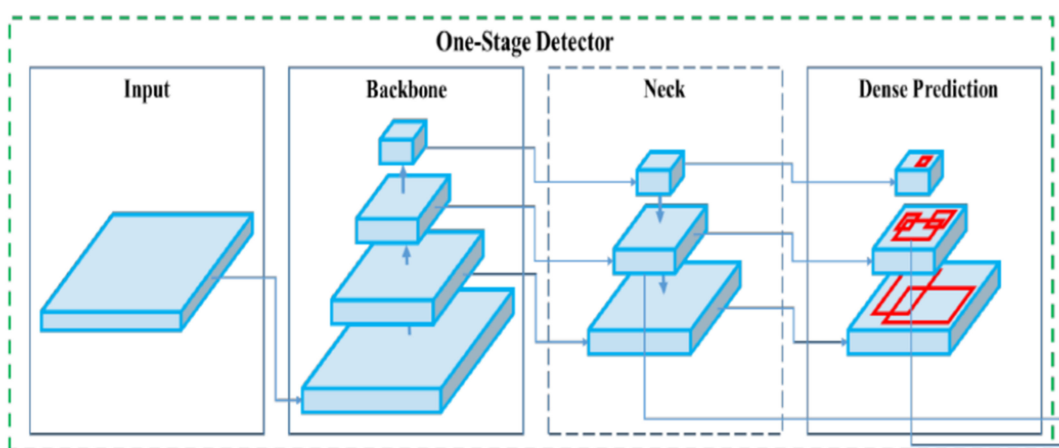
Figure 4. Yolov4 structure (adapted from [24]: Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-J.M. YOLOv4 Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934, with permission from Cornell University).

The CSPDarknet53 backbone involves applying a cross-stage partial network of Wang et al. on the Darknet-53 network developed by Redmon & Farhadi [25,26]. Darknet-53 performed on par or even better in terms of accuracy, took fewer floating-point operations

to train, and performed faster when compared to ResNet-101 and ResNet-153. CSPNet reduced computations by up to 20% with equivalent or higher accuracies when applied to backbones such as ResNet, ResNeXt, and DenseNet.

For the neck, the spatial pyramid pooling (SPP) additional block developed by He et al. eliminates the need for a fixed-size input image [27]. It does this by computing feature maps from entire images once and generating fixed-length representations for training. SPP has been shown to improve the accuracy of detection networks such as R-CNN and vastly improve processing speed when generating detections. The other part of the Yolov4 neck is PANet path-aggregation developed by Liu et al. [28]. This works by shortening the information path between lower layers and topmost feature levels. It does this by linking feature grids and all feature levels to make useful information in each level propagate directly to following proposal subnetworks. PANet path-aggregation boasts higher mean average precisions than Mask R-CNN on COCO datasets and ResNet-50 on the MVD dataset.

Finally, the head or dense predictor uses Yolov3 to produce bounding boxes based on dimension clusters as anchor boxes and thus predicts four coordinates for each box [26]. These predicted bounding boxes are compared to the ground truth bounding boxes and the sum of the squared error loss is used to determine the gradient for regression. Using logistic regression, the predicted bounding box that overlaps with the ground truth the most is chosen. Classes are then predicted using independent logistic classifiers and binary cross-entropy loss. This allows for more accurate class predictions when objects of different classes overlap. Three different scales are used when predicting bounding boxes so that features at different scales can be extracted and combined for a robust feature map. This approach allows for more meaningful semantic information from unsampled features and finer-grained information from earlier feature maps. Lastly, Darknet-53, consisting of 53 convolutional layers is used for the feature extractor. The convolutional layers are successive 3×3 and 1×1 layers with shortcut connections.

2.5. HSI Dataset Preparation

For the present study, a Headwall Nano-Hyperspec visible and near-infrared (VNIR) image sensor which is a type of push-broom or line scanner, optimized for UAVs, was used. With this scanner, the hyperspectral information is obtained from contiguous lines of the sample by relative movement between the sample and the scanner [21]. From this, hyperspectral reflectance information is obtained and can be used to determine the structural and biochemical properties of the sample [29]. The scanner machine was set up inside the ESP glasshouse with some filtering of ultraviolet light (UV) and a bright light source on one side of the scanner to increase the light available to the scanner. The same parthenium weed and non-parthenium plants (grasses) from the RGB dataset were used to create this hyperspectral dataset as well as another monocot (Navua sedge—*Cyperus aromaticus* (Ridley) Mattf. & Kukenth) that was not included in the RGB dataset. One or two parthenium weed plants of the same growth stage were scanned at a time until all individual parthenium weed of each growth stage had been scanned (Figure 5). Each control plant (bluegrass [*B. ewartiana*], kangaroo grass [*T. triandra*] or Navua sedge [*C. aromaticus*]) was also scanned individually.

2.6. HSI Image Processing

The hypercubes generated from the line scanner include 274 spectral bands from 395.794 nm to 1002.02 nm. The first processing step was to convert the raw pixel radiance values to reflectance. This is necessary to account for variations in available/reflective/impinging/absorbed light when capturing the hyperspectral image [30]. This is done by including a material with a known reflectance in the hypercube and using this to generate a ratio between the illuminating light power and the reflected light power (Figure 5). This can

then be applied to the spectral measurements of each pixel in the hypercube following the equation.

$$R(x,\lambda) = \frac{P_r(x,\lambda)}{P_i(x,\lambda)} \qquad (1)$$

where $R(x,\lambda)$ is the reflectance at wavelength λ for the location x, $P_r(x,\lambda)$ is the reflected light power from the object, and $P_i(x,\lambda)$ is the illuminating light power.
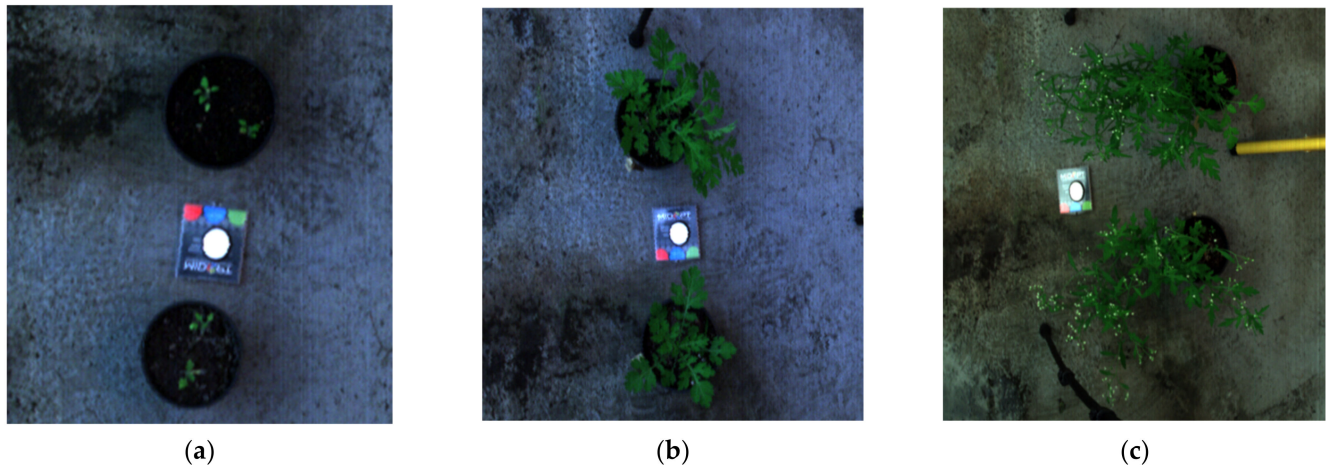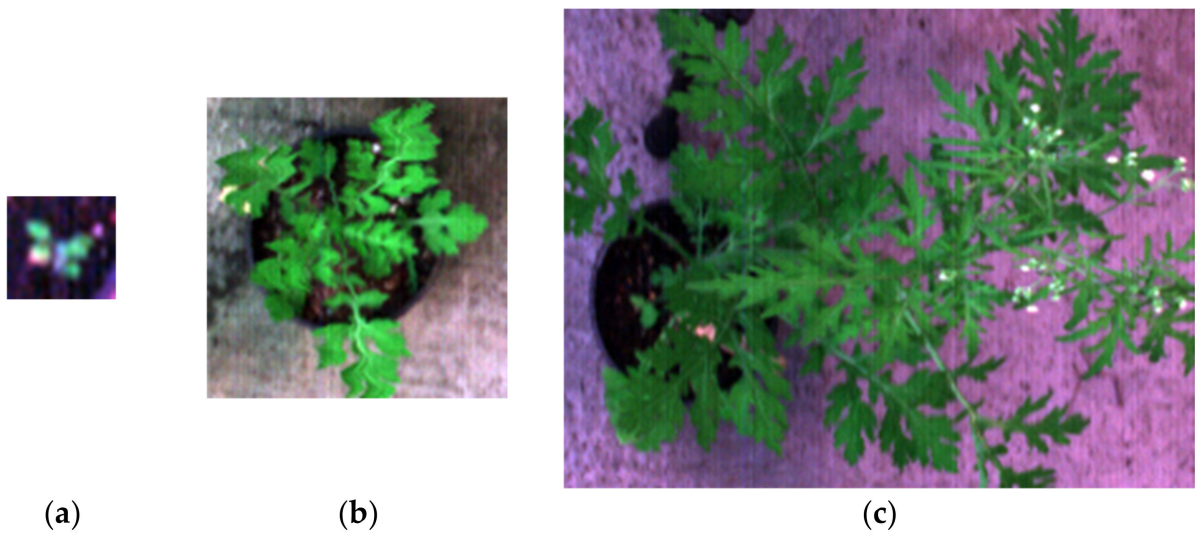

Figure 5. Hypercubes of the three parthenium weed growth stages (**a**) seedling (**b**) rosette (**c**) flowering adult.

