Journal of Information Security and Applications 66 (2022) 103153

Contents lists available at ScienceDirect

Journal of Information Security and Applications

journal homepage: www.elsevier.com/locate/jisa

Virtual sample generation for few-shot source camera identification

Bo Wang [a], Shiqi Wu [a,*], Fei Wei [b], Yue Wang [a], Jiayao Hou [a], Xue Sui [c]

[a] School of Information and Communication Engineering, Dalian University and Technology, Dalian, Liaoning, 116024, PR China
[b] Department of Electrical Engineering, Arizona State University, Tempe, AZ 85281, USA
[c] College of Psychology, Liaoning Normal University, Dalian, Liaoning, 116029, PR China

ARTICLE INFO

Keywords:
Source Camera Identification (SCI)
Few-shot
Virtual sample
Ensemble learning
Semi-supervised

ABSTRACT

The Source Camera Identification (SCI) has achieved remarkable success. However, existing approaches require sufficiently large training sets for high performance on accuracy and robustness. For maintaining high performance given small training sets, we propose a semi-supervised, Mega-Trent-Diffusion (MTD) method to generate virtual samples, such that the training sets can be expanded and unlabeled samples can be fully utilized as well. The stability of our method is improved using ensemble learning. Our theoretical analysis and experiments corroborate the effectiveness of our method beyond others when few-shot is given.

1. Introduction

Digital images are widely used and numerous applications in market can easily modify or process images. However, concerns on the authenticity of digital images arise in cases, e.g., digital images can be taken as evidence or proof for legal issues. Thus, digital image forensics are attracting growing attentions in recent years [1,2].

Researches on digital image forensics are conducted mainly on two aspects: active forensics, which use active means to secure image information, and passive forensics, which perform forensic analysis on digital images without actively adding discriminative information in advance. Passive forensics are more practical in the sense that it does not require pre-processing. In this work, we propose a SCI method based on few-shot in the context of passive forensics approach.

Digital image fingerprints generated by manufacturing imperfections (within the same model) and by structural differences (cross models) are used as features for SCI. There are device-based, model-based and individual-based SCI studies and techniques been proposed in literature [1,3][4]. The specific works are as follows.

First, Device-based SCI determines the source device type of given images, such as cameras, mobile phones, or computers. Lyu et al. [1] proposed to detect device model based on orthogonal mirror filtering. Lyu [2] uses mirror filter to distinguish the statistical characteristics in the frequency domain presented in different directions and angles.

Next, Model-based SCI identifies source camera models of given images. Kharrazi et al. [3] use statistical features, i.e., characteristics of color correlation, color energy ratio, and neighborhood distribution centroid, for identification. Lyu et al. [5] use image quality features and Avcibas et al. [6] use wavelet features to identify camera models. Meng et al. [7] introduce the double-popular feature and optimize the feature selecting method, with an average accuracy rate of 90%. There are researchers who approach from the perspective of lens light distortion. Choi et al. [8,9] propose to describe lens distortion by detecting linear distortion in images. In addition, the CFA interpolation feature generated in the imaging process can be used for SCI. Farid et al. find that the CFA interpolation operation causes correlation between pixels, that is, a peak point of energy appears in the spectrogram. In addition, Swaminathan et al. [10,11]. propose to use a linear interpolation model to estimate the neighborhood CFA interpolation coefficients This method achieves an average accuracy of 85.9% for 19 camera models' images from 9 brands.

Beyond camera device and model identifications, individual-based SCI can be accurate to the source camera individual of a given image. Based on amount and location difference of sensor defect points, Geradts et al. [4] propose to average multiple images to detect image defects produced by CCD modules. Yang et al. [12] applied CCD fingerprint identification technology to the source camera identification field of online social network, which improved the trace ability application field of digital image equipment.

In our work, we focus on identifying the camera model. The framework of source forensics method is depicted in Fig. 1. Existing works, e.g., Xu et al. [13] and Qiao et al. [14], the accuracy and effectiveness relying on sufficiently large training sets. For example, Xu et al. [13] propose to use LBP features to identify camera models, and the number of training samples for each camera model is 150–300. Memon et al. [3]

* Correspondence to: Haishan Building, Dalian University of Technology, Dalian City, Liaoning Province, PR China.
E-mail addresses: bowang@dlut.edu.cn (B. Wang), 57961030@mail.dlut.edu.cn (S. Wu), fwei16@asu.edu (F. Wei), 31909072@mail.dlut.edu.cn (Y. Wang), 32009157@mail.dlut.edu.cn (J. Hou), suixue@lnnu.edu.cn (X. Sui).

