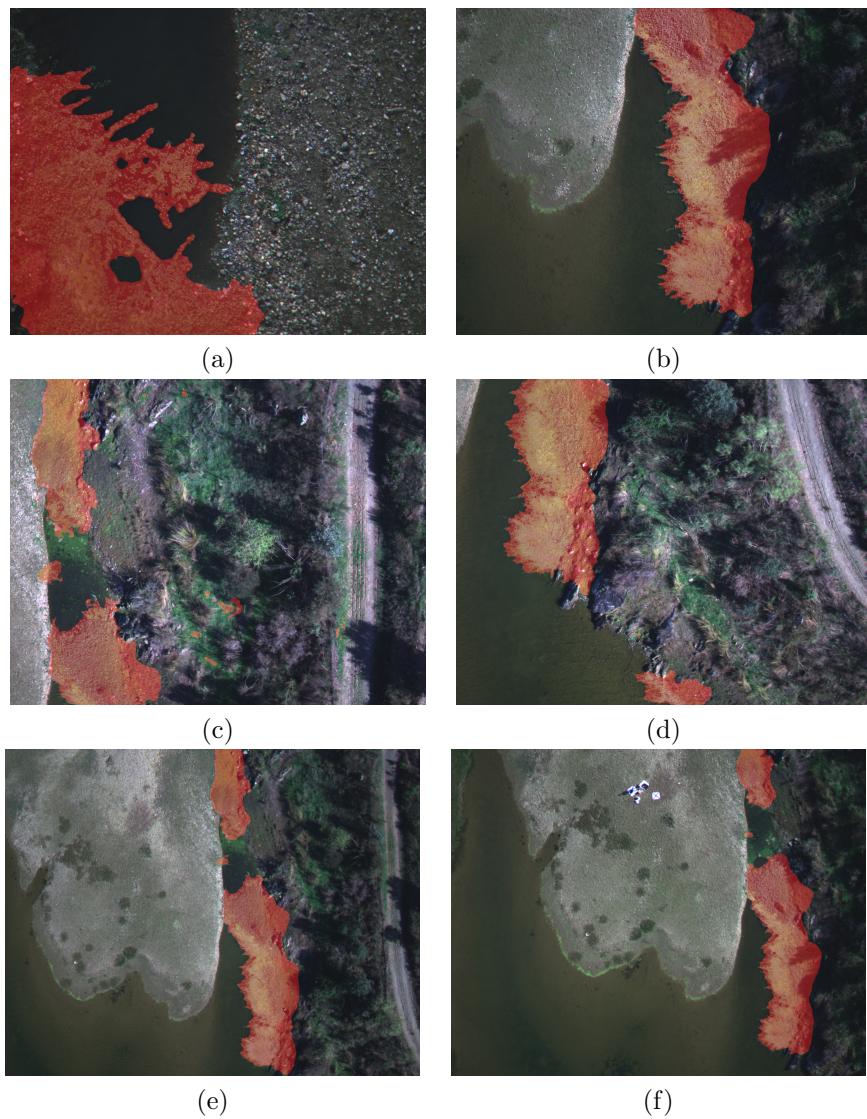


Graphical Abstract

Remote Detection of Exotic Invasive Aquatic Floral Species using Drone-Mounted Multispectral Data

António J. Abreu, Luís A. Alexandre, João A. Santos



Highlights

Remote Detection of Exotic Invasive Aquatic Floral Species using Drone-Mounted Multispectral Data

António J. Abreu, Luís A. Alexandre, João A. Santos

- Creation of a new data set composed of multispectral images focused on Ludwigia p.;
- Use of a drone-mounted multispectral sensor;
- Making the data set publicly available for other researchers;
- Proposal of a new method for detecting Ludwigia p. in multispectral images.

Remote Detection of Exotic Invasive Aquatic Floral Species using Drone-Mounted Multispectral Data

António J. Abreu^{a,b,*}, Luís A. Alexandre^a, João A. Santos^b

^a*NOVA LINCS, Universidade da Beira Interior, R. Marquês de Ávila e Bolama, Covilhã, 6201-001, Portugal*

^b*Zirak srl – Information Technology, Via S. Agostino 6/A, Mondovì (CN), 12084, Italy*

Abstract

Remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance. It is being broadly used to monitor ecosystems, mainly for their preservation. Ever-growing reports of invasive species have affected the natural balance of ecosystems. Exotic invasive species have a critical impact when introduced into new ecosystems and may lead to the extinction of native species. In this study, we focus on Ludwigia peploides, considered by the European Union as an aquatic invasive species. Its presence can negatively impact the surrounding ecosystem and human activities such as agriculture, fishing, and navigation. Our goal was to develop a method to identify the presence of the species. We used images collected by a drone-mounted multispectral sensor to achieve this, creating our LudVision data set. To identify the targeted species on the collected images, we propose a new method for detecting Ludwigia p. in multispectral images. The method is based on existing state-of-the-art semantic segmentation methods modified to handle multispectral data. The proposed method achieved a producer's accuracy of 79.9% and a user's accuracy of 95.5%.

Keywords: Remote Sensing, Drone-Mounted Multispectral Sensor, Invasive Alien Species, Ludwigia peploides, Semantic Segmentation

PACS: 07.05.Mh, 42.68.Wt

2000 MSC: 62M45, 62P12, 68T10

*Corresponding author

Email address: antonio.abreu@ubi.pt (António J. Abreu)

1. Introduction

Exotic invasive species have a critical impact when introduced into new ecosystems, and there is a growing global concern due to their negative ecological and economic ramifications. The introduction of invasive plants often results in the extinction of native plants and a reduction in biodiversity, either by outcompeting or hybridizing with native species (1). This is especially the case in aquatic environments and for aquatic plants (2). Under most scenarios, invasive plants arrive without their co-evolved competitors or parasites, allowing them to spread rapidly, replacing native plants without assuming their ecological roles.

Invasives in wetlands have many adverse effects within and across trophic levels and significantly reduce biodiversity (3; 4). Many invasives may directly compete with other species by secreting allelopathic chemicals that reduce germination and seedling survival or by changing light accessibility (5; 3; 6). Invasives may also significantly impact invertebrate distribution, diversity, and abundance; induce anoxic conditions detrimental to fish and other aquatic life (5; 7); and act as barriers to fish movement (6; 4). They also reduce open water habitats for water birds and other wildlife (4).

RS image analysis is increasingly being used to map invasive plant species. The resulting distribution maps can be used to target the management of early infestations and to model future invasion risks (8; 1). Remote identification of invasive plants based on differences in spectral signatures is the most common approach, typically using hyperspectral data (9). Advances in Unmanned Aerial Vehicle (UAV) and sensor miniaturization are enabling higher spatial resolution species mapping, which is promising for early detection of invasions before they spread over larger areas (10).

Our goal is to monitor the spread of *Ludwigia peploides*, specifically in the Reservoir of the Toulica Dam (Zebreira, Portugal), located in the hydrographic basin of the Aravil river, a tributary of the Tagus, where it was detected in 2020. *Ludwigia peploides* is a species natural to South America that invades rivers, ponds, and rice fields. It can grow in deep waters, as a fully or partially submerged plant, and form floating mantles. When this happens, it prevents the entry of light affecting submerged species and blocking the water lines, affecting navigation, fishing, and recreational use. It competes for space by eliminating native species and producing substances that inhibit the germination and growth of other species. It reproduces vegetatively through stem fragmentation but also seeds. It is an invasive species

that raises concern at the European level (EU Reg. 1143 (11)) and is included in the Portuguese legislation list of invasive species (DL 92/2019 of 10/07 (12)). Although our focus was on this particular location, the developed detection method is general and can be used with data from other locations.

As such, we aimed to develop a system for the remote detection of the *Ludwigia peploides*. To do so, we first created the LudVision data set. It contains various images of *Ludwigia peploides* taken at several altitudes, times of day, and atmospheric conditions. This was to ensure that our data set was as representative as possible. We also propose a new method for detecting *Ludwigia p.* in multispectral images. The method is based on existing state-of-the-art semantic segmentation methods that were modified to handle multispectral data.

