Design and development of a vision system for identifying a workpiece quality and defects using object detection

Traisirin Saengseangfa and Suradet Tantrairatn

School of Mechatronics Engineering, Institute of Engineering

Suranaree University of Technology, 111 university Avenue, Muang District, Nakhon Ratchasima, Thailand

Email: traisirin@gmail.com, suradetj@sut.ac.th,

***Abstract*— In automotive parts manufacture, the quality of the products is the most important issue for a manufacturer. This research presents the design and development of an intelligent vision module classifying the defects on object surfaces using image processing, artificial intelligence, convolutional neural network algorithm as if using human vision. The developed prototype was a neural network-supported computing system for classifying the defects on the surface of black car jacks of 10x10x19 cm (W x L x H) in an environment with specific brightness. The test result was the average precision of 95.6%, showing that it could detect defects on the four types of surfaces previously specified in the Deep Learning process. It could also detect more than one defects on the same object on the belt conveyor system in real time without underfitting or overfitting problem.**

Keywords—Image Processing, Artificial Intelligence, Convolution Neural Network, Deep Learning

I. INTRODUCTION

Thailand depends largely on exports and domestic use in automotive industry [1], aerospace industry, shipyard industry, electrical appliance and electronics industry, etc. In those industry, product quality control process is, without doubt, required. For example, the spray-painting process, as well as many other coating methods [2] depending on the types of the manufactured, is widely performed to enhance the durability of workpieces. Then, there comes the next process, quality control, to inspect conditions of the outer part of the workpieces, whether there are scratches, chafing, bumps and cracks or not.

Quality control [3] inspecting products according to the specified features is an essential process every industry has to go through before products are delivered to consumers. To improve this process's efficiency, technologies to process images is regarded to be useful as, currently, there are algorithms developed in accordance with the development of the computer industry.

In addition, Artificial intelligence (AI) is considered to be suitable to perform analysis of objects located at any places as instructed in place of human vision. Hence, in this research, a prototype able to detect the defects on objects has been developed. It was equipped with NVDIA Jetson Nano device supporting neural networks to classify the detected defects into four types according to previously input data; it was also equipped with a microcontroller to send signals indicating errors detected on the surface of objects.

II. TEHORY AND METHOD

A. Algorithm

• Image Processing

Image processing [4] is photograph manipulation with computers to improve or change the quality of images so that they will be more suitable for further analysis and other purposes. Images are, firstly, retrieved (Image Acquisition), compressed (Image Compression) and enhanced (Image Enhancement) by adjusting the brightness and the color sharpness of each pixel to make them suitable for Deep Learning process. It is commonly used in the process of inspection and quality control in every industry to help reducing operation time and increase efficiency as well as reliability of inspection which is useful for industries requiring high accuracy.

• Artificial Intelligence

Artificial Intelligence [5] is a computing system modeled upon neural networks in human brains. It can learn and optimize processing upon the increasing amount of data through a self-learning process consisting of memorizing, thinking, analyzing, learning and linking complex information at fast speed. It collects large amount of data at high speed and processes them repeatedly through intelligent processing steps.

• Deep Learning

Deep learning [6] is a computer technology developed to be able to imitate human behaviors. It has a computational-thinking process called a neural network (NN) as it works like humans' neural networks. The specific characteristic of deep learning is that, when applied to image processing to classify the defects on target objects, it can simulate conditions and find answers by piling hidden layers, multiple layers of neural networks, to more than two layers to achieve learning and simulation. Therefore, it can be concluded that the more layers of the neural networks there are in processing, the more accurate deep learning can become.

• Convolutional Neural Network

Convolutional Neural Network is one of the deep learning methods that simulates human vision. It is able to differentiate the features of objects such as colors, borders, color contrast and keep them within themselves. As a result, it can recognize what those features are and classify them. Convolutional neural network uses aforementioned hidden layers [7] to identify features and iterates until the answers with accuracy according to the classification come out. Convolutional neural network analyzes the relationship between the extracted features in each layer and the outcomes. How it works is divided into 3 parts as follows. (Fig. 1)

1) Inputs are the parts where data or objects are input; it functions as human vision.

2) Hidden layers are the parts where processing is performed into layers to do training and classification; it functions as human brains.

3) Outputs are the parts that show the results of feature classification resulting from using multiple hidden layers analyzing each image until the answer is output.

