NHL STENDEN
computer vision
& data science

Instance Segmentation of Polymer Flakes through Hyperspectral Imaging

NHL Stenden Lectoraat in Computer Vision & Data Science

This concept paper is under review by the supervisors.

Maurice Ponte, Rami Mazloum

Supervisors: Klass Dijkstra, Maya Aghaei Gavari

Abstract—With the increasing amount of plastic waste being produced worldwide every year, effective recycling of this waste becomes increasingly important. With this research, we aim to introduce Instance Segmentation to this problem. We analyse the resulting Hyper Spectral Images of a Near Infrared Camera using the Mask R-CNN architecture. This was made possible by adding a convolutional layer to the beginning of Mask R-CNN for dimensionality reduction. The Mask R-CNN architecture and modifications have shown to handle dimensionality reduction and processing the Hyper Spectral Images. We found that Mask R-CNN shows promising results in creating instance segmentation masks with an average F1-score of 0.749 across all experiments and polymer types. Improvements could be made to the classification results as these results were most lacking. With visible evidence of the model confusing PE for PS in the second experiment. With this research we have shown promising results and that Instance Segmentation is well worth future research when tackling this subject.

Index Terms—Instance segmentation, Hyperspectral imaging, Deep learning, Mask R-CNN

1 Introduction

Even for the most conscious citizen, plastic is almost unavoidable, and living plastic-free requires a certain amount of access and privilege enjoyed by very few in the world [1]. Packaging materials such as plastic bags, bottles, clingfilm and even milk carton packages produce a very large amount of waste. To make matters worse these items make up a small percentage of all types of plastic, of which the European Union recognizes seven types of different codes: PET, HDPE, PVC, LDPE, PP, PS and others like PC, PA, PMMA, etc. These plastics produce pollution at every step of their lifecycle with them discarded as their most polluting stage, this poses a very serious environmental hazard [1]. In 2018 alone, more than 360 million tons of plastic was produced worldwide. With 62 million tons of which in the European Union, only 15.16% was recycled, equivalent to 9.4 million tons [2].

With the increase of recycling capabilities, a large fraction of this pollution can be reduced. Recycling plastics is proven to be a viable option when it comes to managing, decreasing or even partially solving this 'plastic crisis' [3]. By recycling, plastics lost energy is recovered by converting used plastics back into usable resources and sorting plays a crucial role in this process. To sort the plastic waste, it first needs to be washed and ground up into small flakes. These flakes can then be sorted by their composition. Manual sorting is a though job, time-consuming and error-prone. With the use of automation, these sorting systems can detect and analyze the physical and chemical properties of different types of plastics. These systems have their own shortcomings, they are not able to process all types of plastics simultaneously yet and are sensitive to contaminations (e.g. dirt). [4, 5].

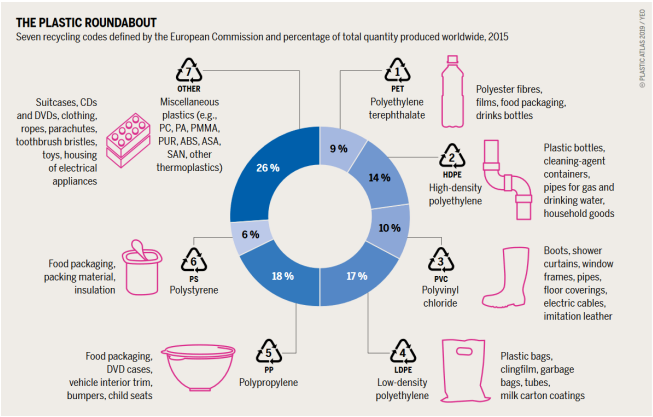

Fig. 1. The plastic roundabout [1].

The goal of this work is to implement instance segmentation for segmenting plastic flakes. In this work, we make use of Hyper Spectral Imaging (HSI). This technology enables obtaining hundreds of spectral wavelengths for each pixel. Earlier research groups have shown HSI to be a valid way to detect plastics [6, 7, 8, 9]. We continue this work by implementing instance segmentation on plastics flakes. This way we can have a per-pixel segmentation of the different types of polymer plastics. This could aid in precision and efficiency at sorting plastics.

