ISJET

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY

Volume 7 No. 2 July-December 2023

ISSN 2586-8527 (Online)
Panyapiwat Institute of Management
Indexed in the Thai-Journal Citation Index (TCI 2)

INTERNATIONAL SCIENTIFIC
JOURNAL OF ENGINEERING AND TECHNOLOGY
(ISJET)

Volume 7 No. 2 July-December 2023

ISSN 2586-8527 (Online)
PANYAPIWAT INSTITUTE OF MANAGEMENT

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)
Volume 7 No. 2 July-December 2023
ISSN 2586-8527 (Online)

Copyright
Panyapiwat Institute of Management
85/1 Moo 2, Chaengwattana Rd.,
Bang Talat, Pakkred,
Nonthaburi, 11120, Thailand
Tel. +66 2855 1560
E-mail: isjet@pim.ac.th
Website: https://ph02.tci-thaijo.org/index.php/isjet/index

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)

Volume 7 No. 2 July-December 2023 ISSN 2586-8527 (Online)

Objective:
International Scientific Journal of Engineering and Technology will be dedicated to serving as a forum to share knowledge on research advances in all fields of sciences: Engineering, Technology, Innovation, Information Technology, Management Information Systems, Logistics and Transportation, Agricultural Science and Technology, Animal Science and Aquaculture, Food Science, and other areas in Sciences and Technology. Submissions are welcomed from both PIM as well as other Thai and foreign institutions.

Scope:
Engineering, Technology, Innovation, Information Technology, Management Information Systems, Logistics and Transportation, Agricultural Science and Technology, Animal Science and Aquaculture, Food Science, and other areas of Sciences and Technology.

Type of Article:
- Research article
- Academic article
- Book review
- Review article

Languages of academic works:
An article written in either English language is accepted for publication.

Reviewing Policy:
1. Any manuscript to be accepted for publication must have been reviewed and approved by at least three peer reviewers in that particular field or related fields. The Journal has a double-blind peer review policy which means that neither the peer reviewer nor the author knows the identity of each other.
2. The submitted manuscript must have never been published in any other periodical, and must not be in the approving process for publication by any other periodical. Also, the author must not plagiarize the work of other people.
3. The article, expression, illustrations, and tables that are published in the Journal are the sole responsibility of the author, and definitely not that of Panyapiwat Institute of Management.
4. The Editorial Board of International Scientific Journal of Engineering and Technology reserves the right to change or revise the name(s) and unit(s) of the author(s) in all cases after the issuance of the letter.
5. The Editorial Board of International Scientific Journal of Engineering and Technology reserves the right to cancel the publication that has been issued a certification of publication in the Journal.
6. The Editorial Board of International Scientific Journal of Engineering and Technology reserves the right for decision making on publishing any article in the Journal.

Frequency of Publication:
Twice a year
- The first issue: January-June
- The second issue: July-December

Publication and Access Charges:
There are no charges to submit and publish all types of articles. Full articles in pdf format can be downloaded free from the journal website at https://ph02.tci-thaijo.org/index.php/isjet/index

ISJET Journal Editorial Board
The office of Research and Development
Panyapiwat Institute of Management
85/1 Moo 2, Chaengwattana Rd.,
Bang Talat, Pakkred, Nonthaburi, 11120, Thailand
Tel. +66 2855 1560
E-mail: isjet@pim.ac.th
Website: https://ph02.tci-thaijo.org/index.php/isjet/index

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)
Volume 7 No. 2 July-December 2023
ISSN 2586-8527 (Online)

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)
Volume 7 No. 2 July-December 2023
ISSN 2586-8527 (Online)

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)
Volume 7 No. 2 July-December 2023
ISSN 2586-8527 (Online)

PANYAPIWAT INSTUTUTE OF MANAGEMENT
85/1 Moo 2, Chaengwattana Rd.,
Bang Talat, Pakkred, Nanthaburi, 11120 Thailand

Dear Colleagues,

We are pleased to present the latest issue of the International Scientific Journal of Engineering and Technology (ISJET), in which we explore cutting-edge research and innovations in the fields of Industry 4.0, artificial intelligence, and other Technological solutions to contemporary industrial challenges.

Research topics include a variety of industrial applications, from low-cost defect detection to fake crowdfunding project identification, improved weed detection, defect reduction in production processes, success factors for Industry 4.0 in Thailand's manufacturing, and strategies for reducing product picking wait times in warehouses.

These works illustrate the pursuit of efficiency, quality, and innovation in the modern industrial landscape, highlighting the transformative power of technology and data-driven decision-making.

Explore these groundbreaking studies to better understand the cutting edge of industrial research. Each represents a significant step forward in addressing the challenges and opportunities of Industry 4.0. We hope you find this collection informative and inspiring as we navigate the dynamic Industry 4.0 landscape together.

With kind regards,

Dr. Wirin Sonsrettee
Associate Editor of Science
isjet@pim.ac.th

INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET)
Volume 7 No. 2 July-December 2023
ISSN 2586-8527 (Online)

CONTENTS

A Low-Cost IIoT-Enabled Computer Vision-based System for Classifying Defect Types and Severity Levels in Industry 4.0

Panisara Kanjanarut[1] and Warut Pannakkong[2]
[1,2]School of Manufacturing Systems and Mechanical Engineering,
Sirindhorn International Institute of Technology, Thammasat University,
Pathum Thani, Thailand
E-mail: panpanisarakj@gmail.com, warut@siit.tu.ac.th

Received: June 11, 2022 / Revised: October 29, 2022 / Accepted: November 11, 2022

Abstract—**Industry 4.0 technologies such as the Industrial Internet of Things (IIoT) and Artificial Intelligence (AI) can assist in automating defect detection and classification processes which are crucial for quality control in the manufacturing industry. However, there is still a barrier to adopting the technologies in Small and Medium Enterprises (SMEs) because of their limited budget. This paper presents a low-cost defect detection and classification system and an interactive real-time dashboard monitoring IIoT data utilizing a single-board computer and mainly open-source software. In the system, workpieces will be classified into non-defective (OK) and defective (NG) workpieces. Then, the NG workpieces will be further classified into defective types and severity levels. The workpiece used in the case study is a sticker on a 4.4 cm diameter bottle cap. The defect types are Off-Color, Missing Details, and Scratches, then each type is divided further into four severity levels. From evaluation, the system can achieve 96% when classifying as OK/NG and 88% accuracy in classifying defective types and levels. The system's reliability is 100%. Based on experts' opinions, the proposed system is relatively low-cost, reliable, and accurate for practical uses. The proposed system can be implemented locally or globally via a cloud server.**

Index Terms—**Computer Vision, Defect Classification, Industrial Internet of Things, Industry 4.0**

I. Introduction

Industry 4.0 is a term coined by the German government project to encourage the 4th generation of manufacturing through the integration of Information and Communication Technologies (ICT), Internet of Things (IoT), Artificial Intelligence (AI), and web services with industrial machines, systems, and processes [1]. It also raises the term Industrial Internet of Things' or IoT. The cyber-physical infrastructure digitizes information and allows communication between relevant parties to improve efficiency [2]. Utilizing these technologies is beneficial for all-size enterprises, but Small and Medium-sized Enterprises (SMEs), whose processes are manual or semi-automated, are generally at an immense investment disadvantage [3]-[6]. Therefore, creating a low-cost technology and system to accompany SME development toward industry 4.0 that is compatible with the existing system becomes a necessary solution [7].

The common practice of Quality Control (QC) in manual or semi-automation processes is to perform an explicitly manual sampling inspection on the workpiece several times using different viewing angles and distances. When a new defective feature is found, it needs to be inspected by an expert and described manually [8]. These have proven to be a problem due to the sampling inspection strategy since not every workpiece will be inspected [9]. Defect Detection and Classification (DDC) using human operators is also labor-intensive, time-consuming, and tiresome. Leading to them losing concentration resulting in the accuracy of operation falling between 60-70% [10].

Recently, Computer Vision (CV) has become wildly adopted in automatic visual inspection for defect detection due to the availability of cheap cameras and the development of AI. CV is a subfield of AI that trains the computer to replicate human vision, allowing computers to accurately identify and classify objects the same way humans do. The advantages of CV are that camera installation does not interrupt the existing production line, and image taking can be done without a pause. In addition, the classification result from the AI model, along with the inspection timestamp, can be stored automatically, reducing the recording time for analysis.

Most existing literature is done with binary classification or defect/non-defect basis. However, there is more to be considered, such as the defect's type and severity. This information, combined with the date-time component, can prove vital for predictive maintenance [11]. Furthermore, detailed information provides ways to increase the efficiency and flexibility of the production system through real-time monitoring and high-speed reporting through the IIoT dashboard [12].

To develop SMEs toward 4.0, IIoT technology is essential. The main objective of IIoT is to connect a physical object to the Internet allowing remote access control and monitoring [13]. With IIoT devices' tremendous programming ability [14], business operations become more agile and transparent. The real-time dashboard will also support effective decision-making [15].

However, for existing literature, a low-cost DDCS that integrates IoT and real-time dashboards while also considering the type and severity level classification has not been developed.

This study aims to develop a low-cost DDCS and an interactive real-time dashboard monitoring IIoT data. An evaluation of the proposed system is also presented.

The rest of the paper is organized as follows, section II is Related Work of the existing literature. section III is Methodology. section IV is the Evaluation Result, and section V is Discussion. Lastly, section VI is the Conclusion and section VII is Future Work.

II. Related Work

The visual inspection process for defect detection and classification can be done in two ways, contact and contactless. For the contact approach, the operation is done manually by a human operator with an aided tool like magnifying glass or microscope [8], [16], [17]. The inspection task is performed by experienced inspectors who can be effect by fatigue and inattentiveness. Hence this method is unreliable and does not provide stable results due to the subjectiveness of the operator [12].

Also, a human could not operate in real-time, contradicting the Industry 4.0 concept [18], [19]. For the contactless approach, the inspection process is done through one or more cameras and CV software. It allows the operator to carry on without stopping the workpiece.

Many CV systems have been presented in literature within the last decade through the availability of cheap digital imaging devices and computational power development. Examples of industrial visual inspection applications in recent literature include mobile phone screen glass [20], steel surface [21], and leather [16]. Both image processing and deep learning approaches are used to obtain promising results. This technology gives digital data ready to be analyzed and stored locally and globally. However, accumulating data locally, for instance on a Programmable Logical Controller (PLC), is not preferable since the storage space is limited. The solution to this is globally stored on a cloud platform enabled by IIoT.

Researchers have recently focused more on integrated information and communication technology [22]. This Internetworking and connectivity of IoT allowed automation in various fields [11]. It also gave analysis measures that contribute towards strategical planning for SMEs. Moreover, knowing the process stability and quality gives the manufacturer time to reduce or control the defects during the process [9]. [23] stated that adopting IoT is one of the solutions to develop SMEs towards Industry 4.0.

From all the literature mentioned above, the algorithm is run on a Personal Computer (PC), which is not cost-effective for classifying and transmitting IoT data. A single-board computer is more practical than a PC for tasks like this. Raspberry Pi is a small yet powerful single-board computer introduced in 2012. Kurkovsky and Williams [24] stated that a Raspberry Pi has enough processing power to implement up to servicing layers needed to have a networked IoT device. Hou et al. [17] also use a Raspberry Pi 3B to run laser chip classification; They suggest that a better single-board computer needs to be considered. In June 2019, the Raspberry Pi 4B was released with a 1.5 GHz 64-bit quad-core ARM Cortex-A72 processor [25]. This improvement is promising

In conclusion, a low-cost IIoT-enables DDCS is proposed and evaluated. The system merges a low-cost non-contact visual inspection with IIoT to develop SMEs toward 4.0 requirements. This solution is important and quite challenging. The system also can be extended to other manufacturing processes.

III. Methodology

This section includes two main sections, the generalized concept of a proposed system and a case study.

A. Proposed System

Fig. 1 shows an overview of the proposed system for deployment in the production line. The camera captures the workpiece image and does image processing, then transmits measurement data to the classification model, database, and dashboard. Soon after, Statistics and a summary of the production line defect type and severity level are shown on the dashboard for monitoring. Such a system allows detailed evaluation and helps identify possible causes of faults.

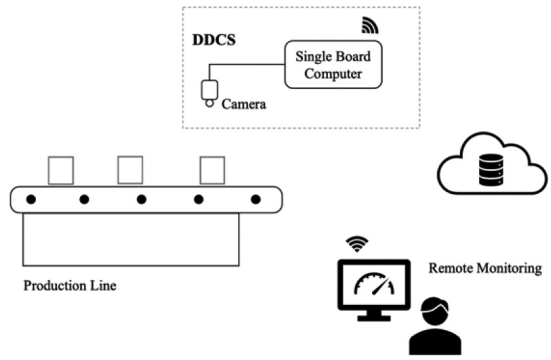


Fig. 1. The proposed system consisted of a camera, a single-board computer, and a remote monitoring system.

B. Case Study

In our case study, the proposed system is implemented in a demonstration set, as shown in Fig. 2. The camera used is a Mitsubishi Electric Vision Sensor (VS20). The single-board computer is a Raspberry Pi 4 Model B with 8 Gb RAM equipped with a 5V fan supply mounted on top powered by its own GPIO pin. The software used to develop a remote monitoring system is Node-Red, which is an open-source program. The database is Google Sheets. A green, yellow, and red LED will be connected to Raspberry Pi to act as a status indicator. Production Line is controlled by a Programmable Logic Controller (PLC). There are three photo sensors used as PLC's input in a ladder diagram. One at the starting of the conveyor is used to signal the conveyor. Another one locates before the camera is used to trigger the camera when the workpieces arrive. The last one is located after the camera signals the pusher to eject the defective workpiece.

1) Experiment Workpieces

The workpiece used in this experiment is a printed sticker on top of a twisted-off plastic cap with a clear glass bottle. For the quality characteristic, the sticker must have the correct color and clear printing with correct details. The types of defects are classified into three types: Off-color, missing details, and scratches, as shown in Fig. 3. Each type of defective workpiece consists of four severity levels (i.e., 1, 2, 3, and 4). The severity levels increase as the number goes up.

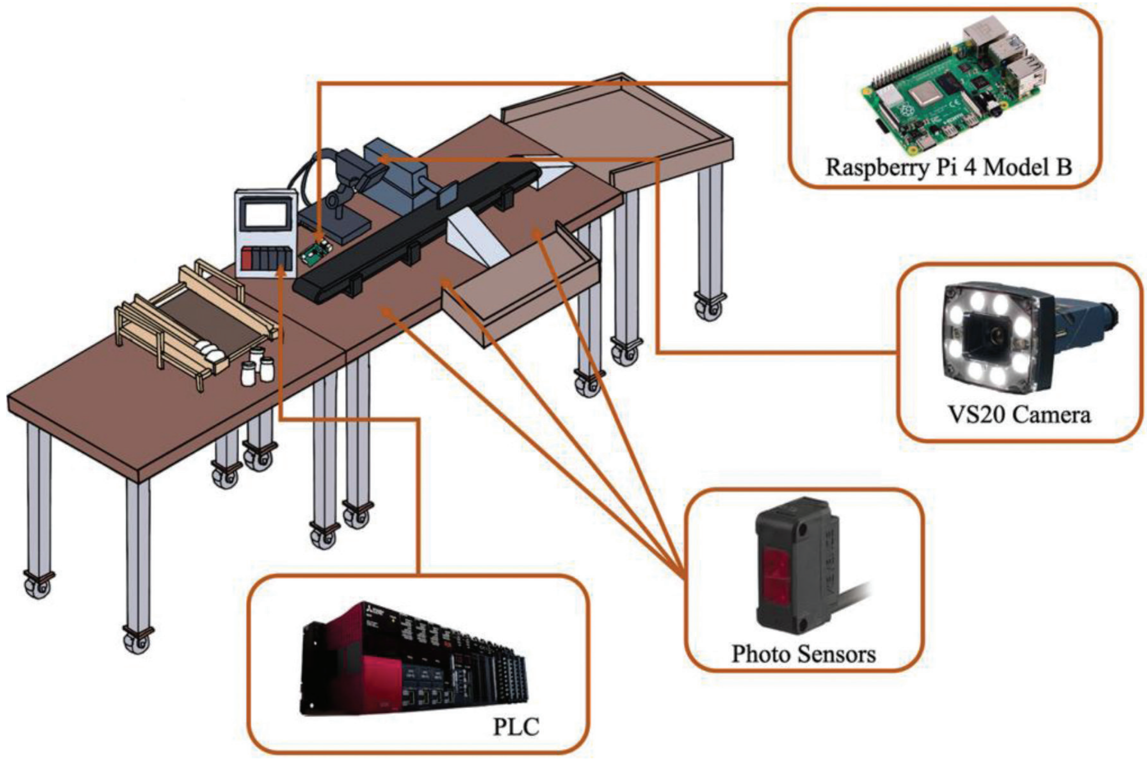


Fig. 2. Components of demonstration set [26]-[28]

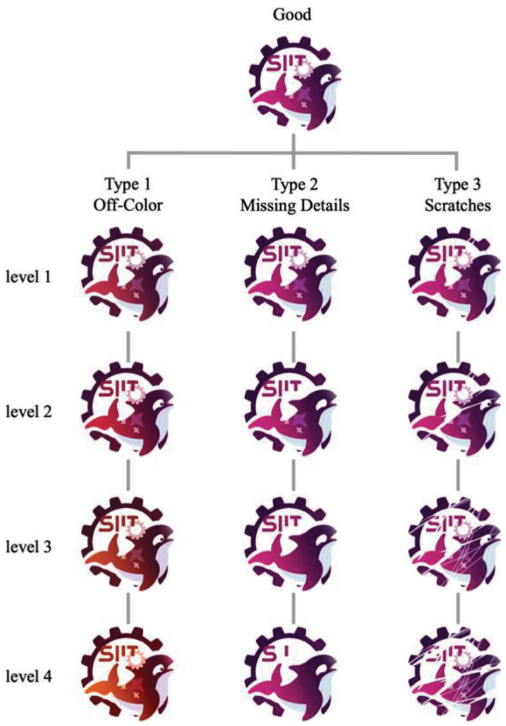


Fig. 3. 13 Types of stickers consisted of good and off-color type 1 to 4, missing details type 1 to 4, and scratches type 1 to 4.

For type 1 defect (off-color), the first two severity levels are less vibrant but are still purple, while the last two severity levels fade to the point that they appear brown. For type 2 defects (missing details), a severity level 1 is missing barely noticeable details like the eye. Severity level 2 is missing one component, while severity levels 3 and 4 lack most of the critical components. For type 3 defects (scratch), severity level 1 contains one or two small scratches. Severity level 2, the scratch is noticeable with minimal effort. Severity level 3, the scratches are evident. For severity level 4, the scratches almost erase the component's details.

Totally, there are 13 different stickers. One is for a good workpiece, and the remaining 12 are defective workpieces. However, severity level 1 is not significantly different from the original workpiece. Therefore, we assume that the user accepts the severity level 1 workpiece, and it can be classified as an OK workpiece along with Good stickers.

2) Vision System

In the demonstration set, the QC process is done using a Cognex Vision Sensor. The vision sensor is a specialized quality-control tool that combines a machine-vision camera with onboard intelligence.

There are three main used tools consisted of pattern, color pixel count, and logic tools. Each tool provides an output signal as pass or fail. The pattern tool determines whether a trained pattern is present or absent. The color pixel count tool determines whether a color feature is present or absent based on the number of matching color pixels in a searching region with the selected colors. Finally, the logic tool is an expression builder to create a logical formula combining the output signal from the other tools and generating a pass or fail signal. The list of tools used is shown in Fig. 4.

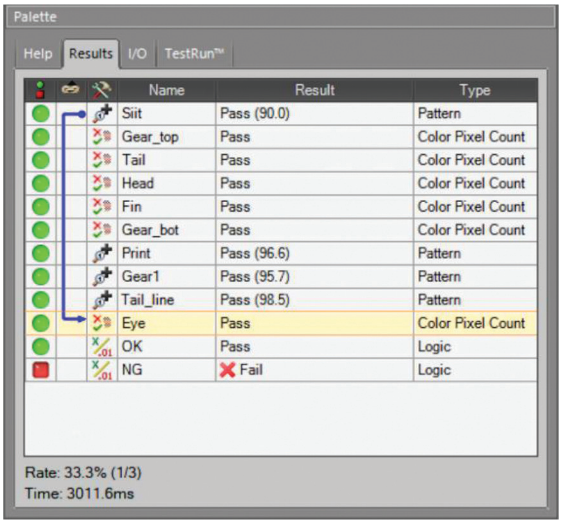


Fig. 4. Image processing tools used to extract information from the workpiece picture.

The output string from the vision system is encoded into UTF-8 format and then transferred to the Raspberry Pi through an ethernet cable. Finally, the output from each tool is selected to be used in the classification model, as shown in Fig. 5.

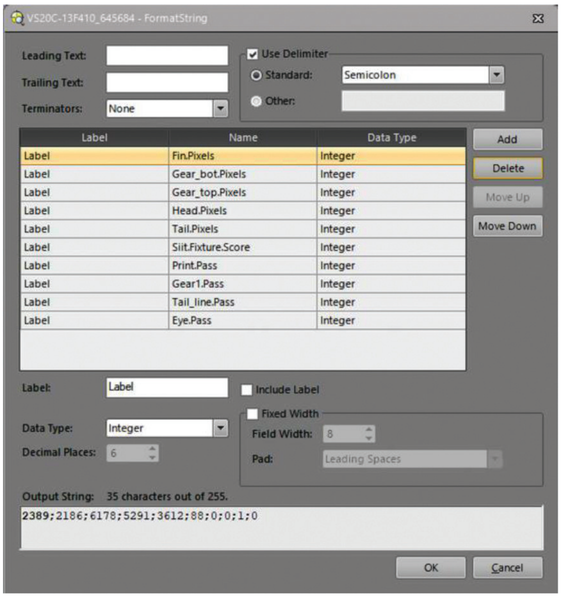


Fig. 5. Outputs from vision sensor to use as inputs for AI model

Outputs selected are the rawest possible outputs from each tool which are the number of pixels from the pixel count tool and the percent similarity score from the pattern tool. However, when the workpiece is missing details. Instead of giving the similarity score of 0, the pattern tool will give an error message and not send any output to the Raspberry Pi. The binary result (.pass) is used to resolve this issue.

3) Raspberry Pi

The communication between Raspberry Pi and VS20 is through a socket protocol coding in Python. The Raspberry Pi becomes a server waiting for a connection. When the data is received, it will go through a classification model and be uploaded to the database and dashboard.

The Raspberry Pi has a General Purpose Imput/Output (GPIO), which can be used to supply current to the output load. In this demonstration, the pin is connected to an LED. A green LED lights up when the connection with Node-RED has been established and will be off if the user calls emergency status from the dashboard, and a yellow LED will lighten up instead. The red LED lights up when the classification model predicts the workpiece is a defective product. A pseudocode for implementing Raspberry Pi is shown in Fig. 6.

```
set up socket server
set up MQTT to Node-RED
load classification model

while True:
    camera connect to Raspberry Pi
    green LED on

    while True:
        received data from the camera
        classification model predicts result
        get timestamp
        upload result and timestamp
        publish result to dashboard

        if result is defect:
            red LED on for 1 s

close the connection
```

Fig. 6. Psudo code

4) Classification Model

Random Forest (RF) is used to classify the workpiece (i.e., good or defect type and severity level) by using outputs from VS20 as its inputs. The training phase of the classification model is done outside using Google Colab, then used Pickle Python module, which is the standard way of serializing objects into a file for Python. Finally, the desterilized model is used to make a new prediction. Results from the classifier are transmitted to Node-RED over the Internet using the MQTT protocol.

In the RF model, there are three hyperparameters which are *n_estimator*, *criterior,* and *bootstrap*. The *n_estimator* specifies the number of trees in the forest. The *criterior* used is gini, which is a type of function to measure the quality of the tree split. The *bootstrap* is used when building the tree. If the *bootstrap* is False, then the whole dataset is used, otherwise, some samples in the dataset are used to train the model. Through a trial and error process, the best configurations of the hyperparameters are *n_estimator* = 500, *criterior* = gini, and *bootstrap* = False.

5) Monitoring

Node-RED is a flow-based development tool for visual programming. It was initially developed by IBM for wiring together hardware devices, APIs, and online services as part of the Internet of Things. AUI interface is then created using JavaScript. Fig. 7 below shows the proposed dashboard.

From the dashboard, detailed statistics related to production status can be monitored remotely. At a glance, it shows the summary of the production state. This information gives a clear summary of the production line and helps determine the progress when working on a specific order.

In the control panel section, the required user's input will be used to calculate the Overall Equipment Efficiency (OEE). It is a performance measurement tool that measures different types of production losses like quality loss, availability loss, and performance loss. The use of OEE can help identify areas for process improvement.

The Node-RED server can be hosted on the Raspberry Pi or the cloud server. Running the server locally will limit the user to be within the same network in order to access the dashboard, while the benefit of hosting on a cloud server is global access but will come with additional costs. Since Node-RED is IBM donated, lunching on the IBM cloud is more convenient.

6) Database

Using Google API, the classification results from the RF model, the date-time components, and the camera outputs can be uploaded to a Google Sheet for later use. The information included are raw data, date, time, and prediction result. The raw data is also split into individual columns as shown in Fig. 8.

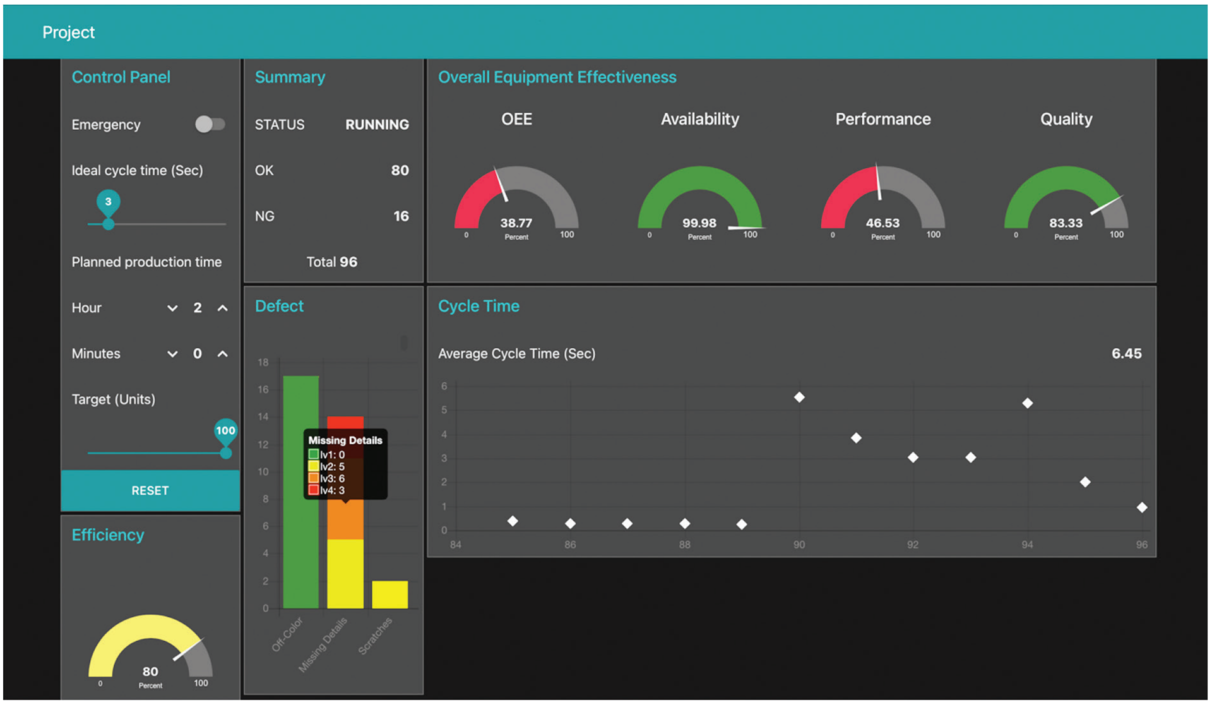
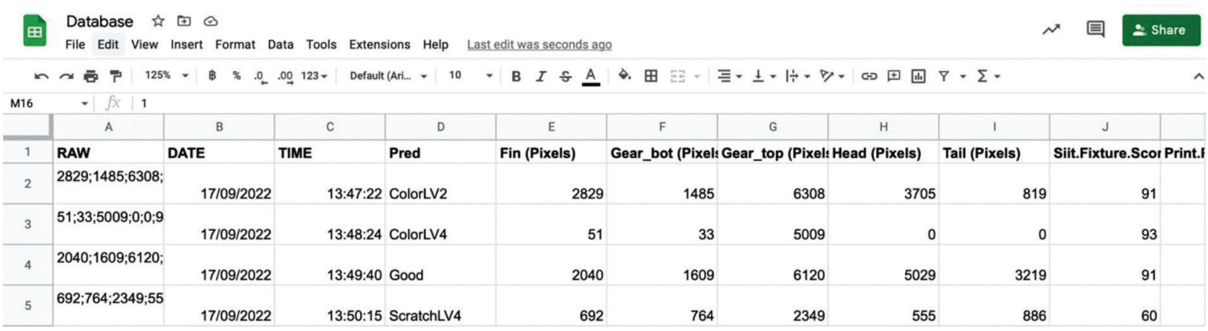


Fig. 7. Dashboard of the case study



	RAW	DATE	TIME	Pred	Fin (Pixels)	Gear_bot (Pixel	Gear_top (Pixel	Head (Pixels)	Tail (Pixels)	Siit.Fixture.Scor	Print.I
2	2829;1485;6308;	17/09/2022	13:47:22	ColorLV2	2829	1485	6308	3705	819	91	
3	51;33;5009;0;0;9	17/09/2022	13:48:24	ColorLV4	51	33	5009	0	0	93	
4	2040;1609;6120;	17/09/2022	13:49:40	Good	2040	1609	6120	5029	3219	91	
5	692;764;2349;55	17/09/2022	13:50:15	ScratchLV4	692	764	2349	555	886	60	

Fig. 8. Database Information

IV. Evaluation

For evaluating the performance of the demonstration set, reliability and accuracy have been tested.