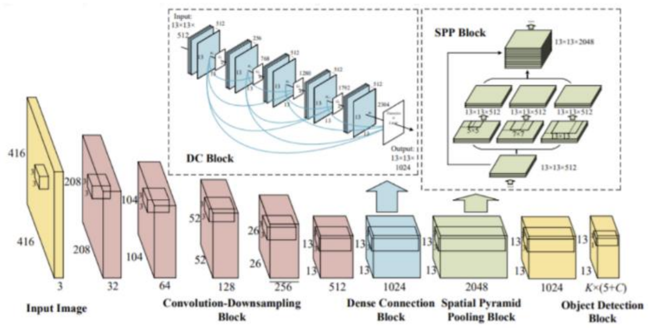


Fig. 1. Convolutional Neural Network

B. *Application of Yolov5 algorithm*

Concerning the YOLO (You Only Look Once) real-time object detection algorithm, which is one of the most effective object detection algorithms that also encompasses many of the most innovative ideas coming out of the computer vision research community. Object detection is a critical capability of autonomous vehicle technology. It is an area of computer vision that is exploding and working so much better than just a few years ago. At the end of this article, we will see a couple of recent updates to YOLO by the original researchers of this important technique. YOLO uses a totally different approach. YOLO is a clever convolutional neural network (CNN) for doing object detection in real-time. The algorithm applies a single neural network to the full image, and then divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities [8]. Therefore, we have decided to use Yolo version 5 to apply with this thesis.

YOLO version 5 algorithm is chosen to apply this thesis because of the following reasons:

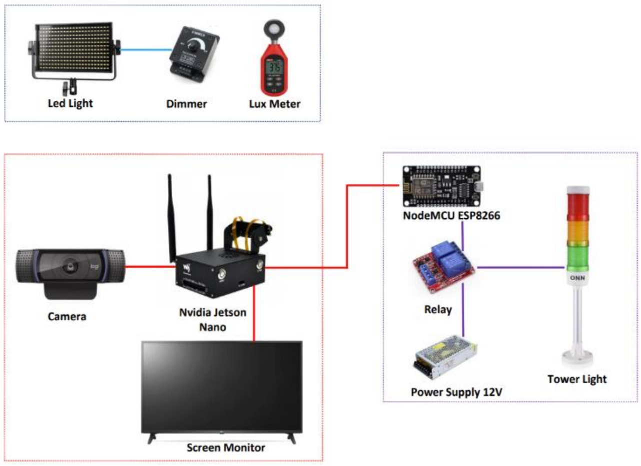
- Speed: This algorithm improves the speed of detection because it can predict objects in real-time.

- High accuracy: YOLO is a predictive technique that provides accurate results with minimal background errors.

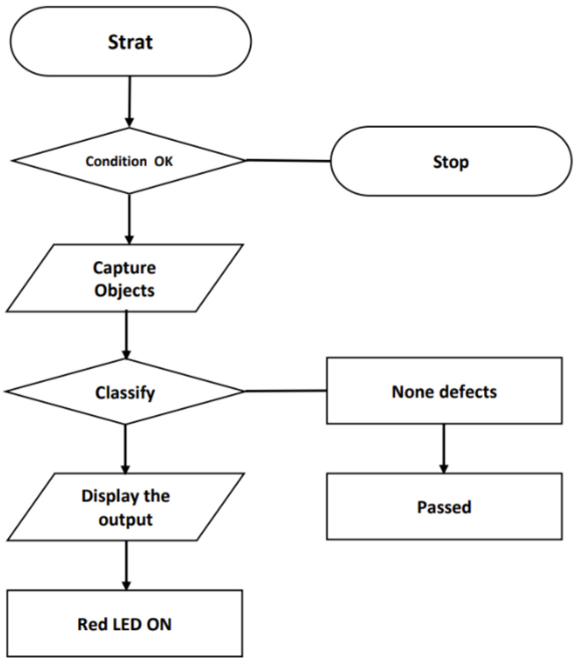
- Learning capabilities: The algorithm has excellent learning capabilities that enable it to learn the representations of objects and apply them in object detection [9].

C. *Processing system and microcontroller*

The prototype was equipped with the Nvidia Jetson Nano computer board with a 128-core GPU and Linux-based system for convolutional neural network algorithms. It received data from a 30-fps Logitech Full-HD camera for processing. A general monitor connected with the computer board with a HDMI cable was used to display the system operation. In addition, LED lights were installed along with an adjustable power supply device for optimal light output. When the main system detected a problem on the surface of an object, it would send commands in serial communication at a speed of 115,200 through a digital pin for the relay to supply current to the coil. As a result, the red light bulb would light up as shown in Fig. 2.



