ITC 1/52

Information Technology
and Control
Vol. 52 / No. 1 / 2023
pp. 5-24
DOI 10.5755/j01.itc.52.1.31549

**Image Feature Fusion Method Based on Edge Detection**

Received 2022/06/02

Accepted after revision 2022/08/23

**https://doi.org/10.5755/j01.itc.52.1.31549**

**HOW TO CITE:** Li, F., Du, X., Zhang, L., Liu, A. (2023). Image Feature Fusion Method Based on Edge Detection. *Information Technology and Control,* 52(1), 5-24. https://doi.org/10.5755/j01.itc.52.1.31549

# Image Feature Fusion Method Based on Edge Detection

**Feng Li, Xuehui Du, Liu Zhang, and Aodi Liu**

Information Engineering University, Zhengzhou 450000, China

**Corresponding author:** xuehui_du@163.com

Deep learning-based image processing algorithms have developed rapidly in the past decade and have shown significant improvements to extract image features when both sufficient computing power and big data are accessible. Thus, rapid advances in applications such as facial recognition and autonomous driving have been one of the implementation areas. On the other hand, edges as a low-level prevalence feature in images with independent semantics are practically adapted to attain better outcomes. However, neural network-based image feature extraction focusing on texture rather than shape leads to insufficient accuracy. To address this issue, an edge feature extraction method utilizing both conventional operators such as HDE and Sobel and a deep learning-based method is proposed to classify and retrieve images with better accuracy outcomes. By doing so, a large amount of data needed to conduct deep learning-based methods is decreased, the transferability of the model is achieved, classification and retrieval accuracies are enhanced, and the data is compressed. All these better results are attained with benchmark data sets. As a result, all these are achieved by proposing a novel method.

KEYWORDS: Edge Detection, HED, Image Classification, Image Retrieval, Mean Shift, Otsu.

## 1. Introduction

When compared with structured data, unstructured data, for instance, images are more intuitive, informative, and user-friendly. However, the semantic information in images is highly abstract since both understanding and managing them with computers need more comprehensive methods. Typically, images are managed manually by giving tags based on the usual processing of structured data. With the development of the Internet and the Internet of Things, the amount of data generated has increased considerably, and manual markings have been impractical to keep up with. Thus, there has been compulsory action needed to conduct automatic identification and management of attributes for image contents. Thus, extracting image features is described as the first step in their representation of them. In conventional feature extraction methods, researchers typically utilize manual-based models to extract image features. Those are called the Harris corner detection operator [26], features from accelerated segment test (FAST) [31, 8] Laplacian of Gaussian (LoG) [11], and difference of Gaussian (DoG) [38]. In addition, other commonly utilized and similar methods are called scale-invariant feature transform (SIFT) [14, 44] histogram of oriented gradient (HOG) [41], and local binary pattern (LBP) [1]. Eventually, all methods serve in various image processing and comprehension tasks and have achieved remarkable results.

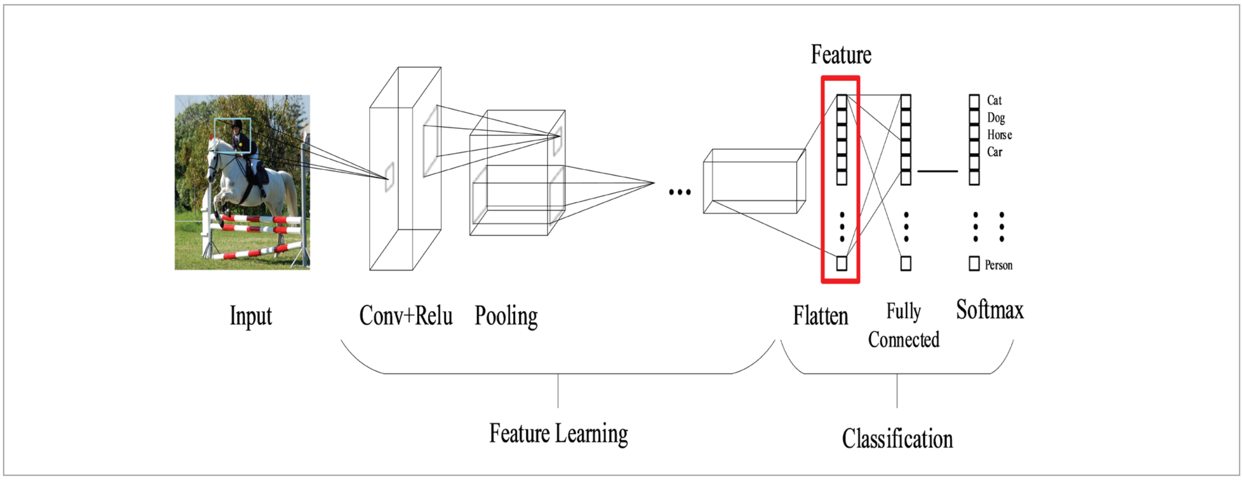
Since Alexnet [12] won the ImageNet large-scale visual recognition challenge (ILSVRC) in 2012, deep learning-based research has attracted considerable attention. Even though deep learning dealing with manual-based low-level feature descriptors is inclined to lead to losses of useful information prematurely, it is more effective to find out task-specific feature descriptions from image pixels directly rather than using manual features (Figure 1). Hence, more effective feature extraction methods promote the progress of related tasks such as image classification and retrieval, target location, and instance segmentation, thus achieving unprecedented success. For image classification, some models [22, 33, 10] have outperformed when compared with manual testing. On the other hand, Neural Networks tend to extract textures rather than shape features [25], which is quite different from the way humans perceive things. A similar problem was found in the implementation of the image classification algorithms. Moreover, positive samples that are significant to human vision tend to be misclassified, which led to grave deviation.

Early image feature extraction algorithms, for example, a bag of words [34,6], exhibited better interpretability. However, due to the insufficient characterization ability of manually established features at pixel levels, they are assessed as not efficient for specific tasks. Currently, the deep learning framework has received widespread attention, ensures complete end-to-end learning, and has strong representability for the extracted features. However, the efficiency of deep learning can only be improved through larger data sets and requires models to be trained as comprehensively as possible, which is difficult to apply in some situations. For example, when there are annotated data, training would still be cumbersome. Moreover, interpretability still becomes an issue.

Deep learning-based methods still try to overcome some issues that have been researched by scholars [7, 4, 36, 23, 9] which are regularization, overfitting, and generalizability issues of the models. While regularization generally deals with reducing the complexity of the models, generalizability copes with generating predictions based on unobservable data, which is called the robustness of the studied model. Some effective regularization approaches are called reduc-

**Figure 1**

High-quality image features extracted by deep learning



ing the complexity of the networks, generating new samples by utilizing new operations such as flipping, and transforming, employing ensembled models, and employing different training approaches such as early stopping criteria and picking better initialization points. On the other hand, generalizability is a significant issue that needs to be tackled since it means that while the constructed models work well on the training partition, it does not function as successfully as on the test partition. Thus, its prediction capability would become weak.

Moreover, several types of research have been recently conducted in various disciplines [37, 39, 30, 29, 2]. Therefore, its implementations can be expanded broadly.

Edges are called low-level feature that is often utilized for contour extraction, segmentation, and image sharpening. When compared with the original image, edge images only retain boundary information of an object, providing a small amount of data, having a strong generalization ability though, and are closely resembled the shapes of original objects. For example, night vision and infrared images have significant edge information. For images with specific shapes, the main content of the image can be accurately judged only by relying on edge features.

Conventional edge detection methods mainly employ edge detection operators, which depict clear physical boundaries and retain much of the original image information. The relative significance of edge detection results is consistent with human vision. Due to the lack of high-level mathematical models to extract semantic features of images, it is challenging to implement these methods to eliminate the interferences led by both noise and texture. Thus, deep learning methods to detect edges can be utilized to address these issues. However, in the case of images deprived of training, deep learning-based edge detection methods, whose performance is unstable, may easily lose significant features. Thus, it is essential to ensure the efficient use of the respective advantages of conventional edge detection and intelligent edge detection approaches concurrently to improve the feature extraction of images.

Since edges generally provide a limited amount of data, the methods such as deep learning-based models cannot generate robust models to make predictions due to not having a large training partition so auxiliary approaches should be employed to remedy the issues of those models.

By doing so, a method extracting edge features is combined with conventional edge detection operators, and the deep learning method is proposed. Thus, it fuses general image and edge image features for such tasks of image classification and retrieval, which leads to generating better interpretability. Therefore, better semantics are ensured. Experimental results show that the proposed method can significantly improve the accuracy and transferability of image classification along with both the accuracy and robustness of image retrieval.

Therefore, the extracted edge features of images when combined with deep learning-based methods utilizing suitable fusion methods could lead to approaches that resolve the issues and generate better outcomes for image classification and retrieval tasks. The proposed method is a novel method whose main contributions are listed as follows:

- The mean shift filtering is used to preprocess the data set to eliminate interferences.
- For image classification and retrieval tasks, an edge detection method combining the HED [3] and Sobel operators [18] is proposed to enable the fusion of extracted edge features and general image features.
- The proposed method improves both the accuracy and robustness of both image classification and retrieval.
- Samples are sorted according to the relative value of the probability of image classification to ensure that they are more suitable for calculating and analyzing the performance of the algorithm.
- Noise reduction is conducted by utilizing morphological processing.
- The transferability of the proposed method with a higher average value is achieved so the generalization ability is increased too.

## 2. Preliminary

### 2.1. Edge Detection

Edge detection is one of the major challenges in image processing. Edges generally provide underlying features that are necessary for other tasks. Decades of research have led to several achievements in the extraction of edge features. Edge features are different from the extracted features when neural networks are conducted. Although edge images contain only a small amount of data, they provide independent semantic meanings of original images. The results of edge image classification can be fused with those of original image classification at the semantic level, which can improve the classification accuracy as well as increase interpretability. In image retrieval tasks, the similarity of edge features can improve accuracy and provide better retrieval results that are more consistent with subjective human feelings.

