On the Performance of Li's Unsupervised Image Classifier and the Optimal Cropping Position of Images for Forensic Investigations

Ahmad Ryad Soobhany, Keele University and Forensic Pathways Ltd., UK

Richard Leary, Forensic Pathways Ltd., UK

KP Lam, Keele University, UK

ABSTRACT

Images from digital imaging devices are prevalent in society. The signatures of these images can be extracted as sensor pattern noise (SPN) and classified according to their source devices. In this paper, the authors assess the reliability of an unsupervised classifier for forensic investigation of digital images recovered from storage devices and to identify the best position for cropping the images before processing. Cross validation was performed on the classifier to assess the error rate and determine the effect of the size of the sample space and the classifier trainer on the performance of the classifier. Moreover, the authors find that the effect of saturation and subsequently the contamination of the SPN in the images affected performance negatively. To alleviate the negative performance, the authors identify the areas of images where less contamination occurs to perform cropping.

Keywords: *Camera Identification, Cross Validation, Digital Image Forensics, Image Classification, Image Cropping, Sensor Pattern Noise*

INTRODUCTION

Digital imaging devices, like digital cameras or mobile phones, are widespread everywhere in society. The images from these devices can be used in the commission of crime. If a group of digital images are extracted from a storage device, for example from a laptop or a phone, it will be helpful for forensic investigators to be able to identify and classify these images according to the source device that created the images. Digital image forensics is an emerging research field that investigates images to identify the source device or to find if the image

DOI: 10.4018/jdcf.2011010101

has been tampered with or undergone any post processing (Sencar & Memon, 2007; Amerini, Caldelli, Cappellini, Picchioni, & Piva, 2010). The images can be identified using their digital signatures, which can be regarded as similar to fingerprints. The fingerprint of an image can be extracted to identify the model or make of the source device or the specific device of the same model that created the image. The type of digital signature considered in this paper is the sensor pattern noise (SPN), which is a high frequency signal (noise) that occurs in the image due to the imperfections in the sensor of the imaging device. The classification of the images can be performed by grouping together images that have similar digital signatures (fingerprints).

Our objective is to assess the performance of an unsupervised classifier proposed in (Li, 2010) by performing a statistical analysis of the classification process. We want to find how reliable this technique is for the forensic investigation of digital images recovered from storage devices. The cross validation technique is used to calculate the estimated prediction error. Another objective is to identify the optimal position for cropping the images from which the fingerprints will be extracted. To achieve this, selections of crops from different parts of the images were taken to investigate the effect of saturation and scene details in the images.

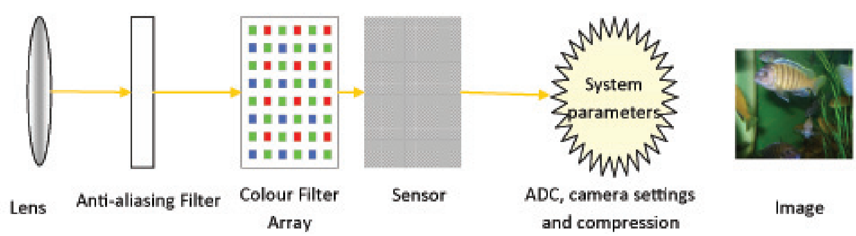
The paper is organised as follows. The image acquisition process through the camera pipeline is described followed by a brief review of the different methods for identifying digital images, in particular sensor pattern noise. Thereafter, the need for image classification is elaborated and two classifiers are described, followed by an overview of different methods of implementing the cross validation technique. The experimental procedures are described and the results are discussed followed by the conclusions.

IMAGE ACQUISITION PROCESS

In the field of digital image forensics (Figure 1), research has been undertaken to help identify the source of digital images by extracting and using their digital signatures. The various image-processing stages or components used in a digital camera leave traces in the resultant images produced and these artefacts can be used to identify the source device. Figure 1 depicts these different stages in the image acquisition process in the camera pipeline.

The light enters the camera through the lens, where the light is focused and the light passes through the anti-aliasing filter. This filter acts as a low-pass filter to prevent spatial frequencies higher than that of the individual pixel in the sensor from passing through, which will otherwise create aliasing (Moiré effect). The colour filter array (CFA) is a filter that will capture the colour components of the light stream. There are different types of CFAs and the Bayer filter, shown in the figure, stores the red, green and blue (RGB) colours where each pixel will store one colour and interpolates the other two colours from the neighbouring pixels. Camera manufacturers use different methods, known as CFA interpolation and demosaicing, to calculate the remaining colours that the pixel does not store. The sensor is the most expensive component of the camera and is

Figure 1. Image acquisition process inside a digital camera



usually of two main types, the complementary metal–oxide–semiconductor (CMOS) and the charge-coupled device (CCD). Traditionally, CCD has been more commonly used but CMOS is being used more often, mainly in mobile (cell) phones cameras, because it consumes less power.