https://doi.org/10.1016/j.jisa.2022.103153

Available online 25 March 2022

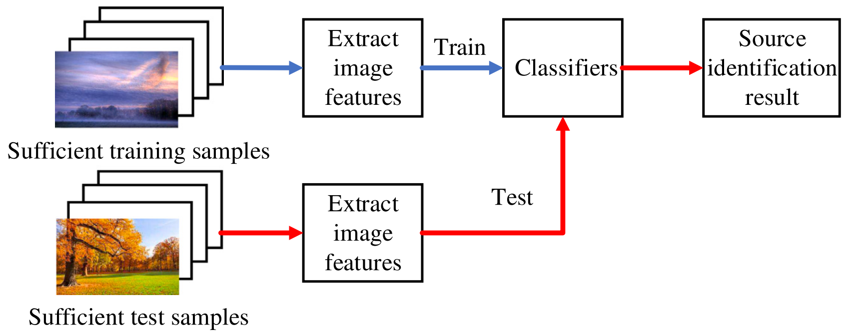

Fig. 1. The framework of source forensics method.

propose a feature set method for SCI, and select 150 images from each model as training samples . Qiao et al. [14] use a Gaussian noise model for SCI, and 100 samples are selected as training samples from each camera model. Roy et al. [15] use a extraction of the discrete cosine transform residual features for SCI, and 100 training samples are selected from each camera model.

However, when the number of training samples decreases, the SVM classifier or other machine learning algorithms cannot be sufficiently trained [16], which will cause the accuracy of identification decreases significantly. In the actual source forensics work, due to the difficulty of obtaining samples, the number of training samples is often far from meeting the needs of identification. Therefore, maximizing the accuracy with insufficient training samples is an urgent issue and the major concern of our work.

In this work, we make the following contributions:

- We propose a MTD method to generate virtual samples as labeled samples. In addition, we introduce ensemble learning to improve the stability of virtual samples.
- We fully utilize the unlabeled samples to improve the accuracy of SCI by semi-supervised learning given few-shot.
- Our experiment results demonstrate the superiority of our method given few-shot.

The rest of this paper is structured as follows. In Section 2, we introduce the related models and definitions. In Section 3, we propose a semi-supervised ensemble learning of virtual samples method for SCI. In Section 4, we demonstrate our experiment setup and discuss the results. We summarize our work in Section 5.

2. Models and definitions

2.1. Local binary pattern

Local Binary Pattern (LBP) is a local operator which describes the image texture features. It has significant advantages, e.g., grayscale invariance, rotation invariance, simple calculation and so on, such that it has been widely used in many fields of computer vision [17,18]. The uniform gray-scale invariant LBP [13] operator is denoted as $LBP_{P,R}$ and defined as follows:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s\left(g_p - g_c\right)2^p \tag{1}$$

where P represents the number of pixels in the neighborhood and R represents its radius. The value of center pixel is denoted as g_c, and the values of the neighborhood pixels in a circle with a radius R are denoted as $\{g_p : p \in [P]\}$. The threshold function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{2}$$

Refer to Eqs. (1) and (2), if we set $P = 8$, $R = 1$, the difference between the gray-level value of the center pixel and its neighborhood pixels can be calculated, and there are 256 different patterns of combination. Xu et al. [13] show that "uniform" local binary patterns are more likely to appear than "non-uniform" local binary patterns. They also show that all "non-uniform" local binary patterns can be integrated into one pattern such that the dimension of features reduces from 256 to 59.

As depicted in Fig. 2, the original image is estimated by the prediction function to obtain the predicted image, and the prediction-error image is obtained by subtracting a predicted image from the original one. We finally extract LBP features from the red and green channels of the original image, the 1st-level diagonal wavelet subband, and prediction-error 2D array [19]. Thus a total dimension of 59×3×2 = 354 improved LBP features are obtained.

2.2. Sample attributes correlation based mega-trent-diffusion (MTD)

Based on global fuzzification and information diffusion, Li et al. [20] introduced the Mega-Trent-Diffusion (MTD) method to fill the blank caused by incomplete samples. Sample attribute-correlation based MTD method considers the correlation between samples attributes when determining the diffusion range of the labeled samples. Li et al. consider the correlation of sample attributes when calculating the range [21]. Taking the correlation of sample attributes into account, the center of labeled samples is calculated by Eq. (3), which is used to calculate the data center of the samples.

$$CL = \begin{cases} \frac{x_{n/2}+x_{n/2+1}}{2}, & n \text{ is even} \\ x_{(n+1)/2}, & n \text{ is odd} \end{cases} \tag{3}$$

Considering the mean is susceptible to outliers, Li et al. decide to use the median of labeled samples instead of the mean to calculate the correlation of sample attributes. As shown in Eq. (3), x represents the labeled samples, n represents the number of the labeled samples. When n is even or odd, a different equation is used to calculate CL. The median of labeled samples differs greatly from the data population in the case of insufficient samples, and the sample median is needed to calculate the correlation of sample attributes. So we replace the correlation with the trend similarity between attributes (TSA):

$$g(h)_{i,j} = \begin{cases} 1, (x_i - CL)(x_j - CL) > 0 \\ 0, (x_i - CL)(x_j - CL) = 0 \\ -1, (x_i - CL)(x_j - CL) < 0 \end{cases} \tag{4}$$

$$S_{i,j} = \frac{1}{k}\sum_{h=1}^{k} g(h)_{i,j}, i \neq j \tag{5}$$

where $g(h)_{i,j}$ represents degree of similarity between the two attributes' dimensions of labeled sample h. If $g(h)_{i,j} = 1$, the similarity trend between the two attributes' dimensions is extremely high, that is, both are on the same side of the center point CL. $S_{i,j}$ represents the similarity between different attributes' dimensions of the labeled samples, and k represents the number of samples in the same class. In order to avoid using mean which is susceptible to outliers, we use

