NHL STENDEN computer vision & data science

Polymer Semantic Segmentation with Hyperspectral Imaging and Artificial Neural Networks

NHL Stenden Centre of Expertise in Computer Vision & Data Science

Tianyi Liu

Supervisors: Klaas Dijkstra, Dijkstra Willem

Abstract—Since the 1950s, plastics have played an essential role in the world, and more than 8.3 billion tones of plastics have been produced. This number keeps increasing every year. Bio-polymers become a solution to the problem that non-degradable plastics are still used to produce frequently-used plastics products. How to sort them becomes a problem. A sorting solution neural network structure called GenNet has already developed by combine hyperspectral images with a neural network. This project is focusing on integrating a new neural network backbone to the GenNet which has already been able to learn and classify normal polymers such as PET, PEF, PP, and PE. The neural network backbone decided to integrate, called HRNet+OCR and HRNet. The HRNet+OCR achieved the highest IOU value 84.5% on semantic segmentation of public dataset Cityscapes, and it's a derivative network from HRNet. Therefore, these two networks are decided to integrate into the GenNet and train and test on a dataset with polymers such as PET, PEF, PP, PE, and bottle made in the same material. The result is that HRNet and HRNet+OCR can not provide better performance than Deep Res U-Net and U-Net++. And the performance of HRNet+OCR will be significantly influenced by choice of the optimizer. By changing optimizer to Adam, HRNet+OCR shows a 46.128% increase on Mean IOU value on the dataset which includes plastics flakes and same material bottles. The HRNet and HRNet+OCR can learn hyperspectral data but may not be the best choice for it.

Index Terms—DeepLearning, Plastics Classification, Hyper Spectral Imagery

1 Introduction

Since the 1950s, plastics have played an essential role in the world, and more than 8.3 billion tones of plastics have been produced. This number keeps increasing every year. Currently, there are still more than 350 million tones of plastics being produced every year, 6.3 billion tones of plastics are plastics waste, and only 9.5% of plastics waste has been recycled[1], but 79% of the waste plastics ended up in landfills or environment. If the demands of plastics keep increasing or unchanged, by 2050, there will be more plastics than fish in the ocean[2]. The reason why a large number of plastics were produced and discarded is that most of the one-time-use products are mainly made from plastics such as packaging, toys, electronics, and bottles.
The plastics ended up in the environment will bring significant health risk for the living organism[3], and this risk will finally come back to human due to the fact that micro-plastics has already been found in the human body[4].

Bioplastics become a solution to the problem that non-recyclable plastics are still used to produce frequently-used plastics products. Bioplastics are defined as being made from renewable resources(bio-based), biodegradable resources, or both of them. Examples of bio-based, non-bio-degradable plastics are bio-PE, bio-PP, and bio-PET[5]. Examples of plastics that are both bio-based and bio-degradable are PLA(polylactic acid), PHA(polyhydroxyalkanoate), and PBS(polybutylene succinate). But until these days, bioplastics are still only a tiny part of all plastics products, which only take 0.02 percent of all the plastics products.[6]
In 2019, 11 multinationals announced that they would work towards using only reusable, recyclable, or compostable packaging by 2025 or earlier [7].
Due to the increasing interest in bio-plastics, how to sort and recycle different kinds of polymers become important. Because the surface information such as color and shapes are similar for different types of plastics, it is hard to learn and classify different normal polymers for the neural network with only surface information.Therefore, an FX17 hyperspectral imaging camera is used to help with sorting and recognizing the polymers.
From previous research, a neural network code structure called GenNet was developed to classify regular plastics by combining the hyperspectral imaging and artificial neural network (ANNs)[8]. The GenNet is a modular which is integrated with four optimizers, two loss functions, a specially made dataloader for hyperspectral data, four preprocessing methods and has a complete process of preporcessing, training, validation ,testing and summarizing. It is more efficient to integrate new neural network into the GenNet rather than integrating components in GenNet such as dataloader and preprocessing methods into other network code structure which is made for classify normal two dimensional images. Therefore, the research questions for this project are:

- What neural network can be integrated to the GenNet
- What is the performance of this neural network backbone?
- What parameters or components will influence the performance of this neural network?

2 State of the art

For different types of polymers, for example PEF and PET,the surface,shape color are quite similar which is hard for neural network to learn and train with. The Hyperspectral imaging can give more information of the material that can not be seen by human eyes and can be used to distinguish difference between different types of polymers. But the Hyperspectral cubes data for different polymers are too complex for analyzing manually cause it contains too much information for every spectrum band it captured. Neural network has the capacity to handle the large amount of data with the help of computer processing power[9]. Therefore, combine Neural Network

Tianyi Liu is a Mecgatronic Engineering student at the HZ University of Applied Sciences, E-mail: ltylty221@gmail.com.
Klaas Dijkstra is a researcher at the NHL Stenden Centre of Expertise in Computer Vision & Data Science, E-mail: klaas.dijkstra@nhlstenden.com.
Willem Dijkstra is a researcher at the NHL Stenden Centre of Expertise in Computer Vision & Data Science, E-mail: willem.dijkstra@nhlstenden.com.

with Hyperspectral image become a solution. In this section,the state of the art of artificial neural networks, hyperspectral imaging and plastic sorting will be discussed.

2.1 Artificial Neural Networks

Artificial neural networks (ANN) are computing systems vaguely inspired by the biological neural networks that human brains[10]. Artificial neural networks can “learn” to perform tasks by considering examples, generally without being programmed with task-specific rules. In artificial neural networks, simple manual nodes, called neurons, connected together to form a similar biological neural network mesh structure.

In the area of analyzing visual imagery, a convolutional neural network (CNN) is the most commonly used cause of its neurons can respond to a part of the surrounding cells in the coverage area[11].

A convolutional neural network consists of one or more convolutional layers and a fully connected layer at the top (corresponding to a classic neural network). It also includes association weights and a pooling layer. A Simplified schematic of a convolutional neural netork is shown in Figure 1.This structure enables convolutional neural networks to take advantage of the two-dimensional structure of the input data. Compared with other deep learning structures, convolutional neural networks can give better results in image and speech recognition. This model can also be trained using a back-propagation algorithm. Compared with other deep and feed-forward neural networks, convolutional neural networks need to consider fewer parameters, making it an attractive deep learning structure[11].

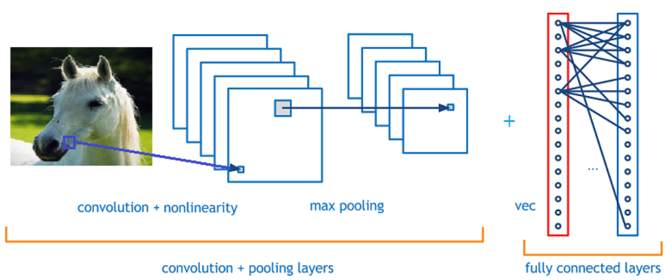


Fig. 1. Simplified schematic of a convolutional neural network[12].

2.2 HRNet+OCR and HRNet

HRNet+OCR structure is a neural network structure founded on a ranking list on website ’paperswithcode.com,’ which is about the performance of semantic segmentation on the ’Cityscapes test’ dataset based on IOU value. HRNet+OCR structure has the highest value of IOU as 84.5%. HRNet+OCR backbone is a derived backbone based on HRNet[13].

The full name of HRNet is High-resolution network. The HRNet has a feature that it will aggregate the output representations at four different resolutions, and then use a 1X1 convolutions to fuse these representations. The output representation will be fed into the classifier[14]. The structure of HRNet is shown in Figure 2

The full name of OCR is Object Contextual Representations. It is

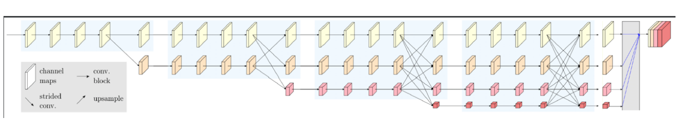


Fig. 2. Structure of HRNet[14]

a function model that can be integrated into HRNet or ResNet.The purpose is that the class label assigned to one pixel is the category of the object that the pixel belongs to. In the normal neural network, the class label will be assigned to the pixel if the pixel has satisfied the feature the network has learned about this class. It may cause some pixel that does not belong to an object such as a point in the background being labeled as an object[13].

