An Efficient Detection Approach of Content Aware Image Resizing

Ming Lu[1, 2, *], Shaozhang Niu[1] and Zhenguang Gao[3]

Abstract: Content aware image resizing (CAIR) is an excellent technology used widely for image retarget. It can also be used to tamper with images and bring the trust crisis of image content to the public. Once an image is processed by CAIR, the correlation of local neighborhood pixels will be destructive. Although local binary patterns (LBP) can effectively describe the local texture, it however cannot describe the magnitude information of local neighborhood pixels and is also vulnerable to noise. Therefore, to deal with the detection of CAIR, a novel forensic method based on improved local ternary patterns (ILTP) feature and gradient energy feature (GEF) is proposed in this paper. Firstly, the adaptive threshold of the original local ternary patterns (LTP) operator is improved, and the ILTP operator is used to describe the change of correlation among local neighborhood pixels caused by CAIR. Secondly, the histogram features of ILTP and the gradient energy features are extracted from the candidate image for CAIR forgery detection. Then, the ILTP features and the gradient energy features are concatenated into the combined features, and the combined features are used to train classifier. Finally support vector machine (SVM) is exploited as a classifier to be trained and tested by the above features in order to distinguish whether an image is subjected to CAIR or not. The candidate images are extracted from uncompressed color image database (UCID), then the training and testing sets are created. The experimental results with many test images show that the proposed method can detect CAIR tampering effectively, and that its performance is improved compared with other methods. It can achieve a better performance than the state-of-the-art approaches.

[1] Beijing Key Lab of Intelligent Telecommunication Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing, 100876, China.

[2] School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan, 114051, China.

[3] Department of Computer Science, Framingham State University, Massachusetts, MA 01772, USA.

* Corresponding Author: Ming Lu. Email: luminghood@126.com.

1 Introduction

With the rapid development of mobile devices, the same image may be displayed on mobile devices of different sizes. In order to meet the diversified requirements of display devices, image resizing based on changing the aspect ratio of the image has become a common image processing application. As such, content aware image resizing (CAIR) [Avidan and Shamir (2007)] arose at a historic moment, which includes seam carving and seam insertion. When changing the aspect ratio of the image, the distortion of the content can be avoided keeping the display effect almost the same as that of the original image. However, CAIR can preserve or even enlarge the semantically important image content, and the semantically insignificant image content can be reduced or even deleted. As such, tampering with an image content can be exploited [Avidan and Shamir (2007)] resulting in a crisis of trust among the public. As shown in Fig. 1, a little girl is running by the sea, and a pigeon is on the beach, any of those can be deliberately removed. The semantic content of the image seems to have changed. To avoid such case, it is critical to find an efficient method to detect the image processing by CAIR.
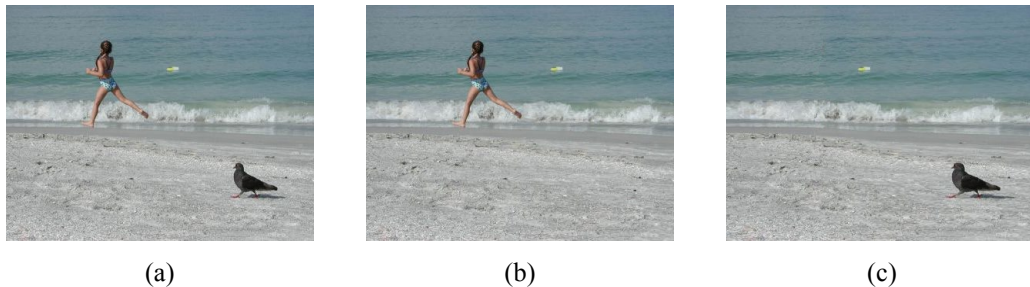

| (a) | (b) | (c) |

Figure 1: An example of content aware image resizing (CAIR) for object removal [Avidan and Shamir (2007)]. (a) the original image; (b) the image after pigeon removal; (c) the image after girl removal

In recent years, some research have been done on the detection of image tampering based on CAIR. The earliest passive forensic method for the detection of image seam carving was proposed by Sarkar et al. [Sarkar, Nataraj and Manjunath (2009)]. They extracted the 324-dimensional Markov features and input them into an SVM classifier system. However, the accuracy of detection was not high and was below 77.3%. Fillion et al. [Fillion and Sharma (2010)] proposed the multi-feature fusion method. The features included the bias of energy distribution, wavelet absolute moments, and the dispersal of seam behavior among others. The accuracy rate of detection was up to 91.3% for resized images with resizing ratio of 30%. For the detection of object removal by seam carving, the accuracy rate of detection was 76%, but the position of the removed object could not be identified accurately. Aiming at the detection of Joint Photographic Experts Group (JPEG) image seam carving, Liu et al. [Liu, Li, Cooper et al. (2012)] proposed a united feature extraction approach based on shift recompression. In 2014, Liu and his team designed an improved approach based on Calibrated Neighboring Joint Density (CNJD) to identify seam carved images. The proposed approach can simultaneously detect steganography and seam carving in JPEG images [Liu and Chen (2014)]. In 2016, a method, based on large-scale feature mining proposed by Liu [Liu (2016)] in dealing with

high-dimensional problems ensemble learning, was used to avoid over-fitting of traditional learning classifiers in detection. In 2017, in order to detect tampered JPEG images under compression attacks, a hybrid detection method based on large-scale feature mining was proposed [Liu (2017); Chang, Shih and Hsu (2013)] proposed another detecting approach of JPEG image seam carving, they exploited the symmetry property of the blocking artifact characteristics matrix (BACM) to distinguish whether an image was subjected to seam carving or not. Ryu et al. [Ryu, Lee and Lee (2014)] developed a detection approach of seam carving by exploiting the noise level and energy bias features to unveil the traces of seam carving and seam insertion, and then a better detection result was acquired, the accuracy was 93.5% for the resizing ratio of 50% by seam carving, but there were some limitations on the detection of object removal for the candidate images. Wei et al. [Wei, Lin and Wu (2014)] proposed a patch analysis approach to detect images seam carving, the approach firstly divided the candidate images into 2×2 mini-patches, then constructed Markov features by connecting patches in sub-direction, vertical direction and diagonal direction, the accuracy of detection was as high as 95.8% for the resizing ratio of 50% by seam carving. In order to improve the patch analysis method [Wei, Lin and Wu (2014); Yin, Yang, Li, et al. (2015)] presented a method based on local binary patterns (LBP) to detect image seam carving. The energy, noise and seam features were extracted from the LBP domain for the candidate images, support vector machine (SVM) was exploited to distinguish whether the images were subjected to seam carving or not. It improved the accuracy of 3.5% on average for the image with different resizing ratios. Sheng et al. [Sheng, Li, Su et al. (2017)] proposed a detection method based on Benford's law for content-aware image resizing, the average accuracy was 89.3% for the image with different resizing ratios. Ye et al. [Ye and Shi (2017)] developed a detection approach based on hybrid features, which included subtractive pixel adjacency model, Markov transition probabilities and local derivative pattern (LDP). Zhang et al. [Zhang, Yin, Yang et al. (2017)] proposed a blind detection method for image seam carving with low scaling ratio. In 2018, a detection approach based on uniform local binary pattern (ULBP) for seam carving was proposed [Zhang, Yang, Li et al. (2018)]. Compared with Ryu et al. [Ryu, Lee and Lee (2014); Wei, Lin and Wu (2014)], the approach improved the performance greatly.

Although the above methods have achieved certain results, the accuracy and performance are still not ideal, and the robustness are poor, which are still far from being a practical application. Therefore, this paper proposed a more accurate and efficient method to detect CAIR tampering. Firstly, the adaptability of threshold *t* for local ternary patterns (LTP) is improved to make LTP operator describe the correlation between a pixel and its neighborhood pixel more accurately. Then, the improved local ternary patterns (ILTP) features and the gradient energy features (GEF) are extracted from the candidate image and concatenated into the combined features for classifier training. Finally, two groups of experiments are designed to verify the important role of ILTP and GEF in the detection of CAIR tampering, to prove the effectiveness of the proposed method.