Conventional edge detection methods mainly rely on manually designed operators for extraction, such as Sobel [42], Robert [35], LoG [4], and Prewitte operators [15]. Several improved algorithms have also emerged, such as Canny [32] and Susan [28]. Such operators employ the first and second derivatives of image data to extract edges and are useful in various optimization methods. In addition, these operators enable fast calculation and clear physical meaning but have some disadvantages since noises corrupt their performances so internal and external edges cannot be separated properly. Snake [13] utilized some control points to form a specific shape as a template called contour lines and achieved harmony through the elastic deformation of the template itself to match the local features of the image. Thus, a complete detection of the object contours in images is realized. However, this method is constrained by its mathematical model and has poorer interpretability for complex shapes. Conventional edge detection operators and their improved algorithms rely on manually constructed rules, which are intuitive and stable. However, it is challenging to implement manually constructed rules to make full use of different types of changes in images, resulting in issues of edge loss, inability to distinguish texture and contour, and imbalance between edge detection and noise resistance. The emergence of deep learning algorithms has provided novel ideas and helps deep neural network models achieve excellent detection effects [3, 45] by learning manually annotated edge image data sets. Thus, edge detection methods based on deep learning have small training sets due to the difficulty in labeling the edges and larger subjective differences among different taggers. When the brightness or texture type in the test image differs greatly from the training set, the generalization ability would become weak.

### 2.1.1. The Sobel Operator

Sobel edge detection operators include detection templates in $x$ and $y$ directions. Let the original image be defined as $I$. If $Gx$ is the convolution of the original image in the $x$-direction, and $Gy$ is the convolution of the original image in the $y$-direction, then both are represented by

$$G_X = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I \tag{1}$$

$$G_Y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I \tag{2}$$

Therefore, the edge image G is defined by

$$G = \sqrt{{G_X}^2 + {G_Y}^2}. \tag{3}$$

When a threshold T is given, binarized edge images can be represented by

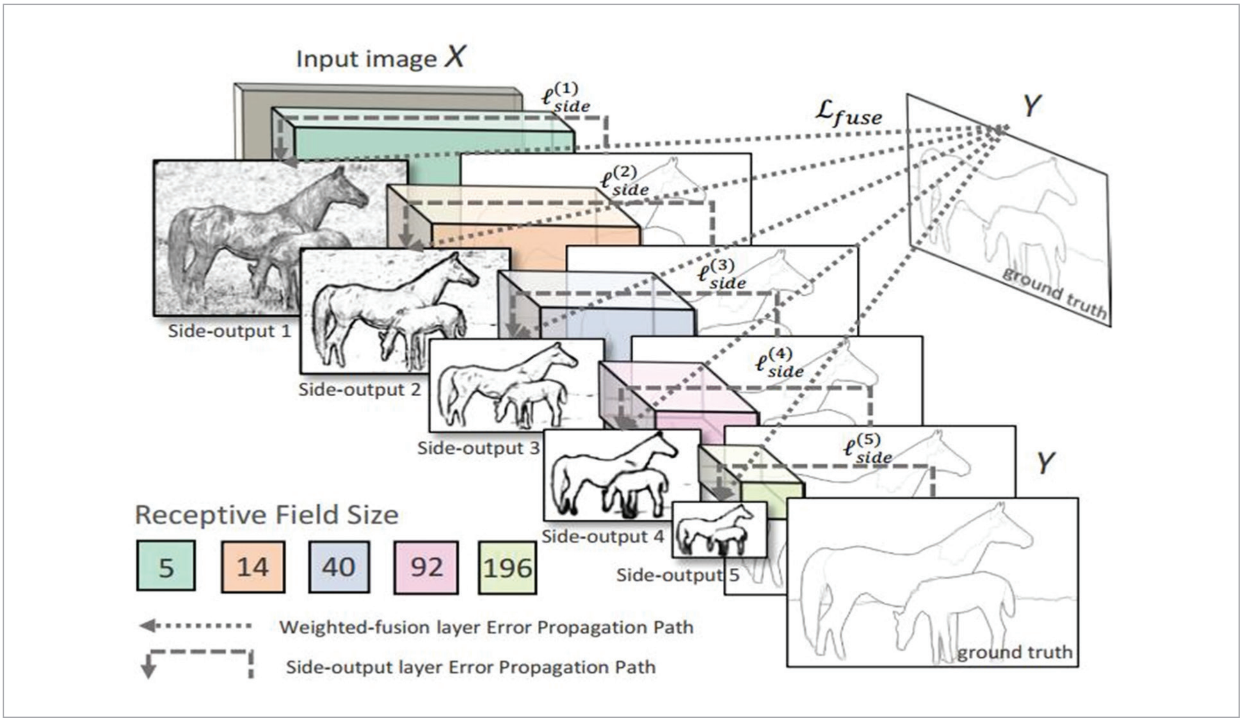
$$G = sign(G, T), \tag{4}$$

where the sign function assigns a value to G regarding the threshold value T. Although there have been many edge detection algorithms in the literature, Sobel operators are one of the widely implemented ones due to their simple, fast, and effective features. When compared with other edge detection operators, Sobel operators are more suitable for providing edge information for deep learning when clear images are readily available because they are simple and retain considerable information in original images.

### 2.1.2. The Detection Algorithm of the HED Edge

The HED edge detection method (Figure 2) utilizes the convolutional neural network (CNN) VGG [16] to train the standard data set BSDS500 for edge detection [40] (500 images in the form of typical size for edge detection algorithms having manually labeled edges by 5 persons). The HED proposed the concept of side output (Figure 2) and derived features of different scales at different convolutional layers. All side

**Figure 2**
The HED algorithm [17]



outputs were fused to obtain edge features, leading to excellent performance. This study was published in ICCV2015 and was nominated for Marr Prize.

The HED detects edges employing deep learning, enabling it to distinguish between textures and contours, and the detected edges are significantly resembling the shape of the object. However, because it utilizes deep learning to detect edges, the generalization ability of the algorithm depends on the size of the training data. The training set BSDS500 has only 500 images, which is a relatively small one, and its generalization ability is weak when encountering different types of images.

### 2.1.3. The Selection Method of the Otsu [27] Threshold

In the case of continuous brightness values, the edge detection results do not directly indicate the exact position of the edge, and generally, binarization operation is required after a given threshold. The maximum inter-class variance method, referred to as the Otsu method, was proposed by Otsu in 1979, which is an adaptive threshold determination method. The Otsu method divides an image into two parts, the background, and the target according to its gray characteristics. A larger interclass variance between the background and the target implies a greater difference between the two parts of the image. When a part of the target is misclassified into the background or a part of the background is misclassified into the target, the resulting interclass variance would become smaller. Therefore, a segmentation that maximizes the variance between classes implies the smallest probability of misclassification. The principle of the algorithm is expressed as follows:

If threshold *T* exists, all pixels can be divided into two types, *C1* and *C2*. The mean values of these two types of pixels are denoted *m1* and *m2*, respectively. The global mean value of the image is denoted by *m*. The probability of a pixel being classified as *C1* and *C2* is denoted by *p1* and *p2*, respectively. The relations are expressed by

$$p1m1 + p2m2 = m \tag{5}$$

$$p1 + p2 = 1. \tag{6}$$

The interclass variance $\sigma^2$ can be expressed by

$$\sigma^2 = p1(m1 - m)^2 + p2(m2 - m)^2. \quad (7)$$

Substituting (5) and (6) into (7), we rearrange them as follows:

$$\sigma^2 = p1p2(m1 - m2)^2. \quad (8)$$

$T$, which can maximize the variance in (8) between classes, is called the optimal threshold.

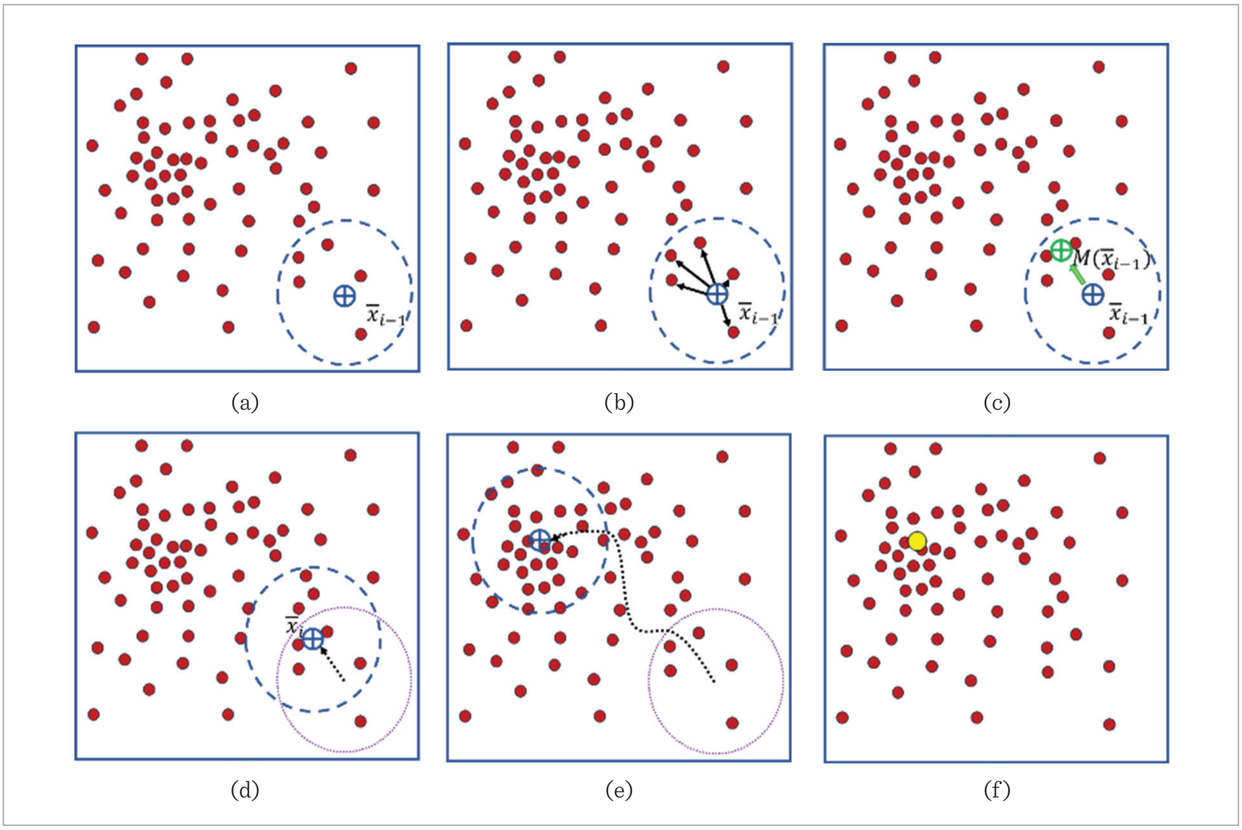
#### 2.1.4. The Mean Shift Filtering [5]

Conventional edge detection operators as well as deep learning-based edge extraction methods are easily corrupted by noises. Therefore, edge detection should typically be combined with both image filtering and morphological processing in practical applications.

