Robust Segmentation Method for Noisy Images Based on an Unsupervised Denosing Filter

Ling Zhang, Jianchao Liu, Fangxing Shang, Gang Li*, Juming Zhao, and Yueqin Zhang

Abstract: Level-set-based image segmentation has been widely used in unsupervised segmentation tasks. Researchers have recently alleviated the influence of image noise on segmentation results by introducing global or local statistics into existing models. Most existing methods are based on the assumption that the distribution of image noise is known or observable. However, real-time images do not meet this assumption. To bridge this gap, we propose a novel level-set-based segmentation method with an unsupervised denoising mechanism. First, a denoising filter is acquired under the unsupervised learning paradigm. Second, the denoising filter is integrated into the level-set framework to separate noise from the noisy image input. Finally, the level-set energy function is minimized to acquire segmentation contours. Extensive experiments demonstrate the robustness and effectiveness of the proposed method when applied to noisy images.

Key words: image segmentation; noisy image; level set; autoencoder

1 Introduction

Scholars have widely studied the segmentation models of noisy images after the introduction of image segmentation, and they have regularly proposed new models for noisy images[1–6], including the introduction of statistics to global or local segmentation models to alleviate the influence of noise on local information calculation[7–10]. These models have certain guiding significance in the segmentation of noisy images. The models are also based on artificial prior knowledge, and noise is removed using filters.

- Ling Zhang, Jianchao Liu, and Gang Li are with College of Software, Taiyuan University of Technology, Taiyuan 030024, China. E-mail: zl2090@126.com; 313013424@qq.com; tx2090@126.com.
- Fangxing Shang and Juming Zhao are with College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China. E-mail: 359809280@qq.com; zhaojumin@tyut.edu.cn.
- Yueqin Zhang is with Information Technology Department, Shanxi Tizones Technology Co., Ltd, Taiyuan 030024, China. E-mail: zhangyq133@163.com.
- ∗ To whom correspondence should be addressed.
  Manuscript received: 2021-02-27; accepted: 2021-03-12

In 2011, Li et al.[11] proposed a new processing scheme, in which the idea of an offset field was integrated into a local segmentation model and ultimately achieved good segmentation results for images with large degrees of gray unevenness. However, the premise of this model is to assume that a grayscale offset field is continuous, smooth, and slowly changing, and noise distribution must be a standard normal distribution with a mean value of 0. These assumptions greatly limit the application of the model.

In 2020, Xu et al.[12] proposed a method for removing noise from noisy images with unknown noise distribution. The basic idea is to treat a collected noise image as a clean image of a supervised training model and then superimpose a layer of noise on it as input. A neural network is used in training to obtain a denoising mapping relationship. This mapping relationship is adopted to remove noise from an unknown image of the original noise distribution.

Previously proposed noise separation models are based on the assumption that the distribution of noise is known for denoising, and no actual noise modeling is performed. Although some models can model noise,

they are based on clean inputs without noisy images. However, real-time collected data are often difficult to assume to be free of noise or have unobservable noise distribution. Therefore, the current study abandons the abovementioned assumption of previous models and proposes an unsupervised noise separation mechanism. First, a layer of random noise is superimposed on an acquired noisy image. The image is used as training data to train a convolution autoencoder to converge and obtain a good noise separation mapping relationship. Second, the convolution kernel obtained in training is used in a denoising filter in a level-set framework to denoise the original image. Finally, by minimizing the energy function based on the level-set segmentation model, the segmentation contour is driven to move toward the target edge. Without relying on known noise distribution, the abovementioned approach is relatively robust for images acquired in real scenes.

2   Related Work

2.1   Gray-scale offset field

Wang and Pan[13] designed a correlation calculation method based on pixel levels, and the method could effectively avoid the influence of noise and other defilement information on the feature extraction process of a model. The basic idea is to calculate the correlation between the value of a pixel and the rest of the pixels in its field. If the correlation is too low, the pixel is different from the rest of the pixels in the field, and the weight of the point is reduced when calculating the energy function. Thus, the model can drive contour evolution on the basis of effective local information. Moreover, the model can effectively adapt to images with gray unevenness.

Li et al.[11] proposed a method for converting an image into digital information and then formalizing such information into a formula to solve the segmentation problem of images with gray unevenness. They believed that actual processed images suffer from noise or defilement by the offset field and thus proposed to model images at the pixel level, as shown in Eq. (1):

$$I(y) = b(y)J(y) + n \qquad (1)$$

where $I(y)$ is the actual processed noisy image, $b(y)$ is the gray value of the offset field at a certain pixel point, and $J(y)$ is the actual gray value of this point. The noise at this point is denoted by n, and the noise is assumed to be additive noise. Although Li et al.[11] proposed modeling, they also emphasized the following restrictions on $b(y)$ and $J(y)$:

(1) In a given image, a change in the overall offset field $b(y)$ needs to follow smooth and slow conditions.

(2) The image is divided into the foreground and background according to the gray value, and the actual gray value $J(y)$ of each pixel in each region should be regarded as a constant.

(3) The probability distribution of noise n cannot interfere with the calculation of the average gray value in a certain area, i.e., the mean value of the n distribution is 0.

In Li et al.'s model[11], the offset field information and noise information in the image are formally applied to the model. Such formalization enables the image segmentation model to use the uneven gray characteristics of the image to obtain its real composition. In this way, the model can extract the image with defacement information and quantify the uneven change of the image's gray level. This idea has been widely used in image segmentation models based on level-set methods by later researchers, and good results have been achieved. However, these models do not model noise despite their good results for noise separation. When the noise intensity is weak, the models obtain good noise separation results. However, they fail to achieve the same when the noise is strong. Moreover, these models are based on input images without noise or with known noise distribution. In sum, existing image segmentation models have many limitations.

2.2   Denoising auto-encoder

The grayscale offset field model solves the problem of an uneven grayscale, but it does not clearly express the distribution of noise. Hence, it does not deface information such as noise in the actual processed image. It instead separates noise from the real image, thereby increasing the difficulty of expressing the nonlinear relationship between complex noise and the actual defaced image being processed. The method in this chapter considers the use of related technologies based on deep learning to model noise. This method can obtain an accurate noise distribution so as to separate noise from real images.

