BananaGAN : Augmenting major banana disease detection using generated diseased pseudostem and rachis images

Javier Alejandro Vergara Zorrilla

Acceptance mark: 4.6

We certify that this Degree Project satisfies, in scope and quality, all the requirements demanded by a Master's Degree Project.

Animesh Acharjee

Dr. Animesh Acharjee
Director

Gloria Inés Alvarez V.

Dr. Gloria Inez Alvarez
Jury

Dr. Luis Eduardo Tobon
Jury

Approved in compliance with the requirements of the Pontificia Universidad Javeriana Cali, to qualify for the Master's degree in Engineering with an emphasis on Electronics.

Camilo Rocha.

HERNÁN CAMILO ROCHA NIÑO Ph. D.
Dean of the Faculty of Engineering and Sciences

JUAN CARLOS MARTÍNEZ ARIAS
Director of postgraduate Engineering and Sciences

Santiago de Cali , 03/03/2021

Maestría en Ingeniería
Facultad de Ingeniería y Ciencias

Pontificia Universidad
JAVERIANA
Cali

Act of corrections, Thesis document

Santiago de Cali, 03/03/2021

Author:
Javier Alejandro Vergara Zorrilla

Título del Trabajo de Grado: BananaGAN : Augmenting major banana disease detection using generated diseased pseudostem and rachis images

Director: Dr. Animesh Acharjee

As indicated in article 2.13 of the Guidelines for Master's Degree Work, I have verified that the student indicated above has implemented all the corrections that the Juries of the Degree Work Project defined to be carried out, as stated in the corresponding Evaluation Act.

Animesh Acharjee

____

Firma del Director del Trabajo de Grado

[redacted]

Research Proposal

BananaGAN : Augmenting major banana disease detection using generated diseased pseudostem and rachis images

Author:
Javier Alejandro Vergara

Director:
Animesh Acharjee, PhD
Co-Director:
Michael Selvaraj, PhD

Electronics and computer sciences department
Pontificia Universidad Javeriana Cali

March 3, 2021

Contents

List of Figures

List of Tables

Acknowledgements

The author acknowledge the cooperation of all the members of the **Bioversity-CIAT alliance**. The author also thanks all the institutions that supported this work and the people who with their hard work and dedication made the data collection possible in the field as well as the guidance of the scientific directors of this project. Finally to my whole family and my wife for supporting all times.

Special Terms

ADP Average detection percentage. 33

AI Artificial intelligence. 3

API Application programming interface. 35

BBD Banana blood disease. 23

BBTV Banana bunchy top virus. 6

BXW Banana xanthomonas wilt. 6

CNN Convolutional neuronal network. 13

COCO Common objects in context. 29

CSI Cross synthetic images. 31

D Discriminator. 25

DL Deep learning. 13

Faster R-CNN faster regions with convolutional neural network. 29

FP False positives. 29

FWB Fusarium wilt of banana. 6

G Generator. 25

GAN Generative adversarial network. 3

GPU Graphics process unit. 18

GSMA Global System for Mobile Association. 6

HP Healthy plant. 31

HW Hardware. 30

IOT Internet of things. 3

IoU Intersection over union. 30

mAP mean Average precision. 14, 29

PCR Polymerase chain reaction. 5

SIG Synthetic images generator. 31

SSD Single shot detector. 14

SW Software. 30

TP True positive. 29

Abstract

Early detection and timely management of crop diseases are essential for reducing yield loss. Traditional manual inspection is often time-consuming, laborious, and biased. Recently, automated imaging techniques have been successfully applied to the detection of crop diseases. Almost this type of research requires a huge amount of images with key typical symptoms from rare classes. The rare class images are the key to differentiated closely related diseased symptoms, but it is mostly internal and difficult to get them. Thus we exploited generative adversarial networks for generating rare classes such as banana pseudostem and rachis images creating new datasets with synthetic images and doing domain disease translation, converting an image with a certain disease into another image with another different disease. These synthetic images were tested in pre-trained disease detection models to see if they are good enough to balance the banana disease datasets and improve the object detection models' overall accuracy and can be applied to other deep learning techniques such as classification and semantic segmentation. mAP score from the trained models with synthetic images was between 64% and 89% accuracy, which conclude that synthetic images are a useful tool as a data augmentation technique.

Keywords— Artificial intelligence, generative adversarial networks, deep learning, disease detection, data augmentation, pseudostem, rachis, synthetic data

Chapter 1

Introduction

In agriculture, there are several environmental threats as drought, pest, and diseases, nutrient deficiency, etc. that affects the crops, reduce their yield and decrease food production. Farmers around the world have the challenge to face those problems, but they do not have enough tools to do it because for years there has been a gap between technology and agriculture. Some few organizations have the power and the technology to make its crops resistant to many threats, but 85% of the world's farms are in smallholder farmers hands[2]. Smallholder farmers rely on empirical knowledge, which is less effective in this overcoming farming challenges[3], technology must be closer to them and enrich their knowledge to improve their food production.

One good example is Banana, it is one of the most important fruit crops in the world in terms of production volume and trade[4]. Only a low percentage of bananas produced are globally traded[5], most of the production is for domestic markets. This crop is a substantial dietary component [6] due to the nutritional facts. Especially in tropics countries, this crop is vital to daily survival, one good example is central Africa, where they eat more bananas than anyone in the world. In some Africa countries, banana represents almost 35% percent of daily calories. The banana crop is also important when it is seeded under mixed complex systems (with cassava or coffee) like in the democratic republic of congo, Benin republic, Colombia, and Panama. In some Latin American countries as Colombia, Panama, and Costa Rica, many typical dishes are made from this fruit, and it is commonly included daily in any meal of the day. But as mentioned several pests and diseases affect banana crops causing significant yield losses[7]. Most of these diseases are presented as physical symptoms specially in leaves, if this is not controlled, pests and diseases will affect the whole crop and even the surrounding farms. An early detection system is needed to help farmers to manage those threats. Also, there are many concerns about the change and the increase of those problems due to climate change[8][9]. To help not only banana farmers but all crops and to close the gap between the agriculture and technology, different areas in science are called to create and develop useful and low-cost technological tools applied to agriculture and easy to use for smallholder farmers to help them with all those threats. Some examples are mobile applications [10] , crop monitoring[11][12], Internet of things(IOT) [13] , remote sensing[14] and artificial intelligence [15]. Artificial intelligence(AI) and specifically machine learning and deep learning applications are helping in disease detection in humans [16], [17] , [18] but also in plants [19] and combined with mobile applications are becoming into great tools in crop diagnosis and management[20]. Deep learning models can be trained with different sources of data, but in diseases diagnosis image based[19] models are used since diseases symptoms are clearly observable. The major problems in machine learning and specially in deep learning applications are the source of information, due to these networks require a massive amount of data to be trained and perform a good generalization, is not easy to obtain a good database with reliable and enough data. As disease diagnosis with deep learning use images, there are some traditional methods to do data augmentation such as flipping or mirroring the images [21], but new technologies are coming to generate artificial images with an approach called generative adversarial networks(GAN)[22] [23] , with these models can be generated artificial data almost real and makes impossible to differentiate between a real and an artificial image for human been[24].

Generative Adversarial Networks (GAN) was mentioned initially by Ian Goodfellow in the international conference on neural information processing Systems in Montreal [25]. GAN uses adversarial methods where a generator and a discriminator compete with each other to learns the data distribution and finally data generation. Due to the good results showed by GAN models and the usability in several different applications, have become one of the most studied technologies in recent years. GAN has proven to be able to generate good quality synthetic data in several applications [26], and can be used to solve well known image processing approaches like image to image translation or domain translation and finally data augmentation. Data augmentation with traditional and new techniques are giving solutions to poor data problems and helps to increase the accuracy of the trained models[27] but is very limited and will lead machine learning models to the overfitting problems. We explored GAN's application to generate synthetic data based on real field images collected from hot spots and solve problems such as fewer labels and unbalanced classifications. To develop this study, context and planning is organized as follows:

- Chapter 1: Introduction.

- Chapter 2: Problem formulation.
- Chapter 3 and 4: Scope and objectives.
- Chapter 5: Justification.
- Chapter 6: Theoretical framework.
- Chapter 7: Resources.
- Chapter 8: Methodology.
- Chapter 9: Materials and methods.
- Chapter 10: Results and discussion.
- Chapter 11: Conclusions and future work.

Chapter 2

Problem formulation

Climate change, pest, and diseases and all the environmental phenomena that currently affect the world, have generated different challenges to deal in the incoming years. One of those challenges is the direct impact on food security [28] at a global level; this has called attention at a general scale to generate resistant crops to all these factors and at same time increasing yield and its nutritional value (plant breeders tasks) but also giving tools to farmers to increase and warrant their food production.