Conventional filtering methods include averaging, median, and Gaussian filtering operations, which can aid the removal of noises but lead to the loss of useful information in the images. Researchers have also proposed many nonlinear filters, such as bilateral filtering [43] and mean shift filtering, which can retain information in the original image while filtering out the noise. Thus, these have been increasingly implemented in various applications of image analysis.

The mean shift filtering algorithm was first proposed by Fukunage in 1975, which is a clustering algorithm replacing the RGB values of all pixels into a class with the center value of the class to achieve data compression and noise reduction. As shown in Figure 3, an initial point is randomly selected (a) to calculate the current point in the filter radius offsetting averages (b, c), and then, a new starting point (d) is selected, and the operation is continued until certain final conditions are met (e). The final mean value is called the clustering center of this class (f), and the sample points utilized in the calculation are marked within

**Figure 3**
The mean shift filtering



(a) (b) (c)

(d) (e) (f)

this class. Then, unlabeled sample points are selected as starting points. Afterward, the process is repeated until all sample points are fully labeled. The final class of the sample point is called the class with the most marks, and the RGB values of the point are replaced by the RGB values of the cluster center of the class to which it belongs. The mean shift operator is widely utilized in image clustering, smoothing, segmentation, and tracking. In this study, it is employed for image preprocessing before conducting edge detection, reducing interferences caused by high-frequency components in images.

### 2.2. Image Classification

Image classification has been developed over many years, and several well-adapted methods [12, 33, 10] have been presented for object recognition in static images. For large data sets, better classification results can be obtained by employing various classification models based on neural networks to train all parameters. Hence, a larger data set implies comprehensive knowledge of image features. For instance, Vit [22] used 300 million images for pre-training, combined these with the attention mechanism, and achieved better results than do CNNs. However, for small data sets, good performances cannot be achieved by training all parameters, and there could exist a chance of overfitting. Transfer learning [24] employs image features learning in large data sets and does fine-tune them in small data sets, achieving satisfactory results, which has been widely implemented.

However, in industrial applications, the main problem is relatively low accuracies, which is less than 100%, that is. It is expected that an obvious positive sample should not be misclassified. The results of the classification should be as close as to human subjective feelings. In other terms, images with obvious features should not be misclassified. However, deep learning frameworks are aimed at improving the classification accuracy of whole training sets rather than single images. Therefore, this problem cannot be resolved effectively in the theoretical framework.

### 2.3. Image Retrieval

Image retrieval can be split into two classes, which are called text-based image retrieval (TBIR) and content-based image retrieval (CBIR). The TBIR searches mainly for transforming an image search into a text search and builds structured labels based on image content depending upon label text. However, the TBIR has poorer scalability and requires considerable manual interventions. On the other hand, in the early stage of the CBIR development, researchers mostly employed the global features of images. Since 2003, due to the excellent performance of the SIFT [14,44] features in image transformation (scale and direction change), image retrieval methods based on local descriptors have been widely researched. In addition, in recent years, CNN-based image representation methods [21, 17] have attracted increasing attention and shown satisfactory performance. The excellent performance of CNNs in image retrieval and other image-processing tasks is mainly attributed to the powerful feature extraction capability of convolutional layers. More effective feature extraction methods can further improve the performance of CNNs. Deep CNN models are employed to extract underlying feature vectors of images, transforming image retrieval into similar vector retrieval. By doing so, strong scalability and high processing efficiency are reached.

## 3. Method

### The Overall Method

The proposed method is illustrated in Figure 4. The edge image is obtained by edge detection of the original image. CNN is employed to convert the original image and edge image into respective feature vectors. Finally, the two features are fused for different tasks, leading to improved performance for the classification and retrieval of images.

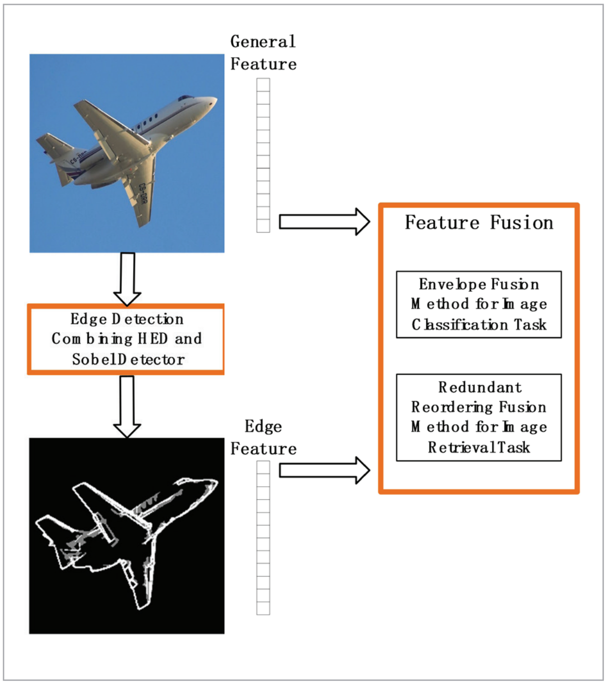
The novelty of this paper is mainly depicted in the two expressions presented inside of the orange frame, that is, extracting edge images and fusing both edge and general features for image classification and retrieval tasks.

### 3.1. An Edge Detection Method Based on Both HED and Sobel Detectors

1 Sobel operators employ the information in neighborhoods of a pixel to detect edges, and the edge detection results reflect the changes in local gray levels, providing rich image details. The HED method obtains the overall cognition of the image through

**Figure 4**

Feature extraction and fusion


General Feature
Feature Fusion
Envelope Fusion Method for Image Classification Task
Redundant Reordering Fusion Method for Image Retrieval Task
Edge Detection Combining HED and SobelDetector
Edge Feature

deep learning, thus combining the HED and Sobel operators brings their respective advantages concurrently to attain better results for edge detections. In addition, some shortcomings of the HED method are addressed in the combination. It shows excellent performance in the training phases, but there exist still some issues when the HED in other images that differ greatly from those in BSDS500 is utilized. Those are expressed as follows:

I. The edge pixel values are continuous, and the data volume of edge images is large, which easily leads to overfitting and weak generalization ability in subsequent deep learning tasks.

II. When the image quality is very high, and the edge is very clear, the detected edge is not sufficiently accurate.

III. The HED method is not affected by dark lines on a bright background.

IV. It is influenced by noise interference.

2 This manuscript proposes a method that improves the effect of edge detection by utilizing mean shift filtering, image inversion, quantization of edge brightness, and morphological processing combined with the HED and Sobel edge detection operators to resolve the above problems,

As a result, the generalizability of the edge detection algorithm and the accuracy of edge position is enhanced concurrently. Thus, the data amount of the edge image is compressed, and the robustness is improved.

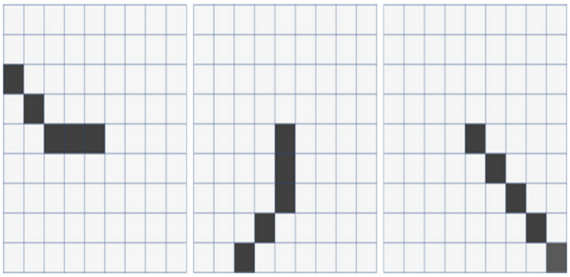
In terms of morphological processing,

3 We propose a new method that utilizes broken line elements to carry out corrosion calculations.

The broken line template unit consists of two end-to-end connected three-pixel short lines either with an angle of 180° or 135° since a rotation freely rotated around the center point is conducted. Overall, there exist 24 specific forms, 3 of which are listed in Figure 5. For the edge image, all templates are utilized for convolution. When the result of the convolution of any template is equal to 5, the edge pixel corresponding to the center point is retained. Otherwise, the edge pixel is discarded. This process is repeated three times to effectively remove noise points in the edge image.

**Figure 5**

A broken line template



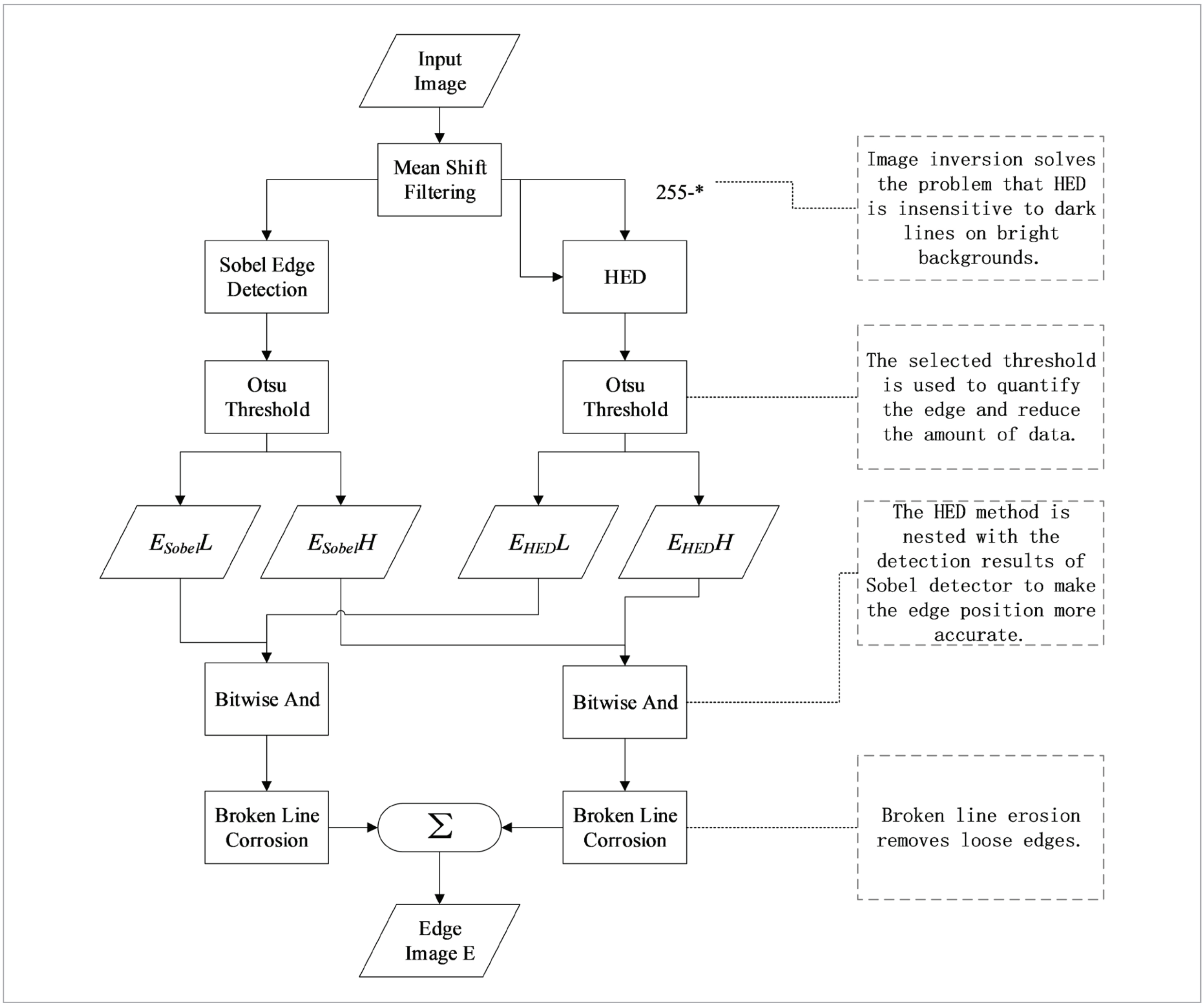
As shown in Figure 6, the original image set is denoted by $I (w, h, 3)$, and $I'$ is obtained after conducting mean shift filtering.

