Received: 19 June 2023 | Revised: 24 September 2023 | Accepted: 24 October 2023

IET Image Processing

DOI: 10.1049/ipr2.12971

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

A DenseNet-based feature weighting convolutional network recognition model and its application in industrial part classification

Kang An | **Xiaoqing Sun** | **Yaqing Song** | **Yebin Lu** | **Qianqian Shangguan**

College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai, China

Correspondence
Yaqing Song and Qianqian Shangguan, College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 201418, China.
Email: yqsong@shnu.edu.cn, shangguan@shnu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 62073245; Natural Science Foundation of Shanghai, Grant/Award Number: 20ZR1440500; Pudong New Area Science and Technology Development Fund, Grant/Award Number: PKX2021-R07; Shanghai Normal University, Grant/Award Number: SK202227

Abstract

Traditional warehousing typically needs machine learning or manual tagging to classify objects. However, this method is less robust and consumes a lot of labour and material resources. Based on DenseNet, this work proposes a feature weighting convolutional network recognition model and designs a set of software and hardware for data acquisition, which is applied to the efficient classification of industrial parts in warehouse management. Firstly, this work modifies DenseNet by embedding SE-Block, and replaces the cross-entropy loss function with the focus loss function to optimize the model structure. Secondly, a multi-view hardware and software acquisition system is designed to complete the functions of part image acquisition, image preprocessing, model training and part recognition. Finally, an industrial parts sorting experiment was designed. Compared with the original DenseNet model, the proposed weighted convolutional network identification model showed that the accuracy of the modified model was increased by 3.09% and the convergence rate was significantly improved. The modified model proposed in this work aims to improve the recognition accuracy of industrial parts in modern warehouse management, so as to modify the classification efficiency of warehouse parts in production.

1 | INTRODUCTION

Warehouse management is the core of the logistics management system. With the expansion of production scale, a large number of industrial parts enter the warehouse, and how to efficiently identify the types of parts brings a greater challenge to the warehouse system. The majority of part management in a traditional warehouse involves manually labelling and placing each item in its proper location, when the parts are broken, it is necessary to find the corresponding parts to replace through manpower. But the increase in the number of parts makes it more difficult to manage, and does not avoid the problem of low efficiency. The parts must therefore be moved about in the warehouse in order to facilitate more organized and effective management. With the advent and development of deep learning [1, 2] and machine vision [3, 4], there are more technical ways to quickly identify the type of parts and thus facilitate the quick finding of parts, saving labour and time costs. The same approach has been applied to intelligent assembly part recognition [5, 6] etc. Thus deep learning has a wide range of application scenarios in the field of industrial part recognition and it is a more common solution in this field today.

There are currently many cases of machine learning in medicine such as Fang et al. combined DCNN with multi-modality images for AD classification [7]; Juan et al. proposed a vestibule segmentation network for CT images under the basic encoder-decoder framework [8]. Machine learning is also applied in agriculture [9, 10]. The method first requires data analysis of the features that play a critical factor in the prediction results, followed immediately by dimensionality reduction of the data, which aims to remove some features that have a low impact on classification while enhancing those that have a high impact on classification. Following the aforementioned activities, the anticipated outcomes are produced from the previously mentioned low-dimensional data, and the prediction algorithm can be adjusted in accordance with the scenario. As mentioned earlier, the machine learning approach relies on a large number of feature engineering, but in an industrial production environment with a large variety of parts, the features that determine the class of parts need to be reanalysed when new parts appear, so using the features observed by machine vision combined with machine learning algorithms is not robust.

In recent years, as deep learning has been further developed, it is applied in industry field. Deep learning is used to solve problems encountered in the production process, such as the application of machine learning to channel modelling, through which the transmission efficiency of wireless communication systems is significantly improved [11, 12]. Meanwhile, more and more scholars have combined machine vision and deep learning and used them in the field of intelligent manufacturing, such as waste classification [13] and defect localization in parts [14]. By combining machine vision with deep learning, industrial part recognition has a wide range of application scenarios.

In the context of the current development of smart manufacturing and the integration of Internet and traditional industries, the concept of smart warehousing was proposed. Smart warehousing is suggested as an effective solution to the rapid classification of industrial parts to achieve storage and access, which is a problem with the traditional warehouse management system due to the increase in the number of parts and the frequent flow of parts that consumes a lot of human and financial resources. At present, there are two primary issues that need to be resolved in the area of industrial part classification: first, how to further modify the recognition model based on the existing network to improve its classification accuracy for the characteristics of industrial parts; second, how to design the appropriate hardware and software collection framework for data collection work given that there isn't a lot of industry-specific public dataset and that training the model needs a lot of data.

Inspired by citations [5] and [21], this work proposes a DenseNet-based feature-weighted convolutional network recognition model for efficient classification of industrial parts. The highlights of this work are:

1. Embedding SE-block in DenseNet to extract feature weights to enhance the object features of useful information while suppressing the features of useless information, thus effectively improving the accuracy of the model; further optimizing the proposed model structure by using the focal loss function for the problem of classification difficulties caused by similar shapes of some parts.
2. To address the problem of lack of open datasets, design the hardware and software system for automatic industrial part acquisition in practical industrial part classification applications, with functions of automatic part data acquisition, preprocessing of acquired data, real-time part identification and model training.
3. Since the collected images are difficult to be directly used in the model training, the traditional image preprocessing method is not targeted and efficient in the specified scene in this work, so the Canny edge detection algorithm is modified and a novel preprocessing process is proposed.

In Section 2, related work on the direction of industrial part classification and edge detection are introduced. In Section 3, the proposed classification model from two aspects: model framework and loss function are described in detail. In Section 4, a complete data acquisition scheme and data preprocessing are presented. Experimental design and results are discussed in Section 5. The conclusion is drawn in the last section.

2 | RELATED WORK

In this section, we summarize prior work on industrial part classification, and then, we introduce related work on edge detection.

2.1 | Industrial part classification

AlexNet [2] led the research craze in this field after winning the ILSVRC championship in 2012. Starting from LeNet [15], the model structure of neural network superposition was proposed for the first time, but the model did not achieve good results due to the limitation of the computing power of computers at that time. After AlexNet, ZFNet [16] explained the good performance of CNN from the perspective of visualization and adjusted the size of convolutional kernels to make the depth features more robust. VGG [17] proposed a regular network structure that standardized the design thoughts of subsequent neural networks and became the backbone of many networks. GoogLeNet [18] used convolutional kernels of different sizes to design convolutional models of depth and width. ResNet [19] and DenseNet [20] both use the thought of hop count connection structure. DenseNet uses a multi-hop-connected structure based on ResNet. These two networks validate the feasibility of deeper networks. This work is a modification of DenseNet to make it more applicable to similar parts classification problems. In recent years DenseNet is also widely used in various fields, Pi et al. proposed propose a method called DNPPro based on densely connected convolutional neural networks to predict the promoter of Nannochloropsis [21]. Jiao et al. present a novel DenseNet framework with attention mechanisms (AM-DenseNet) to extract lung feature of 1 COVID-19 patient [22]. With the improvement of model accuracy, more and more scholars are applying deep learning to other fields such as the power industry [23, 24].

