

Review of Industry Workpiece Classification and Defect Detection using Deep Learning

Changxing Chen, Azween Abdullah, S.H. Kok, D.T.K. Tien

School of Computer Science and Engineering, Taylor's University, Subang Jaya, Malaysia

Abstract—Object detection and classification denotes one of the most extensively-utilized machine vision applications given the high requirements put forward for object classification and defect detection with the rise of object recognition scenes. Notwithstanding, conventional image recognition processing technology encounters specific drawbacks. Its benefits and limitations were duly compared upon selecting several typical conventional image recognition techniques. Resultantly, such recognition approaches required multiple manual participation elements and extensive manpower with restricted object identification. As a branch of machine learning, deep learning has attained more optimal results in the image recognition discipline. In the classification and defect detection of industrial workpieces, over 70 literature reviews of deep learning algorithms across multiple application scenarios for classical algorithm model and network structure assessment based on the deep learning theory. Relevant network model performance was compared and analyzed based on network intricacies parallel to natural image classification. Six research gaps were found based on the reviewed algorithm pros and cons. The corresponding six research proposal in workpiece image classification was highlighted with prospects on the workpiece image classification and defect detection direction development. It provides an empirical solution for the selection of workpiece classification and defect detection deep learning model in the future.

Keywords—Convolutional neural network; image processing; image recognition; defect detection; deep learning

I. INTRODUCTION

In line with the proposal of artificial intelligence [1], optimal computational intelligence performance in mathematical theory and computing power enriches the artificial intelligence theoretical framework and catalyzes artificial intelligence development. Perceivably, artificial intelligence constitutes a part of computer science following the emergence of industry 4.0. Artificial intelligence could enhance the level of organizational astuteness with substantial implications across multiple sectors. Deep learning and artificial intelligence machine learning depict a broad range of application disciplines, such as image recognition, network security, speech recognition, and natural language processing with significant breakthroughs. Various artificial intelligence recognition systems have been consistently developed with distinct functions and forms for economic and social advantages [2]. The fundamentals of artificial intelligence imply machine learning with algorithms. Multiple image recognition techniques require flexible adoption based on distinct application prerequisites to fulfill various image recognition task requirements in the practical application process. Under the computer vision category, image

recognition, which simulates human vision using computers or image-based instruments, facilitates computers to comprehend the recognized entities with algorithms to substitute human eye functions. Conventional image recognition method development proved relatively slow pre-artificial intelligence development. Such recognition techniques were previously based on the object feature descriptor for image recognition and matching with limited discussions on conventional image recognition approaches [3]. The deep learning theory was derived from the conventional neural network under the deep neural network. This theory has eventually become the mainstream of image recognition methods with a distinct object recognition concept simultaneously, it is widely used in pattern recognition [4, 5], image recognition has made breakthroughs upon introducing deep learning into the image processing field while resolving multiple problems that could not be managed by conventional approaches.

Workpiece surface defect is one of the most important factors affecting the product quality of mechanical workpiece. The traditional manual visual inspection method is easy to be affected by manual experience and subjective factors, which lead to inaccurate test results and cannot meet the current inspection requirements and the on-line production requirements of automatic production line. Machine vision inspection has the advantages of high automation, high recognition rate and non-contact measurement. It has gradually become the mainstream method and development trend of surface defect detection. According to the current inspection and classification requirements of workpiece manufacturers, combined with the actual situation of the industrial site, starting from solving the actual impact, workpiece classification and defect recognition are carried out through machine vision. Machine vision classification and detection algorithm can solve the problems of many types and large quantities of workpieces.

A substantial number of factory workpieces (common components in industrial manufacturing) are extensively employed in industrial production. The prerequisites for workpiece recognition speed and accuracy continue rising as opposed to manual workpiece classification strategies with low efficiency and accuracy. In this vein, classification detection denotes high subjectivity. The recent emergence of artificial intelligence technology and computer vision and its application has been extensively employed in industrial sites to catalyze industrial parts classification development. This study summarized the common conventional image recognition approaches, presented specific common image recognition techniques under deep learning, and compared the method performance and effects on other applications.

II. IMPORTANCE OF INDUSTRY WORKPIECE CLASSIFICATION

Workpiece denotes a product manufacturing process component where the machining object in machining or generation (a single part or combination of specific ones) is assembled. The advent of artificial intelligence includes novel application prerequisites for the factory flow production mode given the perpetual improvement of labor cost on the industrial site and highly stringent product quality requirements. Automatic workpiece assembly is highly significant post-artificial intelligence development as the conventional artificial assembly line production mode failed to complement advanced industrial production. As such, industrial automation must be established to optimize manufacturing industry competitiveness. Accelerated automation transformation and the optimization of conventional sectors remain as one of the fundamental points to catalyze industrial development. The artificial intelligence-industrial site integration is inextricably linked for high productivity, novel changes, and industrial development possibilities. Industrial assembly is highly essential in the entire production process as each industrial site requires distinct parts. On another note, a sophisticated industrial production line encompasses workpiece classification and detection. Notably, the robot arm completes the assembly, sorting, and other relevant tasks involving various industrial parts post-classification and detection. Some industrial part sorting proves unsuitable for workers to sort and detect due to industrial site risks when the application scenarios are sufficiently enriched as follows: the monitoring state in the parts-sorting process, sorting process control, and workpiece classification emergency treatment in the industrial site. High requirements are reflected for industrial part detection accuracy and equipment process control stability.

The conventional manual workpiece sorting approach depends on manual operation for parts classification realization. This technique requires high worker's proficiency with substantial product quality implications. Specifically, the equipment of refined workpieces hampers production sorting. As product sorting and detection period in the entire production cycle proves time-consuming, production optimization depends on whether the automatic product line sorting process could be actualized. The means of scientifically controlling the industrial site workpiece classification denotes a complexity that must be regarded and resolved by relevant personnel in encouraging continuous industrial development through automation and intelligence. Machine vision systems, with image processing as one of the pertinent technologies, are increasingly implemented to resolve classification issues. It is deemed necessary to recognize workpieces for target workpiece classification realization. Image recognition is a processing application technology in deep learning and a fundamental task in computer vision. The diversified industrial parts demand substantially challenges the manufacturer's production and classification level given the adverse environment and intricate background within the industrial field. It is considered challenging to recognize the image and resolve the problem with conventional image feature selection based on interference factors: light and workpiece placement background. The industrial workpiece images to be classified

are typically complex and ambiguous in practical production and application, thus rendering it intricate to structure an appropriate workpiece image classification approach. Observably, image classification denotes one of the difficult problems to be resolved in image classification and detection tasks.

III. GENERAL OPEN SOURCE WORKPIECE DEFECT DATASETS

Image datasets to highlight workpiece defects and classification remain lacking to date given the novelty of image defect detection studies. Current online public datasets generally constitute daily necessities, faces, and animals. Most recognition-oriented publications are performed on conventional image classification datasets [6]. Conventional image processing algorithms are typically incorporated into traditional surface defects and classification techniques. Artificial design features and classifiers are commonly implemented compared to the clear classification in computer vision. Specific datasets were adopted to complement the deep learning neural network training. The image formats typically encompass JPG and BMP and JPEG and RGB. Specific datasets in industrial disciplines, such as polished workpieces and customized CNC lathe workpieces do not possess public datasets. The lack of corresponding image training sets would inevitably restrict the promotion of deep learning applications in workpiece recognition. Data incongruence and dataset annotation need to be resolved despite the presence of constructed workpiece dataset pictures.