The sensor is monochromatic, hence the need to use a CFA to extract the colour components of the image. The sensor is where the light photons are converted to electrical signals and after this stage various software processing is performed on the signal. The analogue signals are converted to digital signals by using AD (analogue to digital) converters and various camera settings are then applied. In most low to medium end digital cameras, the raw image is compressed and stored in memory in order to save storage space. Most digital cameras use the JPEG compression to store the digital image. JPEG is a lossy compression technique that eliminates the low frequencies and high frequencies (mostly outside the human visual range) of an image depending on the quality of the output image. The higher the compression ratio is, the smaller the size of the resulting file and poorer the quality of the image.

DIGITAL IMAGE IDENTIFICATION

Digital images can be identified using their metadata (Exif header) that are attached to each image taken by a digital camera. However, this method of identification is not reliable since the metadata can be easily removed or modified by image editing software and when uploading to a website (e.g. social network site). Therefore other methods, which are more robust to editing, have to be used for camera identification. There are several techniques that can be applied to the artefacts left by the stages in the camera pipeline for identifying and linking source devices, such as lens aberration, colour filter array (CFA) interpolation and demosaicing, camera response function (CRF), JPEG compression, sensor pattern noise (SPN), higher order wavelet statistics. Lens produces aberrations due to the design and manufacturing process of these lenses. There are two main distortions that have been studied which are lens radial distortion and chromatic distortion. Lens radial distortion occurs when straight lines from the object are rendered as curved lines on the sensor of the camera and the difference between the distorted line and the straight line can be measured an used to identify the camera (San Choi et al., 2006). Chromatic distortion occurs when light of different wavelengths converge at different positions on the camera sensor, which causes misalignment of the RGB channels. The distorted parameters between the colour channels can be estimated and used to identify source devices (Van et al., 2007).

The different algorithms used by manufacturers to interpolate the other colours not stored for one pixel can be used to identify the model or make of the camera. The identification of the CFA interpolation and demosaicing algorithms present in digital images can be performed by calculating the correlation between the different colour channels in a colour image and estimating the demosaicing algorithm used to produce the image (Bayram et al., 2005; Gunturk et al., 2005; Swaminathan et al., 2007). The camera response function (CRF) maps the scene radiance from the image intensity. The CRF can be estimated by finding the mapping algorithm using a single image, and the imaging device can be identified as the source of that image (Lin et al., 2004; Ng et al., 2007). The CRF estimation method has also been applied to the detection of image splicing (Hsu & Chang, 2007).

Most digital cameras perform JPEG compression before storing the images in memory and quantization tables are used to determine the quantized effect (rounding off) on the high and low spatial frequencies of the image in the frequency domain. Quantization tables vary between camera manufacturers and different camera models from the same manufacturer (Farid, 2006). Digital images are usually recompressed for storage or transmission and in these cases the original source device can be identified (Sorell, 2009). Higher order wavelets statistics were applied, for camera model and make identification, together with binary similarity measures and image quality measures as well as a SVM classifier by Celiktutan et al. (2008). The sensor pattern noise (SPN) consists mainly of the PRNU (photo-response nonuniformity) noise and occurs in the sensor of the camera (Lukas et al., 2006). The SPN can be used to identify source devices and whether the image in question has been tampered (Chen et al., 2008). The next subsection elaborates upon the SPN in more details.

The CFA, CRF, JPEG compression and statistical techniques can be used to identify a particular model or make of a camera whereas the lens aberration and SPN can be used to ascertain distinct devices of the same model. Some of these methods need specific assumptions to be made before processing images but the SPN technique does not require the assumptions. Given that the lenses of higher end digital cameras are exchangeable and lenses can also be changed with relative ease on lower end ones, the lens aberration component of an image will also change which indicates that the camera identification will fail. The sensor, which is relied upon by the SPN technique, is much harder to change as well as being more expensive and hence it is uncommon for a sensor to be changed.

Sensor Pattern Noise

The sensor pattern noise (SPN) occurs due to imperfections that are formed during the manufacturing process of the sensors and due to slight variations in which individual pixels convert light to electrical energy (Fridrich, 2009). A combination of the uniqueness of the imperfections in the silicon material and the different sensitivity of the pixels makes the SPN ideal for differentiating between sensors, i.e. cameras, even if they are made from the same silicon wafer. The SPN, n, which appears as a high frequency signal, can be extracted from an image, I, based on the model proposed in (Lukas et al., 2006) as a high pass filter:

$$n = I - f(I) \tag{1}$$

Where f is a denoising function, which acts as a low pass filter to extract the noise from the image.