Upon inspection, there were some bands that included significant noise, so it was necessary to remove these bands to ensure these bands did not degrade the analysis of the hypercube. These bands included 395.794–426.883 nm and 999.798–1002.02 nm. The hypercubes were then reduced to their regions of interest (ROI) which correspond with the plants in those images. This adjustment procedure eliminates unnecessary pixels that do not include plant samples, reducing the size of the hypercube file and unnecessary complexity. Filtering was performed to remove unwanted noise and generate a more accurate spectral measurement. Filtering methods include Savitzky-Golay spectral filtering with a window length of 25, a polynomial order of 9, and Gaussian filtering with a kernel size of 61. Spectral measurements are represented in the OKLAB colour space. Some examples of the filtered ROIs of parthenium weed plants are seen in Figure 6.

Using custom python code developed by one of the authors (Juan Sandino), the filtered and reduced hypercubes of parthenium weed plants and non-parthenium plants were combined into a larger image in a similar way that the RGB images were manually combined. Non-parthenium plants made up the background of these combined images and the parthenium weeds were superimposed on top. This means that there is no natural plant occlusion which impacts how closely these images mimic a real field. A combined image can be seen in Figure 7.

RGB representations of the combined hypercubes were then generated so that label masks could be made. These masks provided the classifier with training and validation data so that it could generate a trained model. This process involved creating an individual colour mask for each material in the hypercube. Six distinct colour masks were created consisting of flowering (red), rosette (cyan), seedling (magenta), background grass (yellow), concrete (blue), and soil (grey). Effort was made to label as little as possible whilst achieving accurate pixel classifications so that significant test pixels were available to demonstrate classifier performance. The labeled counterpart of Figure 7 can be seen in Figure 8. At this point, the XGBoost classifier could be trained and evaluated.

Figure 6. Filtered regions of interest (ROI) of the three parthenium growth stages: (**a**) seedling, (**b**) rosette, and (**c**) flowering adult.
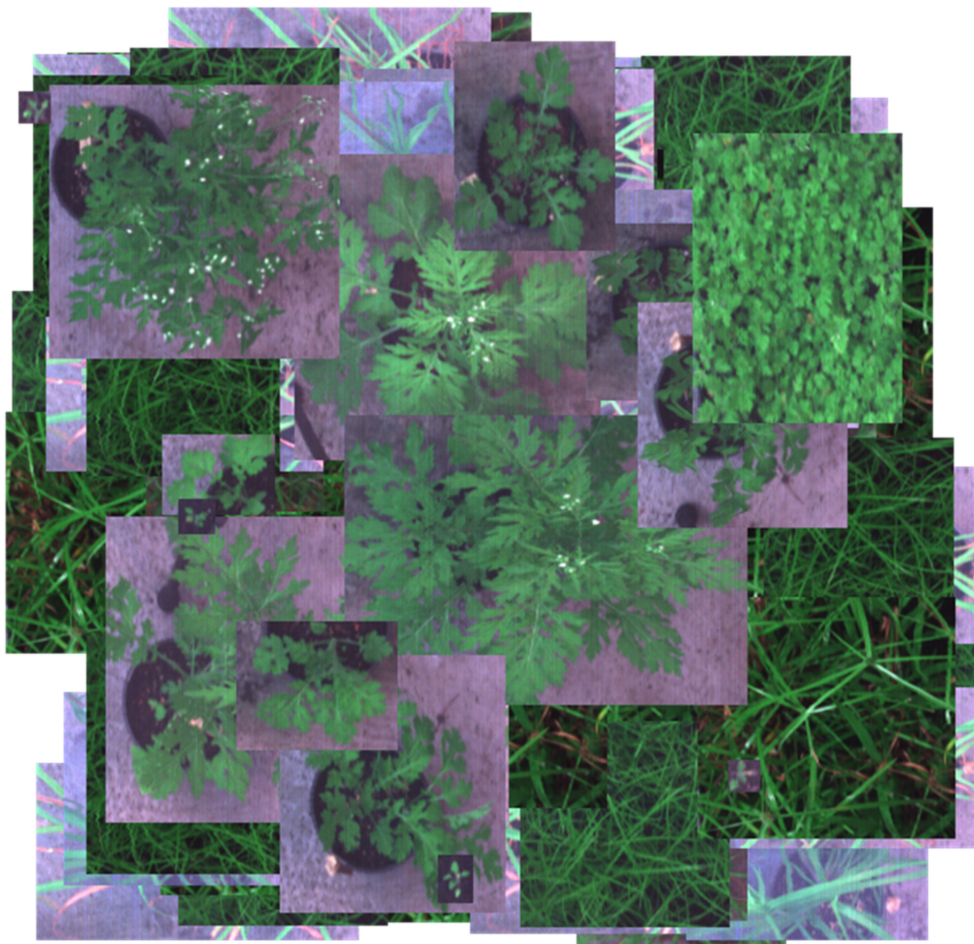
Figure 7. Combined hypercube with seedling, rosette, and flowering parthenium weed plants over non-parthenium plants.
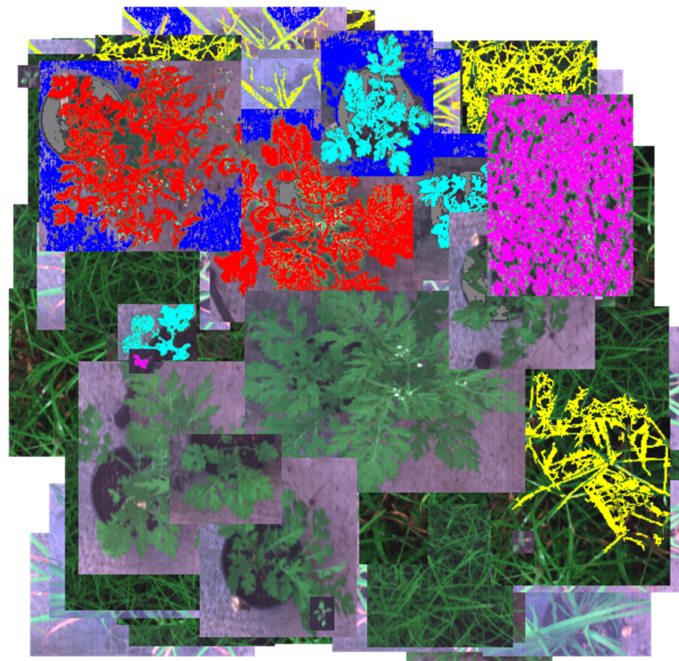
Figure 8. Combined hypercube with label mask.

2.7. XGBoost Modelling

Many different classification methods have been applied to hyperspectral image (HSI) classification with very good results, e.g., [31–33]. XGBoost (version 1.3.3), developed by Chen & Guestrin, is a scalable tree boosting system that boasts impressive performance in many machine learning benchmarking competitions [34]. Based on tree boosting, which has been shown to give state of the art (SOTA) results on many classification benchmarks, XGBoost's scalability allows it to be applied to large datasets without significant computer hardware [35].

This is made possible by XGBoost's unified approach to gradient tree boosting, exploiting several well-developed and novel systems to fit a variety of problem types. These include a novel tree learning algorithm for handling sparse data, a weighted quantile sketch procedure for handling instance weights in approximate tree learning, parallel and distributed computing, enabling quicker model exploration and out-of-core computation allowing hundreds of millions of examples to be processed efficiently [35].

2.8. Performance Metrics

The metrics used to evaluate the performance of the Yolov4 algorithm include true positives (TP), false positives (FP), true negative (TN), false negatives (FN), precision, recall, F1-score, and accuracy. True negatives are excluded as a performance metric for parthenium weed detection as they are not relevant in object detection problems; however, they are relevant for the classification of the flowering/non-flowering and so have been included in the latter. As for the XGBoost classifier, precision, recall and F1-score are the metrics used.