(a) Processing system of the prototype



(b) Working process

Fig. 2. The prototype system and working process

D. Installation of equipment

The Logitech camera was placed 25 centimeters far from the target objects. Above its left and right sides were LED lights whose brightness was fixed at 1,800 lux because of a dimmer supplied with power to convert 220v AC to 12-24v DC. A solid plastic structure of 45x60x35 centimeter (W x L x H) was placed to avoid reflection of light. There was a belt conveyor system with a 20x120x75-cm belt (W x L x H); it was equipped with a 60W motor using a speed controller as shown in Fig 3.

Fig. 3. Installation of equipment

E. Measuring Performance

The prototype's performance efficiency was measured according to Confusion Matrix [10]. The results were classified into four values as follows (Fig 4).

1) True Positive (TP): it refers to the number of predictions where the classifier correctly predicts the positive class as positive.

2) True Negative (TN): it refers to the number of predictions where the classifier correctly predicts the negative class as negative.

3) False Positive (FP): it refers to the number of predictions where the classifier incorrectly predicts the negative class as positive.

4) False Negative (FN): it refers to the number of predictions where the classifier incorrectly predicts the positive class as negative.

The above 4 values, TP, TN, FP, and FN, are used to measure the efficiency with Precision, the ratio of correct positive predictions to the total predicted positives.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

$$Precision = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Fig. 4. Confusion Matrix size 2 x 2 matrix [11]

F. Data Preparation

The data used to train the model deep learning had to be prepared first. They were categorized according to the types previously defined to make labels. The dataset consisted of a total of 1,060 images divided into 4 categories below:

1) Scratch with average size of 3x15 millimeters (W x L)

2) Chafing with average size 10x20 millimeters (W x L)

3) Bump with average size 1x1.5 millimeters (W x L)

4) Crack with average size 0.5x15 millimeters (W x L)

Every image went through a cleaning process (Clean Data) and pre-processing process to ensure their accuracy and suitability accordingly as shown in Fig. 5.

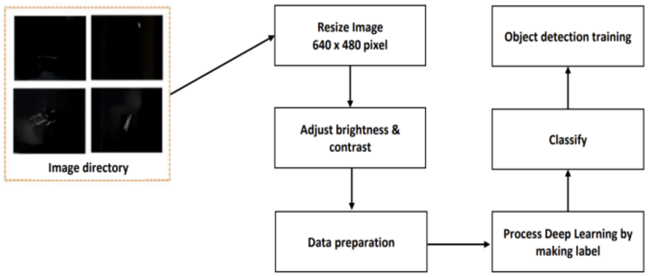


Fig. 5. Data Preparation

Moreover, every image had to be resized to 640 x 490 pixels to prevent processing failure due to difference in pixel sizes. Brightness and contrast of each also had to be adjusted [12] according to the equation below:

$$g(x,y) = \alpha f(x,y) + \beta$$

After that, a target location was set so that the system would learn what was there, leading to add metadata [13] to the images as well. Because this process affected the efficiency in identification and detection directly, a bounding box had to be created to determine the object range, neither bigger nor smaller than the target, maximizing accuracy and minimizing deviation as shown in Fig. 6.

1) p_c was probability for detected objects

2) b_w and b_h were the width and height of the bounding box

3) b_x,b_y was the center point of an object

4) C was a defined class.

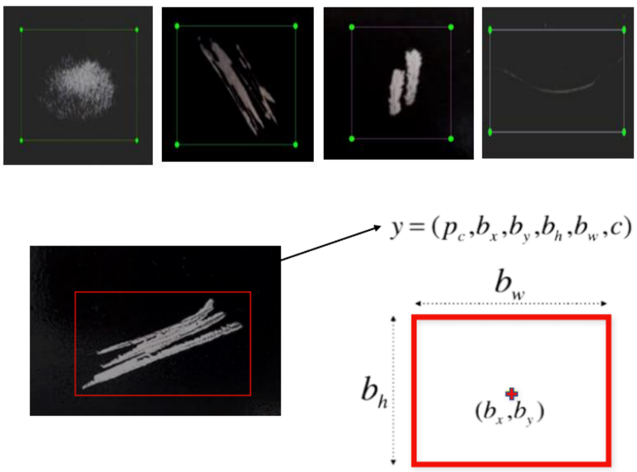


Fig. 6. Adding the Metadata (bounding box regression) [14]