There are several denoising filters that can be used and two of the methods implemented for the signature extraction are a Gaussian filter in the spatial domain and a wavelet domain based approach. The Gaussian filter is two dimensional where the variance of the filter can be varied to choose its cut-off frequency that will determine the level of scene content and sensor noise (Alles et al., 2008). The second approach applies wavelet decomposition to represent the image in different details levels. Then a noise filter, which is a Wiener filter, is applied to the details levels and an image reconstruction is performed to obtain the noise free image. The denoising filter is described in details in Lukas et al. (2006). The wavelet domain filtering approach is claimed to give better results than others (Mihcak et al., 1999; Lukas et al., 2006). The SPN is not the dominant component of the noise residuals that are extracted from the image, thus a smoother image (with less scene details) will provide better magnitude of SPN. However, as reported in (Sencar & Memon, 2008), there are some limitations to using SPN as a fingerprint; it is easily contaminated by details of the scene (which are also high frequency signals with higher magnitude), by saturation due to light sources (flash, sun, light bulb) and rotation. This leads to a high misidentification and misclassification rate. A whole image has to be used for the extraction of the fingerprint in order to get a reasonable identification rate. Furthermore, the extraction process can take a long time and it is reported in Hoglund (2009) that it takes about 30 hours to calculate the reference noise using 200 images of sizes 3072 x 2304 pixels. Instead of using a whole image, the computational cost of the extraction process can be greatly improved by using only part (crop) of the image. A trade-off must be found between speed and accuracy of identification and classification rate.

The SPN from Equation (1) can be improved by using the enhancer developed in (Li, 2009) to attenuate the interference of scene details. There are 5 models that are described and the choice of a model depends on the strength of the SPN components in the image, but all the models attenuate the stronger signals in the image to allow the SPN to be maximised. The enhanced SPN increases the identification rate and allows the use of smaller crops of images to be used for classification.

IMAGE CLASSIFICATION

There has been a fair amount of research performed in the recent years on the identification of source devices using their digital signatures, but the classification of images for forensic investigation purposes has not been investigated to the same depth. Some of the device identification methods mentioned in the previous section performs a certain level of classification, for example by using a support vector machine (SVM). But classification is mainly employed for the purpose of using or combining sets of different features, extracted from digital images known to come from the source device, to obtain the fingerprint of that source device. In the majority of cases when forensic investigators recover digital images from storage devices, they do not possess the source device that created the image. Hence a classification technique has to be employed that do not require any prior knowledge of the source. There are two clustering techniques that use the SPN as the digital signature and that do not require any previous information about the images and they are Bloy (2008) and Li (2010).

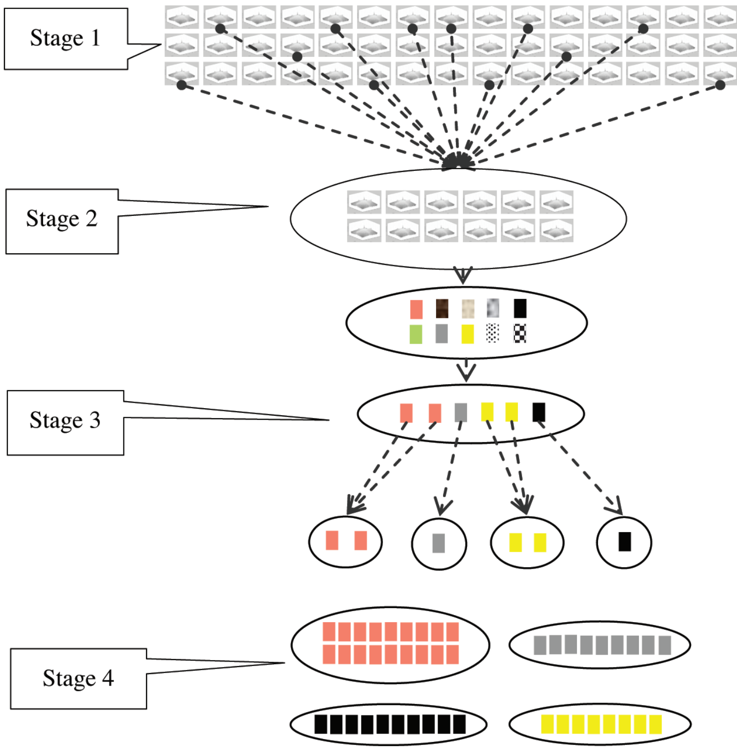
In Bloy (2008), images are classified from a mixed set of images without any previous information about the source devices. The extraction of the signatures is simplified in order to decrease the computing complexity. A median filter was used to perform the denoising instead of using a wavelet based one and only the green channel, of the RBG planes, of the image was used. A threshold value is calculated by correlating a selection of signatures, created using different amount of images, from a given camera with single images from a different camera. The stages of the classifier are summarised below:

1. Iterate randomly through pairs of images in the dataset until one whose correlation is greater than the threshold and average this pair to form the signature.
2. Correlate the rest of the images with the fingerprint, each time assign the signature to the fingerprint until 50 images have been averaged (clustered) or all images have been used.
3. Correlate the remaining unclustered images with the fingerprints obtained in step 2 and checking against the threshold (a maximum of 50 images are used to form the fingerprint but more than 50 images can be associated with the cluster).
4. Repeat step 1 until the stopping condition is reached, i.e. tried enough pairs or all of them without success.

