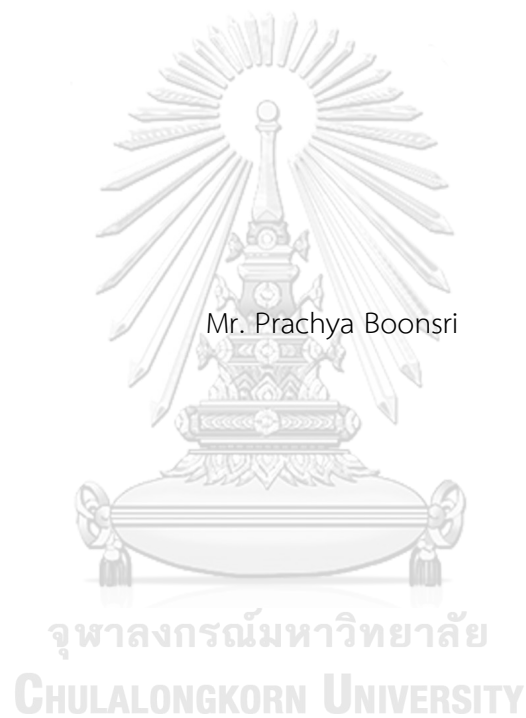


Real-time Object Detection for Screening Cannabis Seed Gender



A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science in Computer Science

Department of Computer Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2022

Copyright of Chulalongkorn University

การตรวจหาวัตถุแบบทันทีเพื่อคัดกรองเพศเมือลี้กัฎษา



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต  
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์  
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย  
ปีการศึกษา 2565  
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title                      Real-time Object Detection for Screening Cannabis Seed  
Gender  
By                                      Mr. Prachya Boonsri  
Field of Study                      Computer Science  
Thesis Advisor                      Associate Professor Yachai Limpiyakorn, Ph.D.

---

Accepted by the FACULTY OF ENGINEERING, Chulalongkorn University in  
Partial Fulfillment of the Requirement for the Master of Science

----- Dean of the FACULTY OF  
ENGINEERING  
(Professor SUPOT TEACHAVORASINSKUN, D.Eng.)

THEESIS COMMITTEE

----- Chairman  
(Assistant Professor SUKREE SINTHUPINYO, Ph.D.)

----- Thesis Advisor  
(Associate Professor Yachai Limpiyakorn, Ph.D.)

----- External Examiner  
(Paskorn Apirukvorapinit, Ph.D.)



ปรัชญา บุญศรี : การตรวจหาวัตถุแบบทันทีเพื่อคัดกรองเพศเมล็ดกัญชา. ( Real-time Object Detection for Screening Cannabis Seed Gender) อ.ที่ปรึกษาหลัก : รศ. ดร.ญาใจ ลีมีปะยกรรม

การรับรู้และความเข้าใจในกัญชาได้ขยายเปิดกว้างมากขึ้นกว่าที่ผ่านมา ผู้ปลูกกัญชาส่วนใหญ่เกือบทั้งหมดต่างให้ความสนใจการปลูกเพื่อให้ได้ช่อดอกขนาดใหญ่จากต้นตัวเมียเป็นหลัก เนื่องจากที่เอชซี ซีบีดี และสารแคนนาบินอยด์อื่นๆ ซึ่งพบในช่อดอกเพศเมีย มีค่าทางการแพทย์และตลาดภาคส่วนอุตสาหกรรมอื่นๆ ดังนั้น การคัดเลือกเฉพาะเมล็ดเพศเมียจึงเป็นขั้นตอนสำคัญในการเพาะปลูกเพื่อให้ได้สารที่เอชซี ซีบีดี ซึ่งให้ผลตอบแทนกำไรมาก เป็นที่น่าเสียดายที่การปลูกกัญชาในประเทศไทยมักเป็นการเพาะเมล็ดปลูกลงในแจ่งผสมปนทั้งต้นเพศผู้และเพศเมีย โดยต้นตัวผู้จะถูกกำจัดจากแปลงเพาะภายหลังในช่วงก่อนออกดอก การเพาะปลูกที่ปฏิบัติกันมาดังกล่าวก่อให้เกิดต้นทุนค่าใช้จ่ายที่สูงขึ้นและผลที่ตามมาของการสูญเสียทางเศรษฐกิจ แนวทางสมาร์ทฟาร์มมิงที่ใช้เทคโนโลยีเอไอจึงถูกนำมาใช้สำหรับการคัดกรองเพศของเมล็ดกัญชาก่อนนำไปเพาะปลูก ชุดภาพข้อมูลเมล็ดกัญชาพันธุ์หางกระรอกซึ่งเป็นสายพันธุ์ไทยที่รู้จักกันอย่างแพร่หลายได้ถูกรวบรวมจากหลายภูมิภาค เทคนิคการเพิ่มข้อมูลได้ถูกนำมาใช้เพื่อเพิ่มจำนวนตัวอย่างและปรับปรุงคุณภาพของภาพถ่าย แบบจำลอง โยโล5 จำนวนสองแบบจำลองสำหรับการตรวจหาวัตถุ ได้ถูกสร้างขึ้นจากชุดข้อมูลเริ่มแรกและชุดข้อมูลที่มีการสร้างข้อมูลเพิ่ม พบว่าแบบจำลองที่เรียนรู้บนชุดข้อมูลภาพที่เพิ่มเติมให้ผลลัพธ์ที่ดีกว่า ได้ค่าตัววัดที่สูงกว่า ได้แก่ ค่าความเที่ยงตรง 96.4%, การเรียกกลับ 97.4% และ เอ็มเอพี 98.7% ด้วยความเร็วในการตรวจหา 7.2 มิลลิวินาที นอกจากนี้ งานวิจัยยังได้นำเสนอแนวทางกำหนดค่าบรรณนิทัศน์กึ่งอัตโนมัติ โลบารรีโอเพนซีวี ได้ถูกใช้เป็นหลักเพื่ออำนวยความสะดวกการประมวลผลเทคนิคการประมวลผลภาพต่างๆ ในการสร้างค่าบรรณนิทัศน์ตั้งต้นแบบอัตโนมัติ ผลลัพธ์ในเบื้องต้นมีความหวัง โดยสมรรถนะแบบจำลองที่ใช้การสร้างค่าบรรณนิทัศน์ตั้งต้นแบบกึ่งอัตโนมัติที่นำเสนอให้ผลลัพธ์ใกล้เคียงกับแบบจำลองที่ใช้การสร้างค่าบรรณนิทัศน์ด้วยมือ อีกทั้งสามารถลดระยะเวลาประมาณกึ่งหนึ่งจากกระบวนการบรรณนิทัศน์เดิม

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์

ปีการศึกษา 2565

ลายมือชื่อนิสิต .....

ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

# # 6470219221 : MAJOR COMPUTER SCIENCE

KEYWORD: Computer vision, Object detection, Image processing, Cannabis seed gender

Prachya Boonsri : Real-time Object Detection for Screening Cannabis Seed Gender.

Advisor: Assoc. Prof. Yachai Limpiyakorn, Ph.D.

Perception and understanding of cannabis are more expansive than they formerly were. Almost all growers are primarily interested in getting harvests of big flower buds from cannabis female plants since THC, CBD and other cannabinoids are found in female flowers and valuable for medical and industrial market segments. Selecting only female seeds to cultivate is thus an important step to produce THC, CBD profitably. Unfortunately, outdoor cultivation in Thailand traditionally grows regular cannabis seeds that grow up of mixed male and female plants. The male plants will be later spot and eliminated during the pre-flowering stage. This incurs the higher cost of investment and the economic loss consequence. A smart farming approach using AI technology is thus introduced for screening cannabis seed genders before cultivation. A dataset of cannabis seed images of Hang Kra Rog, a well-known Thai cannabis cultivar, was collected from several regions. Data augmentation techniques were carried out to increase the sample size and improve the quality of images. The two object detection models, YOLOv5, were constructed using the initial and augmented datasets. The model trained on the augmented image dataset outperformed the other and achieved the higher precision of 96.4 %, recall of 97.4 %, and mAP<sub>0.5</sub> of 98.7 % with detection speed at 7.2 ms. Moreover, an approach of semi-automated image annotation is also presented in this work. The library of OpenCV is mainly used to facilitate operating various image processing technique to generate the initial image annotation. The preliminary result is promising. The performance of the model using the proposed semi-automated annotation achieved comparable performance to the manual annotation model, and reduced half of the time spent on the annotation process.

Field of Study: Computer Science

Student's Signature .....

Academic Year: 2022

Advisor's Signature .....

## ACKNOWLEDGEMENTS

The completion of this thesis could not have been possible without the assistance and participation of so many people whose names may not all be enumerated. Their contributions are sincerely appreciated and gratefully acknowledged.

First of all, I would like to express deep appreciation particularly to the following. Asst. Prof. Sukree Sinthupinyo and Dr. Paskorn Apirukvorapinit for their time, generous effort, expertise and thoughtful advices during the master's defense and presentation. Mainly, I would like to express a sincere gratitude to my professor, Assoc. Prof. Yachai Limpiyakorn, who always gave guidances, awareness, feedback and valuable discussion throughout the journey of this thesis. I am also grateful for her patience and advices which helped me reach all the milestones along the way.

Additionally, thanks also go to the farmers and grower that supplied and helped support the cannabis expertise that served as the inspiration of the thesis. I also appreciate my classmates for pushing me all along and always standing by me.

Last but not least, thank you to friends who assist and carry me all the moments and also my family for always and ever supporting when I need encouragement and energy. Lastly, thank you to myself for every effort and positive attitude that always raise to myself to overcome the valuable moments of this thesis's journey.

Prachya Boonsri

## TABLE OF CONTENTS

	Page
.....	iii
ABSTRACT (THAI).....	iii
.....	iv
ABSTRACT (ENGLISH).....	iv
ACKNOWLEDGEMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES.....	1
LIST OF FIGURES.....	2
Chapter 1 Introduction.....	5
1.1 Statement of the Problems.....	5
1.2 Objective.....	6
1.3 Scope of Study.....	6
1.4 Research Methodology.....	6
1.5 Contribution.....	7
1.6 Publication.....	7
Chapter 2 Literature Review.....	8
2.1 Cannabis.....	8
2.2 Object Detection.....	10
2.2.1 Two-Stage Object Detection.....	11
2.2.2 One-Stage Object Detection.....	11
2.3 YOLO (You Only Look Once).....	11
2.4 Data Augmentation.....	14

2.5 Image Processing.....	15
2.6 Image Annotation .....	16
Chapter 3 Research Methodology .....	18
3.1 Cannabis Seed Gathering and Gender Sorting out .....	18
3.2 Image Capture .....	20
3.3 Image Annotation .....	20
3.3.1 Manual Annotation .....	20
3.3.2 Semi-automated Image Annotation .....	22
3.3.2.1 <i>Region of Interests Cropping</i> .....	23
3.3.2.2 <i>Localization and Rectangle Box Drawing</i> .....	23
3.3.2.3 <i>Adjusting Bounding Box and Class Labeling</i> .....	29
3.4 Train-Test Split .....	30
3.5 Data Augmentation .....	31
3.6 Model Construction .....	34
3.7 Model Performance Evaluation .....	39
Chapter 4 Experimental Results and Discussion .....	42
4.1 Performance Comparison between with and without Augmentation .....	42
4.2 Performance Comparison between Manual and Semi-automated Annotation .....	44
4.2.1 Method Evaluation .....	44
4.2.2 Comparison of Model Performance .....	46
Chapter 5 Conclusion .....	47
REFERENCES .....	48
VITA .....	50



## LIST OF TABLES

Table 1 Summary of manual image annotation of cannabis seeds.....	22
Table 2 Selected augmentation techniques.....	31
Table 3 Model performance of original model on test set.....	41
Table 4 Summary of model performance evaluation of the augmented model.....	41
Table 5 Comparisons of model performance between using augmented dataset versus original dataset.....	42
Table 6 Comparison of manual annotation and automated annotation.....	45
Table 7 Comparison of model performance using manual annotation versus proposed method.....	46

## LIST OF FIGURES

Figure 1 Example of cannabis tree (left) and flower (right) .....	8
Figure 2 Cannabis life cycle.....	9
Figure 3 Traits of cannabis seeds: female (left), male (right) [5].....	10
Figure 4 (a) basic architecture of two-stage detection (b) basic architecture of one-stage detection [6].....	10
Figure 5 YOLO detection system [8].....	11
Figure 6 YOLOv5 benchmark [12].....	12
Figure 7 Example of labeled images of Barley seeds [13] .....	12
Figure 8 Example of inference results [13].....	13
Figure 9 Example of images for model training [14].....	13
Figure 10 Process to automatically label augmented image [17].....	15
Figure 11 Results of image processing by OpenCV using thresholding technique [18] 16	
Figure 12 Process flow of implementing cannabis seed gender detection model .....	18
Figure 13 Traits of research female cannabis seeds, (a) is a deep hole at the head, (b) and (c) is black lines around the head .....	19
Figure 14 Traits of research male cannabis seed, (a) is a shallow hole at the head, (b) and (c) is the appearance of non-black line at the head .....	19
Figure 15 Cannabis seed gathering and gender sorting out .....	19
Figure 16 Various seed placement layouts.....	20
Figure 17 Storage structure of annotated image files .....	21
Figure 18 YOLO labeled output format associated with the image annotation.....	22
Figure 19 Flow of Region of Interest Cropping .....	24
Figure 20 Example results of Rol Cropping.....	24

