Chapter 4

Data Acquisition

The following chapter presents the data acquisition process of our project. This aspect of the solution is very important, due to the fact that most of the work is going to revolve around the images we take. At first, we detail the data acquisition process starting by describing the used tools and the acquired datasets. Then, we delve into the details behind the data preparation.

4.1 Data Acquisition Process

We started the project with no images, and it is our mission to make the acquisition of the dataset; which is always the basic core in a computer vision project. As a consequence, it is important to have a substantial understanding how the color 2D, infrared and 3d images should be taken, the devices and applications to be used.

4.1.1 Data Requirements

In order to successfully merge 3D images, 2D color images and infrared images, we need to take three images of each of the previous image types each time. We are looking for an image having a high number of net lettuces. Images should be taken by the mobile with a position parallel to the ground to decrease dis-torsion.

4.1.2 Data Acquisition Tools

In this subsection, we will describe the characteristics of the devices and applications used to take the three types of images.

4.1.2.1 Galaxy Note 10 Plus

Galaxy Note 10 plus, as shown in the figure n°4.1, has been chosen for its high-resolution cameras compared to the resolution of phone cameras existing in the market. Its camera resolution is equal to 4032x3024 pixels and its pixel size is 14 µm. It contains 3 rear cameras:

- **16 MP Ultra Grand Angle:** widens the field of view from 55° to 123°(corresponds to the minimum zoom), we are talking about focal lengths from about 10 mm to 18 mm
- **12 MP Grand Angle:** allows you to crop and close with optical zoom with focal length larger than 50 mm, as a result a large part of the image is easily sharp. Automatically adjusts to the surrounding light: the f/1.5 aperture allows to bring light to dark places, the f/2.4 aperture allows you to keep a sharp image and the Night mode illuminates scenes that are really very dark.
- **12 MP Telephoto:** or long focal length lens, brings objects closer.
- **ToF (Time of Flight) camera:** or DepthVision camera allows to measure objects and distances in live and augmented reality.

Figure 4.1: Galaxy Note 10 Plus

4.1.2.2 Scanner 3D Application

Galaxy Note 10 plus enables its users to scan objects for instant 3D rendering. The 3D Scanner app includes a circular on-screen guide, asking users to keep the item in this area of the phone and helping them point the camera in the optimal direction for a 3D scan. Once the user scans a full circle around the object and captures all angles, the phone produces a 3D model of the object as shown in the figure n°4.2.

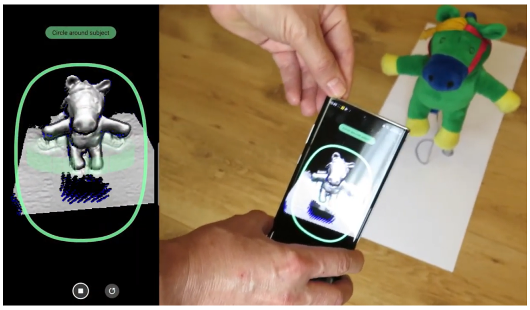

Figure 4.2: 3D Scanner Application

4.1.2.3 Flir One PRO (Android - USB C)

Flir One PRO thermal camera for Android USB C devices, as shown in the figure n°4.4, has thermal and visual cameras with 1440 x 1080 pixel visual resolution. Its pixel size and spectral region are respectively equal to 12 µm and 8 - 14 µm. The Flir One PRO measures temperatures between -20°C and 400°C. This device has been chosen as it connects to an Android smartphone as illustrated in the figure n°4.3 and turns it into a thermal camera and as it provides high quality thermal image.

Figure 4.3: Flir One PRO Connected to Galaxy Note 10 Plus

Figure 4.4: Flir One PRO

4.1.3 Data Acquisition

- We tested using two sticks for the acquisition of images and without stick which one fits our requirements before going to the fields. The stick n°1, shown in the figure 4.5, allows to take images in a parallel position with the ground. However, it is not practical to use it in fields and greenhouses as it is heavy and should be replaced and fixed each time to capture other planted lettuces. We tried to take images without using stick ,as shown in the figure n°4.6, and found out that it is easier than with stick, but it is difficult to avoid the shadow of the user in the images as shown in figure 4.7. In contrast, the stick n°2 enables to avoid the shadow and is easy to use in the fields.

