

Defect Detection of Polyethylene Gas Pipeline Based on Convolutional Neural Networks and Image Processing

Jun-qiang Wang¹, Sixi Zha^{2,3*}, Jia-chen Sun⁴, Yang Wang², Hui-qing Lan^{4,5*}

(1. China Special Equipment Inspection and Research Institute, Beijing 100029, China

2. School of Mechanical Engineering, Xinjiang University, Wulumuqi, 830000, China

3. Guangdong Provincial Key Laboratory of Technique and Equipment for Macromolecular Advanced Manufacturing, South China University of Technology, 510641, China

4. School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China

5. Tangshan Research Institute of Beijing Jiaotong University, Tangshan 063000, China)

Abstract: In this paper, a method based on image recognition was proposed to detect the defects of polyethylene (PE) gas pipeline, especially the deformation due to the indentation. Firstly, the pipeline -detection VGG (PD-VGG) model was established based on the convolutional neural network (CNN), and appropriate model parameters were optimized through model training. The defect recognition rate of the improved model can reach 94.76%. Following, the weighted average graying algorithm was used to separate the defects characterized by deformation. Then, an improved gamma correction algorithm was applied to achieve image enhancement, and the interference of impurities adhered on inter surface of pipeline was also removed by using multi-layer filters. The edge detection of the defect image was completed by using the Canny operator, and following the screening between the target contour and the interference contour by using top-contour. Finally, the algorithm for minimum outer rectangle algorithm was used to fit the defect contour, and the eigenvalues of deformation defects were extracted. The results indicate that the above defect detection method can better extract the deformation contour of the dented pipeline. The high agreement with the experimental results provides a basis for the research of effectively recognizing whether the pipeline has undergone ductile failure only through profile detection of defects.

Keywords: Image recognition; PE pipeline; Defect detection; Convolutional neural network

*Corresponding author.

Sixi Zha, E-mail address: sxzha@xju.edu.cn.

Hui-qing Lan, E-mail address: hqlan@bjtu.edu.cn.

32 **1 Introduction**

33 Nowadays, the advantages of polyethylene (PE) materials have led to their widespread use in areas such as
34 lifeline engineering [1-2]. PE pipes have become a suitable alternative to steel pipes due to their lightweight, low
35 cost, anti-corrosion and easy installation. Because sunlight causes rapid degradation of PE materials, PE pipeline
36 must be buried underground, which makes its inspection process more difficult. Traditional inspection methods
37 suitable for pipeline have complex equipment, high cost, and are not easy to visualize defects. Furthermore,
38 manual inspection methods are inefficient and labor intensive. Therefore, the use of various inspection robots to
39 detect defects in the pipeline has become a research hotspot. Among them, machine vision is a method that
40 utilizes a pipeline robot equipped with a light source and a camera to directly collect, transmit and process images
41 of the inside of the pipeline in real time. It then utilizes image recognition technology for non-destructive, non-
42 contact inspection, which improves the efficiency of pipeline inspection and reduces the influence of human
43 subjectivity. In addition, image processing technology, as a contactless, on-line inspection method, is widely used
44 in the research of pipeline defect detection, and a large number of related studies have been carried out by scholars
45 at home and abroad [3-6]. Before defects are recognized, image preprocessing techniques are usually used to
46 improve the quality of the image or to extract defective regions in preparation for subsequent defect recognition
47 and classification. Image processing techniques mainly involve image noise reduction [7-8], contrast enhance-
48 ment [9], image segmentation [10] and morphological processing [11]. Motamedi [12] et al. performed operations
49 such as gray scaling, filtering and morphology on pipe images to achieve nondestructive detection of defects in
50 urban drainage pipes. Khalifa [13] et al. first performed grayscale transformation on drainage pipe images, then
51 segmented the image by finding the optimal threshold, then enhanced the image by using the open operation, and
52 finally, Laplace edge detection was performed on it to extract the crack eigenvalues. Kirstein [14] et al. combined
53 Canny edge detection, Hough's linear transform and shortest path algorithm to also achieve defect detection of
54 drainage pipes. Alam [15] et al. first grayed out the pipe image, then used the Sobel gradient method to detect its
55 edges, and then removed unwanted tiny objects on the image, and finally performed feature extraction on the
56 processed image to identify pipeline defects. Huynh [16] et al. used the DEE algorithm to achieve automatic
57 detection of small cracks in pipelines. Mashford [17] et al. used the Haar wavelet transform to detect edges of
58 pipeline images. Wei [18] et al. used an industrial camera to capture images of the inside of natural gas pipelines,
59 and used the Laplacian edge detection algorithm to successfully extract the welding seam with defects in them.
60 Altabey [19] built a pipeline crack image segmentation model based on semantic segmentation, and extracted the
61 cracks in high-resolution crack images by using the pipeline image segmentation model. The accuracy rate, recall
62 rate, and F-score of the proposed method are recorded 89.3%, 85.7%, and 80.4%, respectively.

63 Along with image processing techniques, pattern recognition techniques are also important elements which
64 are mainly used to identify and classify defects. First the desired features are extracted in the segmented region

65 and then the classifier will use these features as inputs for further defect recognition, so this step of pattern recog-
66 nition is crucial. Su and Yang [20] proposed an MSED-based image processing method for drainage pipes, where
67 the area of the defective region, aspect ratio, and eccentricity are used as vectors of discriminative features, based
68 on which the shape of the defect is determined. Sinha [21] found that combining images processed based on a
69 morphological approach with a neural network-based classifier can lead to better detection than using any of the
70 methods alone. Zhou [22] presented a defect inspection algorithm of metal surface based on machine vision. The
71 proposed surface defect inspection algorithm first used improved bi-dimensional empirical mode decomposition
72 (BEMD)-based extracting algorithm to perform initial extracting of surface defects through filtering out complex
73 textures on the metal surface, while retaining as much effective information as defects as possible. Moselhi [23]
74 proposed a simplistic image processing means using blurring and edge detection for the acquired images of un-
75 derground sewage pipelines, extracting the basic features of the defects as well as the elongation rate as the
76 feature vectors, and also incorporating the commonly used neural network method to classify the defects.

77 One of the main problems in using image recognition technology is how to effectively extract and recognize
78 defects inside PE gas pipeline. Although most of the current pipeline visual inspection methods can obtain the
79 image of the internal surface in pipeline, the processed image of defect is blurred, and the defect extraction effect
80 will be seriously affected due to noise and uneven lighting phenomenon. For the detection of defects inside PE
81 pipeline, a method is needed to improve the quality and recognition efficiency of defect images.

82 This paper proposes an image-based method for detecting internal defects in small-diameter PE gas pipeline.
83 Firstly, the PD-VGG model was established based on the convolutional neural network, and appropriate model
84 parameters were optimized through model training. The defect recognition rate of the improved model can reach
85 94.76%. Following, the weighted average gray scale algorithm was used to separate the defects characterized by
86 deformation. Then, an improved gamma correction algorithm was applied to achieve image enhancement, and
87 the interference of impurities adhered on inter surface of pipeline was also removed by using multi-layer filters.
88 The edge detection of the defect image was completed by using the Canny operator, and following the screening
89 between the target contour and the interference contour by using top-contour. Finally, the algorithm for minimum
90 outer rectangle algorithm was used to fit the defect contour, and the eigenvalues of deformation defects were
91 extracted. The results indicate that the above defect detection method can better extract the deformation contour
92 of the dented pipeline. The high agreement with the experimental results provides a basis for the research of
93 effectively recognizing whether the pipeline has undergone ductile failure only through profile detection of de-
94 fects.

95 **2 Identification and classification of PE pipeline defects**

96 **2.1 Data set establishment and data enrichment**

97 Using neural network for PE gas pipeline defect image classification, the first thing needed is the dataset,
98 through the collation and generalization of the gas pipeline defect data set in this paper. This dataset contains

99 three classic types of defects in gas pipelines- deformation (caused by external loading), misalignment (caused
 100 by welding errors, as shown in figure 1), and breakage (rupture, crack). The total number of images is 458, and
 101 the specific data are shown in Table 1.