In the original code of HRNet, the developers only provide the SGD as the option for training and testing on the dataset Cityscapes used in their project. In the paper of this project, it is mentioned that for dataset Cityscapes, the SGD optimizer improves the performance of the network by 0.6% on panoptic quality compared with Adam optimizer[14]. What also needs to be mentioned it that in the previous research and project[15], the Adam optimizer provides better performance on hyperspectral data on F1 score and Mean IOU for network backbone Deep Res U-Net and U Net++.

There are other neural network structures and related papers, and code can be found through paperwithcode.com, such as DCNAS, Panoptic-DeepLab.There are also some neural networks specially made to learn hyperspectral data but use the Tensorflow library instead of PyTorch such as HypbridSN and Recurrent 3D-CNN.

2.3 Hyperspectral Cube

Hyperspectral imaging(HSI) is an emerging field in which the advantages of optical spectroscopy as an analytical tool are combined with two-dimensional object visualization obtained by optical imaging. In HSI, each pixel of the image contains spectral information, which is added as a third dimension of values to the two-dimensional spatial image, generating a three-dimensional data cube, sometimes referred to as hyperspectral cube data or as an image cube.

Hyperspectral data cubes can contain absorption, reflectance, or fluorescence spectrum data for each image pixel. It is assumed that HSI data is spectrally sampled at more than 20 equally distributed wavelengths. The spectral range in hyperspectral data can extend beyond the visible range(ultraviolet, infrared)[16].

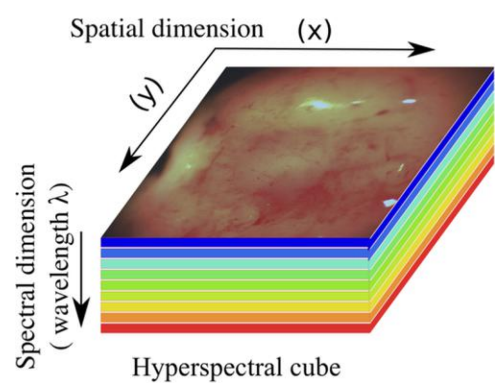


Fig. 3. Schematic representation of hyperspectral imaging hypercube showing the relationship between spectral and spatial dimensions[17]

2.4 Plastic sorting by NIR spectrography

At present, to sort plastics automatically is by using near-infrared (NIR) spectrography. With infrared cameras scanning on a conveyor belt, the system measures the absorption of infrared. The absorption values form a spectrum in which peaks and valleys in the spectrum that are unique to specific types of plastics can be used to identify the type of plastics. NIR spectrography does not measure the spatial dimension, therefore, it cannot distinguish different shapes of the same type of plastics or take a near pixel into the same class[18].

2.5 Semantic Segmentation

Semantic segmentation is the task of classifying every pixel in an image into a class as shown in the image below. Here you can see that all persons are red, the road is purple, the vehicles are blue, street signs are yellow, etc.

Semantic segmentation is different from instance segmentation which

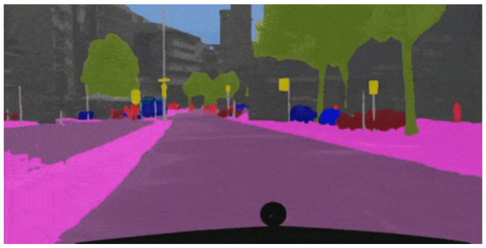


Fig. 4. Semantic Segmentation[19]

is that different objects of the same class will have different labels as in person1, person2 and hence different colours. The picture below very crisply illustrates the difference between instance and semantic segmentation[19]. In figure 4, it also can be seen that image Classification is about recognizing all the categories existed in the images, and object detection is to detect instances of semantic objects of a certain class such as humans, buildings, or cars in images.

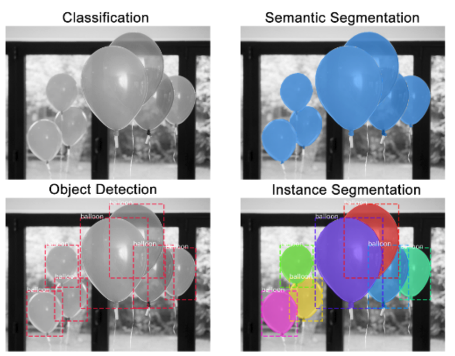


Fig. 5. Different tasks in computer vision[19]

3 MATERIALS AND METHODS

From the previous project[8], a convolutional neural network implementation was developed as a foundation for the following project to build on top by asserting an easy integration of new architectures, loss and gradient functions, data augmentation, and datasets. In order to apply the architecture, it needs to train, validate and test with hyperspectral cubes samples and a two-dimension class-map image in which index values represent what class a pixel belongs to are assigned to each pixel in the image and taken to be the ground truths data–the results the neural network is expected to predict.

3.1 Datasets

In previous projects[8][15].There are two datasets in hyperspectral cubes of 224 channels with pixels as the variable spatial dimension. These datasets are captured by the Specim FX17 camera whose bandwidths used is from 950 to 1700 nanometers in the electromagnetic spectrum, from visible light to near-infrared spectrum.

Dataset	Class
Plastics	PET,PEF,PE,PP
PET-PEF-Cp-shaped	PET,PEF,PET Bottle, PEF Bottle,Co-11p, Co-6p

Table 1. Datasets used in the project

Plastics dataset (White paper background): A dataset which is a union of three dataset from previous project made by a hyperspectral camera Specim FX17 [20][21][22]. The dataset contains 31 hyperspectral cubes, which is a raw snapshot taken on a white paper background for four types of polymers PET, PEF, PE, and PP. There are 27 samples for the training set, and the validation set consists of four samples.There are another three dataset called Plastics-hh, Plastics-hhi and Plastics-hhs which are separately preprocessed by Hyper-Hue, preporcessed by Hyper-Hue with additional intensity information and preprocessed by Hyper-Hue with additional saturation information. The ground truth images of these four dataset are the same.

PET-PEF-CP-shaped dataset: A dataset contains 47 image which include 7 classes PET, PET bottle, PEF, PEF bottle, 6% co-polymer(6% PET in PEF) and 11%co-polymer(11% PET in PEF). Forty-five images in this dataset are used for training, and two images are taken as the validation set and testing set. The dataset has been preprocessed by logarithmic derivative.

3.2 Preprocessing

Due to the hyperspectral cube degradation by the radiance captured by the camera, preprocessing for hyperspectral cube is necessary to normalize, improve (such as reducing the noise) and label the information to let the neural network process them[23].

Because of the visible light spectrum information can hardly provide information of difference between each kind of polymers, and may also cause noise and influence result after train the network, Logarithmic Derivative and Hyper-Hue Intensity and Saturation preprocessing methods are provide from the previous project to remove then Intensity and Saturation information which is part of information of visible light spectrum[8]. A general flowchart of the preprocessing of the current architecture Gennet is shown in the Figure 6.

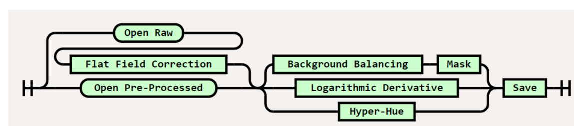


Fig. 6. Pipeline of Pre-processing[8]

For each sample of hyperspectral image, there is two-step will be taken:

1. Flat field correction
2. Logarithmic Derivative or Hyper-Hue Intensity and Saturation.

The purpose of this step 2 is to mitigate or remove the Intensity or Saturation in the color space, which are two important information of a two-dimensional image. In this way, the neural network will learn more on hyperspectral feature instead of two-dimensional color or shaped features. After this step, the dataset will be ready to input into the neural network to train[8].