Figure 21 Flow of localization and rectangle box drawing .....	25
Figure 22 Rol grayscale image (left), Otsu's thresholding image (middle), image with rectangle boxes (right) .....	26
Figure 23 Grayscale image without cropping (left), Adaptive thresholding image (middle), image with rectangle boxes (right) .....	26
Figure 24 Comparison of Histogram correlation in case of normal image .....	27
Figure 25 Comparison of Histogram correlation in case of low light image .....	27
Figure 26 Visualization of image with rectangle boxes and filtering parameters.....	28
Figure 27 Example of rectangle box coordinates in pixel (left) and the basic Label Studio JSON format (right) .....	28
Figure 28 Images with initial bounding boxes in Label Studio .....	29
Figure 29 Example of anomaly bounding boxes.....	30
Figure 30 Dataset storage hierarchy .....	31
Figure 31 Original image (left), Horizontal flip (middle) and Vertical flip (right) .....	32
Figure 32 Original image (left) and Cropped image (right) .....	33
Figure 33 Original image (left) and Gamma Correction image (right).....	33
Figure 34 Original (left) Equalize Histogram image (right) .....	33
Figure 35 Original (left) Raise saturation image (right).....	34
Figure 36 Visualization of box loss on training set (a) and validation set (b) .....	35
Figure 37 Visualization of object loss on training set (a) and validation set (b).....	36
Figure 38 Visualization of classification loss on training set (a) and validation set (b). 36	
Figure 39 F1-Confidence curve of original model.....	37
Figure 40 F1-Confidence curve of augmented model .....	37
Figure 41 Example of inferencing results with confidence threshold of 50%.....	38
Figure 42 Example of inferencing results with confidence threshold of 80%.....	39

<i>Figure 43 IoU formula and interpretation</i> .....	40
Figure 44 Example of inferencing result of seeds with Female (F) class .....	43
Figure 45 Example of inferencing result with seeds in Not Female (NF) class .....	43
Figure 46 Example of inferencing result with seeds in mix classes .....	44
Figure 47 Simulation of bounding boxes matching between manual annotation (ground truth) and semi-automated approach.....	45



## Chapter 1

### Introduction

#### 1.1 Statement of the Problems

Over thousands of years, cannabis is one of the oldest plants well-known as natural product used for recreational or medical purpose [1]. As medicine, it is effective to treat and relieve several sicknesses, such as migraine, fatigue, insomnia cancer pain and etc. [2-4]. Currently the knowledge and perceptions of cannabis have been significantly changing and shifting across many areas with legalization controls. The scientific name of cannabis in Thailand is *Cannabis Sativa L.* that can contain under 0.2% of total tetrahydrocannabinol (THC), which is one of the chemical compounds responsible for psychological effects, also one of the important compounds that found and needed in cannabis along with cannabidiol (CBD) and terpenes [2-4]. Since 2022, cannabis in Thailand has been detached from a list of narcotics banned for use or distribution and made it legal for medical and recreational uses. This has led to trigger cannabis into the spot area of business from the upstream to downstream in the agricultural value chain. Many areas: industrial, academic and especially medical have turned attention to the value and market opportunities that use cannabis as the new mixture, for example, infused drinks and medical product. Due to the potential of its compounds, cannabis has become an economically viable plant. Many growers in Thailand are turning to cannabis as one of the yield crops and probably an alternative economically crop. In Thailand, growing cannabis is mostly sprouting and planting outdoor. The most popular choice is *Hang Kra Rog* cultivar. The growers generally buy *regular seeds* for planting, resulting in the mix of growing up to male and female plants. Conventionally, Thai growers do not start planting by screening seed genders. They will wait until the pre-flowering stage to spot and get rid of any males. This causes extra cost and resources. Even worst, males will pollinate female plants and cause low quality harvest.

This thesis thus presents two folds consisting of 1) an approach of applying a computer vision technology for screening the genders of cannabis seeds prior to planting. The YOLOv5 is selected for implementing the object detector to facilitate the real-time detection of cannabis seed genders. 2) a semi-automated image annotation approach to assist and reduce time and effort spent on the labor-intensive task during the object detection process.

## 1.2 Objective

- 1.2.1 To study and develop a real time object detection model for screening cannabis seed gender
- 1.2.2 To study and present a semi-automated image annotation approach to reduce time and effort spending on the laborious image annotation task.

## 1.3 Scope of Study

- 1.3.1 Gather seeds of *Hang Kra Rog* cultivar from different sources
- 1.3.2 Develop an object detection model based on YOLOv5 to detect binary classification as Female (F) and Not Female (NF)
- 1.3.3 Develop a semi-automated image annotation pipeline to facilitate the resource consuming task
- 1.3.4 Evaluate the model and pipeline performance using the metrics: Precision (P), Recall (R), Mean Average Precision (mAP), and Intersect over Union (IoU)

## 1.4 Research Methodology

- 1.4.1 Study and determine the statement of problems.
- 1.4.2 Literature review and study the related work
- 1.4.3 Gather cannabis seeds and create a dataset of cannabis seed images
- 1.4.4 Design the process and pipeline of development
- 1.4.5 Develop object detection model and semi-automated annotation pipeline
- 1.4.6 Evaluate the model performance and the automated pipeline
- 1.4.7 Analyze and summarize the outcome

1.4.8 Publish the research

1.4.9 Complete thesis writing

## 1.5 Contribution

1.5.1 A real-time predictive model useful for gender screening of cannabis seeds before outdoor planting

1.5.2 The pipeline of semi-automated image annotation for the object detection process that would reduce resources spent on the backbreaking annotation task

## 1.6 Publication

Parts of the thesis had been published in the conferences as follows:

1. P. Boonsri and Y. Limpiyakorn, “Object Detection Model for Gender Screening of Cannabis Seeds”, 9th International Conference on Computer Technology Applications (ICCTA 2023), May 10-12, 2023, Vienna, Austria.
2. P. Boonsri and Y. Limpiyakorn, “Semi-Automated Image Annotation for Cannabis Seed Gender Detection Model”, 3rd IEEE International Conference on Software Engineering and Artificial Intelligence (SEAI 2023), Jun 16-18, 2023, Xiamen, China.

## Chapter 2

### Literature Review

#### 2.1 Cannabis

Cannabis is widely found in various habitats ranging from sea level to the Himalayas foothills [1]. *Cannabis sativa* L. is the species name of Cannabis, also known as Marijuana, Weed or Hemp. Nevertheless, Hemp (*Cannabis sativa* L. subsp. *sativa*) and Marijuana (*Cannabis sativa* L. subsp. *indica*) are different in terms of intoxicating. Hemp contains low levels of Tetrahydrocannabinol (THC) but contains high levels of Cannabidiol (CBD) whereas Marijuana is full of THC. In Thailand, Hemp and Marijuana are legally distinguished by the concentration of THC and CBD. Marijuana is the plant which contains THC over 1% and CBD no more than 2%, that is contrary to Hemp [3]. The term of Cannabis used in this research denotes Marijuana. In Thailand, it is evident that Cannabis had been used as the major medical ingredient in two recipes from 1656–1688, referring to the medical guidebook “*Tamra Phra Osot Phra Naray*” [3, 4].

Cannabis is dioecious plants, i.e., the male and female flowers on different plants. Example of cannabis tree and flower as shown in Figure 1.



Figure 1 Example of cannabis tree (left) and flower (right)



The botany of cannabis is biennial plant, small and straight stem, height can be 1.5–4.5 meters by average. The leaves are distributed on stem, palmately shaped and/or serrate like blade around 6–12 leaflets. However, the botany also depends on the variety of cannabis. Product of cannabis can be either seeds or flowers depending on the objective of cultivation [3] which can turn into business values differently. Cannabis life cycle consists of 6 stages as follows: germination, seeding, vegetative, pre-flowering, flowering and seeds stage as shown in Figure 2.

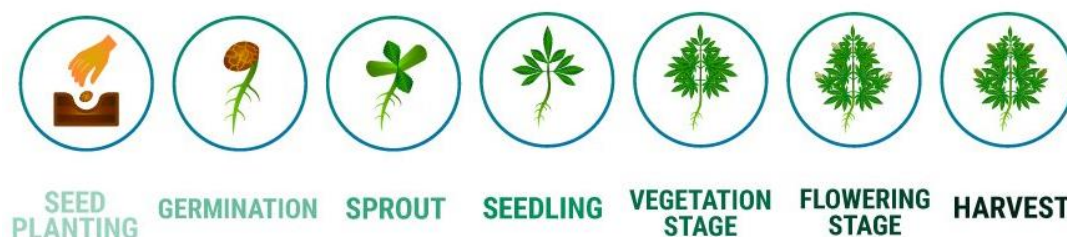


Figure 2 Cannabis life cycle

(<https://greenrushpackaging.com/cannabis-blog/stages-of-marijuana-growth/>)

Hongsuwan et al. [5] studied the distinction of gender of cannabis plants by interviewing six experts. The guidelines gathered from the folk wisdom inform that the genders of cannabis plants could be identified earlier from the traits of seeds. The contour of female seeds looks round chubby, black or dark brown color compared to the paler oval shape of male seeds. Hilum or the top of seeds is an important feature to classify the seed gender. The hilum of female seeds looks like crater with the perfect deep, smooth-edged circle in the middle, and surrounding with black lines. On contrary, the hilum of male seeds looks flat, oval shape and rough-edged like tree root. The distinct traits between female and male cannabis seeds are depicted below in Figure 3. The gender of cannabis plants can be confirmed 6 weeks later after seed germination or in the pre-flowering stage. The appearance of the base of petiole and the mature buds of cannabis are the precise way to distinguish the gender of cannabis plants.



Figure 3 Traits of cannabis seeds: female (left), male (right) [5]

## 2.2 Object Detection

Object detection is a research field of computer vision, deep learning, and artificial intelligence. The discipline is proposed to locate and detect objects from captured visuals through the bounding boxes around particular objects by the computer or machine. The technology was applied to several sectors: autonomous vehicles, face recognition, medical diagnostics, industrial detection etc. The typical architecture of object detection models can be divided into 2 types [6, 7]: 1) two-stage and 2) one-stage as shown in Figure 4.

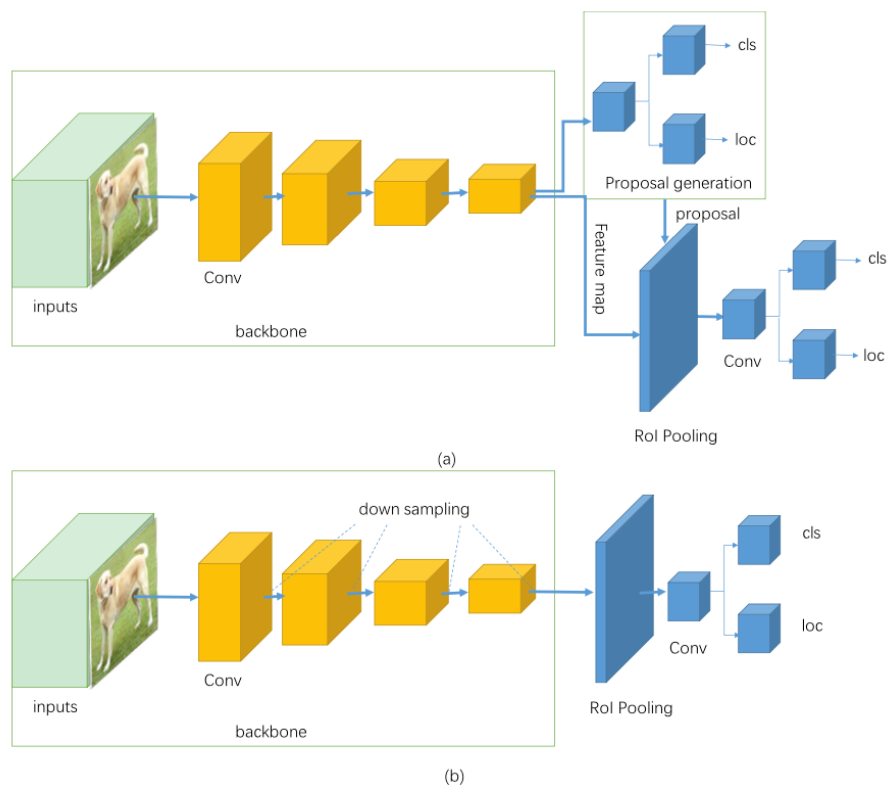


Figure 4 (a) basic architecture of two-stage detection (b) basic architecture of one-stage detection [6]

### 2.2.1 Two-Stage Object Detection

Two-stage Object Detection or region-based architecture separates working process into two tasks: 1) region proposals to extract regions of interest (RoI) chosen as feature; and 2) classifier to classify those candidate regions and bounding-box regression tasks. The two-stage architecture generally provides supreme accuracy. However, the model consists of two parts and requires high computing process. The detection speed is thus slow and not suitable for real-time results. Examples of well-known two-stage object detection models include SSD, R-CNN and family.

### 2.2.2 One-Stage Object Detection

One-stage object detection proposes predicted boxes over the input image directly without region proposals step, making the model run faster for detection, but potentially push the accuracy low. Due to the benefit of speed and accuracy, one-stage object detection model is more suitable for use with mobile or light weight device. YOLO is an example of a famous one-stage detection framework. The algorithm is briefly described in Figure 5.

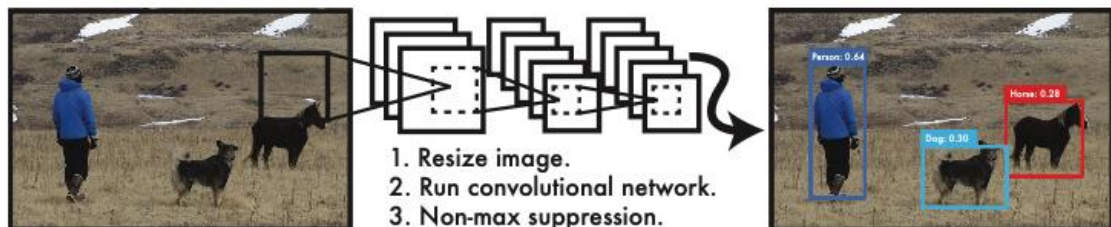


Figure 5 YOLO detection system [8]

## 2.3 YOLO (You Only Look Once)

YOLO was proposed in 2016 by Redmon et al. [8] as one-stage detector. The model is superior to two-stage detector in terms of fast speed with high accuracy. YOLO directly performs regression to detect targets in the image as one-stage. In 2017, Redmon proposed YOLOv2 [9] to improve localization and maintain accuracy. In 2018, Redmon proposed and developed a new version of YOLOv3 [10] which embraced deeper and more complicated network to enable more complex multi-label classification by using Darknet-53 as more robust feature extractor. Later, YOLOv4 was proposed by Bochkovskiy et al. [11] in 2020. In the same year, Ultralytics

announced YOLOv5 [12] with base architecture changed to PyTorch, in order to improve detection speed while maintain detection accuracy. The enhancement also resulted in the smaller size of the ultimate model compared to the previous versions. YOLOv5 utilizes various pretrain weights. The comparison of performance among different benchmarking is shown in Figure 6.