A. Reliability Test

The ability to work over a long period is essential for equipment used in an industrial environment. Hence reliability is a concern.

To check the reliability, testing with an operating speed of 1000 units per hour over consecutive 8 hours has been performed under two ambient temperature conditions (i.e., an air-conditioned room (25) and a normal room (35)).

Over the 8 hours, we monitor the output generated from the Raspberry Pi, such as the classification result and dashboard statistics. As expected, the amount of output at the end of the testing is 8,000.

Moreover, the CPU and GPU temperatures in the Raspberry Pi are measured. The temperatures are shown in Fig. 9.

The temperature could reach a certain point until it throttles the processing capability. The time it takes the processor to calculate is also raised as an issue.

The average temperature is 42.4 on both processors for a normal room. For an air-conditioned room, the average temperature is 36.6 The temperatures are relatively steady around the averages. The testing result implies that the Raspberry Pi can pass this reliability test.

B. Accuracy Test

The accuracy is crucial. Without measuring accuracy, it is difficult to know whether the model is working as expected.

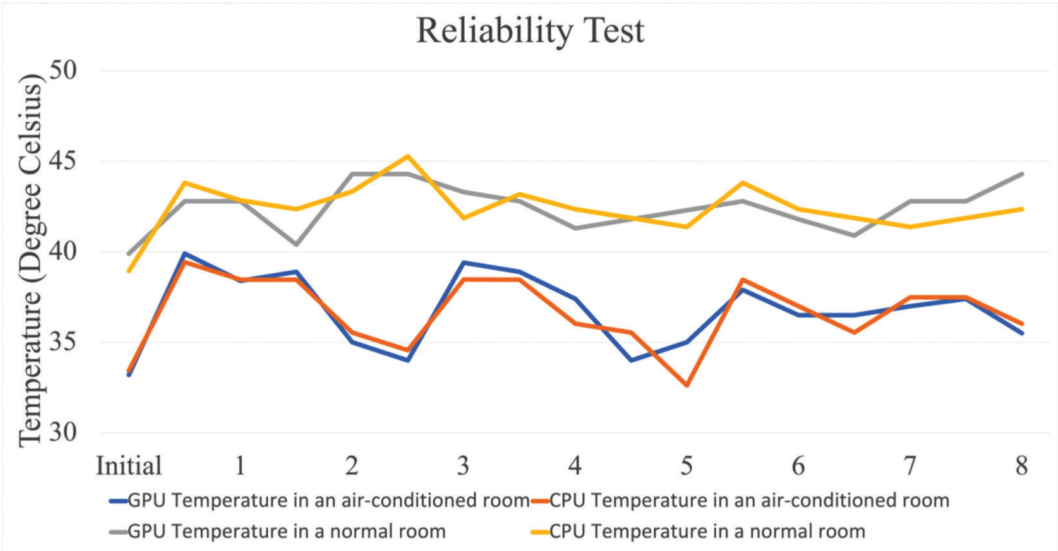


Fig. 9. CPU and GPU temperatures of Raspberry Pi over 8 hours

The test was conducted using 650 workpieces (50 workpieces of each 13 stickers) on the demonstration set. The bottle's placement is done manually at the starting position of the conveyor, and the sticker's rotation angle is random.

The performances measured are classification accuracy, precision, and recall. The results are shown in Fig. 10. The value lower than 80% is in red.

The model accuracy is 87.08%. When comparing precision and recall side by side, for most of the stickers, the results are satisfied. However, recall for Good, Missing Details level 1, and Scratches level 1 is alarming. For precision, Good, Scratches Level 1, and Scratches Level 2 are a concern.

		PREDICTED															
Type		Good	Off Color				Missing Details				Scratch						
	Level	-	1	2	3	4	1	2	3	4	1	2	3	4	Total	Recall	Precision
TRUE Good	-	32									16	2			50	64%	62%
Off Color	1		50												50	100%	100%
	2			50											50	100%	96%
	3			2	48										50	96%	100%
	4					50									50	100%	100%
Missing Details	1	15					17				12	5	1		50	34%	85%
	2							47	1		1	1			50	94%	92%
	3	1						1	44	2		2			50	88%	98%
	4									50					50	100%	96%
Scratch	1	4					2	3			31	10			50	62%	52%
	2						1					49			50	98%	70%
	3											1	49		50	98%	96%
	4												1	49	50	98%	100%

Accuracy	87.08%

Fig. 10. Camera and RF's model's accuracy test

The conclusion is that Good and Scratches level 1 is the main reason why the model is not as accurate. The model tends to predict Missing level 1 as another type and tends to predict other types as Scratches level 2.

Suppose we summarize the model result into OK/NG classification as shown in Fig. 11. The accuracy is significantly improved at 96.31% compared to the result from using the camera alone. The model gets almost 100% in rejecting defective workpieces and 90% accuracy in accepting a good workpiece. A 10% reject when actually a good workpiece is a concern but not as severe as a classified defective workpiece as OK.

		PREDICTED	
		OK	NG
TRUE	OK	179	21
	NG	3	447

		PREDICTED	
		OK	NG
TRUE	OK	89.5%	10.5%
	NG	0.7%	99.3%

Accuracy	96.31%
Recall	89.50%
Precision	98.35%

Fig. 11. Camera and RF's model's accuracy test as OK/NG

V. Discussion

To the best of our knowledge, cost is an aspect that has never been the main discussion for vision systems. The demonstration set presented in this paper solely utilized open-source programming tools. In a situation where the investor already has the camera, the cost of the system is approximately 6,000 THB. Nevertheless, without the VS20 camera, a webcam camera can be implemented without raising costs substantially. The current webcam price in the market is a few thousand. Assuming the webcam price is 5,000 THB, the total system cost becomes 11,000 THB. To hire an employee, would cost 15,000 THB/month. Therefore, the payback period is less than a month, which seems affordable for SMEs.

When compared to other low-cost IIoT applications from various fields, the price is highly dependent on the amount and the quality of sensors integrated into the system. The minimum price is 4,500 THB [30] and the maximum is 41,000 THB [31] while the median falls at 8,700 THB. This makes our application fall in the middle of the range. A box plot for comparing five applications [30]-[34] is shown in Fig.12.

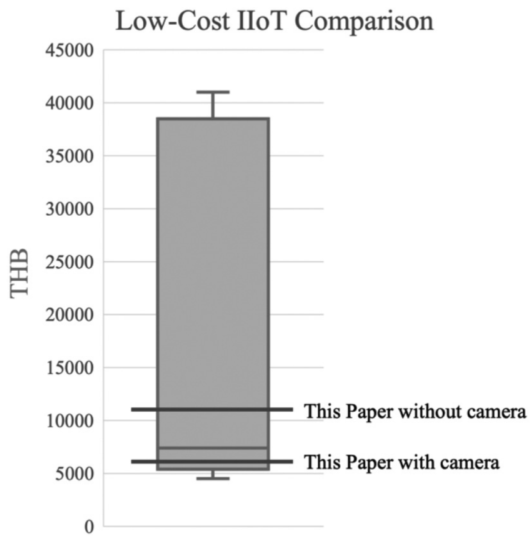


Fig. 12. Box plot of IIoT application cost

After obtaining the performance result, we conducted an initial survey with four people who have more than ten years of experience in the field of manufacturing and CV. The Participants are two Business Owners and two Vision System experts. Their opinion on the proposed system and performance is that pricing is cheaper than expected. If the breakeven point is lower than two years would definitely invest, but accuracy could be better.

From the reliability test, the developed system is highly reliable in both ambient environments. From the accuracy test, considering recall value, the RF model performs excellently except for Good, Missing Details level 1, and Scratches level 1. This is because the difference between these stickers is relatively small. The Missing Details level 1 does not have an eye of the orca, while Scratches level 1 has two minor scratches on the outer gear. A comparison between the three types of stickers is shown in Fig. 13.

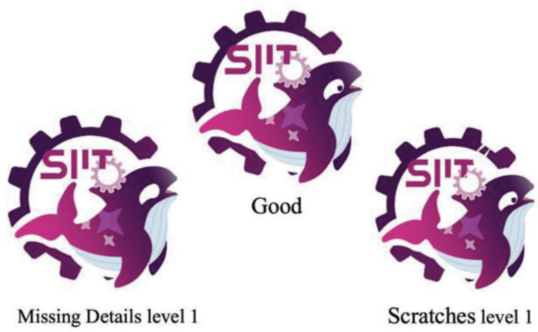


Fig. 13. Comparison between good, missing details level 1 and scratches level 1

The next point of concern is when a defective workpiece is classified as an OK even if the possibility is low at 0.7% (three samples from Fig. 11). The detailed information on the three samples is shown in Table I. The last four columns in Table I and further investigation by comparing the camera binary outputs to the ground truth shown in Table II imply the lack of repeatability since roughly 60% is correct.

TABLE I
GOOD WORKPIECE AS DEFECTIVE WORKPIECE OUTPUT

Predict	True	Print (Binary)	Small Gear (Binary)	Tail Line (Binary)	Eye (Binary)
Good	MissingLV3	0	0	0	1
ScratchLV1	MissingLV2	1	1	1	1
MissingLV1	ScratchLV2	1	1	0	0

TABLE II
GROUND TRUTH OF THE WORKPIECES

True	Print (Binary)	Small Gear (Binary)	Tail Line (Binary)	Eye (Binary)
MissingLV3	0	0	1	0
MissingLV2	1	0	1	0
ScratchLV2	1	1	0	1

VI. CONCLUSION

This study aims to develop a low-cost defect detection and classification system and an interactive real-time dashboard monitoring IoT data. An evaluation of the proposed system is also presented.

This paper proposed a low-cost defect detection and classification system with interactive real-time dashboard monitoring IoT data. The system utilizes 4.0 technology such as AI and IIoT to assist visual inspection process and automated data recording (i.e., date-time component, defective type, and severity level). Furthermore, the automated system allows every workpiece to be inspected, and the use of a single-board computer placed on the surface makes it compatible with the existing production line.

For the case study, the demonstration set has been evaluated. The reliability of Raspberry Pi is high in both 25 and 35 ambient temperatures. The accuracy is 96% on OK/NG classification and 87% on defective type and severity level. To deploy on an actual production line, the accuracy needs to be improved. If camera precision is improved, it should be possible to improve the overall accuracy, as shown when training AI with ground truth.

VII. FUTURE WORK

Another image processing program is also done manually and will need expert assistance when there is a new workpiece. A method to automate this process would be beneficial.

Subsequently, a long-term test should be conducted to determine the reliability and maintenance needed for hardware like the Raspberry Pi. Finding the most suitable machine learning algorithm is challenging and requires a trial-and-error process. Testing other algorithm performance over the same problem is highly interesting. Lastly, the classification should be redefined from multi-class to multi-label problems since defects can be more than one type at the same time.

ACKNOWLEDGMENTS

The authors wish to thank Mitsubishi Electric Thailand for providing us the industry knowledge and supporting us throughout this study. The authors also wish to thank Ms. Sureerat Phadungsakviriya and her team for in coordinating and supporting us to achieve this study. Lastly, the authors wish to thank Associate Professor Pisal Yenradee (D.Eng.) for his valuable guidance and advice.

REFERENCES

[1] Z. Shahbazi and Y. C. Byun, "Integration of Blockchain, IoT, and Machine Learning for Multistage Quality Control and Enhancing Security in Smart Manufacturing," *Sensors*, vol. 21, no. 4, p. 1467, Feb. 2021.

[2] U. Ullah, F. A. Bhatti, A. R. Maud et al., "IoT-Enabled Computer Vision-Based Parts Inspection System for SME 4.0," *Microprocessors and Microsystems*, vol. 87, p. 104354, Nov. 2021.

[3] P. Naruetharadhol, W. A. Srisathan, N. Gebsombut et al., "Industry 4.0 for Thai SMEs: Implementing Open Innovation as Innovation Capability Management," *International Journal of Technol*ogy, vol. 13, no. 1, pp. 48-57, Jan. 2022.

[4] E. Rauch, P. Dallasega, and M. Unterhofer, "Requirements and Barriers for Introducing Smart Manufacture Ring in Small and Medium-Sized Enterprises," *IEEE Engineering Management Review*, vol. 47, no. 3, pp. 87-94, Jul. 2019.

[5] H. Han and S. Trimi, "Towards a Data Science Platform for Improving SME Collaboration Through Industry 4.0 Technologies," *Technological Forecasting and Social Change,* vol. 174, p. 121242, Jan. 2022.

[6] H. Oliff and Y. Liu, "Towards Industry 4.0 Utilizing Data-Mining Techniques: A Case Study on Quality Improvement," in *Proc. CIRP*, 2017, pp. 167-172.

[7] M. Sony and S. Naik, "Critical Factors for the Successful Implementation of Industry 4.0: A Review and Future Research Direction," *Production Planning & Control*, vol. 31, no. 10, pp. 799-815.

[8] J. Villalba-Diez, D. Schmidt, R. Gevers et al., "Deep Learning for Industrial Computer Vision Quality Control in the Printing Industry 4.0," *Sensors*, vol. 19, no. 18, p. 3987, Nov. 2019.

[9] Z. Zhang, Z. Yang, W. Ren et al., "Random Forest-Based Real-Time Defect Detection of Al Alloy in Robotic Arc Welding Using Optical Spectrum," *Journal of Manufacturing Processes*, vol. 2, pp. 51-59, Jun. 2019.

[10] H. I. Çelik, L. C. Dülger, and M. Topalbekiroğlu, "Development of a Machine Vision System: Real-Time Fabric Defect Detection and Classification with Neural Networks," *The Journal of the Textile Institute*, vol. 105, no. 6, pp. 575-585, Jun. 2017.
[11] A. Kanawaday and A. Sane, "Machine Learning for Predictive Maintenance of Industrial Machines Using IoT Sensor Data," in *Proc. The 8th IEEE International Conference on Software Engineering and Service Science (ICSESS) 2017*, 2017, pp. 87-90.
[12] S. Mei, Y. Wang, and G. Wen, "Automatic Fabric Defect Detection with a Multi-Scale Convolutional Denoising Autoencoder Network Model," *Sensors*, vol. 18 no. 4, p. 1064, Apr. 2018.
[13] K. Shafique, B. A. Khawaja, and F. Sabir, "Internet of Things (Iot) for Next-Generation Smart Systems: A Review of Current Challenges, Future Trends and Prospects for Emerging 5G-IoT Scenarios," *IEEE Access*, vol. 8, pp. 23022-23040, Jan. 2020.
[14] D. Shah, J. Wang, and Q. P. He, "An Internet-of-Things Enabled Smart Manufacturing Testbed," *IFAC-Papers OnLine,* vol. 52, no. 1, pp. 562-567, Jan. 2019.
[15] G. Erboz, "How to Define Industry 4.0: Main Pillars of Industry 4.0," in *Proc. International Scientific Conference Managerial Trends in the Development of Enterprises in Globalization Era*, 2017, pp. 761-767.
[16] S. T. Liong, Y. S. Gan, Y. C. Huang et al., "Integrated Neural Network and Machine Vision Approach for Leather Defect Classification," *arXiv*, vol. 1, pp. 1-12, May. 2019.
[17] D. Hou, T. Liu, Y. T. Pan et al., "AI on Edge Device for Laser Chip Defect Detection," in *Proc. The 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, 2019. pp. 0247-0251.
[18] K. Židek, V. Modrák, J. Pitel et al., "The Digitization of Quality Control Operations with Cloud Platform Computing Technologies," *In Industry 4.0 for SMEs*, pp. 305-334, 2020.
[19] S. Groggert, M. Wenking, R.H. Schmitt et al., "Status Quo and Future Potential of Manufacturing Data Analytics—An Empirical Study," in *Proc. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2017, pp. 779-783.
[20] C. Jian, J. Gao, and Y. Ao, "Automatic Surface Defect Detection for Mobile Phone Screen Glass Based on Machine Vision," *Applied Soft Computing*, vol. 52, pp. 348-358, Mar. 2017.
[21] Y. He, K. Song, H. Dong et al., "Semi-Supervised Defect Classification of Steel Surface Based on Multi-Training and Generative Adversarial Network," *Optics and Lasers in Engineering*, vol. 122, pp. 294-302, Nov. 2019.
[22] J. Harkonen, E. Mustonen, J. Koskinen et al., "Digitizing Company Analytics–Digitalization Concept for Valuable Insights." in *Proc. The 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2020, pp. 1012-1016.
[23] D. Grube, A. A. Malik, and A. Bilberg, "Generic Challenges and Automation Solution in Manufacturing SMEs," in *Proc. The 28th DAAAM International Symposium*, 2017, pp.1161-1169.
[24] S. A. Kurkovsky and C. Williams, "Raspberry Pi as a Platform for the Internet of Things Projects: Experiences and Lessons," in *Proc. The 2017 ACM Conference on Innovation and Technology in Computer Science Education*, 2017, pp. 64-69.
[25] Raspberry Pi. (2022, Mar. 8). *Raspberry Pi 4 Model B Specifications*. [Online]. Available: https://shorturl.asia/ichWb
[26] Newark. (2022, Mar. 8). *Single Board Computer, Raspberry Pi 4 Model B*. [Online]. Available: https://shorturl.asia/8FOWd
[27] Mitsubishi Electric (2022, Mar. 8). *Melsensor Vision Sensor*. [Online]. Available: https://www.mitsubishifa.co.th/files/dl/vs20_catalog_eng_rev.pdf
[28] Keyence. (2022, Mar. 8). *Square Retro-Reflective Cable Type.* [Online]. Available: https://shorturl.asia/a9TOQ
[29] Mitsubishi Electric. (2022, Mar. 8). *MELSEC iQ-R Series iQ Platform-compatible PAC*. [Online]. Available: https://shorturl.asia/dfJZx
[30] F. D. M. Marconcini, M. E. V. Segatto, and E. O. T. Salles, "A Low-Cost Smart Surveillance System Applied to Vehicle License Plate Tracking," *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, vol. 21, no. 1, pp.141-156, Mar. 2022.
[31] A. Arunachalam and H. Andreasson "RaspberryPi-Arduino (RPA) Powered Smart Mirrored and Reconfigurable Iot Facility for Plant Science Research," *Internet Technology Letters*, vol. 5, no. 1, p. e272, Jan. 2022.
[32] S. Komarizadehasl, M. Komary, A. Alahmad et al., "A Novel Wireless Low-Cost Inclinometer Made from Combining the Measurements of Multiple MEMS Gyroscopes and Accelerometers," *Sensors*, vol. 22, no. 15, p. 5605.
[33] T. Shaown, M. Shohrab Hossain, and T. M. Mukur, "Low-Cost Health Monitoring System: A Smart Technological Device for Elderly People," in *Proc. 6th International Congress on Information and Communication Technology*, pp. 851-860, 2021.
[34] J. S. Botero-Valencia, J. Valencia-Aguirre, D. Gonzalez-Montoya et al., "A Low-Cost System for Real-Time Measuring of the Sunlight Incident Angle Using IoT," *HardwareX*, vo1. 11, p. e00272, Apr. 2022.

Panisara Kanjanarut is currently a Master's student in the Artificial Intelligence and Internet of Things Program (AIoT) at Sirindhorn International Institute of Technology (SIIT), Thammasat University, Thailand. She received her B.Eng. in Industrial Engineering from SIIT. Her research interests are computer vision in industrial applications, the Industrial Internet of Things (IIoT), and machine learning.

Warut Pannakkong is an Assistant Professor at the School of Manufacturing Systems and Mechanical Engineering (MSME), Sirindhorn International Institute of Technology (SIIT), Thammasat University, Thailand. He received his B.Eng. in Industrial Engineering and M.Eng. in Logistics and Supply Chain Systems Engineering both from SIIT in 2010 and 2014, respectively. He earned his Ph.D. in Knowledge Science from the Japan Advanced Institute of Science and Technology (JAIST) in 2017. His research interests include time series forecasting, data mining, machine learning, computer vision in industrial applications, discrete event system simulation and optimization, production planning, and logistics and supply chain management.

Automatic Detection of Fake Crowdfunding Projects

Qi Li[1] and Jian Qu[2]

[1,2]Faculty of Engineering and Technology, Panyapiwat Institute of Management,
Nonthaburi, Thailand
E-mail: 6372100118@stu.pim.ac.th, jianqu@pim.ac.th

Received: April 28, 2022 / Revised: June 10, 2022 / Accepted: October 18, 2022

Abstract—**There may be fake information in some crowdfunding projects. However, it is difficult for crowdfunding platforms and investors to find fake information in crowdfunding projects. At present, many scholars have studied the methods for identifying fake information, but most of them studied how to distinguish fake information from news articles. Therefore, this research focuses on how to identify fake information that may exist in crowdfunding projects. The detection of fake crowdfunding projects includes functions such as keyword extraction, external knowledge extraction, and classification of real and fake projects. To identify possible fake information in the crowdfunding project, we need to understand more about the crowdfunding project by extracting the keywords of the crowdfunding projects. Therefore, this research compared TF-IDF, CKPE, YAKE, RAKE, TextRank4zh, FastTextRank, HarvestText, and BERT pre-training model methods. We used precision, recall, and F1 scores to measure the effectiveness of the keyword extraction method. Then, we obtained features for judging the authenticity of crowdfunding projects by extracting external knowledge of keywords. Finally, projects were classified using a classification algorithm. The validity of this study for the classification of fake crowdfunding projects achieves 83.77% by the NB method in the dataset.**

Index Terms—**BERT Model, Crowdfunding Project, Information Extraction, Keywords Extraction, Web-Data**

I. Introduction

Today, crowdfunding projects are rising in fervor. As crowdfunding projects come into the public eye, some people will use crowdfunding projects to scam the funds raised for the project, hence the rise in the number of fake crowdfunding projects. According to statistics, the rate of project failures and fraud among investors exceeds 64.6% [1]. After the study on fraudulent crowdfunding projects, we divided fraudulent crowdfunding projects into two types. The first type is that the projects are theoretically and logically achievable, but the initiator of the project maliciously defrauds the crowdfunding funds. This type of fraudulent project is due to the reputation of the project's initiators themselves. The second type is the fake description and design of the crowdfunding project by the initiator of the project. This type of fraudulent project has contradictions and loopholes in the logic of the implementation, which cannot be achieved with the currently available technology. For example, the famous fake project-Triton underwater respirator[1]. The initiator of the project claimed to be able to make Triton, a small artificial fish gill. This respirator produces oxygen from water by electrification. Furthermore, the size is small, and someone can carry it. The developers of Triton claimed their device would allow a user to breathe underwater for 45 minutes at a maximum depth of 15 feet. However, the project is a copy of an idea from a science fiction movie, and Dr. Alistair Dove has confirmed in Deep Sea News that the project is impossible to achieve[2].

On the one hand, it is difficult for the average investor to identify fraudulent information in crowdfunding projects because fraudulent crowdfunding projects are usually deliberately falsified by fraudsters. Another hand, it is difficult for investors with some relevant professional knowledge to identify fraudulent information for crowdfunding projects because crowdfunding projects are usually novel and creative. For example, most of Triton's investors are divers with professional knowledge, but they were still defrauded. Thus, identifying fraudulent crowdfunding projects is a challenging task. Therefore, few people are devoted to the detection of fraudulent crowdfunding projects, and most scholars are devoted to the detection of fake news.

During the research, we found that some ideas from the method of fake news detection can be applied

[1] https://gearjunkie.com/news/triton-artificial-gills-breathe-underwater
[2] https://www.deepseanews.com/2014/01/triton-not-dive-or-dive-not-there-is-no-triton/

to the research of detecting fraudulent crowdfunding projects. For example, a method of fake news detection based on article information. Text information extraction can help us learn more information about crowdfunding projects, which can help us further research the detection of fake crowdfunding projects. The important information in the text can generally be reflected in the keywords. Therefore, we tried to obtain the feature information of crowdfunding projects by extracting the keywords of crowdfunding projects, and this research also compared different keyword extraction methods. However, extracting crowdfunding project information alone cannot help us identify fraudulent crowdfunding projects, so we proposed a comprehensive method to identify fraudulent crowdfunding projects.

The method proposed by the research first extracted the keywords of the crowdfunding project and then retrieved the characteristic information of the crowdfunding project to obtain the judgment basis of the fraudulent crowdfunding project. Finally, we used machine learning algorithms to classify the authenticity of crowdfunding projects.

II. Guidelines for Manuscript Preparation

The methods of keyword extraction generally include unsupervised and supervised methods.

A. Unsupervised Method

The unsupervised method extracted keywords without manual annotation of the corpus. The unsupervised method includes two approaches. (1) The first approach is the method of extracting keywords based on statistical features. Such methods mainly include TF-IDF [2], YAKE [3], and other methods. These methods do not require training data, and mainly use the position of words in the document, co-occurrence frequency [4], Term Frequency (TF), Inverse Document Frequency (IDF), N-gram, and PAT tree as statistical features to select terms as keywords [5]. This type of method can exclude words that are not relevant to the text and is fast to implement, but it cannot reflect the lexical organization structure within the article. Therefore, Qu et al. [6]. proposed to use symmetric conditional probability, chi-square, correlation measure, number of segments, and distance as statistical features to select terms as keywords, reflecting the relationship between words within the article [6]. (2) The second approach is the method that was based on graph networks. These methods built a graph network based on words or phrases. Then these methods used algorithms to calculate important nodes as keywords. Among them, widely used methods are the graph-based ranking model (TextRank) [7] and the rapid keyword extraction algorithm (RAKE) [8]. TextRank algorithm is an important ranking algorithm that can extract keywords, keyphrases, and key sentences from documents. This method inherited the idea of PageRank. Compared with the method of TF-IDF, it can make full use of the relationship between text elements. However, this method still does not solve the problem that high-frequency words excessively affect the results.

In contrast, the RAKE algorithm introduced a concept of degree, did not make any distinction between words and phrases, and used the co-occurrence information of words to determine keywords. The algorithm used Word Frequency, Word Degree, and the ratio of the degree to frequency as features for keyword extraction. However, this method still has a strong dependence on the list of deactivated words. Since the unsupervised keyword extraction method cannot synthesize multiple information of the text in text information extraction, it is not very effective for keyword extraction.

B. Supervised Method

To improve the effectiveness of keyword extraction, supervised keyword extraction methods are proposed. Supervised learning uses the model learned from a set of trained texts to extract keywords. The supervised method has two aspects. The first type is based on traditional supervised learning methods, which include Naive Bayes (NB), Decision Tree (DT), Logistic Regression (LR), K Nearest Neighbors (k-NN), Support Vector Machine (SVM) [9], and methods based on Learning to Rank (LTR). Zhang et al. [9] have proposed an SVM-based method for keyword extraction, which uses “global context information” and “local context information” to extract keywords from documents. However, these methods have certain limitations in adaptability.

To enhance the adaptability of the method, Qu et al. proposed a method that combines statistical features and adaptive rules to extract keywords [10]. The adaptive rule was to modify the basic rule through each term item based on the basic rule (the predefined regular expression matching rules) and then form an adaptive regular expression rule. Another study by Qu [11] proposed that when the search distance was eight, the key characters presented the best recall rate. Thus, a rule-based algorithm that automatically adjusts the candidate generation system can effectively improve the efficiency of candidate selection.

However, the keyword extraction algorithm combining statistical features and adaptive rules relies more on statistical features, and its rules have limitations when applied to other types of text. Therefore, the second type of keyword extraction based on deep learning was proposed. In 2016, Zhang et al. proposed a new deep Recurrent Neural Network (RNN) model, which can jointly handle keyword ranking and keyphrases generation tasks [12]. However, this model is limited in its perf Meng, Zhao et al. proposed a keyphrases generation model

(CopyRNN) with an encoder-decoder framework [13]. This model not only captured the semantic meaning behind the text but also generated missing key phrases based on the text semantics.

However, the RNN method has a strong dependence on the calculation of the sequence in the training process. The distance between time steps may lead to the problem of gradient disappearance, which can be well solved by the BERT model. This model was a typical two-way coding model. It used Transformers as the main framework of the algorithm to capture the two-way relationship in sentences, which improved the model's language representation ability and feature extraction ability [14]. The BERT model has been open-sourced tool by Google, and researchers can use the BERT pre-training model for their natural language processing research. Such as paraphrase recognition, semantic text similarity, repeated question detection, and question-answer retrieval. The BERT model provided a simple model (BERT-base), a complex model (BERT-large), and many embedded models.

This study applied the BERT pre-training model to simplified Chinese text processing and compared the effects of different BERT pre-training models on keyword extraction.

The main methods for the detection of fake crowdfunding projects are divided into two categories: First, the language approach, language patterns linked to false (contradictory); P E Rez -Rosas, who proposed language method to detect conflicts. Second, the network method, using network information to fake (contradictory) connections. Language methods include (1) Data representation methods. It uses the bag-of-words method to aggregate and analyzes the frequency of words or multiple words to reveal contradictions. However, this method not only relies on the language but also the isolated n-gram, which will be out of touch with the contextual information. (2) In-depth syntax method. Various studies have shown that the analysis of words alone is not enough to predict contradictions. In-depth grammatical analysis can predict contradictions. The in-depth grammatical analysis is realized by structural probability context-free grammar, which converts sentences into rewriting rules to describe the grammatical structure of sentences. The accuracy of this method to detect contradictions is 85%-91%, and the specific accuracy depends on the type of rule used. Besides, you can also rely on third-party tools for in-depth syntax analysis to achieve contradiction detection, such as AutoSlog-TS syntax analyzer and other tools. However, this method alone is not enough to detect contradictions, and it needs to be combined with other language methods or network analysis techniques. (3) Discourse analysis, as an alternative method of contradictory clues, analyzes the contradiction through the degree of compatibility between information and information. The accuracy of this method for contradiction detection is 91%, but this method is limited to the application field 4. Rhetorical structure and discourse analysis. Discourse description is realized through the analysis framework of rhetorical structure theory, which identifies examples of rhetorical relations between language elements. But this method is too inferior in the accuracy of contradiction detection. Both of these methods are combined with machine learning to train classifiers to monitor and analyze. Hai et al. proposed semi-supervised learning that is a combination of language methods or network methods and machine learning methods.