Bounding box labels in each class were created for training. There were 1,060 label files in total, one label for one image as shown in TABLE I

TABLE I. AMOUNT LABELS EACH CLASS

No.	Classes	training	test	validation
1	scratch	165	55	45
2	chafing	165	55	45
3	bump	165	55	45
4	crack	165	55	45
	Total	660	220	180

In this process, the data had to be divided correctly to avoid underfitting, wrong outputs of the model, and overfitting, wrong outputs in actual uses despite being correct and precise during training. The latter greatly affects the accuracy of the model when the model's performance is tested with actual objects and the problems occur as shown in Fig. 7.

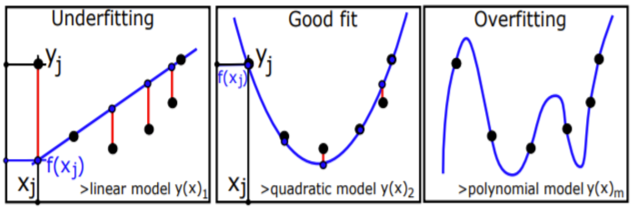


Fig. 7. Underfitting VS. Overfitting [15]

Among the dataset of 1,060 images in total, 660 images were used in the training for the model to learn the classes of different defect objects and 220 images to test the performance of the model. 180 images among those for training were used to test the performance after tuning had been finished.

Next, the already prepared dataset was input into the model for them to be processed. Unable to be input at once, they all were divided into small groups which was called a batch size. Suitable variables were specified to avoid the excessive usage of the CPU and RAM.

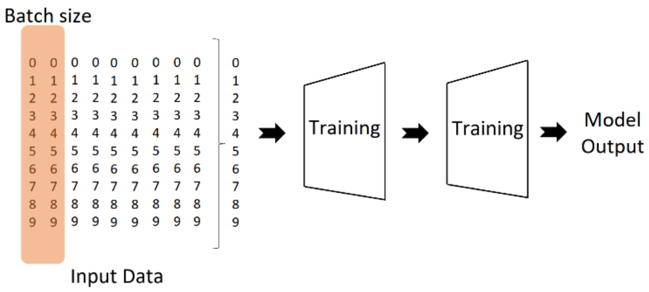


Fig. 8. Input the data to the model [16]

For Yolo's algorithm model, this research used Yolo version 5. The images' parameters were configured to have the 640-pixel size, the same as the images', to avoid problems during training. The batch sizes' parameters were set to 20, preventing excessive usage of CPU. The epoch parameters were set to 100; by training with 53 iterations to get one epoch, the model carried out 5,300 iterations in total as in Table II.

TABLE II. AMOUNT LABELS EACH CLASS

Setting	Parameter
Model	Yolov5S
Dataset	1,060
Image size	640
Batch size	20
Epochs	100 * 3

III. EXPERIMENT RESULTS

A. *The result of developing the model*

Python and some libraries such as Keras, Tensorflow, Pytorch, Matplotlib, OpenCV, Numpy, Pandas, Torchvision, Scipy, and Scikit-learn used for deep learning process were applied to develop the model detecting the defects on objects and classify classes. More details on the installation of the libraries are provided in [17]. The result of the classification of the four classes, which were scratches, chafing, bumps and cracks, resulted in a relatively high precision of 95.6 percent against confidence as shown in Fig. 9.