Rumelhart et al.[14] proposed the concept of autoencoder to process high-dimensional complex data. As the earliest unsupervised neural network learning algorithm, the autoencoder demonstrates strong learning capabilities while promoting the development of neural networks. The self-encoder has a simple steel structure and only contains three layers of neural networks,

namely, the input layer, hidden layer, and output layer. The neurons between the three layers are all fully connected.

Figure 1 presents the schematic of the structure of the autoencoder. The autoencoder completes the mapping of the input image to the feature map through the first two layers. This process is similar to the encoding process in a fully convolutional neural network. The feature map is then completed through the next two layers. The mapping between output images is similar to the decoding process in a fully convolutional neural network. The final output result can be regarded as the prediction of the input by the network obtained by the autoencoder through learning. The square loss function is used as a constraint to achieve accurate prediction results. The form of the loss function is shown in Eq. (2):

$$L(x, y) = \frac{1}{n} \sum (x_i - y_i)^2 =$$
$$\frac{1}{n} \sum [x_i - g(f(x; \theta_1); \theta_2)]^2 \quad (2)$$

where x represents the input data and y represents the output prediction of the output autoencoder. $f(x, \theta_1)$ is the mapping relationship between the output layer and hidden layer. $g(f(x, \theta_1), \theta_2)$ is the mapping relationship between the hidden layer and the output layer, i.e., the formal expression of the output result y. θ_1 and θ_2 represent the parameters of the neural network, respectively.

The process of backpropagation involves keeping the prediction output y close to the input data x through the constraint of the loss function and then using the loss value to continuously adjust the parameters θ_1 and θ_2. The final result is $f(x, \theta_1)$ being able to effectively display the required feature information. At the same time, the key information of the input data can be retained without loss.
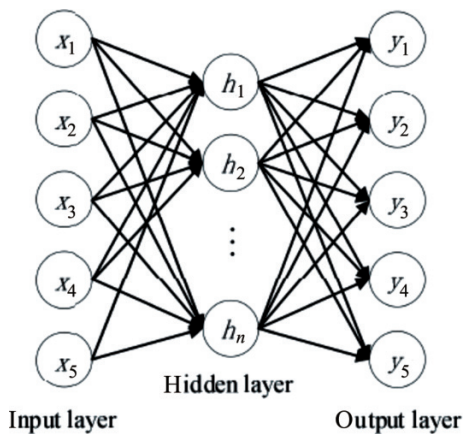
Considering the excellent unsupervised mapping learning ability of the autoencoder, Vincent et al.[15] put forward the idea of applying it to denoising by artificially adding noise to input images so as to train the autoencoder's noise adaptability. The basic idea is to train noisy input data so that the autoencoder can extract effective noise information. Then the reconstructed data output is obtained after the noise is removed through encoding and decoding. Finally, the autoencoder is guided by the difference between the output and input. The encoder updates the parameters to optimize the denoising performance of the model. The loss function still adopts the square difference loss function, and the form is shown in Eq. (3):

$$L(x, y) = \frac{1}{n} \sum (x_i - y_i)^2 =$$
$$\frac{1}{n} \sum [x_i - g(f(x_{i,\mathrm{noise}}; \theta_1); \theta_2)]^2 \quad (3)$$

where $x_{i,\mathrm{noise}}$ represents the result of adding noise to the input data x. $m = f(x_{\mathrm{noice}}; \theta_e)$ represents the mapping process of the autoencoder encoding the noise-defaced data. $y = g(m; \theta_d)$ represents the decoding process of the autoencoder on the encoding result m.

The proposed application of the autoencoder maximizes the unsupervised self-learning ability of neural networks and uses the fully connected method to extract noisy information in the data. However, when the input data are image data, the use of the fully connected method to train the autoencoder yields complex results because of the large number of parameters and low training efficiency. Lecun et al.[16] proposed the concept of the convolutional neural networks by replacing all full connections with convolutional operations on the basis of convolution kernels and successfully introducing neural networks into the field of computer vision; the study ultimately achieved excellent results. Herein, we consider replacing the convolution operation with a fully connected operation in the original autoencoder and constructing a new convolution autoencoder for image data. The new autoencoder is expected to effectively extract image features and preserve image pixel spatial information. It should also improve the model's ability to adapt to noise and other defaced information, reduce the number of parameters, and improve the efficiency of segmentation.

The structure of the convolutional autoencoder is shown in Fig. 2. The proposed method is considered for use in automatic noise separation. The basic principle
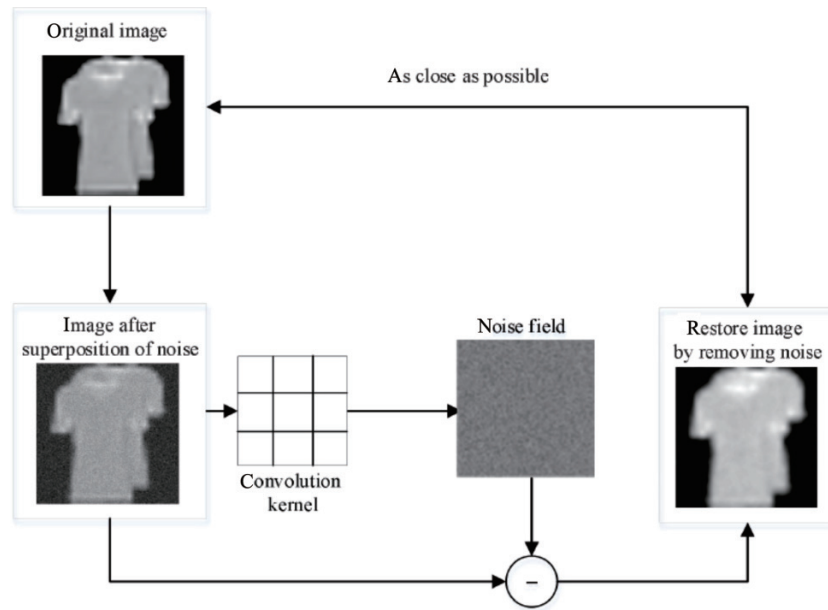

Fig. 1   Diagram of the structure of an autoencoder.