The aim of the unsupervised image classifier in Li (2010) is to cluster a large set of images taken from an unknown number of cameras into groups of images corresponding to their source cameras. The stages of the classifier are shown in Figure 2 without the source devices present.

In Stage 1, the images are cropped and the SPN is extracted and enhanced. For Stage 2, the training set is randomly selected (M) from the dataset and the similarity matrix (S), of size MxM, is created by calculating the correlation between all the SPNs. In Stage 3, the classifier trainer using conditional random fields (CRF) is executed based on S and for each fingerprint a membership pool is created and the class labels are updated iteratively until a predefined stopping condition (no change in membership throughout an entire iteration) is reached. A typical strategy for a stopping condition may limit the amount of class label changes. At this stage clusters of class labels are formed and in the following Stage 4, the rest of the images

Figure 2. Stages of the unsupervised classification of images



are compared to the centroids of the clusters and are subsequently assigned to the respective cluster of which the centroid is closest to the images.

The first classifier needs a level of knowledge about the different cameras, or using some other camera sources, at the initial stage of the classification process in order to calculate the threshold.

Furthermore this technique needs a literal amount of signatures (50) in order to form the fingerprint cluster. The second classifier on the other hand operates without any prior information about the sources and uses the training stage to form the clusters without any limit on the number of signatures to use.

CROSS VALIDATION

The accuracy of a classifier can be estimated by finding its associated prediction error. The main techniques that can predict the errors of classifiers are Holdout (Webb, 1999), Bootstrap (Efron & Tibshirani, 1997) and cross validation, which is probably the simplest and most often used (Hastie et al., 2009). Cross validation can be used to compare between classifiers or, in our case, implemented to predict the performance of a single classifier with different sample sizes. This technique is most commonly used in supervised learning, by dividing the dataset into 3 parts, training set, validation set and test set and can be used to predict the performance of the supervised learning classifier. The training is performed on the training set, then cross validation is applied to calculate the training error, and finally the test error is calculated based on the test set. The training error can be compared against previous values that were learnt by the classifier. For unsupervised learning classifier, since there is no memory of previous datasets or features, the application of cross validation is not straightforward. We only use a training set and a test set, since we already have the clusters formed after the classifier training (stage 3 in Figure 1). The prediction error is calculated by finding the number of class labels that have been misclassified from our dataset.

K-fold cross validation (Figure 3) is a method where the dataset δ of size *N*, is randomly partitioned into *k* mutually exclusive sets (folds) of roughly the same size. The classification is performed on *k* -1 partitions (training set) and the remaining partition is applied to the classifier (test set) as shown in Figure 3 where the grey boxes correspond to the test set and the white boxes to the training set. The number of labels that are falsely classified from the test set will provide the prediction error $\xi(x_i)$ of the *i*th partition x_i. The process is repeated *k* times, where each partition is used as the test set in turn, the overall prediction error CV_{err} is given by:

$$CV_{err} = \frac{1}{K}\sum_{i=1}^{k}\delta(x_i, \xi(x_i)) \tag{2}$$

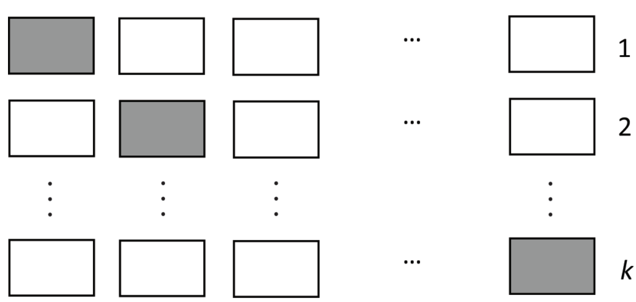
Take the example when the value of *k* is 5 and the size of the dataset is 100, with 20 acquired by each camera. There will be 5 partitions of about 20 items in each and cross validation will be executed 5 times. The value of *k* can be 10 or 20 and if the value of *k* is 2, *k*-fold becomes a variation of the holdout method. In the case when the value of *k* is 1 (also known as leave-one-out), the cross validation will be run 100 times, which is computationally expensive as well as giving a high variance. Although there is a high variance the bias is zero, because all the readings are used for the cross validation. The other values of *k* will give more bias than leave-one-out but their variance is lower. The typical value of *k* = 10 gives a good trade-off for bias and variance as well as being less computationally intensive.