102 Table 1 Raw data for PE pipeline

PE gas pipe raw data				
Category	Normal	Deformation	Misalignment	Breakage
Quantity (sheets)	150	104	96	108

103 In this paper, the collected images were augmented to increase the number of datasets on the one hand,
 104 prevent overfitting on the other hand, and provide better generalization for relevant image recognition. The broad-
 105 ening methods used are angle rotation rotation, noise perturbation, random cropping, horizontal (vertical) flipping,
 106 etc., as shown in figure 1.

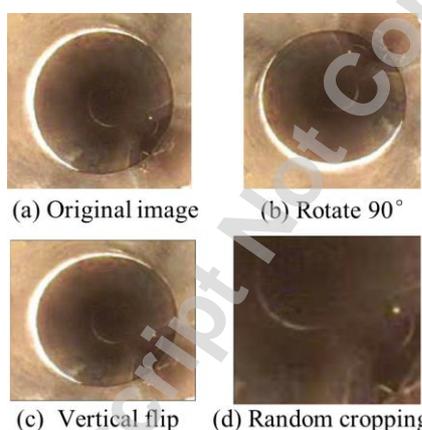


Fig. 1 Data enrichment

2.2 Pipeline defect identification modeling and research

110 Migration learning is a branch of machine learning. Migration learning utilizes the data, parameters and
 111 weights of the source domain to transfer the knowledge to the target domain and complete the new model. At
 112 present, there are not many studies on the defective image recognition of gas pipelines, and the amount of data
 113 owned is not large. This paper is about the image recognition of PE gas pipe defects with small samples, and the
 114 migration learning of VGG-16 will be used to build the PD-VGG model by freezing layer, as shown in figure 2.

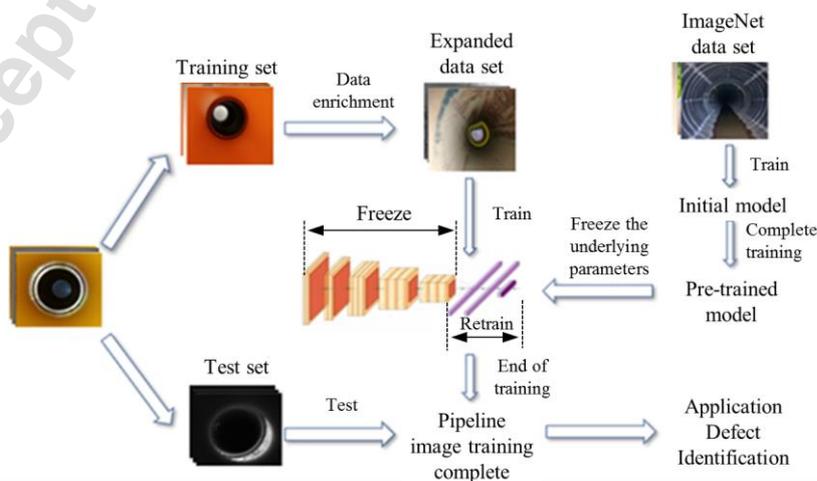


Fig.2 Advantages of migration learning over traditional machine learning

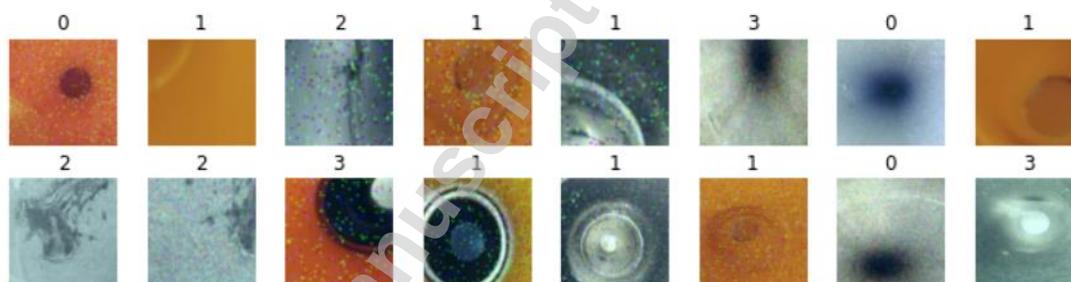
115
 116
 117

118 Generally when the target dataset is small, the number of frozen blocks needs to be decided according to the
 119 similarity of the features of the target dataset and the features of the ImageNet dataset. Because it is impossible
 120 to determine the extent of this similarity, this paper decides to build a Freeze scheme. The scheme is divided into
 121 a total of five times, each time the number of frozen convolutional blocks plus one, the parameter weights of the
 122 already frozen convolutional blocks are not changed, and the unfrozen convolutional blocks follow the estab-
 123 lished densely connected classifiers for training. The specific way is shown in Table2.

124 Table2 Freeze scheme

Scheme	Top	Block1	Block2	Block3	Block4	Block5
F_5	Train	Freeze	Freeze	Freeze	Freeze	Freeze
F_4	Train	Train	Freeze	Freeze	Freeze	Freeze
F_3	Train	Train	Train	Freeze	Freeze	Freeze
F_2	Train	Train	Train	Train	Freeze	Freeze
F_1	Train	Train	Train	Train	Train	Freeze

125
 126 80% of the dataset is divided into the training set and 20% is the test set, and the learning rate is set to be
 127 0.003, Dropout to be 0.3, and 16 samples to be a Batch-Size as shown in figure 3.



128 Fig.3 Images in a Batch

129
 130 The five schemes are assigned to train 15 Epochs, and the results are shown in figure 4. It can be seen that
 131 as the degree of freezing rises, the accuracy of the model also rises, and when freezing five convolutional blocks,
 132 probably because the image features of the target training set have a certain correlation with the image features
 133 of ImageNet, its accuracy reaches more than 85%, while the performance at the other degree of freezing is not
 134 very good, all below 80%, probably because the pipe defect image belongs to a small sample dataset, and its
 135 pipeline defect images belong to a small sample data set, there are relatively few images in the training set, and
 136 there are a lot of parameters that need to be learned and changed in the network model, which leads to unfavorable
 137 phenomena such as overfitting resulting in the model's accuracy can not be increased, because it can be concluded
 138 that the best degree of freezing is F_5.

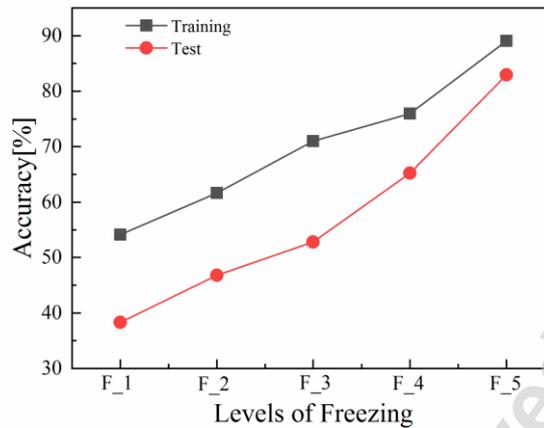


Fig.4 Comparison of Accuracy Rates for Different Levels of Freezing

139
140
141

2.3 Optimization of parameters and comparative analysis

143 After determining the best degree of freezing, the training accuracy of the PD-VGG network model is around
 144 88%, which is not particularly desirable, and it is necessary to improve the overall accuracy by changing the
 145 values of the model parameters and other values, so that the model is more suitable for pipeline defect images.

146 Adding the Dropout method to the PD-VGG model in this paper can avoid the overfitting phenomenon of
 147 the model to a certain extent. When the signal passes forward, Dropout randomly selects a neuron and stops that
 148 neuron with a certain probability, which reduces the overfitting of the neural network to the training data, as
 149 shown in figure 5. In order to make the model perform the best performance, five groups of different random
 150 inactivation percentages are compared and experimented in this paper to determine the most suitable Dropout
 151 parameters for the model, as shown in figure 6. The settings of the Dropout parameters are 0.1, 0.3, 0.5, 0.7, and
 152 0.9. It can be seen that as the probability of neuron inactivation increases, the accuracy of the model is gradually
 153 improving. When the probability is 50%, the improved accuracy reaches the maximum, and then shows a de-
 154 creasing trend. This is because at higher probability, the destruction of parameter update caused by neuron inac-
 155 tivation will result in a rapid decay of the recognition accuracy. Therefore, according to the experiments, this
 156 paper finally chooses the neuron inactivation probability of 50%, which means that the Dropout parameter is 0.5.