Since remote sensing applications typically rely on specialized and expensive equipment, we set on using a general application sensor and platform instead of the highly specialized and expensive equipment traditionally used in RS applications. We wanted to prove that using semantic segmentation models could create high-reliability maps with available application equipment. Our LudVision data set will be provided for public use.

2. Related Work

Over the years, remote sensing has received much attention, especially for the detection of invasive species (13; 1; 14; 15; 16; 17; 18; 19). Specifically, there has been an increasing interest in detecting and monitoring aquatic invasive species (10; 16; 17; 7) due to the severe negative impacts they can have on ecosystems. One prime example of invasive aquatic species are water hyacinths and ludwigia (10; 7; 16; 17). Due to improvements in sensor technology and the availability of UAVs, the use of unmanned aircraft has gained a lot of attention (10; 20; 18; 19). The low costs of acquisition and operation, paired with the ability to be deployed almost instantly without significant planning, have made them a viable alternative compared to more expensive data like Manned Aerial Vehicles or satellites. They are beneficial when high resolution and frequent acquisition is a requirement. The only major downside of drone-mounted sensors is the limited capability to cover broad areas. One of the main ways species can be detected and differentiated is by performing spectroscopy analysis (14; 13; 1; 17; 15). Image spectroscopy relies on the fact that each species has a unique spectral signature which can be

used to identify a species. Multispectral or preferably hyperspectral imagery is needed to extract the spectral signature of a species.

There has been vast interest and research done regarding semantic segmentation, with multiple interpretations and adaptations. Given that our problem needs to be able to keep high-resolution representations, we analyzed a few subsets of models specifically designed for high-resolution representations (21; 22; 23; 24; 25; 26; 27). The concepts applied in the studied models were used to develop our proposed model.

Despite the extensive research and work performed in both remote sensing and semantic segmentation, we could not identify any work done that applied semantic segmentation models to remote sensing. Thus we propose a new semantic segmentation model as the classifier to detect *Ludwigia peploides*.

3. Methods

3.1. Study Site and Targeted Species

The study site is the Reservoir of the Toulica Dam (Zebreira, Portugal), located in the hydrographic basin of the Aravil river, a tributary of the Tagus. In 2020 the IAS *Ludwigia peploides* was spotted in the northeast part of the reservoir and had since spread, forming three big mantles. Recently it started spreading and forming small patches southwest of the initial infestation site. *Ludwigia peploides* is a species natural to South America that invades rivers, ponds, and rice fields. It can grow in deep waters, as a fully or partially submerged plant, and form floating mantles. When this happens, it prevents the entry of light affecting submerged species and blocking the water lines, affecting navigation, fishing, and recreational use. It competes for space by eliminating native species and producing substances that inhibit the germination and growth of other species. It reproduces vegetatively through stem fragmentation but also seeds. The species' stems can grow between 10 cm and 3 m, hence its ability to form large mantles. The leaves are a bright green (that most of the time stands out from the rest of the present vegetation) and can have a lanceolate or oval shape. They measure between 2.5 and 3.8 centimetres, and both the stem and the leaves have different trichomes distributed over the surface. *Ludwigia peploides* also have solitary flowers with yellow petals, which measure from one to 1.5 cm in length, and which develop from tassels emerging from the upper part of the axillary bud. The species blooming period occurs from mid-spring to early fall, and the plant is easily identifiable during this period. This is also the period where the

species grows the most.

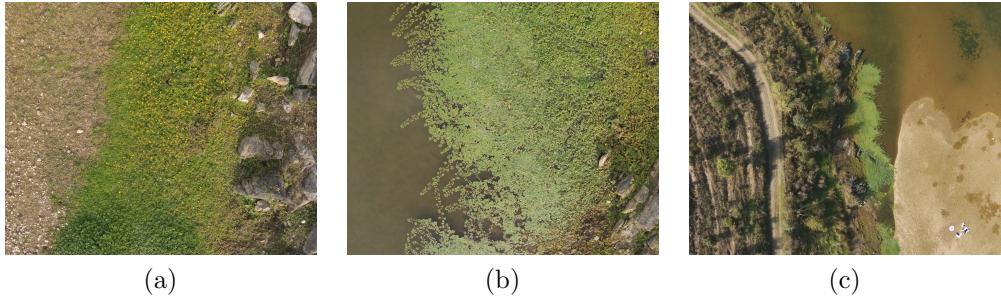


Figure 1: Images from the *Ludwigia peploides* collected at the Reservoir of the Toulica Dam at different altitudes. (a) 15 m, (b) 40 m, and (c) 70 m.

3.2. Sensor and Platform

Given that our study site is small and the infestation is still at its early stages, we do not need to cover large extents, but we need to have a good resolution. We initially trialled the use of satellite images. However, due to the infestation’s low resolution and small coverage area, it was not possible to identify *Ludwigia p.* on the images. Thus, we decided on using a drone-mounted sensor to build our data set. We used the DJI P4 Multispectral. It is a drone based on the DJI Phantom 4, but the P4 has six sensors instead of a traditional RGB camera. One RGB for visible light and five monochrome sensors for multispectral imaging. Table 1 contains the most relevant drone and camera specifications.

3.3. Data Collection

To collect the data for our LudVision data set, we visited the study site twice (October 11th and 20th, 2021). We captured all our data in hover mode because the drone is more stable, allowing us to point the camera straight down.

The data was collected at different altitudes, ranging from 10 m to 70 m. The goal of collecting data from different altitudes is that our model will learn features from a wide range of spatial resolutions. Thus, the final model should be more resilient to altitude variations. The data was also taken at various times to ensure that the model is resilient to variations in solar reflection. Note that no data was collected at solar noon, as it would result

Table 1: DJI P4 Multispectral specifications.

Aircraft	
Hover Accuracy Range	RTK enabled and functioning properly: Vertical: ± 0.1 m; Horizontal: ± 0.1 m RTK disabled: Vertical: ± 0.1 m (with vision positioning); ± 0.5 m (with GNSS positioning) Horizontal: ± 0.3 m (with vision positioning); ± 1.5 m (with GNSS positioning)
Camera	
Sensors	Six 1/2.9" CMOS, including one RGB sensor for visible light imaging and five monochrome sensors for multispectral imaging. Each Sensor: Effective pixels 2.08 MP (2.12 MP in total)
Filters	Blue (B): 450 nm ± 16 nm; Green (G): 560 nm ± 16 nm; Red (R): 650 nm ± 16 nm; Red edge (RE): 730 nm ± 16 nm; Near-infrared (NIR): 840 nm ± 26 nm
Max Image Size	1600 \times 1300 (4:3.25)

in overexposed images due to the light being reflected from the lake's surface. Table 2, contains more detailed information about the collected data.

Table 2: Altitude, time and number of images collected.

Altitude	Time	Number of images
10 m	11h - 12:45h	435
	15:30h - 17h	
15 m	11h - 12h	365
40 m	10h - 12:45h	135
70 m	11h - 12h	27

Every time the drone collects data, it is effectively taking six images, one

RGB image in *.jpg* format and five monochromatic images in *.tif* format, one for each band, as can be seen in figure 2. The RGB image should only be used as a reference as it helps to visualize the scene. The remaining five images are monochromatic and used to train the models.

The included 'sunlight sensor' is a feature of the DJI P4 Multispectral that allows the drone to correct the images according to the solar exposure automatically. Usually, this step needs to be done in the pre-processing phase with specialized software, and before each flight, the sensor would need to be calibrated.

3.4. Data Pre-processing

Given that the drone already performs corrections to the images based on the light conditions, the only pre-possessing left is to merge the individual bands into a single image. This way, we end up with a five-band image that can be fed to the model for training.