The main contributions of this paper are as follows:

(1) The LTP operator is improved to make the threshold *t* have strong adaptive ability, which can more accurately describe the change of correlation between local neighborhood pixels caused by CAIR, it significantly improves the effect of feature extraction.

(2) The GEF feature is proposed to detect the image tampered by CAIR, an effective approach for the detection of CAIR tampering based on ILTP and GEF has been proposed, which exploits the combined features to detect image tampering, and the detection accuracy is higher.

(3) In the experimental part of this paper, the performance comparison experiment between ILTP and conventional LTP for CAIR tampering detection is designed, another set of experiments is designed to compare the performance difference of the method exploits GEF or not.

The article has been organized in the following way: Section 1 introduces the research motivation, related work and the main contributions. The background of the problem is described in Section 2. Section 3 analyzes shortcomings of LBP and conventional LTP in the detection of CAIR tampering, and then the ILTP operator is proposed. The proposed detection framework for CAIR tampering is described in Section 4. In Section 5, a series of experiments are designed to verify the effectiveness of the proposed method. Finally, the work is concluded and the future work is introduced in Section 6.

2 Background

In 2007, Avidan et al. [Avidan and Shamir (2007)] proposed a content-aware image resizing technology, which adjusts the aspect ratio of images by preserving "important" pixels in the image, removing or inserting "unimportant" pixels. The importance of pixels is determined by the energy function as Eq. (1):

$$e_{(I)} = \left| \frac{\partial}{\partial x} I \right| + \left| \frac{\partial}{\partial y} I \right| \tag{1}$$

where *I* is a *m×n* image intensity matrix, *x* and *y* represent the row and column coordinates of a pixel respectively. In fact, it is to extract gradients in *x* and *y* directions according to Sobel operator, and then add their absolute values.

"Seam" is a connected path in the vertical or horizontal direction, which is composed of a series of pixels with lower energy value obtained from Eq. (1). The following discussion in this paper takes vertical seam as an example. For a *m×n* image, a vertical seam can be defined as Eq. (2):

$$S^x = \{S_i^x\}_{i=1}^m = \{i, x(i)\}_{i=1}^m, \ \text{s.t.} \forall i, |x(i) - x(i\text{-}1)| \leq 1. \tag{2}$$

where *x* represents the mapping from [1, 2, ..., *m*] to [1, 2, ..., *n*], *s* represents an eight-connected path passing through the whole image from top to bottom, and each line has only one pixel. The pixels in the seam *s* can be expressed as Eq. (3):

$$I_S = \{I(S_i)\}_{i=1}^m = \{I(i, x(i))\}_{i=1}^m \tag{3}$$

CAIR is to resize the image by continuously selecting the optimal seam and removing or adding pixels in the seam. The energy of pixels in the optimal seam should be as low as possible. According to Eq. (3), *E(s)* represents the cumulative energy, the optimal seam

can be defined as Eq. (4):

$$S^* = \min_s \{E(s)\} = \min_s \{\sum_{i=1}^{m} e(I(S_i))\} \tag{4}$$

the optimal seam can be selected by dynamic programming. For any pixel (i, j) in the seam, only three pixels $(i\text{-}1, j\text{-}1)$, $(i\text{-}1, j)$ and $(i\text{-}1, j+1)$ in the previous line are adjacent to the pixels (i, j). If $M(i, j)$ is used to represent the sum of the energy values from the first line to the pixel (i, j) in a seam, the seam with the lowest energy value shall meet the following Eq. (5):

$$M(i, j) = e(i, j) + \min(M(i-1, j-1), M(i-1, j), M(i-1, j+1)) \tag{5}$$

The selection of optimal seam is a process of search. According to Eq. (5), the search starts from the second line to the last line, the lowest energy pixel in the last line in a vertical seam is locked, and then the search backs up from this point line by line. The optimal seam can be obtained by finding the pixel position of this seam in other lines one by one. Similarly, the optimal seam in the horizontal direction also can be obtained in the same way.

CAIR includes seam carving and seam insertion. Seam carving can reduce the size of image by constantly deleting seams. Fig. 2 shows the image resizing effect realized by seam carving, in which (a) and (d) are the original images from uncompressed color image database (UCID) [Schaefer and Stich (2004)]; (b) shows a 20% horizontal reduction in (a); (c) is the selected optimal seams marker, and the red line in the figure indicates the position of seams; (e) shows a 20% vertical reduction in (d); (f) is the image shown in (d) with 20% horizontal seams in red. Although no new pixel will be added to the image in the whole process, the removal of seam will lead to the correlation change of neighborhood pixels, and whether a vertical or horizontal seam is removed, it will affect the correlation of neighborhood pixels in the vertical, horizontal and diagonal directions.

Seam insertion retarget the image by continuously inserting pixels into the optimal seam. For any pixel in the seam, it will be replaced by two new pixels. If the pixel in the seam is a_i, and the adjacent three pixels in the same line are $\{a_{i-1}, a_i, a_{i+1}\}$. According to Eq. (6), two new pixels b_i and b_{i+1} are used to replace a_i.

$$b_i = \text{round}\left(\frac{a_{i-1} + a_i}{2}\right), \quad b_{i+1} = \text{round}\left(\frac{a_i + a_{i+1}}{2}\right) \tag{6}$$

The above sequence of three adjacent pixels is changed into the sequence of four pixels $\{a_{i-1}, b_i, b_{i+1}, a_{i+1}\}$. If the pixel a_i happens to be at the edge of image, a_i should be reserved and only one pixel is inserted into the seam. If the inserted pixel is b_i, the sequence of two adjacent pixels $\{a_{i-1}, a_i\}$ becomes the sequence of three adjacent pixels $\{a_{i-1}, b_i, a_i\}$.

Seam insertion will introduce new pixels, which will inevitably cause the correlation change of neighborhood pixels around seam. Therefore, the correlation of neighborhood pixels in the eight-connected direction around seam can be used as a clue to detect whether the image has been processed by CAIR.
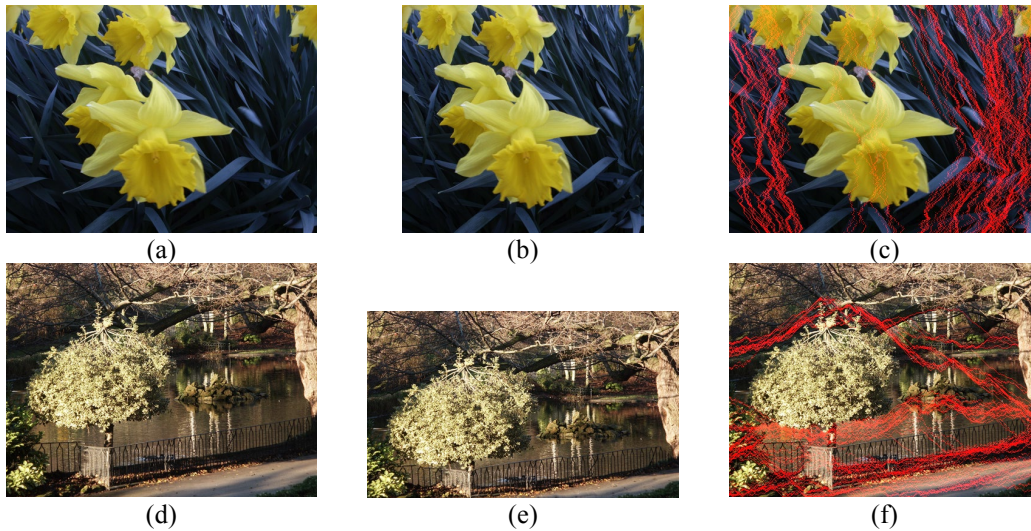
Figure 2: Image resizing effect implemented by seam-carving. (a) the original image from uncompressed color image database (UCID) (ucid00081.tif); (b) the image after 20% horizontal reduction; (c) the original image with 20% vertical seams in red; (d) the original image from UCID (ucid00037.tif); (e) the image after 20% vertical reduction; (f) the original image with 20% horizontal seams in red

3 The Improved Local Ternary Patterns (ILTP)

There are many methods to analyze the correlation between pixels. It can be analyzed by Markov feature, LBP feature and Gray-level Co-occurrence Matrix (GLCM) [Chen, Zhong and Bao (2019)]. Considering that LBP has a considerable advantage in describing the spatial continuity and correlation between pixels, but lacks the ability of magnitude description and is not resistant to noise, the ability to describe the correlation of neighborhood pixels is poor. In this paper, its upgrade pattern LTP was used, and further improved. We proposed the ILTP to describe the correlation of neighborhood pixels.