With the development of artificial intelligence, industrial part recognition methods have started to shift gradually from traditional manual detection to deep learning detection, and with the continuous improvement of algorithms a large number of solutions have been implemented in the production environment with good results. Yang et al. proposed a precise precision part classification method based on non-great suppression, which outperforms the YOLO V3 algorithm in terms of recognition accuracy and robustness [25]. Apostolopoulos et al. proposed a multi-path VGG19, which allows additional local and global feature extraction (multi-level feature extraction) by using multiple processing paths, with an average classification improvement of 6.95% over VGG19 [26]. Cheng et al. combined integration learning with random cropping and developed a random cropping integrated neural network (RCE-NN) to overcome the complex background environment problem [27]. Li et al. proposed an algorithm capable of accurate classification of parts, that is, a convolutional neural network model based on the InceptionNet-V3 pre-trained model through migration learning and reached an accuracy of 99.74% in the test set, and the algorithm can be applied in the intelligent diagnosis and maintenance of the parts with good prospects [28].

In industrial part classification, Hong et al. introduced a method to relabel the detection result of vehicle parts [29]. Prasad et al. proposed a novel deep learning approach to classify the various parts of any operational engine; it can qualitatively classify and henceforth give the corresponding class label of the machinery/engine part under consideration [30]. Abraham et al. developed a two-level machine learning-based system to classify different car parts [31]. As the demand for 3D vision technology in industry continues to rise, 3D point cloud part classification based on deep learning has received widespread attention. Hao et al. introduced some of the mainstream part inspection methods based on 3D point cloud data [32].

The industrial parts collected in closed space in this work have the following differences from traditional data: First, there are many kinds of parts involved in industrial production, and it is difficult for traditional computer vision to solve these problems on such a large scale; second, in industrial production, the objects identified are mostly related in the same production area, such as mechanical parts and production materials. The differences between some parts are very small, which increases the difficulty of recognition and classification to a certain extent; third, the data are collected in a closed environment, which is to reduce the positioning error caused by noise on the subsequent cropped images, so the classification results will rely on the fixed background in the closed environment and reduce the robustness of the model.

Based on the above characteristics, a modified DenseNet model is proposed in this work. Firstly, the DenseNet model can share information globally, and at the same time, there is the problem that different feature information within the model has the same influence factor on the classification results. Secondly, the dataset collected in this work is different from the traditional dataset, because the environment simulated in this work is an industrial closed environment, the background of the images we collected is relatively stable and the differences between different part categories are small, and there is similar feature information interference leading to higher classification difficulty. Finally, by analysing the above features, this work highlights the useful information in classification by embedding the SE-Block module in SE-ResNet. In addition to this, the loss function is replaced with the focal loss function, which makes the model more focused on the difficult classification problem and effectively improves the convergence speed of the model.

2.2 | Edge detection

Edge detection and localization are key steps in the image processing process and have a wide range of applications in fields such as medical image processing and mechanical part recognition. Gong et al. proposed adaptive image segmentation algorithm under the constraint of edge posterior probability. Based on traditional filtering algorithms [33], Shirazi proposed a method to extend a one-dimensional step edge detection filter to two dimensions by complex-valued filtering [34]. The learning-based algorithm detects edges by using a supervised model and hand-crafted features. Some models [35, 36] relied heavily on manual labelling, although at the time these algorithms showed advanced performance on the BSDS500 dataset. In recent years edge detection techniques are also evolving towards deep learning, Lu proposed the vector co-occurrence morphological edge detection operator, which takes the pixel and boundary information both into consideration [37]. Wang proposed an edge detection method using local edge pattern descriptors with multi-scale and multi-resolution properties [38]. He et al. proposed a bidirectional cascade network (BDCN) structure to detect edges at different scales [39]. Recent edge detection algorithms focus on the accurate detection of object boundaries, and in this work, we choose to improve on the Canny algorithm. Fuentes-Alventosa et al. proposed an unsupervised distributed Canny edge detector based on GPU to meet the edge detection real-time requirements [40]. Yang proposed a method for improving Canny edge detection by geo-information mapping [41]. Liu et al. studied an adaptive multi-scale edge detection method based on the Canny algorithm for human dynamic detection in virtual reality scenes, which suppresses impulse noise interference and reduces the possibility of false edges [42].

Inspired by citations [40] and [41], this work makes improvements to the Canny edge detection algorithm. For the contour characteristics of industrial parts, the curvature information of the parts is increased by adding two tilt direction templates to make the edge detection more accurate.

3 | A FEATURE-WEIGHTED CONVOLUTIONAL NETWORK RECOGNITION MODEL BASED ON MODIFIED DENSENET NETWORK

3.1 | Part identification network improvement analysis

In order to adapt to the problem of difficult classification of similar parts encountered in industrial part classification, this work modifies DenseNet by introducing a feature weighting module after each Bottleneck to weight the features that are important for category identification and suppress the features that have a weaker impact on the classification results, so that the classifier can focus more on the difficult categories. Additionally, the focus loss function, which is employed to lessen the weight of simple-to-classify data, has been added in place of the cross-entropy loss function in the loss function section.

3.1.1 | Introduce SE-Block for feature weighting

After introducing SE-Block into ResNet, the model can obtain information about the global sensory field through global pooling. The idea is to utilize a completely connected layer and ReLU function to fuse and nonlinearly delineate the information between each image feature channel, and to rescale the features for each channel's dependency using a fully connected layer and Sigmoid function. This work draws on this method and introduces SE-Block in DenseNet to alter the weights of several significant feature channels.

DenseNet reuses the information within the model making it possible to share the information globally within the model, but this also causes a problem that different feature information within the model plays the same weight on the results of the classification. In this part acquisition system, the acquisition environment simulates a closed environment in industrial production, so the background colours of the captured images are all the same. Since the majority of the components are silver, the classification outcome is not influenced by the colour feature information or other similar feature information. The SE-Block module can weigh the extracted channel features, which can enhance the useful information features and suppress some useless information features. In light of considering the above reasons, this work embeds SE-Block into the DenseNet network. Therefore, considering the above reasons, this work embeds SE-Block into the DenseNet network and proves the rationality of this approach in the experiment.

3.1.2 | Replace the loss function

Due to its ability to be applied during gradient descent computations to prevent the learning rate degradation issue brought on by gradient dispersion, cross-entropy loss function was employed for training models for classification problems. However, the traditional cross-entropy loss function cannot be adapted to all cases.