Flat Field Correction:

Flat field correction will correct sensor inhomogeneity and reduce the noise[24].

$$D_{wc} = \frac{1}{h_n} \sum_{i=1}^{h_n} D_{iwc} \tag{1}$$

$$W_{wc} = \frac{1}{h_n} \sum_{i=1}^{h_n} W_{iwc} \tag{2}$$

$$C_{hwc} = \frac{I_{hwc} - D_{wc}}{W_{wc} - D_{wc}} \forall h \tag{3}$$

For the equation shown above, Equation 1 is to calculate the dark reference, Equation 2 is to calculate the white reference, and Equation 3 is the equation for the flat field correction. For the equations above, C is the flat field corrected image, I is the original image, D for dark reference, W for white reference, h,w for the height and width, and b is for the spectral band.

With the processed image from flat field correction, a normalized base image can be preprocessed by Logarithmic Derivative, Hyper-Hue Saturation Intensity, and Background balancing algorithms.

Logarithmic Derivative:

Logarithmic Derivative will take every pixel and pixel near to it into processing and mitigate the intensity information and keep the essential characteristics of the data. The algorithm will divide the derivative of the image by the original image which will mitigate the Intensity among the pixel near the processed pixel and result in a logarithmic derived cube that can be processed by the neural network[25].

$$\frac{d}{dx}ln(x) = \frac{1}{x} \tag{4}$$

$$\frac{d}{dx}log_b(x) = \frac{1}{ln(b) \times x} \tag{5}$$

The Equation 4 and 5 represent the Logarithmic Derivative[25].

Hyper-Hue Saturation Intensity (HHSI):

Hyper-Hue Saturation Intensity (HHSI) algorithm [26] will separate the Intensity and the saturation information of the data. First, the maximum and minimum value on the hyperchromatic axis will be connected. Then, by projecting the cube to a plane perpendicular to the hyperchromatic axis, the Intensity will be separated. Last, the algorithm will separate the saturation information from the origin of the projection on each axis resultant and leave the Hue for each band. The algorithm will not change the size or dimension of the data. Therefore, the data can still be input to the next step[26][27].

Background Balancing:

Background Balancing will average the value of the background of a sample in all channels. After the Background balancing, the result can be used by computer vision to create the mask[28].

$$P_{hwc} = C_{hwc} \forall w \in \{1,2,3,...\} \tag{6}$$

$$A_c = \frac{1}{WH} \sum_{i=1}^{WH} P_{hwc} \forall c \tag{7}$$

$$B_{hwc} = C_{hwc} \frac{max(A)}{A_c} \forall c \tag{8}$$

Equation 6, 7 and 8 represent the algorithm of Background balancing[28], where C represents the flat field corrected image, A represents the vector with average background values per channel, W, H represents the width and height, w, h represents the indices of the width and height and B represents the background balanced image.

With the preprocess of Logarithmic derivative or Hyper-Hue Saturation Intensity, the result will be in the following formats:

- 224 Channels data with Intensity removed by Logarithmic Derivative.
- 224 Channels data with Saturation removed by Hyper-Hue + Intensity.
- 224 Channels data with Intensity removed by Hypere-Hue + Saturation.
- 223 Channels data with Intensity and Saturation removed by Hyper-Hue.

3.3 Software architecture

3.3.1 GenNet

The architecture from the previous project – GenNet was built based on the PyTorch, Numpy framework. The code was formed in a modular way, and the function of the different components can be modified by parameters easily, which can benefit the following project for experimenting and integrating new functions or architectures. The code is programmed in Python, and each component in the structure is a python class which means it can be used in PyTorch.

In the previous project, DeepResUNet 2018, U-Net 2018, and U-Net 2015 were used as backbones in the architecture to run the test[8][28][29]. Also, DeepLabV+3 neural network backbone was integrated into the GenNet.

Therefore, in this project, The neural network backbone is going to be integrated into the GenNet and test with normal polymers is HRNet–High-Resolution Network and its derived network HRNet+OCR.

After the preprocessing, the data input into the architecture will be an (NxBxHxW) tensor (Number of samples, Channels, Height, Width)where B should be the same as the number of channels of the input data. The architecture will output a spatial dimension class-map which should also be an (NxCxHxW) tensor where C represents the probability of each class for each pixel.

In the neural network, the Loss Function will calculate the degree of fit between the output and its target or ground truth data.[30] For HRNet+OCR, a pixel-wise Cross-Entropy loss is used, which means it is using a 1 X 1 convolution. The Cross-Entropy Loss function can be implemented by PyTorch. In GenNet, the input to the loss function is the (NxCxHxW) output of the network and an (NxHxW) target of the sample tensor.

The optimizer in the neural network will optimize the parameters during training the neural network by processing results from loss functions. In GenNet, the optimizer that can be used in the network can be modified by changing the configuration parameters as long as it is provided by PyTorch. For HRNet, the optimizer to use in orginal code is SGD[31].

3.3.2 Evaluation Metrics

To review the performance of the neural network, the Confusion Matrix will be used. By using the confusion matrix, it can measure how correctly the probability predicted by the neural network and see whether neural network has confused one class with the other and what's performance of the neural network. In GenNet, pixel accuracy,F1 score, mean intersection over union(IOU), frequency weighted intersection over union can be generated for measure and assess the performance of the network.

Accuracy, Precision, Recall and F1 score

The following equations 9, 10, 11 and 12 indicate how to calculate accuracy, precision, recall and F1 score.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{9}$$

$$Precision = \frac{TP}{TP+FP} \tag{10}$$

$$Recall = \frac{TP}{TP+FN} \tag{11}$$

$$F1score = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{12}$$

For a certain class and for a single sample,True Positive(TP) means value of actual class is yes and the value of predicted class is also yes.

True Negative(TN) means the value of actual class is no and value of predicted class is also no.

False Positive(FP) means the actual class is no and predicted class is yes.

False Negative(FN) means the actual class is yes but predicted class in no[32]. A sample is shown in Figure 7.

		Predicted class	
Actual Class		Class = Yes	Class = No
	Class = Yes	True Positive	False Negative
	Class = No	False Positive	True Negative

Fig. 7. True Positive, True Negative, False Positive,False Negative[32]

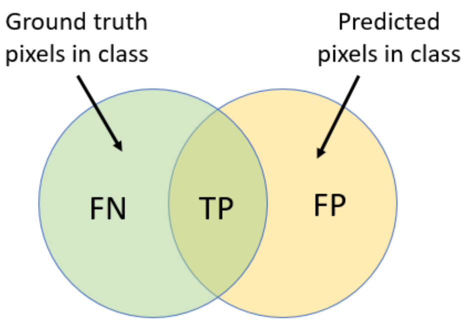


Fig. 8. A sample of the IOU[33]

Intersection-Over-Union:
The Intersection-over-Union (IoU) is an evaluation metric for semantic segmentation networks[33].The IOU value is calculated by the area of overlap between prediction from the network and the ground truth divided by the area of union between the prediction and the ground truth. A sample is shown in Figure 8. Equation 13 shows how to calculate the intersection-over-union.

$$IoU = \frac{TP}{TP+FP+FN} \tag{13}$$

Mean IOU is the average IOU value overall classes. Frequency Weighted IOU value is the IOU value from each class multiply a weight value calculated by the number of samples in this class divided by the total number of samples.

3.3.3 Pipeline

The pipeline of the GenNet is shown in the figure with three main steps: Initialization, Learning, and Testing[8].

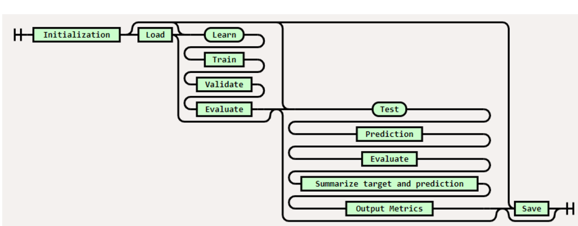


Fig. 9. Pipeline of whole architecture[8]

In Initialization:
1. Define the class, number of class, and it's corresponding pixel values in the masks for the data.
2. Define different usage for dataset: training, validation or testing
3. Define the parameters and components of the network, including Backbone, Loss function, Optimizer, Coders, Metrics evaluator, and a loss graph.
4. Define the way to visualize the results of the experiment.