III. Methods

This research compared several keywords extraction methods, such as TF-IDF, YAKE, RAKE, TextRank4zh, FastTextRank, HarvestText, and BERT pre-training models for a crowdfunding project. In addition, we tested five BERT pre-training models. See Table VIII for details. Furthermore, we performed external information retrieval on the extracted keywords, and the retrieved information is used as a feature for judging fraudulent crowdfunding information. The features are trained by using traditional machine learning to identify fraudulent crowdfunding projects. The flowchart of this research is shown in Fig. 1.

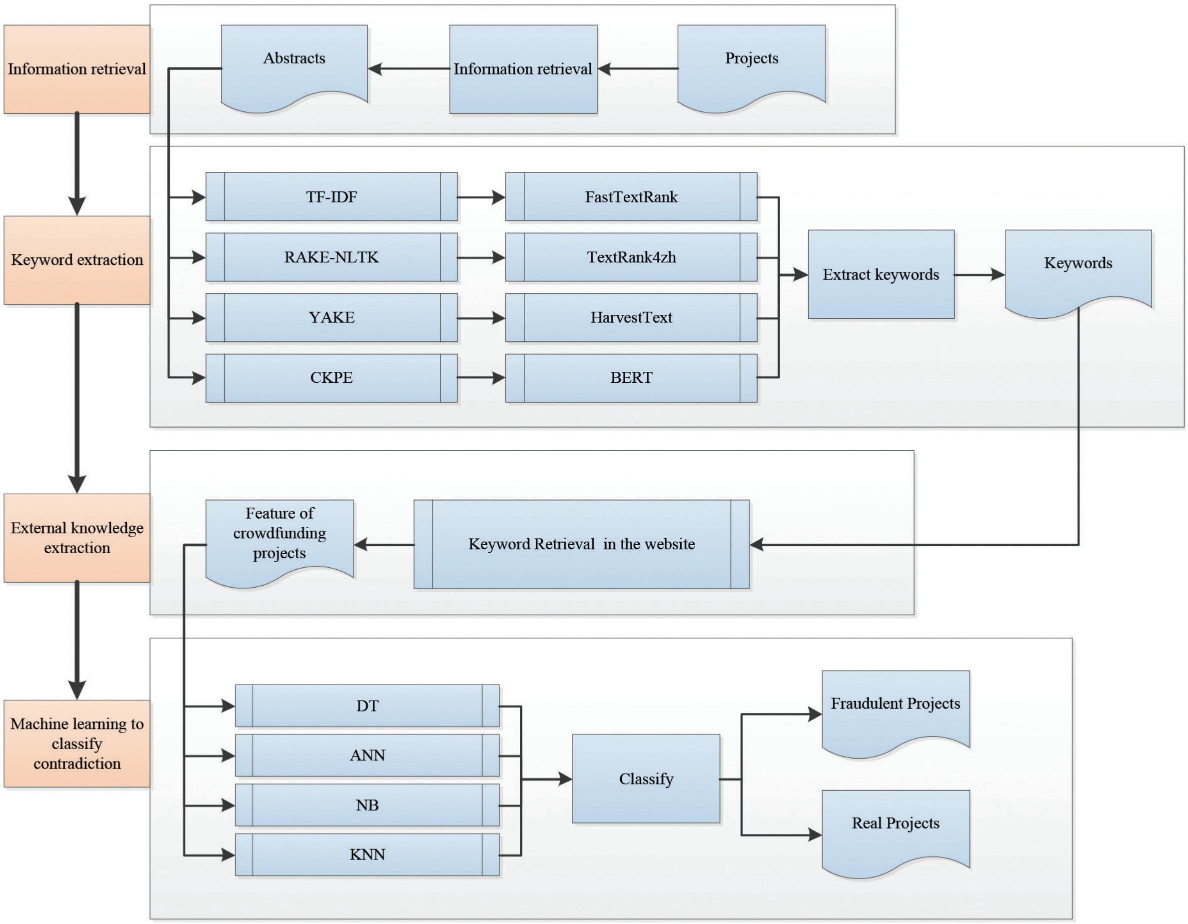


Fig. 1. The flowchart of the research

This study included four experimental steps, information retrieval, keyword extraction, external knowledge extraction, and machine learning for contradiction detection and classification. The specific steps are shown in Fig. 1. Among them, the information retrieval process is the project summary information feedback. In the process of keyword extraction, we compared the traditional statistical method and the deep learning keyword extraction method. The best results of keyword extraction were used for our experiments. We use keyword extraction for network knowledge extraction to obtain the features of crowdfunding projects to judge their authenticity of crowdfunding projects. In the final step, we used the classification algorithms to classify the features of crowdfunding projects obtained by Web-knowledge extraction to predict whether the projects are fake or not.

A. Keyword Extraction

1) TF-IDF

Initially, the TF-IDF algorithm was proposed to calculate the importance of words to documents. According to this algorithmic idea, if the weight of the word was higher when calculating the weight of feature words, the frequency of the word in the text is higher, the more likely it was a keyword [2]. Therefore, after calculating the weights of feature words, the weight values of all words are sorted. Then the method selects the Top K keywords with the largest weights or the keywords with weights greater than a certain threshold. Table I is an example of using TF-IDF to extract the keywords.

TABLE I
EXAMPLE PF EXTRACTING KEYWORDS BY TF-IDF ALGORITHM

Construct	Data
Source Input	MATE X是世界上最酷的可折叠电动自行车。在世界自行车之都丹麦哥本哈根，人们只想踩踏板就对它进行了构想和设计……升级后的制动更加平稳，所需制动力更少。 ***Translation:*** "MATE X is the coolest foldable electric bike in the world. It was conceived and designed in Copenhagen, Denmark, the cycling capital of the world, where people just wanted to pedal …… the upgraded brakes are smoother and require less stopping power."
Output KeyWords	[‘制动’，‘可折叠’，‘自行车’，‘升级’ ‘踏板’，‘显示器’，‘动力’，‘电动’，‘功率’，‘配备’] ***Translation:*** [‘brake’, ‘foldable’, ‘bike’, ‘upgrade’, ‘pedal’, ‘display’, ‘power’, ‘electric’, ‘power’, ‘equipped’]

Initially, the TF-IDF algorithm was proposed to calculate the importance of words to documents. According to this algorithmic idea, if the weight of the word was higher when calculating the weight of feature words, the frequency of the word in the text is higher, the more likely it was a keyword [2]. Therefore, after calculating the weights of feature words, the weight values of all words are sorted. Then the method selects the Top K keywords with the largest weights or the keywords with weights greater than a certain threshold. Table I is an example of using TF-IDF to extract the keywords.

2) YAKE

YAKE is an unsupervised method of automatically extracting keywords based on text features [3]. The keywords extractor, which relies on the statistical characteristics of the text in a single document, selects the most important keywords in the text. This method does not need to be trained on a specific set of text documents, nor does it rely on dictionaries and external corpora. Based on those advantages, we used YAKE for keywords extraction of crowdfunding projects.

YAKE Keywords extractor has been currently used as a keyword's extraction tool. YAKE can be used as a CLI utility in a Docker container or as a REST API server in a Docker container [3]. However, this study used YAKE as a standalone tool and used YAKE as a keyword's extraction tool in python. This method also provided an end-to-end keyphrases extraction pipeline to extract keywords from text documents.

YAKE Keywords extractor can extract keywords for various languages. As shown in Table II, its input is a source input in the example of Table I. The keyphrases column (Table II) is the keyphrases extracted with the YAKE method. The score column (Table II) is the similarity score between the keywords/keyphrases and the text. If the score was lower, the keywords would be more relevant.

TABLE II
EXAMPLE PF EXTRACTING KEYWORDS BY YAKE ALGORITHM

Id	Keyphrases	The Similarity Score
1	多个支持者交付了 mate bikes ***Translation***：Multiple supporters delivered mate bikes	0.0226245446
2	mate bike 在成功的基础上, 我们将所有的时间, 精力, 血液和汗水投入到电动自行车的美丽之中, 这对于任何骑手, 任何旅程, 在任何特定情况下都是完美的 ***Translation***：Based on the success of mate bike, we put all our time, energy, blood, and sweat into the beauty of electric bikes that are perfect for any rider, any journey, and in any given situation.	0.0226245446
…	….	…
9	mate x 拥有功率高达 ***Translation***：mate x has power up to	0.1598191003
10	usd （价值） ***Translation***：usd (value)	0.1652103123

3) RAKE-NLTK

NLTK is a toolkit of natural language processing, which can greatly contribute to the field of natural language processing [15]. The toolkit is very powerful, with powerful functions for text processing, such as text classification, text tagging, stemming, semantic inference, and other functions.

TABLE III
EXTRACT KEYWORDS/KEYPHRASES BY RAKE-NLTK

ID	Keyphrases
1	我们还提供了仅需99 usd （ 价值149 usd ） 即可升级到彩色smart美丽显示器 （ 如图 ） 的选项 ***Translation***：We also have the option to upgrade to a colour smart beautiful display (pictured) for just 99 usd (worth 149 usd).
2	并在所有型号中均 配备了 令人印象深刻的48v电池和控制器 ***Translation:*** And comes with an impressive 48v battery and controller in all models.
…	…
9	该齿轮箱具有改进的易换档性能 ***Translation:*** The gearbox has improved easy shifting performance.
10	该新的计算机显示器配备有背光led作为标准 ***Translation***：This new computer monitor is equipped with backlit led as standard.

In this research, we combined the rapid automatic keywords (RAKE) [8] algorithm with the NLTK toolkit [16]. It formed a powerful keyword extraction method, which was called RAKE-NLTK. As shown in Table III, the keywords/keyphrases are extracted by using this method. Its input is a source input in example Table I.

4) TextRank4zh

The TextRank4zh method is a keyword extraction method [7]. The method first splits the original text into sentences, then filters the stop words of the sentence, and last retains the words of the specified part of speech [7]. The method was to calculate the importance of word nodes according to the principle of the graph to obtain keywords and key phrases.

An example of using the TextRank4zh method to extract keywords is shown in Table IV. The input for this example is the source input of Table I. The weight scores of keywords are listed in the score column (Table IV). According to the TextRank4zh, the higher the weight score, the more important the keywords.

TABLE IV
EXTRACT KEYWORDS BY TextRank4zh AND THE SCORE OF KEYWORDS

Id	Keywords	Score
1	Mate	0.028890332978898956
2	功率 ***Translation***：power	0.024104175996508638
3	…	…
9	电动 ***Translation***：electric	0.01716594583001619
10	构想 ***Translation***： conception	0.016248380651252663

5) FastTextRank

FastTextRank is divided into FastTextRank Word and FastTextRank Sentence. FastTaxtRank Word is a method that divides the article into sentences. It calculates the similarity between words, constructs graphics according to the word similarity, and calculates the importance of each word by an iterative algorithm to obtain keywords [16]. This method is based on the TextRank graphics algorithm. FastTaxtRank Sentence can extract sentences. The method calculates the similarity between sentences by the cosine similarity of word vectors.

TABLE V
EXTRACT KEYWORDS AND KEYPHRASES BY FASTTEXTRANK

Id	Keywords	Keyphrases
1	到 ***Trans-lation:*** reach	我们还升级了Tektro制动系统，包括杠杆和 卡钳，以实现更快，更有效的制动力。 ***Translation:*** We also upgraded the Tektro braking system, including levers and calipers for faster, more effective stopping power.
2	功率 ***Trans-lation:*** power	新款MATE X拥有功率高达750W的强大动力，并在所有型号中均 配备了 令人印象深刻的48V电池和控制器。 ***Translation:*** The new MATEX is packed with power up to 750W and comes with an impressive 48V battery and controller in all models.
…	…	…
9	配备 ***Trans-lation:*** equipped	可折叠踏板是一个全新的设计封装在一个更坚固的踏板架为改进的功率转移。 ***Translation:*** The foldable pedals are an all-new design packaged in a sturdier pedal rack for improved power transfer.
10	USD	在世界自行车之都丹麦哥本哈根，人们只想踩踏板就对它进行了构想和设计。 ***Translation:*** It was conceived and designed in Copenhagen, Denmark, the cycling capital of the world, where people just wanted to pedal.

The word vectors are obtained from each sentence that was being compared. Table V is an example of using FastTextRank to extract keywords and keyphrases, and its input is the source input in Table I.

6) HarvestText

HarvestText is an unsupervised method. The analysis of specific domain text can be processed by the HarvestText method that can integrate domain knowledge (types, aliases) [17]. The HarvestText method extracts keywords based on the TextRank algorithm. The HarvestText method can also use dependency grammar (DG)[1] to extract semantic triple[2] that may represent events.

[1] https://en.wikipedia.org/wiki/Dependency_grammar
[2] https://en.wikipedia.org/wiki/Semantic_triple

TABLE VI
SYNTACTIC STRUCTURE ANALYSIS RESULTS BY HARVESTTEXT MEHTOD

Id	Word literal value/entity name	Part-of-speech	Dependency	Dependent sub-words
0	MATE 是 (MATE is)	v	主谓关系 (subject-predicate relationship)	6
1	世界(world)	n	定中关系 (Ding-China relations)	2
2	上(on)	f’	状中结构 (mesostructure)	6
3	‘酷(cool)	a	主谓关系 (subject-predicate relationship)	6
4	的	u	右附加关系 (right append relationship)	3
5	可(can)	v	状中结构 (mesostructure)	6
6	折叠 (foldable)	v	核心关系 (core relationship)	-1
7	电动 (electric)	b	定中关系 (Ding-China relations)	8
8	自行车 (bike)	o	动宾关系 (verb-object relationship)	6
9	。	w	标点符号 (Punctuation)	6

We analyzed the dependency grammar of crowdfunding projects by the HarvestText method and extracted semantic triple to represent events. The method obtains the key sentences of the crowdfunding projects by combining the triples. To extract facts from sentences is based on dependency grammar. Firstly, the method found meaningful triples in the sentence using a subject, predicate, and other syntactic relationships. Then Semantic triple, which centered on the predicate, were extracted. The extracted sentences that may represent the event included three syntactic structures: Subject-Verb-Object (SVO) structure, Object–Verb–Subject (OVS) structure, and Subject–Object–Verb (SOV) structure. The stop word list in this method is Baidu stop words by default. Syntactic structure analysis results are shown in Table VI.

Table VI is an example of fact extraction using the HarvestText method based on dependency grammar. This example performed part-of-speech tagging and dependency analysis on words in a sentence and then extracted semantic triple that may express events. For example, the original sentence: “MATE X 是世界上最酷的可折叠电动自行车。 (MATEX is the coolest foldable electric bike in the world.)” The proposed semantic triple is: (‘MATE X 是’, ‘折叠’, ‘电动自行车’), (‘MATE X is’, ‘foldable’, ‘electric bike’) as shown in Table VII.

TABLE VII
EXTRACT SEMANTIC TRIPLE BY HARVEST TEXT METHOD

Construct	Data
Input sentence	MATE X 是世界上最酷的可折叠电动自行车。 ***Translation:*** MATEX is the coolest foldable electric bike in the world.
Syntactic relation	[0, 'MATE X是', 'v', '主谓关系', 6] [1, '世界', 'n', '定中关系', 2] [2, '上', 'f', '状中结构', 6] [3, '最酷', 'a', '主谓关系', 6] [4, '的', 'u', '右附加关系', 3] [5, '可', 'v', '状中结构', 6] [6, '折叠', 'v', '核心关系', -1] [7, '电动', 'b', '定中关系', 8] [8, '自行车', '设备', '动宾关系', 6] [9, '。', 'w', '标点符号', 6]
Event triplet	MATE X 是折叠电动自行车 ***Translation:*** MATE X is a folding electric bike.

7) BERT

BERT (Bidirectional Encoder Representations from Transformers) is a pre-training technology for natural language processing proposed by Google. BERT is a deep two-way, the unsupervised model that only uses a plain text corpus for training [18].

This research used the BERT pre-training model to extract keywords from the text. In extracting keywords, firstly, BERT used the Count Vectorizer in Scikit-Learns to remove stop words and extract keywords/keyphrases candidates [18]. Secondly, it converted documents and candidates into vectors by embedding the BERT pre-trained model. Then it calculated the cosine similarity between the candidate vector and the document vector. Finally, keywords/keyphrases were selected according to the cosine similarity [18]. In this paper, the flow chart of applying the BERT pre-training model to extract keywords is shown in Fig. 2.

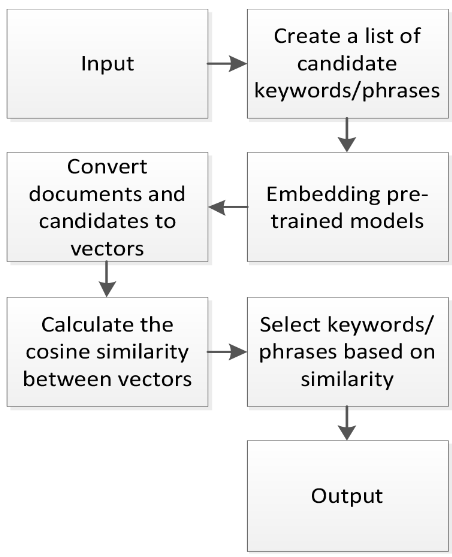


Fig. 2. Flow chart of extracting keywords with BERT pre-training model

When using the BERT pre-training model for keywords extraction, it did not need to train the model and only fine-tune the model to obtain the keywords extracted using different sentence transformer models. When using the CountVectorizer in Scikit-Learns to create a list of keywords/keyphrases candidates, we can customize the size of the candidate words and the list of stop words. When calculating cosine similarity, we can also use different algorithms (maximum and similarity, maximum marginal correlation) to diversify the results [18].

TABLE VIII
BERT PRE-TRAINING MODEL

Label	Model	Annotation
BERT-1	*Distilbert-base-nli-mean-tokens*	Semantic similarity model
BERT-2	*Distiluse-base-multilingual-cased-v1*	Multilingual knowledge distilled version 1 of multilingual Universal Sentence Encoder
BERT-3	*Distiluse-base-multilingual-cased-v2*	Multilingual knowledge distilled version 2 of multilingual Universal Sentence Encoder
BERT-4	*Paraphrase-distilroberta-base-v1*	Paraphrase recognition model
BERT-5	*Paraphrase-xlm-r-multilingual-v1*	Multilingual version of paraphrase-distilroberta-base-v1

Many BERT pre-training models have been proposed, and each model was good in different fields. This study analyzed and compared five models in the multi-language general model. In this study, we shorten the BERT model names from BERT-*1* to BERT-*N*. The information on each BERT pre-training model is shown in Table VIII.

B. External Web-Knowledge Extraction

This research proposed a method to extract the features of crowdfunding projects. The method used the feedback results of retrieving external knowledge as features of whether the crowdfunding projects are fake or not. The feature items we retrieved are listed in Table IX. The Web-knowledge retrieval was done by obtaining external knowledge related to the project keywords on the web [19]. Keywords represent some features of an item, retrieving external knowledge about keywords can help us learn more information about whether the technology of the project can be achieved or not. Since, the initiator of crowdfunding wanted to obtain financial support for the realization of the project through crowdfunding. If the current technology cannot realize the project, the project is more likely to be a fraudulent project. Therefore, we retrieved the keywords that can represent the characteristic information of the crowdfunding project to determine whether the current technology can realize a certain characteristic function of the project.

TABLE IX
FEATURE WE RETRIEVED

Feature-ID	Feature	Example	Translation
1	Keyword and “Category”	可折叠 and “自行车”	Foldable and “bike”
2	“Keyword” and “Category”	“可折叠” and “自行车”	“Foldable” and “bike”
3	“Keyword” and Category	“可折叠” and 自行车	“Foldable” and bike
4	Keyword and Category	可折叠 and 自行车	Foldable and bike
5	“Keyword”	“可折叠”	“Foldable”
6	Keyword	可折叠	Foldable

C. Machine Learning

This research used the results of retrieving external knowledge feedback as features to select and classify projects using traditional machine learning. In this research, the effectiveness of classification with Decision Tree (DT), Naïve Bayes (NB), Artificial Neural Network (ANN), Support Vector Machines (SVM) and other algorithms were compared. The specific results are shown in section IV.

IV. RESEARCH METHODS EXPERIMENT RESULT AND DISCUSSION

A. Data Set

There are two data sets for this research. Data set one contained 120 crowdfunding projects from a list of the highest-funded crowdfunding projects in Wikipedia[1]. For better comparison with dataset two, we only selected the project from Kickstarter and Indiegogo. There are also no restrictions on the types of crowdfunding projects. Data set two contained 100 crowdfunding projects, which came from Kickstarter’s and Indiegogo’s official webpage. For this dataset, we selected 20 fake crowdfunding projects and 80 real crowdfunding projects. We collected a hundred datasets based on the popularity of crowdfunding on Kickstarter and Indiegogo. Only 20 fake crowdfunding project datasets are collected since fake crowdfunding projects are difficult to found. In this study, a fake crowdfunding project refers to a project similar to Triton where the promoters falsely describe and design the project, but there are logical contradictions and loopholes.

[1] https://en.wikipedia.org/wiki/List_of_highest-funded_crowdfunding_projects

B. Results of Keywords Extraction

This research used TF-IDF, YAKE, RAKE-NLTK, TextRank4zh, FastTextRank, HarvestText, and BERT to extract keywords. According to the precision, recall, and F-score of different methods in extracting keywords to compare those method's effectiveness. The calculation formulas of precision, recall and F-score are shown in formulas (1), (2), and (3).

$$precison = \frac{\left|\{rel\,doc\} \cap \{ret\,doc\}\right|}{\{ret\,doc\}} \tag{1}$$

Precision is the ratio of relevant documents to retrieved documents.

$$recall = \frac{\left|\{rel\,doc\} \cap \{ret\,doc\}\right|}{\{rel\,doc\}} \tag{2}$$

Recall is the fraction of the documents that are relevant to the query that are successfully retrieved.

$$F_1\text{-}score = \frac{2 \cdot precision \cdot recall}{(precision + recall)} \tag{3}$$

F_1-*score* Is the weighted harmonic mean of precision and recall, also known as the F_1-*score*.

The symbols in the formula are explained in Table X.

TABLE X
SYMBOL DESCRIPTION

Symbol	Description
Rel Doc	Relevant Documents
Ret Doc	Retrieved Documents

In this research, relevant documents refer to the correctness of the extracted keywords/keyphrases. Retrieved documents refer to keyword candidates extracted by the method.

In order to create a baseline, we hired three master students to evaluate the correctness of the extracted keywords manually. Table IX shows the number of keyword candidates extracted by the traditional methods.

The number of correct keywords is shown in Table X. The number of incorrect keywords is shown in Table XI. The methods include TF-IDF, YAKE, RAKE-NLTK, TextRank4zh, FastTextRank, HarvestText.

The precision, recall, and F-score of the traditional methods are shown in Table XI.

TABLE XI
THE PRECISION, RECALL, AND F1-SCORE OF DIFFERENT METHODS

Methods	Dataset	Precision	Recall	F-score
TF-IDF	A	33.83%	100%	50.56%
	B	52.29%	100%	68.67%
YAKE	A	45.79%	100%	62.81%
	B	60.22%	100%	75.17%
RAKE-NLTK	A	30.98%	100%	47.30%
	B	50.02%	100%	66.68%
TEXTRANK4ZH	A	25.33%	100%	40.43%
	B	42.00%	100%	59.15%
FASTTEXTRANK	A	22.00%	100%	36.07%
	B	28.39%	100%	44.22%
HARVESTTEXT	A	38.77%	100%	55.88%
	B	56.98%	100%	72.60%
CKPE	A	64.43%	100%	78.37%
	B	66.60%	100%	79.95%

It can be seen from Table XI that the effectiveness of the YAKE method for extracting keywords is higher than that of TF-IDF, RAKE-NLTK, TextRank4zh, FastTextRank and HarvestText method. According to the F-score and the Precision values, we can see that the FastTextRank method is less effective than the TextRank4zh method. Because FastTextRank is based on the TextRank4zh method to improve the speed of extracting keywords. Since HarvestText is based on extracting semantic triple to extract keyphrases, its effectiveness is higher than other algorithms, based on statistical rules, location, and other features.

The F_1-*score* of YAKE, RANKE-NLTK, TextRank4zh, FastTextRank, HarvestText, and CKPE methods are shown in Fig. 3.

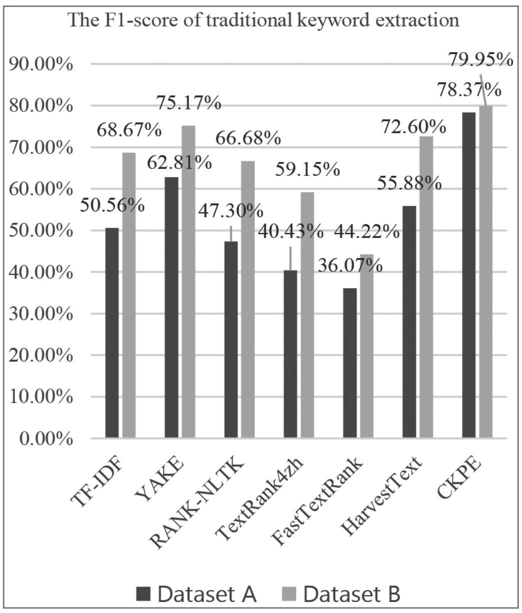


Fig. 3. The F1-score of traditional keyword extraction

TABLE XII
THE PRECISON, RECALL AND F1-SCORE OF THE BERT PRE-TRAINING MODEL

Methods	Dataset	Precision	Recall	F-score
BERT-1	A	45.02%	100%	62.09%
	B	38.73%	100%	55.84%
BERT-2	A	51.98%	100%	68.40%
	B	62.82%	100%	77.16%
BERT-3	A	48.99%	100%	65.76%
	B	62.54%	100%	76.95%
BERT-4	A	34.88%	100%	51.72%
	B	37.32%	100%	54.35%
BERT-5	A	31.84%	100%	48.30%
	B	35.15%	100%	52.02%

It can be seen Table XII the effectiveness of the BERT model for extracting keywords. The effectiveness of the BERT model for extracting keywords varies according to the type of the BERT pre-training model. Among the BERT pre-training model used to extract keywords, the BERT-1 and BERT-2 are more effective. The effectiveness of the BERT-2 model is better than other BERT pre-training models in Table XI. The model of BERT-*N* refers to Table VIII.

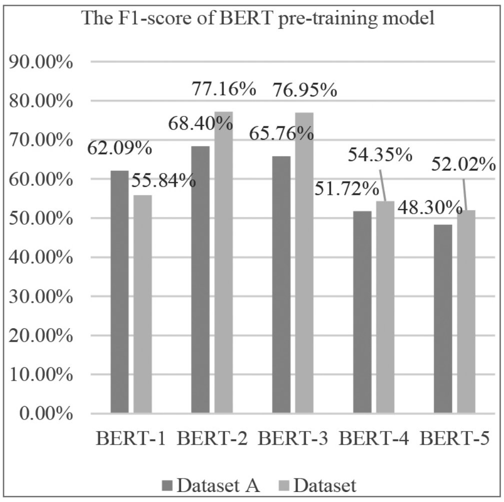


Fig. 4. The F1-score of BERT pre-training model

In addition, we also compared the methods of traditional keyword extraction with the methods of deep learning keyword extraction. The results of the comparison are shown in Fig. 5.

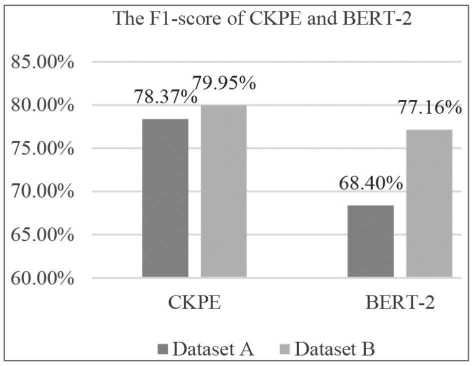


Fig. 5. The F1-score of CKPE and BERT-2

C. *Results of Research*

Among the traditional keyword extraction methods, TF-IDF, YAKE, and CKPE methods have better effects than other methods. Since YAKE extracted long sentences, although the sentence content could help us understand more information on the crowdfunding project, the sentence length was not convenient for later text research. Among the keyword extraction methods based on deep learning, we compared the keyword extraction effect of five BERT embedding models, and the best performance is the BERT-2 model. The results of BERT for Chinese text processing show that the length of the key sentences extracted by the BERT model is too long and some are too short. Sentences that are too long will reduce the computational speed. Sentences that are too short will reduce the validity of the research. Considering the processing speed and validity of the method, we tried to choose a method with a faster calculation speed. Therefore, in this study, we used TF-IDF and CKPE-extracted keywords for external knowledge extraction to obtain features for determining the authenticity of crowdfunding projects [20].

We used machine learning to learn features for classifying fake crowdfunding projects. In the machine learning algorithm, we compared the effectiveness of Decision Tree (DT), Naïve Bayes (NB), Artificial Neural Network (ANN), and Support Vector Machines (SVM) for classifying fake crowdfunding projects. The meta-level and feature selection of the four algorithms and the effectiveness of the optimized algorithms are compared. The specific results are shown in Table XIII.

TABLE XIII
THE RESULTS OF CALSSIFICATION WITH MACHINE LEARNING

Methods	Method	CKPE		TF-IDF	
		A	B	A	B
DT	ML	97.50%	74.22%	97.50%	76.11%
	BE	97.50%	81.56%	97.50%	77.33%
	OP	97.50%	80.52%	97.50%	80.52%
NB	ML	97.50%	74.11%	97.50%	76.22%
	BE	97.50%	79.11%	97.50%	79.44%
	OP	97.50%	83.77%	97.50%	82.11%
ANN	ML	97.50%	77.33%	97.50%	78.22%
	BE	97.50%	80.33%	97.50%	80.33%
	OP	97.50%	81.63%	97.50%	80.61%
SVM	ML	97.50%	80.33%	97.50%	80.33%
	BE	97.50%	80.33%	97.50%	80.33%
	OP	97.50%	80.61%	97.50%	80.61%

Source: ML: Meta-level;
BE: Backward Elimination;
OP: Optimize Parameters (Evolutionary)

Table XIII implies that our proposed method performs better on dataset A, which is because dataset A is a highly imbalanced dataset. Dataset A is from Wikipedia, so we could not achieve its balance between fake and real projects. Therefore, we try to solve this problem by showing with a different dataset which is a more balanced dataset. Dataset B is the balanced dataset we retrieved from Kickstarter and Indiegogo, so the results of our proposed method on dataset B are more indicative of the effectiveness of the method. Among the four methods, DT, NB, ANN, and SVM, we found that the classification algorithm of NB is more effective for the experiment.

