

# Numerical Reader System for Digital Measurement Instruments Embedded Industrial Internet of Things

Natthanan Promsuk<sup>1</sup> and Attaphongse Taparugssanagorn<sup>2</sup>

<sup>1</sup>Department of Computer Engineering, Faculty of Engineering, Chiang Mai University, Chiang Mai, Thailand

<sup>2</sup>School of Engineering and Technology, ICT Department, Telecommunications, Asian Institute of Technology, Pathum Thani, Thailand

Email: nattanun.nn@gmail.com; attaphongset@ait.asia

**Abstract**—In industrial factories, many measuring instruments are used to display, for instance, pressure, voltage, temperature or humidity. Human errors are the main problem and often occur in many processes mostly done manually, such as data acquisition. Therefore, the problem of how we obtain such data automatically and correctly in real-time is important. In this paper, a numeral recognition system (NRS) is proposed based on an optical character recognition (OCR) method. The NRS embedded industrial Internet of things (IIoT) is used to serve a real-time service. Moreover, digital image processing (DIP) together with the multi-layer perceptron (MLP) is applied to efficiently recognize the numeral data. Furthermore, it is very common that the instruments' screens can face the rotation problem. This problem can be solved using the histogram of oriented gradients (HOG) and Hough transform (HT) techniques. In addition, realistic conditions under various noise types are considered such as salt and pepper (SP) noise, Gaussian noise, and Speckle noise. The system performances are evaluated in terms of confusion matrices and accuracies. The strong contribution of our proposed NRS system is that it works excellently in any situations and achieve up to 95.13 percent accuracy. From the actual experiments, we achieve an average about 95 percent accuracy. Although the NRS with the HOG and HT technique takes a bit longer computation time and more memory usage to process the images than another NRS, the system provides better results. Our proposed system is suitable for a real-time service due to low computation time.

**Index Terms**—Optical character recognition, Data acquisition, Industrial internet of things, Screen rotation, Industrial Automation, Multi-Layer Perceptron

## I. INTRODUCTION

In industrial factories, it is very common to see that measuring instruments have a digital display with seven-segment displays (SSDs). Nowadays those measuring instruments with different data formats are found several brands in the market. Therefore, it is challenging to make a system to produce output data in a real-time process. Their output data is often collected manually or by data acquisition (DAQ) cards which sometimes may not be available [1]. Also, the DAQ cards are usually costly. Therefore, a solution to replace the DAQ card is important, for instance, by using a low-cost camera, such as a webcam and a Digital Image Processing (DIP) unit

which can be managed at a lower cost. Moreover, it can avoid having the human eye errors and the unavailability of obtaining real-time data.

“Internet of everything, (IoE)” or “Internet of things, (IoT)” is an emerging concept of new technologies including various types of physical devices, vehicles, buildings, and other items embedded with electronics and defining how things are connected through the Internet [2]. The demand for IoT devices is continuously increased. Especially, in the industrial factories, there are several electronic devices connected with the Internet for transmitting data due to the recent trend of smart industrial automation [3]. Therefore, the concept of the industrial Internet of things (IIoT) are developed. The DIP has been used to solve the two following main problems [4]. The former is the error interpretation. The latter is the data processing for representation, transmission, and storage for autonomous machine perception [4]. Nowadays, many fields of applications, such as agriculture, medical, and industry, have made use of the DIP to process data from images. A neural network (NN) is a system having a structure created similar to the human brain. An NN model uses the human brain system as a reference model because the human brain has a remarkable capability to classify and recognize images, scenes, and objects very quickly irrespective of how complex they are [5]. An NN has been applied to various problems in different ways. For instance, an NN is built to recognize a sequence of handwritten digits [6]. As more examples, NNs have been used for weather prediction [7], oil exploration data analysis [8], facial recognition [9], and speech to text transcription [10].

In industrial factories, several measuring instruments usually use SSDs which are electronic displays for displaying decimal numerals [11]. The output data acquisition and collection from the SSDs cannot be done immediately. One has to wait for people to collect it, otherwise, it has to be processed with different syntax or different databases depending on instrument brands. Furthermore, the calibration process of measurements has been mostly done manually in process control industries, [11]. When we use people to collect the data, we also likely have the human eye error problem. Consequently, it can delay the overall reading process [1]. In addition, it is very common that the instruments' screens can face the

---

Manuscript received September 1, 2020; revised March 11, 2021.  
Corresponding author email: nattanun.nn@gmail.com.  
doi:10.12720/jcm.16.4.132-142

rotation problem. When the screens rotate, the conventional optical character recognition (OCR) based DIP techniques do not properly work.

## II. RELATED WORKS

The OCR techniques have been widely implemented for interpreting scanned images or printed text into machine-encoded text [12]. OCR technology has been implemented to improve and revolutionize the document management process. The OCR helps the people to extract the related information and records it automatically. Moreover, the OCR results are accurate and process adequate information in less time. There are many techniques such as template matching, feature extraction, and NN for character recognition. The normal number recognition fails to interpret the seven-segment number because of the unique appearance of seven-segment digits compared to other numerals. The existing works which applied some DIP were used for recognizing different objects in images. Their methods cannot be applied for seven-segment digits because of the discontinuity in the seven-segment digits. The authors of [11] proposed a system, which is able to detect the SSD but only in the same type of calculator. They separate each digit into twelve equal regions. Then, the system extracts important features from each region to form a feature vector and correlates the vector with the prestored feature vectors. Finally, the system tries to match the formed feature vector with one of the prestored ones. This system is able to work only on the fixed position of screens. Moreover, another limitation of this system is that it can only operate with the same size of the digit and the same character style. In [13], the authors show how to use OCR techniques under various noise environments to detect digit numbers in SSDs. They set the rules to classify the numbers. The performance of their system is improved by building the decision tree for digit recognition. The decision tree is used to identify the digits by deciding according to the percentage of the black pixel density, the width, and the height of the digits' images. However, the drawback of their system is high time-consuming due to several condition rules in the system. In [14], they also apply the OCR techniques to recognize the SSDs in digital energy meter. The maximally stable extremal regions (MSERs) are used as a method of blob detection in their image pre-processing to find the text regions in the grayscale images. It uses the different levels of the black and white color scales because the pixel values in the grayscale images are started from 0 to 255. The MSER can identify the group the pixel to the text regions when the value pixels change. However, their system is limited to just a specific style of SSDs. Moreover, their system can work effectively only with the high contrast image captured during nighttime. All of the works in [11], [13], and [14] similarly apply the feature extraction and OCR techniques, but still without an NN.