1.1 Research Questions

To help define the goal of this research, a main question is set up as following.

- Can instance segmentation hyperspectral images aid in the automation of plastic waste sorting?

To cover the main question fully, these sub-questions must be answered:

Maurice Ponte is a Software Engineering student at the NHL Stenden University of Applied Sciences, E-mail: maurice.ponte@student.nhlstenden.com.

Rami Mazloum is a Software Engineering student at the NHL Stenden University of Applied Sciences, E-mail: rami.mazloum@student.nhlstenden.com.

Klaas Dijkstra is a associate lector at the NHL Stenden Professorship in Computer Vision & Data Science, E-mail: klaas.dijkstra@nhlstenden.com.

Maya Aghaei Gavari is a researcher at the NHL Stenden Professorship in Computer Vision & Data Science, E-mail: maya.aghaei.gavari@nhlstenden.com.

- How can we apply instance segmentation on hyperspectral images?
- Is Instance Segmentation a viable strategy for detecting and classifying plastic flakes in HSI?

2 State of the Art

In this section, the state of the art of instance segmentation and hyperspectral imaging is presented.

2.1 Instance Segmentation

Instance segmentation is a sub-type of segmentation. The main difference with other types of segmentation is that with instance segmentation we segment multiple instances of one class (if available). This results in a per pixel segmentation of objects. The difference can be best observed in figure 2.

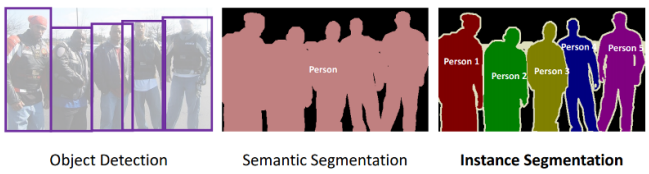

Fig. 2. An example of object detection, semantic and instance segmentation [10].

Instance segmentation can be done using deep neural networks such as Mask-RCNN, as shown in [11, 12, 13]. While making these algorithms is broadly researched, applying this technique to any plastics is a research topic that has little coverage to this day. The work in [14] uses adaptive thresholding combined with connected component analysis. This is not a deep learning approach moreover, the data they use is neatly separated. In [15] they propose a custom pipeline and compare it to the Mask-RCNN algorithm. This custom pipeline achieves quite good results when segmenting fibres with little overlapping compared to the Mask-RCNN model (M-RCNN: 0.57 vs Custom: 0.83 precision). However when fibres overlap the metrics used drop significantly (M-RCNN: 0.23 vs Custom: 0.41 precision).

2.2 Hyperspectral Imaging

A hyperspectral image is the product of a Specim FX17 camera. This camera can capture 224 bands of the Near-Infrared Spectroscopy (NIR) spectrum within a range of between 900nm and 1700nm. To put this in perspective, the human eye can only cover three bands, red, green and blue. A hyperspectral image encapsulates the data from these bands for each pixel in the entire image. This results in a three-dimensional datacube, figure 3.

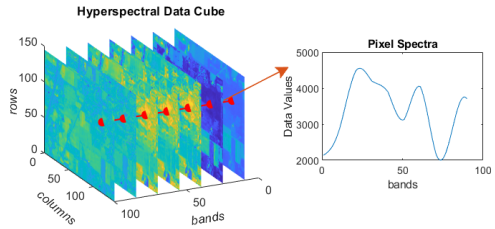

Fig. 3. An example visualization of a hyperspectral data cube and the spectrum for a given pixel [16].

Using Hyperspectral imaging, multiple research groups have shown to be able to detect various types of polymer plastic [6, 7, 8]. This is possible because of the unique absorption levels in various materials. However, these research groups do not implement deep neural networks. Especially on combining HSI with instance segmentation.

3 Materials and Methods

In this section, we give a comprehensive overview of the materials and methods used in this work.

3.1 Methods

In this work, we test the feasibility of instance segmentation on HSI. To achieve this, we use existing Deep Learning models capable of instance segmentation. These models are made to expect 3 channel RGB data as input. This poses a challenge to our research because of the 224 channel data within HSI's. First, we briefly explain the existing models and then give our adjustment.