[29]

The increasing arrival and spread of serious diseases, fungal infections and viruses, due to climate change and land-use change among other factors, pose a serious food security threat. Disease surveillance for theses major diseases carried out through field visits, or human scouting are often complemented by diagnostic tools such as using growth media and polymerase chain reaction (PCR) ([30], [31]). The inspection of early disease detection is assessed as prevalence (present or not) at plot/field, farm, village, or landscape level, and the procedure by itself at a large scale is challenging and time-consuming. These limitations have led scientists to investigate on advanced and novel techniques that could obtain the crop health information rapidly and economically ([32], [29]).

One selected important crop to face the food security problem is Banana *(Musa spp)*[4]; Though a major staple in Africa, Asia, and Latin America, only 13% of bananas produced are globally traded [5], clearly indicating the fruit's importance in domestic markets and for food security. In East and Central Africa, it is a substantial dietary component, accounting for over 50% of daily total food intake in parts of Uganda and Rwanda[6]. In Colombia, Banana is one of the main fruit crops that sustain the country's food security and it as a part of different typical dishes.

A big concern in Banana are pest and diseases, and this has a direct impact on yield and food production. Several diseases affect banana crops around the world threaten food security, such as Banana bunchy top virus(BBTV), banana Xanthomonas wilt(BXW) (see figure 2.1a), Black and Yellow Sigatoka and Fusarium wilt of banana (FWB) and some pest as Banana Weevil(see figure 2.1b). In order to manage the spread of crop diseases and pest, early identification in the field is a crucial step. Fortunately, most of the crop diseases can be managed by detecting the diseases as soon as it appears on the crop. Traditional pest and disease identification approaches rely on the support of agricultural extension specialists, but these approaches are limited in developing countries with low human infrastructure capacity. The major part of banana crops around the world are from smallholders farmers; they do not know, or they use empirical methods to make the diagnosis, but the early identification of the problem is not possible in most of the cases.

(a) BXW in banana fruit bunch

(b) Banana weevil effects in the corm

Figure 2.1: Disease and pest impact in banana

Artificial intelligence (AI) with deep learning models [33] are useful to help identify plant diseases by the plant's appearance and visual symptoms[34]. Deep learning models can be embedded in smartphones to create applications that could alert farmers of the presence of a disease, lead to concerted control actions and thus potentially preventing a disease from spreading over large areas[19]. Even though many of the developing country farmers around the world do not have access to these advanced tools, internet infiltration and smartphone penetration offer new outfits for in-field crop disease detection. The GSMA (Global System for Mobile Association) predicted that global smartphone subscriptions would reach 5 billion by 2020, of which nearly one billion in Africa (GSMA Intelligence, 2016).

Mobile applications are promising to be excellent tools to support farmers in disease detection and crop management. The main problem with AI applications is that it is needed vast amounts of data to train the models and have reliable predictions, since in deep learning techniques specially in object detection models the main data source are images, it also need many images to learn the patterns. In real cases is very difficult to obtain a data set with enough and reliable images, because the images should be taken in different environmental conditions, different hours (morning, noon and afternoon) and by disease experts.

As banana is one of the most important fruit crops around the world, it is needed to create easy tools to detect diseases using deep learning models, but there is a big problem and is the data acquisition as mentioned before, fortunately, traditional data augmentation methods have demonstrated to be an excellent option to increase the number of images[27] but these approaches are just making the models more robust to orientation using the same original information, so creating new artificial information is the real call of duty, generated artificial data could improve not only the performance of the deep learning models but the generalization capabilities, and this is the scope for this work. To evaluate if artificial data is strong enough and a good option in order to increase the accuracy and the performance of deep learning models but also to ease the creation of new technology implemented in banana disease detection for the future, following research questions have been generated:

- Traditional data augmentation techniques are enough to increase the performance of object detection deep learning models in banana disease detection?
- Artificial data helps in disease diagnostic tasks?
- How many images are needed to have a good performance in object detection deep learning models for banana diseases?
- Combining real data with augmentation techniques and artificial data will improve the performance of object detection deep learning models?

Chapter 3

Scope

The scope of this work is to apply data augmentation and artificial data generation techniques to asses the performance of object detection deep learning models in banana disease detection tasks. It aims to:

- To create a banana diseases data set to train object detection deep learning models
- To evaluate novel techniques for data augmentation
- To compare the performance of object detection deep learning models using real, artificial and combined data.
- To summarize the workflow and the results of this work

Chapter 4

Objectives

4.1 General Objective

To evaluate the performance of object detection deep learning models using data augmentation and artificial data generation techniques in banana disease detection.

4.2 Specific Objectives

- To create and label banana diseases data set
- To train and use Generative adversarial networks (GAN) to generate artificial banana diseases images
- To generate two data sets of artificial images using data augmentation techniques and GAN models
- To train object detection deep learning models with the real, synthetic and cross synthetic datasets
- To evaluate and compare the performance of the trained models.

4.3 Expected results

It is expected to have GAN models trained with banana disease images dataset to generate new synthetic images with quality enough to train an object detection deep learning model and finally evaluate if synthetic images are useful to increase the number of images in a banana disease detection dataset

Chapter 5

Justification

Applying machine learning techniques such as deep learning helps farmers with disease classification and detection. In some cases, diseases are presented only in leaves so this approaches can tell to the farmer in a visual way which disease is present in the plant. Fortunately, mayor banana diseases can be treated if are detected in early stages, for instance, mobile applications using deep learning models give a practical tool to the farmers for real in-field detection and additionally crop management advises controlling the current disease. Also, the geolocation information of the images can describe the behavior and the spreading of the diseases through a city or even a country. But the major problem that is not letting this technology going on to the future is the data acquisition.

Data acquisition for training deep learning models is time-consuming, long duration, experts are required and a massive amount of images with several variations. So data augmentation techniques and artificial data generation like GAN models have the duty to help in this task making these models more robust and increasing the performance. This study aims to evaluate if GAN models as new artificial data generation techniques have the power to generate good quality data to help in deep learning training processes. If GAN models can generate good quality data, more real-world applications in the future will be developed based on this new data generation technique. One of GAN models application will be helping farmers in their daily tasks, such as crop evaluation, early diagnosis, and disease detection

Farmers have been distant to novel technologies for years, getting those technologies close to them as in this study, reducing the number of images to be acquired to train deep learning model, making the job easier to create useful technology that in the incoming years will contribute to food security. With the improvement of all these things and technological support can be faced food security problems in the future because with the development of early diagnosis tools and more mobile applications for farmers, they can be lead to have better yield, improve their food production and raise their gains warranting the local food needs. Giving them technological support to face those problems and to improve their crops, making them closer to new technologies and creating tools is an excellent way to make the world more sustainable.

Chapter 6

State of the art

6.1 Previous work

This study is part of a biggest project aiming to use machine learning and deep learning algorithms, remote sensing techniques and different data sources such as smartphone, drones and satellite images in banana crop. Two main studies have been published by the author of this current work and the research phenotyping platform from the international center of tropical agriculture. The first work, ai-powered banana diseases and pest detection [35] collected part of the dataset used in this study, the final dataset have more than 15.000 images and 30.000 image annotations. we developed six disease detection deep learning models with several diseases, ending up with high accuracy models. Those models were used to develop the Tumaini APP mentioned in following chapters. The second study was about banana disease detection using RGB drone images and to classify banana crops under mixed complex systems in Africa called: Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in DR Congo and Republic of Benin [36]. The final objective of this massive project is to create a banana monitoring system using different information sources at a world wide scale. This study will contribute to increase the number of diseased banana images and aiming to improve the accuracy of the developed models.

6.2 Related topics

6.2.1 Automatic Plant diseases detection

Plant disease classification is a very complex task as it relies on experts hands, but there are some systems for automatic diseases detection such as image processing, machine learning, and deep learning to help in these tasks (see figure 6.1). Up to now most of the approaches for plant disease detection were depending on machine learning algorithms, these approaches are easily adapted to certain conditions as light for instance, if the system has a new input with a tiny variation the accuracy of the model will decrease[37]. The development of deep learning, deeper networks, convolutional neuronal networks (CNN) and transfer learning approach creates new models that are not as susceptible as common machine learning techniques.

Figure 6.1: Automatic pest detection in Banana using deep learning

6.2.2 Pattern recognition

Pattern recognition is a big knowledge area and is used in several applications in real life such as modelling, data management, commerce and even in politics. Disease diagnosis and detection are important topics to explore with pattern recognition and they have shown good results[38, 39] also with techniques as machine learning[40, 41]. Recently all these techniques are being applied in agriculture applications, starting with crop modeling [42], using i.e. support vector machines, random forest or K-nearest-neighbor among many others for crop classification [43] and for plant disease detection [44].

6.2.3 Deep learning

Starting from machine learning networks, the necessity to create robust models, capable to manage several training classes, great inference power and good performance, made machine learning networks to grow in the sense of these new networks are deeper. For example in the ImageNet challenge(2012), AlexNet had a good performance[45]. Deep learning (DL) networks are working in important areas as face detection and agriculture, being this last a promising field to apply those techniques. Deep learning is used in Plant stress phenotyping[46], Plant diseases and pests detection in tomato[47] ,Cassava disease detection[20][48], finally deep learning can be embedded into smartphones to create useful applications helping with disease diagnosis[49].