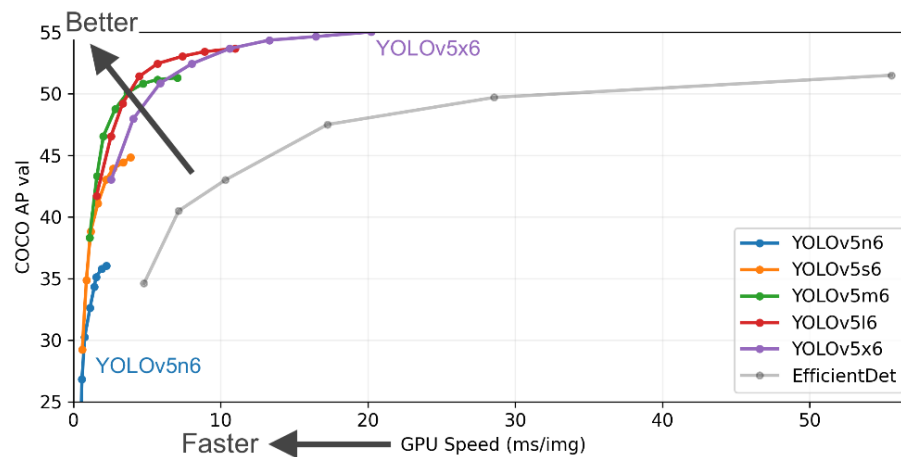


Figure 6 YOLOv5 benchmark [12]

Several research contributes to AI technology for the improvement in the farmer and agricultural sector. Related research of seeds classification using YOLO models includes Barley seed detection and crop seeds classification. Shi et al. [13]. proposed an automated, fast and reliable multi-barley seed detection to assist malting and brewing industries. The model was implemented with YOLOv5x6 to classify nine varieties of Barley seeds. The input images were captured by iPhone. Example of labeled images and the inference results are depicted in Figure 7 and Figure 8, respectively. The model performance achieved the precision of 98.4%, the recall of 98.1%, mAP of 97.5%, and real-time speed detection at 24 ms.



Figure 7 Example of labeled images of Barley seeds [13]

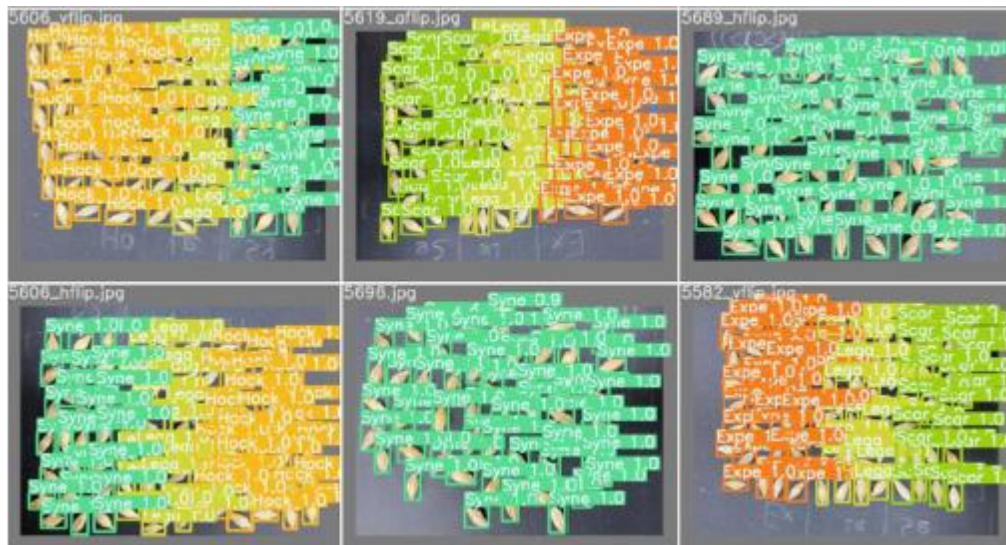


Figure 8 Example of inference results [13]

Kundu et al. [14] proposed an automated approach to classify mixing of crop seeds as maize and pearl millet and also segregate the diseased and healthy seeds of maize for quality testing. The authors developed the deep learning model based on YOLOv5 and conducted data augmentation process to improve the robustness and reliability of proposed model in real time scenario. The model performance achieved the precision and recall of 99% and mAP of 98.3%. Example of images for model training in the research shown in Figure 9.

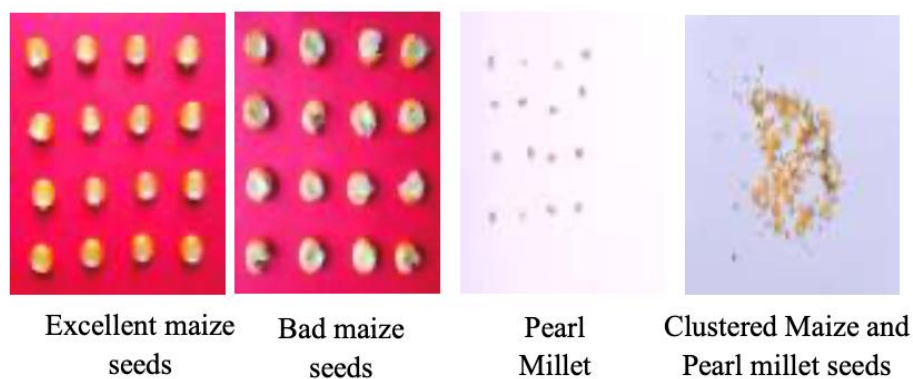


Figure 9 Example of images for model training [14]

In literature, several data augmentation techniques were applied to enlarge the dataset, reduce overfitting, and enhance the model performance. Pang et al. [15] developed an object detector to identify and locate pests based on the improved

YOLOv4 algorithm. The model achieved the average precision of 92.86% and the detection time of a single picture was 12.22 ms. The research work also experimented and analyzed the transfer learning strategy applied for training the model to optimize training iterations and enhance the model performance. The result showed that using the pre-trained weights could improve the convergence speed of the model and enhance locating the pest position in the picture. During the harvesting stage of focusing growth and yields, deep learning techniques achieved better results and helped reducing the losses incurred in the stage. Xie et al. [16] developed the detection model named YOLOv5-Litchi for litchi detection in complex natural environments based on YOLOv5. The YOLOv5-Litchi model outperformed the original YOLOv5 model, achieving the higher recall of 15% and higher mAP of 12.9%. Using the research dataset, the proposed model gained a good balance in accuracy, model size and detection time compared to other detection algorithms consisting of YOLOv4, SSD and Faster-RCNN.

## 2.4 Data Augmentation

Small datasets will lead to overfit during model training, resulting in poor model performance when classifying new or unseen instances. Data augmentation is a common technique applied for increasing the number of samples. The approach consists of several methods to manipulate and generate new training samples from the original dataset, for example, rotating the image, flipping, changing color space, adding noise to the image, modifying the pixel value of the image, etc.

Most studies in agri-technology field that are in computer vision and deep learning task use data augmentation techniques to improve and enhance the model performance [13-16]. CloDSA [17], an open-source library of image augmentation for image localization detection, segmentation and instance segmentation proposed 2 categories of techniques for image data transformation consisting of *position variant* which is modifying the position of the original image for example, rotation and cropping belong to this category. In contrast, position invariant techniques which transformations do not change the position of the object for instance, changing the color palette of the image. Data augmentation for object detection task relies on the

method for object localization and also preserving the class of each bounding box belonging to the original annotation. The augmentation process applied to the object detection problem is shown in Figure 10.

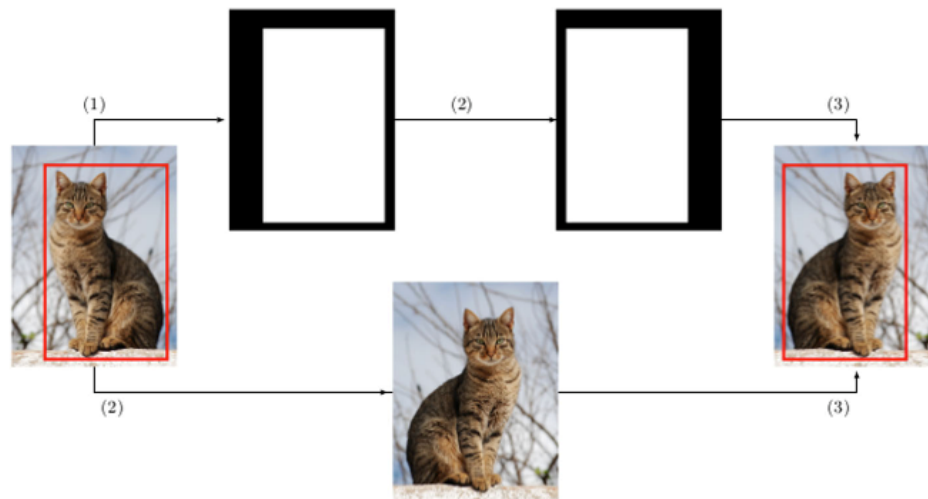


Figure 10 Process to automatically label augmented image [17]

## 2.5 Image Processing

At the core of computer vision is image processing which requires fixed sequences of operations that are performed at each pixel of an original image. A variety of purposes include image analysis, image feature extraction and image pattern recognition. The process covers a wide range of image operations, such as grayscale transformation, edge detection, binary thresholding, feature extraction, etc. Image processing can be applied in several fields, such as medical, agriculture, and industrial system. OpenCV (Open-source Computer Vision Library) [18] is well-known as a fundamental opensource software library for image and computer vision technology which provides extensive algorithms including basic image processing, feature detection, color analysis, contour detection and machine learning. OpenCV provides a common infrastructure for image and computer vision applications. One of the techniques in OpenCV is a common method for image processing, image thresholding which is used to segment each pixel in an image by a threshold value to create binary image between foreground and background. Example of applying thresholding technique for image processing using OpenCV is shown in Figure 11.

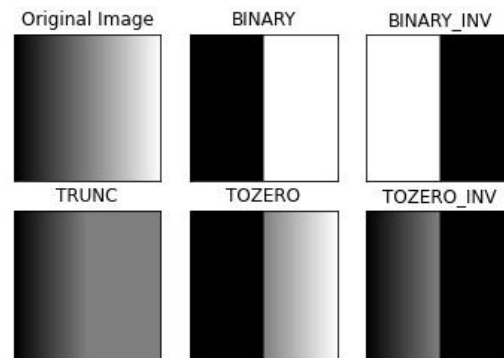


Figure 11 Results of image processing by OpenCV using thresholding technique [18]

## 2.6 Image Annotation

To build a continuous training process of deep neural network model and reduce the resources spent on the most labor-intensive task of object detection which is data/image annotation. Englbrecht et al. [19] developed the solution to fully automate the data annotation process of a custom fluorescent cell nuclei image dataset. The system was based on a combination of image processing steps, for example, thresholding, watershed segmentation, a filtering and post processing tasks applied on a raw image dataset. The research work reported that the time of labeling process can be reduced up to 99.5% and the output of the system provided high quality training data for training segmentation model. Hence, image processing can be used as initial steps to assist object detection models. In literature, image processing is considered as the base method underlying several research works and applications. Huang et al. [20] proposed a text extraction from complicated background and non-uniform lightness situation to identify the text area using several steps of image processing techniques, such as adaptive threshold operation and connected-component. The work achieved high performance of both object identification rate (OIR) and recognition rate (RR) at 81.17% and 91.30% respectively and indicated that adaptive threshold operation performs much better than that of the global threshold particularly, for non-uniform lightness [20]. Contour detection is a technique in image processing to find and detect the boundaries of the objects in the image which have the same color or intensity. It was commonly applied as the based method of image research works. For instance, Ni et al. [21] proposed



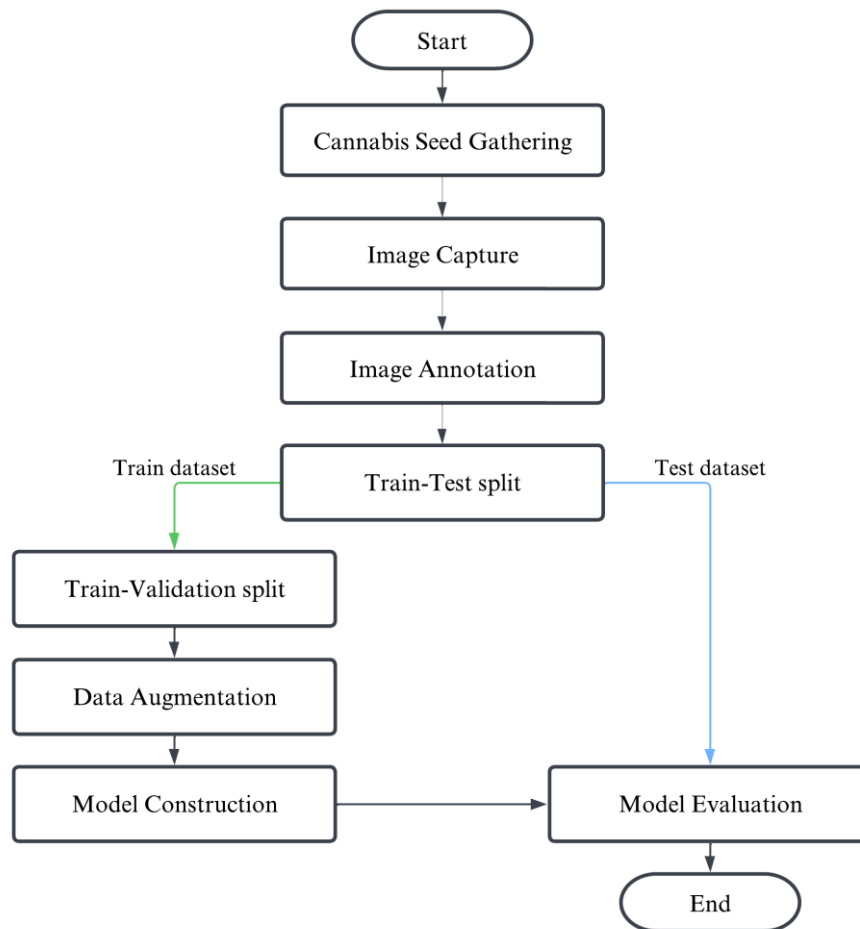
automatic detection and counting of circular shaped using contour detection and circular hough transform. The result of proposed method successfully segmented all the circular shaped objects and outperformed Matlab in processing speed. Similarly, Sakshi et al. [22] developed a method to localize and extract handwritten mathematical expression in which contour detection was the main processed method in the research work.