The application of the OCR with an NN for recognizing the SSDs is initially proposed in [1]. The procedure of using an NN is divided into the two following phases. Firstly, the NN is trained by the back-propagation model and scaled conjugate gradient. Secondly, the trained NN is used to classify the digits. The system is able to achieve high accuracy but only for the same shape and size of digits. Therefore, the system is not flexible enough for a real-life situation. Moreover, the different distances between the camera and display as well as the image skew or rotation can also be the other constraints of their system. The existing works in [1], [11], [13], and [14] which are related to the seven-segment digits could not perform effectively with different styles or formats of seven-segment digits, nonfixed camera range, and screen rotation situation. Moreover, our proposed NRS attains better accuracy compared to the existing methods.

In [15] and [16], they work on the license plate recognition and could achieve high accuracy but with some constraints, for instance, low-resolution images cannot work well and their algorithm only fits a specific number format in each country. The difference between SSD and license plate recognition systems is that the numbers in license plates are the same format in each country but the SSDs have various formats and patterns related to the equipment brands and measuring instrument types. Therefore, our work has to concern about the formats of the SSDs and the rotation screen problem, which can happen. In addition, we deal with the images taken from the low-cost webcams. Last but not least characters and numbers in a license plate and the SSD are different because the SSD has seven lines which are not connected. Therefore, under the rotation screen problem, the SSD recognition system faces a more challenging problem than the license plate recognition system.

## III. OUR CONTRIBUTIONS

The main contribution of this paper is to propose a numeral recognition system (NRS) system based on OCR techniques, NN, and IIoT, which can automatically and accurately detect the numbers in the SSDs. The Multi-Layer Perceptron (MLP) is applied together with the feed-forward artificial NN (FANN) because the advantage of the FANN helps the system easier to perform the data classification. The proposed system is created with the IIoT platforms. In addition, the histogram of oriented gradients (HOG) feature and the Hough transform (HT) techniques are used to solve the screen rotation problem in the SSD recognition. The proposed system is evaluated under different situations including the dead pixels and low-resolution noise which are modeled by Salt and pepper (SP) noise, Gaussian noise, and Speckle noise. There are normally many measuring instruments in a factory. And it can happen that the cameras capturing their displays are rotated from their centers by some reasons. To the best of our

knowledge, our work has novelty in managing the situation possibly-happing with the screen rotation problem and nonfixed camera ranges. The system from the others' works in the literature applying DIP could not collect the correct information from the rotated screens. In contrast, our system can recognize seven-segment digits even from the rotated screens (0 – 180 degrees). In addition, we have developed the NRS model based on the IIoT platform. Our system is also able to recognize any formats of seven-segment digits in nonfixed camera ranges, which is not possible in any other existing works in the literature. Table I summarizes the related works compared to our contributions. We mainly focus on [1], [11], [13], [14].

TABLE I: COMPARISON OF RELATED WORKS AND OUR CONTRIBUTIONS

Reference	Paper discussion	Our contributions
[1]	<p>Their proposed system uses the OCR techniques to detect a camera number in multi-meters with a specific range between the camera and SSD multi-meters. This system is implemented with an NN model.</p>	<p>If the distance between the meters changes, it cannot work due to the meters' shape and digits' size change. Therefore, we proposed the NRS which can effectively work even with the variation of distances and the digits' styles.</p>
[11]	<p>heir proposed system is automatic scanner for digital multi-meter and temperature meter. They used statistical feature extraction for rotation recognition of the digits. This work also can perform one range between camera and SSD.</p>	<p>heir system can work only for two types of meter. Their system cannot handle the effect of the rotation screen problem. In contrast, due to the applied NN model, HOG, and HT our NRS can well perform with various SSD types and can handle the screen rotation problem.</p>
[13]	<p>heir proposed system uses the OCR techniques to detect a number in SSDs. They used the classifying rules according to the DIP to recognize the digit numbers. This system does not apply an NN model.</p>	<p>It focuses only on one format of SSDs. Our system can work with various formats of SSD. Although [13] can also handle the rotation screen problem but only up to 90 degrees while our system can handle it up to 180 degrees.</p>
[14]	<p>T Their proposed system uses the OCR techniques without an NN model to recognize the digits. Their system focuses only on the digital energy screens' meters. They apply the MSER to find the text regions in the SSDs.</p>	<p>It focuses on only one type of digital meter. Their system can work effectively with the digital energy screens' images captured during nighttime. In our NRS, it can work with any type of screens and any styles of SSDs. Moreover, our NRS can handle the rotation angle up to 180 degrees.</p>

#### IV. DIP TECHNIQUES

In this section, the DIP techniques in our proposed NRS using the MLP model are discussed.

##### A. Feature Extraction

A feature extraction is an indispensable technique in the DIP. Different feature extraction techniques are used to collect some important features in images. Then, these features can be useful in recognition or/and classification of images. Before applying a feature extraction technique, the images have to go through different DIP techniques including normalization, thresholding, and binarization. Each feature extraction technique identifies the outstanding characteristics of the image [17].

In our paper, we apply the three feature extraction techniques including area extraction, centroid (center of gravity), and aspect ratio. Their mentioned techniques are implemented to extract digits in the image. The NRS has the six following steps of feature extraction. Firstly, we use the area extraction to collect the number of all objects in the image focusing on a group of black pixels. Secondly, we use the centroid to calculate the center of each object [18]. The objects in the image are found by the area extraction. Thirdly, the system obtains the smallest rectangles (bounding boxes) measured from the centroid to the boundary pixel in the object regions in images. Fourthly, the aspect ratio, which is the ratio between the width and height of the bounding box, is computed with setting an acceptable aspect ratio value between 0.5 and 10. Finally, the system obtains the final bounding boxes which follow the condition of the aspect ratio. Moreover, we also compute the eccentricity ratio, which is the distance between the foci of the ellipse and its major axis length in each object obtained from area extraction [19]. This eccentricity value is between 0 and 1. If the value close to 1 (more than 0.9), it means that the object area is too long and skinny. Therefore, we can interpret that area as noise in the image.

##### B. Adaptive Threshold

An adaptive threshold is used to reduce the weakness of global thresholding using Otsu's method. In Otsu's method, it finds a measure of the spread between the pixel levels in each side of the threshold. The aim of this method is to calculate the minimum threshold which is the sum of a foreground and a background. This method separates an image into two following classes, i.e., a particle region class (class 0) and a background region class (class 1). The variance of intra-class of Otsu's method at time instance  $t$  is expressed as

$$\sigma_b^2 = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2, \quad (1)$$

where  $\mu_0$  and  $\mu_1$  are the means of class 0 and class 1, respectively.  $\omega_0$  and  $\omega_1$  are the probabilities of class 0 and class 1, respectively. The drawback of Otsu's threshold is that this threshold is hard to be found because it requires an image histogram with the bi-modal distribution with a sharp and deep valley between two peaks to be able to define the classes correctly. The result of Otsu's threshold is not satisfactory if the image has a poor contrast between the foreground and the background

[1]. Therefore, an adaptive threshold is used to provide more effective performance in terms of different lighting levels in the same image. An adaptive thresholding is used to obtain a binary image from a grayscale image [1]. This method has the three following steps. Firstly, it calculates the threshold in the small kernel matrix of the image with a mean filtering operation [20]. Therefore, it can collect the different thresholds for different kernel matrix. Secondly, we subtract the original images with different thresholds according to the positions of pixels. Finally, the negative and positive resultant values are converted to black and white pixels.