True positives are simply the number of correct detections; similarly, false positives are incorrect detections. False negatives are the number of objects not detected; however true negatives are unique. True negatives apply when an object, for example, a detected parthenium weed can either be one class or another; in this case flowering or not flowering. True negatives do not apply to object detection as this includes location information and so becomes infinite. They do apply to classification which is the case when determining flowering and non-flowering status as the location of the parthenium weed plants has already been determined. In this case, a true negative is a plant correctly identified as non-flowering, or having no flowers detected, and a true positive is a flowering plant with detected flowers.

"Precision" describes the proportion of positive identifications that are correct. A high precision value says that of the detections made, very few of them are wrong. If all detections made were correct the precision score would be 1.0. "Recall" describes the proportion of actual objects, in this case parthenium weed plants, that were identified. A high recall value suggests that of the detectable parthenium weed plants in the data, almost all of them were identified. If all parthenium weed plants in the dataset were detected, the recall score would be 1.0. "F-1 score" is the harmonic mean of precision and recall; it is a good measure of accuracy as it describes the relationship between precision and recall. If one is high and the other is low, the F-1 score will be negatively impacted. A high F-1 score implies high precision and recall scores, which suggest there are high TP and TN and low FP and FN. Lastly, accuracy represents the number of correctly classified or detected objects over the total number of correct and incorrect detections/classifications. The equations for each of these metrics can be seen in Table 2.

Table 2. A list of performance metrics used to evaluate the Yolov4 and XGBoost algorithm performance with as well as the equations to calculate each.

Performance Metric	Equation
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F1-score	$2 * \frac{Precision*Recall}{Precision+Recall}$
Accuracy (object detection)	$\frac{TP}{TP+FP+FN}$
Accuracy (classification)	$\frac{TP+TN}{TP+FP+TN+FN}$

2.9. Ordination

Another method of analysis used specifically for the hyperspectral data is ordination [36–38]. The primary purpose of this technique is the reduction of a high number of variables (in this case, the spectra bands) that could differentiate parthenium weed from co-occurring plant species; these reduced but unique set of variables represent as many characteristics as possible of the phenomenon analysed in some latent variables [38]. This occurs through a linear transformation of the variables that projects the original ones into a new Cartesian system in which the variables are sorted in decreasing order of variance: therefore, the variable with the greatest variance is projected on the first axis, the second greatest variance on the second axis, and so on. The reduction of complexity occurs by limiting itself to analysing the main (by variance) new variables. Unlike other (linear) transformations of variables practiced in the field of statistics, in this ordination technique it is the same data that determine the transformation vectors [36].

We normalised that spectra data and used common parametric analyses (ANOVA and generalised linear modelling) and multivariate ordination technique of metric multidimensional scaling (MDS) to detect differences in individual and mean spectra signature of each of the three parthenium weed growth stages and non-parthenium (control) plant species (Navua sedge, kangaroo grass and desert bluegrass). For mean spectral signatures, data were grouped into bands of blue (400–500 nm), green (501–600 nm), red (601–700 nm), near infra-red (700–900 nm) and infrared (>900 nm) bands. All ordination analyses were performed on PRIMER version 7.0. software (PRIMER-E Ltd., Plymouth, UK).
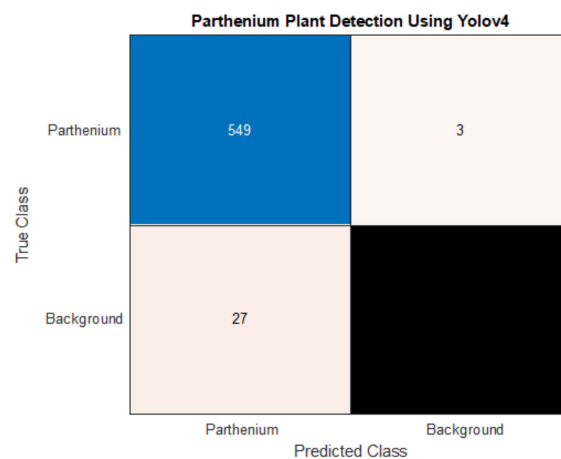
3. Results

3.1. RGB—Yolov4

Parthenium weed detection accuracy was determined by the number of parthenium weed plants with and without a parthenium weed class bounding box on any part of the plant. Additionally, considered was the number of parthenium weed class bounding boxes with no parthenium weed plants inside. A parthenium weed with at least one parthenium weed class bounding box on any part of the plant was considered a true positive. Similarly, a parthenium weed with no bounding box on any part of the plant was considered a false negative. Finally, a parthenium weed class box with no part of a parthenium weed inside

was considered a false positive. Growth stage classification was determined using a flower class detector. It was decided that this would be a classification problem as the goal was not to detect all flowers but to detect any flowers on a given flowering plant. If a flowering plant had at least one or more of its flowers detected, it was considered a true positive. If in that same group of correctly detected flowers, there were objects miss-classified as flowers this was still considered a true positive as the miss-classifications will not add any confusion in determining the flower status of that plant. Similarly, a non-flowering plant with no flower detections on or near the plant that could not be associated with a flowering plant was considered a true negative. If a flowering plant had none of its flowers detected it was considered a false negative. If a non-flowering plant had flower detections on or near it that could not be associated with a flowering plant, it was considered a false positive. Some flower detections were made on non-flower objects, but they were sufficiently far away from any plants to not cause any confusion when determining flower status and so were ignored. Examples of the detections made can be seen in Figure 9.
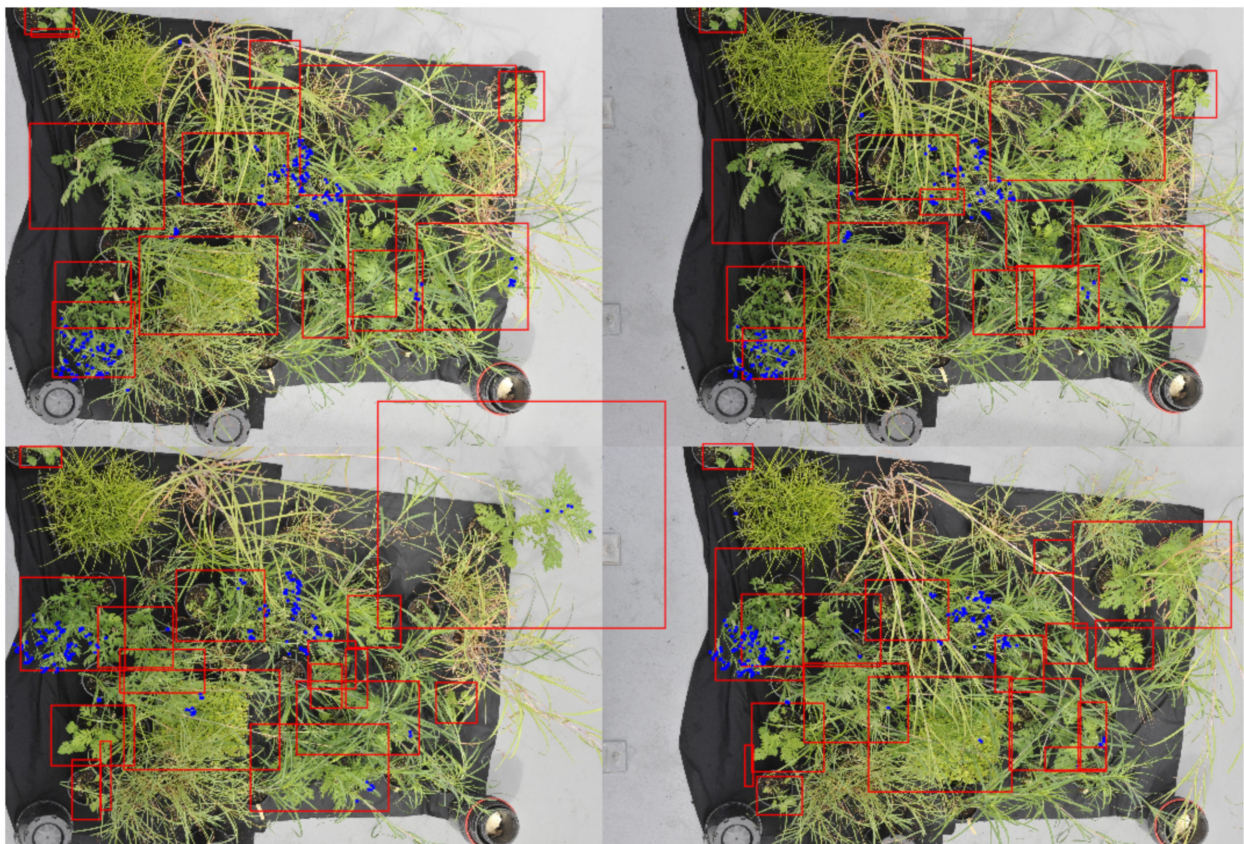

Figure 9. Examples of a simulated orthorectified image with parthenium weed detections in red and flower detections in blue.