Fig. 2　Structure of a convolutional denoising autoencoder.

is to train the convolutional autoencoder by using the original image and the image after the addition of noise to obtain a mapping relationship that can separate noise. Such mapping relationship is then used to remove the noise in the original image.

3　Method

For the real image captured by camera and other sensors, it is difficult to distinguish the real noise from the target. We observe that the target intensity is usually stronger than the noise intensity, that is to say, the observed expected (real) noise n_1 is usually far less than the target x in the potential image, so it is difficult to get the noise distribution directly through observation. If we train a network that is specific to map the new superimposed noise (n_2) image z and the real image as the original noisy image y, the network has the ability to remove noise n_2 from the new noisy image z, then if we use the trained network to denoise the original noisy image y, we can easily remove the observed noise n_1 and get a good original image without noise.

For real images captured by cameras and other sensors, distinguishing real noise from targets is difficult. We observe that target intensity x is usually stronger than noise intensity, i.e., the observed expected (real) noise n_1 is usually far less than the noise x in the potential image. Hence, the noise distribution is difficult to obtain directly through observation. If we train a network to map a new superimposed noise (n_2) image z and a real image as the original noisy image y, then the network

can remove noise n_2 from the new noisy image z. If we use the trained network to denoise the original noisy image y, then we can easily remove the observed noise n_1 and obtain a good original image without noise.

On the basis of the aforementioned convolutional autoencoder, this work designs an image segmentation method that automatically separates noise. Training data are obtained by artificially superimposing noise on the original image and are then used to train the autoencoder of a single convolutional layer together with the collected real image. In this way, the autoencoder can learn the distribution of noise in the noisy image and then compare it with that in the original noisy image to obtain a noise-free image. The parameters obtained in the process are applied to the calculation of image segmentation on the basis of the level-set method to separate the noise offset field while avoiding its impact on the efficiency of image segmentation.

The formal expression of the process is shown in Eq. (4):

$$y = x + n_1; \quad z = y + n_2 \tag{4}$$

where y is the original collected image that can be considered to be composed of the real target x and additive noise n_1, whose distribution is unknown. A layer of noise n_2 is artificially superimposed on y to obtain the required training data z. We train the convolutional autoencoder to obtain the mapping relationship from z to y. Then the convolutional autoencoder can be used to separate the noise n_1 in y.

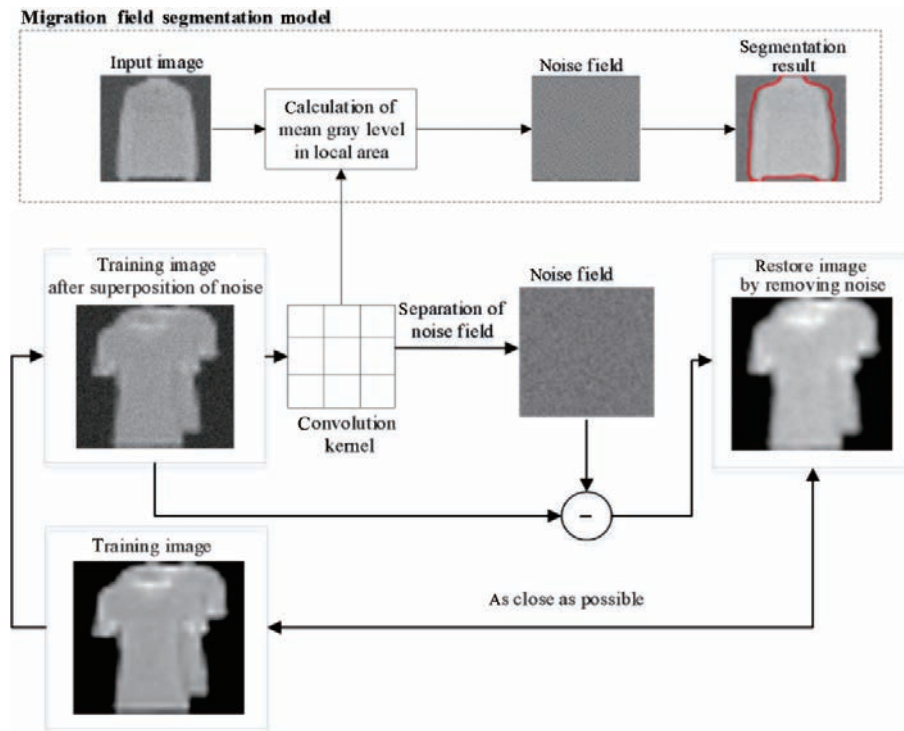Figure 3 shows the structure of the proposed

Fig. 3   Diagram of the proposed method.

segmentation model. By artificially superimposing Gaussian noise pollution information on the input image and then using the resulting image as the input image, we expect that the convolutional autoencoder outputs the image with the superimposed noise removed. During the training process, the square difference loss function is used to constrain the convolutional autoencoder to update the convolution kernel parameters.

In the offset field segmentation model based on the level-set method, the Gaussian kernel is used for feature extraction. Generally, manual prior knowledge is needed to set the weight of the convolution kernel[17]. In this work, we consider using the convolution kernel weight in the trained convolution autoencoder to calculate the local pixel neighborhood mean of the level-set method. This approach avoids not only the influence of noise and other pollution information on the calculation of the pixel neighborhood mean but also the limitation of the artificial knowledge level in setting the weight of the convolution kernel.

4   Experiment

To solve the noise sensitivity problem of the traditional level-set-based image segmentation method[18], this study designs a noise separation mechanism on the basis of the convolution autoencoder. The convolution

kernel weight obtained in the training process is used to calculate the local neighborhood mean in the traditional segmentation method so as to avoid errors in the manual design of convolution kernels and improve segmentation efficiency. Herein, several groups of comparative experiments are designed to describe the proposed method and verify its effectiveness in noise pollution information. The effectiveness of the proposed segmentation method is further confirmed by performing experiments on datasets of simple synthetic noisy images and complex synthetic noisy images and then quantifying the results.

4.1   Dataset and preprocessing

The proposed scheme uses the public image datasets MNIST, Fashion-MNIST, and STL-10 for the experiments. The MNIST and Fashion-MNIST datasets contain 60 000 training samples and 10 000 test samples. The target subjects are handwritten numerals and binary images of common clothing. The STL-10 dataset contains 100 000 training samples, including various objects with complex backgrounds. As no public noisy image dataset is currently available, noise is randomly superimposed on the aforementioned dataset in the experiment to obtain the original image required by the model.

