TOMATO RIPENESS DETECTION USING DEEP LEARNING

YAP YOON HENG

ASSOC. PROF. DR. EILEEN SU LEE MING
School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia
81310 UTM Johor Bahru

UNIVERSITI TEKNOLOGI MALAYSIA

TOMATO RIPENESS DETECTION USING DEEP LEARNING

YAP YOON HENG

A project report submitted in fulfilment of the
requirements for the award of the degree of
Master of Engineering (Computer and Microelectronic Systems)

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

JULY 2022

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

ACKNOWLEDGEMENT

I collaborated on this thesis with many people, including academics, practitioners, and researchers. They have helped me to understand and think more clearly. I want to convey my deep gratitude to Prof. Madya Ts. Dr. Eileen Su Lee Ming, who is my primary thesis supervisor, for her support, direction, criticism, and friendship. This thesis would not have been the same without her ongoing assistance and attention.

In addition, I want to thank my dear family for their unwavering support and encouragement throughout my difficult time at UTM. Their unending support and inspiration gave me the willpower I needed to finish this master's degree. Additionally, I would want to thank my dearest friends and co-workers at Intel. Thank you so much for your unending patience and support as I worked and attended school full-time. Finally, I would like to express my gratitude to everyone who helped me finish my final year project.

ABSTRACT

Tomatoes are fruits with high nutrition and high in fibre; packed with vitamin C, vitamin K1, vitamin B9 and minerals. The global tomato processing market has reached 43.4 million tons in 2021. It is important to determine maturity level of the crops before harvesting to optimize yield. However, manual inspection of ripe tomatoes required huge labour resources and is time consuming. The amount of labour force for fruit harvesting has increased over the years due to increasing demand. Recently, some studies have attempted to evaluate the feasibility of smart agriculture involving machine learning for harvest ripeness detection. However, these works typically used smaller data size, simple dataset with no background or leaves or explored limited machine learning model. Hence, this thesis aimed to identify tomato ripeness detection using two machine learning networks such as Mask RCNN and YOLOv5. Both models were compared based on minimum average precision. The results of these algorithms were benchmarked with previous works in terms of precision and recall. The dataset for this work consisted of 1000 high resolution images (3024 x 4032) with a total of 9063 tomatoes consisting of unripe, half ripe and ripe tomatoes with leafy background to emulate actual environment in a tomato field. The images were annotated with bounding box in VGG image annotator prior to training and testing with the Mask R-CNN, YOLOv5 networks. After that, these images were divided to training, validation and testing set with 80:10:10 ratio and trained using TensorFlow. Parameters such as epochs, step per epochs, learning rate, batch size were tuned to improve training accuracy and reduce training loss. Minimum average precision achieved for Mask R-CNN was 0.903 and YOLOv5 was 0.927. Precision and recall for Mask RCNN was 89.94% and 87.14% respectively. YOLOv5 achieved better precision and recall of 92.72% and 90.87% respectively, which were better compared to Mask RCNN.

ABSTRAK

Tomato merupakan sejenis buah-buahan yang berkhasiat tinggi. Tomato mengandungi nutrisi yang tinggi dalam serat, vitamin C, vitamin K1, vitamin B9 dan mineral. Pada tahun 2021, pasaran pemprosesan global tomato telah mencecah 43.4 juta ton. Oleh itu, penentuan tahap kematangan tomato sebelum menuai adalah sangat penting bagi mengoptimumkan hasil. Kini, ramai yang menggunakan mata kasar dan membuat pemeriksaan secara manual dalam proses penuaian. Keadaan ini memerlukan sumber tenaga manusia yang besar dan amat makan masa. Sejak kebelakangan tahun ini, jumlah tenaga buruh telah bertambah kerana pasaran tomato semakin meningkat. Beberapa kajian yang melibatkan pertanian pintar dalam sektor penuaian tomato telah cuba dilaksanakan dengan menggunakan pembelajaran mesin baru-baru ini. Walau bagaimanapun, kajian-kajian tersebut biasanya menggunakan saiz data yang kecil, ringkas tanpa latar belakang atau daun. Kajian-kajian tersebut juga meneroka rangkaian pembelajaran mesin yang terhad. Justeru, tesis ini bertujuan untuk mengenal pasti pengesanan kematangan tomato menggunakan beberapa seni bina rangkaian pembelajaran mesin seperti Mask RCNN dan YOLOv5. Prestasi rangkaian ini dibandingkan dengan kajian-kajian terdahulu dari segi kejituan purata minimum, kejituan dan perolehan kembali. Set data untuk tesis ini terdiri daripada 1000 imej resolusi tinggi (3024 x 4032) yang melibatkan 9063 biji tomato belum masak, separuh masak dan masak. Imej-imej ini mempunyai latar belakang berdaun. Imej-imej diberi anotasi secara menual dalam menggunakan VGG image annotator. Set data tersebut dibahagikan kepada Latihan, ujian dan pengesahan dalam nisbah 80:10:10. Kejituan purata minimum untuk Mask RCNN mencapai 0.903 dan YOLOv5 mencapai 0.927. YOLOv5 menpunyai kejituan dan perolehan kembali yang lebih tinggi iaitu 92.72% dan 90.87% berbanding dengan Mask RCNN iaitu 89.94% dan 87.14%.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CIELAB	-	Commission Internationale Lab Colour Space
CNN	-	Convolutional Neural Network
FC	-	Fully Connected
FP	-	False Positive
FPN	-	Feature Pyramid Network
FN	-	False Negative
GLCM	-	Gray Level Cooccurrence Matrix
IoU	-	Intersection of Union
JSON	-	JavaScript Object Notations
MRE	-	Mean Relative Error
PAnet	-	Path Aggregation Network
RCNN	-	Region-based Convolutional Neural Network
RGB	-	Red Green Blue
ROI	-	Region Of Interest
RPN	-	Region Proposal Network
SPP	-	Spatial Pyramid Pooling Layer
SSD	-	Single Shot Multibox Detector
SVM	-	Support Vector Machine
TP	-	True Positive
TN	-	True Negative
VGG	-	Visual Geometry Group
YOLO	-	You Only Look Once
mAP	-	Minimum Average Precision