Current industrial dataset usage could accelerate the deep learning algorithm model development while public, reliable, and open-source industrial datasets could compare distinct deep learning detection algorithms. This section briefly elaborates extensively-utilized industrial datasets following the industrial workpiece classification strategy and the broad industrial dataset application for defect detection as shown in Table I.

TABLE I. GENERAL OPEN SOURCE WORKPIECE DEFECT DATASET

Dataset	Literature	Type Quantity	Image Size	Number of Images
NEU surface defect database	[7]	6	200*200	1800
Severstal steel defect dataset	[8]	4	1600*256	12568
DAGM 2007 dataset	[9]	10	Grayscale 8-bit PNG format	11500
Kolektor surface defect dataset	[10]	50	width: 500 px height: from 1240 to 1270 px	399
Rail surface defect dataset	[11]	2	variable size	195

The dataset derived from the NEU surface defect database was gathered and generated by several Northeastern University teachers. Six surface defect types were demonstrated with each type entailing 300 picture samples with a total of 1800 grayscale pictures provided through the bounding box. The picture size was 200*200 pixels. The datasets implied rolling scale

(RS), cracking (CR), pitting surface (PS), plaque (PA), inclusion (in), and scratch (SC). As shown in Fig. 1. Multiple defect types inevitably appeared on the metal workpiece surface during the production process [12]. The dataset was employed to train and classify the surface defect deep learning algorithm [13]. Notably, the artificial feature extraction approach could be integrated with the deep learning algorithm for optimal workpiece classification accuracy amidst insufficient samples.

The Severstal steel defect dataset provided by Severstal steel entails four strip steel surface defect types to locate and classify surface defects on steel plates. A total of 12568 training and 5506 test sets were identified with an image size of 1600*256 [14]. This dataset was analyzed with a per-pixel basis evaluation [15] with the potential as a high-quality defect detection baseline. As such, the deep learning model demonstrated better generalization and higher prediction accuracy in forecasting steel plate surface defects with the Severstal dataset [16]. The deep learning algorithm trained the data and optimized the defect detection learning model for gap identification compared to the other algorithms. In this vein, the deep learning algorithm model was structured while enhancing detection accuracy.

The DAGM 2007 dataset encompasses 10 defect image types with each containing 575 training images and 575 test images: training and test sub-datasets with the same size and a distinct number of label images, as shown in Fig. 2. Every picture encompassed the images saved in grayscale 8-bit PNG format for weak supervised industrial optical detection learning and training. The variance between dataset images proved to be minimal. The recognition algorithm model requirements are considerably high in accurately classifying the defects despite the presence of label files [9]. The accuracy and speed assessment of the fabric defect detection model algorithm using the dataset [17] catalyzed the elevation from the low-resolution feature map to the high-resolution fusion feature counterpart with iterations for optimal prediction outcomes.

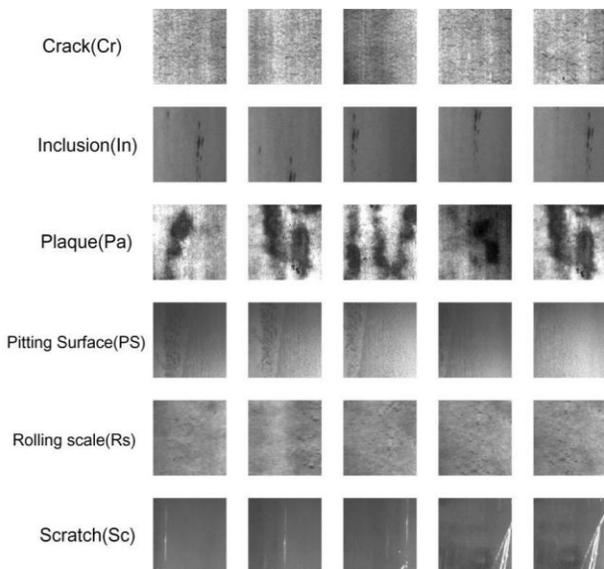


Fig. 1. The Six NEU Types of Surface Defect Images.

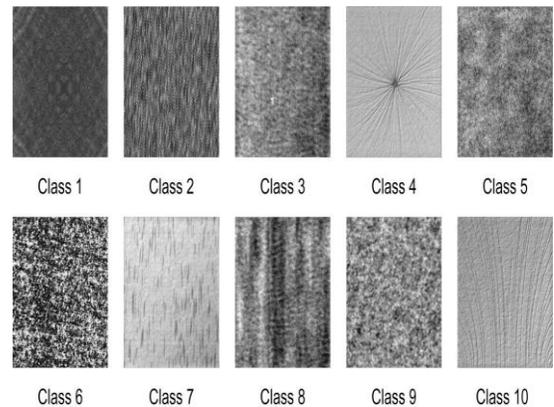


Fig. 2. DAGM Dataset:10 Types of Surface Defect Images.

The Kolektor surface defect dataset primarily gathered defect electronic commutator images, which were subsequently collected under uniform illumination. Labeled pictures were also provided, as shown in Fig. 3. The variance between the pictures proved to be minimal with only one defect picture in every defect type, such as a workpiece with small damages or cracks that are challenging to identify with human eyes. Only 52 defect images were visible in the entire dataset [10], thus significantly facilitating workpiece defect detection tasks.

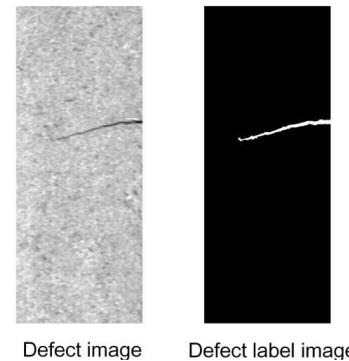


Fig. 3. Kolektor Defect Picture and Labeled Defect Picture.

The rail surface defect dataset implies the train track defects (train track surface crack image) marked by track surface inspection experts. The dataset encompassed 195 challenging images with every image entailing at least one defect and complex and noisy background [18]. Test and prediction datasets could be offered for the deep learning algorithm model to completely detect complex background information defects.

The detection and classification performance of the deep learning algorithm model correlates to image quality, which would then impact performance indicators involving model classification accuracy. Sample incongruence would also influence the classification outcomes with a substantial variance between training and test image quality [12]. Perceivably, the model prediction impact was associated with the model itself and the dataset [13]. More high-quality image data samples could facilitate optimal deep learning model algorithm development for high classification and detection performance.

IV. TRADITIONAL APPLICATION METHODS OF IMAGE RECOGNITION

Image recognition denotes computerized image processing, analysis, and understanding to determine multiple target and object types. This approach implies a practical deep learning algorithm application [19], such as online workpiece recognition for grinding burn and wheel wear following a self-clustering neural network [20]. Feature matching was primarily utilized for workpiece object recognition in the preliminary stage. Computer key visual features [21] were employed for high detection rates. Meanwhile, Salve et al. recommended a means of shape measurement for object recognition [22]. Dalal et al. utilized the histogram of gradient (HOG) descriptor to compute each stage impact on its performance [23]. Effective gray and rotation invariant texture classification techniques were also incorporated in the early stage under the local binary mode [24]. Tuzel et al.'s research integrated region descriptors with target detection and texture classification [25]. As an astute classification system under machine vision, it essentially classifies the peeled open heart fruit core and shell [26].