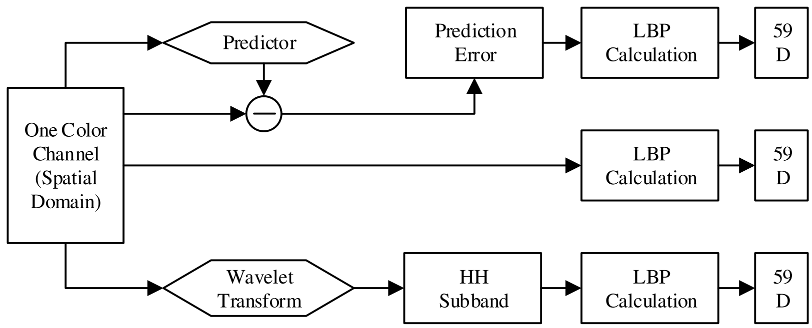

Fig. 2. LBP feature extraction framework for one color channel.

Euclidean distance instead of sample standard deviation to measure the dispersion as shown in Eq. (6).

$$d_i = \sqrt[2]{\sum (x - CL)^2 / k} \tag{6}$$

Incorporating the correlation of sample attributes, the bounds for samples after diffusion are:

$$\begin{aligned} B_i^L &= CL_i - \min_{j \neq i}\{d_i^L \sqrt{-2\ln(\varphi(S_{i,j}))}\} \\ B_i^U &= CL_i + \min_{j \neq i}\{d_i^U \sqrt{-2\ln(\varphi(S_{i,j}))}\} \\ \varphi(S_{i,j}) &= \frac{1}{1+\exp(-10|S_{i,j}|^{-0.5})} \end{aligned} \tag{7}$$

where d_i^L and d_i^U represent the average Euclidean distance of the labeled samples which are less and greater than the data center value respectively. So we get the virtual sample generation range $\left[B_i^L, B_i^U\right]$.

2.3. Semi-supervised learning

Semi-Supervised Learning (SSL) is a method between supervised learning and unsupervised learning [22]. Normally, unlabeled samples are more convenient and less resource-consuming to acquire comparing with labeled samples. Traditional classification methods such as support vector machine (SVM) and neural network often require sufficient labeled samples for training to achieve higher generalization ability [13,14]. Improving the classification performance with few-shot is an urgent problem yet to be solved. When there is few-shot, the SSL method is used to train the sample, then adds unlabeled samples to the original training set (labeled samples) to improve the performance of the classification method according to certain criteria [23].

Semi-supervised Ensemble learning is first proposed in [24], which uses ensemble learning and the semi-supervised processing of unlabeled samples to maximize the interval between decision boundaries. Zhou et al. [25] confirm that semi-supervised processing and ensemble learning can be mutually beneficial. In terms of ensemble learning, sometimes training samples are insufficient. Here, the semi-supervised method can supply the training samples, and increase the diversity of the base learners in ensemble learning. In turn, ensemble learning can improve the semi-supervised learning speed and improve the generalization ability of semi-supervised learning.

Bagging method is a classical method in ensemble learning [26]. Li et al. [27] propose a semi-supervised bagging method based on this, as shown in Fig. 3. The *pseudo lable* is given by Eq. (8).

$$\begin{aligned} &pseudo\ label = y_k \\ &s.t.\ \arg\max_k \sum_{i=1}^{n} \delta\left(y_{A_i} = y_k\right) \end{aligned} \tag{8}$$

where $y_{A_i}, i \in [1, n]$ are the prediction results of different semi-supervised learners A_i. The labels of the m different classes are defined as $y_k, k \in [1, m]$. The δ-function, which is used to count the number of a specific prediction result, is shown in Eq. (9).

$$\delta\left(x = x_1\right) = \begin{cases} 1, x = x_1 \\ 0, x \neq x_1 \end{cases} \tag{9}$$

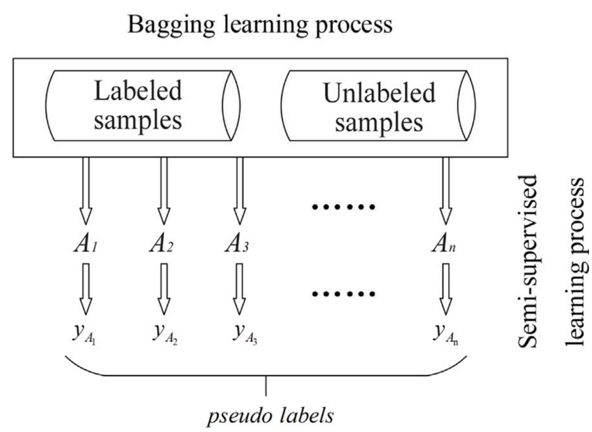

Fig. 3. Semi-supervised bagging method.