4.2 Implement detail

The experiments are all completed in a Windows 10 (64 bit) system with Intel Core i7-9850h CPU@2.60 GHz. The GPU is RTX 2080, and the experimental environment is Python 3.7.4+Python 1.4.0. In the experiment, the convolution core trained by the automatic encoder is extracted by NumPy and imported into MATLAB for local segmentation on the basis of the level set. In this experiment, the learning rate of training the convolutional automatic encoder is 0.001, and the optimizer selects the SGD optimizer.

This study uses F1-score to evaluate the quantitative indicators because the image segmentation problem is the classification of each pixel, i.e., the pixel belongs to either the target domain or the background. Therefore, the classification results of each pixel are combined into a total of four types: true positive (TP), in which the pixel belongs to the target and is predicted to be the target; false negative (FN), in which the pixel belongs to the target and is predicted to be the background; false positive (FP), in which the pixel belongs to the background and is predicted to be the target; and true negative (TN), in which the pixel belongs to the background and is predicted to be the background.

The calculation method is shown in Eq. (5):

$$\text{F1-score} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FN} + \text{FP}} \qquad (5)$$

4.3 Denoising and segmentation result

This section first shows the experimental results of the level-set-based offset field segmentation model for noise-free images and noisy images to illustrate its sensitivity to noise. Before the experiment, a series of preprocessing should be performed to segment an image. The operations include re-standardizing the image to a uniform pixel size and artificially superimposing Gaussian noise with a mean value of 0 on the noise-free image during the comparison experiments.

Figure 4 shows the segmentation results of the noise-free image based on the offset field model of the level set. The first line is the noise-free image to be segmented, the second line is the segmentation result of the model, the third line is the original image's offset field separation result, and the fourth line is the image after removing the effect of the offset field. Through observation, we find that the offset field segmentation model based on the level set can accurately find the target edges for noise-free gray inhomogeneous images and can even identify the weak edges of some targets. By observing the separation results of the offset field, we find that the model can accurately separate the interference information in the image, effectively remove the interference information, and finally obtain a satisfactory segmentation result on the image with an uneven grayscale distribution.
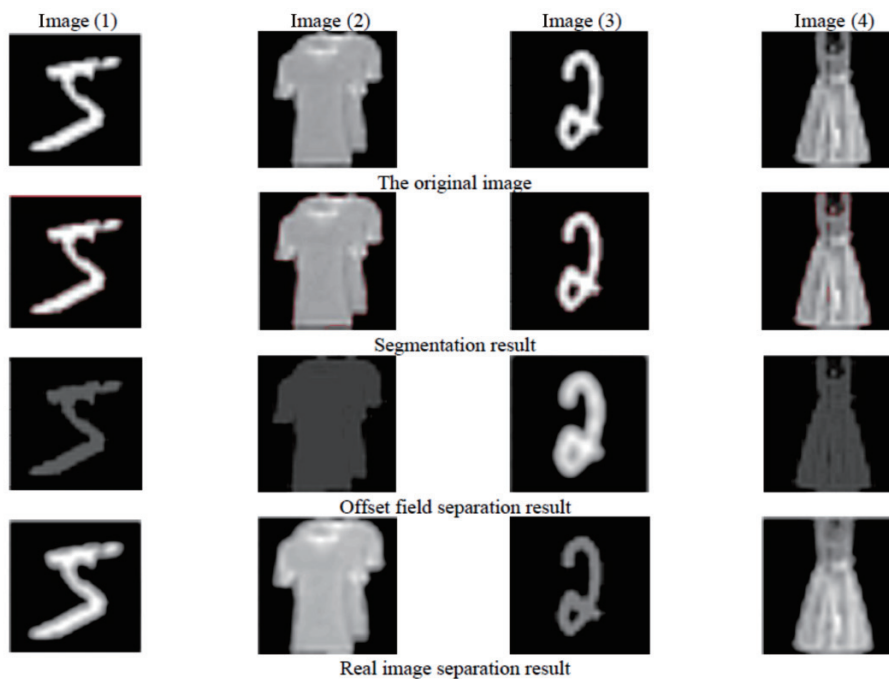

Fig. 4   Bias field in images without noise.

The offset field model based on the level set assumes that the mean value of the noise distribution in the image is 0, i.e., when the mean value of the noise distribution is 0, the influence on the calculation of the local neighborhood gray mean value of the pixel can be ignored. However, the experiments in this work reveal that the influence of noise cannot be ignored in the process of calculating the gray mean value of the local domain. We artificially add Gaussian noise with a mean value of 0 to the image in the above experiment. The experimental results are shown in Fig. 5. The first line is the defaced image superimposed with noise, the second line is the segmentation result of the model, the third line is the separation result of the offset field, and the fourth line is the separation result after removing the offset field. The experimental results show that the offset field model is very sensitive to noise and that it cannot, as the author expected, ignore the interference of noise to the calculation of the gray mean value of the local domain.

4.3.1　Simple image segmentation experiment

To demonstrate the effectiveness of the proposed method, we perform experiments on the simple background MNIST dataset and Fashion-MNIST dataset. Before the experiment, the first choice is to resize the dataset

and then superimpose different intensities of noise on the dataset, i.e., $(\mu = 0, \sigma^2 = 0.1), (\mu = 0, \sigma^2 = 0.2), (\mu = 0, \sigma^2 = 0.3), (\mu = 0, \sigma^2 = 0.4)$, and $(\mu = 0, \sigma^2 = 0.5)$. The convolution kernel weights in the trained convolutional autoencoder are extracted and used to segment the noisy image on the basis of the noise distribution derived by the offset field segmentation model of the level set. The proposed method can effectively extract the parameters of the noise distribution and model the noise image so as to achieve a good noise separation effect.

The experiment is conducted on the MNIST dataset with a simple background. The local binary fitting (LBF) model[19], the migration field model based on the level set, and the proposed model are used to divide the polluted images superimposed with noise of different intensities. An image is randomly selected as the segmentation result under each noise intensity. The segmentation performances of the models are quantified for an intuitive comparison.