4 Then, both Sobel and HED methods are employed to generate edge images $E_{Sobel}$ and $E_{HED}$, respectively. Four-edge images, namely $E_{Sobel}L$, $E_{Sobel}H$, $E_{HED}L$, and $E_{HED}H$, are obtained after running a quantization by the Otsu method.

5 After conducting both summation operation and broken line corrosion, the final edge image $E$ is obtained. The steps of the algorithm are presented as follows:

**Figure 6**

Edge detection method based on both HED and Sobel detectors


Input Image
Mean Shift Filtering
255-*
Image inversion solves the problem that HED is insensitive to dark lines on bright backgrounds.
Sobel Edge Detection
HED
Otsu Threshold
Otsu Threshold
The selected threshold is used to quantify the edge and reduce the amount of data.
$E_{Sobel}L$
$E_{Sobel}H$
$E_{HED}L$
$E_{HED}H$
The HED method is nested with the detection results of Sobel detector to make the edge position more accurate.
Bitwise And
Bitwise And
Broken Line Corrosion
Σ
Broken Line Corrosion
Broken line erosion removes loose edges.
Edge Image E

**Edge Detection Algorithm**

```
Input: I (w, h, 3)
Output: E (w, h, 1)
1. I’ = Mean shift (I, sp=15, sr=20) #” sp” is the physical radius of Mean Shift, “sr” is the RGB radius.
2. # First calculate the left branch.
3. $E_{Sobel}$ = max (Sobel (I’[:, :, 0]), Sobel (I’[:, :, 1]), Sobel (I’[:, :, 2]))
4. $T_{Sobel}$ = Otsu ($E_{Sobel}$)
5. $E_{Sobel}L$ = BW ($E_{Sobel}$, $0.5T_{Sobel}$)#The $E_{Sobel}$ was binarized with a threshold of $0.5T_{Sobel}$.
6. $E_{Sobel}H$ = BW ($E_{Sobel}$, $T_{Sobel}$)
7. #Then calculate the right branch.
8. $E_{HED}$=Max ( HED(I’), HED (255-I’) )
9. $T_{HED}$ = Otsu ($E_{HED}$)
10. $E_{HED}L$ = BW ($E_{HED}$, $0.25T_{HED}$)
11. $E_{HED}H$ = BW ($E_{HED}$, $T_{HED}$)
12. EL = $E_{Sobel}L$.&& $E_{HED}L$
13. EL = Erode (EL)
14. EH = $E_{Sobel}H$.&& $E_{HED}H$
15. EH = Erode (EH)
16. E = 0.5 * (EL + EH)
```

### 3.2. The Feature Fusion Method Based on Both Edge and General Features

By employing the deep learning framework, two classification models, *Model,* and *Modele* are trained to utilize the original image and the edge image data sets, respectively. The classification task utilizes these two models to discriminate the original and edge images. Finally, the outcomes are fused. The retrieval task employs the fully connected (fc) layer features of the *Model*, and then, utilizes the fc layer features of *Modele* to reorder the retrieval results. Research [18] has shown that features of different layers of neural networks have specific complementary features, and the accuracy of image retrieval can be improved by the fusion of features. The proposed fusion method utilizes complementary edge features and general features too. The performances of both image classification and retrieval tasks are improved concurrently.

#### 3.2.1. The Envelope Fusion Method for the Image Classification Task

When deep learning is used for image classification with correctly predicted outcomes, the maximum probability of all classes (the probability of true classification) is generally high. However, when the correctly predicted outcomes are low, the probability of false classification is large. Two classification models corresponding to general image features and edge features are utilized to predict image classification. When the two predictions are inconsistent, the classification accuracy can be improved by judging the classification according to the prediction probabilities.

The original image set is denoted by *I*, and the classification model called *Model* is implemented to attain the class probability vector *v* defined by

$$v = \mathrm{Softmax}(\mathrm{Model(I)}), \tag{9}$$

where softmax is a normalized exponential function.

The corresponding class label is denoted by *Ci* and is defined by

$$Ci = \mathrm{argmax\ index(v)}, \tag{10}$$

where argmax is a function that gives the maximum value of domain x for the function attaining the maximum value in the range.

For edge image *E*, the classification model represented by *Modele* is employed to attain classification probability *ve* and is defined by

$$ve = \mathrm{Softmax}(\mathrm{Modele(E)}). \tag{11}$$

The corresponding class label is denoted by *Ce* and is defined by

$$Ce = \mathrm{argmax\ index(ve)}. \tag{12}$$

For the probabilities *v*[*j*] and *ve*[*j*] of each class, the maximum value is taken to obtain the fused probability vector *V* represented by

$$\mathrm{V[j]} = \max(v[j], ve[j]). \tag{13}$$

The final class label is denoted by C and is expressed by

$$C = argmax\ index(V). \tag{14}$$

When probabilities of the classes need to be obtained, *V* is normalized so that the summation of the probabilities of the classes is equal to 1 and is expressed by

$$V = Normalize(V). \tag{15}$$

#### 3.2.2. The Redundant Reordering Fusion Method for the Image Retrieval

Contemporary applications utilize general image retrieval methods [21] to attain the convolutional layer output of the classification model to extract features from the original image data set, establish the vector database, and *Collection*. To retrieve image *I*, the corresponding feature vector is denoted by *f*. The predefined distance called cosine distance is employed in *Collection* to identify the most similar vector, and the image corresponding to the vector is the most similar image. The performance index, called average precision (AP), of image retrieval, is defined by

$$AP = \frac{1}{N} \times \sum_{i=1}^{N} \frac{1}{position(i)}, \tag{16}$$

where *N* denotes the total number of related images, and *position(i)* represents the position of the *i*th related image in the list of retrieval results. Mean average precision (MAP) is called the average of the AP of multiple queries defined by

$$MAP = \frac{1}{M}\sum_{i=1}^{M} \mathrm{AP}_i. \tag{17}$$

Increasing the number of retrieval images is typically utilized to increase the accuracy of the retrieval by further combining it with other information. As shown in Figure 7, the query image is set to *I*. When the most similar image of the *top(k)* is needed to be retrieved, the following calculations are carried out as follows:

1 The general feature vector *f* (Equation (18)) is employed to retrieve the $top(\lfloor 1.2k \rfloor)$ that shows the number of the most similar images, denoted by $top_+(k)$.

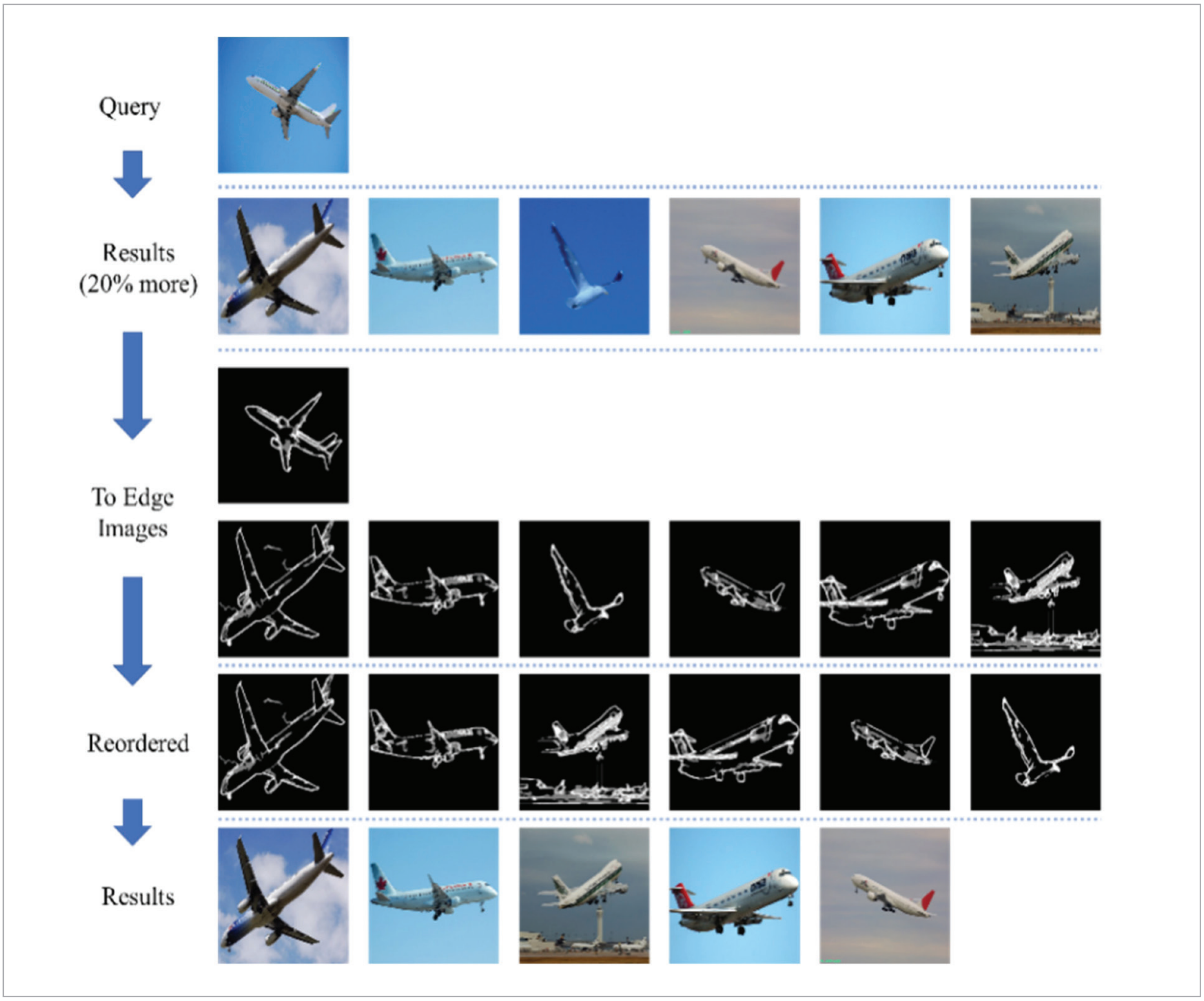
$$f = \mathrm{Model}_{\mathrm{fc}}(\mathrm{I}) \tag{18}$$