K. Xia et al. structured a workpiece sorting system in line with a machine vision industrial robot to complete the sorting operation and fulfill subsequent requirements using image edge detection [27]. Some of the applications depicted in bottleneck identification, which varied from current intuitive approaches, structured a bottleneck identification model following the shortest completion delay time for the overflow load computation of every machine to fulfill each workpiece delivery and optimally determine bottlenecks [28]. Y. Guan et al. employed the affine scale invariant feature transformation (a-sift) technique to identify the rough matching feature points between the assessed and planned workpiece towards workpiece identification by making the identified affine change [29]. Hu's invariant moment was implemented to complement the extracted contour with the target counterpart within the template image for target workpiece identification [30].

Regarding workpiece detection and recognition, conventional approaches typically require manual feature selection and extraction to outline the features as vectors and utilize the similarity measurement function to match the (i) workpiece feature vectors to be identified and (ii) template workpiece [31]. The advent of image recognition remains stunted given its inapplicability in big-scale industries following the low efficiency of conventional sliding window approaches and feature robustness. Table II compares six typical methods and their subsequent categories.

Although conventional image recognition techniques primarily outline objects with artificially-designed features, it is deemed impossible to manually extract rich image feature information from objects with intricate feature designs, thus challenging the recognition problem. As such, a data-driven approach (convolutional neural learning network) proves necessary for image feature data comprehension and processing. The image classification approach with a convolutional neural network could derive the target feature value from the (i) image that is challenging in manual feature extraction or the (ii) image dataset encompassing significant noises compared to the traditional image recognition

counterpart. For example, the workpiece dataset in the industrial field demonstrated good robustness to the training and recognition image upon deriving this eigenvalue with the convolution neural learning network. The extracted feature sequence was simultaneously conveyed to the deep neural learning network, which could further elicit the fuzzy features in the image convolution features and forecast the labeled workpiece image. Workpiece image recognition under the convolutional neural network could integrate the two-step workpiece detection and recognition into one, efficiently determine the novel workpiece information encompassed in the image, and save model space and computation with vital and practical significance for project implementation. Convolutional neural network, a deep neural network with a convolution structure, has recently been incorporated into multiple image recognition scenarios. This network inputs the original image into the network. Every network node conveys the image data post-data pre-processing and outputs the probability distribution on the category label with layer-by-layer weight iterative update and computation.

TABLE II. TRADITIONAL METHODS OF TYPICAL IMAGE RECOGNITION

Methods	Reference	Recognition Object	Advantages	Disadvantages
SVM classifier	[32]	tomato	Detect with the least number of features	Extensive color analysis was required
Moment invariant method for edge detection	[33]	workpiece	Able to measure the roughness of a rotating workpiece	Specific measurement environments were required
Template matching method	[34]	workpiece	Simple, practical, and fast recognition speed	The recognition effect was deemed poor for the workpiece with ambiguous feature points
Feature extraction classification	[35]	workpiece	Effectively resolve the problems of unclear features and inaccurate recognition in a complex environment	The recognition robustness was deemed poor without fulfilling the universality of recognition
ROI extraction method	[36]	Air-bearing workpiece	A block matrix was developed using reference pixels and tolerances	Reference pixels were required
Method based on Zernike Moments	[37]	Complex workpiece	Insightful to detect ignored loading attachments and inaccurate assembly positions	The identified object needed to be located

V. IMAGE RECOGNITION METHOD BASED ON DEEP LEARNING

Deep learning constitutes a subclass of machine learning in traditional techniques. Object image feature extraction heavily relies on the manually-designed feature extractor, which requires expert designer knowledge to conduct intricate parameter adjustment process experiments in the model. Notably, the developed model could only determine objects in a particular environment with low generalization and robustness. The number of image feature parameters permitted in the feature extractor design is restricted following the developers' manual adjustment of model parameters. As a branch of artificial intelligence, deep learning neural network reflected higher adaptability with the advent of artificial intelligence as opposed to conventional machine vision techniques. This network is deemed more extensive in the universality of article recognition given that the deep learning algorithm primarily entails data-driven image feature extraction for a deeper, more efficient, and accurate representation of the image dataset using the image learning of big samples compared to the conventional method. A series of image recognition techniques under deep learning could attain highly precise and optimal recognition to resolve multiple intricate image recognition scenarios. This section emphasizes four classical segmentation approaches based on deep learning: AlexNet [38-40], Yolo [41-43], VGG net [44-46], and ResNet [47-49].

A. AlexNet

Hinton's and Alex Krizhevsky's revolutionary AlexNet neural network (AlexNet) algorithm [38-40] championed the 2012 Imagenet competition. Specifically, Imagenet entails a large image recognition database encompassing marked pictures. AlexNet focuses on the full connection layer function with a total of eight layers: five convolutions and three full-connection. In a three-channel color map with 227 pixels in length and width ($227 \times 227 \times 3$), the image is incorporated into the first layer to be convoluted into $11 \times 11 \times 3$. Every convolution kernel generates a novel pixel while all the convolution kernels subsequently slide through the $227 \times 227 \times 3$ pixel picture with a stripe of four. Following the convolution output layer resolution computation, the convolution pixel layer data is duly computed with a convolution output of $55 \times 55 \times 96$ in the first layer. The total convolution parameters of the first layer imply 35K as only the convolution kernel in the convolution layer entails neural network parameters post-calculation. Meanwhile, the second-layer characteristic map is transmitted to the third counterpart until the seventh-layer output data is fully connected with 1000 neurons from the eighth counterpart. The outcome was generated through softmax, which was utilized as a 1000 input image category for the classification score with the following attributes: (1) AlexNet algorithm converted the traditional neuron activation function $f(x) = (1 + e^{-x})$ to $f(x) = \max(0, x)$ with a rectified linear unit (relu) as an activation function that was extensively utilized in artificial neural network. The typical four-layer network with relus and tanh as an activation function attains the faster convergence speed effect involving relu in the CIFAR-10s experiment dataset compared to the conventional tanh activation function; (2) AlexNet utilized two techniques to

resolve over-fitting issues: data enhancement and dropout. The original picture was cropped to be employed as network input in data enhancement while dropout was utilized to deter over-fitting and promote effective fusing. Regardless, the network model computing cost is exorbitant despite the feasible computation following the use of a graphics processor (GPU) in the training process.

B. YOLO

The You Only Look Once (Yolo) revolutionary neural network algorithm [41-43] resolves object detection (a regression problem) to avoid several reiterated prediction works and complete the input from the original image to the output of the image category following a separate end-to-end network. Yolo entails specific prerequisites to incorporate the image input size into the network, scale the image size to the specified size, classify the picture into $S \times S$ grid, and make predictions in every small grid. Based on Fig. 4, the category probability forecasted by each grid and the confidence predicted by each box were multiplied until the score correlated to every box and category. The non-maximum suppression approach was then utilized to derive the classification outcomes. Essentially, Yolo is deemed beneficial as it disregards the extraction process of region proposal and rapidly identifies objects with minimal background error detection rates and inaccurate background knowledge. This algorithm implies high generalization, which is unlikely to crash when incorporated into fields or unforeseen inputs. Notwithstanding, the error-prone $S \times S$ grid at the frame regression stage leads to inaccurate object positioning. A large missed detection rate is identified in the presence of multiple small targets in a network. The subsequent Yolo version continues to rectify such complexities. Specifically, Yolo V3 elevates detection performance, particularly in the multi-scale fusion approach, to resolve low detection performance post-defect optimization. On another note, a cross-layer connection is presented to optimize small target detection performance.