The experimental results are shown in Fig. 6. The first line is the contaminated image after the superimposition of noise. The second line is the noise offset field separation result. The third line is the noise field separation image. The fourth line is the image after noise
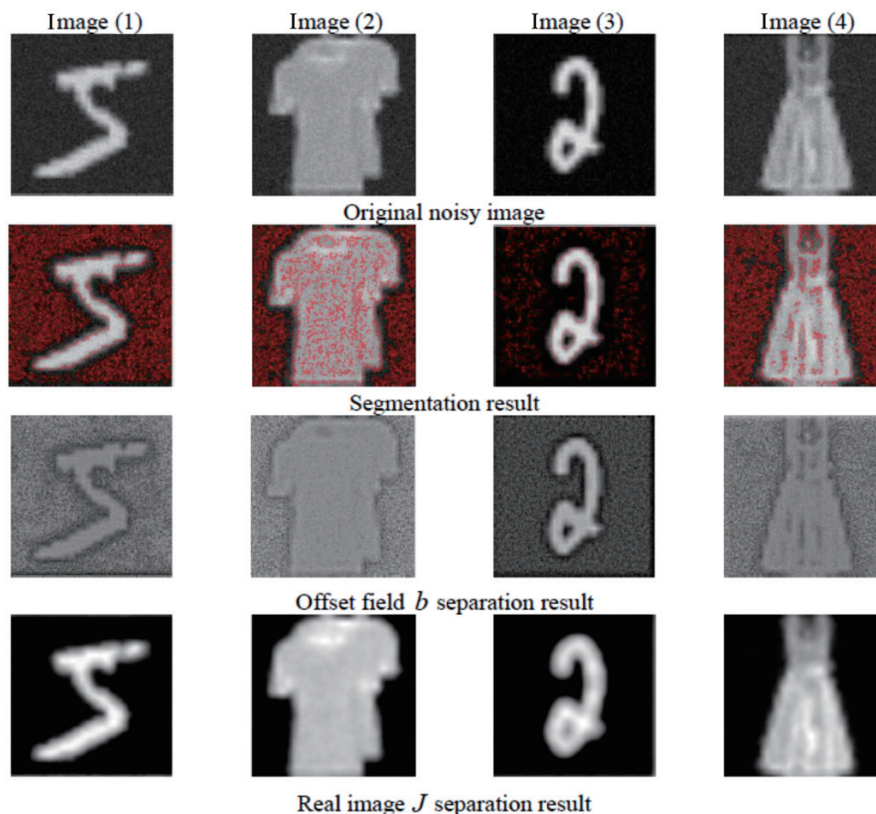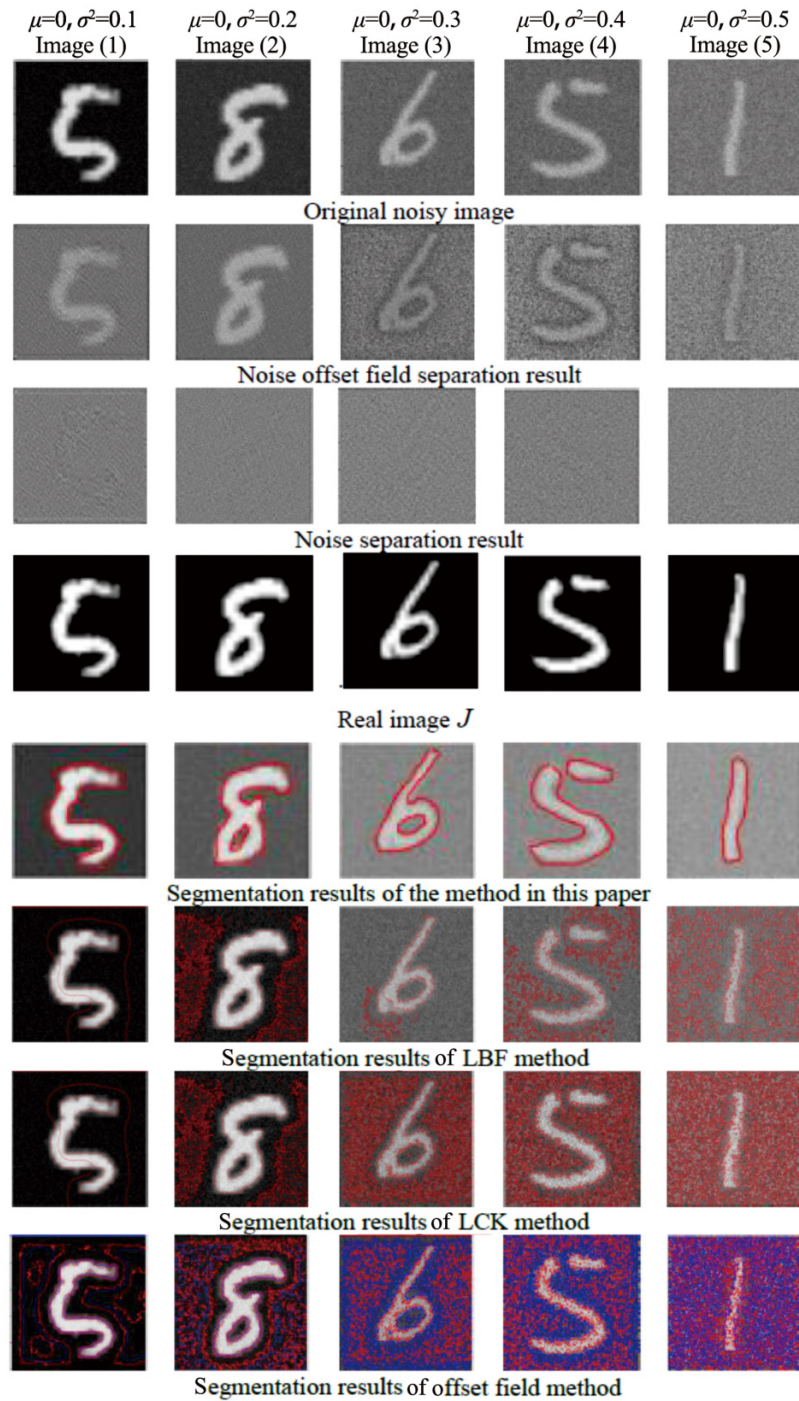

Fig. 5　Bias field in noisy images.