The folds (partitions) can also be stratified, where they will each contain approximately the same amount of labels as compared to the dataset. This ensures that the folds are representative of the distribution of labels in the dataset. Repeating the cross validation multiple times by using different labels for the partitions after each complete run, generally provide a better Monte-Carlo estimate at an added cost (Kohavi, 1995), e.g., running the 10-fold cross validation ten times, hence 100 times in total, will provide a better estimation of the error with less variance. The repeated cross validation was not implemented in this experiment due to the additional computational cost it involves and the single cross validation provided an acceptable Monte-Carlo estimate.

EXPERIMENTAL SETUP

For the purpose of our experiments 1000 pictures were used from 5 source devices comprising of 4 digital cameras and one camera phone. The devices were BlackBerry Curve 8310, Nikon Coolpix E5200, Canon Digital IXUS 500, Olympus C-730 UZ, and Olympus Mju Stylus 1030SW. Since images in real forensic situations come from different cameras with

Figure 3. k partitions blocks for cross validation



different specifications, we did not want to set a specific size for the photos from the cameras, therefore the sizes ranged from 1600 x 1200 pixels to 3648 x 2736 pixels, which still give a wide margin for cropping of the images. All the images were in JPEG format with a compression quality ranging from about 75% to 97% quality. The pictures contain a wide variety of indoor and outdoor sceneries of urban and rural settings, night and day lighting as well as offices and various buildings and some holiday pictures; overall images that would be as close to natural real world situations as possible were captured.

The cropping was performed on the images and the signatures (SPN) were extracted by using a Discrete Wavelet Transform (DWT) followed by a low pass filter and the SPN were enhanced by using one of the models developed in Li (2010). The model chosen was the:

$$n_e\left(\mathrm{i,j}\right) = \begin{cases} e^{-0.5n^2(i,j)/\alpha^2}, & \text{if } 0 \leq n\left(\mathrm{i,j}\right) \\ -e^{-0.5n^2(i,j)/\alpha^2} & \text{otherwise} \end{cases} \quad (3)$$

This has been shown to work in Li (2010) for natural images and the optimal value of α was 7.

To perform the cross validation on the dataset, the folds were randomly selected before the training stage of the classifier. The next step was to create the similarity matrix which, being the most computationally intensive stage of the classification process was no longer a trivial process since for each fold the matrix had to be recalculated. This had the prospect of making the cross validation prohibitively expensive computationally and arises due to the change of size of the matrix by changing the size of the fold and the different images present in the training phase for each partition. To overcome this problem, the similarity matrix was created at the start of the cross validation process for all the SPNs. Row and column deletions were performed in the matrix as the population of the partitions changed and the size of the folds were altered. This method increased the time complexity to create the similarity matrix, but it had to be created only once, which decreased the time complexity during the cross validation process.

The sizes of the folds were chosen as 20, 10, 5, and 2, which provided a wide range of prediction errors. The size of the dataset was decreased from 1000 to 500 and to 250 so as to provide a learning curve that will indicate what effect the sample size has on the classifier. Once the value of k is chosen for the cross validation and one partition set aside for testing, the other partitions are combined to provide data for the training phase and classification phase. In Li (2010), the selected sizes of the classifier trainer set were 120 and 300, where there was little variation of classification performance between the two sizes. The size of the trainer set was altered with different dataset sizes. Moreover, the partitions were stratified, so that each partition contained approximately the same amount of labels.

Image Cropping

The images were cropped to a size of 512 x 512 pixels, which provide a good balance between the computational complexity and the accuracy of the signature. The speed of processing a 1024 x 1024 size image will quadruple when compared to an image of size 512 x 512. Hence cropping the images from their normal sizes greatly reduces the computational complexity. The magnitude and accuracy of the SPN does not deteriorate greatly when the image is cropped to 512 x512 and preliminary tests performed in Li (2010) indicated the size of the cropping can be as low as 256 x 512.

The images were cropped from three positions, the centre, the upper left corner and lower left corner. The preference for choosing these positions was based on the fact that there can be large differences in light levels and scene details between the lower and upper parts of the images as well as the centre. Moreover, the right side and left side of a picture present approximately the same amount of lighting and same average scene details, hence the decision to use only one side of the images to perform the cropping.