3.1.1 Mask-RCNN

Mask R-CNN, extends Faster R-CNN [17] by adding a branch for creating an object mask in addition to the existing branch for bounding box recognition [11]. The first stage is the Region Proposal Network (RPN) and this produces class and box predictions. In parallel to this, the model also produces a binary mask for each Region of Interest (produced by the RPN). This architecture is shown in figure 4.

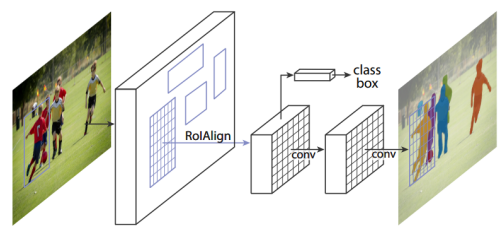

Fig. 4. The Mask-RCNN architecture as shown in [11]

3.1.2 SOLOv2

SOLOv2 is an iteration of the SOLO (Segmenting Objects by LOcations) method. This method does not rely on bounding box predictions, unlike Mask-RCNN. SOLO divides the input image into a uniform grid, *i.e.,* $S \times S$. If the centre of an object falls into a grid cell, that grid cell is responsible for 1) predicting the semantic category as well as 2) segmenting that object instance [18]. This concept is visible in figure 5.

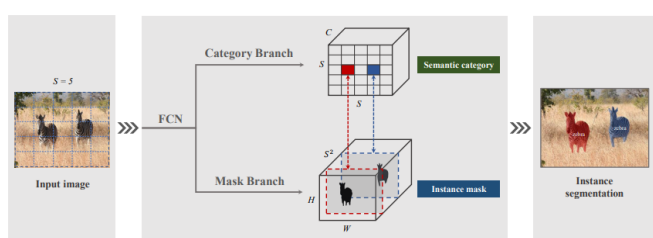

Fig. 5. **SOLO framework.** The illustration shows an input image with a 5x5 grid. If the centre of an object falls into a grid cell, that grid cell is responsible for predicting the semantic category (top) and the instances mask (bottom) [18].

SOLOv2 improves on SOLO by introducing a dynamic scheme for segmenting by separating the mask learning process into two:

convolution kernel learning and feature learning. Furthermore, they introduce a more efficient and effective non-maximum suppression algorithm. With these improvements, SOLOv2 outperforms SOLO by 1,9% while 33% faster. Interestingly the model surpasses many state-of-the-art highly engineered object detection methods when predicting bounding-boxes [13].

3.1.3 Modifications

The above-mentioned algorithms are made to work on three-channel RGB images. This poses a challenge to our work, as our images contain 224 channels of data for each pixel. To use the models, some minor modifications are necessary. The main modification includes an additional convolutional layer on top of the models. We used the following code to add the convolutional layer to the models.

```
class ChannelConv(nn.Module):
    def __init__(self,
                 backbone: torch.nn.Module,
                 in_channels,
                 out_channels=3,
                 kernel_size=1):
        """
        This class encapsulates any model and adds an additional layer to convert
        from in_channels to 3 channels.

        :param backbone: The backbone to wrap
        :param in_channels: The number of features of the input
        :param out_channels: The number of features for the input of model
        """

        super(ChannelConv, self).__init__()
        self.backbone = backbone
        self.conv = torch.nn.Conv2d(in_channels=in_channels,
                                    out_channels=out_channels,
                                    kernel_size=kernel_size, )

    def forward(self, x):
        if isinstance(x, list):
            x = [torch.squeeze(self.conv(
                    torch.unsqueeze(xx, dim=0)),dim=0) for xx in x]
        else:
            x = self.conv(x)

        y = self.backbone(x)
        return y
```

3.2 Computing Environment

To ensure reproducibility. Here we give a full overview of our computing environment (Table 1). Furthermore we use Python version 3.8.5 with Numpy version 1.20.0 and PyTorch version 1.9.