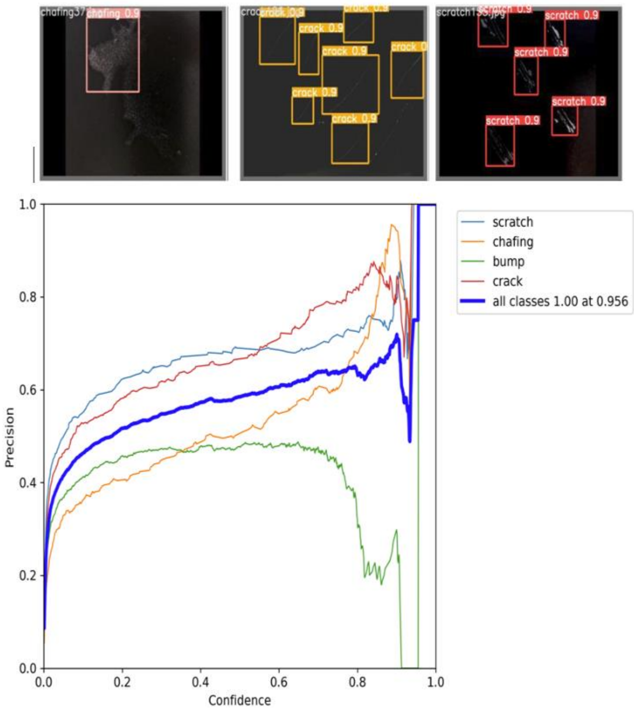


Fig. 9. Precision result from creating the model

According to the graph, the green line representing the Bump class had the lower precision range than the other classes because the bump had the same feature. To fix this, a previous feature was used for the next training and the number of iterations was increased.

B. Testing the model performance with the prepared samples

The model's performance was tested with previously prepared four types of samples which are scratches, chafing, bumps, and cracks with 100 samples each, meaning 400 in total, to see if it could classify them correctly. According to TABLE III showing the test result, the success rate of classification using convolutional neural network algorithm was 96%.

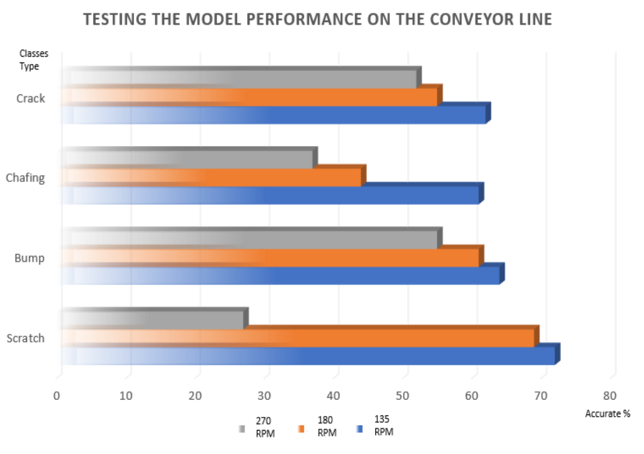
TABLE III. THE RESULT OF TESTING THE MODEL TO CLASSIFY EXAMPLES

No.	Class Type	Amount of the tested sample	Result	Accuracy (%)
1	scratch	100	98	98%
2	bump	100	94	94%
3	chafing	100	97	97%
4	crack	100	95	95%
	average			96%

C. Implementing the model in real-time testing on the conveyor system

The model's performance was tested on the conveyor system and actual car jacks in real time. Two variables, classified features and training data from the computer, were input into the NVIDIA Jetson Nano device with the installed libraries. Frequency of the speed controller of the conveyor's motor was adjusted to 50 Hz, controlling it to spin the belt at 3 speeds: 135 rpm, 180 rpm, and 270 rpm as shown in TABLE IV. The result shows that it could detect each defect on the objects which were running on the conveyor belt in real time, classifying them according to features of the defects correctly without underfitting or overfitting. Nevertheless, when the objects' moving speed increased, the camera's frame rate was affected, causing the accuracy to decrease accordingly.

TABLE IV. TESTING THE MODEL PERFORMANCE ON THE CONVEYOR LINE IN REAL-TIME



IV. EXPERIMENT RESULTS

This research to develop the prototype has integrated image processing algorithms, artificial intelligence performing deep learning and convolutional neural network algorithms to examine, analyze and classify the defects on objects. Firstly, dataset was collected from 4 types of samples, scratches; chafing; bumps; cracks. Then, convolutional neural network algorithms were applied to develop the ability of feature classification of the collected dataset which needed preparation using the principle from the related researches in order to avoid underfitting and overfitting when they were tested with actual objects.

the first test resulted in the average precision of all classes of 95.6%. The second one tested on various samples shows that it could classify the defects on the objects 96% correctly. Lastly, when the processing performance efficiency of the model in real time was tested on the belt conveyor system with a maximum speed of 270 rpm and on car jacks, it could detect and classify the defects on the objects accordingly. As a result, the three tests shows that this research on the design and development of an intelligent vision system for identifying a workpiece quality and defects using artificial intelligence can be an example of how technologies can be applied to encourage development and progress in various industries.

ACKNOWLEDGEMENT

The research would like to thank the King Mongkut's University of Technology North Bangkok and the Suranaree University of Technology for supporting the knowledge to successfully complete this research.