Equation (1) is the formula for calculating the cross-entropy loss in the multi-category task, where θ is the input parameter of the model; n is the number of categories:

$$CE\ (\theta) = -\sum_{i=1}^{n} y_i \log_{(\overline{y_i})} \tag{1}$$

In this work, the focal loss Equation (2) is used to replace the original loss function. The parameters α_i and γ are parameters that need to be adjusted manually. The parameter α is mainly used to solve the problem of unbalanced category data, and in this work the parameter takes a value of 1 because the amount of data in each category is approximately the same. Since some industrial parts are similar in shape, when the classifier classifies the parts, the probability of the parts being recognized into several similar categories is basically the same. $(1-p_i)^{\gamma}$ is used to focus on hard-to-classify samples. It can reduce the loss contribution of easy-to-classify samples, thereby increasing the proportion of losses of hard-to-classify samples. When p_i tends to 1, it means that the sample is easily to be classified, and $(1-p_i)^{\gamma}$ tends to 0, indicating that the contribution to the loss is small, that is, the loss ratio of the easy-to-classify sample is reduced. Therefore, by adjusting the value of γ, the loss

TABLE 1 The details of improved DenseNet network.

Layer	Output size	DenseNet-121	Improved DenseNet
Convolution	112×112	7×7 conv, stride 2	
Pooling	56×56	3×3 average pool, stride 2	
Dense block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3\text{conv} \\ \text{SE}-\text{Block} \end{bmatrix} \times 6$
Transition layer	56×56	1×1 conv	
	28×28	2×2 average pool, stride 2	
Dense block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3\text{conv} \\ \text{SE}-\text{Block} \end{bmatrix} \times 12$
Transition layer	28×28	1×1 conv	
	14×14	2×2 average pool, stride 2	
Dense block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3\text{conv} \\ \text{SE}-\text{Block} \end{bmatrix} \times 24$
Transition layer	14×14	1×1 conv	
	7×7	2×2 average pool, stride 2	
Dense block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3\text{conv} \\ \text{SE}-\text{Block} \end{bmatrix} \times 16$
Classification layer	1×1	7×7 global average pool	
		100D fully connected, softmax	

of easy categories can be reduced while paying more attention to those difficult categories, so the focal loss function is used instead of the cross-entropy loss function in this work. Through experimental analysis, the accuracy of the model is highest when $\gamma = 2$. Based on the above reasons this work replaces the loss function with focal loss function and proves the performance in the experiments.

$$FL\ (p_i) = -\sum_{i=1}^{n} y_i \alpha (1-p_i)^{\gamma} \log_{(p_i)} \tag{2}$$

3.2 | Feature-weighted convolutional recognition network model

Table 1pt presents the specifics of the components' classification model based on the modified DenseNet-121 network, and Figure 1 depicts the model's general structure. The traditional DenseNet model is in column three of Table 1, while the modified DenseNet model is in column four. It is clear that the SE-Block module of the SE-ResNet model has been added in the Dense-Block part of this model. A focal loss function is simultaneously used in place of the cross-entropy loss function.

Since DenseNet generates a large number of channels, and the features of different channels are different in importance

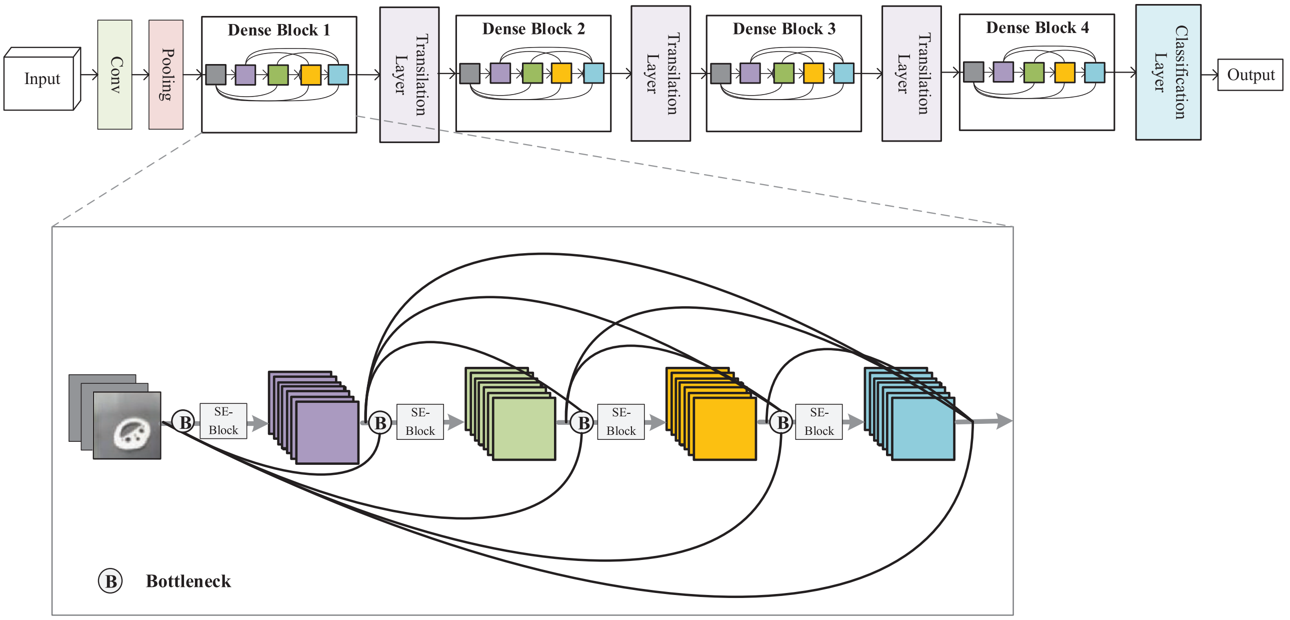

FIGURE 1 Overall structure of the model.

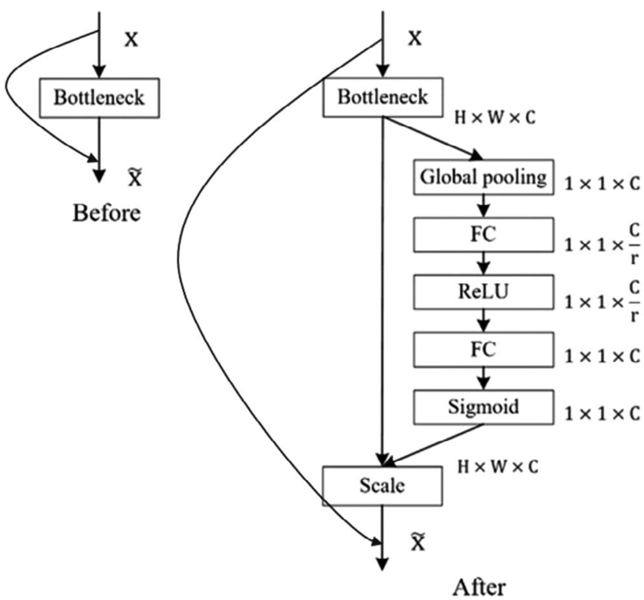

FIGURE 2 The schema of original Bottleneck (left) and the SE-DenseNet module (right).

to the classification task. The network structure of this work is different from the original DenseNet-121. We combine the channel attention mechanism with DenseNet, that is, the channel with more features is weighted in the channel dimension, and the channel with a more favourable classification is given higher weight. It can be seen from Table 1 that the whole network contains four dense blocks, and each dense block consists of many Bottleneck layers. Its structure is BN-ReLU-Conv (1 × 1) -BN-ReLU-Conv (3 × 3). In this work, SE-Block is added after each Bottleneck layer. Figure 2 shows the process of migrating SE-Block into the network.