Table 1. Computing Hardware

CPU	12 Core Intel i9-7960x 2.8 GHz
GPU	GeForce RTX 2070 8 GB GDDR6 VRAM
RAM	16 GB
OS	Debian 10.1 'Buster'

3.3 Dataset

The dataset we use is composed of **72** hyperspectral images created with the Specim FX-17 camera. The images in the dataset contain multiple packaging types (e.g. bottles, flasks) ground up into flakes made of two polymer types (PS and PE).
To annotate our dataset, we converted the data cubes to an RGB representation. This was done by dividing the 224 channels into three equal parts, calculating the mean of these parts and used this result to make a three-channel datacube. This datacube was then converted to an RGB image. After conversion, we annotated the 'RGB'-images by applying polygons to each instance using Supervise.ly[1]. This annotation results in a JSON file containing the polygon information of each instance mask. For Mask-RCNN the annotation box are derived from the mask annotation. The annotation result is shown in figure 6.
Two samples had been discluded from the dataset, the first sample because it was difficult to annotate, the second has a big polymer flake figure which is not representive of the real world problem at hand in. Figure 7 these images can be seen.

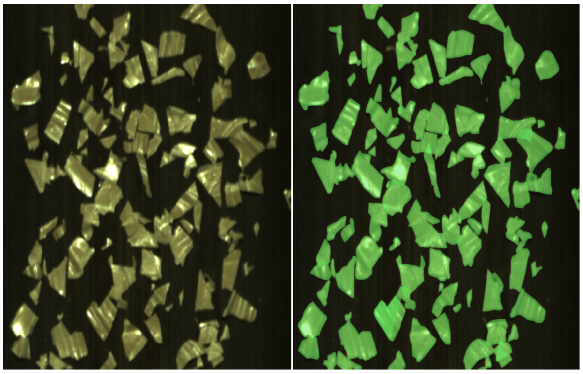

Fig. 6. A 'RGB'-image in our dataset (left) with the resulting annotation (right).

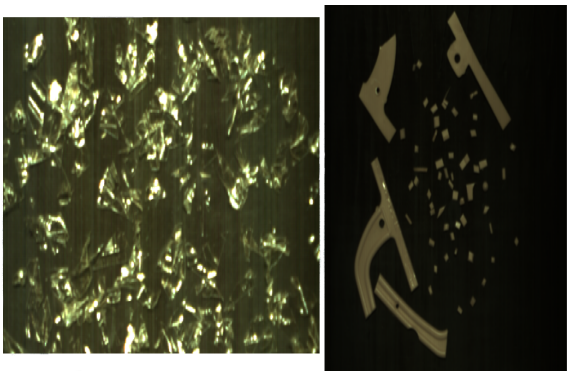

Fig. 7. First sample (left) Second sample(right).

4 Experiments & Results

In this section, we outline the experiments we designed to answer our research questions. The questions answered in this section are:

- How can we apply instance segmentation on hyperspectral images?
- Is Instance Segmentation a viable strategy for detecting and classifying plastic flakes in HSI?

What these research questions impose is a comparison between the two chosen methods of instance segmentation. Multiple experiments are designed to answer these questions:

- How do the networks score when asked to only look for flakes regardless of the type?
- How do the networks score when using multiple types of polymers?
- Does the size of the dataset influence the results?

4.1 Datasets

We composed two versions of our dataset. First is the 'large' dataset (Table 2), this dataset encapsulates all our available data. Second is our 'small' dataset (Table 3) which will encapsulate only a small portion of our data[2].

[1] https://supervise.ly/

[2] A sample refers to one kind of plastic (i.e. bottle, toothbrush) of which multiple different photos were taken in different compositions.

Table 2. Large dataset

	Polymer type	Samples	Images
Training			
	PS	4	12
	PE	5	15
	Total	9	27
Validation			
	PS	4	12
	PE	5	15
	Total	9	27
Testing			
	PS	4	12
	PE	5	15
	Total	9	27
Dataset total		27	81

Table 3. Small dataset

	Polymer type	Samples	Images
Training			
	PS	2	6
	PE	3	9
	Total	5	15
Validation			
	PS	2	6
	PE	3	9
	Total	5	15
Testing			
	PS	4	12
	PE	5	15
	Total	5	27
Dataset total		19	57

4.2 Performance metrics

To measuring the performance of our model a Confusion Matrix will be used. This matrix encapsulates the predictions and ground truth, With this matrix, the following metrics can be calculated: precision, recall and F1-Score.

Precision is the fraction of relevant instances among the predicted instances. The scale of precision is 1 (best) to 0 (worst).