## Chapter 3

### Research Methodology

The object detection model for gender screening of cannabis seeds was developed using YOLOv5 following the major steps as shown in Figure 12.



*Figure 12 Process flow of implementing cannabis seed gender detection model*

#### 3.1 Cannabis Seed Gathering and Gender Sorting out

Initially, the regular seeds of Thai *Hang Kra Rog* cultivar have been gathered from four different regions. Figure 13 and Figure 14 illustrate traits of female and male seeds from the photo images taken in this research. Thanks to the expert that visually inspected and separated the seeds into two groups of Female and not Female as illustrated in Figure 15. Next, the seeds were put on trays following the pre-determined design layouts with the gender mark. The layout could be the whole

tray of female or male seeds, or mixed genders. Each tray may contain 5 to 40 seeds. Examples of seed placement layouts on trays are shown in Figure 16.

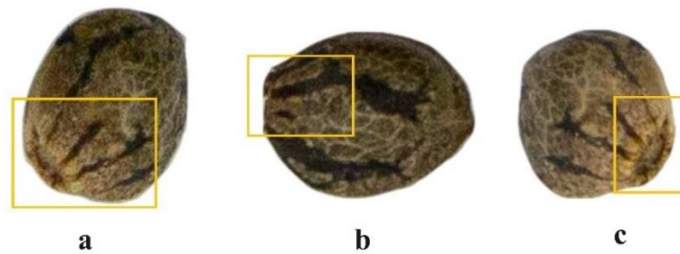


Figure 13 Traits of research female cannabis seeds, (a) is a deep hole at the head, (b) and (c) is black lines around the head

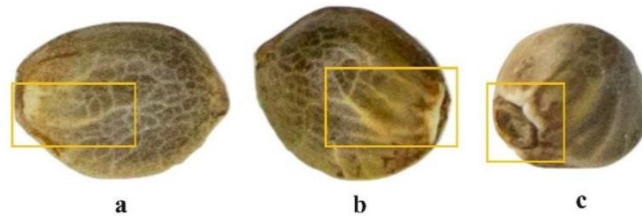


Figure 14 Traits of research male cannabis seed, (a) is a shallow hole at the head, (b) and (c) is the appearance of non-black line at the head

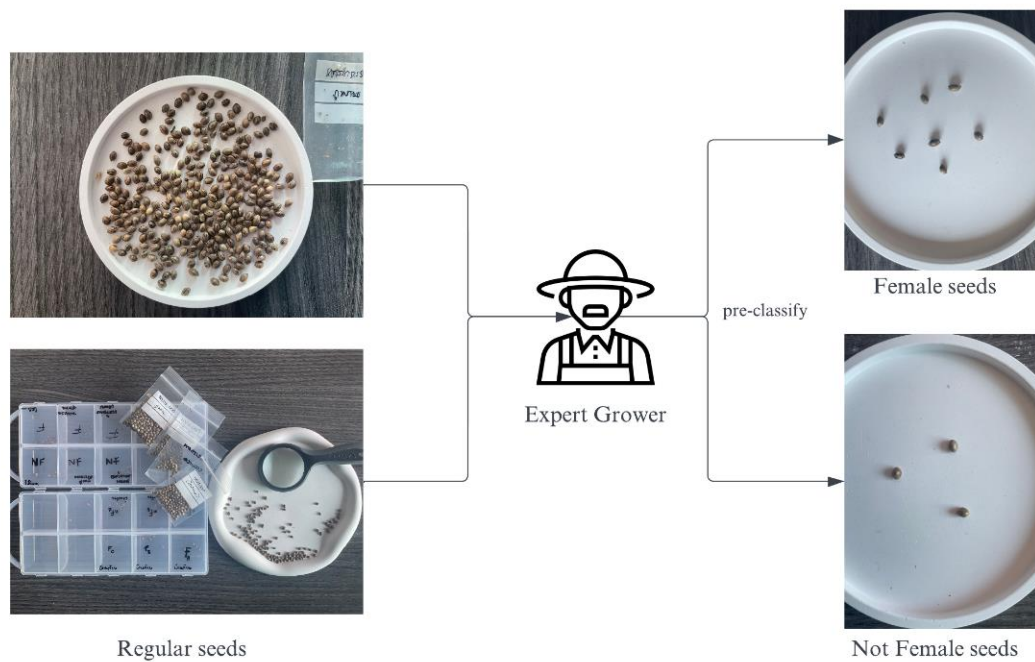


Figure 15 Cannabis seed gathering and gender sorting out



*Figure 16 Various seed placement layouts*

### 3.2 Image Capture

The photo images of trays containing seeds were taken by a smartphone – iPhone 12 Mini, using ratio 16:9 to obtain high resolution in typical conditions of days. The initial 370 photos in .HEIC format were saved and synchronized to Photos, cloud image program of the mac operating system, then exported to .JPEG format without modification.

### 3.3 Image Annotation

#### 3.3.1 Manual Annotation

The output images from the previous step were manually annotated using an open-source, Label Studio [23], which is a platform to manage and facilitate the task of image annotation. Each seed was annotated on two classes: 1) Female (F), and 2) Not Female (NF). The annotated images were stored and organized as folder hierarchy shown in Figure 17.

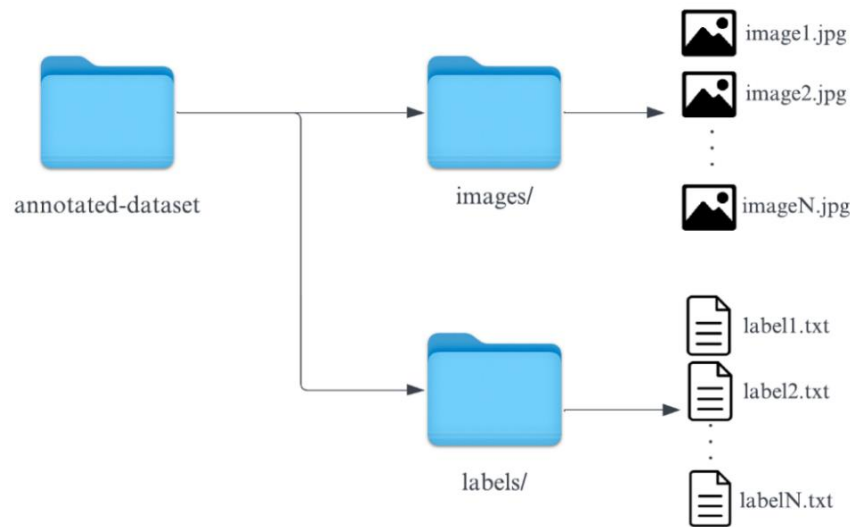


Figure 17 Storage structure of annotated image files

Figure 18 illustrates an example of bounding box associated with the YOLO labeled output format. The annotation output consists of three groups with five values as following:

- 1) *Object Class* denotes the class or category of the object. Here, Female class is assigned to 0, while not Female class is assigned to 1.
- 2) X and Y denote the x and y coordinate of the center of the bounding box, which is normalized between 0 and 1, relative to the grid cell location.
- 3) Width and Height denote the dimensions of width and height of the bounding box that are normalized between 0 and 1, relative to the image size.

The bounding boxes were placed as close to the outline of the seed as possible. Each cannabis seed is bounded by an individual rectangle. In total, 10605 boxes (100%) were created for the dataset of 370 images, separated into 5384 (51%) boxes labeled as a F: Female and 5221 (49%) boxes labeled as NF: Not Female. Observing that the distribution between two classes is balanced dataset. The summary of manual image annotation is described in Table 1.

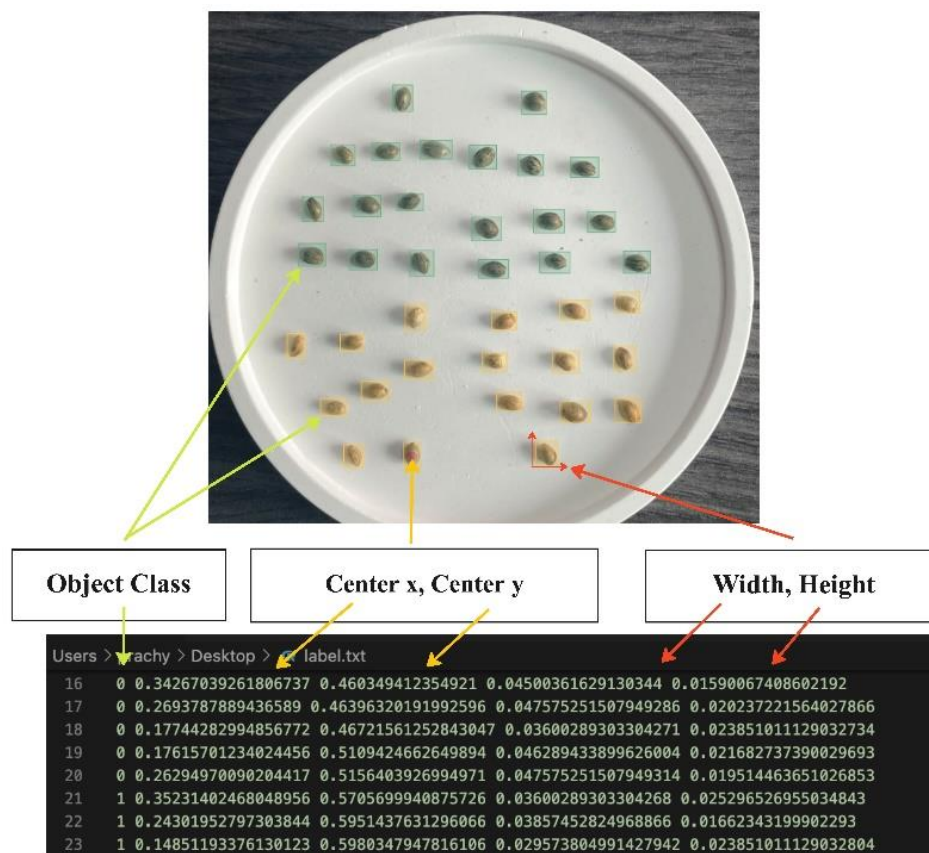


Figure 18 YOLO labeled output format associated with the image annotation

Table 1 Summary of manual image annotation of cannabis seeds

Total Number of Annotations	F: Female	NF: Not Female
10605 (100%)	5384 (51%)	5221 (49%)

### 3.3.2 Semi-automated Image Annotation

Image processing techniques are introduced to enhance the performance of manual annotation considered as a labor-intensive task. A sequence of operations is pipelined to automatically create the initial object bounding boxes. The bounding boxes created manually are used as ground truth to verify the presented approach of automated annotation. And the anomalies will be corrected manually afterwards. Both manual and automated annotations are developed and implemented on cloud environment via Google Colab. Due to the ease of maintaining and integrating with Colab, the dataset and configuration files are stored in Google Drive. The details of the pipelined operations are described in the following subsections.

### 3.3.2.1 Region of Interests Cropping

Starting from determining a particular region of an image that contains seeds. The background and unrelated pixels will be removed resulting in reduced size of the original images. As a result, the total size of the original dataset decreases from 2 GB to 0.6 GB. Figure 19 illustrates the workflow of cropping the region of interests (RoI) that consists of the following tasks.

- 1) The color of original images is transformed from RGB to grayscale [20, 21] to emphasize the depth and strength of the shadows and increase contrast.
- 2) Smooth out noise using median blurring method [20] that is fast and edge preserving
- 3) Apply Histogram Correlation [24] comparison to separate under light images
- 4) Otsu's Binarization Thresholding is applied to enable cropping only the region of interest, i.e., to distinguish between the white tray (region containing seeds) and other unnecessary parts.
- 5) Perform contour detection [21, 22] to detect the border of white tray
- 6) Circle Hough transform is specifically used to determine the circular shape [21] which expected to be the white tray from the max area result of contour detection.
- 7) Cropping only the max area circle which is the interesting region of circular white tray. Some results of the operation are illustrated in Figure 20.

### 3.3.2.2 Localization and Rectangle Box Drawing

The output of the previous step consists of grayscale images with RoI cropping and without RoI cropping (under light condition). Figure 21 illustrates the sequence of operations responsible for localization and rectangle box drawing. The RoI cropping images branch to Otsu's Binarization thresholding that enables automatic computation of threshold for object segmentation as shown in Figure 22. While the under light condition images proceed to Adaptive thresholding [20] which outperforms in low light condition to distinguish objects in pixel-based images as shown in Figure 23. The subsequent steps are described as follows.