C. Histogram of Oriented Gradients (HOG)

The main concept of the HOG algorithm is to replace the image pixels with the gradient direction of adjacent pixels because the HOG feature tries to rectify the robustness of the illumination variance [21]. In the HOG algorithm, it is divided the image into smaller cells. Then, the image borders in each cell are presented in their histograms by gradients calculation. The HOG is able to calculate the magnitude and direction of gradient as:

$$g = \sqrt{g_x^2 + g_y^2} \tag{2}$$

and

$$\theta_g = \arctan\left(\frac{g_y}{g_x}\right) \tag{3}$$

where  $g$  is the gradient magnitude,  $\theta_g$  is a direction of the gradient,  $g_x$  is the horizontal gradient value, and  $g_y$  is the vertical gradient value. In [22], the HOG feature was initially proposed. It shows good results in pedestrian detection with complex background in the image. Then, the HOG feature was applied to detect the human and count the total of humans in the video (VDO) files [23].

In our paper, the HOG values are considered as the inputs of the MLP model. The dimension of the HOG must be calculated. The HOG dimension is the same value as the number of input nodes, which is computed as  $\lfloor \text{BlockPerImage} \times N_{bin} \times \text{BlockSize} \rfloor$ , where  $\text{BlockPerImage}$  is the number of blocks in each image,  $N_{bin}$  is the number of bins in the histogram,  $\text{BlockSize}$  is the number of cells in block, and  $\lfloor \cdot \rfloor$  is a floor function. In our NRS, we use  $8 \times 8$  cell size,  $2 \times 2$  block size, nine bins, and twenty-four blocks per image. Consequently, the HOG dimension is 864.

D. Hough Transform (HT)

The HT is one of the feature extraction techniques used in DIP, image analysis, and computer vision. It is an important tool to detect straight lines in images even in the presence of noise and missing data [24]. The straight-line elements are collected in the array called Hough space. Although we can express the line segment in several forms the most convenient form is a parametric notion, which is expressed as:

$$x \cos(\theta) + y \sin(\theta) = \rho \tag{4}$$

where  $\rho$  is the normal line length from the origin to this line and  $\theta$  is the rotation angle of  $\rho$  with respect to the x-axis for any  $(x,y)$  on this line.

In [24], the authors show how to implement the HT to the images. It expresses the corresponding formula which is used to find the Hough space elements. Their results are the creating line in the different images. In [25], the authors show how to implement the HT to a line detection system. Their work was done applying this technique for line detection in the road images. In our paper, the HT is applied for finding the longest straight line in the image to identify the horizontal boundary of the screen. Moreover, the NRS can get the rotation angle of the SSDs from the HT technique.

V. IMAGE NOISE

We create the system under realistic situations with the three following types of noise, i.e., the SP noise, Gaussian noise, and Speckle noise.

A. Salt and Pepper Noise (SP Noise)

SP noise, a.k.a., impulse noise is a form of noise which typically happens in an image. This noise is caused by many factors, such as the analog-to-digital (A/D) converter errors, data transmission errors, and pixels disturbing in the camera sensors [26]. This noise can be caused by sharp and sudden disturbances in the image. This effect of SP noise is called the dead pixels. Fig.1(a) shows the digital meter image without noise and Fig.1(b) shows the example of SP noise with a noise density of 0.01 in the image. The noisy image caused by SP noise is expressed as:

$$N(x, y) = I(x, y) + \eta_{SP}(x, y) \tag{5}$$

where  $I(x, y)$  denotes the original image and  $\eta_{SP}(x, y)$  denotes the SP noise with the noise density ( $d$ ) = 0.01. This affects approximately  $dM_{img}$ , where  $M_{img}$  is the number of pixels of  $I(x, y)$ .

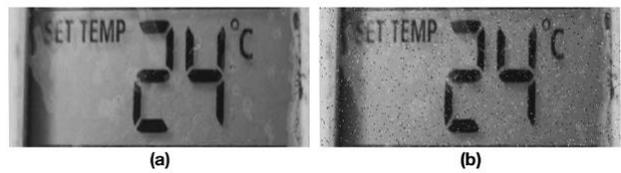


Fig. 1. Comparison between (a) Digital meter image without noise and (b) Digital meter image with SP noise with noise density of 0.01

B. Gaussian Noise

This noise is a white idealized noise form which is caused by random fluctuations in the signal [27]. The Gaussian noise in digital images mostly occurs during image acquisition from sensors or webcam cameras. This type of noise normally affects the gray values in the digital image. It means that Gaussian noise is caused by poor illumination and/or transmission, and/or high temperature. Fig. 2 shows the comparison between digital meter image without noise in Fig. 2(a) and digital meter

image with the Gaussian noise with a noise density of 0.01 in Fig. 2(b). The noisy image caused by the Gaussian noise is expressed as:

$$N(x, y) = I(x, y) + \eta_{\text{Gaussian}}(x, y), \quad (6)$$

where  $I(x, y)$  denotes the original image and  $\eta_{\text{Gaussian}}(x, y)$  denotes the SP noise with  $d = 0.01$ .

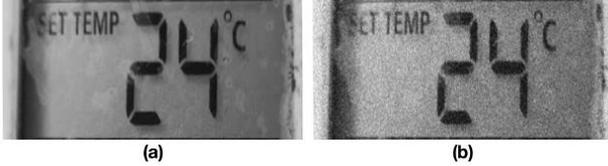


Fig. 2. Comparison between (a) Digital meter image without noise and (b) Digital meter image with Gaussian noise with noise density of 0.01

### C. Speckle Noise

Speckle noise, a.k.a., multiplicative noise normally occurs in ultrasonic or radar images, such as the medical ultrasonic images and the synthetic aperture radar (SAR) images [28]. This noise can be generated by the random value multiplications with image pixel values. In Fig. 3, it compares the images with and without Speckle noise. The digital meter image without the noise shows in Fig. 3(a) and the digital meter image with Speckle noise with a noise density of 0.01 shows in Fig.3(b). The noisy image caused by the Speckle noise is expressed as:

$$N(x, y) = I(x, y)\eta_{\text{Speckle}}(x, y) + I(x, y) \quad (7)$$

where  $\eta_{\text{Speckle}}(x, y)$  denotes the Speckle noise which is uniformly distributed random noise with mean 0 and variance 0.01.