The specific steps of semi-supervised bagging method are as follows.

(1) KNN is used as the basic classifier to train the basic classifier based on all labeled samples. Using bagging method, unlabeled samples are extracted from the original samples and sent to the basic classifier for training. A total of n groups of samples are extracted and trained to obtain n groups of pseudo labels;

(2) Using the semi-supervised method, all labeled samples and n groups of pseudo labeled samples are combined into n groups of new training samples to train n semi-supervised classifiers A_i;

(3) Finally, the remaining unlabeled samples are sent to the n semi-supervised classifiers for testing, and the predicted label results of the n semi-supervised classifiers are y_{A_i}, where different predicted labels are marked as y_k. By using the δ-function to count the times of these different prediction labels, we can select the prediction labels with the most statistical times as the final labels. Then we can get the final classification result.

2.4. Ensemble learning

Ensemble learning is a form of "expanding others" by constructing multiple weak classifiers and combining them into a powerful classifier to effectively accomplish their tasks, it comprehensively determines the learning results by combine multiple learners obtained through training [28]. A weak classifier has a slightly better classification effect than a random guess. Since the classification results of a single base classifier are often not up to the ideal classification standard, the weak classifiers obtained from the initial training are usually assembled when the base classifier is used [29]. By learning and complementing each other, the classifiers can ultimately improve the accuracy of the classification and achieve the desired experimental results, which is shown in Fig. 4.

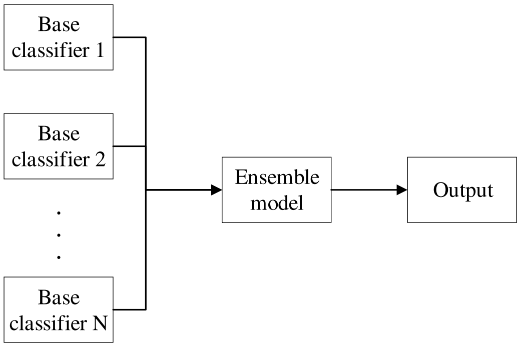

Fig. 4. Ensemble learning model.

According to current results, ensemble learning methods can be divided into "Homogeneous" and "Heterogeneous" ensemble learning. "Homogeneous" ensemble learning constructs the base learner with learning algorithms of only one type. For example, in our method all the base learners utilize the support vector machine (SVM). In addition, the learning algorithms for regression, such as decision trees, etc., can also be used in "Homogeneous" ensemble learning. "Heterogeneous" ensemble learning can be considered as a broad definition. In order to achieve the same goal, various learning algorithms can be used to construct the base learner. There are many strategies for ensemble classifiers, such as averaging, voting, and weighted average method [30]. For the prediction of classification problems, the voting method is simple and effective. In addition, the result of our base classifier is the class of the image. So we choose the voting method in our work. The voting method is also "Majority rule", which means if there are n classifiers, the class obtained by a large number of classifiers is the final class of the image.

3. Proposed method

We know that semi-supervised ensemble learning can solve the problem of inaccurate classification with few-shot. By filtering the appropriate unlabeled samples into existing training sets under certain conditions, one can increase the number of training samples and reduce the number of samples required for ensemble learning. We propose a virtual sample generation method based on semi-supervised ensemble learning to improve the accuracy of SCI with few-shot, which is depicted in Fig. 5, and we named it MTD-SEM.

We extend the training set by generating virtual samples based on MTD method. The upper and lower limits of the specific boundary for virtual samples generation are determined by the distribution of the existing training samples. Similarly, after determining the range $[B_i^L, B_i^U]$, the virtual samples' values are generated based on the uniform distribution within $[B_i^L, B_i^U]$. Considering the randomness of generating virtual samples in the range $[B_i^L, B_i^U]$ according to uniform distribution is too strong, so we use the Triangular Membership Function to judge the possibility of the generated virtual samples [31].

In Fig. 6, MF represents the distribution of samples, and tp is the virtual sample values randomly generated in $[B_i^L, B_i^U]$ according to uniform distribution. Based on the generated tp value, the corresponding MF value is calculated according to Eq. (10), and the MF value is used as the possibility of its occurrence. At the same time, random number r is generated on [0,1] according to uniform distribution. If $r < MF$, the generated virtual sample tp is more likely to exist and can be retained as a suitable virtual sample. Repeat the above process until enough virtual samples are generated.

$$MF(tp) = \begin{cases} \frac{tp-L}{CL-L}, & L \le tp \le CL \\ 0, & tp < L, tp > U \\ \frac{U-tp}{U-CL}, & CL < tp \le U \end{cases} \tag{10}$$

where L is B_i^L, U is B_i^U in Eq. (10).