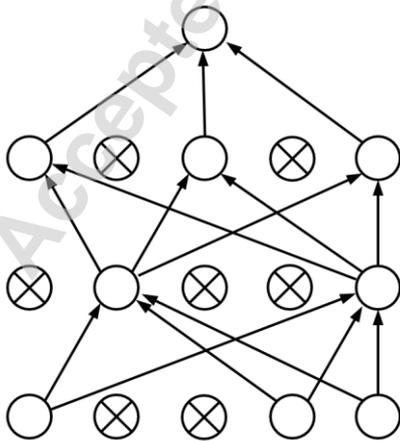


Fig. 5 Addition of Dropout neural network

157

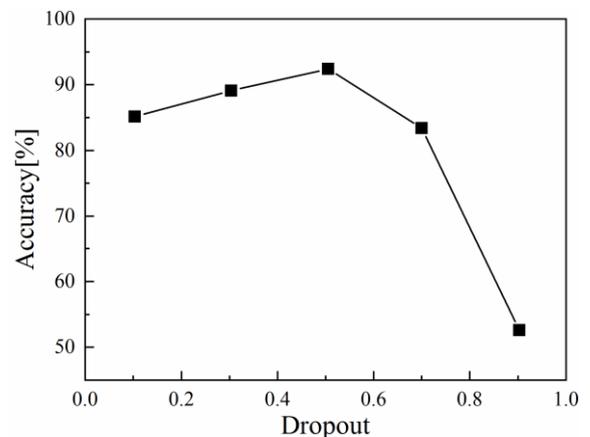


Fig. 6 Effect of stochastic inactivation on model accuracy

158 Because the PE gas pipeline defect image dataset belongs to small sample dataset, generally speaking, the

159 number of training rounds does not need to be a lot of convergence to achieve a certain degree of accuracy, this
 160 paper, Epoch selected 15 rounds, from figure 7, we can see that, in the case of the same number of rounds, with
 161 the increase in the Batch-Size, the accuracy of the model is also increasing gradually, in the Batch-size of 32, its
 162 training accuracy is the highest, and after convergence, with the increase of the number of batches, the accuracy
 163 shows a decreasing trend, which proves that the model does not reach the established accuracy under the large
 164 number of batches, and the model does not converge, for the consideration of the convergence speed, this paper
 165 sets the Batch-size to 32.

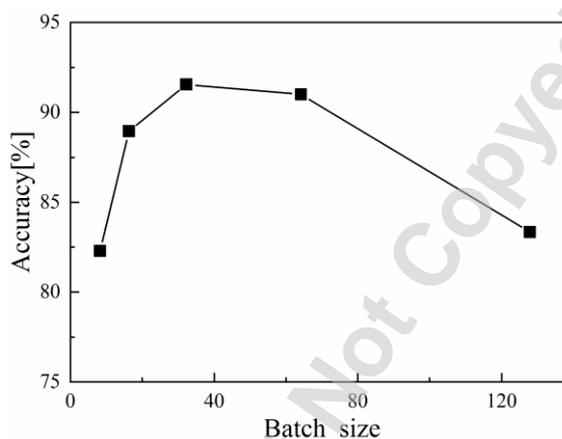


Fig. 7 Impact of number of batches under the same Epoch on modeling

166
167
168

169 In order to achieve the best performance of the PD-VGG model and to seek the optimal learning rate, this
 170 paper uses multiple sets of experiments to continuously adjust the learning rate of the model and to check the
 171 effect of the learning rate on the accuracy of the model and on the number of rounds in which there is no longer
 172 a significant change in the accuracy, and the data from five sets of data were selected from the multiple sets of
 173 experiments to be compared, as shown in Table 3.

174
175

Table 3 Effect of learning rate on the model

Learning rate	0.3	0.03	0.003	0.0003	0.00003
Accuracy rate (%)	80.37	83.17	88.98	93.44	93.58
No. of rounds (pcs)	10	10	15	25	>30

176
177

178 According to the above information, it can be seen that the model accuracy improved by roughly 8% during
 179 the reduction of the learning rate from 0.3 to 0.003, and then the trained Epoch increased to 5 rounds. In the
 180 reduction to 0.0003, the accuracy reached a more satisfactory 93.44%, but the trained Epoch increased to 25
 181 rounds. After lowering it again by an order of magnitude, the change in accuracy was not significant, but the
 182 Epoch increased a lot, increasing the training time a lot. This is probably because a smaller learning rate slows
 183 down convergence and increases the time to find the optimum. Although a learning rate of 0.00003 presents better
 184 results than 0.0003, the former requires a large number of training rounds and more training time. Therefore, the
 185 learning rate of 0.0003 is finally chosen in this paper.

186 Based on the results of the above sets of experiments, this paper finalizes the PD-VGG model. Then the
 optimization parameters in the image recognition process of PE gas pipeline defects are determined. Among them,

187 Dropout random deactivation is determined as 0.5, the number of Batch-Size batches is determined as 32, and
 188 the learning rate of the optimizer is determined as 0.0003. The designed PD-VGG is compared with AlexNet,
 189 VGG-16, SVM, and VGG-19 to verify the reliability and effectiveness of the algorithm in this paper, as shown
 190 in Table 4.

191 Table 4 Comparison of training results for each network

Network models	Training Accuracy (%)	Training loss	Test Accuracy (%)	Test Loss
AlexNet	82.25	0.9068	69.27	1.046
VGG-16	88.98	0.8818	79.85	0.9603
VGG-19	87.45	0.9096	72.36	1.008
SVM	93.37	0.8122	82.74	0.9172
PD-VGG	94.76	0.7974	86.89	0.8765

192
 193 As can be seen from the information in Table 4, the AlexNet model has the lowest accuracy and poor gen-
 194 eralization ability when dealing with the PE gas pipeline defect image dataset. The PD-VGG training established
 195 in this paper has the highest accuracy and the smallest gap between training and testing. This largely reduces the
 196 degree of overfitting. In comparison, it can be found that the network model in this paper handles the best results.

197 In order to more intuitively observe how good the improved PD-VGG model is, neural network performance
 198 metrics are used to judge it. The neural network performance metrics are the ones that can intuitively understand
 199 the performance of a model is good or bad, including accuracy, recall, precision, specificity, and F-measure value.
 200 They are all calculated based on the positive and negative samples in the sample data.

201 **Accuracy** is how much the model judges correctly in all samples. The calculation is shown in equation 1.

$$202 \quad Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

$$203 \quad Recall = \frac{TP}{TP + FN} \quad (2)$$

$$204 \quad Precision = \frac{TP}{TP + FP} \quad (3)$$

$$205 \quad Specificity = \frac{TN}{FP + TN} \quad (4)$$

$$206 \quad F - measure = \frac{(\alpha^2 + 1) * Precision * Recall}{\alpha^2 (Precision + Recall)} \quad (5)$$

207 Where, TP , FP , TN , and FN denote numbers of true positives, false positives, true negatives, and false negatives,
 208 respectively. Accuracy and F-measure are used to evaluate overall performance. F-measure is calculated so that
 209 α takes the value 1.

210 Table 5 Fine-tuned performance parameters

Types	Accuracy	Precision	Recall	Specificity	F-measure
Deformation	0.996	1.0	0.979	1.0	0.989

Misalignment	0.974	0.957	0.917	0.989	0.936
Breakage	0.952	0.885	0.987	0.934	0.933
Normal	0.974	1.0	0.891	1.0	0.942

211 In the performance parameters of the various types of parameters is naturally closer to "1" the better, as can
212 be seen in the table, almost all the performance parameters of the identification of the types of parameters tend
213 to be close to 1, the best has been equal to 1, the performance of the worst indicators is only 0.891, and the ideal
214 state of the difference of 0.109, but in general, the PD-VGG network model adapted to the of the PE gas pipe
215 defect image dataset.

216

217 **3 Research on image processing algorithm based on pipe defor-** 218 **mation**

219 By recognizing the images of PE gas pipelines, the basic classification of defect images is completed. In this
220 paper, we have selected a kind of defects that need to distinguish the degree of damage and have a big impact on
221 the life of the pipe - deformation defects, which refers to the appearance of polyethylene gas pipeline in the
222 service process by the extrusion of external factors or ground settlement caused by the appearance of the change
223 in morphology. The general observation of the degree of deformation is to compare the deformation ratio, i.e. the
224 ratio of the deformed pipe diameter to the original pipe diameter. However, it is found that the environmental
225 factors are not easy to extract the pipe diameter, so the deformation depth of the pipe will be used in the post-
226 processing to carry out the related research.