4 | DESIGN OF SOFTWARE AND HARDWARE SYSTEM FOR AUTOMATIC INDUSTRIAL PARTS COLLECTION

In this work, we use supporting hardware and software for image acquisition. The acquisition framework mainly provides a confined space for part acquisition, and the software designed to improve the acquisition efficiency includes functions such as image acquisition and image processing, and is also used in subsequent model training. In order to acquire a large number of high-quality industrial parts datasets, this work uses traditional image processing methods such as Gaussian blur and edge detection to locate and segment the captured raw images. In this process, in order to improve the quality and acquisition efficiency of the acquired images, the following two aspects are modified: the edge detection algorithm is optimized to enable it to obtain more part contour information; a complete acquisition process is designed to automatically locate and segment the parts in the images. The aforementioned effort aims to lessen the impact of backdrop in the photographs on categorization, remove redundant information, and crop the images into consistent size image data.

4.1 | Design of hardware

Figure 3 is the image acquisition frame, 1 is the aluminium frame; 2 is a black acrylic plate, mainly used to reduce the reflection of light, so that the camera can capture the specific features and contours of the parts; 3 is the acquisition camera. The whole acquisition frame mainly has five angles of camera, which are top, front, back, left and right. It can shoot the parts from multiple angles, so as to speed up the efficiency of collecting datasets. The soft light cover is not shown in the picture. It is mainly

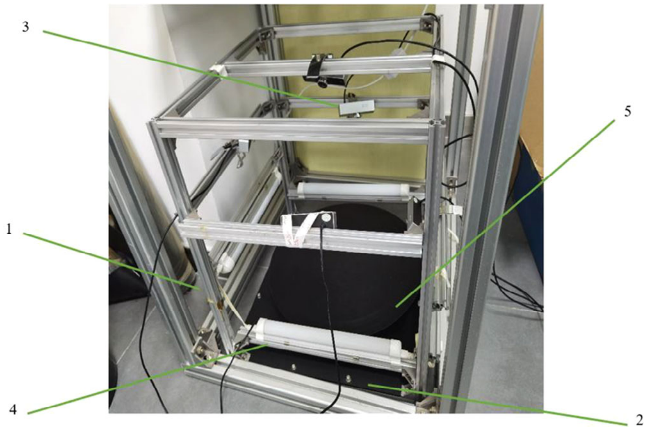

FIGURE 3 Image acquisition frame: (1) aluminium frame, (2) acrylic plate, (3) camera, (4) LED, (5) bottom turntable.

used to place the parts in a closed environment to avoid the interference of external strong light on the picture shooting, so that the picture can only be taken to the approximate shape. 4 is an adjustable brightness LED lamp. Because the external light cannot be irradiated in a closed environment, it is necessary to provide a light source. Because the light of the LED is bright, it will cause a loss in the quality of the image. Therefore, a layer of the lamp shade is set on the LED lamp to form a diffuse reflection and further improve the quality of the image. 5 is the bottom turntable. The parts are placed on the turntable, and the angle of the parts can be automatically adjusted after the motor is turned on. Only the top view is typically utilized in the classic meaning of machine vision. In this work, parts are photographed from a variety of angles to help extract part information more effectively.

4.2 | Design of software

As the efficiency of collecting data manually is low, this work designs an image acquisition software. The software runs on a Windows system and is written by PyQT. Its architecture is shown in Figure 4. There are four functional modules included: image acquisition, image processing, management model and image classification. Firstly, the image acquisition module can control the camera of the acquisition frame to shoot the parts on the turntable, and the taken photos can be placed in the specified folder. Secondly, the image-preprocessing module can crop and augment the images under the specified folder. Then, the management model can train the model and save the location of the model. Finally, the image classification module can classify images in real time. Figure 5 is an example of some industrial parts collected by this acquisition system.

4.3 | Dataset preprocessing

Due to the low resolution of the image data collected by the acquisition system, and the proportion of some parts in the edge of the images or parts in the images being too small, the collected original images cannot be directly put into the model for training, and the position of the part needs to be located first to facilitate segmentation. The correct positioning of the parts depends on the binary image generated by Canny edge detection. Therefore, it's crucial to understand how to maximize the impact of Canny edge detection. However, the part photos gathered by the aforementioned acquisition program will contain noise information such as shadows and white spots in the black backdrop owing to wear, aging, illumination, and other factors. The edge extraction algorithm is sensitive to noise, and the noise will affect the positioning of the part in the background, so it is

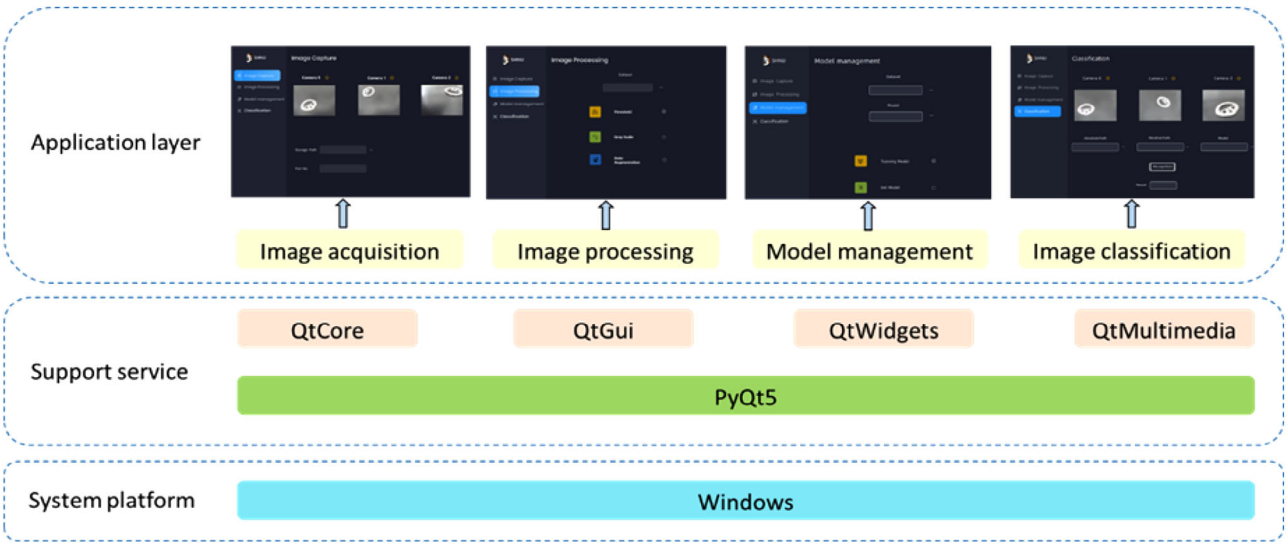

FIGURE 4 Software architecture.