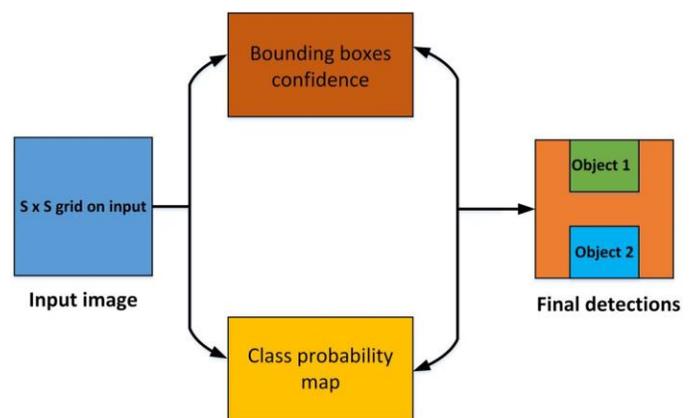


Fig. 4. Yolo Diagram Algorithm Process.

C. VGG Net

The VGG net [44-46] model is characterized by a substantial number of layers, including multiple network layers ranging in depth between 11 and 19. The deep learning-model performance correlation is examined to enhance the overall recognition performance by improving the network layer

depth. This model aims to transform the convolution of the larger core layer into multiple smaller-layered convolution cores. Vggnet-16 and Vggnet-19 are extensively employed to render the entire network to be highly effective. The VGG also denotes a five-layer convolution and two full-connection layers for image feature extraction and one full connection layer for feature classification akin to the AlexNet framework. Fig. 5 illustrates the Vggnet-16 network structure diagram. The convolution layer kernel in the Vgg net structure is 3*3. Three groups of 3 *3 convolution layers connected with a 1*7*7 kernel were employed with the same effect. In terms of model benefits, the number of parameters is duly reduced. The original parameter ($C*7*7$) was transformed into $3*C*3*3$ for a convolution layer with C kernels. Despite the presence of more parameters and deeper levels, VGG requires lesser iterations to initiate convergence given the depth and small filter size function as the post rule while the pre-initialization operation is performed on some layers. Such advantages could increase the non-linear correction layer, mitigate gradient disappearance and over-fitting issues, and optimize the model training speed. The network structure attributes simultaneously facilitated it to regulate the number of parameters while eliciting more image features to prevent over-computation and structure intricacy.

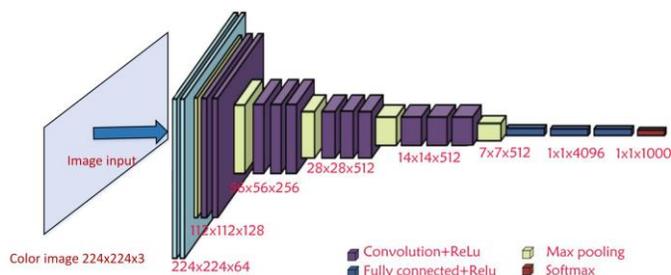


Fig. 5. Network Structure of VGG16 Adapted from [45].

D. ResNet

The proposed ResNet [47-49] algorithm model resolves the deepening of network layers despite the initially-enhanced accuracy. This precision would worsen if the number of network layers continues to increase. In other words, the network model could alleviate the degradation problem in network training. Regarding the conventional deep learning algorithm, the layers to which the gradient could not be conveyed are not trained when the number of network layer increases. Thus, the effect is not as robust as the shallow network with adequate layers as the error rate would rise with the increase in layers. The ResNet algorithm model puts forth the residual module in resolving the degradation problem. Based on the notion underpinning this method, the network at layer N is derived from the network at layer N-1 with conversion. It is connected to the upper-layer network for

gradient propagation, which subsequently resolves the gradient disappearance caused by the neural network passing through depth. Thus, the residual structure is presented. Following Fig. 6, the output layer $H(x) = F(x)$ is changed to $H(x) = F(x) + x$ where the network loss function $f(x, w)$ is extended to the multi-layer neural network using the back propagation gradient value formula. The front-layer network gradient becomes smaller with the increased number of layers 'n' in the neural network and the return of errors. Thus, the gradient would not disappear even with multiple network layers. On another note, ResNet could develop the residual module in the form of a small kernel with the other layers utilizing full convolution excluding the full connection layer for classification, which could substantially optimize the calculation speed. This residual block structure method is employed for reference as ResNet entails multiple algorithm types with a network depth of 50, 101, and 152. The ResNet model performance significantly varies with distinct sizes. Overall, this model needs to be structured based on the actual application context.

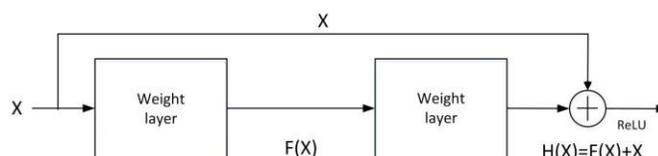


Fig. 6. Residual Block Structure.

VI. PERFORMANCE COMPARISON

Based on recent research, image recognition technology could flexibly select multiple algorithms for recognition based on various application scenarios in actual recognition tasks regarding deep learning. Some might even need to integrate different recognition methods in obtaining the most optimal recognition accuracy. It is deemed pivotal to develop a set of robust and stable recognition algorithms with high market value and application possibilities following the intricate recognition environment, which is subject to more illumination or other interference. Fundamentally, relevant literature adopts four parameters to compute the following metrics (accuracy, precision, recall, and F1-score) as the evaluation indices of algorithm advantages and disadvantages. As the primary problem to be solved, accuracy could further optimize the empirical depth in the image recognition field with a highly positive effect on the efficient incorporation of multiple technologies and the development of relevant disciplines. This study primarily summarized the deep learning technology application to image recognition. Table III presents (i) the recent comparison of model precision and loss of value in deep learning algorithms, (ii) evaluations on the technological complexities encountered in the workpiece image classification application, and (iii) specific study gaps.