In Learning:
1. Training: Input samples and target data in the dataset to network. An (NxCxHxW) result will output in which C is the number of classes. The result will input to the loss function to calculate the difference between the current prediction and target prediction. A zero gradient will be created, and the backward propagation will be implemented, and the parameters in the network will be adjusted through the optimizer. This process will be repeated through many epochs to reduce the loss.
2. Validating: Input the data in the validation set and compare the results and target data by the loss function. It will tell the performance of the network for predicting the results.
3. Evaluating: The GenNet will save the model weights by judging the F1 Score and validation loss for each epoch. This model will keep updating as long as one with better performance comes up during the training. At the end of training, the model weights has the least validation loss, the model weights have the highest F1 score and the model weights generated at the last epoch of training will be saved.

In Testing:
1. Input the data in testing set to the trained network with the selected model weights saved during training. The network will output the prediction image, and each pixel in this image contain the value representing the class it is predicted to be.
2. Evaluating the model with confusion matrix.
3. Summarizing the output, target, accuracy, F1 score, intersection-over-union and confusion matrix.
4. Display summary information of accuracy, F1 score and intersection-over-union.
5. Output and save the Confusion matrix.

4 EXPERIMENTS

After integrating the HRNet and HRNet+OCR to the GenNet, the performances of HRNet and HRNet+OCR on hyperspectral data are need to be tested. In this chapter, The plan of experiments and the reason why to execute these will be stated.

4.1 Purposes

1.Test the performance of HRNet and HRNet+OCR on classifying polymers.
2.Test the influence of the different preprocessing methods applied on the dataset.
3.Test the influence of optimizers on performance of HRNet and HRNet+OCR.
4.Test the performance on classifying objects made by the same materials.

4.2 Experiment 1:Test HRNet with different optimizers

In this experiment, HRNet is tested with different optimizer SGD and Adam on two dataset PET-PEF-CP-shaped and Plastics. The purpose of this experiment is to study the influence of optimizer on the performance of the HRNet. In the original code of HRNet, the developers only provide the SGD as the option for training and testing on the dataset Cityscapes used in their project. In the paper of this project, it is mentioned that for dataset Cityscapes, the SGD optimizer improves the performance of the network by 0.6% on panoptic quality compared with Adam optimizer[14]. But from the previous research and project, a fact that Adam optimizer provides better performance on hyperspectral data on F1 score and MIOU for networks Deep Res U-Net and U Net++. Therefore, testing the influence of optimizer becomes a necessary experiment. This experiment is executed with the following parameters :

- Optimizer learning rate:0.001
- Loss Function: Cross Entropy
- Epoches:300
- Test Dataset: PET-PEF-CP-shaped

- Optimizer: SGD, Adam

4.3 Experiment 2:Test different neural networks performance

Reference networks are needed to test and compare with the HRNet to test the performance of the HRNet and whether it is suitable for semantic segmentation on hyperspectral data. The reference networks used in this experiment are Deep Res U-Net and U Net++. This experiment is executed with the following parameters :

- learning rate:0.001,
- Loss Function: Cross Entropy
- Scheduler: Inactivated
- Epoches:300
- Test Dataset: PET-PEF-CP-shaped,Plastics
- Optimizer: SGD

4.4 Experiment 3: Test the influence of different preprocessing methods for HRNet

In the previous project, two different preprocessing methods Logarithmic Derivative and Hyper-Hue [8]are provided and used on a base dataset Plastics. The results indicate that, in some cases, Hyper-Hue provides U-Net and U-Net++ better predictions, and in most cases, the Logarithmic Derivative method provides good results. Therefore, by training the HRNet on the dataset used in the previous project, it can also be found whether different preprocessing methods change the performance of HRNet. This experiment is executed with the following parameters :

- learning rate:0.001,
- Loss Function: Cross Entropy
- Epoches:300
- Test Dataset: Plastics, Plastics-hh, Plastics-hhi, Plastics-hhs
- Optimizer: SGD

4.5 Experiment 4: Test HRNet+OCR with different optimizer

HRNet+OCR is a derived network base on HRNet. The purpose of this experiment is to test the performance of HRNet+OCR to make a comparison with HRNet and test the influence of different optimizers. The reason why perform this experiment is that HRNet+OCR has a better performance on the public semantic segmentation dataset Cityscapes, but the best performance is with the SGD optimizer[14].According to the previous projects, Adam optimizer can improve the performance of Deep Res U-Net and U Net++. Therefore, it is necessary to test the influence of different optimizers and test the performance of the HRNet+OCR. This experiment is executed with the following parameters :

- learning rate:0.001,
- Loss Function: Cross Entropy
- Epoches:300
- Test Dataset: PET-PEF-CP-shaped
- Optimizer: SGD, Adam

The following dataset will be used to train the HRNet. :

- Plastics (including 5 types of polymers, preprocessed by Flat Field Correction and Logarithmic Derivateive)
- Plastics-hh(including 5 types of polymers, preprocessed by Flat Field Correction and Hyper-Hue with result of 223 channel data)
- Plastics-hhi(including 5 types of polymers, preprocessed by Flat Field Correction, and Hyper-Hue + Intensity with result of 224 channels data)
- Plastics-hhs(including 5 types of polymers, preprocessed by Flat Field Correction, and Hyper-Hue + Saturation with result of 224 channels data)
- PET-PEF-PP-shaped(5 types of polymer but included PET,PEF made bottle as 2 additional classes to classify, removed several channels to reduce the influence of noise)

4.6 Experiment products

- Confusion matrixes
- Mean IOU value
- F1 score
- Prediction on testing set and ground truth image

5 RESULTS

First, the GenNet is available to record average accuracy, accuracy for each class, average F1 score, F1 score for each Class, Mean IOU value, Frequency Weighted IOU value are recorded after training is finished. The average F1 score and accuracy are very high at the early epochs of training, which sometimes are more than 0.90 in accuracy and F1 score. But the Mean IOU is very low. In most cases, it's around 0.1 in the early epochs of training. This phenomenon may because the only object that the network can recognize at this moment is the background, and background usually takes most of the area of an image. This phenomenon also happens to the Frequency Weighted IOU value, because of the background takes the most area of an image, and it has the largest weight when calculating.
Figure 10 and 11 show two samples of summary of metrics of HRNet on testing dataset with different model weights from 2 epochs of training and 210 epochs of training. The average F1 score with only 2 epochs of training is higher than the one with 210 epochs of training and is equal to the F1 score of the first class – the background. In some cases, the pixel accuracy starts at around 0.8 after the first epoch of training by changing the learning rate of the optimizer.

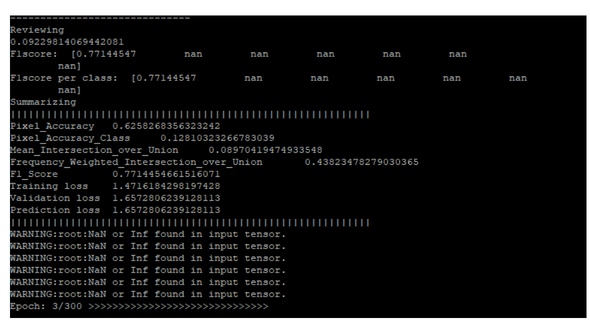
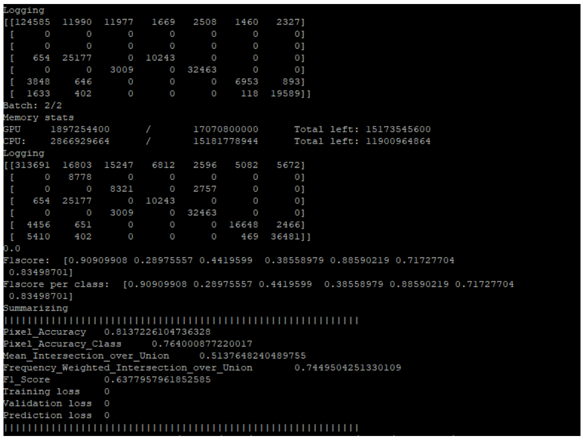