227

228 **3.1 Graying of pipeline deformation defect images**

229 Color images can truly reflect the most realistic conditions inside the pipe at this moment, such as the pres-
230 ence of defects, color changes in the pipe body, and the size of pipe deformation. It can maximize the information
231 of the acquired image in front of the technicians. However, color images contain a large amount of color infor-
232 mation and consume a lot of memory space for storage. The use of color images in the recognition process also
233 affects the computing speed, which results in a decrease in processing speed. According to the actual situation of
234 detection, the structural defect detection of PE gas pipeline only needs to be discriminated by the edge information
235 of the pipeline and the gray scale change of the pipeline image, and the color information of the image will not
236 play a decisive role. In this paper, there is no need to retain the color information while performing the recognition,
237 but only need to know whether the edge information in the defective image of the pipeline and the gray scale
238 information of the pixels in the image are drastically changed or not. Therefore, it is necessary to convert the
239 color image into gray scale image. Generally, image graying is processed by component method, maximum value
240 method, mean value method and weighted average method. Through comparison, this paper chooses the weighted
241 average method to gray scale the image. Weighted average method according to the formula (6) for the three
242 components of RGB weighted average can get a more reasonable grayscale image. Empirical formula is as fol-
243 lows,

$$Gray(i, j) = 0.299 \times R(i, j) + 0.578 \times G(i, j) + 0.114 \times B(i, j) \quad (6)$$

244 All the above equations represent the grayscale value $Gray(i, j)$ at coordinate (i, j) , $R(i, j)$ and $G(i, j)$
 245 and $B(i, j)$ represent the three color component luminance values of the pixel point, respectively.

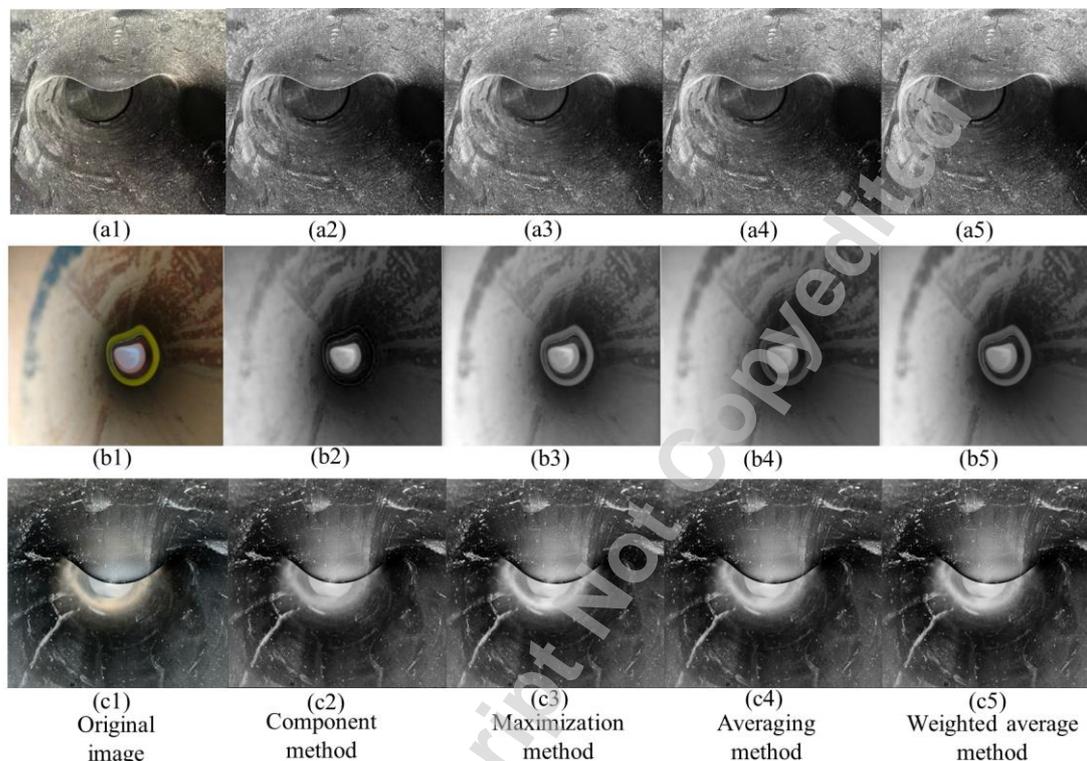


Fig. 8 Deformation defects of PE pipes

246
 247
 248

249 From the figure 8 we can see that the resultant images after processing of the component method are all
 250 relatively darker, the average method and the maximum value method perform similarly in the pipeline image,
 251 while the weighted average method better retains the details of the original image, and the information of the
 252 original image is more similar to that of the original image, in order to better carry out the following image
 253 processing, so in this paper, we will use the weighted average method as the algorithm for the grayscaling of the
 254 image.

255 3.2 Gas pipeline image enhancement based on improved gamma correction

256 The image enhancement process is to enhance the contrast between the defective part of the image and the
 257 background part, so that the pipeline defective image is more recognizable. As shown in figure 9, the following
 258 conclusions can be drawn after using the image enhancement algorithm. The histogram equalization image not
 259 only enhances the contrast between the background and the defects, but also highlights the noise in the figure,
 260 and the brightness of the equalized image is not uniform compared with the original defect image. The histogram
 261 equalization algorithm is unselective of pixel points when processing the image, which will enhance the noise
 262 intrusiveness after processing and cause unnecessary trouble to the subsequent feature extraction. Therefore, in
 263 this paper, histogram equalization will not be used to enhance the image.

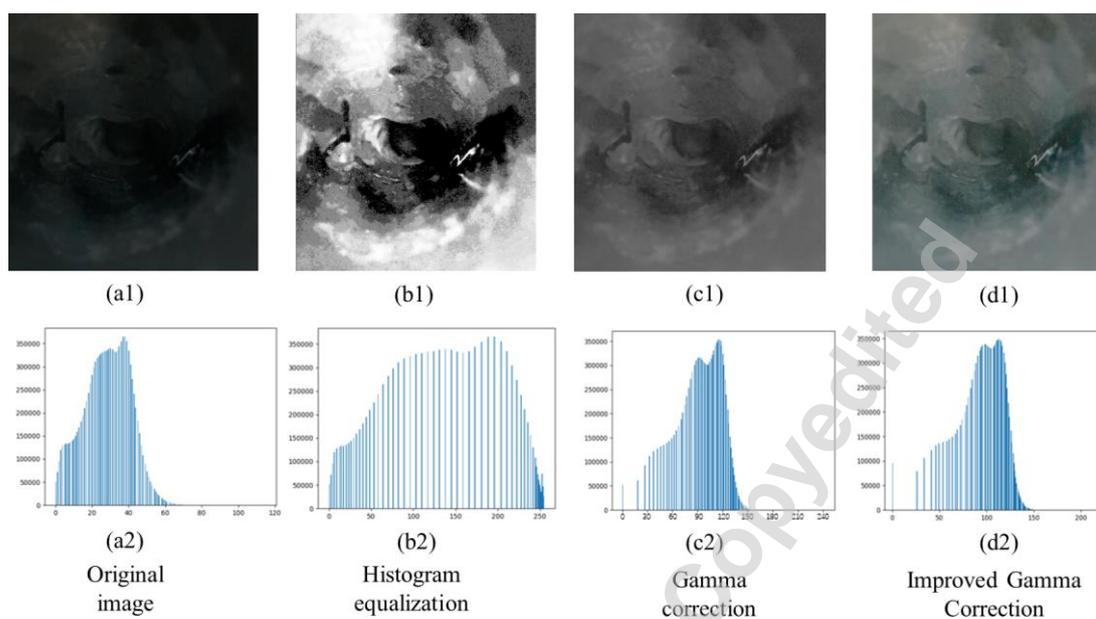


Fig.9 Comparison of image enhancement algorithms

264
 265
 266

267 The image enhanced using the improved gamma correction algorithm is shown in figure 9(d), from which
 268 it can be seen that there is no distortion in the defect picture after gamma correction. At the same time, the defects
 269 are more distinct from the background for better viewing. Moreover, the edges of the defects are more prominent,
 270 the difference with the background is greater, it becomes clearer, and the contrast between the background of the
 271 pipeline and the defects present in the pipeline is obvious. The improved gamma correction algorithm not only
 272 reconciles the uneven illumination of the original image but also suppresses the interference points that affect the
 273 recognition. Therefore, this paper decides to use improved gamma correction to enhance the contrast between the
 274 pipe background and pipe defects.