TABLE III. DEEP LEARNING ALGORITHM COMPARISONS

Literature	Model	Object	Accuracy	Loss rate
[50]	LeNet-5	Oil–water two-phase flow	88%	High (loss rate)
	VGG-16	Oil–water two-phase flow	98.3%	Median (loss rate)
	AlexNet	Oil–water two-phase flow	99.3%	Low (loss rate)
[41]	YOLO	Insulator	90%	3.8% (class fault)
	Fast R-CNN	Insulator	84%	4.1% (class fault)
[48]	ResNet-18	Apple leaf	99%	8.9%
	ResNet-34		97%	1.23%
	VGG		89%	2.43%
[51]	VGG-16	Chest X-ray images	97%	NA
	ResNet		86.7%	
	ResNet-50		98%	
	AlexNet		96.5%	
	ResNet-50+SVM		95.38%	
[40]	Alexnet	Power equipment	83.55%	NA
[52]	YOLO	Leucocyte	93.7%	NA
[53]	YOLO	Resistors and capacitors	92.57%	0.87% (error rate)
[47]	ResNet50	PatchCamelyon	98.8%	98.5% (AUC-ROC score)
	Vgg-16		95.9%	95.2% (AUC-ROC score)
	Vgg-19		97%	96.1% (AUC-ROC score)
[54]	ResNet-50	Bacterial image	99.9%	0.02% (loss rate)
	ResNet-34		99.3%	0.04% (loss rate)
[55]	AlexNet	Wireless capsule endoscopy	95%	NA
[56]	AlexNet	MNIST	98%	NA
[57]	GoogLeNet	Colour images	NA	90.31% (FMI)
	Alexnet		NA	87.69% (FMI)
	VGG16		NA	89.78% (FMI)
[58]	Deep Transfer Learning	Tongue images	95.92%	NA
[59]	LeNet	Microorganism image	98.66%	NA
[60]	transfer learning approach	Flower image	98.6%	98% (recall)
[61]	VGG16	Fundus images	91.3%	NA
[62]	RS-oriented error-tolerant deep learning	Noisy remote sensing image	95.9%	NA
[63]	deep learning and image recognition	Brake pad contour image	98%	NA
[64]	self-supervised super sample decomposition for transfer learning	COVID-19	99.8%	99.7% (sensitivity)
[65]	Meta-learning	Classical image	92.48%	NA

Notes: NA - Not Available

The accuracy rate in Table III denotes the percentage of the number of correctly predicted samples in the total number of samples or the proportion of precisely-predicted outcomes in the total number of samples. The formula of accuracy is shown in (1), essentially, TP implies a positive prediction with an actual prediction that is true. The TN denotes a negative and true forecast. The FP demonstrates a positive forecast with an actual prediction that is false. The FN denotes the forecast to be negative with an actual prediction that is false. The FP and FN reflect prediction error values, which means that the pre-determined target is not discovered. Several relevant detection indicators (accuracy) imply the proportion of accurate forecasting outcomes in all positive predictions. Recall rate denotes the proportion of accurately-predicted findings in all positive occurrences. In deep learning classification and detection tasks, the precision of appropriate techniques is typically assessed based on statistical findings. High TP and TN reflect high precision and optimal detection impacts of the deep learning model.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Based on Table III, the deep learning algorithm model accuracy is fundamentally between 80% and 99%. The designed model evaluation accuracy is significantly enhanced compared to conventional techniques, thus implying the application effect of the deep learning model to be more ideal. Despite the diversification of model evaluation methods in theoretical and practical research, such studies remain considerably scattered given the primary utilization of the accuracy evaluation index for model assessment. Accuracy implies the accurately-forecasted image proportion in objective evaluation and classification. Although the high precision value in outcome detection or classification assumably reflects a high recall value, this evaluation index might prove contradictory in some cases. For example, an accurately-forecasted outcome in prediction would demonstrate a 100% precision value albeit with a significantly low recall value. Meanwhile, all the outcomes returned with a recall of 100% would denote substantially low precision. Hence, most studies do not employ such evaluation predictors for model assessment.

Despite there being no unified recognition technique for small target workpiece identification in workpiece recognition, relevant researchers have recommended multiple detection methods under deep learning to resolve the problems associated with small target detection. Notwithstanding, different study objects could ascertain whether precision or recall is high based on the required judgment following relevant research.

Detection approaches have undergone continuous optimization to manage multiple classification and detection problems involving deep learning image datasets and fulfill the application requirements in actual scenarios. The high-dimensional semantic features of image data could be elicited by convolution and non-linear layers, which is much better than traditional detection performance approaches. Image is more vulnerable to noise, thus significantly increasing workpiece classification and interpretation complexities in the industrial field. It is deemed pivotal to determine how to fully

utilize the evaluation index information of in-depth learning in serving the workpiece identification and classification application requirements and alleviate data processing intricacies for optimal workpiece classification, detection, and interpretation within the industrial field.

VII. RESEARCH GAP

The research gaps are summarized in Table III.

1) The recognition algorithm still requires a specific optimization level, specifically through image feature extraction, to further enhance the image recognition rate following experimental data, computing equipment, and research time issues. Such image features could be conveniently extracted by the convolution filter following [50]. Empirically, the accuracy of the three aforementioned deep learning network algorithms proves relatively high. Simultaneously, a small number of target image misrecognition would not impact the traffic pattern recognition outcomes. Literature [61] could substantially optimize the recognition and classification rate with the image feature extraction approach.

2) The number of deep learning network parameters consumes much time, power, and hardware resources in the actual training process, which complicates the neural network application. The benefits of speed and accuracy in algorithm recognition are highlighted in [41]. Recognition accuracy could be further optimized given the increased number of trained images despite its time consumption and high cost.

3) In terms of computing power limitation, a higher image resolution could enhance the number of extracted image features through the algorithm given the low image resolution in the training dataset. Following [47], an optimized classical depth learning algorithm model could significantly enhance low image resolution performance. Test time augmentation (TTA) was adopted to improve the image while the derived image prediction accuracy data proved better than other depth learning algorithm models.

4) Concerning workpiece classification and recognition algorithms, the newly-proposed algorithm performance could be assessed and examined to accelerate the iterative update speed of the theoretical algorithm. The number and type of database samples failed to fulfill the prerequisites and match the workpiece defects generated by practical applications following the inadequate public datasets of some types based on the current workpiece defect database. Most of the algorithms could not be fairly compared given the absence of an agreed database standard [66]. Despite the presence of extensively utilized workpiece defect datasets, such as NEU [7], UCI [67], and Rail surface [68], the odious industrial setting poses substantial complexities to the workpiece defect image dataset of the actual industrial production line.

5) Deep learning training requires sufficient training datasets in the workpiece image dataset production. It is deemed necessary to ascertain the means of training a model that could precisely detect images through a restricted number of sample datasets. Data optimization served to enhance

segmentation accuracy. The deep learning benefits minutely differed across multiple evaluation indicators with high precision. Fundamentally, the evaluation indices were not extremely low.

6) Several deep learning algorithm models might not apply to specific application scenarios given the emergence of more relevant counterparts. It is deemed feasible to develop a novel deep learning algorithm to enhance the impact of workpiece classification and recognition. For example, [69] integrated the computational advantages of a 2D FCN network and the ability to resolve 3D spatial consistency without influencing segmentation accuracy. Given the palpable limitations, the resolution of recognized pictures and speed calculation require improvement as network calculation precision serves to increase with the number of computations. This accuracy could be improved [40] through integrations with other methods albeit with low precision in some image recognition scenarios. Despite the improved effects, some room for improvement is still available. Regarding parameter optimization, [65] proposed that tests could be performed on non-trained tasks. Meanwhile, classification could be realized even with minimal training: a direction worthy of the effort.

VIII. RESEARCH PROPOSAL

1) Workpiece image classification algorithms could strive to fuse image feature extraction into one step and enhance recognition accuracy. Regarding the recognition target limitation, the extracted image required a strong expression while the recognition and classification rate proved relatively low. This situation adversely influenced the recognition of intricate or unclear images given its complexity in feature image extraction. The algorithm model could be learned in depth from the noise dataset with the possible identification of low feature target images in line with [62], thus broadening the deep learning application range. Specifically, [63] employed the edge feature extraction approach to determine the internal and external features of the recognition target image contour. The image features were distinguished post-feature point extraction. The deep learning algorithm model was subsequently presented for feature training towards high recognition accuracy. Meanwhile, [56, 57] utilized the deep learning method to fuse image features and resolve the target image feature ambiguity or relatively indistinguishable recognition rates.