Since we determine whether a crowdfunding project is fake or not based on whether the innovative technology proposed by the project can be achieved, the method we propose is only applicable to technology-based crowdfunding projects. In this research, we used the accuracy of human identification as a baseline. Since much fake information about crowdfunding projects is difficult or even impossible to identify artificially, the accuracy of this baseline was 75%.

V. CONCLUSION

There is possible fake information in some crowdfunding projects. Unfortunately, the crowdfunding platform cannot easily identify such fake information in the crowdfunding project. To identify possible fake information in the crowdfunding project, we proposed a method to detect fake crowdfunding projects. In our proposed method, the keyword extraction method is firstly used to extract keywords from the project abstract. Then we search for external knowledge based on keyword extraction, and the feedback information from external knowledge retrieval is used as the feature basis for judging the authenticity of crowdfunding projects. Finally, the machine learning features are used to classify the real and fake crowdfunding projects. We need to extract the keywords of the crowdfunding project and analyze the information of the crowdfunding project. Many studies are devoted to analyzing the successful or failed factors of crowdfunding projects, but they neglect to study the authenticity of crowdfunding projects. In the process of analyzing the authenticity of the crowdfunding project, we need to extract text keywords. At present, there are many methods for extracting keywords from Chinese text. We want to compare the effectiveness of keyword extraction methods for crowdfunding projects through experiments.

The experimental results show that the BERT method is more effective than other methods for keyword extraction. Among the different BERT embedded models, the most effective one is BERT-2. In the classification results, the best result is the NB algorithm, which has an efficiency of 83.77% in the classification of fake crowdfunding projects. In the future, we plan to train the BERT pre-training model to improve the effectiveness of the BERT model in extracting keywords for crowdfunding projects.

ACKNOWLEDGMENT

The first author conducted the experiment and drafted the manuscript. The last author guided and advised the experiment and co-drafted the manuscript. The first and last authors each contributed 50% equally to this work. The last author is the corresponding author.

The first author received scholarship support from CPALL for conducting this research in PIM.

REFERENCES

[1] S. Pandey, S. Goel, S. Bansla et al., "Crowdfunding Fraud Prevention Using Blockchain," in *Proc. 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)*. 2019, pp. 1028-1034.

[2] C. Zhang, "Automatic Keyword Extraction from Documents Using Conditional Random Fields," *Journal of Computational Information Systems,* vol. 4, no. 3, pp. 1169-1180, 2008.

[3] R. Campos, V. Mangaravite, A. Pasquali et al., "YAKE! Keyword Extraction from Single Documents Using Multiple Local Features," *Information Sciences,* vol. 509, pp. 257-289, 2020.

[4] C. Wartena, R. Brussee, and W. Slakhorst. "Keyword Extraction Using Word Co-Occurrence," in *Proc. 2010 Workshops on Database and Expert Systems Applications*, 2010, pp. 54-58.

[5] F. Bayatmakou, A. Ahmadi, and A. Mohebi. "Automatic Query-Based Keyword and Keyphrase Extraction," in *Proc. 2017 Artificial Intelligence and Signal Processing Conference (AISP)*, 2017, pp. 325-330.

[6] J. Qu, T. Theeramunkong, N. Le Ming et al. "A Flexible Rule-Based Approach to Learn Medical English-Chinese OOV Term Translations from the Web," *International Journal of Computer Processing of Languages*, vol. 24, no. 2, pp. 207-236, 2012.

[7] R. Mihalcea and P. Tarau, "Textrank: Bringing Order into Text," in *Proc. The 2004 Conference on Empirical Methods in Natural Language Processing*, 2004, pp. 404-411.
[8] S. Rose, D. Engel, N. Cramer et al., "Automatic Keyword Extraction from Individual Documents," *Text Mining: Applications and Theory*, vol. 1, 2010, pp. 1-20.
[9] K. Zhang, H. Xu, J. Tang et al., "Keyword Extraction Using Support Vector Machine," in *Proc. International Conference on Web-Age Information Management*, 2006, pp. 85-96.
[10] J. Qu, NL. Minh, and A. Shimazu, "Web Based English-Chinese OOV Term Translation Using Adaptive Rules and Recursive Feature Selection," In *Proc. The 25th Pacific Asia Conference on Language, Information and Computation*, 2011, pp.1-10.
[11] J. Qu, T. Theeramunkong, C. Nattee et al., "Web Translation of English Medical OOV Terms to Chinese with Data Mining Approach," *Science Technology Asia*, vol. 16, no. 2, pp. 26-40, 2011.
[12] Q. Zhang, Y. Wang, Y. Gong et al., "Keyphrase Extraction Using Deep Recurrent Neural Networks on Twitter," in *Proc. The 2016 Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 836-845.
[13] R. Meng, S. Zhao, S. Han et al., "Deep Keyphrase Generation," in *Proc. The 55th Annual Meeting of the Association for Computational Linguistics*, 2017, pp. 582-592.
[14] J. Devlin, M. W. Chang, K. Lee et al., "Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding," in *Proc. The 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2019, pp. 4171-4186.
[15] E. Loper and S. Bird, "Nltk: The Natural Language Toolkit," in *Proc. The ACL Interactive Poster and Demonstration Sessions*, 2004, pp. 214-217.
[16] D. Adimanggala, F. A. Bachtiar, and E. Setiawan, "Evaluasi Topik Tersembunyi Berdasarkan Aspect Extraction Menggunakan Pengembangan Latent Dirichlet Allocation," *Jurnal RESTI*, vol. 5, no. 3, pp. 511-519, 2021.
[17] J. Jiang, "Information Extraction from Text," *Boston, MA: Springer*, 2012, pp. 11-41.
[18] P. Sharma and Y. Li. (2019, Aug. 2). *Self-Supervised Contextual Keword and Keyphrase Retrieval with Self-Labelling*. [Online]. Available: https://shorturl.asia/vwUqZ
[19] J. Qu, N. Phaphoom, C. Wangtragulsang et al., "Social Media Contact Information Extraction," in *Proc. 2018 International Conference on Information Technology (InCIT)*, 2018, pp. 1-5.
[20] Q. Li and J. Qu, "Product Ontology Construction for Crowdfunding Pojects", *presented at the 7th International Conference on Business and Industrial Research (ICBIR) 2022*, Bangkok, Thailand, May 19-20, 2022.

Qi Li is currently studying for the Master of Engineering (Engineering and Technology), Panyapiwat Institute of Management, Thailand. Her research interests are natural language processing and information retrieval.

Jian Qu is a full-time lecturer at the Faculty of Engineering and Technology, Panyapiwat Institute of Management. He received Ph.D. with Outstanding Performance award from Japan Advanced Institute of Science and Technology, Japan, in 2013. He received B.B.A with Summa Cum Laude honors from Institute of International Studies of Ramkhamhaeng University, Thailand, in 2006, and M.S.I.T from Sirindhorn International Institute of Technology, Thammast University, Thailand, in 2010. His research interests are natural language processing, intelligent algorithms, machine learning, machine translation, information retrieval, and image processing.

Dense-YOLO: An Improved Weed Detection Platform Based on MSRCP

MingYuan Wang[1] and Watis Leelapatra[2]
[1,2]Department of Computer Engineering, Khon Kaen University, Thailand
E-mail: wangmingyuan@kkumail.com, watis@kku.ac.th

Received: August 24, 2022 / Revised: October 10, 2022 / Accepted: October 18, 2022

***Abstract*—For real-time weed detection needs and the flexibility of deploying models in embedded devices. We proposed an improved object detection platform, named Dense-YOLO which is based on Multi-scale retinex with chromaticity preservation (MSRCP) and YOLOv4 architecture. First, we use MSRCP to preprocess original images to provide a foundation for subsequent feature extraction. Second, Depthwise Separable Convolution (DSC) is used to reduce parameters, making it suitable for development on embedded devices. Third, we used k-means++ to optimize the clustering of anchor size. Fourth, DenseNet-121, PANet, and SPP modules together constitute Dense-YOLO. Last, we analyze the effectiveness of focal loss. Compared with YOLOv4, mAP is improved by 7.26%, three-quarters of the parameters are removed and 6.1 higher in FPS.**

***Index Terms*—Weed Detection, Neural Network, Lightweight Platform**

I. Introduction

Weeds are one of the biggest threats in agriculture, they compete with crops for water, sunshine, and nutrients, significantly affecting crop productivity and quality [1]. The traditional weeding methods are not suitable for modern agriculture, because of time-consuming, labor-intensive, and low efficiency. Mechanical, biological, thermal, electrical, and laser weeding technologies are rapidly attracting attention from scientists and research organizations [2]. Weed suppression is generally carried out at the seedling stage, and object detection is the first step before weeding [3].

In the research of weed detection, the wide applications of traditional computer vision mainly include two steps: feature extraction and pattern recognition [4]. For example, Pulido et al. [5] can improve the classifier's performance to above 90% based on Gray Levelco Occurrence Matrix (GLCM) and Support Vector Machines (SVM) [6]. Although this approach achieves high accuracy, it takes a long time to classify, has poor generalization, and is difficult to apply in practice.

Now, the emergence of neural networks and the popularity of deep learning have provided greater space for the development of object detection algorithms, such as LeNet [7], which was first used for document recognition. In recent years, many SOTA algorithms have been proposed, such as AlexNet [8], VGG [9], GoogLeNet [10], and ResNet [11]. Object detection algorithms have evolved from the traditional VJ Framework (used in face recognition) [12], HOG [13], and DPM (based on SVM) [14] to convolutional neural networks (CNNs), such as two-stage algorithms: SPPNet [15], Faster R-CNN [16] which is improved on Fast R-CNN [17], and one-stage algorithms: SSD [18]-[21], YOLO [22]-[25] that is a kind of end-to-end architecture and is famous as can balance with high accuracy and detection speed. Based on YOLOv3-tiny [24], Sharpe et al. [26] studied the influence of different annotation methods (entire plant annotation and partial annotation of leaf blade) and found the latter can have a better performance in detecting goosegrass. Gao et al. [27] also used YOLOv3-tiny as a foundation and presented an approach to improve the mean average precision (mAP) of convolvulus sepium and sugar beet to 82.9%. From the above research, we can achieve high accuracy through the Graphics Processing Unit (GPU). However, there are many embedded robots and devices that need lightweight algorithms that have low computational complexity and better detection performance during forward inference.

In this paper, Dense-YOLO: a novel lightweight weed detection platform is proposed, and it is based on YOLOv4 [25] architecture to detect garlic sprouts and weeds. During detection, it can trade-off between accuracy, model scale, and speed. Through experiments, we found that multi-scale retinex (MSR) [28] with chromaticity preservation (MSRCP) [29] can be a better choice to preprocess our original images because of its good performance. Meantime, we use k-means++ [30] cluster anchors instead of the original k-means method in our weed dataset can improve detection precision. DSC and Focal Loss [31] also be applied in our new network to be responsible for reducing the whole network's parameters and solving the class imbalance encountered during training, respectively. Our proposed methods and Dense-YOLO have the characteristics of low parameters, high accuracy, small size, and good inference speed, that provide a new way of applications in embedded weeding robots.

II. Related Work

A. MSRCP

Land et al. [32] proposed the Retinex (consists of retina and cortex) theory that explained how the human visual system perceives a scene and divided the visual system model into two parts:

$$I(x,y) = L(x,y) \cdot R(x,y) \quad (1)$$

Where $I(x,y)$ denotes the final images received by eyes, and the low frequency information is denoted by $L(x,y)$ reflects the effect of lightness. $R(x,y)$ denotes scene reflectance and can reflect scene properties using high-frequency information. According to Retinex, Single-Scale Retinex (SSR), MSR, Multi-Scale Retinex with Color Restoration (MSRCR) and MS-RCP were proposed one after another and be applied in image augmentation. While MSRCP can preserve better image chromaticity and color distribution by applying Retinex theory on intensive channels and mapping to each channel according to the original RGB scale. The intensity channel can be computed as follows:

$$I_{nt} = (I_R + I_G + I_B)/3 \quad (2)$$

For those images with a correct color distribution and white lightning, MSRCP can avoid color distortion and restore better images while preserving the original color distribution. This provides an advantage for feature extraction in CNNs.

B. Depthwise Separable Convolution (DSC)

Different with standard convolution, DSC is constituted by depthwise convolution and pointwise convolution. As shown in

Fig. 1., the first step is to process the original $H \times W \times C$ images using $n \times n$ filters that number is same as input images, so the output is $(H-n+1) \times (W-n+1) \times C$ feature maps that will be the input images for the following phase. Second, pointwise convolution is in charge of mixing the information between channels. $1 \times 1 \times m$ filters will be used in this process, while m is the number of needed output feature maps. After both kinds of convolution, the final output feature maps shape is $(H-n+1) \times (W-n+1) \times m$.

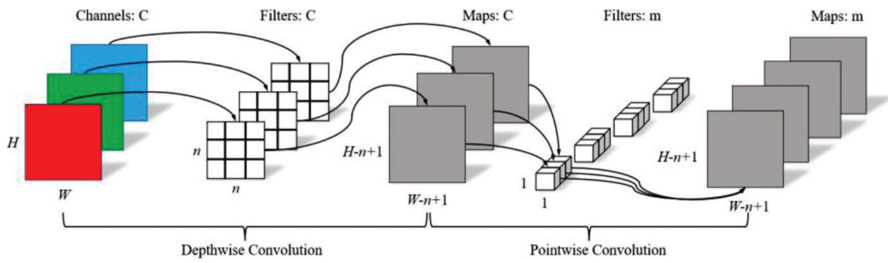


Fig. 1. Depth wise separable convolution process

The computational cost of standard convolution and DSC can be computed by Equations (3)-(4). Where, W_n and W_d denote the parameters of standard convolution and DSC respectively. W and H are the shapes of images, C is the number of channels. The size of the filter is represented by n. As shown in Equation (5), DSC has smaller parameters than standard convolution in the operation of convolution and is more suitable for deep learning networks.

$$W_n = m \times n^2 \times C \times (H-n+1) \times (W-n+1) \quad (3)$$

$$W_d = (n^2 + m) \times C \times (H-n+1) \times (W-n+1)$$

$$e = \frac{W_d}{W_n} = \frac{1}{m} + \frac{1}{n^2} \approx \frac{1}{n^2}, (m \gg n) \quad (4)$$

C. DenseNet

Dense convolutional network (DenseNet) [33] was proposed by Huang et al. in 2017 and is used in deeper convolution networks. The core point of DenseNet is to connect every layer to each other layer in a feed-forward fashion. As a traditional neural network that has L connections, there is $L \times (L+1)/2$ direct connections between each layer and its subsequent layer in DenseNet. The feature maps of all preceding layers are used as inputs for each layer, and its own feature maps are used as inputs for all following layers. DenseNet not only can avoid vanishing gradient but also can reduce the scale of the network, and reuse and combine feature maps.

As illustrated in Fig. 2 each Dense Block contains enormous nodes that consist of normalization, activation functions, and convolutional layers, and are connected to the nodes before it. To increase inference efficiency, 1 × 1 convolution is used as a bottleneck layer to reduce the number of input feature maps.

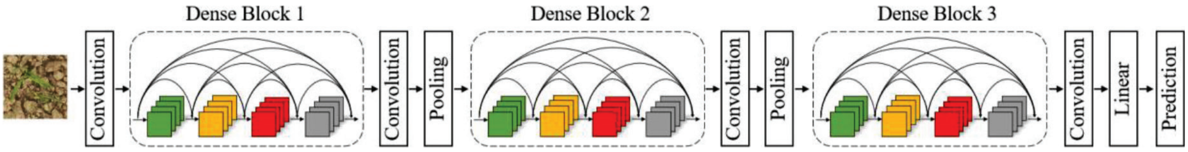


Fig. 2. DenseNet model structure

D. YOLOv4

From 2016-2018, Redmon et al. successively introduced YOLOv1 [22], YOLOv2 [23] and YOLOv3 [24] and realized object detection and classification prediction in a single network. This approach can detect objects rapidly while maintaining sufficient accuracy, which is suitable for real-world applications. Bochkovskiy et al. proposed YOLOv4 [25] in 2020, based on YOLOv3 and conducted some improvements to enhance feature extraction. Cross Stage Partial Darknet53 (CSPDarknet53), Spatial Pyramid Pooling (SPP), Path Aggregation Network (PANet), and YOLO Head comprise the YOLOv4 structure. Unlike YOLOv3, CSPDarknet53 is used instead of Darknet53, Mish activation replaces the LeakyReLu activation to improve the stability of gradient propagation. Three different size pooling layers in SPP can increase receptive field and fuse more feature maps. In PANet, multi-scale feature fusion occurred frequently to extract further features with various scales. After YOLO Head, we can get the anchor adjustive sizes, object location, classification confidence score and Intersection over Union (IoU) score. Finally, Non-Maximum Suppression (NMS) is used to filter badly bounding boxes. Equation (6) is to computed confidence score which reflect the credibility of predicted object.

$$C_i^j = P_r(Object) \times IoU_{pred}^{truth} \quad (5)$$
$$P_r(Object) \in [0, 1]$$

Where, C_i^j denotes the confidence score of the jth bounding box in the ith grid. $P_r(Object)$ reflects the probability that one bounding box contains the target object.

from the advantages of YOLOv4 that can trade off detection accuracy and speed. Compared with the original YOLOv4, there are many improvement architectures proposed. For detecting apple flowers, based on YOLOv4, Wu et al. [34] used the channel pruning algorithm to achieve a huge reduction in model size, inference time, and network parameters. Dong et al. [35] focused on some attention methods to increase weed detection accuracy, such as shuffle attention and transformer module. Multi-scale fusion and pooling methods are also applied by Che et al. [36] to improve the dishes detection performance. According to different fields, YOLOv4 is improved and applied in their own research which is an inspiration for this paper.

III. Methodology

Fig. 3 is the proposed framework which is designed to achieve better performance in weed detection and has lower parameters in the whole network. Four techniques will be introduced in this framework: MSRCP is used as an enhancement method for original images. K-means++ can optimize the initial anchor size. Focal Loss aims to address the common issue of class imbalance. Dense-YOLO is an improved one-stage object detection based on DenseNet and YOLOv4. Their detail will be presented in the following sections.

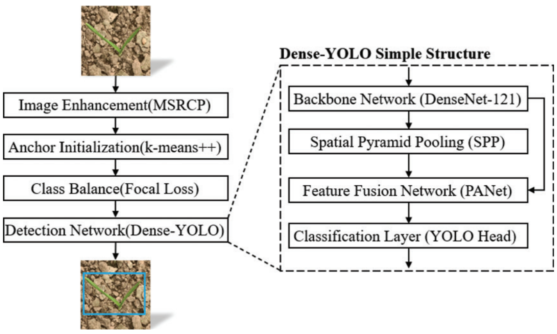


Fig. 3. The diagram of proposed framework

A. Image Enhancement

In the real world, those captured images have different lightness, dark regions, and poor contrast because of the effect of sunlight or the change of weather. Indeed, the result of applying data augmentation can restore a clearer and better image from an original image with dark regions or poor quality in other areas. In this paper, we tested different methods to find out the best one to support our next experiments, including Contrast Limited Adaptive Histogram Equalization (CLAHE) [37], homomorphic filtering based on Hue-Saturation-Intensity (HSI) [38], and MSRCP. CLAHE can reduce the noise problem and poor contrast by clipping several pixels that exceed the contrast factor to redistribute the whole histogram. According to the HSI of the original color image, the homomorphic filtering approach can not only improve the unequal distribution issue but also maintain the

RGB information. To preserve the original color distribution, MSRCP applies Retinex to the intensity channel without color restoration and maps the data to each channel using the original RGB scale. As shown in Fig. 4 comparative images were carried out in this paper to intuitively feel the impact of each algorithm on the original image. The contrast and color distribution after HFHSI processing have greatly improved. The lightness has increased using CLAHE, however, the contrast is relatively low. As for MSRCP, the leaves of the image are clearly distinguished from the background. Although the image is unnatural, the details are clearer, such as the classification texture, and the whole lightness is improved greatly, the saturation is better, and the contrast is enhanced.

As shown in Table I, we also conducted comparative tests based on YOLOv4 to find out the optimal image enhancement method that is suitable for the CNNs. MSRCP achieves a good performance both in vegetable and weed AP value and mAP is higher by 2.75% than the original image. So, we will use MSRCP as our image enhancement method for further experiments.

(a) Origin (b) HFHSI (c) CLAHE (d) MSRCP

Fig. 4. The comparison of different image enhancement algorithm

TABLE I
COMPARISON OF DIFFERENT IMAGE ENHANCEMENT ALGORITHMS

Method	AP/%		Precision/%	Recall/%	mAP/%
	Vegetable	Weed			
Origin	90.70	79.81	92.87	82.64	85.26
HFHSI	91.13	76.03	92.22	83.14	83.58
CLAHE	90.48	75.90	94.08	82.30	83.19
MSRCP	93.07	82.95	92.43	87.65	**88.01**

B. Dimension Cluster

Instead of clustering anchors by standard k-means that are used in the original YOLOv4, we run k-means++ clustering on the training set bounding boxes to automatically find good anchors and avoid the shortcoming caused by random initial cluster center. Based on k-means++, a specific way of choosing anchors is used in this paper and we detail the whole process:

1) Prepare our bounding boxes dataset (all the class) and the number of needed centers k=9.

2) Take one random center from dataset N as the initial center O_1.

3) Take a new center O_x from dataset N, choosing x with $P(x)$.

$$D(x) = 1 - IoU(x, O_x) \tag{6}$$

$$P(x) = \frac{D(x)^2}{\sum_{x \in N} D(x)^2} \tag{7}$$

Where, $IoU(x, O_x)$ denotes IoU between data x and center O_x. $D(x)$ is the distance of data x with its adjacent center O_x. $P(x)$ is the probability of being the next center.

4) Repeat step 3, until we have taken 9 centers altogether.

5) Iterate compute $D(x)$, $x \in N$, classify $x \in D(x)_{\min}$ to the new cluster center C_i that can get by Equation (9).

$$C_i = \frac{\sum_{x \in C_i} x}{|C_i|} \tag{8}$$

6) When the value of C_i is stable, the output is the optimal cluster results.

We implemented the above algorithm and conducted combined experiments on our bounding boxes dataset and get the best performance anchors group: (130, 124), (83, 236), (195, 174), (305, 197), (259, 266), (308, 322), (570, 814), (1764, 449), and (2370, 1268).

C. Dense-YOLO

Based on YOLOv4, we proposed a novel end-to-end network: Dense-YOLO. Instead of using CSPDark53 as the backbone, we introduced DenseNet-121 as the backbone to rich the feature extraction and stable gradient propagation, as shown in Fig. 5. Through SPP structure, maxpooling layers of different sizes can extract more features from backbone network and increase receptive field. The purpose of the transition section which consists of 1×1 convolution and average pool is to change the number of channels and condense feature maps.

After transition section, we need to bridge the backbone and PANet, feature maps with different shapes are divided to three parts: 52×52×256, 26×26×512, and 13×13×1024. As like the original YOLOv4, those feature maps are responsible for detecting different size objects respectively. To lightweight our network, we replace all the 3×3 standard convolutions with DSC, Finally the number of parameters is decreased to one-quarter of the original YOLOv4. This strategy increases the feasibility of the object detection method being deployed on embedded robots or devices. In the improved backbone, like DenseNet-121, four Dense Blocks are used to improve the learning ability of feature extraction. The number of Dense units in each Dense Block is 6, 12, 24, and 16 respectively. ResNets and DenseNet convolutional feed-forward networks layer transition are shown by Equations (10) and (11). X_l is the output of l^{th} layer. $H(\cdot)$ denotes the function of operations such as convolution, Batch Normalization (BN) [39], rectified linear activation (ReLU) [40], or pooling. Although they all belong to residual networks, DenseNet combines the feature maps of all preceding layers and can reuse features throughout the networks to obtain a more accurate model.

$$X_l = H_l(X_{l-1}) + X_{l-1} \tag{9}$$

$$X_l = H_l(X_0, X_1, \cdots, X_{l-1}) \tag{10}$$

In the part of feature fusion, Dense-YOLO method keeps the PANet structure to extract feature maps with different scales to increase object detection speed, as the same YOLOv4 method. In the end, three head parts are used to predict classes and bounding boxes of weeds and vegetables.

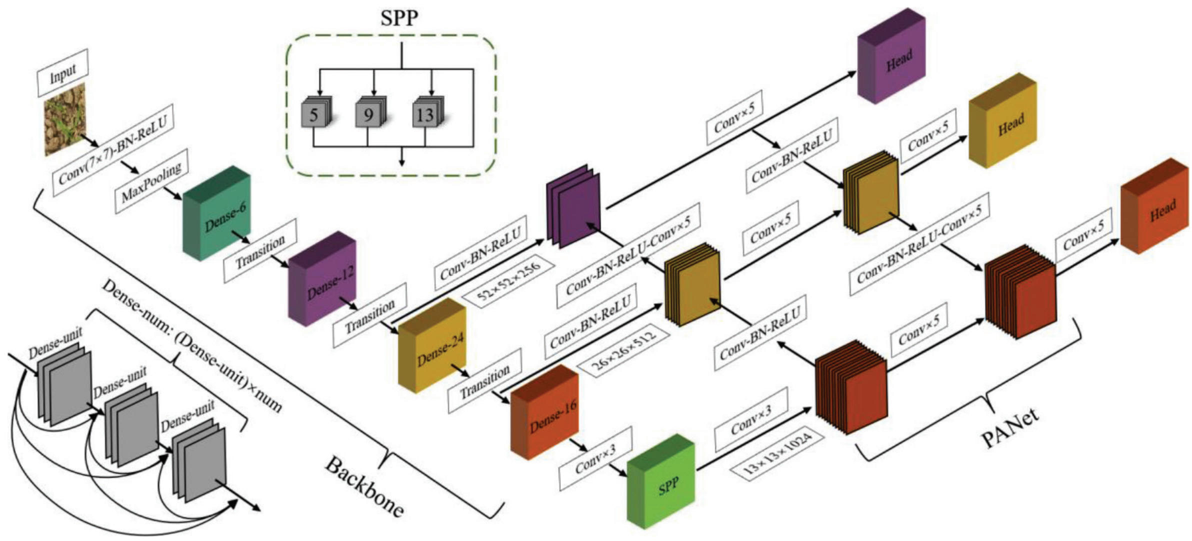


Fig. 5. Dense-YOLO model structure

D. Class Balance

Actually, class imbalance is a common phenomenon, and there are a lot of vegetables and small weeds in the real dataset. Using standard cross entropy as a loss function may cause iteration to slow or deviate from the correct direction when training on the dataset with foreground-background class imbalance. So, we introduce Focal Loss in our model to make the network focus on hard, misclassified examples during training. Based on standard cross entropy, Focal Loss introduced a weighting factor $\alpha \in [0, 1]$ to address class imbalance as shown in Equation (12).

$$Loss_{Focal} = -\alpha(-P(y|x))^{\gamma} \log(P(y|x)) \tag{11}$$

Where α for the positive class and $1-\alpha$ for the negative class. $\gamma \geq 0$ is focusing parameter, it can adjust the loss speed of easy examples and address the network to focus on the hard negatives. $P(y|x) \in [0, 1]$ denotes the estimated probability for the example x with label y in our model. We implement the Focal Loss method to Dense-YOLO to limit the dominant role of easy examples, avoid learning useless features, and improve training efficiency.

IV. Experiments

A. Dataset

A large-scale dataset is necessary for supervised learning and deep learning. That's why we established our own dataset for weed detection through capturing images from farmland at a height of 40 cm. We capture images during stem elongation stage because there are more features that can be distinguished, and occlusion has less impact on each object. Our dataset contains 1800 original images and two categories: vegetable and weed. In the real world, our dataset has following properties: (a) occlusion with each other, (b) incomplete objects on the edge of image, (c) other objects coexist with our target object, and (d) dark regions, as shown in Fig. 6.

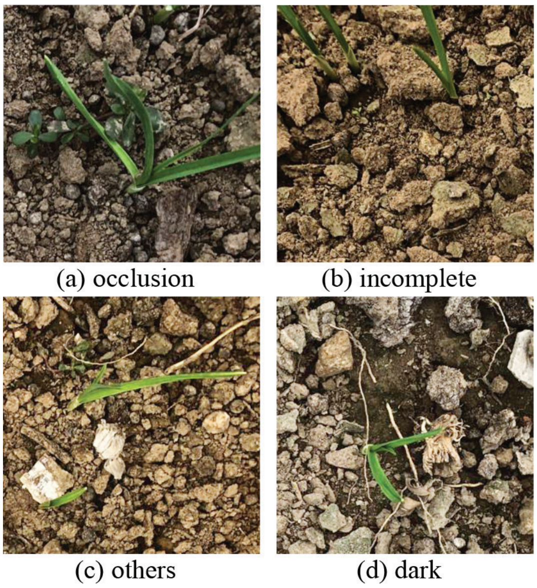


Fig. 6. Example images of vegetable and weed

Data augmentation is widely used in State-of-the-Art (SOTA) detection algorithms and can boost a network's performance and robustness. We add rotation, flipping, scaling, blurring, and cropping into augmentation strategies to expand image number and prevent overfitting. As shown in Table II, the original images are expanded to 5000, and 90% of them are used as training and validation, and 10% are to test performance.