6.2.4 Object detection

Classification in deep learning tasks aims to predict a class inside an image, for instance, if it is wanted to differentiate between if what there are in the images is a cat or a dog, the model can predict which animal it is. Object detection not just classify between a cat or a dog but detects where inside the image is the detected object (see figure 6.2), in this way, the system may tell the farmer where the disease is in a real scenario[47].Two examples of this technology in a real scenario are in Cassava and tomato, where object detection and transfer learning were used to detect diseases in the cassava and tomato leaves[49][47] and implemented as a mobile application[50]. Currently, convolutional Neural Networks (CNN) are considered as the foremost method for object detection and there are some well performed methodologies such as Faster R-CNN and Single Shot Detectors (SSD). The evaluation of the performance is different from classification techniques as confusion matrix, there are some techniques to do the evaluation as mean average precision mAP score.

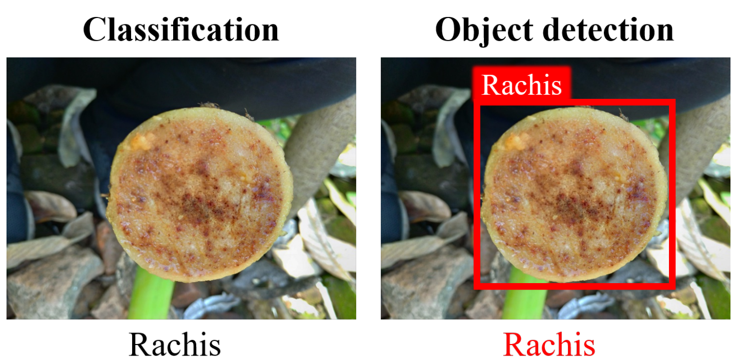

Figure 6.2: Difference between classification and object detection.

6.2.5 Transfer learning

Transfer learning is becoming into an exciting way to face situations where the number of training data is not enough to train a deep learning model [51], is well known that for deep learning is needed huge amount of data to train the network, , i.e. in a classification task, to train a classifier from zero will be needed thousands of images. What transfer learning is giving is a new technique to take pre-trained networks and adapt them to do a new task [52]. For instance, an existing network trained to classify cars can be re-trained to classify trucks without a training process from scratch, reducing the training time and getting faster results.

6.2.6 Generative adversarial networks

GAN models are part of artificial intelligence, they are part of non supervising learning approaches which creates a system with two neuronal networks that compete with each other, were created by Ian Goodfellow et al. in 2014 [22]. GAN models are currently being used in different applications: Style based generator architecture [53] that is a work from NVIDIA's team capable to generate artificial information based on two images and creating a new one with the mixing of features from both images. Another application is creating super resolution images[54]. GAN is a promising technique to help in different areas such as artificial data generation to dataset augmentation or image domain translation.

6.2.7 Traditional data augmentation techniques

Some machine and deep learning problems requires a data set with hundred or thousand of images, which is hard to collect in most of the cases, but with few images can be generate new samples to train the models applying different techniques[27] such as:

- Vertical and horizontal flipping
- Image rotations
- Brightness decrease and brightness increase
- Contrast enhancement and contrast reduction
- Sharpness enhancement
- Addition of random Gaussian noise

Those traditional techniques make models more robust and invariant to some variables as light condition and angle rotation, increasing the accuracy of the model.

Chapter 7

Resources

7.1 Human Resources

7.1.1 Director

Dr. Animesh Acharjee, UK Centre for Computational Biology, University of Birmingham, UK.

7.1.2 Co-Director

Dr. Michael Gomez Selvaraj, International Center for Tropical Agriculture (CIAT), Palmira-Colombia.

7.1.3 Consultants

- PhD student Henry Ruiz, Texas A&M University, Texas-United states.
- MSc. Milton Valencia, International Center for Tropical Agriculture (CIAT), Palmira-Colombia.
- MSc. Manuel Alejandro Valderrama, International Center for Tropical Agriculture (CIAT), Palmira-Colombia.

7.1.4 Research Group

International Center for Tropical Agriculture (CIAT)

7.2 Hardware

- GPU NVIDIA Tesla V100 32GB
- Laptop ASUS ROG Intel Core i7, 32GB RAM DDR4
- Smartphone
- Hard drive 8 TB

7.3 Software

- Operating system Linux
- Data annotation software CVStudio
- Image processing library OpenCV
- Programming language Python 3.7
- Model training FalconCV using as backend Tensorflow with the tensorFlow API and Keras

Chapter 8

Methodology

8.1 Research Methodology

This work is an analytic research to evaluate the performance of deep learning models trained with different sources of information (real and artificial data).

8.1.1 Activities

To achieve the objective of this work and following each sub-objective, have been proposed the following probes and activities:

Activity 1. To create and label real banana diseased dataset

1. To collect banana images from different places.
2. To filter the images.
3. To split and arrange the images in a structured format (Class based).
4. To label the images with the correct bounding box and class label.

Materials:

- Smartphone or RGB camera
- Laptop:
  Processor: Intel Core i7-6700HQ Quad Core Processor (6M Cache, 2.6GHz - 3.5GHz)
  Ram: 32GB RAM DDR4 2133MHz
  Hard Drive: 500GB Solid State Drive + 2TB 5400rpm Hard Disk Drive
  Graphics Card: NVIDIA GeForce GTX 960M 4GB

- Storage space
- Label image Software

Activity 2. To train Generative adversarial networks generating artificial banana diseased images

1. To enrich the knowledge about GAN models
2. To explore different GAN models.
3. To select one GAN methodology and apply the methodology generating new data.
4. To test the quality of the generated data.

Materials:

- Laptop:
  Processor: Intel Core i7-6700HQ Quad Core Processor (6M Cache, 2.6GHz - 3.5GHz)
  Ram: 32GB RAM DDR4 2133MHz
  Hard Drive: 500GB Solid State Drive + 2TB 5400rpm Hard Disk Drive
  Graphics Card: NVIDIA GeForce GTX 960M 4GB
- Server: Processor: Intel Xeon E5-2667 v4 @ 3.20 GHz x16, GPU: NVIDIA Tesla M60, OS: Windows Server 2016 x64
- Python Software

  Numpy

  Scipy

  Pandas

  Matplotlib

  Augmentor

  PyTorch

  Keras

  TensorFlow

Activity 3. To generate two data sets of artificial images using GAN models

1. To apply the selected techniques to create the new data set.

2. To create a data set with GAN artificial images.

3. To label the images with the correct bounding box and class label .

Materials:

- Laptop:
  Processor: Intel Core i7-6700HQ Quad Core Processor (6M Cache, 2.6GHz - 3.5GHz)
  Ram: 32GB RAM DDR4 2133MHz
  Hard Drive: 500GB Solid State Drive + 2TB 5400rpm Hard Disk Drive
  Graphics Card: NVIDIA GeForce GTX 960M 4GB.

- Label image Software

Activity 4. To train deep learning models with the real, augmented, artificial and combined data sets

1. To train deep learning models with real data.

2. To train deep learning models with synthetic data.

3. To train deep learning models with cross synthetic data.

Materials:

- Laptop:
  Processor: Intel Core i7-6700HQ Quad Core Processor (6M Cache, 2.6GHz - 3.5GHz)
  Ram: 32GB RAM DDR4 2133MHz
  Hard Drive: 500GB Solid State Drive + 2TB 5400rpm Hard Disk Drive
  Graphics Card: NVIDIA GeForce GTX 960M 4GB

- Server: Processor: Intel Xeon E5-2667 v4 @ 3.20 GHz x16, GPU: NVIDIA Tesla M60, OS: Windows Server 2016 x64

- Python Software

  Numpy

  Scipy

  Pandas

  Matplotlib

  Augmentor

PyTorch

Keras

TensorFlow

Activity 5. To evaluate and compare the performance of the trained models

1. To generate evaluation metrics (such as mAP score) for each trained model.
2. To compare the results
3. To write the conclusions

Materials:

- Laptop:
  Processor: Intel Core i7-6700HQ Quad Core Processor (6M Cache, 2.6GHz - 3.5GHz)
  Ram: 32GB RAM DDR4 2133MHz
  Hard Drive: 500GB Solid State Drive + 2TB 5400rpm Hard Disk Drive
  Graphics Card: NVIDIA GeForce GTX 960M 4GB
- Server: Processor: Intel Xeon E5-2667 v4 @ 3.20 GHz x16, GPU: NVIDIA Tesla M60, OS: Windows Server 2016 x64
- Python Software

  Numpy

  Scipy

  Pandas

  Matplotlib

  Augmentor

  PyTorch

  Keras

  TensorFlow

Chapter 9

Materials and methods

9.1 Overall system description

The synthetic data generation approach to train disease detection models consists of two main models, based on an important banana plant part (Pseudostem and rachis). Each model (plant part) has its own classes extracted from different banana diseases and healthy plants. The pseudostem model, affected by Banana xanthomonas wilt (BXW) and Fusarium wilt of banana (FWB) and the rachis model, affected by banana blood disease (BBD). Both models with check images of their healthy appearance as mentioned before. Once synthetic images are generated, a deep learning object detection model is trained to test the synthetic data in a real disease detection scenario. An overview of the complete pipeline is illustrated in Figure 9.1.