RESULTS

The first set of results for one run of cross validation when k=10, was giving a high percentage of errors when cropping from the centre of the images. The position of the cropping was then changed from the centre to the upper left corner and lower left corner of the image. All three different positional cropping of the SPNs were used to classify the same set of images to determine the position with the lowest classification errors. The average percentage errors for the centre, upper-left corner and lower-left corner were 2.2, 2.5, and 0.81 respectively. The results showed that, for our dataset, the best position to perform cropping of an image for classification is the lower left corner. The centre of the image usually has a high light intensity (saturation) in the event the flash is activated when the picture is taken at night or in low light, similarly the upper left corner of the image can have high intensity of light since light sources from light bulbs or sunlight, are most often present in that space of the picture. Outdoor pictures taken at night will affect the top of the image too, due to the low light intensity in this area of the image which will adversely affect the multiplicative nature of the SPN. The saturation of the pixels in these areas of the images usually corrupts the SPN of the image. The lower left corner of the image usually has normal light intensity and less scene details than the centre of the image. Hence we chose to crop the images from the lower left corner. These findings are different from Li (2010) because our dataset had a proportion of images taken during the night and indoors, when the flash was fired as well as in bright sunlight where the top part of the image was saturated. The edges of the images can be providing some further help to enhance the SPN due to the edge effect.

Cross validation was performed with sizes of folds of 2, 5, 10, and 20 and it was found that the variance of the error rates increases when the number of folds decreases.

Figure 4 shows the error and variance for different folds when cross validation was performed on 1000 images. Two fold and five fold cross validation produced higher values of variance whereas twenty fold produced the less variance of errors, but the variance of ten and twenty fold are close to each other. It can be seen that simply adding more folds beyond k = 10 does not necessarily yield drastically better results. It was also observed from the results that when the size of the sample space was increased the variance tends to decrease, because the classifier has more images to train in order to find the clusters.

In Figure 5, the error rates for all three sample sizes (1000, 500, 250) are shown with different sizes of folds. The error rates for k = 10 and k = 20 are quite similar for all three curves and there is large difference when the number of folds goes towards k = 2. Taking the computational cost when k = 20 into consideration, it is preferable to use k = 10 for our dataset. It is interesting to note that when the sample size is 500, the error rates are at their lowest values. The reason for these low error rates is because the images in this sample size were chosen at random and they contained the least amount of corrupted images. For a smaller sample size, 250, the error rate increases drastically when the value of k is less than 5.

The error rates in Table 1 shows that the unsupervised classifier performs better when the size of the training set is less than 50% of the size of the sample space but more than approximately 125 images. When the size of the sample space is 1000 and 250, the error rates obtained were higher when the training set was nearly as large as the size of the sample space. This may occur due to over training in stage 3 of the unsupervised classifier as depicted in Figure 2, where the intra class and extra class boundaries of the clusters appear to overlap and thus indistinguishable.

The error rates included in Table 1 are the estimated prediction error rates (P_{est}). They are the maximum likelihood estimate of the true error rate (P_{true}) of the classifier. The relationship between P_{est} and P_{true} with a 95% confidence interval (95% being considered high enough to be accepted by a forensic investigator) for dif-

Figure 4. Variance and error rate with respect to number of folds for 1000 images

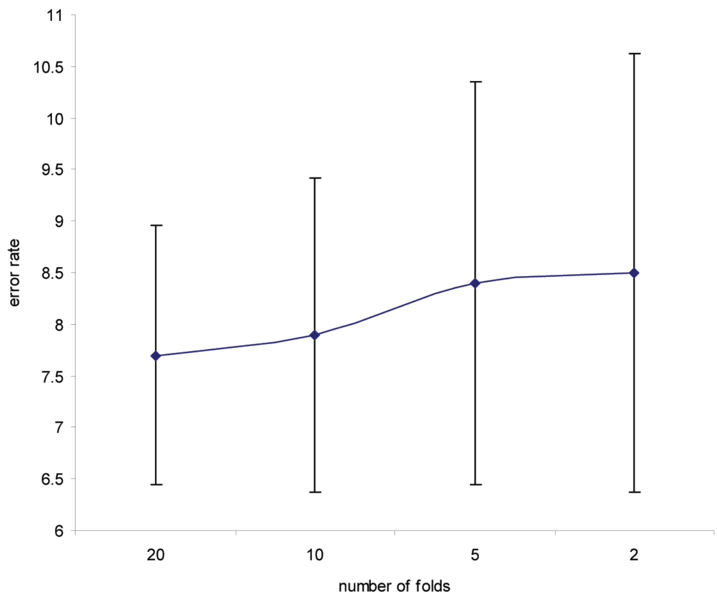


Table 1. Percentage error rates for different sample sizes and varying classifier trainer size when k =10

Sample Size	Classifier Trainer Size			
	900	*500*	*250*	*125*
1000	0.83	0.79	0.74	0.91
500		0.54	0.58	0.88
250			1.84	1.12