TABLE II
Statistical Data of Augmentation

Origin	Augmented	Train	Validation	Test
1800	5000	4050	450	500

B. Evaluation

Popular measures used to evaluate detection performance include Precision (P), Recall (R), Average Precision (AP), mean mAP, Frame Per Second (FPS), and the number of parameters in the whole network. Using Equation (13) to calculate the P and Equation (14) to get R which is the proportion of predicted items to the total number of true items. The P-R curve can be obtained by P and R, and the area enclosed by the P-R curve with the x and y axes is AP. mAP is the value by summing the AP of each category and then averaging.

$$P = \frac{TP}{TP + FP} \tag{12}$$

$$R = \frac{TP}{TP + FN} \tag{13}$$

$$mAP = \frac{\sum_{i=1}^{k} AP_i}{k} \tag{14}$$

Where *TP* denotes the positive examples are correctly identified as positive samples. *FP* is the negative examples that are incorrectly identified as positive samples. *FN* indexes the positive samples are incorrectly identified as negative samples. *k* is the number of detection classes.

C. Clustering Analysis

Dense-YOLO is an anchor-based detector, anchors with suitable size are the crucial factor for yielding good result performance. The size distribution of bounding boxes differs from other common datasets, so it is necessary to cluster before training. Instead of k-means method, we use k-means++ to conduct clustering analysis in each class. We set k as 9 because Small (S), Medium (M), and Large (L) are three groups for detecting different-sized objects and each group has three pairs of data. Table III presents the results. Moreover, to study the effects of different group combinations, mixed experiments are conducted to find a better anchor setting. Table IV displays group settings and comparison results. In total, there are 9 groups from A to I.

Table V is our experimental setup: all training and tests are conducted with 3060Ti GPU and use this GPU to execute multi-scale training in the batch size of 64 while mini-batch size is 8 or 4 depending on GPU memory limitation. This work uses Python 3.6 and Pytorch1.10 as the programming language and deep learning platform. During training, the Adam method is the optimizer to implement gradient backpropagation. The total epoch, momentum, learning rate, and decay are 200, 0.937, 0.001, and 0.0005 for all methods, respectively.

TABLE III
EXPERIMENTAL RESULTS OF K-MEANS++

	Weed			Vegetable		
	S	**M**	**L**	**S**	**M**	**L**
Clustering by Category	130,124	305,197	271,457	150,142	506,234	570,814
	83,236	259,266	546,240	287,197	534,466	1764,449
	195,174	308,322	419,348	252,420	895,389	2370,1268
Together Clustering	(148,140) (280,208) (259,421) (498,238) (534,468) (895,389) (570,814) (1707, 421) (2265,1256)					

TABLE IV
MIXED COMPARISON EXPERIMENTS GROUPS SETTING AND DETECTION RESULTS

Combination Setting							**Experimental Results**				
	Weed			**Vegetable**			**Weed**	**Vegetable**	**Precision/%**	**Recall/%**	**mAP/%**
Group	**S**	**M**	**L**	**S**	**M**	**L**	**AP/%**				
A	**Together Clustering Result**						78.86	93.07	93.66	88.81	85.97
B	√	-	-	-	√	√	80.00	92.70	93.83	88.81	86.35
C	-	√	-	√	-	√	81.76	93.82	93.90	89.98	87.76
D	-	-	√	√	√	-	84.10	94.08	93.15	90.82	89.09
E	√	√	-	-	-	√	85.41	93.72	94.07	89.98	**89.56**
F	-	√	√	√	-	-	84.69	92.70	91.79	89.65	88.69
G	√	-	√	-	√	-	79.38	92.70	93.51	88.98	86.04
H	√	√	√	-	-	-	79.35	93.40	94.37	89.48	86.37
I	-	-	-	√	√	√	84.63	93.91	93.16	90.15	89.27

TABLE V
THE HARDWARE AND DETAILS OF MODEL TRAINING EXPERIMENTAL SETUP

Setup	**GPU**	**Programming Language**	**Platform**	**Optimizer**	**Epoch**	**Momentum**	**Learning Rate**	**Decay**
Details	3060Ti	Python 3.6	Pytorch 1.10	Adam	200	0.937	0.001	0.0005

Based on YOLOv4, we choose MSRCP algorithm to process original images before feeding into the neural network. From Table IV, different anchors combination does affect the result and group E achieves the best performance with 89.56% mAP. Compared with the original k-means method, The relative mAP has improved by 1.55%. The reason is the suitable combination makes it easy for detector to adjust bounding boxes and finish regression. Next, we will use the anchor size of group E to continue experiments.

D. Lightweight Network

To reduce the total parameters and increase interference speed during detection, Howard et al. proposed MobileNetV1 [41], MobileNetV2 [42] and MobileNetV3 [43]. Because DSC was introduced into their network to reduce parameters, while trade-off between accuracy, model size, and detection speed. In this paper, we also use DSC to replace the 3×3 standard convolution to lightweight our network and make it possible for applying to embedded devices. Additionally, we used DSC in YOLOv4 rather than traditional convolution to test the efficacy of DSC.

TABLE VI
COMPARISON EXPERIMENTS

Convolution Method	AP/%		Precision/%	Recall/%	mAP/%	Time/ms	FPS	Model Size/MB	Parameter
	Weed	Vegetable							
Standard Convolution	85.41	93.72	94.07	89.98	89.56	56.9	17.6	244	64,009,375
DSC	75.67	91.05	92.76	87.65	83.36	44.7	22.4	137	35,757,727

TABLE VII
COMPARISON OF DIFFERENT BACKBONES

Backbone	AP/%		Precision/%	Recall/%	mAP/%	Time/ms	FPS	Model Size/ MB	Parameter
	Weed	Vegetable							
CSPDarkNet53	75.67	91.05	92.76	87.65	83.36	44.70	22.40	137.00	35,757,727
MobileNetV1	75.46	94.09	94.17	88.98	84.77	19.90	50.30	51.10	12,271,999
MobileNetV2	74.12	91.45	93.31	88.48	82.78	22.20	45.00	46.50	10,381,119
MobileNetV3	73.10	88.49	90.78	85.48	80.80	24.10	41.50	53.70	11,309,039
GhostNet	74.78	92.24	92.91	87.48	83.51	31.70	31.50	42.50	11,008,515
ResNet-50	83.81	94.43	92.84	90.98	89.12	36.80	27.20	127.00	33,261,183
VGG-16	86.15	95.80	94.23	92.65	90.98	63.00	15.90	89.90	23,517,567
DenseNet-121	84.67	95.06	93.52	91.49	89.86	41.50	24.10	61.90	16,018,879

As shown in Table VI, DSC reduces half of the parameters and computational cost in the whole network, but the mAP also has a big drop at the same time. Don't worry, this will be improved through our proposed Dense-YOLO which is also based on DSC.

E. Backbone Improvement

To slim our network and keep high accuracy, we evaluate and compare the performances of different architectures as feature backbone. Besides MobileNets, GhostNet [44], ResNet-50, VGG-16 and DenseNet-121 also participated in the comparison experiments. There are three major observations from Table VII:

1) The situation is changed with different backbone-we can further reduce the total parameters by combining different feature extraction backbone. This indicates that the model size problem is well addressed in this operation, and we manage to obtain more accuracy gains from modified backbone.

2) We also note that the number of parameters is not the decisive factor for precision, recall and mAP. A new network that trade-off accuracy and model size can be obtained from a suitable backbone. More importantly, the detection speed or FPS increased along with the reduction of parameters, this is significant for real-time models.

3) Compared to the original YOLOv4 backbone (CSPDarkNet53), VGG-16 achieves the top-1 mAP 90.98%, however other measures are not ideal. While DenseNet-121 not only performs better with 89.86 mAP, but also has a lower parameter and faster detection speed. Because of this, we use DenseNet-121 as the backbone of Dense-YOLO, so that the network parameters can be dramatically reduced to a quarter of the YOLOv4, while the FPS is increased by 6.5, and the mAP doesn't drop down.

F. Focal Loss Test

As we discuss in class balance section, here we introduce Focal Loss into Dense-YOLO to see its effect during training. Table VIII is the results before and after using Focal Loss. This comparison successfully verifies the effectiveness of Focal Loss on imbalance dataset, we achieve 92.52% of mAP, a 2.66% improvement.

TABLE VIII
COMPARISON OF USING FOCAL LOSS AND NOT

	AP/%		Precision	Recall	mAP
	Weed	Vegetable			
Traditional	84.67	95.06	93.52	91.49	89.86
Focal Loss	96.05	88.99	96.10	90.48	**92.52**

G. Comparison with SOTA

To demonstrate the superiority of Dense-YOLO, comparisons are made with other state-of-the-arts. We compared detection algorithms with different structures such as SSD, CenterNet [45], EfficientDet [46], YOLOv4-Tiny [47] and RetinaNet since we want to design another excellent structure. As shown in Table IX. Tests show that Dense-YOLO can substantially reduce model size and increase speed while maintaining the stability of mAP compared to original YOLOv4. YOLOv4-Tiny and Efficient Det-D0 can minimize model size and improve FPS, unfortunately, they also cause a huge decrease in mAP. Surprisingly, Dense-YOLO is 26.3% higher in mAP than YOLOv4-Tiny, and compared to original YOLOv4, it is not only 7.26% higher in mAP, but also three-quarters of the parameters are removed and 6.1 higher in FPS.

Why this proposed model can get the best result? The reasons can be attributed to the techniques we discussed above. MSRCP improves the quality and feature distribution compared to the original images. K-means++ makes the initial anchor size closer to most of the target objects. Class imbalance can be weakened through Focus Loss, thus the mAP has a further increment. Meanwhile, DenseNet-121 can deepen layers of the backbone network to fit the more complex model. While the use of DSC reduces the compute parameters and model size.

TABLE IX
COMPARISON WITH OTHER STATE-OF-THE-ARTS

Model	AP/%		mAP/%	Time/ms	FPS	Model Size/MB	Parameter
	Weed	Vegetable					
YOLOv4	90.70	79.81	85.26	54.30	18.40	244.00	64,009,375
YOLOv4-Tiny	52.26	80.19	66.22	8.40	118.50	22.50	5,882,634
SSD	58.73	84.44	71.59	26.10	38.40	91.20	23,745,908
Center Net	69.90	89.16	79.53	32.60	30.70	125.00	32,718,597
Efficient Det-D0	75.18	92.51	83.85	36.40	27.50	16.10	3,875,723
Retina Net	61.23	82.78	72.00	86.40	11.60	138.00	36,350,582
Dense-YOLO	96.05	88.99	**92.52**	40.70	24.50	61.90	16,018,879

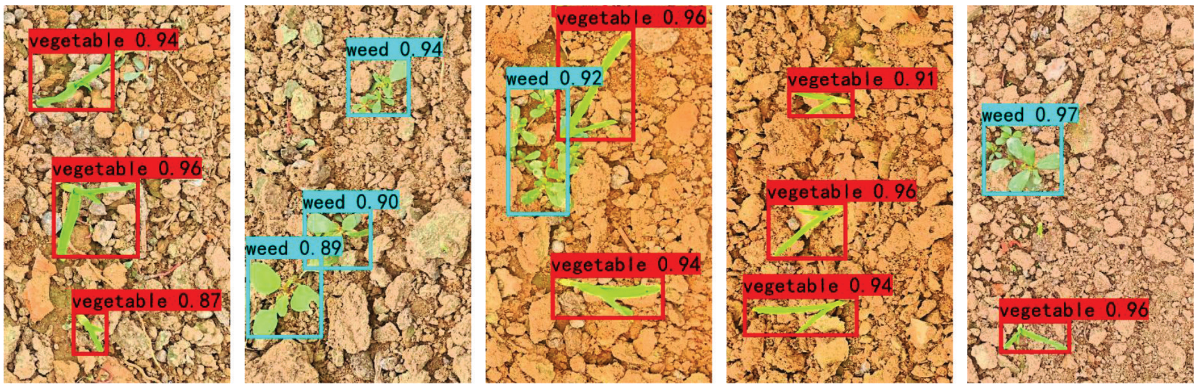


Fig. 7. Detection results from Dense-YOLO

V. CONCLUSION

For the needs of weeding robots or other embedded devices, this paper proposed a compacted, lightweight and speed detection method in terms of network structure, we term it Dense-YOLO. MSRCP algorithm is introduced to preprocess images to increase color contrast and remove dark regions. To lightweight the whole network, we replaced all 3×3 standard convolution with DSC and achieved a noticeable effect in the parameter. We used k-means++ algorithm to optimize the clustering of anchor size, instead of the traditional method. After extensive comparison analysis, we used DenseNet-121, PANet and SPP modules together to constitute the network of Dense-YOLO, a plug-and-play model for smart agriculture and is a portable neural architecture to promote weed detection in both flexibility and accuracy. In the end, we also demonstrated the effectiveness of Focal Loss in solving the class imbalance problem.

We are pleased to present our encouraging results and will continue to focus on computer vision and neural network research in future work, such as the network pruning to further slim our model.

REFERENCES

[1] D. D. Patel and B. A. Kumbhar, “Weed and its Management: A Major Threats to Crop Economy,” *J. Pharm. Sci. Biosci. Res*, vol. 6, pp. 453-758, 2016.

[2] L. Wen, X. Liming, and X. Jiejie, “Research Status of Mechanical Intra-Row Weed Control in Row Crops,” *Journal of Agricultural Mechanization Research*, vol. 39, pp. 243-250, 2017.

[3] B. Liu and R. Bruch, “Weed Detection for Selective Spraying: A Review,” *Current Robotics Reports*, vol. 1, pp. 19-26, Mar. 2020.

[4] S. Shanmugam, E. Assunção, R. Mesquita et al., “Automated Weed Detection Systems: A Review,” *KnE Engineering*, pp. 271-284, Jun. 2020.

[5] C. Pulido, L. Solaque, and N. Velasco, “Weed Recognition by SVM Texture Feature Classification in Outdoor Vegetable Crop Images,” *Ingeniería e Investigación*, vol. 37, pp. 68-74, Apr. 2017.

[6] S. R. Gunn, "Support Vector Machines for Classification and Regression," *ISIS Technical Report*, vol. 14, pp. 5-16, May. 1998.

[7] Y. Lecun, L. Bottou, Y. Bengio et al., "Gradient-Based Learning Applied to Document Recognition," in *Proc. IEEE*, pp. 2278-2324, 1998.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, pp. 84-90, May. 2017.

[9] K. Simonyan and A. Zisserman. (2014, Sep. 4). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. [Online]. Available: https://arxiv.org/abs/1409.1556

[10] C. Szegedy, W. Liu, Y. Jia et al., "Going Deeper with Convolutions," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1-9.

[11] K. He, X. Zhang, S. Ren et al., "Deep Residual Learning for Image Recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770-778.

[12] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, pp. 137-154, May. 2004.

[13] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886-893.

[14] P. F. Felzenszwalb, R. B. Girshick, D. McAllester et al., "Object Detection with Discriminatively Trained Part-Based Models," *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1627-1645, Sep. 2010.

[15] K. He, X. Zhang, S. Ren et al., "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, pp. 1904-1916, Jan. 2015.

[16] S. Ren, K. He, R. Girshick et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Advances in Neural Information Processing Systems*, Jun. 2015.

[17] R. Girshick, "Fast R-Cnn," in *Proc. IEEE International Conference on Computer Vision*, 2015, pp. 1440-1448.

[18] C. Fu, W. Liu, A. Ranga et al. (2017, Jan. 23). *DSSD: Deconvolutional Single Shot Detector*. [Online]. Available: https://arxiv.org/abs/1701.06659

[19] Z. Li and F. Zhou. (2017, Dec. 4). *FSSD: Feature Fusion Single Shot Multibox Detector*. [Online]. Available: https://arxiv.org/abs/1712.00960

[20] J. Jeong, H. Park, and N. Kwak. (2017, May. 26). *Enhancement of SSD by Concatenating Feature Maps for Object Detection*. [Online]. Available: https://arxiv.org/abs/1705.09587

[21] W. Liu, D. Anguelov, D. Erhan et al., "SSD: Single Shot Multibox Detector," *European Conference on Computer Vision*, pp. 21-37, Otc. 2016.

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

[23] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7263-7271.

[24] J. Redmon and A. Farhadi. (2018, Apr. 6). *YOLOv3: An Incremental Improvement*. [Online]. Available: https://arxiv.org/abs/1804.02767

[25] A. Bochkovskiy, C. Wang, and H. M. Liao. (2020, Apr. 23). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. [Online]. Available: https://arxiv.org/abs/2004.10934

[26] S. M. Sharpe, A. W. Schumann, and N. S. Boyd, "Goosegrass Detection in Strawberry and Tomato Using a Convolutional Neural Network," *Scientific Reports*, vol. 10, pp. 1-8, Jun. 2020.

[27] J. Gao, A. P. French, M. P. Pound et al., "Deep Convolutional Neural Networks for Image-Based Convolvulus Sepium Detection in Sugar Beet Fields," *Plant Methods*, vol. 16, pp. 1-12, Dec. 2020.

[28] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-Scale Retinex for color image enhancement," in *Proc. 3rd IEEE International Conference on Image Processing*, 1996, pp. 1003-1006.

[29] A. B. Petro, C. Sbert, and J. Morel, "Multiscale Retinex," *Image Processing on Line*, vol. 4, pp. 71-88, Apr. 2014.

[30] D. Arthur and S. Vassilvitskii, "k-means++: The Advantages of Careful Seeding," in *Proc. Symp. Discrete Algorithms*, 2007, pp. 1027-1035.

[31] T. Y. Lin, P. Goyal, R. Girshick et al., "Focal Loss for Dense Object Detection," in *Proc. IEEE International Conference on Computer Vision*, 2017, pp. 318-327.

[32] E. H. Land, and J. J. McCann, "Lightness and Retinex Theory," *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1-11, Jan. 1971.

[33] G. Huang, Z. Liu, L. Van Der Maaten et al., "Densely Connected Convolutional Networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700-4708.

[34] D. Wu, S. Lv, M. Jiang et al., "Using Channel Pruning-Based YOLO v4 Deep Learning Algorithm for the Real-Time and Accurate Detection of Apple Flowers in Natural Environments," *Computers and Electronics in Agriculture*, vol. 178, p. 105742, Nov. 2020.

[35] H. Dong, X. Chen, H. Sun et al., "Weed Detection in Vegetable Field Based on Improved YOLOv4 and Image Processing," *Journal of Graphics*, vol. 43, pp. 559-569, Mar. 2022.

[36] X. Che and H. Chen, "Muti-Object Dishes Detection Algorithm Based on Improved YOLOv4," *Journal of Jilin University*, pp. 1-7, Nov. 2021.

[37] K. Zuiderveld, "Contrast Limited Adaptive Histogram Equalization," *Graphics Gems*, pp. 474-485, May. 1994.

[38] Y. Zhang and M. Xie, "Colour Image Enhancement Algorithm Based on HSI and Local Homomorphic Filtering," *Computer Applications and Software*, vol. 30, pp. 303-307, Dec. 2013.

[39] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *Proc. International Conference on Machine Learning*, 2015, pp. 448-456.

[40] X. Glorot, A. Bordes, and Y. Bengio, "Deep Sparse Rectifier Neural Networks," in *Proc. the Fourteenth International Conference on Artificial Intelligence and Statistics*, 2011, pp. 315-323.

[41] A. G. Howard, M. Zhu, B. Chen et al. (2017, Apr. 17). *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. [Online]. Available: https://arxiv.org/abs/1704.04861

[42] M. Sandler, A. Howard, M. Zhu et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510-4520.

[43] A. Howard, M. Sandler, G. Chu, et al., "Searching for MobileNetV3," in *Proc. IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1314-1324.

[44] K. Han, Y. Wang, Q. Tian et al., "GhostNet: More Features from Cheap Operations," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1580-1589.

[45] X. Zhou, D. Wang, and P. Krähenbühl. (2019, Apr. 16). *Objects as Points*. [Online]. Available: https://arxiv.org/abs/1904.07850

[46] M. Tan, R. Pang and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 10781-10790.

[47] Z. Jiang, L. Zhao, S. Li et al., (2020, Nov. 9). *Real-Time Object Detection Method Based on Improved YOLOv4-Tiny*. [Online]. Available: https://arxiv.org/abs/2011.04244

MingYuan Wang received the B.Eng. degree in Mechanical Design, Manufacturing and Automation from Henan University of Science and Technology, Henan, China, in 2013. Past work experience on acting product designer at YTO Group Corporation from 2014 to 2020. He current is a graduate student at Department of Computer Engineering, Khon Kaen University, Thailand. His current research interests include computer vision and its applications to other science and engineering fields.

Watis Leelapatra received the B.Eng in Computer Engineering from Khon Kaen University, Khon Kaen, Thailand, and the M.S. degree in Computer Science from the Case Western Reserve University, Cleveland, USA, and the D.Eng. Degree in Computer Science from Asian Institute of Technology, respectively. He current is a lecturer at the Department of Computer Engineering, Khon Kaen University.

Defect Reduction in Crispy Coconut Rolls Production Process

Ploypailin Phrikthim[1] and Paitoon Siri-O-Ran[2]

[1,2]Faculty of Engineering and Technology, Panyapiwat Institute of Management, Nonthaburi, Thailand.
E-mail: ploypailinphr@pim.ac.th, paitoonsir@pim.ac.th

Received: April 30, 2022 / Revised: August 1, 2022 / Accepted: September 26, 2022

***Abstract*—The objective of this research is to reduce the defect in the production process of crispy coconut rolls. By using the quality control tool (QC Tool) to find the cause and improve the quality of the production process. This research has used the Check Sheet to find the point outside the control line by using the waste control chart (P-Chart) and the Pareto diagram to distinguish the importance of sequential problems with Pareto's 80: 20 laws for selecting the most defects in the out of shape. Will focus on this one type of defect and apply this waste reduction in crispy coconut rolls. The problem was analyzed with a fishbone diagram to set up measures to solve the problem.**

The improvement result was able to reduce waste caused by the crispy coconut rolls production process from the previous loss of 5,028 kg accounting for 8.6%, decreased to 2,949 kg accounting for 3.4%, a decrease from June to November 2021 can reduce waste due to amorphous piece of works by 2,079 kg, representing a percentage of waste that can be reduced by 58.65%, representing an annual loss of 415,734 baht per year.

***Index Terms*—Waste, 7 QC Tools, Efficiency**

I. Introduction

A. Statement of The Problems

In Thailand, the food processing industry is now regarded as a significant manufacturing sector. Food production is expected to expand by 4.5% in 2021, while exports are expected to increase by 7.1%, totaling 1.05 trillion baht (National Food Institute). Due to higher competition and increased production costs. This forces manufacturers to develop production processes in order to compete on quality control and cost reduction.

This is a food processing industry case study. There is a problem with quality control. The majority of the waste in the process and all waste data from the crispy coconut rolls machinery was collected for three months from June to August because of the forming process of crispy coconut rolls. From Table I the data show total production volume of the production line was found to be 412,195 kg.

During the forming process of crispy coconut rolls. The machine that produced the most waste was the third machine, which produced a total of 8,834 kg of waste.

The study shows that the crispy coconut roll forming process generates the majority of the waste in the production process, so the researcher believes that reducing waste in the forming process will increase profits and lower costs. The amount of waste generated in each machine during the crispy coconut rolls production process from June to August 2021 is shown in Table I.

TABLE I
THE AMOUNT OF WASTE GENERATED IN EACH MACHINE

Machine	Total Defect (kg)	Total Finished Goods (kg)	Total (kg)	% Defect
M/C1	6,547	76,576	83,123	22
M/C2	4,290	79,758	84,048	14
M/C3	8,834	63,448	72,282	30
M/C4	4,453	84,496	88,949	15
M/C5	5,640	78,154	83,794	19
Total	**29,764**	**382,432**	**412,195**	**100**

Using statistical quality control concepts to determine the problem's core cause and systematically eradicate the problem's primary cause quality control tools have been utilized in several studies to reduce the amount of waste generated throughout the manufacturing process as well as increase efficiency. As a result of the research study, reducing waste from the crispy coconut rolls manufacturing process can be done by reducing out-of-form workpieces can assist in lowering production costs the original loss of 5,028 kg of product value was reduced to 2,949 kg as a result of this approach. If the reduced cost is 103,934 baht per three months, the marketing value is 415,734 baht per year.

B. Scope of Study

• This research focuses on the study data of a case study company.

• This research uses a chart to control the proportion of defects in the study of quality data.

• Only investigate the No. 3 crispy coconut rolls-making machine. Timeframe for research: June to November 2021.

C. Objectives

This research aims to study the quality of production and reduce waste in the crispy coconut rolls manufacturing process after applying knowledge management processes in a case study in the food industry.

II. Literature Review

A. Quality Control

Quality control is first known in the industrial business. It refers to management in order to control raw materials and production to prevent mistakes and damage to final products. After that, quality control is adopted with statistical data, and statistical principles and methods are applied to quality control. Therefore, quality control is insensitive in many industries [1].

William Edwards Deming was the first one to introduce concepts of quality controls to Japan and apply Statistics Quality Control (SQC). SQC mean to use theories and statistical method in every step of production in order to produce and save cost at the same time. Deming proposed deming's 14 points and insisted that his quality control theory can be applied well the both in service and industrial sectors.

B. Theories of Control Chart

Control Chart is a tool developed for production process control. In the Production process. There is variation caused by 4 main factors:

• Material
• Machinery
• Method of work
• Man-Made error

To find cause-and-effect Analysis Tools: A fishbone diagram is a cause-and-effect discovery tool that helps visualization for categorizing the potential causes of a problem.

Element of Control Chart

One of the most effective tools that are widely used in the control phase is the control chart. This tool is effective in monitoring the improvement levels and process there are many types of such charts and all depend on the overall data that have been collected or have to be collected. When applying the statistical approach, the control chart can be different for binominal distribution and poisson distribution for binominal distribution, the collected data is based on two possible outcomes: the existence of the defect and the non-existence of the defect. However, the poison distribution has multiple outcomes, as there is one margin where the defect in the production units is stated in addition, the poisson distribution-based control chart also includes errors [2].

The control chart consists of three main lines as shown in Fig. 1 to Fig. 4.

• Central Line (CL)
• Upper Control Line (UCL)
• Lower Control Line (LCL)
• Control Chart (CC)

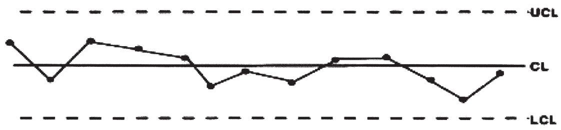


Fig. 1. Element of control chart (normal)

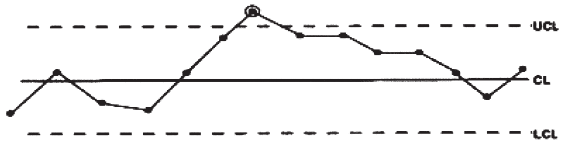


Fig. 2. Element of control chart (special cause)

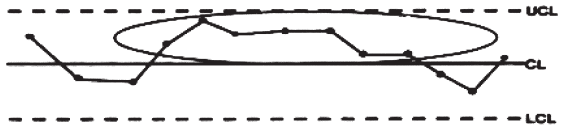


Fig. 3. Element of control chart (trend)

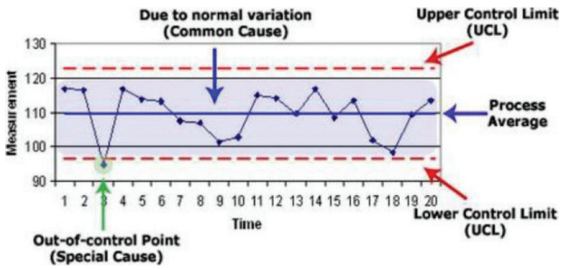


Fig. 4. Control chart

To understand product waste management, it has been suggested that waste in manufacturing should help find the root cause of the problem, waste reduction and production control Quality Control Tools (QC Tools) are used. Data was collected using the check sheet, which included surveying the waste condition and obtaining data from the inspection department on the amount of waste generated during the manufacturing process, as well as stating the problem [3]. To prioritize, use a Pareto chart and the 80:20 rule to select the most wasteful product, and then use a fish-bone diagram to analyze the problem and devise measures. Solve problems using brainstorming and compare all data before and after improvement [4]. Once the problem's main cause has been recognized, it's time to plan and design a solution, using the 5W 1H concept as a guideline. There must be performance standards in place to avoid the recurrence of problems [5].

III. Research Methodology

The steps that will be taken for research completion are as follows

- Study the relevant theories and data collection tools
- Define the problem and research data process
- Designing the experiment to find parameters and Testing
- The data analyzing
- Improve and decision-process
- Report the result in Results and Discussion [6]

A. Process Analysis

To investigate and analyze the problem, the process analysis is first conducted to identify the problem in the manufacturing process of crispy coconut rolls.

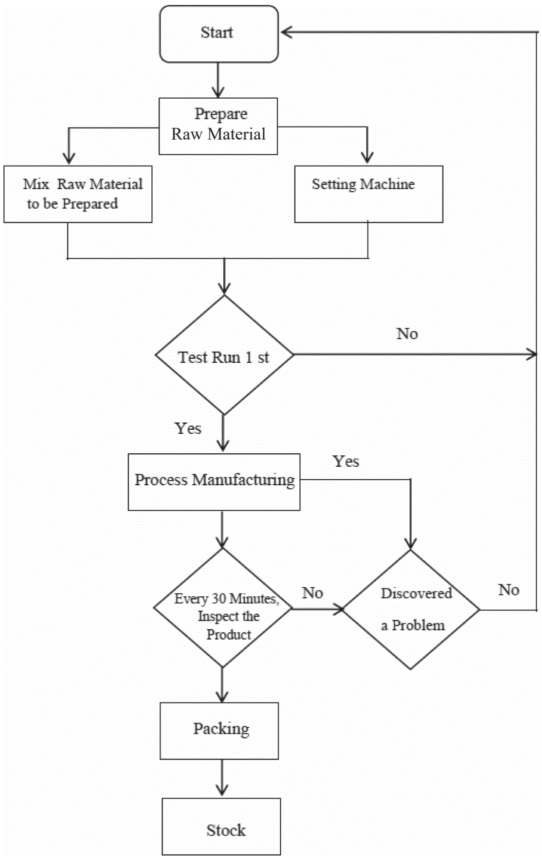


Fig. 5. Process flow of crispy coconut rolls

From the illustration shown in Fig. 5, The first step in the process is to prepare the ingredients for and production then mix the raw materials according to the production formula to make them vicious and meet the starch water requirements. For in-process operation start the machine and setting system then test run the product. The retort process has to be closely inspected by inspectors every 30 minutes to ensure that defective products do not pass to the packing process and customer. The defect after the retort process can be classified into four types of defects.