Merging the individual bands is not as straightforward as simply stacking the five bands and exporting them as a final image. The sensors are positioned in a 3×2 array, meaning they all have a unique perspective. Furthermore, we must account for lens imperfections and distortions due to manufacturing tolerances. This way, if we just stacked the bands, the final image would be distorted. Figure 3 contains an example of bands stacked before alignment (the image is distorted and looks overexposed) and bands stacked after alignment (the image is perfectly aligned without distortions).

The alignment was performed using the homography tools from *OpenCV* in *Python*. Once all bands are aligned, they are stacked, and the final image is trimmed to 1400×1100 pixels and exported as a TIFF file.

3.5. Annotation Process

We use the *Hasty.ai* tool for the annotation process. *Hasty.ai* is an excellent tool for annotating data with organic and complex shapes, which is precisely our case. It has Artificial Intelligence enabled tools that allow for assisted annotation, speeding up the entire process. This tool allowed us to create the ground truth for our dataset, which was used to train and validate the proposed model.

3.6. Testing for Spectral Radiance

After completing our data's pre-processing and annotation steps, we assessed the spectral radiance. Our data was captured to allow us to leverage

photophysiological measurements. Thus, we measured the radiance for Ludwigia p., rock, surrounding vegetation, and water. As expected, Ludwigia p. has a unique and distinct spectral signature, especially in the non-visible bands (Red Edge and Near Infra-red), as shown in figure 4. This supports our initial argument that a multispectral sensor would be enough to capture significant differences in the spectral radiance, allowing for the detection of Ludwigia peploides.

4. A New Method for Detection of Ludwigia in Multispectral Images

We analyzed both preliminary concepts regarding RS and state-of-the-art semantic segmentation methods and related work carried out by other authors. After completing our research, we concluded that we should take a "hybrid" approach to our problem by combining both RS concepts and semantic segmentation. The reasoning behind our approach is that despite having a multi-spectral data set, the number of bands in our data set is minute when compared to other studies. We only have five bands, whereas other authors had access to hyperspectral data, with considerably more bands. Thus, we believe that relying solely on the information of only five bands will not be enough to use models like the ones used by other authors (10; 17).

We propose the use of state-of-the-art semantic segmentation models that will be modified to take our data as input. By using semantic segmentation models, we can leverage both the capabilities of these models to recognize objects and the multi-spectral nature of our data. Fundamentally, the model will have the same behaviour as usual but have access to the information of two additional bands, which should help it find meaningful features to detect the presence of Ludwigia peploides.

We started by performing a series of tests on an existing semantic segmentation model HRNet (24) model in its HRNet + Object Context Recognition (OCR) (23) implementation. The architecture of the model is represented in figure 5 and its respective fusion model is represented in figure 6. A few modifications had to be made to train the model on our data, especially on the input layer. Given that the model was originally designed to use traditional RGB imagery, we had to modify the input layer to be able to accommodate our data which has five bands (RGB plus a Near InfraRed and a RedEdge

band). The results of these initial tests were used as a baseline to validate our new proposed model.

We chose this model for two reasons: First, the model can extract high-resolution representations, which are needed for position-sensitive tasks, like semantic segmentation. The high-resolution representations are attained by connecting high-to-low resolution convolution streams in parallel rather than in series. Then, the representations are fused by a custom fusion module, which exchanges information across multi-resolution representations. This means that lower-resolution representations also contain information obtained from the high-resolution representations. This allows the model to extract better features for the images at lower resolutions. Given that images taken at high altitudes tend to have low resolution, it makes this model fitting for our data; Second, the mentioned implementation integrates the OCR module (24). This context module uses a set of pixels lying in the object instead of a set of surrounding sparsely sampled pixels. This is achieved by differentiating the same-object-class contextual pixels from the different-object-class contextual pixels and structuring the contextual pixels into object regions to exploit the relations between pixels and object regions. This allows setting the context for each pixel more accurately. Because *Ludwigia p.* typically forms large mantles at the water's surface, the module will have a larger area to extract the context for a given pixel. Our data set only has two labels (*Ludwigia p.* and background), making it easier to differentiate same-object-class contextual pixels from the different-object-class contextual.

After establishing a baseline with HRNet (24), we focused on ways we could improve and adapt the model to our data/problem. One of our primary goals is to detect *Ludwigia p.* at higher altitudes (ideally on satellite imagery). Thus, we tried to find ways to simulate the appearance of satellite data using our data set. When compared to our data, satellite imagery has two distinct differences: it usually has a lower resolution and covers a broader area. So, we had to find ways to both lower our data's resolution and, if possible, enlarge the perceived Field of View (FoV) of the convolution layers.

To achieve the down-sampling of our data, we increased the *stride* value on the fusion models for all connections made between *stage 1* with the remaining *stages*. This effectively reduces the perceived resolution of our images, making them more closely resemble higher altitude images. To enlarge the perceived FoV, we decided to use *atrous* convolutions (also known as dilated convolutions). This type of convolution allows enlarging the con-

volution's FoV by increasing the *dilation rate* value. If the *dilation rate* is set to zero, then we effectively have a "conventional" convolution layer. An added benefit of using dilated convolutions is that they save computation and memory costs, as increasing the value of the *dilation rate* allows for a larger receptive field which means viewing more data points.

After implementing the modifications mentioned above, the model is expected to extract low-resolution features, resulting in better performance, especially in data captured at higher altitudes. Its also expected that the model to be more efficient due to the use of dilated convolutions and the down-sampling process.

We performed the following changes to the model:

- We kept the same number of stages, as we did not see substantial benefits in either increasing or decreasing their number;
- We increased the *stride* value on the fusion module from 2 to 3 for all connections made between *stage* 1 with the remaining *stages*. All remaining connections kept the *stride* value at 2.
- We replaced the conventional convolution layers from *stages* 2 and 3, with dilated convolutions, with *dilation* = 2 and *padding* = 2. The convolution layers in *stages* 1 and 4 where kept unchanged.

The architecture of the resulting model is represented in figure 7.

5. Results

A series of tests were made to validate our model's performance and compare it to the baseline results. To evaluate our model, we used the following metrics:

- **Producer's accuracy:** is the map accuracy from the point of view of the mapmaker (the producer). This is, how often the real features on the ground are correctly shown on the classified map. It is also the number of reference sites classified accurately, divided by the total number of reference sites for that class.
- **User's accuracy:** is the accuracy from the point of view of a map user. The user's accuracy essentially tells how often the class on the map will be present on the ground. This is referred to as reliability.

Table 3: Results obtained with HRNet and our proposed model.

	HRNet	Our model
Producer’s Acc (10/15 m)	0.877	0.959
Producer’s Acc (40 m)	0.767	0.899
Producer’s Acc (70 m)	0.768	0.799
User’s Acc (10/15 m)	0.925	0.850
User’s Acc (40 m)	0.861	0.916
User’s Acc (70 m)	0.850	0.955
Class IoU (10/15 m)	0.835	0.820
Class IoU (40 m)	0.749	0.830
Class IoU (70 m)	0.720	0.769

- **IoU:** is a metric that determines the extent of overlap between two areas. In the case of semantic segmentation, it is used to determine the overlap between the ground truth and the model’s predictions. This metric ranges from 0 to 1 (or 0 to 100%), where 1 means a perfect overlap. We can also calculate the class and mean Intersection over Union (IoU) for multi-class segmentation. As the names suggest, class IoU is calculated by evaluating the overlap for each class, and the mean IoU is the mean value of all classes. Given that we have two classes (*ludwigia* and *background*, we will use class IoU).

Separate training and validation data sets were used. Table 3 contains the results of the best configurations on both HRNet (24) and our proposed model. As can be seen, our model has improved performance compared to HRNet, especially at higher altitudes. Furthermore, we verified that the training times of the model were cut in half. The total training time dropped from 130 h to 57 h (training times of the baseline experiments on HRNet and running the same experiments for validating our model).

If we look closely at the results, we can make the following conclusions. The producer’s accuracy has seen an overall increase across all attitudes. Given that producer’s accuracy refers to how often the actual features on the ground are correctly shown on the classified map, it means that the classification is more reliable. Thus, when our model classifies a pixel as being Ludwicia p., it is more likely that the prediction is accurate compared to HRNet.