9.2 Raw dataset description

The raw dataset consists of 3500 pre-screened images by banana experts; it is composed of two main banana plant parts, the pseudostem divided into three classes based on the disease, BXW, and healthy class (as described in [35]) and FWB. The second banana plant part is the rachis, divided into BBD and a healthy class, as shown in table 9.1, and examples of each class in the dataset are presented in Figure 9.2. This dataset has been collected under real field conditions and taking key symptoms of each disease under different environmental conditions, disease stage, and multiple locations. In addition to the described dataset, 150 images per class were added to test each detection model generated under real images.

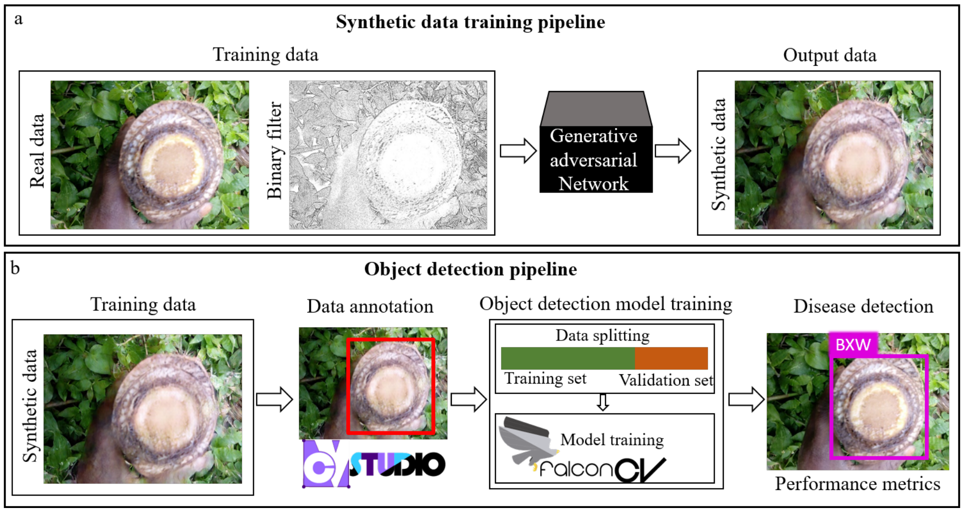

Figure 9.1: Overall pipeline description. a) Synthetic data generation pipeline: In the first stage, a binary filter is applied to the real data to create a paired dataset; with the paired dataset, a generative adversarial network is trained, and using this trained network, a new synthetic dataset is generated as the output. b) Object detection pipeline: The synthetic dataset generated in the synthetic data training pipeline is annotated to create the disease label and split into a training and validation dataset. The training dataset is used to train an object detection model. The output is a trained model to detect banana diseases.

9.3 Synthetic data generation

9.3.1 Data preprocessing and preparation

To train a supervised generative adversarial network (GAN) is needed a paired dataset to apply image-to-image translation. To create this dataset, some image processing techniques are applied [55] to the raw images like applying a greyscale filter, edge detection, background removal, or binarization. The selection of the technique to apply is based on the problem to solve; in this case, we applied a binarization filter with adaptative gaussian thresholding using python and the image processing library OpenCV [56], due to a binarization filter will help to extract key disease features and patterns which may ease the training process and the image generation. Also, the binarization process depends on a threshold that is not 100% accurate to represent the original image; this will lead the GAN to

Classes/ Model	Healthy (HP)	Banana blood disease (BBD)	Banana xan-thomonas wilt (BXW)	Fusarium wilt (FWB)	Total
Pseudostem	700		700	700	2100
Rachis	700	700			1400
Test images	300	150	150	150	750

Table 9.1: Description of banana diseases dataset used in this study.

generate a new image with variations from the original one, which is important to create variety in the new synthetic dataset and help in the data augmentation. Once the filter is applied, data is split into a paired dataset, a training dataset A (binarized images), and a training dataset B (raw images) as shown in Figure 9.3.

9.3.2 Generative adversarial network

We selected the pix2pixHD [1] GAN architecture that is an improvement of the pix2pix [55] conditional GAN framework to do an image-to-image translation, In the case of GANs, it is integrated a generator G and a discriminator D, where for our task the G is transforming the image from the dataset A (binarized image) to the dataset B (raw original image), and the D detects whether is a real or a generated image. This network is a supervised network that works with paired data as described before, so the input is a pair of binarized image-raw images (Figure 9.3) and, the output a new synthetic image generated from the binarized image.

9.3.3 Pix2PixHD GAN model architecture

Pix2pixHD was created in order to improve the resolution (up to 2048 x 1024) of the output images, the proposed method to improve the results and the architecture of this GAN model are: (1) Multi-scale generator to stabilize training and change the resolution from 1024x512 to 2048x1024, (2) Multi-scale discriminator to ensure scene consistency at all level and improves visual results, (3)Improved adversarial loss for a stable training and better visual results, (4) feature encoder network to learn representation of an object instance as a feature vector, to allow interactive editing

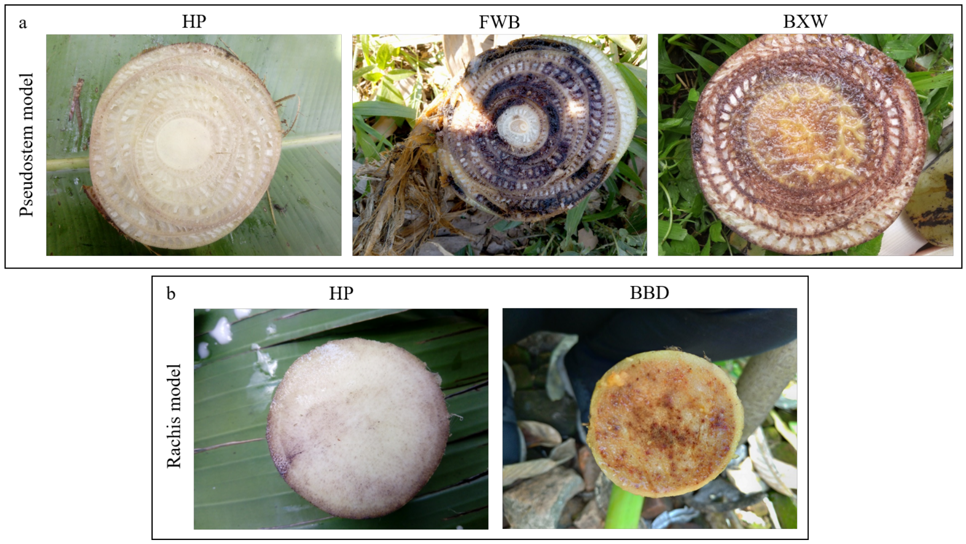

Figure 9.2: Dataset per-class example. a) Pseudostem model example of a Healthy plant(HP), Fusarium Wilt of banana(FWB) and banana xanthomonas wilt(BXW) classes. b) Rachis model example of a Healthy plant(HP) and banana blood disease(BBD) classes.

Multi-scale generator

In this multi-scale generator as can be observed in Figure 9.4, First is trained an autoencoder style network G1, at smaller resolution (1024x512), then appends generator G2, to generate a higher resolution (2048x1024) and this approach leads to a stable training, finally G2 takes coarse output from G1 and refines it to produce final result

Multi-scale discriminator

In the discriminator network, 3 discriminators have been used to operate on different resolutions, to ensure resemblance at all levels, 70x70 Patch GAN architecture evaluates image in 70x70 patches, resulting in smaller kernels, efficient memory usage and yielded better(as observed in Figure 9.5). If Ck -¿ 44Convolution-InstanceNorm-LeakyReLU layer with k filters and stride 2. Finally the distribution of the architecture is C64-C128-C256-C512

Figure 9.3: Example of the training paired dataset, on the left the binarized image (dataset A), and on the right, the raw image (dataset B).

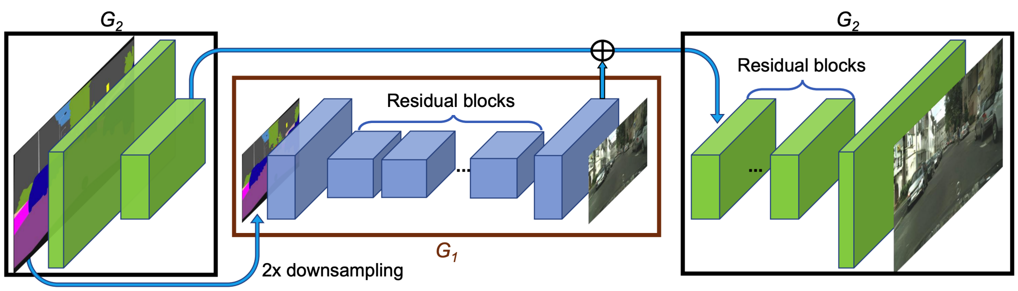

Figure 9.4: Multi-scale generator network, image taken from [1].