Fig. 6   Bias field in noisy image models.

separation. The fifth line is the segmentation result of the proposed method. The sixth line is the segmentation result of the LBF model. The seventh line is the segmentation of the LCK model. The eighth line is the offset field segmentation result. The experimental results show that when the superimposed noise intensity is low, obvious noise pollution can be observed. However, traditional segmentation methods cannot effectively segment this type of polluted image, which is affected by noise and thus leads to segmentation failure. The proposed method can effectively extract the noise field and obtain a noise-free image after operation. At the same time, the trained convolution kernel weights are applied to the level-set-based offset field segmentation model, and good segmentation results can be obtained.

To further prove the effective segmentation of the

proposed method for noise-stained images, we select the Fashion-MNIST dataset, whose target is more complex than that of the MNIST dataset. Similar to the MNIST dataset, the chosen dataset is preprocessed and artificially overlaid with noise. Then segmentation experiments are performed with different segmentations.

The experimental results are shown in Fig. 7. The first line is the defaced image after the introduction of artificial noise. The second line is the noise offset field separation result. The third line is the noise field separation image. The fourth line is the image after noise separation. The fifth line is the proposed model's segmentation result. The sixth line is the segmentation of the LBF model. The seventh line is the segmentation result of the LCK model. The eighth line is the offset field segmentation result. The experimental results

Fig. 7　Segmentation results of the compared models with multivariance Gaussian noise.

intuitively indicate that although the LBF model selects local pixel information as its segmentation benchmark, it is too sensitive to noise. Regardless of the intensity of the superimposed noise, the model cannot effectively achieve segmentation and obtain good segmentation results. The LCK model based on the correlation calculation in the local pixel domain cannot effectively separate noise information pixels from noncontaminated pixels when noise information exists. Hence, the model cannot obtain accurate segmentation results. The noise field separation mechanism based on the proposed convolutional automatic encoder can effectively extract the probability parameters of the noise distribution and reconstruct the noise field by using the modified parameters. Therefore, it can effectively separate the influence of noise contamination information on the segmentation process and obtain accurate segmentation results.

The segmentation results are quantified to obtain an intuitive understanding. The F1-score is used to evaluate the segmentation results, and it is visualized as an error bar graph in Fig. 8. The marked points shown in Fig. 8 represent the F1-score of the image showing the segmentation result. The upper and lower short lines represent the highest and lowest scores of the segmentation result of a certain model under noise intensity, respectively. Figure 8 shows that regardless of the intensity of superimposed noise, all the models, except the proposed one, cannot effectively divide the noisy image. Noise obviously exerts a great influence on these models. Although the segmentation accuracy of the proposed model declines when the noise intensity
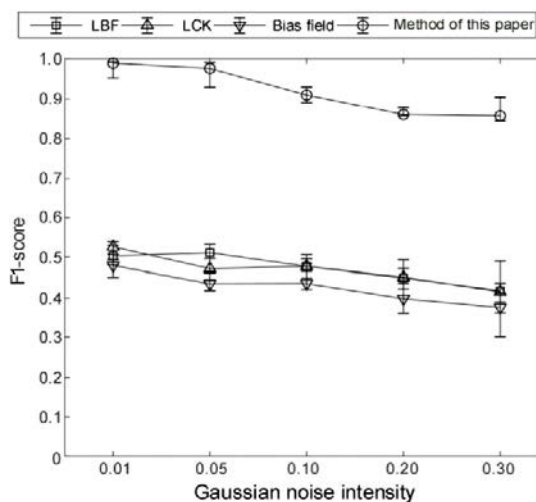

Fig. 8   Metrics of segmentation results on a synthetic image with Gaussian noise.

is strong, its score is still higher than that of the other segmentation models. When the noise intensity is low, the proposed method can obtain good segmentation results. Although some details of the target in the original image, such as its blurred edges, are affected when noise is separated, the information is not enough to affect the final segmentation results. Hence, the adaptability to noise of the proposed method is stronger than that of the other models.

4.3.2   Experiment of complex image segmentation

As mentioned previously, experiments are conducted on the MNIST and Fashion-MNIST datasets with relatively simple targets to illustrate the effectiveness of the proposed method in the segmentation of noise-damaged images. In this section, the STL-10 dataset, a natural image dataset with relatively complex targets, is selected for further experiments to further illustrate the effectiveness of the proposed method for noise-damaged images. Similarly, the dataset needs to be preprocessed by resetting the size and manually adding Gaussian noise. The noise distribution is as follows: ($\mu = 0$, $\sigma^2 = 0.1$), ($\mu = 0$, $\sigma^2 = 0.2$), ($\mu = 0$, $\sigma^2 = 0.3$), ($\mu = 0$, $\sigma^2 = 0.4$), and ($\mu = 0$, $\sigma^2 = 0.5$). The other noise separation mechanisms are the same as those in the previous experiments. The convolution autoencoder is used to fit the parameters of the noise distribution so as to reconstruct the noise image and achieve the purpose of removing noise.

The dataset of simple images selected in the previous section does not contain backgrounds, and the target is relatively simple. In segmentation, attention should be mainly paid to the gray values of pixels in the target area. The STL-10 dataset contains complex backgrounds and target information. In segmentation, the focus should be not only on the feature extraction of the target but also on the complex and diverse background information. In addition to the artificially superimposed noise information, high requirements are proposed for the segmentation model in the extraction of target characteristics and the calculation of the mean gray value.