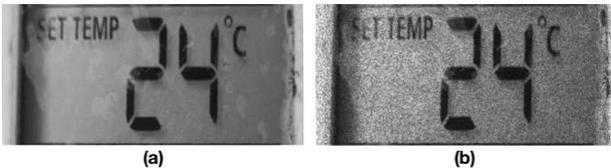


Fig. 3. Comparison between (a) Digital meter image without noise and (b) Digital meter image with noise density of 0.01

The SP noise in Fig. 1(b) shows that the noise components are randomly scattered in the image due to the sudden and sharp interference signal in the system. The effect of Gaussian noise is almost similar to Speckle noises as shown in Fig. 2(b) and Fig. 3(b), respectively.

## VI. SYSTEM MODEL

An NRS is developed to recognize the numerical value in the SSDs. The similar of our work are proposed in [1], [11], [13], and [14]. Although their work has an acceptable accuracy rate, their system can only work for the same pattern of the SSD. Therefore, we apply the HOG as the NN input together with the MLP model, which is one of the NN models, on the NRS for improving the complexity and the system performance. An NN represents a human brain system which is able to

simulate the learning process [29]. In our paper, we use an IIoT platform to resolve the limitations of the acquisition and accessibility of data by transmitting the images with using the low-cost webcams, to get rid of the problem in human eye error, and to process it in real-time. The processes of an IIoT platform in the proposed NRS have three modules. In the first module, low-cost webcams take images of the SSDs. This module observes the screen rotation problem and solves this problem using the HOG and the HT techniques. In addition, the images convert from the grayscale images to the binary images. In the second part, the binary images are sent by the Internet to the receiver part. This part can be affected by the noise. The most commonly noise for the wireless communication is the additive white Gaussian (AWGN) noise at the receiver's front-end, and man-made noise [30]. In the last module, the server is used to operate NN model to recognize the seven-segment digits. The mechanism of the proposed NRS with the IIoT platforms is shown in Fig. 4.

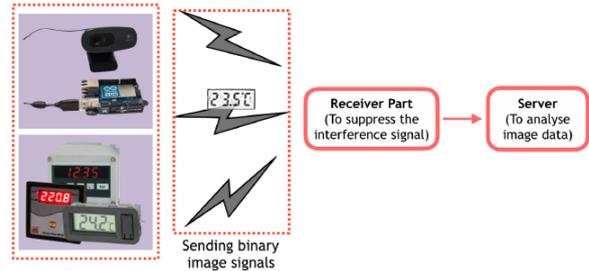


Fig. 4. Proposed IIoT in NRS

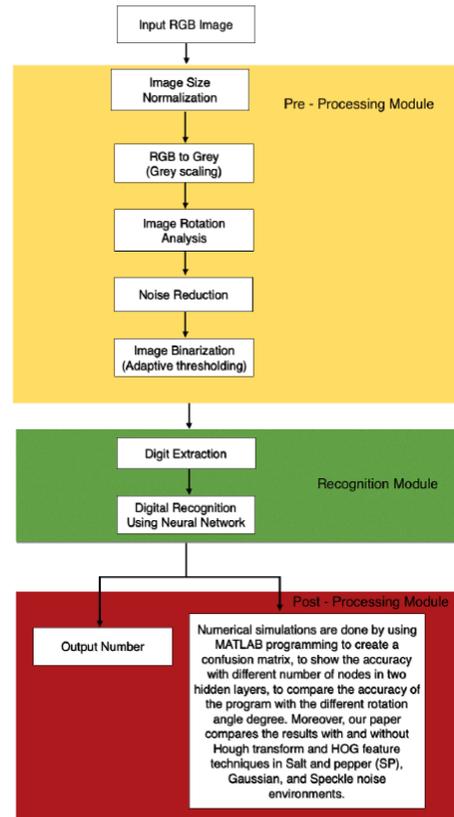


Fig. 5. Flowchart of functions in the proposed NRS

The functions in the overall proposed system are explained in the flowchart shown in Fig. 5. The NRS is developed for identifying SSDs. The red green blue (RGB) image obtained from a low-cost camera, e.g., a webcam. The RGB image is pre-processed by the following processes: image size normalization (The dimensions of the normalized image matrix are 500 rows and 1000 columns.), converting from the RGB color model to the grayscale model, image rotation analysis, noise reduction, and image binarization or adaptive thresholding. After that, the processed image is sent to the recognition module which is done by a digit extraction and a proposed NN method. As previously explained, the digit extraction consists of three common feature extractions which are centroids, aspect ratio, and area. Finally, the resultant performances are evaluated by the post-processing module in terms of confusion matrices, accuracy, and system complexity in terms of computation time and memory usage.

### VII. RECOGNITION MODULE WITH MULTI-LAYER PERCEPTRON (MLP)

An MLP is developed extending from a single-layer perceptron, which does not have a hidden layer. The hidden layer collects the group of neurons which contain an activation function. In the activation functions, we used the sigmoid-function because the dataset of this system is binary images which consist of 0 and 1 for black and white pixel's color. The hidden nodes depend on the types of input data and the number of input nodes. Many researches have tried to obtain the optimal number of nodes in the hidden layer based on trial-and-error and rule-of-thumb methods [31]. Our work shows how to use the MLP with two hidden layers. Most of the MLP models apply a back propagation (BP) algorithm which can be defined as a gradient descent method to minimize the squared output error [18]. Fig. 6 shows the BP in our MLP model with two hidden layers. In Fig. 6, the number of nodes of the input layer is 864 nodes due to the dimension of the HOG feature whose calculation is previously shown in Subsection 4.3. Hidden layers 1 and 2 have 64 and 32 nodes, respectively. These are the results of the trial-and-error method which is explained in the next paragraph. The number of nodes of the output layer is 11 according to the numerical digits 0-9, and noise or symbols in number 10.

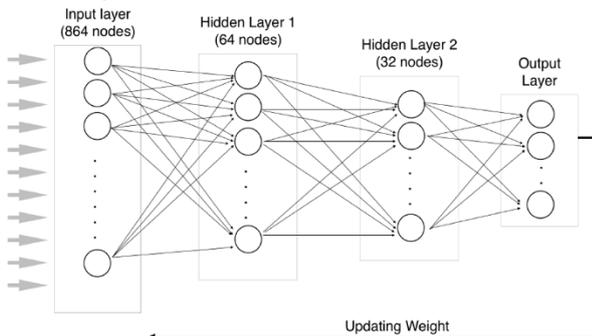


Fig. 6. Our BP in MLP model with two hidden layers

The critical task when we apply the NN model is how to select the optimal number of nodes in each hidden layer. Because if the unnecessary nodes are present in the NN, then the overfitting can occur. Therefore, we use the trial-and-error method to verify the best number of hidden nodes as it is normally used [32]. As can be seen in Fig. 7, we increase the number of nodes according to the trial-and-error method in the first hidden layer as  $2^i$ , where  $i$  is the iteration number and the second hidden layer increases  $2^{i-1}$ . Finally, we select 64 and 32 nodes in the first and second hidden layers, respectively as they provide the best accuracy.