Figure 4.5: Stick N°1

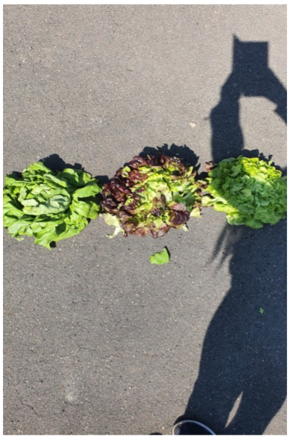

Figure 4.6: Without Stick

Figure 4.7: Stick N°2

- Images will be taken with the wide angle camera and photo mode when it indicates that this is the best shot. Images contain the maximum of net lettuces. We find out that the best shot is taken at one meter distance.
- When we moved to the greenhouses of Limagrain producers in Angers, France, who plant lettuce, we discovered that the 3D scanner application could not create the 3D model because it does not detect the object. One of the reasons the overlapping of the lettuces as shown in figures n°4.8 and n°4.9. An reason is that in some greenhouses there is a tarpaulin on the ground that obstruct the object detection as shown in figure n°4.10. As a result, we focused only on taking infrared and color images.

Figure 4.8: Overlapping Iceberg Plants

Figure 4.9: Overlapping Batavia Plants

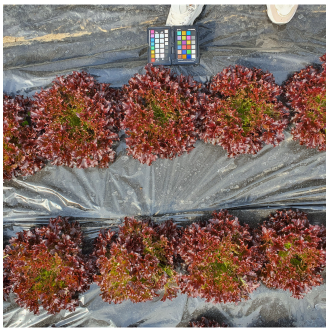

Figure 4.10: Tarpaulin the Ground

4.2 Data Preprocessing

Since there is an offset between the thermal camera in Flir one PRO and color camera in Galaxy Note 10 plus as shown in figure n°4.11, there is a shift between the thermal image and color image as can be seen in figures n°4.12 and n°4.13. However, we should have pair images with the same aligned geometry to have a better merging

Figure 4.11: Used Device

Figure 4.12: Raw RGB Image

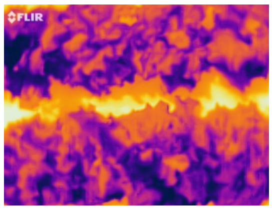

Figure 4.13: Raw IR Image

First of all, in order to avoid quality loss, images are converted to PNG format. Then with the aim for having corrected shift, we tried to find out the shift using scikit-image library as shown in the code in the figure n°4.14.

```
RGB = cv2.imread("C:/Users/Chayma.MOUSSA/Desktop/IR_1.jpg", 1)
RGB_resized = cv2.resize(RGB, (1080,1440 ))
plt.imshow(RGB_resized)

IR = cv2.imread("C:/Users/Chayma.MOUSSA/Desktop/IR_3.jpg", 1)
IR_gray = cv2.cvtColor(IR, cv2.COLOR_BGR2GRAY)

#register the translation between the IR image and the RGB image
shift, error, diffphase = register_translation(IR_gray, RGB_resized)

print(shift)
print(error)
print(diffphase)
#apply the offset to the RGB image
offset_image = fourier_shift(np.fft.fftn(RGB_resized), shift)
offset_image = np.fft.ifftn(offset_image)
plt.imshow(offset_image.real)
```

Figure 4.14: Correcting the shift with scikit-image library

However, it does not return the same aligned geometry images. Later, we opt for an online tool named *overlay.imageonline* that gives better aligned images as indicated in figures n°4.15 and n°4.16. Later, images are resized to the same size.

Figure 4.15: Raw RGB Image

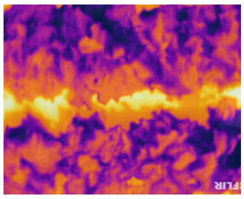

Figure 4.16: Raw IR Image

4.3 Data Preparation

Data preparation procedure for large data sets is an essential process for our project. Starting by data annotation and data augmentation, we will detail later on the description of the data sets.

4.3.1 Data Annotation

CVAT assists in the key guided computer vision tasks of object identification, image categorization, and image segmentation. Users can annotate data for each of these scenarios using CVAT as shown in the figure n°.17. Annotations can also be generated in different formats, such as MS COCO format, YOLO, or PascalVOC detailing the job and the annotations.

The introductory phase in data preparation is to annotate all the pictures. Those annotations acts like an information base for the model learning, validation and test phases. The model will point to come as near to all these annotations as possible. As a consequence, it's critical to engender correct identifiers. This section clarifies the Computer Vision Annotation Tool for the annotation work. This Computer Vision Annotation Tool(CVAT) is being utilized to name information for computer vision applications.

CVAT facilitate the key guided computer vision tasks of object identification and image categorization. Users are able to annotate their datasets for each of these scenarios utilizing CVAT. Annotations can also be produced totally in various formats, for example a dataset YOLO, COCO , or PascalVOC specifying the work and the annotations.

We choose to annotate distinctive images consequently we did not annotate all images in the dataset as shown in tables n°4.2, 4.3, 4.4, 4.9, 4.11, 4.13; the number of annotated images is different from the total number of images.