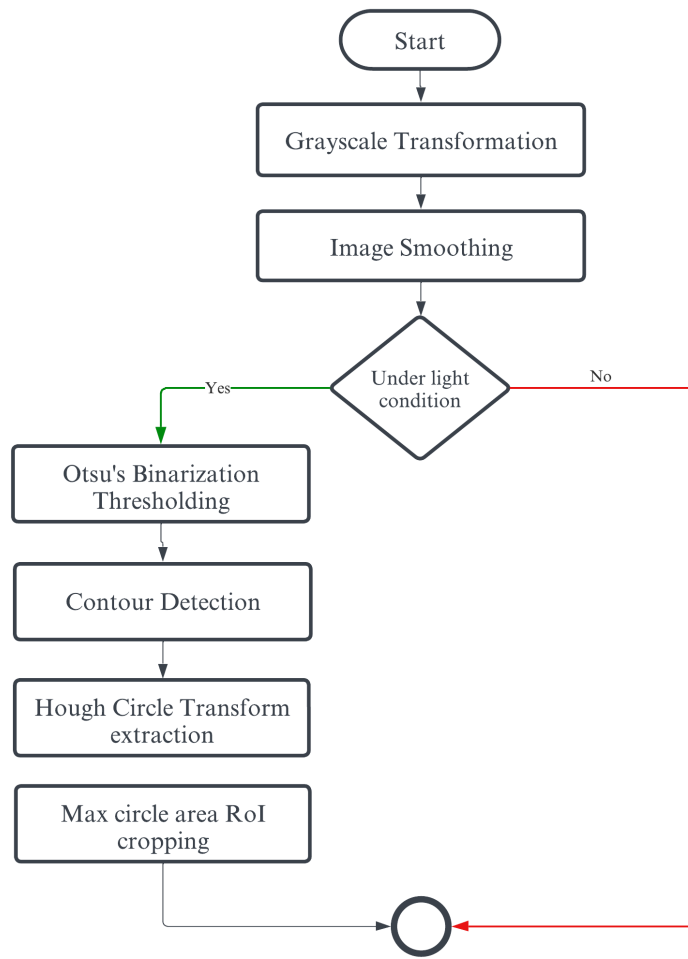


Figure 19 Flow of Region of Interest Cropping



Figure 20 Example results of ROI Cropping



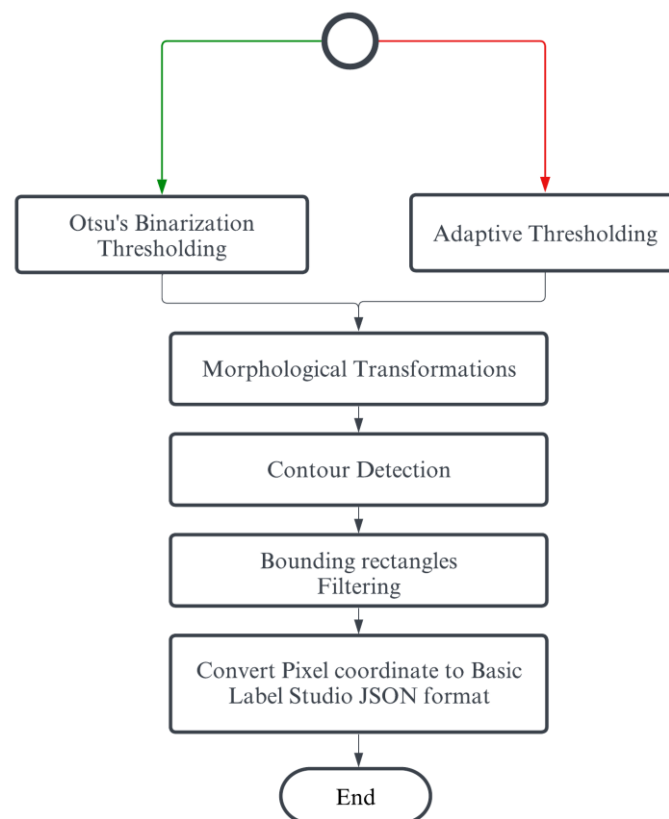


Figure 21 Flow of localization and rectangle box drawing

- 1) Morphological transformation includes dilation and erosion. Dilation is applied for filling the edge of cannabis seeds, while erosion is used for noise removal and only substantive objects remain.
- 2) Contour detection applies `cv2.RETR_TREE` for finding all promising contour lines and reconstructing a full hierarchy of nested contours. The results of contour detection are then performed using `minEnclosingCircle()` to retrieve the center coordinate of the detected results.
- 3) Predefined radius is used to expand the size of detected circle to approximately cover the regular seed.
- 4) Rectangle Boxes are created based on the predefined circle, then filtering is then performed to sort out those boxes with size of too big or too small that are then evaluated by Histogram Correlation whether the box contains a seed. The correlation between pixel intensity distribution of the low light image and that of the base image of low light seed will be calculated and

evaluated against the predefined threshold to anticipate the existence of seed Figure 24 and Figure 25 illustrate the comparisons of Histogram correlation in cases of normal image and low light image, respectively.

- 5) Finally, the rectangular coordinates in pixel values are converted into basic Label Studio JSON format recognized as locations of seeds in each image.

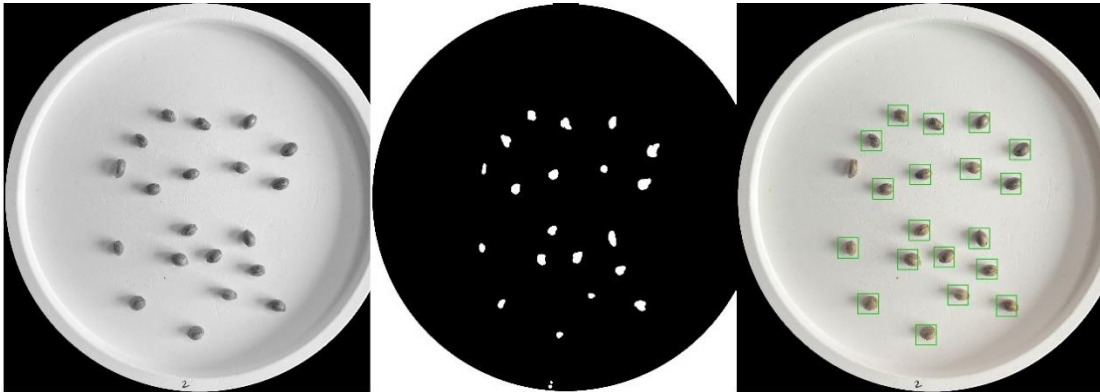


Figure 22 Rol grayscale image (left), Otsu's thresholding image (middle), image with rectangle boxes (right)

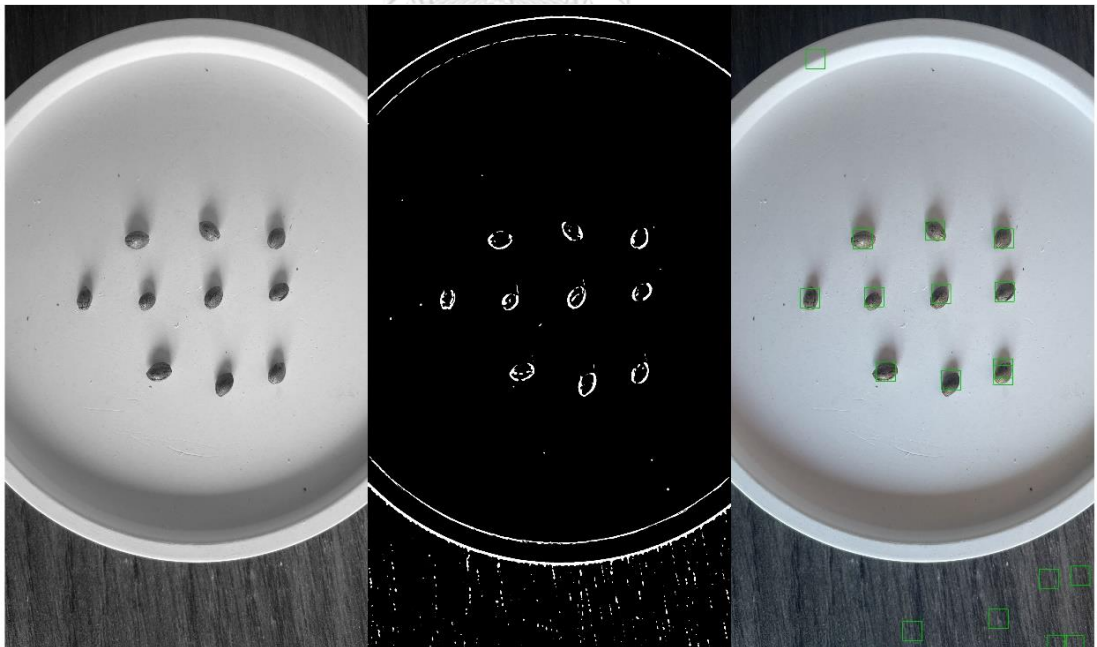


Figure 23 Grayscale image without cropping (left), Adaptive thresholding image (middle), image with rectangle boxes (right)

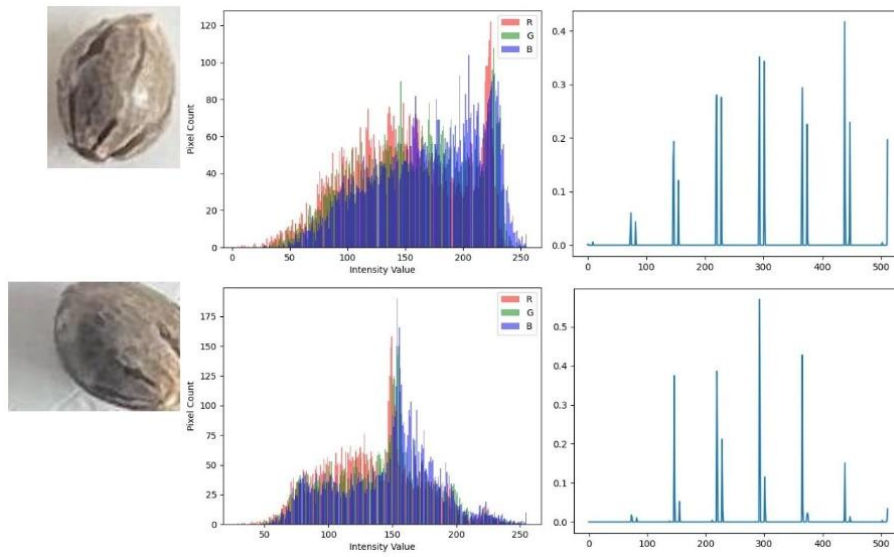


Figure 24 Comparison of Histogram correlation in case of normal image

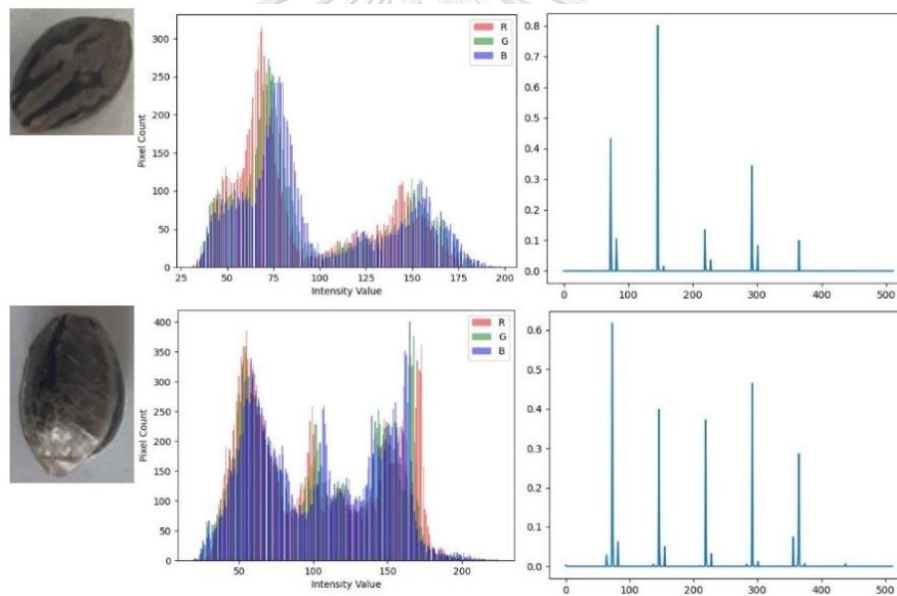


Figure 25 Comparison of Histogram correlation in case of low light image

Figure 26 shows the visualization of image with rectangle boxes and filtering parameters consists of score which calculated by Histogram correlation method, aspect which is the ratio of height divided by width of rectangle box from contour detection and area is the area of rectangle, multiply its height by its width. These parameters are used to filter valid bounding boxes.

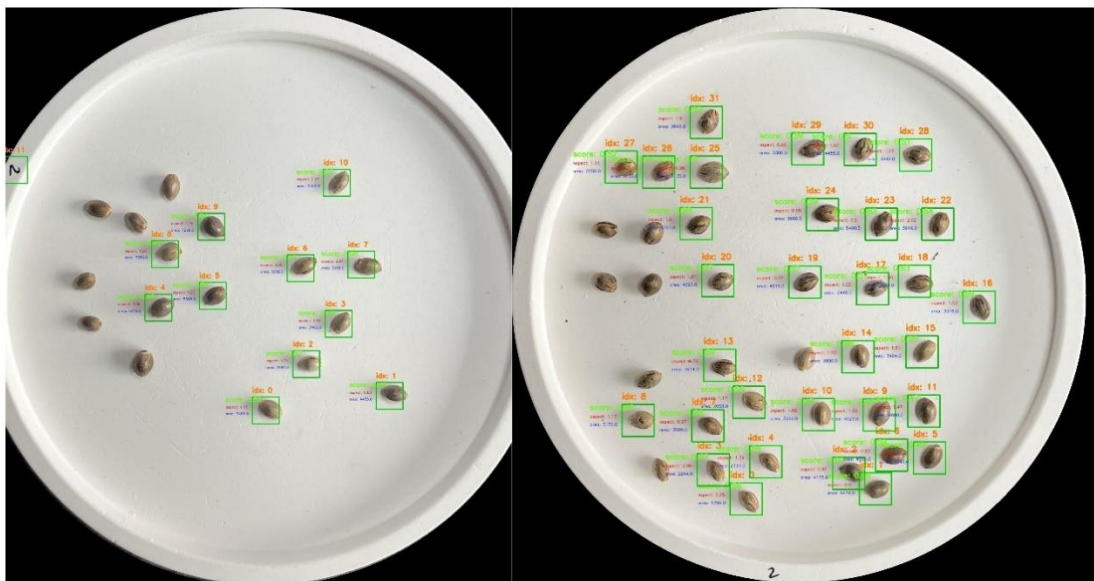


Figure 26 Visualization of image with rectangle boxes and filtering parameters

Label Studio is used to automate the initial bounding boxes. The coordinates of image in pixel value need to be converted to basic Label Studio JSON format. The pixel coordinate as a result of the operation and the basic Label Studio JSON format are illustrated in Figure 27.

```

1f_3nf_72_1of5.txt
Figures Paper > final > crois > pixels2 > 1f_3nf_72_1of5.txt
1 2 1541 1702 130 130
2 2 915 1589 130 130
3 2 1166 1551 130 130
4 2 310 1507 130 130
5 2 1400 1507 130 130
6 2 1240 1356 130 130
7 2 548 1321 130 130
8 2 1415 1142 130 130
9 2 541 1026 130 130
10 2 813 1004 130 130
11 2 1145 1000 130 130
12 2 1387 961 130 130
13 2 1594 923 130 130
14 2 1177 773 130 130
15 2 819 756 130 130
16 2 1325 733 130 130
17 2 1520 665 130 130
18 2 792 522 130 130
19 2 1285 489 130 130
20 2 1463 486 130 130
21 2 671 368 130 130
22

1f_3nf_72_1of5.json
Users > prachy > Library > CloudStorage > GoogleDrive-mrboonsri.p@gmail
1 {
2   "annotations": [
3     {
4       "result": [
5         {
6           "original_width": 2218,
7           "original_height": 2189,
8           "value": {
9             "x": 69.47700631199278,
10            "y": 77.75239835541342,
11            "width": 5.861136158701533,
12            "height": 5.938784833257195,
13            "rotation": 0,
14            "rectanglelabels": [
15              "Null"
16            ]
17          }
18        },
19        "from_name": "label",
20        "to_name": "image",
21        "type": "rectanglelabels",
22        "origin": "manual"
23      ]
24    }
25  ]
26 }