3.1 The application of LBP

LBP is an operator for extracting texture features of images proposed by Professor Ojala at the University of Oulu in Finland. LBP code is computed by comparing each pixel with its neighboring pixels and can effectively describe the local texture structure of the image. Because of its superior performance, it has been applied to various image analysis [Ojala, Pietikainen and Harwood (1996)].

In a 3×3 window, there are one central pixel I_c and eight neighboring pixel I_0, I_1, ... I_7, local texture Lt can be defined as Eq. (7):

$$Lt \sim (I_0 - I_c, I_1 - I_c, ..., I_7 - I_c) \tag{7}$$

the central pixel is taken as the threshold, and the eight neighborhood pixels are compared with it separately. If the neighborhood pixel value is greater than the central pixel value, the corresponding position is marked as 1; otherwise, it is marked as 0. The

process is binary processing, expressed as Eq. (8):

$$S(x) = \begin{cases} 1 , x \geq 0 \\ 0 , x < 0 \end{cases} \qquad (8)$$

the eight neighborhood pixels will get eight binary values, which are concatenated to get a binary number, starting from the position of the top-left pixel in a clockwise direction. The corresponding decimal value of the generated binary number is the LBP value of the central pixel. The LBP value can be expressed as Eq. (9):

$$LBP(x_c, y_c) = \sum_{k=0}^{7} S(I_k - I_c) 2^k \qquad (9)$$

generalized by Eq. (9), the LBP value of any pixel in an image can be calculated by Eq. (10):

$$LBP_{P,R}(x_c, y_c) = \sum_{k=0}^{P-1} S(I_k - I_c) 2^k. \qquad (10)$$

By comparing the central pixel and its neighboring pixels, LBP, to some extent, can describe the correlation of the local neighborhood pixels. When a vertical seam is removed, all pixels on the right side of the adjacent region need to be shifted one position to the left to fill the gap, which causes the eight connected neighborhood pixels to change, resulting in the change of LBP value. It can be detected that the image has undergone CAIR tampering. As shown in Fig. 3, the LBP values of seam and its adjacent pixels change significantly in a small region (red circle marking in Fig. 3(a)) through which the seam passes, and the larger the gradient, the more obvious it is. Fig. 3(b) shows the change of the small area marked by the red circle in Fig. 3(a) by seam carving, Fig. 3(c) indicates the change of the pixel value corresponding to the small area, Fig. 3(d) illustrates the change of LBP value of each pixel in the small area.
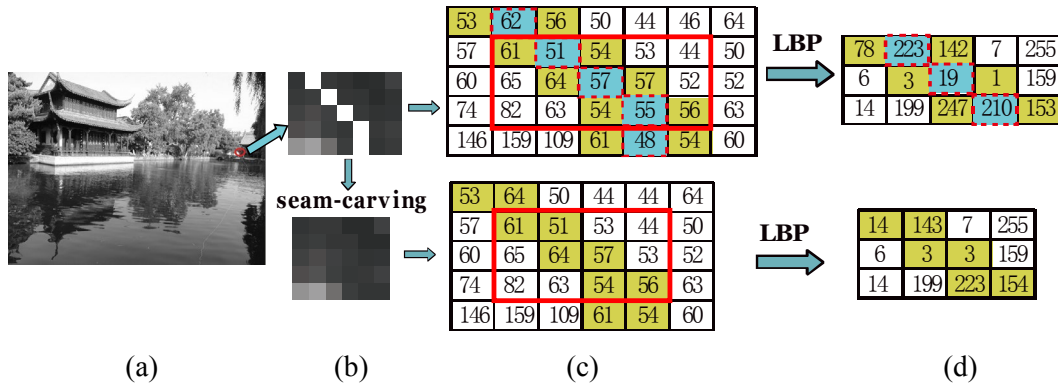
Figure 3: Change of local binary patterns (LBP) values caused by seam carving. (a) the original image; (b) a small region through which a seam passes; (c) pixels in the small region; (d) LBP value

Although LBP is simple to calculate and robust to single gray level change, by comparing the binary correlation between the central pixel and its neighborhood pixels in a certain region. However, there are some shortcomings for LBP. First of all, LBP is only a sign pattern rather than a magnitude pattern, which does not describe the degree of difference

between pixels and neighborhood pixels, and the magnitude information is also one of the important features to local neighborhood pixels; secondly, LBP has nothing to do with the intensity of gray change, which makes the extraction of texture features vulnerable to noise, resulting in the reduction of detection accuracy. Therefore, when the image contains noise, especially when the noise intensity is large, its visual feature performance will decline rapidly. In order to eliminate the influence of noise on detection and improve the ability of feature identification, this paper proposed the Improved Local Ternary Pattern (ILTP) operator to extract features for the detection of CAIR.

3.2 The improvement of LTP adaptability

LTP is another improved pattern of LBP descriptor [Tan and Triggs (2010)]. It can describe the gradient direction information of the image more specifically by ternary coding, and improve the robustness of the descriptor to noise partially, calculated by Eq. (11):

$$LTP_{P,R}(x_c, y_c) = \sum_{k=0}^{P-1} S(I_k - I_c)3^k, \ S(x) = \begin{cases} 1, & x \geq t \\ 0, & -t < x < t \\ -1, & x \leq -t \end{cases}. \tag{11}$$

It can be seen from Fig. 4 that LTP operator has certain robustness to noise, to some extent, it can distinguish the smooth region with noise from the texture region, while LBP operator cannot. Therefore, LTP operator can describe the correlation of local neighborhood pixels more accurately and eliminate the interference of noise.
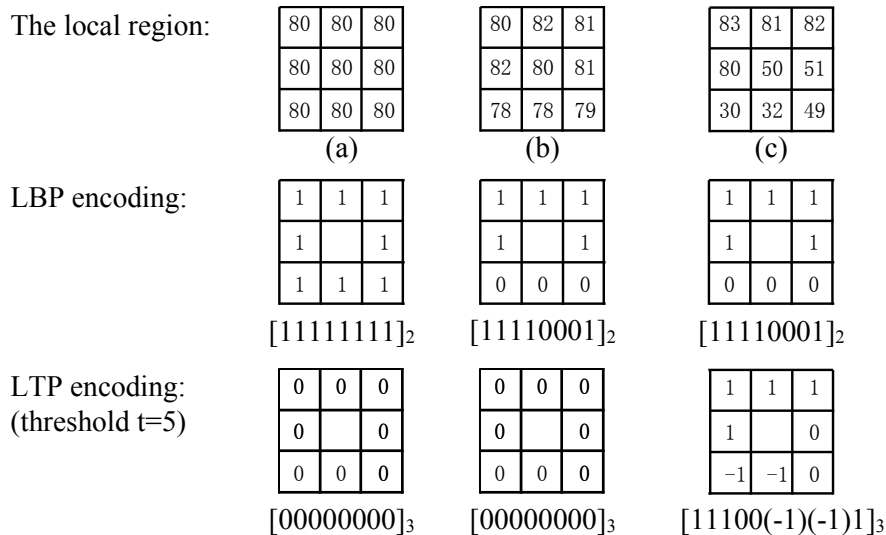
The local region:

80	80	80
80	80	80
80	80	80

(a)

80	82	81
82	80	81
78	78	79

(b)

83	81	82
80	50	51
30	32	49

(c)

LBP encoding:

1	1	1
1		1
1	1	1

$[11111111]_2$

1	1	1
1		1
0	0	0

$[11110001]_2$

1	1	1
1		1
0	0	0

$[11110001]_2$

LTP encoding:
(threshold t=5)

0	0	0
0		0
0	0	0

$[00000000]_3$

0	0	0
0		0
0	0	0

$[00000000]_3$

1	1	1
1		0
−1	−1	0

$[11100(-1)(-1)1]_3$

Figure 4: Comparison of LTP and LBP. (a) the smooth region; (b) the smooth region with noise; (c) the texture region

However, in the conventional LTP operator, the threshold value *t* is obtained by artificial estimation according to the image feature of a specific field, and the ability of anti-noise and magnitude description is relatively poor. Therefore, this paper proposed the adaptive threshold, which can better reflect the change degree of correlation between a pixel and its neighborhood pixels. The threshold value *t* is determined by the dispersion degree,

which will change with the different neighborhoods. Therefore, it is better to quantify the dispersion degree as a threshold. Although the dispersion degree can well reflect the distribution of intensity difference between the central pixel and the neighboring pixels, it cannot reflect the overall pixel intensity level in the local neighborhood, nor the intensity difference among the neighboring pixels, which are the key indicators related to the granularity of local texture quantization.

In order to solve the above problems, this paper proposed to combine the dispersion degree with the standard deviation which can reflect the overall intensity level and difference of pixels in the local neighborhood. The more adaptive threshold can be obtained by complementing the advantages of dispersion degree and standard deviation. The specific implementation adopts incremental step local sampling method, the steps are as follows:

(1) The average value of intensity differences between the center pixel and its neighborhood pixels is calculated incrementally, as Eq. (12):

$$\overline{\Delta g} = \left(\sum_{i=0}^{p-1} \Delta g_i \right) \Big/ p \tag{12}$$

where p is the number of neighborhood pixels with incremental sampling, and the incremental step is 2^i, Δg_i is the intensity difference between the center pixel and the ith neighborhood pixel, $\overline{\Delta g}$ is the average value of intensity differences between the center pixel and its neighborhood pixels.

(2) The texture fluctuation degree of the neighborhood pixels relative to the center pixel is calculated incrementally, as Eq. (13):

$$w = \left(\sum_{i=0}^{p-1} \left(\Delta g_i - \overline{\Delta g} \right)^2 \right) \Big/ p \tag{13}$$

(3) The dispersion degree of the center pixel and its neighborhood pixels is calculated as Eq. (14):

$$d = \sqrt{w} \tag{14}$$

(4) The mean value μ of local neighborhood pixel intensity is generated, according to Eq. (15):

$$\mu = \frac{1}{p} \left(\sum_{i=0}^{p-1} g_i + g_c \right) \tag{15}$$

(5) The standard deviation σ of local neighborhood pixel intensity is obtained, according to Eq. (16):

$$\sigma = \sqrt{\sum_{i=0}^{p-1} \left(g_i - \mu \right)^2 \Big/ p} \tag{16}$$

(6) The adaptive threshold t is obtained by the dispersion degree d and the standard deviation σ, as Eq. (17):

$$t = d + \frac{\overline{\Delta g}}{\mu} \bullet \sigma \tag{17}$$

For LTP operator, the key to the ability of magnitude description lies in the selection of

threshold value. The threshold value should neither be too small nor too large. The improper selection of threshold value directly affects the LTP operator's recognition of local texture changes, while the adaptive threshold can generate different values for different local textures by dynamically adjusting the threshold value. In this way, the problem of threshold selection in conventional LTP is solved and the change of local texture is recognized more accurately.
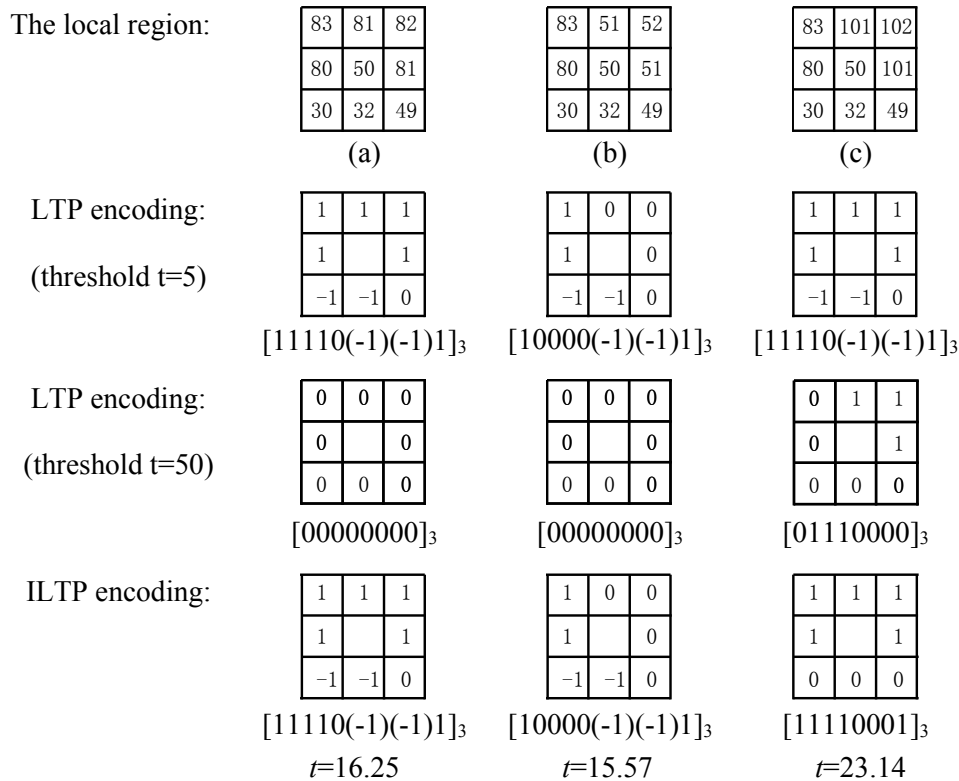
The local region:

83	81	82
80	50	81
30	32	49

(a)

83	51	52
80	50	51
30	32	49

(b)

83	101	102
80	50	101
30	32	49

(c)

LTP encoding:
(threshold t=5)

1	1	1
1		1
−1	−1	0

$[11110(-1)(-1)1]_3$

1	0	0
1		0
−1	−1	0

$[10000(-1)(-1)1]_3$

1	1	1
1		1
−1	−1	0

$[11110(-1)(-1)1]_3$

LTP encoding:
(threshold t=50)

0	0	0
0		0
0	0	0

$[00000000]_3$

0	0	0
0		0
0	0	0

$[00000000]_3$

0	1	1
0		1
0	0	0

$[01110000]_3$

ILTP encoding:

1	1	1
1		1
−1	−1	0

$[11110(-1)(-1)1]_3$
t=16.25

1	0	0
1		0
−1	−1	0

$[10000(-1)(-1)1]_3$
t=15.57

1	1	1
1		1
0	0	0

$[11110001]_3$
t=23.14

Figure 5: Comparison of the improved LTP and the conventional LTP. (a) the original region; (b) the change region with small dispersion degree; (c) the change region with large dispersion degree

As shown in Fig. 5, when the threshold value *t*=5, due to the small threshold value, although it is easy to recognize the change texture with small dispersion degree, it is unable to recognize the change region with large dispersion degree. On the contrary, when *t*=50, the threshold value is too large, and it is unable to recognize the texture change region with small dispersion degree, but it has strong recognition ability for the texture change region with large dispersion degree. The adaptive threshold solves the above contradiction, for different local neighborhood images, to generate different threshold values, so as to effectively identify different texture changes.