The researcher has observed operations collecting data on the types of waste in the production process. From the illustration, as shown in Fig. 6 check sheet tool was only used to record the data, and the data from the check sheet was used for more analysis for the next step.

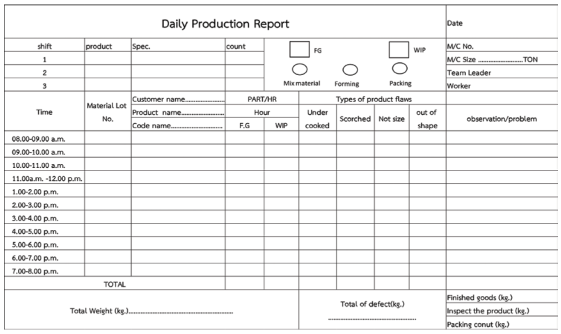

Daily Production Report
Date
shift | product | Spec. | count | FG | WIP
1
2
3
Mix material | Forming | Packing
M/C No.
M/C Size ....TON
Team Leader
Worker
Time | Material Lot No. | Customer name.................. Product name.................. Code name.................. | PART/HR Hour F.G WIP | Types of product flaws: Under cooked, Scorched, Not size, out of shape | observation/problem
08.00-09.00 a.m.
09.00-10.00 a.m.
10.00-11.00 a.m.
11.00a.m. -12.00 p.m.
1.00-2.00 p.m.
2.00-3.00 p.m.
3.00-4.00 p.m.
4.00-5.00 p.m.
5.00-6.00 p.m.
6.00-7.00 p.m.
7.00-8.00 p.m.
TOTAL
Total Weight (kg.)...
Total of defect(kg.) ....
Finished goods (kg.)
Inspect the product (kg.)
Packing conut (kg.)

Fig. 6. Check sheet example

The next step summarizes the amount of waste detected and records it on the check sheet. From June to August 2021, problems were detected and the total number of detections was indicated.

TABLE II
The Total Number of Type of Defect

Type waste	Number of Defects (kg)	% Percent Defect (kg)	%Total Percent Defect (kg)
Out of Shape	5,028	57	57
Not the Right Size	2,178	25	82
Undercooked	823	9	91
Scorched.	805	9	100
Total	**8,834**	**100**	**-**

From Table II, the data show the total number of each type of defect.

Use the waste data from crispy coconut roll production that does not meet all of the standards to classify the waste. The Pareto chart shows Fig. 7 represents it type of defects that are prioritized and need to be focused total piece number of defective products.

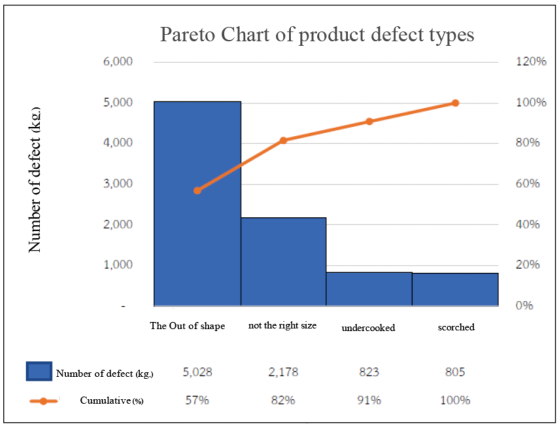


Fig. 7. The number of each defective type of crispy coconut rolls

The pareto chart of crispy coconut rolls shown in Fig. 7 represents four types of defects. The out-of-shape contributed 57%, not the right size contributed 25% undercooked contributed 9%, and scorched contributed 10%, respectively. Fig. 8 shows a picture of the defect type.

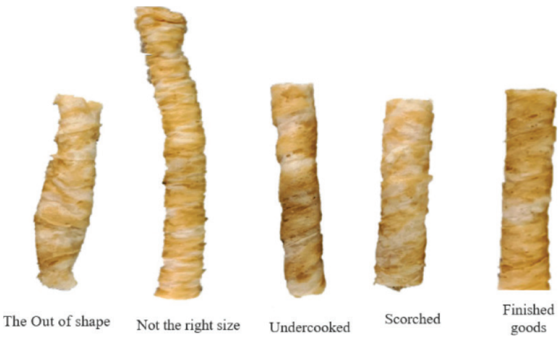


Fig. 8. Shows a picture of the defect type

Pareto charts show the ordered frequency counts of values for the different levels of a categorical or nominal variable. The charts are based on the “80/20” rule. This rule says that about 80% of the problems are the result of 20% of causes.

The Pareto chart represents the total waste of one major type of defect; the shape contributed 57% of total waste. In conclusion, this study research will focus on this one type of defect and the company can apply this waste reduction in crispy coconut rolls study to another defect type in the future.

B. Cause and Effect Analysis Tools

The researcher analyzes and determines the causes of the problems related to defect types.

The first step is to conduct an analysis using a cause and effect diagram. That is a cause and effect discovery tool that helps visualization for categorizing the potential causes of a problem. Typically cause and effect diagram combines the practice of the brainstorming process and encourages broad thinking, keeping users from limited thinking patterns that can lead to getting stuck, the process of asking why something happened repeatedly at each stage helps drill down to one or more root causes.

From the cause and effect diagram, it was found that the cause of the waste was due to the deformed work piece, which was the waste caused by people, machines, raw materials, and methods. Which proceeds to find the cause and the solution as follows:

Manpower

• Be inexperienced and unaware of how to operate a machine.

• Lack of Trained Experts shown in Fig. 9.

Fig. 9. The operational or functional labor of people engaged.

Machine the more misuse or lack of maintenance a machine shown in Fig. 10.

•The machine is to heat.

Fig. 10. Equipment used for production.

Material is a production process and contributes to service delivery processes.

• Blade wear

• Change the formula of the raw materials used in the manufacturing process.

• The sensor is not working.

• The roll spool is worn shown in Fig. 11.

Method

• The device settings aren’t standardized.

• Keeping the work environment clean shown in Fig. 12.

• Work checks that aren’t appropriate


Fig. 11. The technical repair roll is worn out.


Fig. 12. Keeping the work environment clean

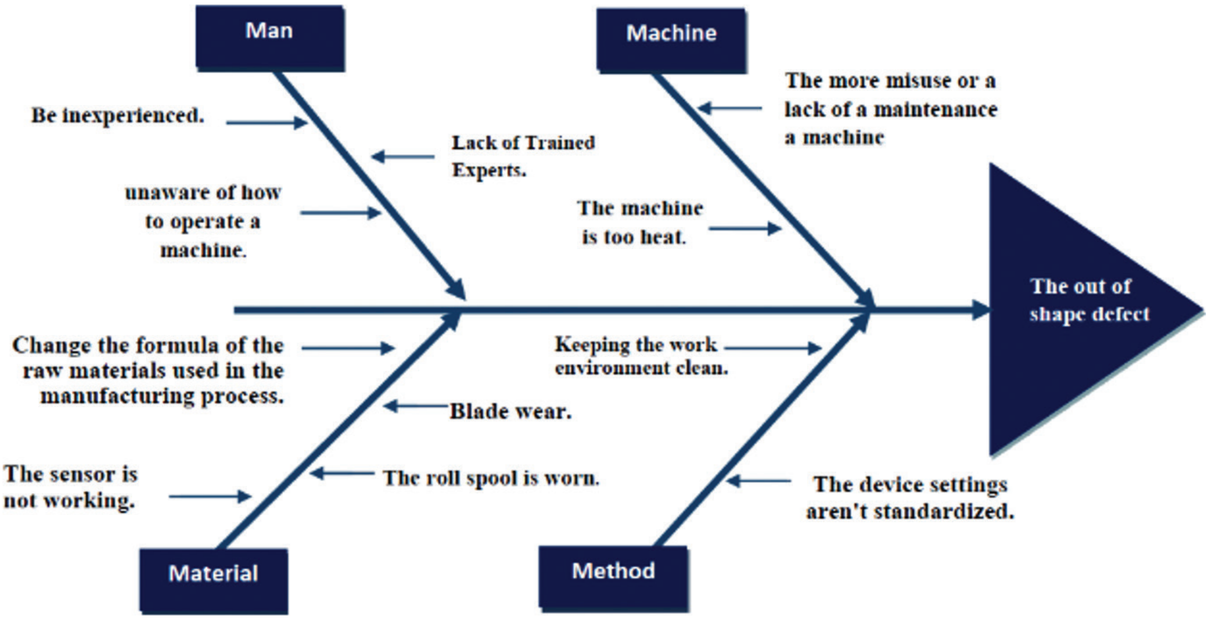


Fig. 13. Cause and effect diagram of the out of shape defect

When many causes are identified, the researcher has to analyze the importance of the detail of each cause can be controlled by controlling a plan or recommended action as identified in Table III.

This article uses a cause-and-effect diagram [7] shown in Fig. 13. It is a tool for brainstorming to find the cause of the forming process of crispy coconut rolls to find a solution to the problem.

TABLE III
PROPOSE A SOLUTION TO THE PROBLEM AND ACTION CAUSES OF THE DEFECT

Causes	Recommended Action
Be inexperienced and unaware	Staff training and qualification program
Lack of Trained Experts.	Staff training and qualification program
The more misuse or a lack of maintenance of a machine	Staff training and problem-solving
The machine is too heat.	Preventive maintenance
Blade wear	Preventive maintenance
Change the formula of the raw materials	Production planning on an Ongoing Basis
The sensor is not working.	Check the performance and condition of the machine
The roll spool is worn.	Check the performance and condition of the machine
The device settings aren’t standardized.	Check the performance and condition of the machine
Keeping the work environment clean.	Assign cleaning duties
Work checks that aren’t appropriate.	Staff training and problem-solving

Improvement guidelines, meetings with relevant production process working groups to put them into practice, improvement trials with the production process in mind, and continuous monitoring of operations are all part of the process.

IV. Results

Apply knowledge management process. When corrective action is taken in the production process according to predetermined remedial guidelines, the researcher has adopted a solution to improve all factors and all causes of problems together to make improvements in an integrated manner. The pre-improvement waste data from June to August 2021 was compared to the post-improvement waste generation data from September to November 2021, which accounted for the proportion of the waste generated. The waste was generated by all production processes during that time.

The waste in June 2021 will be out of shape. Out of shape from the crispy coconut roll production process amounted to 1,849 kg or 9.37 percent of the total weight produced (19,730 kg).

The data was used to create a control chart for the manufacturing process shown in Fig. 14. The results of the crispy coconut roll production process, which was quality checked. A total of 5 points outside the upper control line were discovered to be wasteful on June 5th, 8th, 9th, 14th, and 24th, 2021.

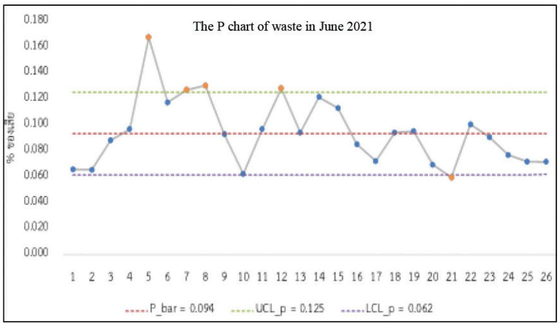


Fig. 14. The P-Chart shows points outside the upper control line in June 2021 before the improvement.

The waste in July 2021 will be out of shape. Out of shape from the crispy coconut roll production process amounted to 1,617 kg or 8.56 percent of the total weight produced (18,894 kg).

The data was used to create a control chart for the manufacturing process shown in Fig. 15. The results of the crispy coconut roll production process, which was quality checked. A total of 3 points outside the upper control line were discovered to be wasteful on July 2nd, 4th, 9th, and 21st, 2021.

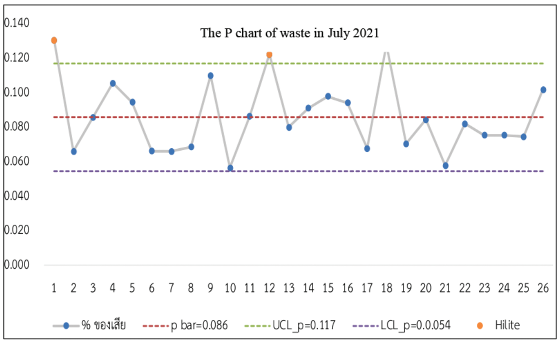


Fig. 15. The P-Chart shows points outside the upper control line in July 2021 before the improvement.

The waste in August 2021 will be out of shape. Out of shape from the crispy coconut roll production process amounted to 1,562 kg or 7.79 % of the total weight produced (20,051 kg).

The data was used to create a control chart for the manufacturing process shown in Fig. 16. The results of the crispy coconut roll production process, which was quality checked. A total of 4 points outside the upper control line were discovered to be wasteful on August 7th, 9th, 19th, and 24th, 2021.

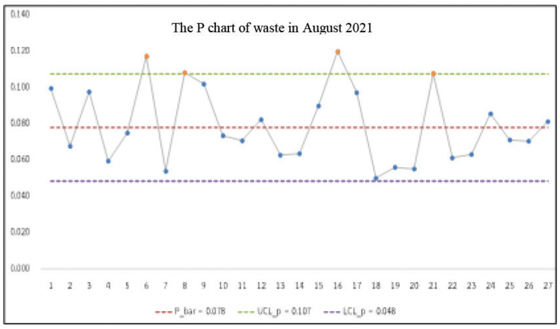


Fig. 16. The P-Chart shows points outside the upper control line in August 2021 before the improvement.

TABLE IV.
THE AMOUNT OF WASTE THE OUT OF SHAPE FROM JUNE-AUGUST 2021

Month 2021	Total Finished goods (kg)	Total Defect (kg)	Percent Defect (Average)
June	19,730	1,849	9.37
July	18,894	1,617	8.56
August	20,051	1,562	7.79
Total	58,675	5,028	8.57

Table IV, the amount of waste from the outshape of June-August 2021. Data on the occurrence of out-of-shape defects from June to August 2021. Prior to the improvement, wastes resulting from the manufacturing process were discovered in chronological order. Obviously, 5,028 kg of waste were produced out of total production of 58,675 kg, resulting in an 8.7% waste percentage.

Data on post improvement waste in the manufacturing industry from September to November 2021, Fig. 17 to Fig. 19 depicts the data. The P chart shows non-points outside the control line.

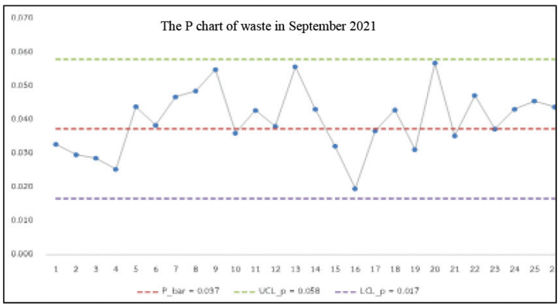


Fig. 17. The P chart shows non-points outside the control line in September 2021 after the improvement.

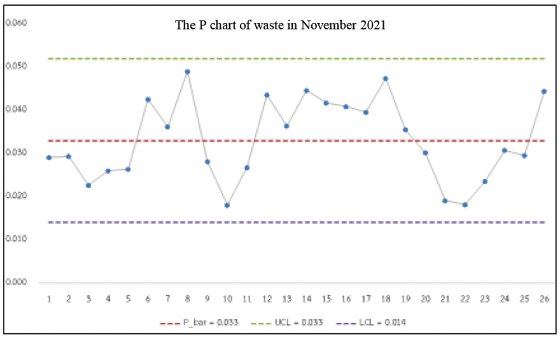


Fig. 18. The P chart shows non-points outside the control line in October 2021 after the improvement.

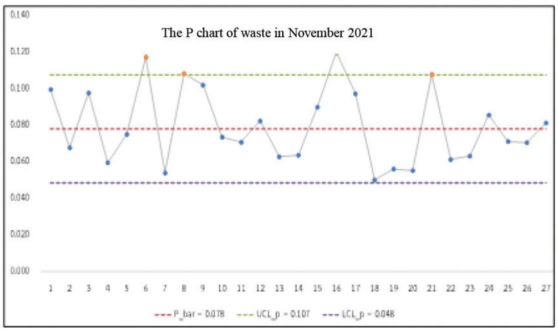


Fig. 19. The P chart shows non-points outside the control line in November 2021 after the improvement.

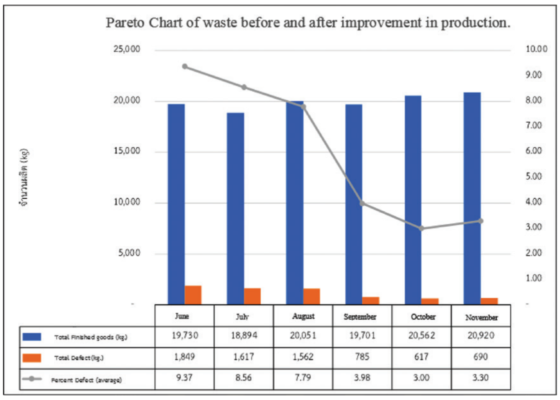


Fig. 20. Graph comparing the proportion of waste before and after improvement in production.

TABLE V.
THE AMOUNT OF WASTE THE OUT OF SHAPE OF SEPTEMBRR - OCTOBRT 2021

Month 2021	Total Finished goods (kg)	Total Defect (kg)	Percent Defect (Average)
September	19,701	785	3.98
October	20,562	617	3.00
November	20,920	690	3.30
Total	**61,183**	**2092**	**3.42**

From September to November 2021, shows a summary of defects caused by the product being out of shape. Following the improvements, it was discovered that there were wastes from the manufacturing process in order. A total finished goods of 2,092 kg with a defect waste percentage of 3.4%.

Summary of the findings Following the revision and improvement plan, it was discovered that the piece of work was out of shape due to the crispy coconut roll manufacturing process. As shown in Fig. 20, the data was sorted in order from June to November 2021.

V. Discussion

The Scope of the research study is only crispy coconut rolls work-piece type of defect is out of shape. Therefore, the factors for production might be different from this analysis.

Monitoring using the control chart to control the process, a control chart is developed to monitor the process's performance. Because the study is related to the defective rate of one major defect type, the P-chart is the appropriate control chart to monitor the proportion of defects that occurred in the production line.

VI. Conclusion

This experimental study collects data on pretest-posttest improvement with the goal of controlling the quality of crispy coconut rolls in the food processing industry by comparing the defect amount to the product quality of manufacturing before and after implementing quality control through the control chart.

Table VI includes the following items: a comparison of defective products between before and after using a control chart in September-October 2021.

TABLE VI
Comparison of Defective Product Types That are out of Shape Between Before and After

Month 2021		Total Defect (kg)	% Defect (Average)
Before	June	19,730	9.37
	July	18,894	8.56
	August	20,051	7.79
After	September	19,701	3.98
	October	20,562	3.00
	November	20,920	3.30

The information was compared. In September-November 2021, defects saw a significant reduction in waste, which was reduced by 2,079 kg on average as a percentage from 9.37 percent to 3.0 percent, a decrease of 6.37 percent.

Based on the amount of waste that was reduced from 5,028 kg to 2,949 kg, a total of 2,079 kg can be saved. It can be reduced by 58.65% in terms of waste percentage. When the reduced waste is compared to the lost product marketing opportunity in June-November 2021, the loss of product marketing opportunity is determined.

When the finished product is ready, customers will be able to buy it in kilograms at a price of 50 baht. As a result of this process, the product's value loss will be reduced from the original loss of 5,028 kg The weight has been reduced to 2,949 kg If the annual loss is 415,734 baht, the marketing value in three months is 103,934 baht.

VII. Suggestion

1) Study the use and application of the control chart in the quality control process in other food product lines.

2) Study the control chart in a group of defect types that are out of shape so that the data received can be compared with the amount of waste before and after improvement.

Quality control tools can be applied to other production processes as a result of the study of defect reduction in the crispy coconut rolls production process using quality control tools. to lessen the amount of waste produced.

References

[1] S. Piandee. (2021, Jul. 29). *7 Quality Tools (7 QC Tools)*. [Online]. Available https://www.gotoknow.org/posts/446393

[2] D. C. Montgomery, *Introduction to Statistical Quality Control*. New Jersey: John Wiley & Son, Inc, 2009, p.179.

[3] T. Meechamnan, "Waste Reduction of Black Dot Problem in Injection Process: Case Study of Plastic World Co., ltd.," M.S. thesis, Dept. Engineering Management. Eng., Dhurakij Pundit Univ., Bangkok, Thailand, 2020.

[4] W. Sutjaritjun, "Waste Reduction in Injector Process Case Study: Cleanliness out of Specification," M.S. thesis, Dept. Engineering Management. Eng., Dhurakij Pundit Univ., Bangkok, Thailand, 2018.

[5] S. Yaveera, "Application of Quality Tools in Egg Production in Small Chicken Farms," M.S. thesis, Dept. Industria. Eng., Chiang Mai Univ., Chiang Mai, Thailand, 2017.

[6] Indeed. (2021, Jul. 27). *Definition of Quality Control*. [Online]. Available https://www.indeed.com/career-advice/career-development/qc-process

[7] R. Paiboon, "Plastic Lump Defect Reduction in Recycled Plastic Pellet Manufacturing," M. S. thesis, Dept. Engineering Program. Eng., Chulalongkorn Univ., Bangkok, Thailand, 2017.

PloypailinPhrikthim received industrial engineering from Panyapiwat Institute of Management, Thailand, in August 2017 and currently studying for a master's degree in the Faculty of Engineering and Technology. Used to work in production engineering at Charoen Pokphand Foods, in 2017-2018. My current job is Student Innovator and Entrepreneur Affairs Officer, Strategic and Innovation Management Department at Panyapiwat Institute of Management, Nonthaburi, Thailand.

Paitoon Siri-O-Ran received a doctor of engineering in industrial engineering from Kasetsart University, Thailand, in 1994. Used to be a professor in industrial engineering, from 2004 – 2011 at Southeast Asia University, Thailand, and The current job is Head of industrial engineering faculty at Panyapiwat Institute of Management, Nonthaburi, Thailand.

Key Success Factors for Implementing Industry 4.0 of Thailand Manufacturing

Varuj Mingmankong[1], Chawalit Jeenanunta[2*], Rujira Chaysiri[3], and Yasushi Ueki[4]

[1,2,3]Sirindhorn International Institute of Technology, Thammasat University
[4]Institute of Developing Economies, Chiba, Japan
E-mail: varuj33124@gmail.com, chawalit@siit.tu.ac.th, rchaysiri@siit.tu.ac.th, yasushi_ueki@ide.go.jp

Received: February 4, 2022 / Revised: March 3, 2022 / Accepted: June 17, 2022

Abstract—**Thailand's manufacturing is a crucial component to drive the economy. However, most manufacturers in Thailand face obstacles and challenges to reach Industry 4.0 in their production process, which will help improve their productivity and efficiency. This paper aims to identify the key factors for Thailand's manufacturing firms to implement Industry 4.0. The success factors were identified by literature review and the Delphi method. First, we conducted a semi-structured interview with 2 experts and used the Delphi method to analyze the potential factors. The Analytic Hierarchy Process (AHP) is applied to find the ranking of the top 5 factors. There are 13 experts completing the questionnaire, and they are at the managerial level related to Industry 4.0. Analytic Hierarchy Process (AHP) was applied to find the relative weight of success factors to rank the importance of the success factors and give concrete guidance to implement Industry 4.0. The top 5 factors are leadership vision, support from top management, knowledge of technology by an employee, aligning Industry 4.0 with organization strategy, and the process of digitalization of the industry.**

***Index terms*—Industry 4.0, Key Success Factor, Analytic Hierarchy Process**

I. Introduction

Thailand's manufacturing is a crucial component of the Thai economy. Thailand manufacturing has 90.1% of the export value of significant products [1]. However, 60% of the development in Thailand's manufacturing was manual (25%) or a combination of computer numerically controlled and manual (35%), while only 5% of Thailand's manufacturing is fully automatic [2]. Regarding Thailand manufacturing's importance to Thailand's economic development, identifying and ranking key success factors to help Thailand's manufacturing reach a higher level is essential.

The current manufacturing age has evolved from mass production to customized production with digitalization in its operation [3]. Industry 4.0 integrates the physical and digital world using Cyber-Physical System (CPS) to create real-time productivity. Industry 4.0 provides real-time data to monitor, which helps improve productivity, performance, and product quality in the production process. Industry 4.0 is a cross-disciplinary concept integrating manufacturing with information technology and the Internet of Things (IoT). Industry 4.0 is characteristic of the automated manufacturing process, with a collaboration between man and machine in the center [4].

However, only a few firms could implement these concepts to bring their factories to Industry 4.0. Thus, we are interested in learning how those successful firms could do it. This paper aims to identify key success factors for Thailand's manufacturing to implement Industry 4.0. The paper is organized in the following structure. First, the potential success factors are identified from the literature review. The methodology for the Delphi method and AHP are then explained and followed by the results, discussion, and conclusion.

II. Literature Review

In this section, potential factors were identified from a comprehensive literature review. Articles related to the study of the implementation of Industry 4.0 were reviewed, and 19 factors were identified. The list of these potential success factors is summarized in Table I below.

TABLE I
POTENTIAL SUCCESS FACTORS FOR THAILAND MANUFACTURING IMPLEMENTING INDUSTRY 4.0

Number	Factors Name	Literature Support
1	The expected benefit of technology	[5] [6] [7] [8]
2	Real-time link between physical production and digital factory	[9]
3	Sufficient and capable technological infrastructure	[10]
4	Support from top management	[11]
5	Knowledge of technology by top management	[12]
6	Leadership vision	[13]
7	Align the Industry 4.0 with organization strategy	[14]
8	Degree of automation and the variety of product	[5]
9	Knowledge of technology by an employee	[5] [12] [14]
10	Employee readiness	[15]
11	High-cost investment	[9]
12	The process of digitalization of the industry	[16]
13	Government support	[10]
14	The importance of a company's competitiveness in the market to offer the latest technology	[15]
15	Impact of new technology on logistics and supply chain	[15]
16	Internet and connectivity in the shop floor	[17]
17	Potential Risk	[7]
18	The importance of a company's competitiveness in the market in order to offer the latest technology	[15]
19	Strategy to compromise with employee	[9]

III. METHODOLOGY

In this research, key success factors for Thailand's manufacturing approach to Industry 4.0 are identified using the AHP method to analyze the potential success factors from literature reviews and expert interviews. A literature review and semi-structured interviews were conducted to identify the potential factors. Delphi method is undertaken to find the potential factors by interviewing experts. Delphi provides more accurate decisions from a structured group of individuals. The expert answer questionnaire rates each possible factor from 1 to 5, ranking from the least to the most important. The factors with a total score higher than 9 will be selected as potential factors.

The Analytic Hierarchy Process (AHP) method was used to analyze the weights for each potential factor for implementing Industry 4.0. Thomas developed the AHP to quantify the weights of decision criteria [18]. Each of the experts utilized their experiences to compare the relative importance of each pair of items using a pairwise comparison questionnaire.

The detailed Analytic Hierarchy Process methodology involves 5 steps, as follows:

Step 1: Data collection

The experts who provide the comparison judgment of the pairwise questionnaire must have experience in implementing Industry 4.0. The experts should be a managerial-level industrial company that plans or adapts Industry 4.0 in their organization or an expert in the field of Industry 4.0. A respondent also needs to have job experience of more than 5 years in the field.

A pairwise comparison questionnaire has the format as shown in Fig. 1. The respondents answer the pairwise comparison to select the factors that represent the important degree of each factor to other factors. The respondents would rate by using a scale from 1 to 9, representing the degree of importance. The detailed Thomas scale is shown in Table II [18].

Step 2: Construct a pairwise comparison matrix

The pairwise comparison data will be represented in the square matrix $C^E = [c_{ij}^E]$, where i, j = 1…n represents the number of factors and E = 1…m represents the number of experts. The initial matrix has the principal diagonal containing values of 1 because each factor is as important as itself. Let's suppose that expert 3 decides that factor 1 is strongly preferred over factor 2. In the matrix, c_{12}^3 is rated as 5 and 1/5 in c_{21}^3.

TABLE II
PAIRWISE COMPARISON SCALE FOR AHP PREFERENCES [18]

Values/Rates	Description
1	Equally importance
2	Equally preferred, but with certain moderate differentiation tendencies
3	Moderately preferred
4	Preferred towards strongly preferred
5	Strongly preferred
6	Strongly preferred towards obviously preferred
7	Obviously preferred
8	Obviously preferred towards extremely preferred
9	Extremely preferred

	Factor A	Absolutely more important	←							Equally important	→							Absolutely more important	Factor B
1	Expected benefit from technology ผลประโยชน์ที่คาดว่าจะได้รับจากเทคโนโลยีใหม่	9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9	Internet and connectivity in the shop floor อินเตอร์เน็ตและสัญญาณภายในหน้างาน

Fig. 1. Example of a Pairwise Comparison of Questionnaires

Step 3: Determine the importance coefficients (weights) of the factors for each expert

For the square matrix C^E, the weights are consistent if $c_{ik}^E = c_{ij}^E c_{jk}^E$ for all i, j, and k. To determine the consistency matrix, find a vector W^E of order n such that $C^E W^E = \lambda^E W^E$. We call W^E an eigenvector, and λ^E is an eigenvalue. For a consistent matrix, $\lambda^E = n$.