Improved adversarial loss

Instead having a traditional GAN objective function as the loss function, was added a feature matching terms to improve visual similarity between predictions and ground truth. In the discriminator was added a discriminator feature matching loss with difference in activations at every level. Finally an additional feature matching loss on activations from pretrained VGG network

Feature encoder network

The aim of this feature encoder network. is to learn a feature vector at the object instance level, With added input of 3d feature vector, averaged per instance, tries to learn input instance vector vs visual observation. After joint training, they perform k-means clustering with 10 final clusters over the entire training set per semantic category. Cluster centers are provided as visual options in the final

Figure 9.5: Patch size variation, image taken from [1].

editable framework

9.3.4 Parameters of GAN models

Each GAN model was trained with 200 iterations at the starting learning rate of 0.0002 and an additional number of 200 iterations to linearly decay learning rate to zero based on the recommendations from pix2pixHD authors[1], the weights of the network were initialized using a Gaussian distribution with mean 0 and standard deviation 0.02. As the input in this case is not a label map as in the original one, the label nc parameter was in 0 which will use RGB colors as the input and a resize and crop parameter set to fineSize = 1024.

9.3.5 Synthetic images testing

Initially, output synthetic images were tested using the android app called Tumaini (TumainiAPP can be found in the google play store) specially developed for banana disease detection using object detection deep learning models.

9.4 Object detection pipeline

9.4.1 CVstudio for data annotation

Image annotation process was done using our recently open-source software called CVstudio in the raw and synthetic datasets, this software allows to the user create bounding boxes and class labels and save it as an XML file using the PASCAL VOC format [57], also is possible to use pre-trained object detection models to detect and create automatic bounding boxes. An overview of CVstudio software is shown in Figure 9.6.

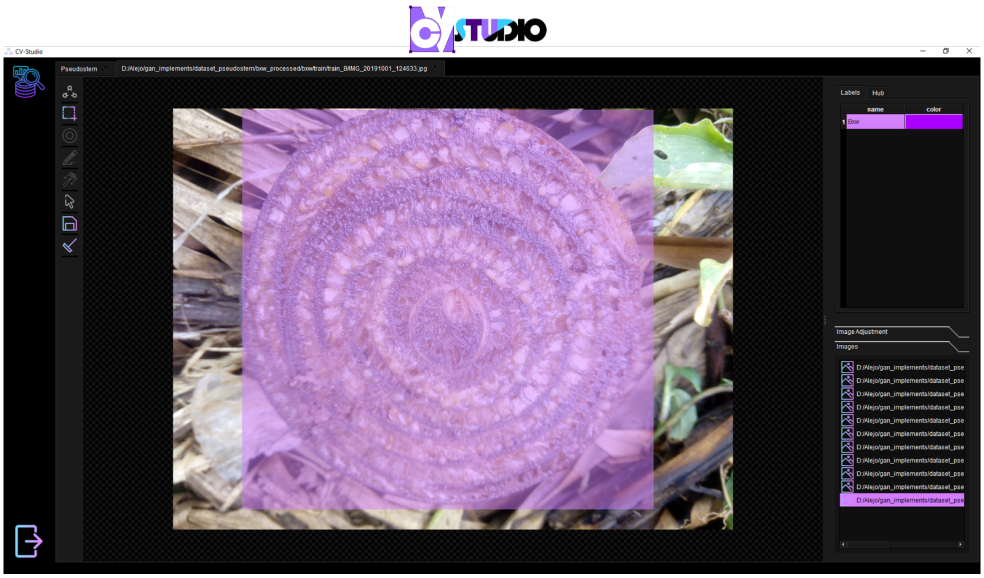

Figure 9.6: CVstudio software overview and labeling process, purple rectangle is an example of a bounding box which is the region of interest for the model.

9.4.2 Model training with FalconCV

FalconCV [58] is a recently deep learning (DL) library to ease the training process of DL models, in the backend FalconCV is using TensorFLow [59] and its object detection API[60], we used a pre-trained version of Faster R-CNN with InceptionV2 trained with COCO (Common objects in context [61]) data set as the base model to do the transfer learning and train our disease detection model with raw and synthetic data, InceptionV2 has shown good accuracy for banana disease detection ([35]). The transfer learning was done with 20.000 epoch for the training and a data splitting of 70% for training and 30% for testing in all trained models.

9.4.3 Performance testing

In order to evaluate the performance of the models to see if using raw and synthetic data is possible to train disease detection models with accurate predictions, we used the mean average precision score (mAP). mAP is a used metric to evaluate the performance of an object detection model, is calculated as the average for the precision of all classes, the precision is calculated as the number of true positives (TP) over the TPs plus false positives (FP) (TPs + FPs is also known as total positive results). Finally, a TP or FP is extracted using the intersection over union (IoU) which is calculated as the overlap (or the intersection over the union) between the predicted bounding box and the labeled bounding box (ground truth), in this case, was considered an IoU greater than 0,5 to be taken as a TP. Additional to mAP we also calculated the precision and recall per class. Also, the loss function is a function that penalizes an incorrect prediction in the training process, this loss function was extracted using Tensorboard [62] that is a visualization toolkit of TensorFLow. This loss function allows us to monitor the training process of each model, as much as it decreases the model is being well trained, but also to control the overfitting.

9.5 Software and hardware specification

The hardware (HW) used to develop and re-train GAN and disease detection models was an NVIDIA Tesla V100 GPU with 32GB, Linux as the operative system and python 3.7 as the programming language to train the models and to create the image processing scripts, the overall software (SW) and hardware used in this study are presented in table 9.2.

Hardware and software	**Specifications**
GPU	NVIDIA Tesla V100 32GB
Operative system	Linux
Data annotation software	CVstudio
Model training	Pix2PixHD (GAN), FalconCV with TensorFlow and TensorFlow object detection API
Image processing library	OpenCV
Programming language	Python 3.7
Visualization toolkit	TensorBoard from TensorFlow

Table 9.2: Hardware and software specifications.

Chapter 10

Results and discussion

10.1 Synthetic data generation

10.1.1 Generative adversarial network

Using the Pix2PixHD architecture we generated five GAN models, three models in pseudostem case (BXW, FWB, and HP) and two for Rachis (BBD and HP), each model took around 2 days to finish the training process using all parameters, HW and SW described before, all trained GAN models were named as synthetic images generator (SIG).

10.1.2 New synthetic images generation

As can be observed in Figure 10.1, each class was used to generate two new synthetic datasets, the first one using the same SIG as the class and the second one using the opposite class (cross synthetic images, CSI); this method is also known as domain translation(in this case translating from a disease image to a healthy image and vice-versa. In the example shown in Figure 10.1 was used the BBD raw dataset (700 BBD images) then using the BBD SIG, 700 synthetic BBD images were generated, and using the Healthy SIG, 700 HP cross synthetic images were generated. In the pseudostem model, as there are three classes, we used each class paired with its same SIG to generate the first new dataset; examples of the generated images can be found in Figure 10.2, but in the case of the second dataset (CSI), was used the HP class with each disease SIG to generate the new BXW and FWB CSI datasets and the FWB class with the HP SIG to generate the HP CSI. FWB was selected because the BXW images have bacterial ooze in most of them, making the translation harder.

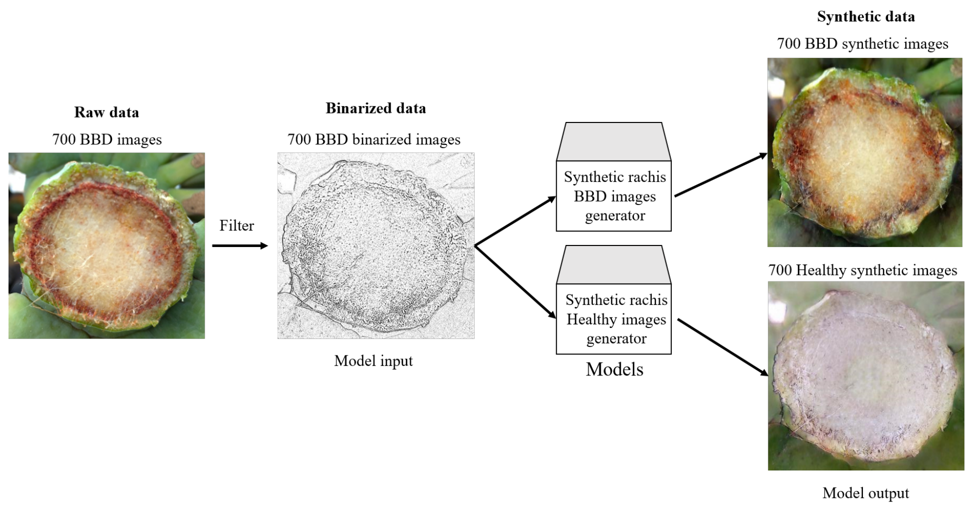

Figure 10.1: Example of the synthetic image generation workflow for banana blood disease raw images. The binarized image from a BBD image is passed to two models, the first one, trained with the same class as the input image (BBD image in this case) being the output a new bbd image reconstructed from the binary image and the second one, trained with a different class from the input image (model trained using healthy plants in this example) and being the output a new healthy image generated from a BBD binary image this is also known as domain translation and is called in this study as cross synthetic image.