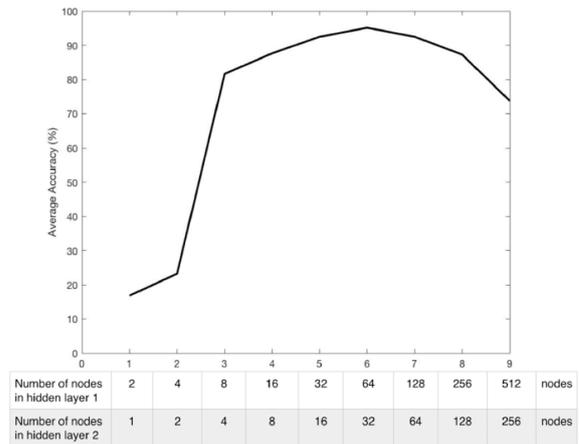


Fig. 7. Accuracy (in %) of hidden nodes in two hidden layers graph

#### A. Back Propagation (BP)

A BP model is widely performed in the supervised learning approaches [25]. This method is used in an ANN to calculate a gradient that is needed in the calculation of the weights in the network. The BP with a conjugate gradient algorithm can minimize the total error ( $E_{total}$ ) expressed as:

$$E_{total} = \sum_{i=1}^N \frac{1}{2} (target_i - out_i)^2 \quad (8)$$

where  $N$  is the number of nodes in the output layer.  $target_i$  and  $out_i$  are the target and output values of each node, respectively.

$$\frac{\partial E_{total}}{\partial W_i} = \frac{\partial E_{total}}{\partial out_{o1}} \frac{\partial out_{o1}}{\partial net_{o1}} \frac{\partial net_{o1}}{\partial W_i} \quad (9)$$

and

$$W_i^{new} = W_i - \eta \frac{\partial E_{total}}{\partial W_i} \quad (10)$$

where  $W_i^{new}$  is an updated weight of weight  $W_i$ .  $\eta$  is a learning rate.  $net_{o1}$  is a network input value at output node 1. Typically,  $\eta$  is set to 0.5.

### VIII. RESULTS AND DISCUSSION

In this section, the simulations and the corresponding results comparing the system with and without the rotation process are presented and analyzed.

A. Improvement by Using HOG Feature and HT Techniques

This work compares the performances between the NRS with and without the HOG feature and HT techniques. The HOG feature and HT techniques are used to solve the effect of the screen rotation problem. An NN is implemented to recognize the SSDs. However, the inputs of the NN model of both cases are different. In the system without the HOG feature and HT techniques, the number of pixels is used as inputs, while the number of HOG features is the input of the system with the HOG feature and HT techniques. Firstly, we show the benefit of using these techniques in the normal case, i.e., no screen rotation problem. Later, such a screen rotation problem is taken into account. Fig. 8 shows the preliminary results as the confusion matrix of the original NRS, while Fig. 9 shows the confusion matrix of the NRS with the HOG feature and HT techniques. The NRS with the HOG feature and HT techniques outperforms the original NRS even in the normal situation (no screen rotation problem) as we can see from the overall accuracy of 95.13% and 90.26%, respectively. Note that the percentages of the accuracy are computed by dividing the correctly identified objects and the total number of objects.

Object	Number of object	Number of times object recognized as										Symbols		
		0	1	2	3	4	5	6	7	8	9			
0	26	26	-	-	-	-	-	-	-	-	-	-	-	-
1	18	-	18	-	-	-	-	-	-	-	-	-	-	-
2	24	-	-	24	-	-	-	-	-	-	-	-	-	-
3	16	-	-	-	16	-	-	-	-	-	-	-	-	-
4	11	-	-	-	-	11	-	-	-	-	-	-	-	-
5	17	-	-	-	-	-	17	-	-	-	-	-	-	-
6	10	-	-	-	-	-	-	10	-	-	-	-	-	-
7	10	-	-	-	1	1	-	-	8	-	-	-	-	-
8	50	-	-	2	-	-	-	-	-	48	-	-	-	-
9	16	-	-	-	-	3	-	-	-	-	13	-	-	-
Symbols	69	-	4	-	-	2	4	2	2	4	1	-	-	50

Fig 8. Confusion matrix of the original NRS (no screen rotation problem)

Object	Number of object	Number of times object recognized as										Symbols		
		0	1	2	3	4	5	6	7	8	9			
0	26	25	-	-	-	-	-	-	-	-	-	-	-	1
1	18	-	15	-	-	-	-	-	-	-	-	-	-	3
2	24	-	-	24	-	-	-	-	-	-	-	-	-	-
3	16	-	-	-	16	-	-	-	-	-	-	-	-	-
4	11	-	-	-	-	11	-	-	-	-	-	-	-	-
5	17	-	-	-	-	-	16	-	-	-	-	-	-	1
6	10	-	-	-	-	-	-	10	-	-	-	-	-	-
7	10	1	-	-	-	-	-	-	7	-	-	-	-	2
8	50	-	-	-	-	-	-	-	-	48	-	-	-	2
9	16	-	-	-	-	1	-	-	-	-	15	-	-	-
Symbols	69	-	-	1	-	-	1	-	-	-	-	-	-	67

Fig. 9. Confusion matrix of the NRS with the HOG feature and HT techniques (no screen rotation problem)

B. NRS Results and Discussion for SSD

We propose an NRS based on an OCR method together with an MLP. Moreover, the HOG feature and HT techniques are used to solve the problem of screen

rotation. SP noise, Gaussian Noise, and Speckle noise are included in the analysis since they are typical noises that can be seen in an image. A Wiener filter, which is an adaptive filter and the most important technique to remove unfocussed optics or linear motion is used. In Fig. 10, we can see from the performance that the Wiener filter can significantly improve the performance under all the types of noise.

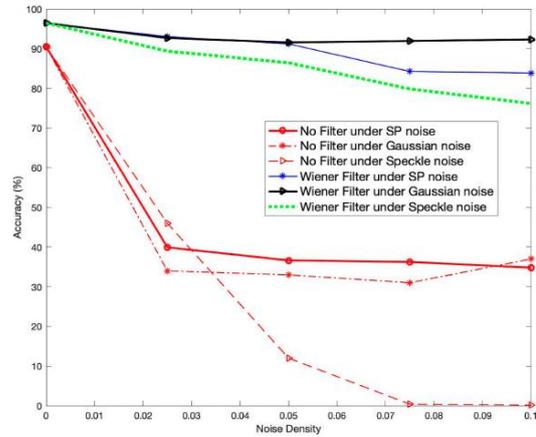


Fig. 10. Accuracy (in %) of NRS with and without Wiener filter under the SP noise, Gaussian noise, and Speckle noise with the different noise densities.