REFERENCES

[1] THAILAND AUTOMOTIVE INSTITUTE, "Statistics of Thai automotiveindustry.",[Online].Available: https://data.thaiauto.or.th/.[Accessed:18-August-2020]

[2] Dr.Panadda, "Surface Engineering.", Material and Surface Technology Damage Analysis Laboratory

[3] Narathip,"Quality Control.", Faculty of Technology, UDON THANI RAJABHAT UNIVERSITY,2014

[4] Akekachai Pannawan and Supattarachai Sudsawat, "Automated part inspection by image processing system in vehicle part manufacturing", ISSN 1512-7805 Printed in Thailand.

[5] Kwanchanok,"AI:Artificial Intelligence.",[Online].Available: https://www.lib.ku.ac.th/2019/index.php/covid-19/1045-artificialintelligence[Accessed:30-August-2020]

[6] Napparut Pattanasarn and Nattavut Sriwiboon, "Image Processing for Classifying the Quality of the Chok-Anan Mango by Simulating the Human Vision using Deep Learning.", Department of Informatics and Computer, Faculty of Science and Health Technology, Kalasin University,2020.

[7] Ashwin Bhandare, Maithili Bhide,Pranav Gokhale, Rohan Chandavarkar, "Applicatoin of Convolutional Neural Networks.",Department of Computer Engineering, Pune Institute of ComputerTechnology Savitribai Phule Pune University, Pune,2016.

[8] ODSC, "Overview of the Yolo Object Detection Algorithm.",[Online].Available: https://medium.com/@ODSC/overview-of-the-yolo-object-detection-algorithm-7b52a745d3e0[Accessed:28-September-2021]

[9] DATA, "CNN Yolo Algorithm.",[Online].Available: http://datahacker.rs/object-detection-yolo-algorithm/[Accessed:28-September-2021]

[10] Mark Harlan Goadrich, "Learning Ensembles of First-Order Clauses That Optimize Precision-Recall Curves.", UNIVERSITY OF WISCONSIN–MADISON,2007.

[11] Stardet,"Confusion Matrix.",[Online].Available: https://www.stardat.net/post/confusion-matrix[Accessed:30-April-2021]

[12] 305393 Special Topic in Human Computer Interaction, "How to adjust the Brightness and Contrast of the image by using OpenCV with Visual Studio.",[Online].Available: https://305393advancedtopicsinhci.wordpress.com/2017/02/15/how-to-adjust-the-brightness-and-contrast-of-the-image-by-using-opencv-with-visual-studio/[Accessed:20-April-2021]

[13] Sitthitouch, "Generating an architectural model using object recognition to observe human movement in public space case study:

3-dimensional parametric model of bus station platforms.", Ref. code: 25605816030075HOB,2018

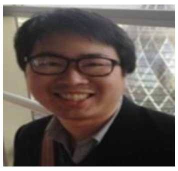
[14] Grace Karmini,"Introduction to Yolo Algorithm for object detection.",[Online]:Available: https://www.section.io/engineering-education/introduction-to-yolo-algorithm-for-object-detection/[Accessed:28-September-2021]

[15] L.Shao,A.Mahajan,T.Schreck, and D.J.Lehmann, "Interactive Regression Lens for Exploring Scatter Plots.", Eurographics Conference on Visualization (EuroVis) 2017

[16] Mr.PL,"Deep Learning EP.3 with Keras with Image Classification.",[Online].Available: https://medium.com/mmp-li/deep-learning-ep-3-keras-with-image-classification-2d5d95ee5613[Accessed:15-August-2021]

[17] AyushExel,"Install Libraly.",[Online].Available: https://github.com/ultralytics/yolov5.[Accessed:10-May-2021]

TRAISIRIN SAENGSEANGFA received his Bachelor's degree in Industrial Engineering Technology in 2018 from Kasem Bundit University, Thailand. His final project fucused on creating a vision system for identifying a workpiece quality and defects using Artificial Intelligence.



SURADET TANTRAIRAN is a lecturer at the School of Mechanical Engineering Suranaree University of Technology, Thailand. He received his Bachelor's degree and Master's degree in Aerospace Engineering in 2005 and 2007 from the Kasetsart University and a Ph.D. in Automtic Control and System Engineering in 2016 from University of Sheffield, UK. His research interested include Flight Dynamics Modelling, Flight Simulation, Aircraft System Identification or Estimation and Selection, Aircraft State Estimation, Flight Control, Adaptive Control, Fault Tolerance Flight Control, Aircraft Fault Detection and Isolation, Aircraft State Monitoring Intelligent Control, and Autonomous UAS.