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

Recall is the fraction of relevant instances predicted. The scale of recall is 1 (best) to 0 (worst).

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

F1-Score uses precision and recall to determine a mean value of accuracy. The scale of the F1 score is 1 (best) to 0 (worst).

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

For the evaluation of the correctness of mask predictions, we use the metric Intersection over Union (IoU) [19]. The Intersection over Union is calculated by looking at the overlap in masks between the prediction and the ground truth.

$$IoU = \frac{TP}{TP+FP+FN} \tag{4}$$

4.3 Experiment 1

We change the classes of our plastic flakes in the dataset to one generic class for this experiment. Consequently, we train the model to do background separation because the model does not have to take the spectral range of the different polymers into account. We do this to see if the model can correctly process Hyper Spectral Images, produce good instance masks and answer our first research question. Table 4 shows hyperparameters used.

Table 4. Hyperparameters for experiment 1

Batch size	1
Number of tiles	1
Learning rate	0.0001
Number of classes	1
Epochs	250

The resulting metrics for this experiment are shown in Table 5 and Table 6. We can see quite high metrics for the box and mask metrics which are promising results for the remainder of the experiments. In Figure 8 we can see how the mask metrics translate into actual instance masks.

Table 5. Experiment 1: Box metrics

Polymer type	F1-Score	Precision	Recall
Flake[3]	0.861	0.889	0.834

Table 6. Experiment 1: Mask metrics

Polymer type	IoU	F1-Score	Precision	Recall
Flake	0.865	0.927	0.910	0.946

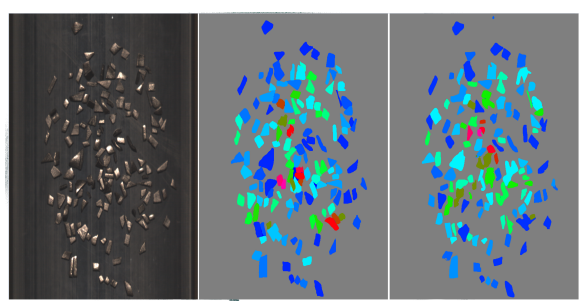

Fig. 8. Instance seqmentation results of experiment 1.LTR 'RGB' representation-Target mask - Result masks

4.4 Experiment 2

For this experiment, we train the model on all available classes in the dataset. The results of this experiment tell us how well the model can classify different polymer types and also create good instance masks. Table 7 shows hyperparameters used.

Table 7. Hyperparameters for experiment 2

Batch size	1
Number of tiles	1
Learning rate	0.0001
Number of classes	2
Epochs	250

The resulting metrics for this experiment are shown in Table 8 and Table 9. We can see a quite big difference in Intersection over Union

between the two classes. This becomes apparent when examining the classification results in some of the images figure 9. The green flakes in the image on the right are classified as PS and the blue ones are classified as PE. This while the whole image consists of PS Flakes as can be seen in the centre image. But again the instance masks are pretty good, see Figure 10.

Table 8. Experiment 2: Box metrics

Polymer type	F1-Score	Precision	Recall
PE	0.743	0.847	0.662
PS	0.697	0.622	0.792

Table 9. Experiment 2: Mask metrics

Polymer type	IoU	F1-Score	Precision	Recall
PE	0.660	0.795	0.852	0.745
PS	0.484	0.652	0.541	0.820

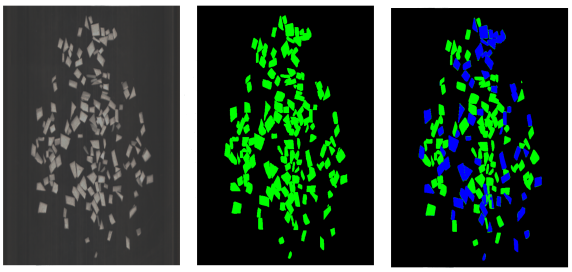

Fig. 9. Classification results of experiment 2.LTR 'RGB' representation-Target - Result .