There are many ways to calculate the eigenvector or the importance coefficients (weights) of the factors for each expert. The importance coefficients of the factors i are calculated based on equation (2), as follows:

$$S_i^E = \sum_{j=1}^{n} c_{ij}^E, \text{ for all i, and E} \quad (1)$$

$$\lambda_i^E = \frac{S_i^E}{\sum_{i=1}^{n} S_i^E}, \quad \text{for all i, and E} \quad (2)$$

Note that S_i^E is the element i in the score summation S_i^E and λ_i^E is the element i in the eigenvector λ_i^E. The following condition must be observed:

$$\sum_{i=1}^{n} \lambda_i^E, \text{ for all E} \quad (3)$$

Step 4: Determine the Consistency Index (CI) and Consistency Ratio (CR) of the matrix

The objective is to determine whether the decision-makers have been consistent in their choices [18]. For example, suppose the decision-makers affirm that the Leadership Vision factor is more important than the Employee Readiness factor. In addition, the Employee Readiness factor is more important than the Government Support factor. It would be inconsistent to affirm that the Government Support factor is more important than the Leadership Vision factor (if A>B and B>C it would be inconsistent that A<C).

However, the condition $c_{ik}^E = c_{ij}^E c_{jk}^E$ does not hold because the human judgment is inconsistent. Thus, we could find W^E vector that satisfies $C^E W^E = \lambda_{max}^E W^E$ and $\lambda_{max}^E \geq n$. If $\lambda_{max}^E = n$, the judgment of expert E is consistent. The next step is to calculate the λ_{max}^E as follows:

$$R_i^E = \frac{\sum_{j=1}^{n} \lambda_i^E c_{ij}^E}{\lambda_i^E}, \text{ for all i, and E} \quad (4)$$

Where R_i^E represents the estimator for λ_{max}^E. Since there are n estimates for λ_{max}^E. The mean of the estimators, λ_{max}^E is calculated based on equation (4), as follows:

$$\lambda_{max}^E = \frac{\sum_{i=1}^{n} R_i^E}{n}, \text{ for all E} \quad (5)$$

The difference between λ_{max}^E and n shows the inconsistency of the judgment. The Consistency Index (CI) for expert E can be calculated by (λ_{max}^E -n)/ (n-1). The Consistency Ratio (CR) for expert E is then calculated by CI^E/RI, where RI is the index of consistency for random judgments [18]. If $CR^E < 0.10$, then the matrix or the judgment for expert E is consistent. Otherwise, the expert must re-evaluate the conflicting judgment.

TABLE III
RANDOM CONSISTENCY INDEX FOR THE CORRESPONDING NUMBER OF FACTORS [18]

n	1	2	3	4	5	6	7	8	9
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45

Step 5: Determine the weighted values of comprehensive experts

Once the Consistency Ratio (CR) for all experts is consistent, each factor's weights for all experts are the normalized sum of the eigenvectors.

$$\lambda_i = \frac{\sum_{E=1}^{m} \lambda_i^E}{m}, \text{ for all i} \quad (6)$$

IV. RESULTS AND DISCUSSION

Delphi method is conducted to find the potential factors by interviewing 2 experts. Each expert is selected because there has experience in implementing Industry 4.0 in many organizations and obtained an opportunity to talk with many managerial levels in many organizations. They have high experience in implementation work and understand what organizations need for Industry 4.0 to be implemented. The demographics of the two experts are shown in Table IV.

TABLE IV
EXPERT DETAIL IN THE DELPHI METHOD

Expert No.	Field of Expert	Position
1	Industrial IoT platform provider	Sales Engineer
2	Industrial IoT platform provider	Solution Engineer

Initially, there are 19 factors identified as potential success factors from the literature. The experts in the specific field were interviewed to discuss and rate these potential factors in the next step. The experts are a sales engineer and a solution engineer in an Industrial IoT platform provider. The experts were selected due to their adeptness in implementing various projects related to Industry 4.0. Both experts also had experience in conversations with the top managerial level in many organizations. Thus, they have a deep understanding of what is important for implementing Industry 4.0 to succeed in the organization.

After the short interview with two experts using the Delphi method, the result is shown in Table V. Some factors had been dropped out because the total score is less than 9. 9 factors out of 19 were selected as the potential success factors to implement Industry 4.0, and they are grouped into 4 categories. In Table VI, there are 9 factors that remained, Expected benefit from technology, Real-time link of physical production and digital factory, support from top management, Leadership vision, Align the Industry 4.0 with organization strategy, knowledge of technology by an employee, Employee readiness, the process to the digitalization of the industry, and Government support.

TABLE V
EXPERT SCORING FROM THE DELPHI METHOD

Number	Factors Name	Expert 1	Expert 2	Total Score
		Score	Score	
1	The expected benefit of technology	5	5	10
2	Real-time link between physical production and digital factory	5	5	10
3	Sufficient and capable technological infrastructure	4	4	8
4	Support from top management	5	5	10
5	Knowledge of technology by top management	4	4	8
6	Leadership vision	4	5	9
7	Align the Industry 4.0 with organization strategy	4	5	9
8	degree of automation and the variety of product	4	4	8
9	Knowledge of technology by an employee	5	4	9
10	Employee readiness	5	5	10
11	High-cost investment	4	4	8
12	The process of digitalization of the industry	5	5	10
13	Government support	4	5	9

TABLE V
EXPERT SCORING FROM THE DELPHI METHOD (CON.)

Number	Factors Name	Expert 1 Score	Expert 2 Score	Total Score
14	The importance of a company's competitiveness in the market to offer the latest technology	4	4	8
15	Impact of new technology on logistics and supply chain	2	1	3
16	Internet and connectivity in the shop floor	5	3	8
17	Potential Risk	3	3	6
18	The importance of a company's competitiveness in the market in order to offer the latest technology	4	4	8
19	Strategy to compromise with employee	4	3	7

TABLE VI
THE SUCCESS FACTORS FOR THAILAND MANUFACTURING APPROACH INDUSTRY 4.0

No.	Category	Factors
1	Technology	The expected benefit of technology
2		Real-time link between physical production and digital factory
3	Leadership	Support from top management
4		Leadership vision
5	Organization	Align the Industry 4.0 with organization strategy
6		Knowledge of technology by an employee
7		Employee readiness
8		The process of digitalization of the industry
9		Government support

The following steps are the AHP method applied to identify the key success factors to implement Industry 4.0 for Thailand manufacturing.

Step 1: Data collection

The detailed descriptions of the potential factors are from the literature review and the Delphi method, as presented in Table VII. In this step, the pairwise comparison questionnaire is constructed and sent to the target experts to obtain the data. The research target must be an expert in Industry 4.0 technology or managerial level in a manufacturing field planned or used to implement Industry 4.0 in their organization.

TABLE VII
POTENTIAL FACTORS DESCRIPTION

Factors	Factor's Name	Descriptions
a	The expected benefit of technology	The expected benefit from technology which will help in improving ROI or performance in the organization
b	Real-time link between physical production and digital factory	The data is obtained from the machine in production in real time.
c	Support from top management	Support from top management in terms of financial and project approval
d	Leadership vision	Leadership has a deep understanding of Industry 4.0 and is ready for a challenge.
e	Align the Industry 4.0 with organization strategy	The project related to Industry 4.0 must be aligned with the organizational strategy
f	Knowledge of technology by an employee	An employee should understand and have some skills related to Industry 4.0
g	Employee readiness	An employee is ready for a change within an organization
h	The process of digitalization of the industry	The process of obtaining the data from physical production to visualization
i	Government support	The government support in terms of policy, law, or BOI

According to the result of the questionnaire, 12 experts (92%) are managerial level in the organization, and 1 expert (8%) is an expert of Industry 4.0 technology, as presented in Table VIII. 8 experts (62%) have planned or using technology related to Industry 4.0, 1 expert (8%) sell the equipment related to Industry 4.0, 3 experts (22%) are solution provider related to Industry 4.0, and 1 expert (8%) is wanting to study more detail in Industry 4.0, as presented in Fig. 2.

According to Table VIII, all experts have a year experience more than 5 years. 2 experts (16%) have 6-10 years of experience, 4 experts (31%) have 10-15 years of experience, 2 experts (16%) have 15-20 years of experience, and 5 experts (37%) have 20-30 years of experience.

TABLE VIII
EXPERT POSITION AND YEARS OF EXPERIENCE

Expert	Position	Years of Experience
1	General Manager	15
2	Business Development Manager	10
3	Chief Executive Officer	11
4	Chief Executive Officer	30
5	General Manager	25
6	Chief Executive Officer	26
7	Engineering Manager	10
8	Senior Manufacturing Manager	18
9	Managing Partner	17
10	Chief Executive Officer	6
11	Head of Solution Engineer	8
12	Chief Executive Officer	25
13	Engineering Manager	20

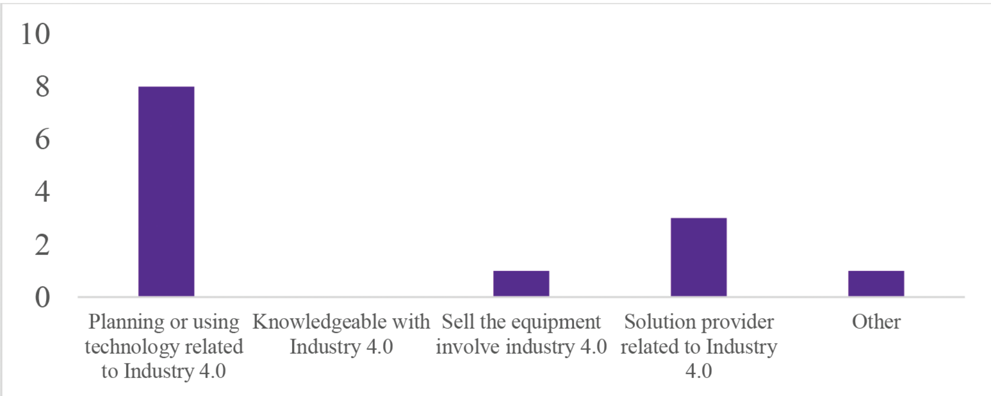


Fig. 2 Experts' knowledge in Industry 4.0

Step 2: Construct a pairwise comparison matrix.

In Table IX, the matrix represents the weight comparison between every pair of factors of the first expert. The matrix diagonal value is 1 because comparing itself would be equally important. To fill in the matrix, if factor a is slightly less important than factor b, which equals 3. Factor b compared to factor a should be inverse, which is 1/3.

TABLE IX
MATRIX OF COMPARISON WEIGHT BETWEEN EACH FACTOR FROM THE FIRST EXPERT

Factor	a	b	c	d	e	f	g	h	i
a	1.00	0.33	0.13	0.33	0.20	0.17	1.00	1.00	0.17
b	3.00	1.00	0.33	1.00	0.33	0.20	0.14	0.20	0.33
c	8.00	3.00	1.00	1.00	5.00	0.14	9.00	9.00	7.00
d	3.00	1.00	1.00	1.00	3.00	0.14	1.00	1.00	5.00
e	5.00	3.00	0.20	0.33	1.00	1.00	0.20	1.00	5.00
f	6.00	5.00	7.00	7.00	1.00	1.00	7.00	7.00	5.00
g	1.00	7.00	0.11	1.00	5.00	0.14	1.00	1.00	5.00
h	1.00	5.00	0.11	1.00	1.00	0.14	1.00	1.00	5.00
i	6.00	3.00	0.14	0.20	0.20	0.20	0.20	0.20	1.00

Step 3: Determine the importance coefficients (weights) of the factors for each expert

Equations (1) and (2) are used to calculate the importance coefficients, while equation (3) is also observed, as shown in Table X.

TABLE X
THE IMPORTANCE COEFFICIENTS (WEIGHTS) OF THE FACTOR FROM THE FIRST EXPERT

Factor	a	b	c	d	e	f	g	h	i	Score Summation (S)	The Importance Coefficients or Eigenvector (λ)
a	1.00	0.33	0.13	0.33	0.20	0.17	1.00	1.00	0.17	4.33	0.024
b	3.00	1.00	0.33	1.00	0.33	0.20	0.14	0.20	0.33	6.54	0.036
c	8.00	3.00	1.00	1.00	5.00	0.14	9.00	9.00	7.00	43.14	0.239
d	3.00	1.00	1.00	1.00	3.00	0.14	1.00	1.00	5.00	16.14	0.089
e	5.00	3.00	0.20	0.33	1.00	1.00	0.20	1.00	5.00	16.73	0.093
f	6.00	5.00	7.00	7.00	1.00	1.00	7.00	7.00	5.00	46.00	0.255
g	1.00	7.00	0.11	1.00	5.00	0.14	1.00	1.00	5.00	21.25	0.118
h	1.00	5.00	0.11	1.00	1.00	0.14	1.00	1.00	5.00	15.25	0.084
i	6.00	3.00	0.14	0.20	0.20	0.20	0.20	0.20	1.00	11.14	0.062
Total	-	-	-	-	-	-	-	-	-	**180.54**	**1.000**

Step 4: Determine the Consistency Index (CI) and Consistency Ratio (CR) of the matrix

In this step, the program called AHP-OS is used to calculate CR from the matrix [19]. The first round of CR for each expert is represented in Table XI.

TABLE XI
CONSISTENCY RATIO FOR EACH EXPERT

Expert	Consistency Ratio	
	1st Round	2nd Round
1	0.458	0.092
2	0.205	0.076
3	0.220	0.096
4	0.148	0.070
5	0.214	0.093
6	0.410	0.096
7	0.179	0.094
8	0.126	0.055
9	0.288	0.056
10	0.154	0.080
11	0.073	0.073
12	0.060	0.060
13	0.124	0.052

According to the table, only expert numbers 11 and 12 had CR lower than 0.1 in the first round. Therefore, the other 11 experts were interviewed again to adjust conflict judgments. We interviewed the experts by confirming their opinions on comparing the selected factor with other factors. During the interview, the AHP-OS program is used to check the CR in real time. The interview will continue until the CR is lower than 0.1. The adjusted CR is shown in the second-round column in Table XI.

Step 5: Determine the weighted values of comprehensive experts.

In this step, the compressive weights for each factor are calculated. Equation (6) is used to calculate the comprehensive weight of the factors. The results for both rounds of the interview are shown in Table XII.

TABLE XII
COMPREHENSIVE WEIGHT OF ALL FACTORS

Factor	Comprehensive Weight	
	1st Round	**2nd Round**
a	0.0720	0.0625
b	0.0443	0.0479
c	0.1625	0.1625
d	0.1784	0.1868
e	0.1139	0.1147
f	0.1139	0.1172
g	0.1069	0.0983
h	0.1024	0.1067
i	0.1057	0.1033
Summation	1.0000	1.0000

After computing comprehensive weight, the ranking of each factor is obtained, presented in Table XIII, Fig. 3, and Fig. 4.

TABLE XIII
FACTOR RANKING

Ranking	Factor	
	1st Round	**2nd Round**
1	Leadership vision	Leadership vision
2	Support from top management	Support from top management
3	Align the Industry 4.0 with organization strategy	Knowledge of technology by an employee
4	Knowledge of technology by an employee	Align the Industry 4.0 with organization strategy
5	Employee readiness	The process of digitalization of the industry

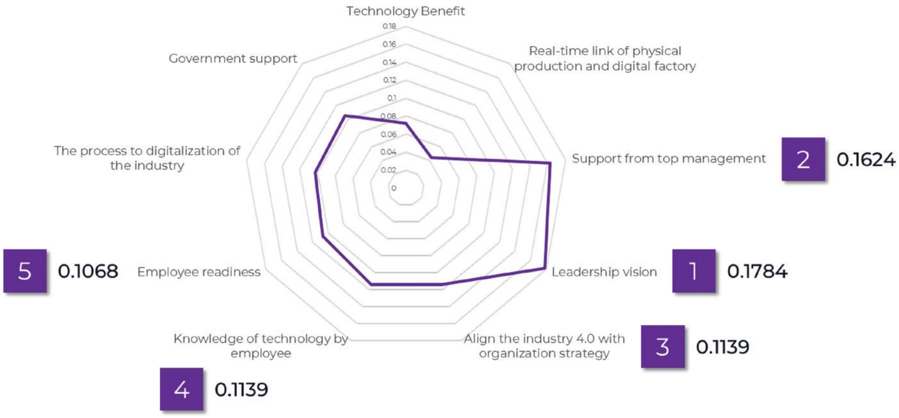


Fig. 3. The comprehensive weight for each factor in the 1st Round

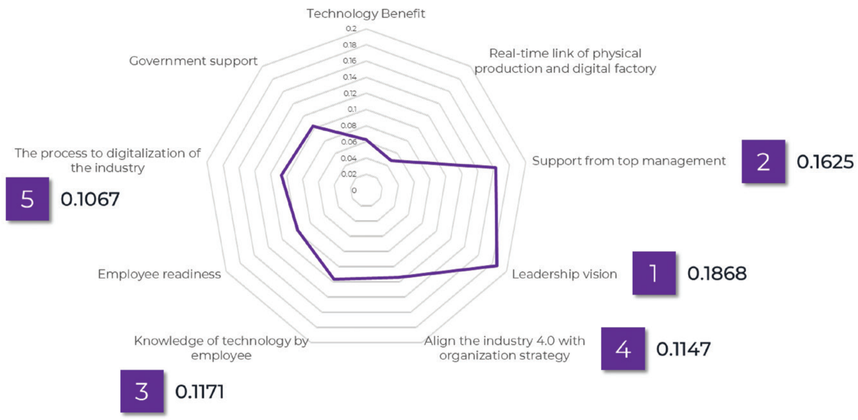


Fig. 4. The comprehensive weight for each factor in the 2nd Round

From Table XIII, there is a change of ranking after the experts adjusted their ranking. Rank 3 and 4 factors have been switched placed after the adjustment. Moreover, the Rank 5 factor of the second round is replaced by the new factor, the process of digitalization of the industry. However, Leadership and support from top management do not change after the adjustment. It indicates that these two factors are important in respondents' views. If we look at the comprehensive weight of these two factors, the weights are in the top two highest weights before and after adjustment. Thus, it can indicate that leadership vision and support from top management are the most important factors when implementing Industry 4.0 into the organization

From Fig. 4, the results indicate that leadership vision (0.1868) is the most critical factor for Thailand's manufacturing to implement Industry 4.0. The top manager must have the vision to challenge new technology and accept the change within an organization. The second factor is support from top management (0.1625), which is needed when starting something new within the organization. The third factor is knowledge of technology by employees (0.1171), which will help the implementation become more accessible and more effective. The fourth factor is aligning Industry 4.0 with organization strategy (0.1147). Aligning the project with the organization's strategy will create a clear implementation objective. The fifth factor is the process of digitalization of the industry (0.1067) which a well-indicated process to digitalize would create a steady work process with a clear information flow. The first step would be to educate the leaders and top managers about Industry 4.0. There are many courses related to Industry 4.0 in Thailand, such as courses from TGI or SIMTEC. Lastly, an employee with skills and knowledge in Industry 4.0 would help the project implementation become easier. The other choice is outsourcing an expert on the solution related to Industry 4.0 to help implement. With help from an expert or skilled employee, it is possible to create a steady process to digitalize which will lead to the excellent step for implementing Industry 4.0.

V. Conclusion

In this research, the key success factors are identified for Thailand manufacturing firms to implement Industry 4.0. The potential factors were identified using a literature review and semi-structured interviews with 2 top experts using the Delphi method. Then, the Analytic Hierarchy Process (AHP) is applied to find the top 5 factors. There are 13 experts who are managerial level related to Industry 4.0 completed the questionnaire. Analytic Hierarchy Process (AHP) was applied to find the relative weight of important success factors to rank the importance of the success factors and give concrete guidance to implement Industry 4.0.

Based on this research, we recommended that the factors that should be considered when implementing Industry 4.0 for Thailand manufacturing are leadership vision, support from top management, knowledge of technology by an employee, aligning Industry 4.0 with organization strategy, and the process to a digitalization of the industry.

For starters, the top management needs to have knowledge about Industry 4.0. It would be risking to invest their money in the unknown.

Acknowledgment

This research is fully supported by the Center of Excellence in Logistics and Supply Chain System Engineering and Technology (COE LogEn), Sirindhorn International Institute of Technology (SIIT), Thammasat University. We would like to show our gratitude to all the experts who helped with the questionnaire construction. Without their assistance, we would not have completed a questionnaire for this research. We are truly thankful for their support.

References

[1] NESDC. (2020, Apr.). *Thai Economic Performance in Q4 and 2020 and Outlook for 2021*. [Online]. Available: https://www.nesdc.go.th/nesdb_en/article_attach/article_file_20220221084015.pdf

[2] A. Kohpaiboon, "Industry 4.0 Policies in Thailand," *ISEAS-Yusof Ishak Institute*, no. 2, pp. 1-17, Feb. 2020.

[3] P. Kumar, J. Bhamu, and K. S. Sangwan, (2021). "Analysis of Barriers to Industry 4.0 Adoption in Manufacturing Organizations: An ISM Approach," *Procedia CIRP*, vol. 98, pp. 85-90, Jan. 2021.

[4] G. Nick, T. Kovács, A. Kő et al., "Industry 4.0 Readiness in Manufacturing: Company Compass 2.0, a Renewed Framework and Solution for Industry 4.0 Maturity Assessment," *Procedia Manufacturing*, vol. 54, pp. 39-44, Aug. 2021.

[5] F. Yu and T. Schweisfurth, "Industry 4.0 Technology Implementation in Smes - A Survey in The Danish-German Border region," *International Journal of Innovation Studies*, vol. 4, pp. 76-84, Sep. 2020.

[6] C. Liebrecht, M. Kandler, and M. Lang, "Decision Support for the Implementation of Industry 4.0 Methods: Toolbox, Assessment and Implementation Sequences for Industry 4.0," *Journal of Manufacturing Systems*, vol. 58, Part A, pp. 412-430, Jan. 2021.

[7] H. Yüksel, "An Empirical Evaluation of Industry 4.0 Applications of Companies in Turkey: The Case of a Developing Country," *Technology in Society*, vol. 63, p. 101364, Nov. 2000.

[8] S. Jaskó, A. Skrop, T. Holczinger et al., "Development of Manufacturing Execution Systems in Accordance with Industry 4.0 Requirements: A Review of Standard- and Ontology-Based Methodologies and Tools," *Computers in Industry*, vol. 123, p. 103300, Dec. 2020.

[9] V. A. Wankhede and S. Vinodh, "Analysis of Industry 4.0 Challenges Using the Best Worst Method: A Case Study," *Computers & Industrial Engineering*, vol. 159, Sep. 2021.

[10] R. Kumar, R. D. Raut, K. Nayal et al., "To Identify Industry 4.0 and Circular Economy Adoption Barriers in the Agriculture Supply Chain by Using ISM-ANP," *Journal of Cleaner Production*, vol. 293, p. 126023, Apr. 2021.

[11] M. Sony and S. Naik, "Critical Factors for the Successful Implementation of Industry 4.0: A Review and Future Research Direction," *Production Planning & Control*, vol. 31, no. 10, pp. 799-815, Nov. 2019.

[12] D. Dikhanbayeva, A. Tokbergenova, Y. Lukhmanov, "Critical Factors of Industry 4.0 Implementation in an Emerging Country: Empirical Study," *Future Internet*, vol. 13, no. 6, p. 137, May. 2021.

[13] A. R. Bakhtari, V. Kumar, M. M. Waris et al., "Industry 4.0 Implementation Challenges in Manufacturing Industries: An Interpretive Structural Modelling Approach," *Procedia Computer Science*, vol. 176, pp. 2384-2393, Oct. 2020.

[14] M. Cugno, R. Castagnoli, and G. Büchi, "Openness to Industry 4.0 and Performance: The Impact of Barriers and Incentives," *Technological Forecasting and Social Change*, vol. 168, p. 120756, Jul. 2021.

[15] F. Nwaiwu, M. Duduci, F. Chromjakova, "Industry 4.0 Concepts Within the Czech SME Manufacturing Sector: An Empirical Assessment of Critical Success Factors," *Business: Theory and Practice*, vol. 21, no. 4, pp. 58-70, Feb. 2021.

[16] A. Ribeiro, A. Amaral, and T. Barros, "Project Manager Competencies in the Context of the Industry 4.0," *Procedia Computer Science*, vol. 181, pp. 803-810, Jan. 2021.

[17] H. Meissner, R. Ilsen, and J. C. Aurich, "Analysis of Control Architectures in the Context of Industry 4.0," *Procedia CIRP*, vol. 62, pp. 165-169, Dec. 2017.

[18] T. L. Saaty, "The Analytic Hierarchy Process: Planning, Priority Setting, Resources Allocation," New York: McGraw-Hill, 1980, pp. 1-287.

[19] K. D. Goepel. (2019, Sep, 19). AHP Calculator-AHP-OS. [Online]. https://bpmsg.com/ahp

Varuj Mingmankong holds a B.Sc. in Engineering Management and M.Engineering. In Logistics and Supply Chain Systems Engineering from the University of Thammasat University. His research interests are in the application of lean automation and innovation. He joined a company related to industrial IoT platforms as a solution engineer.

Chawalit Jeenanunta holds a B.Sc. in Mathematics and Computer Science and a M.Sc. in Management Science from the University of Maryland. He received his Ph.D. in Industrial and Systems Engineering from Virginia Polytechnic Institute and State University. He joined Sirindhorn International Institute of Technology, Thammasat University, Thailand as a lecturer, and now he is an associate professor. He was a chair of the Management Technology curriculum, Head of the School of Management Technology, and Deputy Director for Building, Ground, and Properties. He is also the head of the Center of Excellence in Logistics and Supply Chain Systems Engineering (LogEn) and the Center for Technology Transfer of Industry 4.0.

His research interests are in applications of operations research and innovation. He received research funding from the Thailand Research Fund, the Institute of Developing Economies and Japan External Trade Organization (IDE-JETRO), the Economic Research Institute for ASEAN and East Asia (ERIA), PTT Logistics and Electricity Generating Authority of Thailand (EGAT).

Rujira Chaysiri received a Bachelor of Arts in Mathematics from the University of Virginia, a Master of Science in Operations Research from Columbia University, and a Doctor of Philosophy in Systems Engineering from the University of Virginia, USA. He is currently working as a lecturer in the School of Management Technology, Sirindhorn International Institute of Technology, Thammasat University, Thailand. His research focuses on creating analytical models that are used to obtain exact or near-optimal solutions for operational problems in the industry.

Yasushi Ueki is Senior Research Fellow of Institute of Developing Economies (IDE), Japan External Trade Organization, Chiba, Japan. He is also Guest Senior Researcher of Research Institute of Automobile and Parts Industries, Waseda University, Tokyo, Japan and Research Fellow of Economic Research Institute for ASEAN and East Asia (ERIA), Jakarta, Indonesia. He received Ph.D. in international public policy from Osaka University, Japan in 2004. After joining IDE in 1999 as research fellow, he served United Nations Economic Commission for Latin America and the Caribbean (ECLAC) in Santiago, Chile as an expert from 2002 to 2005 and ERIA as an economist from 2014 to 2018. His recent research focuses on technologies, management, and policies for industrial development and innovation.

Time Reduction in Picking the Product a Case Study of Roof Tile Warehouse

Pakdee Jaisue[1] and Paitoon Siri-O-Ran[2]

[1,2]Faculty of Engineering and Technology Panyapiwat Institute of Management, Nonthaburi, Thailand.
E-mail: pakdeejai@gmail.com, paitoonsir@pim.ac.th

Received: March 24, 2022/ Revised: August 10, 2022/ Accepted: August 11, 2022

***Abstract*—The warehouse has a problem with waiting time for picking products. Also, the case study has a problem with finding goods. The stored area of goods has stored at a random location. Operators work redundantly and make it difficult to find products. This paper aims to improve working time by using simulation software. FlexSim simulation software is used to simulate the models and analyze the performance by input data layout with different conditions. Generating 3D models of scenarios for simulation software. In the first scenario, create the random laying goods location model. The data used to be an input based on actual work. In addition, the other two scenarios model solutions to improve search times faster and more conveniently. The result of the simulation is 10039.01 seconds for 5 orders in a random model. Second, the classification model has 9274.78 seconds of time working and is stored together on the same side. Third, the distribution group model stores goods in separate locations. This model has a working time of 8920.49 seconds.**

***Index Terms*— ABC Analysis, FlexSim Simulation Software**

I. Introduction

A warehouse is a building for the storage of all resources and goods from industries such as manufacturing, distribution, transport business, etc. The storage management is important for the logistic operation process. The competitive advantage of good management is to satisfy customers and make more potential for the company.

The distribution is the place to dwell goods or materials before being transported to a network stocking location or end-users. It is important to manage and occupy a position within the warehouse.

In the case study, the roof-tile warehouse for the home construction industry has a problem with time searching for material. This research requires improving and reduce the time process by using simulation software based on FlexSim to optimize storage location to decrease the timing of finding goods in the storage.

The researcher took the data from the working process within the warehouse from the case study. Using information like the category of goods, working process, and daily ordered requirements for the case study.

Due to the time limitation, the researchers used a simulation software tool to show improvements. Considered outcome data from Flexsim simulation to be the result of improvement.

II. Background

A. Flexsim Introduction

Among the things of simulation software such as FlexSim [1], Siemens [2], and Arena Simulation Software [3] are often used for generating the model for determining the various solutions in industrial and engineering systems. Anyway, FlexSim is one of the software to provide the replication of realistic system situations. Zhu et al., [4] have published information research about this software is the first in the world that software integrates the C++IDE and uses graphical objects to generate simulation environments depending on input information. This simulation software provides experience in simulation manufacturing environments and health care. The interface for users considered for friendliness, analysis of input and output data during simulation, can generate the complex model and analysis data immediately based on computer processing efficiency and output statistics producing standard capability such as utilization, cycle time, and wait time [5]. Along with the present FlexSim software was developed to be easy to use and has more potential analysis of input data. This software splits a special option that analyzes and offers a solution experiment model for finding the best various solution that provides those who has a license only.

The visualization of this program helps users understand the complex process. The 3D interface has friendliness with users as shown in Fig. 1 the 3D objects.