2) Given the presence of issues involving extensive training and prediction periods, future workpiece image classification algorithm studies could consider how to mitigate network model redundancy, optimize the number of network layers, and shorten the computation time while simultaneously ensuring recognition accuracy to some extent. Based on [48], shallow networks could also reflect optimal recognition accuracy and low error rates with even better recognition impacts than deep networks. Palpable target detection errors or reiterated identification in employing the same algorithm for target detection and limited training time resembles the

drawbacks highlighted in conventional neural networks following [52, 53].

3) Deep learning algorithm could further enhance recognition accuracy with optimal hardware and image acquisition. AS affirmed by [59], low resolution and multiple datasets could be utilized for the deep learning algorithm to realize classification. The usage of more epochs improved training accuracy while surface accuracy could be optimized by the number of iterations. Regarding the disadvantage, the loss value would be too high in the iterative process. The same deep learning method demonstrated minimal variance in recognition accuracy under multiple levels, specifically in small target images, with improved computer algorithms [54].

4) Further enhance network generalization (particularly in restricted datasets), optimize small dataset detection, and integrate the conventional recognition. The means of developing an algorithm with strong applicability require further examination for thousands of object types. Regarding limited sample collection, a more widely-disseminated collection database was established in the public counterpart following [51] and a dataset in line with their actual application scene by gathering six chest X-ray image databases. The gathered database should be relevant to the research. Detection algorithm with convolution neural network to improve microfeature extraction and segmentation capacity and model accuracy in the follow-up. As deep learning technology was employed for dataset pre-processing albeit with relatively ideal experimental effects.

5) It is also rendered possible to optimize moving image recognition, develop a recognition model towards dataset expansion, and enhance the recognition model adaptability to the actual industrial setting in the process of moving image recognition. For example [55] could further determine the range of textures and features, such as color while [58] employed the deep transfer learning technique to assess tongue images and resolve the complexities in gathering adequate marker image samples. The recommended approach implied better classification accuracy. Essentially, [70] outlined the mapping relationship between image classification input feature vector and image category and structured a moving image recognition model.

6) The incorporation of deep learning algorithm benefits into specific fields or different but relevant disciplines or problems has garnered more attention towards completing or improving the learning effect of target fields or tasks. Ensemble learning is a promising and experimentally-proven technology. Based on [60], deep learning approaches significantly influence intricate tasks, such as image feature extraction, segmentation, and semantic classification. Meanwhile, [64] mitigated the network complexity by pruning the redundant depth algorithm model parameters to derive a small and efficient classification model, enhance the runtime reasoning speed of neural network, and elicit the favored classification effect with minimal calculation and

computational workload in the future. The incorporation of meta-learning into deep learning denotes a viable method.

IX. CONCLUSION

This study reviewed the recent development of deep learning in image recognition, emphasized specific deep learning image method types, summarized and compared similar algorithm performance, examined the advantages and disadvantages of every deep learning algorithm based on different application scenarios, and flexibly selected deep learning classification methods to effectively improve the recognition effect. highlighting six research proposals for workpiece image recognition and defect detection. Empirical solutions are provided for the selection of future deep learning models for artifact classification and defect detection.

Many actual influencing factors and complex situations are present in workpiece recognition. In other words, it is challenging to apply the model derived by conventional methods to the actual circumstance. Research on the deep convolution neural network given its prevalence in computer vision tasks has made a significant breakthrough, thus proving the potential of deep learning in image classification. There are also limitations in the research of workpiece recognition, in terms of theoretical experiments on images, there is no research on image acquisition and recognition of workpieces in practical application scenarios, and it is necessary to build recognition systems to carry out research on practical workpiece recognition applications based on multiple factors such as light, angle and placement position.

In the future work, more methods will be reviewed to enhance the generalization ability of the model and improve the practical application ability of the model in the industrial field, which is also the improvement direction proposed by the future in-depth learning research work. At the same time, it can effectively and reducing the dataset is also one of the priorities of the future work.

REFERENCES

- [1] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955," *AI magazine*, vol. 27, no. 4, pp. 12-12, 2006.
- [2] P. TianY, WangX, "Deep Learning Strong Parts for Pedestrian Detection," presented at the IEEE International Conference on Computer Vision, 2016.
- [3] S. I. Krizhevsky, A. Hinton, G. E., "ImageNet classification with deep convolutional neural networks," in *Proceedings of the International Conference on Neural Information Processing Systems* vol. 1097-1105, ed, 2012.
- [4] R. Kaluri and C. Reddy, "An Enhanced Framework for Sign Gesture Recognition using Hidden Markov Model and Adaptive Histogram Technique," *International Journal of Intelligent Engineering and Systems*, vol. 10, no. 3, pp. 11-19, 2017.
- [5] R. Kaluri and C. H. P. Reddy, "Optimized feature extraction for precise sign gesture recognition using Self-Improved Genetic Algorithm," 2018.
- [6] S. S. Pandarge and V. R. Ratnaparkhe, "Image Processing based Linear Discriminant and Quadratic Discriminant Classifier for Feature Extraction Models," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 15-17 May 2019 2019, pp. 1502-1505, doi: 10.1109/ICCS45141.2019.9065388.
- [7] B. Zhao, M. Dai, P. Li, R. Xue, and X. Ma, "Defect Detection Method for Electric Multiple Units Key Components Based on Deep Learning," *IEEE Access*, vol. 8, pp. 136808-136818, 2020, doi: 10.1109/ACCESS.2020.3009654.
- [8] Z. Xinzi, "BSU-Net: A Surface Defect Detection Method Based on Bilaterally Symmetric U-Shaped Network," in *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, 25-27 Dec. 2020 2020, pp. 1771-1775, doi: 10.1109/ICMCCE51767.2020.00388.
- [9] H. Zhou, B. Jang, Y. Chen, and D. Troendle, "Exploring Faster RCNN for Fabric Defect Detection," in *2020 Third International Conference on Artificial Intelligence for Industries (AI4I)*, 21-23 Sept. 2020 2020, pp. 52-55, doi: 10.1109/AI4I49448.2020.00018.
- [10] Y. Li and J. Li, "An End-to-End Defect Detection Method for Mobile Phone Light Guide Plate via Multitask Learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-13, 2021, doi: 10.1109/TIM.2021.3053987.
- [11] D. Zhang, K. Song, Q. Wang, Y. He, X. Wen, and Y. Yan, "Two Deep Learning Networks for Rail Surface Defect Inspection of Limited Samples With Line-Level Label," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 10, pp. 6731-6741, 2021, doi: 10.1109/TII.2020.3045196.
- [12] T. Benbarrad, L. Eloutouate, M. Arioua, F. Elouaai, and M. D. Laanaoui, "Impact of Image Compression on the Performance of Steel Surface Defect Classification with a CNN," *Journal of Sensor & Actuator Networks*, Article vol. 10, no. 4, pp. 73-73, 2021, doi: 10.3390/jsan10040073.
- [13] X. Feng, X. Gao, L. Luo, C. Humbert, and R. B. Navarro, "X-SDD: A New Benchmark for Hot Rolled Steel Strip Surface Defects Detection," *Symmetry* (20738994), Article vol. 13, no. 4, p. 706, 2021, doi: 10.3390/sym13040706.
- [14] J. Božič, D. Tabernik, and D. Skočaj, "End-to-end training of a two-stage neural network for defect detection," in *2020 25th International Conference on Pattern Recognition (ICPR)*, 10-15 Jan. 2021 2021, pp. 5619-5626, doi: 10.1109/ICPR48806.2021.9412092.
- [15] F. Akhyar, C. Y. Lin, G. S. Kathiresan, B. Surianarayanan, and C. Y. Hsu, "Detectors++: The Robust Baseline for a Defect Detection System," in *2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, 15-17 Sept. 2021 2021, pp. 1-2, doi: 10.1109/ICCE-TW52618.2021.9603063.
- [16] M. Abu, A. Amir, Y. H. Lean, N. A. H. Zahri, and S. A. Azemi, "The Performance Analysis of Transfer Learning for Steel Defect Detection by Using Deep Learning," *Journal of Physics: Conference Series*, vol. 1755, no. 1, p. 012041, 2021/02/01 2021, doi: 10.1088/1742-6596/1755/1/012041.
- [17] H. Dong, K. Song, Y. He, J. Xu, Y. Yan, and Q. Meng, "PGA-Net: Pyramid Feature Fusion and Global Context Attention Network for Automated Surface Defect Detection," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7448-7458, 2020, doi: 10.1109/TII.2019.2958826.
- [18] M. Niu, K. Song, L. Huang, Q. Wang, Y. Yan, and Q. Meng, "Unsupervised Saliency Detection of Rail Surface Defects Using Stereoscopic Images," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2271-2281, 2021, doi: 10.1109/TII.2020.3004397.
- [19] X. Caiyun, "Overview of image recognition technology," (in China), *Computer knowledge and technology*, vol. (10), pp. 2446-2447, 2013,9.
- [20] J. Shi, X. Zhang, B. Zhong, and H. Ren, "A study of on-line identification of grinding burn and wheel wear based on self-clustering neural network," in *IEEE International Conference on Industrial Technology*, 1996.
- [21] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 8-14 Dec. 2001 2001, vol. 1, pp. I-I, doi: 10.1109/CVPR.2001.990517.
- [22] S. G. Salve and K. C. Jondhale, "Shape matching and object recognition using shape contexts," in *2010 3rd International Conference on Computer Science and Information Technology*, 9-11 July 2010 2010, vol. 9, pp. 471-474, doi: 10.1109/ICCSIT.2010.5565098.
- [23] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, 2005.