```

Figure 27 Example of rectangle box coordinates in pixel (left) and the basic Label Studio JSON format (right)

### 3.3.2.3 Adjusting Bounding Box and Class Labeling

The output in JSON format from the previous stage will be imported to Label Studio [23] for manual class labeling: Female (F) or Not Female (NF). Prior to assign the relevant class to a bounding box, visual inspection is performed to verify the quality of bounding boxes obtained from the previous stage. Figure 28 shows the imported images to Label Studio.

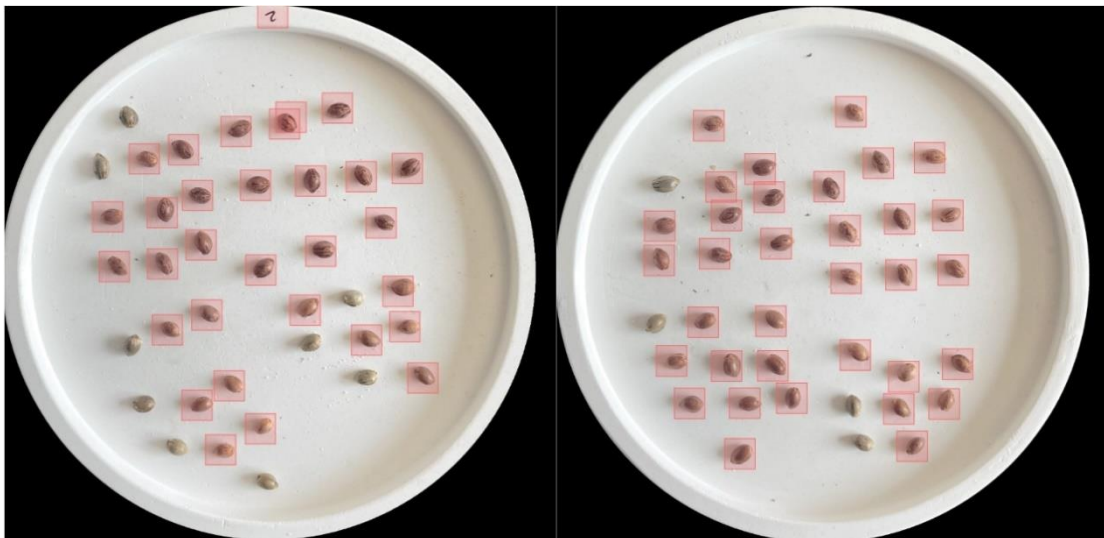


Figure 28 Images with initial bounding boxes in Label Studio

Figure 29 depicts examples of faulty bounding boxes. The red rectangle denotes the first case of fault, the red circle denotes the second case of fault, and the orange rectangle denotes the third case of fault. The anomalies will be managed:

- 1) Remove those boxes without objects inside.
- 2) Fill in a bounding box for existing object but not detected (missing).
- 3) Make fine position adjustments to the box location such as incomplete covering and mis-localized (scaling).

The output images follow the YOLO label format ready as input for training the object detection model.

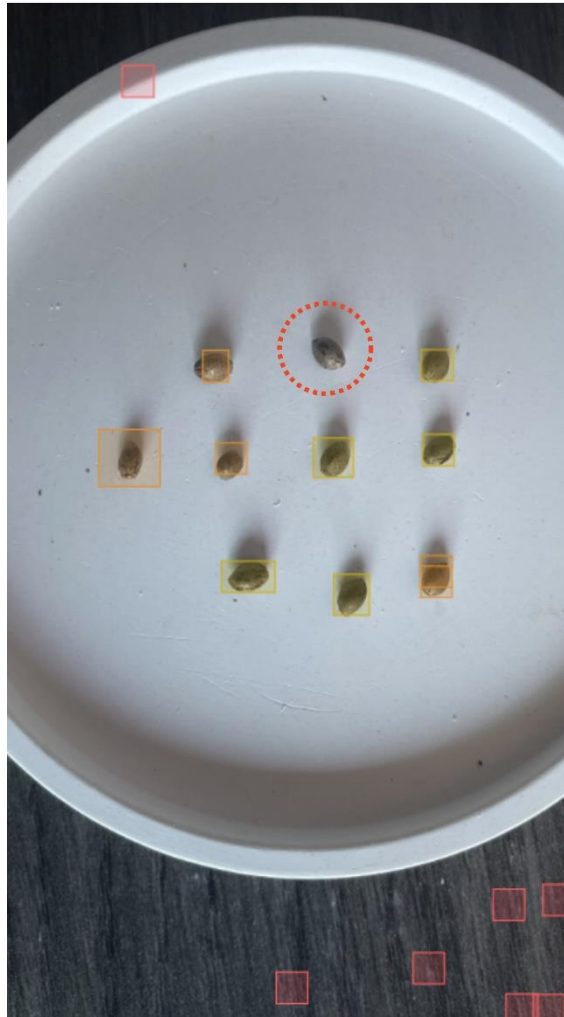


Figure 29 Example of anomaly bounding boxes

### 3.4 Train-Test Split

A total of 370 annotated images were randomly selected and split into the training, and test sets by the ratio 80:20. Next, the train set was split into train: validate using the ratio 80:20. The validation set was created as an unbiased dataset to assess the model in each iteration. The total number of instances contained in the training, test, and validation sets was 236, 74, and 60 images, respectively. Each type of dataset is separately stored in its own folder as shown in Figure 30.

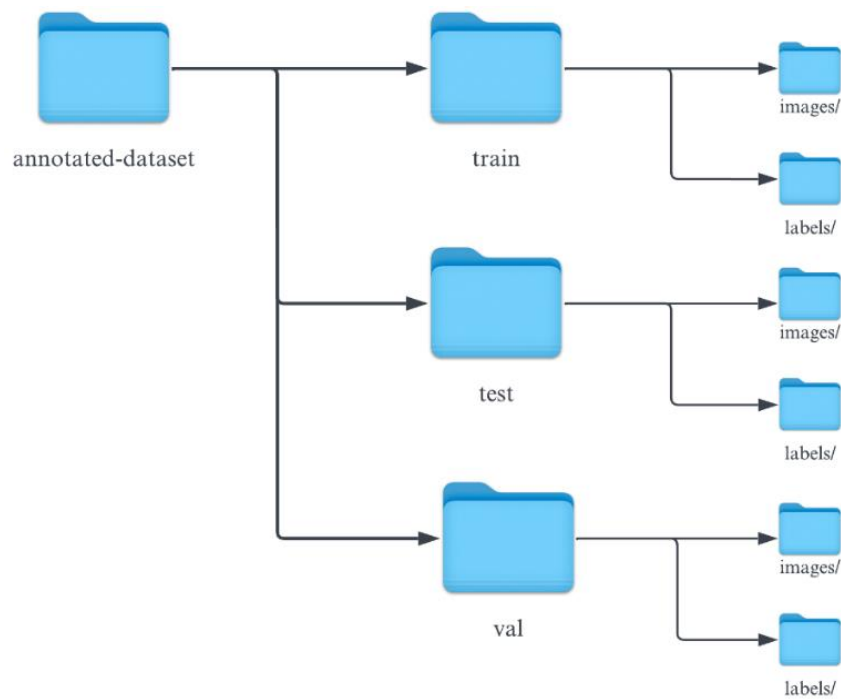


Figure 30 Dataset storage hierarchy

### 3.5 Data Augmentation

Data augmentation is performed to expand the training data for small dataset. An open-source library for image augmentation, CloDSA [17], is used to augment the images in the training set. Starting with applying the Positioning techniques: horizontal flipping, vertical flipping, cropping; and then applying the Enhancing techniques, the total number of images was increased to 2832. The chosen augmentation techniques are summarized in Table 2 and described as following:

Table 2 Selected augmentation techniques

Augmentation Techniques	
Positioning	Coloring and Enhancement
Horizontal Flipping	Gamma correction
Vertical Flipping	Equalize Histogram
Cropping	Raise Saturation
Preserving Original	

- Horizontal and Vertical flipping: The methods to generate new view of image by a mirror-reversal across an axis as horizontal or vertical, example as shown in Figure 31.
- Cropping: The technique to create a random subset of original images that can focus only on certain object's areas, for example as shown in Figure 32.
- Gamma correction: A nonlinear operation used to manipulate luminance and intensity of an image from the differences between the way a camera captures, and the way of human's visual system processes light, example as shown in Figure 33.
- Equalize histogram: The technique used to enhance contrast of an image by stretching out the intensity range of the whole image, example as shown in Figure 34.
- Raise saturation: The technique to increase intensity of the original image to make color more vibrant and distinctive, example as shown in Figure 35.



*Figure 31 Original image (left), Horizontal flip (middle) and Vertical flip (right)*





Figure 32 Original image (left) and Cropped image (right)

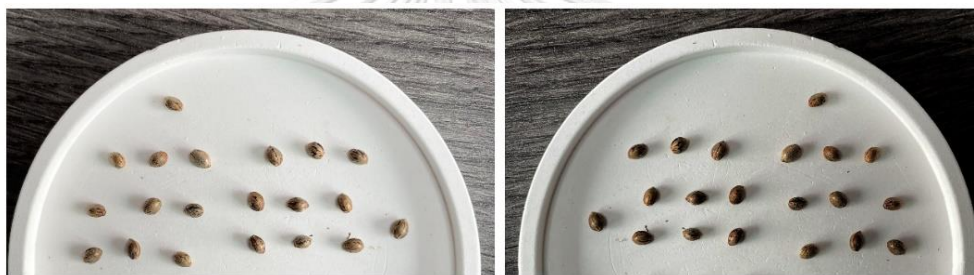


Figure 33 Original image (left) and Gamma Correction image (right)



Figure 34 Original (left) Equalize Histogram image (right)

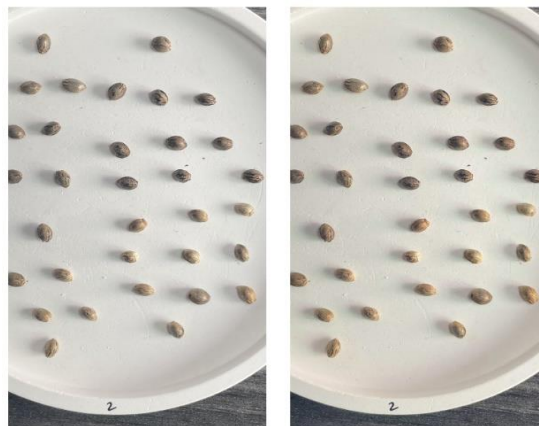


Figure 35 Original (left) Raise saturation image (right)

### 3.6 Model Construction

Due to its outstanding performance in agriculture tasks [13-16], the one-stage object detection model, YOLOv5 is selected for implementing the object detector to facilitate the real-time detection of cannabis seed genders. The prominent point of YOLOv5 is significantly good at object detection challenge in terms of speed and accuracy. Additionally, the model had been specially improved to work with small objects. To increase robustness and speed up training process, transfer-learning method is applied with selected pre-trained model. The pre-trained model YOLOv5n is selected as it is the smallest and most lightweight model in the YOLOv5 series [12].

The model was trained using Google's research computing service, that is, Google Colab environment with specification of GPU = NVIDIA Tesla T4, RAM size = 54.8GB. Due to the ease of maintaining and integrating with Colab, the image dataset (\*.zip) was stored in Google Drive. YOLOv5 model was downloaded from Ultralytics GitHub repository. The pre-trained model *yolov5n* was initiated for training. The hyperparameters were configured to minimize the training loss, consisting of learning rate = 0.01, momentum = 0.9, batch size = 16, the input image resolution was 640 x 640 pixel. The maximum epoch was set to 500 epochs, and early stopping was set to 100 in case of halting model training where no further improvement on validation set. The defaults suggested by the developer of YOLOv5 were applied for the rest.

Regarding YOLO, there are three kinds of losses. The value of loss is calculated in each epoch during training the model.

- 1) *Box loss* indicates how well of the predicted bounding box to cover an object exactly. Figure 36 illustrates two box loss graphs indicating the prediction error during training the two models (original vs. augmented) measured on the training data that are gauged by the prediction error measured on the validation dataset.
- 2) *Object loss* represents the confidence of object presence in predicted bounding boxes. Figure 37 illustrates the two graphs of object loss on training set during training the two models of original and augmented, compared to that on validation set.
- 3) *Classification loss* denotes the probability of predicting the correct class of all given objects. Figure 38 illustrates the two graphs of classification loss on training set during training the two models of original and augmented, compared to that on validation set.