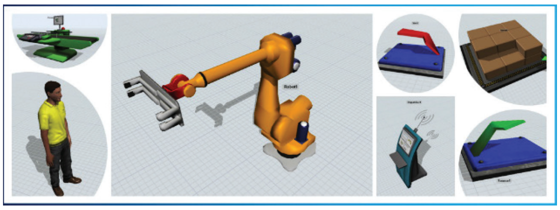


Fig. 1. The 3D objects

From the illustration, as shown in Fig. 1. The 3D object on FlexSim software creates main items and objects that are used more often for example:

• The source the object that controls and distributes item or service into the system or the model we generate

• Queue waiting for space for the item, material, goods, component, etc. This object is used for waiting in the process or dwelling on the product.

• Sink this object is meant to be the end process and shows the basic statistics including the quantity.

• Processor this object is meant to be the one process or representative of the machine in one process. And also it can input the data such as processing time to make the material into the system under different time conditions.

• Operator a representative for man-work. The routine time of working can be input to properties of the objects.

• TaskExecuters can represent many forms such as track, Forklift, AGV, plane, robot, crane, operator, and dispatcher. The task executors can be moved.

• Dispatcher is used for controlling the TaskExecuter to work at the same objects such as five operators working at the processor.

• Rack and floor storage are used for storing the items in process or material before into process. The rack can set up the data about dwell and quantity.

B. Theory

Plant layout is the works area within the factory that defines it to be. That area arranging machines and utilization of facilitating planning production. The position of goods in that area needs to be organized to help better control working capital costs.

The ABC analysis is one of the warehouse management techniques that define the rank of goods based on demand, cost, and risk data. The products are kept stocked based on profitability or sales volume. The goods are separate to classify of rank ABC.

Class A has a high value with 70%-80% of annual consumption value but has 10%-20% of total goods in inventory. The data of the A class must have high accuracy.

Class B has a medium value with 15%-20% of annual consumption value and 30% of total goods. This class regularly needed to check and record all data.

Class C has a low value with 5% of annual consumption value but has a total volume of goods in inventory high to 50%. This class could be general parts or goods that are used daily.

C. Literature Review

Tokgöz [5] reviews a software simulation named FlexSim. To investigate whether undergraduate industrial engineering students can be used and utility in simulation courses for students. The researcher has generated 6 sample scenarios for project design experience using simulation software. This paper design 6 example projects about health care, production, purchasing, inventory management, and gaming system, and then let junior and senior-level IE students use the simulation software FlexSim in real-life manufacturing. This paper explained the advantages and disadvantages of researchers and their example experience that collecting for three years. The main advantage is explained that the program is used fully, and friendly using for education and analysis of the industrial problems from input information. The main disadvantage is that reliability may not be possible and the result can be less accurate compared to the mathematics model's analysis due to random numbers in the program at that time.

Kanse and Patil [6] investigate the effect of discrete plant layout on production capacity and throughput time in the manufacturing process by using the simulation software FlexSim to generate scenarios interaction that can increase productivity and arrange a new plant layout for four products from the scope of this research. The project is about manufacturing plant layout optimization using simulation. This research shows different results. The four-component results compare with the existing layout. Finally, this paper shows the tables of four production quantity per month increases means the process time decreases and productivity increases.

Fidan et al. [7] re-designing the layout to improve production capacity and help with the disorganized material flow in the production process. Name Utilization of FlexSim Software to Identify the Suitable Layout Planning of Production Line. The researchers experimented with the simulation software FlexSim to increase capacity and re-designing the layout. the researcher collects all data need to simulate in software and generates 2 types of scenarios compared with the actual layout. The result of an experiment is improved production capacity from 11 units per day to 17 units per day which mean this paper can reveal throughput capacity is increased by 64.7%. Calculation from actual data they are collecting.

Koster et al. [8] published a special issue at the production research name Warehouse Design and Management. This paper pays attention to new technologies and methods. How do they impact inventory design and management? In particular, the e-commerce impact on the warehouse operation. This article approach allows for a thorough search of the design area and identifying candidate designs for consideration by designers. The design approach is demonstrated by designing the space picked forward. This paper offers a way to support hierarchical design decisions for warehouses. It considers the separation of design problems into a series of subproblems and use a formal model of the system to integrate solutions to these sub-problems. The conclusion of this paper shows warehousing research doesn't just focus on traditional topics such as improving efficiency in simple low-level picking situations or classic automation but on new automation technology more and more attention is being paid to such as high-density storage systems and AVS/RS systems or other factors (such as the human factor) that improve warehouse efficiency.

III. Methodology

First, we studied the existing process of material distribution like the process in and out of the area within the warehouse. Second, determine the problem that disturbs the work in process. third, solve the problem with simulation software. Finally, calculate the cost which might be reduced if implementing those solutions.

A. *Actual Working*

For case a study the large distribution area of 6,400 square meters stores 6 types of house construction materials as common types of cement, floor tiles, roof tile, steel, wood, sanitary, and doors in the same type. The case study in this research is the roof tile for house instruction which has a spacious area of 40 meters long and 16 meters wide. The process of picking the product up and the route of transport of the products.

The products are stored within the warehouse in a random area. All the materials have labels to identify the products, although the products do not clarify the area spot to store.

The number of products was no. 1-6 shown as Fig. 2 the type of products requirement is no. 1, 2, 4. The product was laid on the pallet on the floor storage.

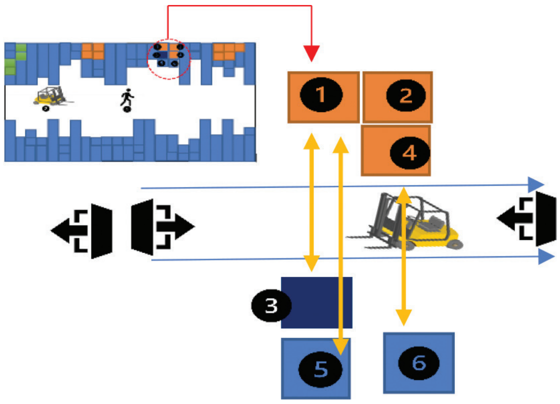


Fig. 2. How to pick up the goods

The routing for picking up product no.1-6 as shown in Fig. 2, operator no. 2 with transporter shows up with the transporter and then brings it to the exit way. However, operator no. 2 must carry pallets no. 3, 5, and 6 out of the way before picking up the required materials and then lift the pallets no. 3, 5, and 6 back to the flour storage when operator no. 2 carries pallets no. 1, 2, and 4 out already.

The product shown in Fig. 2 is the components overlapping on a pallet. The quantity input for generating the scenarios will count on a pallet for convenience and realism.

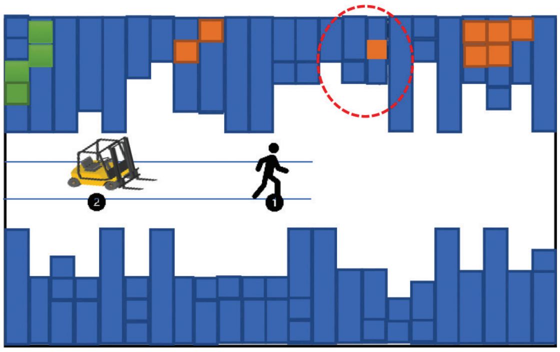


Fig. 3. Actual layout

As illustrated in Fig. 3 the actual route process for picking up the products from one order. First, operator no. 1 searches for the product requirement, and when finding the target at blocks no. 1, 2, and 4.

The actual routing process flow as shown in Fig. 3 the causing delay and waiting time that the operator must carry 3 products out to pick up the material requirement and the transporter returns to this area to carry the pallet back to the blocks again.

Because of random store products, the mean capital costs including labor and fuel are higher. This research will organize the new layout spot to facilitate transportation. And arranging the 3 types of product placement.

The researcher determines 5 orders for study in this paper and collects the working time to pick up the product following each order.

B. Determine the Data

The simulation model needs the data to generate the scenario and then collect the output analysis to show the best solution for this problem.

This research defines a Forklift as being a transporter and collecting the work time for comparison. Max speed of the transporter is 10 kg/hr. The data in Table I shows the input data has arrival time and type of products that relate to the quantity of the products. The input data for generating the model are the same for this paper.

TABLE I
INPUT TIME OF EACH QUANTITY

Arrival time (Sec)	Item Name Product	Quantity (pallet)	Type
1000	Product	120	1
2000	Product	80	2
3000	Product	80	3
4000	Product	120	1
5000	Product	60	2

* Arrival time is the time for the product appears on the model.
* Item Name Product is the identified name for the program.
* Quantity is the quantity of the product that must be input into the system model.
* Type is the classification of products that are defined by the simulation.

From Table I the data show input for generating the scenarios and quantity for the model run simulation. The quantity of products is 460 pallets, determine the type of each product is 1, 2, and 3 for classes A, B, and C as prioritized.

The next step in facilitating transport is classifying the ranks of the products and determining product along with sales valuable.

TABLE II
ABC CLASSIFICATION

Products Name	% Value	% Quantity	Class
Roof-Tile Sized 1	56%	8%	A
Roof-Tile Sized 2	26%	9%	A
Roof-Tile Sized 3	15%	32%	B
Roof-Tile small 4	3%	51%	C

* Product names are finished goods to study in this paper that are stored in the storage area,
* % Value is the percent total of the sale value.
* % Quantity is the total quantity for study.
* Class for identifying classification

As shown in Table II. ABC analysis technique distinguishes 3 types of products. A-class has roof-tile-sized no. 1 and size no. 2 for 82 percent of the total value. The red area in storage is classified as product class A. B-class represents the green area for 15 percent of the total value. The last 3 percent of the total value is a blue area was C-class.

The input data is the percent of the quantity that can conclude the product color is A-class is red for 17 percent, B-class is green for 32 percent, and C-class is blue for 51 percent at shown in Table II.

C. Design New Plant Layout

Design a new plant layout that facilitates transportation. The actual layout doesn't have space for a transporter. The researcher designs the new floor storage

The illustration is shown in Fig. 4. New layout is a new placement area designed for goods and adds more blank areas for transport. The material is stored by sitting back-to-back. The transporter drives to an area at each front.

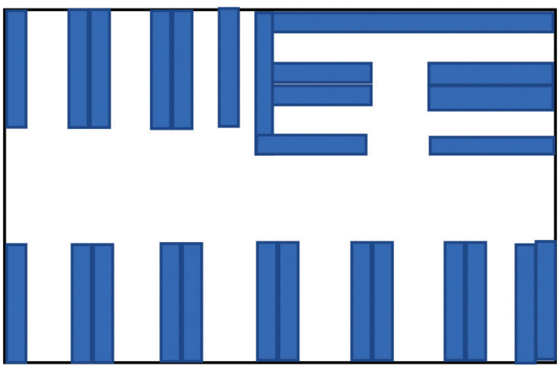


Fig. 4. New layout

This paper generates 3 models with the same data input and plant layout, but different positioning of the product based on ABC analysis for 2 models.

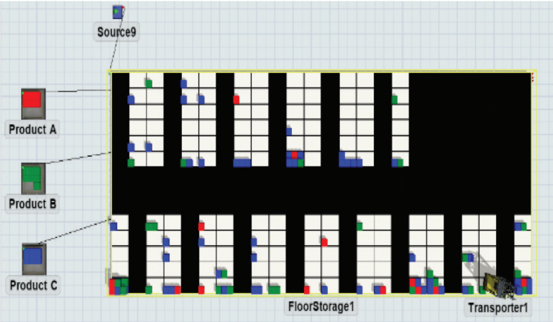


Fig. 5. Random location

From the illustration as shown in Fig. 5 random location. The simulation occupies a product with a random space area. The product appears on the layout and the Forklift will transport the product into Queue objects at specific goods.

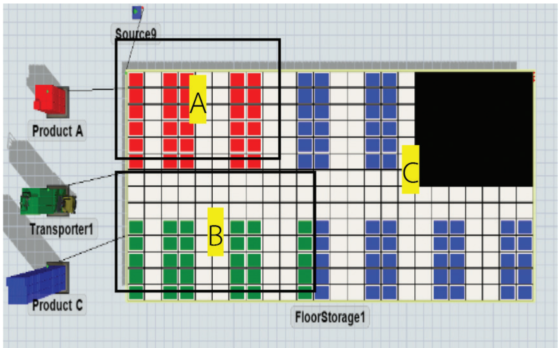


Fig. 6. Group class location

Second, from the illustration, as shown in Fig. 6. Group class location. the position of the product is based on ABC analysis data. The A-class defines to be the red area near each other. The green area is class B the position opposite of A-class and the rest of the area is for C-class.

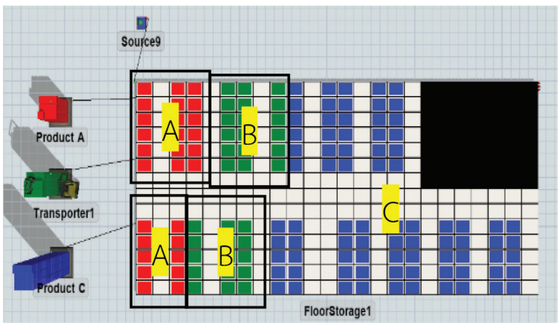


Fig. 7. Distribution group location

From the illustration as shown in Fig. 7. Distribution group location. The good's location is different from layout 2. The classification A moves opposite each other, and the transporter will notice from both sides and do the same for B and C class locations.

IV. Results

Comparison

Three scenarios simulation has runtime output when finished working. The results of forklift runtime show the different working times as shown in Table 3. The runtime of each model.

TABLE III
THE RUNTIME OF EACH LAYOUT

Models	Run Time (Sec)
Random location	10039.01
Group class location	9274.78
Distribution class location	8920.49

* Models that generate different goods locations to simulate the working time
* Run Time is the working time when input to the system until out of the system.

From the data, in Table III the runtime of each layout shows the working time that starts from the transporter picking up products at floor storage and transporting them to the customer.

The random location means the 3 products classification random location to storing and has 10039.01 seconds for finish work based on the actual workflow.

The group class location means the products have been grouped and stored in the area on the same side. The result from the analysis output of the Forklift is 9274.78 seconds for the finishing process.

The distribution class location means the location of the product has a separate group for the storing, but the location has near each other. The result of the runtime is 8920.49 seconds.

The performance efficiency after improvement can calculate from the working time before and after improvement. Comparison of the working time before and after improvement as follows:

The formula for efficiency performance is

$$\frac{\text{Actual runtime} - \text{runtime after improving}}{\text{Actual runtime}} \times 100$$

A comparison of the performance efficiency between the random location model and group class location is as the follows:

$$= \frac{10039.01 - 9274.78}{10039.01} \times 100$$

$$= 7.61\% \qquad (1)$$

And the working time of the random location model compared with a working time of distribution class location is as follows:

$$= \frac{10039.01 - 8920}{10039.01} \times 100$$

$$= 11.15\% \qquad (2)$$

From percent performance at (1) and (2) shows the efficiency after improvement. The different goods location of the two models has more efficiency compared to the random location.

V. Conclusion

In the simulation output, we concluded the runtime or working time that the transporter picks products for the customer at the queue object. The large area and massive products have been laid on the pallet. The simulation software gives information about the effect of different layouts and modifications on a good's location. However, it shows statistics of reduced runtime in comparison with different simulation models.

TABLE IV
Comparison with Random

Models	Group (%)	Distribution (%)	Variance (%)
Random	7.61	11.15	3.54

* Model for prototype to compare with the other
* Group is the model random location.
* Distribution is model name distribution class location.

Table IV shows the grouping class model and distribution model compared with the random location model. The variance of the two models is 3.54%. The best solution for this problem is distribution group location which improves performance by more than 11.15 percent.

However, the solution from the idea to solve this problem came from the working time based on the software environment and actual data of the warehouse in the case study.

The reliability of the information of the existing plan model on software is familiar from the real situation but doesn't guarantee the result was the fact when adjusting the solution into the real works as depending on various factors.

References

[1] FlexSim. (2022, Jan. 18). *FlexSim Express*. [Online]. Available: https://account.flexsim.com /flexsim-express/

[2] Siemens. (2021, Jan. 20). *Siemens Engineering Innovation*. [Online]. Available: https://www.plm. automation.siemens.com/global/en/products/simcenter/

[3] Arena Simulation Software. (2022, Jan. 20). *Rockwell Automation*. [Online]. Available: https://shorturl.asia/d5KyY

[4] X. Zhu, R. Zhang, F. Chu et al., "A Flexsim-based Optimization for the Operation Process of Cold-Chain Logistics Distribution Centre," *Journal of Applied Research and Technology*, vol. 12, no. 2, pp. 270-278, Apr. 2014.

[5] E. Tokgöz, "Industrial Engineering and Simulation Experience Using Flexsim Software," *Computers in Education Journal*, vol. 8, no. 4, Dec. 2017.

[6] B. Kanse and T. Patil, "Manufacturing Plant Layout Optimization Using Simulation," *International Journal of Innovative Science and Research Technology*, vol. 5, no. 10, Oct. 2020.

[7] F. Smaili, H. Ibishi, X. Mehmeti et al. (2020, Oct. 31). *Utilization of FlexSim Software to Identify the Suitable Layout Planning of Production Line*. [Online]. Available: https://shorturl.asia/DMbtC

[8] R. Koster, A. Johnson, and D. Roy, "Warehouse Design and Management," *International Journal of Production Research*, vol. 55, no. 21, pp. 6327-6330, Jul. 2017.

Pakdee Jaisue received industrial engineering from Panyapiwat Institute of Management, Thailand, in August 2017 and currently studying for a master's degree in the Faculty of Engineering and Technology. Used to work as an intern in engineering at the corrugated paper boxes factory, in 2017-2018. The current job is as a lecturer in the industrial engineering faculty at Panyapiwat Institute of Management (EEC), Chonburi, Thailand.

Paitoon Siri-O-Ran received a industrial engineering doctorate from Kasetsart University, Thailand, in 1994. Used to be a professor in industrial engineering, from 2004-2011 at Southeast Asia University, Thailand, and The current job is Head of the industrial engineering faculty at Panyapiwat Institute of Management, Nonthaburi, Thailand.

PAPER FORMAT (IEEE Style)

I. Format

- Your paper must use a paper size corresponding to A 4 which is 210 mm (8.27 inch) Wide and 297 mm (11.69 inch)
- Your paper must be in two column format
- Articles not more than 15 pages in length, single-sided A4 paper, margins (top, bottom, left, right) are 1 inch (2.54 cm)
- Abstract and References and content set to double columns,
- English font is Times New Roman, as follows:

TABLE I
Font Sizes for Papers

Content	Font Size	Labelling
Title (Single column)	18 (CT)	bold
Authors (Single column)	11 (CT)	bold
Authors Information (Single column)	10 (CT)	regular
Abstract	10 (LRJ)	bold
Index Terms (Keywords)	10 (LRJ)	bold
Content	10 (LRJ))	regular
Heading1	10 (CT)	regular (Capitalization)
Heading 2	10 (LJ)	regular
Table Title (Place above the Table)	8 (CT)	regular
Table content	8 (CT)	regular
Figure caption (Place below the figure)	8 (LJ)	regular
Reference Head	10 (CT)	regular (Capitalization)
Reference	8 (LJ)	regular
Author Profiles	10 (LRJ)	bold author name/ profile regular

CT=Centre Text, LJ=Left Justified, RJ=Right Justified, LRJ=Left & Right Justified

II. Composition of the Article

A. Article title

B. Authors information, Write (all) the author's name, affiliation, department, city, country and E-mail (set to Single Column) all.

C. Abstract, Must be under 200 words and not include subheadings or citations. Define all symbols used in the abstract. Do not delete the blank line immediately above the abstract.

D. Index Terms, Enter key words or phrases in alphabetical order, separated by commas.

E. Content

1) Academic article, should include: Introduction, Content, and Conclusion.

2) Research article, should include: introduction, literature review, Materials methods, Results, Discussion, and conclusion.

Clearly summarize the important findings of the paper. It should contain such as objectives, methods and major results.

F. Introduction

The Introduction section of reference text expands on the background of the work (some overlap with the Abstract is acceptable). The introduction should not include subheadings.

G. Pictures, table, etc., Must be use in numerical order in the article, provided the source correctly, cannot use other people' copyright.

Chart should be colored contrastingly or in black and white.

H. Reference

1) Cited in the main text. Indicate the number in the [ ] mark at the end of the text or the name of the referring person. Let the numbers be in the same line of content as [1].

2) Cited after the article. Put all bibliographical reference after articles, and order according to the author's name, please refer IEEE format. The footer reference format is as follows.

III. Rererences

References in research articles and scholarly articles. For academic and research journals, INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET). The technology defines referrals according to the IEEE format. All references should be listed at the end of the paper using the following.

Basic format for books:

J. K. Author, "Title of chapter in the book," in *Title of His Published Book,* xth ed. City of Publisher, Country if not USA: Abbrev. of Publisher, year, ch. x, sec. x, pp. xxx–xxx.

Examples:

[1] G. O. Young, "Synthetic structure of industrial plastics," in *Plastics,* 2nd ed., vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15-64.

[2] W.-K. Chen, *Linear Networks and Systems.* Belmont, CA: Wadsworth, 1993, pp. 123-135.

Basic format for periodicals:

J. K. Author, "Name of paper," *Abbrev. Title of Periodical*, vol. x, no. x, pp. xxx-xxx, Abbrev. Month. year.

Examples:

[3] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," *IEEE Trans. Electron Devices*, vol. ED-11, no. 1, pp. 34-39, Jan. 1959.

[4] E. P. Wigner, "Theory of traveling-wave optical laser," *Phys. Rev.*, vol. 134, pp. A635-A646, Dec. 1965.

[5] E. H. Miller, "A note on reflector arrays," *IEEE Trans. Antennas Propagat.*, to be published.

Basic format for reports:

J. K. Author, "Title of report," Abbrev. Name of Co., City of Co., Abbrev. State, Rep. xxx, year.

Examples:

[6] E. E. Reber, R. L. Michell, and C. J. Carter, "Oxygen absorption in the earth's atmosphere," Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1988.

[7] J. H. Davis and J. R. Cogdell, "Calibration program for the 16-foot antenna," Elect. Eng. Res. Lab., Univ. Texas, Austin, Tech. Memo. NGL-006-69-3, Nov. 15, 1987.

Basic format for handbooks:

Name of Manual/Handbook, x ed., Abbrev. Name of Co., City of Co., Abbrev. State, year, pp. xxx-xxx.

Examples:

[8] *Transmission Systems for Communications*, 3rd ed., Western Electric Co., Winston-Salem, NC, 1985, pp. 44-60.

[9] *Motorola Semiconductor Data Manual*, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.

Basic format for books (when available online):

Author. (year, month day). *Title.* (edition) [Type of medium]. *volume (issue).* Available: site/path/file

Example:

[10] J. Jones. (1991, May 10). *Networks.* (2nd ed.) [Online]. Available: http://www.atm.com

Basic format for journals (when available online):

Author. (year, month). Title. *Journal.* [Type of medium]. *volume (issue),* pages. Available: site/path/file

Example:

[11] R. J. Vidmar. (1992, Aug.). On the use of atmospheric plasmas as electromagnetic reflectors. *IEEE Trans. Plasma Sci.* [Online]. *21(3),* pp. 876-880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar

Basic format for papers presented at conferences (when available online):

Author. (year, month). Title. Presented at Conference title. [Type of Medium]. Available: site/path/file

Example:

[12] PROCESS Corp., MA. Intranets: Internet technologies deployed behind the firewall for corporate productivity. Presented at INET96 Annual Meeting. [Online]. Available: http://home.process.com/Intranets/wp2.htp

Basic format for reports and handbooks (when available online):

Author. (year, month). Title. Comp any . City, State or Country. [Type of Medium]. Available: site/path/file

Example:

[13] S. L. Talleen. (1996, Apr.). The Intranet Archi-tecture: M anaging information in the new paradigm. Amdahl Corp., CA. [Online]. Available:http://www.amdahl.com/doc/products/bsg/intra/infra/html

Basic format for computer programs and electronic documents (when available online):

ISO recommends that capitalization follow the accepted practice for the language or script in which the information is given.

Example:

[14] A. Harriman. (1993, June). Compendium of genealogical software. *Humanist.* [Online]. Available e-mail: HUMANIST @NYVM.ORG Message: get GENEALOGY REPORT

Basic format for patents (when available online):

Name of the invention, by inventor's name. (year, month day). *Patent Number* [Type of medium]. Available: site/path/file

Example:

[15] Musical toothbrush with adjustable neck and mirror, by L.M.R. Brooks. (1992, May 19). *Patent D 326 189* [Online]. Available: NEXIS Library: LEXPAT File: DESIGN

Basic format for conference proceedings (published):

J. K. Author, "Title of paper," in *Abbreviated Name of Conf.*, City of Conf., Abbrev. State (if given), year, pp. xxxxxx.

Example:

[16] D. B. Payne and J. R. Stern, "Wavelength-switched pas- sively coupled single-mode optical network," in *Proc. IOOC-ECOC,* 1985, pp. 585-590.

Example for papers presented at conferences (unpublished):

[17] D. Ebehard and E. Voges, "Digital single sideband detection for interferometric sensors," presented at the 2nd Int. Conf. Optical Fiber Sensors, Stuttgart, Germany, Jan. 2-5, 1984.

Basic format for patents:

J. K. Author, "Title of patent," U.S. Patent x xxx xxx, Abbrev. Month. day, year.

Example:

[18] G. Brandli and M. Dick, "Alternating current fed power supply," U.S. Patent 4 084 217, Nov. 4, 1978.

Basic format for theses (M.S.) and dissertations (Ph.D.):

J. K. Author, "Title of thesis," M.S. thesis, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

J. K. Author, "Title of dissertation," Ph.D. dissertation, Abbrev. Dept., Abbrev. Univ., City of Univ., Abbrev. State, year.

Examples:

[19] J. O. Williams, "Narrow-band analyzer," Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, 1993.

[20] N. Kawasaki, "Parametric study of thermal and chemical nonequilibrium nozzle flow," M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.

Basic format for the most common types of unpublished references:

J. K. Author, private communication, Abbrev. Month, year.

J. K. Author, "Title of paper," unpublished.

J. K. Author, "Title of paper," to be published.

Examples:

[21] A. Harrison, private communication, May 1995.

[22] B. Smith, "An approach to graphs of linear forms," unpublished.

[23] A. Brahms, "Representation error for real numbers in binary computer arithmetic," IEEE Computer Group Repository, Paper R-67-85.

Basic format for standards:

Title of Standard, Standard number, date.

Examples:

[24] IEEE Criteria for Class IE Electric Systems, IEEE Standard 308, 1969.

[25] Letter Symbols for Quantities, ANSI Standard Y10.5-1968.

First A. Author and the other authors may include biographies at the end of regular papers. Biographies are often not included in conference related papers. The first paragraph may contain a place and/or date of birth (list place, then date). Next, the author's educational background is listed. The degrees should be listed with type of degree in what field, which institution, city, state, and country, and year the degree was earned. The author's major field of study should be lower-cased.

The second paragraph uses the pronoun of the person (he or she) and not the author's last name. It lists military and work experience, including summer and fellowship jobs. Job titles are capitalized. The current job must have a location; previous positions may be listed without one. Information concerning previous publications may be included. Try not to list more than three books or published articles. The format for listing publishers of a book within the biography is: title of book (city, state: publisher name, year) similar to a reference. Current and previous research interests end the paragraph.

The third paragraph begins with the author's title and last name (e.g., Dr. Smith, Prof. Jones, Mr. Kajor, Ms. Hunter). List any memberships in professional societies. Finally, list any awards and work for committees and publications. If a photograph is provided, the biography will be indented around it. The photograph is placed at the top left of the biography, and should be of good quality, professional-looking, and black and white (see above example). Personal hobbies will be deleted from the biography. Following are two examples of an author's biography.

Second B. Author was born in Greenwich Village, New York City, in 1977. He received the B.S. and M.S. degrees in aerospace engineering from the University of Virginia, Charlottesville, in 2001 and the Ph.D. degree in mechanical engineering from Drexel University, Philadelphia, PA, in 2008. From 2001 to 2004, he was a Research Assistant with the Princeton Plasma Physics Laboratory. Since 2009, he has been an Assistant Professor with the Mechanical Engineering Department, Texas A&M University, College Station. He is the author of three books, more than 150 articles, and more than 70 inventions. His research interests include high-pressure and high-density nonthermal plasma discharge processes and applications, microscale plasma discharges, discharges in liquids, spectroscopic diagnostics, plasma propulsion, and innovation plasma applications. He is an Associate Editor of the journal *Earth*, *Moon*, *Planets*, and holds two patents.

Mr. Author was a recipient of the International Association of Geomagnetism and Aeronomy Young Scientist Award for Excellence in 2008, the IEEE Electromagnetic Compatibility Society Best Symposium Paper Award in 2011, and the American Geophysical Union Outstanding Student Paper Award in Fall 2005.

Third C. Author received the B.S. degree in mechanical engineering from National Chung Cheng University, Chiayi, Taiwan, in 2004 and the M.S. degree in mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 2006. He is currently pursuing the Ph.D. degree in mechanical engineering at Texas A&M University, College Station.

From 2008 to 2009, he was a Research Assistant with the Institute of Physics, Academia Sinica, Tapei, Taiwan. His research interest includes the development of surface processing and biological/ medical treatment techniques using nonthermal atmospheric pressure plasmas, fundamental study of plasma sources, and fabrication of micro- or nanostructured surfaces.

Mr. Author's awards and honors include the Frew Fellowship (Australian Academy of Science), the I. I. Rabi Prize (APS), the European Frequency and Time Forum Award, the Carl Zeiss Research Award, the William F. Meggers Award and the Adolph Lomb Medal (OSA).

Remark: More detail information, Pleases read Preparation of Papers for INTERNATIONAL SCIENTIFIC JOURNAL OF ENGINEERING AND TECHNOLOGY (ISJET), https://ph02.tci-thaijo.org/index.php/isjet/index

Panyapiwat Institute of Management (PIM)
85/1 Moo 2, Chaengwattana Rd,
Bang Talat, Pakkred, Nonthaburi 11120, Thailand
Tel. +66 2855 1560
https://www.tci-thaijo.org/index.php/isjet/index
https://isjet.pim.ac.th
E-mail: isjet@pim.ac.th