4 The proposed framework for the detection of CAIR

In order to effectively detect CAIR using the proposed method in this paper, a detection scheme for CAIR is designed. The framework of detection is mainly divided into two parts, one is the training of classifier, the other is the detection of the candidate images, as shown in Fig. 6. The detection of CAIR tampering is a binary classification problem. In this paper, support vector machine (SVM) [Cortes and Vapnik (1995)] is used as a classifier, it should be the supervised learning model of association learning algorithm, which is used to analyze data and recognize patterns. It is usually used for classification and regression analysis. When the number of training samples is fixed, SVM classifier can provide better classification performance than deep learning. At present, it is easy to operate and does better performance for a small number of samples. The training stage: after image preprocessing, LTP features and the gradient energy features are extracted respectively, then the above features are concatenated. Finally, the concatenated features are used to train and test classifier. The detection stage: the number of samples is select reasonably, the ratio of training samples to test samples reaches 3:1, which can make the detection effect better.
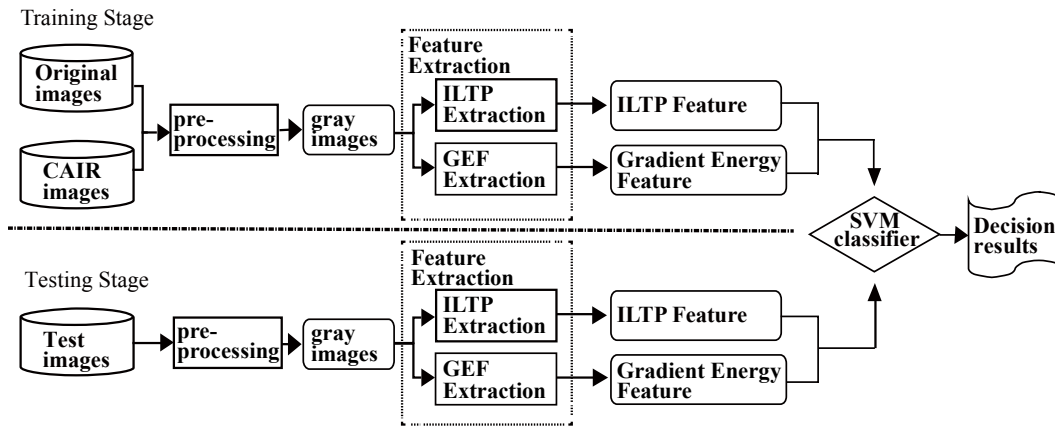

Figure 6: The proposed approach framework for CAIR detection

4.1 Image preprocessing

The proposed CAIR tamper detection method runs in gray-scale domain. Therefore, in the preprocessing stage, a color image is converted into a gray-scale one before further processing, the equation shown in Eq. (18) is used.

$$I=0.299R+0.587G+0.114B \tag{18}$$

4.2 Feature extraction

CAIR will inevitably lead to changes in the correlation between neighborhood pixels around seam, and feature extraction is to find the features that best reflect this correlation. The ILTP features can well describe the correlation between pixels, and can avoid the disturbance of noise. This paper combines the ILTP features and the gradient energy features to train the classifier to achieve better classification results.

4.2.1 ILTP feature extraction

The process of ILTP feature extraction is shown in Fig. 7, the image is divided into *m×n* blocks with the same size. The ILTP operator is used to extract the ILTP features of each block. We find the maximum and minimum values of all measured values, determine the range through them, count the number of times that the feature values fall in each bin, and form a statistical histogram. Finally, the statistical histograms of each block are integrated into the whole ILTP feature vectors. In order to simplify the calculation and reduce the dimension of feature vectors, ILTP features are decomposed into upper and lower ULBP code [Ojala, Pietikainen and Maenpaa (2002)], the upper pattern is described as histogram $H_{0,LTPU}$, and the lower pattern is described as histogram $H_{1,LTPL}$, obtain overall histogram $[H_{0,LTPU}\ H_{1,LTPL}]$. If each image is divided into 8×8 blocks, and the histogram block matrix as follows:

$$\begin{pmatrix} H_{0,1} & H_{0,2} & \cdots & H_{0,8^2} \\ H_{1,1} & H_{1,2} & \cdots & H_{1,8^2} \end{pmatrix} \tag{19}$$

The feature training set is constructed according to the block matrix of histogram.
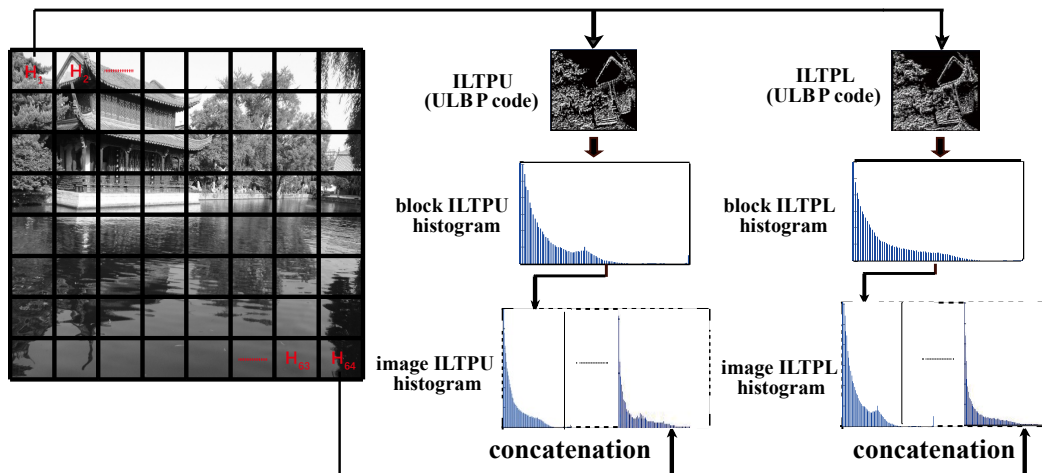

Figure 7: A schematic diagram of improved local ternary patterns (ILTP) feature extraction process

4.2.2 The extraction of gradient energy feature

As described in Section 2, as CAIR is to remove or insert low energy seams so as to retarget the image size, it is expected that the energy distribution of the image with CAIR is different from that of the original image. Therefore, the change of energy distribution can also be used as a clue to detect the image with CAIR. It is generally believed that the accumulation of local energy in the gradient direction is certain, and if it changes, it is considered that tampering has occurred. In this paper, the gradient energy feature is proposed to describe the change of local energy distribution in image block. The gradient energy feature is the energy accumulation in the direction of gradient in the local

neighborhood, which can be defined as Eq. (20):

$$GE\left(N_{p,r}, \theta\right) = M\left(i, j\right) \cdot \sum_{g \in N_{p,r}} \left(g_x \cos\theta + g_y \sin\theta\right)^2. \tag{20}$$

where g_x, g_y is the gradient value of pixel g (i, j) in x, y direction; $N_{p,r}$ is the selected local neighborhood, p is the number of pixels in the local neighborhood, r is the radius of the local neighborhood; M (i, j) is the sum of accumulated energy of pixel g (i, j) in the local neighborhood; GE $(N_{p,r}, \theta)$ is the gradient energy of pixel g in the local neighborhood, the range of the orientation θ is 0°-180°.