As for the user’s accuracy, it decreased for lower altitude images (10/15

m), but it increased for higher altitudes (40 and 70 m). Given that our goal is to classify images at higher altitudes (because it allows us to cover a larger area and possibly upscale our model to satellite images), we consider this a good result. It is essential for us that the model performs better at higher altitudes, even if it means losing some performance at lower altitude classification.

Class IoU, follows the same trend as the user's accuracy. Our model has slightly lower performance at low altitude images, but it outperforms HRNet at higher altitudes.

6. Discussion

6.1. Model Validation

Due to the modifications, we made to the original HRNet, our new model is better fitted for our application. This means that there is no need to create a custom model from scratch. It is possible to create a new semantic segmentation method by modifying existing models so that the new model is better suited for a given application. This allows reducing the time needed to develop said, custom model.

Analyzing all the metrics in general, we conclude that it outperforms the HRNet model, which already demonstrated excellent results in correctly classifying *Ludwigia peploides*. Our model's gain in performance is especially noticeable in high-altitude images, which is a desirable outcome. Our proposed model is able to detect the presence of the targeted species with great accuracy.

Furthermore, we verified that the training times of the model were cut in half. This not only means that our model is more capable of identifying *Ludwigia p.*, but it is also faster to deploy due to the shorter training time.

By looking at both the metrics and model output, we can conclude that our model produces high-reliability maps and has the tendency to output more false positives than false negatives. It is better to have a model that occasionally identifies other species as *Ludwigia p.* than to have a model that sometimes fails to identify the IAS. This is crucial because in real-world scenarios, not identifying a single strand of *Ludwigia p.* may lead to identifying a site falsely as not being infested. Due to the capability of the IAS spreading rapidly from one single seed, this can lead to a severe infestation in the future. *Ludwigia p.* reproduces aggressively, and one misclassified plant can lead to a wholly covered body of water in a brief period.

6.2. Analyzing the Model’s Output

After analyzing and comparing quantitative results, we shifted our focus to analyzing the actual outputs of the model. Although evaluating the model using quantitative results is imperative, a visual analysis of the predictions allows us to understand the scenarios where the model is and is not accurate.

Considering the quantitative results and our observations, we further conclude that the model performs well. As seen in figures 9 the model’s predictions are very close to the respective ground truths. Figure 10 presents the overlap of the model’s predictions on the respective images, where it is clear that the model is able to correctly and reliably identify the targeted species at multiple altitudes.

6.3. Comparing Our Results to Similar Studies

To further validate our model, we compared its performance with similar studies. We compare our model’s performance to the performance achieved by Bolch in (10). His work, setup, and objectives are very similar to ours, and he also assesses the performance of water primrose (*Ludwigia spp*), which is very similar to *Ludwigia p.*). He achieved a producer’s accuracy of 100% on water primrose and a user’s accuracy of just 50%. Our model at 70 m (the highest altitude on our data set), achieved a producer’s accuracy of 79.9% (a 20.1% decrease compared to (10)), and a user’s accuracy of 95.5% (an increase of 45.5%). We have to keep in mind that the sensor used in (10) is superior to ours and has more bands (their sensor has 270 bands, while our sensor only has five bands). This further proves not only our method’s validity but also that great classification results can be achieved using data sets built with data captured from general application sensors.

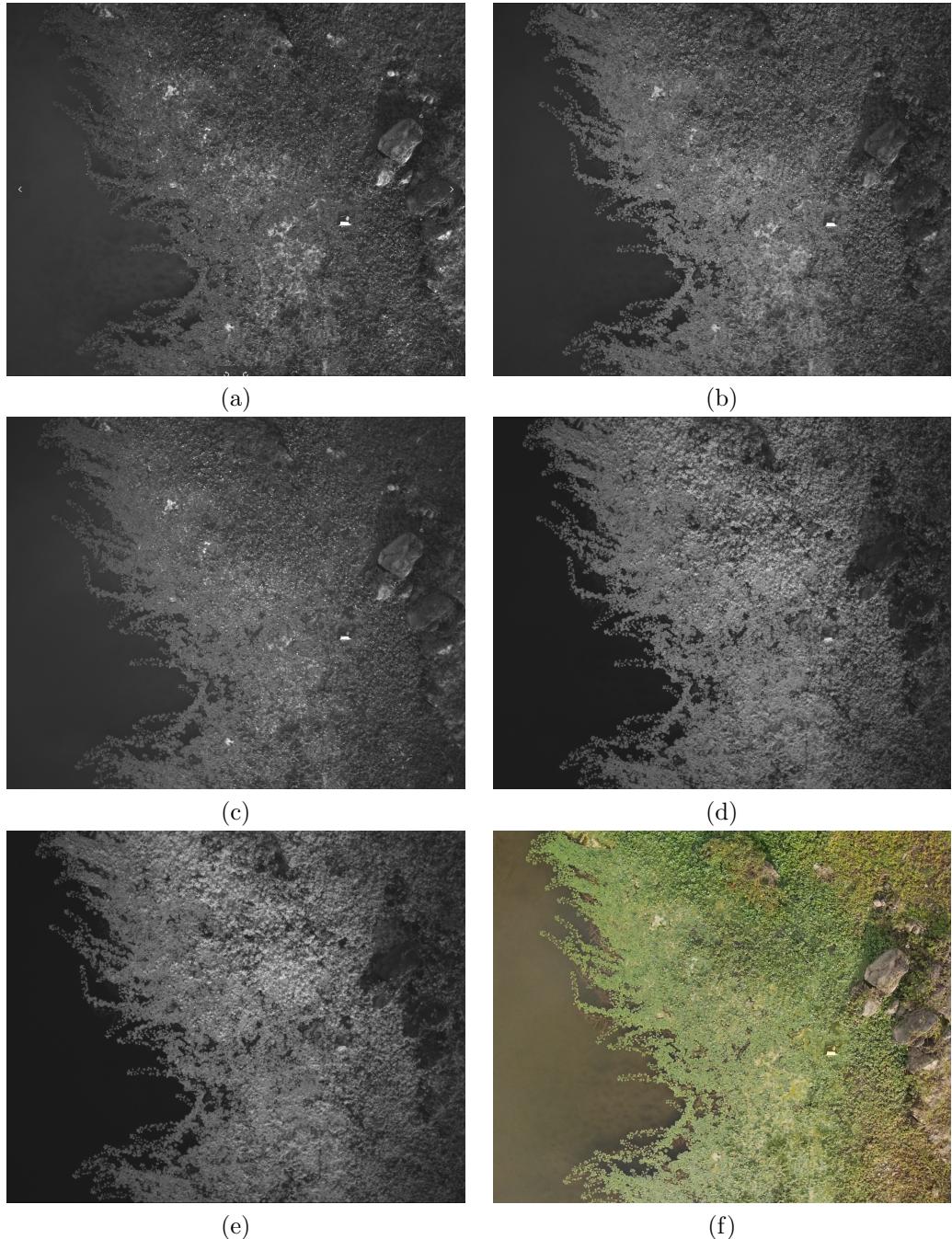


Figure 2: Example of a red band (a), green band (b), blue band (c), red edge band (d), near infra-red band (e), and RGB image (f) collected by our drone.



Figure 3: Example of an image created by stacking bands before alignment (a) and stacking bands after alignment (b).