Number	Cameral 0	Cameral 1	Cameral 2
0			
11			
45			
46			
58			

FIGURE 5 Sample parts.

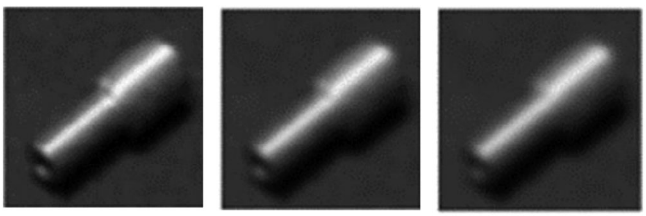

FIGURE 6 The processing of Gaussian filter with different radius.

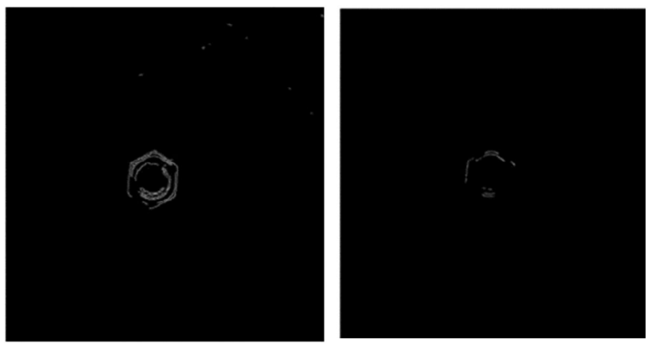

FIGURE 7 Canny edge detection diagram. (a) Gaussian blur radius is 3 × 3. (b) Gaussian blur radius is 7 × 7.

necessary to remove the noise first. This work uses Gaussian blur for noise reduction. When the radius of Gaussian blur is larger, the image is smoother and the denoising effect is better, but this will also cause the edge of the part to be excessively smooth and cannot be detected. Figure 6 shows parts with different filter sizes.

Figure 7 is the experimental picture of the parts under different Gaussian filter radii obtained by Canny operator edge detection. The parts taken are silver parts with greater luster. From Figure 7a, it can be seen that when the Gaussian blur radius is 3 × 3, more details of the part contour can be displayed, but there will be more noise, which will lead to the inability to correctly locate the position of the part. When the fuzzy radius is further increased, the contour of the part in Figure 7b becomes more and more blurred. When locating the position of the part, because the contour of the edge disappears, it is easy to cut out the information of the parts in the picture together in the process of cutting. Therefore, in order to ensure that the edge obtains more information, the Canny operator edge detection is modified. At the same time, in order to correctly locate the information, an algorithm is proposed to effectively locate the position of the part and successfully cut the part from the picture.

The traditional Canny operator has certain limitations. This is because the shape of the part is diverse, and most of the parts have a certain radian. The direction template of the traditional Canny operator edge detection is only vertical and horizontal, which will lose part of the information with radians. Therefore, in order to solve this problem, this work increases the radian information of the part by adding two inclined direction templates.

-1	0	1
-2	0	2
-1	0	1

(a)

-2	-1	0
-1	0	1
0	1	2

(b)

-1	-2	-1
0	0	0
-1	2	1

(c)

0	1	2
-1	0	1
-2	-1	0

(d)

FIGURE 8 Template direction.

As shown in Figure 8, 8a and 8c are Canny's templates, which can be used to calculate the gradient values of each pixel in the image in the horizontal and vertical directions. 8b and 8d are newly added templates, representing the directions of 45° and 135°, respectively.

In Figure 9a-c are respectively the traditional Canny operator, the Canny operator with the direction template added and the original figure. It can be seen from the figure that compared with the traditional Canny operator, the edge contour of the Canny operator after adding the direction template is clearer. It can be seen from the selected parts of the red box and the blue box, the contour in Figure 9b is clearer than that in Figure 9a. From

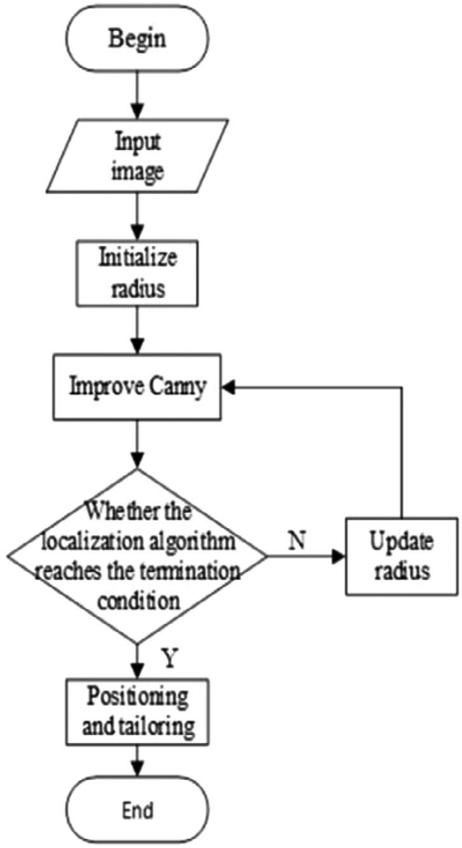

FIGURE 10 Flow chart of picture cropping.

a global point of view, the contour of Figure 9b is also clearer than that of Figure 9a.

Through the above experimental comparison, due to the different material sizes of the parts, it is impossible to select a fixed Gaussian blur radius to smooth the image. Therefore, this work proposes a method to locate the position of the part and cut it, as shown in Figure 10 as a specific flow chart.

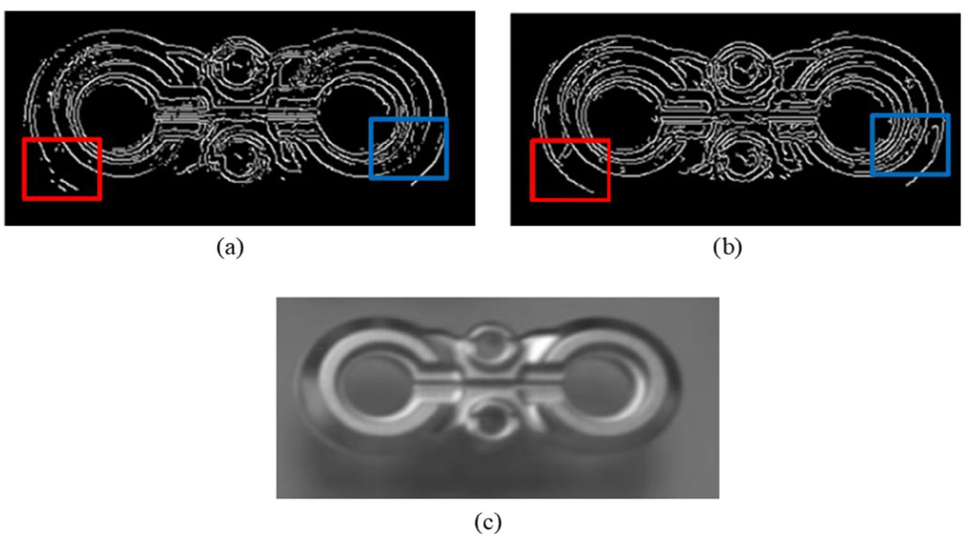

FIGURE 9 Comparison graph before and after the modification of Canny operator. (a) Traditional Canny. (b) Adding direction templates. (c) Original picture.