In addition to the above features, the key gradient energy features extracted in this paper are listed as follows Eqs. (21) to (25):

(1) The mean of gradient energy:

$$GE_{mean}\left(N_{p,r}, \theta\right) = \frac{1}{r^2\left(g_x^2 + g_y^2\right)} \sum_{g \in N_{p,r}} \left[M\left(i, j\right) \cdot \sum_{g \in N_{p,r}} \left(g_x \cos\theta + g_y \sin\theta\right)^2\right] \tag{21}$$

(2) The standard deviation of gradient energy:

$$GE_{std}\left(N_{p,r}, \theta\right) = \sqrt{\frac{1}{r^2\left(g_x^2 + g_y^2\right)} \sum_{g \in N_{p,r}} \left[GE\left(N_{p,r}, \theta\right) - GE_{mean}\left(N_{p,r}, \theta\right)\right]^2}. \tag{22}$$

(3) The max gradient energy:

$$GE_{max}\left(N_{p,r}, \theta\right) = \max_{g \in N_{p,r}} \left\{M(i, j) \cdot \sum_{g \in N_{p,r}} \left(g_x \cos\theta + g_y \sin\theta\right)^2\right\}. \tag{23}$$

(4) The min gradient energy:

$$GE_{min}\left(N_{p,r}, \theta\right) = \min_{g \in N_{p,r}} \left\{M(i, j) \cdot \sum_{g \in N_{p,r}} \left(g_x \cos\theta + g_y \sin\theta\right)^2\right\}. \tag{24}$$

(5) The difference of gradient energy features in the local neighborhood:

$$GE_{diff}\left(N_{p,r}, \theta\right) = GE_{max}\left(N_{p,r}, \theta\right) - GE_{min}\left(N_{p,r}, \theta\right). \tag{25}$$

The above five key gradient energy features need to be extracted from different directions. Due to the destroy of the energy distribution in the eight connected directions by CAIR, so typical directions include horizontal ($\theta=0°$), vertical ($\theta=90°$), main diagonal direction ($\theta=45°$), and sub diagonal direction ($\theta=135$).

4.2.3 Feature combination

We address the problem of combining the ILTP and the gradient energy feature. The candidate image is divided into $m \times n$ blocks with equal size and non-overlapping each other. According to the above method, ILTP features are extracted from each block, and the feature histogram is generated. Then, the candidate image is converted to ILTP image, ILTPU image and ILTPL image are generated. Further, the gradient energy

features are extracted from the above two images for each block, and the gradient energy feature histogram is obtained. Finally, the ILTP features and the gradient energy features in each block are concatenated into the combined feature for the classifier.

5 Experimental results and analysis

UCID (uncompressed color image database) [Schaefer and Stich (2004)], which is widely used in the industry, is adopted in the experiment. The image database contains 1338 uncompressed true color images, with two sizes of 512×384 and 384×512. In the experiment, 800 images are randomly selected as the original image training set, and the remaining 538 images are selected as the original image test set to verify the detection effect of seam carving and seam insertion respectively. There is no public image database for CAIR forensics, all the CAIR tampered images used in the experiment are created by the method of reference [Avidan and Shamir (2007)]. There are two parts in simulation experiment, one is the detection of image seam carving, and the other is the detection of image seam insertion. In this paper, LIBSVM [Chang and Lin (2011)] developed by Chih-Chung Chang and Chih-Jen Lin of National Taiwan University is used as classifier. The optimal parameters c and γ are obtained by 3-folds cross validation, the non-tampered image is labeled as "0", and the tampered image is labeled as "1" by 2-value classification. The hardware environment used in this experiment is Intel (R) core (TM) i5-7500 CPU 3.40 GHz, the operating system is Windows 7 (64 bit), and the software platform is matlab 2015a.

5.1 The verification of seam carving detection effect

In the detection of images with seam carving, according to the different resizing ratios in horizontal directions, 8 seam carved copies of each image are generated, namely 3%, 5%, 7%, 9%, 10%, 20%, 30% and 40%. In this way, 6400 tampered images are generated from 800 original training images, constitute the tampered image training set, then the original image training set and tamper image training set are used to train the classifier. In order to ensure the sufficiency of the test set, the above method is used to make a tamper image test set with 4304 (538×8=4304) pictures. In this way, with 538 original test images added, the number of samples in the image test set reaches 4842. In order to fully investigate the detection effect of the approach proposed in this paper on different resizing ratios, the experiment is divided into two groups: the detection of the seam carved image with low resizing ratios and the one with high resizing ratios, the resizing ratio below 10% can be considered as low resizing ratios. In the same group, the images of each resizing ratio are detected separately, and then all the images to be detected are mixed for testing, each group is tested three times, and the mean value is taken.

In order to further test the performance of the method proposed in this paper, the method proposed in reference [Zhang, Yang, Li et al. (2018)] represents the state-of-the-art technology, which has been implemented and tested for fair comparison. In addition, in order to investigate the respective contribution of the improved LTP feature and the gradient energy feature in CAIR tampered detection. The detection method based on conventional LTP features combined with gradient energy features (LTP+GEF) and the detection method based on only improved LTP features (ILTP) are respectively

implemented and compared with the method proposed in this paper. The detection effect comparison among the proposed method in this paper and the approach of reference [Zhang, Yang, Li et al. (2018)] and the above two methods implemented in this paper for auxiliary verification of the contribution of ILTP and GEF is shown in Tabs. 1 and 2. Tab. 1 shows the comparison of accuracy among the four detection methods for images with low resizing ratios. Obviously, the accuracy of the proposed approach is obviously higher than that of other approaches.

Table 1: Comparison of method accuracy (low resizing ratios %)

Resizing Ratios	The method [Zhang, Yang, Li et al. (2018)]	LTP+GEF (this paper)	ILTP (this paper)	Proposed method
3	64.28	66.06	70.13	75.54
5	75.12	76.83	78.86	80.79
7	78.32	80.51	82.66	86.75
9	80.87	82.82	86.82	90.27
MIX	79.39	81.66	85.01	89.67

Table 2: Comparison of method accuracy (high resizing ratios %)

Resizing Ratios	The method [Zhang, Yang, Li et al. (2018)]	LTP+GEF (this paper)	ILTP (this paper)	Proposed method
10	83.53	87.06	85.24	88.54
20	93.92	95.83	94.36	96.79
30	95.94	97.51	96.07	97.75
40	97.57	98.99	98.78	99.27
MIX	95.35	98.15	97.58	98.67

As can be seen from Tab. 2, the proposed approach can achieve the highest accuracy among the four detection approaches for images with high resizing ratios. The comparison of ROC (Receiver Operating Characteristic Curve) [Fawcett (2006)] curve is shown in Fig. 8.
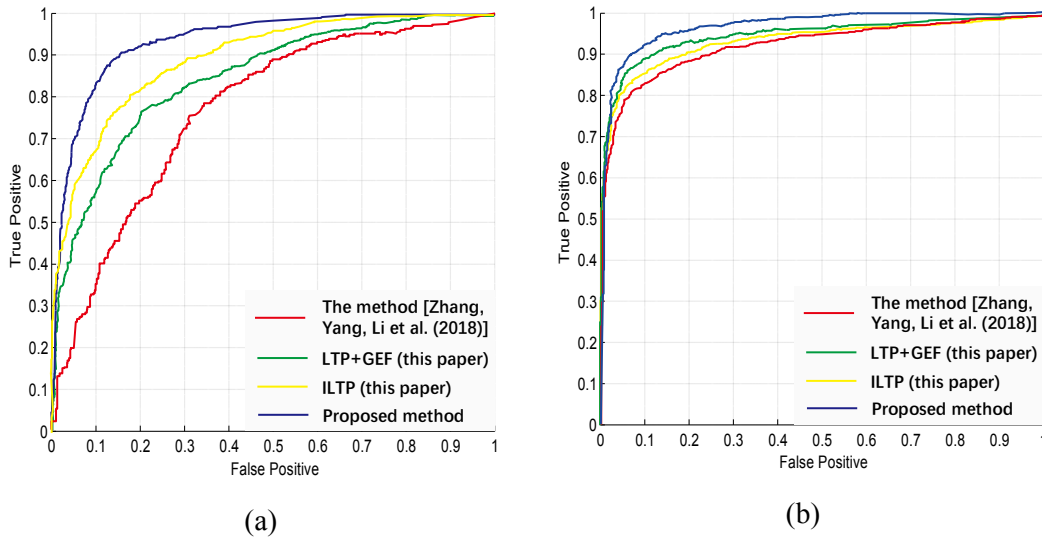
Figure 8: ROC curve comparison of different detection methods. (a) ROC curve comparison (resizing ratio 5%); (b) ROC curve comparison (resizing ratio 10%)