Fig. 10. Summary of metrics results. Use model weights of HRNet after two epochs of training on PET-PEF-CP-shaped dataset and test on testing dataset.



```
Logging
[[124585  11990  11977   1669   2508   1460   2327]
 [     0      0      0      0      0      0      0]
 [     0      0      0      0      0      0      0]
 [   654  25177      0  10243      0      0      0]
 [     0      0   3009      0  32463      0      0]
 [  3848    646      0      0      0   6953    893]
 [  1633    402      0      0      0    118  19589]]
Batch: 2/2
Memory stats
GPU     1897254400      /       17070800000      Total left: 15173545600
CPU:    2866929664      /       15181778944      Total left: 11900964864
Logging
[[313691  16803  15247   6812   2596   5082   5672]
 [     0   8778      0      0      0      0      0]
 [     0      0   8321      0   2757      0      0]
 [   654  25177      0  10243      0      0      0]
 [     0      0   3009      0  32463      0      0]
 [  4456    651      0      0      0  16648   2466]
 [  5410    402      0      0      0    469  36481]]
0.0
F1score:  [0.90909908 0.28975557 0.4419599  0.38558979 0.88590219 0.71727704
 0.83498701]
F1score per class:  [0.90909908 0.28975557 0.4419599  0.38558979 0.88590219 0.71727704
 0.83498701]
Summarizing
Pixel_Accuracy   0.8137226104736328
Pixel_Accuracy_Class     0.764000877220017
Mean_Intersection_over_Union     0.5137648240489755
Frequency_Weighted_Intersection_over_Union      0.7449504251330109
F1_Score         0.6377957961852585
Training loss    0
Validation loss  0
Prediction loss  0
```

Fig. 11. Summary of metrics results. Use model weights of HRNet after 210 epochs of training on PET-PEF-CP-shaped dataset and test on testing dataset.

Second, after performing all the training and record results, it is found that for some experiments, some predictions that use the model weights with the least validation loss don't have a high F1 score and Mean IOU value. Instead, the predictions that use the model weights generated from the last few epochs of training provide a higher score on the F1 score and Mean IOU.
Therefore, in this chapter, the F1 score and Mean IOU value are recorded to compare the assess performance of each experiment. And the F1 score and Mean IOU from the model weights that have the least validation loss and one generated at last epoch of training are both recorded in Table 2 and Table 3.
Notice: The predictions and confusion matrix showed, and the results mentioned in this section are all generated from the network using the model weights that have the least validation loss unless otherwise stated.

Fig. 12. Ground truth image of PET-PEF-CP-shaped dataset.The green is PET, yellow class is PET bottle, red is PEF ,blue is PEF bottle, white 6% co-polymer and the orange is 11% co-polymer

The ground truth of the testing set of PET-PEF-CP-shaped is shown in Figure 12.The ground truth of the testing set of Plastics and its color map to classes is shown in Figure 13. The ground truth image is the same as the ones of the datasets Plastics-hh, Plastics-hhi, and Plastics-hhs.

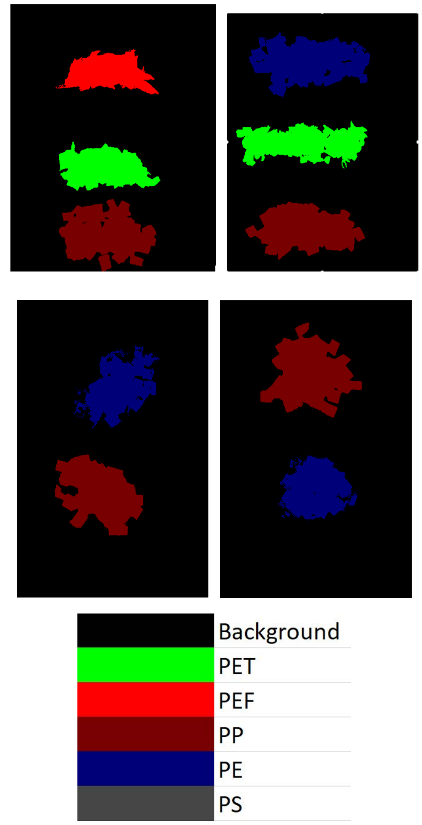


Fig. 13. Ground truth image of testing set of Plastics dataset and the color map of classes.The ground truth image and it's label are the same as Plastics-hh, Plastics-hhi, Plastics-hhs datasets.

Table 2 records all the result from different experiment and the model weights used to predict the testing sets are the model weights have the least validation loss during the training. Table 3 use the model weights generated at the last epoch of training.

Backbone	Dataset	Optimizer	F1 score	Mean IOU
HRNet	PET-PEF-CP-shaped	Adam	**0.63652**	**0.54901**
		SGD	0.64417	0.51320
	Plastics	SGD	**0.84169**	**0.74574**
	Plastics-hh	SGD	0.59438	0.49066
	Plastics-hhi	SGD	0.72035	0.58397
	Plastics-hhs	SGD	0.67420	0.54386
Deep Res U-Net	PET-PEF-CP-shaped	SGD	**0.97260**	**0.94696**
	Plastics	SGD	**0.97626**	**0.95420**
U-Net++	PET-PEF-CP-shaped	SGD	0.80172	0.70572
	Plastics	SGD	0.95744	0.91966
HRNet+OCR	PET-PEF-CP-shaped	Adam	**0.78055**	**0.67304**
		SGD	0.40651	0.18052

Table 2. F1 score and Mean IOU value achieved from all the experiments. In this table, the results come from predication on each dataset's testing set by the different networks with the model weights that have the least validation loss.

Backbone	Dataset	Optimizer	F1 score	Mean IOU
HRNet	PET-PEF-CP-shaped	Adam	**0.73932**	**0.62533**
		SGD	0.52452	0.40599
	Plastics	SGD	**0.78606**	**0.67680**
	Plastics-hh	SGD	0.55069	0.45577
	Plastics-hhi	SGD	0.64424	0.50225
	Plastics-hhs	SGD	0.60139	0.48535
Deep Res U-Net	PET-PEF-CP-shaped	SGD	**0.93498**	**0.88119**
	Plastics	SGD	**0.97439**	**0.95060**
U-Net++	PET-PEF-CP-shaped	SGD	0.64838	0.57765
	Plastics	SGD	0.95159	0.90964
HRNet+OCR	PET-PEF-CP-shaped	Adam	**0.71063**	**0.60509**
		SGD	0.34982	0.14381

Table 3. F1 score and Mean IOU value achieved from all the experiments. In this table, the results come from predication on the testing set of each dataset by the different networks with the model weights that generated at the last epoch of training.

5.1 Experiment 1:Test HRNet with different optimizer

The prediction from this experiment are shown in Figure 14 and 15. In this experiment, the neural network backbone HRNet was trained with different optimizer Adam and SGD. As shown in Table 2,When using the model weights have the least validation loss to predict on the testing dataset, F1 score, and Mean IOU are very similar for both optimizers. HRNet with Adam optimizer has higher Mean IOU value which is 0.54901,0.03579 higher than HRNet with SGD optimizer but also has a lower F1 score which is 0.63652, 0.00765 lower than HRNet with SGD optimizer.
When looking to the results in Table3(last epoch model weights), HRNet with Adam optimizer has a much better result. The F1 score is 0.73932, and the Mean IOU is 0.62533. With the same epoch of training, HRNet with Adam optimizer may tend to have a better performance than with SGD optimizer.

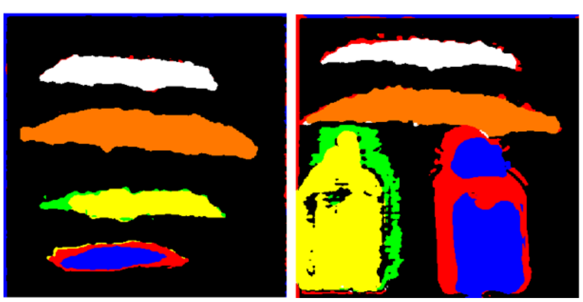


Fig. 14. Prediction on testing set of PET-PEF-CP-shaped dataset from HRNet trained with Adam optimizer.