FIGURE 11 The appearance of parts.

In this flow, after Gaussian blur and Canny, the white part in gray image is the main body of the part, and the black part is the background. The value of the white pixel is set to 255, and the value of the black pixel is set to 0. Then the program traverses all rows and columns in the image, calculates the number of points with a value of 255, saves this value as w_{pre}, and saves the Gaussian radius r_{pre}(default 9) used at this time. Judging whether r_{pre} is greater than 0, if it is greater than 0, r_{pre}-2 is assigned to r_{next}, r_{next} is used as a new Gaussian radius and Canny is used again. The image is traversed, and the number of points with a value of 255 is assigned to w_{next}. At this point, w_{next} is compared with the previous w_{pre}. If the ratio of w_{next} to w_{pre} is 1.1 and above, it shows that the value of wpre is too large and the information of the part losses. So, w_{next} needs to be assigned to wpre, and then re-execute this step. If the ratio of w_{next} to wpre is less than 1.1, it shows that the Gaussian filter will not have a great impact on the information of the part, and the image can be cut next.

Following the above steps, the program traverses the images, sets the horizontal coordinates of the white pixel points in the leftmost column to x_1, the leftmost column to x_2, the vertical coordinates of the top row to y_1, and the bottom row to y_2. To get the complete part and the subsequent expansion of the dataset, the x-axis and the y-axis are expanded by 30pixe points, that is each image will be segmented into $|x_1 - x_2|+30*|y_1 - y_2|+30$. If this size exceeds the image boundary, the image boundary is used as the crop boundary, and then scaling images to a fixed size.

In this work, 100 kinds of parts are selected as the original collected images, and the software designed in this work is used to collect images. There are a total of three angles of the camera. Because different parts have different shapes, for example, some parts have similar sides, and some of the shapes seen from the side are not similar, different parts will have different numbers of datasets. Most of the parts were picked up by three cameras and numbered between 50 and 60. Through a series of data augmentation, the number has been increased to between 500 and 600. The term data augmentation mainly includes flipping, rotating, deforming, cutting, increasing noise, etc., which effectively suppress the generation of overfitting problems. Figure 11 shows some pictures of the parts after data augmentation.

5 | EXPERIMENT

5.1 | Experimental environment

The PyTorch framework serves as the foundation for the model framework, picture enhancement, and image reading in this study. Numpy is used to handle image data, and Tensorboard is used to visualize experimental results. The experiment makes

TABLE 2 Experimental environment.

Index	Parameter
CPU	Intel Core i9
GPU	RTX3090
RAM	64GB
Frame	PyTorch
Language	Python3.7

TABLE 3 Comparison of three models to evaluate.

	Original DenseNet model	Adding SE-Block	Replacing loss function
Model A	✓	×	×
Model B	✓	✓	×
Model C	✓	✓	✓

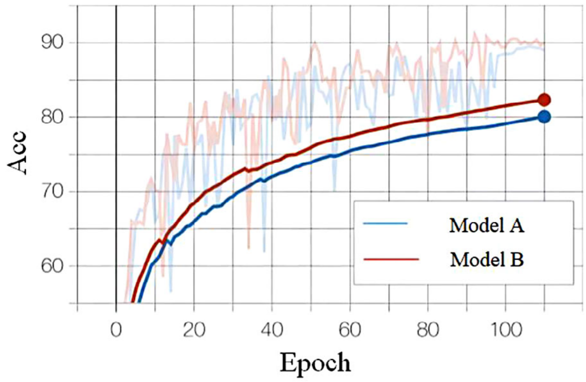

FIGURE 12 Comparison of the accuracy between SE-Densenet and the orginal one.

use of 10-fold cross-validation, in which 600 photos are randomly divided into 10 portions with a 9:1 train-to-test ratio, and the findings of several groups are averaged to lessen variation. Table 2 depicts the experimental setting.

5.2 | Experimental result

In this work, the model is trained under the PyTorch open-source framework. According to the modified DenseNet model above, the model is used to classify 100 classified industrial parts. In the training, the network model adopts the Adam algorithm and its learning rate is selected as 0.01.

The experiments are conducted on three models. Model A is the original DenseNet network. Model B is the SE-DenseNet that only added SE-Block. Model C is the proposed modified DenseNet in which SE-Block is added and the loss function is replaced. Table 3 shows the comparison of the three models to evaluate.

Figure 12 reveals the accuracy comparison diagram of Model A (original DenseNet network) and Model B (SE-DenseNet).

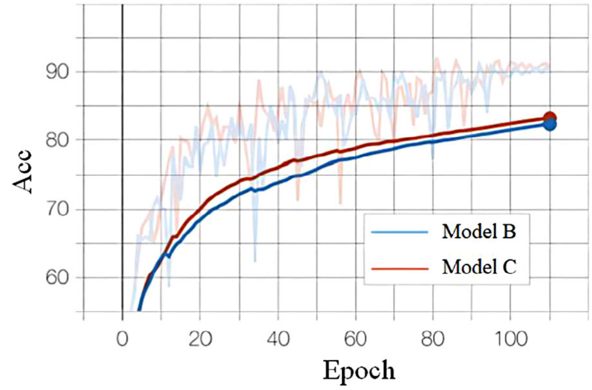

FIGURE 13 Ablation experiment on precision of loss function.

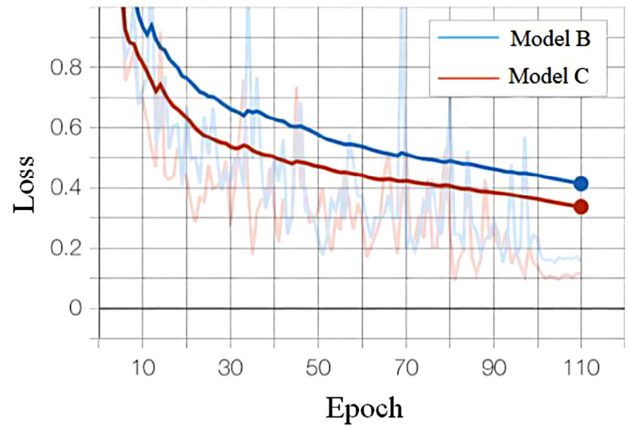

FIGURE 14 Ablation experiment on the loss of loss function.