The test results show that the proposed approach can detect seam carved images well, the detection accuracy of the seam carved images with low resizing ratios is more than 75.54%, and that of the high resizing ratios is more than 88.54%. As shown in Tabs. 1 and 2, compared with the method [Zhang, Yang, Li et al. (2018)], the latter three methods use a local texture description operator with magnitude description ability and anti-noise to some extent, so the detection accuracy of the latter three methods is higher than the method [Zhang, Yang, Li et al. (2018)]. In the low resizing ratios test, the method (ILTP) is better than the method (LTP+GEF), while method (LTP+GEF) is better than method (ILTP) in high resizing ratios test. The main reason is that the improved LTP is used in the method (ILTP), and the threshold t with strong adaptive ability can describe the local texture more accurately, refine the granularity of texture description. Therefore, it can describe the small changes of the pixel correlation in the local neighborhood more accurately, which plays a greater advantage in the low resizing ratios test. With the increase of resizing ratio, a large number of original texture structures will be lost, which will interfere with ILTP feature extraction, and the accuracy will be improved slowly. Due to the gradient energy feature, the method (LTP+GEF) can accurately describe the change of pixel correlation in a large range of local neighborhoods, which is helpful for the high resizing ratios detection. But in the detection of low resizing ratios, the performance inferior to method (ILTP). The proposed method takes the advantages of the above two methods, the experimental results show that the performance of proposed method is much better than other methods. Fig. 8 shows the ROC curves [Fawcett (2006)] when the resizing ratio is 5% and 10% respectively. From the ROC curve in Fig. 8, the above argument can also be verified. As we all know, the larger AUC (Area Under Curve) [Fawcett (2006)] is, the better performance of the ROC model. The AUC of ROC curve for the proposed method is greatest in the four methods, which shows that the performance is superior

5.2 The verification of seam insertion detection effect

The selection and creation of tampered image training set and test set are similar to Section 5.1. However, the seam insertion method is used to enlarge the image horizontally. Considering the influence of different resizing ratios on the performance of the method, the experiment is still divided into two groups: low resizing ratio detection and high resizing ratio detection. In each group, the tampered image is tested separately for each resizing ratio, and then mixed test is carried out. As there is little research on seam insertion detection at present, reference in Ryu et al. [Ryu, Lee and Lee (2014)] is a relatively advanced method in the existing methods, so the proposed method in this paper is compared with it. At the same time, we also apply the methods (LTP+GEF) and ILTP mentioned in Section 5.1 to the detection of seam insertion, and give a comparison with the method proposed in this paper. The comparison of test results is shown in Tabs. 3 and 4.

Table 3: Comparison of method accuracy (low resizing rate %)

Resizing Ratios	The method [Ryu, Lee and Lee (2014)]	LTP+GEF (this paper)	ILTP (this paper)	Proposed method
3	97.16	97.23	98.56	98.79
5	97.35	97.59	98.71	98.82
7	97.35	97.67	98.78	98.86
9	97.36	97.99	98.83	98.98
MIX	97.35	97.86	98.55	98.65

Table 4: Comparison of method accuracy (high resizing rate %)

Resizing Ratios	The method [Ryu, Lee and Lee (2014)]	LTP+GEF (this paper)	ILTP (this paper)	Proposed method
10	97.76	97.83	98.24	98.54
20	97.76	97.81	98.38	98.79
30	97.76	97.80	98.57	98.75
40	97.65	97.78	98.79	98.27
MIX	97.68	97.70	98.66	98.67

Tabs. 3 and 4 show the performance of detection for the images seam insertion with low resizing ratios and high resizing ratios, it is apparent that the accuracy of the proposed method is higher than other methods, which proves that the proposed approach can detect image with seam insertion well. The method (ILTP) is always better than the method (LTP+GEF) regardless of low or high resizing ratios in the seam insertion detection. The main reason is that when seams are continuously inserted, a large number of relatively smooth regions will be generated. These relatively smooth regions need to be detected by fine-grained local texture description tools (such as the LTP descriptor with adaptive threshold). However, the gradient energy feature cannot accurately detect the change of pixel correlation in the local neighborhood, so there is no advantage of the method (LTP+GEF) in the detection of image seam insertion. In the Fig. 9, the ROC curves

indicate the excellent performance of the proposed method, when the resizing ratio is 5% and 10%. The ROC curve comparison of test results is shown in Fig. 9.
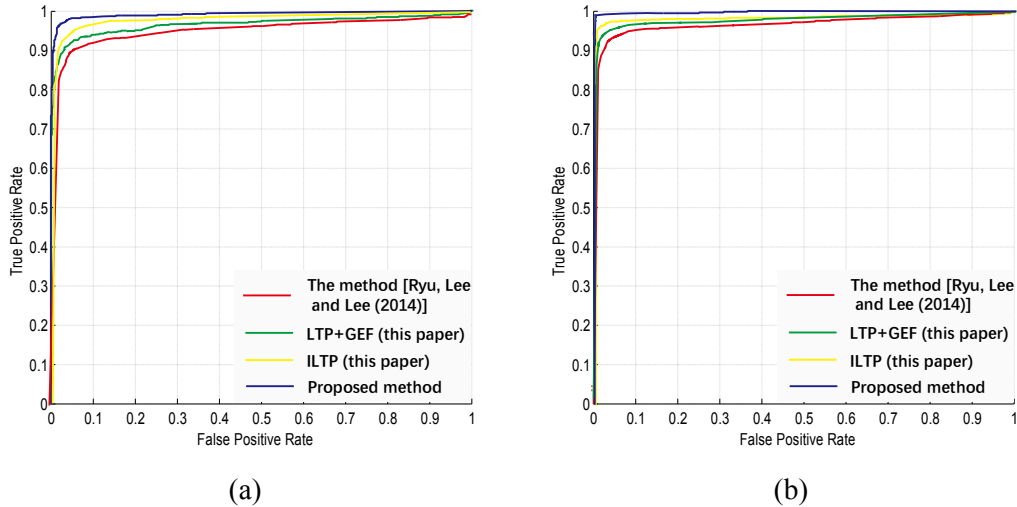

(a)   (b)

Figure 9: ROC curve comparison of different detection methods. (a) ROC curve comparison (resizing ratio 5%); (b) ROC curve comparison (resizing ratio 10%)

This experiment not only tests the accuracy of the detection results, but also investigate the method's ability to locate the tampered area. For example, Fig. 10 shows the localization effect of the tampered area in the image with 5%, 10% and 20% resizing ratios by seam insertion.
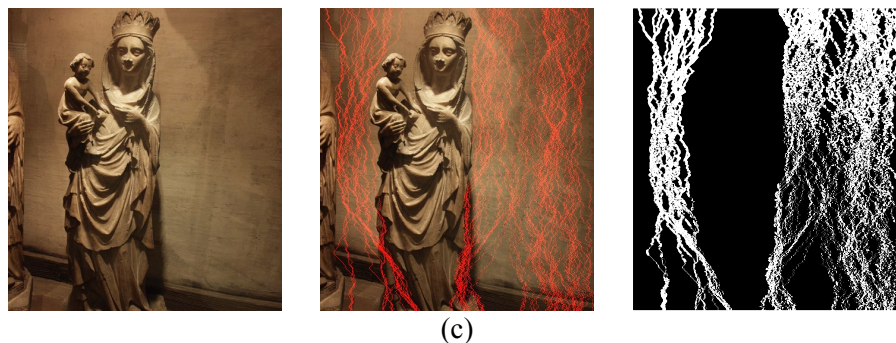

(a)

(b)