- [24] T. Ojala, M. Pietikinen, and T. Menp, "Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns," European Conference on Computer Vision, 2000.
- [25] O. Tuzel, F. Porikli, and P. Meer, "Region Covariance: A Fast Descriptor for Detection and Classification," Springer, Berlin, Heidelberg, 2006.
- [26] H. Nouri-Ahmadabadi, M. Omid, S. S. Mohtasebi, and M. S. Firouz, "Design, development and evaluation of an online grading system for peeled pistachios equipped with machine vision technology and support vector machine," *Information Processing in Agriculture*, vol. 4, no. 4, p. 9, 2017.
- [27] X. Ke and Z. Weng, "Workpieces sorting system based on industrial robot of machine vision," in *International Conference on Systems & Informatics*, 2016.
- [28] D. Wen, H. Li, and Z. Aixia, "A bottleneck resource identification method for completing the workpiece based on the shortest delay time," in *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC)*, 8-10 Aug. 2011 2011, pp. 9-13, doi: 10.1109/AIMSEC.2011.6010180.
- [29] Y. Guan, J. Tian, and Y. Tang, "Workpiece Identification and Localization Based on Monocular Vision," in *2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2019.
- [30] C. Huang, C. Dan, and X. Tang, "Implementation of Workpiece Recognition and Location Based on Opencv," in *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, 2016.
- [31] Y. Wei, "Research on recognition and pose estimation methods for low texture workpiece," Changzhou University, [D] 2021.
- [32] A. Sdk, B. Se, C. Sb, and B. Bk, "A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier," *Microprocessors and Microsystems*, vol. 76, 2020.
- [33] B. M. Kumar and M. M. Ratnam, "Machine vision method for non-contact measurement of surface roughness of a rotating workpiece," *Sensor Review*, vol. 35, no. 1, pp. 10-19, 2015.
- [34] Z. Liu, P. Wan, L. Ling, L. Chen, and W. Zhou, "Recognition and Grabbing System for Workpieces Exceeding the Visual Field Based on Machine Vision," *Jiqiren/Robot*, vol. 40, no. 3, pp. 294-300 and 308, 2018.
- [35] X. X. L. Wang H, Zuo Y B, "HL-S workpiece recognition algorithm based on multi-feature fusion," *Journal of Electronic Measurement and Instrument*, vol. 33(12), pp. 94-99, 2019.
- [36] P. Kunakornvong and P. Sooraksa, "Machine Vision for Defect Detection on the Air Bearing Surface," in *2016 International Symposium on Computer, Consumer and Control (IS3C)*, 4-6 July 2016 2016, pp. 37-40, doi: 10.1109/IS3C.2016.20.
- [37] B. Guotang, C. Yuansong, Q. Renchao, and H. Wei, "An Defect Detection Method Based on Zernike Moments for Industrial CT Image," in *2007 8th International Conference on Electronic Measurement and Instruments*, 16-18 Aug. 2007 2007, pp. 2-831-2-835, doi: 10.1109/ICEML.2007.4350809.
- [38] Q. Wang, J. Xu, T. Sun, Z. Lv, and G. Zong, "Research on Road Adhesion Condition Identification Based on an Improved ALexNet Model," *Journal of Advanced Transportation*, Article pp. 1-14, 2021, doi: 10.1155/2021/5531965.
- [39] C. Wang, L. Yu, X. Zhu, J. Su, and F. Ma, "Extended ResNet and Label Feature Vector Based Chromosome Classification," *IEEE Access*, vol. 8, pp. 201098-201108, 2020, doi: 10.1109/ACCESS.2020.3034684.
- [40] Y. Guo, Z. Pang, J. Du, F. Jiang, and Q. Hu, "An Improved AlexNet for Power Edge Transmission Line Anomaly Detection," *IEEE Access*, vol. 8, pp. 97830-97838, 2020, doi: 10.1109/ACCESS.2020.2995910.
- [41] S. Wang, L. Niu, and N. Li, "Research on Image Recognition of Insulators Based on YOLO Algorithm," in *2018 International Conference on Power System Technology (POWERCON)*, 6-8 Nov. 2018 2018, pp. 3871-3874, doi: 10.1109/POWERCON.2018.8602149.
- [42] G. Li, Z. Song, and Q. Fu, "A New Method of Image Detection for Small Datasets under the Framework of YOLO Network," in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 12-14 Oct. 2018 2018, pp. 1031-1035, doi: 10.1109/IAEAC.2018.8577214.
- [43] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 27-30 June 2016 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
- [44] M. Ye et al., "A Lightweight Model of VGG-16 for Remote Sensing Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6916-6922, 2021, doi: 10.1109/JSTARS.2021.3090085.
- [45] H. Ge, Z. Zhu, R. Liu, and X. Wu, "Application of Multiscale Facial Feature Manifold Learning Based on VGG-16," *Journal of Sensors*, Article pp. 1-7, 2021, doi: 10.1155/2021/7129800.
- [46] M. Rezaee, Y. Zhang, R. Mishra, F. Tong, and H. Tong, "Using a VGG-16 Network for Individual Tree Species Detection with an Object-Based Approach," in *2018 10th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)*, 19-20 Aug. 2018 2018, pp. 1-7, doi: 10.1109/PRRS.2018.8486395.
- [47] Z. Zheng, H. Zhang, X. Li, S. Liu, and Y. Teng, "ResNet-Based Model for Cancer Detection," in *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, 15-17 Jan. 2021 2021, pp. 325-328, doi: 10.1109/ICCECE51280.2021.9342346.
- [48] X. Li and L. Rai, "Apple Leaf Disease Identification and Classification using ResNet Models," in *2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT)*, 13-15 Nov. 2020 2020, pp. 738-742, doi: 10.1109/ICEICT51264.2020.9334214.
- [49] A. Mahajan and S. Chaudhary, "Categorical Image Classification Based On Representational Deep Network (RESNET)," in *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, 12-14 June 2019 2019, pp. 327-330, doi: 10.1109/ICECA.2019.8822133.
- [50] D. Meng, Y. Hongyi, C. Xiaoyan, and W. Xinqiang, "Oil-in-Water Two-Phase Flow Pattern Identification From Experimental Snapshots Using Convolutional Neural Network," *IEEE Access*, article vol. 7, pp. 6219-6225, 01/01/ 2019, doi: 10.1109/ACCESS.2018.2888733.
- [51] E. Cortés and S. Sánchez, "Deep Learning Transfer with AlexNet for chest X-ray COVID-19 recognition," *IEEE Latin America Transactions*, vol. 19, no. 6, pp. 944-951, 2021, doi: 10.1109/TLA.2021.9451239.
- [52] Y. C. Yang and W. Chen, "An Improved YOLO Leucocyte Classification and Recognition Method," in *2021 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, 27-28 March 2021 2021, pp. 618-621, doi: 10.1109/ICITBS53129.2021.00157.
- [53] J. Huang and Y. Lu, "A Method for Identifying and Classifying Resistors and Capacitors Based on YOLO Network," in *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*, 19-21 July 2019 2019, pp. 1-5, doi: 10.1109/SIPROCESS.2019.8868519.
- [54] T. Shaily and S. Kala, "Bacterial Image Classification Using Convolutional Neural Networks," in *2020 IEEE 17th India Council International Conference (INDICON)*, 10-13 Dec. 2020 2020, pp. 1-6, doi: 10.1109/INDICON49873.2020.9342356.
- [55] S. Sunitha and S. S. Sujatha, "An Improved Bleeding Detection Method for Wireless Capsule Endoscopy (WCE) Images Based on AlexNet," in *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, 13-14 May 2021 2021, pp. 11-15, doi: 10.1109/ICSPC51351.2021.9451699.
- [56] Y. Lu, "Image Classification Algorithm Based on Improved AlexNet in Cloud Computing Environment," in *2020 IEEE International Conference on Industrial Application of Artificial Intelligence (IAAD)*, 25-27 Dec. 2020 2020, pp. 250-253, doi: 10.1109/IAAI51705.2020.9332891.
- [57] B. LakshmiPriya, N. Pavithra, and D. Saraswathi, "Optimized Convolutional Neural Network based Colour Image Fusion," in *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 3-4 July 2020 2020, pp. 1-4, doi: 10.1109/ICSCAN49426.2020.9262439.
- [58] C. Song, B. Wang, and J. Xu, "Classifying Tongue Images using Deep Transfer Learning," in *2020 5th International Conference on Computational Intelligence and Applications (ICCIA)*, 19-21 June 2020 2020, pp. 103-107, doi: 10.1109/ICCIA49625.2020.00027.