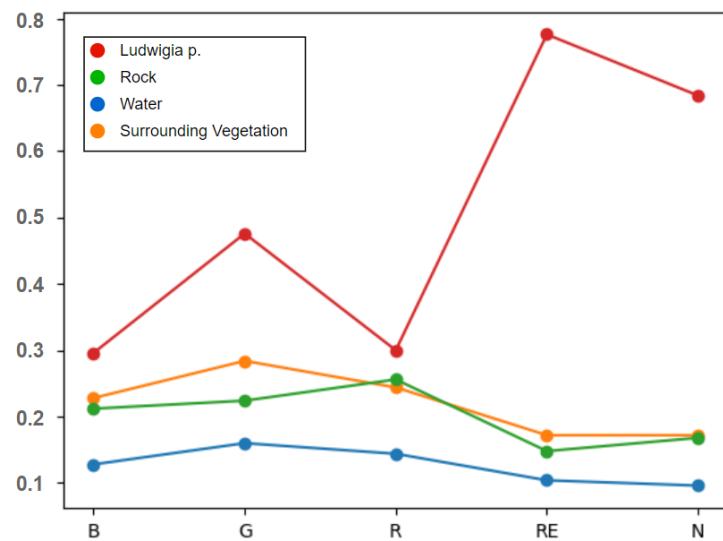


Figure 4: Reflectance values (in %) for each band corresponding to Ludwigia p., water, surrounding vegetation, and rock.

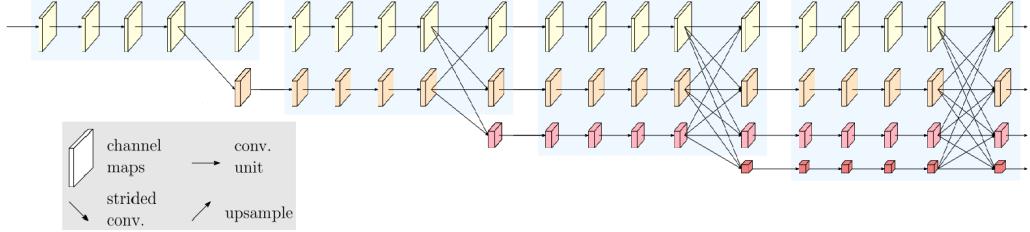


Figure 5: An example of a high-resolution network. Only the main body is illustrated, and the stem (two stride-2 3×3 convolutions) is not included. There are four stages. The 1st stage consists of high-resolution convolutions. The 2nd (3rd, 4th) stage repeats two-resolution (three-resolution, four-resolution) blocks (24).

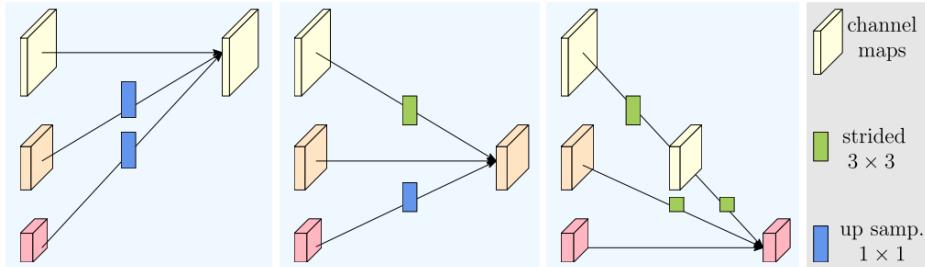


Figure 6: Representation of how the fusion module aggregates the information for high, medium and low-resolutions from left to right, respectively (24).

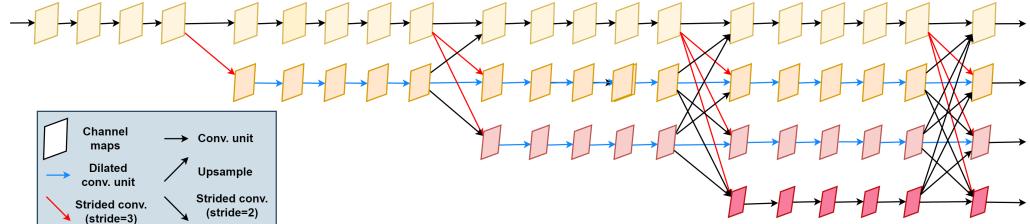


Figure 7: Architecture of our model, that is based on HRNet (24).