Additional images were created, converting from BXW class to FWB class and vice-versa trying not to only turn from a diseased image to a healthy one but also to do disease image translation. These cross synthetic images were not included in the cross synthetic dataset because most of the images do not clearly represent the diseases symptoms once the translation is done, the results from the CSI of the disease to disease process (BXW to FWB or FWB to BXW) are shown in Figure 10.4, the figure clearly shows that symptoms in each case are not clearly represented and are not useful to be included in the datasets used to train the models. It is still an initial approach to do disease image translation. The complete dataset after filtering the 700 images per class due to some artifacts described in the next section is presented in table 3. Synthetic images are very similar to the raw images because they were generated with the same raw training images but with a few differences, as shown in Figure 10.3, which can increase the number of images in an object detection or classification dataset. In some cases,

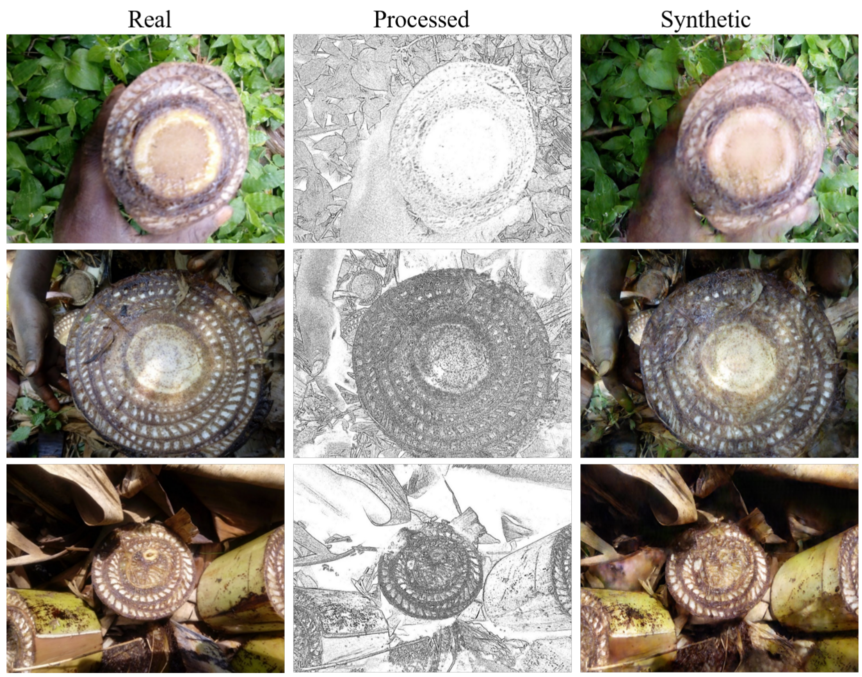

Figure 10.2: Examples of generated synthetic images. On the left column, all the original images; on the center column, all binarized images from the left column. Finally, in the right column, the result from passing the binarized image through the GAN model. There are some variations in the synthetic images from the original images that can be observed, which creates variation in the dataset.

synthetic images do not have enough difference from the original raw images, so they must not be included in a training dataset for classification or object detection tasks due to the possibility of overfitting. In the case of CSI, they are a completely new dataset due to the domain translation, as shown in Figure 10.3.

10.1.3 Synthetic images testing and artifacts

To test if the synthetic images were correctly generated, we used the android app for banana diseases detection Tumaini, where they report more than 90% accuracy in their object detection models. These results are shown in Table 10.1, where we got an average detection percentage (ADP) of 97% in synthetic images and 80% in CSI. Even though 700 images were generated for all classes, a filtering process was done based on the Tumaini app detections and a scan of each image, where

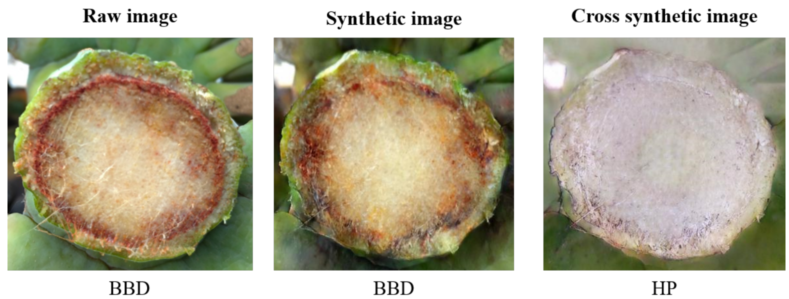

Figure 10.3: Comparison between raw, synthetic and cross synthetic images.

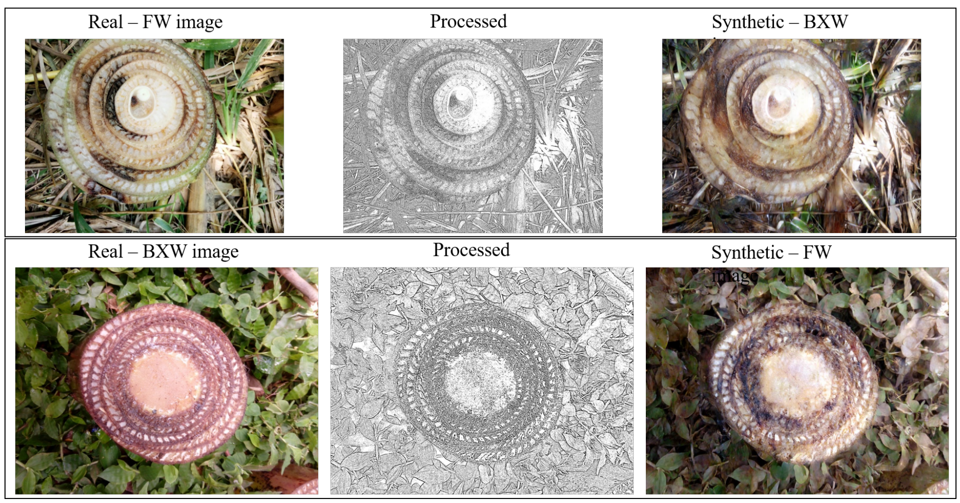

Figure 10.4: Examples of diseases images translation (domain translation).

bad detections are also related to some distortions and artifacts presented on the images; some of these artifacts and distortions are normal due to the GAN network like discoloration, noise, background discoloration, and darkening, as shown in Figure 10.5.

10.2 Object detection models

10.2.1 Data annotation using CVstudio

CVstudio for data annotation is easier and faster than other annotation tools due to its interactive and intuitive graphical user interface and because the user

Model	HP	ADP	BBD	ADP	BXW	ADP	FWB	ADP	Total
Pseudostem (Raw images)	700	100%			700	100%	700	100%	2100
Rachis (Raw images)	700	100%	700	100%					1400
Pseudostem (Synthetic images)	690	98%			687	98%	691	98%	2068
Rachis (Synthetic images)	685	97%	679	97%					1364
Pseudostem (CSI)	602	86%			584	83%	563	80%	1749
Rachis (CSI)	560	80%	553	79%					1113

Table 10.1: Description of complete dataset and average detection percentage(ADP).

can upload a trained object detection (or segmentation) model with your object in interest, and the software will automatically detect it. This detection tool is based on the accuracy of the uploaded model; in the pseudostem dataset using the Raw pseudostem trained model, we found an average confiability of 80% in the synthetic image dataset for all classes for this annotation process. For rachis using raw dataset trained model; we found an average confiability of 85% for all classes, all the annotations were checked, and remaining images were done manually. This tool reduced annotation time from hours to a few minutes in large datasets if the user has a pre-trained model.

10.2.2 Model training with FalconCV

FalconCV is using in the backend TensorFlow 2.0 and the object detection API. A new model can be retrained using transfer learning and modifying all training parameters from the training script, installing and configuring an environment with TensorFlow and object detection is very difficult and time-consuming, Falcon is reducing and facilitating this process.

Figure 10.5: Artifacts examples presented in the synthetic images.

10.2.3 Performance testing

Different performance metrics were extracted in order to test and evaluate the model behaviour

Loss function

After training all models with an average of four hours per model, we extracted the loss function (shown in Figure 10.6), finding a particular difference between models trained with raw data and models trained with synthetic data. Models trained with synthetic data had a faster convergence and stabilization than those trained with raw data, this behavior is because despite the object of interest is well generated, the image background tends to be blurred (Figure 10.7) which highlight the object, making the training process to converge faster. As all raw datasets are under real field conditions and not under a controlled environment, these models are harder to train; then, if the object is highlighted because of the blur in the background and other artifacts in the image, this will ease the detection process. Also, it is important to have in mind that both raw data models have similar behaviour (using different data) which can be considerate as a validation and not as an independent event, the same situation is happening with synthetic data models. All models finished with good results, as shown in Figure 10.6 despite the faster convergence of synthetic data models.