The performance of the Yolov4 detector can be seen in Figures 10 and 11. The Yolov4 detector included a class for parthenium weed plants and a class for individual flowers with the parthenium weed class being an object detection problem and the flowering class being a classification problem. The Yolov4 detector achieved a detection accuracy of 94% for parthenium weed plants and a classification accuracy of 86% for flowering or non-flowering plants. All these detections were made on images unique to both the training and validation sets using the same plants. These plants had undergone two weeks of growth before being imaged for the test set and so were effectively unique to the plants used in the training and validation sets.

Figure 10. Confusion matrix depicting predicted parthenium weed/background class numbers vs. true parthenium weed/background class numbers. The background includes control plants and no plants. The resulting performance metrics are indicated in Table 3.
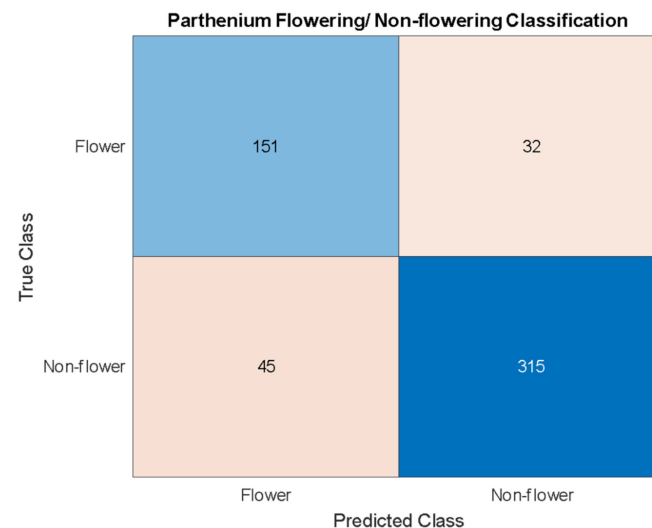
Figure 11. Confusion matrix depicting predicted flowering/non-flowering class numbers vs. true flowering/non-flowering class numbers. The resulting performance metrics are indicated in Table 3.

Table 3. Performance metrics of the Yolov4 algorithm as a parthenium weed detector and as a classifier of the weed into flowering vs. non-flowering status.

	Metric							
Detection Task	**Precision**	**Recall**	**F1-Score**	**TP**	**FP**	**TN**	**FN**	**Accuracy**
Object detection Parthenium weed	0.95	0.99	0.97	549	27	N/A	3	0.95
Object classification Parthenium Flower vs. Non-flower	0.83	0.77	0.8	151	32	315	45	0.86

3.2. Hyperspectral—XGBoost

To evaluate the performance of the XGBoost classifier, a combined hypercube was made with four parthenium flowering plants, six rosettes, and four seedlings as well as all non-parthenium hypercubes included in it. The pixel classification performance can be seen in Table 4. The associated labels can be seen again in Figure 8, and the predicted pixels of the combined hypercube can be seen in Figure 12. For the hypercube, the classifier achieved a performance of 99–100% on the validation subset of labeled training data, represented

by the colour masks. This translated to a significant amount of all labeled and unlabeled pixels in each hypercube receiving a correct classification. In the combined hypercube, all parthenium weed plants across the three growth stages can be easily identified.

**Table 4.** Performance metrics of the XGBoost as a hyperspectral pixel classifier.

Object Class	Precision	Recall	F1-Score	No. of Labeled Pixels
Parthenium weed-Flowering	1.00	1.00	1.00	5556
Parthenium weed- Rosette	0.99	0.99	0.99	1806
Parthenium weed-Seedling	1.00	1.00	1.00	4580
Grass	1.00	1.00	1.00	3510
Concrete	0.99	0.99	0.99	3916
Soil	0.99	0.98	0.99	911

3.3. Parthenium Weed Hyperspectral Signatures

In order for the XGBoost algorithm to differentiate each growth stage of parthenium weed from each other and from the control plants (grasses and a sedge), the spectral signatures of the parthenium weed and non-parthenium (control) plants were captured. Figure 13 shows all of the parthenium weed spectral signatures across the three growth stages. While the signatures are certainly quite similar, which is expected of plants of the same species, there are subtle differences between the growth stages that are enough to differentiate them from each other using the XGBoost algorithm.
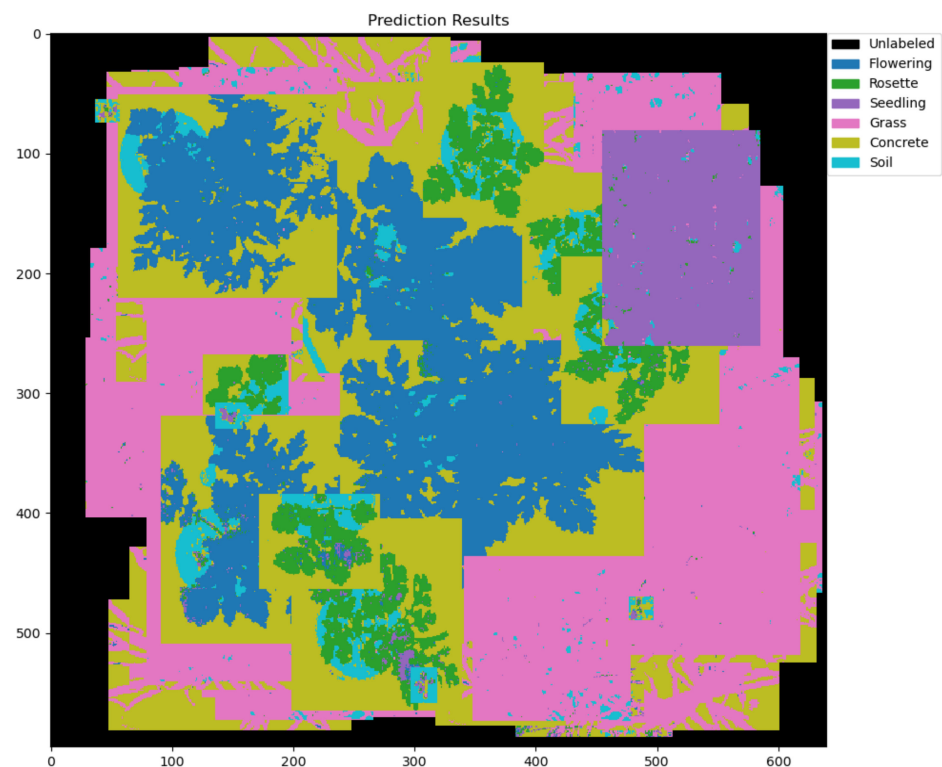