Besides, our method is based on the semi-supervised Bagging method, which can be divided into two phases: ensemble learning and semi-supervised ensemble learning. The flow chart of the method is shown in Figs. 3 and 4. First, we extract each class samples' LBP features, and the training samples are randomly extracted from each class by 5, 10, 15, 20 and 25 samples. Then, based on the training samples, 5 sets of virtual samples are generated using sample attributes correlation based MTD method. The virtual samples are randomly generated in the range based on the trend diffusion theory, and have strong randomness, which leads that the classification results of the base classifier is unstable. Therefore, we generate 5 sets of virtual samples based MTD method to improve the difference between the base classifiers and improve their generalization ability. Then these sets of virtual samples are added to the original sample set as the new 5 sets of training sets. Based on the new training sets, 5 base classifiers (SVMs) can be trained. The ensemble classifier 1 is obtained by integrating 5 base classifiers using a relative majority voting method.

The semi-supervised method can select a certain number of unlabeled samples as training samples. In this paper, we use posterior probability as the standard for selecting samples. Each unlabeled sample gets its posterior probability belonging to each class by ensemble classifier 1. The greater the posterior probability, the greater the probability that the sample belongs to this class. So we arranged the samples' posterior probabilities of each class in descending order, and the first few unlabeled samples with the highest posterior probability in each class are retained. The corresponding class is the pseudo label of the unlabeled sample.

The semi-supervised ensemble learning method is different from the semi-supervised method in that the selected unlabeled samples are not directly added to the training classifier in the original training sample, but the selected unlabeled samples are used as the new training samples first, and generate new virtual samples based MTD method. Considering the instability of the virtual samples generated by the MTD method, similarly, the new training samples are added to the original training set together with the new virtual samples to obtain a new training samples. Finally, the final SCI classifier is obtained again based ensemble learning, named as the ensemble classifier 2.

4. Experiments

In this section, the SCI experiment based on the semi-supervised ensemble learning of virtual samples is carried out on the image database, which is commonly used in forensics. The experimental results of this method are compared with the SCI results based on virtual sample generation method and virtual sample ensemble learning method, etc.

4.1. Experimental setup

In order to ensure the reliability of the experimental results, the well-known "Dresden Image Database" [32] is used in this paper. The images taken by 16 models are arbitrarily selected as experimental samples. As shown in Table 1, each camera model has 180 image samples and the resolution of each camera is given. Blocks of size 512 × 512 are intercepted from the center of each image. Features used in our experiments are LBP features. A total of 180 samples per camera, of which the training samples range from 5 to 25, and the remaining samples that exclude the training samples are selected as the test sample. The experiments are repeated 20 times and the results are averaged to remove the effect of randomness.

4.2. Results and discussion

In the first part of the experiment, we need to assign pseudo labels to unlabeled samples in the semi-supervised process. Therefore, for

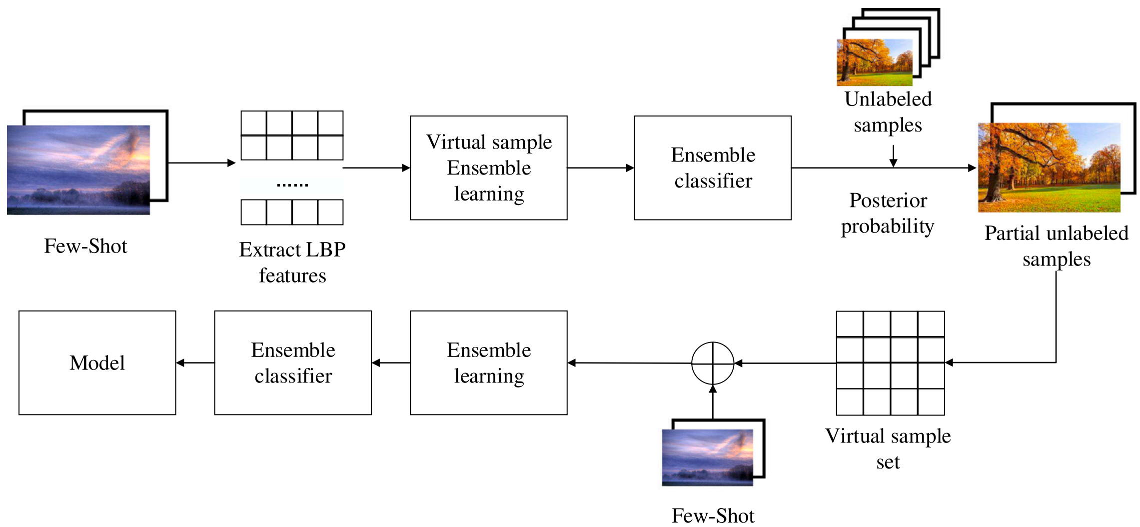

Fig. 5. Training flow chart of source camera identification based on virtual sample semi-supervised ensemble learning.