- [59] T. Treebupachatsakul and S. Poomrittigul, "Microorganism Image Recognition based on Deep Learning Application," in 2020 International Conference on Electronics, Information, and Communication (ICEIC), 19-22 Jan. 2020 2020, pp. 1-5, doi: 10.1109/ICEIC49074.2020.9051009.
- [60] N. Alipour, O. Tarkhaneh, M. Awrangjeb, and H. Tian, "Flower Image Classification Using Deep Convolutional Neural Network," in 2021 7th International Conference on Web Research (ICWR), 19-20 May 2021 2021, pp. 1-4, doi: 10.1109/ICWR51868.2021.9443129.
- [61] X. Cheng, X. Feng, and W. Li, "Research on Feature Extraction Method of Fundus Image Based on Deep Learning," in 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), 20-22 Nov. 2020 2020, pp. 443-447, doi: 10.1109/AUTEEE50969.2020.9315604.
- [62] Y. Li, Y. Zhang, and Z. Zhu, "Learning Deep Networks under Noisy Labels for Remote Sensing Image Scene Classification," in IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, 28 July-2 Aug. 2019 2019, pp. 3025-3028, doi: 10.1109/IGARSS.2019.8900497.
- [63] J. W. Chen, "Analysis of Subsampled Image Size for Detection and Identification of Brake Pad Contours by Using Deep Learning," in 2020 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), 23-25 Oct. 2020 2020, pp. 174-176, doi: 10.1109/ECICE50847.2020.9301977.
- [64] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "4S-DT: Self-Supervised Super Sample Decomposition for Transfer Learning With Application to COVID-19 Detection," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 7, pp. 2798-2808, 2021, doi: 10.1109/TNNLS.2021.3082015.
- [65] C. So, "Exploring Meta Learning: Parameterizing the Learning-to-learn Process for Image Classification," in 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), 13-16 April 2021 2021, pp. 199-202, doi: 10.1109/ICAIIIC51459.2021.9415205.
- [66] X. Fang, Q. Luo, B. Zhou, C. Li, and L. Tian, "Research Progress of Automated Visual Surface Defect Detection for Industrial Metal Planar Materials," Sensors, vol. 20, no. 18, 2020.
- [67] A. Abdullahi, N. A. Samsudin, M. R. Ibrahim, M. S. Aripin, and Z. A. Othman, "Towards IR4.0 implementation in e-manufacturing: Artificial intelligence application in steel plate fault detection," Indonesian Journal of Electrical Engineering and Computer Science, vol. 20, no. 1, p. 430, 2020.
- [68] J. Gan, Q. Li, J. Wang, and H. Yu, "A Hierarchical Extractor-Based Visual Rail Surface Inspection System," IEEE Sensors Journal, vol. PP, no. 23, pp. 1-1, 2017.
- [69] J. Duan et al., "Automatic 3D Bi-Ventricular Segmentation of Cardiac Images by a Shape-Refined Multi- Task Deep Learning Approach," IEEE Transactions on Medical Imaging, vol. 38, no. 9, pp. 2151-2164, 2019, doi: 10.1109/TMI.2019.2894322.
- [70] Z. Wang, C. Tang, X. Sima, and L. Zhang, "Research on Application of Deep Learning Algorithm in Image Classification," in 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), 14-16 April 2021 2021, pp. 1122-1125, doi: 10.1109/IPEC51340.2021.9421185.