$$top_{+}(\mathrm{k}) = \mathrm{search}(f, \mathrm{collection}, \lfloor 1.2\ \mathrm{k} \rfloor) \tag{19}$$

2 The divergence (*Div*) of *I* and $top_+$ (Equation (20)) is calculated, where *ve* is obtained from Equation (11), and $max_{(i)}th$ represents the (i)th largest value in all dimensions of the vector. The complementarity between edge features and general features is more significant when the discrete degree of class probability is large. Then, the value of *Div* can be utilized to decide whether to carry out feature fusion.

**Figure 7**

Redundant reorder image retrieval



$$Div = \left| \left( \frac{1}{10} \sum_{i=1}^{3} \frac{max_{(i)th}(ve+0.01)}{\sqrt{max_{(i+1)th}\ (ve+0.01)}} \right) \right| \quad (20)$$

3 When both *Div(I)* and *max (Div(top₊))* are greater than the threshold L (set as 0.5 herein), $top_+(k)$ is reordered. The corresponding edge image is calculated, and the edge feature vector *fe* is obtained (Equation (21)). Afterward, the edge feature vectors are utilized to calculate the similarity value with the query image *I*. The images are finally reordered.

$$fe = \mathrm{Modele_{fc}(E)} \quad (21)$$

$$top_+(k) = \mathrm{Reorder}(top_+(k), fe) \quad (22)$$

4 *top(k)* is extracted from $top_+(k)$ as the final retrieval result.

$$Result = \mathrm{top}(top_+\mathrm{(k)}, k). \quad (23)$$

# 4. Experiments

## 4.1. Edge Detection

The proposed method is employed to detect image edges, and the experimental results are presented below.

As shown in Figure 8, images with many dark lines are selected to detect edges using the HED and the proposed methods. The edge detection method has an image color reversal step. The problem could not be resolved to utilize the HED since the difficulty in detecting dark lines on a bright background is addressed.

As shown in Figure 9, an image is randomly selected from the VOC2012 data set, and edge detection is performed utilizing the Canny operator, the Sobel operator, the HED, and the proposed method.

When compared with both Canny and Sobel operators, the proposed method is corrupted less by noise and high-frequency components while fully retaining image edge information. When compared with the HED method, the position of the edges is more accurate, and the amount of data is less (only three brightness levels: 0, 0.5, 1.)

**Figure 8**

When compared with the HED, the proposed method exhibits better performance to detect dark lines on a bright background (Original image (top), the HED method (middle), and the proposed method (bottom))

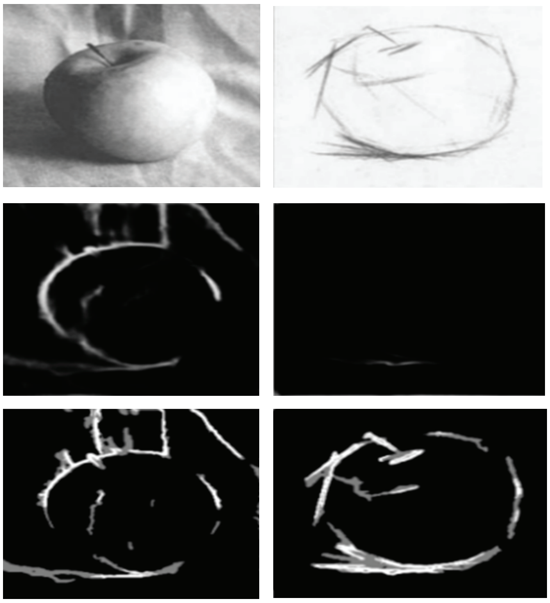
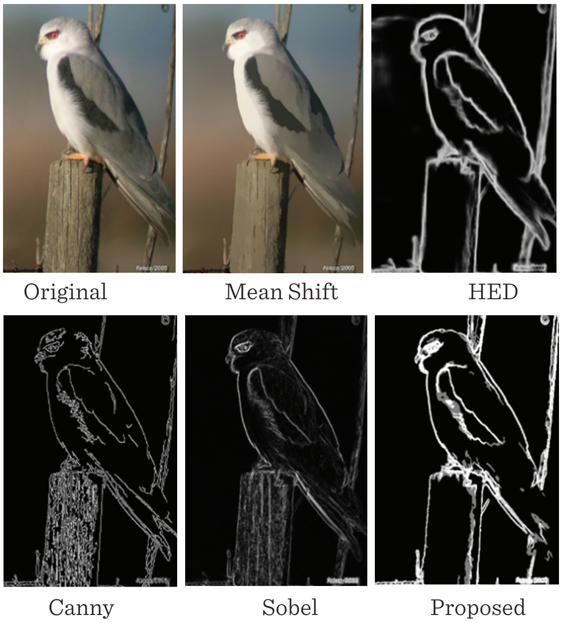


**Figure 9**

Edge detection effects



Original Mean Shift HED

Canny Sobel Proposed

### 4.2. The Accuracy Experiment of the Image Classification

Resnet50 [33] was utilized to conduct image classification on the Office-Caltech 10 data set, which contains four different domains, namely, Amazon, Caltech, DSLR, and Webcam. Each domain contains 10 types of images. This data set can not only verify the accuracy of image classification but also reflect the performance of transferring the model trained in one domain to other domains.

#### 4.2.1. The Experiment of the Image Classification

The training was conducted on each domain and classification tests were performed on all four domains. We implement this method to employ only general image features as a baseline. Then, we replaced the general features with edge features and repeated the experiment by utilizing the proposed method. The classification results are presented in Table 1.

**Table 1**

The image classification results of four different data domains (%)

| Training Domain | Feature | Amazon | Caltech | DSLR | Webcam |
|---|---|---|---|---|---|
| **Amazon** | Baseline | 99.37 | 60.28 | 52.87 | 38.31 |
| | Edge (3.1) | 98.54 | 59.57 | 55.41 | 50.84 |
| | The Proposed | 99.69 | 65.45 | 61.78 | 48.14 |
| **Caltech** | Baseline | 81.73 | 97.95 | 71.34 | 65.76 |
| | Edge (3.1) | 73.38 | 95.10 | 61.15 | 58.64 |
| | The Proposed | 84.86 | 98.93 | 73.25 | 69.15 |
| **DSLR** | Baseline | 32.25 | 32.06 | 99.36 | 82.71 |
| | Edge (3.1) | 32.67 | 34.82 | 97.45 | 76.61 |
| | The Proposed | 41.75 | 39.54 | 100 | 88.14 |
| **Webcam** | Baseline | 25.89 | 27.87 | 79.62 | 99.32 |
| | Edge (3.1) | 47.08 | 37.40 | 94.90 | 100 |
| | The Proposed | 44.47 | 38.11 | 97.45 | 100 |

Table 1 shows that when only proposed edge images (Edge Image E in 3.1) are utilized to classify, the accuracy is slightly lower than that of the baseline. However, due to the complementarity of edge features and general features, the fusion results based on the proposed method demonstrated significantly improved accuracy, which are 7.88%, 5.97%, 7.32%, and 4.83% on Amazon, Caltech, DSLR, and Webcam, respectively, with an average of 6.5%. Thus, it showed better transferability in all cases.

#### 4.2.2. The Comparison of the Proposed Edge Detection Method and Sobel Detector for the Image Classification

The proposed edge detection method is replaced by the Sobel operator, and the corresponding experimental results are shown in Figure 10.

The edge image obtained employing the Sobel detector retains considerable information from the original image, while in the case of the edge obtained utilizing the proposed method, some non-object contour edges are lost. Thus, it also has fewer data but a stronger generalization ability. As shown in Figure 10, when the Sobel operator is implemented to detect edges, the classification accuracy increases slightly after running fusion. That the edge image features obtained utilizing the proposed method are more complementary to the general image features presented (Figure 11), which is more effective for improving the classification outcomes.

#### 4.2.3. Experimental Analysis of the Image Classification

As shown in Figure 12, the class probability is analyzed. While the horizontal axis represents the samples intercepted by the black horizontal line in the left figure, the vertical axis represents the class probability. When the classification model called *Model* is implemented to predict the class label of a sample, a probability vector *V* is obtained, in which the probability of true classification is indicated by a red line (*S*), and the other dimensions represent the probability of a false classification, and the green line (*N*) in the figure indicates the maximum probabilities of all false classifications. When $S/N>1$, it implies that the model can correctly classify the samples (i.e., the part between two points A and B on the graph).