Figure 12. Output of combined hypercube with predicted colour masks.
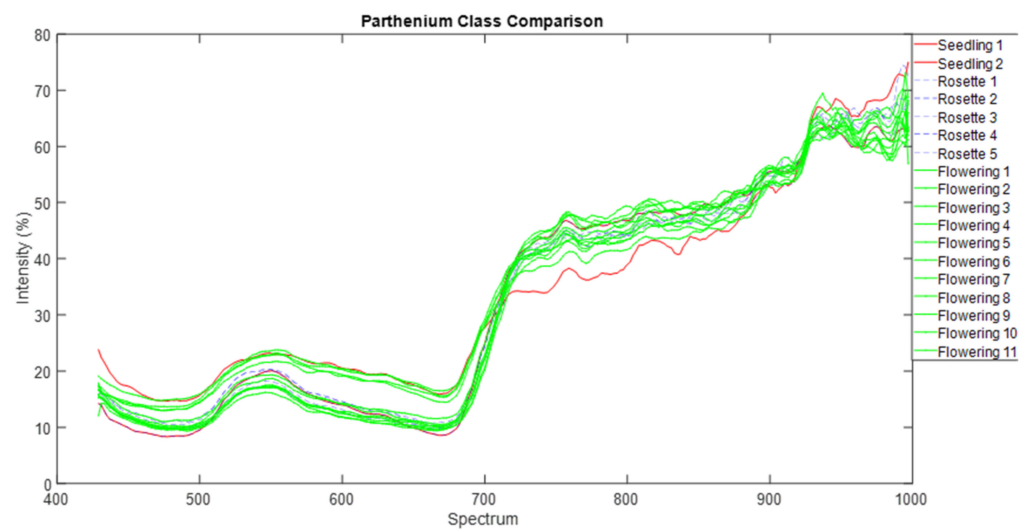
Figure 13. Spectral signature of all parthenium weed plants used in this investigation grouped in their respective growth stages.
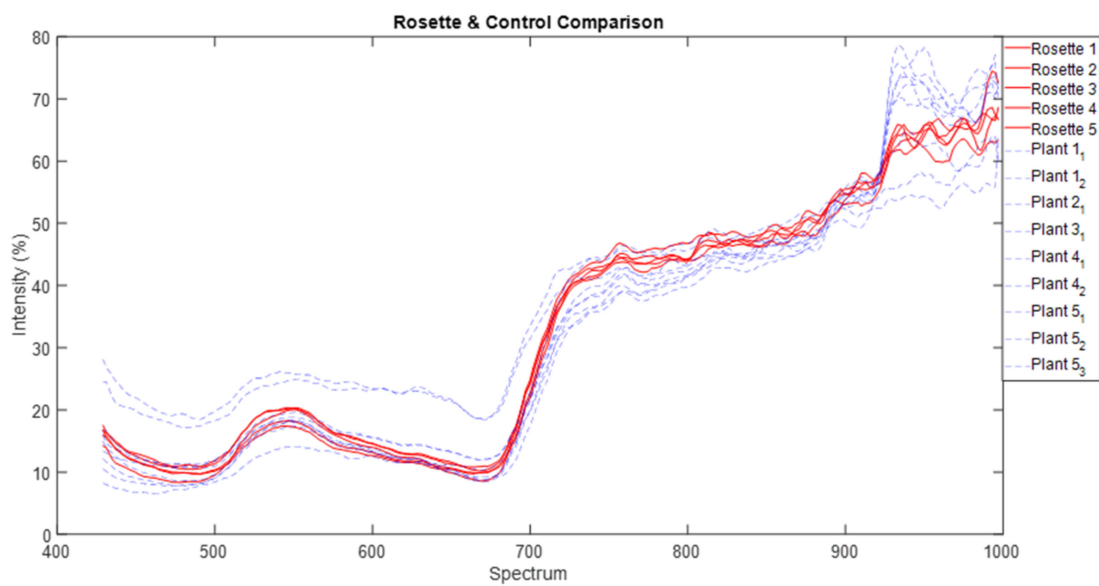
Figure 14 shows the spectral signatures of the parthenium weed seedlings compared to the control plants that commonly co-occur with the weed in a grazing (pasture) landscape. Spectra differences are much more pronounced than that depicted in Figure 13, and clearly show differences in the spectral signatures of parthenium weed seedlings and non-parthenium plants. These differences exist across the spectral signatures from 400 nm to 1000 nm.
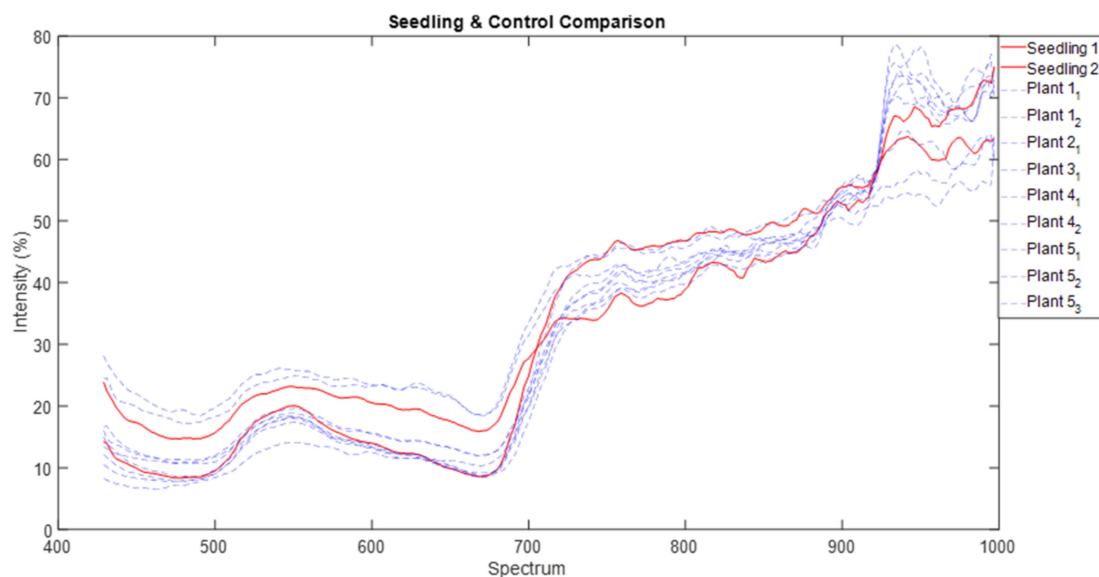
Figure 14. Spectral signature of the seedling stage of parthenium weed compared to the non-parthenium/control plants (Plant 1–Plant 5).

Figure 15 shows the spectral signatures of the parthenium weed rosettes compared to the control plants. Similar to the seedlings, there are obvious differences between parthenium weed rosettes and the spectral signatures of the control plants. While the signatures are similar between 700 nm and 900 nm, between 400 nm and 700 nm and between 900 nm and 1000 nm, the differences are obvious.

Figure 15. Spectral signature of the rosette stage of parthenium weed compared to the non-parthenium/control plants (Plant 1–Plant 5).

Finally, Figure 16 shows the spectral signatures of the flowering parthenium weed plants compared to the control plants. The differences between flowering parthenium weed and control plants are less pronounced than that depicted in Figure 15 above (i.e., parthenium weed rosette and control plants), however there are significant differences, particularly between 900 nm and 1000 nm.
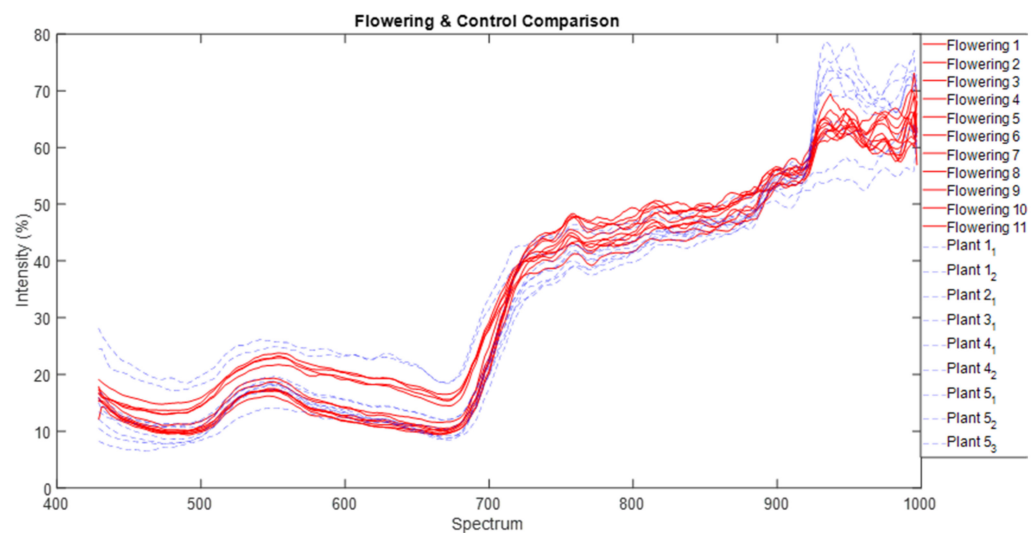

Figure 16. Spectral signature of the flowering adult stage of parthenium weed compared to the non-parthenium/control plants (Plant 1–Plant 5).

3.4. Hyperspectral Ordination

Results from ordination are shown in Figure 17. Two axes were enough to capture the majority (90.5%) of the variation in the dataset (Axis I—60.8%, Axis II–29.7%). Like Figures 13, 15 and 16, wavelengths responsible for major differentiation of parthenium weed plants from control plants (grasses and sedges) are mainly in the blue (400–500 nm) and the infrared (>900 nm) ranges (*F-ratio* between plant species = 14.21, 3.67, 3.90, 1.71 and 10.35 for blue, green, red, near infra-red and infra-red bands, respectively; all with $p < 0.05$, except NIR (700–900 nm)) which was non-significant). However, note that while Axis I indeed captured the majority of the variation in the dataset, the distribution along