275 3.3 Defect image blurring and edge detection

276 In order to better extract the contour from the pipe defect image, for the pipeline internal image detection of
 277 the actual working conditions, in-depth comparative analysis, the results are shown in figure 10. Figure 10(a)~(f)
 278 are the original image, mean filter, box filter, Gaussian filter, median filter and bilateral filtering. From this figure,
 279 it can be seen that there are more impurities on the pipe wall, and there is some noise interference in the captured
 280 image, which will inevitably cause more or less influence and interference on the later deformation defect edge
 281 extraction. The median filter retains the details of the original image the best, and most of the details of the tube
 282 wall are retained, but this is not what is expected in this study. Linear filters such as Gaussian, while smoothing
 283 the impurity information of the pipe wall, are inferior to bilateral filters. The latter blurs out the unwanted inter-
 284 ference by smoothing, yet retains the contour information of the deformation defects well. Bilateral filtering will
 285 be investigated in subsequent image processing work.

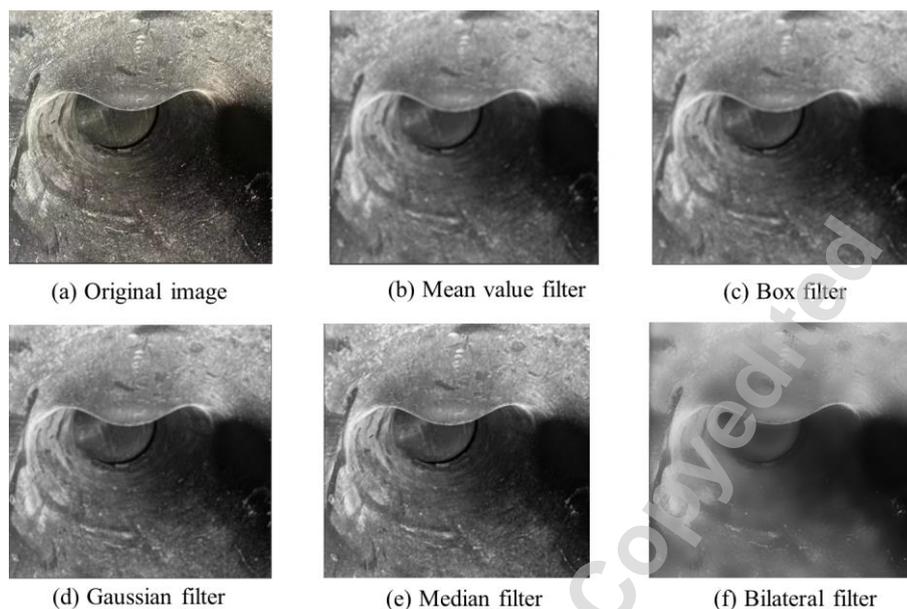


Fig. 10 Comparison of Filtering Algorithms

286
287
288

289 After the initial screening and preprocessing of the acquired image, in order to subsequently calculate the
 290 amount of geometric features to be detected image and used for pattern recognition classification, the need for
 291 pipeline defects edge extraction. Due to the PE gas pipe internal smooth, so this time the prominent edge is a
 292 very important feature of the whole image, this feature for image segmentation and determine which type of
 293 defects provide an important basis. Generally in industrial vision, often used edge detection algorithms for Sobel,
 294 Scharr, Laplacian, Canny. based on the deformation of defective images, the effect of these types of edge detec-
 295 tion algorithms are shown in figure 11.

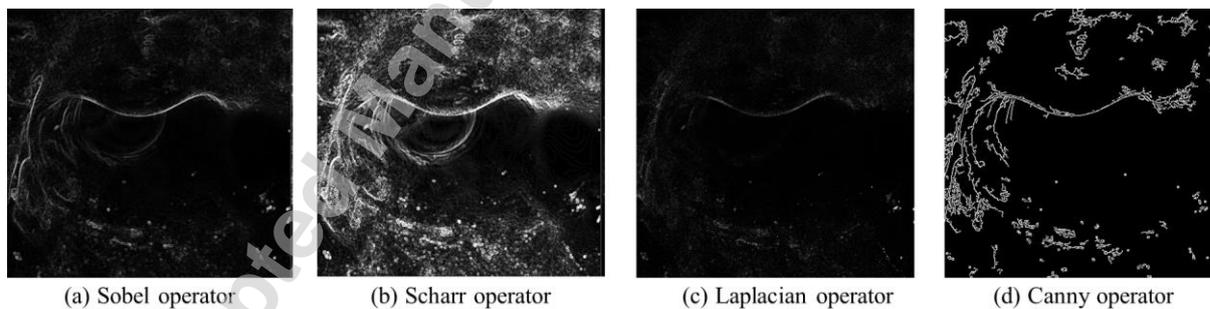


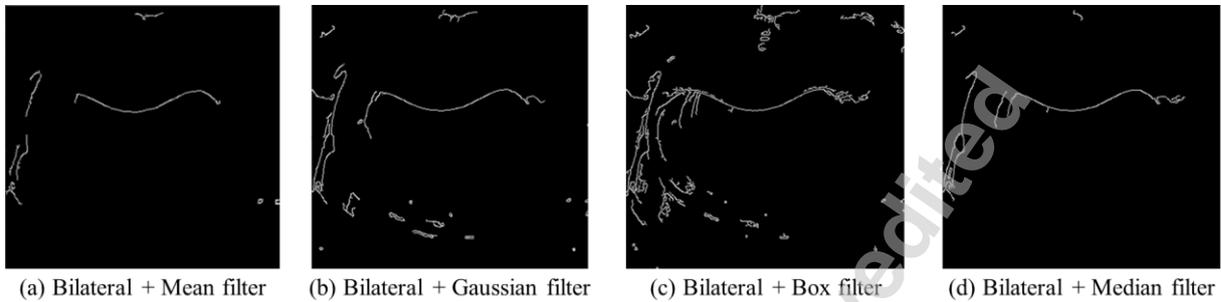
Fig.11 Comparison of Edge Detection Algorithms

296
297
298
299

300 From the figure, it can be seen that the Sobel and Laplacian contour extraction is too diluted in the defor-
 301 mation defect part of the contour relative to other operators, and the Scharr operator, although the defect contour
 302 is well preserved, retains too much interference from impurities in the pipe wall, which is not suitable for edge
 303 extraction. While Canny operator reflects the deformation contour better, compared with Scharr operator, the
 304 interference is removed more, but there are still more non-essential contours, so the later study will start from the
 direction of reducing the interference.

305 After preliminary research, bilateral filtering and Canny operator is more suitable for PE gas pipeline defects
 306 contour extraction, bilateral filtering for edge retention is better, but interference, noise removal is not very com-
 307 plete, however, the mean, box and other linear filtering for non-contour and other interference removal is very
 308 good, so try to double filtering to eliminate the interference of the pipe wall. Double filtering is to use the bilateral

309 filter as the base, and add another layer or even multiple layers of filters on this basis to realize better removal of
310 interference noise. The specific effect of Canny edge detection of deformation defects after double filtering is
311 shown in figure 12.