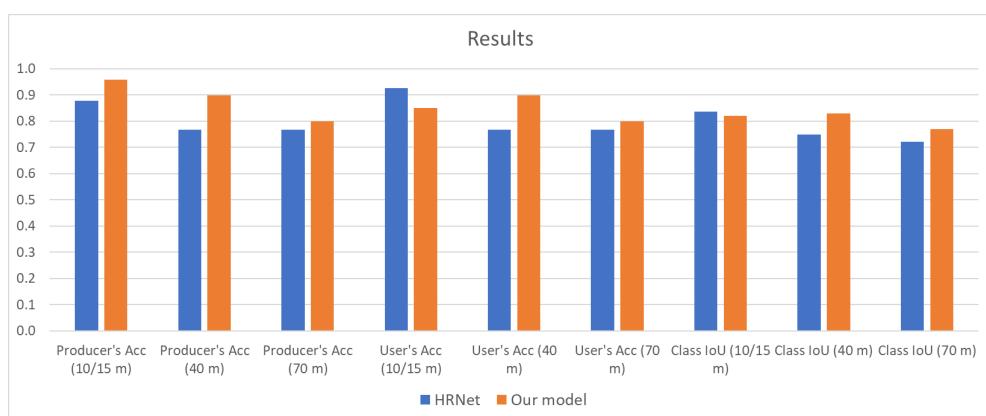


Figure 8: Performance comparison between HRNet and our model

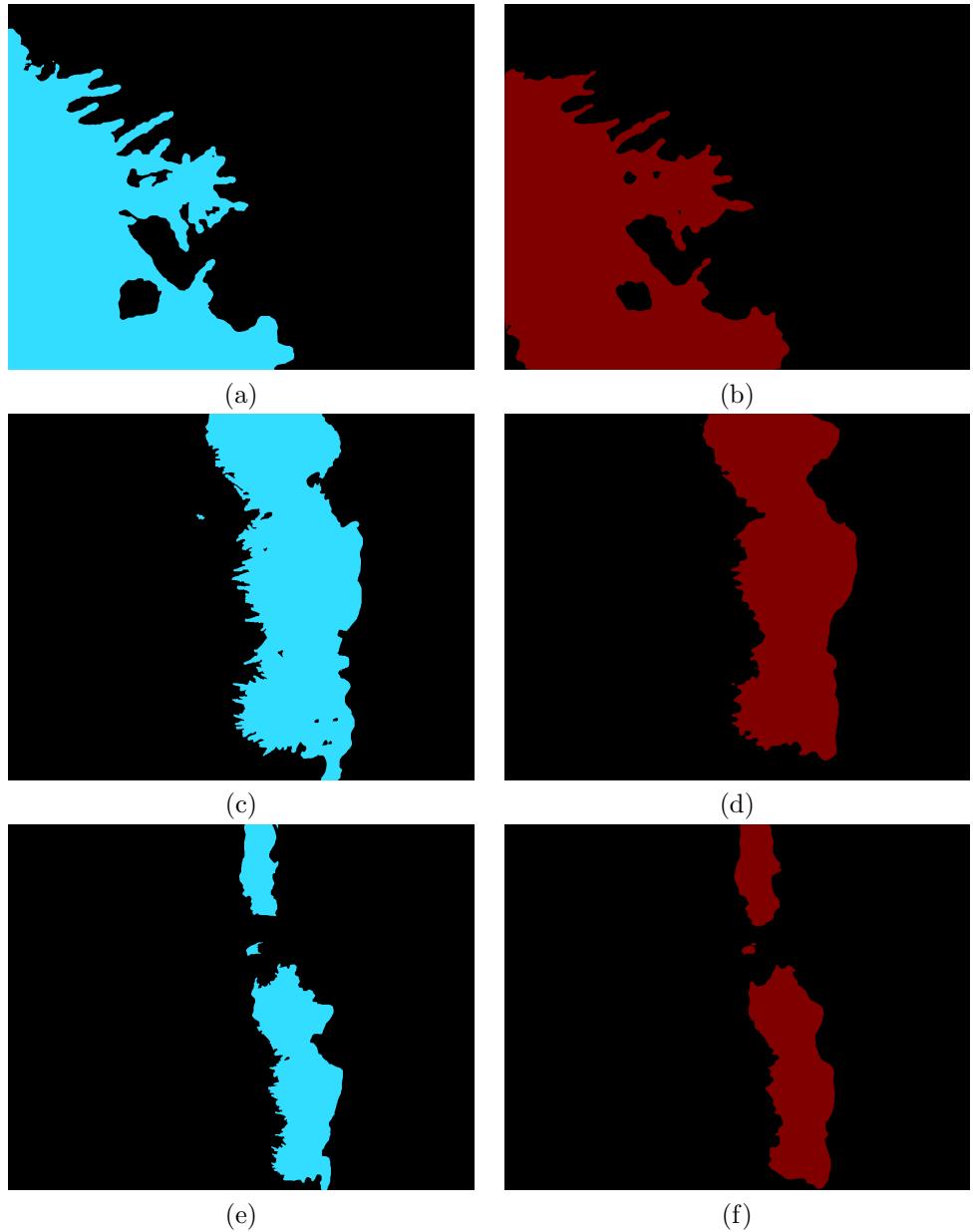


Figure 9: Ground truths and their respective model predictions. Images (a) and (b) are an example of the ground-truth and model predictions corresponding to images taken at 10/15 m, (c) and (d) at 40 m, and (e) and (f) at 70 m.

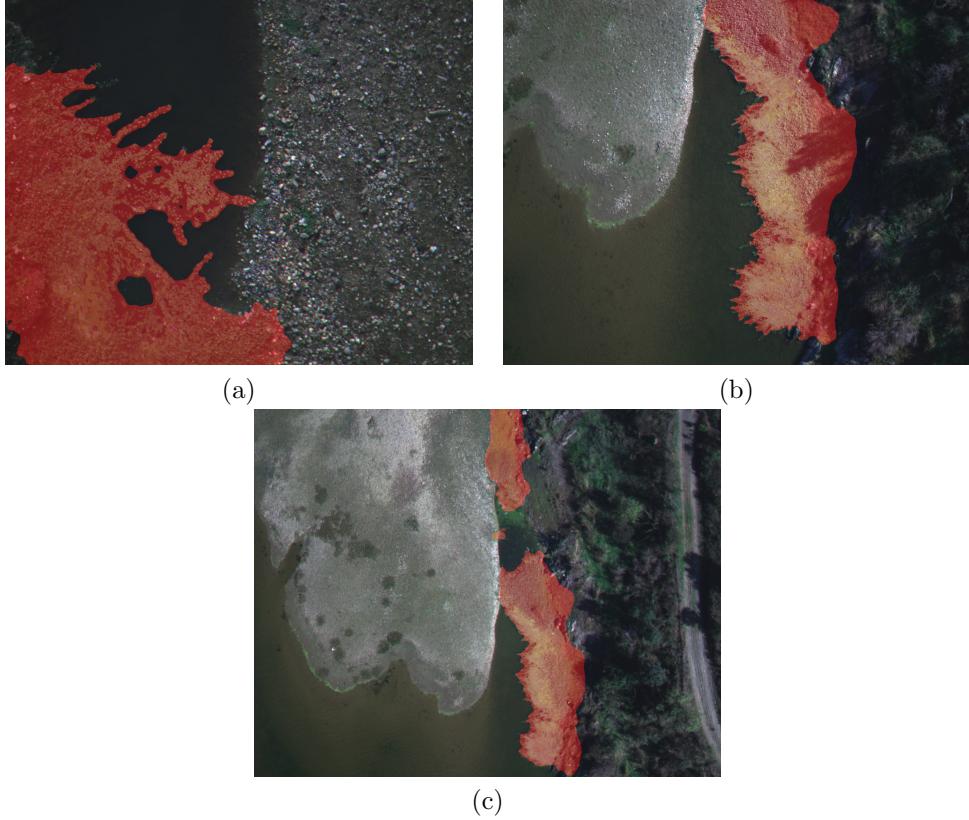


Figure 10: Predictions overlapped on their respective images.

7. Conclusion

The detection of IAS is a very pertinent topic. We set out to accomplish this type of detection using a low-cost sensor and created a new data set that is freely available. We also created a new approach to the detection of *Ludwigia*. p. by modification of a semantic segmentation method. We showed how segmentation models could be used to identify the targeted species. By using semantic segmentation models, we were able to leverage both the capabilities of these models to recognize objects and the multispectral nature of our data. Before being able to start training and testing models, we need first to have a data set. Unfortunately, we could not find any publicly available data sets and had to build our own. We captured data at multiple altitudes, periods of the day, and atmospheric conditions. We wanted to make sure that

our data set was as representative as possible and contained variations of the atmospheric conditions. After collecting the necessary images, we started the processing and annotation steps of the data set. Once annotated, we could use our data set to train and test semantic segmentation models. We first used a preexisting model (HRNet (24)), which was modified to handle our data. Once we validated the concept of using a semantic segmentation model to detect *Ludwigia p.*, we established a baseline. We proposed a set of modifications to the HRNet model to increase its performance, especially at high altitudes, and reduce the training time. The proposed model needs considerably less training time while increasing performance in the scenarios we want. Comparing our results to those other authors achieved, we have comparable performance using simpler data.

8. Supplementary Materials

The complementary academic thesis of this work can be provided on request from the corresponding author.

9. Author Contributions

António Abreu: Conceptualization, Data collection planning and implementation, Data processing and annotation, Model development, experimentation and validation, Data analysis and interpretation, Validation, Visualization, Led manuscript writing with primary contributions from Luís A. Alexandre. **João Santos:** Conceptualization, Supervision, Data collection planning and implementation. **Luís A. Alexandre:** Funding acquisition, Conceptualization, Supervision, Writing – reviewing & editing, Visualization.

10. Funding

This research was funded by Zirak srl – Information Technology and by NOVA LINCS (UIDB/04516/2020) with the financial support of FCT-Fundaçao para a Ciéncia e a Tecnologia, through national funds.

11. Data Availability Statement

The data presented in this study are available on request from the corresponding author. The data are not publicly available due to their large size.

12. Acknowledgments

The authors would like to thank Hasty.ai for providing a subscription, free of charge, that allowed to fully access all the features. It made the annotation process easier and faster, allowing the authors more time to focus on more critical aspects of this work.

13. Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] E. A. Bolch, M. J. Santos, C. Ade, S. Khanna, N. T. Basinger, M. O. Reader, E. L. Hestir, Remote detection of invasive alien species, *Remote Sensing of Plant Biodiversity* (2020) 267–307doi:10.1007/978-3-030-33157-3_12.
URL <https://link.springer.com/chapter/10.1007/978-3-030-33157-3\12>
- [2] B. Gallardo, M. Clavero, M. I. Sánchez, M. Vilà, Global ecological impacts of invasive species in aquatic ecosystems, *Global Change Biology* 22 (1) (2016) 151–163. doi:10.1111/GCB.13004.
URL <https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.13004>
- [3] A. Malik, Environmental challenge vis a vis opportunity: The case of water hyacinth, *Environment International* 33 (1) (2007) 122–138. doi:10.1016/J.ENVINT.2006.08.004.
- [4] L. Thouvenot, J. Haury, G. Thiebaut, A success story: water primroses, aquatic plant pests, *Aquatic Conservation: Marine and Freshwater Ecosystems* 23 (5) (2013) 790–803. doi:10.1002/AQC.2387.
URL <https://onlinelibrary.wiley.com/doi/full/10.1002/aqc.2387>
- [5] S. Danelot, C. Robles, N. Pech, A. Cazaubon, R. Verlaque, Allelopathic potential of two invasive alien Ludwigia spp., *Aquatic Botany* 88 (4) (2008) 311–316. doi:10.1016/J.AQUABOT.2007.12.004.