C. Accuracy of Original NRS and NRS with HOG Feature and HT Techniques

We implement the original NRS and the NRS with HOG feature and HT techniques under the screen rotation which rotated between 0-180 degrees. The accuracy of both systems under the screen rotation problem (0-180 degrees) is shown in Fig. 11. In the original NRS, when the screen is rotated with 180 degrees, the accuracy is higher than the other rotation angles (except 0 degree) because some digits such as 0, 1, and 8 show the same characters in 0 and 180 degrees. We can see from the figure that using the NRS with the HOG feature and HT techniques can improve the accuracy in all rotation degrees. It means that the results in terms of accuracy of the NRS with the HOG feature and HT techniques are better than the results of the original NRS.

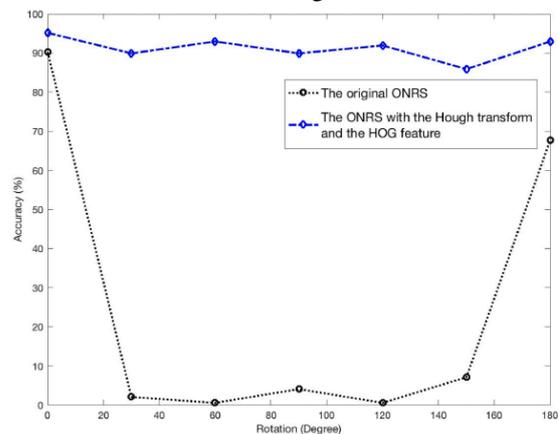


Fig. 11. Comparison between the accuracy (in %) of the original NRS and the NRS with HOG feature and HT techniques

**D. Results under Screen Rotation Problem**

The screen rotation problem can usually occur with cameras in industries. Therefore, an NRS with the HOG feature and HT techniques is proposed to solve this problem. The proposed NRS can very well recognize not just the SSDs but also the symbol and the noise components in the SSDs with screen rotation problem as shown in Fig. 12(a)-(f). The symbol and the noise components are collected as number 10, which is predefined in the database.

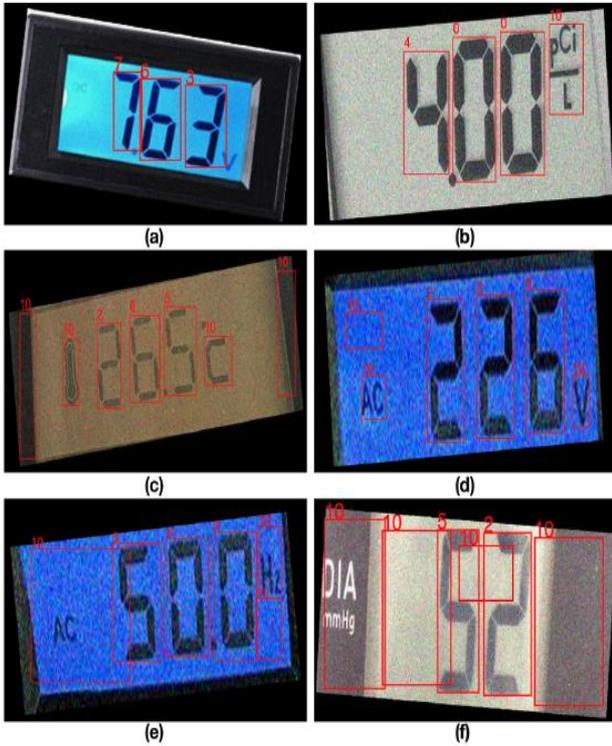


Fig. 12. The NRS results for screen rotation problem under the Gaussian noise density of 0.01. (a) The NRS result which can detect the SSDs (b) and (c) The NRS results which can detect the symbol in number 10 (d), (e), and (f) The NRS results which can detect the symbols and noise in number 10

**E. Accuracy of NRS under Random Screen Rotation Problem (0-180 degrees)**

The accuracy (in %) of the proposed NRS under three types of noise including SP noise, Gaussian noise, and Speckle noise for the screen rotation problem is shown in Fig. 13. The rotation angles are randomly varied with a uniform distribution from 0-180 degrees. Five iterations are simulated to show the average accuracy of the proposed system, which are 89.84%, 89.38%, and 89.74% for the cases under SP noise, Gaussian noise, and Speckle noise, respectively. Overall, they are almost the same

**F. Accuracy of NRS under Random Screen Rotation Problem (0-180 degrees)**

The accuracy (in %) of the proposed NRS under three types of noise including SP noise, Gaussian noise, and Speckle noise for the screen rotation problem is shown in Fig. 13. The rotation angles are randomly varied with a

uniform distribution from 0-180 degrees. Five iterations are simulated to show the average accuracy of the proposed system, which are 89.84%, 89.38%, and 89.74% for the cases under SP noise, Gaussian noise, and Speckle noise, respectively. Overall, they are almost the same.

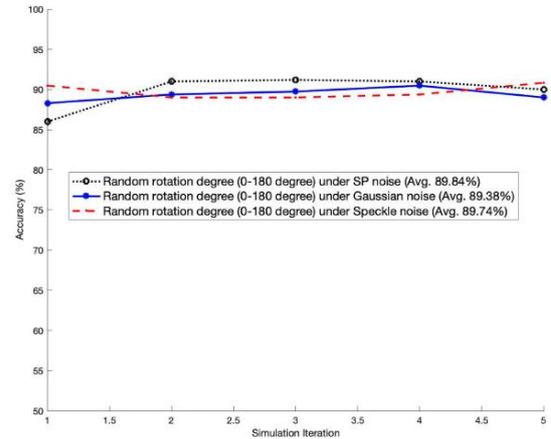


Fig. 13. Accuracy (in %) of the system under three noise environments for the random screen rotation problem (0-180 degrees) versus the number of iterations of the simulations

**G. System Complexity in Terms of Computation Time and Memory Usage of NRS**

The NRS with and without the HOG feature and HT techniques are tested on an Intel Core i5 processor and 8 GB RAM based computer. The MATLAB version R2017a is used. This paper used datasets from many different resources on the Internet such as ImageNet, Mendeley, and Google dataset. In addition, some images were taken by our low-cost camera. In total 497 images were used. We have 51 test set images which contain 198 seven-segment numbers and 69 symbols. Table II. shows the system complexity as the computation time and the memory usage under three types of noise with 0.01 noise density. The comparison of NRS with and without the HOG feature and HT techniques shows that under SP noise the system takes the longest time and the most memory. In contrast, the system operating under the Speckle noise takes the least time and memory. The system with the HOG feature and HT techniques takes a longer time than the original NRS in general. Although the proposed NRS with the HOG feature and HT techniques is more complicated taking longer computation time, it is not significant. Table II. shows that our system took at most 1.86 seconds for analyzing an image which is acceptable as a real-time processing since the maximum delay requirement for IIoT architecture must be less than two seconds [33].