this axis of the three parthenium weed growth stages was wide (i.e., scattered) rather than concentrated, especially for flowering parthenium weed and the seedlings.
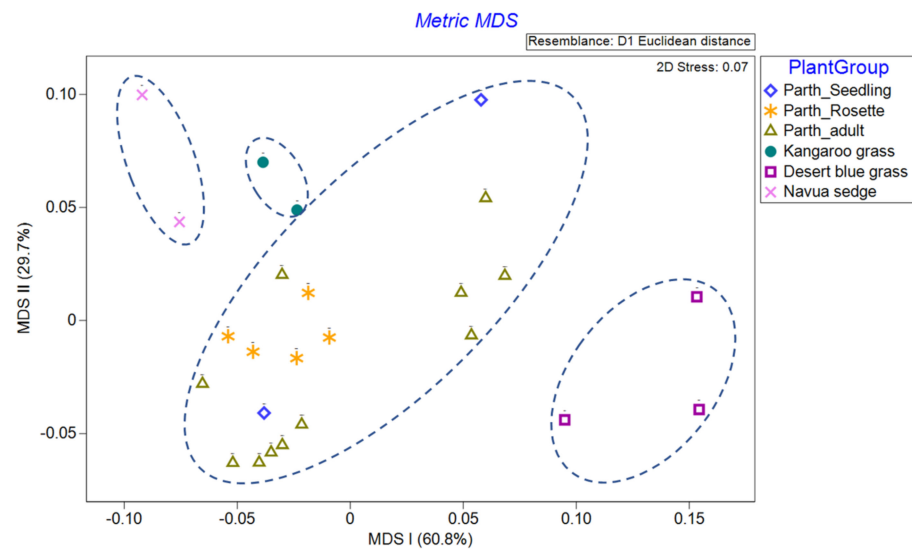
Figure 17. Results of ordination performed on the spectral signatures of each plant, showing separate groupings of each plant species (circled). Within the parthenium grouping, plants in the rosette stage are highly aggregated (asterisks), while those of flowering plants appeared in two subgroups (open triangles).

4. Discussion

The findings from the current study suggest that machine learning techniques can be applied effectively and efficiently to invasive weed detection in managed and pastoral environments. We have used hyperspectral and RGB imaging techniques to create two distinct datasets with the same sample plants, and hence estimate the effectiveness of two machine learning algorithms. Object detection using Yolov4 was applied to the RGB dataset, while pixel classification using XGBoost was applied to the hyperspectral dataset. Yolov4 achieved a detection accuracy for the invasive parthenium weed of 95% and was able to classify flowering and non-flowering growth stages of the weed with an accuracy of 86%. XGBoost achieved a pixel classification accuracy of 99% with 100% of the seedling, rosette, and flowering plants being identified with a colour mask over the body of the plant.

As well as a detection accuracy of 95%, seen in Table 3, the precision and recall for parthenium weed detection is 95% and 99%, respectively. While both scores are high, a very high recall suggests that there are very few false negatives which is the case with only three false negatives in the test dataset. In this case, a high recall is very desirable as false negatives mean undetected and potentially uncontrolled plants that may remain in an affected area targeted for local eradication, reach reproductive maturity and continue to add to the parthenium weed seed bank [1,39]. Low numbers of false positives, whilst also desirable, are comparatively less so. False positives (4.7% in this study) may trigger parthenium weed control methods to be applied when in reality, it is not necessary, though this may be a reasonable compromise when eradication of the weed is the goal (see Regan et al. (2011) and Panetta (2015). for in-depth analysis of this treatise [40,41]).

The precision and recall for flowering/non-flowering classification (83%) are lower than the actual parthenium weed detection (95%), but this is also a more difficult problem. The approach taken for this classification is centered around object detection with the objects being parthenium weed flowers. A single flower detection on a parthenium weed is enough to classify that parthenium weed as a flowering plant. Thus, the problem is one of both classification and object detection. Because flowers still need to be detected by the Yolov4 algorithm they need to be treated as a unique class like the parthenium weed class. However, the flowers are much smaller and so have much less information from which

the algorithm can extract relationships. An accuracy score of 86% reported in this work is not insignificant but does not achieve the goal of a very accurate flowering/non-flowering differentiation. The results do suggest that this higher accuracy is achievable, and this will likely be done with data of a lower GSD, affording this class much more detail from which to extract relationships. This would also improve the parthenium weed detection ability.

Precision and recall scores achieved by XGBoost for classes in the hyperspectral dataset are all between 98% and 100% (Table 4). These are for pixel classification, not for object detection; however, if all the pixels of an object are classified in the correct class, then the object is effectively detected. The colour masks achieved by XGBoost for each object are also more useful than a bounding box as it gives the true size and structure of the plant. What XGBoost does particularly well is to differentiate all three growth stages of parthenium weed- seedling, rosette, and flowering plants. Growth stages were differentiated in the RGB dataset using the presence of flowers. However, Yolov4 could not determine enough of a difference when flowering, rosette, and seedling plants were split into unique classes; there just wasn't enough unique relationships in the RGB data. XGBoost on the other hand looks at the spectral signatures of each growth stage and has reliably determined unique relationships between each stage. It was interesting to note that using metric MDS ordination, though Axis I captured the majority of the variation in the spectra dataset, however the distribution along this axis of the three parthenium weed growth stages was wide (i.e., scattered) rather than concentrated, especially for flowering parthenium weed plants and the seedlings. While we lack enough replicates for the seedlings (only two individuals were scanned), and as such it is safe to remove/discount this growth stage from our analyses for lack of power, the display of two discrete groups seen for the flowering (adult) plant stage might not be unexpected. Flowering parthenium weed plants (adults) that we used are probably in different phenological stages of reproduction and/or leaf retention/shedding, and this might have led to the variation in spectral signature captured and seen. Hyperspectral reflectance curves are often affected by plant phenological stage [42,43]. We are left with the rosette growth stage having consistent spectral signature, suggesting that this growth stage might have the greatest utility in terms of differentiation (i.e., maximum separability) from other plants (be it native or other weeds) under field conditions.

In sum, the comparatively better results using hyperspectral data were not unexpected, but our findings will suggest this method is preferred to achieve the best accuracy. Note however that the XGBoost method is limited by the much more expensive equipment necessary to capture this kind of data, thus likely putting it out of reach of the average consumer.

Machine learning techniques are increasingly being effectively applied to weed detection in row-cropping and pastoral environments (e.g., [12,44–48]). This present investigation further strengthens the latter body of work, especially for Parthenium weed in the field (see also [49–51]). No doubt natural environments can be protected if weed populations are reduced through proper weed management options [52]. Often proper weed management involves significant inputs of herbicides, but its use can be dramatically reduced if herbicide applications are optimized using SSWM techniques, including the use of aerial imagery to map weed distribution at varying scales that will assist in management optimization [53]. This paper provides a preliminary indication that the above machine learning methods can detect parthenium weed to a significant accuracy in pastoral environments and thus can be used for SSWM. This work reported herein also further strengthens the potential of using Yolov4 (including its more recent versions of Yolov7) and XGBoost for identification and classification of weed

Caveats

The data used for this study was developed in a controlled (glasshouse) environment to model pasture field environments as closely as possible. Thus, the methods have not been applied to real-world environments and so any generalization to real-world environments must be tentative and treated with caution. Parthenium weed was the

only weed investigated in this study and while efficient management of this weed alone would be of significant benefit, it cannot be said that the same methods could be applied successfully to other species of weeds (but see [54]). Problems were also encountered when trying to detect parthenium weed in its seedling stage using RGB images and during object detection; consequently, it is likely to be difficult to detect weeds of this size in real field environments. This same problem was somewhat encountered in the hyperspectral data, albeit to a lesser extent. This may be due, in part to the quality of data capture technology used, pre- and post-processing methods utilized, and training hardware employed, all of which could be improved further upon. Yolov4 was the only machine learning method used for each data type, and while this method answers the question posed for this study, our investigation does not determine that these methods are the only methods or the best methods. Several other machine learning algorithms, including Region-based CNN, single-shot-Detection (SSD), random forest (RF), support vector machine (SVM) and k-nearest neighbours (KNN), might need to be comparatively tested [55,56].

Data quality is consistently an issue for the task reported in this work. Previous attempts to train RGB detectors in the field have yielded very limited performance and detection accuracy (P. Trotter, Osunkoya OO, pers. Observation and Communication). These attempts involved capturing data in the field using a drone (UAV) and following pre-processing and training protocols similar to what was done in this investigation. The main difference between our reported work (in a control environment and which gave high accuracy (95%) and the previous ones (in pastoral fields, which gave lower accuracy of 35%) is the quality of the data. Previous attempts involved data with GSDs that are far greater than what was used herein, amplified and complicated by the fact that the previous work used data derived from the field which, by nature, is a more heterogeneous environment of fast changing atmospheric condition (especially solar illumination, cloud, and shadow), with notable variation in terrain (slope and aspect) and soil quality. Hence, field image data captured via UAV can be less reliable and more challenging for training [57,58]. This highlights the importance of data quality, specifically the GSD of the images.

Despite these field challenges, this study provides strong evidence that parthenium weed can potentially be detected in both RGB and hyperspectral imagery using machine learning methods in a real pastoral environment. Because the simulated environment used in this study was developed to imitate natural/human-made landscapes as closely as possible, our promising findings suggest that with further work and refinements, these methods might be applicable to field environments. Future work should also expand the list of target weeds to determine how broad and applicable our techniques and findings are. There are many more machine learning algorithms that exist other than the ones involved in this study (e.g., RF, SVN, and KNN), and so increasing the number of AI algorithms should be encouraged in future research to potentially provide a greater number of solutions and detection accuracy [55,56]. Lastly, with time better hardware and software technology will evolve in the AI environment and should be used to improve detection success.