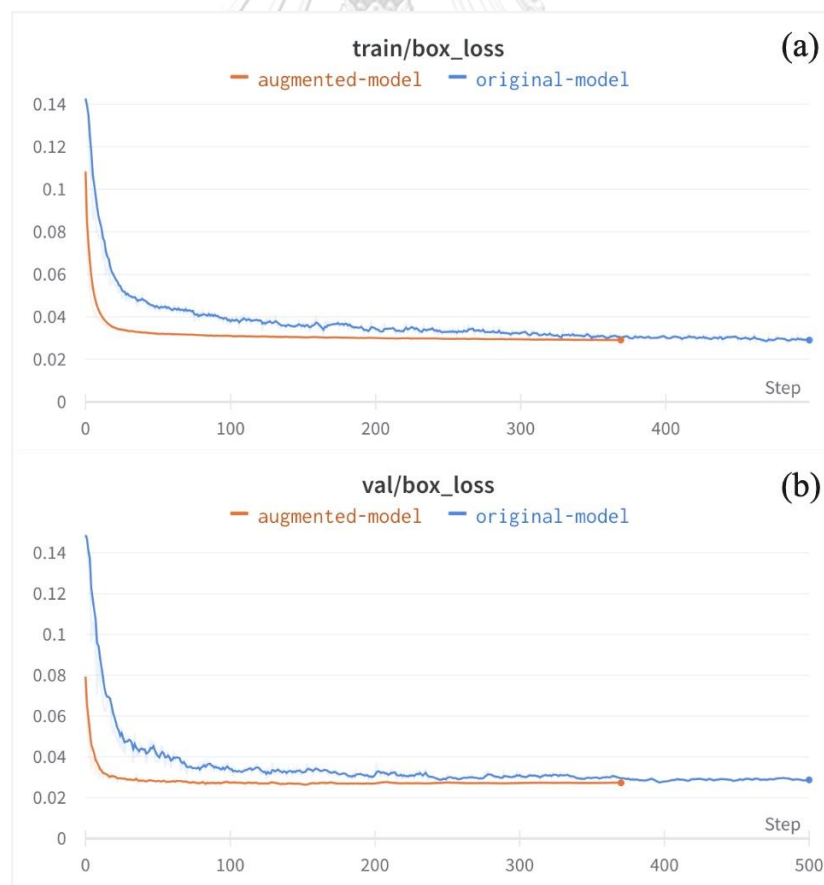


Figure 36 Visualization of box loss on training set (a) and validation set (b)

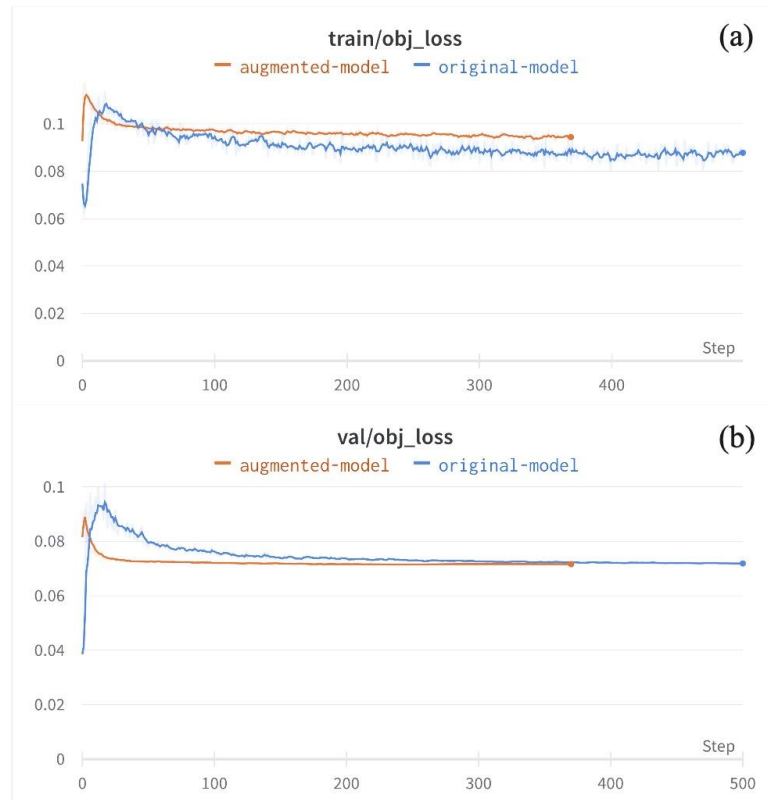


Figure 37 Visualization of object loss on training set (a) and validation set (b)

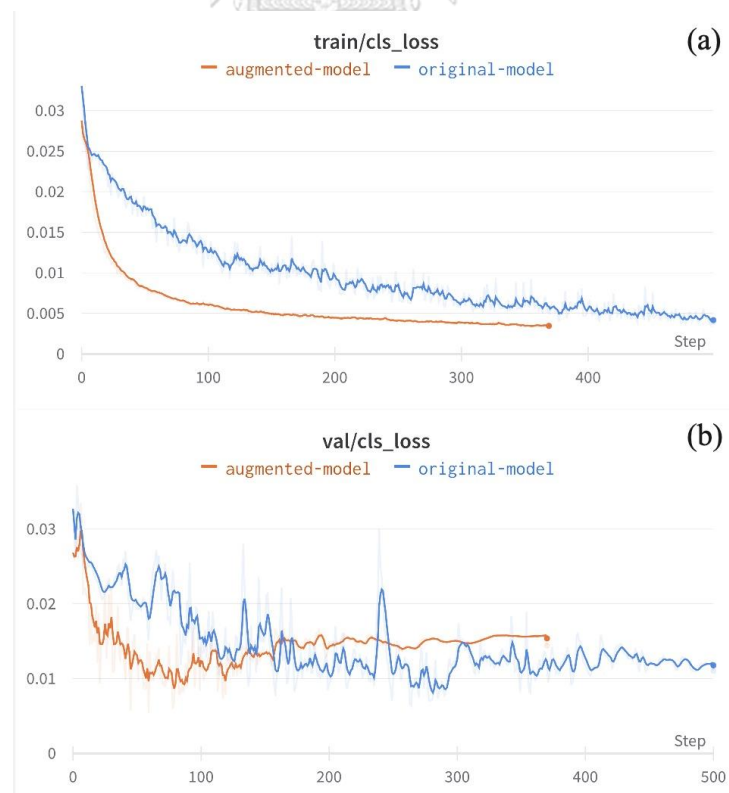


Figure 38 Visualization of classification loss on training set (a) and validation set (b)

The results from Figure 36-38 show that all losses on the augmented model tend to decrease rapidly, lower and gain faster decreasing convergence compared to the loss graphs of the original model. The augmented model was stopped earlier than the original model, at around 370 epochs while there was no early stopping in the original model, indicating the model trained with augmented data could generalize better and more robust than the model trained without augmented data.

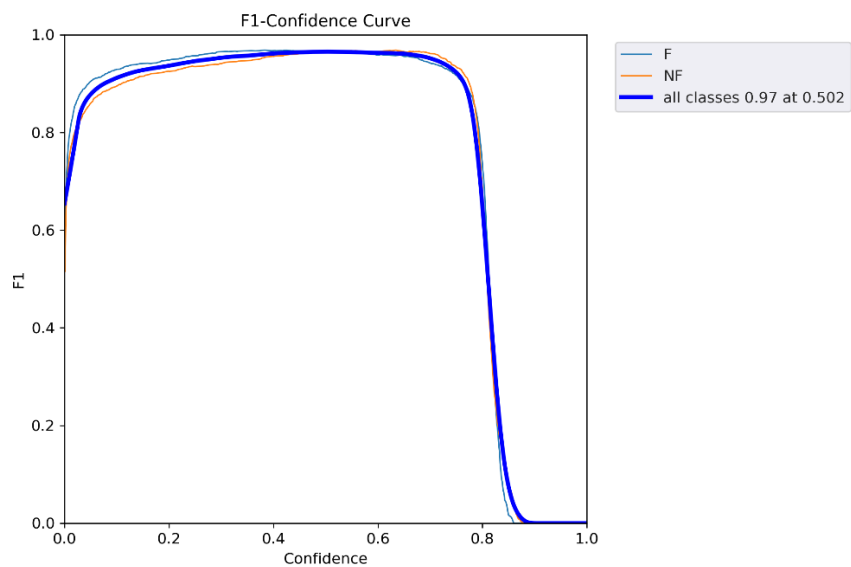


Figure 39 F1-Confidence curve of original model

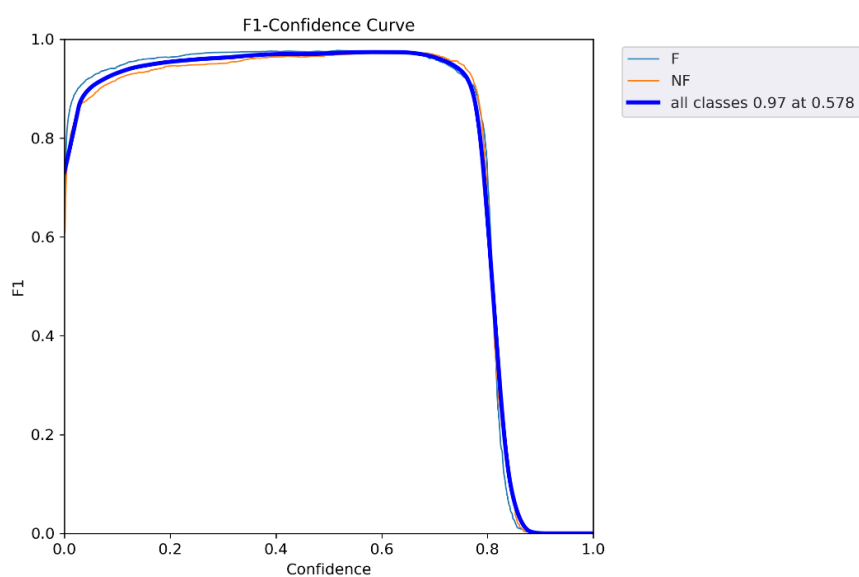
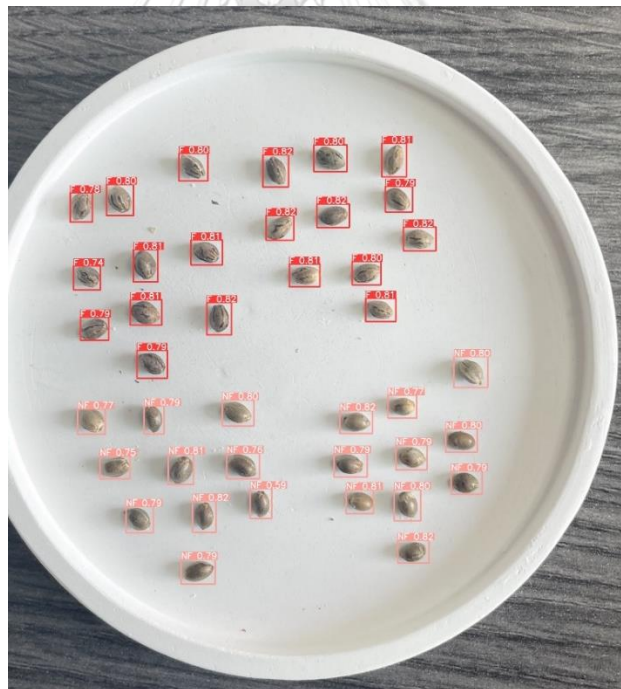


Figure 40 F1-Confidence curve of augmented model

The F1-Confidence curve indicates the trade-off between precision and recall on different confidence threshold values. Figure 39 illustrates the F1 confidence curve resulted from the original model training. The graph shows that the highest confidence threshold that gives the highest F1 score and is the best balance between precision and recall was set around 0.5. Whereas the highest confidence threshold is 0.578 for the augmented model as shown in Figure 40. Examples of inferencing results using different confidence thresholds setting are depicted in Figure 41, 42. Setting the lower confidence threshold enables detection of all objects as shown in Figure 41. There is more possibility to include false predictions though. On contrary, setting the higher confidence threshold will cause more predictions ignored and there is more possibility to fail on detection as shown in Figure 42. Hence, the two models trained on the original and augmented datasets will set the confidence threshold to 50% .



*Figure 41 Example of inferencing results with confidence threshold of 50%*

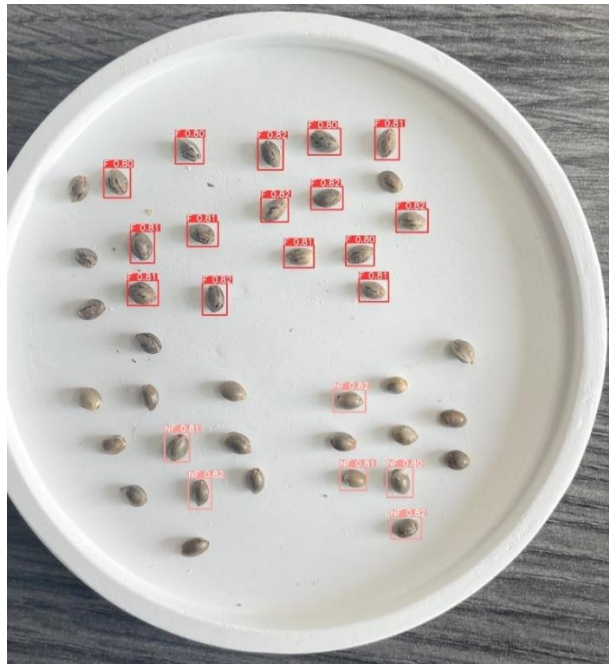


Figure 42 Example of inferencing results with confidence threshold of 80%

Through training, losses were calculated and updated along with the best weight being yielded in each iteration in case there is improvement over previous iteration, best weight will be stored and obtained after end of training. Best weight is going to be used to evaluate against test set for model performance evaluation further.