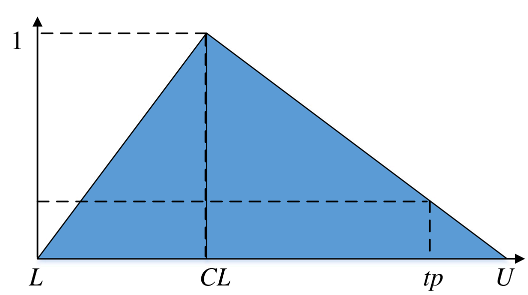

Fig. 6. The MF value of virtual sample (tp).

Table 1
Dataset in experiments.

Camera module	Abbr.	Size
Canon_Ixus70	C1	3072 × 2304
Casio_EX-Z150	C2	3264 × 2448
FujiFilm_FinePixJ50	F1	3264 × 2448
Kodak_M1063	K1	3664 × 2748
Nikon_CoolPixS710	N1	4352 × 3264
Nikon_D70	N2	3008 × 2000
Nikon_D200	N3	3872 × 2592
Olympus_mju_1050SW	O1	3648 × 2736
Panasonic_DMC-FZ50	P1	3648 × 2736
Praktica_DCZ5.9	P2	2560 × 1920
Rollei_RCP-7325XS	R1	3072 × 2304
Samsung_L74wide	S1	3072 × 2304
Samsung_NV15	S2	3648 × 2736
Sony_DSC-H50	SD1	3456 × 2592
Sony_DSC-T77	SD2	3648 × 2736
Sony_DSC-W170	SD3	3648 × 2736

Table 2
The five highest pseudo label accuracy of unlabeled samples.

Number	5	10	15	20	25
Accuracy	90.00%	94.03%	96.11%	96.25%	97.92%

each class of samples, we selected five unlabeled samples with the highest posterior probability of assigning pseudo labels as new training samples. Finally, The distribution accuracy of pseudo labels is shown in Table 2.

The results show that as the number of training samples increases, the accuracy of the judgment gradually increases. When the number of samples is 5, 10, 15, 20, and 25, respectively, the corresponding accuracy is 90.00%, 94.03%, 96.11%, 96.25%, 97.92%, and the lowest judgment accuracy is also 90.00%, which shows that the unlabeled samples we screened are suitable.

Table 3
Average accuracy of source camera identification.

Algorithm	5	10	15	20	25
LBP	45.64%	71.24%	78.97%	83.29%	85.45%
MTD	49.74%	71.93%	80.15%	84.32%	86.94%
MTD-EM	50.10%	71.33%	80.00%	84.07%	87.04%
SEMI	52.24%	72.22%	80.45%	84.35%	86.99%
MTD-SEM	61.07%	74.11%	81.22%	84.73%	87.52%

Table 3 shows the accuracy of image source identification in different situations. Without the virtual samples (LBP), the accuracy of SCI is 45.64%, 71.24%, 78.97%, 83.29%, and 85.45%, which is achieved based on the image source identification model trained with the raw dataset of 5, 10, 15, 20, and 25 training samples respectively. In order to make full use of the labeled sample information, we introduce virtual samples generated by MTD method (MTD). The accuracy of SCI slightly improves, and when the number of samples is 5, the improvement of accuracy is maximized, which is 4.10%. Due to the instability of the virtual samples generated by the MTD method, we subsequently apply the ensemble learning method (MTD-EM) proposed in this paper [33] and the accuracy of SCI is slightly improved. The final experimental results also prove that our results are better than their methods in the same experimental configuration.

In order to verify the effectiveness of semi-supervised method, we conduct a separate set of semi-supervised experiments (SEMI). When the number of samples is 5, 10, the accuracy is 52.24% and 72.22%, which improves significantly. In the end, we combine virtual samples based on ensemble learning with semi-supervised (MTD-SEM). When the number of sample ranges from 5 to 25, the classification accuracy of the MTD-SEM method is the highest, and the accuracy far surpasses the baseline of LBP method, which is 15.43%, 2.87%, 2.25%, 1.44%, 2.07%.

When the number of labeled samples is 25, the average confusion matrix of MTD-SEM method obtained by SVM classification over 20 iterations is shown in Table 4. The asterisks in the table represent the classification probability below 0.01%. As demonstrated, our MTD-SEM method achieves a highest accuracy of 94.98% for SD2 camera, and we find that nearly half accuracy of SCI is up to 92.00%. Except for the SD1 camera and the SD3 camera, most other cameras accuracy of SCI

Table 4
Average confusion matrix obtained by SVM classification over 20 iterations.