5. Conclusions

The present work adds to the growing literature on machine learning towards optimising the discrimination, mapping, and monitoring of invasive alien plants [9,10,47,51,54,59]. We have shown that RGB images of Parthenium weed obtained in simulated pastoral environments can be processed successfully with a YOLOv4-CNN classification protocol-achieving an overall accuracy of 95% for detection, and 86% for differentiation of flowering and non-flowering stages of the weed. Using an XGBoost classifier for the pixel classification of the hyperspectral dataset, we achieved a classification accuracy of 99% for parthenium weed growth stages. Our ordination of the hyperspectral signatures suggested that the wavelengths responsible for major differentiation of parthenium weed from co-occurring control plants (grasses and sedges) are mainly in the blue (400–500 nm) and the infrared (>900 nm) ranges- a finding similar to that reported by Ullah et al. (2021) for the same weed in Pakistan [51]. It is worthwhile noting that ML discrimination of growth

stages of the focal weed has rarely been reported, but the work reported herein indicates that it is achievable.

For scalability of our findings to the field, improvements need to be made in use of UAV and deployment of high specification (e.g., HS/MS) cameras that can achieve low ground sampling distances (GSDs) while maintaining flight heights; this will allow a more efficient, high resolution spatial image capture [59]. A detector must also be trained on data that are captured in the field. The performance of this detector in terms of training and testing vastly depends on the capability of available computer system; thus, in future applications, computer systems specifically designed for object detection and AI should be used. Image and data augmentation techniques for generation of training data could be optimised as this may also likely improve performance. Better image annotation techniques, such as Semantic Segmentation [59], should be used because in natural fields, established Parthenium weed and co-occurring plants present a much more complicated scenarios that can make bounding box annotations inefficient and limit performance. Lastly, and overall, with further improvement in image processing (e.g., using Mask/faster R-CNN) and ML protocols (e.g., transitioning from YOLOv4 [the version available in the course of this work] to more recent YOLOv7 or even YOLO-R [You Only Learn One Representation]), the findings reported herein have the potential for extension and application in real-world pastoral and cropping environments to manage better many more weeds and revolutionize how people deal with the ever-growing problem of invasive alien species (see [60]).

References

1. Dhileepan, K.; Callander, J.; Shi, B.; Osunkoya, O.O. Biological control of parthenium (*Parthenium hysterophorus*): The Australian experience. *Biocontrol Sci. Technol.* **2018**, *28*, 970–988. [CrossRef]
2. Osunkoya, O.O.; Lock, C.; Buru, J.C.; Gray, B.; Calvert, M. Spatial extent of invasiveness and invasion stage categorization of established weeds of Queensland Australia. *Aust. J. Bot.* **2020**, *68*, 557–573. [CrossRef]
3. Dhileepan, K. Reproductive variation in naturally occurring populations of the weed *Parthenium hysterophorus* (Aster-aceae) in Australia. *Weed Sci.* **2012**, *60*, 571–576. [CrossRef]
4. Mao, R.; Nguyen, T.L.T.; Osunkoya, O.O.; Adkins, S.W. Spread pathways of the invasive weed *Parthenium hysterophorus* L.: The potential for water dispersal. *Austral. Ecol.* **2019**, *44*, 1111–1122. [CrossRef]
5. Osunkoya, O.O.; Lock, C.B.; Dhileepan, K.; Buru, J.C. Lag times and invasion dynamics of established and emerging weeds: Insights from herbarium records of Queensland, Australia. *Biol. Invasions* **2021**, *23*, 3383–3408. [CrossRef]
6. Adkins, S.; Shabbir, A. Biology, ecology and management of the invasive parthenium weed (*Parthenium hysterophorus* L.). *Pest Manag. Sci.* **2014**, *70*, 1023–1029. [CrossRef]

7. Aktar, W.; Sengupta, D.; Chowdhury, A. Impact of pesticides use in agriculture: Their benefits and hazards. *Interdisc. Toxicol.* **2009**, *2*, 1–12. [CrossRef]
8. Castaldi, F.; Pelosi, F.; Pascucci, S.; Casa, R. Assessing the potential of images from unmanned aerial vehicles (UAV) to support herbicide patch spraying in maize. *Precis. Agric.* **2017**, *18*, 76–94. [CrossRef]
9. Kiala, Z.; Odindi, J.; Mutanga, O. Determining the Capability of the Tree-Based Pipeline Optimization Tool (TPOT) in Mapping Parthenium Weed Using Multi-Date Sentinel-2 Image Data. *Remote Sens.* **2022**, *14*, 1687. [CrossRef]
10. Royimani, L.; Mutanga, O.; Odindi, J.; Zolo, K.S.; Sibanda, M.; Dube, T. Distribution of *Parthenium hysterophorus* L. with variation in rainfall using multi-year SPOT data and random forest classification. *Remote Sens. Appl. Soc. Environ.* **2018**, *13*, 215–223. [CrossRef]
11. Su, W.-H. Advanced Machine Learning in Point Spectroscopy, RGB- and Hyperspectral-Imaging for Automatic Discriminations of Crops and Weeds: A Review. *Smart Cities* **2020**, *3*, 767–792. [CrossRef]
12. Zhang, W.; Hansen, M.F.; Volonakis, T.N.; Smith, M.; Smith, L.; Wilson, J.; Ralston, G.; Broadbent, L.; Wright, G. Broad-Leaf Weed Detection in Pasture. In Proceedings of the 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), Chongqing, China, 27–29 June 2018; pp. 101–105. [CrossRef]
13. Shendryk, Y.; Rossiter-Rachor, N.A.; Setterfield, S.A.; Levick, S.R. Leveraging High-Resolution Satellite Imagery and Gradient Boosting for Invasive Weed Mapping. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4443–4450. [CrossRef]
14. Tkačik, G.; Garrigan, P.; Ratliff, C.; Milčinski, G.; Klein, J.M.; Seyfarth, L.H.; Sterling, P.; Brainard, D.H.; Balasubramanian, V. Natural Images from the Birthplace of the Human Eye. *PLoS ONE* **2011**, *6*, e20409. [CrossRef]
15. Park, W.; Pak, S.; Shim, H.; Le, H.A.N.; Im, M.; Chang, S.; Yu, J. Photometric transformation from RGB Bayer filter system to Johnson–Cousins BVR filter system. *Adv. Space Res.* **2016**, *57*, 509–518. [CrossRef]
16. Wiesner-Hanks, T.; Wu, H.; Stewart, E.; DeChant, C.; Kaczmar, N.; Lipson, H.; Gore, M.A.; Nelson, R.J. Millimeter-Level Plant Disease Detection From Aerial Photographs via Deep Learning and Crowdsourced Data. *Front. Plant Sci.* **2019**, *10*, 1550. [CrossRef]
17. Pathak, A.R.; Pandey, M.; Rautaray, S. Application of deep learning for object detection. *Procedia Comput. Sci.* **2018**, *132*, 1706–1717. [CrossRef]
18. Wu, X.; Sahoo, D.; Hoi, S.C.H. Recent advances in Deep Learning for Object Detection. *arXiv* **2019**, arXiv:1908.03673. [CrossRef]
19. Liu, L.; Ouyang, W.; Wang, X.; Fieguth, P.; Chen, J.; Liu, X.; Pietikäinen, M. Deep Learning for Generic Object Detection: A Survey. *Int. J. Comput. Vis.* **2020**, *128*, 261–318. [CrossRef]
20. Amigo, J.M.; Santos, C. Preprocessing of Hyperspectral and Multispectral Images. In *Data Handling in Science and Technology*; Amigo, J.M., Ed.; Elsevier: Amsterdam, The Netherlands, 2020; Volume 32, pp. 37–53. [CrossRef]
21. Gowen, A.A.; Feng, Y.; Gaston, E.; Valdramidis, V. Recent applications of hyperspectral imaging in microbiology. *Talanta* **2015**, *137*, 43–54. [CrossRef]
22. Shwetank; Neeraj; Jitendra; Vikesh; Jain, K. Pixel Based Supervised Classification of Hyperspectral Face Images for Face Recognition. *Procedia Comput. Sci.* **2018**, *132*, 706–717. [CrossRef]
23. Lin, T.Y.; Maire, M.; Belongie, S.; Bourdev, L.; Girshick, R.; Hays, J.; Perona, P.; Zitnick, C.L.; Dollár, P. Microsoft COCO: Common Objects in Context. In *Computer Vision—ECCV 2014*; Springer: Cham, Switzerland, 2014. [CrossRef]
24. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-J.M. YOLOv4 Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
25. Wang, C.-Y.; Liao, H.-Y.M.; Wu, Y.-H.; Chen, P.-Y.; Hsieh, J.-W.; Yeh, I.-H. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. *arXiv* **2019**, arXiv:1911.11929.
26. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. In *Computer Vision—ECCV*; Springer: Cham, Switzerland, 2014; Volume 8691, pp. 346–361. [CrossRef]
28. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path Aggregation Network for Instance Segmentation. *arXiv* **2018**, arXiv:1803.01534.
29. Jacques, S.L. Polarized Light Imaging of Biological Tissues. In *Handbook of Biomedical Optics*; Boas, D.C., Pitris, C., Ramanujam, N., Eds.; CRC Press: Boca Raton, FL, USA, 2016; pp. 649–669.
30. Klein, M.E.; Aalderink, B.J.; Padoan, R.; De Bruin, G.; Steemers, T.A. Quantitative Hyperspectral Reflectance Imaging. *Sensors* **2008**, *8*, 5576–5618. [CrossRef] [PubMed]
31. Ozdemir, A.; Polat, K. Deep Learning Applications for Hyperspectral Imaging: A Systematic Review. *J. Inst. Electron. Comput.* **2020**, *2*, 39–56. [CrossRef]
32. Audebert, N.; Le Saux, B.; Lefèvre, S. Deep Learning for Classification of Hyperspectral Data: A Comparative Review. *IEEE Geosci. Remote Sens. Lett.* **2019**, *7*, 159–173. [CrossRef]
33. Pathak, D.K.; Kalita, S.K.; Bhattacharya, D.K. Classification of Hyperspectral Image using Ensemble Learning methods: A comparative study. In Proceedings of the IEEE 17th India Council International Conference (INDICON), New Delhi, India, 10–13 December 2020; pp. 1–6. [CrossRef]
34. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]
35. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]