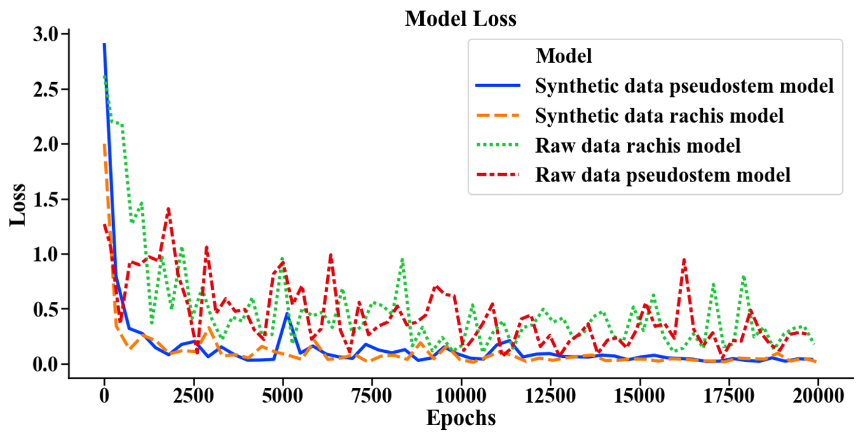

Figure 10.6: Loss function comparison for raw and synthetic object detection models.

Precision and recall

To evaluate the model's accuracy, initially, we calculated the IoU over the data and extracted the precision and recall per class (shown in table 10.2). In table 10.2 can be observed that those model trained with raw images have good precision and recall values (near to 1) which indicates a good behavior, models trained with synthetic images have good recall values which indicate that the model is detecting all possible positive cases, but lower precision values than previous raw images models which indicate that this model is less accurate, finally, in CSI model, precision and recall is similar to synthetic images models which are expectable due to both are generated information.

In this case, what the precision is telling (as it is per class) is that, if the value is near to 1 the model is being able to detect a high percentage of your results which are relevant. On the other hand, what recall is saying is the percentage of total relevant results correctly classified by the model. As and example, in the raw images model, using the BXW class, precision was 0.92 which means that the model is detecting 92% of the possible important results, then the recall of 1 is saying that 100% of the possible positive cases are being detected correctly. Compared to the worst class (FWB) from the CSI model, where the precision was 0.62 which means that this model is detecting 62% of possible important results and the recall off 0.94 concluding that from all detections made, 94% were accurate detected

In conclusion, all models have good values in precision, all above 62% concluding that even worst classes of CSI models (worst performed compared to the

Figure 10.7: Blurred background and artifacts in a generated image.

others) are being able to detect at least 62% of possible important cases in a bad scenario. But from those important detections, more than 94% (recall for FWB class in the pseudostem CSI model) of the relevant results are correctly classified. Also, pseudostem models have better behavior than rachis models in a small percentage, despite the number of classes.

mAP score

As a final performance metric to determine if the comparison between training an object detection model using raw and synthetic data is useful, we extracted the mAP score in each model; the results can be observed in table 10.3. As reported in Selvaraj et. al 2019, [35], raw models have a mAP score of 0.99 giving accurate detections. Synthetic models reported to be 0.89 for pseudostem and 0.88 for rachis, which is lower than raw models but still accurate enough to give good detections. The difference in the accuracy is derived from the noise generated in the synthetic images. Finally, CSI models with 0.67 and 0.64 of mAP score for pseudostem and rachis respectively, these results are lower than previous models with almost 20% in difference (comparing with synthetic models), despite lower accuracy in comparison, 0.65 mAP score in average for both models is not a bad result taking into account that these models were trained using CSI. The low accuracy in CSI models is also derived from the noise in the image generation and the domain translation from a diseased plant to another or from a healthy plant to a diseased one, having in mind that all models were tested using real field images, to check their behaviour under a real scenario. In conclusion, if the lower accuracy was the one from the CSI rachis model, can be said that in the worst scenario your

Model	**Class**	**Precision**	**Recall**
Pseudostem (Raw images)	BXW	0.92	1
	FWB	0.91	0.99
	HP	0.95	0.99
Pseudostem (Synthetic images)	BXW	0.83	0.99
	FWB	0.62	0.99
	HP	0.80	1
Pseudostem (CSI)	BXW	0.63	0.96
	FWB	0.62	0.94
	HP	0.73	0.98
Rachis (Raw images)	BBD	0.91	1
	HP	0.96	0.99
Rachis (Synthetic images)	BBD	0.85	1
	HP	0.66	1
Rachis (CSI)	BBD	0.70	0.95
	HP	0.64	0.95

Table 10.2: Precision and recall per class @ 0.5 intersection over union (IoU).

original dataset will be increased in a 64%.

Data source	**Model**	**Accuracy**
Raw images	Pseudostem	0.99
Raw images	Rachis	0.99
Synthetic images	Pseudostem	0.89
Synthetic images	Rachis	0.88
CSI	Pseudostem	0.67
CSI	Rachis	0.64

Table 10.3: mAP score performance metric to compare models behavior.

Chapter 11

Conclusions and future directions

Traditional data augmentation techniques like mirroring and splitting images are useful to increase the accuracy in DL models. Still, those techniques have not enough variation to increase the variability in the dataset, which leads DL algorithms to have better performance; the more variability (with real coverage of your features, in this case, diseases symptoms), the more generalization in your predictions. In this study, we generated synthetic images to test novels data augmentation techniques using GAN power to train object detection models in a real disease detection case applied to three major banana diseases in different plant parts. Two different kinds of synthetic images were generated using the Pix2PixHD trained GAN model, the first one using increasing the original raw images dataset from 3500 images to 9794 (6294 between synthetic images and CSI), an increase of almost 180% in images number. Few of the generated images presented different artifacts, which are bad to use to train object detection models, especially most artifacts were presented in the CSI generated dataset. All generated images were tested in Tumaini mobile app, ending up with more than 80% ADP, and bad detection images were removed from the dataset to train the object detection different models. After testing the three datasets (raw, synthetic, and CSI), the models' accuracy was between 64 and 99%, being CSI models lower than others and synthetic models results were comparable to raw models. The results lead this study to conclude that using synthetic images to train object detection models has similar performance to training with raw real images, even using the domain translation technique, reducing the time, disease spreading risk and the laborious work in image acquisition tasks. Future work will be focused on characterizing and go deeper into the image artifacts to reduce the number of affected images and increasing the number of images for synthetic datasets. Generative adversarial networks are daily evolving, testing new GAN architectures will be useful to create new synthetic datasets and improve the variability of them and also image resolution to increase object detection models accuracy, a good start is to use CycleGAN [63] to see how the image translation is performing using a different GAN model. One of the challenges in this study was to generate disease to disease translation images; developing a successful approach to generate this kind of images will increase the number of trainable images in the dataset. Used images were taken in a real field scenario, which makes difficult the generation task because of the background and variability in the images, but finally, it was tested that synthetic images can be generated with quality enough, but there are some diseases with complex visual symptoms and parts of the plant with very irregular shapes that are difficult to generate with this GAN approach; in future work, other approaches can be tested to see if it is possible to generate more complex images under the same real field scenario. m In this study, a binary filter was used to create the paired binary-raw images dataset, other image processing techniques are also used as mentioned before; generating synthetic information using the developed approach with different filters and image processing techniques will increase the variation in the training dataset and reducing the number of acquired images needed, also some considerations are important to have in mind, one of them is the difference between the generated image and the original one because in some cases the difference is small enough to say that there are two equal images in the dataset which will be very bad for an object detection model training, leading the model to overfit if there are several repeated images. Using a dataset where most images have a similar background, in terms of color and objects like green leaves background will train your network to generate synthetic images with a similar background, which is not what is wanted in a training dataset; one solution is to use multi-season data where the background of the image and even the symptoms of the plant have several variations, this will end with a complete dataset to train GAN models where images under different context and recreating different seasons can be generated Developed techniques in this study can be applied to other DL areas such as machine learning, semantic segmentation, and different real-life problems to fill the gap in images and data collection. In the incoming years, artificial data generation will be an important part of machine learning and deep learning real applications and will increase the number of these applications in real life. The recent increase in the development of disease detection apps for smartphones in several crops like banana, cassava, tomato, potato, etc. creates the need to collect thousands of images; as mentioned collecting diseased images is difficult in different ways; fortunately, the used approach here can be applied to all disease detection mobile apps and several different crops, while the symptoms of the disease are visual.

Bibliography

[1] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 8798–8807, 2018.

[2] O. Nagayets, "Small farms: current status and key trends," *The future of small farms*, p. 355, 2005.

[3] C. Hillnhuetter and A.-K. Mahlein, "Early detection and localisation of sugar beet diseases: new approaches," *Gesunde Pflanzen*, vol. 60, no. 4, pp. 143–149, 2008.

[4] F. Faostat, "Food and agriculture organization statistical database," *Retrieved Feb*, 2014.

[5] T. Lescot, "World plantain and banana production systems," 2013.