Average TP=87.52%		Predicted															
		C1	C2	F1	K1	N1	N2	N3	O1	P1	P2	R1	S1	S2	SD1	SD2	SD3
Actual	C1	91.25%	2.44%	*	*	1.15%	0.50%	*	1.36%	0.22%	0.14%	0.14%	2.51%	0.29%	*	*	*
	C2	1.51%	91.18%	0.07%	0.07%	0.72%	*	*	0.29%	0.57%	0.57%	*	3.73%	*	0.43%	0.22%	0.65%
	F1	0.22%	*	91.04%	0.14%	0.07%	0.29%	*	0.29%	0.93%	3.66%	1.51%	0.36%	1.36%	0.07%	*	0.07%
	K1	*	0.22%	0.22%	89.18%	0.57%	0.57%	4.73%	0.29%	0.50%	0.93%	0.43%	0.07%	0.65%	1.15%	0.07%	0.43%
	N1	1.15%	4.09%	*	*	89.68%	1.00%	0.07%	0.07%	0.22%	1.72%	0.14%	1.86%	*	*	*	*
	N2	0.22%	*	0.65%	0.22%	0.50%	87.46%	5.02%	0.29%	*	0.57%	0.57%	2.80%	1.72%	*	*	*
	N3	0.29%	*	0.22%	2.94%	0.07%	1.72%	90.47%	*	*	*	0.14%	0.86%	*	1.08%	1.51%	0.72%
	O1	1.36%	0.14%	0.14%	0.14%	0.79%	0.36%	0.43%	91.04%	0.50%	0.22%	*	4.16%	*	0.36%	0.36%	*
	P1	0.07%	1.15%	0.50%	0.29%	*	0.07%	1.22%	0.36%	93.55%	0.43%	*	1.00%	0.79%	0.07%	0.07%	0.43%
	P2	0.29%	1.22%	1.36%	0.29%	0.14%	2.29%	*	*	0.22%	90.25%	1.15%	0.29%	2.29%	0.22%	*	*
	R1	*	0.14%	2.94%	0.07%	0.07%	1.94%	*	*	*	0.86%	91.33%	0.50%	1.51%	0.65%	*	*
	S1	0.72%	1.51%	0.79%	0.36%	0.29%	0.57%	0.22%	1.29%	0.65%	0.14%	0.50%	92.47%	0.43%	0.07%	*	*
	S2	*	*	0.65%	*	*	0.50%	*	*	*	2.65%	2.51%	0.22%	92.62%	0.79%	*	0.07%
	SD1	0.22%	*	0.36%	*	*	*	0.29%	*	0.43%	0.14%	0.43%	0.22%	0.43%	62.80%	2.01%	32.69%
	SD2	*	*	0.29%	*	*	*	0.14%	*	0.07%	0.07%	0.22%	0.36%	0.07%	3.08%	94.98%	0.72%
	SD3	*	*	0.22%	*	*	*	0.29%	*	0.43%	*	0.14%	*	0.14%	36.99%	0.79%	61.00%

is higher than 90.00%. The SD1 camera and the SD3 camera accuracy of SCI is 62.80% and 61.00%. These two cameras are confused with each other when classifying. The reason is the two models adopt similar image post-processing algorithm.

5. Conclusion

In this paper, we propose to combine a semi-supervised virtual sample generation method with the ensemble learning to identify image source camera given insufficient labeled samples. We use Mega-Trend-Diffusion (MTD) method to generate virtual sample, and apply ensemble learning to improve the instability. Through semi-supervised method we can make full use of unlabeled samples. Our experiments compare LBP method, MTD method, MTD-EM method and SEMI method. The results demonstrate that our proposed method MTD-SEM is superior to the existing methods given insufficient labeled samples. In future, we will further consider improving classification accuracy by introducing new features.

CRediT authorship contribution statement

Bo Wang: Conceptualization, Supervision, Project administration. **Shiqi Wu:** Data curation, Writing – original draft, Methodology. **Fei Wei:** Software, Validation. **Yue Wang:** Visualization, Investigation. **Jiayao Hou:** Writing – review & editing. **Xue Sui:** Writing – review & editing.

Declaration of competing interest

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

Acknowledgments

This work is supported by the National Natural Science Foundation of China (No. U1936117, No. 62106037, No. 62076052, No. 61772111), the Science and Technology Innovation Foundation of Dalian (No. 2021JJ12GX018), the Fundamental Research Funds for the Central Universities, China (DUT21GF303, DUT20TD110, DUT20RC(3)088), and the Open Project Program of the National Laboratory of Pattern Recognition (NLPR), China (No. 202100032).