The results of image segmentation are shown in Fig. 9. The first line is the contaminated image after the superimposition of noise. The second line is the noise migration field separation results. The third line is the separation noise field image. The fourth line is the image after noise separation. The fifth line is the segmentation result of the proposed model. The sixth line
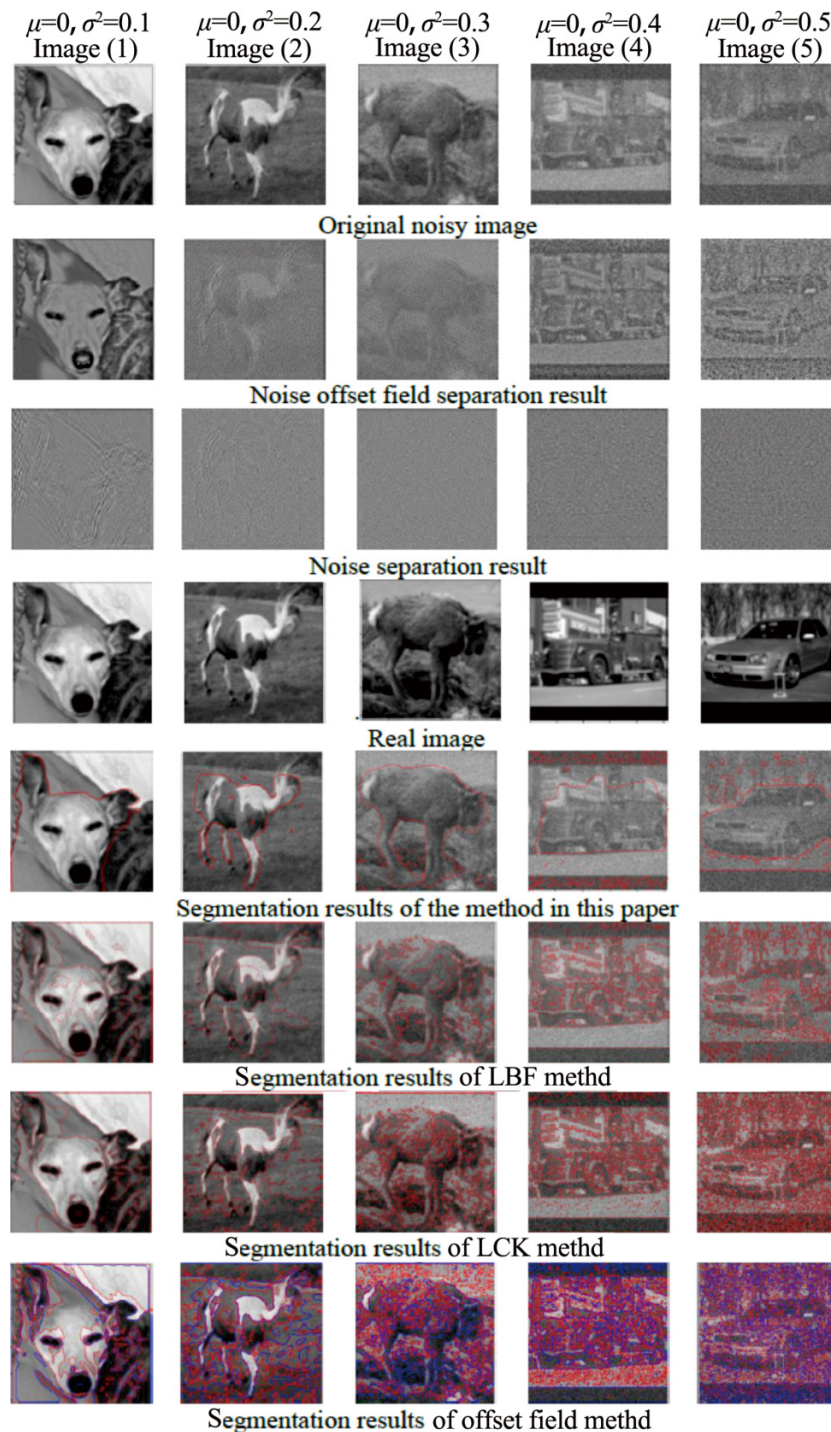
Fig. 9 Segmentation results of compared models with multivariance Gaussian noise.

is the segmentation result of the LBF model. The seventh line is the segmentation result of the LCK model. The eighth line is the segmentation result of the migration field model. The experimental results show that when the superimposed noise intensity is weak, the segmentation models based on the level set, such as the LBF and LCK models, can also roughly describe the contour of the target. However, the overall image segmentation results indicate the existence of false segmentation, i.e., in the case of complex backgrounds, these models are interfered by noise and cannot easily separate backgrounds and targets. When the noise intensity is too high, the segmentation results of the proposed method also show unsatisfactory regions. Nevertheless, the segmentation results obtained by the proposed method are excellent. When the noise intensity is low, the

proposed method can effectively separate the noise field and finally obtain good segmentation results.

This experiment still uses the F1-score to evaluate the segmentation results of the model and visualizes the results as an error bar graph shown in Fig. 10. Each marker in the image represents the F1-score of the image in the experiment. The upper and lower short lines denote the highest and lowest scores of a certain model under intense noise, respectively. Figure 10 shows that although the segmentation accuracy of the proposed model decreases with the increase of noise intensity, it is still higher than those of the other models. Regardless of the noise intensity, the other models cannot obtain accurate segmentation results for complex images with noise.

Through the comparative experiments, the following conclusions can be drawn. The segmentation model based on the traditional level-set algorithm can obtain good segmentation results for simple images without noise and background. When the image contains noise, noisy pixels cannot be easily distinguished. Segmentation failure is caused by dots and normal pixels. When a complex image with noise is encountered, the segmentation ability of the model is further reduced. As for the proposed method, it can effectively separate the noise field under different noise intensities regardless of whether the image has a simple or complex background. Moreover, it can obtain excellent segmentation results despite its segmentation when the noise intensity is high. Although the proposed method experiences a reduction in accuracy, it can still obtain better segmentation results than other models. The experiments in this work fully
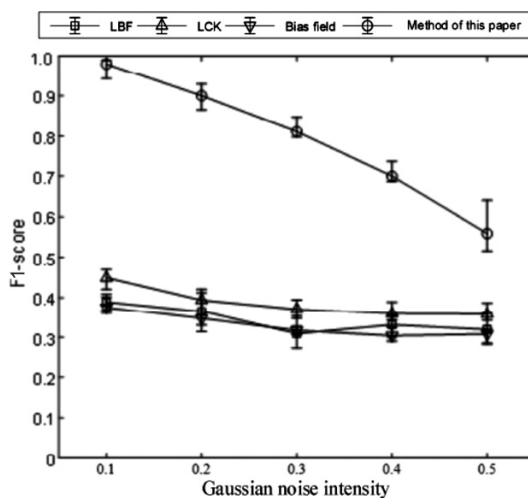

Fig. 10 Segmentation results of compared models with multivariance Gaussian noise.

illustrate the effectiveness of the proposed method in the segmentation of noisy images.