**H. Our Testing Experiments with the NRS**

In this part, we describe the experiment steps and show the usage of our NRS in Fig. 14(a)-(d). We took 96 additional images of the rotated SSDs and used them for the additional testing process. Each step of the experiment is explained as follows. First, the low-cost camera captures the input images. Fig. 14(a) shows one

example of them. It consists of the seven-segment digits and other undesired text information. Second, the images are normalized and further proceeded by the pre-processing stage applying the adaptive thresholds as shown in Fig. 14(b). In Fig. 14(b), the HT is implemented to solve the screen rotation problem. Next, the result of the feature extraction is depicted in Fig. 14(c). In this figure, the problem of discontinuity in each seven-segment digit is solved. Finally, in the last step, each digit is recognized by the NN model and the corresponding output is shown in Fig. 14(d). We can see that the other undesired text information is shown as number 10. The results of NRS from testing in each digit are shown in Table III. The average accuracy is about 95 percent, which is similar to the earlier results.

TABLE II: COMPLEXITY COMPARISON BETWEEN NRS WITH AND WITHOUT THE HOG FEATURE AND HT TECHNIQUES FOR SSDS

NRS	Types of Noise	Computation Time and Memory Usage (per image)
NRS without HOG feature and HT Technique	SP Noise	1.77 seconds and 28.72 MB
	Gaussian Noise	1.58 seconds and 28.54 MB
	Speckle Noise	1.40 seconds and 27.71 MB
NRS with HOG feature and HT Technique	SP Noise	1.86 seconds and 47.54 MB
	Gaussian Noise	1.75 seconds and 45.51 MB
	Speckle Noise	1.61 seconds and 41.27 MB

TABLE III: PERFORMANCE OF NRS FROM THE EXPERIMENTS

Digit	Total Samples	Correctly Recognized Samples	Accuracy (in %)
Digit 0	61	59	96.72
Digit 1	39	36	92.31
Digit 2	34	32	94.12
Digit 3	44	43	97.73
Digit 4	37	36	97.30
Digit 5	33	31	93.94
Digit 6	56	55	98.21
Digit 7	35	32	91.43
Digit 8	40	37	92.50
Digit 9	24	22	91.67
Digit 10 (Noises and Symbols)	116	110	94.83
Total	516	493	95.54

IX. CONCLUSION

The numeral recognition system (NRS) based on an optical character recognition (OCR) method was proposed. Firstly, we collected the images of seven-segment displays (SSDs) from various types of instrument displays, such as calculators, multi-meters, and industrial equipment. Secondly, the digital image processing (DIP) techniques including morphological operations, filtering, and adaptive threshold setting, were used to prepare more suitable images for pattern recognition. Then, a multi-layer perceptron (MLP), which

is one of the neural network (NN) models, was used to classify the digit data. The NRS was implemented in an industrial Internet of things (IIoT) platform which made the collected data possible to be used immediately without a cumbersome process for dealing with different equipment brands. The Salt and pepper (SP) noise, Gaussian noise, and Speckle noise were taken into account in the study since they are typical noise types that occurred on an image. An adaptive Wiener filter is used for noise mitigation. Moreover, the screen rotation problem was considered and solved by applying the histogram of oriented gradients (HOG) and Hough transform (HT) techniques. Finally, the MLP was used for classifying the numbers in the images. We achieved the average accuracy about 95 percent from the actual experiments. The system complexity in terms of computation time and memory usage was verified. Under the SP noise, the system took longer time than when it operated under the Gaussian and Speckle noises. The NRS with the HOG and HT techniques could well enhance the accuracy while taking just a bit more computation time and memory usage than the NRS without the HOG and HT techniques. Due to low computation time, in general, our proposed system is suitable for a real-time service. IN the possible use cases application of this work in the future, it can include remote data acquisition and system actuation in factories or remote data acquisition and laboratory equipment parameter adjustment for online laboratory classes.

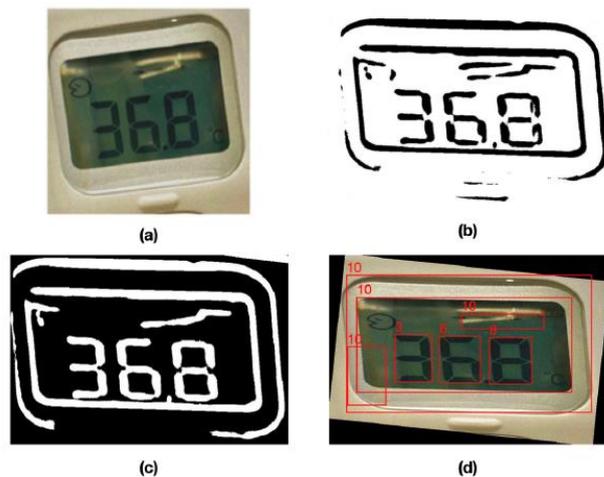


Fig 14. (a) The input image (b) The pre-processing image (c) The image with the feature extraction techniques (d) Recognized output image

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Both authors conducted the research. Natthanan Promsuk analyzed the data and wrote the paper. Attaphongse Taparugssanagorn advised and supported in the research. Both authors had approved the final version.