CHAPTER 1

INTRODUCTION

1.1 Research Background

Tomato processing is one of the most internationally diversified agricultural sectors. This is because tomatoes are nutritious and beneficial to human health. Tomatoes are consumed worldwide due to this factor. Tomato production quantities are much higher compared to other crops grown worldwide. It is six times higher than rice and three times more than potatoes [1].

Tomato are common fruits that have high nutrition in fibre, vitamin C, vitamin K1, vitamin B9 and minerals. Usually, tomatoes can come in different maturity and colours such as red, yellow and green. Red represents ripe, yellow represents half ripe and green represents unripe [2]. The global tomato processing market has reached 43.4 million tons in 2021. It is important to determine maturity level of the crops before harvesting to optimize yield. However, manual inspection of ripe tomatoes required huge labour resources and is time consuming. The amount of labour force for tomato harvesting has increased over the years due to increasing demand. The global tomato processing market expects to reach 54.5 million tons in 2027 [3]. As a result, tomato farming is crucial in rural and suburban areas of emerging nations since it can boost the local economy.

When immature or unripe fruits are harvested, the quality of these fruits are poor. Usually, they are incapable of ripening. Immature fruits are susceptible to internal deterioration and decay. Similarly, if fruits are harvested late, the chances of getting rotten fruits are very high. Thus, improper harvest timing will result in drastic postharvest loss. In order to reduce the losses of preharvest and postharvest fruits qualitatively and quantitatively, it is important to understand the maturity condition of fruit. The loss of quality of tomatoes is one of the major challenges faced in tomato agriculture sector [4]. Typically, farmers use their personal experience to detect type of disease and maturity level of tomatoes. Manual inspection is used to determine the tomato crop's level of maturity. This in turn leads to a dependency on manual labour, which is inconsistent. To overcome these problems, smart harvesting is being promoted in recent years.

1.2 Problem Statement

Recently, some studies have attempted to evaluate the feasibility of smart agriculture involving machine learning for harvest ripeness detection. However, these works typically used smaller data size, simple dataset with no background or leaves or explored limited machine learning networks. From the literature review, most works focused on single method to detect tomato ripeness such as SVM, CNN, decision tree, GLCM, Mask RCNN, Faster RCNN and YOLOv3. Also, some of the studies detect tomatoes without background. It is less effective during harvesting process because the tomatoes are actually hanging on tomato trees in actual environment, and typically surrounded by leaves.

1.3 Objectives

The aims of this project are:

(a) To configure two machine learning networks to detect tomato ripeness.

(b) To evaluate the performance of the two types of machine learning architecture in tomato ripeness detection.

(c) To validate the algorithm with bigger dataset - 1000 images depicting a total of 9063 unripe, half ripe and ripe tomatoes with leafy background within the images.

1.4 Scope of Work

The project's scopes are as follows:

(a) To use publicly accessible database and to work only on tomatoes.

(b) Dataset would be labelled manually using VGG annotator tool.

(c) Two different object detection networks, which are Mask RCNN and YOLOv5, would be implemented and compared.

(d) The performance of these two different architectures would be evaluated based on performance accuracy and mAP of tomato ripeness classification.

(e) Mask RCNN and YOLOv5 would be trained using TensorFlow.

(f) Intel Tiger Lake core i7-11800H and 6GB RTX 3060 GPU would be used as computational platform to train the network.

REFERENCES

[1] Background, "The global tomato processing industry," Tomato Neiss, June 2020. [Online]. Available: http://www.tomatonews.com/en/background_47.html. [Accessed 25 January 2022].

[2] A. Bjarnadottir, "Healthline," Healthline, 25 March 2019. [Online]. Available: https://www.healthline.com/nutrition/foods/tomatoes#vitamins-and-minerals. [Accessed 25 January 2022].