The shadow part is the actual accuracy curve of the two models, and the curve is the line segment smoothed by Smooth = 1. It can be seen from the figure that the two models oscillate at the beginning of training and the amplitude is large. After many iterations, the accuracy gradually approaches to be gentle. Model A achieves 88.47 % accuracy after 110 rounds of training, while the accuracy of Model B reaches 91.09 %, with a 2.62 % accuracy improvement. In addition, it can be seen from the diagram that Model B rises faster than the original network Model A.

Figure 13 represents the result of ablation experiment on precision of loss function. The accuracy of Model B reaches 91.09 % after 110 rounds of training, while the accuracy of Model C reaches 91.56 %, with an improvement of 0.47 %. We can see the curve after Smooth = 1 smoothing. Model C rises faster than Model B, and the accuracy of the shadow part is also higher than Model B. This is mainly because some industrial parts have similar parts in shapes, some categories of parts are difficult to classify. By replacing the loss function, the model can pay more attention to categories with similar part shapes.

Figure 14 shows the result of the loss comparison between Model B and Model C. It can be seen from the figure that the downward trend of Model C is significantly faster than that of

TABLE 4 Comparison of loss in ablation experiments.

	Model B	Model C
Loss	0.1014	**0.152**

TABLE 5 Test accuracy of three models.

	Model A	Model B	Model C
Accuracy	88.47	91.09	**91.56**

TABLE 6 The accuracy comparison between the model in this work and other models.

Model	Testing accuracy	Training accuracy	Epoch
AlexNet	80.42	84.21	200
ResNet	86.23	91.45	120
VGGNet-16	82.76	85.4	200
Modified DenseNet	91.56	96.8	120

Model B, so the model can reach the optimal point faster. And the curve after Smooth = 1 decreases significantly faster than the former. As shown in Table 4, the loss of Model B and Model C is 0.1014 and 0.152 respectively. It can be found that the value of the loss is reduced, indicating that the model convergence effect is better (The figure only shows part of the change process, because cross-validation is used in this work, only when the result of the test set is poor, it can be explained that the model is overfitting. It can be seen that the model with the replaced loss function has a better result).

Table 5 gives the test accuracy of three models. The accuracy of the three networks on the test set are 88.47, 91.09, and 91.56 respectively. Experiments show that Model C, the modified DenseNet network, has better accuracy.

In addition, this work also compares with other models. As shown in Table 6, the test accuracy, training accuracy and training times of the model are respectively. In this training, the model uses the same parameters. It can be seen from the experimental results that the model in this work can achieve higher accuracy with fewer epochs. AlexNet and VGGNet-16 have no residual structure, so the relative accuracy is low, and the speed of convergence to the best accuracy is slow. The model in this work improves the model for some features in the part, and has a greater accuracy improvement than other models. Within the allowable range, our model can achieve better results with fewer training epochs.

6 | CONCLUSION AND FORESIGHT

In this work, a classification model based on a modified DenseNet network is proposed. In order to optimize the feature information weight of Dense Block in DenseNet, this work uses SE-Block to re-weight the weight of feature information. In this way, the weight of feature information that plays a decisive role in classification can be effectively enhanced, which can effectively improve the accuracy of the model. Considering that there is a certain similarity between the targets, the focus loss function is used instead of the cross-entropy loss function, so that the model can focus on the difficult classification, thus further improving the accuracy of the model. This modified network is applied to solve the large-scale classification problem of industrial parts. This work develops a collection framework that offers a closed environment for portion collection and integrates the UI interface based on the collection framework to address the issue of low efficiency and quality of data collection. The software is written by PyQT, and its functions include image acquisition, image processing, management model and image classification. Aiming at the problem that the collected image pixels are too large to be directly put into the model training and the proportion of the part body in the image is too small, the Gaussian smoothing denoising is used, and then the Canny operator edge detection is used to determine the position of the part in the image. Finally, the image is cut according to the position. The purpose of Gaussian smoothing is to prevent noise interference positioning. Aiming at the radius selection problem of Gaussian filtering, this work proposes an algorithm that can automatically select an effective radius value. For Canny operator edge detection, this work improves it by adding 45 ° and 135 ° direction templates. After locating the part, it is cut and scaled to a uniform size, and then put into the model training. Finally, the modified network was subjected to ablation experiments and comparison experiments with other networks on the dataset collected in this work. The experimental results show that: The use of a feature weighting module and focus function improves the classification accuracy to a certain extent. The network proposed in this work has the best performance in classification results compared with other classification networks.

The research direction of this work can be further modified. In this work, the parts with obviously different shapes or large appearance gaps are used, that is, the classification problem of the general direction is dealt with. For some parts with similar shapes and only different sizes or details, how to collect data and how to effectively identify them is a direction worth studying in the future industrial field. In this work, we did not compare our network with the latest ones, only compared with several classic networks. In the follow-up study, we will further improve our experiment by comparing the modified model with recent architectures. We will also explore more innovative approaches that are not limited to the work on attention modules or loss functions.

AUTHOR CONTRIBUTIONS

An Kang: Funding acquisition; Methodology; Supervision; Writing—original draft; Writing—review and editing. **Sun Xiaoqing**: Investigation; Visualization; Writing—original draft; Writing—review and editing. **Song Yaqing**: Conceptualization; Methodology; Supervision; Writing—original draft; Writing—review and editing. **Lu Yebin**: Resources; Software; Validation; Writing—original draft. **Shangguan; Qianqian**: Methodology; Project administration; Supervision; Writing—review and editing.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant No. 62073245), the Natural Science Foundation of Shanghai (20ZR1440500), and Pudong New Area Science and Technology Development Fund (PKX2021-R07) and Shanghai Normal University (SK202227).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Yaqing Song https://orcid.org/0000-0003-0914-0395