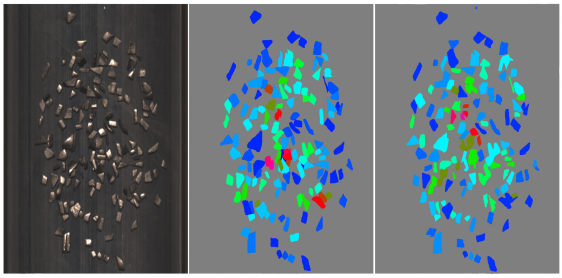

Fig. 10. Instance seqmentation results of experiment 2.LTR 'RGB' representation-Target mask - Result masks.

4.5 Experiment 3

For this experiment, we train the model on a smaller portion of our dataset. This experiment will tell us how well the model performs on smaller amounts of data. Table 10 shows hyperparameters used.

Table 10. Hyperparameters for experiment 3

Batch size	1
Number of tiles	1
Learning rate	0.0001
Number of classes	2
Epochs	500

The resulting metrics for this experiment are shown in Table 11 and Table 12. We can see that when doubling the amount of epochs used to train the model, it preforms similairly as in experiment 2.

Table 11. Experiment 3: Box metrics

Polymer type	F1-Score	Precision	Recall
PE	0.865	0.927	0.910
PS	0.865	0.927	0.910

Table 12. Experiment 3: Mask metrics

Polymer type	IoU	F1-Score	Precision	Recall
PE	0.627	0.771	0.790	0.752
PS	0.429	0.600	0.525	0.702

5 Discussion, Conclusion and Future work

In this section, we discuss the results of our experiments and a conclusion is made. We also propose future work.

5.1 Discussion

In this work, one convolutional neural network was used to test the feasibility of applying Instance Segmentation to HyperSpectral images. This was done to answer the following research question: Can instance segmentation hyperspectral images aid in the automation of plastic waste sorting? We used two sub-questions to provide an answer: How can we apply instance segmentation on HyperSpectral Images? and Is Instance Segmentation a viable strategy for detecting and classifying plastic flakes in HyperSpectral Images?

Relating the first sub-question, we used an additional convolutional layer to be able to process the 224 bands of HyperSpectral Images. Secondly, we used the Mask R-CNN instance segmentation architecture to produce instance masks. The combination of these made it possible to apply instance segmentation to HyperSpectral images.

The second sub-question can be answered by looking at the results of the second and third experiments. Lets first examine the box results in Table 13. Experiment 2 and 3 clearly have comparable box metric results when examined next to each other. This result shows that instance segmentation is a viable strategy for detecting plastic flakes in HyperSpectral Images. As mentioned earlier, the classification results are something that could be improved upon (see Figure 9). This also becomes apparent when looking at the resulting mask metric IoU of experiment 2 and 3 in Table 14. But when looking at an individual instance mask (Fig. 10) we see that this is very capable of creating good instance masks.

Table 13. Box metrics of experiment 2 and 3

	F1-Score		Precision		Recall	
Polymer	**Exp. 2**	**Epx. 3**	**Exp. 2**	**Epx. 3**	**Exp. 2**	**Epx. 3**
PE	0.743	0.865	0.847	0.927	0.662	0.910
PS	0.697	0.865	0.662	0.927	0.792	0.910

Table 14. IoU metrics of experiment 2 and 3

	IoU	
Polymer	**Exp. 2**	**Epx. 3**
PE	0.660	0.627
PS	0.484	0.429

5.2 Conclusions

With the findings of the discussion, we can conclude our research questions. Instance Segmentation on HyperSpectral Images can be achieved when applying the appropriate modifications to the chosen

network architecture, in our work the Mask R-CNN with modified convolutional layer.

Promising results are shown by the experiments for detecting, classifying and creating instance masks of plastic flakes using HyperSpectral Imaging and a Convolutional Neural Network. We have shown not only that the model could create good instance masks on much data, but it also had good performance on a smaller portion of data. Using these instance masks, separation of polymers is possible on the polymer types used in this work. Furthermore, the resulting instance masks can be used in recycling plants to create statistics on the plastic flakes (e.g. average polymer flake area of measurement).

This leads us to conclude that Instance Segmentation on Hyperspectral Images on Plastic Flakes is a promising method to apply in recycling plants for the recycling of plastic flakes. Although more research is needed to confirm this statement.