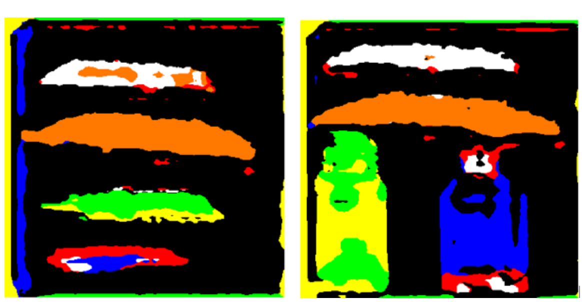


Fig. 15. Prediction on testing set of PET-PEF-CP-shaped dataset from HRNet trained with SGD optimizer.

It is also can be found in the prediction from the HRNet (Figure 14 and 15)and it's confusion matrix(Figure 16and 17) that HRNet will confuse PET with PET bottle and PEF with PEF bottle.

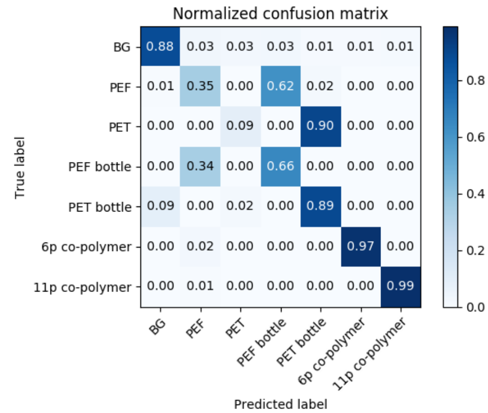


Fig. 16. Confusion matrix of prediction on PET-PEF-CP-shaped testing set. Using HRNet trained with Adam optimizer.

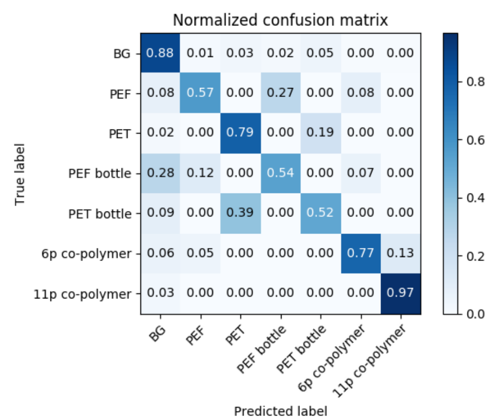


Fig. 17. Confusion matrix of prediction on PET-PEF-CP-shaped testing set. Using HRNet trained with SGD optimizer.

5.2 Experiment 2: Test different neural networks performance

From Table 2 and 3, it is found that Deep Res U-Net has the best performance on both PET-PEF-CP-shaped and Plastics datasets. Its results on the PET-PEF-CP-shaped dataset are F1 score of 0.97260 and the Mean IOU of 0.94696. The highest scores of U-Net++, which are F1 score of 0.80172 and Mean IOU of 0.70572.
The highest scores of HRNet are F1 score of 0.73932, and Mean IOU 0.62533, and the highest scores of HRNet+OCR are F1 score of 0.78055, and Mean IOU of 0.67304 . The highest score of HRNet and HRNet+OCR are achieved when using Adam optimizer.
The prediction from Deep Res U-Net is Figure 18, prediction from U-Net++ is Figure 19 and prediction from HRNet are Figure 14 and 15(From Experiment 1). Compare with the prediction from U-Net++ and from HRNet which is gained in Experiment 1, prediction from Deep Res U-Net has less confusion on flakes and objects made in the same material.

Fig. 18. Prediction on testing set of PET-PEF-CP-shaped dataset from Deep Res U-Net trained with SGD optimizer.

Fig. 19. Prediction on testing set of PET-PEF-CP-shaped dataset from U Net++ trained with SGD optimizer.

For the Plastics dataset, Deep Res U-Net, U-Net++, and HRNet have a similar situation that Deep Res U-Net has the best performance on F1 score and Mean IOU and HRNet has the lowest score on F1 score and Mean IOU.
Figure 20 is the prediction on Plastics dataset from Deep Res U-Net, Figure 22 is prediction from U-Net++ and Figure 21 is from HRNet.

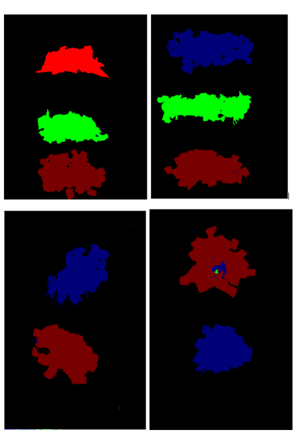


Fig. 20. Prediction on testing set of Plastics dataset from Deep Res U-Net trained with SGD optimizer.

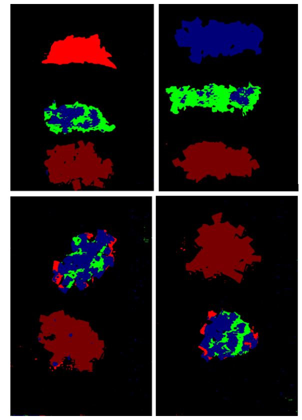


Fig. 21. Prediction on testing set of Plastics dataset from HRNet trained with SGD optimizer.

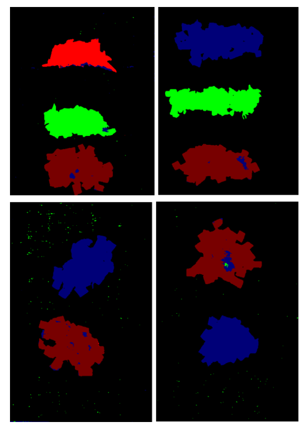


Fig. 22. Prediction on testing set of Plastics dataset from U Net++ trained with SGD optimizer.

Figure 23 is the confusion matrix of prediction on PET-PEF-CP-shaped dataset from U-Net++.

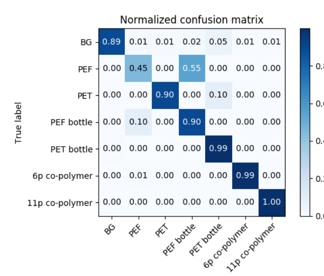


Fig. 23. Confusion matrix of prediction on PET-PEF-CP-shaped testing set. Using U-Net++ trained with SGD optimizer.

It is also found from confusion matrix of U-Net++ Figure 23 and confusion matrix of HRNet Figure 14 have similar phenomenon that they will confuse the PET with PET bottle and PEF with PEF bottle in PET-PEF-CP-shaped dataset.

5.3 Experiment 3: Test the influence of different preprocessing methods for HRNet

From Table 2 and 3, it is clear that HRNet has a higher score on F1 score and Mean IOU for predicting the testing set in Plastics dataset,

which is preprocessed by Logarithmic Derivative. Compare the results from rest datasets preprocessed by Hyper-Hue, the dataset which adds the intensity information achieved a better result than the one only preprocessed by Hyper-Hue and the one preprocessed by Hyper-Hue and added Saturation information.
Comparing the prediction made by HRNet on these four dataset in Figure 21, 24, 25 and 26, it is found that image preprocessed by Hyper-Hue have more noise in the background and is more easy to confuse PE with PET or PEF.

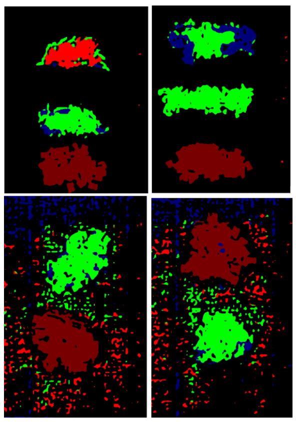


Fig. 24. Prediction on testing set of Plastics-hh dataset from HRNet trained with SGD optimizer.