REFERENCES

1. Fitch, F.B., McCulloch Warren, S., Walter, P.: A logical calculus of the ideas immanent in nervous activity. J. Symb. Log. 9(2), 49–50 (2014)
2. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. Commun. ACM 60(6), 84–90 (2017)
3. Lowe, D.,G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)
4. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification withdeep convolutional neural networks. Adv. Neural Inf. Processing Syst. 60(6), 84–90 (2017)
5. Doltsinis, S., Krestenitis, M.; Doulgeri, Z.: A machine learning framework for real-time identification of successful snap-fit assemblies. IEEE Trans. Autom. Sci. Eng. 1(17), 513–523 (2020)
6. Xue, G., Gan, G., Peng, Z., et al.: Research on the technology of automatic sorting and distribution of parts needed for assembly of aircraft parts. In: IEEE 10th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT). Cape Town, South Africa, pp. 61–65 (2019)
7. Fang, X., Liu, Z., Xu, M.: Ensemble of deep convolutional neural networks based multi-modality images for Alzheimer's disease diagnosis. IET Image Proc. 14, 318–326 (2020)
8. Juan, C., Li, Z., Yao, Z.: Deeply supervised vestibule segmentation network for CT images with global context-aware pyramid feature extraction. IET Image Proc. 17, 1267–1279 (2022)
9. Xin, L., Yu, G., Rui, G., et al.: An automatic plant leaf stoma detection method based on YOLOv5. . IET Image Proc. 17, 67–76 (2022)
10. Junos, M.H., Khairuddin, A.S.M., Thannirmalai, S., et al.: An optimized YOLO-based object detection model for crop harvesting system. IET Image Proc. 15(9), 2112–2125 (2021)
11. Alghorani, Y., Chekkouri, A.S., Chekired, D.A., et al.: Improved S-AF and S-DF relaying schemes using machine learning based power allocation over cascaded rayleigh Fading channels. TechRxiv 22(12), 7508–7520 (2021)
12. Li, D., Xu, Y., Zhao, M., et al.: Knowledge-driven machine learning and applications in wireless communications. IEEE Trans. Cognit. Commun. Networking 8(2), 454–467 (2022)
13. Koskinopoulou, M., Raptopoulos, F., Papadopoulos, G., et al.: Robotic waste sorting technology: toward a vision-based categorization system for the industrial robotic separation of recyclable waste. IEEE Rob.. Autom. Mag. 28(2), 50–60 (2021)
14. Lin, X., Wang, X., Li, L.: Intelligent detection of edge inconsistency for mechanical workpiece by machine vision with deep learning and variable geometry model. Appl. Intell. 50, 2105–2119 (2020)
15. Yann, L.: Backpropagation applied to handwritten zip code recognition. Neural Comput. 1(4), 541–555 (1989)
16. Zeiler, M.,D., Fergus, R.: Visualizing and understanding convolutional networks. Computer Vision—ECCV 2014, Springer, Berlin, pp. 818–833 (2014)
17. Russakovsky, O., Deng, J., Su, H., et al.: ImageNet large scale visual recognition challenge. Int. J. Comput. Vision 115(3), 211–252 (2015)
18. Szegedy, C., Wei, L., Jia, Y., et al.: Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA, pp. 1–9 (2015)
19. He, K., Zhang, X., Ren, X., et al.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA, pp. 770–777 (2016)
20. Huang, G., Liu, Z., Van Der Maaten, L.: Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, pp. 4700–4708 (2017)
21. Pi, J., Zhen, Z., Lin, J., et al.: Promoter prediction in nannochloropsis based on densely connected convolutional neural networks. Methods 204, 38–46 (2022)
22. Jiao, S., An, Z., Qing, Y., et al.: AM-DenseNet: A novel DenseNet framework using attention mechanisms for COVID-19 CT image classification. International Conference on Cloud Computing and Intelligent Systems (CCIS). Chengdu, China, pp. 474–479 (2022)
23. Shan, Y., Mi, W., Hua, G., et al.: Short-term power load forecasting based on VMD-DenseNet. Comput. Appl. Software 40, 34–70 (2023)
24. Bouchiba, N., Kaddouri, A.: Performance comparison of fuzzy logic and deep learning algorithms for fault detection in electrical power transmission system. In: 2021 IEEE Electrical Power and Energy Conference (EPEC). Toronto, Canada, pp. 101–106 (2021)
25. Yang, J., Li, S., Gao, Z., et al.: Real-time recognition method for 0.8 cm darning needles and KR22 bearings based on convolution neural networks and data increase. Appl. Sci. 8(10), 1857 (2018)
26. Apostolopoulos, I.D., Tzani, M.A.: Industrial object and defect recognition utilizing multilevel feature extraction from industrial scenes with Deep Learning approach. J. Ambient Intell. Hum. Comput. 14, 10263–10276 (2023)
27. Cheng, B., Wu, W., Tao, D., et al.: Random cropping ensemble neural network for image classification in a robotic arm grasping system. IEEE Trans. Instrum. Meas. 69(9), 6795–6806 (2020)
28. Li, Q., Chen, G.: Recognition of industrial machine parts based on transfer learning with convolutional neural network. PLoS One 16(1), e0245735–e0245735 (2021)
29. Hong, S., Yu, C., Gil, Y., et al.: Relabeling method for improving vehicle part detection. In: 2022 13th International Conference on Information and Communication Technology Convergence (ICTC). Jeju Island, Korea, pp. 806–808 (2022)
30. Prasad, M., Rajora, S., Gupta, D., et al.: Fusion based En-FEC transfer learning approach for automobile parts recognition system. In: IEEE Symposium Series on Computational Intelligence (SSCI). Bangalore, India, pp. 2193–2199 (2018)
31. Abraham, C., Christian, M., John, P., et al.: Damage identification of selected car parts using image classification and deep learning. In: International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). Manila, Philippines, pp. 1–5 (2021)
32. Hao, Q., Zhen, L., Guang, G., et al.: Research progress of part point cloud detection based on deep learning, In: 2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT). Jilin, China, pp. 797–801 (2023)
33. Gong, M., Lan, J., Yang, C., et al.: Adaptive image segmentation algorithm under the constraint of edge posterior probability. IET Comput. 11, 702–709 (2017)
34. Shirazi, S., Safabakhsh, R.: Omnidirectional edge detection. Comput. Vision Image Understanding 113(4), 556–564 (2009)
35. Arbeláez, P., Maire, M., Fowlkes, C., et al.: 'Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell 33(5), 898–916 (2011)
36. Lim, J.J., Zitnick, C.L., Dollar, P.: Sketch tokens: A learned mid-level representation for contour and object detection. In: Computer Vision and Pattern Recognition (CVPR), Portland, USA, pp. 3158–3165 (2013)

37. Lu, Y., He, C., Yu, Y., et al.: Vector co-occurrence morphological edge detection for colour image. Int. J. Comput. Vision 15(13), 3063–3070 (2021)
38. Wang, Y., Zhang, N., Yan, H., et al.: Using local edge pattern descriptors for edge detection. Int. J. Pattern Recogn. Artif. Intell. 32(3), 1850006 (2017)
39. He, J., Zhang, S., Yang, M., et al.: DCN: Bi-directional cascade network for perceptual edge detection. IEEE Trans. Pattern Anal. Mach. Intell. 44(4), 2089–2107 (2020)
40. Fuentes-Alventosa, A., Gomez-Luna, J., Medina-Carnicer, R.: GUD-Canny: A real-time GPU-based unsupervised and distributed Canny edge detector. J. Real-Time Image Process. 19(3), 591–605 (2022)
41. Yang, L., Li, M., Wu, T., et al.: Geo-information mapping improves Canny edge detection method. IET Image Proc. 17, 1893–1904 (2023)
42. Liu, J., Chen, D., Wu, Y., et al.: Image edge recognition of virtual reality scene based on multi-operator dynamic weight detection. IEEE Access 8, 111289–111302 (2020)

How to cite this article: An, K., Sun, X., Song, Y., Lu, Y., Shangguan, Q.: A DenseNet-based feature weighting convolutional network recognition model and its application in industrial part classification. IET Image Process. 1–13 (2023). https://doi.org/10.1049/ipr2.12971