312 (a) Bilateral + Mean filter (b) Bilateral + Gaussian filter (c) Bilateral + Box filter (d) Bilateral + Median filter
313 Fig.12 Canny edge detection after double filtering
314

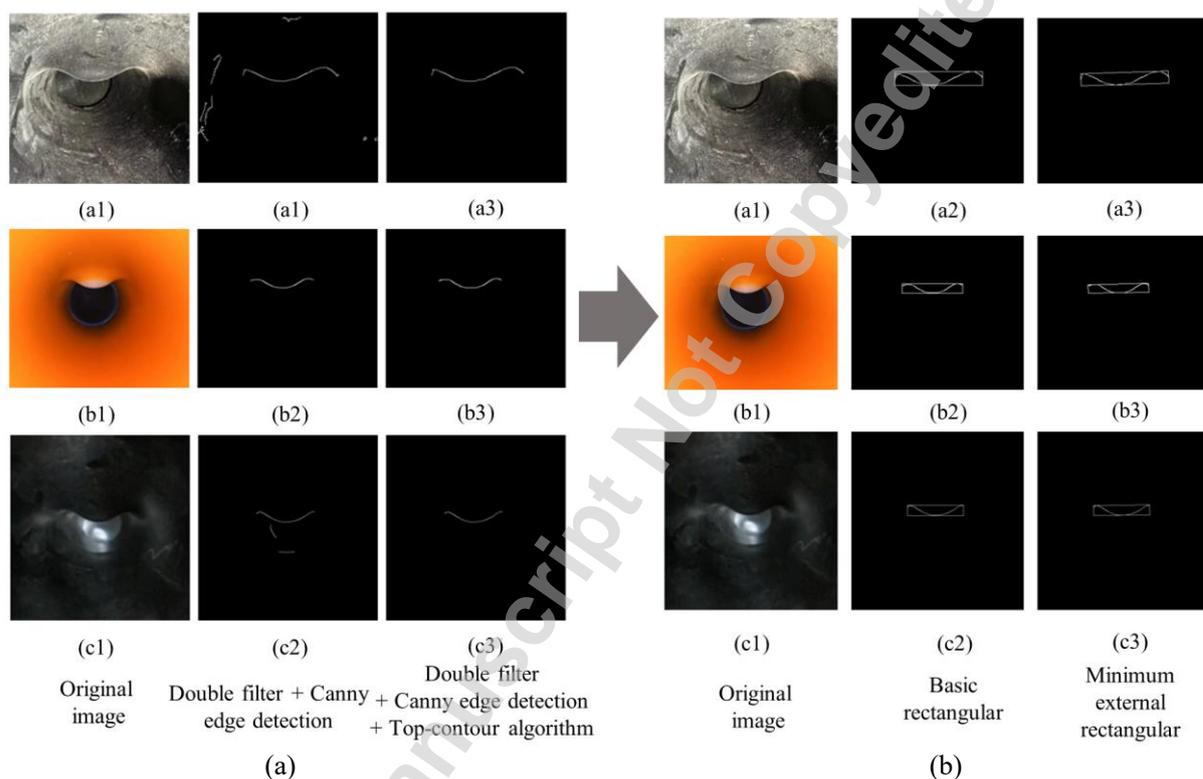
315 From the figure, it can be clearly seen that the "Bilateral + Median" combination is the least ideal for the
316 processing of interference, and the "Bilateral + Gaussian" and "Bilateral + Box" groups are relatively better, but
317 some other non-defective contours are recognized as strong edges or weak edges with connections, and the re-
318 moval of the interference part is incomplete. Although the two groups of "Bilateral+Gaussian" and "Bilat-
319 eral+Box" are relatively better, some of the other non-defective contours are recognized as strong edges or weak
320 edges with connections, and the interference part of the connectivity domain is larger, so the removal is incom-
321 plete. The combination of "bilateral + mean" is the best group for the deformation of the defective contours, the
322 interference part of the removal of more and not a wide range of connectivity, the deformation of the defective
323 contour part of the retention of the good, easy to carry out the later work. Double filtering is the best processing
324 method for PE gas pipeline deformation defect contour detection, and finally the Canny edge detection with
325 double filtering is chosen as the pre-deployment of edge extraction.

326 3.4 Pipe deformation defect image contour extraction

327 Canny edge detection can detect edges, but the detected edges are not a continuous whole. Contour extrac-
328 tion, on the other hand, can connect the needed edges into a whole for later computation. Two operators findCon-
329 tours and drawContours are used most in contour extraction, the former can find the existence of the contour
330 information in the input image, the latter can find the contour information in a certain format to draw out.

331 In this paper, we propose the top-contour algorithm, which compares the entire extracted contour to a nu-
332 merical value, traverses it and then sorts it by contour size to customize the interception of one or more contours.
333 Because the current acquisition of deformation defects in the image edge detection after the defect edge contour
334 in the whole image belongs to the largest or larger class, the output for a maximum contour as a top-contour
335 traversal results, so that after a screening contour, that is, to save the energy of manually selecting the contour,
336 but also to facilitate the direct drawing of the contour. In order to test the universality of this paper will be different
337 environments under the PE gas pipeline deformation defects in the image of the corresponding contour extraction
338 process, the specific processing results shown in figure 13 (a). Figure 13 (a) in a, b, c represents the deformation
339 defect images taken in different environments, from 1 to 3 is the original input image, after double filtering and
340 canny edge detection of the image (too clean and smooth pipeline is only a single layer of bilateral filtering), top-
341 contour algorithm processed image. As can be seen from the figure, for pipes that are not smooth and clean, such

342 as a1 and c1, there are still these unwanted interference and impurities in the contours of the image after double
 343 filtering and canny edge detection, but after using the top-contour algorithm to traverse all the detected contours
 344 in the image and extract the established maximal contour, the interferences due to the pipe wall or the quality of
 345 the image are all removed, and only the remaining Only the needed deformation defect contours are drawn by
 346 the drawContours operator, which not only saves the time of manually selecting the contours, but also facilitates
 347 the extraction of deformation defect contour feature values and operations at a later stage.



348
 349
 350 Fig. 13 Deformation defect contour extraction effect (a) and comparison of defective outer rectangular boxes
 351 (b) in different environments
 352

353 4 Defect feature extraction and calculation

354 4.1 Deformation defect contour feature information extraction

355 After the preliminary process, the contour of the pipe deformation defects has been extracted, in order to
 356 better observe the deformation condition of the defect contour, the next need to solve the problem is to extract
 357 the relevant information of the contour, in order to study the depth data of the deformation defects. In this paper,
 358 we use the method of contour fitting to approximate the defect contour by contour, and extract relevant feature
 359 information through the fitted contour.

360 In this paper, a rectangular frame is chosen as the outer frame of the deformation defect contour. There are
 361 two specific methods of external connection, the first is the basic external rectangular frame, and the other is the
 362 minimum external rectangular frame. In this paper, the two methods are used in different environments and dif-
 363 ferent deformation depths to try to calculate, as shown in figure 13 (b).
 364

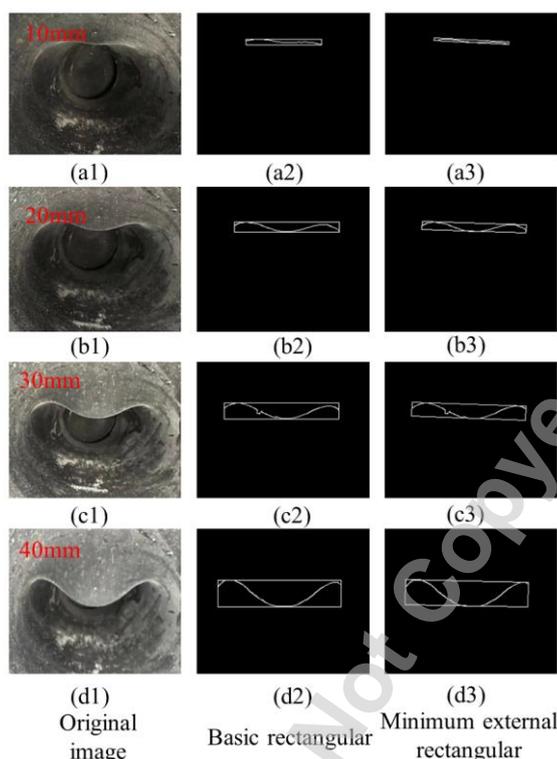


Fig.14 Comparison of the outer rectangular box of defects at different depths

Figure 14 (a) to (d) are the images of the deformation defects at different depths (from 10mm to 40mm), and 1 to 3 are the original defects, the base outer rectangular box, and the minimum outer rectangular box, respectively.