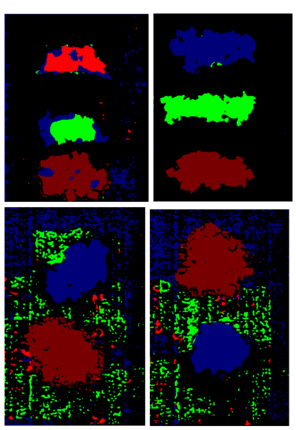


Fig. 25. Prediction on testing set of Plastics-hhi dataset from HRNet trained with SGD optimizer.

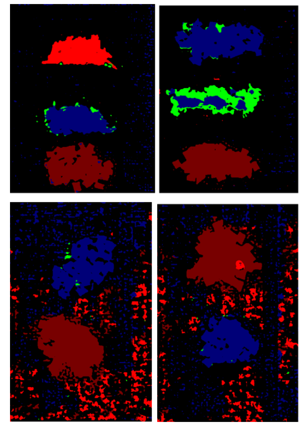


Fig. 26. Prediction on testing set of Plastics-hhs dataset from HRNet trained with SGD optimizer.

This phenomenon can also be found in the confusion matrix for all four dataset predicted by HRNet in Figure 27,28, 29 and 30. For dataset Plastics-hhs, which is preprocessed by Hyper-Hue and added saturation information and dataset Plastics, which is preprocessed by Hyper-Hue and remove both intensity and saturation information, HRNet confuses the PET with PE. With additional intensity information after being preprocessed by Hyper-Hue, the prediction has some improvement but confuses some PEF with PE. It is all shown that for the dataset Plastics which is preprocessed by Logarithmic Derivative, HRNet also confuse the PET with PE.

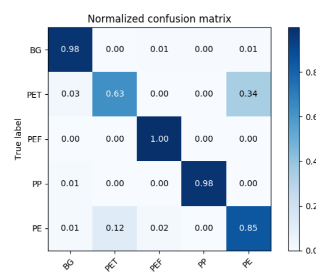


Fig. 27. Confusion matrix of prediction on Plastics testing set. HRNet with SGD optimizer.

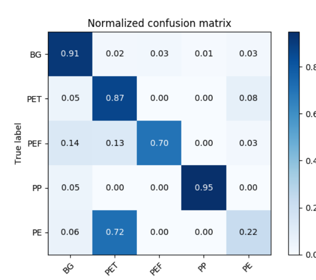


Fig. 28. Confusion matrix of prediction on Plastics-hh testing set. Using HRNet trained with SGD optimizer.

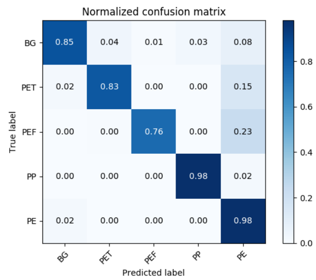


Fig. 29. Confusion matrix of prediction on Plastics-hhi testing set. Using HRNet trained with SGD optimizer.

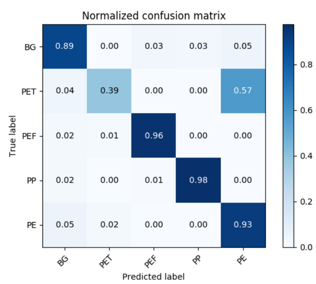


Fig. 30. Confusion matrix of prediction on Plastics-hhs testing set. Using HRNet trained with SGD optimizer.

5.4 Experiment 4: Test HRNet+OCR with different optimizer

By comparing the results shown in Table 2 and 3,the F1 Score and Mean IOU value of HRNet+OCR with Adam Optimizer, which are F1 score of 0.78055 and Mean IOU of 0.67304, have exceeded those of HRNet+OCR with SGD a lot, which are F1 Score of 0.40651 and Mean IOU of 0.18052. HRNet+OCR with Adam has a much better performance. And it is also found that compare to the best results the HRNet has on the PET-PEF-CP-shaped data set, which are F1 score of 0.73832 and Mean IOU of 0.62533 when using Adam optimizer and model weights from last epoch of training, HRNet+OCR with Adam optimizer is also better. HRNet+OCR lead HRNet around 0.05 both for F1 score and Mean IOU.
. Figure 31 and 32 show the prediction made by HRNet+OCR with Adam optimizer and SGD optimizer separately.The HRNet+OCR can recognize bottle object more and construct the object more intergrally.

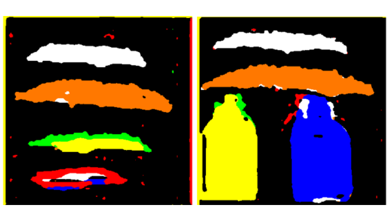


Fig. 31. Prediction on testing set of Plastics-hhs dataset from HRNet+OCR trained with Adam optimizer.

Fig. 32. Prediction on testing set of Plastics-hhs dataset from HRNet+OCR trained with SGD optimizer.

Figure 33 and 34 are the confusion matrix from HRNet+OCR with Adam and SGD. It also shows that HRNet+OCR still will confuse PET with PET bottle and PEF with PEF bottle.

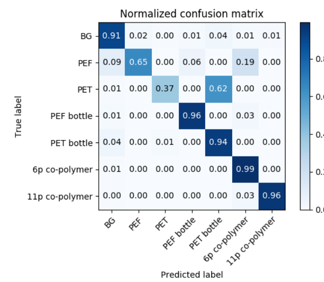


Fig. 33. Confusion matrix of prediction on PET-PEF-CP-shaped testing set. Using HRNet+OCR trained with Adam optimizer.

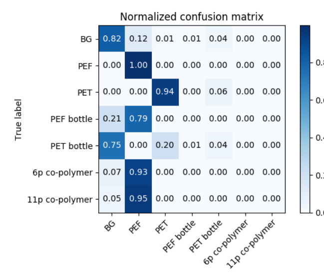


Fig. 34. Confusion matrix of prediction on PET-PEF-CP-shaped testing set. Using HRNet+OCR trained with SGD optimizer.

6 CONCLUSION AND DISCUSSION

According to Experiment 1 and Experiment 4, HRNet and HRNet+OCR with Adam optimizer have better performance on the PET-PEF-CP-shaped dataset and Plastics dataset. This phenomenon indicates that for HRNet, compare to the SGD optimizer which is the only Optimizer provided and limited to use in the original code, Adam Optimizer provides HRNet and HRNet+OCR better performance on hyperspectral data.
In experiment 2, by comparing the confusion matrix, F1 score, and Mean IOU value on the testing set from Deep Res U-Net, HRNet, and U-Net++. It is found that Deep Res U-Net has the best performance on both PET-PEF-CP-shaped and Plastics dataset and HRNet have the lowest score on both dataset. In the test performed on the PET-PEF-CP-shaped dataset, it is founded that U-Net++ and HRNet have a very similar phenomenon that they can hardly distinguish the flakes and object made in the same material. But this isn't a problem for Deep Res U-Net.

In Experiment 3, by training HRNet on the same set of samples preprocessed by different preprocessing methods, the HRNet shows its the best performance on the F1 score and Mean IOU on the dataset preprocessed by Logarithmic Derivative instead of dataset preprocessed by Hyper-Hue. Another thing that also is found that by comparing the confusion matrixes from Plastics-hh and Plastics-hhs, it is found the HRNet is hard to distinguish PE and PET. But for Plastics-hhi dataset, this phenomenon has a little improvement which indicates that intensity information may be important for HRNet to learn the hyperspectral data.