To obtain the sample distribution as shown in Figure

**Figure 10**

The comparison of the proposed method and use of the Sobel detectors (accuracy (%))

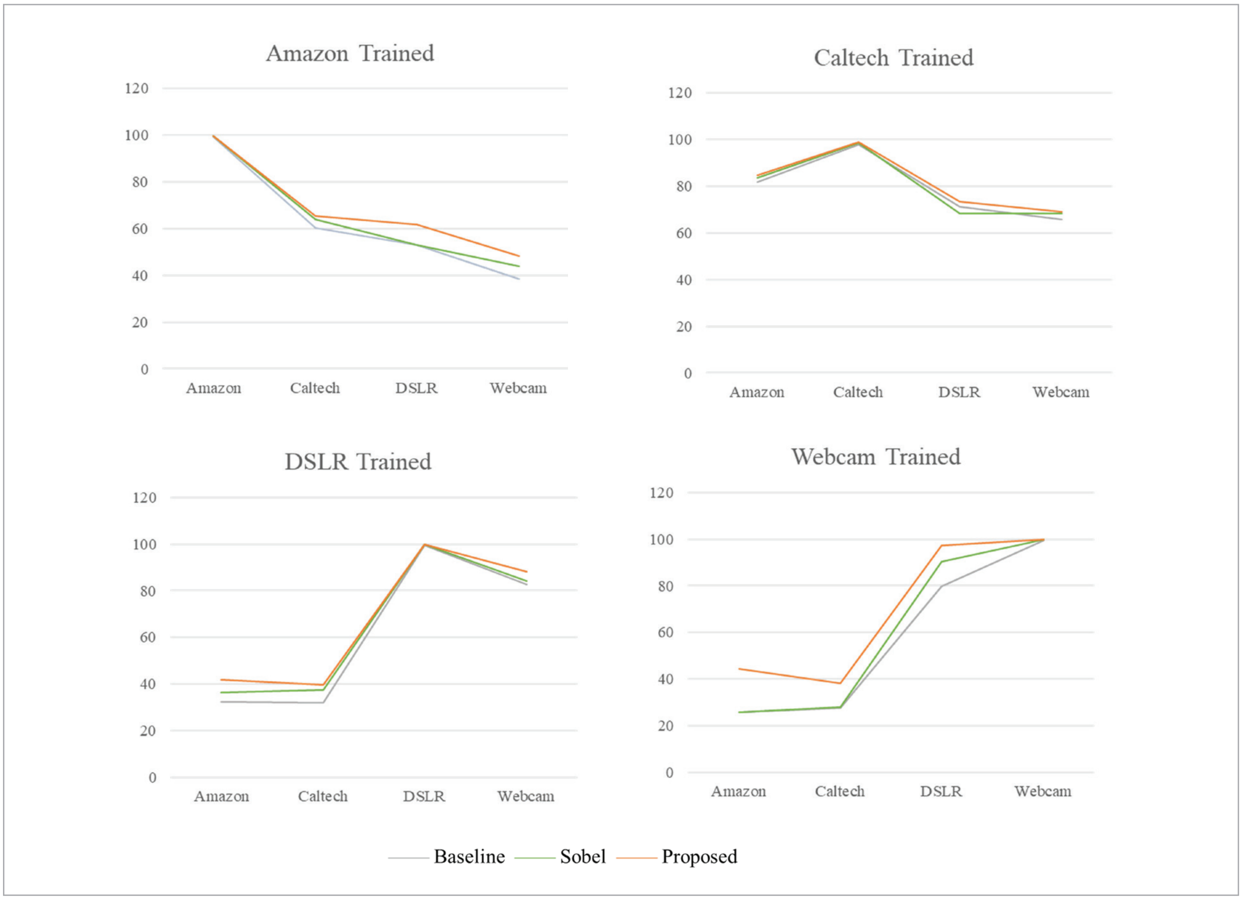
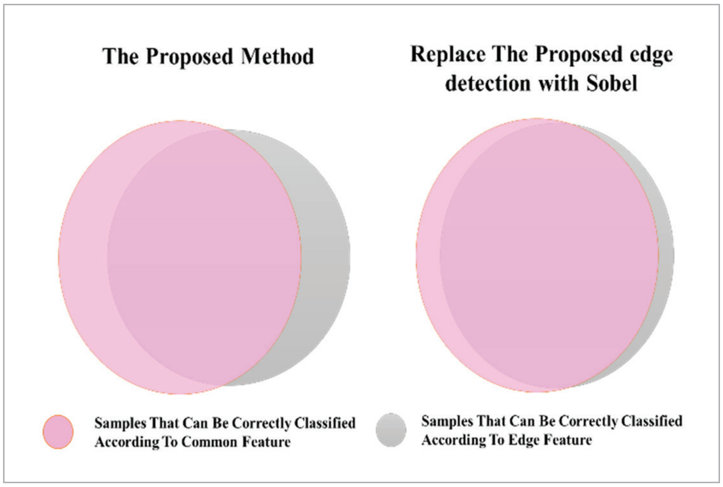


**Figure 11**

Effective improvement of the accuracy of image classification using the proposed method


The Proposed Method
Replace The Proposed edge detection with Sobel
Samples That Can Be Correctly Classified According To Common Feature
Samples That Can Be Correctly Classified According To Edge Feature

12, all samples are arranged according to the value of $S/N$. First, the image to the left of the dashed line in Figure 12 is calculated, and all samples to be measured (ILSVRC2012 is utilized here, class=0 (tench)) are randomly divided into two groups. For groups 1 and 2, *Model* (Resnet50) is utilized to classify some (≤ 10%) randomly selected samples from group 1 and arrange them in the ascending order of $S/N$ (the horizontal axis in Figure 13 (a) representing the sorted samples). Both $S$ and $N$ show some fluctuations. After conducting 5000 experiments and averaging $S$ and $N$ (Figure 13 (b)), it is shown that the images of $S$ and $N$ are smoother, and the monotonicity is opposite. Group 1 is changed to Group 2, and the ascending order of $S/N$ is altered to descending order.

**Figure 12**

When $S/N$>1, the correct classification of the samples (|$AB$|)

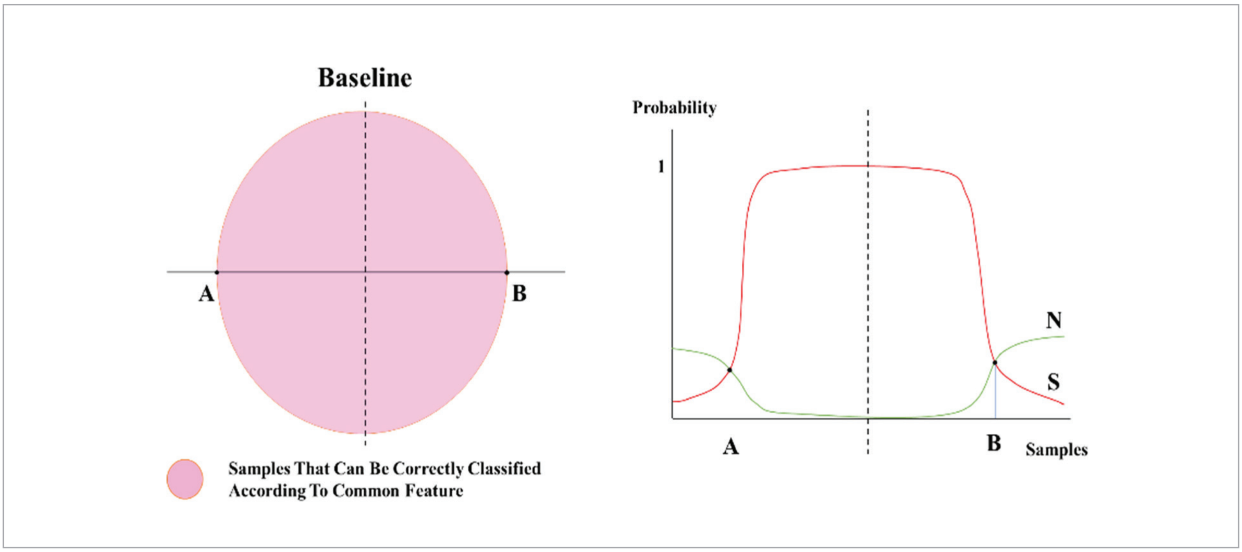
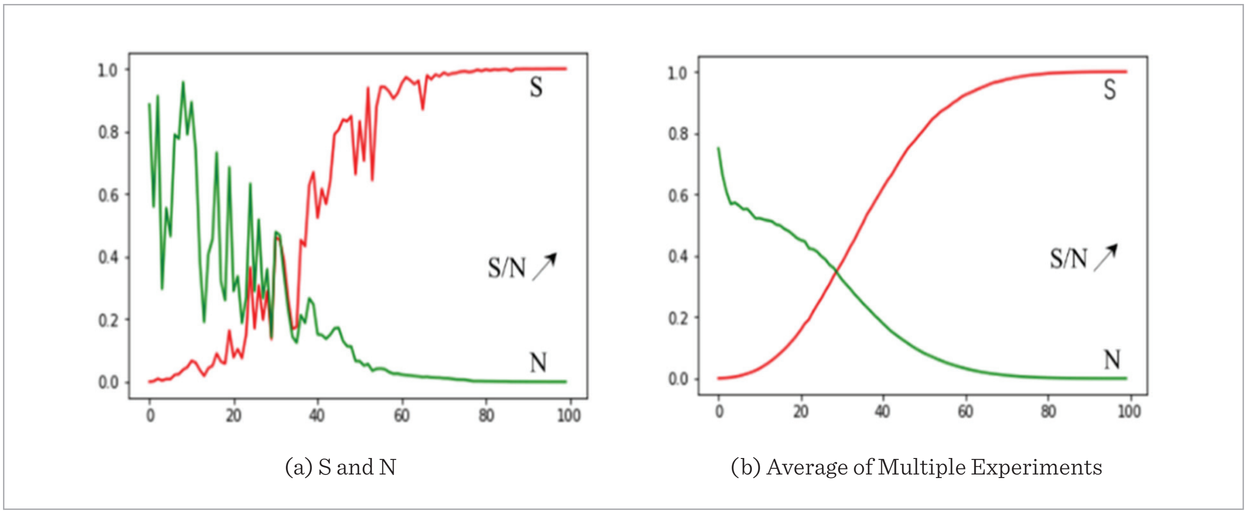

Baseline
A
B
Samples That Can Be Correctly Classified According To Common Feature
Probability
1
N
S
A
B
Samples

**Figure 13**

Samples sorted in ascending order of $S/N$ (model=Resnet50, data set=ILSVRC2012, class=0 (label=" tench"))



(a) S and N

(b) Average of Multiple Experiments

The image to the right of the dashed line in Figure 12 is obtained utilizing the same method.