In our case, we have multiple objects in each image. Some of the images have around 6 to 10 lettuces. In 369 images we have 2298 lettuce objects as indicated in table n°4.8.

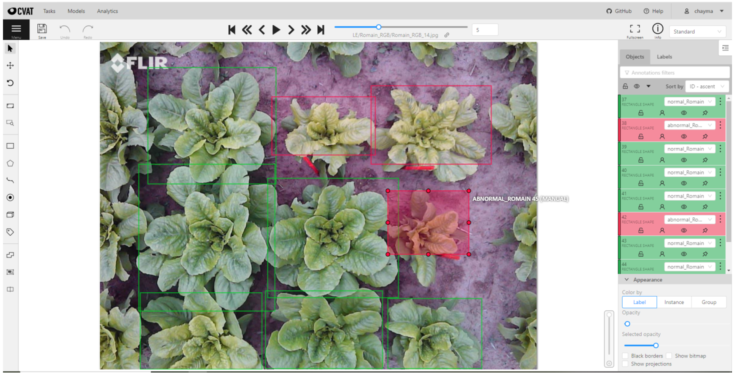

Figure 4.17: CVAT Annotation Software Process

4.3.2 Data Augmentation

After annotating the data sets, we split our data into training (70%), validation (20%) and testing (10%) dataset. Later, we will apply the data augmentation techniques for the training set. Actually, data augmentation is an effective method in deep learning and particularly when dealing with pictures. This technique upgrades the execution of the model by artificially expanding the number of examples and operates as a regulation device by randomising the pictures at each stage which confines the model from retaining the pictures and in this way overfitting. For each picture in the training kit, we use the following conversions for each epoch:

1. **Rotating:** we rotate the pictures with an angle of 90° in different directions; clockwise, counter-clockwise, upside down. This improves the performance of the model to be rotational constant since the real application pictures may be rotated in different directions.

2. **Flipping:** we flip the picture along the vertical, horizontal and in both axes. This is a supplement way to include a rotational invariant element to the model in this instance.

3. **Saturation augmentation:** is similar to hue except that it adjusts how vibrant the image is. A fully desaturated image is grayscale, partially desaturated has muted colors, and a positive saturation shifts colors more towards the primary colors. Adjusting the saturation of an image can help model perform better when colors in the wild are different (for example, if a different white-balance is set, different lighting is in effect, or even if it's foggy outside). It is comparable to tint but that it alters how dynamic the picture may be. A completely desaturated picture is grayscale, mostly desaturated has quieted colors, and a positive immersion shifts colors more towards the essential colors. Adjusting the immersion of a picture can help when colors within the wild are distinctive (for case, in case a diverse white-balance is set, diverse lighting is in impact, or indeed in case it's foggy exterior).

4.3.3 Dataset Description

The acquired dataset is formed of images of six typologies of lettuce taken in different environments and times and under uncontrolled conditions; inside greenhouses and outside. This database also includes a color reference ; important device is employed in image processing; e.g, the color reference is always utilized to assess the color appearance. In the first acquisition seances, the color images were taken by the Galaxy Note 10 plus, but later we decided to take them by the visual camera of the Flir one PRO. In this case, we avoid the shift issue between the pair images.

The lettuce producers send us color images including different lettuce typologies, which we added to our data set taken by Samsung Galaxy A51 and Samsung A3. We have as a result a color raw image dataset, named D1, taken by four different devices and taken under uncontrolled conditions as schematized in the figure n°4.18.

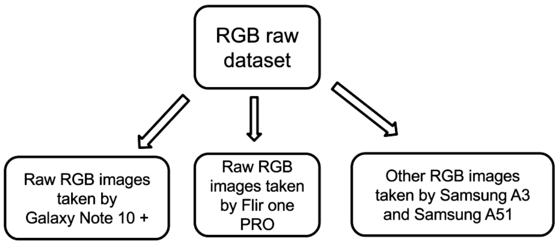

Figure 4.18: Acquisition of D1

Table n°4.1 provides an overview of the raw RGB data set of the different lettuce typologies. Global data set includes 370 pairs of raw images and 415 of color images.

For all raw color images, we applied on it the three data augmentation techniques for every lettuce typology as shown in the tables n°4.2, 4.3, 4.4, 4.5, 4.6, 4.7. We grouped all the typologies together as indicated in the table n°4.8.

The data sets that are going to be fused and RGB data set that is going to be compared with the fused data sets , named D2, include the corrected shift pair images and the pair images taken by Flir one PRO as shown in the figure n°4.19.