References

[1] Lyu S, Farid H. How realistic is photorealistic? IEEE Trans Signal Process 2005;53(2):845–50.
[2] Lyu S. Natural image statistics for digital image forensics (Ph.D. thesis), Hanover, New Hampshire: Thesis of Doctor of Philosophy, Dartmouth College; 2005.
[3] Kharrazi M, Sencar HT, Memon N. Blind source camera identification. In: 2004 International conference on image processing, 2004. vol. 1, IEEE; 2004, p. 709–12.
[4] Geradts ZJ, Bijhold J, Kieft M, Kurosawa K, Kuroki K, Saitoh N. Methods for identification of images acquired with digital cameras. In: Enabling technologies for law enforcement and security. vol. 4232, International Society for Optics and Photonics; 2001, p. 505–12.
[5] Lyu S, Farid H. Detecting hidden messages using higher-order statistics and support vector machines. In: International workshop on information hiding. Springer; 2002, p. 340–54.
[6] Avcibas I, Memon N, Sankur B. Steganalysis using image quality metrics. IEEE Trans Image Process 2003;12(2):221–9.
[7] Meng F, Kong X, You X. A new feature-based method for source camera identification. In: IFIP international conference on digital forensics. Springer; 2008, p. 207–18.
[8] Choi KS, Lam EY, Wong KK. Automatic source camera identification using the intrinsic lens radial distortion. Opt Express 2006;14(24):11551–65.
[9] Choi KS, Lam EY, Wong KK. Source camera identification using footprints from lens aberration. In: Digital photography II. vol. 6069, 2006, p. 172–9.
[10] Swaminathan A, Wu M, Liu KR. Nonintrusive component forensics of visual sensors using output images. IEEE Trans Inf Forensics Secur 2007;2(1):91–106.
[11] Swaminathan A, Wu M, Liu KR. Non-intrusive forensic analysis of visual sensors using output images. In: 2006 IEEE international conference on acoustics speech and signal processing proceedings. vol. 5, IEEE; 2006, p. V.
[12] Yang W-C, Lin T-H. Source camera identification in LINE social network via CCD fingerprint. In: 2020 13th International congress on image and signal processing, biomedical engineering and informatics. IEEE; 2020, p. 322–6.
[13] Xu G, Shi YQ. Camera model identification using local binary patterns. In: 2012 IEEE international conference on multimedia and expo. IEEE; 2012, p. 392–7.
[14] Qiao T, Retraint F, Cogranne R, Thai TH. Source camera device identification based on raw images. In: 2015 IEEE international conference on image processing. IEEE; 2015, p. 3812–6.
[15] Roy A, Chakraborty RS, Sameer VU, Naskar R. Camera source identification using discrete cosine transform residue features and ensemble classifier. In: CVPR workshops. 2017, p. 1848–54.
[16] Tan Y, Wang B, Li M, Guo Y, Kong X, Shi Y. Camera source identification with limited labeled training set. In: International workshop on digital watermarking. Springer; 2015, p. 18–27.
[17] Wang B, Yin J, Tan S, Li Y, Li M. Source camera model identification based on convolutional neural networks with local binary patterns coding. Signal Process, Image Commun 2018;68:162–8.
[18] Wang B, Zhong K, Shan Z, Zhu M, Sui X. A unified framework of source camera identification based on features. Forensic Science International 2020;307:110109.
[19] Gärdenfors P. Conceptual spaces: the geometry of thought. MIT Press; 2004.
[20] Li D-C, Wu C-S, Tsai T-I, Lina Y-S. Using mega-trend-diffusion and artificial samples in small data set learning for early flexible manufacturing system scheduling knowledge. Comput Oper Res 2007;34(4):966–82.
[21] Li D-C, Lin W-K, Chen C-C, Chen H-Y, Lin L-S. Rebuilding sample distributions for small dataset learning. Decis Support Syst 2018;105:66–76.
[22] Zhu XJ. Semi-supervised learning literature survey. Tech. rep., University of Wisconsin-Madison Department of Computer Sciences; 2005.

[23] Le TB, Kim SW. On incrementally using a small portion of strong unlabeled data for semi-supervised learning algorithms. Pattern Recognit Lett 2014;41(MAY 1):53–64.
[24] Bennett KP, Demiriz A, Maclin R. Exploiting unlabeled data in ensemble methods. In: Proceedings of the eighth ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2002, p. 289–96.
[25] Zhou Z-H. When semi-supervised learning meets ensemble learning. Front Electr Electron Eng China 2011;6(1):6–16.
[26] Skurichina M, Duin RP. Bagging, boosting and the random subspace method for linear classifiers. Pattern Anal Appl 2002;5(2):121–35.
[27] Li Y, Su L, Chen J, Yuan L. Semi-supervised question classification based on ensemble learning. In: International conference in swarm intelligence. Springer; 2015, p. 341–8.
[28] Dietterich TG, et al. Ensemble learning. In: The handbook of brain theory and neural networks. vol. 2, Massachusetts: MIT Press Cambridge; 2002, p. 110–25.
[29] Wang XZ, Xing HJ, Li Y, Hua Q, Dong CR, Pedrycz W. A study on relationship between generalization abilities and fuzziness of base classifiers in ensemble learning. IEEE Trans Fuzzy Syst 2015;23(5):1638–54.
[30] Rokach L. Ensemble-based classifiers. Artif Intell Rev 2010;33(1–2):1–39.
[31] Lin L-S, Li D-C, Pan C-W. Improving virtual sample generation for small sample learning with dependent attributes. In: 2016 5th IIAI international congress on advanced applied informatics. IEEE; 2016, p. 715–8.
[32] Gloe T, Bhme R. The dresden image database for benchmarking digital image forensics. 2010, ACM.
[33] Wu S, Wang B, Zhao J, Zhao M, Zhong K, Guo Y. Virtual sample generation and ensemble learning based image source identification with small training samples. Int J Digit Crime and Forensics 2021;13(3):34–46.