- [6] I. Stiers, N. Crohain, G. Josens, L. Triest, Impact of three aquatic invasive species on native plants and macroinvertebrates in temperate ponds, *Biological Invasions* 2011 13:12 13 (12) (2011) 2715–2726. doi:10.1007/S10530-011-9942-9.
URL <https://link.springer.com/article/10.1007/s10530-011-9942-9>
- [7] S. Nehring, D. Kolthoff, The invasive water primrose Ludwigia grandiflora (Michaux) Greuter & Burdet (Spermatophyta: Onagraceae) in Germany: First record and ecological risk assessment, *Aquatic Invasions* 6 (1) (2011) 83–89. doi:10.3391/ai.2011.6.1.10.
URL <http://www.bkg.bund.de>
- [8] B. A. Bradley, Remote detection of invasive plants: A review of spectral, textural and phenological approaches, *Biological Invasions* 16 (7) (2014) 1411–1425. doi:10.1007/s10530-013-0578-9.
URL <https://link.springer.com/article/10.1007/s10530-013-0578-9>
- [9] K. S. He, D. Rocchini, M. Neteler, H. Nagendra, Benefits of hyperspectral remote sensing for tracking plant invasions, *Diversity and Distributions* 17 (3) (2011) 381–392. doi:10.1111/j.1472-4642.2011.00761.x.
URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1472-4642.2011.00761.x>
- [10] E. A. Bolch, E. L. Hestir, S. Khanna, Performance and Feasibility of Drone-Mounted Imaging Spectroscopy for Invasive Aquatic Vegetation Detection, *Remote Sensing* 2021, Vol. 13, Page 582 13 (4) (2021) 582. doi:10.3390/RS13040582.
URL <https://www.mdpi.com/2072-4292/13/4/582>
- [11] EUR-Lex - 32014R1143 - EN - EUR-Lex, last accessed in: 2021-10-02. [Online]. Available at.
URL <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex\%3A32014R1143>
- [12] Decreto-Lei 92/2019, 2019-07-10 - DRE, last accessed in: 2021-10-02. [Online]. Available at.
URL <https://dre.pt/home/-/dre/123025739/details/maximized>

- [13] H. Gholizadeh, M. S. Friedman, N. A. McMillan, W. M. Hammond, K. Hassani, A. V. Sams, M. D. Charles, D. A. R. Garrett, O. Joshi, R. G. Hamilton, S. D. Fuhlendorf, A. M. Trowbridge, H. D. Adams, Mapping invasive alien species in grassland ecosystems using airborne imaging spectroscopy and remotely observable vegetation functional traits, *Remote Sensing of Environment* 271 (mar 2022). doi:10.1016/J.RSE.2022.112887.
- [14] J. Baron, D. J. Hill, Monitoring grassland invasion by spotted knapweed (*Centaurea maculosa*) with RPAS-acquired multispectral imagery, *Remote Sensing of Environment* 249 (nov 2020). doi:10.1016/J.RSE.2020.112008.
- [15] J. Dai, D. A. Roberts, D. A. Stow, L. An, S. J. Hall, S. T. Yabiku, P. C. Kyriakidis, Mapping understory invasive plant species with field and remotely sensed data in Chitwan, Nepal, *Remote Sensing of Environment* 250 (dec 2020). doi:10.1016/J.RSE.2020.112037.
- [16] V. R. Tóth, P. Villa, M. Pinardi, M. Bresciani, Aspects of Invasiveness of *Ludwigia* and *Nelumbo* in Shallow Temperate Fluvial Lakes, *Frontiers in Plant Science* 0 (2019) 647. doi:10.3389/FPLS.2019.00647.
URL <https://www.frontiersin.org/articles/10.3389/fpls.2019.00647/full>
- [17] S. Khanna, M. J. Santos, J. D. Boyer, K. D. Shapiro, J. Bellvert, S. L. Ustin, Water primrose invasion changes successional pathways in an estuarine ecosystem, *Ecosphere* 9 (9) (2018) e02418. doi:10.1002/ECS2.2418.
URL <https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ecs2.2418>
- [18] H. Yao, R. Qin, X. Chen, Unmanned Aerial Vehicle for Remote Sensing Applications—A Review, *Remote Sensing* 2019, Vol. 11, Page 1443 11 (12) (2019) 1443. doi:10.3390/RS11121443.
URL <https://www.mdpi.com/2072-4292/11/12/1443/htmhttps://www.mdpi.com/2072-4292/11/12/1443>
- [19] T. Z. Xiang, G. S. Xia, L. Zhang, Mini-Unmanned Aerial Vehicle-Based Remote Sensing: Techniques, applications, and prospects, *IEEE Geoscience and Remote Sensing Magazine* 7 (3) (2019) 29–63.

doi:10.1109/MGRS.2019.2918840.
URL <https://arxiv.org/abs/1812.07770>

- [20] T. Kattenborn, J. Lopatin, M. Förster, A. C. Braun, F. E. Fassnacht, UAV data as alternative to field sampling to map woody invasive species based on combined Sentinel-1 and Sentinel-2 data, *Remote Sensing of Environment* 227 (2019) 61–73. doi:10.1016/J.RSE.2019.03.025.
- [21] A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, *Advances in Neural Information Processing Systems* 25 (2012).
URL <http://code.google.com/p/cuda-convnet/>
- [22] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, Y. Sun, T. He, J. Mueller, R. Manmatha, M. Li, A. Smola, U. Davis, ResNeSt: Split-Attention Networks (apr 2020). arXiv:2004.08955.
URL <https://arxiv.org/abs/2004.08955v2>
- [23] Y. Yuan, X. Chen, X. Chen, J. Wang, Segmentation Transformer: Object-Contextual Representations for Semantic Segmentation, arXiv (2019) 1–23arXiv:1909.11065.
URL <https://arxiv.org/abs/1909.11065v6>
- [24] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang, W. Liu, B. Xiao, Deep High-Resolution Representation Learning for Visual Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43 (10) (2019) 3349–3364. arXiv:1908.07919, doi:10.1109/TPAMI.2020.2983686.
URL <https://arxiv.org/abs/1908.07919v2>
- [25] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam, Rethinking Atrous Convolution for Semantic Image Segmentation (jun 2017). arXiv:1706.05587.
URL <https://arxiv.org/abs/1706.05587v3>
- [26] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*

- 11211 LNCS (2018) 833–851. arXiv:1802.02611, doi:10.1007/978-3-030-01234-2_49.
URL <https://arxiv.org/abs/1802.02611v3>
- [27] L. C. Chen, M. D. Collins, Y. Zhu, G. Papandreou, B. Zoph, F. Schroff, H. Adam, J. Shlens, Searching for Efficient Multi-Scale Architectures for Dense Image Prediction, Advances in Neural Information Processing Systems 2018-Decem (2018) 8699–8710. arXiv:1809.04184.
URL <https://arxiv.org/abs/1809.04184v1>
- [28] P. J. Weisberg, T. E. Dilts, J. A. Greenberg, K. N. Johnson, H. Pai, C. Sladek, C. Kratt, S. W. Tyler, A. Ready, Phenology-based classification of invasive annual grasses to the species level, Remote Sensing of Environment 263 (sep 2021). doi:10.1016/J.RSE.2021.112568.
- [29] EHS - Hyperspectral Remote Sensing Scenes, last accessed in: 2021-12-10. [Online]. Available at.
URL https://www.ehu.eus/ccwintco/index.php?title=Hyperspectral\Remote_Sensing_Scenes
- [30] F. Yu, V. Koltun, Multi-Scale Context Aggregation by Dilated Convolutions, 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings (nov 2015). arXiv:1511.07122.
URL <https://arxiv.org/abs/1511.07122v3>
- [31] K. He, G. Gkioxari, P. Dollar, R. Girshick, Mask R-CNN, Proceedings of the IEEE International Conference on Computer Vision 2017-Octob (2017) 2980–2988. doi:10.1109/ICCV.2017.322.
- [32] M. Tan, Q. V. Le, EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, 36th International Conference on Machine Learning, ICML 2019 2019-June (2019) 10691–10700. arXiv:1905.11946.
URL <https://arxiv.org/abs/1905.11946v5>
- [33] S. K. Roy, S. R. Dubey, S. Chatterjee, B. B. Chaudhuri, FuSENet: Fused squeeze-and-excitation network for spectral-spatial hyperspectral image classification, IET Image Processing 14 (8) (2020) 1653–1661. doi:10.1049/IET-IPR.2019.1462.
URL <https://www.researchgate.net/publication/340403143\>