[6] S. Abele and M. Pillay, "Bacterial wilt and drought stresses in banana production and their impact on economic welfare in uganda: Implications for banana research in east african highlands," *Journal of Crop Improvement*, vol. 19, no. 1-2, pp. 173–191, 2007.

[7] G. Blomme, M. Dita, K. S. Jacobsen, L. Pérez Vicente, A. Molina, W. Ocimati, S. Poussier, and P. Prior, "Bacterial diseases of bananas and enset: current state of knowledge and integrated approaches toward sustainable management," *Frontiers in plant science*, vol. 8, p. 1290, 2017.

[8] K. A. Garrett, G. Forbes, L. Gómez, M. Gonzáles, M. Gray, P. Skelsey, and A. H. Sparks, "Cambio climático, enfermedades de las plantas e insectos plaga.," 2013.

[9] E. Hamada, R. Ghini, *et al.*, "Impacts of climate change on plant diseases and pests in brazil.," *Revista Mexicana de Ciencias Agrícolas*, no. Especial 2, pp. 195–205, 2011.

[10] C. Z. Qiang, S. C. Kuek, A. Dymond, and S. Esselaar, "Mobile applications for agriculture and rural development," 2012.

[11] J. A. Berni, P. J. Zarco-Tejada, L. Suárez, and E. Fereres, "Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 722–738, 2009.

[12] E. R. Hunt, W. D. Hively, S. Fujikawa, D. Linden, C. S. Daughtry, and G. McCarty, "Acquisition of nir-green-blue digital photographs from unmanned aircraft for crop monitoring," *Remote Sensing*, vol. 2, no. 1, pp. 290–305, 2010.

[13] W. Baoyun, "Review on internet of things [j]," *Journal of electronic measurement and instrument*, vol. 12, pp. 1–7, 2009.

[14] P. C. Doraiswamy, S. Moulin, P. W. Cook, and A. Stern, "Crop yield assessment from remote sensing," *Photogrammetric engineering & remote sensing*, vol. 69, no. 6, pp. 665–674, 2003.

[15] H. Murase *et al.*, "Artificial intelligence in agriculture.," *Computers and Electronics in Agriculture*, vol. 29, no. 1/2, 2000.

[16] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115, 2017.

[17] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, *et al.*, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," *arXiv preprint arXiv:1711.05225*, 2017.

[18] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221–248, 2017.

[19] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in plant science*, vol. 7, p. 1419, 2016.

[20] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," *Frontiers in plant science*, vol. 8, p. 1852, 2017.

[21] M. D. Bloice, C. Stocker, and A. Holzinger, "Augmentor: an image augmentation library for machine learning," *arXiv preprint arXiv:1708.04680*, 2017.

[22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in Neural Information Processing Systems 27* (Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, eds.), pp. 2672–2680, Curran Associates, Inc., 2014.

[23] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.

[24] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," *arXiv preprint arXiv:1812.04948*, 2018.

[25] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets. in international conference on neural information processing systems," 2014.

[26] L. Lan, L. You, Z. Zhang, Z. Fan, W. Zhao, N. Zeng, Y. Chen, and X. Zhou, "Generative adversarial networks and its applications in biomedical informatics," *Frontiers in Public Health*, vol. 8, p. 164, 2020.

[27] J. G. A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification," *Computers and electronics in agriculture*, vol. 153, pp. 46–53, 2018.

[28] T. Wheeler and J. Von Braun, "Climate change impacts on global food security," *Science*, vol. 341, no. 6145, pp. 508–513, 2013.

[29] P. R. Steward, C. Thierfelder, A. J. Dougill, and I. Ligowe, "Conservation agriculture enhances resistance of maize to climate stress in a malawian medium-term trial," *Agriculture, Ecosystems & Environment*, vol. 277, pp. 95–104, 2019.

[30] G. Blomme, K. Jacobsen, W. Ocimati, F. Beed, J. Ntamwira, C. Sivirihauma, F. Ssekiwoko, V. Nakato, J. Kubiriba, L. Tripathi, *et al.*, "Fine-tuning banana xanthomonas wilt control options over the past decade in east and central africa," *European journal of plant pathology*, vol. 139, no. 2, pp. 271–287, 2014.

[31] W. Ocimati, H. Bouwmeester, J. C. Groot, P. Tittonell, D. Brown, and G. Blomme, "The risk posed by xanthomonas wilt disease of banana: Mapping of disease hotspots, fronts and vulnerable landscapes," *PloS one*, vol. 14, no. 4, p. e0213691, 2019.

[32] A. Heim, "Food environment research among an indigenous community in namibia–a new approach to explore food security of rural people in developing countries," *Journal of Hunger & Environmental Nutrition*, pp. 1–20, 2019.

[33] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[34] A. Camargo and J. Smith, "Image pattern classification for the identification of disease causing agents in plants," *Computers and Electronics in Agriculture*, vol. 66, no. 2, pp. 121–125, 2009.

[35] M. G. Selvaraj, A. Vergara, H. Ruiz, N. Safari, S. Elayabalan, W. Ocimati, and G. Blomme, "Ai-powered banana diseases and pest detection," *Plant Methods*, vol. 15, no. 1, p. 92, 2019.

[36] M. G. Selvaraj, A. Vergara, F. Montenegro, H. A. Ruiz, N. Safari, D. Raymaekers, W. Ocimati, J. Ntamwira, L. Tits, A. B. Omondi, *et al.*, "Detection of banana plants and their major diseases through aerial images and machine learning methods: A case study in dr congo and republic of benin," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 169, pp. 110–124, 2020.

[37] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, B. Kamel, and A. Moussaoui, *Deep Learning for Plant Diseases: Detection and Saliency Map Visualisation*, pp. 93–117. 06 2018.

[38] A.-h. Zhang, H. Sun, S. Qiu, and X.-j. Wang, "Nmr-based metabolomics coupled with pattern recognition methods in biomarker discovery and disease diagnosis," *Magnetic Resonance in Chemistry*, vol. 51, no. 9, pp. 549–556, 2013.

[39] B. Felson, "A new look at pattern recognition of diffuse pulmonary disease," *American Journal of Roentgenology*, vol. 133, no. 2, pp. 183–189, 1979.

[40] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," *Artificial Intelligence in medicine*, vol. 23, no. 1, pp. 89–109, 2001.

[41] P. Sajda, "Machine learning for detection and diagnosis of disease," *Annu. Rev. Biomed. Eng.*, vol. 8, pp. 537–565, 2006.

[42] M. Bannayan and G. Hoogenboom, "Using pattern recognition for estimating cultivar coefficients of a crop simulation model," *Field Crops Research*, vol. 111, no. 3, pp. 290–302, 2009.

[43] G. Camps-Valls, L. Gómez-Chova, J. Calpe-Maravilla, E. Soria-Olivas, J. D. Martín-Guerrero, and J. Moreno, "Support vector machines for crop classification using hyperspectral data," in *Iberian Conference on Pattern Recognition and Image Analysis*, pp. 134–141, Springer, 2003.

[44] T. Rumpf, A.-K. Mahlein, U. Steiner, E.-C. Oerke, H.-W. Dehne, and L. Plümer, "Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance," *Computers and electronics in agriculture*, vol. 74, no. 1, pp. 91–99, 2010.

[45] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25* (F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, Curran Associates, Inc., 2012.

[46] A. K. Singh, B. Ganapathysubramanian, S. Sarkar, and A. Singh, "Deep learning for plant stress phenotyping: Trends and future perspectives," *Trends in Plant Science*, vol. 23, no. 10, pp. 883 – 898, 2018.

[47] A. Fuentes, S. Yoon, S. Cheol Kim, and D. Sun Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, p. 2022, 09 2017.

[48] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Computational Intelligence and Neuroscience*, vol. 2016, p. 11, 2016.

[49] A. Ramcharan, P. McCloskey, K. Baranowski, N. Mbilinyi, L. Mrisho, M. Ndalahwa, J. Legg, and D. P. Hughes, "A mobile-based deep learning model for cassava disease diagnosis," *Frontiers in Plant Science*, vol. 10, p. 272, 2019.

[50] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. Hughes, "Using transfer learning for image-based cassava disease detection," *arXiv preprint arXiv:1707.03717*, 2017.

[51] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[52] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pp. 242–264, IGI Global, 2010.

[53] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," *CoRR*, vol. abs/1812.04948, 2018.

[54] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. P. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," *CoRR*, vol. abs/1609.04802, 2016.

[55] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.

[56] G. Bradski and A. Kaehler, "Opencv," *Dr. Dobb's journal of software tools*, vol. 3, 2000.

[57] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International journal of computer vision*, vol. 88, no. 2, pp. 303–338, 2010.

[58] D. L. Henry Ruiz, "Falconcv, an open-source transfer learning library that offers developers an interface to interact with some of the most popular computer vision frameworks," jun 2020–.

[59] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pp. 265–283, 2016.

[60] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, *et al.*, "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7310–7311, 2017.

[61] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*, pp. 740–755, Springer, 2014.

[62] D. Mané *et al.*, "Tensorboard: Tensorflow's visualization toolkit, 2015."

[63] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.