36. Osunkoya, O.O.; Akinsanmi, O.A.; Lim, L.S.A.; Perrett, C.; Callander, J.; Dhileepan, K. *Parthenium hysterophorus* L. (Aster-aceae) had limited impact on major soil nutrients and enzyme activity: Is the null effect real or reflects data insensitivity? *Plant Soil* **2017**, *420*, 177–194. [CrossRef]

37. Gambardella, C.; Parente, R.; Ciambrone, A.; Casbarra, M. A Principal Components Analysis-Based Method for the Detection of Cannabis Plants Using Representation Data by Remote Sensing. *Data* **2021**, *6*, 108. [CrossRef]

38. Huang, Y.; Li, J.; Yang, R.; Wang, F.; Li, Y.; Zhang, S.; Wan, F.; Qiao, X.; Qian, W. Hyperspectral Imaging for Identification of an Invasive Plant *Mikania micrantha* Kunth. *Front. Plant Sci.* **2021**, *12*, 787. [CrossRef] [PubMed]

39. Navie, S.C.; Panetta, F.D.; McFadyen, R.E.; Adkins, S.W. Germinable soil seed bank of central Queensland rangelands invaded by the exotic weed *Parthenium hysterophorus* L. *Weed Biol. Manag.* **2004**, *4*, 154–167. [CrossRef]

40. Regan, T.J.; Chades, I.; Possingham, H.P. Optimally managing under imperfect detection: A method for plant invasions. *J. Appl. Ecol.* **2011**, *48*, 76–85. [CrossRef]

41. Panetta, F.D. Weed eradication feasibility: Lessons of the 21st century. *Weed Res.* **2015**, *55*, 226–238. [CrossRef]

42. Pena-Barragan, J.M.; Lopez-Granados, F.; Jurado-Exposito, M.; Garcia-Torres, L. Spectral discrimination of *Ridolfia segetumand* sunflower as affected by phenologicalstage. *Weed Res.* **2006**, *46*, 10–21. [CrossRef]

43. Pfitzner, K.; Bartolo, R.; Whiteside, T.; Loewensteiner, D.; Esparon, A. Hyperspectral Monitoring of Non-Native Tropical Grasses over Phenological Seasons. *Remote Sens.* **2021**, *13*, 738. [CrossRef]

44. Louargant, M.; Jones, G.; Faroux, R.; Paoli, J.-N.; Maillot, T.; Gée, C.; Villette, S. Unsupervised Classification Algorithm for Early Weed Detection in Row-Crops by Combining Spatial and Spectral Information. *Remote Sens.* **2018**, *10*, 761. [CrossRef]

45. Revanasiddappa, B.; Arvind, C.; Swamy, S. Real-time early detection of weed plants in pulse crop field using drone with IoT. *Technology* **2020**, *16*, 1227–1242.

46. Wu, Z.; Chen, Y.; Zhao, B.; Kang, X.; Ding, Y. Review of Weed Detection Methods Based on Computer Vision. *Sensors* **2021**, *21*, 3647. [CrossRef]

47. Kiala, Z.; Mutanga, O.; Odindi, J.; Peerbhay, K.Y.; Slotow, R. Automated classification of a tropical landscape infested by Parthenium weed (*Parthenium hysterophorus*). *J. Remote Sens.* **2020**, *41*, 8497–8519. [CrossRef]

48. Olsen, A. Improving the Accuracy of Weed Species Detection for Robotic Weed Control in Complex Real-time Environments. Ph.D. Thesis, James Cook University, Douglas, QLD, Australia, 2020.

49. Kganyago, M.; Odindi, J.; Adjorlolo, C.; Mhangara, P. Selecting A Subset of Spectral Bands for Mapping Invasive Alien Plants: A Case of Discriminating *Parthenium hysterophorus* Using Field Spectroscopy Data. *Int. J. Remote Sens.* **2017**, *38*, 5608–5625. [CrossRef]

50. Kiala, Z.; Mutanga, O.; Odindi, J.; Peerbhay, K. Feature selection on sentinel-2 multispectral imagery for mapping a landscape infested by parthenium weed. *Remote Sens.* **2019**, *11*, 1892. [CrossRef]

51. Ullah, S.; Shakir, M.; Iqbal, M.S.; Iqbal, A.; Ali, M.; Shafique, M.; Rehman, A.; Godwin, J. Identifying optimal waveband po-sitions for discriminating *Parthenium hysterophorus* using hyperspectral data. *Ecol. Inform.* **2021**, *64*, 101362. [CrossRef]

52. Song, Y.; Fiaz, M.; Kim, D.W.; Kim, J.; Kwon, C.H. Increasing forage yield and effective weed control of corn-soybean mixed forage for livestock through using by different herbicides. *J. Anim. Sci. Technol.* **2019**, *61*, 185–191. [CrossRef]

53. Weis, M.; Keller, M.; Rueda Ayala, V. Herbicide reduction methods. In *Herbicides—Environmental Impact Studies and Management Approaches*; Alvarez-Fernandez, R., Ed.; Intech: Rijeka, Croatia, 2012.

54. Wang, Q.; Cheng, M.; Huang, S.; Cai, Z.; Zhang, J.; Yuan, H. A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed *Solanum rostratum Dunal seedlings*. *Comput. Electron. Agric.* **2022**, *199*, 107194. [CrossRef]

55. Islam, N.; Rashid, M.; Wibowo, S.; Xu, C.Y.; Morshed, A.; Wasimi, S.A.; Moore, S.; Rahman, S.M. Early Weed Detection Using Image Processing and Machine Learning Techniques in an Australian Chilli Farm. *Agriculture* **2021**, *11*, 387. [CrossRef]

56. Zhang, R.; Wang, C.; Hu, X.; Liu, Y.; Chen, S.; Su, B. Weed location and recognition based on uav imaging and deep learning. *Int. J. Precis. Agric. Aviat.* **2020**, *3*, 23–29. [CrossRef]

57. Adão, T.; Hruska, J.; Padua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sens.* **2017**, *9*, 1110. [CrossRef]

58. Sarić, R.; Nguyen, V.D.; Burge, T.; Berkowitz, O.; Trtílek, M.; Whelan, J.; Lewsey, M.G.; Čustović, E. Applications of hyperspectral imaging in plant phenotyping. *Trends Plant Sci.* **2022**, *27*, 301–315. [CrossRef]

59. Sa, I.; Popović, M.; Khanna, R.; Chen, Z.; Lottes, P.; Liebisch, F.; Nieto, J.; Stachniss, C.; Walter, A.; Siegwart, R. Weed-map: A large-scale semantic weed mapping framework using aerial multispectral imaging and deep neural network for precision farming. *Remote Sens.* **2018**, *10*, 1423. [CrossRef]

60. Ambrozowicz, M. Weed Detection Using Machine Learning a Game Changer in Agriculture. Available online: https://picterra.ch/blog/weed-detection-using-machine-learning/ (accessed on 14 October 2020).