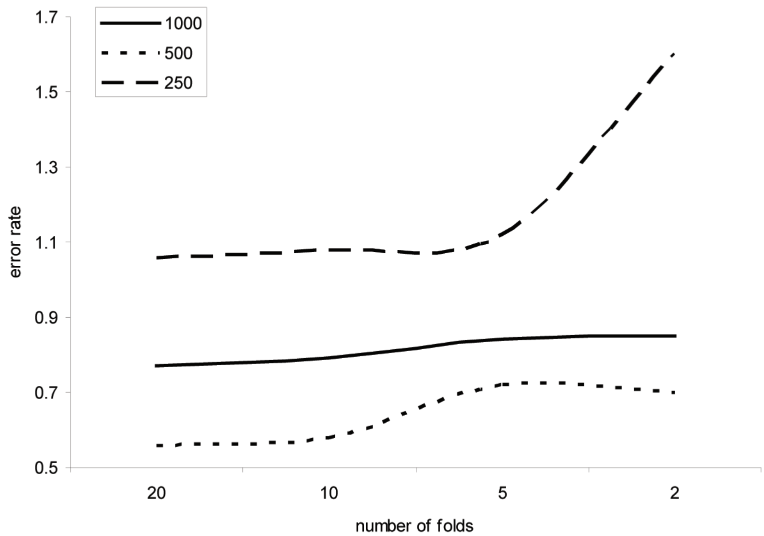
ferent sample size has been calculated. The average P_{est} including all the tests was below 2%, giving P_{true} between 0.5% to 6% and 1.6% to 3% for sample size of 125 and 1000 respectively.

Furthermore, an additional test was performed to check if the classifier will perform with a small dataset of 50 images comprising of 7, 18, 15 and 10 pictures from the BlackBerry Curve 8310, the Nikon Coolpix E5200, the Canon Digital IXUS 500, and the Olympus Mju Stylus 1030SW respectively. The signatures were clustered in their respective groups according to their camera source together with an outlier group that consisted images that had high level of saturation and darkness.

CONCLUSION

A statistical analysis of an unsupervised image classifier was performed in order to assess its performance for the purpose of forensic investigation of images. Using cross validation it was found that the unsupervised classifier performs with error rates less than 6% with a 95% confidence interval when *N* is 125. However, the performance was achieved using optimal values

Figure 5. Effect of number of folds on percentage error rate for different sample sizes of images



such as fold size, size of the sample space, cropping position. It has been showed that performing a ten fold cross validation provides the best trade-off between computational cost and variance of the errors. The best size of the training set to use in order to give lower error rates was also determined.

It was shown that using the lower corner for cropping of an image will provide the best classification rate results for a dataset chosen in similar conditions to this experiment. The further tests performed on the set of 50 images showed that the classifier will cluster smaller groups of images than predicted. Based on the results from the dataset, the classifier performs reliably for the given size of the sample space and can be justifiably used for forensic investigation of images under the conditions set in this experiment.

However, the error rates that were measured are predominantly due to saturation. More research needs to be carried out in the area of SPN extraction, which is our future work in this field.

ACKNOWLEDGMENTS

This work was jointly funded by the Acorn fund of Keele University, UK and Forensics Pathways Ltd, UK. The unsupervised classifier has led to a pending UK patent (Application Number 0902406.5) for Forensic Pathways Ltd. Thanks go to John Butcher and Max Legg for participating in the image acquisition process and the reviewing stage. The unsupervised classifier has led to a pending UK patent (Application Number 0902406.5) for Forensic Pathways Ltd. Thanks go to John Butcher and Max Legg for participating in the image acquisition process and the reviewing stage.

REFERENCES

Alles, E. J., Geradts, Z. J. M. H., & Veenman, C. J. (2008). Source camera identification for low resolution heavily compressed images. In *Proceedings of the International Conference on Computational Sciences and Its Applications* (pp. 557-567).

Amerini, I., Caldelli, R., Cappellini, V., Picchioni, F., & Piva, A. (2010). Estimate of PRNU noise based on different noise models for source camera identification. *International Journal of Digital Crime and Forensics*, *2*(2), 21–33. doi:10.4018/jdcf.2010040102

Bayram, S., Sencar, H., Memon, N., & Avcibas, I. (2005). Source camera identification based on CFA interpolation. In *Proceedings of the IEEE International Conference on Image Processing*, *3*, 69–72.

Bloy, G. J. (2008). Blind camera fingerprinting and image clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *30*(3), 532–534. doi:10.1109/TPAMI.2007.1183

Celiktutan, O., Sankur, B., & Avcibas, I. (2008). Blind identification of source cell-phone model. *IEEE Transactions on Information Forensics and Security*, *3*(3), 553–566. doi:10.1109/TIFS.2008.926993

Chen, M., Fridrich, J., & Goljan, M. (2008). Determining image origin and integrity using sensor noise. *IEEE Transactions on Information Forensics and Security*, *3*(1), 74–90. doi:10.1109/TIFS.2007.916285

Efron, B., & Tibshirani, R. J. (1997). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.

Farid, H. (2006). *Digital Image Ballistics from JPEG Quantization* (Tech. Rep. No. TR2006-583). Hanover, NH: Dartmouth College.