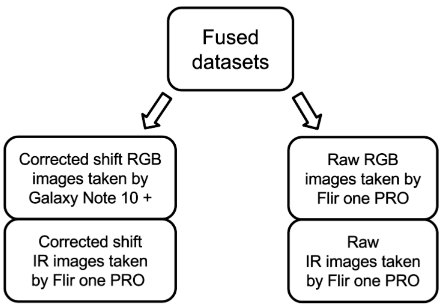

Figure 4.19: Acquisition of D2

We applied only rotation and flip as data augmentation techniques. In fact, these data sets will be used for merging the pair images and if we apply saturation, this will affect the quality of the fused images and we cannot find out if the classification with color image is better than the fused images or not. Tables from n°4.9 to n°4.14 illustrate detailed data overviews of each lettuce typology used for fusion infrared and color images. Table 4.15 show an overview of all lettuce typologies data set used for the fusion.

As the number of abnormals lettuce is less than normal lettuces as illustrated in the detailed data overview tables of all the typologies; the number of abnormals are higher than the normals, as a result we have an unbalanced data set.

Conclusion

In this chapter, we outlined in details all the steps in the data acquisition and data preparation process. In the next chapter, we will describe the modeling step and then we will show the results obtained.

Table 4.1: Dataset

Typology	Images with RGB images taken by Galaxy Note 10+		Images with RGB images taken by Flir One Pro		Other RGB images
	RGB	IR	RGB	IR	
Batavia	59	59	57	57	32
Multifeuille	83	83	10	10	4
Beurre	0	0	63	63	0
Chene	0	0	62	62	0
Romain	12	12	5	5	1
Iceberg	8	8	11	11	8

Table 4.2: Detailed Data Overview of Batavia

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images	148	257	303
Images annotated	108	257	303
Objects	623	1726	2594
Normals	510	1488	2219
Abnormals	113	238	375

Table 4.3: Detailed Data Overview of Multifeuille

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images count	97	203	233
Images annotated	92	188	218
Objects	754	1333	1860
Normals	683	1233	1460
Abnormals	71	100	140

Table 4.4: Detailed Data Overview of Beurre

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images	64	139	154
Images annotated	62	137	152
Objects	382	708	800
Normals	272	501	558
Abnormals	110	207	242

Table 4.5: Detailed Data Overview of Chene

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images	63	133	151
Images annotated	63	133	151
Objects	300	787	895
Normals	240	637	720
Abnormals	60	150	175

Table 4.6: Detailed Data Overview of Romaine

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images	17	38	41
Images annotated	17	38	41
Objects	121	247	314
Normals	107	211	256
Abnormals	14	36	58

Table 4.7: Detailed Data Overview of Iceberg

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images	27	60	65
Images annotated	27	60	65
Objects	118	378	428
Normals	109	348	389
Abnormals	9	30	39

Table 4.8: Detailed Data Overview of All Typologies[+Experts Images]

	Without data augmentation	With data augmentation	
		Rotation & Flip	Rotation & Flip & Saturation
Total images	416	930	994
Images annotated	369	884	948
Objects	2298	5179	6564
Normals	1921	4418	5557
Abnormals	377	761	1007

Table 4.9: Detailed Data Overview of Beurre

	Without data augmentation	With data augmentation: Rotation & flip
Total images	64	139
Images annotated	62	137
Objects	382	708
Normals	272	501
Abnormals	110	207

Table 4.10: Detailed Data Overview of Chene

	Without data augmentation	With data augmentation: Rotation & Flip
Total images	63	133
Images annotated	63	133
Objects	300	787
Normals	240	637
Abnormals	60	150

Table 4.11: Detailed Data Overview of Batavia

	Without data augmentation	With data augmentation: Rotation & Flip
Total images	74	165
Images annotated	72	163
Objects	275	780
Normals	233	660
Abnormals	42	120

Table 4.12: Detailed Data Overview of Multifeuille

	Without data augmentation	With data augmentation: Rotation & Flip
Total images	83	195
Images annotated	83	195
Objects	400	1129
Normals	355	1029
Abnormals	45	100

Table 4.13: Detailed Data Overview of Romaine

	Without data augmentation	With data augmentation: Rotation & Flip
Total images	12	25
Images annotated	11	24
Objects	49	145
Normals	39	117
Abnormals	10	28

Table 4.14: Detailed Data Overview of Iceberg

	Without data augmentation	With data augmentation: Rotation & Flip
Total images	19	61
Images annotated	19	61
Objects	96	260
Normals	80	216
Abnormals	16	44

Table 4.15: Detailed Data Overview of All Typologies

	Without data augmentation	With data augmentation: Rotation & Flip
Total images	317	707
Images annotated	314	701
Objects	1365	3709
Normals	1119	3042
Abnormals	246	667