From these two figures, it can be found that the minimum external rectangular box deflection angle formed by the defect contour at different depths in different environments is not much relative to the X-axis. The area and width of the rectangular box are increasing as the depth of the defect is getting larger and larger. The minimum outer rectangle is better suited to deformed defects though, and can be changed as the angle of the defect varies. One of the problems is that it needs to cycle through the angle calculations, which takes more running time compared to the base rectangular box. In this paper, a set of experiments was done to compare the area and runtime of the two rectangular boxes, as shown in figure 15.

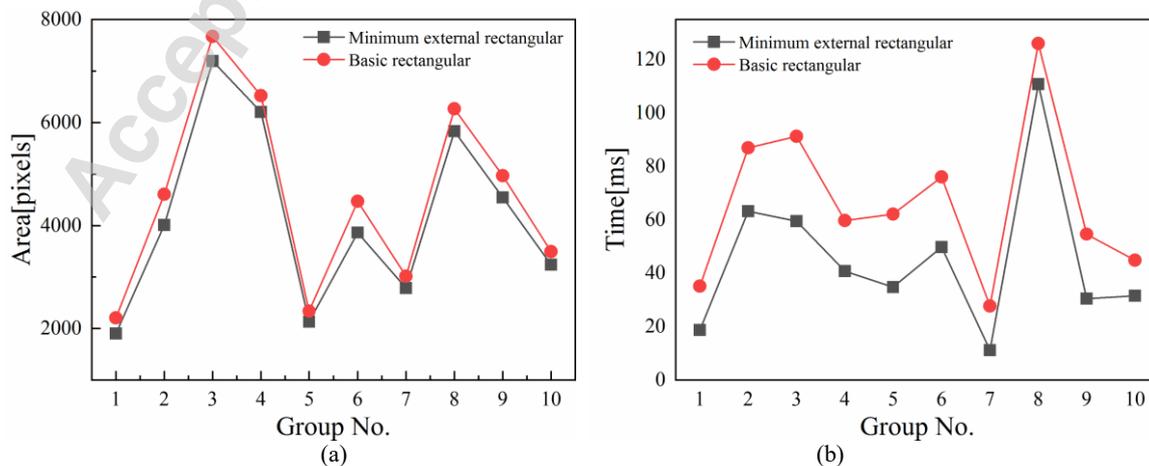


Fig. 15 Comparison of area calculation (a) and runtime (b) of different algorithms

As shown in figure 15, this paper selected ten groups of test patterns, the data are averaged over many runs.

383 As can be seen from the figure, it can be seen that although the computation time of the minimum outer rectangle
384 is more time-consuming than that of the rectangle without rotation, the time difference is within 9 milliseconds,
385 which is within the acceptable range. Also the area enclosed by it is smaller than the latter value. For non-enclosed
386 contours such as deformation defects, the smaller the enclosed area formed, the better, so the minimum outer
387 rectangle algorithm is finally chosen as the contour fitting.

388 4.2 Calculation of eigenvalues of defective contours

389 After the contour fitting, the eigenvalues of the deformed defect contour and its fitted contour can be calcu-
390 lated, and its attribute features and enclosing object features are of great significance for describing the infor-
391 mation of the image, and this paper will carry out the calculation of the eigenvalues from several aspects.

392 The area of minimum outer rectangle algorithm can be obtained from the return value, the larger the area,
393 the greater the degree of deformation in the pixel case, and the same shooting distance of the target defect, as
394 shown in equation 7, for the area of minimum outer rectangle algorithm, W is the width of the rectangle, H is the
395 height.

$$396 S_w = W * H \quad (7)$$

397 The area of the defect itself, which is calculated according to Green's formula (Type II curve integral),
398 characterizes the connection and relationship between the curve C in the two-dimensional image with respect to
399 the coordinate curve integral and the dual integral enclosed by that curve itself, as shown in equation 8.

$$400 \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = \int_C P dx + Q dy \quad (8)$$

401 For the perimeter of the defect, it is based on the L2 distance algorithm to find its perimeter, as shown in
402 equation 9.

$$403 C = \sum_{k=0}^n \frac{(x_{k+1} + x_k)(y_{k+1} - y_k)}{2} \quad (9)$$

404 Based on the deformation defects, the formation of minimum outer rectangle algorithm, its aspect ratio for
405 the consideration of the depth of the defect is very important, the lower its value represents the depth of the deeper,
406 proving that the pipeline due to deformation of the degree of damage is greater, and its calculation method is
407 shown in equation 10, AR is the aspect ratio, W is the width, H is the height.

$$408 AR = W / H \quad (10)$$

409 Extend is a commonly used value for contour features, representing the ratio of the area of the target contour
410 and the area of the outer boundary, calculated as shown in equation 11, for the area of the contour itself, for the
411 area of the outer contour.

$$412 Extend = \frac{S_T}{S_w} \quad (11)$$

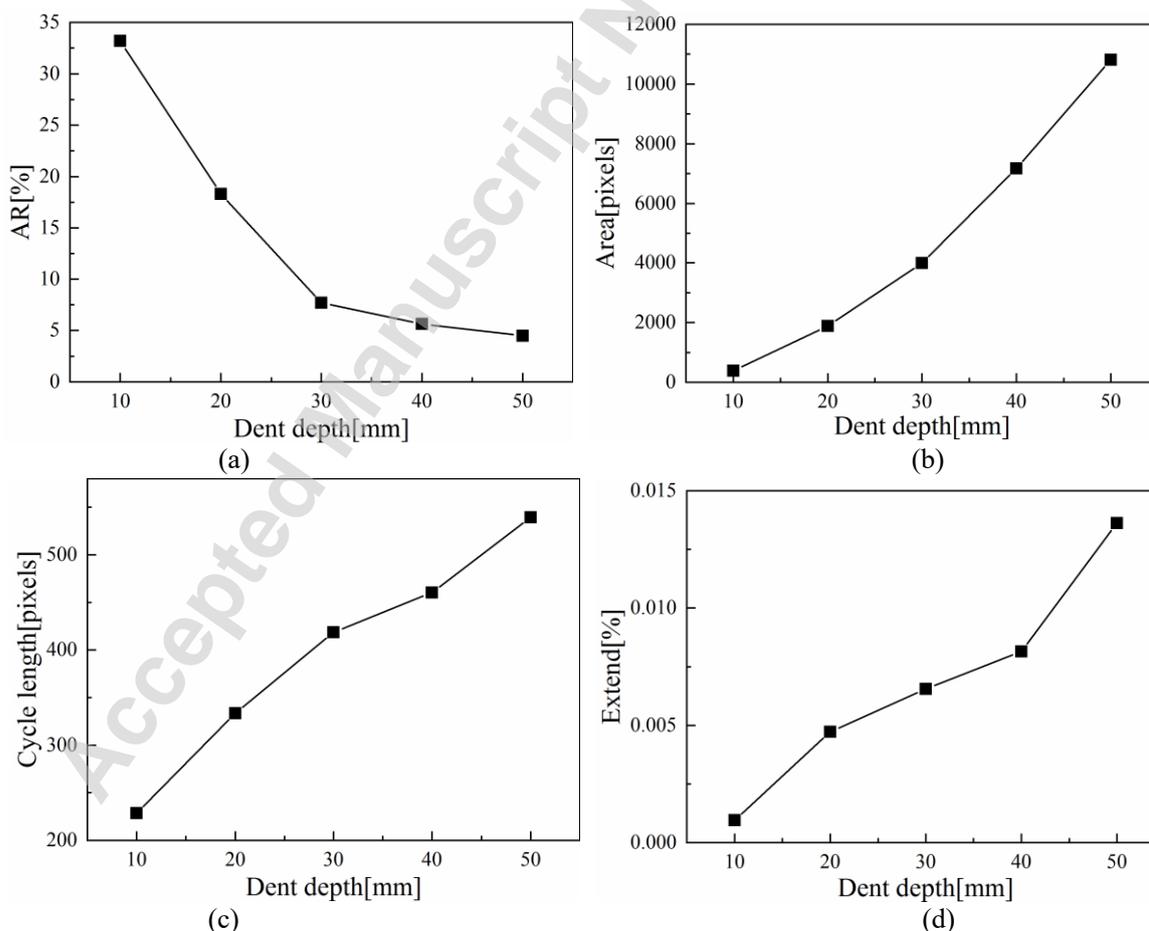
413 In this paper, twenty images with a shooting spacing of about 30cm were selected, whose defect depths

414 varied from 10mm to 50mm, and the relevant calculations were performed on the image contours. At the same
 415 time, the values were averaged to observe how the eigenvalues differed at different depths, as shown in table 6,
 416 where the units of area and perimeter are pixels.