The analysis of the proposed method is shown in Figure 14. When compared with Figure 12, the classification probability of *Modele* is found to be higher. The red and green dashed lines indicate the probability of true classes and the maximum probabilities of false classes, respectively. According to Equation (13), both $A$ and $B$ move to the right to reach $A'$ and $B'$ after running the fusion method. In deep learning-based image classification, the decline of $S$ is steep while the increase of $N$ is smooth. They are presented by expressions as follows:

$$|AA'| < |BB'| \tag{24}$$

$$|A' B'| = |AB| - |AA'| + |BB'| > |AB|. \tag{25}$$

**Figure 14**

Increased accuracy of classification using the proposed fusion method

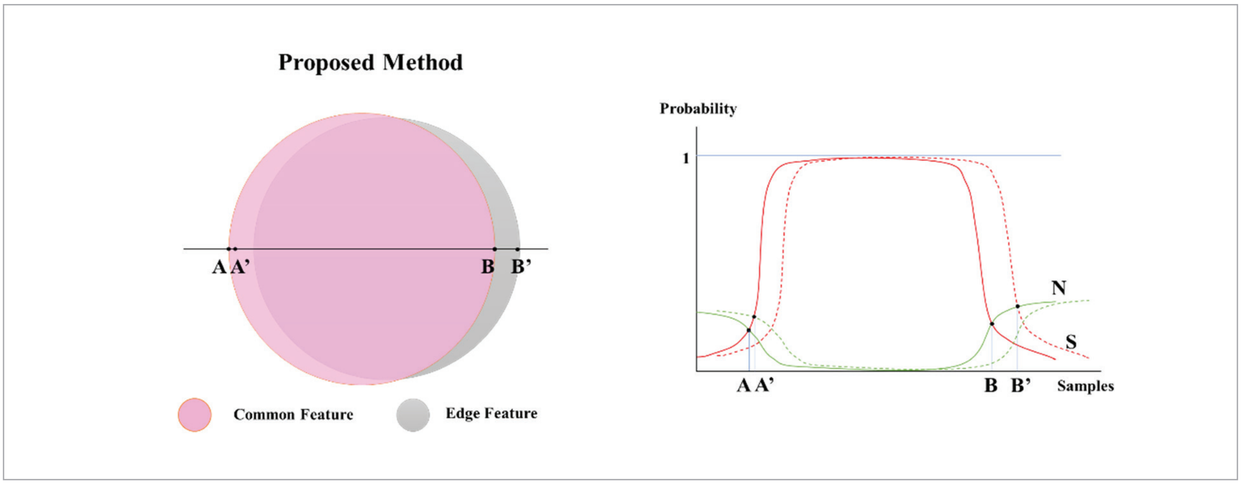
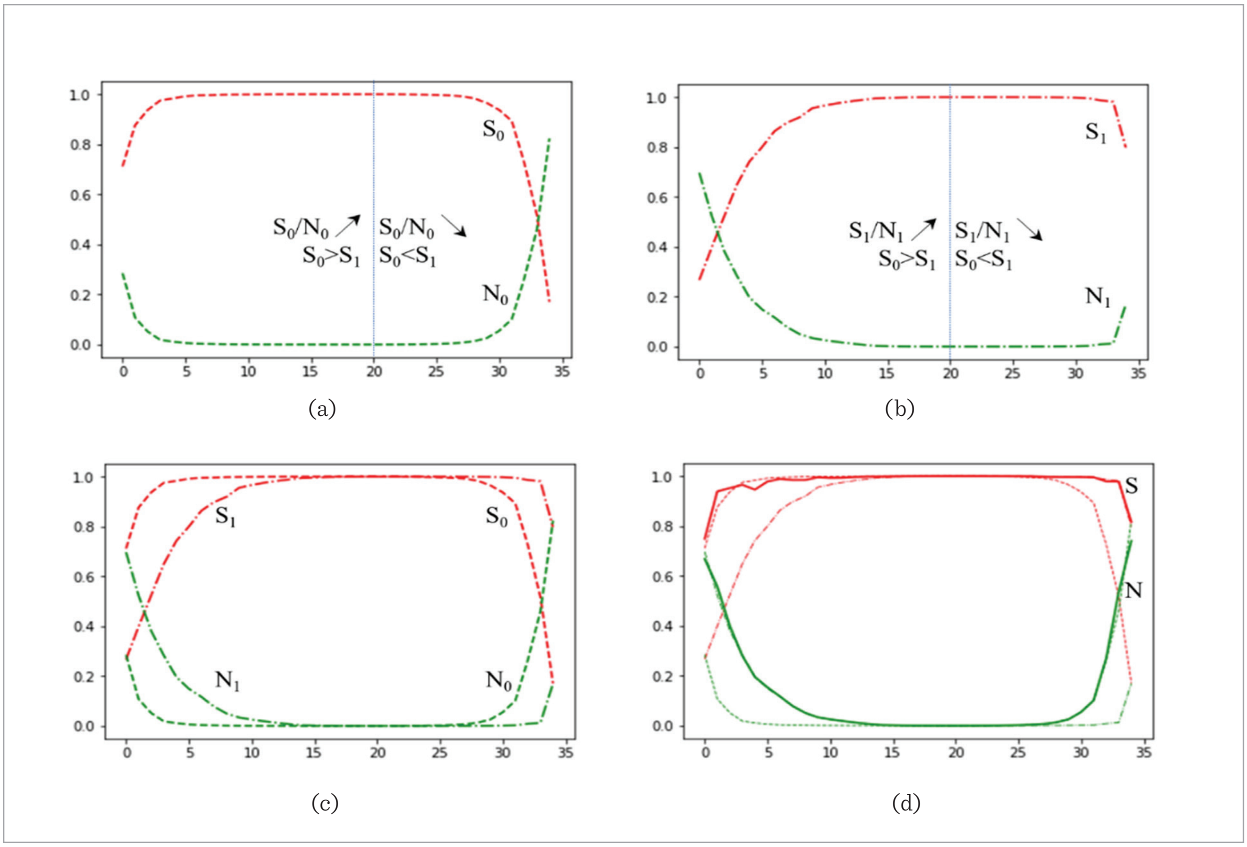


**Figure 15**

The backpack images of the Office-Caltech 10 data set used for the classification experiment



(a)

(b)

(c)

(d)

Therefore, the accuracy of the classification increases.

Figure 15 shows the results of the classification experiment, which is Experiment 4.2.1 utilizing the backpack images from the Office-Caltech 10 data set, where the red and green lines represent *S* and *N*, respectively. Subscripts 0 and 1 indicate classifications using *Model* and *Modele* (the two features), respectively. When $S_0$>$S_1$, images are assigned to group 1 and the remaining images are assigned to group 2, and then, the method shown in Figure 13 is adopted where (a) represents the result of classification utilizing *Model* (baseline), (b) depicts the result of classification employing *Modele*, (c) presents the complementarity of *Model* and *Modele* based on the classification results, and (d) adds red and green solid lines based on the outcomes of (c) to represent the values of *S* and *N* after running the fusion of (a) and (b), respectively. After conducting fusion, the intersection points of *S* and *N* in (d) are observed on the outside of the graph, and the final classification accuracy rate would become significantly higher, which is almost 100% and is greater than that of A or B.

### 4.3. The Accuracy Experiment of the Image Retrieval

The pascal-VOC2012 data set is utilized to conduct experiments, and the retrieval results are the top (5) ones. As query images are to be retrieved, 1000 images are randomly selected, and the remaining images are utilized to obtain vector sets via the feature extraction model called $Model_{fc}$. The image retrieval method is also implemented by employing only general image features as the baseline. To test the robustness of the image retrieval method, adversarial perturbation [19] and Gaussian noise are added to the query images to be tested. The experimental results are shown in Table 2.

**Table 2**

The image retrieval utilizing the Pascal-VOC2012 data set (MAP)

| | Clear | Gaussian Noise ($\sigma^2$=0.02) | Adversarial Perturbation ($l_\infty$=0.05) |
|---|---|---|---|
| Baseline | 0.8392 | 0.8413 | 0.6941 |
| The Proposed | 0.8530 | 0.8422 | 0.7013 |

Experimental results show that the proposed method can improve the accuracy of image retrievals. When Gaussian noise and adversarial perturbation [20] are added to the images, the image retrieval is corrupted and the accuracy decreases. However, the proposed edge detection method employs both quantization and broken line corrosion operation, thus greatly limiting the depth pattern structure in Gaussian noise and adversarial perturbation. Therefore, the robustness of the image retrieval method is improved.

Even though better outcomes are attained by conducting the proposed method, there exist some limitations stemming from the implemented methods that are the building blocks of the proposed method. 1. We did not study the impact of other operators on the proposed method to make a comparison. 2. Since the proposed method is based on data, the superiority cannot be generalized.

## 5. Conclusion

The complementarity of features has always been a popular research topic. In addition to semantically separable features, spatially separable features also have complementarity features.

In this research, an edge detection method combining the HED and Sobel detectors is proposed. To obtain better performances in deep learning tasks, the proposed method employs mean shift filtering, color reversal, quantization of edge brightness, broken line corrosion, and morphological processing. By doing so, the effect of edge detection is improved when combined with the HED and Sobel edge detection operators.

Thus, the two issues, called the generalizability and the accuracy of edge position, underlined in the literature are resolved considerably. Moreover, the amount of the edge image data is compressed and resulted in better robustness.

The proposed edge detection method outperforms the methods of Canny, Sobel, and HED. While the proposed method is better than both Canny and Sobel concerning being affected less by noise and high-frequency components and completely fully retaining image edge information, it outperforms the HED regarding having the more correct position of the edges with less amount of data.

For image classification and retrieval tasks, the edge features and general image features are fused. Experimental results suggest that the proposed method can reduce the amount of edge data and increase the complementarity of features while providing sufficient shape information. The accuracy and robustness of image classification and retrieval are improved by fusing general image and edge image features.

Different from the end-to-end paradigm of deep learning, the proposed method utilizes the fusion of semantically separable features. After reordering the samples, monotonic and stable probability graphs are obtained, which facilitate further analysis and better understanding.

In future studies, the fusion method dealing with spatially separable features and its application in both target detection and instance segmentation will be investigated based on the proposed method.