5.3 Future work

For future work, we would like to provide multiple research directions to strengthen our conclusion. The first is regarding the classification results of this work. We propose a study where a multitude of polymers are added to the dataset, this could possibly improve the classification results as the model has more variance in spectral range and thus leading to better classification results. Secondly, we propose examining preprocessing methods to improve classification results as in this study none were used.

Secondly, we propose a benchmark study of a multitude of neural networks. In this study, only one was used but more were studied and explained (see **??**). This was due to the time limitations of our research and the failed attempt of adding the convolutional layer to this architecture. While this was disappointing, the examined architecture was very promising and even more architectures were referenced [**?**, 13, 12]. It would be interesting to see how these architectures would hold up when compared to each other.

5.4 Acknowledgements

This project is a collaboration between NHL Stenden Professorship Computer Vision & Data Science and NHL Stenden Professorship Circular Plastics. We thank our supervisors Klaas Dijkstra and Maya Aghaei for their excellent guidance throughout this research and Femke Jaarsma for her contributions during weekly meetings.

References

[1] M Franklin L Fuhr. Plastic atlas: Facts and figures about the world of synthetic polymers. 2019*.

[2] Hannah Ritchie and Max Roser. Plastic pollution. *Our World in Data*, 2018. https://ourworldindata.org/plastic-pollution.

[3] Angeli Mehta. The plastic sorting challenge. 2020.

[4] J. M. Garcia A.R. Rahimi. Chemical recycling of waste plastics for new materials production. 2017.

[5] Biswajit Ruj, Vivek Pandey, Priyajit Jash, and V Srivastava. Sorting of plastic waste for effective recycling. *Int. J. Appl. Sci. Eng. Res.*, 4, 01 2015.

[6] Bin Hu, Silvia Serranti, Norbert Fraunholcz, Francesco Di Maio, and Giuseppe Bonifazi. Recycling-oriented characterization of polyolefin packaging waste. *Waste Management*, 33(3):574–584, 2013. Special Thematic Issue: Urban Mining.

[7] Yan Zheng, Jiarui Bai, Jingna Xu, Xiayang Li, and Yimin Zhang. A discrimination model in waste plastics sorting using nir hyperspectral imaging system. *Waste Management*, 72:87–98, 2018.

[8] Daniel Caballero, Marta Bevilacqua, and José Amigo. Application of hyperspectral imaging and chemometrics for classifying plastics with brominated flame retardants. *Journal of Spectral Imaging*, 8(1):a1, 2019.

[9] Zilong Zhong, Jonathan Li, Lingfei Ma, Han Jiang, and He Zhao. Deep residual networks for hyperspectral image classification. pages 1824–1827, 2017.

[10] DigitalSreeni. 73 - image segmentation using u-net - part1 (what is u-net?), 2019.

[11] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B. Girshick. Mask R-CNN. *CoRR*, abs/1703.06870, 2017.

[12] Daniel Bolya, Chong Zhou, Fanyi Xiao, and Yong Jae Lee. YOLACT++: better real-time instance segmentation. *CoRR*, abs/1912.06218, 2019.

[13] Xinlong Wang, Rufeng Zhang, Tao Kong, Lei Li, and Chunhua Shen. Solov2: Dynamic and fast instance segmentation. *arXiv: Computer Vision and Pattern Recognition*, 2020.

[14] J. Lorenzo-Navarro, M. C. Santana, M. Gómez, A. Herrera, and Pedro A. Marín-Reyes. Automatic counting and classification of microplastic particles. In *ICPRAM*, 2018.

[15] V. Wegmayr, A. Sahin, B. Sæmundsson, and J. M. Buhmann. Instance segmentation for the quantification of microplastic fiber images. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 2199–2206, 2020.

[16] mathworks.com. Hyperspectral data cube and a pixel spectra.

[17] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In Corinna Cortes, Neil D. Lawrence, Daniel D. Lee, Masashi Sugiyama, and Roman Garnett, editors, *NIPS*, pages 91–99, 2015.

[18] Xinlong Wang, Tao Kong, Chunhua Shen, Yuning Jiang, and Lei Li. SOLO: Segmenting objects by locations. In *Proc. Eur. Conf. Computer Vision (ECCV)*, 2020.

[19] Adrian Rosebrock. Intersection over union (iou) for object detection.