417 Table 6 Eigenvalues of deformation defects at different depths

Depth	AR	Minimum External Rectangular Area	Cycle length	Extend
10mm	33.28	412.36	229.58	9.416e-4
20mm	18.36	1893.30	333.48	4.76e-3
30mm	7.72	4007.28	419.54	6.52e-3
40mm	5.60	7193.88	460.17	8.14e-3
50mm	4.52	10805.24	539.62	1.36e-2

418 In order to better observe the relationship between different deformation defects and the feature information
 419 of their own contours and outer rectangular boundaries, a line graph of their relationship is plotted, as shown in
 420 figure 16.



421
422

423
424

425 Fig. 16 Comparison of eigenvalues of deformation defects for different depth cases

426
427

428 From Figure 16, it can be seen that as the depth continues to deepen, its width-to-height ratio shows a down-
 429 ward trend, the most serious decline in the depth of 30mm, proving that in this case, the deformation defects of
 the degree of damage is very large, and there may be the phenomenon of yielding of the pressurized portion of

430 the region; and minimum outer rectangle algorithm and contour perimeter shows a rising trend, Extend is also
431 the same, but the growth trend of the trend in the point of passing through the 30mm from a slow change to a
432 steeper, and the results proved that deformation defects are more obvious in the change at about 30mm, but to be
433 further investigated.

434 **5 Conclusion**

435 In this paper, an image processing method is used to conduct an in-depth study on the image defect recog-
436 nition and feature detection algorithm, and a set of algorithms applicable to PE gas pipeline defect detection is
437 proposed. The following conclusions are drawn: based on the built image acquisition system and the constructed
438 image library, the improved CNN model is used to recognize and classify the defective images of PE gas pipelines.
439 The data set of specific defects of gas pipeline is established, and data augmentation is carried out, and PD-VGG
440 is used for migration learning and freezing network layer to complete the model building. Finally, the parameter
441 optimization is carried out for several model parameters to improve the accuracy of the training model, which
442 proves the applicability of the model to the defective images of gas pipelines. OpenCV-Python is utilized to write
443 algorithms for related image processing. Through the noise analysis of PE gas pipeline defects, the image is first
444 grayscaled by the weighted average method, then the contrast is enhanced by gamma correction, and finally the
445 double filtering method is used to remove the influence of noise on the subsequent steps. Edge detection for
446 deformed defect images was accomplished using Canny operator. In contour extraction, it was found to be too
447 cumbersome to find the desired contour, and the top-contour algorithm was used to complete the filtering between
448 the deformed defective contour and the interfering contour.

449 **Declaration of conflicting interests**

450 The authors declare that they have no known competing financial interests or personal relationships that
451 could have appeared to influence the work reported in this paper.

452 **Funding**

453 This research was funded by Xiong'an New Area Science and Technology Innovation Project, National Key
454 R&D Program of China (No. 2022XAGG0147), Young Scientific and Technological Talent Project of China
455 Special Equipment Inspection and Research Institute (No. KJYC-2023-02) and Opening Project of Guangdong
456 Provincial Key Laboratory of Technique and Equipment for Macromolecular Advanced Manufacturing (No.
457 20240575).

458

Reference

- [1] Zha, S., Lan, H-q., Huang, H. (2022). Review on lifetime prediction of polyethylene pipes: limitations and trends. *International Journal of Pressure Vessels and Piping*. 198:104663.
- [2] Zha, S., Lan, H-q. (2021). Fracture behavior of pre-cracked polyethylene gas pipe under foundation settlement by extended finite element method. *International Journal of Pressure Vessels and Piping*. 189:14270.
- [3] Duran, O., Althoefer, K., & Seneviratne, L. D. (2002). Automated sewer inspection using image processing and a neural classifier. *Proceedings of the 2002 International Joint Conference on Neural Networks*. IJCNN'02, New York, 1126-1131.
- [4] Niemueller, T. (2006). Automatic detection and segmentation of cracks in underground pipeline images seminar. *Medical Image Processing Summer Semester*, Institut für medizinischen Informatik, Aachen, Germany.
- [5] Shinde, M. D., & Wane, K. (2016). An Application of image processing to detect the defects of industrial pipes. *International Journal of Advanced Research in Computer and Communication Engineering*. 5(3), 979- 981.
- [6] Mashalizadeh, A. M., Delavari, H., & Razaghian, F. (2014). Defect Detection in Drilling Pipes, using Combination of Artificial Neural Networks and Machine Vision Techniques. *Computer Science, Engineering*. 2(8), 48-57.
- [7] Wang Y, Dai Y, Liu X, et al. (2018). Study on the method of colour image noise reduction based on optimal channel-processing. *IET Image Processing*, 2018, 12(9):1555-1559.
- [8] Shao, J., Du, D., Chang, B., & Shi, H. (2012). Automatic weld defect detection based on potential defect tracking in real-time radiographic image sequence. *NDT & E International*, 46, 14-21.
- [9] Zahran, O., Kasban, H., El-Kordy, M., & El-Samie, F. E. A. (2013). Automatic weld defect identification from radiographic images. *NDT & E International*, 57, 26-35.
- [10] Guo, X., Tang, C., Zhang, H., & Chang, Z. (2012). Automatic thresholding for defect detection. *ICIC Express Letters*, 6(1), 159-164.
- [11] Alaknanda, Anand, R. S., & Kumar, P. (2006). Flaw detection in radiographic weld images using morphological approach. *NDT & E International*, 39(1), 29-33.
- [12] Motamedi, M., Faramarzi, F., & Duran, O. (2012). New concept for corrosion inspection of urban pipeline networks by digital image processing. *Conference of the IEEE Industrial Electronics Society*, IEEE, New York, 1551-1556.
- [13] Khalifa, I., Aboutabl, A. E., Barakat, G.S.A. (2013). A new image -based model for predicting cracks in sewer pipes. *International journal of advanced computer science and applications*, 4(12):65-71.
- [14] Kirstein, S., Muller, K., Walecki-Mingers, M., & Deserno, T. M. (2012). Robust adaptive flow line detection in sewer pipes. *Automation in Construction*, 21, 24-31.
- [15] Alam, M. A., Ali, M. N., Syed, A. A., Sorif, N., & Rahaman, M. A. (2015). An algorithm to detect and identify defects of industrial pipes using image processing. *Proceedings of the 8th International Conference on Software, Knowledge, Information Management and Applications*. Dhaka, Bangladesh, 1-6.
- [16] Huynh, P., Ross, R., Martchenko, A. (2015). Dou-edge evaluation algorithm for automatic thin crack detection in pipelines. *Proceedings of 2015 IEEE International Conference on Signal and Image Processing Applications*. Kuala Lumpur, Malaysia, 191-196.
- [17] Mashford, J., Rahilly, M., Lane, B., Marney, D., Burn, S. (2014). Edge detection in pipe images using classification of haar wavelet transforms, *Applied Artificial Intelligence*, 28:7, 675-689.
- [18] Wei, L. I., Li, L., & Da, J. Y. (2003). Application of wavelet and laplacian operator in pipe welding flaw image detection. *Microcomputer Development*. 12, 71-74.
- [19] Altabay, W. A., Kouritem, S. A., Abouheaf, M. I., & Nahas, N., (2022). Research in image processing for pipeline crack detection applications," *2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Maldives, Maldives, 2022, 1-6.

- [20] Su, T. C., & Yang, M. D. (2014). Application of morphological segmentation to leaking defect detection in sewer pipelines. *Sensors*, 14(5), 8686-8704.
- [21] Sinha, S. K., Fieguth, P. W. (2006). Neuro-fuzzy network for the classification of buried pipe defects. *Automation in Construction*, 15(1), 73-83.
- [22] Zhou, A., Zheng, H., Li, M., & Shao, W. (2020). Defect Inspection Algorithm of Metal Surface Based on Machine Vision. *12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, Phuket, Thailand, 45-49.
- [23] Moselhi, O., Shehab-Eldeen, T. (1999). Automated detection of surface defects in water and sewer pipes. *Automation in Construction*. 8(5), 581-588.

Accepted Manuscript Not Copyedited