5 Conclusion

The traditional offset field model cannot effectively divide noisy images. Hence, this study proposes the use of a convolution autoencoder to fit the noise field distribution in noisy images. As the noise in collected data cannot be effectively obtained by observing its distribution characteristics, another layer of noise should be superimposed on the noisy image to train the convolution autoencoder in the simulation of the noisy image and original image. The convolution autoencoder is used to remove the noise in the original image, and the weights of the convolution kernel trained in the convolution autoencoder are used in the offset field model based on the level set. This approach improves the offset field model. While the noise detracts from the adaptability of the image, it also avoids the error caused by the design of the Gaussian kernel because of artificial prior knowledge. Moreover, the learning ability of the convolution autoencoder is integrated into the level-set framework to further improve the learning ability and adaptability of the proposed model to different intensities of noise. Through multiple sets of comparative experiments, we prove that the proposed method can effectively adapt to noise distribution of different intensities and effectively extract the parameters of the noise distribution. The results of this study can aid the model segmentation of targets contaminated with noise.

References

[1] Y. Wang, G. Yan, H. Zhu, S. Buch, and Z. Zhong. VC-Net: Deep volume-composition networks for segmentation and visualization of highly sparse and noisy image data, *IEEE Transactions on Visualization and Computer Graphics,* vol. 99, p. 1, 2020.

[2] S. J. Niu, Q. Chen, D. S. Luis, and L. R. Daniel, Robust noise region-based active contour model via local similarity factor for image segmentation, *Pattern Recognition*: *The Journal of the Pattern Recognition Society*, vol. 61, pp. 104–119, 2017.

[3] N. Tag, F. Zhou, Z. Gu, H. Zheng, Z. Yu, and B. Zheng, Unsupervised pixel-wise classification for chaetoceros

image segmentation, *Neurocomputing*, vol. 318, no. 27, pp. 261–270, 2018.

[4] L. Deng, M. Yang, Z. Liang, Y. He, and C. Wang, Fusing geometrical and visual information via superpoints for the semantic segmentation of 3D road scenes, *Tsinghua Science and Technology*, vol. 25, no. 4, pp. 498–507, 2020.

[5] L. Guo, L. Chen, C. L. P. Chen, and J. Zhou, Integrating guided filter into fuzzy clustering for noisy image segmentation, *Digital Signal Processing*, vol. 83, pp. 235–248, 2018.

[6] G. Li, H. F. Li, and L. Zhang, Novel model using kernel function and local intensity information for noise image segmentation, *Tsinghua Science and Technology*, vol. 23, no. 3, pp. 303–314, 2018.

[7] J. Liu, M. Li, J. X. Wang, F. X. Wu, T. M. Liu, and P. Yi, A survey of MRI-based brain tumor segmentation methods, *Tsinghua Science and Technology*, vol. 19, no. 6, pp. 578–595, 2014.

[8] X. F. Wang, H. Min, L. Zou, and Y. G. Zhang, A novel level set method for image segmentation by incorporating local statistical analysis and global similarity measurement, *Pattern Recognition*, vol. 48, no. 1, pp. 189–204, 2015.

[9] L. Zhang, X. G. Peng, G. Li, and H. F. Li, A novel active contour model for image segmentation using local and global region-based information, *Machine Vision and Applications*, vol. 28. no. 1, pp. 75–89, 2017.

[10] T. Ivanovska, R. Laqua, L. Wang, A. Schenk, and J. H. Yoon, An efficient level set method for simultaneous intensity inhomogeneity correction and segmentation of MR images, *Computerized Medical Imaging and Graphics*, vol. 48, pp. 9–20, 2016.

[11] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI, *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2007–2016, 2011.

[12] J. Xu, Y. Huang, L. Liu, F. Zhu, and L. Shao, Noisy-as-clean: Learning unsupervised denoisingfrom the corrupted image, arXiv preprint arXiv: 1906. 06878, 2020.

[13] L. Wang and C. Pan, Robust level set image segmentation via a local correntropy-based K-means clustering, *Pattern Recognition*, vol. 47, no. 5, pp.1917–1925, 2013.

[14] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[15] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P. A. Manzagol, Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, *Journal of Machine Learning Research*, no. 11, pp. 3371–3408, 2010.

[16] Y. Lecun, L. Bottou, and Y. Bengio, Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324,1998.

[17] X. Xie, C. Wang, A. Zhang, and X. Meng, A robust level set method based on local statistical information for noisy image segmentation, *Optik-International Journal for Light and Electron Optics*, vol.125, no. 9, pp. 2199–2204, 2014.

[18] X. Chen, Y. Wang, and X. Wu, Local image intensity fitting model combining global image information, *Computer Application*, vol. 38, no.12, pp. 3574–3579, 2018.

[19] C. Li, C. Y. Kao, J. C. Gore, and Z. Ding, Implicit active contours driven by local binary fitting energy, presented at IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 2007.

Ling Zhang is currently a lecturer at Taiyuan University of Technology, China. She received the PhD degree from Taiyuan University of Technology in 2016. She is a senior member of China Computer Federation (CCF). Her research interests include image processing and machine learning.
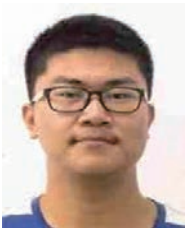
Jianchao Liu is pursuing the master degree at Taiyuan University of Technology, China. His research interests include image processing and computer vision.

Fangxin Shang is pursuing the master degree at Taiyuan University of Technology, China. His research interests include deep learning and image processing.

Gang Li is an associate professor at Taiyuan University of Technology, China. He received the PhD degree from Taiyuan University of Technology in 2018. He is a senior member of China Computer Federation (CCF). His research interests include image processing and visual information processing.

Juming Zhao received the PhD degree from Taiyuan University of Technology in 2015. She is a professor of Information and Computer Science at Taiyuan University of Technology. Her main research interests include information physics system and weak singal processing.

Yueqin Zhang is currently a senior engineer at Shanxi Tizones Technology Co., Ltd, China. She received the master degree from Beijing University of Science and Technology in 1997. Her research interests include data mining and intelligent information processing.