[3] "Global Tomato Processing Market," Imarc, 11 July 2020. [Online]. Available: https://www.imarcgroup.com/global-tomato-processing-market. [Accessed 25 January 2022].

[4] Y. P. Huang, T. H. Wang and H. Basanta, "Using Fuzzy Mask R-CNN Model to Automatically Identify Tomato Ripeness," vol. 8, p. 11, 2020.

[5] P. Kamavisdar, S. Saluja and S. Agrawal, "A Survey On Image Classification," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 2, no. 1, pp. 1005-1009, 2013.

[6] C. Ma, X. Yi and S. Xu, "Research on Image Classification Method Based on DCNN," in *International Conference on Computer Engineering And Application (ICCEA)*, Shanghai, 2020.

[7] A. Garcia-Garcia, S. Orts-Escolano and S. Oprea, "A survey On Deep Learning Techniques For Image And Video Semantic Segmentaion," *Applied Soft Computing*, pp. 41-65, 2018.

[8] D. Guanlin, "Research on Semantic Segmentation Algorithm Based on Deep Learning Control Tools," in *International Conference on Computer Communication and Network Security*, Ganzhou, 2020 .

[9] S. Choudhuri, N. Das and M. Nasipuri, "Object Localization on Natural Scenes," ResearchGate, Kolkata, 2017.

[10] E. Etemad and Q. Gao, "Object Localization by Optimizing Convolutional Neural Network Detection Score Using Generic Edge Features," in *IEEE International Conference on Image Processing*, Halifax, 2017.

[11] A. M. Hafiz and G. M. Bhat, "A Survey on Instance Segmentation: State of the Art," Int J Multimed Info Retr, Srinagar, 2020.

[12] A. Satishchandra and S. Chickerur, "Analysis of Instance Segmentation Using Mask-RCNN," in *2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies* , Hubballi, 2019.

[13] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *ArXiv*, 2016.

[14] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature Pyramid Networks for Object Detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[15] M. Kuo, H. T. Chan and C. H. Hsia, "Study on Mask R-CNN with Data Augmentation for Retail Product Detection," in *International Symposium on Intelligent Signal Processing and Communication Systems* , 2021.

[16] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016.

[17] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *IEEE conference on computer vision and pattern recognition,* p. 7263–7271, 2017.

[18] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," arXiv preprint arXiv:1804.02767, 2018.

[19] A. Bochkovskiy, C.-Y. Wang and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv:2004.10934, 2020.

[20] U. Nepal and H. Eslamiat, "Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs," *Sensors*, pp. 1-15, 2022.

[21] M. B. Garcia, S. Ambat and R. T. Adao, "A Machine Learning Approach for Tomato Ripening Stage Identification Using Pixel-Based Color Image Classification," in *IEEE*, Philipines, 2019.

[22] E. E. Lavindi, E. J. Kusuma and G. F. Shidik, "Neural Network based on GLCM, and CIE L*a*b* Color Space to Classify Tomatoes Maturity," in *iSemantic*, Indonesia, 2019.

[23] L. Zhang, J. Jia and G. Gui, "Deep Learning Based Improved Classification System for Designing Tomato Harvesting Robot," *SPECIAL SECTION ON AI-DRIVEN BIG DATA PROCESSING: THEORY, METHODOLOGY, AND APPLICATIONS,* vol. 6, p. 10, 2018.

[24] K. Yamamoto, W. Guo, Y. Yoshioka and S. Ninomiya, "On Plant Detection of Intact Tomato Fruits Using Image Analysis and Machine Learning Methods," *Sensors,* vol. 14, no. 7, pp. 12191-12206, 2014.

[25] Megha.P.Arakeri and Lakshmana, "Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry," *7th International Conference on Communication, Computing and Virtualization,* vol. 79, pp. 426-433, 2016.

[26] C. Hu, X. Liu, Z. Pan and P. Li, "Automatic Detection of Single Ripe Tomato Automatic Detection of Single Ripe Tomato Intuitionistic Fuzzy Set," vol. 7, p. 14, 2019.

[27] G. Liu, J. C. Nouaze and P. L. T. Mbouembe, "YOLO-Tomato: A Robust Algorithm for Tomato Detection Based on YOLOv3," in *MDPI*, Shouguang, 2020.

[28] Laboro, "Engineer Column," Laboro AI, 2017. [Online]. Available: https://laboro.ai/activity/column/engineer/laboro-tomato/. [Accessed 5 December 2021].

[29] W. A. S. Darma, N. Suciati and D. Siahaan, "A Performance Comparison of Balinese Carving Motif Detection and Recognition using YOLOv5 and Mask R-CNN," *International Conference on Informatics and Computational Sciences (ICICoS) ,* p. 52, 2015.

[30] L. Cai, T. Long, Y. Dai and Y. Huang, "Mask R-CNN based Detection and Segmentation for Pulmonary Nodule 3D Visualization Diagnosis," in *IEEE*, ChongQing, 2019.