## REFERENCES

- [1] S. Ghosh and S. Shit, "A low cost data acquisition system from digital display instruments employing image processing technique," in *Proc. International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, New Delhi, India, 2014, pp. 1065–1068.
- [2] N. Promsuk, A. Taparugssanagorn, and J. Vartiainen, "Probability of false alarm-based interference suppression methods in Internet of Things (IoT) systems," *Computer Networks* vol. 144, pp. 201–215, October 2018.
- [3] N. Promsuk and A. Taparugssanagorn, "Long short term memory network-based interference recognition for industrial internet of things," *Journal of Communications*, vol. 15, no. 12, pp. 876–885, December 2020.
- [4] C. Gonzalez, R. Woods, and S. Eddins, "Digital image processing using matlab," *Pearson Education*, Second Indian Reprint
- [5] P. Sharma and A. Singh, "Era of deep neural networks: A review," in *Proc. 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Delhi, 2017, pp. 1–5.
- [6] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," *Advances in Neural Information Processing Systems*, vol. 1, pp. 396–404, 1990.
- [7] S. S. Baboo and I. K. Shereef, "An efficient weather forecasting system using artificial neural network," *International Journal of Environmental Science and Development*, vol. 1, p. 321, October 2010.
- [8] M. Nikraves, L. A. Zadeh, and F. Aminzadeh, "Soft computing and intelligent data analysis in oil exploration," *Elsevier*, 2003.
- [9] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, pp. 98–113, January 1997.
- [10] G. Hinton, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, pp. 82–97, November 2012.
- [11] R. P. Ghugardare, S. P. Narote, P. Mukherji, and P. M. Kulkarni, "Optical character recognition system for seven segment display images of measuring instruments," in *Proc. TENCON 2009- 2009 IEEE Region 10 Conference*, IEEE, 2009, pp. 1–6.
- [12] X. Zhai, F. Bensaali, and R. Sotudeh, "Ocr-based neural network for anpr," in *Proc. IEEE International Conference on Imaging Systems and Techniques Proceedings*, IEEE, Manchester, UK, 2012, pp. 393–397.
- [13] P. H. Kulkarni and P. D. Kute, "Optical numeral recognition algorithm for seven segment display," in *Proc. Conference on Advances in Signal Processing (CASP)*, Pune, India, IEEE, 2016, pp. 397–401.
- [14] K. Kanagarathinam and K. Sekar, "Text detection and recognition in raw image dataset of seven segment digital energy meter display," *Energy Reports*, vol. 5, pp. 842–852, July 2019.
- [15] Y. Wen, Y. Lu, J. Yan, Z. Zhou, K. M. V. Deneen, and P. Shi, "An algorithm for license plate recognition applied to intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, pp. 830–845, December 2011.
- [16] N. D. V. Dalarmelina, M. A. Teixeira, and R. I. Meneguette, "A realtime automatic plate recognition system based on optical character recognition and wireless sensor networks for its," *Sensors*, vol. 20, p. 55, January 2020.
- [17] G. Kumar and P. K. Bhatia, "A detailed review of feature extraction in image processing systems," in *Proc. Fourth International Conference on Advanced Computing & Communication Technologies*, Rohta, India, IEEE, 2014, pp. 5–12.
- [18] M. Yuwono, J. Sepulveda, and A. A. Handojoseno, "Centroid extraction from hartmann-shack images using swarm clustering approach," in *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, CA, USA, 2012, pp. 1446–1449.
- [19] T. Luhmann, "Eccentricity in images of circular and spherical targets and its impact on spatial intersection," *The Photogrammetric Record*, vol. 29, no. 148, pp. 417–433, June 2014.
- [20] S. Sun and H. Zhao, "Kernel averaging filter," in *Proc. Congress on Image and Signal Processing*, IEEE, Hainan, China 2008, pp. 681–685.
- [21] H. C. Yang and X. A. Wang, "Cascade face detection based on histograms of oriented gradients and support vector machine," in *Proc. 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, IEEE, Krakow, Poland, 2015, pp. 766–770.
- [22] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, CA, USA, 2005, pp. 886–893.
- [23] T. Surasak, I. Takahiro, C. H. Cheng, C. E. Wang, and P. Y. Sheng, "Histogram of oriented gradients for human detection in video," in *Proc. 5th International Conference on Business and Industrial Research (ICBIR)*, Bangkok, Thailand, 2018, pp. 172–176.
- [24] L. Chandrasekar and G. Durga, "Implementation of hough transform for image processing applications," in *Proc. International Conference on Communication and Signal Processing*, Melmaruvathur, India, 2014, pp. 843–847.
- [25] B. Fang and L. Qian, "Optimized generalized hough transform for road marking recognition application," in *Proc. International Conference on Security, Pattern Analysis, and Cybernetics*, Shenzhen, 2017, pp. 361–367.
- [26] E. J. Leavline and D. A. A. G. Singh, "Salt and pepper noise detection and removal in gray scale images: an experimental analysis," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 6, no. 5, pp. 343–352, October 2013.
- [27] A. M. Hambal, Z. Pei, and F. L. Ishabailu, "Image noise reduction and filtering techniques," *International Journal of Science and Research*, vol. 6, no. 3, pp. 2033–2038, March 2017.
- [28] A. Maity, A. Pattanaik, S. Sagnika, and S. Pani, "A comparative study on approaches to speckle noise reduction in images," in *Proc. International Conference on Computational Intelligence and Networks*, Malaysia, 2015, pp. 148–155.

- [29] S. P. Mahasagara, A. Alamsyah, and B. Rikumahu, "Indonesia infrastructure and consumer stock portfolio prediction using artificial neural network backpropagation," in *Proc. 5th International Conference on Information and Communication Technology*, 2017, pp. 1-4.
- [30] J. Vartiainen, R. Vuhtoniemi, A. Taparugssanagorn, and N. Promsuk, "Interference suppression and signal detection for LTE and WLAN signals in cognitive radio applications," *International Journal on Advances in Telecommunications*, vol. 10, no. 1-2, pp. 1–10, 2017.
- [31] S. Karsoliya, "Approximating number of hidden layer neurons in multiple hidden layer bpnn architecture," *International Journal of Engineering Trends and Technology*, vol. 3, no. 6, pp. 714–717, November 2012.
- [32] F.S. Panchal and M. Panchal, "Review on methods of selecting number of hidden nodes in artificial neural network," *International Journal of Computer Science and Mobile Computing*, vol. 3, pp. 455-464, November 2014.
- [33] P. Ferrari, A. Flammini, E. Sisinni, S. Rinaldi, D. Brandão, and M. S. Rocha, "Delay estimation of industrial IoT applications based on messaging protocols," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 9, pp. 2188-2199, September 2018.

Copyright © 2021 by the authors. This is an open access article distributed under the Creative Commons Attribution License ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial, and no modifications or adaptations are made.



**Natthanan Promsuk** received the Bachelor of Engineering (Computer Engineering) degree from Chiang Mai University, Chiang Mai, Thailand in 2014. He obtained the degree of Master of Engineering in 2017. Also, he obtained the degree of Doctor of Engineering (in Telecommunications) in 2020 from the Asian Institute of Technology (AIT), Pathum Thani, Thailand. Currently, he is a

faculty member in Chiang Mai University in the department of Computer Engineering. His research interest includes wireless communications, industry 4.0, wireless sensor network, Artificial Intelligence (AI), data acquisition, and computer network.



**Attaphongse Taparugssanagorn** received the Bachelor of Engineering degree in from the Chulalongkorn University, Bangkok, Thailand in 1997. After that he worked as an engineer for Siemens Ltd, Bangkok in Telecommunications transmission department for two years. He then obtained the Master of Science (in Electrical Engineering) degree in 2001

from the Technische Universität Kaiserslautern, Kaiserslautern, Germany. He gained an experience as a researcher in the institute of communications, the University of Stuttgart until 2003. Then, he joined Centre for Wireless Communications, University of Oulu, Oulu, Finland at the end of 2003 and completed his Doctor of Technology degree there in 2007. In 2008, he worked as a visiting postdoctoral researcher in Yokohama National University, Yokohama, Japan. After he returned home in Bangkok in 2011, he joined the Asian Institute of Technology, Thailand as an adjunct faculty. At the end of 2012, he has also joined at the National Electronics and Computer Technology Center, Thailand. In August semester 2015 until present he has been a full-time faculty member at the Asian Institute of Technology. Currently, he is an associate professor in the department of ICT. His research interest includes signal processing, statistical signal processing (detection and estimation techniques), and Artificial Intelligence (AI) (machine learning) for various applications, for example, wireless communications, fall detection, pattern recognition, coexistence problems, AI and Internet of Things (IoT) based systems, medical/healthcare/ music/sport ICT, and many others.