Fridrich, J. (2009). Digital image forensic using sensor noise. *IEEE Signal Processing Magazine*, *26*(2), 26–37. doi:10.1109/MSP.2008.931078

Gunturk, B. K., Glotzbach, J., & Altunbasak, Y. (2005). Demosaicking: Color filter array interpolation in single chip digital cameras. *IEEE Signal Processing Magazine*, *22*(1), 44–54. doi:10.1109/MSP.2005.1407714

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer-Verlag.

Hoglund, T. (2009). Digital camera identification: A brief test of a method based on the sensor noise. *Journal of Forensic Identification*, *59*(5), 475–501.

Hsu, Y. F., & Chang, S. F. (2007). Image splicing detection using camera response function consistency and automatic segmentation. In *Proceedings of the IEEE International Conference on Multimedia and Expo* (pp. 28-31).

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Vol. 2, pp. 1137-1145).

Li, C.-T. (2010). Source camera identification using enhanced sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, *5*(2), 7–11.

Li, C.-T. (2010). Unsupervised classification of digital images using enhanced sensor pattern noise. In *Proceedings of the IEEE International Symposium on Circuits and Systems* (pp. 3429- 343).

Lin, S., Gu, J., Yamazaki, S., & Shum, H.-Y. (2004). Radiometric calibration from a single image. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2*, 938–945.

Lukas, J., Fridrich, J., & Goljan, M. (2006). Digital camera identification from sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, *1*(2), 205–214. doi:10.1109/TIFS.2006.873602

Mihcak, M. K., Kozintsev, I., & Ramchandran, K. (1999). Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, *6*, 3253–3256.

Ng, T.-T., Chang, S.-F., & Tsui, M.-P. (2007). Using geometry invariants for camera response function estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8).

San Choi, K., Lam, E. Y., & Wong, K. K. Y. (2006). Source camera identification using footprints from lens aberration. In *Digital Photography II. Proceedings of the Society for Photo-Instrumentation Engineers*, *6069*, 172–179.

Sencar, H. T., & Memon, N. (2007). *Overview of state-of-the-art in digital image forensics*. Singapore: World Scientific Press.

Sorell, M. J. (2009). Conditions for effective detection and identification of primary quantisation of re-quantized JPEG images. *International Journal of Digital Crime and Forensics*, *1*(2), 13–27. doi:10.4018/jdcf.2009040102

Swaminathan, A., Wu, M., & Liu, K. J. R. (2007). Nonintrusive component forensics of visual sensors using output images. *IEEE Transactions on Information Forensics and Security*, *2*(1), 91–106. doi:10.1109/TIFS.2006.890307

Van, L. T., Emmanuel, S., & Kankanhalli, M. S. (2007). Identifying source cell phone using chromatic aberration. In *Proceedings of the IEEE International Conference on Multimedia and Expo* (pp. 883-886).

Webb, A. (1999). *Statistical pattern recognition*. Oxford, UK: Hodder Arnold.

Ahmad Ryad Soobhany received his B.Eng. (Hons) degree in Electrical & Electronics Engineering from University of Mauritius, Mauritius and his M.Phil. degree in Computer Science from Keele University, UK. He is currently working as a Software Engineer at Forensic Pathways Ltd, UK and is a PhD candidate at Keele University, UK. His research interests include digital forensics, image processing, blind source separation, intelligent agents, machine learning, computer vision and pattern recognition.

Richard Leary, MBE, LLB (Hons) is the Managing Director of Forensic Pathways. He is also Chairman of the Regional ICT Advisory Board to the Regional Development Agency in the UK. He has a Ph.D in forensic evidence and is a Senior Research Fellow at University College London, Visiting Professor at Warwick University and Guest Lecturer at Birmingham City University. Richard is also a lawyer. He has investigated organized serious crime and advised government of the use of intelligence systems and processes. He is an acknowledged expert in informatics and the use of interactive computing and information systems to aid the process of discovery, investigation and risk management. He has particular expertise in informatics and evidence and intelligence management. He has worked in the UK, Europe, USA and the Middle East. He is currently participant of two FP7 projects – Odyssey and DIVeFor.

KP Lam joined Computer Science at Keele University in 1997/98 and has extensive experience of working with industry. His research interests have a significant overlap in the fields of digital multimedia, computer vision/visual analytics and biometric security technologies for distributed systems. He was a major grant holder of the EPSRC (UK) funded project (entitled Element Specific X-ray Imaging for Security applications, 2005-07) jointly with the School of Chemistry and Physics, and Criminology, and latterly, the UWSP/AWM (UK)Proof of Concepts funded project that concerned the joint development of a digital media based authentication protocol/system with a company in the West Midlands. Currently, he is working on the EPSRC/Keele funded project (3ME) to explore and develop innovative methods of visualising and analysing biomedical data/images including MRI scans, X-rays/FT-IR spectra, and live cells.