### 3.7 Model Performance Evaluation มหาวิทยาลัย

The model performance is evaluated by the three metrics consisting of Precision, Recall, and mean Average Precision (mAP). In object detection task, both classification and localization are considered. The overall model performance is measured based on two dimensions. Inferencing result is the output of the model represented by three attributes: the object class, the bounding box and the confidence score. The metric that is used to compare the predicted bounding box versus the ground truth bounding box is Intersection over Union (IoU). IoU measures the overlapped area between ground truth and predicted bounding boxes divided by the union area between ground truth and predicted bounding boxes. The threshold is commonly 50% or in range of 50%-95%. The formula and example of IoU are illustrated in Figure 43.

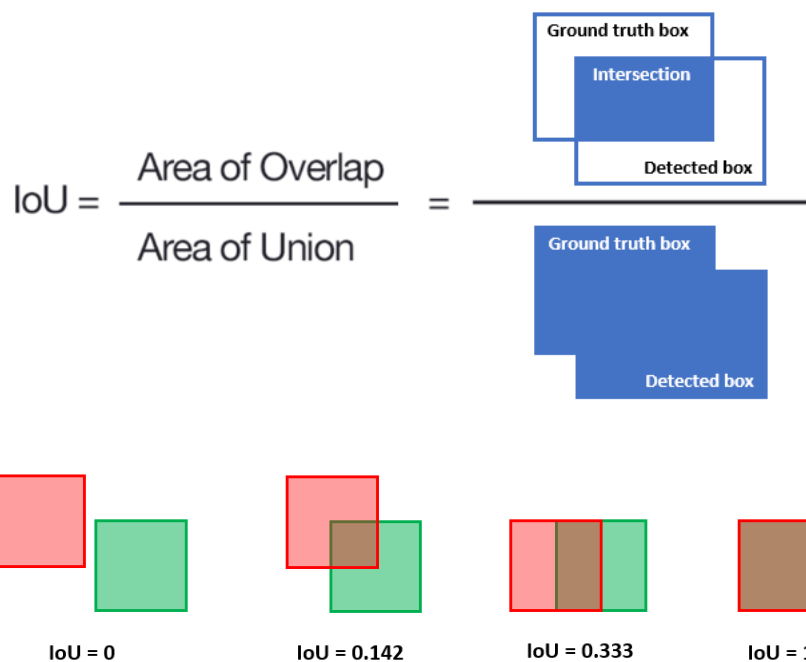


Figure 43 IoU formula and interpretation

(<https://www.baeldung.com/cs/object-detection-intersection-vs-union>)

True positive (TP) is detected when the predicted bounding box with the confidence score of predicted class has the IoU value over the threshold and False Positive (FP) is detected when the predicted result has the IoU value less than the threshold. False Negative (FN) means the model failed to detect the object in the image. The calculation of Precision (P) and Recall (R) is shown in (1) and (2) below.

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

Average precision (AP) denotes the area enclosed by the precision-recall curve (PR-curve). The formula is shown in (3). The mean value of AP from all classes is mAP. The larger value of mAP, the better performance of the model. The formula is shown in (4).

$$AP = \int_0^1 P(R) * dR \quad (3)$$

$$mAP = \frac{1}{k} * \sum_{i=1}^k AP_i \quad (4)$$



where  $k$  = number of classes in the dataset

and  $AP_i$  denotes the calculation of area under PR-curve of class  $i$ .

Table 3 summarizes the model performance evaluation on test set of the original model, achieving Precision of 96.3%, Recall of 96.2%, and mAP at IoU threshold 50% = 98.2%. Table 4 summarizes the model performance evaluation on test set of the augmented model, achieving the values of Precision, Recall, and mAP as 96.4%, 97.4% and 98.7%, respectively.

*Table 3 Model performance of original model on test set*

Class	Precision	Recall	mAP_0.5
All	0.963	0.962	0.982
F: Female	0.973	0.952	0.985
NF: Not Female	0.953	0.972	0.980

*Table 4 Summary of model performance evaluation of the augmented model*

Class	Precision	Recall	mAP_0.5
All	0.964	0.974	0.987
F: Female	0.982	0.959	0.988
NF: Not Female	0.945	0.989	0.985

## Chapter 4

### Experimental Results and Discussion

#### 4.1 Performance Comparison between with and without Augmentation

The prediction results are summarized in Table 5. All metric values of Precision, Recall and mAP are reported for comparison of the model performance between the model trained with original dataset versus augmented dataset

*Table 5 Comparisons of model performance between using augmented dataset versus original dataset*

Class	Precision		Recall		mAP_0.5	
	Original	Augmented	Original	Augmented	Original	Augmented
All	0.963	<b>0.964</b>	0.962	<b>0.974</b>	0.982	<b>0.987</b>
F	0.973	<b>0.982</b>	0.952	<b>0.959</b>	0.985	<b>0.988</b>
NF	0.953	<b>0.945</b>	0.972	<b>0.989</b>	0.980	<b>0.985</b>

The results revealed that the model trained with augmented dataset outperformed and achieved the higher performance values compared to the model trained with original dataset in all metrics, that is, achieving Precision of 96.4% Recall of 97.4% and  $mAP_{0.5} = 98.7\%$ . Figure 44-46 illustrate the inferencing result from the model trained with augmented dataset. The tiny rectangle surrounding each seed is the predicted bounding box and labelled with the predicted object class associated with the value of the confidence score. In this work, the confidence threshold is set to 0.5 for the both models as described in section 3.5.

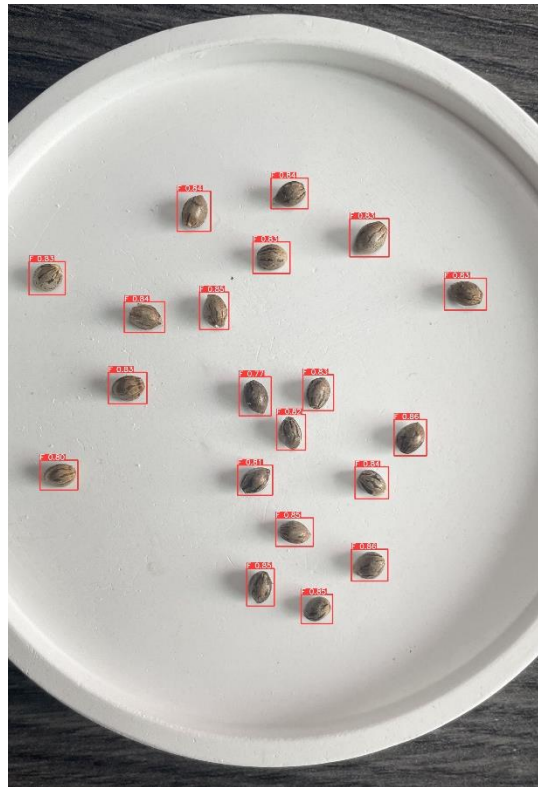


Figure 44 Example of inferring result of seeds with Female (F) class



Figure 45 Example of inferring result with seeds in Not Female (NF) class



*Table 6 Comparison of manual annotation and automated annotation*

	Manual	Automated
Number of localized boxes	10605	10600
Automated localization	N/A	7 min
Image annotation (hrs.)	12	6
Intersection over Union (IoU)	1.000	0.844

Figure 47 illustrates the simulation of bounding boxes matching between manual annotation (ground truth) and semi-automated approach. The boxes which are outlined with blue color represent the bounding boxes of ground truth and the boxes which outlined in orange represent the bounding boxes from the semi-automated approach while the number above those boxes is the index of bounding boxes used for visualization.



*Figure 47 Simulation of bounding boxes matching between manual annotation (ground truth) and semi-automated approach*

#### 4.2.2 Comparison of Model Performance

Table 7 reports the measurement data of model performance trained on the original dataset. The evaluation metrics include Precision, Recall, and mAP\_0.5. The results revealed that the model trained with the images fully annotated by humans outperformed the model trained with the images annotated by the proposed approach. However, the performance of the model constructed by the semi-automated image annotation could achieve comparable results while using only half of the overall time of manual method.

*Table 7 Comparison of model performance using manual annotation versus proposed method.*

Method	Precision	Recall	mAP_0.5
<b>Manual</b>	<b>0.963</b>	<b>0.962</b>	<b>0.982</b>
<b>Automated Localization + Adjustment</b>	<b>0.865</b>	<b>0.860</b>	<b>0.855</b>
<b>Automated Localization</b>	<b>0.654</b>	<b>0.678</b>	<b>0.531</b>

## Chapter 5

### Conclusion

Computer vision has played an important role in agriculture, which is one of the sectors that is affected by digital and emerging technology. Agri-technology is considered data intense approaches to drive agricultural productivity, improve quality, reduce loss, and ultimately increase profitability [25]. This thesis proposes two folds. First, is an AI approach for screening the gender of cannabis seeds. We focus on outdoor cultivation planting with regular seeds. Automated screening the gender of cannabis seeds upfront prior to plant could help reduce loss, lower waste of buying extra seeds and cost to spot and eliminate the male plants in later pre-flowering stage, in addition to lessen quality of female flowers. The experimental results are promising with better precision of 96.4%, recall of 97.4%, and mAP<sub>0.5</sub> of 98.7% with detection speed at 7.2 ms. Second, to build the continuous model development and assist the most time and labor-intensive task in the process of object detection, an approach of semi-automated image annotation was proposed. The first automation part of the segmentation of images to locate cannabis seeds still requires further enhancement. The latter part of manual adjustment of the bounding boxes including the task of class labeling helps achieving the comparable prediction results compared to the model with manual annotation, while merely spending half of the time used for manual annotation.

Further direction of research work would be divided into two parts as follows: 1) Real world applications and implementation and 2) the improvement of automated pipeline for image annotation. More techniques and more accurate methods would be applied to support localization in case of low contrast images due to non-studio photography. 3). Integration of development, to build the fully pipeline of model training, semi-automated annotation approach and model construction should be connected to establish a continuous development.

## REFERENCES

1. Thomas, B. and M. ElSohly, *The Botany of Cannabis sativa L.* 2016. p. 1-26.
2. Woodbridge, M., *A Primer To Medicinal Cannabis An Introductory Text To The Therapeutic Use of Cannabis Acknowledgment.* 2019.
3. Department of Agriculture, T.D., *Cannabis Sativa L. Production Manual for Farmer in Medical and Industrial Purposes.* 2021.
4. Narongvit, T., *Cannabis "Magic medicine of the world".* 2022: Happy Books.
5. Hongsuwan, N., et al., *The Wisdom of Analyzing External Attributes to Determine the Sex of the Cannabis Plant in the Area of Sakon Nakhon Province.* Udon Thani Rajabhat University Journal of Sciences and Technology, 2022. **10**(1).
6. Jiao, L., et al., *A Survey of Deep Learning-Based Object Detection.* IEEE Access, 2019. **7**: p. 128837-128868.
7. Deng, J., et al., *A review of research on object detection based on deep learning.* Journal of Physics: Conference Series, 2020. **1684**(1): p. 012028.
8. Redmon, J., et al., *You Only Look Once: Unified, Real-Time Object Detection,* in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2016, IEEE Computer Society. p. 779-788.
9. Redmon, J. and A. Farhadi, *YOLO9000: Better, Faster, Stronger.* 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: p. 6517-6525.
10. Redmon, J. and A. Farhadi, *YOLOv3: An Incremental Improvement.* ArXiv, 2018. **abs/1804.02767.**
11. Bochkovskiy, A., C.-Y. Wang, and H.-Y.M. Liao, *YOLOv4: Optimal Speed and Accuracy of Object Detection.* ArXiv, 2020. **abs/2004.10934.**
12. Jocher, G., et al., *ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation.* 2022.
13. Shi, Y., et al., *Multi-Barley Seed Detection Using iPhone Images and YOLOv5 Model.* Foods, 2022. **11**(21).



14. Kundu, N., G. Rani, and V. Dhaka, *Seeds Classification and Quality Testing Using Deep Learning and YOLO v5*. 2021. p. 153-160.
15. Pang, H., et al., *A real-time object detection model for orchard pests based on improved YOLOv4 algorithm*. *Sci Rep*, 2022. **12**(1): p. 13557.
16. Xie, J., et al., *Litchi Detection in a Complex Natural Environment Using the YOLOv5-Litchi Model*. *Agronomy*, 2022. **12**(12): p. 3054.
17. Casado-García, Á., et al., *CLoDSA: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks*. *BMC Bioinformatics*, 2019. **20**(1).
18. Bradski, G., *The OpenCV Library*. *Dr. Dobb's Journal of Software Tools*, 2000.
19. Englbrecht, F., I.E. Ruider, and A.R. Bausch, *Automatic image annotation for fluorescent cell nuclei segmentation*. *PLoS One*, 2021. **16**(4): p. e0250093.
20. Huang, C.-M., Y.-K. Lin, and R.-W. Chang, *Apply Adaptive Threshold Operation and Conditional Connected-component to Image Text Recognition*. *Computer Science and Information Technology*, 2014. **2**: p. 87-94.
21. Ni, J., et al., *Automatic detection and counting of circular shaped overlapped objects using circular hough transform and contour detection*. 2016. p. 2902-2906.
22. Sakshi and V. Kukreja, *Segmentation and Contour Detection for handwritten mathematical expressions using OpenCV*. 2022. p. 305-310.
23. Tkachenko, M., et al., *Label Studio: Data labeling software*. 2020. p. %&.
24. Patacchiola, M. *The Simplest Classifier: Histogram Comparison*. 2016 [cited 2023; Available from: <https://mpatacchiola.github.io/blog/2016/11/12/the-simplest-classifier-histogram-intersection.html>].
25. Liakos, K.G., et al., *Machine Learning in Agriculture: A Review*. *Sensors (Basel)*, 2018. **18**(8).

## VITA

**NAME** Prachya Boonsri

**DATE OF BIRTH** 07 July 1994

**PLACE OF BIRTH** Khonkaen

**INSTITUTIONS ATTENDED** B. Eng. (Information Engineering) 1st class honor, King Mongkut's Institute of Technology Ladkrabang  
M.Sc. (Computer Engineer) Chulalongkorn University

**HOME ADDRESS** 393/1193 Elio Del Nest Condominium, Fl. 23, Tower F,  
Udomsuk Rd. Bangna, Bangkok, Thailand 10260