**Funding:** This study did not receive any funding.

**Data availability:** Data will be provided upon request to the authors.

**Conflict of Interest:** None.

## References

1. Abd Ghani, M.K., Mohammed, M.A., Arunkumar, N., Mostafa, S.A., Ibrahim, D.A., Abdullah, M.K., Jaber, M.M., Abdulhay, E., Ramirez-Gonzalez, G., Burhanuddin, M. Decision-Level Fusion Scheme for Nasopharyngeal Carcinoma Identification Using Machine Learning Techniques. Neural Computing and Applications, 2020, 32(3), 625-638 (26-1). https://doi.org/10.1007/s00521-018-3882-6
2. Canny, J. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, (6), 679-698. https://doi.org/10.1109/TPAMI.1986.4767851
3. Dalal, N., Triggs, B. Histograms of Oriented Gradients for Human Detection. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). IEEE, 2005, 1, 886-893. https://doi.org/10.1109/CVPR.2005.177
4. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv preprint arXiv: 2010, 11929, 2020.
5. Fukunaga, K., Hostetler, L. The Estimation of the Gradient of a Density Function with Applications in Pattern Recognition. IEEE Transactions on Information Theory, 1975. 21(1), 32-40. https://doi.org/10.1109/TIT.1975.1055330
6. Geirhos, R., Rubisch, P., Michaelis, C., Bethge, M., Wichmann, F. A., Brendel, W. Imagenet-Trained Cnns Are Biased Towards Texture, Increasing Shape Bias Improves Accuracy and Robustness. arXiv preprint arXiv:1811.12231, 2018.
7. Goodfellow, I., Shelens, J., Szegedy, C. Explaining and Harnessing Adversarial Examples. In International Conference Learning Representations. 2015.
8. Gordo, A., Almazán, J., Revaud, J., Larlus, D. Deep Image Retrieval: Learning Global Representations for Image Search. In European Conference on Computer Vision. Springer, 2016, 9910, 241-257. https://doi.org/10.1007/978-3-319-46466-4_15
9. Harris, C., Stephens, M. A Combined Corner and Edge Detector. In Alvey Vision Conference. Manchester, UK, 1988, 15 (50), 10-5244. https://doi.org/10.5244/C.2.23
10. He, K., Zhang, X., Ren, S., Sun, J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), v2016. https://doi.org/10.1109/CVPR.2016.90
11. Ibrahim, D. A., Zebari, D. A., Mohammed, H. J., Mohammed, M. A. Effective Hybrid Deep Learning Model for Covid-19 Patterns Identification Using Ct Images. Expert Systems, 2022, e13010 DOI: 10.1111/exsy.13010. https://doi.org/10.1111/exsy.13010
12. Kass, M., Witkin, A., Terzopoulos, D. Snakes: Active Contour Models. International Journal of Computer Vision, 1988, 1 (4), 321-331. https://doi.org/10.1007/BF00133570
13. Krizhevsky, A., Sutskever, I., Hinton, G. E. Imagenet Classification with Deep Convolutional Neural Networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, Nevada, 2012, (10-2), 1097-1105.
14. Lal, S., Rehman, S. U., Shah, J. H., Meraj, T., Rauf, H. T, Damaševičius, R., Mohammed, M. A., Abdulkareem, K.

H. Adversarial Attack and Defence through Adversarial Training and Feature Fusion for Diabetic Retinopathy Recognition. Sensors, 2021, 21 (11), 3922. https://doi.org/10.3390/s21113922

15. Liu, Y., Cheng, M.-M., Fan, D.-P., Zhang, L., Bian, J.-W., Tao, D. Semantic Edge Detection with Diverse Deep Supervision. International Journal of Computer Vision, 2022, 130 (1), 179-198. https://doi.org/10.1007/s11263-021-01539-8

16. Lowe, D. G. Object Recognition from Local Scale-Invariant Features. In Proceedings of the Seventh IEEE International Conference on Computer Vision. IEEE, 1999, 2, 1150-1157. https://doi.org/10.1109/ICCV.1999.790410

17. Lowe, D. G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 2004, 60 (2), 91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

18. Marr, D., Hildreth, E. Theory of Edge Detection. Proceedings of the Royal Society of London. Series B. Biological Sciences, 1980, 207 (1167), 187-217.https://doi.org/10.1098/rspb.1980.0020

19. Martin, D. R., Fowlkes, C. C., Malik, J. Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26 (5), 530-549. https://doi.org/10.1109/TPAMI.2004.1273918

20. Mastoi, Q.-U.-A., Wah, T. Y., Mohammed, M. A., Iqbal, U., Kadry, S., Majumdar, A., Thinnukool, O. Novel Derma Fusion Technique for Ecg Heartbeat Classification. Life, 2022, 12 (6), 842. https://doi.org/10.3390/life12060842

21. Ojala, T., Pietikainen, M., Maenpaa, T. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24 (7), 971-987. https://doi.org/10.1109/TPAMI.2002.1017623

22. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9 (1), 62-66. https://doi.org/10.1109/TSMC.1979.4310076

23. Paris, S., Durand, F. A Fast Approximation of the Bilateral Filter Using a Signal Processing Approach. In European conference on computer vision. Springer, 2006, 3954, 568-580. https://doi.org/10.1007/11744085_44

24. Prewitt, J. M. Object Enhancement and Extraction. Picture Processing and Psychopictorics, 1970, 10 (1), 15-19.

25. Roberts, L. Machine Perception of Three-Dimensional Solids. In Optical and Electro-Optical Formation Processing, Optical and Electro-Optical Information Processing, J. T. Tippett, et al., Eds., May 1965.

26. Rosten, E., Drummond, T. Machine Learning for High-Speed Corner Detection. In European Conference on Computer Vision. Berlin, Springer, 2006, 3951, 430-443. https://doi.org/10.1007/11744023_34

27. Rosten, E., Porter, R., Drummond, T. Faster and Better: A Machine Learning Approach to Corner Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 32 (1), 105-119 DOI: 10.1109/TPAMI.2008.275. https://doi.org/10.1109/TPAMI.2008.275

28. Salvador, A., Giró-i-Nieto, X., Marqués, F., Satoh, S. I. Faster R-Cnn Features for Instance Search. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, 9-16. https://doi.org/10.1109/CVPRW.2016.56

29. Simonyan, K., Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409, 1556, 2014

30. Smeulders, A. W., Worring, M., Santini, S., Gupta, A., Jain, R. Content-Based Image Retrieval at the End of the Early Years. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22, (12), 1349-1380. https://doi.org/10.1109/34.895972

31. Smith, S. M., Brady, J.M. Susan-A New Approach to Low-Level Image Processing. International Journal of Computer Vision, 1997, 23 (1), 45-78. https://doi.org/10.1023/A:1007963824710

32. Sobel, I., Feldman, G. A 3× 3 Isotropic Gradient Operator for Image Processing. Duda, R. and Hart, P. (Eds.). Pattern Classification and Scene Analysis, 1973, 271-272.

33. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A. A. Inception-V4, Inception-Resnet and the Impact of Residual Connections on Learning. In Thirty-First AAAI Conference on Artificial Intelligence, 2017, 3951. https://doi.org/10.1609/aaai.v31i1.11231

34. Tramer, F., Carlini, N. Brendel, W., Madry, A. On Adaptive Attacks to Adversarial Example Defenses. Advances in Neural Information Processing Systems, 2020, 33, 1633-1645.

35. Tsai, C.-F. Bag-of-Words Representation in Image Annotation: A Review. International Scholarly Research Notices, 2012. https://doi.org/10.5402/2012/376804

36. Wang, M., Deng, W. Deep Visual Domain Adaptation: A Survey. Neurocomputing, 2018, 312. https://doi.org/10.1016/j.neucom.2018.05.083

37. Xie, S., Tu, Z. Holistically-Nested Edge Detection. In Proceedings of the IEEE International Conference on Computer Vision, 2015, 1395-1403. https://doi.org/10.1109/ICCV.2015.164

38. Young, R. The Gaussian Derivative Model for Spatial Vision. I- Retinal Mechanisms. Spatial Vision, 1987, 2 (4), 273-293. https://doi.org/10.1163/156856887X00222

39. Yu, W., Yang, K., Yao, H., Sun, X., Xu, P. Exploiting the Complementary Strengths of Multi-Layer Cnn Features for Image Retrieval. Neurocomputing, 2017, 237, 235-241. https://doi.org/10.1016/j.neucom.2016.12.002

40. Zebari, D. A., Ibrahim, D. A., Zeebaree, D. Q., Mohammed, M. A., Haron, H., Zebari, N. A., Damaševičius, R., Maskeliūnas, R. Breast Cancer Detection Using Mammogram Images with Improved Multi-Fractal Dimension Approach and Feature Fusion. Applied Sciences, 2021, 11 (24), 12122. https://doi.org/10.3390/app112412122

41. Zhao, M., Jha, A. Liu, Q. Millis, B. A., Mahadevan-Jansen, A., Lu, L., Landman, B. A., Tyska, M. J., Huo, Y. Faster Mean-Shift: Gpu-Accelerated Clustering for Cosine Embedding-Based Cell Segmentation and Tracking. Medical Image Analysis, 2021, 71, 102048. https://doi.org/10.1016/j.media.2021.102048

42. Zheng, Q., Yang, M., Tian, X., Jiang, N., Wang, D. A Full Stage Data Augmentation Method in Deep Convolutional Neural Network for Natural Image Classification. Discrete Dynamics in Nature and Society, 2020, https://doi.org/10.1155/2020/4706576

43. Zheng, Q., Yang, M., Yang, J., Zhang, Q., Zhang, X. Improvement of Generalization Ability of Deep Cnn Via Implicit Regularization in Two-Stage Training Process. IEEE Access, 2018, 6, 15844-15869. https://doi.org/10.1109/ACCESS.2018.2810849

44. Zheng, Q., Zhao, P., Li, Y. Wang, H., Sang, Y. Spectrum Interference-Based Two-Level Data Augmentation Method in Deep Learning for Automatic Modulation Classification. Neural Computing and Applications, 2021, 33 (13), 7723-7745. https://doi.org/10.1007/s00521-020-05514-1

45. Zheng, Q., Zhao, P., Zhang, D., Wang, H. Mr-Dcae: Manifold Regularization-Based Deep Convolutional Autoencoder for Unauthorized Broadcasting Identification. International Journal of Intelligent Systems, 2021, 36 (12), 7204-7238 https://doi.org/10.1002/int.22586