In Experiment 4, HRNet+OCR is trained on both PET-PEF-CP-shaped and Plastics datasets with two different Optimizer SGD and Adam. The result is similar to Experiment 1, that Adam optimizer improves the performance of HRNet+OCR. And the performance of HRNet+OCR with Adam Optimizer is better than HRNet with Adam Optimizer on PET-PEF-CP-shaped dataset.
Therefore, answers to the research questions and the conclusion is as follows:
A neural network called HRNet and its derivative network backbone HRNet+OCR is available to integrate into the GenNet without much change on Dataloader, Loss function, or Optimizer in the current structure. The HRNet+OCR backbone has a good result on the original project[14], which is performed on a public dataset Cityscapes with 84.5% IOU list at the top of the billboard for semantic segmentation job on this dataset. But the HRNet and HRNet+OCR cannot provide better performance on learning and predicting hyperspectral data than Deep Res U-Net which is already integrated to the GenNet.
Some reasons may cause this result:
First, the structure of HRNet is by fuse four different predictions in four different resolutions into one final prediction by a 1x1 convolution layer. This operation may magnify the noise and error in each prediction.
Second, the HRNet and U-Net++ have a similar phenomenon that they confuse the objects made in the same materials. This indicates that HRNet can learn hyperspectral data, but from the confusion matrixes come from the HRNet, it may prefer to classify plastics flake as an object – a bottle made by the same material. This may mean that the structure of HRNet and U-Net++ takes more weight on the two-dimensional information such as edges, angles, and channel of image that contain the information of colors. Even the network has learned the feature of plastic material through channels containing hyperspectral information. It seems that it cannot combine the hyperspectral features with the two-dimensional features, and it prefers to use the two-dimensional features to predict the image. This may also explain that why prediction on dataset preprocessed by Hyper-Hue and add additional intensity information achieves little improvement in Table2 and 3.
Also, the function of OCR is to assign the class label only to the pixel that belongs to the object in the same category. This is to reduce the chance to label a pixel in the background or in other objects as a wrong class. Therefore, this function enhances the two-dimensional features by judging whether a pixel is in an object.
Another point is that the choice of Optimizer may influence a lots on how a neural network learns hyperspectral data. This phenomenon is exceptionally distinct on HRNet+OCR backbone which has enhanced the ability to learn both two-dimensional features and hyperspectral features.
For the future project, if another neural network backbone is needed, the network derives from or combines with the DeepLabV3 network, or the DeepLab series network could be a good way to search for. Because of the Deep Res U-Net has a much better performance than U-Net++ and HRNet and it's performance was also analyze in the previous project[29]. And also the neural network made specially to learning hyperspectral data are another good way to look cause of currently the neural network specially made for hyperspectral data are only a few and coded in Tensorflow instead of PyTorch such as HybirdSN which can also be found on paperwithcode.com.
In this project, due to the situation, bio-polymers are not tested. Whether bio-polymer can also be classified is still an uncertain subject need to be studied.

REFERENCES

[1] J. R. Jambeck. K. L. Law R. Geyer. Production, use, and fate of all plastics ever made., 2017.
[2] WORLD ECONOMIC FORUM. The new plastics economy, rethinking the future of plastics., 2016.
[3] British Plastics Federation. Plastics additives., 2020.
[4] Micro-plastics are entering the human body and are dangerous to human health., 2019.
[5] European Bioplastics e.V. What are bioplastics?, 2020.
[6] Johannes Becker Lars B¨orger Jens Hamprecht Sebastian Koltzenburg Robert Loos Michael Bernhard Schick Katharina Schlegel Carsten Sinkel Gabriel Skupin Andreas K¨unkel and Motonori Yamamoto. Polymers, biodegradable., 2016.
[7] Ellen Macarthur Foundation. Eleven companies take major step towards a new plastics economy, 2018.
[8] Carlos Alejandro Guerrero Martinez. Polymer segmentation through hyper-spectral imaging and convolutional neural networks, 2019.
[9] Neural networks:advantages and applications, April 18,2019.
[10] Yu-Hsiu Lin. Chia-Ching Kung Ming-Han Chung I-Hsuan Yen Yung-Yao Chen. Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart homes., 2019.
[11] Nagornov N.N.-Lyakhov P.A. Valuev-G.V. Chervyakov N.I. Valueva, M.V. Application of the residue number system to reduce hardware costs of the convolutional neural network implementation. 2020.
[12] Ays¸egu¨l Uc¸ar O¨ zal Yıldırım Yakup demir Musab Cos¸kun. Face recognition based on convolutional neural network, 2017.
[13] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, Wenyu Liu, and Bin Xiao. Deep high-resolution representation learning for visual recognition. *TPAMI*, 2019.
[14] Ke Sun, Bin Xiao, Dong Liu, and Jingdong Wang. Deep high-resolution representation learning for human pose estimation. In *CVPR*, 2019.
[15] Bart Riesebos. Pet and pef classification using hyper-spectral imaging and semantic segmentation networks, 2019.
[16] N.MacKinnon. D.L.Farkas F.Vasefi. Hyperspectral and multispectral imaging in dermatology, 2016.
[17] Yijing Xie. Eli Nabavi Robert Bradford Shakeel R Saeed Sebastien Ourselin-Tom Vercauteren Jonathan Shapey. Intraoperative multispectral and hyperspectral label-free imaging: A systematic review of in vivo clinical studies., 2019.
[18] Jeffrey Gotro. Recycle and disposal of plastic food packaging waste 8: More about waste plastic sorting technologies., 2017.
[19] Semantic segmentation-popular architectures., 2019.
[20] A. Jand. Identify plastics using deep learning and hyperspectral imaging, 2017.
[21] A. Stellingwerf and J. Hu. Description of polymer classification by applying hyperspectral imaging and deep learning techniques, 2017.
[22] H. Kloosterman. Hyperspectral band selection, 2018.
[23] P. Ghamisi G. Licciardi B. Rasti, P. Scheunders and J. Chanussot. Noise reduction in hyperspectral imagery: Overview and application. 2018.
[24] Leia explains: What is the effect of flat field correction and why is it absolutely essential for the use of line scan cameras?, 2020.
[25] Frédéric Itthirad Mohamed Bouabdellah Brigitte Closs Michel Jourlin, Josselin Breugnot. Logarithmic image processing for color images, 2014.
[26] S.-H. Lee H. Liu and J. S. Chahl. Transformation of a high-dimensional color space for material classification. 2017.
[27] Armin Haeberle Sulaiman Vesal Vincent Christlein Andreas Maier Christian Riess AmirAbbas Davari, Nikolaos Sakaltras. Hyper-hue and emap on hyperspectral images for supervised layer decomposition of old master drawings, 2018.
[28] Helena Kloosterman. Hyperspectral band selection, 2019.
[29] Bart Riesebos. Pet and pef classification using hyper-spectral imaging and semantic segmentation networks, 2019.
[30] Loss and loss functions for training deep learning neural networks, 2019.
[31] Jingdong Wang Yuhui Yuan, Xilin Chen. Object-contextual representations for semantic segmentation, 2019.
[32] Accuracy, precision, recall —& f1 score: Interpretation of performance measures, 2016.
[33] Intersection over union (iou) for object detection, 2016.

6.1 APPENDIX

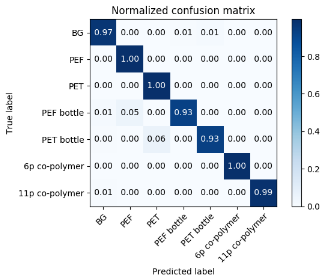


Fig. 35. Confusion matrix of prediction on PET-PEF-CP-shaped testing set.Using Deep Res U-Net trained with SGD optimizer.

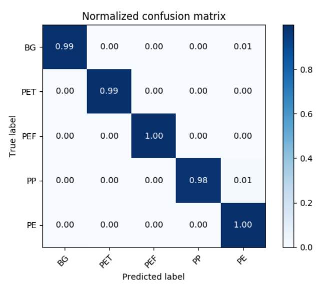


Fig. 36. Confusion matrix of prediction on Plastics testing set. Using Deep Res U-Net trained with SGD optimizer.

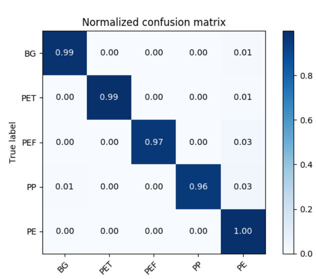


Fig. 37. Confusion matrix of prediction on Plastics testing set. Using U-Net++ trained with SGD optimizer.