_FuSENet_Fused_Squeeze-and-Excitation_Network_for\
_Spectral-Spatial_Hyperspectral_Image_Classification

- [34] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications (apr 2017). arXiv:1704.04861.
URL <https://arxiv.org/abs/1704.04861v1>
- [35] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, Y. Wei, Deformable Convolutional Networks, Proceedings of the IEEE International Conference on Computer Vision 2017-Octob (2017) 764–773. arXiv:1703.06211, doi:10.1109/ICCV.2017.89.
URL <https://arxiv.org/abs/1703.06211v3>
- [36] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei, O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. Fei-Fei, ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision 115 (3) (2014) 211–252. arXiv:1409.0575, doi:10.1007/s11263-015-0816-y.
URL <https://arxiv.org/abs/1409.0575v3>
- [37] F. Chollet, Xception: Deep Learning with Depthwise Separable Convolutions, Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017-Janua (2016) 1800–1807. arXiv:1610.02357, doi:10.1109/CVPR.2017.195.
URL <https://arxiv.org/abs/1610.02357v3>
- [38] P. Krähenbühl, V. Koltun, Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials, Advances in Neural Information Processing Systems 24: 25th Annual Conference on Neural Information Processing Systems 2011, NIPS 2011 (oct 2012). arXiv:1210.5644.
URL <https://arxiv.org/abs/1210.5644v1>
- [39] Y. Yuan, X. Chen, X. Chen, J. Wang, Segmentation Transformer: Object-Contextual Representations for Semantic Segmentation, arXiv (2019) 1–23arXiv:1909.11065.
URL <https://arxiv.org/abs/1909.11065v6>

- [40] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-Decem (2015) 770–778. arXiv:1512.03385, doi:10.1109/CVPR.2016.90. URL <https://arxiv.org/abs/1512.03385v1>
- [41] R. G. Pontius, Quantification Error Versus Location Error in Comparison of Categorical Maps (2000).
- [42] J. Wang, K. Sun, T. Cheng, B. Jiang, C. Deng, Y. Zhao, D. Liu, Y. Mu, M. Tan, X. Wang, W. Liu, B. Xiao, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, MARCH 2020 Deep High-Resolution Representation Learning for Visual Recognition (2019). arXiv:1908.07919v2. URL <https://github.com/HRNet>.
- [43] E. L. Hestir, S. Khanna, M. E. Andrew, M. J. Santos, J. H. Viers, J. A. Greenberg, S. S. Rajapakse, S. L. Ustin, Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem, *Remote Sensing of Environment* 112 (11) (2008) 4034–4047. doi:10.1016/J.RSE.2008.01.022.
- [44] M. J. Santos, E. L. Hestir, S. Khanna, S. L. Ustin, Image spectroscopy and stable isotopes elucidate functional dissimilarity between native and nonnative plant species in the aquatic environment, *New Phytologist* 193 (3) (2012) 683–695. doi:10.1111/J.1469-8137.2011.03955.X. URL <https://www.readcube.com/articles/10.1111%2Fj.1469-8137.2011.03955.x>
<https://www.readcube.com/articles/10.1111/j.1469-8137.2011.03955.x>
- [45] A. Hirano, M. Madden, R. Welch, Hyperspectral image data for mapping wetland vegetation, *Wetlands* 2003 23:2 23 (2) (2003) 436–448. doi:10.1672/18-20. URL <https://link.springer.com/article/10.1672/18-20>
- [46] J. B. Adams, D. E. Sabol, V. Kapos, R. Almeida Filho, D. A. Roberts, M. O. Smith, A. R. Gillespie, Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon, *Remote Sensing of Environment* 52 (2) (1995) 137–154. doi:10.1016/0034-4257(94)00098-8.

- [47] S. Khanna, M. J. Santos, S. L. Ustin, P. J. Haverkamp, An integrated approach to a biophysiological based classification of floating aquatic macrophytes, <https://doi.org/10.1080/01431160903505328> 32 (4) (2011) 1067–1094. doi:10.1080/01431160903505328.
URL <https://www.tandfonline.com/doi/abs/10.1080/01431160903505328>
- [48] L. W. Lehnert, H. Meyer, W. A. Obermeier, B. Silva, B. Regeling, B. Thies, J. Bendix, Hyperspectral data analysis in R: The hsdar package, Journal of Statistical Software 89 (2019). doi:10.18637/JSS.V089.I12.
URL https://www.researchgate.net/publication/325142486\Hyperspectral_Data_Analysis_in_R_the_hsdar_Package
- [49] G. S. Rowan, M. Kalacska, A review of remote sensing of submerged aquatic vegetation for non-specialists, Remote Sensing 13 (4) (2021) 1–50. doi:10.3390/RS13040623.
URL <https://www.mdpi.com/2072-4292/13/4/623>
- [50] J. L. Bandfield, P. R. Christensen, M. D. Smith, Spectral data set factor analysis and end-member recovery: Application to analysis of Martian atmospheric particulates, Journal of Geophysical Research: Planets 105 (E4) (2000) 9573–9587. doi:10.1029/1999JE001094.
URL <https://onlinelibrary.wiley.com/doi/full/10.1029/1999JE0010944>
- [51] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs, IEEE Transactions on Pattern Analysis and Machine Intelligence 40 (4) (2016) 834–848. arXiv:1606.00915, doi:10.1109/TPAMI.2017.2699184.
URL <https://arxiv.org/abs/1606.00915v2>
- [52] L. C. Chen, Y. Yang, J. Wang, W. Xu, A. L. Yuille, Attention to Scale: Scale-aware Semantic Image Segmentation, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-Decem (2015) 3640–3649. arXiv:1511.03339, doi:10.1109/CVPR.2016.396.
URL <https://arxiv.org/abs/1511.03339v2>

- [53] T. Chakraborty, U. Trehan, SpectralNET: Exploring Spatial-Spectral WaveletCNN for Hyperspectral Image Classification (apr 2021). arXiv:2104.00341.
 URL <https://arxiv.org/abs/2104.00341v1>
- [54] Q. Zhu, W. Deng, Z. Zheng, Y. Zhong, Q. Guan, W. Lin, L. Zhang, D. Li, A Spectral-Spatial-Dependent Global Learning Framework for Insufficient and Imbalanced Hyperspectral Image Classification, IEEE Transactions on Cybernetics (may 2021). arXiv:2105.14327, doi:10.1109/tcyb.2021.3070577.
 URL <https://arxiv.org/abs/2105.14327v1>
- [55] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, P. Luo, SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers (may 2021). arXiv:2105.15203.
 URL <https://arxiv.org/abs/2105.15203v3>
- [56] S. K. Roy, G. Krishna, S. R. Dubey, B. B. Chaudhuri, HybridSN: Exploring 3D-2D CNN Feature Hierarchy for Hyperspectral Image Classification, IEEE Geoscience and Remote Sensing Letters 17 (2) (2019) 277–281. arXiv:1902.06701, doi:10.1109/lgrs.2019.2918719.
 URL <https://arxiv.org/abs/1902.06701v3>
- [57] B. Somers, G. P. Asner, Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests, Remote Sensing of Environment 136 (2013) 14–27. doi:10.1016/J.RSE.2013.04.006.
 URL <https://www.sciencedirect.com/science/article/pii/S0034425713001338>
- [58] S. Khanna, M. J. Santos, E. L. Hestir, S. L. Ustin, Plant community dynamics relative to the changing distribution of a highly invasive species, Eichhornia crassipes: A remote sensing perspective, Biological Invasions 14 (3) (2012) 717–733. doi:10.1007/S10530-011-0112-X.