(c)
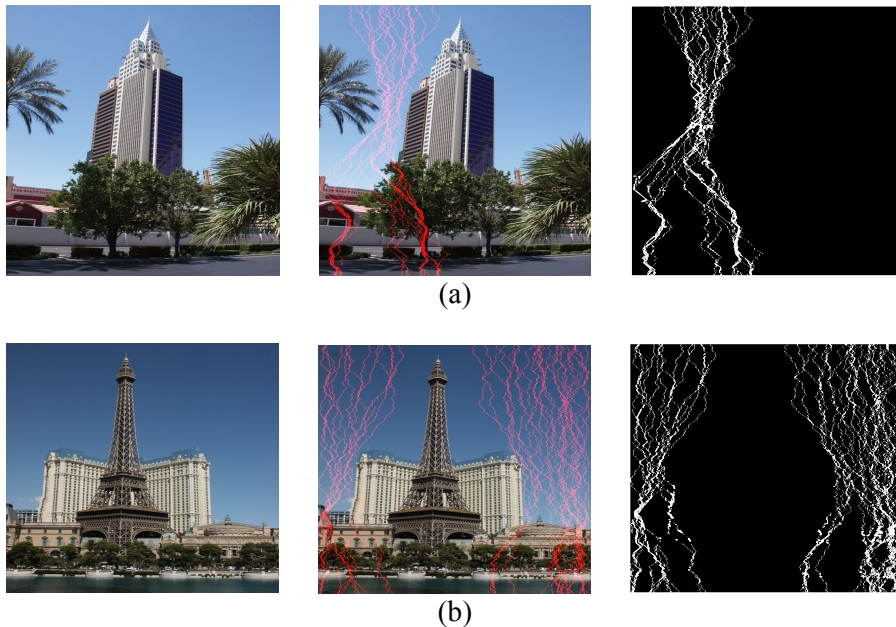
Figure 10: Display of the effect of tampered area location. (a) Detection and location results of 5% resizing ratio; (b) Detection and location results of 10% resizing ratio; (c) Detection and location results of 20% resizing ratio

6 Conclusion

In this paper, we proposed a detection method for CAIR tampering. Based on the change of pixels correlation in local neighborhood, we extracted ILTP features and gradient energy features, trained LIBSVM classifier with the combined features. The combined features can effectively reflect the essence of CAIR, make the method suitable for the detection of both seam carving and seam insertion. By setting the adaptive threshold t in the LTP operator, the anti-noise ability and magnitude recognition ability of LTP were enhanced. The features extracted by ILTP operator were more accurate for describing the correlation of pixels in the seam neighborhood, so the detection accuracy was improved, especially for the detection of low resizing ratios, it was significantly improved compared with other approaches. Through the comparison of experimental results, the accuracy of the proposed method is higher than that of other three methods for comparison, no matter for the detection of low resizing ratios or high resizing ratios. ILTP features and gradient energy features are complementary, and the lack of one will affect the accuracy of detection results. At present, the proposed method only can detect whether the image has been processed by CAIR, but the tampering of the semantic cannot be accurately judged. Steganalysis is a powerful tool for digital image forensics, a steganalysis algorithm for color image based on channel gradient correlation proposed in the reference Kang et al. [Kang, Liu, Yang et al. (2019)] has provided another way for us to solve the semantic tampering of image. In the future work, we will consider the combination of steganalysis technology to explore a better method to solve the forensics problem for the tampering with the semantic content of the image by CAIR.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Avidan, S.; Shamir, A. (2007): Seam carving for content-aware image resizing. *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 10-16.

Avidan, S.; Shamir, A. (2007): Seam carving for content-aware image resizing. http://www.faculty.idc.ac.il/arik/SCWeb/imret/index.html.

Chang, C. C.; Lin, C. J. (2011): LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems & Technology*, vol. 2, no. 3, pp. 2-27.

Chang, W. L.; Shih, T. K.; Hsu, H. H. (2013): Detection of seam carving in jpeg images. *International Joint Conference on Awareness Science and Technology & Ubi-Media Computing*, pp. 632-638.

Chen, X.; Zhong, H.; Bao, Z. (2019): A GLCM-feature-based approach for reversible image transformation. *Computers, Materials & Continua*, vol. 59, no. 1, pp. 239-255.

Cortes, C.; Vapnik, V. (1995): Support-vector networks. *Machine Learning*, vol. 20, no. 3, pp. 273-297.

Fawcett, T. (2006): An introduction to ROC analysis. *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874.

Fillion, C.; Sharma, G. (2010): Detecting content adaptive scaling of images for forensic applications. *Proceedings of SPIE, Media Forensics and Security II*, pp. 888-896.

Liu, Q.; Li, X.; Cooper, P. A.; Hu, X. (2012): Shift recompression-based feature mining for detecting content-aware scaled forgery in JPEG images. *Twelfth International Workshop on Multimedia Data Mining*, pp. 10-16.

Liu, Q.; Chen, Z. (2014): Improved approaches with calibrated neighboring joint density to steganalysis and seam-carved forgery detection in JPEG images. *ACM Transactions on Intelligent Systems and Technology*, vol. 5, no. 4, pp. 1-30.

Liu, Q. (2016): An approach to detecting JPEG down-recompression and seam carving forgery under recompression anti-forensics. *Pattern Recognition*, vol. 65, pp. 35-46.

Liu, Q. (2017): Exposing seam carving forgery under recompression attacks by hybrid large feature mining. *International Conference on Pattern Recognition, IEEE.*
https://doi.org/10.1109/ICPR.2016.7899773.

Ojala, T.; Pietikainen, M.; Harwood, D. (1996): A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, vol. 29, no. 1, pp. 51-59.

Ojala, T.; Pietikainen, M.; Maenpaa, T. (2002): Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 24, pp. 971-987.

Ryu, S. J.; Lee, H. Y.; Lee, H. K. (2014): Detecting trace of seam carving for forensic analysis. *IEICE Transactions on Information and Systems*, vol. E97D, no. 5, pp. 1304-1311.

Sarkar, A.; Nataraj, L.; Manjunath, B. S. (2009): Detection of seam carving and localization of seam insertions in digital images. *11th ACM Workshop on Multimedia and Security-MM & Sec'09*, pp. 107-116.

Schaefer, G.; Stich, M. (2004): UCID-an uncompressed color image database. *Proceedings of SPIE, Storage & Retrieval Methods & Applications for Multimedia*, vol. 5307, pp. 472-480.

Sheng, G.; Li, T.; Su, Q.; Chen, B.; Tang, Y. (2017): Detection of content-aware image resizing based on Benford's law. *Soft Computing*, vol. 21, pp. 5693-5701.

Tan, X.; Triggs, B. (2010): Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing,* vol. 19, no. 6, pp. 1635-1650.

Kang, Y.; Liu, F.; Yang, C.; Xiang, L.; Wang, P. (2019): Color image steganalysis based on channel gradient correlation. *International Journal of Distributed Sensor Networks*. https://doi:10.1177/1550147719852031.

Wei, J. D.; Lin, Y. J.; Wu, Y. J. (2014): A patch analysis method to detect seam carved images. *Pattern Recognition Letters*, vol. 36, pp. 100-106.

Ye, J.; Shi, Y. Q. (2017): A local derivative pattern-based image forensic framework for seam carving detection. *Digital Forensics and Watermarking*, pp. 172-184.

Yin, T.; Yang, G.; Li, L.; Zhang, D.; Sun, X. (2015): Detecting seam carving based image resizing using local binary patterns. *Computers & Security*, vol. 55, pp. 130-141.

Zhang, D.; Yin, T.; Yang, G.; Xia, M.; Li, L. et al. (2017): Detecting image seam carving with low scaling ratio using multi-scale spatial and spectral entropies. *Journal of Visual Communication and Image Representation*, vol. 48, pp. 281-291.

Zhang, D.; Yang, G.; Li, F.; Wang, J.; Sangaiah, A. K. (2018): Detecting seam carved images using uniform local binary patterns. *Multimedia Tools and Applications*. https://doi:10.1007/s11042-018-6470-y.